

TESIS

MODELACIÓN, SIMULACIÓN DE FUNCIONAMIENTO Y EVALUACIÓN DE
PRESTACIONES DE PROTOCOLOS DE RED CON REDES DE PETRI
DESARROLLO DE HERRAMIENTAS ACADEMICAS DE ENSEÑANZA

Alumno: Santiago C. Pérez

Director: Mario Distéfano UTN FRMendoza

Sub-Director: Luis Marrone UNLP

Trabajo Final presentada para obtener el grado deMAGISTER EN REDES DE DATOS
Facultad de InformáticaUniversidad Nacional de La Plata

Abril de 2006

TESIS

Alumno: Santiago C. Pérez Maestría de Redes de Datos UNLP Año 2006

MODELACIÓN, SIMULACIÓN DE FUNCIONAMIENTO Y EVALUACIÓN DE PRESTACIONES DE PROTOCOLOS DE RED CON REDES DE PETRI DESARROLLO DE HERRAMIENTAS ACADÉMICAS DE ENSEÑANZA

INDICE

<u>i) CAPITULO INTRODUCTORIO</u>	6
ii) OBJETIVO	6
iii) LA MOTIVACIÓN	7
iv) METODOLOGIA	7
v) RESULTADOS ESPERADOS	8
vi) VISION GENERAL DE LOS CAPITULOS	8
<u>CAPITULO I</u>	
<u>ASPECTOS TEÓRICOS SOBRE LAS REDES DE PETRI</u>	9
I.1) INTRODUCCIÓN	9
I.2) REDES de COLAS	9
I.2.1 Redes de colas básicas	9
I.2.2 Redes de colas	10
I.2.3 Redes de colas con capacidad infinita y ruteo dependiente del estado	12
I.2.4 Redes de colas extendidas	13
I.2.5 Notación de Kendall-Lee	13
I.3) REDES DE PETRI NO TEMPORIZADAS	13
I.3.1 Redes y Sistemas Lugar/Transición	15
I.3.2 Sistemas de Red: Marcado y Juego de marcas	17
I.3.3 Semántica de Intercalación: Observaciones secuenciales	18
I.3.4 Localidad y la Síntesis del Modelo	19
I.3.5 Dependencia causal, conflictos y concurrencia	20
I.3.6 Subclases y Extensiones del Formalismo Básico	21
I.3.7 Subclases sintácticas	22
I.3.7.1 Redes Ordinarias y Sistemas de Red Elementales	22
I.3.7.2 Subclases Topológicas	23
I.3.8 Arcos Inhibidores y Prioridades	24
I.3.9 Redes y Sistemas Coloreados	24
I.4) REDES DE PETRI TEMPORIZADAS	25
I.4.1 Temporización de transición	26
I.4.2 Reducción del No Determinismo	27
I.4.2.1 Reducción no determinística	27
I.4.2.2 Reducción estocástica	28
I.4.2.3 Reducción determinística	29
I.4.2.4 Mezcla de técnicas de reducción diferentes	29
I.4.3 Transiciones Temporizadas e Inmediatas	29
I.4.4 Especificación de Servidores, conflictos y políticas (Apéndice B)	31
I.4.5 Observación en el tiempo: Ecuaciones de ocurrencia	32
I.4.6 Análisis Operacional	32
I.4.7 Observación y especificación de modelos temporizados	33

I.4.8 Casos de Markov y semi-Markov (Apéndice C).....	33
I.4.9 Redes de Petri Estocásticas Generalizadas (GSPNs).....	33
I.4.10 Redes Estocásticas Bien formadas (SWNs).....	34
I.4.11 Redes de Petri Determinística y Estocástica (DSPNs).....	34
I.5) PROPIEDADES LOGICAS DE LOS SISTEMAS P/T Y SU ANÁLISIS.....	35
I.5.1 Propiedades lógicas básicas.....	35
I.5.2 Técnicas de análisis básicas para sistemas de red.....	38
I.5.3 REDES DE PETRI ESTOCÁSTICAS EXPONENCIALES-SPN (Apéndice D).....	39
I.6) REDES DE PETRI COLOREADAS (CPNs).....	39
I.6.1 Definición formal.....	40
I.6.2 Redes Bien Formadas.....	41
I.6.3 Técnicas de Análisis para Sistemas Coloreados Bien Formados.....	41
I.7) REDES DE PETRI DE SISTEMAS DE COLA DE POBLACIÓN FINITA (FPQSPNs).....	42
I.7.1 Introducción.....	42
I.7.2 Definición formal.....	42
I.7.3 Propiedades de Tiempo.....	43
I.7.4 Matriz de Incidencia.....	43
I.7.5 Algoritmo de Desplegado.....	45
I.7.6 Graficación de las Características 2D de los Modelos de Redes de Petri.....	46
I.8) REDES DE PETRI VERSUS REDES DE COLAS.....	48
I.8.1 Introducción.....	48
I.9) COMPARACION ENTRE TIPOS DE REDES DE PETRI.....	49
I.9.1 Introducción.....	49
I.9.2 Hacia un formalismo de Redes de Petri útil prácticamente.....	49
I.9.3 Formalismos de Redes de Petri.....	50
I.9.4 Formalismos de Redes de Petri y los protocolos de comunicaciones.....	50
I.10) CONCLUSION.....	51

CAPITULO II

<u>MODELACION DE PROTOCOLOS DE COMUNICACIONES</u>	52
II.1) PROTOCOLOS DE RED (definición general de los mismos).....	52
II.1.1 Introducción.....	52
II.1.2 Medidas de performance de los Protocolos de Comunicación.....	52
II.1.2.1 Carga ofrecida.....	53
II.1.2.2 Rendimiento.....	54
II.1.2.3 Ancho de banda.....	54
II.1.2.4 Retardo.....	54
II.1.3 Características del tráfico de red.....	55
II.1.3.1 Influencia del Modelo OSI para el Proceso de Arribo de Paquetes.....	55
II.1.4 Modelación de Redes de Petri de los Procesos de Arribo de Paquetes.....	56
II.1.4.1 Proceso de Arribo de Paquetes Determinístico.....	56
II.1.4.2 Proceso de Arribo de Paquetes Poisson.....	56
II.1.4.3 Proceso de Arribo de Paquete Poisson con Recuperación.....	56
II.1.4.4 Proceso de Arribo de Paquetes Poisson Modulado con Markov.....	57
II.1.4.5 Proceso de Arribo de Paquetes Poisson con Recup. Modulado con Markov.....	58
II.2) MODELACIÓN DE ALOHA.....	59
II.2.1 Introducción.....	59
II.2.2 Aloha puro.....	59
II.2.2.1 Modelación de Aloha puro.....	60

II.2.2.2 Canal de Comunicaciones.....	60
II.2.3 Aloha ranurado.....	61
II.2.3.1 Propiedades comunes	61
II.2.3.2 Modelación de Aloha ranurado.....	62
II.3) MODELACIÓN DE CSMA/CD.....	63
II.3.1 Introducción.....	63
II.3.2 Conducta Ethernet.....	63
II.3.3 CSMA/CD como Sistema de Población Infinita	65
II.3.4 Modelación de CSMA/CD.....	65
II.3.4.1 Confianza en el modelo	66
II.3.4.2 Modelación de CSMA persistente	67
II.3.5 Modelación de CSMA/CD (considerado como un sistema M/D/1).....	68
II.3.6 Modelación de CSMA/CD (considerado como un Sistema M/M/1).....	69
II.3.7 CSMA/CD con Distribución de longitud de Paquete Bimodal	70
II.3.8 CSMA/CD como Sistema de Cola con Recup. y con Perdida de Paquetes	71
II.3.8.1 CSMA/CD como Sistema de Cola con Recuperación.....	71
II.3.8.2 CSMA/CD con Perdida de Paquete	72
II.3.9 Análisis de Ethernet	73
II.3.9.1 Binary Exponencial Backoff.....	73
II.4) MODELACIÓN DE 802.11	74
II.4.1 Introducción	74
II.4.2 Mecanismos MAC en Wireless LANs.....	74
II.4.3 El DCF del IEEE 802.11.....	75
II.4.4 Modelación 802.11 en Acceso Básico como CSMA no persistente.....	77
II.5) MODELACIÓN DE ALOHA PURO CON COLOREADAS	79
II.5.1 Introducción	79
II.6) CONCLUSION	80

CAPITULO III

<u>SIMULACION DE PROTOCOLOS DE COMUNICACIONES</u>	82
III.1) PRINCIPIOS DE SIMULACIÓN.....	82
III.1.1 Modelos de Sistemas	83
III.2) SIMULADORES DE REDES DE PETRI.....	83
III.2.1 Simuladores existentes	84
III.3) SIMULADOR DE ANIMACIÓN PARA EVALUACION DE LA DINAMICA	85
III.4) SIMULACIÓN DE ANIMACIÓN DE ALOHA.....	90
III.4.1 Simulación de animación de Aloha Puro	90
III.4.2 Simulación de animación de Aloha Ranurado	94
III.5) SIMULACIÓN DE ANIMACIÓN DE CSMA/CD.....	95
III.5.1 Simulación de animación de CSMA	95
III.5.2 Simulación de animación de CSMA/CD.....	99
III.6) CONCLUSION	101

CAPITULO IV

<u>EVALUACION DE PERFORMANCE DE PROTOCOLOS DE COMUNICACIONES</u>	103
IV.1) PRINCIPIOS DE EVALUACIÓN DE PERFORMANCE	103
IV.1.1 Introducción	103
IV.1.2 Performance de los protocolos de acceso al medio.....	104
IV.2) SIMULADOR PARA EVALUACIÓN DE PERFORMANCE	105

IV.2.1 Introducción	105
IV.2.2 Editor de Red de Petri	106
IV.2.2.1 Comandos del Menú en el Editor de Red de Petri	109
IV.2.3 Simulador para Evaluación de Performance	110
IV.2.4 Matriz de la Red de Petri.....	113
IV.3) EVALUACIÓN DE PERFORMANCE DE ALOHA	114
IV.3.1 Evaluación de Performance de Aloha Puro.....	114
IV.3.2 Evaluación de Performance de Aloha Ranurado.....	118
IV.4) EVALUACIÓN DE PERFORMANCE DE CSMA/CD	119
IV.4.1 Evaluación de Performance de CSMA.....	119
IV.4.2 Evaluación de Performance de CSMA/CD	121
IV.5) EVALUACION DE PERFORMANCE DE ETHERNET	129
IV.6) EVALUACION DE PERFORMANCE DE CSMA no persistente.....	130
IV.7) COMPARACION ENTRE MET. ANALITICOS Y STPNPLAY FPQSPN.....	131
IV.7.1 Introducción	131
IV.7.1.1 Aloha puro.....	131
IV.7.1.2 Aloha ranurado	135
IV.7.2 CSMA.....	136
IV.7.2.1 CSMA no persistente.....	136
IV.7.2.2 CSMA persistente	139
IV.7.2.3 CSMA p-persistente	141
IV.7.2.4 CSMA Mp-persistente.....	142
IV.7.2.5 CSMA Mp-persistente óptimo	142
IV.8) CONCLUSION	143

CAPITULO V

<u>CONCLUSIONES</u>	145
----------------------------------	-----

APENDICE A: TABLAS CON VALORES NUMERICOS DEL SIMUL. STPNPlay ...	147
---	-----

APENDICE B: ESPECIFICACIÓN DE SERVIDORES, CONFLICTOS Y POLÍTICAS

B.1 Resolución de Conflictos, y Conflictos y Precedencia	152
--	-----

APENDICE C: CASOS DE MARKOV Y SEMI-MARKOV

C.1 Casos de Markov y semi-Markov	155
---	-----

APENDICE D: REDES DE PETRI ESTOCÁSTICAS EXPONENCIALES.....

D.1 Introducción.....	157
D.2 El Proceso Estocástico Asociado con una Red de Petri Estocástica	157
D.2.1 Máquina de Estado Finito y Grafo Marcado de las Redes de Petri Estocástica	157
D.2.2 Redes de Petri Estocástica con una estructura general	159

ANEXO A: BIBLIOGRAFÍA	163
------------------------------------	-----

ANEXO B: GLOSARIO	165
--------------------------------	-----

ANEXO C: HERRAMIENTAS ACADEMICAS DE ENSEÑANZA

C.1 Introducción	167
C.2 Láminas de Presentación en Multimedia	168

TESIS

Alumno: Santiago C. Pérez Maestría de Redes de Datos UNLP Año 2006

MODELACIÓN, SIMULACIÓN DE FUNCIONAMIENTO Y EVALUACIÓN DE PRESTACIONES DE PROTOCOLOS DE RED CON REDES DE PETRI

DESARROLLO DE HERRAMIENTAS ACADEMICAS DE ENSEÑANZA

i) CAPITULO INTRODUCTORIO

Los Sistemas Dinámicos de Eventos Discretos (DEDS) se basan en los sistemas dinámicos, donde los estados y eventos discretos tienen un rol fundamental. Los DEDS han crecido en interés, en la medida que la importancia de la automatización y control lo hacen en la tecnología moderna. Los ámbitos de aplicación como sistemas de fabricación flexibles, sistemas de transporte, sistemas distribuidos, y **redes de telecomunicaciones** usan intensivamente los conceptos y técnicas de los DEDS.

Los DEDS típicos exhiben evoluciones paralelas, las cuales conducen a conductas complejas debido a la presencia de fenómenos de sincronización y recursos compartidos. Las Redes de Petri (RdP) son un formalismo matemático apropiado para modelación de DEDS concurrentes. Los modelos de red son frecuentemente observados como especificaciones autodocumentadas, a causa de que su naturaleza gráfica facilita la comunicación entre diseñadores y usuarios. Los fundamentos matemáticos del formalismo permiten corrección lógica y análisis de eficiencia.

Más que un simple formalismo, las RdP son una familia de ellos, desde bajo a alto nivel, cada una de ellos más apropiados para diferentes propósitos. En cualquier caso, ellos pueden representar conductas muy complejas desde la simplicidad del modelo real, consistiendo de unos pocos objetos, relaciones y reglas.

Las redes de Petri son un modelo formal (matemático) para describir estados y acciones. Disponen de una representación gráfica, como el caso de un grafo dirigido, en el que intervienen dos clases de nodos: lugares (representados por circunferencias) y transiciones (representadas por segmentos rectilíneos) y todos unidos por arcos.

Las características fundamentales de las RdP son su simplicidad (intervienen pocas y simples entidades matemáticas), su generalidad (capacidad de modelar secuencias, decisiones, concurrencia, sincronizaciones, etc.), su adecuación (capacidad de expresar diferentes estructuras de concurrencia) y su flexibilidad (permite un modelado progresivo, refinamientos sucesivos y una estructura modular)

Son valores adicionales de las RdP su representación gráfica, la existencia de técnicas de validación (como la vivacidad, limitaciones, bloqueos, etc.), la interpretación estocástica (para la evaluación de rendimientos) y la existencia de herramientas de software de diseño y análisis, que implementan las técnicas conocidas.

ii) OBJETIVO

Las técnicas de análisis cualitativas y cuantitativas utilizadas para la evaluación de las redes de telecomunicaciones, y específicamente las de los protocolos de red, tienen constantes progresos. Un resultado concreto es la producción continua de trabajos diversos para modelar y analizar la corrección y la performance del funcionamiento de diversos protocolos.

Sin embargo, las Redes de Petri han demostrado ser la herramienta base para la introducción de descripciones exactas de aplicaciones complejas, y que proveen la posibilidad de explotar la estructura de sus modelos para desarrollar un mejor entendimiento de la conducta dinámica de los sistemas y para deducir técnicas de análisis efectivas.

El trabajo se propone establecer de una manera uniforme, y con el uso de las Redes de Petri, las características, similitudes y diferencias cualitativas y cuantitativas, entre algunos de los protocolos de red de canal compartido utilizados en el ámbito de las redes LAN y WAN. Es decir, se utilizarán las Redes de Petri como modelo matemático, compilando las novedades teórico-prácticas de esta materia, y se seleccionarán herramientas apropiadas de modelación, simulación y de evaluación de performance, para investigar propiedades o características de los protocolos, que otras del tipo analíticas no ponen totalmente de relieve, o que permitan demostrar que ambos enfoques son complementarios. Además, se realizarán experiencias prácticas para validar alguno de los modelos de Red de Petri. Finalmente, a partir de ello, se buscará establecer una secuencia alternativa de enseñanza-aprendizaje de los protocolos de red de canal compartido, usando este instrumento matemático.

iii) LA MOTIVACIÓN

Dada las características de los modelos de red, como su naturaleza gráfica, se facilitaría la interpretación, comunicación y transferencia entre los interesados que usan y discuten sobre los protocolos de red. Además, dado los fundamentos matemáticos del formalismo es posible efectuar un enfoque integrado de validación de los parámetros de eficiencia para la comparación entre los protocolos. Y dejando de manifiesto desde el punto de vista de la modelación y de sus resultados, que los protocolos superiores han seguido un curso de perfeccionamiento del protocolo Aloha, como modelo base, en los aspectos de detección de portadora y/o de colisiones, para el incremento de la performance de uso del canal.

Al efecto, se han seleccionado un grupo de protocolos de red de canal compartido, que en secuencia de complejidad creciente (Aloha puro, Aloha ranurado, CSMA persistente, CSMA no persistente, CSMA/CD, Ethernet y una aproximación para 802.11) tendrán un tratamiento común para la modelación (con Redes de Petri temporizadas y no temporizadas), simulación y evaluación de prestaciones.

Finalmente, dada la unicidad de criterio mencionada previamente, y la disponibilidad de herramientas software para la simulación, es posible desarrollar para la secuencia de protocolos propuestos, una unidad de enseñanza temática para dichos contenidos, que resultará del propio trabajo, como se mencionó en los objetivos.

iv) METODOLOGIA

Se plantea un trabajo de investigación, con las siguientes actividades:

- Recopilación de información relativa a modelos de Redes de Petri con aplicación al modelado de protocolos de red,
- Compilación de herramientas software de simulación para la animación y evaluación de performance de los modelos de los protocolos,
- Definición de la secuencia de modelación a partir del modelo de Red de Petri para Aloha puro,
- Validación funcional de los modelos con las herramientas software de simulación.
- Comparación de los resultados (carga ofrecida, rendimiento, colisiones, etc.). obtenidos entre los modelos de Red de Petri para los protocolos seleccionados,

- Comparación de los resultados obtenidos (carga ofrecida y rendimiento) con los resultados de origen analítico,
- Desarrollo de una unidad de enseñanza-aprendizaje sobre protocolos de red de canal compartido, basada en Redes de Petri, y las herramientas disponibles.

v) RESULTADOS ESPERADOS

Entre los resultados alcanzables con el presente trabajo se consideran:

- Un enfoque alternativo para la enseñanza-aprendizaje de los protocolos de red de canal compartido,
- La disponibilidad de resultados complementarios de origen analítico y vía la modelación con Redes de Petri,
- La disponibilidad de herramientas software de simulación para la enseñanza con animación de los protocolos,
- La obtención de resultados prácticos para la determinación de la confiabilidad y verificación de los modelos,
- La generación de una unidad de enseñanza basada en Redes de Petri.

vi) VISION GENERAL DE LOS CAPITULOS

El trabajo de tesis está organizado en los siguientes capítulos:

- El Capítulo I es una introducción al estado actual de las Redes de Petri (Petri Nets). Se describe el formalismo desde los conceptos básicos hasta la presentación del formalismo conocido como FPQSPN (Finite Population Queueing Systems Petri Net), introducido por sus autores para la evaluación de rendimiento de protocolos.
- En el Capítulo II se establecen los conceptos de medidas de performance, y se presentan los modelos FPQSPN para los protocolos de comunicaciones seleccionados. Los modelos están ordenados en complejidad desde Aloha.
- El Capítulo III presenta los principios generales de simulación, y posteriormente, los aplicables a Redes de Petri. Además, se introduce un simulador llamado HPSim utilizable para los aspectos de animación. Se discute como sería una secuencia de la animación cuando los protocolos evolucionan entre sus estados.
- En el Capítulo IV se introducen los conceptos de performance del canal de comunicaciones. Además, se presenta el simulador llamado STPNPlay, utilizable para la evaluación de rendimiento de protocolos de comunicaciones, para los modelos FPQSPN. Se efectúa una comparación entre las curvas de rendimiento entre el método analítico y las obtenidas usando el simulador, y se valida la herramienta y los modelos utilizados.
- En el Capítulo V se exponen las conclusiones y aportes del presente trabajo.
- Finalmente, se incluyen Apéndices como complementación a los contenidos, un Anexo con la Bibliografía, un Anexo de Glosarios, y un Anexo sobre Herramientas Académicas de Enseñanza.

CAPITULO I

ASPECTOS TEÓRICOS SOBRE LAS REDES DE PETRI

I.1) INTRODUCCIÓN

Los Sistemas Dinámicos de Eventos Discretos (DEDS) están basados sobre los sistemas dinámicos, donde los estados y eventos discretos tienen un rol fundamental. Los DEDS se están volviendo de interés creciente en la medida que la importancia de la automatización y control crecen en la tecnología moderna. Los dominios de aplicación como sistemas de fabricación flexibles, sistemas de transporte, sistemas distribuidos, y redes de telecomunicaciones usan intensivamente los conceptos y técnicas de los DEDS. La optimización del diseño, la planificación (control de performance), el monitoreo, y la implementación (posiblemente tolerante a fallos) son los usos críticos que requieren el uso de la modelación, y de análisis cualitativo y cuantitativo. Los métodos formales para la especificación, verificación y evaluación de performance de los DEDS se están así volviendo incrementalmente importantes cuando crece el tamaño y la complejidad del sistema. Las Redes de Petri han sido utilizadas como la base para la introducción de las extensiones de tiempo que permiten una descripción exacta de aplicaciones complejas, que provee la posibilidad de explotar la estructura de los modelos para desarrollar un mejor entendimiento de la conducta de los sistemas dinámicos, para deducir técnicas de análisis efectivas.

I.2) REDES de COLAS (QUEUEING NETWORKS) [1][2][3]

Las redes de colas han sido ampliamente usadas en la evaluación de performance de los DEDS como los sistemas de computación, redes de comunicaciones y sistemas de fabricación. Las redes encoladas son frecuentemente vistas como técnicas de modelación y análisis. Esto es a causa que la modelación y los aspectos de análisis han sido usualmente tratados juntos. La parte de análisis ha resultado en el desarrollo de varios métodos analíticos diferentes, que son de amplio interés y que también han conducido el camino para el desarrollo de métodos similares en el contexto de las Redes de Petri.

La discusión general de las redes considera los conceptos de colas simples, colas multiclase, redes de colas, redes de colas multiclase, redes de colas con capacidad finita y ruteo de estado dependiente.

I.2.1 Redes de colas básicas

Las colas simples consisten de una cola, también llamada buffer, y uno o más recursos llamados servidores (figura n° 1). Las entidades llamadas clientes o trabajos generados por procesos de arribo externo se unen a la cola para recibir un servicio de uno de los servidores. A la finalización del servicio, este cliente deja el sistema. Para caracterizar completamente la conducta de este sistema, se pueden definir los siguientes conceptos: el número de servidores, la disciplina del servicio, la capacidad de la cola, los procesos de arribo y los procesos de servicio.

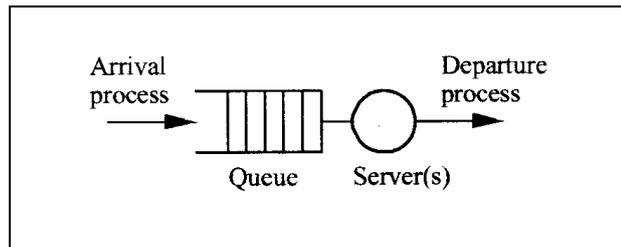


figura n° 1 Un simple sistema de cola

El número de servidores puede ser uno, un número finito mayor que uno, o infinito. La disciplina de servicio, o regla de planificación, define la política por la cual los servidores son asignados a los varios clientes presentes en el sistema. La capacidad de la cola es el número total de clientes que pueden estar simultáneamente presentes en el sistema (esperando o recibiendo servicio). Los procesos de arribo caracterizan la forma en que los clientes arriban al sistema. Los procesos de servicio caracterizan los requerimientos de los clientes en términos del servicio.

La importancia de la evaluación de performance es que proveen información cuantitativa sobre la conducta de un sistema de evento discreto. Los parámetros principales de performance de interés son: el tiempo de respuesta promedio, el tiempo de espera promedio, la longitud de cola promedio, la productividad, y la utilización de los servidores. Además, de los valores promedios, puede ser de interés conocer la varianza de las cantidades promedio, o más generalmente a cerca de sus distribuciones. Por ejemplo, puede ser importante conocer la varianza del tiempo de espera, además de conocer su valor promedio.

Un camino para determinar estas cantidades es registrar la conducta de un sistema durante un periodo largo de tiempo y derivar las cantidades de interés de esta observación. Una alternativa es construir un modelo estocástico de este sistema real. Un sistema puede ser modelado con una cola simple.

Es importante destacar que cuando se trabaja con los parámetros de performance se prioriza la representación de la conducta típica de un cliente y no la conducta específica de cualquiera en particular. Se llama caracterización de la carga de trabajo o tráfico, a las distribuciones de los tiempos de interarribo y de servicio, las cuales deben corresponder con la del sistema real.

En las representaciones de colas simples puede ser apropiado distinguir entre diferentes tipos de clientes. Esta es la noción de clases de clientes, o colas multiclase. En este caso, se pueden definir parámetros de performance total y parámetros de performance dependiente de la clase.

Finalmente, en algunos sistemas reales los clientes no arriban de a uno en el tiempo, sino en grupos, y/o los clientes no son atendidos de a uno en el tiempo, sino en grupos. Esta característica puede ser incorporada en los modelos de colas. Consideremos por ejemplo los procesos de arribo. La forma usual para caracterizar un proceso de arribo "batch" es definir los instantes de arribo de los batches y el tamaño de los batches.

1.2.2 Redes de colas

Una generalización natural de las colas simples es considerar las redes de colas. Una red de colas consiste de un conjunto de colas también llamadas estaciones. Cada estación consiste de una cola y un conjunto de servidores (simple servidor, múltiple servidor o infinitos servidores). La suposición usual es que todas las colas tienen capacidad infinita. La

caracterización de las disciplinas de servicios y los procesos de servicios en cada estación es la misma como en el caso de colas simples.

Hay dos tipos básicos de redes de colas llamadas redes de colas abiertas y redes de colas cerradas (open queueing networks and closed queueing networks) (figuras n° 2, 3 y 4).

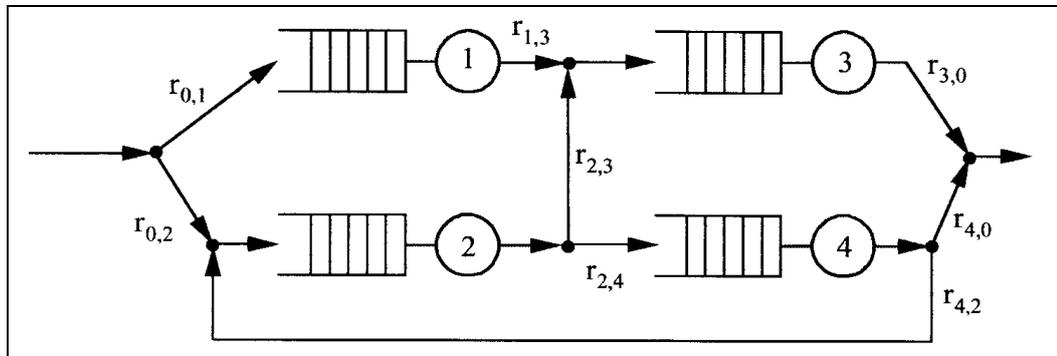


figura n° 2 Ilustración de una red de colas abierta

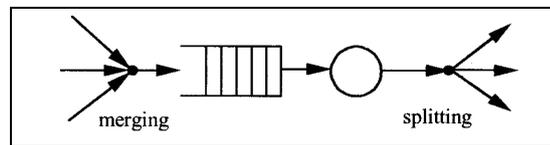


figura n° 3 Ilustración de los mecanismos de mezcla y división

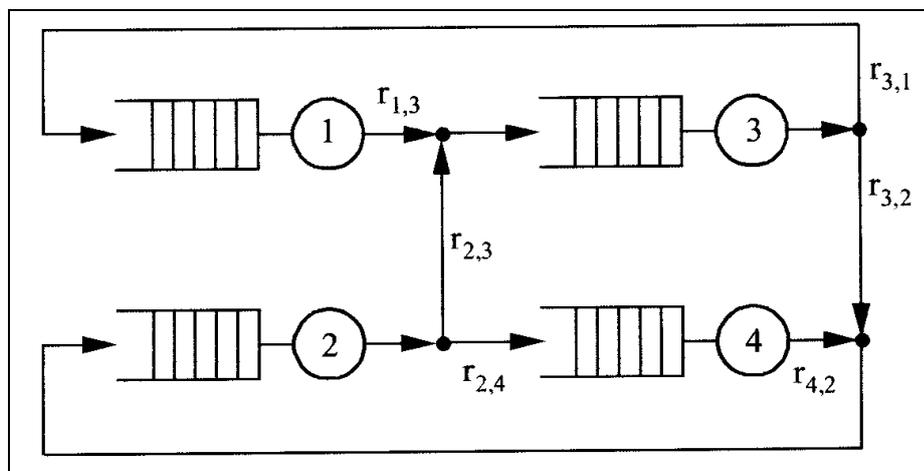


figura n° 4 Ilustración de una red de cola cerrada

En una red de colas abierta, los clientes arriban desde el exterior de acuerdo a un proceso de arribo dado y visitan un conjunto de estaciones después de lo cual ellos dejan el sistema.

En una red de colas hay dos clases de parámetros de performance que son de interés, llamados parámetros de performance local y global, que se refieren a las colas en particular o a toda la red, respectivamente.

En una red de cola cerrada hay un número fijo de clientes, llamados la población de la red, que viajan a través de la red, y por lo tanto, los clientes nunca entran ni dejan el sistema.

La caracterización de los procesos de arribo es la misma como en el caso de las colas simples. La característica adicional que necesitan ser especificadas es el ruteo de los clientes, es decir, como los clientes viajan a través de la red. El ruteo está normalmente caracterizado en una forma probabilística. El ruteo en las redes de colas da origen a dos mecanismos llamados la mezcla y la división de flujos de clientes, también llamados superposición y descomposición de flujos, respectivamente.

En una red de colas hay dos clases de parámetros de performance que son de interés, llamados parámetros de performance local y global, que se refieren a las colas en particular o a toda la red, respectivamente.

Como en el caso de las colas simples, puede haber situaciones para las cuales se requiere especificar la noción de clases de clientes. Esto puede deberse a la necesidad de distinguir los clientes con respecto a los procesos de arribo, los procesos de servicio y las disciplinas de servicio en las colas diferentes. Además, en el caso de red de colas, la distinción puede también ser con respecto al ruteo de los clientes, es decir, clases diferentes de clientes pueden tener diferentes ruteos probabilísticos a través de la red. Esto conduce a la definición de diferentes tipos de redes de colas multiclase: redes de colas multiclase abiertas, redes de colas multiclase cerradas y redes de colas multiclase mixtas. Este último caso sería la clase más general ya que incluye redes abiertas y cerradas como situaciones especiales (figura n° 5).

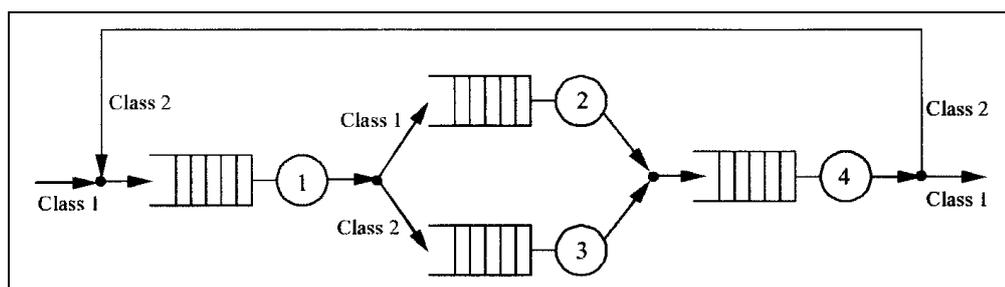


figura n° 5 Ilustración de una red de cola multiclase

En una red de colas multiclase, se puede combinar la noción de parámetros de performance total y clase-dependientes, además de los parámetros de performance local y global.

Finalmente, en algunas situaciones puede ser de interés permitir que los clientes cambien mientras viajan a través de la red, es decir, un cliente dejando una estación como un cliente de una clase conmuta a otra clase (con alguna probabilidad). Tales redes son llamadas redes de colas multiclase con conmutación de clase.

I.2.3 Redes de colas con capacidad infinita y ruteo dependiente del estado

Las redes de colas con colas de capacidad infinita no permiten modelar sistemas que tengan buffers finitos. En las redes de colas con capacidad finita debe especificarse el número máximo de clientes que pueden estar simultáneamente presentes en cada estación, incluyendo aquellos en servicio. Para tales redes, se necesita especificar la conducta de la red cuando una o más colas están llenas. Cuando una cola está llena durante algún periodo de tiempo, la conducta del sistema puede estar especificada en las siguientes formas diferentes:

mecanismo de bloqueo antes del servicio, mecanismo de bloqueo después del servicio, mecanismo de bloqueo y recirculación, y mecanismo de pérdida.

Estos diferentes mecanismos nos permiten modelar diferentes conductas encontradas en los sistemas reales. En particular, los sistemas de fabricación son frecuentemente modelados usando los mecanismos de servicio después del bloqueo y también de servicio antes del bloqueo. Mientras que las redes de comunicaciones son frecuentemente modeladas usando los mecanismos de bloqueo y recirculación, o el mecanismo de pérdida.

Los mecanismos de ruteo dependiente del estado pueden también ser introducidos en las redes de colas. Esto corresponde a la situación en la cual el ruteo de los clientes depende sobre el estado real de la red, por ejemplo, el número de clientes actualmente presentes en cada cola. Uno de tales mecanismos es llamado ruteo de cola más corto, que sucede cuando un cliente completa su servicio en una estación, y se une a la cola que tiene el número más pequeño de clientes, entre un número predeterminado de colas. El caso más general puede ser definido en términos de un mecanismo de ruteo dependiente del estado probabilístico.

I.2.4 Redes de colas extendidas (Extended Queueing Networks)

El formalismo de red de colas básico es útil para modelar algunas importantes clases de sistemas reales. Sin embargo, no nos permite representar sistemas reales que incluyan mecanismos de sincronización entre diferentes clientes, y mecanismos de ensamblado/desensamblado. Esto puede ser manipulado usando un formalismo más general llamado formalismo de red de colas extendido.

I.2.5 Notación de Kendall-Lee [4]

Según la notación propuesta por Kendall-Lee, cada sistema de colas se representa con seis características, de acuerdo al siguiente esquema: $1/2/3/4/5/6$, donde:

1. Indica la naturaleza del proceso de llegada (clientes), siendo M cuando el tiempo entre llegadas es independiente y distribuido idénticamente, y siguen una distribución exponencial, D cuando el tiempo entre llegadas es independiente y distribuido idénticamente, y determinista, y G cuando el tiempo entre llegadas es independiente y distribuido, y de distribución general.
2. Indica la naturaleza de los tiempos de servicio (servidores), y se representan como en 1.
3. Indica el número de servidores en paralelo.
4. Indica la disciplina de la cola, siendo FCFS (First Come First Served- Primero Llegar- Primero en ser Atendido), LCFS (Last Come First Served- Último Llegar Primero en ser Atendido), y SIRO (Service in random order-Servicio en Orden Aleatorio).
5. Indica el número máximo permitido de clientes en el sistema, incluyendo los que esperan y los que están siendo atendidos.
6. Indica el tamaño de la población de la que se toman los clientes.

I.3) REDES DE PETRI NO TEMPORIZADAS (UNTIMED PETRI NETS) [5][6][7]

Los DEDS típicos exhiben evoluciones paralelas, las cuales conducen a conductas complejas debido a la presencia de fenómenos de sincronización y recursos compartidos. Las Redes de Petri (RdP) son un formalismo matemático apropiados para modelación de DEDS concurrentes. Las RdP han sido aplicadas satisfactoriamente a campos tales como redes de

comunicación, sistemas de computación, sistemas de fabricación de partes discretas, etc. Los modelos de red son frecuentemente observados como especificaciones autodocumentadas, a causa de que su naturaleza gráfica facilita la comunicación entre diseñadores y usuarios. Los fundamentos matemáticos del formalismo permiten corrección lógica y análisis de eficiencia. Además, estos modelos pueden ser automáticamente implementados usando una variedad de técnicas de hardware y software, y pueden ser usados para propósitos de monitoreo, una vez que el sistema está trabajando. En otras palabras, las RdP pueden ser usadas a lo largo de todo el ciclo de vida de un sistema.

Más que un simple formalismo, las RdP son una familia de ellos, desde bajo a alto nivel, cada una de ellos más apropiados para diferentes propósitos. En cualquier caso, ellos pueden representar conductas muy complejas desde la simplicidad del modelo real, consistiendo de unos pocos objetos, relaciones y reglas. Más precisamente, un modelo de un sistema dinámico consiste de dos partes:

1. Una estructura de red, un grafo dirigido bipartito inscripto, que representa la parte estática del sistema. Las dos clases de nodos son llamadas lugares y transiciones, gráficamente representadas como círculos y cajas, respectivamente. Los lugares corresponden a las variables de estado del sistema y las transiciones a sus transformaciones. El hecho que ellos sean representados en el mismo nivel es una de las características destacadas de las RdP comparadas a otros formalismos. Las inscripciones pueden ser muy diferentes, conduciendo a varias familias de redes. Si las inscripciones son simplemente números naturales asociados con los arcos, llamados pesos, se obtienen las redes lugares/transiciones. En este caso, los pesos permiten la modelación de servicios y arribos masivos. Una subclase interesante históricamente y conceptualmente de redes lugares/transiciones se obtiene cuando todos los pesos son uno. Estas redes se dicen son ordinarias y ellas conducen, por ejemplo, a una directa pero importante generalización de modelos de autómatas. Inscripciones más elaboradas, asociadas con lugares, transiciones y arcos, conducen a los llamados formalismos de RdP de Alto Nivel.
2. Un marcado, gráficamente representado por marcas dentro de los lugares, que representan un estado total distribuido sobre la estructura. El marcado de un lugar (variable de estado) es su valor de estado. Un sistema de red es una estructura de red junto con un marcado inicial. La dinámica del sistema (la conducta del sistema) está dada por las reglas de evolución del marcado. La regla básica permite la ocurrencia de una transición cuando los valores de estado de entrada cumplen alguna condición expresada por las inscripciones de los arcos. La ocurrencia de una transición cambia los valores de sus variables de estado adyacentes, de acuerdo nuevamente a las inscripciones de los arcos.

La separación de este planteo permite razonar a los modelos basados en redes en dos diferentes niveles: estructural y conductional. Desde la primera forma podemos derivar algunas conclusiones rápidas como la conducta posible del sistema modelado. Los razonamientos puramente conductionales pueden ser más concluyentes, pero pueden requerir computaciones costosas, o tal vez no puedan ser factibles. El razonamiento estructural puede ser observado como una abstracción de la conducta. Por ejemplo, en lugar de estudiar si un sistema dado tienen un espacio de estado finito, podríamos investigar si el espacio de estados es finito para cada estado inicial posible, o podríamos estudiar si hay un estado inicial que garantiza actividad infinita mas que decidir esto por un estado inicial dado.

La interpretación de un modelo precisa la semántica de objetos y su conducta, eventualmente haciendo explícita la conexión de este modelo al mundo exterior dentro de un tipo de aplicaciones. Es decir, la interpretación considera el entorno en el cual el modelo será

ejercitado. Una interpretación puede dar un significado físico a las entidades de la red (lugares, transiciones, marcas), condiciones de evolución y, posiblemente, definirá las acciones generadas por las evoluciones. Los grafos interpretados, como las Redes de Petri, son comunes en la modelación de sistemas.

El formalismo de Red de Petri básico puede ser interpretado para describir diferentes perspectivas de un sistema dado a lo largo de su ciclo de vida. La variedad de interpretaciones es una de las mayores ventajas de modelar con redes. Los sistemas de redes provistas con interpretaciones apropiadas puede ser usadas a lo largo del diseño y operación, y los conceptos y resultados básicos son compartidos (pueden ser reusados) por las diferentes interpretaciones, conduciendo a alguna economía en el análisis y síntesis de los modelos. Una particularmente interesante familia de interpretaciones de red se obtiene cuando el tiempo y las probabilidades se asocian con el modelo. El formalismo básico es convencionalmente conocido como Red de Petri autónomo ó básico, cuando la dinámica del sistema de red está gobernada solo por las reglas de evolución de la red, y no por eventos externos, temporizaciones, etc. También el nombre Red de Petri no temporizada se usa cuando la clase de una de las interpretaciones ulteriores tiene en cuenta alguna cosa para hacer con el tiempo. La evolución autónoma de un modelo de red se dice ser completamente no determinista en el sentido que ni la resolución de conflictos ni el tiempo de ocurrencia de una transición habilitada están especificada del todo.

I.3.1 Redes y Sistemas Lugar/Transición (Place/Transition P/T)

Inicialmente corresponde introducir formalmente los sistemas de red lugar/transiciones, enfatizando la distinción entre la (estructura de la) red y el sistema (la red con un marcado y sus reglas de evolución), y como son capaces para modelar DEDES. Además, se debe describir en algún detalle las características de modelación de los sistemas de P/T, llamadas la localidad de estados y acciones, y la representación de los diversos fenómenos encontrados en sistemas paralelos y distribuidos. Estas características de modelación permiten representar en una forma natural situaciones típicas tales como las sincronizaciones, exclusión mutua, conflictos, etc, y aplicar las metodologías de síntesis top-down y bottom-up al diseño de los sistemas.

En los sistemas de Red de Petri, el estado está descrito en una conducta distribuida por medio de un conjunto de variables de estado. Los lugares (representados como círculos) son el soporte de estas variables de estado. Las acciones individuales, que transforman las variables de estado a la que están conectadas, son modeladas por transiciones (representadas como barras o cajas). En el formalismo P/T, los lugares y las transiciones están relacionados a través de una relación de flujo ponderada.

Una red P/T es una cuádrupla:

$$N = (P, T, F, W)$$

Donde:

1. P y T son conjuntos no vacíos finitos disjuntos, los lugares y las transiciones, respectivamente.
2. $F \subseteq (P \times T) \cup (T \times P)$ es la relación de flujo (conjunto de arcos dirigidos).
3. $W: F \rightarrow N^+$ asigna un peso o multiplicidad a cada arco.

El grafo de la figura n° 6 es una estructura de red. Los arcos están rotulados con números naturales, los pesos o multiplicidades de arcos. Por convención, los arcos no rotulados

están ponderados en uno. Todos los pesos de los arcos en la red ejemplo son uno, excepto para $W(p_6,t_6)$ y $W(t_6,p_1)$ que son dos. Un lugar p es un lugar de entrada (salida) de transiciones t si hay un arco yendo desde p a t (t a p). En la figura, $\{p_3,p_5\}$ son lugares de entrada de t_3 , mientras que $\{t_2,t_5\}$ son las transiciones de salida de p_2 . Las relaciones adyacentes entre nodos, las cuales determinarán luego la relación entre los estados y las acciones, son fijas para la estructura de la red, desde que ellas son estáticas.

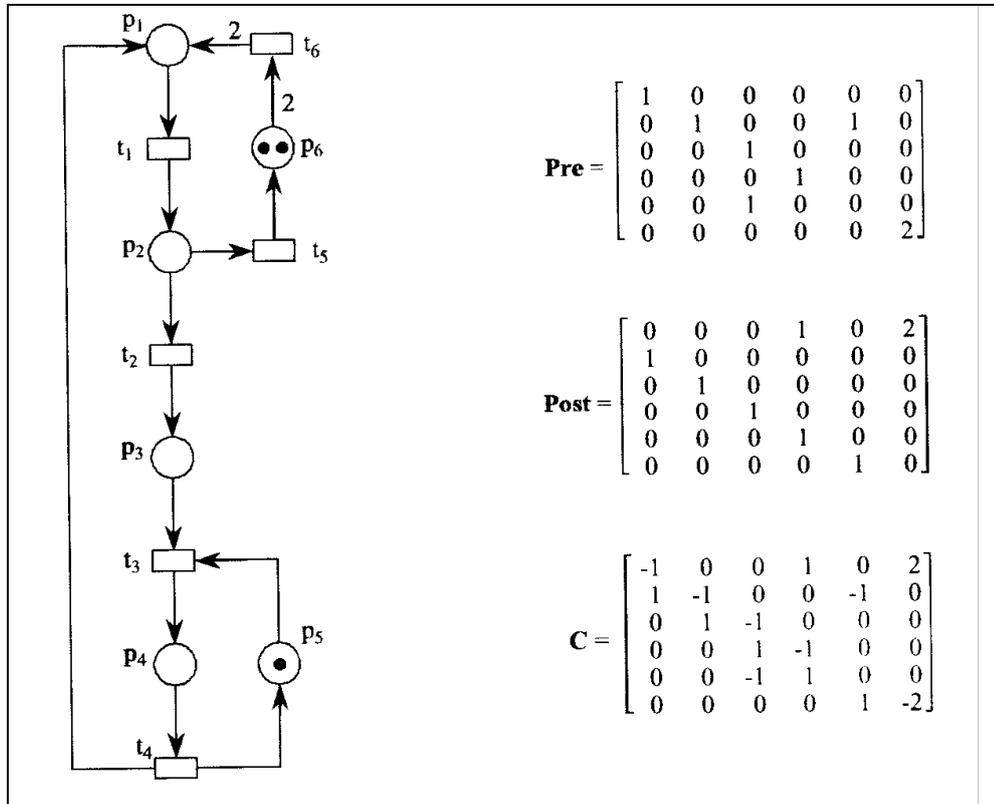


figura n° 6 Un sistema de red Lugar/Transición (P/T) y sus matrices incidentes

Hay una representación alternativa de las redes P/T donde la relación de flujo está descrita en forma matricial. Esto naturalmente conduce a una interesante ecuación de estado, como descripción de la evolución del sistema.

Una red P/T es una cuádrupla:

$$N = (P, T, \text{Pre}, \text{Post})$$

Donde:

1. P y T son conjuntos como antes.
2. $\text{Pre}, \text{Post} \in \mathbb{N}^{P \times T}$ son las matrices de pre-incidencia y post-incidencia.

Con la notación matricial, hay un arco yendo desde el lugar p a la transición t cuando $\text{Pre}[p,t] > 0$, y el valor de $\text{Pre}[p,t]$ es precisamente el peso del arco. Igualmente, para post-incidencia, $\text{Post}[p,t] = W(t,p)$ cuando es diferente de cero.

En el ejemplo de la figura anterior, t_4 es un fork (bifurcación), t_3 es un join (unión), p_2 es un distribuidor y p_1 es un colector. Un par formado de un lugar p y una transición

t es llamado un autoloop si p es entrada y salida de t. Las figuras n° 7a, 7b y 7c muestran ejemplos de autoloops, y la figura n° 7d presentación una modificación de la red de la figura n° 7c para la eliminación del autoloop. Una red se dice pura si no tiene loop. Las redes puras son completamente caracterizadas por una simple matriz: $C = \text{Post} - \text{Pre}$. Esta es llamada la matriz de incidencia de la red pura. Las entradas positivas (negativas) en C representan la función de incidencia post (pre). Si la red no es pura, C “no ve” los loops (así C no es una matriz de incidencia sino justamente una matriz de flujo de marcas).

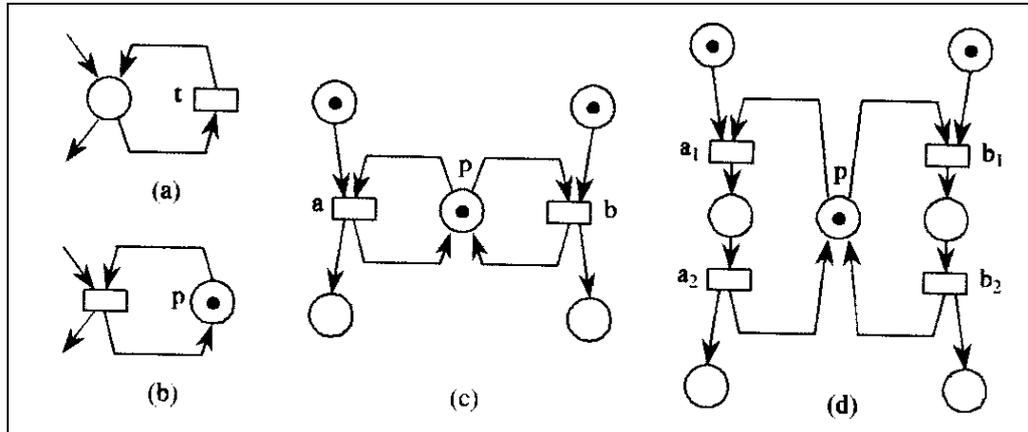


figura n° 7 Auto loops

I.3.2 Sistemas de Red: Marcado y Juego de marcas

La estructura de una red es algo estática. Asumiendo que la conducta de un sistema puede ser descrito en términos del estado y sus cambios, la dinámica sobre una estructura de red es creada al definir su estado inicial y las reglas de evolución del estado.

El marcado de una red N es un vector indexado de lugar m que asigna un entero no negativo (número de marcas) a cada lugar. Un sistema de red P/T tiene el par $S=(N,m_0)$, o más explícitamente, $S=(P,T,F,W,m_0)$ o $S=(P,T,Pre,Post,m_0)$ donde N es una red P/T y m_0 es su marcado inicial.

El número de marcas en un lugar representa el estado local del lugar, es decir, el valor de la variable de estado representada por ese lugar, que en el formalismo es un entero (de modo que puede ser interpretado como un contador o depósito). El estado del sistema de red total está definido por la colección de los estados locales de los lugares. Por lo tanto, el vector m es el vector estático del DEDS descrito por el sistema de red. Gráficamente, ponemos $m[p]$ puntos negros (tokens) en el círculo que representa el lugar p . El marcado del sistema de red en la figura es $m_0=[0,0,0,0,1,2]$. Por cuestiones de conveniencia, frecuentemente se usa una notación alternativa para marcado, por ejemplo, $m=p_5 + 2p_6$.

La evolución del estado distribuido está definida a través de una regla de disparo u ocurrencia, informalmente llamada como el “juego de marcas”. Esto es a causa que las estructuras de red pueden ser vistas como una clase de chequeadores, las marcas como marcadores, y las reglas de disparo como la regla del juego. Las transiciones representan movimientos potenciales en el juego de marcas. Debe observarse que las marcas no se mueven, sino destruidos y creados en cada movimiento, posiblemente en un numero diferente.

El marcado en un sistema de red evoluciona como sigue:

1. Una transición se dice estar habilitada en un marcado dado cuando cada lugar de entrada tiene al menos tantas marcas como el peso de los arcos uniéndolos. Formalmente, t está

habilitada en m si $m \geq \text{Pre}[P,t]$. El número de habilitaciones simultáneas de una transición t en un marcado dado m se llama su grado de habilitación, y es denotado por $e(m)[t]$.

2. La ocurrencia, o disparo, de una transición habilitada es una operación atómica que remueve de (suma a) cada lugar de entrada (salida) un número de marcas igual al peso de los arcos uniendo el lugar (transición) a la transición (lugar). Formalmente, la ocurrencia de t en el marcado m , denotado por $m \xrightarrow{t} mt$, viendo el marcado $mt = m + C[P,t]$.

La precondition de una transición puede ser vista como los recursos requeridos para que la transición se dispare. La poscondición representa los recursos producidos por el disparo de la transición. En un modelo teórico o abstracto, las transiciones habilitadas pueden no estar forzadas a dispararse. Esta es una forma de no determinismo. En la modelación práctica, la interpretación gobierna parcialmente el disparo de transiciones habilitadas (dependiendo de si o no ocurre un evento externo asociado a una transición habilitada). Además, debe observarse que no está preciso si la ocurrencia de una transición toma algún tiempo, dado que el tiempo no ha sido introducido aún.

I.3.3 Semántica de Intercalación: Observaciones secuenciales

Una forma común de describir la conducta de un sistema P/T es por medio de sus observaciones secuenciales. Así, decimos que el observador se supone ve solo eventos simples, es decir, una transición ocurriendo en un momento. La semántica de intercalación de un sistema de red está dada por todas las posibles secuencias de las ocurrencias de transiciones individuales que podrían ser observadas desde el marcado inicial. Si dos transiciones a y b están habilitadas simultáneamente y la ocurrencia de una no hace deshabilitar la otra, en principio ellas podrían ocurrir en el mismo momento, pero el observador secuencial verá a seguido por b , o viceversa. El nombre semántica de intercalación viene de esta forma de ver ocurrencias simultáneas.

Consideremos a S como un sistema P/T

1. Una secuencia de ocurrencia o disparo de m es una secuencia $t_1, \dots, t_k, \dots, t_n$, tal que $m \xrightarrow{t_1} m_1 \xrightarrow{t_2} \dots \xrightarrow{t_k} m_k \dots$. Si la secuencia de disparo llega al marcado m' , este es indicado como $m \xrightarrow{t} m'$.
2. El lenguaje de S , denotado como $L(S)$, es el conjunto de todas las secuencias de ocurrencias desde m_0 .
3. La conjunto de alcanzabilidad de S , denotado como $RS(S)$, es el conjunto de todos los marcados alcanzables desde m_0 al disparar alguna secuencia en $L(S)$.
4. El grafo de alcanzabilidad de S , denotado por $RG(S)$ es un grafo rotulado donde los vértices son los marcados alcanzables, y hay un arco rotulado desde el vértice m a el vértice m' si $m \xrightarrow{t} m'$.

Integrando la ecuación de evolución desde m_0 a un m arbitrario obtenemos:

$$m_0 \xrightarrow{t} m \xrightarrow{t} m = m_0 + C.$$

Esta es conocida como la ecuación de estado del sistema, sobre la cual se basan algunos métodos de análisis estructural. Las Redes de Petri, como introducidas hasta ahora, son una formalismo matemático, en el mismo sentido como son las ecuaciones diferenciales. Mientras estas últimas son útiles para describir sistemas dinámicos continuos, la primera forma son introducidos para la descripción de sistemas dinámicos de eventos discretos (DEDS).

I.3.4 Localidad y la Síntesis del Modelo

Está claro de las definiciones de la estructura de red, la representación de estado y las reglas de evolución, que existe un principio de localidad en los estados y las acciones. La localidad de la representación del estado y de las acciones es otra característica crucial de las Redes de Petri, la cual le permite la modelación de conductas distribuidas complejas. Su importancia para la síntesis de modelos reside en el hecho que las redes pueden ser localmente modificadas, o bien, refinadas, sin ninguna alteración del resto del modelo. También es posible que diferentes redes, modelando diferentes partes de un sistema, puedan estar compuestas por algunos nodos compartidos (representando actividades y estados comunes). En otras palabras, las redes pueden estar sintetizadas usando las aproximaciones top-down y bottom-up (figura n° 8). La síntesis top-down es cualquier procedimiento que comienza con un modelo abstracto inicial, y conduce al modelo final a través de pasos de refinamiento. En una aproximación bottom-up, los módulos son producidos, posiblemente en paralelo por diferentes grupos de diseñadores, luego compuestos.

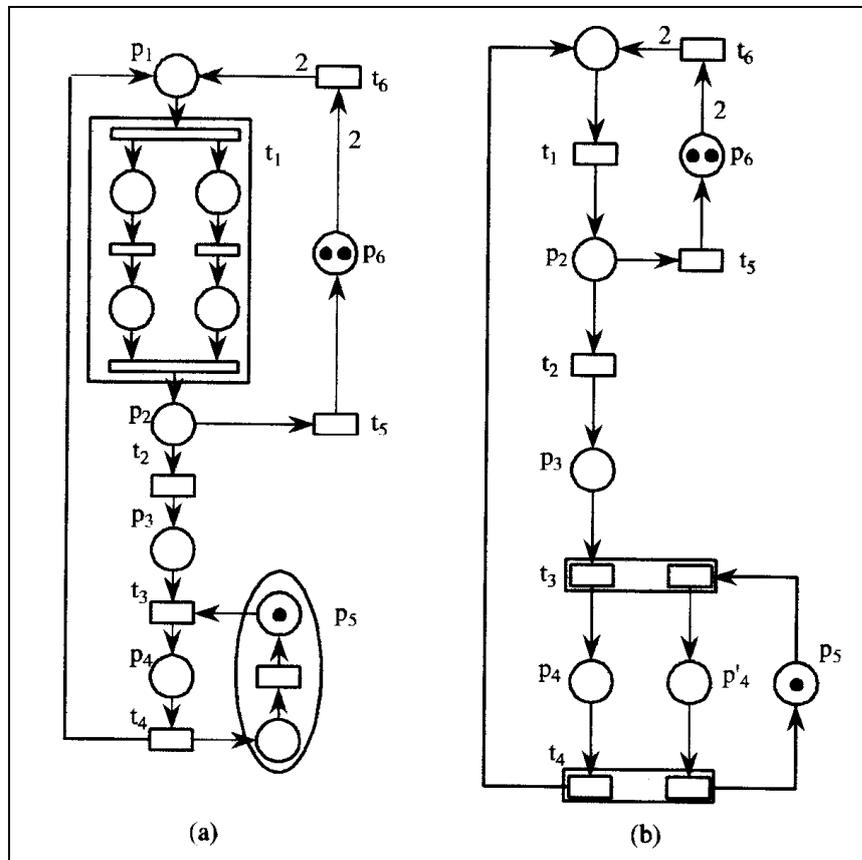


figura n° 8 Sintetización del modelo del la multicomputadora PLC de la figura n° 6
 (a) un modelo refinado (b) Como una sincronización de dos submodelos

I.3.5 Dependencia causal, conflictos y concurrencia.

Es posible adecuar el formalismo para modelar DEDS con actividades de concurrencia y paralelas. Informalmente, los tres fenómenos básicos a ser representados son la dependencia causal, es decir, algunas acciones requieren que otras sean realizadas primero; los conflictos, es decir, algunas acciones son alternativas, una decisión sobre cual ocurrirá debe tomarse, y concurrencia, es decir, hay acciones que deben ocurrir simultáneamente. Estos fenómenos son realmente representados por modelos de Redes de Petri en una forma muy natural (figura n° 9 y 10). Ellos son la base para la modelación de esquemas típicos en sistemas paralelos y distribuidos, como la exclusión mutua, ensamblado, aleatoriedades, etc. La dependencia causal y conflictos son nociones clásicas en los sistemas secuenciales, como los autómatas finitos.

- a) Concurrencia. Dos transiciones son concurrentes en un marcado dado si ellas pueden ocurrir simultáneamente. En el caso de la semántica de intercalación, la concurrencia de dos o más acciones a y b, se representa por la posibilidad de realizarse en cualquier orden (primero a y luego b, o viceversa). No obstante, la presencia de todas las secuencializaciones posibles de las acciones, no implica que ellas estén verdaderamente en concurrencia. Dos acciones pueden ocurrir en cualquier orden, pero ellas tal vez no puedan ocurrir simultáneamente. Es importante distinguir entre concurrencia de transiciones y la concurrencia de actividades de sistema, lo cual es nuevamente una manera de interpretación del modelo. En el caso de que las transiciones implementen las actividades del sistema, ambas cosas coinciden: la ocurrencia de transiciones tomarían algún tiempo, de modo que sus ocurrencias pueden solaparse en el tiempo. En este sentido, ellas podrían ocurrir simultáneamente.
- b) Dependencia causal. Las dependencias causales están representadas por el ordenamiento de las acciones inducidas por la relación de flujo. Las dependencias aparecen en la forma de secuencia, como en los sistemas secuenciales, pero también en la forma de sincronizaciones. La construcción de red muy básica usada para modelar dependencias causales es un lugar conectando dos transiciones. Las transiciones conectadas a través de un lugar se dice están en relación causal estructural. La estructura de red muy básica usada para modelar las sincronizaciones es una transición con más que un lugar de entrada, es decir, una transición join. Puede también decirse que un arco múltiple desde un lugar a una transición es una clase de sincronización, a causa que algunas tokens individuales deben ensamblarse en frente de la transición a habilitar.
- c) Conflictos: Los conflictos en los sistemas secuenciales son claramente la situación en la cual dos acciones están habilitadas de modo que una debe ser elegida para que ocurra. Las cosas se vuelven más complicadas en el caso de sistemas concurrentes, donde el hecho que dos transiciones estén habilitadas no necesariamente implica que deba elegirse una. Algunas veces, la definición secuencial es apropiada, pero en otros no. La construcción de red muy básica usada para modelar conflictos es un lugar con más que una transición de salida, es decir, un lugar distribuidor. De hecho, los lugares distribuidos son necesarios para modelar conflictos, pero lo inverso no es verdad. Las transiciones de salida de un lugar distribuidor se dicen estar en relación de conflicto estructural. Muy frecuentemente en la literatura, nuestros conflictos estructurales son simplemente conflictos, pero se prefiere sumar el adjetivo

estructural para mejorar la distinción de la conducta, de la dinámica, noción de conflicto efectivo, que depende del marcado. Se quiere hacer notar, que un conflicto estructural hace posible la existencia de un conflicto efectivo, pero no lo garantiza.

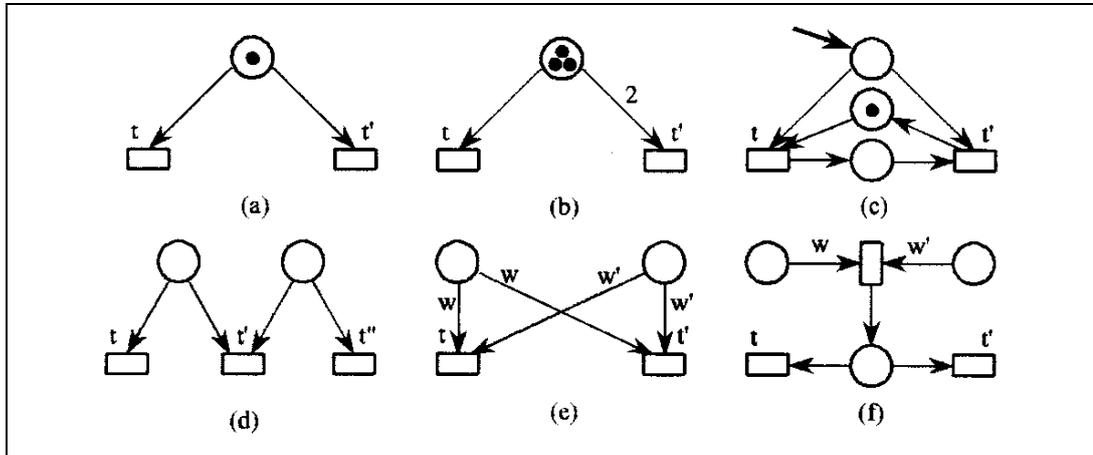


figura n° 9 Conflictos y conflictos estructurales

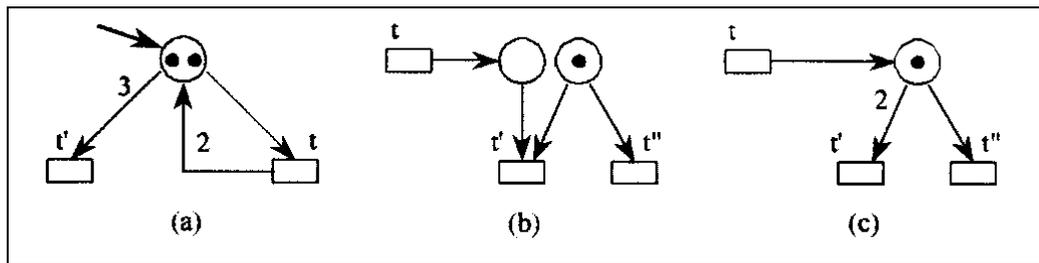


figura n° 10 Conflictos no igual y confusión

I.3.6 Subclases y Extensiones del Formalismo Básico

En este punto, es importante hacer una distinción entre potencia y conveniencia de modelación o expresión, en orden para apreciar mejor los méritos relativos de los diferentes formalismos, restricciones y extensiones. Diremos que un formalismo tiene mayor potencia de modelación teórica que otra cuando la primera puede modelar sistemas que no puede el último. Por ejemplo, la Red de Petri tiene mayor potencia de modelación que el autómatas finito dado que ellos pueden modelar sistemas con un espacio de estado finito. Cuando los sistemas que dos formalismos pueden modelar son los mismos, es decir, cuando ellos tienen la misma potencia de modelación, la distinción de estos formalismos es una materia de la conveniencia. Además, al tamaño y claridad de un modelo en cada formalismo, otros aspectos son importantes en juzgar la conveniencia relativa de dos formalismos; es decir, la habilidad para construir modelos por refinamiento y composición, la habilidad para representar adecuadamente datos y control, la analizabilidad, o el grado de parametrización. El orden para un formalismo, como más o menos conveniente que otro, no es siempre fácil o aún posible, dado que la conveniencia comprende aspectos que son típicamente contradictorios, tales como la compactación y la analizabilidad. En algunos casos, existen las transformaciones completas (preferiblemente sintácticas) para

trasladar modelos desde un formalismo a otro, lo cual puede ser usado para aplicar técnicas de análisis desarrolladas por un formalismo a un modelo producido en un formalismo abreviado, lo cual es entonces visto como más conveniente. Aunque tales transformaciones son impracticables, y debe ser altamente deseable desarrollar la analizabilidad directa de formalismos abreviados, así incrementando la conveniencia total.

I.3.7 Subclases sintácticas

Las subclases de los sistemas de red pueden ser definidos como restringiendo la conducta (comportamiento), o la estructura (topología), o sintaxis que describe matemáticamente el modelo. Una posible forma de obtener subclases sintácticas es restringir las inscripciones (por ejemplo, redes con cada peso igual a uno son ordinarias) o la topología, usualmente ayuda en limitar el juego entre los conflictos y las sincronizaciones. Lo último puede ser obtenido por dar una restricción general típicamente sobre lugares distribuidor y/o transiciones join, o por dar reglas para construir modelos.

I.3.7.1 Redes Ordinarias y Sistemas de Red Elementales

Una red de P/T es ordinaria cuando cada arco pesa igual a uno. Por lo tanto, la ocurrencia de una transición consume un token desde cada lugar de entrada, y produce un token para cada lugar de salida. En términos de modelación, si la transición realiza algún servicio y los tokens representan clientes encolados en lugares, sirviendo simultáneamente un grupo de clientes, o al arribo simultáneo de un grupo de clientes a una cola, no puede ser modelada, al menos directamente. De hecho, se puede implementar una red P/T con pesos por una ordinaria usando transformaciones (figura n° 11).

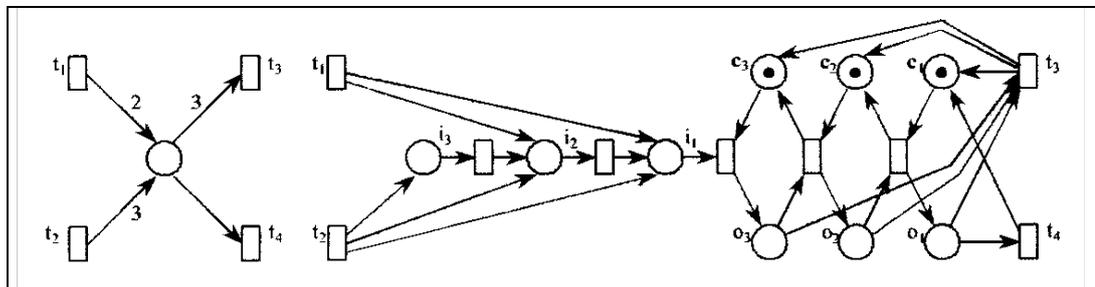


figura n° 11 Implementación ordinaria de una red ponderada distinto a 1

En el caso particular que el límite de cada lugar es uno (es decir, el sistemas en 1-limitado, o seguro), los lugares pueden ser interpretados como condiciones booleanas, las cuales se mantienen o no dependiendo sobre si ellos están marcados o no. Una transición es un evento que pueden ocurrir cuando se mantiene pre-condiciones. Después de su ocurrencia, las precondiciones se vuelven falsas y las poscondiciones se hacen verdaderas. Si queremos adherir a esta interpretación, podemos modificar la regla de ocurrencia, de modo que no necesitamos preocuparnos acerca de la 1-limitación. Una transición t está habilitada cuando sus precondiciones se mantienen y sus poscondiciones no. Esto es usualmente conocida como sistemas de red elemental (EN) (figura n° 12). La idea de tomar en cuenta el marcado de los lugares de salida cuando definiendo la habilitación de una transición puede ser extendida a redes P/T generales por asignar una capacidad a cada lugar. No obstante, las capacidades son una conveniencia de modelación menor dado que ellas pueden ser simuladas usando lugares

complementarios (es decir, el inverso de un lugar dado p marcado con la capacidad de p menos su marcado) preservando la semántica de intercalación y concurrencia.

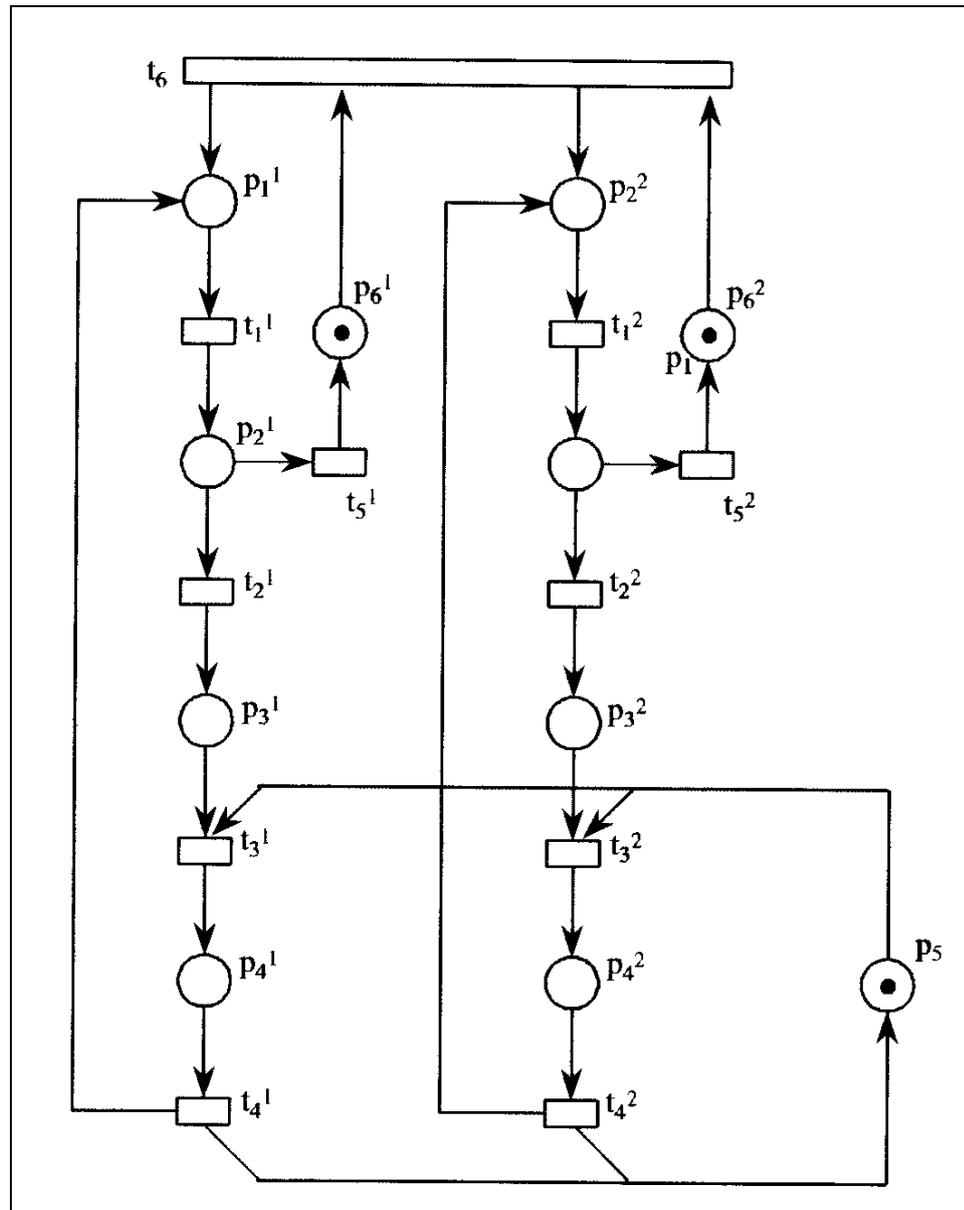


figura n° 12 Un modelo EN de la multicomputadora PLC de la figura n° 6

I.3.7.2 Subclases Topológicas

Históricamente las subclases de redes ordinarias han recibido especial atención a causa de los potentes resultados obtenidos tempranamente con su uso. Algunas de ellas aparecen como subclases de sus generalizaciones ponderadas. Si observamos la potencia de modelación, claramente algunas subclases tienen menos que otras si las primeras son apropiadamente incluidas en las últimas. También las generalizaciones ponderadas tienen más potencia de modelación que sus contrapartes ordinarias dado, que en general, las implementaciones ordinarias de pesos no preservan la pertenencia de clase topológica.

Dentro de las subclases de redes ordinarias se enumeran: Redes Join-Free y Máquinas de estado; Redes Distributor-Free y Grafos marcados, Redes Equal Conflict y Elección Libre, Redes Elección Asimétrica o Simple, Redes Subclases Modular.

I.3.8 Arcos Inhibidores y Prioridades

De acuerdo a sus definición, los sistemas de red P/T no permiten modelar tests zero, es decir, transiciones que están habilitadas sólo si algún lugar está vacío. Los arcos inhibidores y prioridades incrementan claramente la potencia de modelación teórica, aún cuando resulten en modelos menos analizables.

Una red P/T con arcos inhibidores y prioridades es una sextupla:

$$N = (P, T, Pre, Post, Inh, pri)$$

Donde:

1. $(P, T, Pre, Post)$ es una red P/T.
2. $Inh \ N^{IP \times IT}$ es la matriz de incidencia de inhibición.
3. $pri \ N^{IT}$ es el vector de prioridad.

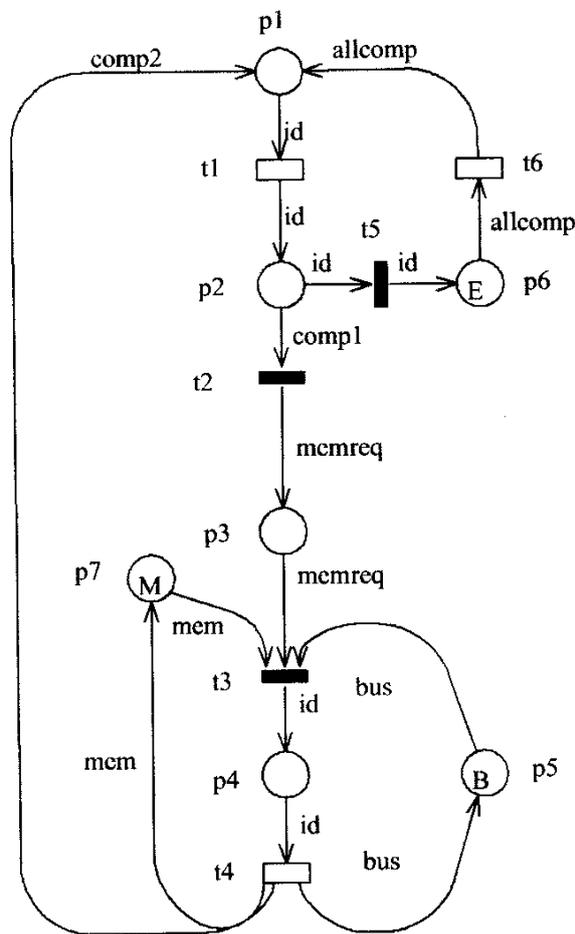
Las redes donde $Inh = 0$ o $pri = k.1$ no tienen arcos ni prioridades, respectivamente, y la matriz o vector correspondiente no son explícitos. Los arcos inhibidores y las prioridades modifican la condición de habilitación de la regla de ocurrencia: Una transición t está habilitada en m si $m \geq Pre[P, t]$, $m \geq Inh[P, t]$, y no hay t' tal que $pri[t'] > pri[t]$ la que está habilitada.

Los arcos inhibidores son graficados como arcos dirigidos encabezados por un círculo desde los lugares a las transiciones, rotulados con el correspondiente peso de arco (a menos que sea uno). Las transiciones de prioridad más baja (convencionalmente prioridad cero) son graficados como cajas, mientras que el resto son graficadas como barras, rotuladas con el nivel de prioridad cuando es mayor que uno.

I.3.9 Redes y Sistemas Coloreados [8]

Las Redes y Sistemas Coloreados son una extensión de las Redes de Petri. Por medio de la introducción de inscripciones elaboradas asociadas con lugares, transiciones y arcos, permite escribir modelos compactos de sistemas complejos. Entre los varios formalismos de Redes de Petri de Alto Nivel (HLPN), restringiremos a las Redes de Petri Coloreadas (CPN), y en particular a la subclase de Redes Bien-Formadas (WN).

La nueva característica introducida por las CPNs es la posibilidad de tener marcas distinguidas (este justifica el adjetivo coloreada: la marca puede estar representada gráficamente como puntos de diferentes colores en lugar de ser todos puntos negros; realmente el color de la marca puede ser cualquier clase de información). Cada marca en las CPNs puede tener información cuyo tipo depende sobre el lugar donde está ubicada, dado que la definición de un CPN debe incluir la especificación de un dominio de color para cada lugar, que es el tipo de datos asociados a las marcas en el lugar.



$$M = m1 + m2 + m3$$

$$B = b1 + \dots + b_{nb}$$

$$E = e1 + e2 + e3$$

figura n° 13 Modelo CPN de un PLC multiprocesador

Si el número posible de colores es finito, las CPNs tienen la misma potencia de modelación teórica como los sistemas de red P/T, desde una transformación algorítmica permite obtener un modelo equivalente (por lejos más grande) de cualquier modelo CPN dado.

El estado, o marcado, en las CPNs está representada por el multiconjunto de marcas coloreadas con cada lugar. Como en los sistemas P/T, el cambio de estado está realizado por disparo de transiciones. Dado que las marcas en las CPNs están distinguidas, alguna información adicional se necesita para definir las marcas coloreadas que dibujadas desde los lugares de entrada y puestas en los lugares de salida de una transición dada cuando la dispara. Esta información, asociada con cada arco, define un multiconjunto sobre el dominio color del lugar conectado al arco (figura n° 13).

I.4) REDES DE PETRI TEMPORIZADAS (TIMED PETRI NETS) [9][10][11][12][13]

Las Redes de Petri fueron originalmente propuestas como un modelo causal, que explícitamente despreciaba el tiempo. La introducción del tiempo en el modelo de red básico implica dificultades en la definición de un modelo coherente. En particular, la especificación de la conducta en el tiempo para evaluar la performance requiere un gran tratamiento de interpretación a ser sumado sobre el modelo de sistema básico. Aproximamos el problema en dos pasos: primero definimos la idea de observación de la conducta de un modelo autónomo en el tiempo; segundo definimos un modelo no autónomo en la cual cada actividad modelada por los retardos de disparo de transición es especificada en una manera apropiada.

I.4.1 Temporización de transición

Uno de los objetivos principales para un modelo de Red de Petri temporizada de un sistema distribuido es describir las actividades del sistema para entender sus interacciones mutuas y sus temporizaciones.

Tomando conceptos de los modelos de colas, podemos identificar los servidores que transportan tales actividades en una conducta asíncrona e independiente (excepto para interacciones representadas explícitamente en el nivel de Red de Petri). Ampliamente hablando, una actividad puede ser definida como una cantidad de trabajo realizada por un servidor para producir algunos resultados comenzando desde una precondition dada. La terminación de tal actividad toma tiempo, y la cantidad de tiempo necesaria depende de la velocidad del servidor, como así también de la cantidad de trabajo.

La finalización de una actividad puede ser vista como un cambio en el estado del sistema: “los resultados producidos por la actividad se vuelve disponible”. Dado que los cambio de estado en la Redes de Petri están determinadas por disparos de transición, es natural e intuitivo relacionar la finalización de las actividades al disparo de las transiciones. Sobre el otro extremo, las transiciones pueden dispararse solo después de sus habilitaciones. Un mapeo natural de las actividades sobre los modelos de Red de Petri consiste en identificar servidores con transiciones habilitadas: la habilitación de una transición corresponde a poner un servidor a trabajar en una actividad, y el disparo de la transición corresponde a la finalización de la actividad, con la producción de resultados.

Más que un servidor pueden estar iniciando actividades de un tipo dado. Por ejemplo, en un mercado genérico m , el grado de habilitación $ei(m)$ dice el número de servidores que pueden trabajar concurrentemente para llevar adelante las actividades del tipo representado por la transición t_i T (es decir, consideramos una noción de auto-concurrencia). El tiempo necesario para completar la actividad es modelado por el intervalo de tiempo entre la habilitación y el disparo de una instancia de transición. Con este mapeo en mente, podemos fácilmente mantener la idea de disparo atómico para las transiciones (que es una suposición crucial para los modelos de Red de Petri no temporizadas), dado que el disparo en si mismo puede ser pensado como un evento instantáneo que concluye una actividad temporizada. Usando la terminología de las Redes de Colas, las transiciones pueden así ser interpretadas como estaciones que pueden tener hasta ebi servidores, donde ebi es llamado el límite de habilitación de la transición t_i .

Aparece un problema cuando consideramos el caso de las transiciones conflictivas. Por ejemplos, si se analiza un conflicto de free-choice entre dos transiciones. Cuando una marca es depositada en el lugar de entrada, ambas transiciones están habilitadas. Con nuestra interpretación, dos servidores son iniciados (simultáneamente) para trabajar sobre las dos actividades conflictivas modeladas por las transiciones. Sin embargo, solo una de ellas puede llegar a su finalización (modelada por el correspondiente disparo de transición). La otra actividad debe malograrse por el conflicto. ¿Qué hacemos con la porción de trabajo realizado por el servidor antes de malograrse?. Algunas políticas de memoria y muestreo pueden ser definidas para especificar la conducta de un modelo de Red de Petri conteniendo transiciones temporizadas conflictivas.

I.4.2 Reducción del No Determinismo

El punto inicial para la definición de las Redes de Petri Temporizadas es un sistema de red no determinístico, en el cual algunos detalles relacionados a la conducta son dejados sin definición. Una interpretación temporizada del sistema de red consiste entonces en especificar la conducta “en el tiempo” en tal forma que:

1. Es compatible con el modelo no temporizado y no determinístico,
2. La “cantidad de no determinismo” es reducido al tener en cuenta las condiciones de temporización.
3. La conducta del sistema es especificada en forma suficientemente precisa para derivar medidas de performance (el grado de precisión requerido depende de las técnicas de análisis que son usadas).

El problema de reducir el no determinismo puede ser atacado en algunas formas diferentes. Existen tres técnicas diferentes que pueden ser combinadas unas con otras, y cuya combinación cubre todas las clases de “redes de Petri temporizadas” propuestas en las diversas bibliografías sobre el particular hasta ahora:

N) reduce la cantidad de no determinismo, aún manteniendo el modelo no determinístico (pero a una extensión menor);

S) introduce una medida estocástica para el no determinismo del modelo subyacente;

D) define algoritmos determinísticos para sustituir especificaciones no determinísticas.

El no determinismo puede asumir dos formas (levemente) diferentes en una sistema de red temporizado: la elección entre transiciones conflictivas y la “duración” de la habilitación para las transiciones (independientemente del hecho que ellas pueden ser persistentes o en conflicto con otras transiciones). Ambos tipos de no determinismo pueden ser resueltos por cualquiera de las tres técnicas básicas mencionadas (o combinaciones de ellas).

I.4.2.1 Reducción no determinística

El ejemplo clásico de esta aproximación es la definición de los así llamados sistemas de red de “intervalo temporizado”, originalmente introducido en el contexto de verificación de protocolos de comunicaciones involucrando mecanismos de time-out. Un tiempo de disparo mínimo y uno máximo son asociados con cada transición t como parte de la interpretación de la red ($t_m(t)$ y $t_M(t)$ respectivamente). La regla de disparo es redefinida en el tiempo como sigue: cualquier transición t debe permanecer habilitada al menos $t_m(t)$ unidades de tiempo antes de ser capaz para dispararse; una transición que permanece habilitada por $t_M(t)$ unidades de tiempo es forzada a dispararse a la finalización del periodo máximo. Por supuesto, la inequidad $t_m(t) < t_M(t)$ debe siempre ser satisfecha por la definición de tiempo para tener sentido. Nada se dice con respecto al tiempo de disparo real en el caso $t_m(t) < t_M(t)$: puede asumir cualquier valor dentro del intervalo en un forma totalmente no determinística. Por ello, en el caso que es no persistente, t puede ser deshabilitada ante de disparar y antes de alcanzar su máximo tiempo de habilitación $t_M(t)$.

Además, los sistemas de red no interpretados pueden ser redefinidos como redes de intervalo temporizado con $t_m(t)=0$ y $t_M(t)=\infty$ para todo t , sin cambiar su conducta no determinística. Por el contrario, la conducta de un modelo temporizado con $t_m(t)>0$ y/o

$t_m(t) < \infty$ para algún t , puede ser diferente de uno de un sistema de red no interpretado, siempre desde un punto de vista puramente cualitativo.

Claramente, este tipo de reducción no determinística sola, no hace suficientemente precisa a la conducta especificada del modelo para computar la performance “exacta” (a menos que $t_m(t) = tM(t)$ para todas las transiciones). A lo sumo los límites (superior e inferior) pueden ser computados, debido a la falta de conocimiento sobre las instancias de ocurrencias reales para el disparo de transiciones. De hecho, esta clase de modelos temporizados han sido usados esencialmente para exámenes de las propiedades cualitativas, como una alternativa a los sistemas de red no interpretados para tener en cuenta las condiciones de temporización.

I.4.2.2 Reducción estocástica

Esta es la forma principalmente usada para la reducción del no determinismo en el caso que uno quiera usar el modelo interpretado para la evaluación de performance. La idea es preservar la “variabilidad posible” que es inherente a el sistema de red subyacente no determinístico, y al mismo tiempo hacer la descripción de la conducta suficientemente precisa para ser capaz de aplicar las técnicas de análisis estadísticas para computar la performance. En este caso, la reducción es completa, en el sentido que se da una descripción matemáticamente precisa, y nada se deja indefinido.

Esta clase de reducción es obtenida al sumar una “medida probabilística” al espacio de variabilidad del modelo no determinístico. Una función de densidad de probabilidad (pdf) con soporte infinito $[0, \infty)$ se asocia con cada instancia de retardo de disparo de transición, y una probabilidad discreta es asignada a cada transición conflictiva para la elección dentro de cada conjunto conflictivo efectivo. El valor real de cada instancia de retardo de disparo se vuelve así una variable aleatoria, que en principio puede asumir cualquier valor en el rango de $[0, \infty)$, como en sistema de red subyacente. Sin embargo, diferentemente del sistema de red no interpretado, ahora podemos aplicar análisis estadístico para estimar probabilidades de ocurrencia de algunos eventos. Se nota que de hecho en el modelo no interpretado no tiene sentido introducir el concepto de probabilidad y/o frecuencias esperadas de eventos debido a su naturaleza puramente no determinística que hace asumir cualquier clase de conducta típica o esperada para los tiempos de disparo o superar las elecciones de conflictos.

Desde un punto de vista cualitativo, la interpretación estocástica no hace cambiar la conducta del modelo con respecto al sistema de red no interpretado. La suma de una medida de probabilidad puede solo complementar tal conducta con información adicional que permite la aplicación de la idea de conducta esperada de la observación del modelo. Si pensamos por ejemplo en la relación de alcanzabilidad, el análisis de alcanzabilidad del sistema de red no interpretado podemos decir si un mercado dado puede ser alcanzado desde el mercado inicial o no. La propiedad de alcanzabilidad no es afectada por la interpretación estocástica. Sin embargo, la suma de una interpretación estocástica le permite a uno hacer preguntas adicionales: “¿Como comúnmente el evento a alcanzar es alcanzado después de t unidades de tiempo desde el comienzo de la observación, cuando el modelo pueda asegurar que estaba en su mercado inicial”.

Se nota que la interpretación estocástica puede seguir dos diferentes aproximaciones. Podemos asumir que conocemos los procesos estocásticos que gobiernan la conducta del sistema de red (asumimos un modelo matemático) o queremos estimar sus parámetros basados sobre la observación de la conducta real de tal modelo (medimos experimentalmente el sistema o su modelo de simulación).

I.4.2.3 Reducción determinística

Esta es nuevamente una forma de reducir el no determinismo en una forma completa. Esta vez, la reducción es obtenida eliminando la posibilidad de variación. En algún sentido, esto podría ser visto como un caso especial de la reducción no determinística, en la cual $t_m=t_M$. Sin embargo, se prefiere mantener este caso separado para entender totalmente todas las implicancias de esta case de reducción. Por ejemplo, el hecho que la conducta completa del modelo interpretado es especificada en una forma determinística permite a uno computar todos los resultados de performance (al menos todos los resultados de tiempo transitorio por simple simulación). Sobre el otro extremo, el modelo interpretado puede disponer de más propiedades seguras y diferentes propiedades de vivacidad cuando es comparada al sistema de red subyacente.

Las políticas de resolución de conflictos determinísticos pueden obviamente ser definidas además a la definición de retardo de disparo de transiciones fijas determinísticamente.

I.4.2.4 Mezcla de técnicas de reducción diferentes

Algunas veces uno puede encontrar en la literatura propuestas de interpretaciones temporizadas para redes de Petri en las cuales los retardos de disparo y la elección estocástica se mezclan entre las transiciones conflictivas. Esto es probablemente el ejemplo más común de mezcla de técnicas diferentes para la reducción del no determinismo.

Otro ejemplo, es la definición de soporte finito para los retardos de disparo de transición en el marco de la interpretación estocástica. Además, esto puede ser visto como la superposición de la primer técnica (para reducir el soporte desde $[0, \text{infinito})$ a $[t_m, t_M]$) y de la segunda técnica (para proveer una medida de probabilidad sobre el rango de variabilidad). Esta forma de presentar la interpretación deja en claro que esta interpretación hereda todas las propiedades de las dos técnicas de reducción, llamadas para permitir el análisis de performance exacto, y en general, cambio de la conducta cualitativa con respecto al sistema de red subyacente sin interpretación.

I.4.3 Transiciones Temporizadas e Inmediatas

Intuitivamente, antes de discutir todos los detalles de las interpretaciones de temporización de las transiciones, tiene sentido introducir una primera distinción estricta entre eventos de tiempo cero y actividades de tiempo no cero. Las motivaciones para hacer tal distinción son varias. Desde el punto de vista de la modelación, algunas veces es útil implementar en forma de sistema de red, algún pequeño algoritmo para computar resultados más complejos que no pueden ser expresados directamente en términos de primitivas de Redes de Petri. En estos casos, no queremos hacer que el disparo de estas “transiciones adicionales” perturbe la conducta en el tiempo de nuestro modelo, de modo que los podemos especificar como “transiciones inmediatas”, implementando todos los pasos de nuestro algoritmo en tiempo cero.

En otras palabras, la duración de algunas actividades podría ser despreciable comparada con otras. En estos casos, podría ser conceptualmente más simple y/o más conveniente desde el punto de vista del análisis definir las actividades “cortas” como realmente teniendo tiempo cero.

En cualquier caso, la interacción entre las transiciones inmediatas y temporizadas debe estar claramente definida. La forma usual de definir esta interacción es introducir el concepto de prioridad. Esto está intuitivamente justificado en la forma siguiente. Si observamos

la conducta de dos transiciones $t1$ y $t2$, una inmediata y la otra caracterizada por un retardo de disparo $\epsilon > 0$ (figura n° 14). Una vez que las dos están habilitadas simultáneamente, debido al disparo de una transición $t3$, la transición inmediata se disparará siempre antes que la temporizada. Por ello, podríamos intuitivamente decir que $t1$ tiene prioridad sobre $t2$ debido al hecho que ella es inmediata más que temporizada.

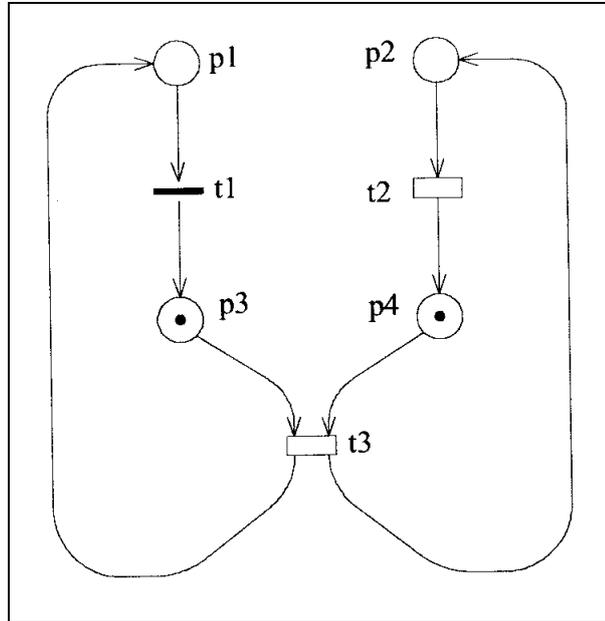


figura n° 14 Transiciones inmediatas y temporizadas

Esta idea intuitiva de prioridad de transiciones inmediatas sobre las transiciones temporizadas puede ser formalizada (y extendida) como una interpretación adicional y separada del sistema de red básico, independientemente del concepto de temporización. Esta introducción de prioridades en sistemas de red no temporizadas nos permite estudiar la interacción entre transiciones temporizadas e inmediatas sin realmente tener que definir la interpretación de la red, y ver cuales son los efectos sobre las propiedades cualitativas del sistema de red. El sistema de red con prioridades graficado en la figura n° 15 está limitado debido a la interpretación de prioridad, independientemente de la interpretación de tiempo real, mientras que la red subyacente P/T no interpretada es estructuralmente no limitada. Después de la introducción de una interpretación de prioridad consistente, la introducción de una interpretación de tiempo (siempre incluyendo transiciones inmediatas) no produce ningún cambio más que las propiedades cualitativas del modelo. Además, para establecer esta consistencia entre la interpretación del tiempo y el uso de las prioridades, se requiere que:

- las transiciones temporizadas tengan prioridad 0, y sean graficadas como una caja rectangular en el representación de la red;
- las transiciones inmediatas tengan prioridad mayor que 0, y sean graficadas con barras delgadas en la representación de la red.

Bajo estas condiciones, las especificaciones de prioridad y temporización son complementarias para la definición de la conducta observable en el tiempo; por lo tanto, las transiciones inmediatas siempre disparan antes que las temporizadas. Un ejemplo clásico que puede ser encontrado en la literatura de un modelo temporizado y priorizado consistente son las Redes de Petri Estocásticas Generalizadas (GSPNs Generalized Stochastic Petri Nets - ver punto I4.9).

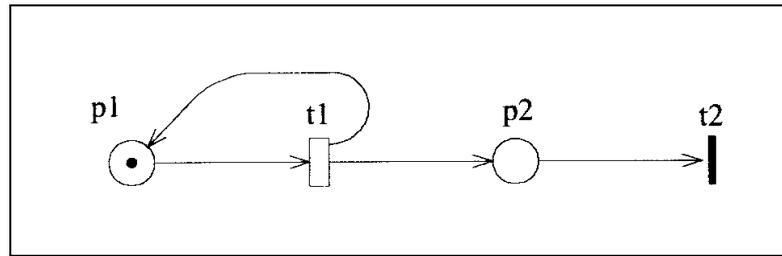


figura nº 15 Un sistema no limitado que se vuelve seguro debido a las interpretaciones de seguridad

I.4.4 Especificación de Servidores, conflictos y políticas

Servidores finitos e infinitos. Consideremos dos ejemplos alternativos de los sistemas de red de un multiprocesador PLC (Programmable Logical Controller) con sus versiones coloreadas y no coloreadas (ver figura nº 13 para la versión coloreada). Los dos modelos son propuestos para representar exactamente la misma conducta en diferentes niveles de detalle: el primer modelo nos permite mantener un seguimiento de que procesador esta haciendo que; el segundo modelo corresponde a una vista más abstracta del sistema, en la que puede aún observarse la actividad de los diferentes procesadores, pero no podemos mantener un seguimiento de su identidad. Además, para tener esta clase de “equivalencia” entre los modelos coloreados y no coloreados, necesitamos asumir la semántica servidor-infinito para relacionar las actividades para habilitar las transiciones. En un modelo temporizado, el grado de habilitación de cada transición temporizada representa el número de actividades concurrentes modeladas por la habilitación de las transiciones en un marcado dado.

Asumiendo la semántica infinita servidor-infinito para las transiciones temporizadas significa que queremos modelar el equivalente de una cola servidor-simple, tenemos que estar seguros que la transición temporizada que representa la finalización de la actividad del servidor tiene un grado de habilitación máximo de 1. Esto puede obtenerse en una forma sistemática, sumando un lugar conteniendo siempre una marca, conectado con el arco de entrada y un arco de salida a la transición del “server”. Como una notación abreviada en la representación gráfica de los sistemas de red, puede evitarse poner el lugar adicional para reducir el grado de habilitación máximo y usar además la inscripción “1-server” (o “k-server” en general) asociado con la transición.

Políticas de Colas. Cuando consideramos el caso de transiciones k-server que podrían estar habilitadas más que k veces (si no sumamos la condición server al sistema de red), entonces el uso de políticas de cola similar a la explícitamente en los modelos de cola podrían ser útiles. En principio, si tenemos un límite de k servers disponibles para llevar a cabo actividades en paralelo, y tenemos $n > k$ “clientes” demandando actividades, entonces lo mejor que podemos hacer es elegir k clientes y comenzar su actividad mientras que los otros $n - k$ clientes permanecen esperando por un servidor que se libere. La política de cola es precisamente el algoritmo que seguimos para seleccionar k salidas de n clientes para darles el “privilegio” de comenzar sus servicios de una forma directa. En algunos modelos de colas, las políticas de cola pueden ser especificadas, algunas de ellas teniendo en cuenta el orden de arriba en clases (por ejemplo, según prioridades), otros implementando políticas más “democráticas” con el objeto de no hacer distinciones entre clientes, etc.

En nuestro caso, dado que las marcas (del mismo color) en el mismo lugar son indistinguibles, no tiene sentido presentar el problema en el caso de transiciones persistentes: la

conducta que puede ser observada en el nivel de detalle de las cantidades operacionales no podría ser afectada por elecciones diferentes. Por el contrario, si pueden ocurrir los conflictos, entonces la elección de una política de cola particular afectaría la política de resolución de conflictos: asumiendo por ejemplo FCFS significaría especificar una reducción determinística del inherente no determinismo a la elección en el nivel de la Red de Petri. Usualmente, las políticas de colas explícitas no son definidas para las Redes de Petri temporizadas, y el problema es resuelto en el nivel de elección de la semántica para el caso de transiciones conflictivas (es decir, en un nivel más fino de la conducta del sistema de red). En general, las disciplinas de colas de prioridad, orden aleatorio y PS son directas para implementar usando esta aproximación (especialmente en el caso de temporización markoviana), mientras que las disciplinas de cola dependientes del estado como FCFS (First Come First Served), LCFS (Last Come First Served), primero el trabajo más largo, etc. resultan en representaciones de sistemas de red complejos.

Conceptos complementarios en Apéndice B. (Resolución de conflictos, y Conflictos y Precedencia)

I.4.5 Observación en el tiempo: Ecuaciones de ocurrencia

La conducta en el tiempo de un modelo de red P/T, o aún coloreada, puede ser completamente representada en términos de secuencias de tiempos de ocurrencias de los eventos de interés en la red, es decir, la creación y destrucción de marcas en lugares, y la habilitación y disparo de transiciones. Al efecto, podría construirse un diagrama de tiempo de ocurrencias. Además, para formalizar el estudio de esta clase de descripción conductional extremadamente detallada, se comienza por definir las ecuaciones de ocurrencias, para usarlas en la descripción de la conducta en el tiempo de un modelo de Red de Petri, en una forma matemáticamente precisa.

I.4.6 Análisis Operacional

El análisis operacional de los modelos de Redes de Petri puede ser considerado como un caso particular del uso de las ecuaciones de ocurrencias, en la que los eventos y sus relaciones son explotadas para derivar algunas medidas de performance de alto nivel y exactas relaciones entre ellas. Los principios básicos del análisis operacional de sistemas de eventos discretos son:

- definir cantidades básicas que en principio son observables y medibles en un experimento corriendo desde tiempo 0 a un tiempo dado,
- definir cantidades derivadas en términos de otras básicas,
- probar simples relaciones algebraicas entre las cantidades básicas y derivadas que se mantienen válidas para cualquier posible experimento realizado usando un modelo.
- Dar una interpretación física consistente a las cantidades operacionales.

Algunas de las definiciones básicas para cantidades operacionales son:

- marcado instantáneo: es el número de marcas presentes en el lugar p en el tiempo T ,
- marcado promedio: es la relación entre el área generado por el marcado en función del tiempo, hasta una cierta coordenada de tiempo, dividida por el valor de esa coordenada de tiempo,
- grado de habilitación instantáneo: es el número de servidores que están corrientemente activos, que están realizando trabajo de disparo de transición,

- trabajo de habilitación total: es el tiempo activo total para todos los servidores asociados con el disparo de transición durante un tiempo T ,
- grado de habilitación promedio: es el número promedio de servidores activos asociados con el disparo de transición durante un tiempo T ,
- cuenta de disparo total: es el número total de disparos de una transición t observados desde un tiempo 0 a T ,
- rendimiento: es la frecuencia de disparo real de la transición t observada durante un tiempo T , y
- tiempo de disparo promedio: es el tiempo promedio para que un servidor complete un disparo durante un tiempo T .

I.4.7 Observación y especificación de modelos temporizados

Con la introducción de las ecuaciones de ocurrencia y del análisis operacional, direccionamos el problema de la observación de la conducta de un modelo temporizado, con el objeto de medir algunos de sus números relacionados a la performance, como por ejemplo un rendimiento de transición particular. Esto se suma al problema de especificar la conducta en el tiempo para un sistema de Red de Petri no temporizado, no interpretado. Los dos aspectos (especificación y observación-computación de las medidas de performance) están, sin embargo, fuertemente relacionadas, en el sentido que las medidas de cantidades operacionales deben ser consistentes con la interpretación/especificación del modelo.

El punto crítico en establecer una relación consistente entre la especificación de temporización y las medidas de performance es el concepto de intervalo de habilitación de transición. Si se especifica el retardo de habilitación de la transición, o bien determinísticamente o estocásticamente, entonces se puede computar una cantidad operacional, sin realmente tener que efectuar una medida experimental de nuestro modelo, y ser capaz de aplicar leyes operacionales para computar otras medidas de performance directamente de la especificación de nuestro modelo interpretado temporizado.

En este punto, la idea detrás del uso de las Redes de Petri temporizadas para las medidas de performance debería ser intuitivamente clara.

I.4.8 Casos de Markov y semi-Markov (En Apéndice C-Son Redes particulares simplificadas)

I.4.9 Redes de Petri Estocásticas Generalizadas (GSPNs Generalized Stochastic Petri Nets)

Un ejemplo muy popular de Red de Petri temporizada, son las GSPNs ó Redes de Petri Estocásticas Generalizadas. En este caso, la reducción del no determinismo es estocástico, con la restricción adicional de considerar solo tiempos de disparo exponencialmente distribuidos para transiciones temporizadas. Esta elección tiene fuerte impacto sobre la definición de la semántica de temporización, simplificando drásticamente la definición de un modelo de competencia para las transiciones temporizadas.

Las GSPNs permiten los conflictos para las transiciones temporizadas sin ninguna restricción (elección libre o no) con una semántica puramente de competencia. La elección natural para la política de deshabilitación en el caso de transiciones múltiple-server es por supuesto aleatoria, lo que es perfectamente consistente con las propiedad sin memoria de todas las temporizaciones de transiciones. En la definición temprana de la dependencia de marcado GSPNs general, se permitió definir el parámetro frecuencia de disparo para cada transición habilitada. En la definición revisada, solo la dependencia de habilitación se permitió, para forzar al modelador a incrementar la exactitud de la representación de red no temporizada subyacente,

para que concentre la modelación explícita de actividades en términos de servidores de transición activa.

En las GSPNs están también permitidas las transiciones inmediatas, y en caso de conflicto se asume una elección aleatoria para definir un orden de disparo entre los así llamado "sets" de conflictos extendidos". Esta definición implica una política de preselección definida correctamente, entre transiciones inmediatas en conflicto. Las GSPNs implican el uso de una estructura de prioridad consistente para las transiciones, en la cual todas las transiciones temporizadas tienen el nivel de prioridad más bajo, mientras que las transiciones inmediatas están definidas para tener prioridad mayor que las transiciones temporizadas. El uso de niveles de prioridad diferente para diferentes transiciones permite definir la resolución de conflicto determinístico que resalta la naturaleza estocástica de la resolución de conflicto dentro de los conjuntos de conflicto extendido. La combinación de reglas de preselección determinística y estocástica para la solución de conflictos entre transiciones inmediatas es regulada por algunos criterios de corrección que proveen suficientes condiciones para hacer seguro que la conducta del modelo está completa y correctamente especificada en el nivel de red, sin tener que necesariamente estudiar las secuencias de disparo de transiciones antes de definir las probabilidades de elección.

La potencia de modelación y conveniencia de las GSPNs para la representación de conductas complejas es también incrementada por la adopción de una definición extendida para el sistema de red subyacente que incluye arcos inhibidores múltiples, como así también prioridades de nivel múltiple para las transiciones inmediatas.

Concentrados en la interpretación estocástica, un simple parámetro numérico tiene que estar asociado con cada transición: la frecuencia de disparo asociado con cada transición temporizada, de servidor infinito, distribuida exponencialmente, y un peso de disparo asociado con cada transición inmediata que podría estar habilitada en conflicto con otras transiciones en el mismo nivel de prioridad.

I.4.10 Redes Estocásticas Bien Formadas (SWNs Stochastic Web-formed Nets)

Otro ejemplo popular de la Red de Petri temporizada es una extensión de red GSPN Coloreadas, conocidas con el nombre de Redes Estocásticas Bien Formadas. La idea que conduce a este modelo es extender la semántica de temporización GSPN, pero aplicable sobre una estructura de un modelo de red de alto nivel. La motivación principal para la propuesta de este modelo de GSPNs Coloreado, llamado SWNs, es explotar la simetría de la red que puede ser fácilmente expresada en una manera concisa usando sistemas de Red Bien Formadas, en lugar de sistemas de red P/T, no sólo para dibujar modelos de red más compactos, sino también para reducir la complejidad de análisis de los modelos estocásticos. Desde el punto de vista de la especificación de la semántica de tiempo, un modelo SWN está completamente especificado por asignar una simple parámetro de frecuencia numérica o de peso a cada instancia de color habilitada de cada transición. Además, para preservar las propiedades de simetría, también en la conducta del tiempo, el parámetro de frecuencia o peso se asume ser idéntico para todas las instancias color pertenecientes a la misma subclase de color estático.

I.4.11 Redes de Petri Determinística y Estocástica (DSPNs Deterministic and Stochastic Petri Nets)

Este modelo también fue propuesto como una extensión del popular modelo GSPN. La motivación para este caso es permitir una mezcla de dos diferentes mecanismos de temporización para las transiciones en el mismo modelo: transiciones estocásticas con

distribución sin memoria para retardos de disparo y transiciones determinísticas con retardo de disparo fijo.

De acuerdo a nuestra taxonomía de reducción no determinística, este es un modelo mixto, en el cual la reducción determinística completa se asume para algunas transiciones, mientras que reducción estocástica es asumida para otras transiciones. En principio, una definición arbitraria del subconjunto de transiciones temporizadas determinísticas podría tener un fuerte impacto sobre la conducta cualitativa del sistema de red subyacente. Sin embargo, la definición de DSPN fue introducida en tal forma, de modo de evitar esta clase de errores entre los modelos interpretados y los no interpretados.

Una condición suficiente que garantiza la consistencia cualitativa en la conducta después de la introducción de la interpretación, que fue asumida como parte de la definición de DSPNs, es que la interpretación de tiempo determinística está permitida solo para las transiciones con un servidor simple que nunca son habilitadas concurrentemente con otras transiciones temporizadas determinísticamente. Además, esto implica que siempre una transición temporizada determinística esta habilitada, todas las otras transiciones temporizadas podrán comportarse de acuerdo a la interpretación estocástica con distribución de probabilidad de soporte infinito. Así, cualquier valor es asignado como un tiempo de disparo fijo para tal transición determinística, dado que es mayor que cero y menos que infinito, cada otra transición temporizada tendrá una chance para dispararse antes con una probabilidad no nula, y una chance a dispararse después con un probabilidad no nula. Por lo tanto, todas las intercalaciones de disparo están aún permitidas, como en el caso del sistema de red no determinístico, no interpretado.

Las transiciones temporizadas determinísticamente deben ser ulteriormente especificadas en términos de la política de memoria. Tal especificación puede se evitada sólo en caso de transiciones persistentes. Debido a problemas de complejidad en el estudio matemático del modelo, la definición original de DSPN permitió sólo políticas de memoria de habilitación asociada con las transiciones determinísticas, aunque desde el punto de vista de la especificación, memoria de paso de tiempo no tendría problema adicional.

I.5) PROPIEDADES LOGICAS DE LOS SISTEMAS P/T Y SU ANÁLISIS [14][15][16]

I.5.1 Propiedades lógicas básicas

Sólo unas pocas propiedades son significativas para cualquier sistema concurrente, no sólo para aquellos modelados con las Redes de Petri. No obstante, estos estamentos usando los conceptos y objetos de Redes de Petri las hacen especialmente “fácil para entender” en varios casos. Las propiedades a ser consideradas son:

- 1) Limitación, caracterizando el número finito del espacio de estado,
- 2) Vivacidad, relacionado a la disparabilidad potencial en todos los marcados alcanzables. La liberación de bloqueos mortales es una condición más débil en la cual sólo la actividad infinita global (por ejemplo la disparabilidad) de modelo de sistema de red está garantizada, aún si alguna parte de él no trabaja.
- 3) Inversibilidad, caracterizando la recuperación del marcado inicial desde cualquier marcado alcanzable.
- 4) Exclusión Mutua, tratando con la imposibilidad de submercado simultáneo (p-mutex) o concurrencia de disparo (t-mutex).

En la figura n° 16, el disparo de t_2 conduce a $m = p_3 + p_4$. Disparando ahora t_4 , se alcanza $m_1 = p_1 + p_3$. Repitiendo w veces la secuencia t_2t_4 , el marcado $m_w = p_1 + wp_3$ se alcanza. Así, el marcado de p_3 puede ser arbitrariamente grande; y el lugar p_3 se dice ser

ilimitado. En la práctica, la capacidad de elemento físico representado por p3 debería ser finito, y un overflow puede aparecer, lo que es una situación patológica.

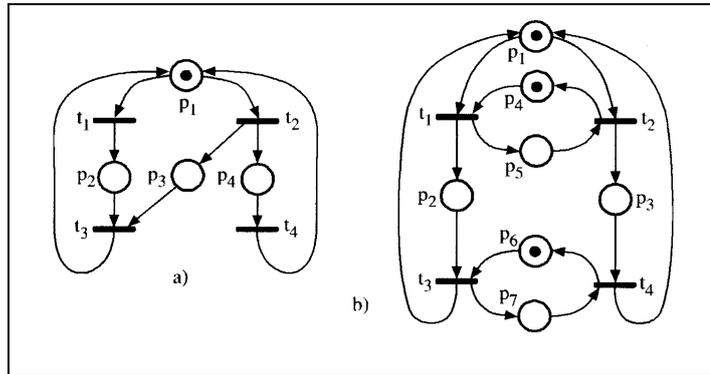


figura n° 16 Sobre las conductas patológicas cualitativas
 (a) un sistema de red ilimitado, bloqueable mortalmente (no vivo) e irreversible
 (b) al incrementar el marcado inicial (mo[p5] = 1) el sistema de red vivo puede alcanzar un estado de bloqueo mortal.

El número máximo de marcas que un lugar puede contener es su límite de marcado. Un lugar está limitado si su límite es finito. Un sistema de red está limitado si cada lugar es limitado. La limitación de un sistema (con todos los lugares limitados) es una propiedad conductacional generalmente requerida.

Para cualquier marcado inicial sobre la estructura de la red de la figura n° 17, se pueden definir las siguientes leyes de conservación de marcas:

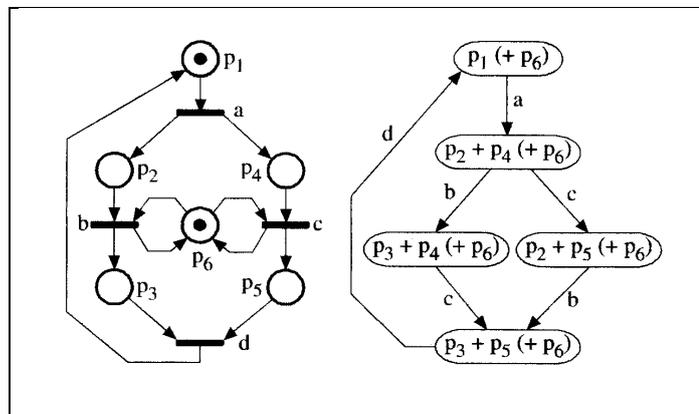


figura n° 17 Un sistema de red limitado, vivo y reversible, y su grafo de alcanzabilidad

$$\begin{aligned}
 m[p1] + m[p2] + m[p3] &= mo[p2] + mo[p3] = k1(mo) \\
 m[p1] + m[p4] + m[p5] &= mo[p1] + mo[p4] = k2(mo) \\
 m[p6] &= mo[p6] = k3(mo)
 \end{aligned}$$

donde mo es el marcado inicial y m cualquier marcado alcanzable. Por lo tanto,

$$\begin{aligned}
 m[p1] &\leq \min(k1(mo), k2(mo)) \\
 m[pi] &\leq k1(mo), i = 2,3
 \end{aligned}$$

$$m[p_i] \leq k_2(m_0), j = 4,5$$

$$m[p_6] = k_3(m_0)$$

Las desigualdades de arriba significan que para cualquier m_0 , el sistema de red está limitado. Esta propiedad, más fuerte que la limitación, es llamada limitación estructural a causa que se mantiene independientemente del marcado inicial (solo el número finito de m_0 es asumido).

Un escenario alternativo se da cuando disparamos en la figura n° 16-a la transición t_1 desde el marcado. Después de eso, ninguna transición puede ser disparada, y una situación de bloqueo mortal total se ha alcanzado. Un sistema de red se dice estar libre de bloqueos mortales si siempre (es decir, desde cualquier marcado alcanzable) al menos una transición puede ocurrir. Una condición más fuerte que la libertad de bloqueos mortales es la vivacidad. Una transición t es potencialmente disparable en un marcado m dado, si hay una secuencia de disparo de transición que conduce a un marcado m' en el cual t está habilitado. Una transición está viva si ella es potencialmente disparable en todos los marcados alcanzables. En otras palabras, una transición está viva si ella nunca pierde la posibilidad de disparo (es decir, de realizar alguna actividad). Un sistema de red está viva si todas las transiciones están vivas.

Para cualquier marcado inicial sobre la figura n° 16-a podemos definir que ninguna vivacidad se mantiene (de hecho, un bloqueo mortal total siempre será alcanzado). La no vivacidad para los marcados iniciales arbitrarios refleja una patología de la estructura de red: no vivacidad estructural. Una red es estructuralmente viva si hay al menos un marcado inicial vivo.

Un marcado está es estado hogar si es alcanzable desde cualquier otro marcado alcanzable. El marcado inicial del sistema de la red en la figura n° 16-a no está en estado hogar: después del disparo de la transición t_3 o t_4 no es posible recuperar este marcado inicial. En el caso particular que el marcado inicial sea un estado hogar, el sistema de red es reversible, de modo que siempre es posible retornar al marcado inicial.

La vivacidad, la limitación y la reversibilidad son justamente tres “buenas” propiedades conductacionales importantes para estudiar un sistema de red. Aunque estas características son propiedades conceptuales para las redes en general, debe evaluarse en cada modelo de red la satisfacción ó no de las mismas con el objeto de determinar su correspondencia con el sistema real, ó su viabilidad práctica. La figura n° 18 demuestra que ellas son independientes una de otra, en cada uno de los 8 casos posibles.

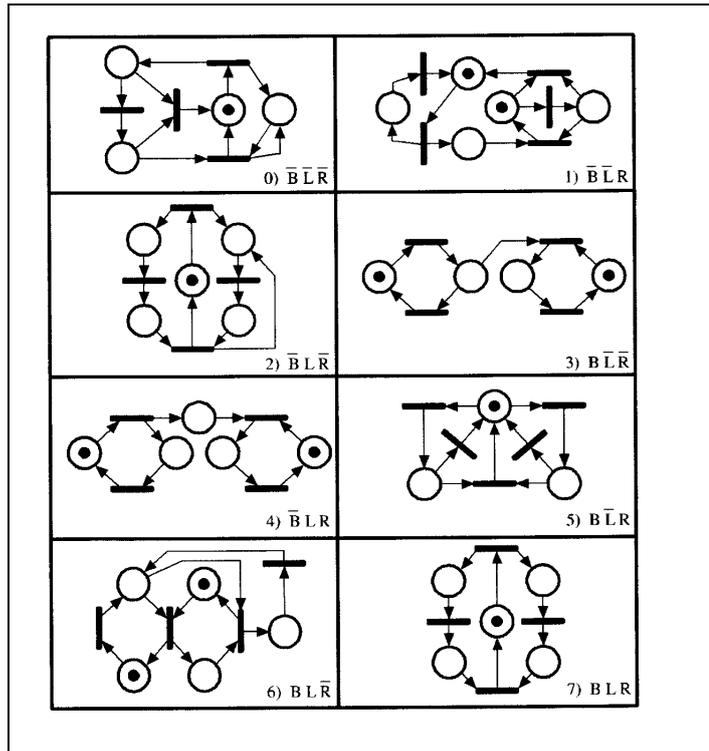


figura n° 18 La limitación (B), la vivacidad (L) y la reversibilidad (R) son propiedades independientes

La última propiedad básica es la exclusión mutua. Esta propiedad captura condiciones como la imposibilidad de un acceso simultáneo de dos robots a un simple almacén. Dos lugares (transiciones) están en exclusión mutua si nunca pueden estar marcados (disparados) simultáneamente. Por ejemplo, en el sistema de red de la figura n° 16 podemos escribir:

$$m[p1] + m[p2] + m[p3] = 1$$

de modo que p1, p2 y p3 están en exclusión mutua.

I.5.2 Técnicas de análisis básicas para sistemas de red P/T

Convencionalmente, las técnicas de análisis para las Redes de Petri se clasifican como: 1) Enumeración, 2) Transformación y 3) Análisis Estructural. Los métodos de simulación también han sido aplicados para estudiar los sistemas modelados con redes P/T. Ellos proceden desplazando las marcas sobre el modelo de sistema de red bajo ciertas estrategias. En general, los métodos de simulación no permiten probar las propiedades, pero ellos podrían ser una gran ayuda para entender el sistema modelado o para fijar los problemas manifestados durante la simulación. Los métodos de simulación son extremadamente útiles cuando el tiempo es asociado con la evolución de la red (sistemas temporizados), o cuando deseamos conocer la repuesta del sistema descrito con una red en una ambiente que también es definido por la simulación.

1) Los métodos de Enumeración: están basados sobre la construcción de un grafo de alcanzabilidad (RG reachability graph) el que representa, individualmente, el marcado de la red y el simple disparo de transiciones entre ellos. Si el sistema de red está limitado, el grafo de alcanzabilidad es finito y las diferentes propiedades cualitativas pueden ser fácilmente verificadas. Si el sistema de red no está limitado, el RG es infinito y su construcción no es posible. En este caso, pueden construirse los grafos finitos conocidos como grafos de cobertura.

A pesar de su potencia, la enumeración es frecuentemente difícil para aplicar, aún en pequeñas redes, debido a su complejidad computacional creciente y proporcional al tamaño de la red. (No obstante hay propuestas como las FPQSPN en I.7 que acotan sustancialmente este problema).

2) El análisis por Transformación: procede transformando un sistema de red $S = (N, mo)$ en un sistema de red $S' = (N', mo')$ preservando el conjunto de propiedades a ser verificadas (es decir, S' satisface las propiedades si S las satisface). El objetivo final es verificar las propiedades en S' en una forma más fácil que en S . El espacio de estado de S' puede ser mayor que de S , pero S' puede pertenecer a una subclase para la cual la enumeración de estado puede ser evitada.

Los métodos de reducción son una clase especial de métodos de transformación en los cuales se construye una secuencia de sistemas de red preservando las propiedades a ser estudiadas. La construcción se hace de tal forma que el sistema de red nuevo es más pequeño (menos marcados o el grafo de alcanzabilidad tiene menos lugares o transiciones) que el previo en la secuencia.

La aplicabilidad de los métodos de reducción está limitada por la existencia de sistemas de red irreducibles. Prácticamente hablando, las reducciones obtenidas son normalmente considerables, y pueden permitir que las propiedades sean verificadas directamente. A causa de la existencia de sistemas irreducibles, este método debe ser complementado por otros.

3) Las Técnicas de Análisis Estructural investigan las relaciones entre la conducta de un sistema de red y su estructura, mientras que el marcado inicial actúa básicamente como un parámetro. En esta última clase de técnicas de análisis podemos distinguir dos subgrupos:

- 1) Álgebra Lineal/Técnicas basadas en la programación lineal: que están basados sobre la ecuación de estado de red. En ciertos análisis ellas permiten un diagnóstico rápido sin la necesidad de enumeración.
- 2) Técnicas basadas en grafos: en que la red es vista como un grafo bipartito y se aplican algunos razonamientos ad hoc (frecuentemente derivados de las reglas de disparo). Estos métodos son especialmente efectivos en analizar subclases restringidas de redes ordinarias.

Los tres grupos de técnicas de análisis no son excluyentes entre si, sino complementarias. Normalmente, el diseñador puede usarlas de acuerdo a la necesidad de los procesos bajo análisis. Obviamente, aunque tenemos que distinguir entre los métodos de transformación/reducción y de análisis estructural, debemos resaltar que las técnicas de reducción más populares actúan básicamente sobre el nivel de estructura de la red, y así pueden ser consideradas también como técnicas estructurales.

En lo que concierne al análisis cualitativo de los sistemas interpretados, debería señalarse que las conclusiones acerca de las propiedades del modelo autónomo subyacente pueden ser solo suficientes, necesarias o ni suficientes ni necesarias. Para subclases de redes particulares, bajo suposiciones razonables sobre la conducta, existen las condiciones de ambiente necesarias y suficientes.

I.5.3 REDES DE PETRI ESTOCASTICAS EXPONENCIALES-SPN (En Apéndice D)[17][18]

Los sistemas de Redes de Petri Estocásticas pueden ser convertidos en cadenas de Markov, y a partir de ello, sus análisis pueden ser realizados para computar interesantes índices de performance. La construcción de la cadena de Markov asociada con un sistema de Red de Petri Estocástica es la base para la derivación subsecuente del modelo probabilístico asociado con una Red de Petri Estocásticas Generalizada.

I.6) REDES DE PETRI COLOREADAS (CPNs COLOURED PETRI NETS) [19][20]

Las Redes de Petri Coloreadas fueron introducidas en el punto I.3.9. Las mismas son de amplia utilización para diferentes propósitos prácticos. La principal razón de este tipo de Redes de Petri de Alto Nivel (HLPNs) es que los modelos de red resultantes tienen una representación gráfica y una semántica bien formada que permiten análisis formal. El paso desde las Redes de Petri de Bajo Nivel (LLPNs) a las Redes de Petri de Alto Nivel puede ser comparado al paso desde los lenguajes assembler a los modernos lenguajes de programación con un concepto elaborado de tipo. Como se discutió en el punto introductorio, en las Redes de Petri de Bajo Nivel hay sólo una clase de marcas, y esto significa que el estado de un lugar está descrito por un entero (y en algunos casos por un valor booleano). En las Redes de Petri de Alto Nivel, cada marca puede transportar información compleja o datos, con lo que pueden describir el estado entero de un proceso o una base de datos.

Desde esta perspectiva, las Redes de Petri Coloreadas y el conjunto de herramientas software que las soportan, como CPN Tools, constituyen una instrumento importante para el diseño, especificación, simulación, validación e implementación de aplicaciones industriales, entre las que se incluyen los protocolos de comunicaciones.

I.6.1 Definición formal

Una CPN no-jerárquica se define formalmente como una Tupla de la forma CPN = $(\Sigma, P, T, A, N, C, G, E, I)$, donde

- Σ es un conjunto finito de tipos no vacíos, también llamados conjuntos de colores.
- P es un conjunto finito de lugares.
- T es un conjunto finito de transiciones.
- A es un conjunto finito de arcos tal que: $P \cap T = P \cap A = T \cap A = \emptyset$
- N es una función de nodo, definida como $N : A \rightarrow (PxT) \cup (TxP)$
- C es una función de color, definida como $C : P \rightarrow \Sigma$
- G es una función de guarda. Está definida de T hacia expresiones lógicas
- E es una función de expresión de arco.
- I es una función de iniciación

El set de colores (Σ) determina los tipos, operaciones y funciones que pueden usarse en las notaciones de la red (expresiones en arco, guardas, expresiones de inicialización, etc.). Se asume que cada tipo tiene al menos un elemento. Los conjuntos P, T y A son conjuntos finitos disjuntos de plazas, transiciones y arcos, como se definieron en las Redes de Petri Generales. Los conjuntos de lugares, transiciones y arcos deben ser finitos.

La función de nodos N mapea cada arco en un par (x_1, x_2) donde x_1 es el nodo origen y x_2 es el nodo destino. Los dos nodos tienen que pertenecer a conjuntos distintos, es decir, si el primero es un lugar el segundo debe ser una transición y viceversa. En una CPN se permite tener varios arcos entre un par ordenado de nodos, de modo que A es un conjunto separado. La utilización de arcos múltiples proporciona una facilidad para el modelado de sistemas complejos.

C define función de color que mapea cada lugar p, a un tipo C(p). Intuitivamente esto significa que cada marca en p debe tener un valor de datos que pertenezca a C(p), es decir que corresponda a un tipo de Σ .

La función guarda G mapea cada transición t, en una expresión lógica donde todas las variables tienen tipos que pertenecen a Σ . La función guarda permite insertar

precondiciones para las transiciones, las cuales pueden indicar los requisitos bajo los cuales una transición puede ser disparada. Si se omiten las expresiones de guarda, estas se evaluarán como verdadero.

La función de expresión de arco E mapea cada arco a en una expresión del tipo $C(p(a))_{MS}$. Esto significa que cada expresión de arco debe evaluarse como multi-set de acuerdo al set de colores asignado al correspondiente lugar.

La función inicialización I mapea cada lugar p , dentro de una expresión del tipo $C(p)_{MS}$. Con la evaluación de las expresiones de iniciación se obtiene la marca inicial. Cuando se omite la expresión de iniciación, se evalúa como conjunto vacío (\emptyset).

A diferencia de las RdeP donde interesa conocer cuantas marcas que puede haber en un lugar determinado, ahora interesa conocer específicamente cuales son las marcas que puede contener un lugar. Para una RdeP considerada como CPN se requiere un solo conjunto de tipos en Σ con dos elementos (negro y blanco) indicando la presencia o ausencia de una marca. De esta forma la función de color C proporcionará dos valores (negro y blanco) para cada uno de los lugares de la red. La función de guarda G para cada transición se deberá omitir para que se evalúe Verdadero. A cada arco se le coloca una expresión de arco que entregue un token con valor negro cada vez que se requiera una marca del lugar adyacente. Finalmente, para obtener la marca inicial se evalúa la función iniciación para obtener color negro solo en los lugares donde se requiera una marca inicial.

I.6.2 Redes Bien Formadas

Las Redes Bien Formadas (WN) son una subclase de las CPNs, cuya peculiaridad es una sintaxis muy estructurada para la definición de los dominios color de los lugares y las transiciones, y de las funciones de arcos. La motivación para tal sintaxis es la posibilidad de explotar automáticamente la simetría intrínseca del modelo para generar eficientemente un grafo de alcanzabilidad agregado.

El punto inicial en la definición estructurada de la sintaxis de las WN coloreadas es el conjunto de clases de colores básicos $\{C_1, \dots, C_2\}$. Una clase color básico C_i es un conjunto no vacío y finito de colores; intuitivamente una clase de color básico puede ser definida como un conjunto de colores identificando objetos de la misma naturaleza. Una clase de color básico es ordenado si una función sucesor está definida sobre sus elementos, tal que induce una ordenación circular sobre los elementos de clase. Un ejemplo de clase ordenada es la clase de procesadores conectados en una topología ring. Las clases de colores básicas son disjuntas; sin embargo, una clase puede ser particionada en algunas subclases estáticas que representan objetos del mismo tipo pero con conducta diferente. Por ejemplo, la clase de color básica de procesadores podría particionarse en dos subclases estáticas disjuntas, una conteniendo los procesadores rápidos y la otra conteniendo los procesadores lentos.

I.6.3 Técnicas de Análisis para Sistemas Coloreados Bien Formados

El uso de las redes P/T para representar sistemas complejos puede dar como resultados modelos con muchos lugares y transiciones. Parte de esta complejidad puede ocultarse cuando para la representación se usan modelos de más alto nivel, tales como las redes coloreadas. Además, las mismas propiedades para los sistemas modelados con una red P/T son aplicables a las redes coloreadas; por ejemplo, los conceptos de vivacidad y de reversibilidad.

Sin embargo, la primera estrategia para decidir tales propiedades para una red coloreada podría ser transformar la red en una red P/T equivalente, por un proceso llamado

desplegado (usado también en la conversión de las redes FPQSPN), y luego aplicar las técnicas conocidas. Realmente hay dos importantes desventajas en esta aproximación: se multiplicará el número de lugares, transiciones y arcos, y por lo tanto, el costo de computar los resultados será alto; y también la manipulación de estos resultados puede no ser tan fácil.

La idea es entonces adaptar a las redes coloreadas las técnicas de análisis definidas para las redes P/T, lo que realmente significa que debemos ser capaces de trabajar directamente en el nivel de red en el caso coloreado. Esto implica manipular funciones de color además de valores numéricos. Si no hay restricciones en la definición de las funciones, nuevamente el costo será muy alto. Pero aún si las funciones son restringidas, los resultados de análisis definidos para las redes P/T no pueden extenderse automáticamente a las redes coloreadas.

En estos casos, la sintaxis de las redes bien formadas puede ser explotada para derivar técnicas de análisis eficientes. Por ejemplo, se puede tomar ventaja de la simetría intrínseca de un sistema coloreado, lo que está reflejado en las clases de color básicos, para derivar una representación reducida del grafo de alcanzabilidad.

I.7) REDES DE PETRI DE SISTEMAS DE COLAS DE POBLACIÓN FINITA (FPQSPNs FINITE POPULATION QUEUING SYSTEMS PETRI NETS) [21]

I.7.1 Introducción

Una reciente contribución a la Teoría de las Redes de Petri, introduce una clase de Redes de Petri como un resumen de una GSPN, que es apropiada para modelar y analizar sistemas vistos como sistemas de colas donde la población es finitamente grande. La propuesta mantiene la transparencia del modelo de las Redes de Petri de bajo nivel. Asociado a esta técnica, se presentó un algoritmo para el juego de marcas de eventos discretos, que resuelve la duda sobre la complejidad de las funciones relacionadas a las transiciones para modelos de Redes de Petri complejos y grandes. El algoritmo está basado sobre la condición de habilitación de transición parcial.

Las Redes de Petri de Sistemas de Cola de Población Finita (FPQSPN) son Redes de Petri de Alto Nivel apropiadas para modelar sistemas, donde algunos elementos funcionales idénticos (la población de usuarios) comparten recursos comunes. Un diagrama de FPQSPN se supone contiene un subdiagrama de población y un subdiagrama representando los recursos. Los recursos compartidos pueden representar un centro de llamada de un operador de teléfono celular, un canal de comunicaciones o una sección crítica, donde el resto del modelo FPQSPN representa un proceso de programa accediendo a la sección crítica, el canal de comunicaciones o simplemente un usuario demandando al centro de llamadas.

I.7.2 Definición Formal

Una FPQSPN es una 9-tupla:

$$N = (P, T, Pre, Post, Mo, so, t, tt, k)$$

Donde:

1. P y T son conjuntos no vacíos finitos disjuntos, los lugares y las transiciones, respectivamente.
2. Pre y Post son las matrices de incidencia anterior y posterior, de tamaño (PxT).

3. M_0 es un vector que representa el marcado inicial.
4. t es el tiempo asociado a la transición.
5. τ es el tipo de tiempo asociado a la transición.
6. τ so determina si el lugar o la transición están compartidos.
7. k es un número natural de clientes (en términos de sistemas de colas, es tamaño de la población).

El conjunto de arcos es notado por F , don $F = \text{Pre} \cup \text{Post}$. En FPQSPN hay dos tipos de lugares y transiciones: unos compartidos y otros no compartidos. El algoritmo de desplegado, asociado con esta propuesta, despliega cada uno de los lugares o transiciones no compartidos al conjunto de k lugares, donde k es el número de clientes. Los lugares y transiciones compartidos son dejados sin cambios.

Gráficamente, los objetos compartidos serán distinguidos por una doble línea. Los arcos son dibujados en la misma forma, indiferente que ellos conecten un lugar o transición compartida a otros no compartidos, o ellos conecten objetos que son compartidos o no compartidos. Durante el desplegado de FPQSPN a una Red de Petri de bajo nivel, la población es reemplazada por un conjunto de subdiagramas idénticos.

La objetivo principal para la definición de las FPQSPN fue definir una especificación de Red de Petri, que es fácil para implementar, y que es tan gráfica y transparente como los modelos de Redes de Petri de bajo nivel, y que es potente modelando protocolos de comunicación representados con sistemas de cola de población finita. Las FPQSPN son un resumen de las GSPN presentadas previamente.

I.7.3 Propiedades de Tiempo

Pueden haber transiciones inmediatas, transiciones de tiempo determinísticos o transiciones de tiempo estocásticas en FPQSPN. El tiempo asociado a la transición significa el tiempo entre el momento en que la transición está habilitada y el momento en que la transición es disparada. Una transición se vuelve disparable si y sólo si estará continuamente habilitada durante su tiempo asociado (la marca en los lugares entrantes residen hasta que el retardo de la transición se agota).

Un conflicto entre una transición inmediata y una transición disparable de tiempo no cero se resuelve de la misma manera que un conflicto entre dos transiciones inmediatas. En el caso de un conflicto entre una transición esperando tiempo y una transición disparable, la transición disparable es disparada primero. De igual forma, cuando dos transiciones de tiempo no cero se encuentran en conflicto estructural, habilitadas en el mismo momento, la transición con tiempo real más corto se dispara primero. En FPQSPN, el número de disparos simultáneos de una transición está limitado a uno (semántica de simple servidor). Igualmente, la transición puede estar habilitada sólo una vez. La transición no puede ser habilitada nuevamente durante su tiempo de periodo de espera.

I.7.4 Matriz de Incidencia

En la figura n° 19 se muestra un modelo FPQSPN representando una sección crítica. Los lugares P_1 y P_2 , y las transiciones T_1 , T_2 y T_3 representan un cliente; y los lugares P_3 y P_4 , y la transición T_4 representan un recurso compartido. La transición fuente T_1 es una transición de tiempo estocástico con tiempo de distribuido exponencialmente. Las transiciones T_3 y T_4 representan el paso a través de la sección crítica. Aquí, la transición de tiempo estocástico con distribución uniforme T_3 representa algunas acciones realizadas sobre el lado del

cliente y la transición de tiempo determinístico T4 representa las acciones sobre el lado del servidor. De acuerdo a la caracterización de Kendall (I.2.5), el modelo de la figura representa un sistema de cola M/G/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución general, y 1 servidor). La transición T1 representa procesos de arribo de paquetes Poisson. Las transiciones T3 y T4 representan tiempo de servicio estocástico.

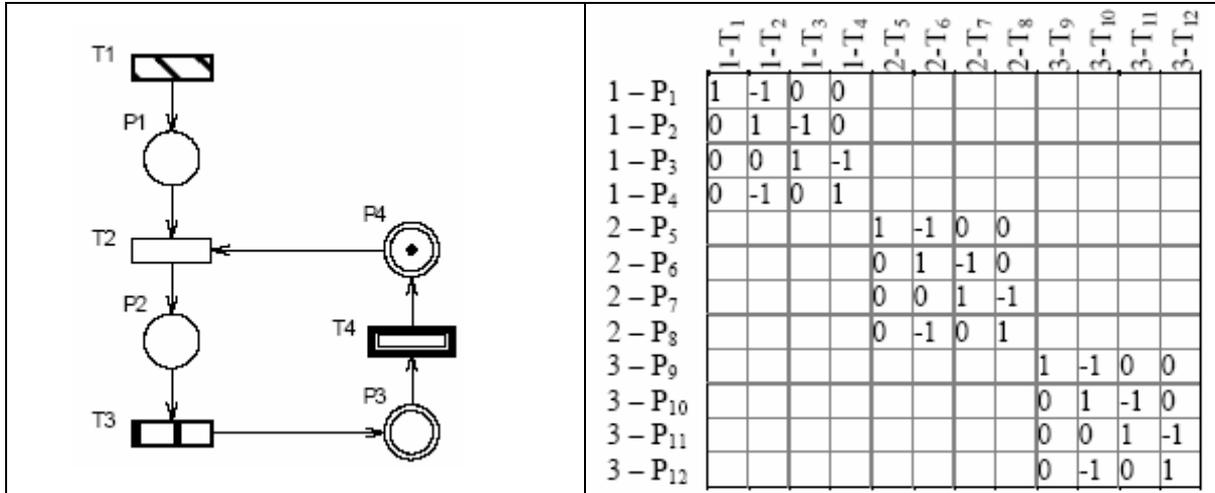


figura n° 19 Modelo FPQSPN de una sección crítica

figura n° 20 Tres modelos independientes

Las dimensiones de la matriz de incidencia están dadas como número de lugares multiplicado por k x número de transiciones multiplicado por k. En el cliente o consumidor nth, la posición del arco entre el lugar P_i y la transición T_j en la matriz de incidencia de la red desplegada puede ser calculada de acuerdo a la siguiente fórmula:

$$pos = [i + np * n, j + nt * n] \quad (n = 1 \dots k)$$

donde np es el número de lugares del modelo plegado, nt es el número de transiciones en el modelo plegado, y k es el número de clientes.

En la figura n° 20, hay una matriz de incidencia del modelo de Red de Petri, el cual está desplegado desde la observación de los objetos compartidos. Los tres modelos independientes aparecen en tal caso.

El lugar compartido puede ser usado para representar un recurso común o una variable global. Por ejemplo, cuando se modela un protocolo de comunicaciones, un conjunto de lugares compartidos pueden ser usados para los estados del canal de comunicaciones. Un par de lugares compartidos P3 y P4 en el modelo de la sección crítica indica si la sección crítica está libre ó ocupada.

La transición compartida puede ser usada para representar cualquier acción global o (en las Redes de Petri temporizadas) alguna constante de tiempo global en el modelo. Por ejemplo, cuando se modela los protocolos de comunicaciones, un conjunto de transiciones compartidas de tiempo no cero puede ser usado para pronunciar constantes de tiempo del canal de comunicaciones. O por ejemplo, la transición compartida de tiempo determinístico T4, en la figura n° 17 representa el tiempo de paso a través de la sección crítica.

I.7.5 Algoritmo de desplegado

Un algoritmo de desplegado genera la Red de Petri de bajo nivel, como una versión del diagrama. Cada objeto es procesado de la siguiente manera:

1. Si el objeto es un lugar i th no compartido, o una transición i th no compartida, crea k lugares indexados por $i + np*j$ ($j = [0...n-1]$), o n transiciones indexadas por $i + nt*j$ ($j = [0...n-1]$).
2. Si el objeto es un lugar o transición i th compartido, simplemente crea un lugar o transición indexada por i .
3. Si el objeto es un arco entre dos objetos no compartidos, crea un conjunto de arcos.
4. Si el objeto es un arco entre un objeto no compartido y un objeto compartido, crea k arcos con los mismos números sobre el lado del objeto compartido y sobre el lado del objeto no compartido.
5. Si el objeto es un arco entre dos objetos compartidos, solo se copia el arco.

En la figura n° 21 hay una versión desplegada del modelo de la figura n° 19, y su matriz de incidencia está en la figura n° 22. Solo objetos simples son creados en el caso de objetos compartidos.

La matriz de incidencia generada por el algoritmo de desplegado contiene $(k-1)*nshp$ filas libres y $(k-1)*nsht$ transiciones libres, donde k $nshp$ es el número de lugares compartidos, y $nsht$ es el número de las transiciones compartidas. Por algunas razones, es preferible tener la matriz de incidencia en una forma reducida.

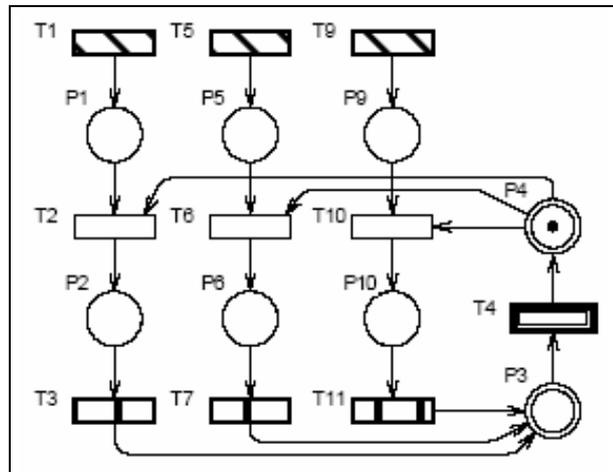


figura n° 21 Versión desplegada del modelo

	1-T ₁	1-T ₂	1-T ₃	1-T ₄	2-T ₅	2-T ₆	2-T ₇	2-T ₈	3-T ₉	3-T ₁₀	3-T ₁₁	3-T ₁₂
1 - P ₁	1	-1	0	0								
1 - P ₂	0	1	-1	0								
1 - P ₃	0	0	1	-1	0	1			0	1		
1 - P ₄	0	-1	0	1	-1	0			-1	0		
2 - P ₅					1	-1	0	0				
2 - P ₆					0	1	-1	0				
2 - P ₇					0	0	0	0				
2 - P ₈					0	0	0	0				
3 - P ₉									1	-1	0	0
3 - P ₁₀									0	1	-1	0
3 - P ₁₁									0	0	0	0
3 - P ₁₂									0	0	0	0

figura n° 22 Matriz de incidencia del modelo desplegado

Las dimensiones de la matriz de incidencia está dado por la siguiente expresión, después de remover las filas y columnas vacías.

$$\text{Tamaño } [A] = [k*(np - nshp) + nshp, k*(nt - nsht) + nsht]$$

Donde np es el número de lugares, y nt es el número de transiciones en el modelo plegado. La matriz de incidencia reducida de la figura n° 22 está en la figura n° 23. Por conveniencia, la matriz de incidencia reducida significa una matriz, donde las líneas o columnas vacías.

	T ₁	T ₂	T ₃	T ₄	T ₅	T ₆	T ₇	T ₈	T ₉	T ₁₀
P ₁	1	-1	0	0						
P ₂	0	1	-1	0						
P ₃	0	0	1	-1	0	1		0	1	
P ₄	0	-1	0	1	-1	0		-1	0	
P ₅					1	-1	0			
P ₆					0	1	-1			
P ₇								1	-1	0
P ₈								0	1	-1

figura n° 23 Matriz de incidencia reducida del modelo desplegado

I.7.6 Graficación de las Características 2D de los Modelos de Redes de Petri.

En varios casos, graficamos los parámetros de sistemas como una función de cierta variable. Por ejemplo, las características del rendimiento de los protocolos de comunicaciones se usan para ser graficados como curvas interpoladas entre algunos puntos. El análisis con el juego de las marcas nos provee información acerca de la conducta del modelo en un simple punto de operación. En tales casos, usualmente hacemos algunas tareas de simulación, una por cada punto de la curva. Así, cuando se grafica una característica que contiene n puntos, se hacen n pasos de simulación. En cada uno de ellos, se cambian los tiempos promedio de la transición designada (cambio del punto de operación) y se hace la simulación. Otra posibilidad

de cambiar el punto de operación es cambiar el peso del arco, o en algunos casos, el cambio del marcado inicial.

Cuando se simula un modelo FPQSPN, primero se lo despliega. Se aplican reglas especiales para la transición designada. Cuando la transición designada es una transición no compartida, es desplegada a un conjunto de k transiciones, donde k es el número de nodos. En cada paso. El simulador asigna el tiempo a todas las transiciones designadas.

También se tiene que dedicar una atención especial a la selección de las condiciones de los pasos de simulación. Las condiciones de los pasos tienen que ser elegidos como para no distorsionar la forma de la característica. La influencia de la condición de paso de simulación es importante sobre la confianza del modelo.

Por ejemplo, en el modelo de Red de Petri de un simple protocolo de comunicaciones de la figura n° 24, llamamos la transición designada a la transición fuente de tiempo no cero T1. El tiempo promedio de la transición de tiempo estocástico con distribución exponencial T1, que representa el ritmo de arribo promedio del proceso de Poisson, determina el punto de operación. La curva característica en la figura n° 25 representa un rendimiento y consiste de algunos valores, donde cada valor corresponde a la simulación del juego de marcas en un punto de operación dado. El simulador genera en forma automática la curva a partir del modelo del protocolo representado con FPQSPN.

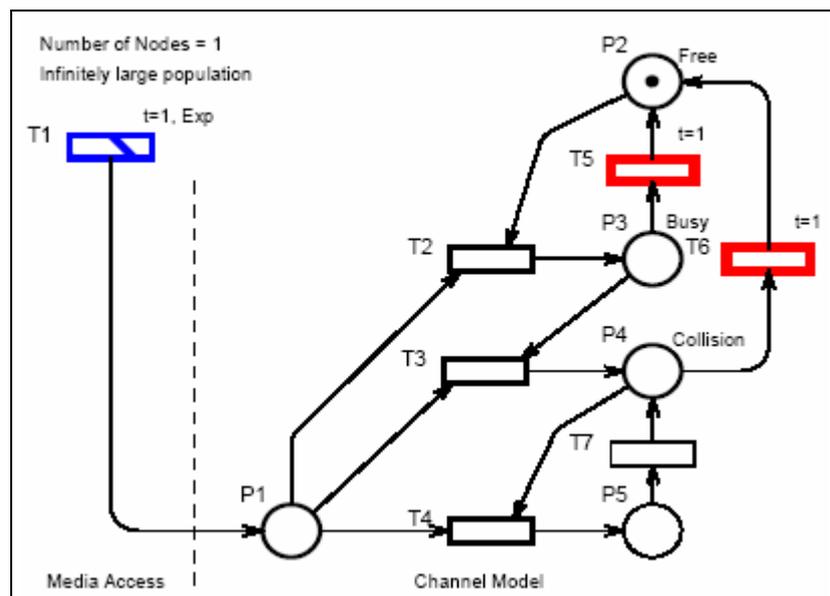


figura n° 24 Modelo FPQSPN de un protocolo de comunicación

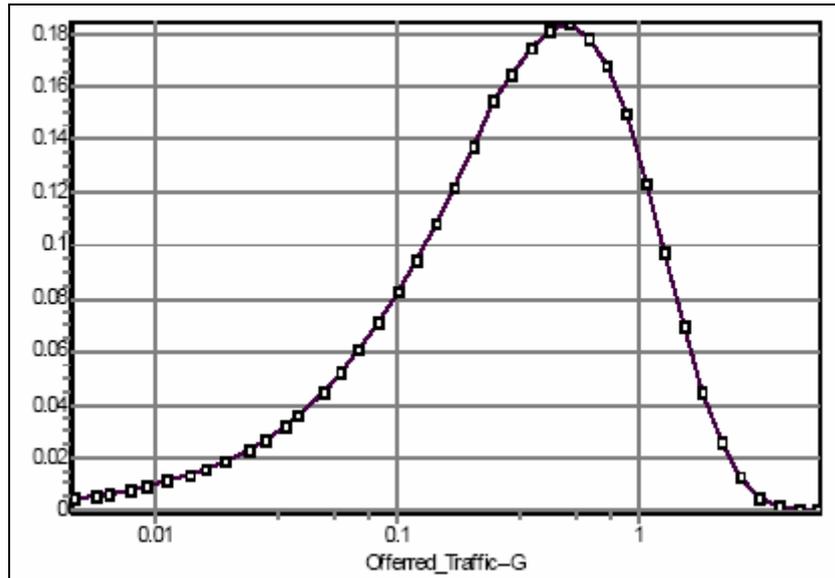


figura nº 25 Rendimiento del canal graficado como una función del tráfico ofrecido

I.8) REDES DE PETRI VERSUS REDES DE COLAS [22]

I.8.1 Introducción

Tanto en la utilización del formalismo para Redes de Petri como para Redes de Colas se encuentran varias extensiones que han sido propuestas y usadas para incrementar la potencia de modelación para la representación de los sistemas concurrentes con sincronizaciones. Han existido una serie de intentos publicados para comparar las Redes de Petri con las Redes de Colas (incluyendo la noción de tiempo), y han sugerido generalmente que existe una equivalencia sustancial entre estas dos clases de formalismos.

Tal conclusión puede ser más o menos convincente desde un punto de vista metodológico, pero en opinión del autor G. Chiola [22], sin ningún impacto práctico. Además, tal equivalencia sustancial no puede ser establecida para cualquier par de formalismos realmente propuesto, soportados por las herramientas de software relativas. En las Tablas nº 1 y 2, se presenta un sumario planteado por el mismo Chiola para la comparación, estableciendo las similitudes y diferencias, tratando de establecer o identificar algunas de las fortalezas y debilidades de tales formalismos, con la idea de que desde los ámbitos de investigación y/o desarrollo puedan proponerse algunas extensiones de los mismos que superen las dificultades.

Redes de Petri	Redes de Colas
Flujo de marcas	Flujo de clientes
Modelación abstracta	Modelación abstracta
Marcado paramétrico	Población paramétrica
Análisis operacional	Análisis operacional
Flujo de servidores equivalentes	Flujo de servidores equivalentes
Simulación discreta	Simulación discreta

Tabla nº 1 Sumario de comparación: Similitudes

Redes de Petri	Redes de Colas
Posible modelación detallada	Solo modelación abstracta
Ninguna semántica especial	Semántica predefinida
Sistemas no paramétricos	Población siempre paramétrico

Tabla nº 2 Sumario de comparación: Diferencias

I.9) COMPARACION ENTRE LOS TIPOS DE REDES DE PETRI [22]

I.9.1 Introducción

Durante los quince últimos años, se han publicado un gran número de propuestas para la utilización de los formalismos basados en Redes de Petri para la evaluación de performance. La mayoría de estas propuestas son sustancialmente diferentes una de otras desde algunos puntos de vista; el único punto de unión es el hecho que todos los formalismos introducen de alguna forma el concepto de tiempo que fue apropiadamente evitado en la definición original de las Redes de Petri.

Por lo tanto, el termino “Redes de Petri Temporizadas” ha asumido un amplio sentido para indicar una clase de modelos más que una elección específica propuesta por un autor particular. Por lo tanto, hasta ahora cada autor se ha mantenido usando su propia versión desarrollada del formalismo Redes de Petri Temporizadas más que intentar cualquier esfuerzo de unificación con otras propuestas levemente diferentes. Parte de esta falta de unificación es probablemente debido a los diferentes requerimientos de modelación que aparecen diferentes dominios de aplicación y desde diferentes formaciones culturales, más que a actitudes viciosas deliberadas.

I.9.2 Hacia un formalismo de Redes de Petri útil prácticamente

Los formalismos basados en Redes de Petri se han introducido en respuesta a las necesidades reales en el dominio de evaluación de performance de los sistemas paralelos y distribuidos, donde no haya mecanismos incluidos en las Redes de Colas.

Las variaciones históricas de las Redes de Petri Temporizadas derivan de:

- La gran flexibilidad de formalismo de Redes de Petri en si mismo, que puede aconsejar diferentes reinterpretaciones y extensiones,
- El uso del formalismo dentro de diferentes dominios de aplicación,
- La elección de diferentes métodos y técnicas de modelación y análisis.

Ninguno de los formalismos de Redes de Petri temporizados propuestos ha sido específicamente diseñado como un formalismo para sustituir las Redes de Colas. Además, algunos de ellos fueron diseñados para complementar las descripciones de las Redes de Colas para aplicaciones grandes y complejas.

Las Redes de Petri permiten tanto la modelación a nivel de red y a nivel de estado, usando el mismo formalismo. Este provee libertad en la definición que debe ser expresada por la estructura de la red y que puede se expresada por la interpretación. Esta libertad puede ser explotada para sintonizar la potencia de modelación del formalismo a las necesidades de diferentes dominios de aplicación.

Desde el punto de vista de la modelación a uno le gustaría expresar todos los mecanismos básicos que caracterizan un dominio de aplicación dado por la interpretación, de modo que la estructura del modelo se mantenga simple.

I.9.3 Formalismos de Redes de Petri.

En la Tabla n° 3, se presenta un sumario para la comparación entre diferentes formalismos de Redes de Petri, estableciendo sus similitudes y diferencias, tratando de establecer o identificar algunas de las fortalezas y debilidades de tales formalismos para su aplicación en el ámbito del análisis de los Protocolos de comunicaciones.

Redes de Petri	Propiedad/Objetivo
No temporizada (NP)	Modelo Causal sin tiempo
Temporizada Estocástica (SPN)	Modelo con Interpretación Temporal y Política de Resolución de Conflicto
Temporizada Markoviana (MPN)	Modelo Particular de SPN en que los tiempos de servicios de transición son variables aleatorias independientes y con distribución de probabilidad exponencial y tasa de encaminamiento de política de competición
Temporizada Generalizada (GSPN)	Modelo Particular de MPN que incluye transiciones inmediatas
Temporizada Coloreada (CPN)	Modelo Compacto de Redes de Petri Complejas
Temporizada FPQSPN	Modelo Compacto de GSPN de Sistemas vistos como Colas.

Tabla n° 3 Sumario de comparación: Similitudes

I.9.4 Formalismos de Redes de Petri y los protocolos de comunicaciones.

En distintos trabajos se han usado las Redes de Petri para el análisis de los protocolos de comunicaciones. Sus aproximaciones difieren en la clase de Red de Petri que ellos han usado y también en los métodos de análisis. Y dependiendo sobre el método de análisis, los autores han validado protocolos de comunicaciones o han verificado algunas propiedades. Los autores difieren en si usan Redes de Petri de bajo nivel o de alto nivel. Las Redes de Petri Coloreadas (CPN) han demostrado tener una gran potencia de modelación para los protocolos de comunicaciones, pero los modelos son menos transparentes debido a las expresiones matemáticas asociadas a los objetos de la Red de Petri. Además, en CPNs las herramientas de simulación existentes, los modelos no se convierten en Redes de Petri de bajo nivel antes de la simulación, sino en un lenguaje de programación especializado [23]. Otros autores [24] han usado Redes de Petri Estocásticas o Generalizadas para la modelación y análisis de ciertas clases de sistemas de colas

I 10) CONCLUSION

Como fue planteado en los objetivos y la metodología a seguir en el trabajo de tesis, se han compilado las novedades teórico-prácticas de las Redes de Petri, y se ha seleccionado el tipo más apto de Red de Petri para la modelación, entre otros aspectos. Para ello, se ha efectuado una compilación de información sobre características de distintos tipos de Redes de Petri con aplicación práctica en general, y específicamente al modelado de los protocolos de comunicaciones.

En este Capítulo se han presentado los aspectos teóricos sobre Redes de Colas, las Redes de Petri No Temporizadas y Temporizadas. Entre las Redes de Petri Temporizadas se han planteado las técnicas de reducción del indeterminismo, y los diferentes tipos de Redes de Petri Temporizadas. Además, se han planteado las propiedades lógicas y análisis de las Redes de Petri, y las técnicas generales para el cálculo de índices útiles desde su observación el tiempo. Finalmente, se han presentado las tablas comparativas de las Redes de Petri y las Redes de Colas, y de las Redes de Petri entre sí.

A los efectos de los objetivos del trabajo de tesis, se han seleccionado las FPQSPN (Finite Population Queuing Systems Petri Nets). Se ha presentado este tipo de Red de Petri y sus características asociadas, propuesto desde sus orígenes para su aplicación el área de los protocolos de comunicaciones. FPQSPN es una clase de Redes de Petri que es apropiada para modelación simplificada de sistemas de cola con población finita. Un modelo de Red de Petri de un sistema de cola con población finita, típicamente consiste de una parte compartida y una parte no compartida, que representa la población. En FPQSPN, todos los clientes son representados por un simple subdiagrama, que es desplegable en k submodelos antes de la simulación.

La ventaja principal para su adopción es que FPQSPN permite modelación eficiente de los sistemas de cola de población finita, y mantiene la transparencia de las Redes de Petri de Bajo Nivel. Además, a la propuesta de FPQSPN se asocia un algoritmo eficiente para el cálculo computacional de rendimiento y graficación de funciones de algunas variables, usando representaciones matriciales.

Estas características de FPQSPN lo hacen especialmente apropiado para la aplicación en el estudio de protocolos de comunicaciones, y por lo tanto para.

- Un enfoque alternativo para la enseñanza-aprendizaje de los protocolos de red de canal compartido usando FPQSPN, los modelos para cada caso, y su simulador asociado,
- La disponibilidad de resultados complementarios de origen analítico y vía la modelación con Redes de Petri FPQSPN,
- La disponibilidad de herramientas software de simulación para la enseñanza con animación de la dinámica y el cálculo de rendimiento de los protocolos,
- La obtención de resultados prácticos para la determinación de la confiabilidad y verificación de los modelos, al contrastarlo con los métodos analíticos,
- La generación de una unidad de enseñanza basada en Redes de Petri.

CAPITULO II

MODELACION DE PROTOCOLOS DE COMUNICACIONES

II.1) PROTOCOLOS DE RED (definición general de los mismos)[21][25][26]

II.1.1 Introducción

El análisis de performance de los protocolos de comunicación es un esfuerzo para cuantificar la conducta de la red y representarla a través de algunas medidas numéricas significativas. Puede definirse el análisis de performance como el desarrollo y estudio de los modelos matemáticos que predicen la performance de las redes en algún sentido bien definido.

Desde un punto de vista de análisis de performance, hay al menos 4 elementos principales que deberían ser incorporados en el modelo:

- El protocolo de red es un conjunto de convenciones y standards que gobiernan el tratamiento y el formato de datos en todas las partes de la red. Por ejemplo, el método de acceso al medio define las reglas para compartir recursos facilitados por la capa física. La selección de un método MAC apropiado es vinculado con las propiedades físicas del canal. Así, los métodos CSMA proveen buen rendimiento sobre los canales terrestres, no son utilizables para canales satelitales debido a su retardo de propagación intrínseca de cerca de 0.26 seg., lo que plantea una diferencia importante entre los canales de radio satelitales de los enlaces terrestres.
- La topología de la red y la configuración de la red son una descripción gráfica del arreglo de los componentes de la red diferentes y sus interconexiones. La teoría gráfica es una herramienta potente para describir la topología de redes metropolitanas y WAN, lo que típicamente consiste de una combinación de diferentes estructuras topológicas básicas. Las topologías de red básicas como la bus o la estrella no pueden describir la topología de redes modernas, pero son aún aplicables para análisis de métodos MAC no determinísticos.
- Las medidas de performance.
- Finalmente, las características del tráfico de paquetes de la red han sido objeto de estudio considerable en la literatura [27]. Varios modelos estocásticos de tráfico de red han sido propuestos. No solo en el análisis de cola, el modelo de tráfico es llamado el proceso de arribo.

II.1.2 Medidas de performance de los Protocolos de Comunicación

La idea general del análisis de performance de una red es cuantificar la conducta de la red en términos de algunas medidas numéricas significativas. La clave para definir una medida de performance general es considerar la red o el dispositivo de red, como un sistema de cola abierta general, como se muestra en la figura n° 26 (y luego asociar el modelo FPQSPN).

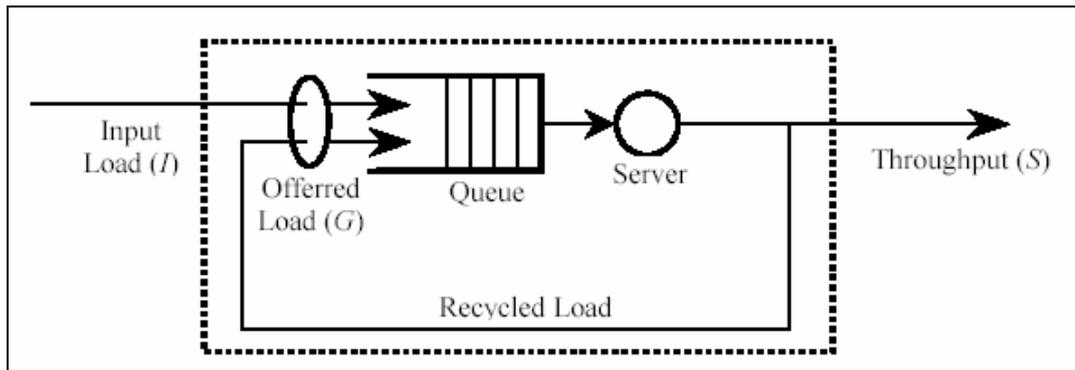


figura n° 26 Un modelo de cola general para una red o dispositivo de red

Dependiendo del modelo de red específico o la operación del dispositivo, el sistema de cola tiene un “proceso de arribo”, un “proceso de partida” y un “proceso de reciclado”, un tamaño de cola y una disciplina de servicio (protocolo MAC), etc. Podemos identificar dos parámetros que afectan la performance del sistema para esta representación muy general:

- carga de entrada (I) es el número promedio de nuevos trabajos que arriban al sistema desde el exterior del sistema sobre un intervalo de tiempo,
- carga ofrecida (G).

Los arribos, que representan usuarios demandando un servicio, pueden ser un proceso con recuperación o sin recuperación. Los procesos con recuperación implican que las demandas de usuarios insatisfechos demandan el servicio nuevamente después de algún periodo de tiempo (el tiempo de recuperación), y la aproximación sin recuperación considera que los usuarios insatisfechos se pierden. En los modelos sin recuperación, la carga reciclada (paquetes replanificados) es considerada ser paquetes arribados nuevamente. Tales sistemas no tienen definido la carga de entrada. El tráfico ofrecido es considerado ser la carga de entrada. Desde el punto de vista de la capa MAC, un paquete puede ser retransmitido a causa de que el canal estaba ocupado o hubo una colisión cuando el nodo experimentó transmitir el paquete, o un receptor no recibió correctamente el paquete debido a un error sobre el medio.

Cuando la capa MAC tiene implementado una unidad de detección de colisión o error, retransmite el paquete. De otro modo, las capas superiores del modelo OSI de ISO demanda la retransmisión del paquete. Los protocolos de comunicación con detección de colisión, como el CSMA/CD, son frecuentemente modelados como sistemas de cola con recuperación.

II.1.2.1 Carga ofrecida

La carga ofrecida (G) es el número promedio de trabajos ofrecidos al sistema de cola sobre un intervalo de tiempo. Cuando los paquetes tienen longitud constante T , el intervalo de tiempo es considerado ser T . El tráfico ofrecido al canal consiste no solo de los nuevos paquetes, sino también de los paquetes previamente colisionados.

La carga ofrecida es algunas veces definida como la fracción del ancho de banda de la red, que el nodo usaría si tuviera completo acceso a la red. La carga ofrecida es también definida como el número de hosts esperando por la transmisión.

II.1.2.2 Rendimiento

Se define rendimiento (S) como la fracción del ancho de banda de la red nominal que es usada por los datos transmitidos exitosamente. También puede definirse, para una longitud constante de los paquetes, como el número promedio de paquetes exitosamente transmitidos por el tiempo de transmisión de paquete t.

El rendimiento es definido por algunos autores por la utilización del canal, y otros definen la utilización del canal como el rendimiento expresado en porcentaje. Finalmente, le llaman rendimiento obtenible máximo, para un modo de acceso dado, a la capacidad del canal.

El rendimiento es también definido como el número promedio de trabajos que dejan los límites del sistema de cola, sobre un intervalo de tiempo. El intervalo de tiempo puede ser una medida de tiempo standard, o un periodo de modelo específico como el tiempo medio de transmisión de paquete. Debe notarse que en estado estable, la carga de entrada y el rendimiento son iguales si ningún trabajo es descartado. De otro modo, los paquetes se acumularían en el sistema. En las redes de computadoras o dispositivos de red, donde el trabajo es típicamente un paquete de dato, el rendimiento es usualmente expresado en paquetes por unidad de tiempo. Para paquetes de tamaño fijo, o cuando la distribución del tamaño de paquetes es conocida, el rendimiento en bits por unidad de tiempo es justamente otra forma de representar la misma cantidad. En algunos sistemas comerciales, el rendimiento significa el número de unidades transferidas (bits, bytes, datagramas) por unidad de tiempo.

II.1.2.3 Ancho de banda

El ancho de banda es el rendimiento máximo teórico bajo métodos de control de acceso al medio ideales, sin solapamientos o espacios entre paquetes. El rendimiento teórico máximo es algunas veces diferente a la capacidad del canal.

II.1.2.4 Retardo

El retardo, o también llamado tiempo de transporte, es el intervalo de tiempo promedio desde el momento que un nuevo trabajo entra al sistema, hasta el momento que el trabajo deja el sistema. Otra medida es el tiempo de cola (tiempo de espera o tiempo de espera de cola), definida como el tiempo promedio que un trabajo gasta en la cola. Por ejemplo, para un segmento Ethernet, este es el tiempo que un paquete que arriba debería esperar antes de encontrar una chance para la transmisión. En Aloha puro el tiempo de espera es siempre cero. Para una red libre de colisiones, el tiempo de servicio es la suma del tiempo de transmisión del paquete (T) y el tiempo de propagación (el tiempo que necesita un paquete para propagarse sobre la red). Puede definirse el retardo (D) como el retraso promedio entre el tiempo desde que un paquete es generado y el tiempo cuando el es exitosamente recibido. D es una suma del tiempo de espera y el tiempo de servicio.

$$D = (G/S - 1).R + 1 + a$$

El tiempo de servicio es una suma del tiempo de transmisión del paquete T, considerado ser 1, y el retardo de propagación a. La probabilidad de transmisión exitosa de paquete está definido en términos de S y G como $P_s = S/G$. El tiempo de espera está dado por $R*(G/S - 1)$, donde R es el tiempo medio entre dos transmisiones consecutivas y la expresión en paréntesis es el número promedio de retransmisiones.

En los métodos MAC, que no son capaces de detectar colisiones, el paquete deja el sistema cuando la transmisión comienza, pero por ejemplo en Ethernet, la transmisión del paquete puede ser abortada si una colisión es sensada. Así, un paquete puede pasar algunas veces el límite del sistema. Una medida de tiempo útil para Ethernet, es el intervalo de tiempo entre el momento que un trabajo entra al servidor por primera vez, y el momento que él deja el límite del sistema. En un segmento Ethernet esto corresponde al tiempo de adquisición del canal, también llamado tiempo de acceso.

La varianza de retardo, también llamada el jitter de retardo, es también una medida importante en la aplicaciones de tiempo real.

II.1.3 Características del tráfico de red

Varios de los estudios existentes sobre performance LAN están basados sobre modelos de tráfico simplificados. Los procesos de Poisson, un simple modelo matemático usado para describir el arribo de clientes en las redes de telefonías de circuito conmutado, es aún una suposición subyacente en los modelos de tráfico usados en varios de los estudios sobre redes de datos de paquetes conmutados. Varios de los estudios también asumen paquetes de longitud fija, una suposición que está lejos de la realidad en una LAN moderna, que está típicamente distribuida en ambientes de computación. Los protocolos de control de flujo, usados en las redes de computadoras modernas, producen distribución bimodal de longitudes de paquetes debido a los paquetes de reconocimiento.

Varios estudios han sugerido el uso de los Procesos de Poisson Modulados de Markov (MMPP – Markov Modulated Poisson Process) [28]. Este es un proceso doblemente estocástico en el cual los arribos ocurren en la forma Poisson, con un ritmo que varía de acuerdo a una cadena de Markov de estado finito, independiente del proceso de arribo. Entre otros modelos, están los Procesos de Poisson Conmutados Generalizados (GSPP – Generalized Switched Poisson Process) [29], algunos casos especiales como los Procesos de Poisson Conmutados (SPP – Switched Poisson Process) y los Procesos de Poisson Interrumpidos (IPP – Interrupted Poisson Process).

II.1.3.1 Influencia del Modelo OSI de ISO para el Proceso de Arribo de Paquetes

La distribución de la carga de red depende fuertemente sobre el tipo de capas superiores del modelo OSI. Para el tráfico de área amplia WAN, los procesos de Poisson son válidos sólo para modelar el arribo de sesiones de usuarios (conexiones Telnet, conexiones de control FTP, etc) que fallan como modelos exactos para otros procesos de arribo WAN. Para pequeñas transferencias de ráfaga máquina generadas, tales como SMTP y NNTP, los procesos de arribo de conexión no son bien modelados con Poisson, lo que no es sorprendente dado que ambos tipos de conexiones son iniciados por máquina y pueden ser activados con timers.

Así, los procesos que describen los arribos de los usuarios demandando algún servicio de aplicación son procesos aleatorios, los cuales pueden ser bien aproximados por Procesos de Poisson. Cuando el usuario trabaja con la aplicación, la aplicación genera demandas a las capas inferiores del modelo OSI. El proceso que describe los arribos de demandas a las capas inferiores, puede ser visto como el proceso de usuario por la transformación de la capa de aplicación. La transformación corresponde al tipo de aplicación. Igualmente, por ejemplo, la transformación de la capa de sesión cambia con cada tipo de esquema de la capa de sesión. La capa OSI puede sumar su propio proceso de arribo determinístico o estocástico, independiente del tráfico de las capas superiores. El tráfico propio de las capas es generado por servicios de

protocolos de mantenimiento, y deberían ser sumados a los procesos transformados desde la capa superior.

El tráfico real sobre los canales de comunicación de las redes de hoy en día no le corresponde a una distribución de variables aleatorias. La mejor forma de modelar el proceso de tráfico es descomponer el tráfico de acuerdo a las capas del modelo OSI. Otra aproximación es encontrar la mejor aproximación no determinística.

II.1.4 Modelación de Redes de Petri de los Procesos de Arribo de Paquetes

Pueden plantearse varios tipos de tráficos de red, y mostrar como modelar algunos tipos de tráfico de paquetes usando Redes de Petri.

II.1.4.1 Proceso de Arribo de Paquetes Determinístico

El proceso determinístico de arribo de paquetes se aplica para tráfico generado por un proceso automático, por ejemplo, paquetes de información en algunos protocolos de ruteo. Otro ejemplo de tal tráfico puede ser el tráfico generado por un sistema de control de tiempo real. Tal tipo de tráfico de paquete puede ser modelado con las Redes de Petri FPQSPN con una transición fuente de tiempo determinística. El tiempo asociado a la transición representa el tiempo entre dos arribos de paquetes consecutivos.

II.1.4.2 Proceso de Arribo de Paquetes Poisson

Se modela el proceso de arribo de paquetes Poisson por una transición con distribución exponencial. El tiempo medio asociado a la transición representa el valor medio de tiempo entre dos arribos de paquetes. La secuencia de valores de tiempo, asociada con una cierta secuencia de disparo, representa los tiempos interarribo de paquetes. El número de paquetes generados por unidad de tiempo, está dado por la distribución Poisson. Si asumimos que cada usuario genera al menos un requerimiento por unidad de tiempo, la frecuencia de paquetes generados significa el número de usuarios demandando servicio por unidad de tiempo.

II.1.4.3 Proceso de Arribo de Paquete Poisson con Recuperación

En los sistemas de cola, los usuarios demandando el servicio son satisfechos sin retardos, o son colocados en una cola, o en un sistema de colas de recuperación, donde reintentan el servicio nuevamente después de algún periodo de tiempo. Los usuarios reintentan el servicio nuevamente, cuando el sistema tiene cola de capacidad finita, la que está actualmente llena o cuando el usuario temporalmente cancela el requerimiento. Algunos métodos de acceso al medio, como CSMA, replanifican los paquetes cuando el canal está ocupado. Algunas medidas de performance y carga consideran a los reintentos de los usuarios como nuevos requerimientos arribados y algunos de ellos no.

En las transiciones T1, T5 y T11 de la figura n° 27, representan procesos de arribo de usuario. Las transiciones T2, T6 y T14 representan procesos de arribos demandados a la capa física. En el ejemplo de un sistema sin recuperación, un usuario espera en cola (lugar P1) hasta que él es satisfecho. Otra opción es que el usuario insatisfecho deje el sistema. En tal caso, el usuario es considerado ser un usuario arribado nuevamente después de su re-arribo. Finalmente, la última opción es que el usuario espera en el lugar P14 por un periodo de tiempo aleatorio cuando el encuentra el sistema ocupado. Luego, el usuario retorna al lugar P14. Cuando se usan modelos de protocolos de comunicaciones, las transiciones T12 y T14 nunca están en conflicto.

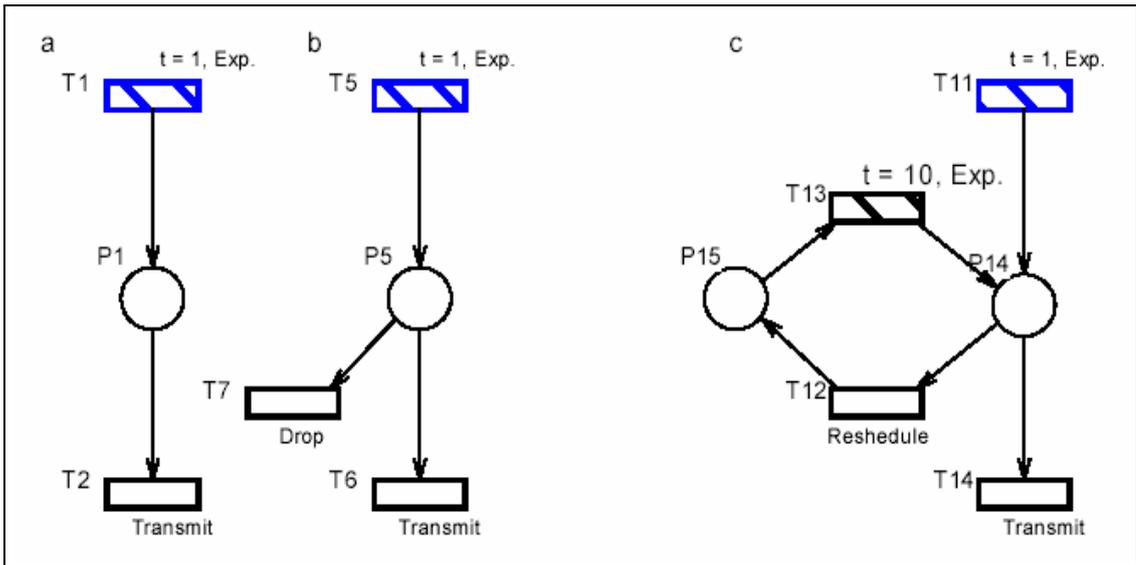


figura n° 27 Procesos de arribo Poisson en Redes de Petri

II.1.4.4 Proceso de Arribo de Paquetes Poisson Modulado con Markov

Un proceso de Poisson modulado con Markov (MMPP) de orden NM consiste de una Cadena de Markov de estado NM, en la cual cada estado i ($i = 1, 2, \dots, NM$) está asociado con un proceso de Poisson de ritmo dado. En otras palabras, MMPP es un proceso de Poisson variable en el tiempo, cuyo ritmo es cambiado (modulado) de acuerdo a la cadena de Markov.

Un ejemplo de MMPP es un proceso que conmuta entre periodos de silencio y periodos en el cual los requerimientos arriban con un ritmo dado. Tal ejemplo se aplica para modelar sistemas de población finita, como las redes de computadoras reales, cuando los nodos tienen buffers de salida de capacidad uno. En tal caso, un nuevo paquete no puede arribar cuando hay un paquete en el buffer de salida del transmisor. En la figura n° 28 se muestra los modelos MMPP de dos estados.

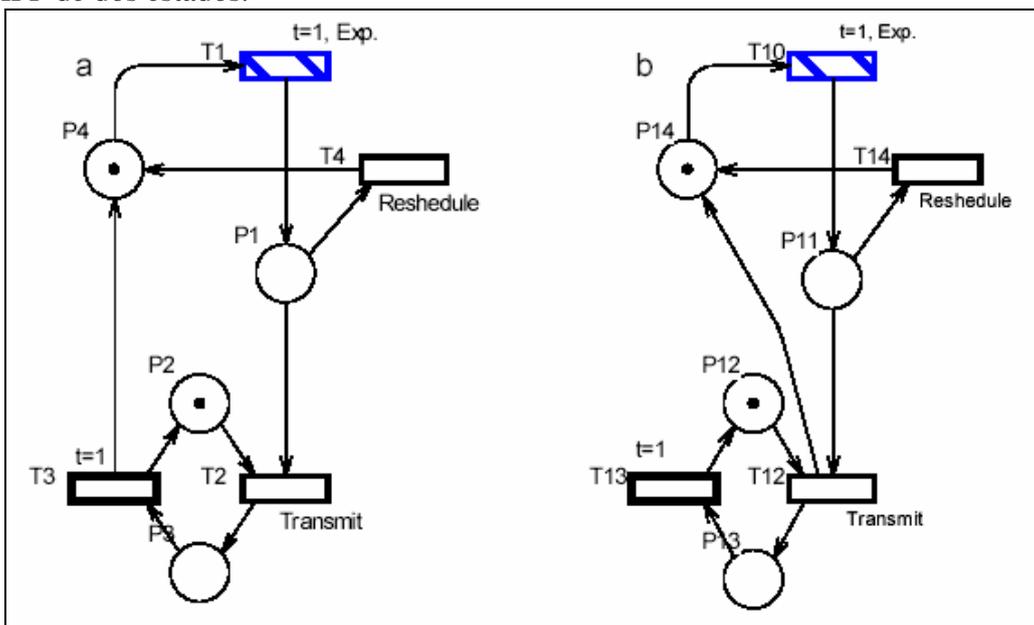


figura n° 28 Arribos MMPP de dos estados en FPQSPN

En la literatura se espera que el buffer de transmisión esté lleno hasta que la transmisión del paquete ha finalizado. El modelo de Red de Petri de tal proceso de arribo se presenta en la figura a. La cola de salida tiene capacidad cero. Eso significa que la única cola es el buffer de transmisión de capacidad uno. Después que la transición T1 es disparada (nuevo arribo a la QS), una marca es removida desde el lugar P4, y sumada al lugar P1. El proceso de arribo es deshabilitado hasta que el requerimiento es transmitido o replanificado (transiciones T2 o T4 se disparan). La transición inmediata T2 representa el comienzo de la transmisión del paquete, y el tiempo de la transición de tiempo determinístico T3 corresponde al tiempo de finalización de la transmisión. Un ejemplo de un sistema de cola con tal proceso de arribo es un sistema G/D/I (proceso de llegada con distribución general, proceso de tiempos de servicio con distribución determinística, y 1 servidor)

El proceso de arribo en la figura b es similar al de la figura a, con excepción que un nuevo paquete puede arribar también durante la transmisión del previo (la capacidad de la cola de salida y del buffer de transmisión son iguales a uno).

II.1.4.5 Proceso de Arribo de Paquetes Poisson con Recuperación Modulado con Markov

Los modelos de la figura n° 29 difieren de los de la figura previa en los procesos de recuperación, los cuales retornan los clientes insatisfechos a la cola. Aquí un paquete es retransmitido hasta que el es exitosamente transmitido (algunos sistemas, como Ethernet realizan solo un número limitado de retransmisiones). El nodo de la figura a nunca acepta un siguiente paquete hasta que finaliza la transmisión del previo. El nodo experimenta poner un paquete inmediatamente después de su arribo. Cuando el canal está ocupado, el nodo experimenta transmitir después de un periodo de aleatorio dado por el tiempo distribuido exponencialmente de la transición T4. En el sistema de la figura b, un nuevo paquete puede arribar durante los reintentos de paquetes no transmitidos previamente. El nodo tiene una cola de salida de capacidad infinita. Notamos que las transiciones T2, T5 y T15 tienen también otras transiciones de entrada des los lugares indicando el estado del canal (servidor en SQ).

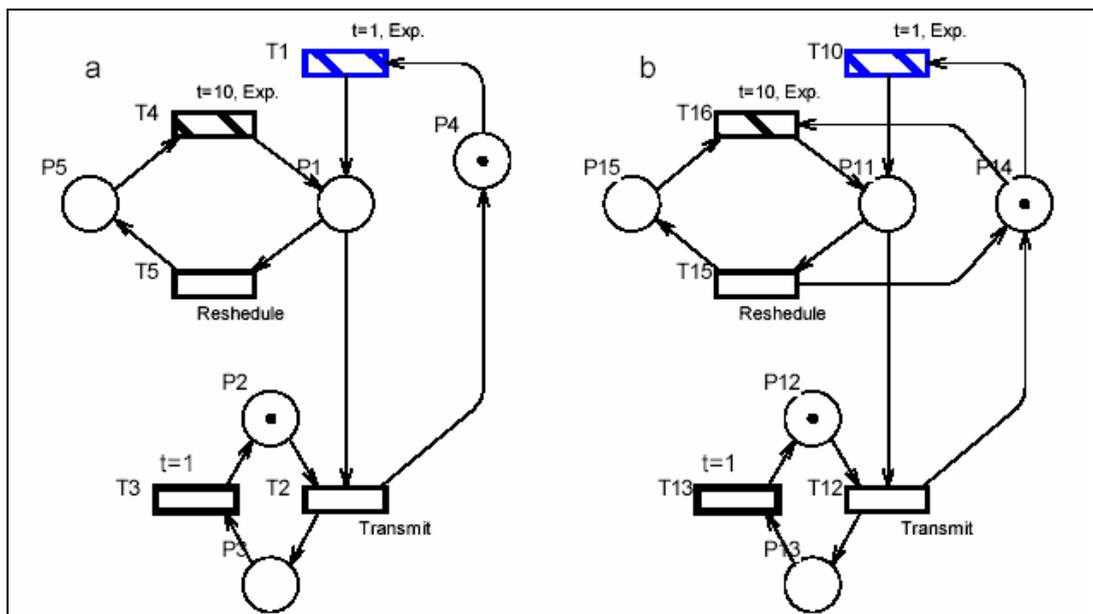


figura n° 29 Arribos MMPP de dos estados con recuperación en FPQSPN

II.2) MODELACIÓN DE ALOHA

II.2.1 Introducción

Aloha, en su versión pura y ranurada, ha sido el primer esquema de acceso aleatorio desarrollado en ARPA del Departamento de Defensa, a los finales de los 60 y el principio de los 70. Aloha fue desarrollado para usarse en comunicaciones wireless y satelitales.

El análisis teórico basado sobre la teoría de colas toma ventaja de la “ley de números grandes”. La ley establece que con cada alta probabilidad, la demanda de la población en cualquier instante de tiempo iguala aproximadamente a la suma de las demandas promedio de la población. Gracias a esta suposición, podemos aproximar los arribos de paquetes desde la población total por un simple proceso de arribo de paquete (por ejemplo, Poisson) con su correspondiente valor medio. De acuerdo a la notación de los sistemas de cola, se considera al canal Aloha como un sistema de cola M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor)

El tráfico ofrecido al canal consiste no sólo de nuevos paquetes, sino también de los paquetes previamente colisionados. Basados en las suposiciones hechas previamente, el valor de G ($G = \lambda.T$) representa también el número de usuarios experimentando transmitir sus paquetes en el tiempo (λ es el ritmo medio de los arribos al sistema de cola). Además, desde la literatura se han hecho las siguientes suposiciones:

- La fuente de tráfico consiste de un número infinito de usuarios quienes colectivamente forman una fuente de Poisson independiente con un ritmo de generación de paquetes medio agregado de λ paquetes por segundo.
- Cada paquete tiene longitud constante demandando T segundos para la transmisión. S (en estado estable $S = G$) es el número promedio de paquetes transmitidos por tiempo de transmisión (es el ritmo de entrada normalizado a T).
- Si un nodo usa algún esquema de reconocimiento en las capas superiores del modelo OSI, el paquete de reconocimiento es transmitido después de un tiempo relativamente grande comparado con T .
- Los tiempos interarribos del proceso definidos por los tiempos de arranque de todos los paquetes, incluyendo retransmisiones, son independientes y distribuidos exponencialmente.
- Cada usuario en la población finita se asume tiene a lo sumo un paquete requiriendo transmisión en cualquier momento (incluyendo cualquier paquete bloqueado previamente). Así, el usuario nunca transmite dos o más paquetes consecutivamente sin pausa.

Bajo condiciones de estado estable (todos los paquetes generados por una colección de usuarios son exitosamente transmitidos), S puede ser considerado como un ritmo de rendimiento de canal, o utilización de canal (la capacidad máxima teórica de canal es $1/T$ paquetes por segundo cuando $S = 1$).

II.2.2 Aloha puro

En Aloha puro los usuarios pueden transmitir en cualquier momento que ellos lo necesiten. Si ellos fallan al intentar escuchar su transmisión exitosa, después de un retardo de propagación, ellos saben que ocurrió una colisión y retransmiten el paquete después de retardos de retransmisión aleatorios.

II.2.2.1 Modelación de Aloha puro

Se asume que los tiempos interarribos del proceso definidos por el tiempo de arranque de los paquetes, incluyendo las retransmisiones son distribuidos exponencialmente independientes, de modo que los tiempos interarribos entre paquetes consecutivos forma una variable aleatoria distribuida exponencialmente con media $1/G$.

En la figura n° 34, se nota que hay un sistema de cola con población infinitamente larga, pero el parámetro “Número de Nodos” igual a uno. Esto es así, a causa de que este parámetro es una propiedad de las Redes de Petri. Este parámetro aparecerá en la mayoría de los modelos. Hay una aproximación diferente en nuestro análisis de los sistemas de cola de población infinita y finita. En los modelos donde consideraremos una población finitamente grande, el parámetro está relacionado al tamaño de la población. Cuando este parámetro es igual a uno, sabemos que no usamos mecanismos de modelado y desplegado provistos por FPQSPNs.

II.2.2.2 Canal de Comunicaciones

Desde el punto de vista de Aloha, consideraremos que al canal como un dispositivo que tiene una o más entradas y ninguna salida. La salida del canal no es un objeto del análisis de performance de protocolos. Todas las entradas del canal son asumidas idénticas, y el tiempo de transmisión de paquete, el cual está dado en la práctica por la longitud del paquete y por la velocidad en baudios del canal, es una constante.

Como mencionado arriba, un paquete es exitosamente transmitido, cuando el valor actual de la variable aleatoria que representa el tiempo interarribo entre paquetes iguala al menos al intervalo vulnerable. Si cualquier otro paquete es ofrecido durante el intervalo vulnerable ocurre una colisión. Si el canal está en estado de colisión en el tiempo t_i , y luego un paquete es ofrecido al medio, la colisión persiste hasta el tiempo t_i+1 . El intervalo vulnerable iguala al tiempo de transmisión de paquete bajo una condición para todos los nodos, que nunca comience su transmisión cuando su parte receptora está recibiendo un paquete (asumimos que el retardo de propagación a es constante para todos los nodos). De otro modo, el intervalo vulnerable es una suma de retardo de propagación a , dado por propiedades físicas del canal, y el tiempo de transmisión del paquete.

El proceso de arribo de paquetes en el modelo de Red de Petri de Aloha puro de la figura n° 30, es representado por una transición de tiempo de fuente estocástica T_1 , a causa de que la transición es un generador de una variable aleatoria distribuida exponencialmente. Las transiciones T_5 y T_6 son transiciones de tiempo determinístico con $t = 1$. El tiempo t_5 de la transición T_5 representa el tiempo de transmisión de paquete, y el tiempo t_6 de T_6 representa el retardo entre el tiempo de arribo del último paquete participando en una colisión, y el instante en que la colisión se resuelve. Una marca en el lugar P_2 representa un estado libre (inactivo) del canal, el lugar P_3 representa la transmisión de un paquete, y el lugar P_4 representa una colisión. Los lugares P_2 , P_3 , P_4 y P_5 forman una p-invariante (componente conservativo) S_i . El lugar P_5 y la transición T_7 evitan el autoloop entre P_4 y T_4 . Por conveniencia el lugar P_5 y la transición T_7 son presentadas justo para asegurarse que el modelo opere correctamente en todos los simuladores. En el caso del disparo de una transición autoloop, los simuladores deberían resetear los timers de las transiciones temporizadas que están en conflicto con la transición autoloop. Para asegurarnos que nuestros modelos funcionen apropiadamente aún en los casos cuando los simuladores no hagan esto, evitaremos los autoloops en la mayoría de los modelos presentados.

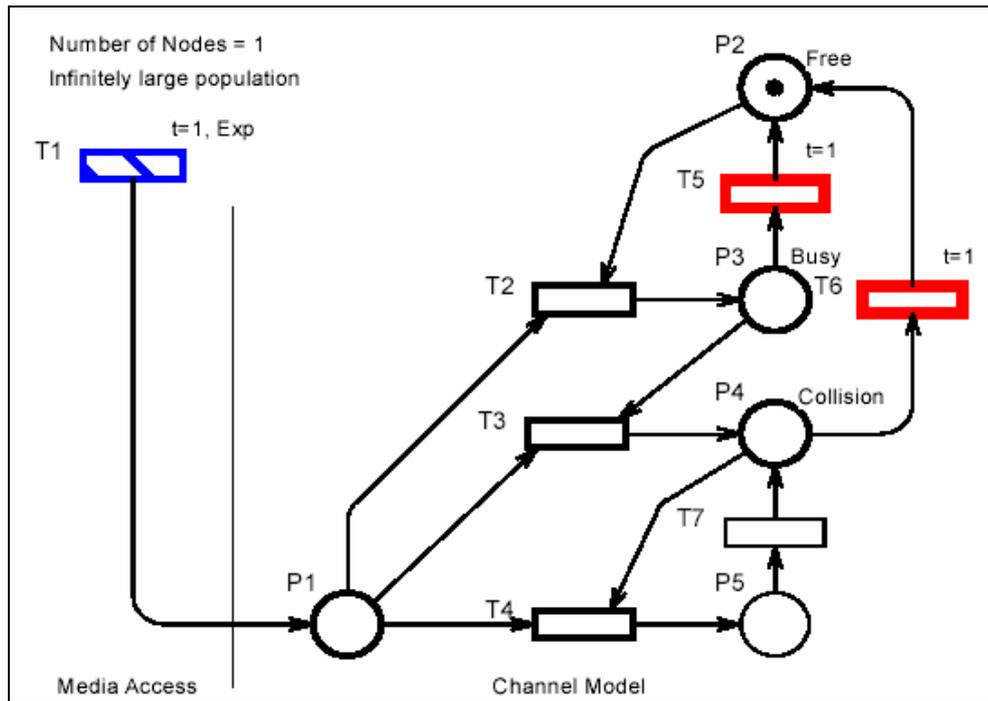


figura n° 30 Modelo FPQSPN del protocolo Aloha Puro

Cuando ofrecemos un paquete a un canal libre (en el modelo ponemos una marca en el lugar P1), la transición T2 se vuelve habilitada. Después del disparo, la transición consume la marca desde P1 y produce una marca en el lugar P3. La marca reside en P3 hasta que la transición de tiempo determinístico T5 es disparada, o hasta que un nuevo paquete arriba. En este último caso, la transición T3 es disparada, y el canal cambia su estado a colisión. Si ningún otro paquete arriba durante el tiempo t6, el canal retorna al estado libre (la transición T6 es disparada), de otro modo la transición T4 se dispara y la marca en el lugar P4 “es movida a” P5. En este momento, la transición T6 está deshabilitada y su tiempo es inicializado después del disparo de T7.

II.2.3 Aloha ranurado

En Aloha ranurado, el tiempo es dividido en segmentos, cuya duración iguala exactamente al tiempo de transmisión de paquete T. La transmisión de paquete siempre comienza al principio de un slot de tiempo. La división del tiempo en segmentos causa que la probabilidad de ocurrencia de colisiones se reduzca y una colisión afecte siempre un intervalo de tiempo de longitud T. Dos o más paquetes se solapan completamente en caso de una colisión.

II.2.3.1 Propiedades comunes

Importantes parámetros del canal son el número de transmisiones por transmisión exitosa (G/S), y el factor crucial $(G/S - 1)$, que es el número de retransmisiones por transmisión exitosa. El tiempo entre la transmisión de un paquete y la recepción del correspondiente reconocimiento está dado por la fórmula:

$$R = 1 + 2a + \delta + \alpha$$

Además, llamaremos a D el tiempo promedio desde que un paquete es generado hasta que es exitosamente recibido:

$$D = (G / S - 1)(R + d) + d + 1 + a$$

Donde α denota el tiempo de transmisión del paquete de reconocimiento que es usualmente transmitido a través de un canal separado, y dejamos denotar a δ al retardo de propagación de un camino. En las fórmulas anteriores se asume que los paquetes de reconocimiento son siempre correctamente recibidos, y el retardo de procesamiento de los paquetes de reconocimientos es despreciable. Finalmente, δ significa el retardo de transmisión promedio y d iguala a 0.5 para Aloha ranurado, y de otro modo d iguala a cero. La capacidad del canal máxima de Aloha pura es 0.1839 ($G = 0.5$), y la capacidad del canal máxima de Aloha ranurado es 0.3678 ($G = 1$).

II.2.3.2 Modelación de Aloha ranurado

La parte derecha del modelo de Red de Petri de Aloha ranurado en la figura n° 31 representa un canal de comunicaciones idéntico al de Aloha puro. El lugar $P6$ representa un buffer de transmisión inmediata. Se nota que hay usualmente múltiples buffers en un nodo (llamamos buffer inmediato a aquel buffer desde el cual el dato es directamente transmitido al canal). Las transiciones $T8$, $T10$, $T11$, $T12$, y los lugares $P6$, $P9$, $P10$ y $P11$ forma un “dispositivo ranurador” que asegura que la transmisión sea siempre realizada en el comienzo del tiempo de slot. Las transiciones $T11$ y $T10$ son transiciones de tiempo determinístico. Los tiempos $t11 = 0.00001 + t10 = 1$ forman un retardo de longitud 1.00001, que iguala al tiempo de transmisión de paquete más un pequeño retardo asegurando que un conflicto no ocurra exactamente al final de la transmisión. El tiempo asociado a la transición $T11$ actúa también como una función de prioridad (la transición $T5$ tiene una prioridad sobre $T11$) que asegura que todos los paquetes, representados por las marcas en el lugar $P6$ son transmitidos al comienzo de los slots de tiempo.

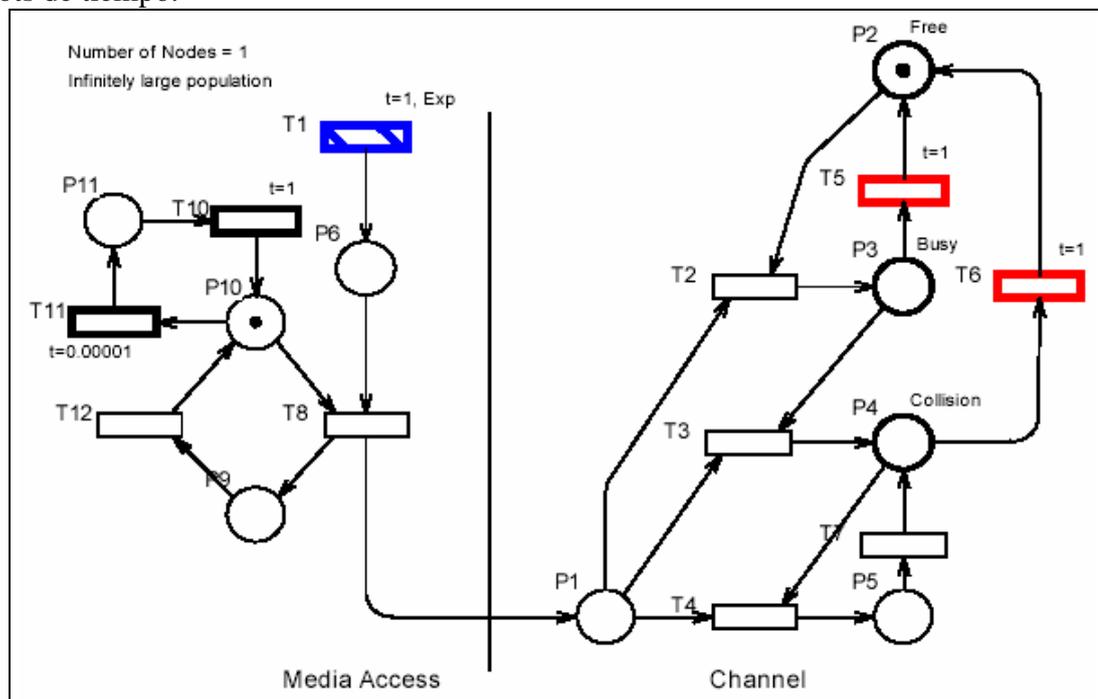


figura n° 31 Modelo FPQSPN del protocolo Aloha Ranurado

El número de marcas producidos en el lugar P6 por slot de tiempo es una variable aleatoria con distribución de Poisson. Cuando una marca reside en el lugar P10, el marcado del lugar P6, en el extremo del slot de tiempo representa el número de paquetes transmitidos simultáneamente a través del canal. La transición T12 y el lugar P9 evitan el autoloop entre P10 y T8. Asumiendo $t_5 = t_6 = 1$, el tiempo medio asociado a la transición T1 representa la recíproca del tráfico asociado G

La única instancia de tiempo, cuando la colisión puede ocurrir es el breve intervalo en el comienzo de un slot de tiempo. El canal está siempre libre en el momento entre dos slots de tiempo consecutivos. La frecuencia de disparo de la transición T2 representa la relación de los slot de tiempo “full”, en el cual el canal no está vago. La suma de las frecuencias de disparo de las transiciones T3 y T4 representan la relación de colisión del canal (la porción del ancho de banda del canal que es “gastado” por las colisiones).

Bajo estas condiciones, y con estos modelos de Aloha se podría presentar la animación de los mismos usando simulación (Capítulo III).

II.3) MODELACIÓN DE CSMA/CD

II.3.1 Introducción

La LAN Ethernet fue desarrollada por DEC, Xerox e Intel, pero se volvió un standard formalizado por el Instituto de Ingeniería Eléctrica y Electrónica (IEEE) como IEEE802.3.

Una red Ethernet es un sistema broadcast y multiacceso para computadoras locales con control distribuido. Las Ethernets son una de tecnologías LAN (Local Area Network) más populares de hoy en día. Entre sus características atractivas está su facilidad de mantenimiento y administración, facilidad de reconfiguración de red (en la mayoría de los casos las estaciones pueden ser desconectadas desde un punto, y reconectadas en otro punto sin la necesidad de “bajar” la red), acceso para un medio simple y pasivo (especificación IEEE 802.3 5 y 802.3 2), o medios activos relativamente baratos facilitados por los concentradores de red, y cable de par trenzado o cable de fibra óptica. Una ventaja adicional es la ausencia de un controlador central asignando acceso al canal. La desventaja de Ethernet transicional es su relativamente baja velocidad de 10 Mbps, fue resuelto por las extensiones de Ethernet como la 802.3u Fast Ethernet y la IEEE 802.3z Gigabit Ethernet operando a 1 Gbps.

Es necesario destacar la operación de la capa MAC Ethernet llamada CSMA/CD, la cual difiere de CSMA por el hecho que se escucha el canal durante la transmisión.

II.3.2 Conducta Ethernet

Ethernet usa una técnica de Control de Acceso al Medio (MAC) basada en la competencia llamada Acceso Múltiple con Sensado de Portadora y Detección de Colisión (CSMA/CD). CSMA/CD como todas las técnicas basadas en la competición, es muy simple en principio. Se espera hasta que el bus está libre y entonces se transmite. En general, el algoritmo de CSMA/CD es formalmente descrito por los siguientes pasos:

- Paso 1: Si el medio está inactivo (libre), entonces transmite. De otro modo, se va al paso 2.
- Paso 2: Si el medio está ocupado, se continúa escuchando hasta que el canal es sensado libre; luego se transmite inmediatamente.

- Paso 3: Si una colisión es detectada durante la transmisión, entonces se transmite una señal breve de jam para asegurarse que todas las estaciones sepan que ha habido una colisión y luego cesa la transmisión.
- Paso 4: Después de transmitida la señal de jam, se espera una cantidad aleatoria de tiempo, y se experimenta transmitir nuevamente; se repite el Paso 1.

Si una estación Ethernet quiere poner un paquete, se aplica primero una técnica de sensado de portadora, forzando a diferir la transmisión si alguna otra estación está transmitiendo un paquete. La estación esperará hasta que el paquete haya pasado antes de transmitir su propio paquete. Las colisiones, no obstante, pueden ocurrir, a causa que dos o más estaciones pueden sensar el canal como inactivo y comenzar la transmisión simultáneamente. Sin embargo, dado que cada transmisor continúa monitoreando el canal durante la transmisión, puede proveerse detección de colisión cuando la señal del canal no concuerda con su propia salida. En este caso, cada estación que interrumpe la transmisión, usa un procedimiento para forzar el consenso de colisión, para asegurar que todas las otras estaciones colindantes hayan visto la colisión (sensando una señal jam), y luego para. La retransmisión es entonces planificada para algún tiempo posterior, pero para evitar colisiones repetidas, cada estación espera por un periodo aleatorio de tiempo antes de retransmitir. Finalmente, para evitar la sobrecarga del canal (haciendo inestable el sistema), un algoritmo de control de congestión se usa y garantiza que el rango de los intervalos de retransmisiones se incremente apropiadamente en según la carga del sistema. El “truncated Binary Exponential Back-off” (BEB), y las modificaciones de los Métodos de Arbitración Logarítmica Binaria (BLAM) son usados como un mecanismo de control de congestión en Ethernet.

Algunos parámetros son propios de Ethernet, como:

- Bit rate: Bit rate significa velocidad de transmisión. El Bit Rate de Ethernet (IEEE 802.3) es 10 Mbps, o 100 Mbps para Fast Ethernet (IEEE 802.3u), o 1 Gbps para Gigabit Ethernet (IEEE 802.3z).
- Retardo de Propagación máximo: Un retardo round-trip entre dos transmisores es conocido como un retardo de propagación. El retardo de propagación máximo está definido en 464 bits.
- Tiempo de jam máximo: Un transmisor que detecta una colisión continúa para transmitir un patrón “jam” de 32 bits. En los modelos, el tiempo de jam máximo implica el tiempo de resolución de colisión, es que es denotado como b .
- Tiempo de slot: El límite superior en el tiempo de adquisición de la red es llamado a tiempo slot. Está definido en 512 bits.
- Límite de intentos: En la circunstancia de una colisión, la retransmisión de la trama es experimentada hasta un número máximo definido, que se conoce como el límite de intentos o número máximo de reintentos. El valor límite de intentos de Ethernet es 16.
- Tamaño de trama mínimo: El tamaño de trama mínimo de Ethernet es 64 bytes.
- Tamaño de trama máximo: La adopción de tamaño de trama máximo significa que la longitud de los buffers de trama usados para la transmisión y la recepción puede ser cuantificada. En Ethernet, 1518 limita el tiempo de acceso promedio.
- Número de estaciones: Existe un número máximo de estaciones por segmento o red. Ethernet fija 100 estaciones por segmento y 1024 en una Ethernet multi-segmento.
- Espacio Inter. Trama: Es una constante Ethernet standard igual a 96 bits.
- Límite backoff: Hay dos versiones de algoritmo backoff: un Backoff Exponencial Binario truncado (BEB) y Backoff Lineal. Backoff significa que la unidad MAC retarda un número integral aleatorio de slots. En general, el Límite Backoff de Ethernet es 10.

- Estructura del paquete: La estructura de paquete Ethernet es de 64 bits de preámbulo, 48 bits de la dirección Ethernet destino, 48 bits de la dirección Ethernet origen, 16 bits de longitud de paquete, 368 bits a 12 kbits de datos, y 32 bits CRC.

El mecanismo de detección de colisión de Ethernet es levemente diferente para una variedad de tipos de medios físicos (por ejemplo, en el cable coaxial el nodo monitorea el cable exactamente en el mismo punto que el transmite, pero un nodo de par trenzado monitorea la señal desde los concentradores de la red).

II.3.3 CSMA/CD como Sistema de Población Infinita

Consideraremos que la longitud media de paquete es de 768 bytes, y todas las constantes son normalizadas a este valor. Además, el retardo de propagación se considera igual a 0.01; y el tiempo de jam normalizado a la longitud de paquete como igual a 0.0052. El gap Inter-trama normalizado a 0.0156.

En la cuestión de las colisiones, el Ethernet requiere dos precauciones: (1) El protocolo reduce la probabilidad de aparición de colisiones al sensar el canal antes de la transmisión (CSMA); (2) Reduce el gasto de la capacidad del canal por disponer de las colisiones en un breve tiempo constante. En Ethernet de 10 Mbit el tiempo entre la aparición de colisiones y su disposición es igual a 3.2 microsegundos. Así, en los modelos de CSMA/CD, la disposición de las colisiones se representa por transiciones de tiempo determinístico T6. Inicialmente no se modelará las detecciones de colisiones.

II.3.4 Modelación de CSMA/CD

Inicialmente consideraremos la modelación CSMA del canal de comunicaciones. El origen de las colisiones en los métodos CSMA es el retardo de propagación. Esto es a causa que los nodos nunca ven el estado real del canal, sino que ven el estado del canal a segundos antes. En este modelo del canal de comunicaciones, el retardo de propagación es una constante para todos los nodos.

El significado de los lugares P1, P2, P3, P4, P5 y las transiciones T2, T3, T4, T5, T6 de la figura n° 32 es idéntica a los correspondientes lugares y transiciones en el modelo de tres estados de Aloha. El modelo muestra su “estado visible por los nodos” por dos lugares complementarios P10 y P11 (el término lugares complementarios significa un componente consistente de un par de lugares marcados por cero o uno). Una marca en el lugar P11 representa “estado libre visible”, y una marca en el lugar P10 representa “estado activo visible”. El canal se vuelve “activo visible” a segundos después de un nodo arbitrario entra en transmisión (representado por el disparo de la transición T2). El canal se vuelve libre visible nuevamente, a segundos después que el nodo transmitiendo actualmente finaliza su transmisión.

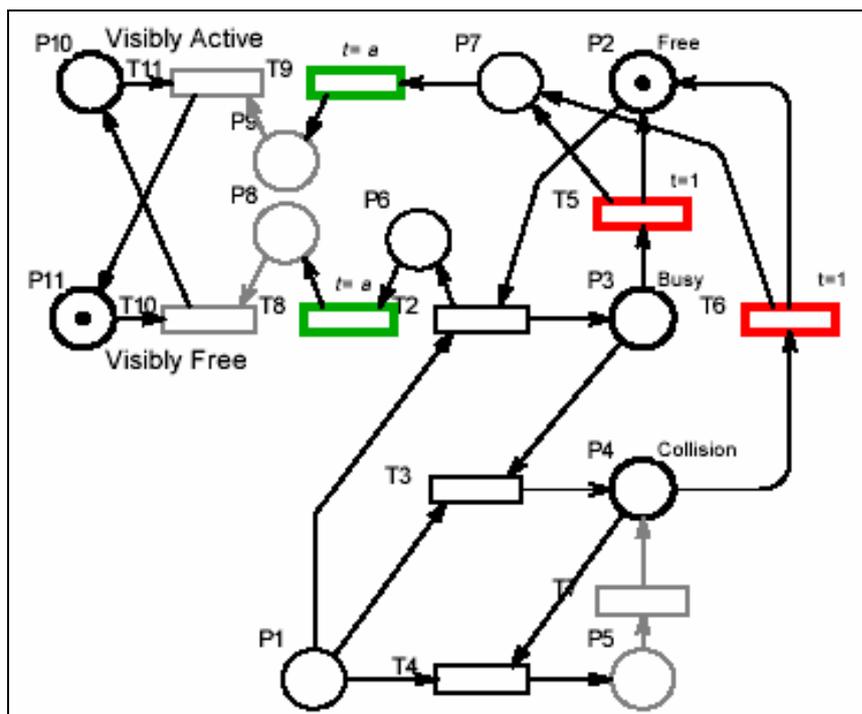


figura n° 32 Modelo FPQSPN del canal CSMA

Cuando el canal está en estado de colisión, se vuelve libre a segundos después que el último nodo participando en el conflicto para su transmisión (representada por el disparo de la transición T6). Así, una marca en el lugar P10 aparece a segundos después del comienzo del periodo ocupado. Una marca en el lugar P11 aparece a segundos después del comienzo de un periodo inactivo. Las transiciones T10 y T11 aseguran que el estado del tiempo de las transiciones T8 y T9 no se reseteen cada vez que un nodo lee el estado del canal. Cuando se lee el estado, una marca en el lugar P10 y P11 es removida y retornada en un loop de tiempo cero. Los lugares P2, P3, P4, P5 forman un componente conservativo. En un estado estable, se mantiene esa cuenta de disparos de $T2 = T5 + T6$, y $T6 = T3$.

Se espera que dos más nuevos nodos arbitrarios comiencen sus transmisiones en exactamente el mismo tiempo. Aún en el caso de una versión sloteada de protocolo, uno de los nodos comienza una instancia un tiempo al menos infinitamente breve antes que el segundo. Pero en el modelo es posible tener transmisiones simultáneas (en la simulación es necesario a causa que el tiempo es medible solo con precisión finita). Cuando dos o más nodos parecen comenzar su transmisión en el mismo momento (en tal situación dos o más marcas aparecen en el lugar P1), el canal cambia su estado desde el estado libre al estado de colisión por el disparo de la secuencia de transiciones (T2, T3, n*T4).

II.3.4.1 Confianza en el modelo

Bajo la semántica de disparo de transiciones de servidor infinito, el modelo del canal de comunicaciones opera correctamente para un valor arbitrario de retardo de propagación. La operación correcta significa que cuando hay más que una marca en el lugar P6 o P7., las marcas son consumidas en el mismo orden que ellas fueron producidas, y después que el tiempo de a unidades que ellas fueron producidas, el marcado de los lugares P10 y P11 muestra un tiempo de a unidades el estado viejo del canal. Bajo la semántica de disparo de transición de

simple servidor y $a < 1$, el modelo también opera correctamente; el marcado de los lugares P6 y P7 iguala a cero o a uno. Para $a > 1$, cuando el marcado de P6 y P7 puede ser mayor que uno, el modelo falla a causa que una marca producida en P6 o P7, cuando hay otra marca allí, es consumida un tiempo de a unidades después que la primera marca es consumida. Así, la segunda marca causará el cambio de estado visible del canal, luego que a segundos después que el cambio ocurrió. Cuando no se usa la semántica de servidor infinito se puede resolver este problema reemplazando las transiciones T8 y T9 por conjuntos de transiciones, de modo que las transiciones actúen como transiciones de semántica de disparo multiservidor.

II.3.4.2 Modelación de CSMA persistente

El CSMA persistente obtiene el máximo rendimiento razonable del canal al no dejar al canal libre si algún nodo tiene datos para transmitir. Un nodo que tiene dato para transmitir, opera de la siguiente manera:

- Cuando el canal es sensado libre, transmite el paquete,
- Cuando el canal es sensado ocupado, el nodo espera hasta que el canal se vuelve libre, y entonces transmite el paquete.

La modelación la realizaremos considerando que se trata de un sistema de cola M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor). En el modelo FPQSPN de CSMA persistente (figura n° 33), el proceso de arribo de paquete está representado por la transición T1. Asumiendo $t_5, t_6=1$, el tiempo promedio asociado a la transición T1, representa la recíproca del tráfico ofrecido G .

Un nuevo requerimiento que arriba al sistema de cola (un nodo arbitrario sobre la red recibe un paquete para transmisión) incrementa el marcado del lugar P14 en uno $\{M(P14)>0\}$. Cuando el canal está libre visiblemente $\{M(P11)=1\}$, la transición T14 se vuelve habilitada y la marca “se mueve” desde el lugar P14 al lugar P1. El disparo de la transición T14 representa el arranque de la transmisión de un paquete. Si el canal está activamente visible, después del arribo de un nuevo requerimiento, $\{MP(10) = 1\}$, el lugar P14 está marcado al menos por uno (el nodo almacena el paquete en su buffer de salida) hasta que el canal se vuelva visiblemente libre. La replanificación del paquete está representada por una marca en el lugar P14. La transición T12 y el lugar P12 son usados sólo para evitar el autoloop entre la transición T14 y el lugar P11. Cuando T14 está disparada, la transición T10 previene el “reseteo” del tiempo de transición T8. El disparo de la transición T14 no tiene influencia en las propiedades de tiempo del modelo del protocolo.

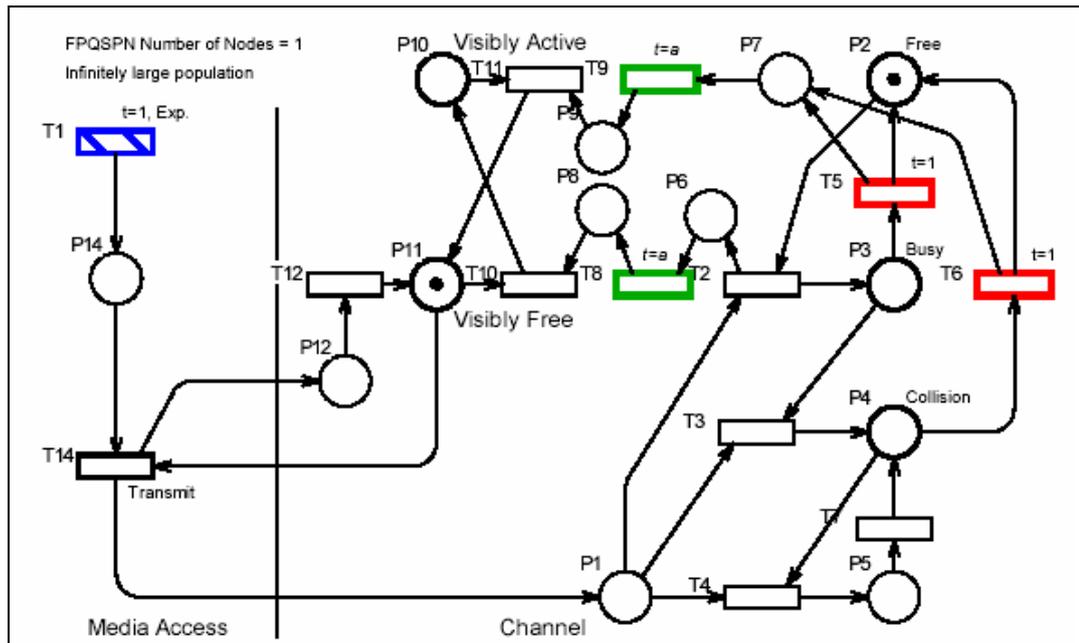


figura n° 33 Modelo FPQSPN del protocolo CSMA Persistente

El modelo representa un número infinitamente grande de nodos conectados al canal (note que la variable Número de Nodos = 1 significa que el modelo de Red de Petri contiene sólo un submodelo representando la población y un submodelo representando el canal). Cada nodo puede tener a lo sumo un paquete para transmitir en el mismo momento. Los nodos tienen sus buffer de salida de tamaño uno. El lugar P14 almacena los paquetes antes de la transmisión. Debido a la naturaleza estocástica del proceso de disparo de la transición T1, el marcado del lugar P14 no está limitado, a causa que cada marca en el lugar P14 representa un nodo, el cual tiene un paquete para transmisión. Colocar una marca en el lugar P14 marcado actualmente con cero, significa que un nodo justo recibió un paquete y entra en el algoritmo MAC. Colocar una marca en el lugar P14 marcado actualmente con distinto de cero, significa que un nodo ha recibido un paquete, pero espera hasta que el canal se vuelva libre. El marcado real del lugar P14 representa el número de nodos, lo cuales están realmente participando en el algoritmo MAC.

Puede demostrarse desde el análisis de cola, el rendimiento de los métodos CSMA es dependiente del retardo de propagación a.

II.3.5 Modelación de CSMA/CD (considerado como un sistema M/D/1)

El modelo FPQSPN de CSMA/CD, considerado un sistema de cola M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor), de la figura n° 34, es un modelo similar al modelo CSMA persistente. Ambos modelos difieren en la transición T6, en la transición T15 y el lugar P15, y en la ausencia de la transición T7 y el lugar P5. El modelo de la figura se subdivide en capa MAC y el canal de comunicaciones. Esta distinción es inexacta a causa de que la transición de resolución de colisiones (T6) es una función de los nodos y la transición T15 representa el espacio Inter-trama.

Un marca aparece en el lugar P9 a segundos después que el canal queda libre. De hecho, el canal se vuelve disponible para la transmisión después de $(a + \text{Inter-trama})$ segundos,

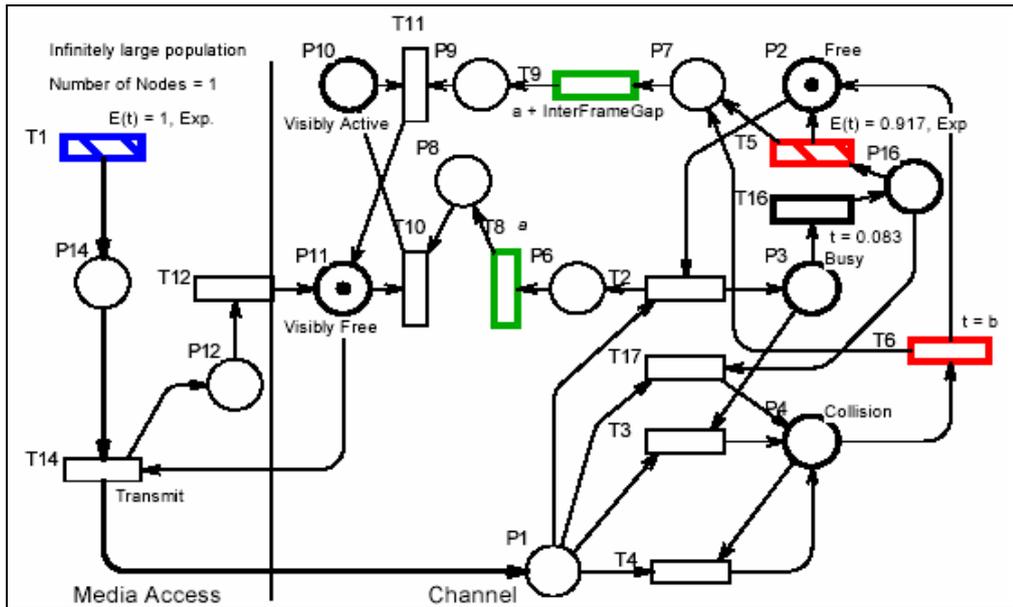


figura n° 35 Modelo FPQSPN del protocolo CSMA/CD M/M/1

II.3.7 CSMA/CD con Distribución de longitud de Paquete Bimodal

Existen diversos antecedentes sobre análisis extensivos del efecto de las distribuciones de tamaño de paquete bimodal sobre el rendimiento, la capacidad y el retardo de las redes CSMA/CD. Los análisis comienzan con tres parámetros la longitud de los paquetes cortos (L_s), la longitud de los paquetes largos (L_l) y la fracción de los paquetes que son cortos (FrS). Los resultados muestran que cuando solo una pequeña fracción de los paquetes es largo, el rendimiento máximo de la red se aproxima al máximo rendimiento que puede ser obtenido usando el tamaño de paquete de longitud fija, que iguala la longitud de paquete medio.

El modelo de CSMA/CD, modelado como sistema de cola M/G/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución general, y 1 servidor), correspondiente para paquetes de dos tamaños (figura n° 36) difiere del modelo de tamaño fijo, en las transiciones T15, T16, T17 y el lugar P15. Además, las transiciones T5 y T15 representan la transmisión del paquete, y son transiciones de tiempo determinístico. El tiempo de la transición T5 representa la longitud de los paquetes largos (igualado a 1), y el tiempo de la transición T15 igual a la longitud de los paquetes cortos, normalizado a la longitud de los paquetes largos. Ambas transiciones son transiciones con distribución uniforme, La relación entre los tiempos de dichas transiciones determina la fracción de paquetes que son cortos.

Puede demostrarse que el rendimiento es fuertemente dependiente de la fracción de paquetes que son cortos. Las curvas de rendimiento particular difieren unas de otras a causa de la influencia de la longitud de paquete medio, lo cual influye sobre el retardo relativo de propagación del canal a . La diferencia entre el rendimiento bajo distribución de longitudes de paquetes bimodal y bajo longitud de paquete fija correspondiente es pequeña cuando se compara a la influencia del número de nodos o la influencia del retardo del propagación del canal a . La influencia es más fuerte para G grandes ($G > 10$).

La carga bimodal influye sobre el rendimiento del canal a causa de que la fracción de los paquetes cortos decremento los longitud de paquete media. La influencia de la distribución de paquete bimodal en si misma puede ser despreciable para pequeños G ($G < 10$).

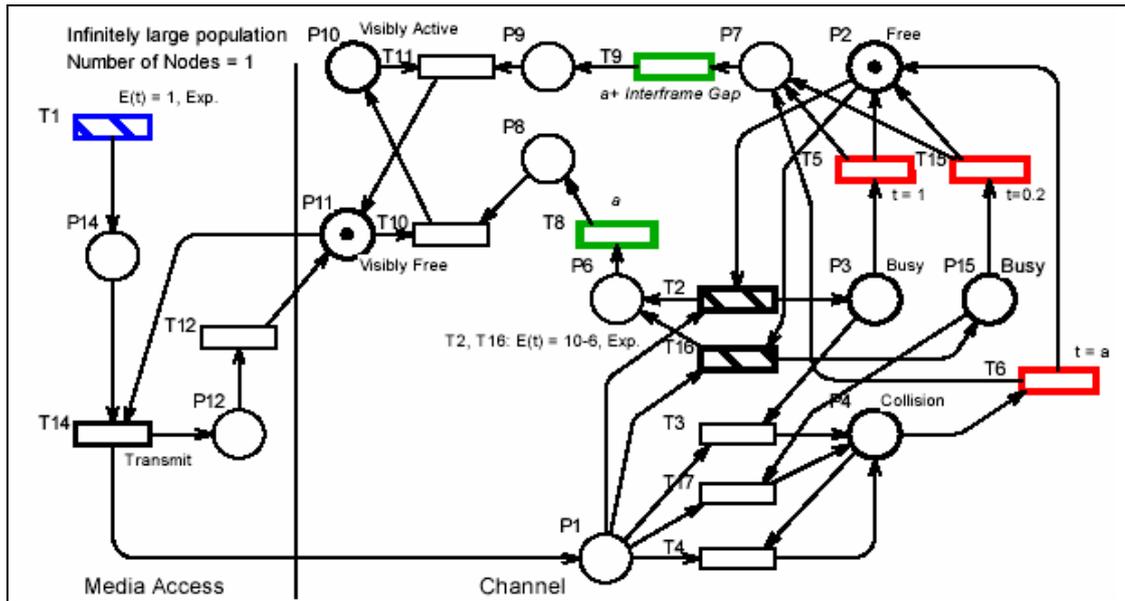


figura n° 36 Modelo FPQSPN del protocolo CSMA/CD M/G/1 con dos longitudes de procesos de arribo de paquete

II.3.8 CSMA/CD considerado como Sistema de Cola con Recuperación y con Pérdida de Paquetes

II.3.8.1 CSMA/CD como Sistema de Cola con Recuperación

En el modelo básico, el tráfico ofrecido a los nodos (la carga de entrada I) está dado por las frecuencias de disparo de $T1$ en todos los nodos. La frecuencia de disparo de $T1$ está dada por la recíproca del tiempo medio de transición $T1$, y se usa la frecuencia de $T1$ como variable independiente para obtener resultados. Cuando no se considera la pérdida de paquetes, el estado estable se mantiene para el modelo con recuperación, cuando la carga de entrada iguala al rendimiento del canal S . Por esta razón, el tráfico ofrecido no puede ser usado como variable independiente.

El modelo de Red de Petri, considerado como un modelo de cola $M/D/1$ (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor), correspondiente al Sistema de cola con recuperación (figura n° 37) difiere del modelo básico en que introduce las modificaciones en las transiciones y lugares para incorporar la recuperación de los paquetes. El parámetro fundamental es la inclusión del intervalo de recuperación.

Diversos trabajos determinan que:

- Indiferente del número de nodos sobre el canal, para valores pequeños de carga de entrada (valores grandes de tiempo medio de la transición $T1$), el rendimiento del canal es más grande para pequeños valores de retardo de recuperación que para valores grandes de retardo de recuperación. El mayor retardo de recuperación medio incrementa el tiempo de cola medio del sistema. Así, los valores pequeños de retardo de recuperación medio da mejor performance para valores pequeños de tráfico ofrecido.

- Para valores grandes de carga de entrada el canal se vuelve inestable para valores pequeños de retardo de recuperación medio.
- Cuando el retardo de recuperación medio es suficientemente grande, el canal es estable aún cuando la carga es máxima.
- Dado un valor de retardo de recuperación medio; el canal cargado por varios nodos (por ejemplo 100) es inestable para carga de entrada grande, mientras que el mismo canal cargado por pocos nodos es estable, para aún un valor más pequeño del retardo de recuperación medio.

La red CSMA/CD es inestable para la cual tiene que estar establecida el mecanismo de control de congestión, el cual optimiza el retardo de recuperación medio, de modo que el canal sea estable y el rendimiento óptimo.

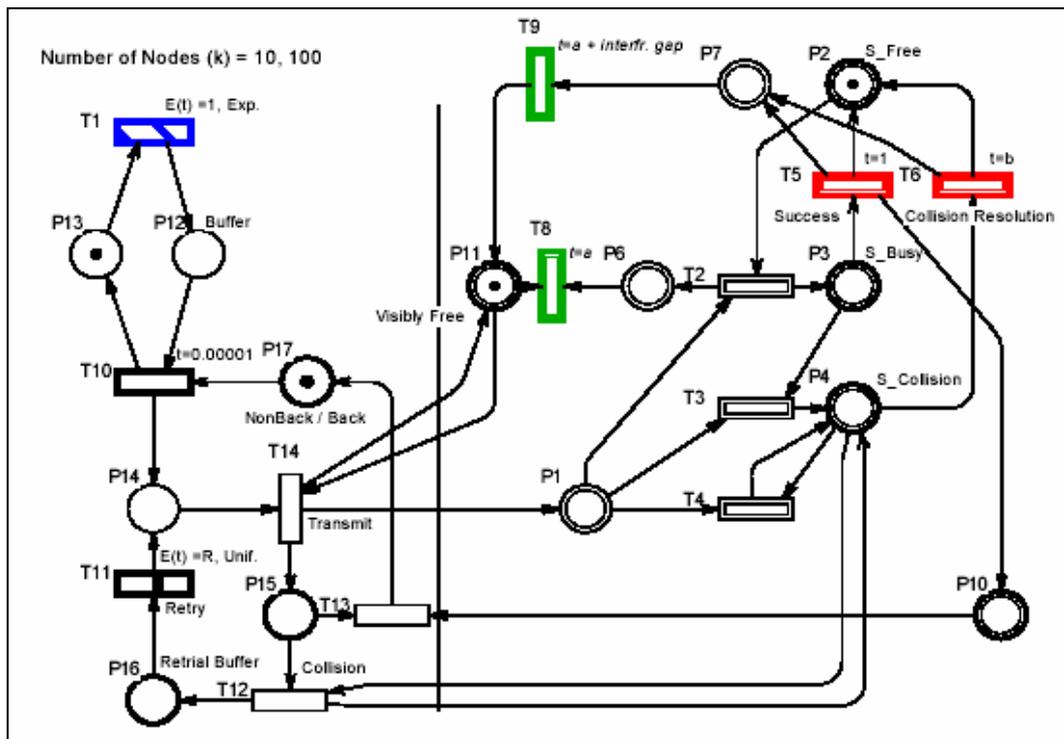


figura n° 37 Modelo FPQSPN de CSMA/CD como un Sistema M/D/1/∞/k con recuperación

II.3.8.2 CSMA/CD con Pérdida de Paquete

Ethernet usa CSMA/CD persistente como su protocolo de medio de acceso primario. Un nodo Ethernet pierde el paquete, que fue no exitoso durante sus 15 intentos de transmisión. Esta precaución, que es una parte del mecanismo de control de congestión usado en Ethernet, asegura que el paquete, cuya transmisión es difícil, no congestione el canal por siempre. Debido a la pérdida de paquetes, para el modelo básico, no se mantiene el rendimiento del canal S igualado a la carga de entrada I .

El modelo correspondiente (figura n° 38), considerado como un sistema de cola M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor), es similar al anterior, excepto que un lugar debe representar la cuenta de colisiones del paquete procesado actualmente. Puede demostrarse que de la comparación de los modelos sin y con pérdida de paquetes, bajo exactamente las mismas

condiciones, el mecanismo de pérdida de paquetes mejora la conducta del canal. Además, el mecanismo de pérdida de paquete es un factor que estabiliza el canal.

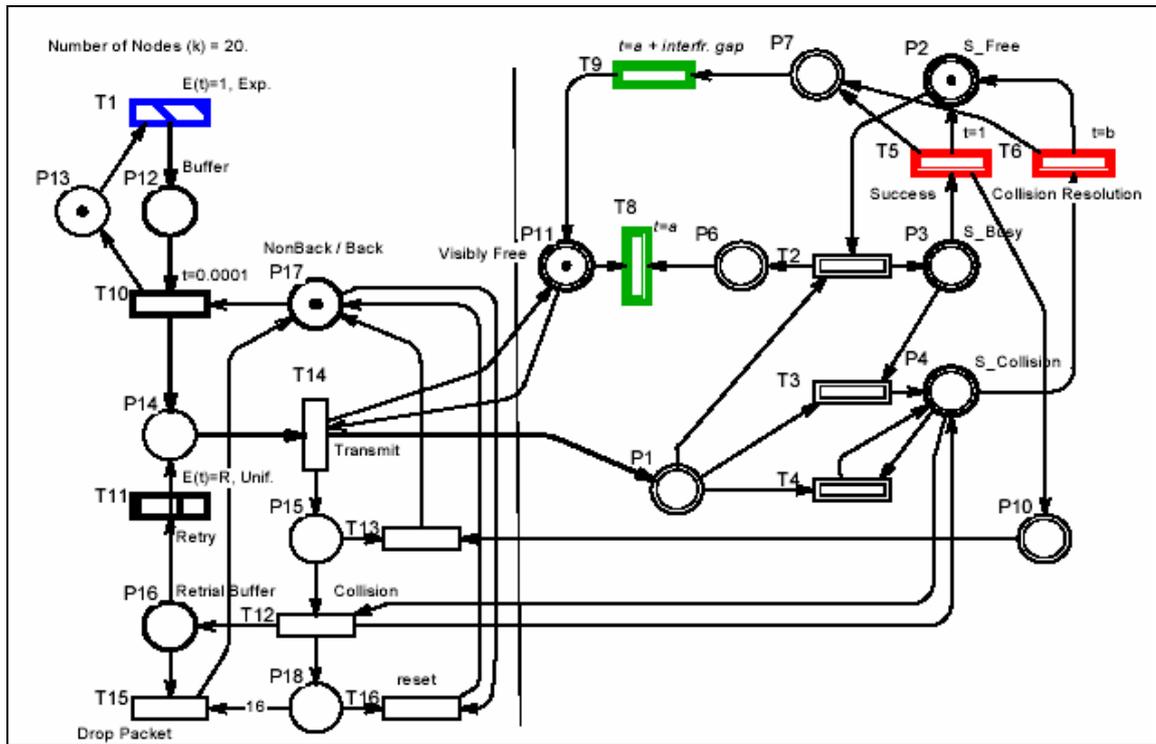


figura n° 38 Modelo FPQSPN del protocolo Ethernet como un Sistema M/D/1/∞/k con recuperación sin control de congestión

II.3.9 Análisis de Ethernet

La estabilidad es uno de los objetivos teóricos fundamentales en el análisis de protocolos de comunicaciones. La estabilidad está influida por el mecanismo de control de congestión, que controla la intensidad de reciclar los paquetes previamente colisionados. Existen diversas definiciones para enfocar el fenómeno de estabilidad. Por ejemplo, se considera que el canal es estable si cada estación es capaz de transmitir paquetes exitosamente en el ritmo en el cual los generan. Otra forma de definición de estabilidad de un protocolo de comunicación es que el rendimiento S es una función creciente de la carga de entrada I .

II.3.9.1 Binary Exponencial Backoff

El truncated Binary Exponencial Backoff (BEB) es un mecanismo de control de congestión usado en Ethernet. En este algoritmo, una estación duplica su intervalo de retransmisión medio con cada colisión sucesiva del mismo paquete. La retransmisión es retardada un múltiplo entero aleatorio del tiempo de slot. Sobre una LAN Ethernet, el tiempo de slot está definido de 512 bits, correspondiente al tamaño mínimo de una trama Ethernet, pero puede haber un retardo arbitrariamente largo antes de que un paquete sea transmitido, aún si la red no está fuertemente cargado. Para evitar este problema, la duplicación para después de las primera diez colisiones sucesivas, y el proceso es truncado. Después de 16 colisiones sucesivas, el intento de transmisión es abandonado y el paquete se pierde.

El modelo de Ethernet (figura n° 39), considerado como un sistema M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor) difiere del anterior por la implementación del algoritmo BEB.

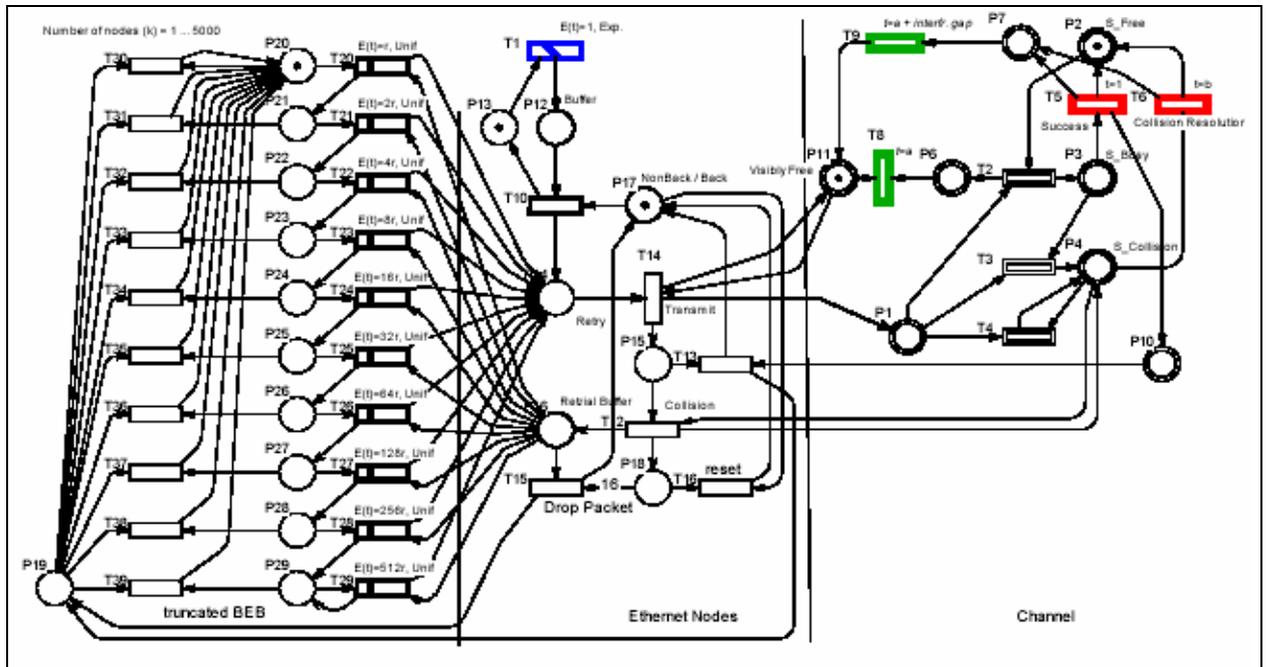


figura n° 39 Modelo FPQSPN del protocolo Ethernet como un Sistema M/D/1/∞/k

II.4) MODELACIÓN DE 802.11 [30][31][32][33]

II.4.1 Introducción

En los años recientes, el mercado para comunicaciones wireless se ha expandido explosivamente en el mundo, en campos tales como la telefonía celular, redes de datos móviles y redes de área local wireless (WLANs). Las WLANs son apropiadas para las comunicaciones de alta velocidad sobre áreas pequeñas, donde el cableado para las redes convencionales es difícil o antieconómico. Para establecer interoperabilidad entre sistemas WLAN de diferentes vendedores, la IEEE estableció la versión final de un standard internacional para las WLANs, a los finales de 1997. El estándar IEEE 802.11 especifica en detalle la capa de control de acceso al medio (MAC) y la capa física (PHY) para operar WLANs en la banda de 2.4-2.5 Ghz en las velocidades de 1 o 2 Mbps. Desde ese momento, han aparecido otras opciones para velocidades más altas, con la misma y otras frecuencias de señalización de radio, (802.11b, 802.11a, 802.11g entre otros).

II.4.2 Mecanismos MAC en wireless LAN

El mecanismo MAC primario en la IEEE 802.11, referido como la función de coordinación distribuida (DCF), es una variante de CSMA no persistente. Todas las estaciones de una celda compiten por el mismo canal. En el así llamado esquema de acceso básico, una estación que desea transmitir una trama de dato, primero sensa el canal y transite sólo si el canal

ha estado libre por algún periodo de tiempo. De otro modo, la estación difiere la transmisión y va a backoff. El receptor de la trama de datos transmite un reconocimiento al transmisor. Bajo condiciones ideales, las colisiones pueden ocurrir solo si dos estaciones comienzan a transmitir en casi el mismo instante de tiempo: el periodo vulnerable está compuesto por el tiempo que la señal se propaga sobre el canal, el tiempo que el hardware necesita para detectar una señal y el tiempo de conmutación desde el modo de recepción a transmisión. Por lo tanto, no hay disponible ningún mecanismo para detección de colisiones durante la transmisión. Una estación fuente reconoce una colisión por un reconocimiento erróneo o por una transmisión en progreso después que la propia transmisión ha finalizado.

Una variante al esquema básico recibe el nombre de MACA (Acceso Múltiple con Prevención de Colisiones - Múltiple Access with Collision Avoidance), también denominada por otros autores bajo el nombre de CSMA/CA (Acceso Múltiple con Sensado de Portadora y Prevención de Colisiones), haciendo referencia a una variante de CSMA que se propone evitar las colisiones. Este protocolo MAC intenta evitar las colisiones, cuando una estación tiene una trama que transmitir, enviando una trama pequeña de aviso denominada RTS (Request To Send – Solicitud de Envío). La trama contiene información sobre la longitud de la trama que se pretende transmitir y el nodo destino. El nodo destino, al recibir la trama RTS, y si está en condiciones de recibir la transmisión, responde con otra trama denominada CTS (Clear To Send - Libre para Envío). El CTS también indica la longitud de la trama que se va a recibir.

En estas condiciones, para una transmisión entre dos nodos, cualquier estación que escucha un RTS está evidentemente cerca del emisor del mismo, y permanecerá en silencio durante un tiempo suficiente para que el CTS del destinatario del RTS, se transmita de regreso sin conflicto. Una estación que escuche el CTS (pero no el RTS) está bastante cerca del destino, y deberá permanecer en silencio durante la siguiente transmisión de datos, cuya longitud puede determinar examinando la trama CTS. A pesar de estas precauciones aún pueden ocurrir colisiones. Un emisor sin éxito (es decir, uno que no escucha un CTS en el intervalo de tiempo esperado) espera un tiempo aleatorio y reintenta. El algoritmo empleado es el retroceso exponencial binario como en Ethernet.

Además, de estos mecanismos se han propuestos otros como el MACAW (MACA Wireless) introduciendo mejoras adicionales.

II.4.3 El DCF del IEEE 802.11

En el standard 802.11 (original), se han definido las dos versiones mencionadas en el apartado anterior de la función de coordinación distribuida (DCF): el acceso básico basado sobre handshake de dos vías, y el Request-to-Send/Clear-To-Send (RTS/CTS) basado sobre handshake de cuatro vías. Una estación puede usar RTS/CTS para tramas de datos que exceden un umbral configurable. En ambos casos, sólo el primer paquete va a competir por el medio y el acceso está basado sobre dos periodos de tiempo: el Espacio Intertrama DCF (DIFS) y el Espacio Intertrama Corto (SIFS) con $SIFS < DIFS$.

El acceso básico funciona como sigue:

Una estación que desea transmitir una trama de datos dirigida

- a. Sensa el canal,
- b. Si el canal ha estado libre por más que DIFS, la estación transmite la trama y espera por un reconocimiento positivo (ACK),
- c. La estación va a backoff si 1) el canal no ha estado libre por un periodo igual a DIFS, y 2) ningún reconocimiento ha arribado en tiempo (correspondiente a una colisión), ó c) la trama es consecutiva a una transmisión previa de la misma estación.

Después, al final de la recepción exitosa de la trama, la estación receptora espera un periodo igual a SIFS, y pone el reconocimiento ACK.

En RTS/CTS dos o más paquetes son intercambiados: el paquete RTS contiene información acerca de la longitud de la trama, el paquete CTS es un reconocimiento del paquete RTS y contiene la misma información. Los espacios entre transmisiones sucesivas son igual a SIFS. La figura n° 40 muestra un handshake RTS/CTS exitoso. De la información contenida en los paquetes RTS y CTS, las otras estaciones puede deducir el tiempo que el medio estará ocupado. Las intenciones de usar el mecanismo RTS/CTS más complejo son dos. Primero, dado que sólo el RTS breve compite por el medio, las colisiones gastan menos ancho de banda. El RTS/CTS puede por lo tanto ser visto como un mecanismo de detección de colisión virtual. Segundo, las colisiones adicionales causadas por las terminales hidden son también evitadas por el intercambio RTS/CTS.

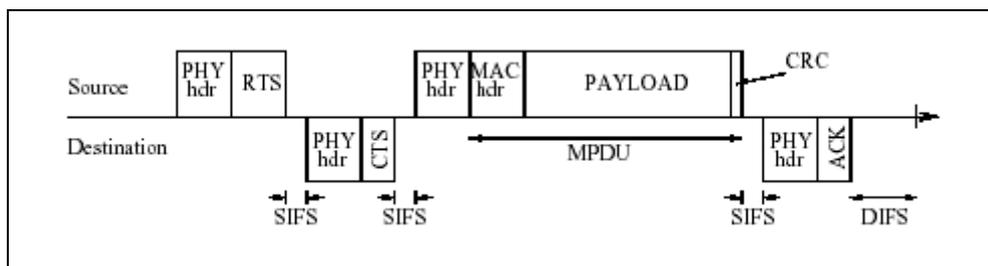


figura n° 40 Handshake RTS/CTS exitoso

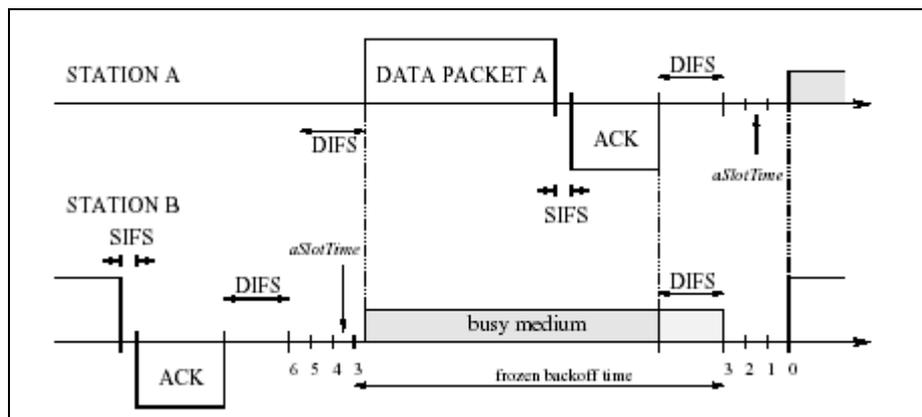


figura n° 41 Procedimiento Backoff con Acceso Básico

En 802.11 se usa un backoff exponencial binario slotteado. Tan pronto como el canal es monitoreado libre por un DIFS, la estación selecciona un contador backoff aleatorio, el cual es decrementado cuando el medio está libre. Cuando una transmisión es detectada, el contador es “congelado” hasta que el medio está nuevamente libre y una DIFS ha transcurrido. El contador backoff es elegido como sigue:

$$\text{BackoffTime} = \text{Random}() \times \text{aSlotTime},$$

Donde $\text{Random}()$ es un entero uniformemente distribuido en $[0, \text{CW}]$. La ventana de competición CW es un entero a $(\text{aCWmin} + 1) \cdot 2^{\text{bc}} - 1$, donde aCWmin es el valor inicial y bc es una variable que es inicializada con cero, e incrementado antes, por un procedimiento backoff repetido por un

trama pendiente. bc puede crecer hasta un valor máximo bc_{max} correspondiente a aCW_{max} , después que el permanece sin cambios hasta el se resetea a cero para la transmisión exitosa siguiente.

La figura nº 41 ilustra el procedimiento backoff. Después de la finalización de una transmisión previa, la estación B espera por un DIFS y entonces elige un contador backoff igual a 6 antes de transmitir el paquete consecutivo. Mientras, una nueva trama es generada en la estación A. La estación A sensa el canal libre por un DIFS, y transmite la trama. Durante la transmisión, la estación B detiene el decremento de su contador backoff. El contador es decrementado nuevamente cuando el canal ha sido sensado libre por un DIFS.

El standard 802.11 describe unos pocos mecanismos adicionales que no se han planteado. Estos incluyen las características de Espacio Intertrama Extendido (EIFS) y la Función de Sincronización de Tiempo (TSF), como así también características opcionales, tales como el Función de Coordinación Puntual (PCF), administración de potencia, y fragmentación de unidades de datos grandes. Además, en la discusión anterior se ha asumido condiciones ideales del canal, es decir, no son considerados el problema de terminal oculta, errores de bits, y las fluctuaciones de la calidad del canal; y se ignoran la posible movilidad de las estaciones y el efecto cerca/lejos, que permite capturar la recepción posiblemente exitosa de transmisiones simultáneas.

II.4.4 Modelación 802.11 en Acceso Básico como CSMA no persistente.

Es oportuno tener en cuenta las suposiciones planteadas en II.2.1, y la modelación del canal discutido en II.3.4 y la figura nº 30.

Como se mencionó previamente, el mecanismo MAC primario en la IEEE 802.11, referido como la función de coordinación distribuida (DCF), es una variante de CSMA no persistente. Por ello, se desarrollará el modelo desde esa perspectiva.

En el CSMA no persistente la idea para reducir la interferencia entre los paquetes es siempre replanificar un paquete que encuentra el canal en estado ocupado en su arribo. Un nodo que tiene un paquete para transmitir, opera de la siguiente manera:

- Si el canal es sensado libre, el nodo transmite el paquete inmediatamente,
- Cuando el nodo sensa el canal ocupado, el planifica la retransmisión del paquete de acuerdo a la distribución de retardo de retransmisión. El procedimiento MAC se repite en este nuevo punto en el tiempo.

El CSMA no persistente difiere del CSMA persistente por replanificar un paquete después de sensar el canal ocupado. Mientras el CSMA persistente transmite el paquete inmediatamente después que el canal está libre, el CSMA no persistente replanifica la transmisión del paquete. Para pequeños G , este procedimiento extiende el tiempo medio que un paquete gasta en el sistema. Para grandes G , el procedimiento reduce la probabilidad de colisión después que el canal está libre, de modo que el rendimiento máximo del canal es grande comparado al CSMA persistente (estos aspectos se discutirán en el Capítulo IV).

CSMA no persistente es inherentemente un sistema con recuperación. Uno puede distinguir dos mecanismos de recuperación en este protocolo. El primero es el reciclado de paquetes, el cual encuentra el canal ocupado. El segundo es un reciclado potencial de paquetes recibidos con error por las capas superiores del modelo OSI. Aquí se presentará la modelación de la conducta desde el punto de vista del canal.

El modelo FPQSPN de la figura nº 42 representa al CSMA no persistente, considerando que se trata de un sistema cola M/D/1.

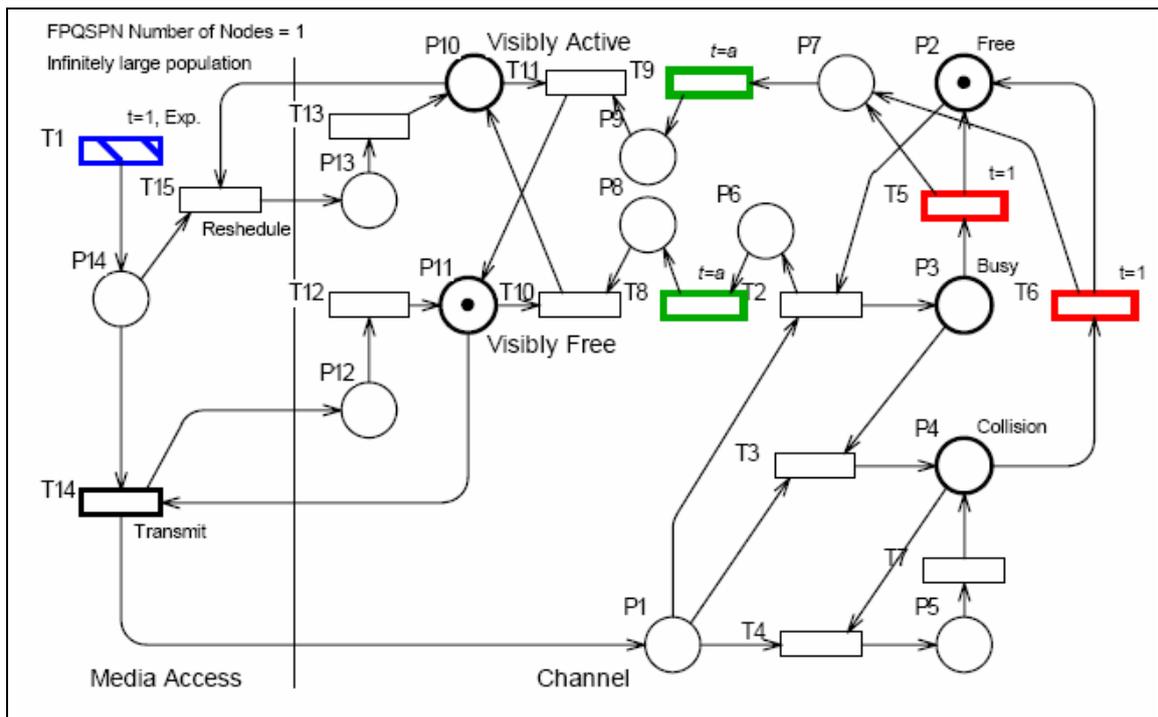


figura nº 42 Modelo FPQSPN de CSMA no persistente

Un nuevo requerimiento que arriba al canal libre (un nodo arbitrario sobre la red recibe un paquete para transmisión) marca el lugar P14 con 1. En este momento, cuando el canal está visiblemente libre, la transición T14 “compromete el paquete” al canal (el consume una marca desde el lugar P14 y produce una en P1). El disparo de la transición T14 representa el comienzo de la transmisión del paquete. Cuando colocamos una marca en P14 y el canal está visiblemente activo, (hay una marca en P10), la transmisión es replanificada. Se espera que el retardo de replanificación del paquetes sea “largo” comparado al tiempo de transmisión del paquete, de modo que el paquete replanificado puede ser considerado con un nuevo paquete arribado. La transición T15, que representa la replanificación de un paquete “deja caer la marca” en el lugar P14. La transición es disparada si hay tráfico sobre el canal en el momento cuando un paquete está listo para colocarse. Se nota que el valor del tráfico ofrecido G representa el número de intentos del algoritmo MAC por unidad de tiempo. En estado estable, la cuenta de disparos de la transición T1 es igual a la cuenta de disparos de la transición T14 más la cuenta de disparos de la transición T15.

Como en el caso de CSMA persistente, el modelo representa un número infinitamente grande de nodos conectados al canal (note que la variable Número de Nodos = 1 significa que el modelo de Red de Petri contiene sólo un submodelo representando la población y un submodelo representando el canal). Cada nodo puede tener a lo sumo un paquete para transmitir en el mismo momento. Los nodos tienen sus buffers de salida de tamaño uno. El lugar P14 almacena los paquetes antes de la transmisión. Debido a la naturaleza estocástica del proceso de disparo de la transición T1, el marcado del lugar P14 no está limitado, a causa que cada marca en el lugar P14 representa un nodo, el cual tiene un paquete para transmisión. Colocar una marca en el lugar P14 marcado actualmente con cero, significa que un nodo justo recibió un paquete y entra en el algoritmo MAC. Colocar una marca en el lugar P14 marcado actualmente con distinto de cero, significa que un nodo ha recibido un paquete, pero espera hasta que el canal se vuelva libre. El marcado real del lugar P14 representa el número de nodos, lo cuales están realmente participando en el algoritmo MAC.

Como se discutió precedentemente en los otros modelos, se considera que al canal es un dispositivo que tiene una o más entradas y ninguna salida. La salida del canal no es un objeto del análisis de performance de protocolos. Todas las entradas del canal son asumidas idénticas, y el tiempo de transmisión de paquete, el cual está dado en la práctica por la longitud del paquete y por la velocidad en baudios del canal, es una constante. Igualmente que un paquete es exitosamente transmitido, cuando el valor actual de la variable aleatoria que representa el tiempo interarribo entre paquetes iguala al menos al intervalo vulnerable. Si cualquier otro paquete es ofrecido durante el intervalo vulnerable ocurre una colisión. Si el canal está en estado de colisión en el tiempo t_i , y luego un paquete es ofrecido al medio, la colisión persiste hasta el tiempo t_i+1 .

El intervalo vulnerable iguala al tiempo de transmisión de paquete bajo una condición para todos los nodos, que nunca comience su transmisión cuando su parte receptora está recibiendo un paquete (asumimos que el retardo de propagación a es constante para todos los nodos). De otro modo, el intervalo vulnerable es una suma de retardo de propagación a , dado por propiedades físicas del canal, y el tiempo de transmisión del paquete.

Por ello, el proceso de arribo de paquetes en el modelo de Red de Petri de la figura n° 42, es representado por una transición de tiempo de fuente estocástica T1, a causa de que la transición es un generador de una variable aleatoria distribuida exponencialmente. Las transiciones T5 y T6 son transiciones de tiempo determinístico con $t = 1$. El tiempo t_5 de la transición T5 representa el tiempo de transmisión de paquete, y el tiempo t_6 de T6 representa el retardo entre el tiempo de arribo del último paquete participando en una colisión, y el instante en que la colisión se resuelve. Una marca en el lugar P2 representa un estado libre (inactivo) del canal, el lugar P3 representa la transmisión de un paquete, y el lugar P4 representa una colisión.

La caracterización del protocolo 802.11 con RTS/CTS y las modelaciones realizadas usando Redes de Petri [10] coincide con los estudios analíticos en los siguientes aspectos,

- el deterioro del rendimiento observado en el mecanismo de acceso básico para G superiores es compensado en el mecanismo RTS/CTS,
- el mecanismo de sensado virtual hace al protocolo muy similar al CSMA/CD no persistente,
- FHSS obtiene un rendimiento más bajo que DSSS, probablemente debido al periodo vulnerable más largo,
- el rendimiento para DSSS con acceso básico permanece aceptable, si la velocidad está en un orden de magnitud mayor que el estándar, pero que sería demasiado bajo para bitrates más grandes.
- El rendimiento para DSSS con acceso básico y RTS/CTS para $B = 2$ Mbps y únicamente tramas grandes, pueden utilizar el canal eficientemente; y que el intercambio de tramas adicionales hace a RTS/CTS levemente honeroso para tramas pequeñas.

II.5) MODELACIÓN DE ALOHA PURO CON COLOREADAS [19][20][23]

II.5.1 Introducción

Como en el punto II.2.1 y II.2.2, se plantea la modelación del canal de Aloha puro mediante el formalismo de las Redes de Petri Coloreadas. Como discutido en I.9.4, las Redes de Petri Coloreadas (CPN) son útiles para la modelación para los protocolos de comunicaciones, pero son menos transparentes que FPQSPN. Además, en CPNs, las herramientas de simulación existentes, permiten la generación de los modelos, como se muestra a continuación, pero están limitados para evaluar parámetros de rendimiento con precisión.

El gráfico de la figura n° 43 se compone de dos partes:

- La primera es un generador de mensajes que sigue la distribución de Poisson compuesto por el lugar T y la transición $T0$. De la transición $T1$ se lleva la cuenta de los mensajes emitidos. El lugar T tiene asignado un color denominado *data* que es tipo cadena temporizado. Dicho lugar tiene una marca inicial con la cadena *Juán*, temporizado con tiempo inicial cero, representado por $@0$. Los dos arcos representados por p son variables del tipo *DATA*. La transición $T0$ tiene asociada un tiempo de retardo representado por una función de distribución de Poisson llamada *posdelay(0.35)*; esta función envía un paquete por segundo donde se ha hecho un ajuste que tiene que ver con la velocidad del procesador ya que las transiciones no son realmente instantáneas.. El lugar B representa la cola de mensajes disponibles a la entrada del canal transmisor.
- La segunda parte representan los estados del canal de comunicaciones. El lugar C equivale a canal libre representado por una marca que junto con un mensaje del lugar B dispara la transición $T2$ pasando el mensaje mas la marca habilitante al lugar D donde se retarda una unidad de tiempo que representa el tiempo de ocupación del canal, transcurrido ese tiempo y no habiendo mensaje en B se habilita la transición $T5$ y el mensaje es transmitido con éxito. Estando el mensaje en D y otro en B , se dispara $T3$ se produce una colisión ocupando una marca e en el lugar F , si otro mensaje está disponible en B se dispara $T6$ produciendo una doble colisión. La marca e de colisión y doble colisión son devueltas al lugar C contabilizando dichos sucesos. Cabe señalar que el retardo $@+1$ afecta no solo al mensaje exitoso sino también a las colisiones.

Debajo del gráfico se encuentra un recuadro con las declaraciones de las variables del sistema de acuerdo a la herramienta CPN ML, que es el lenguaje del simulador utilizado. Es importante señalar que al contrario de otras herramientas de modelación, el uso del formalismo de coloreadas permite representar por un lado los mensajes (variable p) y por otro las marcas de habilitación (variable e). En ese sentido, el modelo resultante permitirá la animación, posibilitando la visualización de la circulación de los mensajes en diferentes situaciones. Todo esto permite una mejor comprensión del funcionamiento de los canales de comunicación para distintos tipos de formatos mensajes.

II 6) CONCLUSION

Dentro de los objetivos y la metodología a seguir en el trabajo de tesis, se ha adoptado una secuencia de modelación de protocolos de red de canal compartido usando las FPQSPN, desde Aloha puro, dejando de manifiesto desde el punto de vista de la modelación, que protocolos superiores han seguido un curso de perfeccionamiento desde aquel, en los aspectos de detección de portadora y/o de colisiones, para el incremento de la performance de uso del canal.

En este Capítulo se han presentado los aspectos teóricos sobre los Protocolos de Red, sus medidas de performance y características de tráfico de red. Posteriormente, los modelos FPQSPN de protocolos en el siguiente orden: Aloha puro, Aloha ranurado, CSMA persistente, CSMA no persistente, CSMA/CD, y Ethernet). Finalmente, y a título ilustrativo se ha presentado el modelo con CPN (Redes de Petri Coloreadas) del modelo de Aloha puro.

Dado que las FPQSPN son apropiadas para la modelación y análisis de sistemas vistos como sistemas de colas (como los protocolos de comunicaciones), donde la población es finitamente grande, y para mantener la transparencia de las Redes de Petri de Bajo Nivel, es relativamente fácil observar la creciente complejidad de los modelos desde Aloha puro.

Efectivamente, desde la representación gráfica del modelo FPQSPN de Aloha puro puede observarse que la modelación es del canal y que los estados (lugares) básicos del mismo son: libre, ocupado, y colisionado. Y a este modelo básico con tres estados del canal, se han agregado nuevos estados (lugares) que corresponden al funcionamiento ranurado, detección de portadora o detección de colisiones, para los restantes protocolos superiores, cuyo objeto es el incremento de la performance.

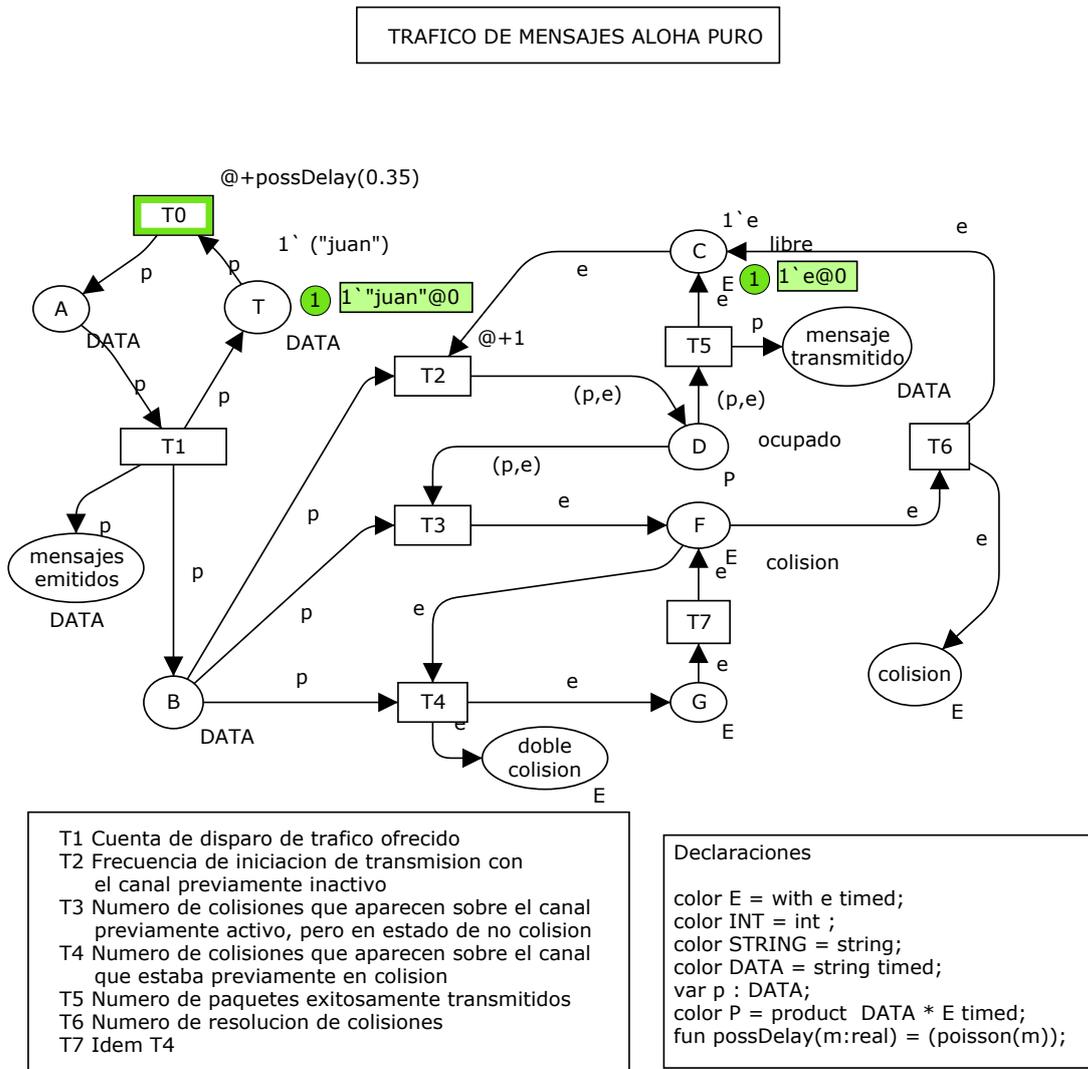


figura nº 43 Modelo con Coloreadas de Aloha Puro

CAPITULO III

SIMULACION DE PROTOCOLOS DE COMUNICACIONES

III.1) PRINCIPIOS DE SIMULACIÓN [34]

Se considera una simulación (o simulación de computadora) como la ejecución de una abstracción (o modelo del sistema) de un sistema real existente o no existente por medio de software (de computadora). Como una alternativa a las técnicas de análisis, la simulación aparece como un buen compromiso entre modelos matemáticos-analíticos, frecuentemente demasiados abstractos con respecto a las suposiciones de la conducta del sistema, y la necesidad de construir sistemas reales experimentales, frecuentemente demasiados caros. La simulación sirve como un complemento a la aproximación de modelación matemática-analítica, en casos donde el sistema real no existe, de modo que las medidas no pueden aplicarse, y una abstracción realista de la conducta del sistema real hipotético no puede ser expresada o analíticamente evaluada. Sobre el otro extremo, complementa las mediciones si ellas son demasiado costosas, dañinas, inapropiadas o imposibles.

Para entender la operación del sistema real, para evaluar su performance, para seleccionar y sintonizar parámetros de diseño, para verificar y evaluar diseños, para comparar alternativas, etc. el analista está generalmente interesado en su conducta, es decir, en la salida que se genera basada en una cierta entrada (figura n° 44).

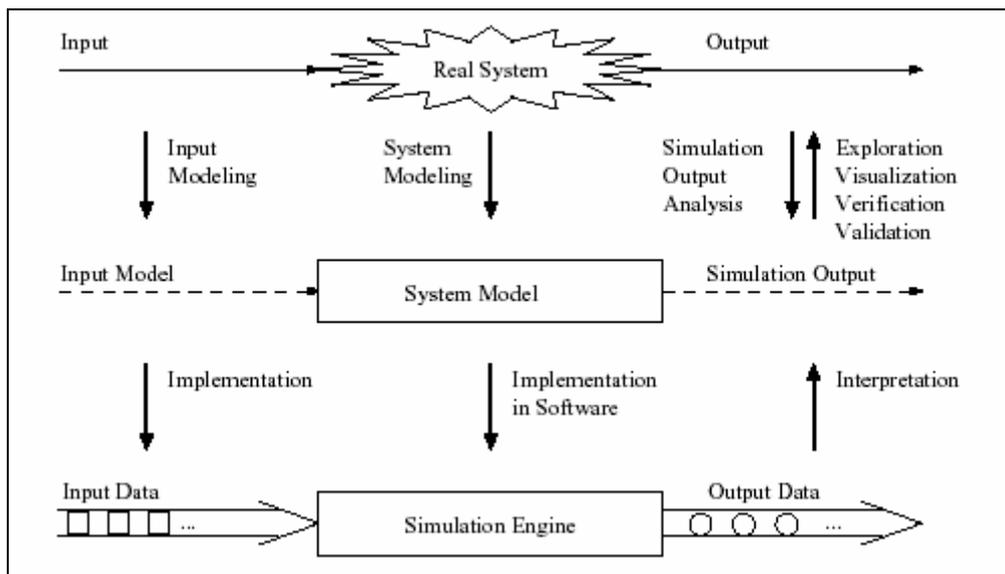


figura n° 44 Simulación de Computadora

Para el desarrollo del software de simulación es importante expresar adecuadamente un sistema real en un modelo del sistema, proceso de abstracción que se llama modelación del sistema. La entrada al sistema real requiere una representación apropiada para o dentro del modelo del sistema. La modelación de la entrada para una simulación del sistema es llamada en la literatura modelación de entrada. El análisis de la salida del modelo usualmente involucra investigación estadística de los datos de salida, la exploración de grandes espacios de datos, la visualización apropiada, y la verificación y validación de los experimentos de simulación (el análisis de la salida de simulación).

Una vez que un modelo de sistema, con el modelo de entrada ha sido desarrollado o definido, la implementación de esos modelos en software se vuelve actual. Generalmente, un motor de simulación o programa de simulación está implementado tomando datos de entrada para activar la ejecución, y para la generación de los datos de salida para ser interpretados como la salida de la simulación.

III.1.1 Modelos de Sistemas

Básicamente, un modelo de sistema es una especificación del fenómeno presente en el sistema físico real (o al menos alguno de sus componentes) en términos de un conjunto de estados y eventos, en un nivel apropiado de abstracción o detalle. En su software de implementación (motor de simulación), los estados son usualmente realizados como variables de estado, cuyos valores definen el estado actual del sistema simulado, mientras que los eventos son usualmente componentes de software (frecuentemente llamados “procedimientos de eventos”) que cambian los valores de las variables de estado, indicando el cambio de estado del sistema. Desde un punto de vista de la tecnología de software, realizar una simulación significa efectuar una “mímica” de la ocurrencia de los eventos de acuerdo a como ellos evolucionan, y reconociendo sus efectos representados por los estados.

La simulación, como un área de investigación y disciplina de ingeniería, ha establecido un espectro amplio de diferentes clases de modelos de sistemas. En las simulaciones de modelos de sistemas dinámicos la evolución de los estados del sistema es considerado en el tiempo, mientras que el tiempo no es una variable en una simulación de modelos de sistema estáticos. En las simulaciones orientadas al tiempo, la distinción primaria es con respecto a la métrica del tiempo y los parámetros de estado. El tiempo en los modelos de sistema de tiempo continuo, es considerado real, y el estado del sistema está definido en todo momento (es decir, el número de marcas en los lugares de una Red de Petri). Un modelo de sistema de tiempo discreto implementa el tiempo como una variable discreta, y el estado del sistema está definido en los instantes de tiempo respectivos únicamente. En forma apropiada, las variables de estado tienen valores continuos en un modelo de sistema de estado continuo, y los valores enteros en un modelo de sistema de estado discreto.

Con respecto a la relación entrada-salida de una simulación podemos distinguir modelos de sistema determinístico donde la salida de simulación puede ser predicha con certeza, desde los modelos de sistema probabilísticos, donde una y la misma entrada puede producir salidas diferentes. Si la salida de simulación tiene una dependencia lineal de la entrada, la literatura habla de modelos de sistemas lineales y modelos de sistemas no lineales en los otros casos.

III.2) SIMULADORES DE REDES DE PETRI

Se disponen de numerosas utilidades de simulación, con distintos alcances y funcionalidades. En general, son herramientas para el diseño, análisis y simulación de sistemas concurrentes y distribuidos. A causa de la complejidad de estos sistemas, es imposible usar técnicas ad hoc para obtener análisis cualitativo (corrección) y cuantitativo (performance). Las Redes de Petri Estocásticas Generalizadas son un formalismo que tiene la potencia necesaria para representar tales sistemas temporizados complejos, mientras que retienen la capacidad para realizar ambos tipos de análisis. Los GSPNs también tienen una representación gráfica conveniente. Usando este tipo de representación, los simuladores pueden permitirle al usuario:

- Crear y manipular interactivamente GSPNs,
- Formular modelos jerárquicos basados en subredes,

- Animar interactivamente el modelo resultante,
- Realizar análisis de corrección vía exploración del espacio de estado y análisis estructural, y
- Realizar análisis de performance vía análisis Markoviano y simulación.

Los simuladores disponen, en realidad, de tres componentes principales: el editor de Red de Petri gráfico, el analizador y el simulador. De la revisión de los paquetes de Redes de Petri existentes, los editores modernos permiten al usuario crear y modificar una GSPN en un ambiente interactivo amigable e intuitivo. Algunas de las características de varios editores incluyen la selección y manipulación de elementos de Redes de Petri en grupos o individualmente, una barra de herramientas para operaciones frecuentemente usadas, y la capacidad para exportar la red. Aparecen algunos problemas con los editores de Red de Petri existentes, cuando se necesita la especificación de sistemas complejos, que resultan en una muy gran GSPNs. Esto es a causa que no hay soporte para la modularidad o el reuso de componentes. Por ello, los nuevos productos soportan unidades jerárquicas llamadas subredes para resolver este problema. Los niveles más altos ven a las subredes como cajas negras, y se comunican con ellos a través de puntos de interacción. Alternativamente, puede recurrirse a formalismos para sistemas complejos usando simuladores específicos para Redes de Petri Coloreadas, o para las Redes de Petri FPQSPN.

El analizador de los paquetes existentes más elaborados provee análisis de corrección y performance. El análisis de corrección verifica las propiedades funcionales tales como la limitación, vivacidad, bloqueos mortales posibles, y la existencia de estados hogar. Habitualmente se usan dos técnicas para determinar la corrección funcional: el análisis invariante y la generación del grafo de alcance. El análisis invariante es una técnica basada estructuralmente que involucra la aplicación de álgebra lineal, en un algoritmo mínimo optimizado. La generación del grafo de alcance involucra una exploración de todos los estados posibles alcanzables; el grafo de estado resultante es entonces analizado, usando la teoría de grafos general. El analizador también realiza un análisis temporal para estudiar la performance del sistema. Esto involucra la generación de la cadena de Markov subyacente a la Red de Petri para realizar un análisis Markoviano. Los resultados incluyen número promedio de marcas por lugar, distribución de las marcas total, probabilidades de habilitación de transiciones, rendimiento de las transiciones y la distribución de estado estable.

El análisis del estado estable es imposible para las redes ilimitadas, o con un espacio de estado muy grande. Se incluye entonces un simulador que involucra una ejecución paso a paso del modelo de Red de Petri., mientras registra información acerca de su conducta temporal. En cada paso, el estado es alterado para primero determinar las transiciones habilitadas, y luego disparando una transición de acuerdo a la distribución de probabilidad de disparo de las transiciones habilitadas. Los resultados de performance provistos por el simulador son: número promedio de marcas por lugar, rendimiento de transición y trazado de salida. Un gran número de disparos se requiere para construir un intervalo confiable para estos resultados.

III.2.1 Simuladores existentes

En muy amplia la gama de simuladores, de libre disponibilidad y comerciales, con las siguientes características:

- Rango de Redes de Petri Soportadas
Redes de Petri de Alto Nivel,
Redes de Petri Temporizadas,
Redes de Petri Orientadas a Objeto,

- Redes de Petri Estocásticas,
- Redes de Petri Lugar/Transición,
- Redes de Petri Coloreadas,
- Redes de Petri Bien Formadas,
- Redes de Petri Estocásticas Generalizadas,
- Redes de Petri Jerárquicas,
- Redes de Petri FPQSPN, etc.
- Rango de Componentes Disponibles
 - Editor Gráfico,
 - Juego de Marcas,
 - Animación,
 - Simulación Rápida,
 - Análisis de Performance Simple,
 - Formato para intercambio de Archivos,
 - Espacio de Estados,
 - Invariantes de Lugar,
 - Invariantes de Transición,
 - Análisis Estructural,
 - Chequeo de Alcanzabilidad,
 - Chequeo de Bloqueo Mortal,
 - Análisis de Performance Avanzado,
 - Generación de Código,

Sólo a título ilustrativo, se pueden mencionar los siguientes productos: ALPHA/Sim, Artifex, CPN Tools, Design/CPN, HPSim, JARP, Netlab, Opera, Petrigen, Petri NET, PetriSim, PIPE, StpnPlay, TimeNET, Visual Object Net ++, WinSim, entre otros.

III.3) SIMULADOR DE ANIMACION PARA LA EVALUACIÓN DE LA DINAMICA [35]

Para el presente trabajo se ha seleccionado el simulador HPSim. Se trata de un simulador desarrollado por Henryk Anschuetz (Alemania), en el contexto de un estudio, completamente en C++ sobre Visual Studio. La adopción del mismo se debe a su simplicidad para su utilización, por la facilidad de instalación y bajos requerimientos hardware y software, soporte de ambientes operativos Windows, su libre disponibilidad y porque cubre los aspectos mínimos para permitir la animación de protocolos, aunque no provee análisis de performance avanzada. En HPSim es relativamente fácil construir los modelos básicos FPSQPN.

El simulador puede accederse gratuitamente para fines educativos y de investigación desde el sitio WEB: http://www.winpesim.de/petrinet/e/hpsim_e.htm.

El HPSim es una herramienta gráfica para Redes de Petri, y provee:

- Simulador: Animación de juego de marcas, simulación paso a paso o continua, simulación estocástica, salida de los vectores de marcado en un archive.
- Red Simulada: Red Lugar/Transición (rótulos no coloreados), Lugares con capacidad limitada, diferentes tipos de arcos con esos, transiciones temporizadas.

Sus componentes son:

- Graphical Editor
- Token Game Animation

- Fast Simulation
- Simple Performance Analysis

El HPSim puede ejecutarse en los ambientes operativos:

- PC, MS Windows 95
- PC, MS Windows 98
- PC, MS Windows NT
- PC, MS Windows 2000
- PC, MS Windows XP

Se muestra a continuación algunas imágenes del ambiente de trabajo de HPSim. En la figura n° 45 se muestra el simulador, dejando a la vista los menús desplegables: File, Edit, View, Window, Tools, Zoom, Simulation, y Extra, y con una Red de Petri ejemplo ya cargada.

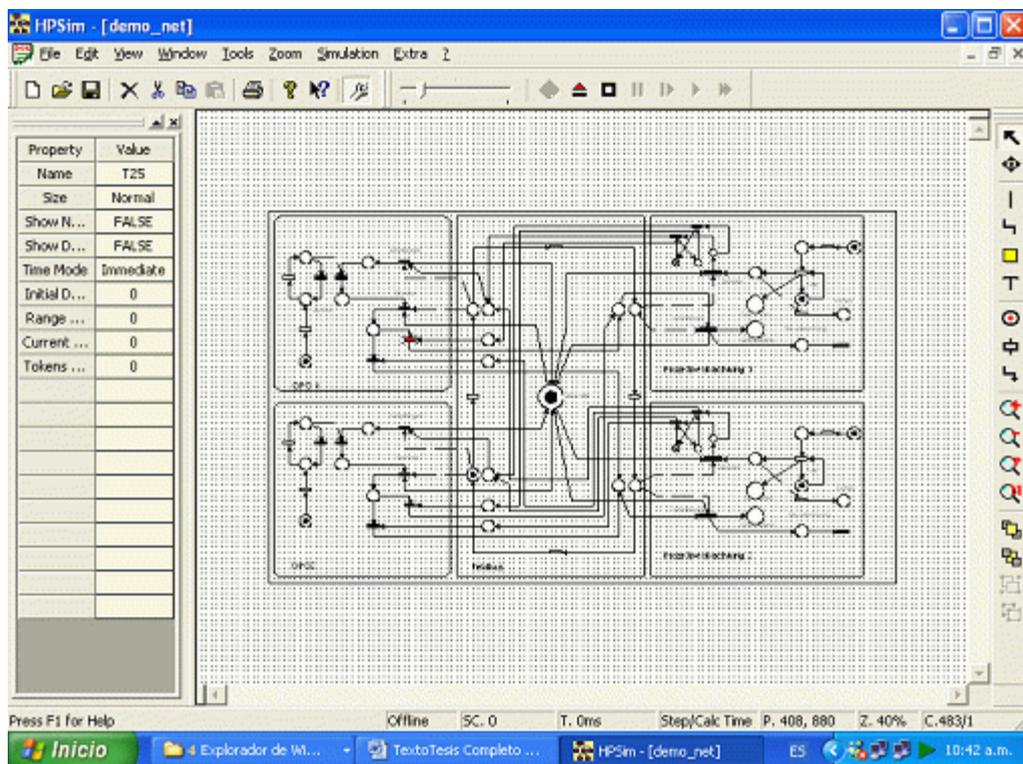


figura n° 45 Ambiente de trabajo del simulador HPSim

En la figura n° 46 se presenta desplegado el menú File, desde el cual se puede abrir, guardar, exportar, importar o imprimir el archivo que contiene la Red de Petri de trabajo; o cerrar el diseño actual o iniciar un nuevo diseño. Desde la figura n° 47 a 51 se muestran sucesivamente los desplegados del Menú Edit, del Menú Tools, del Menú Simulation, del Menú Simulation en Mode Sim, y el Ambiente de Simulación en Modo Paso a Paso, respectivamente.

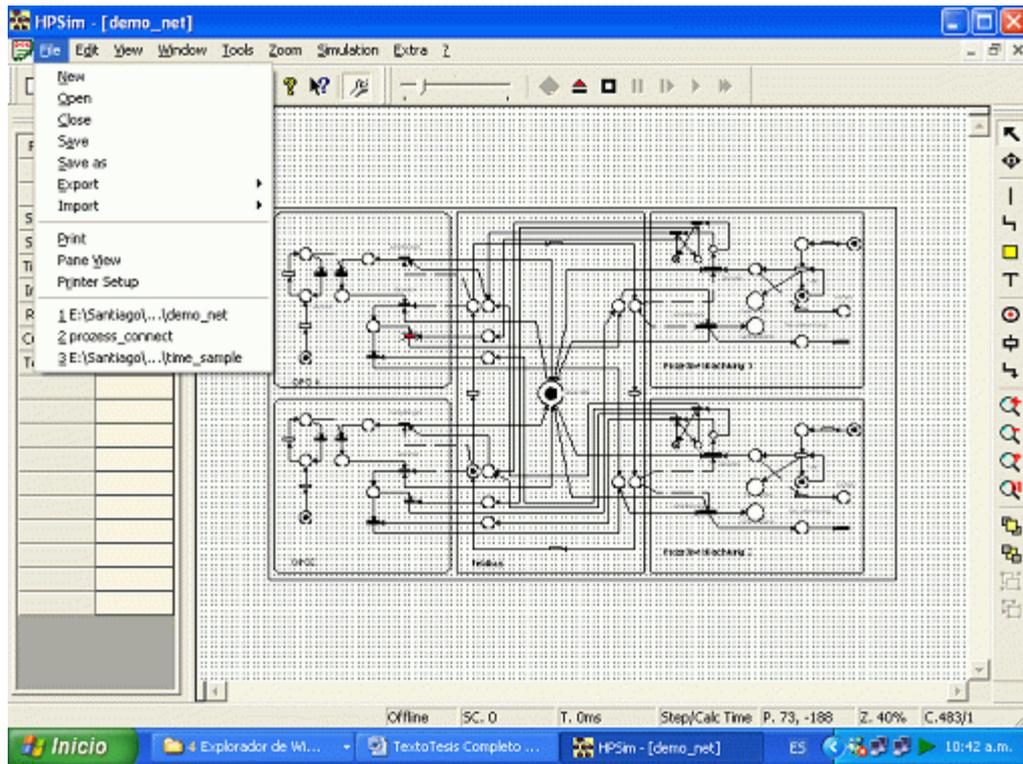


figura n° 46 Desplegado del Menu File

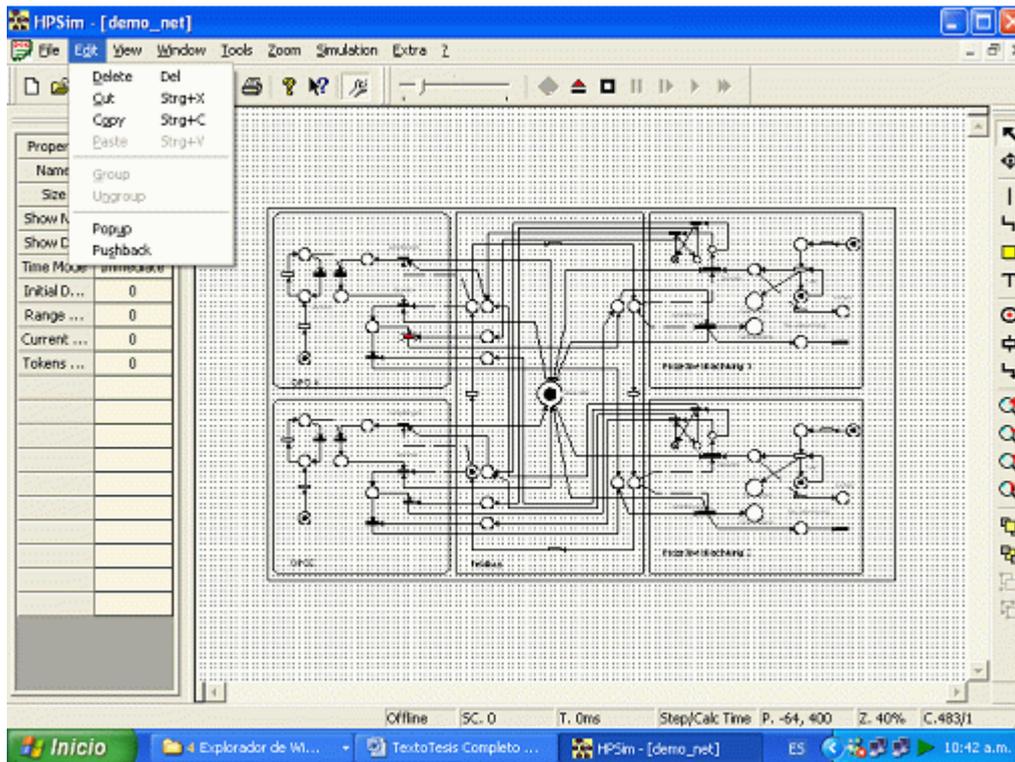


figura nº 47 Despliegado del Menú Edit

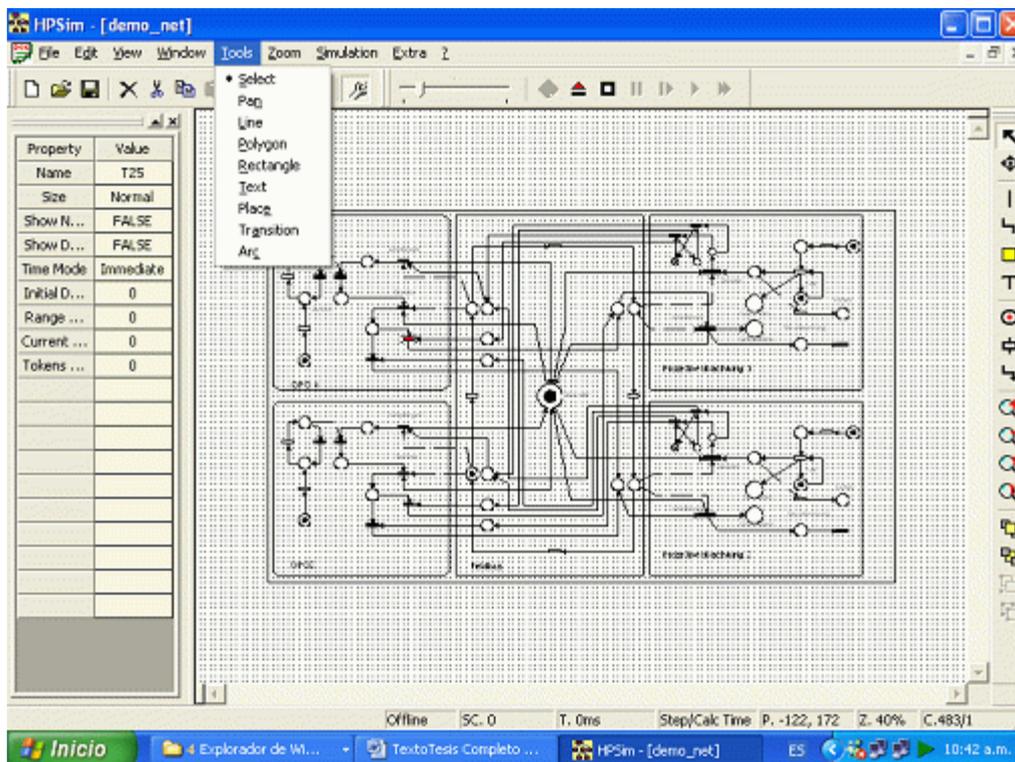


figura nº 48 Despliegado del Menu Tools

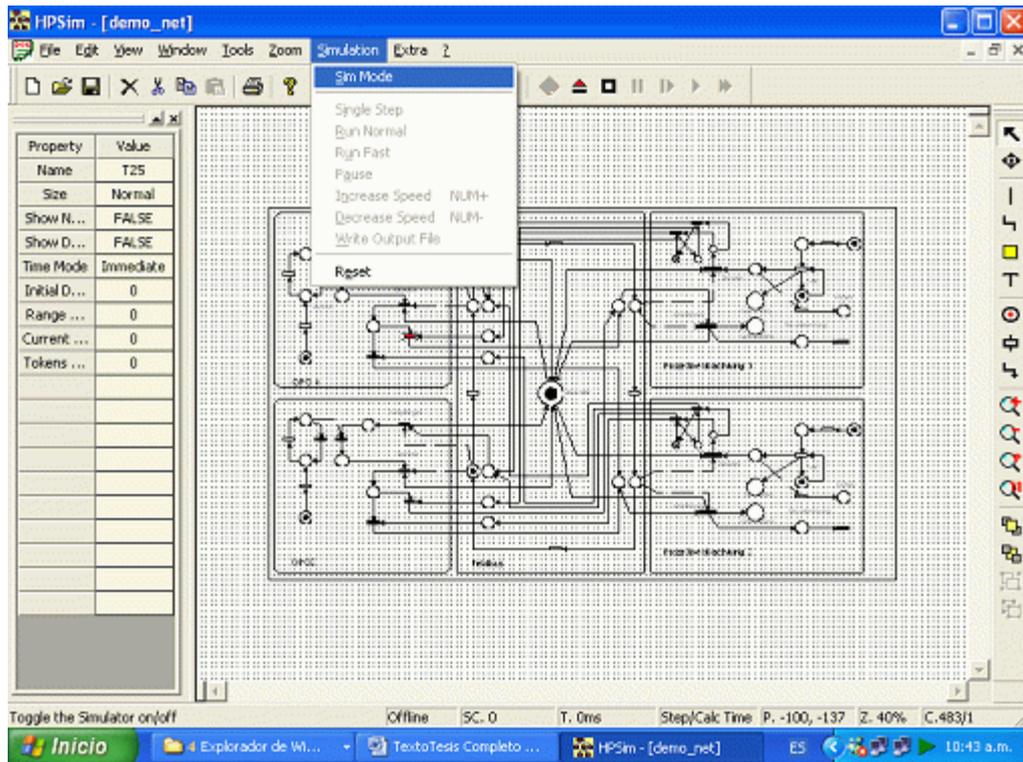


figura nº 49 Desplegado del Menú Simulation

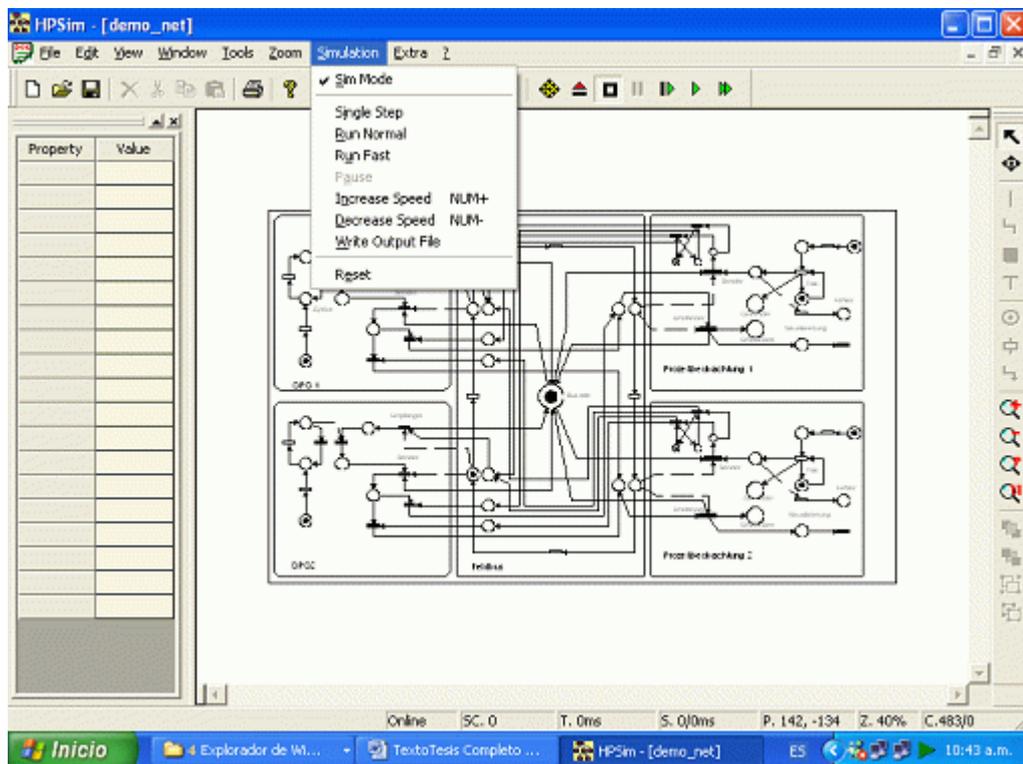


figura nº 50 Desplegado del Menu Simulación en Mode Sim

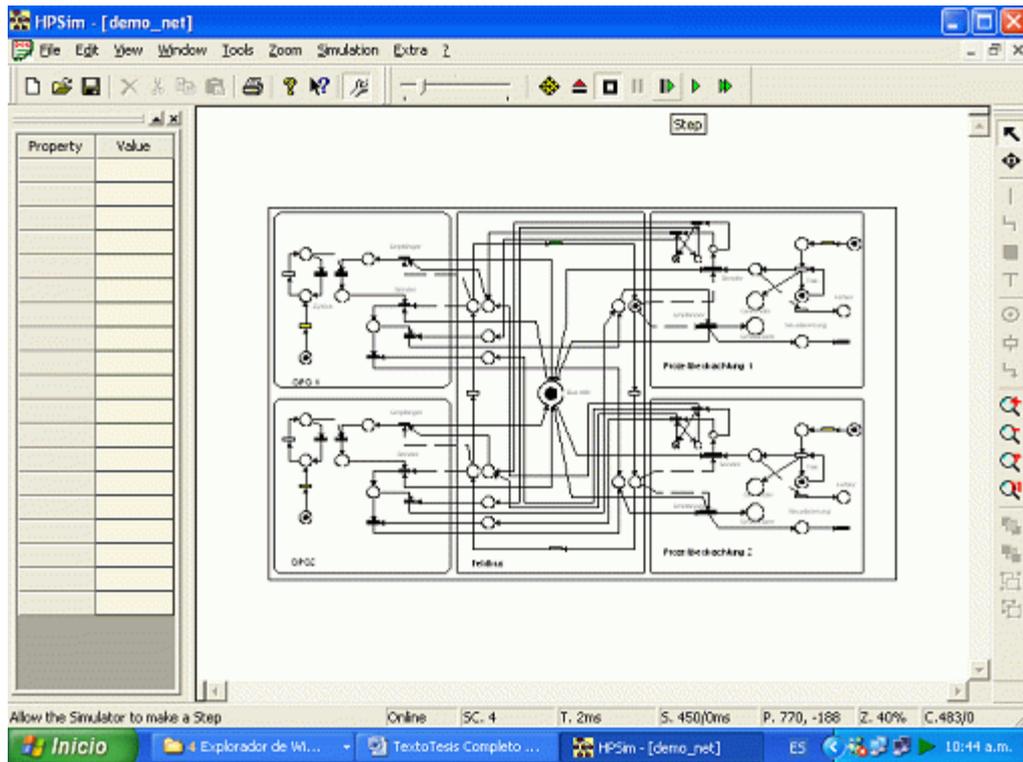


figura n° 51 Ambiente de Simulación en Modo Paso a Paso

III.4) SIMULACIÓN DE ANIMACIÓN DE ALOHA

III.4.1 Simulación de animación de Aloha Puro

A partir del modelo de Aloha puro presentado en II.2.2 (figura n° 30), las frecuencias de disparo de las transiciones tienen el siguiente significado:

- La cuenta de disparo de la transición T1 significa el tráfico ofrecido G,
- La frecuencia de disparo de la transición T2 representa la frecuencia de iniciaciones de la transmisión sobre el canal previamente inactivo,
- La frecuencia de disparo de la transición T3 representa el número de colisiones que aparecen sobre el canal que estaba previamente activo, pero en estado de no colisión, y
- La frecuencia de disparo de la transición T4 (la misma que T7) representa el número de colisiones que aparecen sobre el canal que estaba previamente en estado de colisión.

El número de paquetes exitosamente transmitidos iguala a la cuenta de disparo de la transición T5. La transición T6 es la única forma, que la marca retorna al lugar inicial P2 después de una colisión. Así, la cuenta de disparo de T6 representa el número de resolución de colisiones. Si asumimos $t_5 = t_6 = 1$, el tiempo medio asociado a la transición T1 representa la recíproca del tráfico ofrecido G.

Bajo estas condiciones, a continuación se describe y grafica la secuencia de estados posibles del canal para el modelo de Aloha puro, considerando la animación o movimiento de marcas que tendría el modelo usando el simulador HPSim. Se supone que hemos

arrancado el simulador HPSim, cargado el modelo de Red de Petri correspondiente a Aloha puro, ingresado en el menú de simulación y lanzado la animación. Nuestra intención es mostrar lo que haría el simulador, comenzando desde el estado de Canal Libre.

Canal Libre: En la dinámica propia de la simulación, se establece un punto de arranque o reset, que se configura en el simulador, y que ubica la marca en uno de los estados. En nuestro caso la marca se coloca en el lugar que representa el estado libre, desde el cual sea produce una evolución del sistema hacia los distintos estados posibles del modelo de Aloha puro, descrito precedentemente. La figura n° 52, con una marca en P2, corresponde al canal libre de tráfico.

Estado del Canal Libre

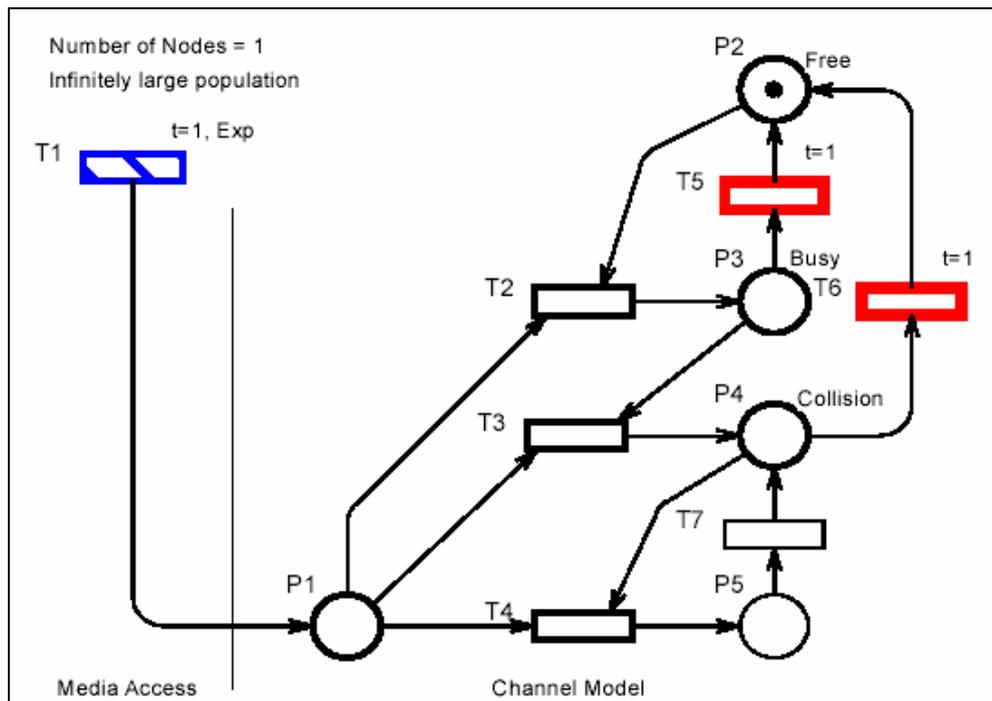


figura n° 52 Animación con Estado del Canal Libre

Canal Ocupado: Cuando el simulador está en marcha, emitirá sucesivamente en el tiempo, marcas desde la transición T1 con distribución exponencial. En decir, este caso ofrecemos un paquete a un canal libre (figura n° 53). El simulador pone una marca en el lugar P1 y se habilita la transición T2. Después del disparo, la transición consume las marcas de P1 y P2, y produce una marca en el lugar P3 (figura n° 54), cambiando el estado del canal de libre a ocupado. La marca reside en P3 hasta que se dispara la transición de tiempo deterministico T5 (considerado el tiempo de transmisión), o hasta que una nueva paquete arriba a P1.

Estado del Canal Ocupado

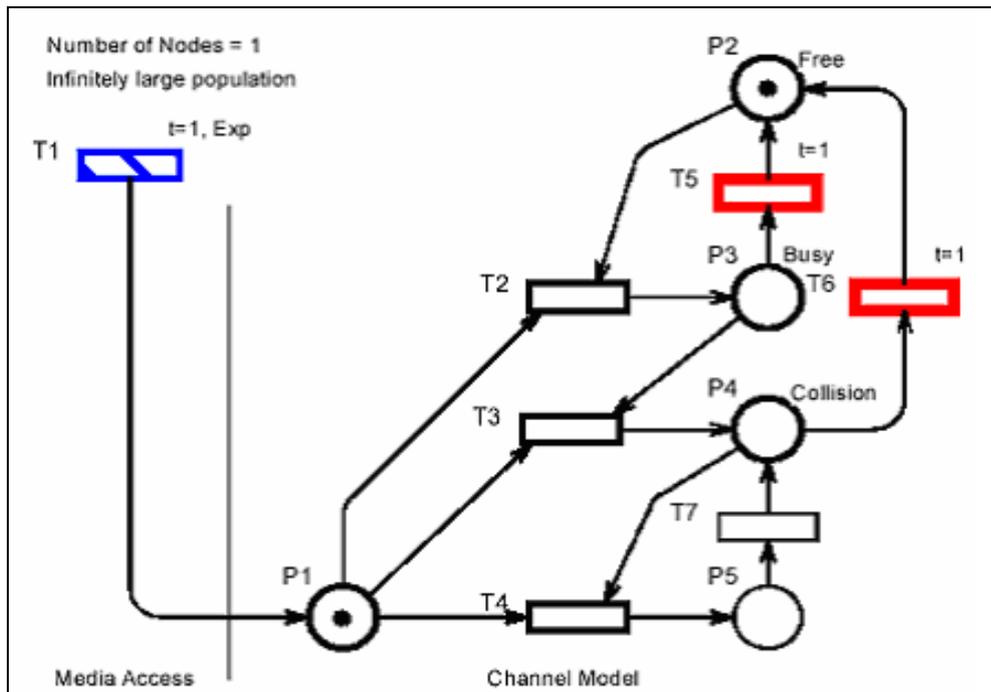


figura nº 53 Animación con Estado del Canal Libre

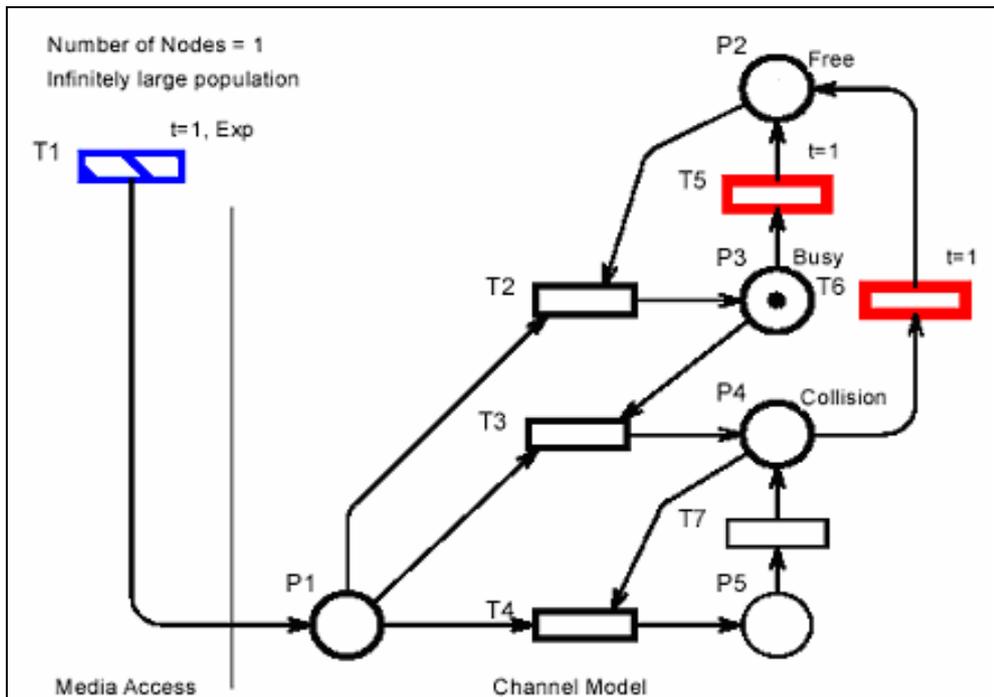


figura nº 54 Animación con Estado del Canal Ocupado

Canal en Colisión o Libre: Si el canal se encuentra ocupado y aparece un nuevo paquete en P1, la transición T3 se dispara y el canal cambia su estado de ocupado a colisión (figura nº 55). Si ningún otro paquete arriba durante el tiempo T6, el canal retorna al estado libre, luego de dispararse la transición T6 (figura nº 56).

Estado del Canal en Colisión

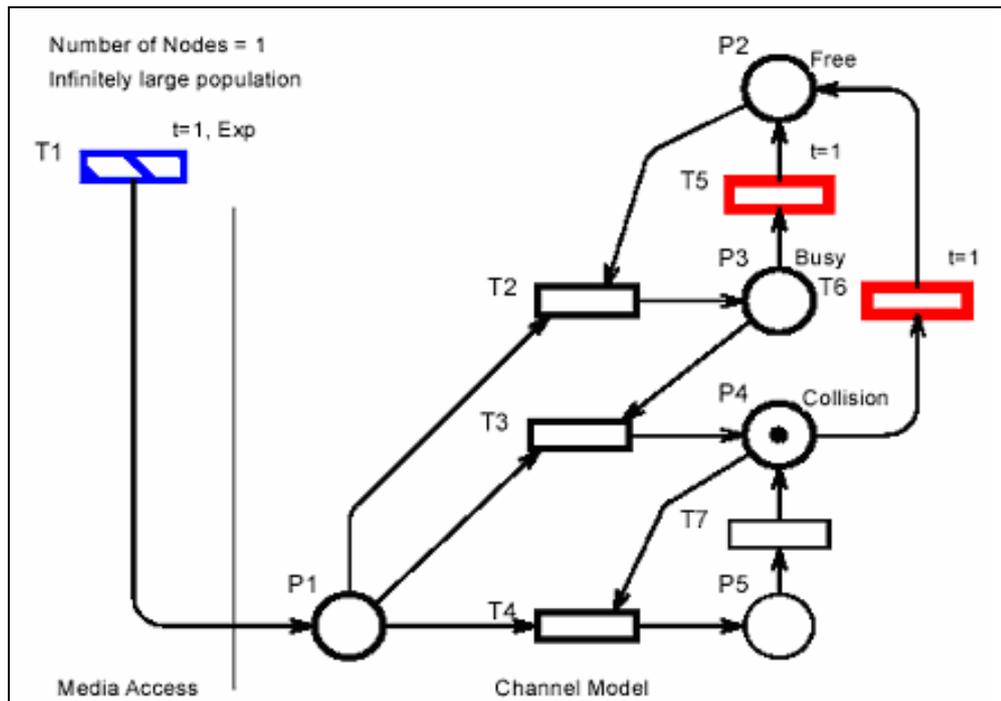


figura nº 55 Animación con Estado del Canal en Colisión

Estado del Canal Libre

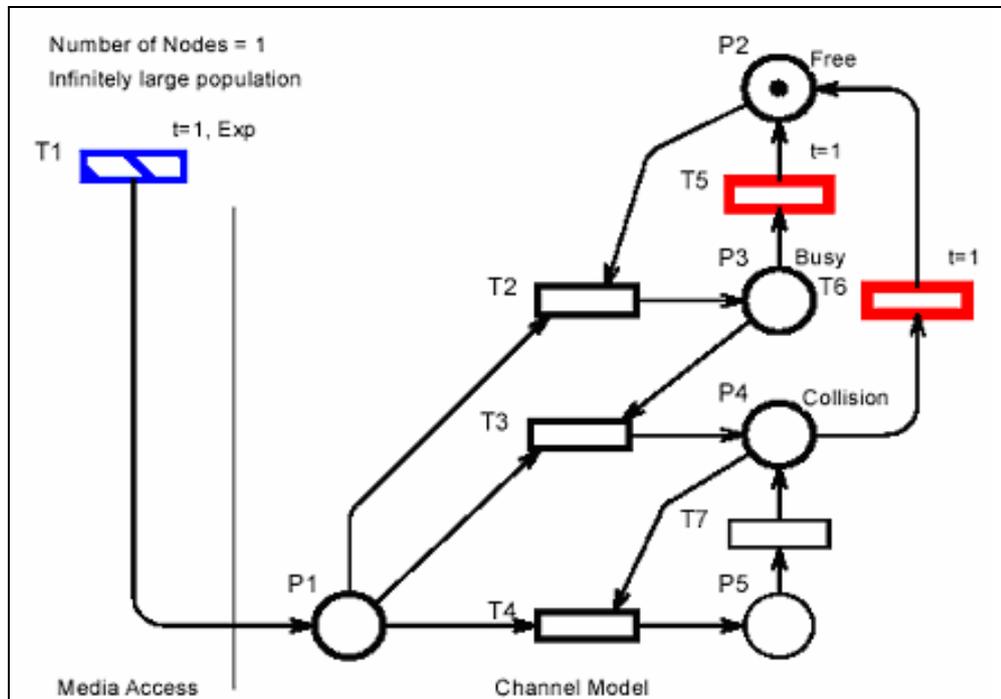


figura nº 56 Animación con Estado del Cana Libre

Estado del Canal Continúa en Colisión: En realidad este no es un nuevo estado; es el estado de colisión. Si el canal se encuentra en colisión, y aparece un nuevo paquete en P1

(figura nº 57), la transición T4 se dispara y la marca en el lugar P4 se mueve a P5. En este momento, la transición T6 está deshabilitada y su tiempo es inicializado después del disparo de T7.

Estado del Canal Continúa en Colisión

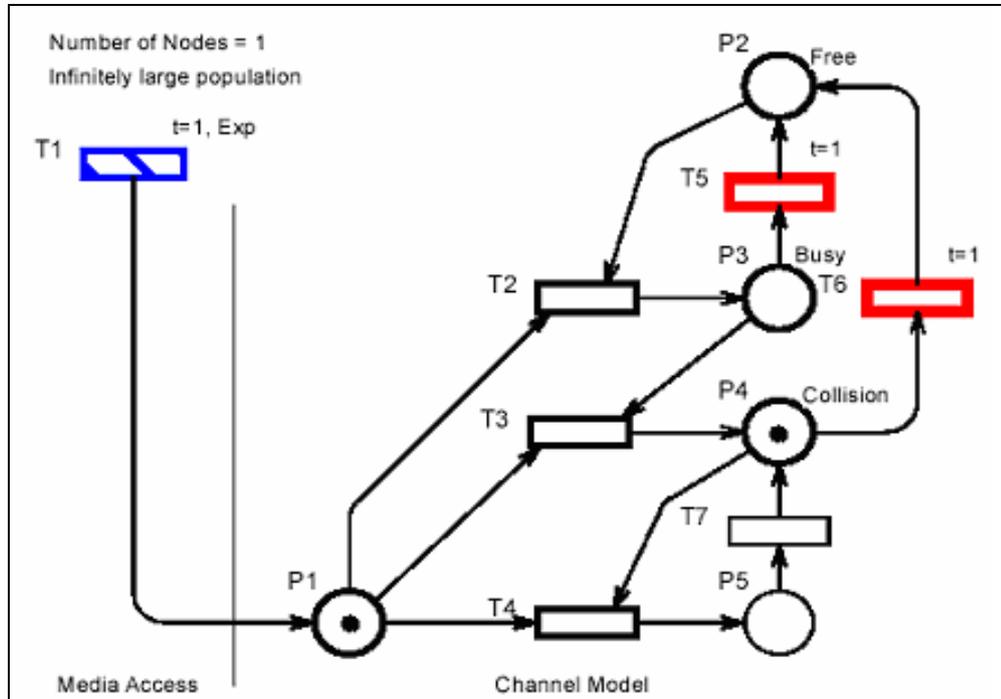


figura nº 57 Animación con Estado del Canal Continúa en Colisión

III.4.2 Simulación de animación de Aloha Ranurado

Teniendo en cuenta la descripción del modelo de Aloha ranurado realizada previamente en II.2.3, se concluye que la parte derecha del modelo de Red de Petri representa un canal de comunicaciones idéntico al de Aloha puro. Por lo tanto, en la figura nº 58, con una marca en P2, corresponde a un canal libre de tráfico. Y así cada uno de los estados ocupado, en colisión, y continúa en colisión discutidos precedentemente para Aloha puro son válidos para Aloha ranurado, por lo que sobreentendemos su análisis.

Nuevamente, se supone que hemos arrancado el simulador, cargado el modelo de Red de Petri correspondiente a Aloha ranurado, ingresado en el menú de simulación y lanzado la animación.

Estado del Canal Libre

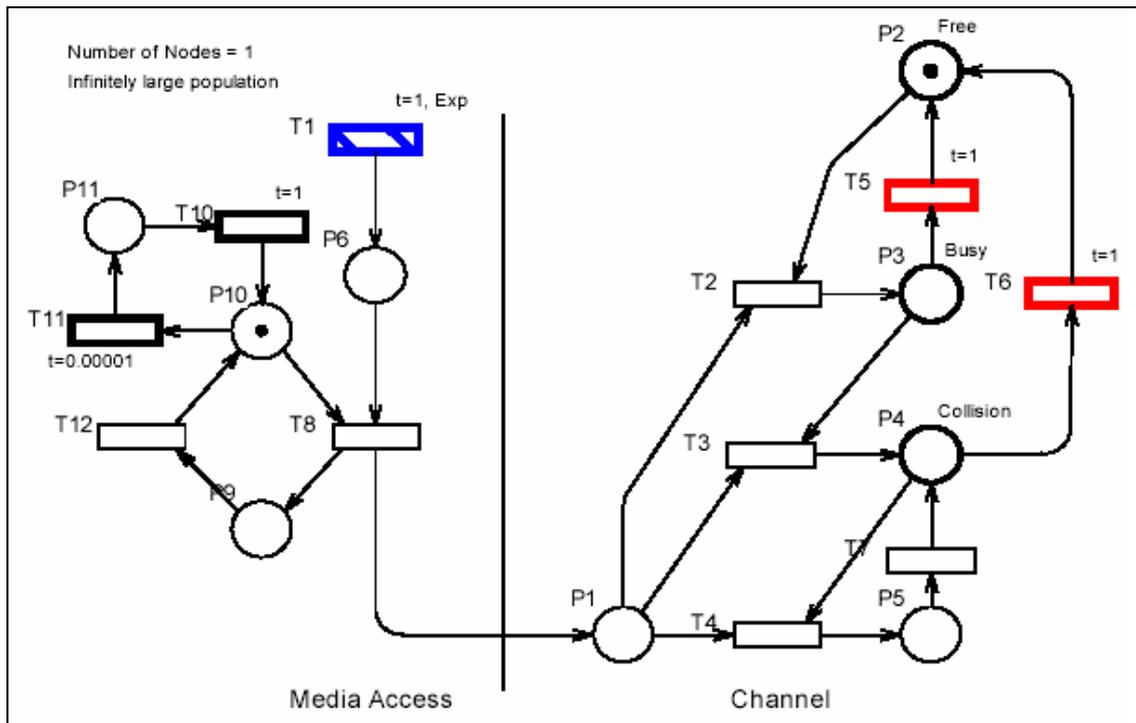


figura nº 58 Animación con Estado del Canal Libre

III.5) SIMULACIÓN DE ANIMACIÓN DE CSMA/CD

III.5.1 Simulación de animación de CSMA

Canal Libre: Se recuerda que el significado de los lugares P1, P2, P3, P4 y P5, y las transiciones T2, T3, T4, T5 y T6 es idéntico al significado de los lugares y transiciones correspondientes al modelo de tres estados de Aloha. El modelo muestra sus “estados visibles por medio de los nodos” P10 y P11. Una marca en el lugar P2 representa el estado libre y una marca en P11 representa “estado visiblemente libre” (figura nº 59).

Estado del Canal Libre

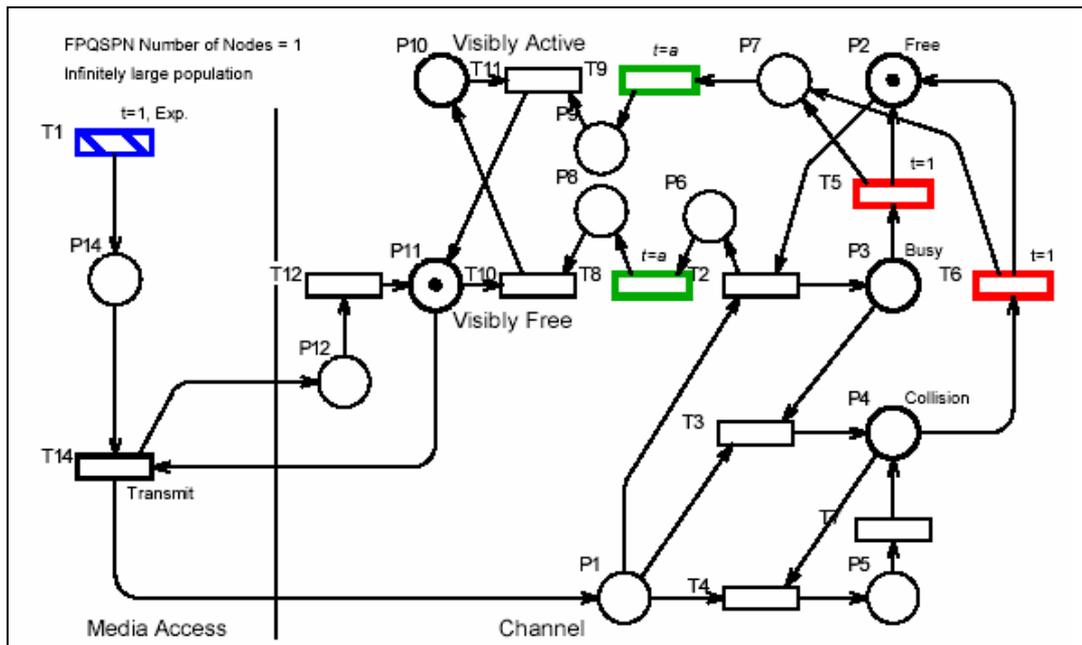


figura nº 59 Animación con Estado del Canal Libre

Canal Ocupado: El canal se vuelve ocupado al llegar una marca al lugar P1, indicando que un nodo ha entrado en transmisión, y producirse el disparo de la transición T2. En dicho momento se coloca una marca en el lugar P3, y el canal se vuelve “visiblemente activo” a segundos después al colocarse una marca en el lugar P10 (figura nº 60).

Estado del Canal Ocupado

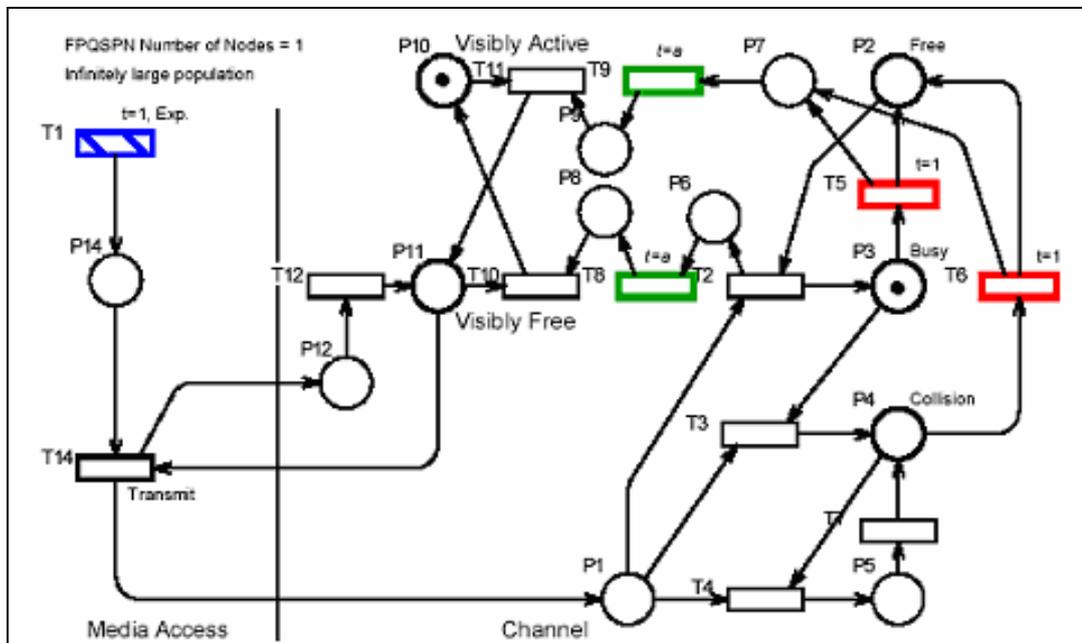


figura nº 60 Animación con Estado del Canal Ocupado

Canal Libre: El canal se vuelve nuevamente libre, a segundos después que el nodo que está transmitiendo corrientemente finaliza su transmisión (disparo de T5) (figura nº 61).

Estado del Canal Nuevamente Libre

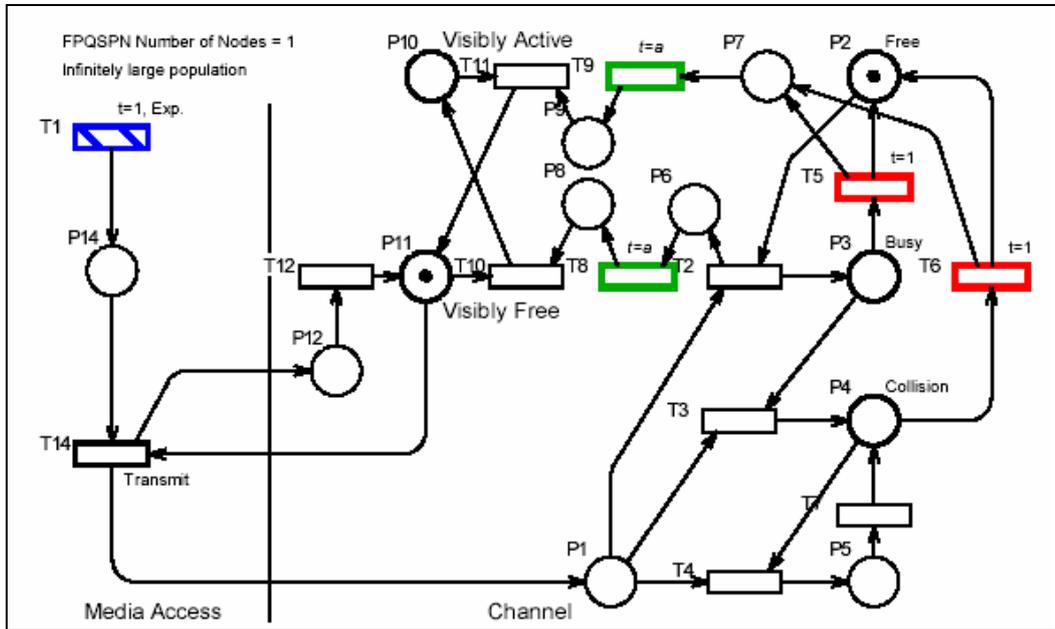


figura nº 61 Animación con Estado del Canal Nuevamente Libre

Canal en Colisión: Si el canal se encuentra ocupado y aparece un nuevo paquete (figura nº 62), la transición T3 se dispara y el canal cambia su estado de ocupado a colisión (figura nº 63).

Estado del Canal Ocupado

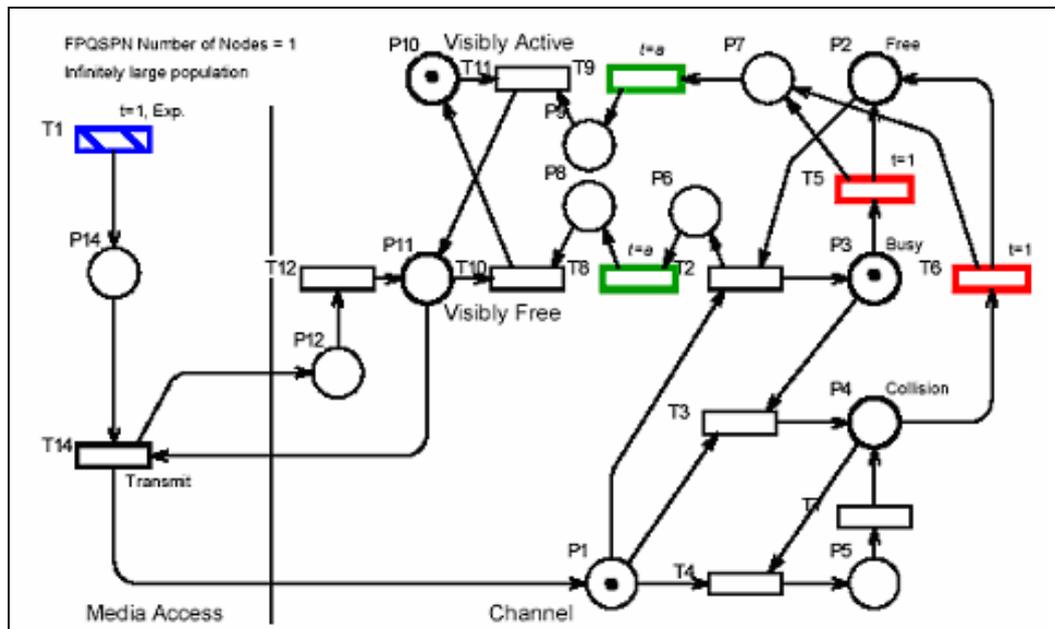


figura nº 62 Animación con Estado del Canal Ocupado

Estado del Canal en Colisión

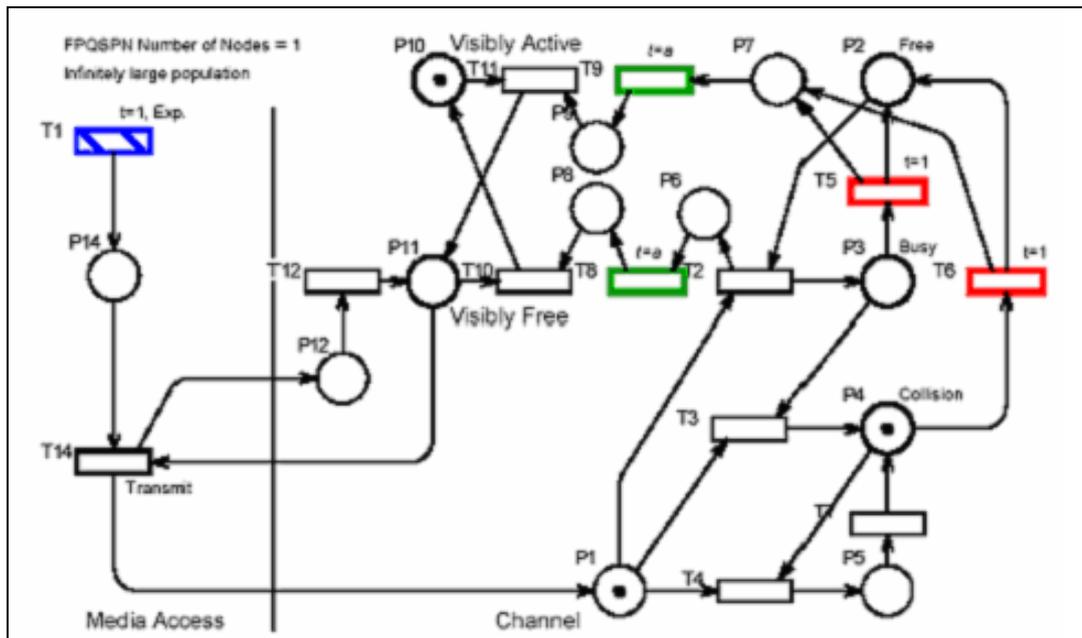


figura ° 63 Animación con Estado del Canal en Colisión

Canal Libre: Cuando el canal está en estado de colisión, se vuelve libre a segundos después que el último nodo participante en el conflicto para su transmisión (representado por el disparo de la transición T6)(figura n° 64).

Estado Libre

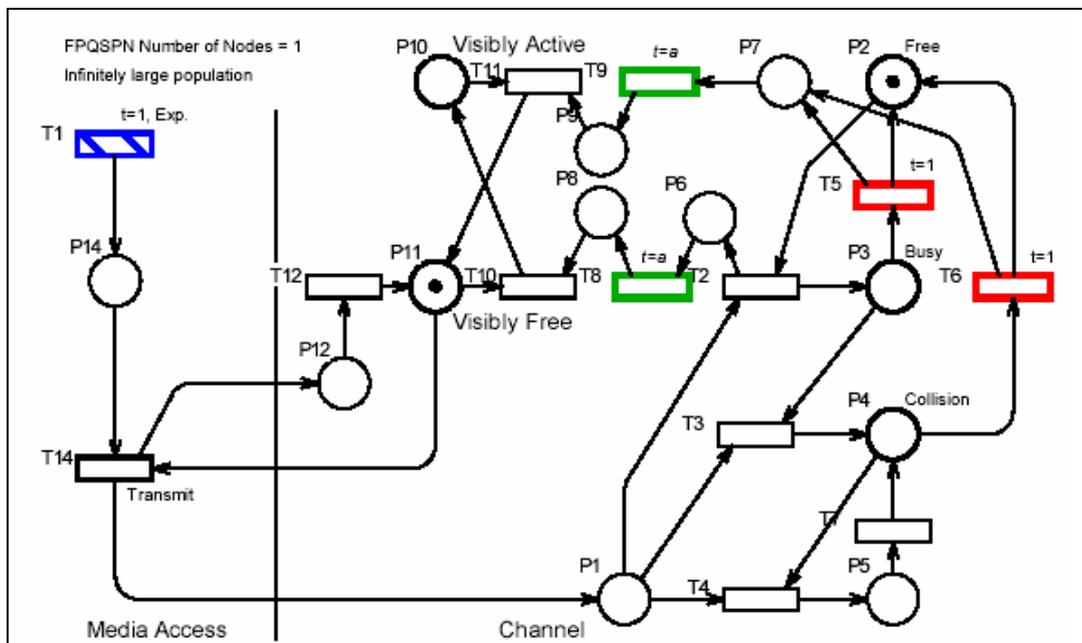


figura n° 64 Animación con Estado del Canal Libre

Estado del Canal Ocupado

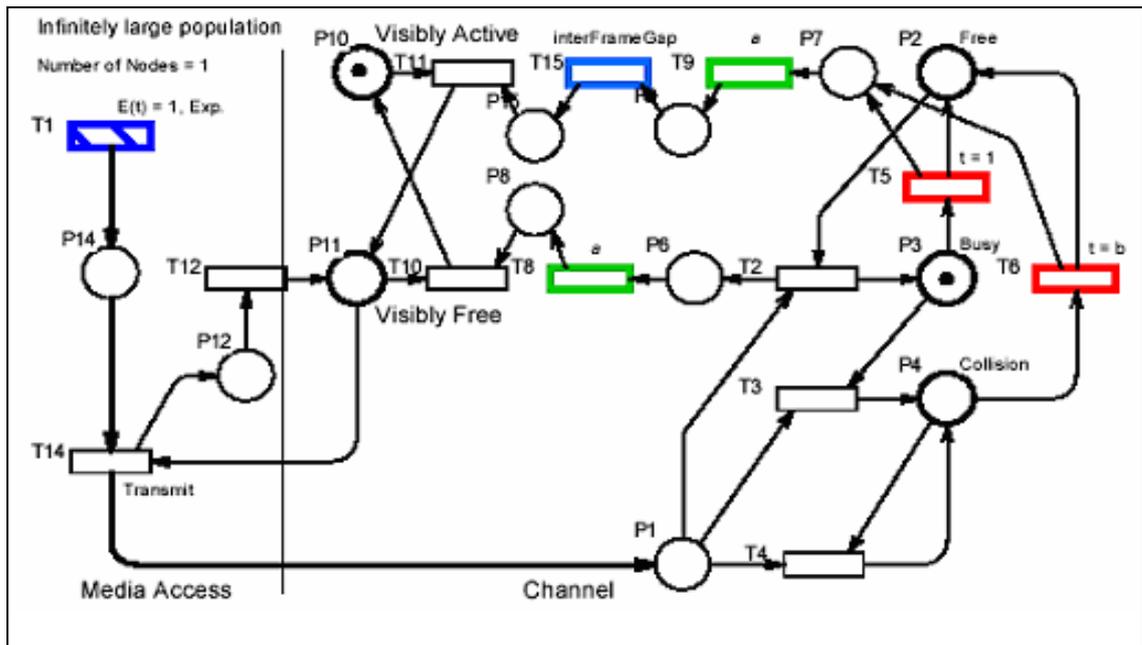


figura nº 66 Animación con Estado del Canal Ocupado

Canal en Colisión: Si el canal se encuentra ocupado y aparece un nuevo paquete (figura nº 67), la transición T3 se dispara y el canal cambia su estado de ocupado a colisión (figura nº 68).

Estado del Canal en Colisión

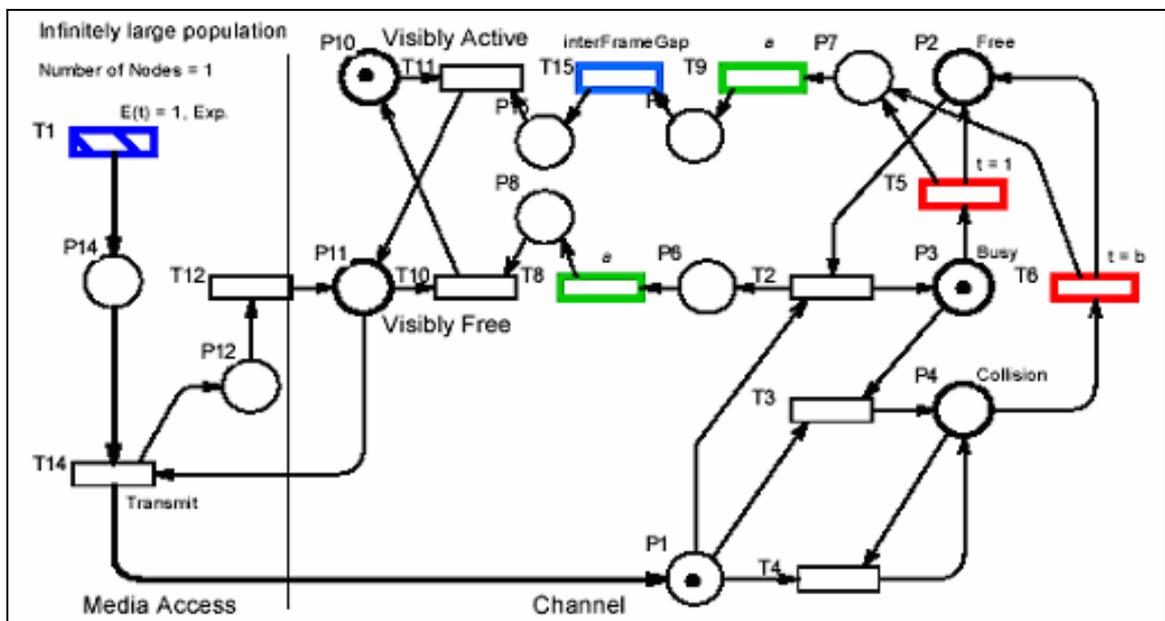


figura nº 67 Animación con Estado del Canal Ocupado

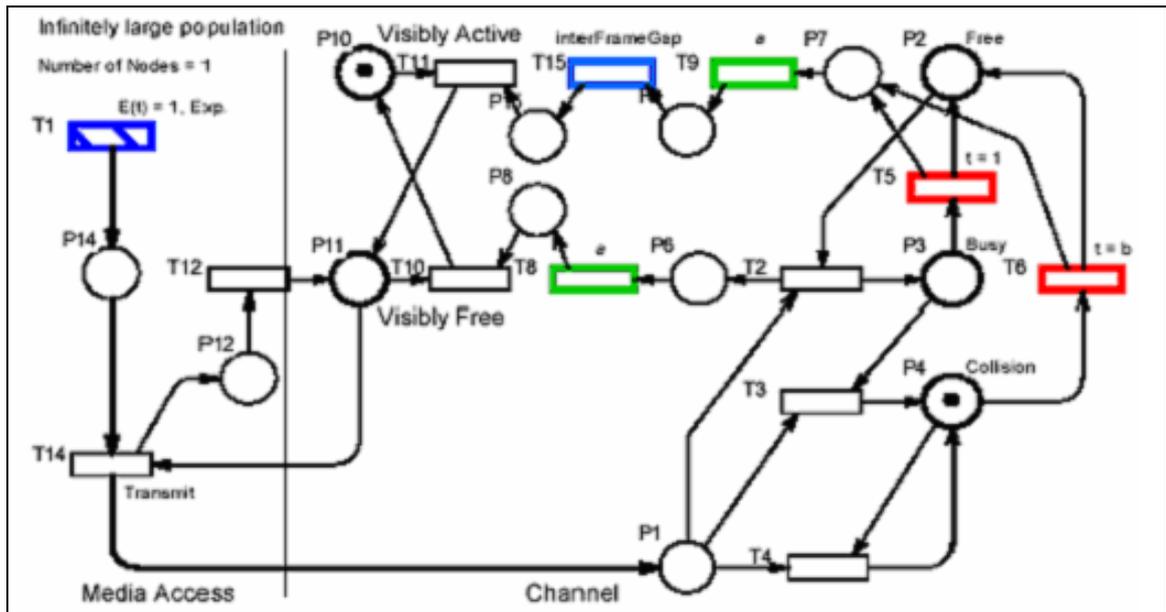


figura nº 68 Animación con Estado del Canal en Colisión

Canal Libre: Cuando el canal está en estado de colisión, se vuelve libre [a + inter-frame-gap] segundos después que el último nodo participante en el conflicto para su transmisión (representado por el disparo de la transición T6)(figura nº 69).

Estado del Canal Libre

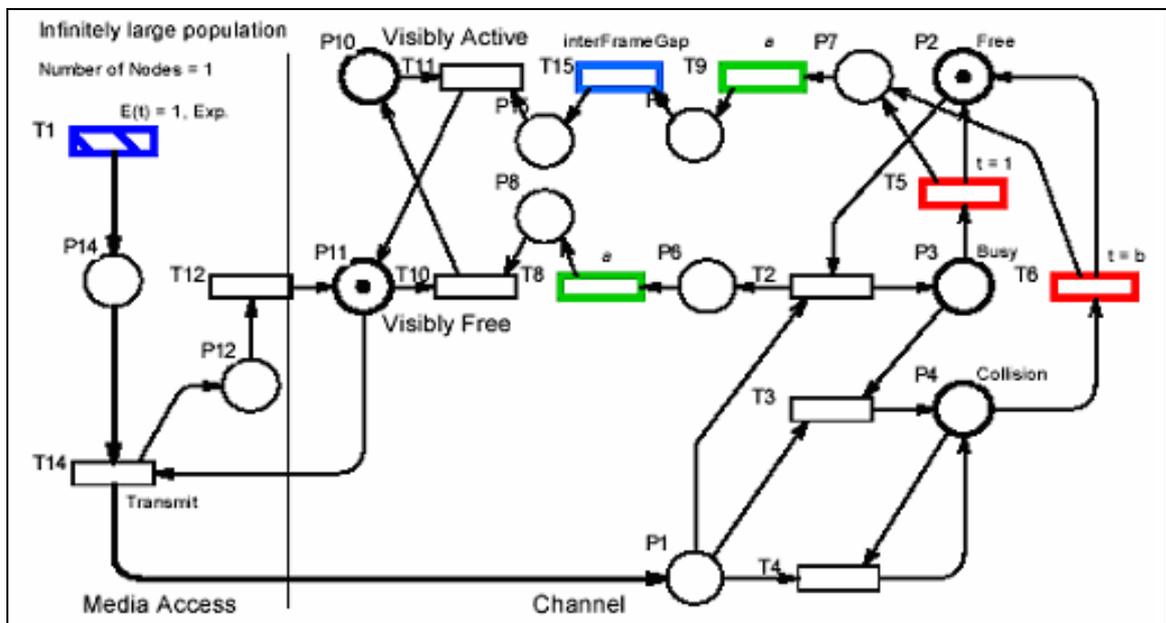


figura nº 69 Animación con Estado del Canal Libre

III.6) CONCLUSION

Dentro de los objetivos y la metodología a seguir en el trabajo de tesis, se ha seleccionado una herramienta para la simulación de la animación de protocolos, para observar la dinámica, investigar propiedades o características.

En este Capítulo se han presentado los aspectos teóricos sobre Principios de Simulación y una compilación de simuladores de Redes de Petri comerciales o de libre disponibilidad, definiendo sus características, funcionalidades y áreas de aplicación. Además, se ha esquematizado la secuencia de estados posibles que podría verse en un simulador de animación de dinámica para los protocolos.

De los compiladores, se ha adoptado el simulador HPSim por su simplicidad, por la facilidad de instalación y bajos requerimientos hardware y software, soporte de ambientes operativos Windows, su libre disponibilidad y porque cubre los aspectos mínimos para permitir la animación de protocolos. Y aunque no provee análisis de performance avanzada puede usarse para un rudimentario estudio de la cuenta y movimiento de las marcas en los diferentes estados.

CAPITULO IV

EVALUACIÓN DE PERFORMANCE DE PROTOCOLOS DE COMUNICACIONES

IV.1) PRINCIPIOS DE PERFORMANCE EN CANALES COMPARTIDOS [36]

IV.1.1 Introducción

De acuerdo a la clasificación tecnológica de las redes, éstas pueden ser redes broadcast o redes formadas por enlaces punto a punto. En el caso de las redes punto a punto, la información se envía a la máquina situada al otro lado del enlace, que está claramente identificada y el medio de transmisión normalmente está siempre disponible.

En las redes broadcast existe una nueva dificultad. Como el canal es compartido, es necesario habilitar mecanismos que permitan a cada host utilizar el medio para enviar tramas a la máquina destino. El hecho de compartir el canal generará conflictos o incluso pérdida de tramas en algunos casos, por lo que los protocolos deberán establecer mecanismos adecuados para resolver estos conflictos y permitir que los nodos retransmitan las tramas que no hayan podido ser enviadas correctamente.

Debido a esta mayor complejidad es que en las redes broadcast se suele dividir la capa de enlace dos subcapas: una inferior, que se ocupa de controlar la función de acceso al medio de transmisión, denominada subcapa MAC y la superior, llamada subcapa de control de enlace o LLC, que corresponde a las funciones de la capa de enlace comunes a todo tipo de redes.

Existen dos soluciones principales para resolver la problemática de cómo asignar el canal o medio físico compartido a los nodos de la red, una es la asignación estática del canal y la otra es la asignación dinámica. Las formas de asignación estática más utilizadas son Multiplexación por División de Frecuencias (FDM) y Multiplexación por División de Tiempo (TDM). La idea es repartir el canal entre las estaciones que lo requieren, y una vez asignado un ancho de banda o un slot de tiempo a una estación ésta lo tendrá reservado para su uso mientras no lo libere, independientemente de que lo use o no. Estos métodos no son aplicables a las redes broadcast debido a que la naturaleza del tráfico es impulsivo o de ráfagas, lo que hace que se pierdan las frecuencias o slots asignados cuando un host no desea transmitir, haciendo el rendimiento muy pobre cuando el número de estaciones es mucho menor que el número de asignaciones de frecuencia o tiempo. Por otro lado, si el número de hosts es mayor que el de asignaciones habrán estaciones que no podrán transmitir ni recibir datos.

La asignación dinámica del canal consiste en reservar el medio para transmitir sólo cuando sea necesario. Como el medio de comunicación es único las estrategias de competencia para acceder a él son variadas y asumen la ocurrencia de colisiones (que corresponde a la transmisión simultánea de dos tramas, los que se traslapan y producen una señal distorsionada en el medio); sin embargo, existen esquemas que evitan las colisiones. Otras estrategias utilizan el tiempo continuo para transmitir, es decir, transmiten cuando lo necesitan, y otras usan slots de tiempo o transmiten por intervalos. Existen métodos que introducen mejoras que permiten al adaptador de red detectar la colisión y esperar un tiempo para retransmitir la trama. Los que no detectan colisión deben esperar la respuesta de si la transmisión fue exitosa o no para reenviar la trama si es necesario. En resumen, las reglas que especifican como se resuelve la situación de acceder el medio compartido se llaman protocolos de acceso al medio. La asignación estática tiene la desventaja de que se puede subutilizar el canal cuando las transmisiones son impulsivas o de ráfagas de datos. La asignación dinámica se realiza sólo cuando sea necesario transmitir, pero pueden ocurrir colisiones producto de la contienda por el

canal y se necesitan definir estrategias para llegar a un acuerdo entre los nodos en cuanto a la utilización del canal.

IV.1.2 Performance de los protocolos de acceso al medio

Por lo expuesto, la selección del protocolo a utilizar define, a partir de sus características, un rendimiento máximo. De los distintos modelos de Redes de Petri FQPSPN para los protocolos presentados en el Capítulo II, se deduce gráficamente que hay una complejidad creciente desde el modelo básico de Aloha puro. El incremento de la complejidad del protocolo seguramente tiene asociado la búsqueda de una mejora en el rendimiento del canal. (Se discutirán los métodos para calcular el rendimiento de ciertos protocolos con uso de un simulador de FPQSPN y en forma analítica a partir del apartado IV.2)

Brevemente, repasaremos en forma descriptiva el comportamiento de alguno de estos protocolos:

Aloha Puro: Cuando un host desea transmitir, simplemente emite una trama, sin preocuparse en ningún momento del estado del canal. Una vez finalizado, queda en espera de recibir la confirmación de que la información ha sido recibida correctamente por el destinatario, quien verifica esto usando el campo CRC de la trama. Si pasado un tiempo no se recibe confirmación, el emisor supone que ha ocurrido una colisión y espera un tiempo aleatorio y a continuación reenvía la trama. El problema principal del protocolo es que el envío de trama por parte de los nodos se hace en forma caótica y basta que dos tramas colisionen, o se solapen solamente en un bit para que ambos sean inútiles y deban retransmitirse, puesto que los nodos sólo se percatarán del problema después de haber terminado la transmisión. Por otro lado, el segundo trama podría colisionar con un tercero, y así sucesivamente, de ahí que en una red ALOHA cuando el tráfico crece las colisiones aumentan de manera no lineal y el rendimiento decae rápidamente. El rendimiento máximo de ALOHA es de 18,4%, que se consigue con un tráfico ofrecido de $G=0,5$.

Aloha Ranurado: Es una mejora a ALOHA que consiste en dividir el tiempo para la emisión de las tramas en intervalos de duración constante del tamaño de una trama. Además, los nodos deben sincronizarse para saber cuando empieza cada intervalo. Esto reduce la probabilidad de colisión, ya que limita el efecto de colisión a un intervalo concreto, y no se pueden encadenar colisiones. En Aloha ranurado, la eficiencia máxima es de 36,8% y se consigue con un tráfico ofrecido de $G=1$.

CSMA: Estos protocolos antes de transmitir observan si alguien ya está transmitiendo, lo que permite hacer un uso más eficiente del canal ya que no se interrumpe la transmisión que está en proceso. El nombre genérico de estos protocolos es de Acceso Múltiple con Detección de Portadora. Tiene algunas variantes como el protocolo CSMA 1-persistente que funciona de la siguiente forma: cuando tiene que transmitir una trama, primero escucha el canal y si está libre envía la trama, caso contrario, espera a que se libere y en ese momento la envía. Se denomina CSMA 1-persistente porque existe la probabilidad 1, es decir, certeza de que la trama se transmitirá cuando el canal esté libre. En una situación real con alto tráfico es muy posible que cuando un nodo termine de transmitir existan varios esperando enviar sus datos, y con CSMA 1-persistente todas las tramas serán emitidos a la vez y colisionarán, pudiéndose repetir el proceso varias veces con la consiguiente degradación del rendimiento. Cabe señalar que una colisión ocurrirá aunque no empiecen a transmitir exactamente a la vez, basta simplemente con que dos nodos empiecen a transmitir con una diferencia de tiempos menor que la distancia que los separa, ya que en tal caso ambos detectarán el canal libre en el momento de iniciar la transmisión. Se deduce entonces, que en este tipo de redes el retardo de propagación de la señal puede tener un efecto importante en el rendimiento. El rendimiento obtenido con este protocolo

puede llegar al 55% con un tráfico de $G=1$. Otra variante es el CSMA no persistente que funciona de la siguiente manera: antes de enviar se escucha el canal, si el canal está libre se transmite la trama. Si está ocupado, en vez de quedar escuchando, se espera un tiempo aleatorio después del cual se repite el proceso. El protocolo tiene una menor eficiencia que CSMA 1-persistente para tráficos moderados, pues introduce una mayor latencia; sin embargo se comporta mejor en situaciones de tráfico intenso ya que evita las colisiones producidas por las estaciones que se encuentran a la espera de que termine la transmisión de una trama en un momento dado, y puede alcanzar el 82 % con una carga $G=10$. Finalmente para CSMA p-persistente, el protocolo utiliza intervalos de tiempo y funciona de la siguiente manera: cuando el nodo tiene algo que enviar primero escucha el canal, si está ocupado espera un tiempo aleatorio. Cuando el canal está libre se selecciona un número aleatorio con distribución uniforme entre 0 y 1, si el número seleccionado es menor que p la trama es transmitida, si se espera el siguiente slot de tiempo para transmitir y repite el algoritmo hasta que la trama es transmitido o bien otro nodo utiliza el canal, en cuyo caso se espera un tiempo aleatorio y empieza de nuevo el proceso desde el principio. La eficiencia del protocolo es, en general, superior a la de CSMA 1-persistente y CSMA no persistente.

CSMA/CD: Un problema con los protocolos anteriores es que una vez se ha empezado a transmitir una trama se sigue transmitiendo aún cuando se detecte una colisión. Como es más eficiente parar de transmitir y esperar un tiempo aleatorio para volver a hacerlo, ya que la trama será errónea de todas formas, los protocolos de Acceso Múltiple con Detección de Portadora y Detección de Colisiones o CSMA/CD, implementan esta mejora. En una red CSMA/CD la única circunstancia en la que puede producirse una colisión es cuando dos hosts empiezan a transmitir a la vez, o con una diferencia de tiempo lo bastante pequeña como para que la señal de uno no haya podido llegar al otro antes de que éste empiece a transmitir.

IV.2) SIMULADOR PARA EVALUACION DE PERFORMANCE [37]

IV.2.1 Introducción

Se seleccionó la herramienta interactiva para la modelación, el análisis de rendimiento y la simulación de sistemas usando Redes de Petri FPQSPN, que se corresponde con los modelos del Capítulo II. Se trata del simulador STPNPlay y modelos desarrollados por Josef Capek [21], del Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague (Checoslovaquia). Es una herramienta de simulación para modelación interactiva que funciona en el entorno Windows. El énfasis en su desarrollo ha sido la exactitud y la velocidad del algoritmo de simulación para calcular performance, y la portabilidad de los modelos del sistema y los resultados a otras herramientas y plataformas. El simulador puede accederse gratuitamente para fines educativos y de investigación desde el sitio WEB: <http://dce.felk.cvut.cz/capekj/StpnPlay/>. Esta herramienta no permite la animación del protocolo, por lo que es imposible ver la dinámica del protocolo como en HPSim.

El simulador STPNPlay fue principalmente desarrollado para modelar y analizar protocolos de comunicaciones. La versión 0.65 permite generar características de performance (soporta coordenadas lineales y logarítmicas) de protocolos de comunicaciones (por ejemplo, modelo de carga de Poisson, población infinita), características de performance de red consistente del número de nodos máximo (máximo 1000 nodos CSMA) bajo varias funciones de carga. El algoritmo de simulación rápida dispara 500.000 transiciones por segundo, y su complejidad es independiente del número de elementos de la Red de Petri.

Cuando se analizan protocolos de comunicación, el usuario crea primero un modelo en Redes de Petri de Sistemas de Cola de Población Finita (FPQSPN), el que es una abreviación para permitir modelar múltiples nodos conectados a un canal de comunicaciones

compartido. Luego, el usuario designa las transiciones que representan la carga del sistema, e identifica las transiciones que representan las variables dependientes deseadas de las características evaluadas. Finalmente, el usuario configura los parámetros de simulación y comienza los procesos de simulación. Algunos procesos de simulación que tienen la misma o diferentes prioridades, pueden correrse en el mismo momento.

El programa STPNPlay consiste de tres módulos: el Editor de Red de Petri, que contiene las funciones para crear y editar el diagrama de Red de Petri; el simulador de la Red de Petri que permite la evaluación de performance de los modelos de Red de Petri, y el módulo de matriz para las operaciones con las matrices de incidencia (figura n° 70).

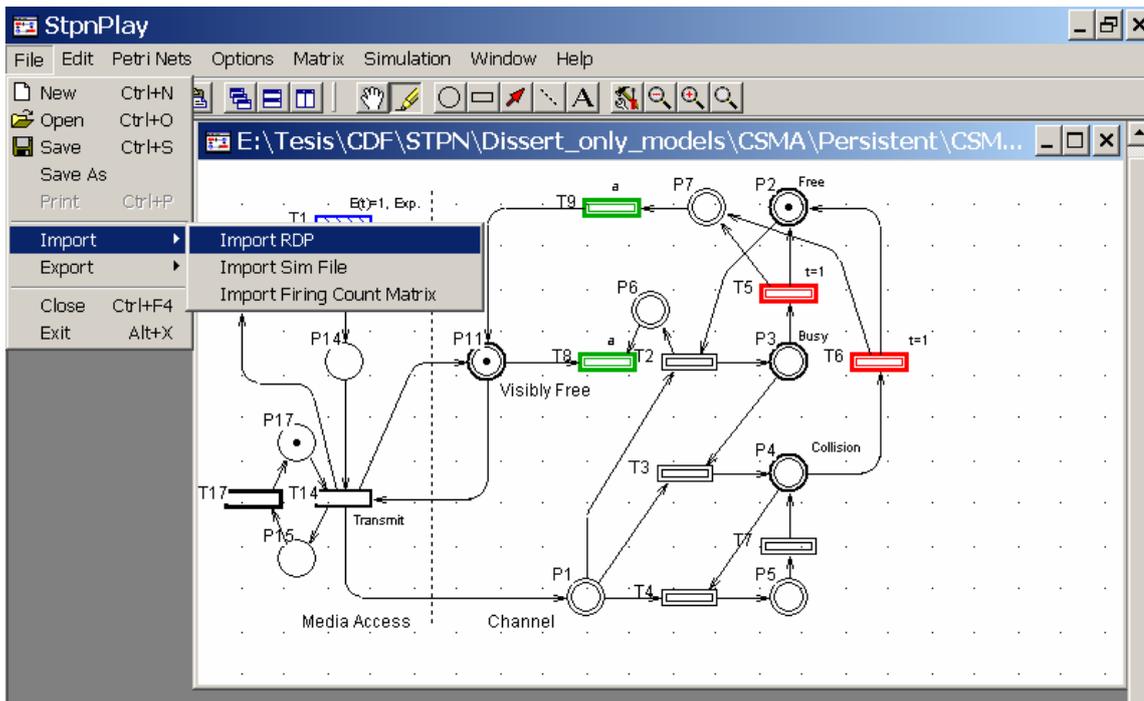


figura n° 70 STPNPlay: Menú de File

IV.2.2 Editor de Red de Petri

El Editor de la Red de Petri (figuras n° 71 y 72) permite dibujar objetos de la Red de Petri como lugares, transiciones y arcos. El Editor también permite dibujar dos tipos de objetos auxiliares: las líneas usadas para separar partes del diagrama y objetos de texto. Los objetos de Red de Petri tienen las siguientes propiedades:

- Transición:
 - Número: Número de transición en la red. El número es asignado automáticamente cuando se crea el objeto. El usuario puede modificarlo a otro número no existente.
 - Tiempo: Discutido luego.
 - Color: Es el color del rectángulo que representa el objeto. El usuario puede directamente seleccionar un color Standard o usar un color arbitrario de la paleta de Windows.
 - X-size, Y-size: Las dimensiones del rectángulo que representa el objeto.
 - Ancho: El ancho de línea del rectángulo
 - Compartido: El usuario puede seleccionar si el objeto aparece en nodos múltiples de la Red de Petri, o si el objeto representa una parte común para todos los nodos.

Estilo de llenado: El estilo es usado para llenar el rectángulo de las transiciones. Se soportan múltiples estilos standard.

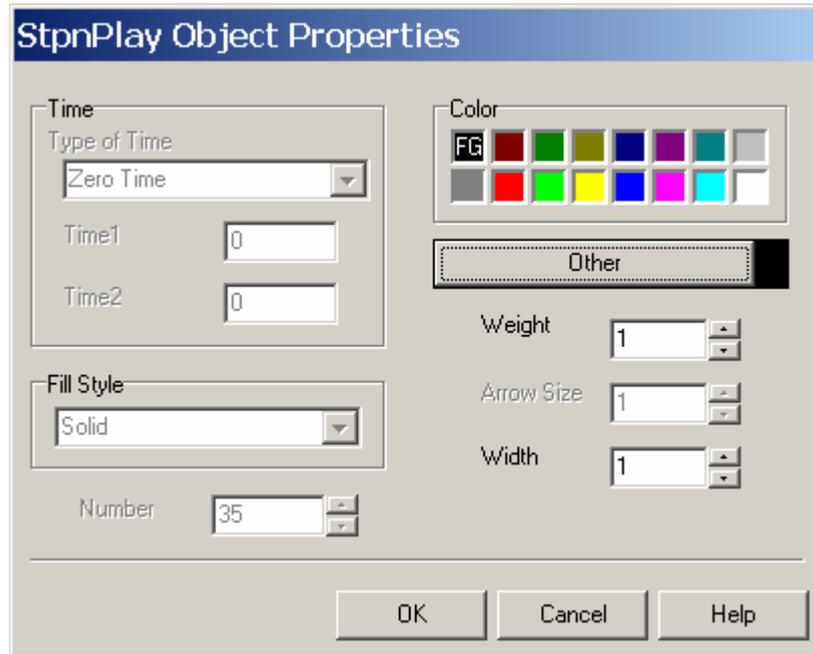


figura nº 71 STPNPlay: Propiedades de transiciones

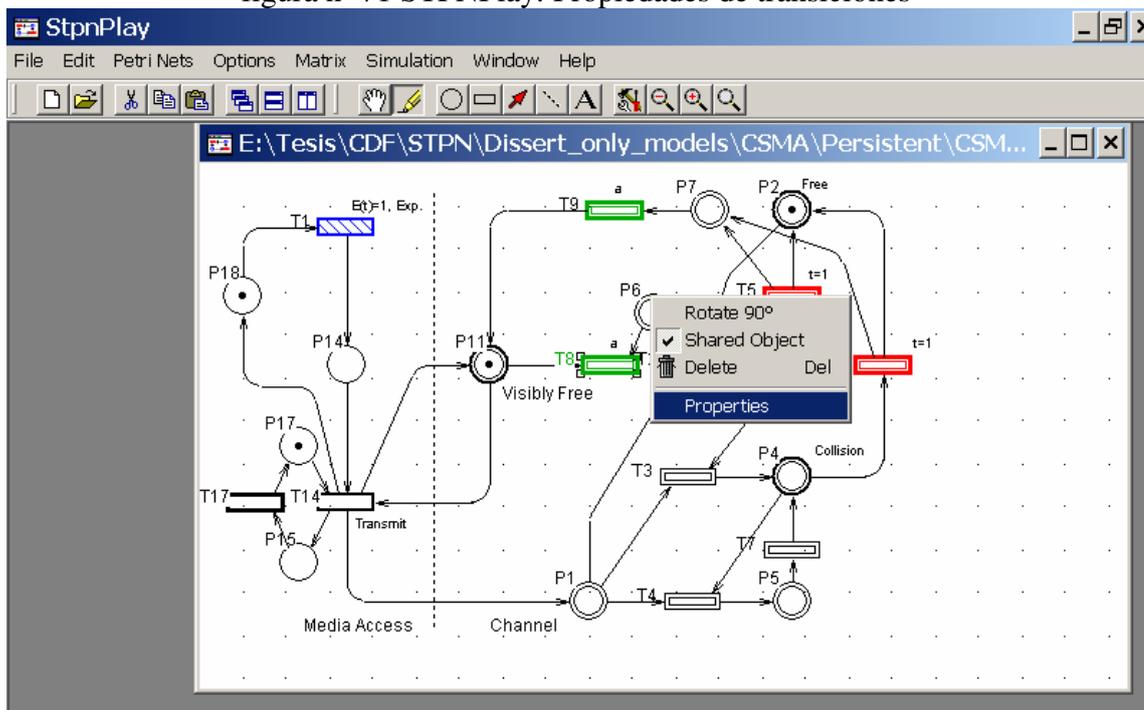


figura nº 72 STPNPlay: Objetos Compartidos

- Lugar:
 Número: El número del lugar en el modelo. El número es asignado automáticamente cuando se crea el objeto.
 Tiempo: Discutido luego.
 Color: Es el color del círculo que representa el objeto.
 Diámetro: Es el diámetro del círculo que presenta el objeto.

Ancho: El ancho de la línea del círculo.

Marcado: El marcado inicial del lugar.

Compartido: EL usuario puede elegir si el objeto aparece en múltiples nodos de la Red de Petri, o si el objeto representa una parte común para todos los nodos.

- Arco:

Tiempo: Discutido luego.

Color: Es el color de la línea y flecha que representan el objeto.

Ancho: Es el ancho de la línea.

Peso: Es el peso del arco en la Red de Petri.

El usuario puede modificar puntos de arcos ya dibujados.

- Línea:

Color: Es el color de la línea y flecha que representa el objeto. El usuario puede directamente seleccionar un color Standard, o usar un color arbitrario de la paleta de Windows.

Ancho: Es el ancho de la línea.

Estilo: La línea puede ser dibujada usando todos los estilos soportados por GUI Windows.

- Texto:

Font: El usuario puede configurar todas las propiedades font provistas por GUI Windows.

El usuario puede recuperar todos los objetos borrados usando el comando undelete, y simplemente clickeando el Mouse el objeto será recuperado (figura n° 73).

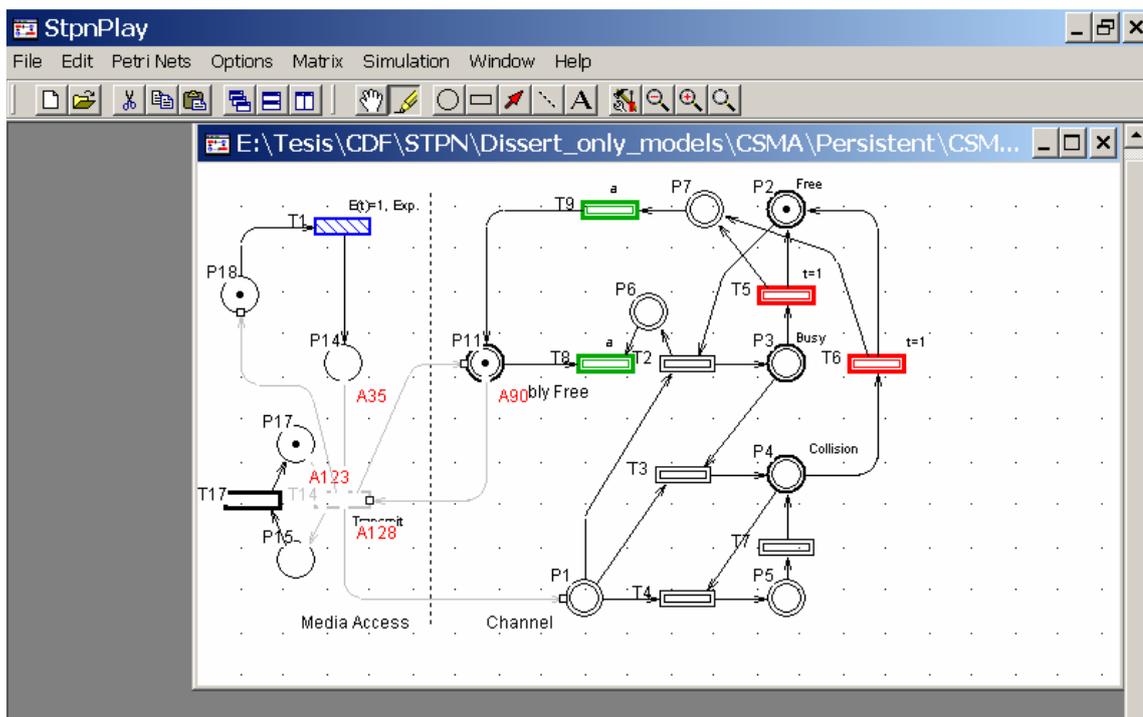


figura n° 73 STPNPlay: Modo Undelete

El editor de Red de Petri soporta zoom continuo. Los pasos del zoom pueden ser arbitrarios dado que el zoom está en punto flotante (figura n° 74).

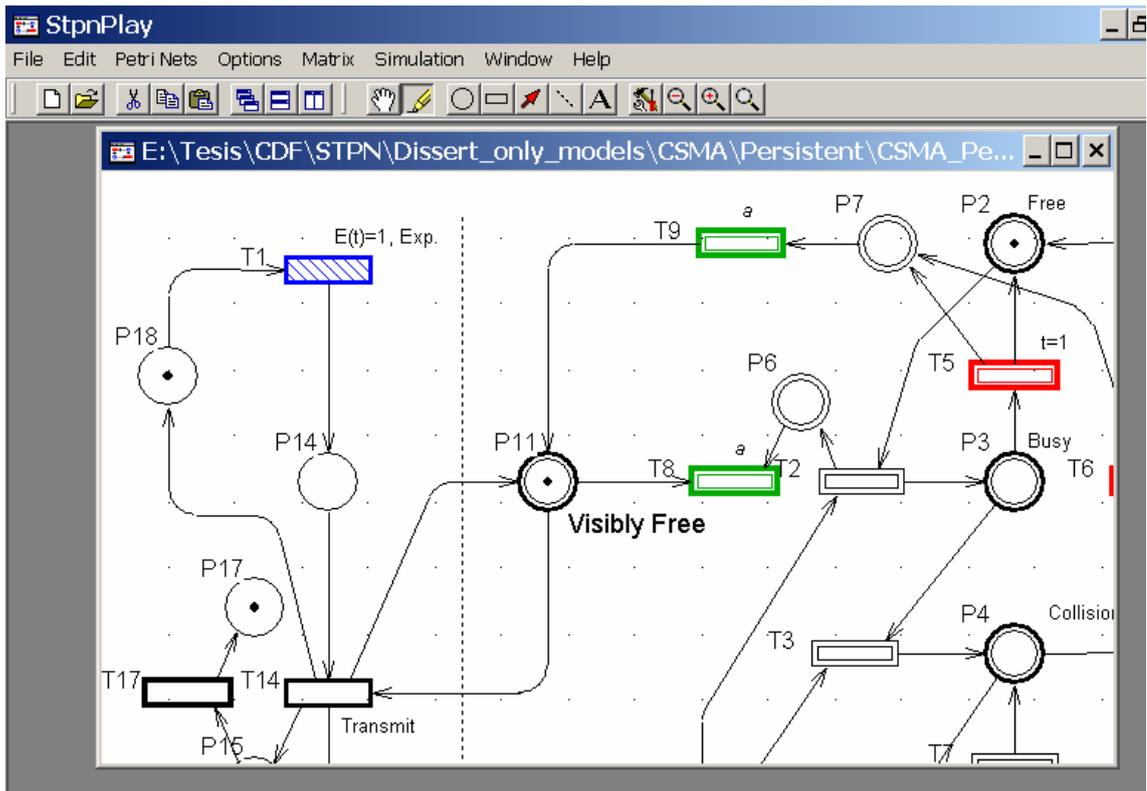


figura nº 74 STPNPlay: Capacidad de zoom

IV.2.2.1 Comandos del Menú en el Editor de Red de Petri

Los comandos están divididos en algunos sub-menús (figura nº 75). La estructura del menú es diferente para los módulos particulares del programa.

El menú File contiene comandos para crear nuevos diagramas de Red de Petri. El usuario puede abrir, cerrar, guardar e imprimir diagramas de Red de Petri existentes. El usuario puede exportar diagramas de Redes de Petri en formato texto, y en PetriMaker, y exportar el diagrama a metafile Windows (wmf). En este menú, el usuario puede también cargar Redes de Petri desplegadas directamente en el simulador, o cargar una matriz de disparo para mostrar el chart de la simulación. En el modo simulación, el usuario puede exportar una matriz de disparo en varios formatos.

El menú Edit contiene la función para cortar y pegar.

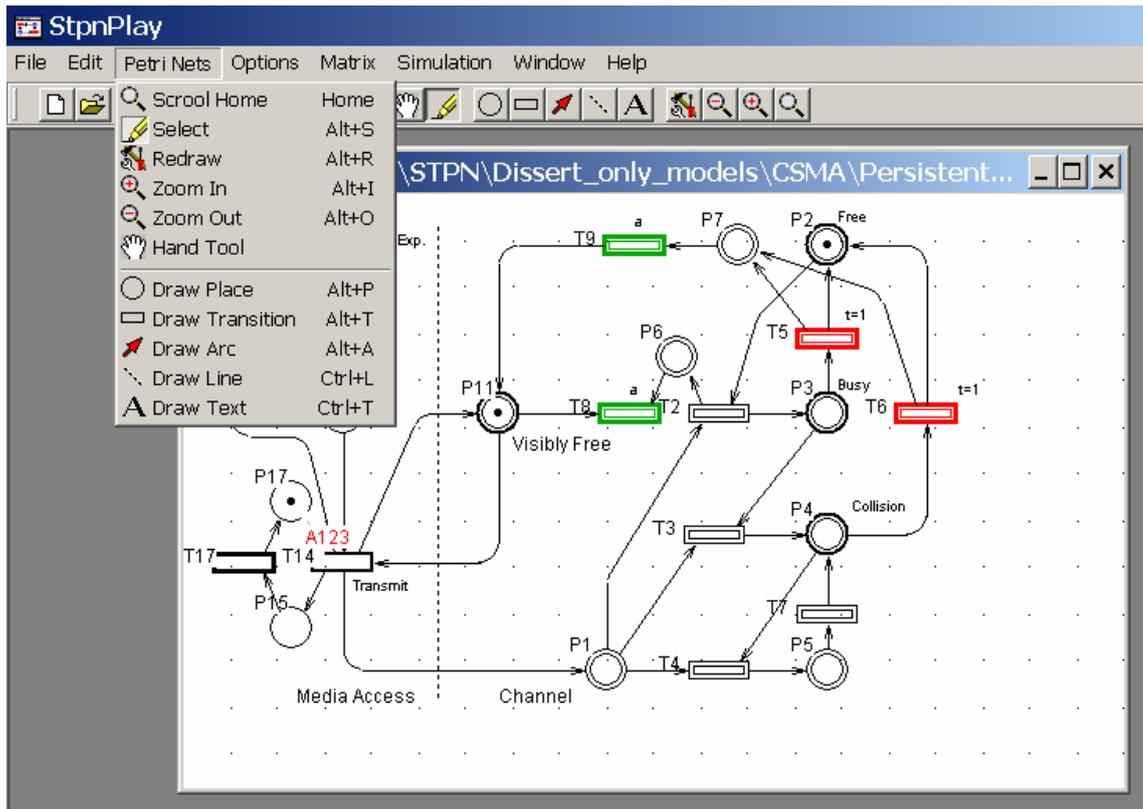


figura n° 75 STPNPlay: Menu de Redes de Petri

El menú Petri Net contiene comandos para crear y seleccionar objetos de tipo lugares, transiciones, arcos, líneas y texto. Hay también comandos para ampliar, reducir y redibujar el diagrama. Vía este menú es también posible configurar las propiedades de la aplicación como el número de nodos, el zoom, propiedades por defecto de todos los objetos de la Red de Petri, la simulación por defecto y propiedades del gráfico.

El menú Matriz contiene comandos para crear la matriz pre-condición, post-condición y de incidencia de una Red de Petri Estocástica plegada o desplegada.

El menú Simulation permite ingresar en el modo simulación.

IV.2.3 Simulador de Evaluación de Performance

En el simulador de la Red de Petri, el usuario puede importar y exportar la Red de simulación desplegada; configurar las propiedades de la simulación; comenzar, parar y poner en pausa la simulación; y configurar las propiedades del chart producido por el proceso de simulación (figura n° 76).

El usuario puede ampliar y reducir el chart; copiar el chart en distintos formatos y mostrar detalles de partes interesantes de las características.

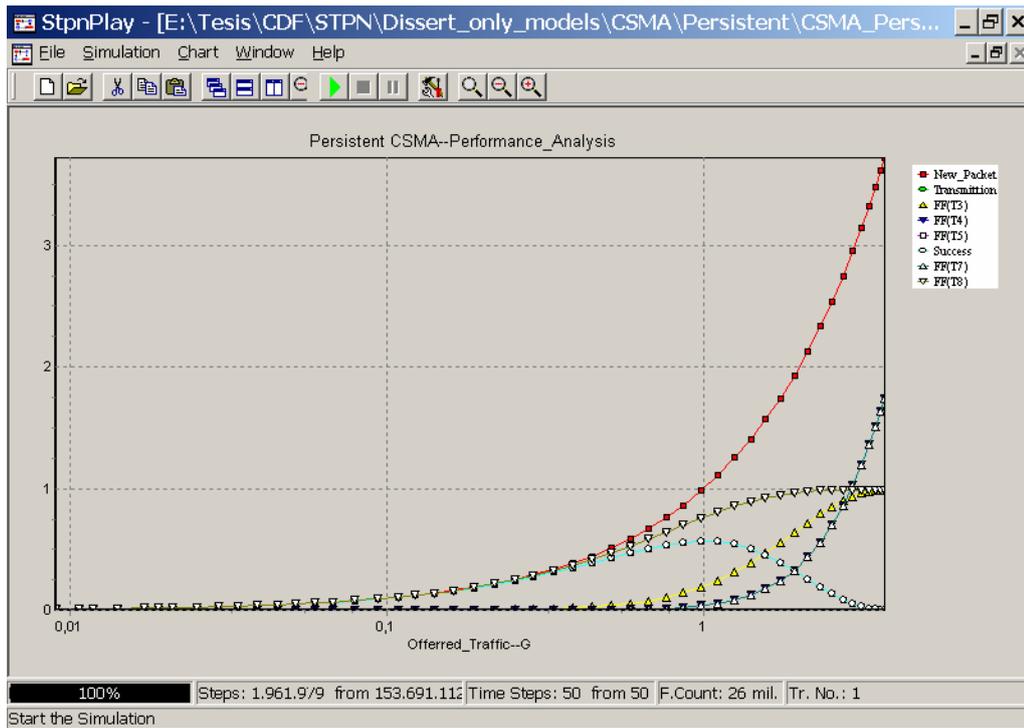


figura n° 76 STPNPlay: Chart-3D en coordenadas logarítmicas

Como se muestra en la figura n° 77, el usuario puede seleccionar la salida de la simulación: la matriz de disparo, la matriz de disparo de la red, que representan resultados de simulación promedio generados por todos los nodos conectados al canal; o la secuencia de disparo guardada en un archivo. Los eventos transitorios (números de transiciones disparadas) y eventos tangibles pueden ser guardados en la secuencia de disparo.

Las transiciones pueden tener tiempo asociado, que es asignado según la distribución de número múltiplo aleatorio; así el usuario puede modelar varios tipos de sistema de cola de acuerdo a la clasificación de Kendall.

Durante la simulación, el tiempo medio de las transiciones designadas se cambia en pasos múltiplo (vector de tiempo). En cada paso, el simulador analiza la Red en el juego de marcas GSPN, y computa el vector de cuenta de disparos. El usuario puede elegir si el vector de tiempo forma progresión aritmética o geométrica. El usuario puede también configurar el tiempo máximo o mínimo, y el número de pasos de tiempo de simulación. El tiempo asociado a la transición designada es proporcional a la variable independiente de las características evaluadas.

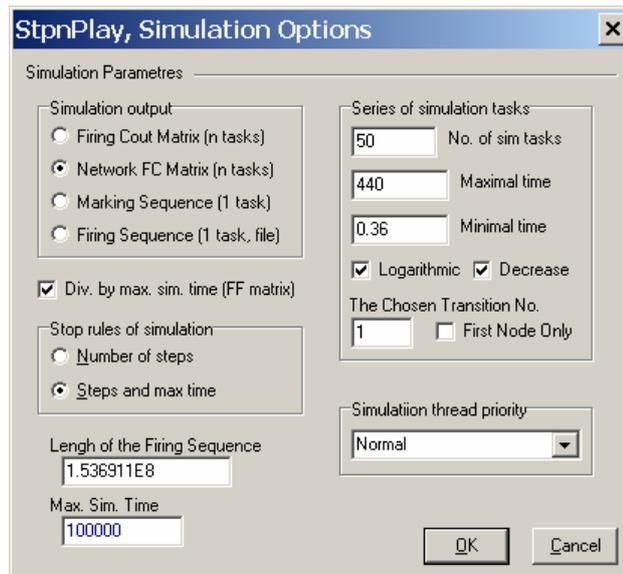


figura nº 77 STPNPlay: Opciones de Simulación

En cada paso, el simulador finalizar en una cierta longitud de secuencia de disparo, en un cierto tiempo de simulación máximo, o en una combinación de ambos. Esto permite analizar las características dependientes del tiempo e independientes del tiempo del sistema simulado.

Las opciones de formar del chart se muestran en la figura nº 78. El usuario puede mostrar hasta ocho características en un chart simple. La variable independiente puede ser el tiempo simulado, número de pasos de tiempo o frecuencia de disparo de una transición arbitraria en el modelo. Las variables dependientes pueden ser el tiempo o la frecuencia de disparo de la transición arbitraria.

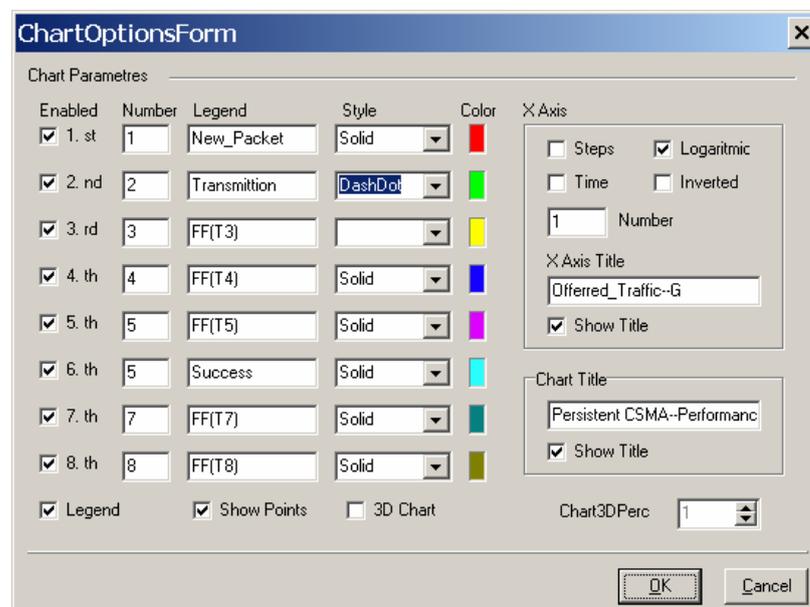


figura nº 78 STPNPlay: Propiedades del Chart

Como se muestra en la figura nº 79, el usuario puede exportar la matriz de resultados de la simulación.

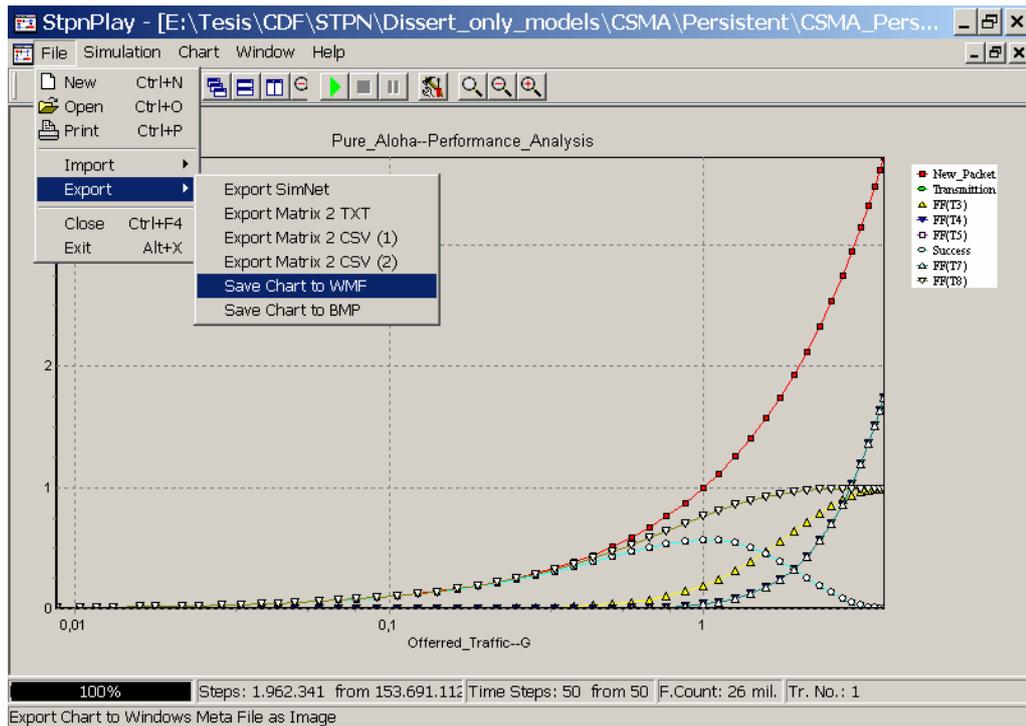


figura n° 79 Chart-2D en coordenadas lineales-Menu de exportación de la simulación

IV.2.4 Matriz de la Red de Petri

El usuario puede generar la matriz de pre-condición y de post-condición para el modelo plegado, y para el modelo desplegado conteniendo múltiples nodos (figura n° 80).

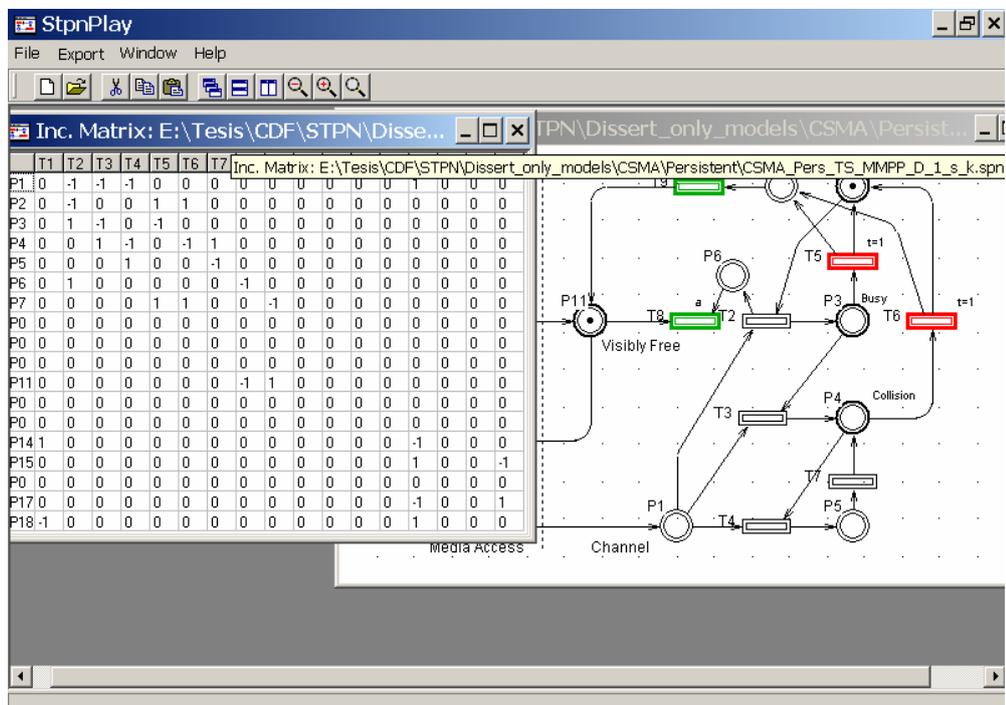


figura n° 80 STPNPlay: Matriz de incidencia

IV.3) EVALUACIÓN DE PERFORMANCE DE ALOHA

IV.3.1 Evaluación de Performance de Aloha Puro

La simulación de Aloha puro es visto como un sistema de cola M/D/1 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor). Como se discutió en el punto III.4.1 para el modelo de Aloha puro de II.2.2, las frecuencias de disparo de las transiciones tienen los siguientes significados:

- la cuenta de disparo de la transición T1 es el tráfico ofrecido G,
- la frecuencia de disparo de la transición T2 representa el frecuencia de iniciaciones de la transmisión sobre el canal previamente inactivo,
- la frecuencia de disparo de la transición T3 representa el número de colisiones que aparecen sobre el canal previamente activo, pero en estado de no colisión,
- la frecuencia de disparo de la transición T4 (la misma que T7) representa el número de colisiones que aparecen sobre el canal, con estado previo de colisión.
- la cuenta de disparo de la transición T5 es igual al número de paquetes exitosamente transmitidos, y
- la cuenta de disparo de la transición T6 representa el número de resolución de colisiones.

Se supone que se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a Aloha puro de II.2.2, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador.

Entonces con el simulador STPNPlay, y asumiendo que $t_5, t_6 = 1$, lo que significa que el tiempo medio asociado a la transición T1 representa la reciproca de tráfico ofrecido G. En la figura nº 81, las frecuencias de disparo de las transiciones en el modelo del canal son graficadas como funciones del tráfico ofrecido $G \leq (0.01, 5)$. Como se muestra en la figura, la curva del rendimiento del canal y la curva que representa el número de paquetes que encuentra el canal libre en el comienzo de la transmisión (frecuencia de disparo de la transición T2), se aproxima a la curva que representa el tráfico ofrecido para valores pequeños de G.

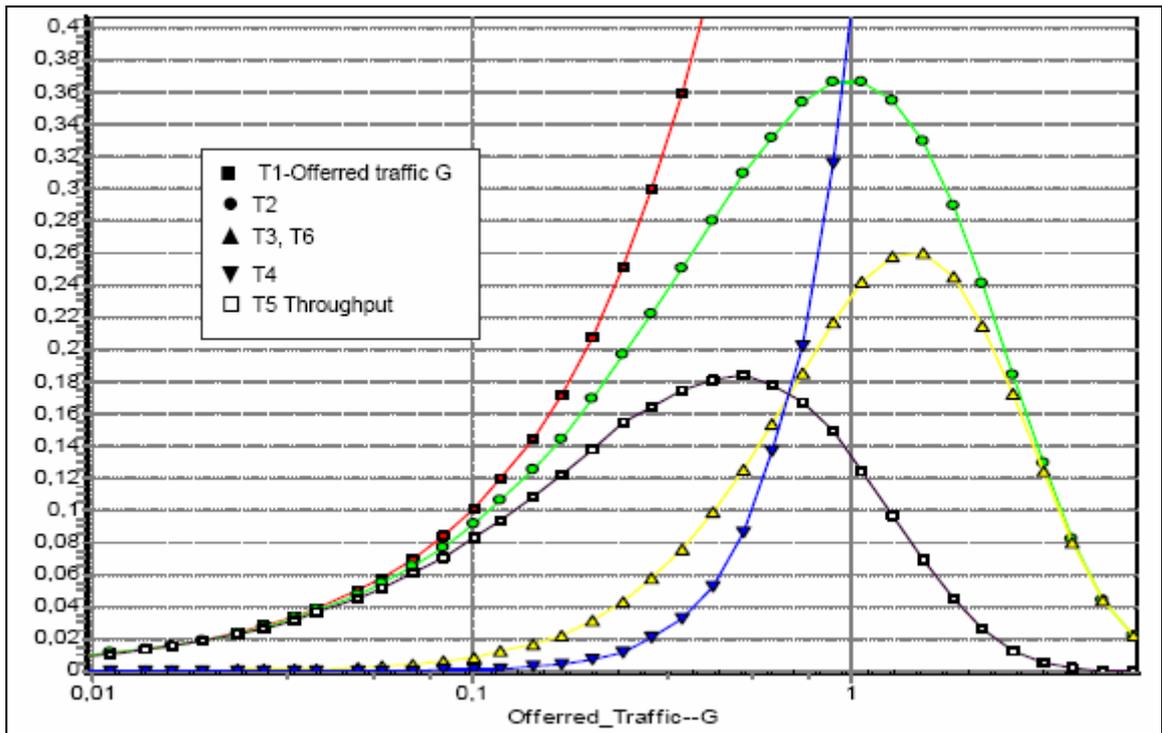


figura nº 81 Rendimiento de Aloha puro – Frecuencias de disparo de transición

El número de paquetes que causaron una colisión (frecuencia de disparo de T3) y el número de paquetes que encontraron el canal en estado de colisión (frecuencia de T4), y la relación de colisión definida como el número de paquetes colisionados versus el número de paquetes ofrecidos, es baja para pequeños valores de G . Cuando se incrementa G , la relación de colisión se incrementa. El rendimiento máximo $S=0,18$ se obtiene en $G = 0,5$. En este punto en promedio, un paquete arribado encuentra con probabilidad de 0,12 el canal ocupado (disparo de la transición T3), y con probabilidad 0,08 encuentra el canal en estado de colisión (transición T4). Para $G > 10$, casi todos los paquetes encuentran el canal en estado de colisión al comienzo de sus transmisiones.

En el modelo desplegado, el proceso de arribo de paquetes está representado por un conjunto de k transiciones de tiempo estocástico T1 con tiempo distribuido exponencialmente. Se asume que no hay correlación entre las variables aleatoriamente generadas por los disparos de las transiciones. Los procesos de arribo generados por toda la población es la suma de los procesos de Poisson de cada uno de ellos. El valor medio del tráfico ofrecido está dado por la suma de las recíprocas de los tiempos medios de las transiciones T1 en todos los nodos de la red desplegada. Un solo nodo no genera colisiones sobre el canal. El nodo usa casi todo el ancho de banda del canal entero (figura nº 82). Para $G > 1$, es $S = 1$, a causa de que el nodo siempre tiene un paquete listo y nunca comienza la transmisión durante su propia transmisión.

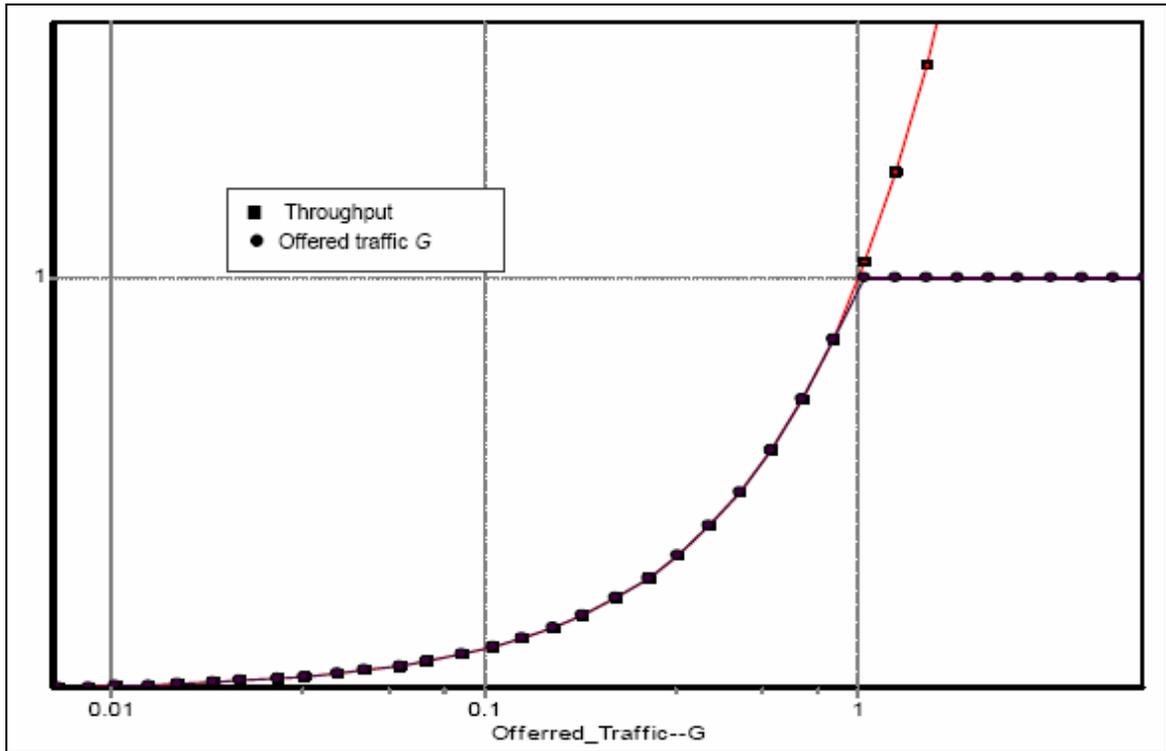


figura nº 82 Rendimiento de Aloha puro – Un nodo sobre el canal

Las frecuencias de disparo de la red consistente de dos nodos están graficadas en la figura nº 83. Dos nodos producen el tráfico ofrecido máximo $G = 2$ debido a que cada nodo produce $G = 1$. El ritmo de las colisiones son la suma de las frecuencias de disparo de T3 y T4, y su máximo valor es igual a 2. El rendimiento máximo del canal es 0.322. Para grandes valores de tráfico ofrecido, el rendimiento del canal cargado por dos nodos es menor comparado a la carga generada por un número grande de nodos, a causa que ambos nodos transmiten tan frecuentemente que ellos aseguran el solapamiento de uno con el otro. Para tráfico ofrecido $G > 2$, el marcado del lugar P6 en ambos nodos se aproxima a infinito.

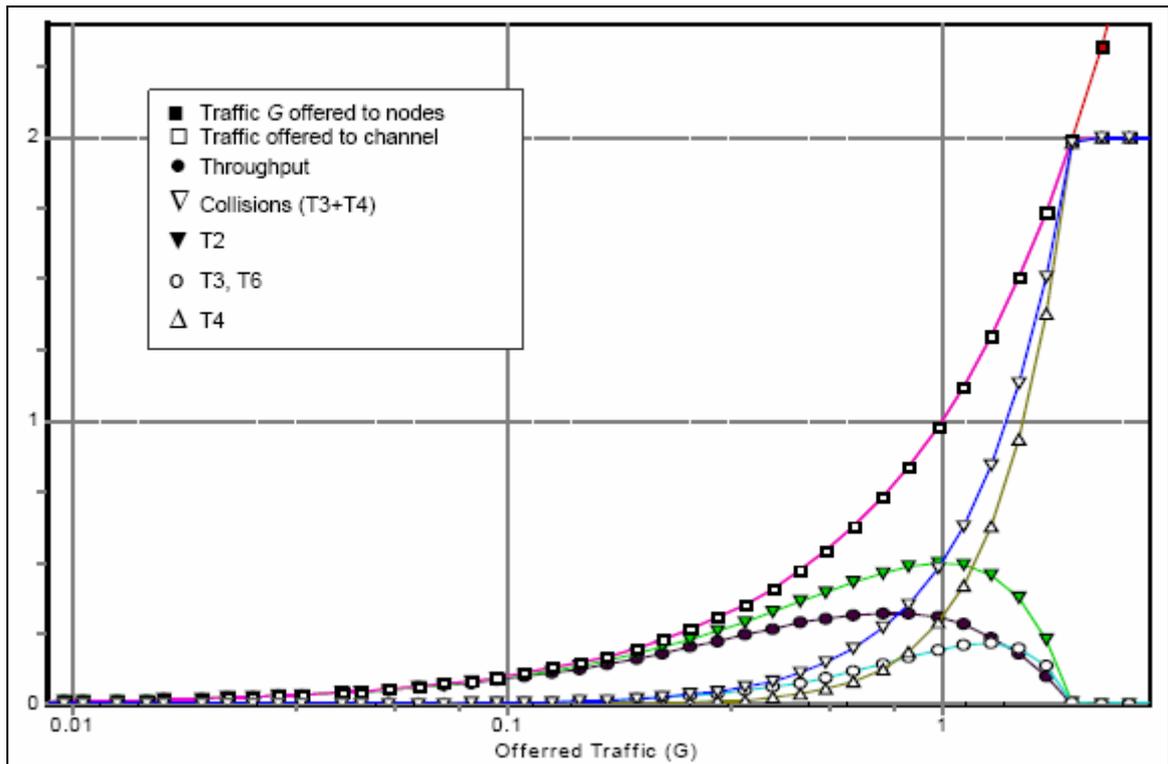


figura nº 83 Rendimiento de Aloha puro – Dos nodos sobre el canal

En las figura nº 84 se presentan los rendimientos para distintas cantidades de número de nodos para Aloha puro.

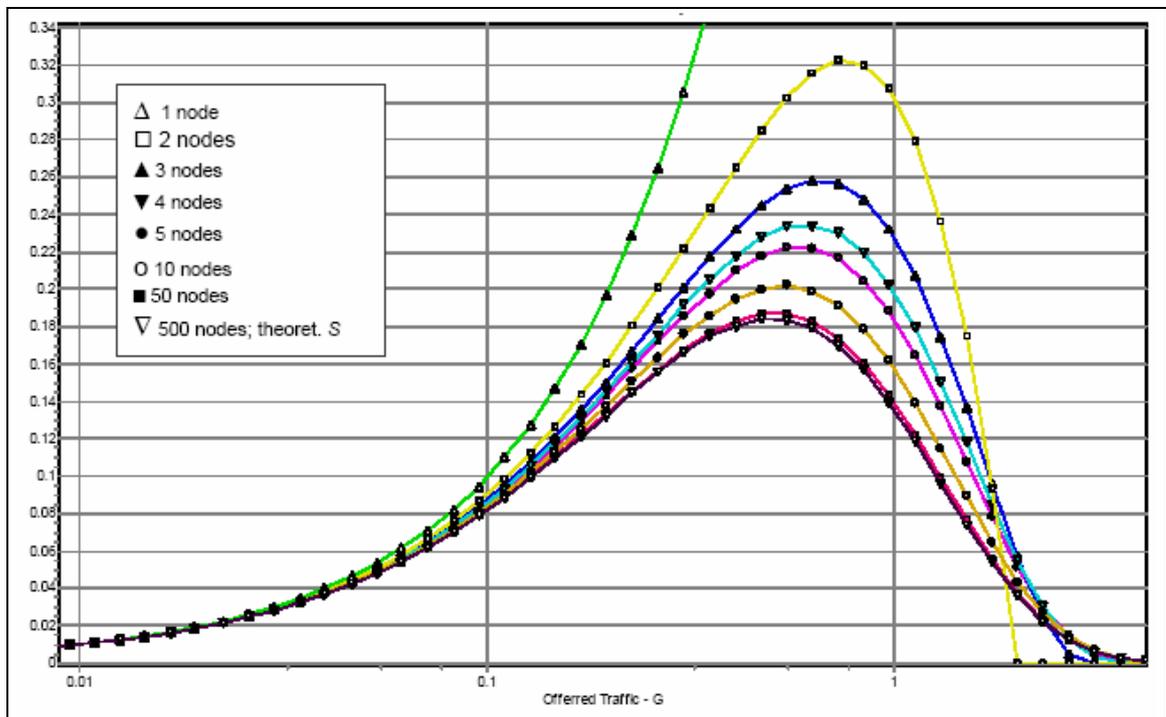


figura nº 84 Rendimiento de Aloha puro – Distinto número de nodos sobre el canal

IV.3.2 Evaluación de Performance de Aloha Ranurado

Como en IV.3.1 suponemos que se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a Aloha ranurado de II.2.3, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador. Asumimos que $t_5, t_6 = 1$; el tiempo medio asociado a la transición T1 representa la recíproca del tráfico ofrecido G .

El rendimiento del canal y las frecuencias de disparo de las transiciones T2, T3 y T4 son graficadas como funciones del tráfico ofrecido $G < (0.05, 10)$ en la figura n° 85. Como se indico previamente, la única instancia de tiempo, cuando las colisiones pueden ocurrir, es el breve intervalo (en el modelo es un intervalo infinitamente breve) en el comienzo del tiempo de slot. El canal está siempre libre en el momento entre dos tiempos de slot consecutivos. La frecuencia de disparo de la transición T2 representa la relación de slots de tiempo en el que el canal no está libre. La suma de las frecuencias de disparo de las transiciones T3 y T4 representa la colisión del canal (la porción del ancho de banda del canal que es perdido por las colisiones).

El Aloha rasurado obtiene un rendimiento máximo mayor que el Aloha puro al reducir el tiempo consumido en las colisiones. Mientras en Aloha puro el tiempo gastado en las colisiones es hasta dos unidades de tiempo de largo (dos tiempos de transmisión de paquetes, una colisión es Aloha ranurado gasta exactamente una unidad de tiempo. Así, en Aloha ranurado un paquete está siempre almacenado en el buffer de salida (lugar P6) hasta el comienzo de slot de tiempo siguiente. Para población infinita, el retardo promedio entre el momento de arribo de un paquete a la capa MAC y el de cuando su transmisión comienza (el tiempo de encolado promedio) es igual a medio tiempo de slot (recordar el supuesto que un nodo tiene a lo sumo un paquete en su cola).

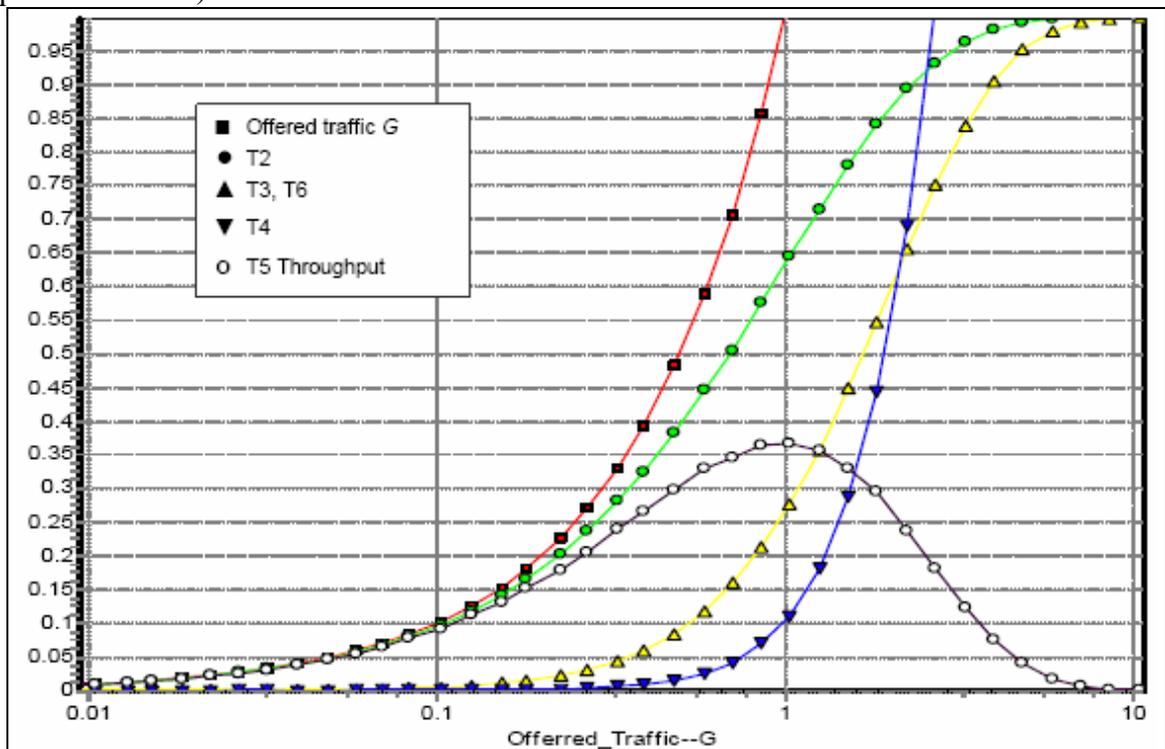


figura n° 85 Rendimiento de Aloha ranurado – Frecuencias de disparo de transición

En las figura n° 86 se presentan los rendimientos para distintas cantidades de número de nodos para Aloha ranurado.

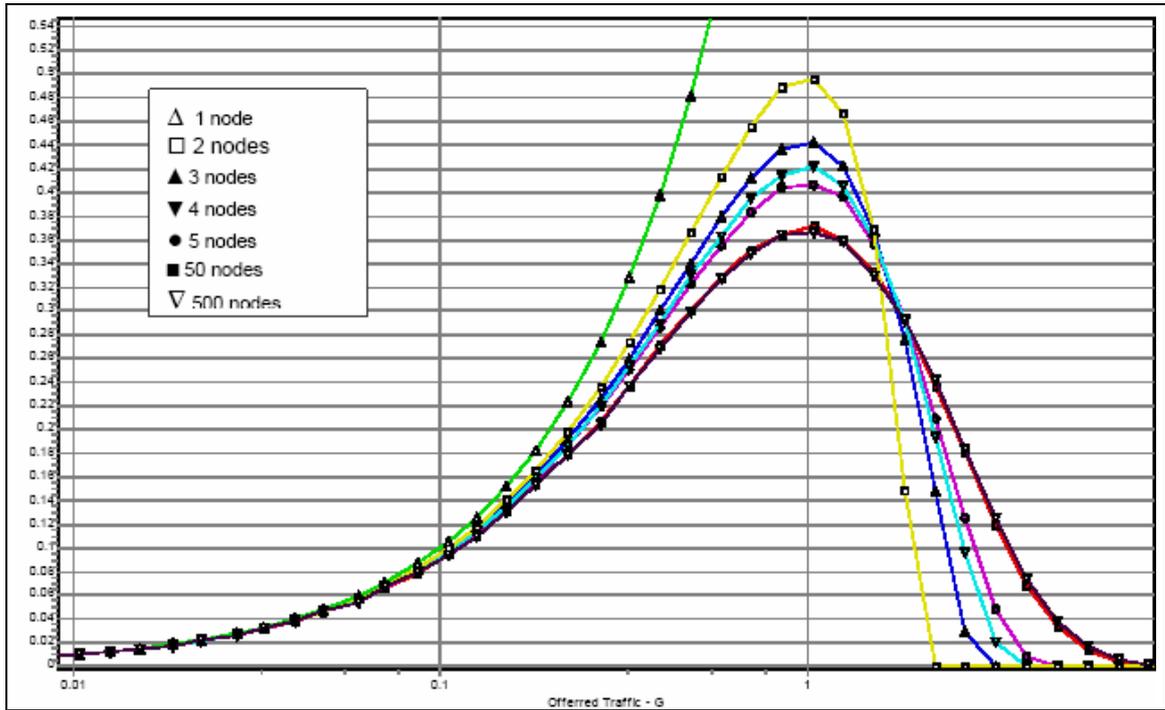


figura n° 86 Rendimiento de Aloha ranurado – Distinto número de nodos sobre el canal

IV.4) EVALUACIÓN DE PERFORMANCE DE CSMA/CD

IV.4.1 Evaluación de Performance de CSMA (persistente)

Nuevamente, suponemos que se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a CSMA persistente de II.3.4.2, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador.

Para $a = 0.01$, las frecuencias de disparo de las transiciones T1, T2, T4, T5 y T6 son graficadas en la figura n° 87, como una función de $G < (0.01, 10)$. El rendimiento de los métodos CSMA es dependiente del retardo de propagación a .

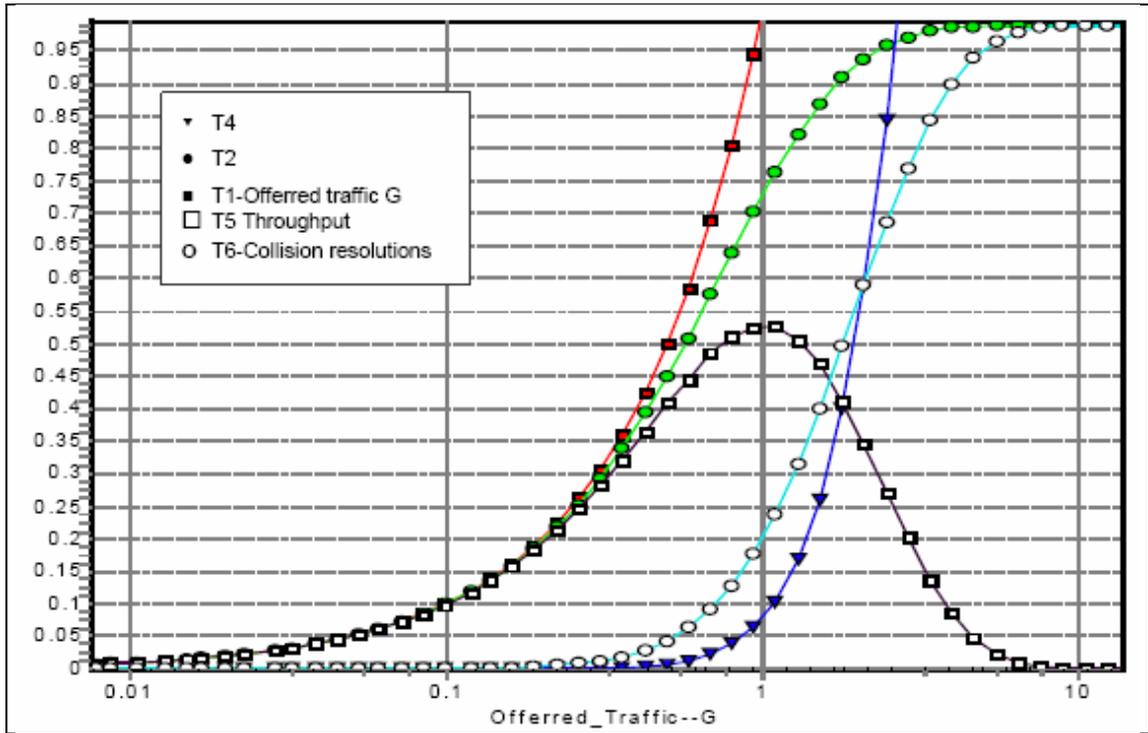


figura nº 87 Rendimiento de CSMA Persistente – Frecuencias de disparo de transición para $a = 0.01$

El número de nodos que participan en los procesos MAC influye en el rendimiento del canal. En la figura nº 88 se grafican las características de S versus G, para varios números de nodos, y $a = 0.01$; y en la figura nº 89 para $a = 0.99$.

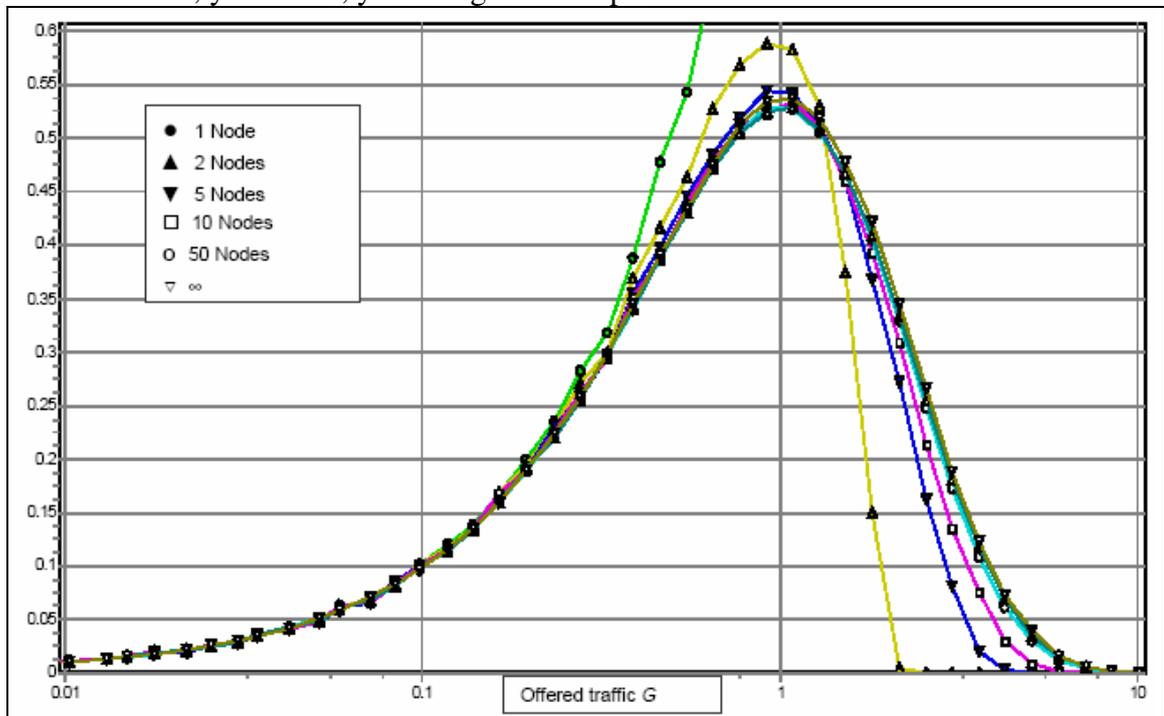


figura nº 88 Rendimiento de CSMA Persistente – Distinto número de nodos sobre el canal para $a = 0.01$

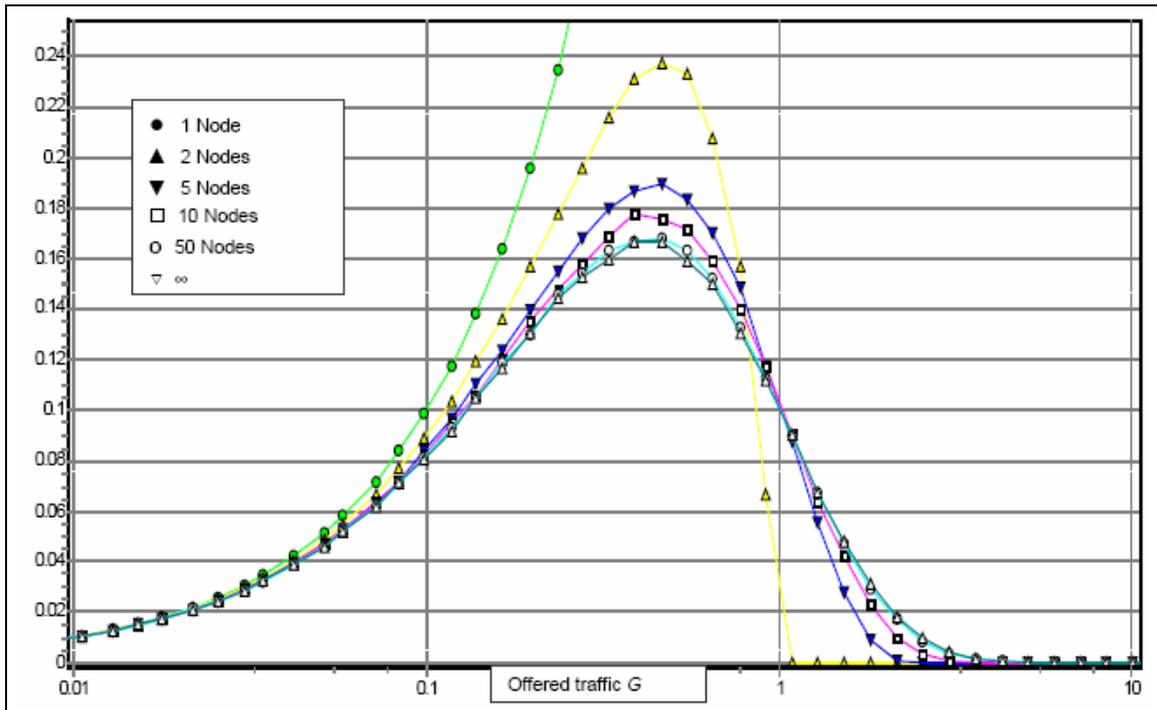


figura nº 89 Rendimiento de CSMA Persistente – Distinto número de nodos sobre el canal para $a = 0.99$

Para el caso de un simple nodo (figura nº 90), la parte horizontal de la curva de rendimiento del canal, representa el estado en que el nodo tiene siempre al menos un paquete en su buffer de salida. En tal situación, no hay estado estable y el retardo de transmisión de paquete medio diverge al infinito. El rendimiento máximo del canal cargado por un simple nodo de disminuye cuando el retardo de propagación es grande. El rendimiento más pequeño para a grande se origina por los periodos libres de tamaño a entre dos paquetes consecutivos. Después de la finalización de la transmisión, el nodo espera cerca de a segundos hasta que el canal se vuelve libre.

IV.4.2 Evaluación de Performance de CSMA/CD

En este caso se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a CSMA/CD de II.3.5, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador.

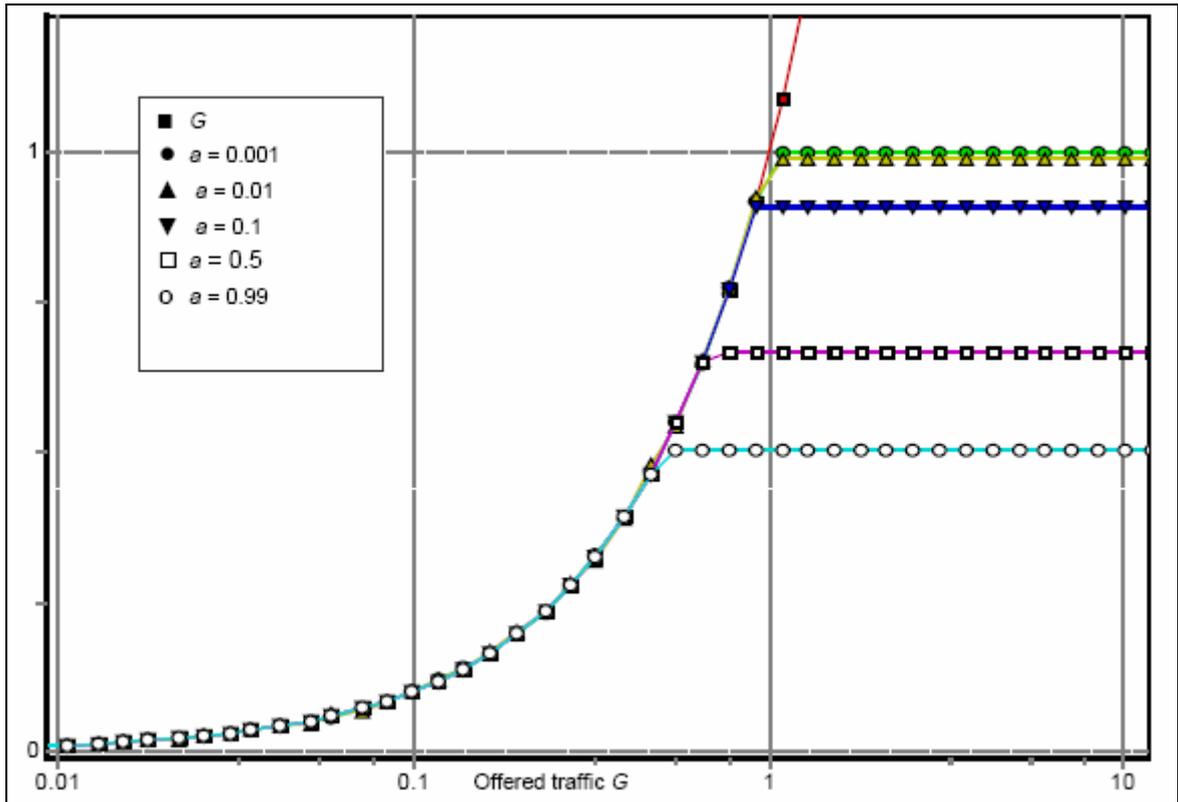


figura nº 90 Rendimiento para CSMA Persistente – Un nodo sobre el canal y para distintos a

Algunas características de rendimiento de CSMA/CD se presentan en las figuras nº 91 y 92. En la primera, es interesante ver que el rendimiento del canal máximo corresponda a un tráfico ofrecido relativamente grande ($G \approx 20$). En la segunda podemos ver como el retardo de propagación del canal a influye sobre el rendimiento de CSMA/CD. El retardo de propagación no solo influye sobre el rendimiento máximo de CSMA/CD, sino también sobre el punto de operación (valor de G) en el cual el canal tiene el rendimiento máximo.

Las características de rendimiento CSMA/CD, considerado como un sistema de cola $M/D/1$ (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinística, y 1 servidor), $M/M/1$ (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución exponencial, y 1 servidor) y $M/U/1$ (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución general, y 1 servidor) se grafican en la figura nº 93, como una función del tráfico ofrecido G , usando el modelo de Red de Petri correspondiente a CSMA/CD de II.3.6. Dicha figura confirma la corrección de la suposición de longitudes de paquete constante en los modelos simples. Puede observarse que pequeña es la influencia de las diferentes distribuciones de longitud de paquete para las curvas de rendimiento del canal CSMA/CD, y que pequeña es la diferencia entre los rendimientos máximos del canal. Finalmente, se observa que la diferencia entre la curva de rendimiento del canal cargado y longitudes de paquetes uniformemente distribuidas, y la curva $M/D/1$ es más grande que la diferencia entre las características de rendimiento de $M/M1$ y $M/D/1$

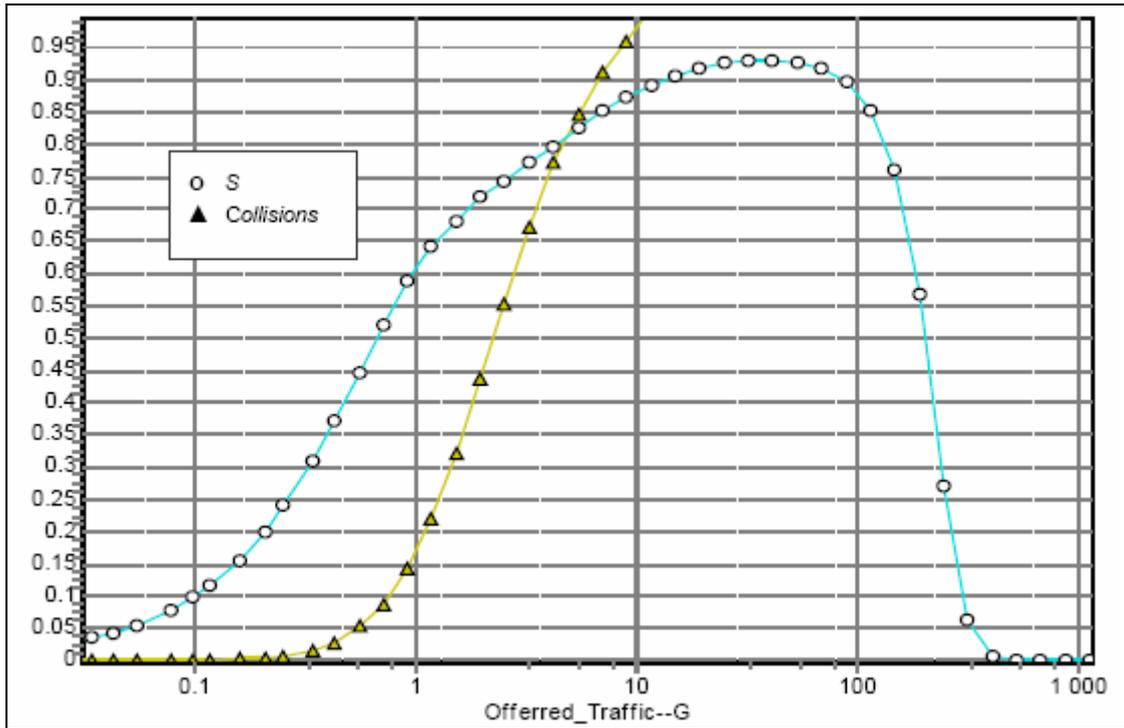


figura n° 91 Rendimiento de CSMA/CD como M/D/1 – Para $a=0.01$ y $b=0$

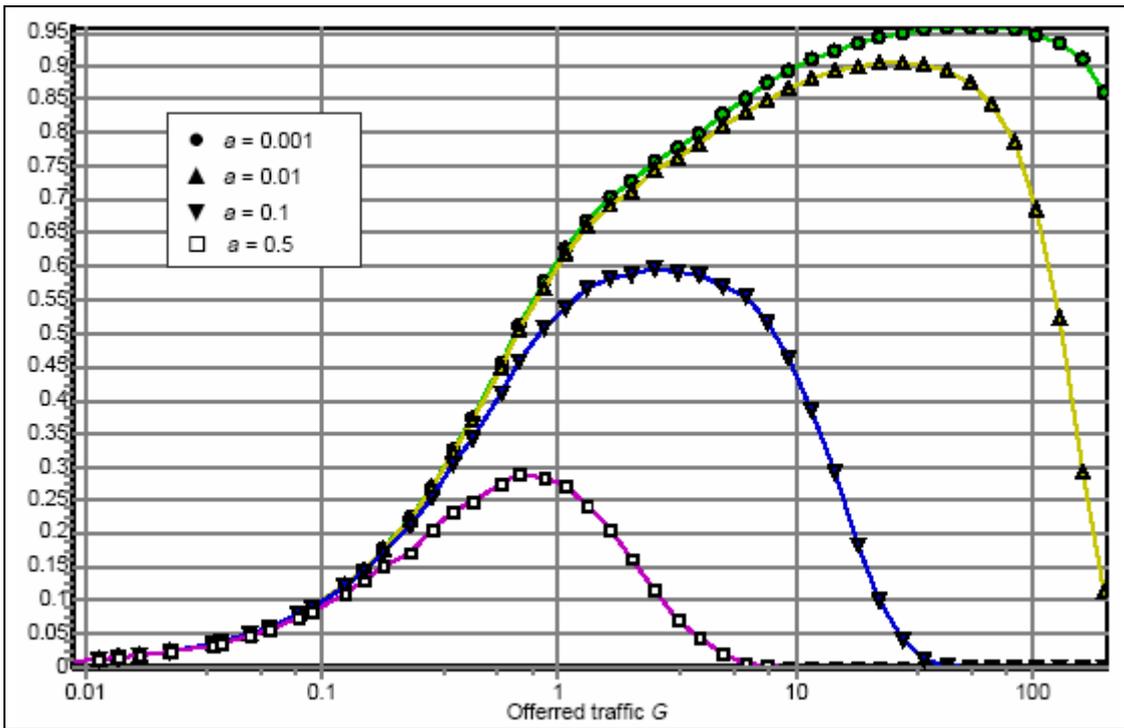


figura n° 92 Rendimiento de CSMA/CD como M/D/1 – Para distintos a y $b=0.1$

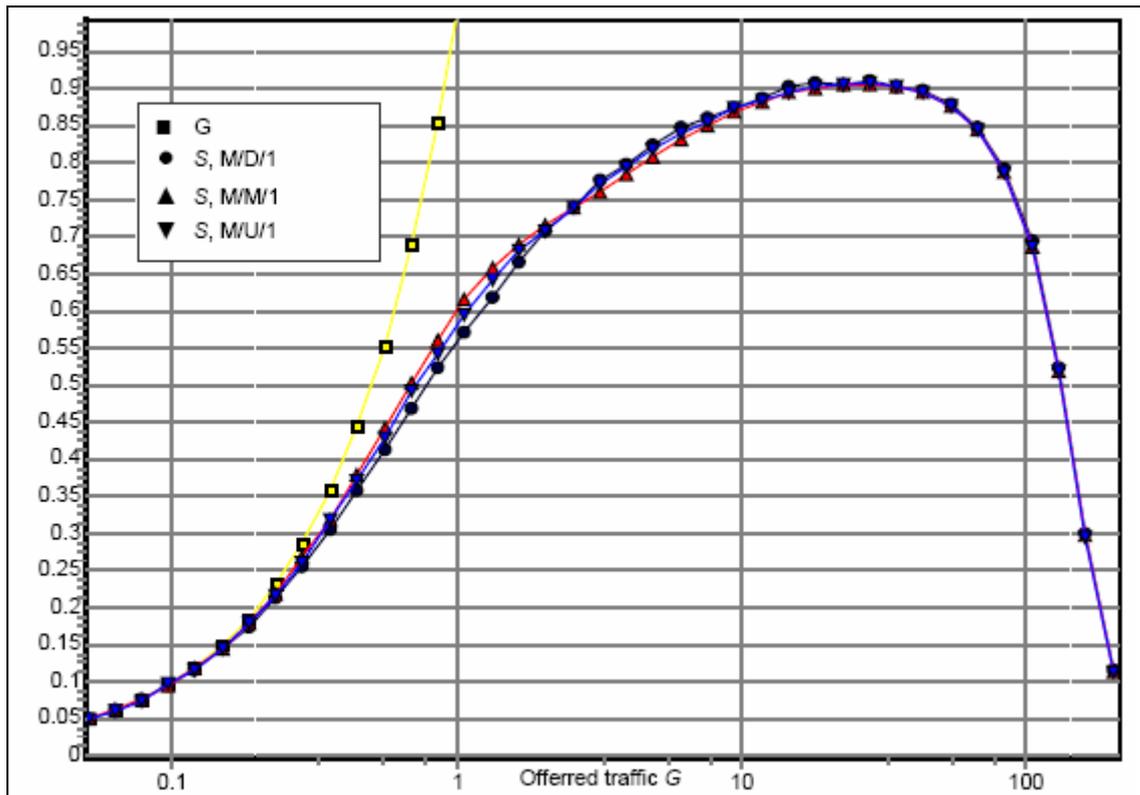


figura nº 93 Rendimiento de CSMA/CD – Distintas distribuciones de retardo, $a = 0.01$ y $b = 0.01$

Para el caso de CSMA/CD bajo distribución de longitud de paquete bimodal, usando el modelo de Red de Petri correspondiente a CSMA/CD de II.3.7, en la figura nº 94, se grafica el rendimiento del canal como una función del tráfico ofrecido para algunos valores de FrS , $L_s = 0.1$, $a = 0.01$ y $b = 0.01$. Los valores de FrS fueron calculados desde los tiempos promedio de las transiciones T_5 y T_{15} . Las curvas en el chart fueron computadas como funciones de la frecuencia de disparo de la transición T_1 , y manualmente desplazadas a la posición correspondiente sobre el eje G . Como se puede observar, el rendimiento simulado depende fuertemente de la fracción de paquetes que son cortos. Las curvas de rendimiento particular difieren entre sí a causa de que FrS influye sobre la longitud de paquete promedio, que a su vez afecta el retardo de propagación relativo del canal.

La influencia de la distribución de la longitud de paquete sobre el rendimiento del canal, tiene sólo una pequeña importancia. En la figura nº 95, se compara las curvas de rendimiento para $FrS = 0.5$ y para $FrS = 0.99$, a las correspondientes curvas de rendimiento del canal cargado por longitud de paquete fija. La diferencia entre los rendimientos bajo distribución de paquete bimodal y bajo longitud de paquete fija es pequeña cuando se la compara a la influencia del número de nodos o la influencia de a .

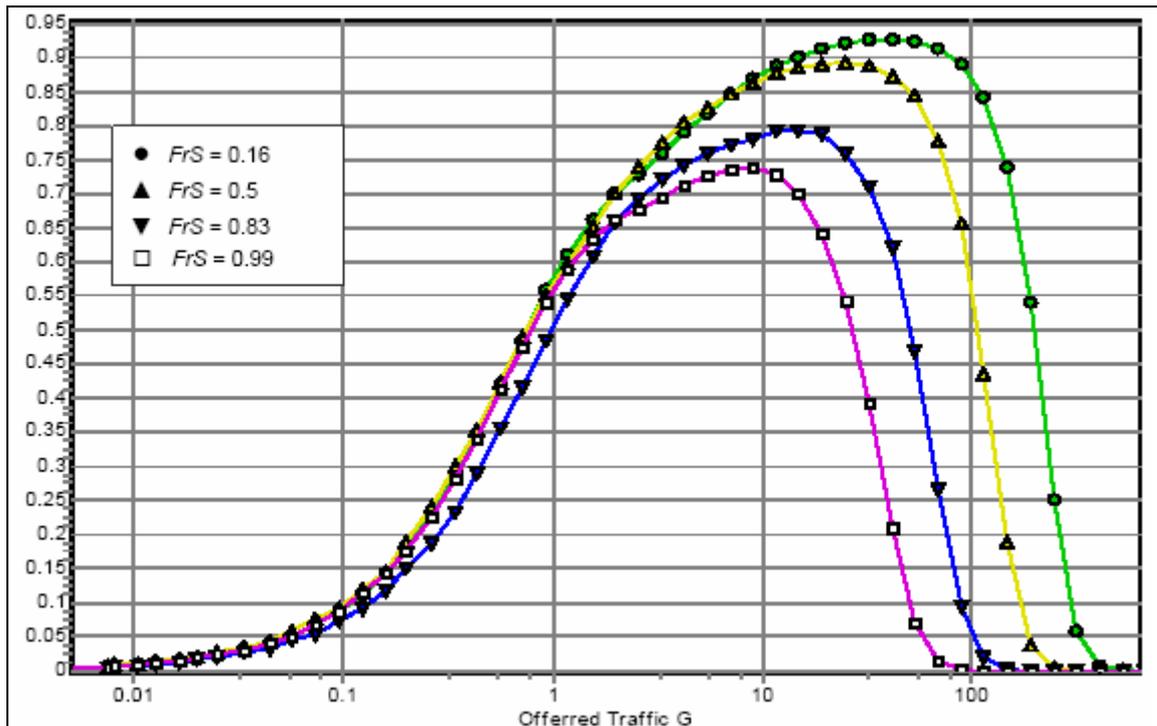


figura n° 94 Rendimiento de CSMA/CD – Dos longitudes de paquete para algunos valores de FrS, $a = 0.01$ y $b = 0.01$

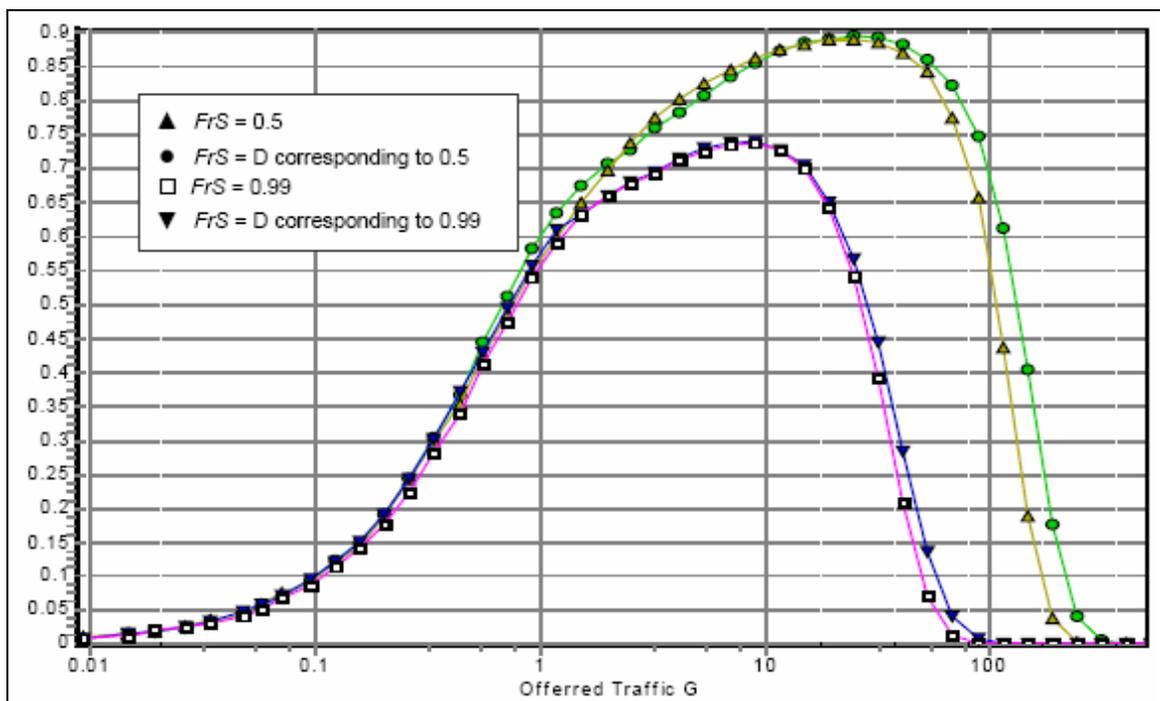


figura n° 95 Rendimiento CSMA/CD – Comparación para dos longitudes de paquete y longitud fija para algunos valores de FrS, $a = 0.01$ y $b = 0.01$

En la figura n° 96, se grafican las características de rendimiento como una función del tráfico ofrecido, para los sistemas M/D/1/5 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución determinístico, 1 servidor, 5 nodos o clientes

sobre el canal) , M/G/1/5 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución general, 1 servidor, 5 nodos o clientes sobre el canal) y M/M/1/5 (proceso de llegada con distribución exponencial, proceso de tiempos de servicio con distribución exponencial, 1 servidor, 5 nodos o clientes sobre el canal), $a = 0.01$ y $b = 0.008$. Las características de rendimiento de los sistemas de M/M/1 y M/G/1 fueron adquiridas de la misma forma que para el análisis de población infinita. Las curvas de la figura confirman la corrección de suponer las longitudes de paquete constante en los modelos simples de población finita. Puede observarse que pequeña es la influencia de las distribuciones de longitud de paquete diferentes para las curvas de rendimiento de canal CSMA/CD cargado por cinco nodos. Además, que la mayor influencia, que es menos del 10 %, para valores de $G \approx 1$. La diferencia entre los rendimientos de canal máximo es más pequeña que 5 % para distribuciones de longitud de paquete diferentes.

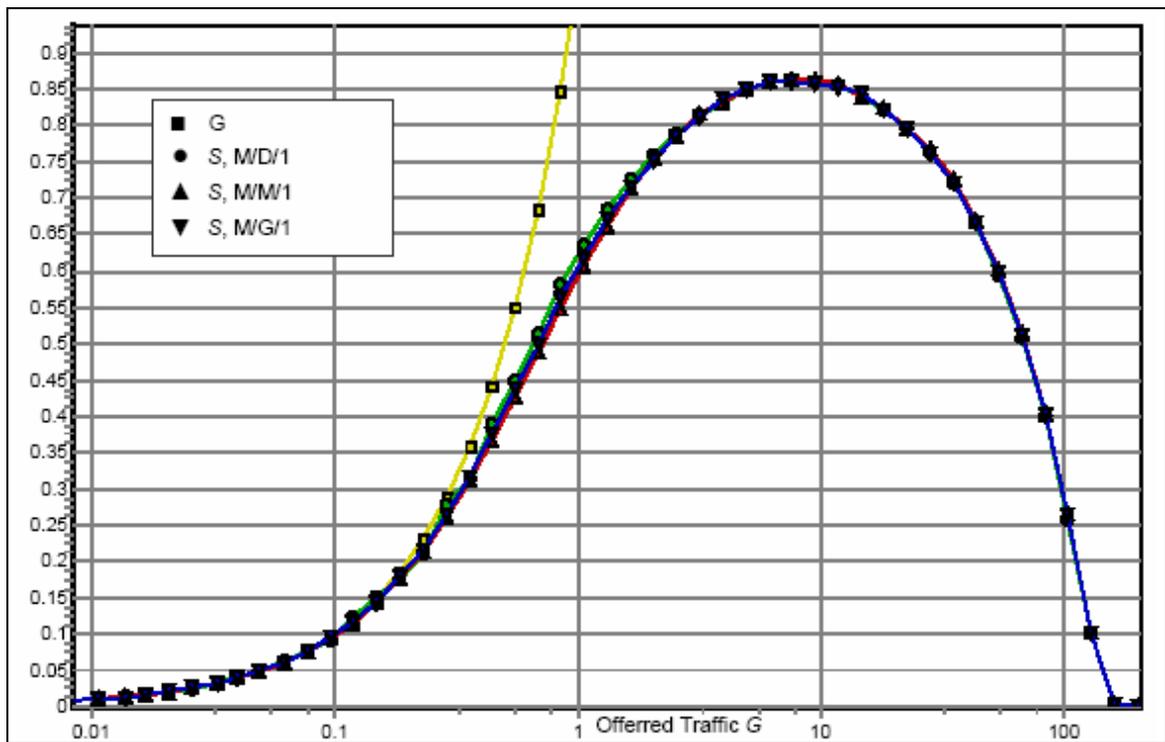


figura nº 96 Rendimiento de CSMA/CD – Comparación para varios tipos de distribución de retardo de transmisión de paquete, $a = 0.01$ y $b = 0.008$

En las figuras nº 97 a 100, se grafican las características de rendimiento de CSMA/CD, como una función del tráfico ofrecido G , considerado como un sistema de cola M/D/1/∞/k, para algunos valores de retardo de propagación a , y algunos números de nodos k . Cuando se compara la influencia del número de nodos, la influencia del retardo de propagación, y la influencia de la distribución de longitud de paquete, se puede decir que lo más importante es la influencia de a , mientras que la influencia de la longitud de paquete es despreciable. La influencia del número de nodos sobre el canal CSMA/CD es grande para pocos nodos.

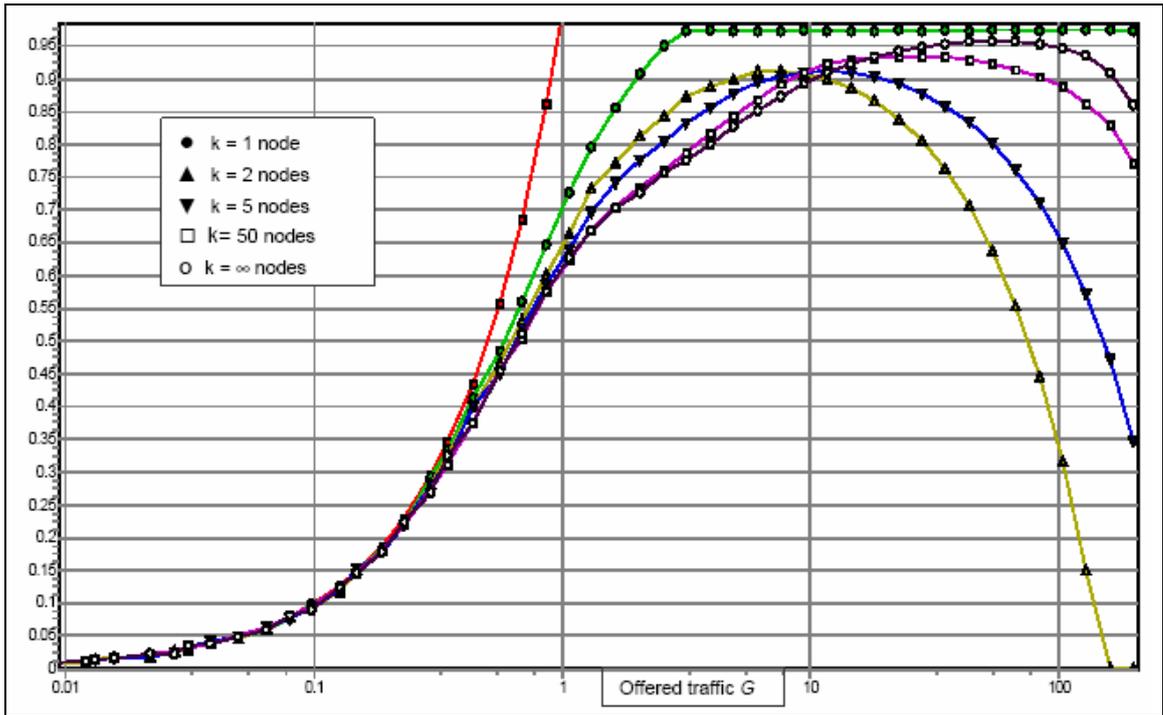


figura nº 97 Rendimiento CSMA/CD M/D/1/∞/k – Distinto número de nodos para $a = 0.001$ y $b = 0.008$

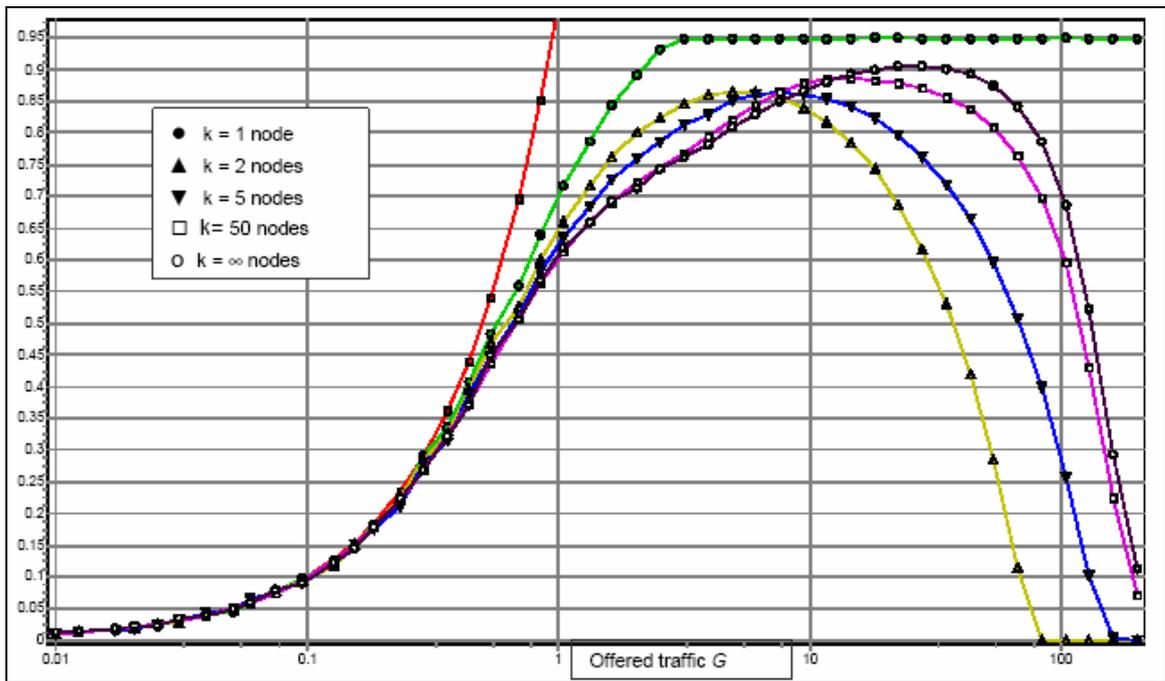


figura nº 98 Rendimiento CSMA/CD M/D/1/∞/k – Distinto número de nodos para $a = 0.01$ y $b = 0.008$

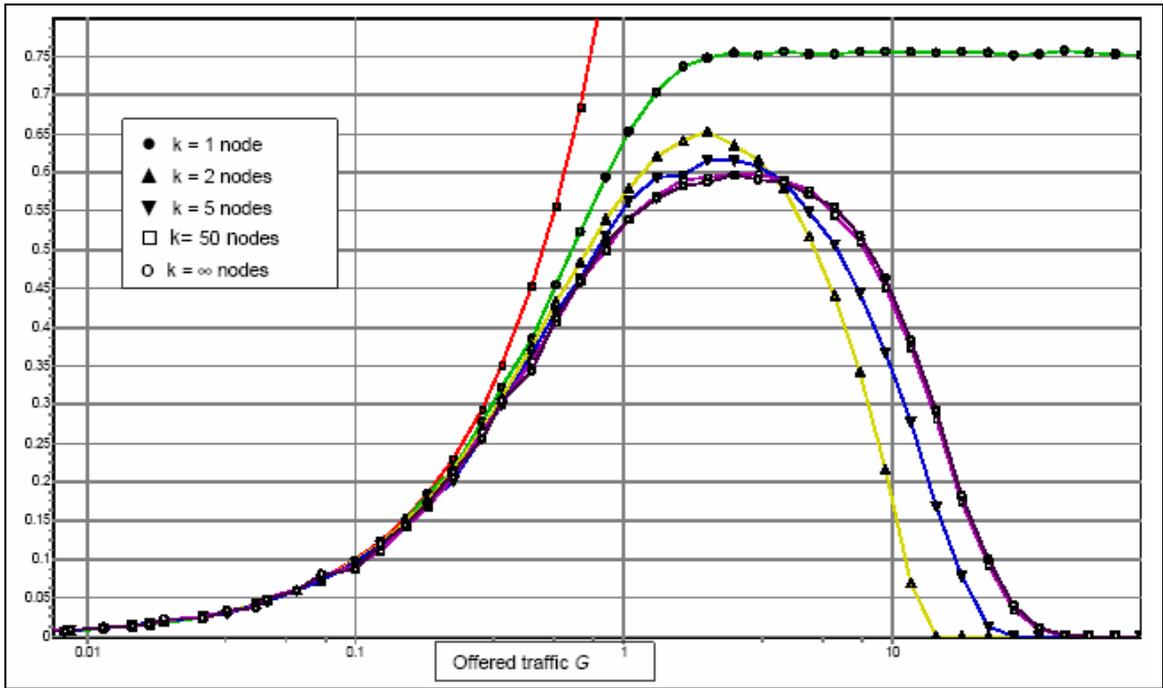


figura nº 99 Rendimiento CSMA/CD M/D/1/∞/k – Distinto número de nodos para $a = 0.1$ y $b = 0.008$

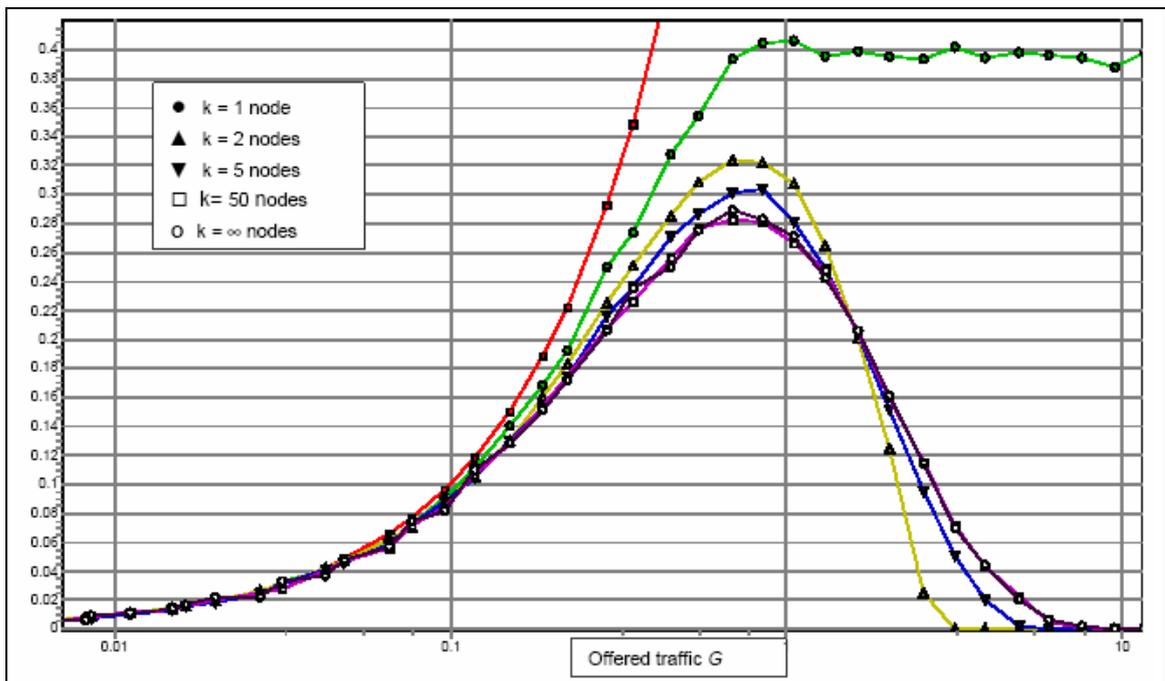


figura nº 100 Rendimiento CSMA/CD M/D/1/∞/k – Distinto número de nodos para $a = 0.99$ y $b = 0.008$

IV.5) EVALUACION DE PERFORMANCE DE ETHERNET

En este caso se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a Ethernet de II.3.9, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador.

En las figuras n° 101 y 102 se presentan las frecuencias de disparo de algunas transiciones graficadas como función de la carga de entrada I (la frecuencia de disparo de la transición T1). En promedio, las frecuencias de disparo de T20, T22, T24, T26 y T28 muestran las veces que un paquete fu retransmitido, una o más veces, tres o más veces, y así sucesivamente. La frecuencia de disparo de la transición T28 representa el número promedio de paquetes que fueron transmitidos ninguna o más que nueve veces. Se observa que para valores pequeños de la carga de entrada, el número de paquetes recuperados, representado por las frecuencias de disparo de T20 a T28 es pequeño. Para valores grandes del tráfico ofrecido, la frecuencia de disparo de T20 es casi la frecuencia de disparo de T1, lo que significa que la mayoría de las transiciones que vuelven al sistema fueron retransmitidas. Debido a que un paquete que arriba al sistema es transmitido exitosamente o descartado, la frecuencia de disparo de T5 más la frecuencia de disparo de T15 es igual ala frecuencia de disparo de T1.

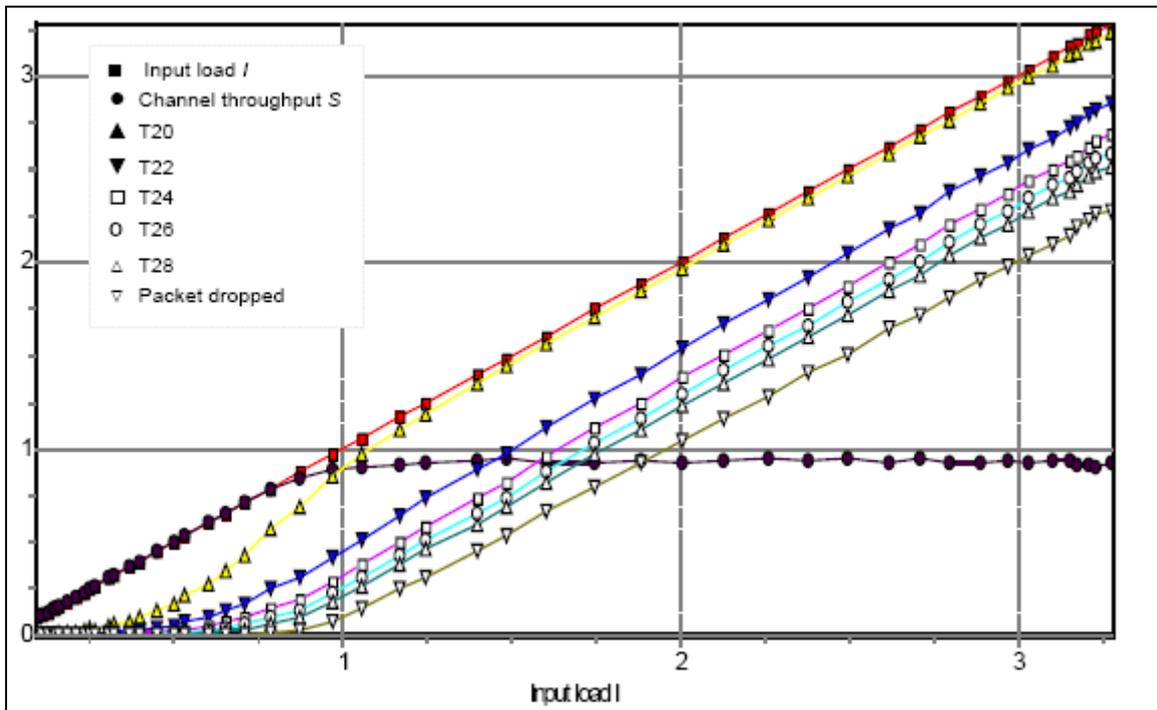


figura n° 101 Rendimiento de Ethernet – Frec. de disparo de transiciones seleccionadas para 500 nodos

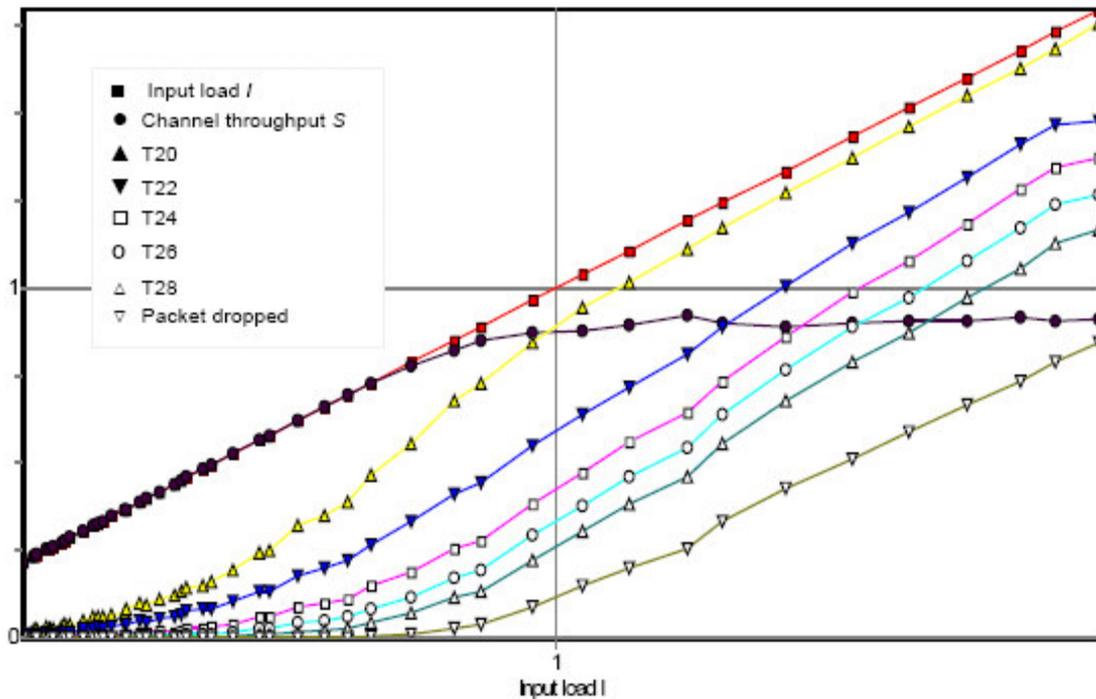


figura nº 102 Rendimiento de Ethernet – Frec. de disparo de transiciones seleccionadas para 1500 nodos

IV.6) EVALUACION DE PERFORMANCE DE CSMA no persistente

Nuevamente, se supone que se ha arrancado el simulador STPNPlay, cargando el modelo de Red de Petri correspondiente a CSMA no persistente de la figura nº 42, ingresado en el menú de simulación y lanzado la graficación de las curvas de rendimiento. Nuestra intención es mostrar el resultado gráfico que daría el simulador.

En la figura nº 103 se grafica las frecuencias de disparo de las transiciones T1, T2, T4, T5 y T6 como función de G variando desde 0.01 hasta 500, para valores $a = 0.01$.

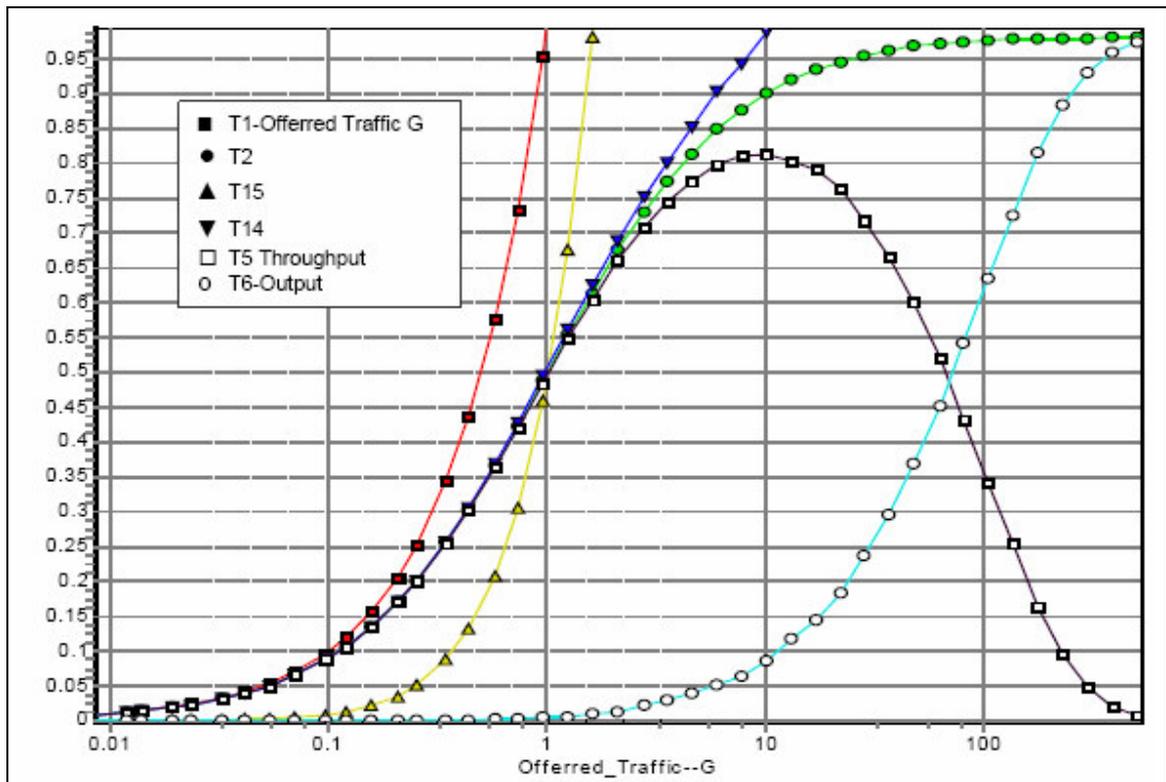


figura nº 103 Rendimiento de CSMA no persistente – Frecuencias de disparo de transición para $a = 0.01$

IV.7.) COMPARACION ENTRE LOS METODOS ANALITICOS Y EL SIMULADOR STPNPLAY USANDO FPQSPN [21] [36][37]

IV.7.1 Introducción

Las propiedades o características de los protocolos, analizadas a partir de los modelos FPQSPN y los simuladores de animación y rendimiento, permite concluir que este enfoque modelando con Redes de Petri es complementario con las de origen analítico. Además, que pone en evidencia aspectos sobre los protocolos que los métodos analíticos muestran parcialmente o son de imposible estudio.

Sin embargo, es imprescindible efectuar la comparación que asegure la validez funcional de los modelos FPQSPN y del simulador STPNPlay. Con dicho objeto, en este apartado se presentarán las comparaciones de los resultados obtenidos con el simulador STPNPlay (rendimiento en función de la carga) y los resultados de origen analítico. La calidad de la comparación asegurará la confiabilidad del simulador STPNPlay y la verificación de los modelos de los protocolos estudiados.

IV.7.2.1 Aloha Puro

En Aloha puro, los usuarios pueden transmitir en cualquier momento que ellos lo necesiten. Si fallan al escuchar su transmisión exitosa después de un retardo de propagación,

saben que ocurrió una colisión, y retransmite los paquetes después de un retardo aleatorio. Bajo condiciones de estado estable, la probabilidad P_s que un paquete ofrecido arbitrario sea exitoso está dado por la fórmula:

$$P_s = S / G$$

El proceso de ofrecimiento de paquetes al canal se asume que es una variable aleatoria con distribución Poisson:

$$P(X = x) = (G^x / x!) \cdot e^{-G} \quad \text{para todo } x=0,1,\dots$$

Cada vez que dos marcos tratan de ocupar el canal al mismo tiempo, habrá una colisión y ambos se alterarán. Si el primer bit de un marco nuevo se traslapa al último bit de un marco casi terminado, ambos marcos se destruirán por completo, y tendrán que retransmitirse después. El tiempo de marco es el tiempo necesario para transmitir un marco estándar de longitud fija (es decir, la longitud del marco dividida entre la tasa de bits). En este punto, se supone que la población infinita de usuarios genera marcos nuevos según una distribución de Poisson, con una media de S marcos por tiempo de marco. (La suposición de población infinita es necesaria para asegurar que S no disminuya a medida que se bloquean los usuarios.) Si S es mayor que 1, la comunidad de usuarios está generando marcos a una velocidad mayor que la que puede manejar el canal, y casi cada marco sufre una colisión. Para un rendimiento razonable esperamos que S esté entre 0 y 1.

Además de marcos nuevos, las estaciones también generan retransmisiones de marcos que previamente sufrieron colisiones. Supongamos además que la probabilidad de k intentos de transmisión por tiempo de marco, viejos y nuevos combinados, también es Poisson, con una media G por tiempo de marco. Claramente G es mayor que S . Con carga baja (es decir S casi 0), habrá pocas colisiones y, por lo tanto, pocas retransmisiones, por lo que G es casi S . Con carga alta habrá muchas colisiones, por lo que G es mayor que S . Con todas las cargas, el rendimiento es sólo la carga ofrecida G , por la probabilidad de que una transmisión tenga éxito (es decir, $S = G \cdot P_0$), donde P_0 es la probabilidad de que un marco no sufra colisión).

Un marco no sufrirá colisión si no se envían otros marcos durante una ranura de un tiempo de marco desde su inicio (ver figura nº 104). ¿En que condiciones llegará sin daño el marco sombreado?. Sea t el tiempo requerido par enviar un marco. Si cualquier otro usuario generó un marco entre el tiempo t_0 y $t_0 + t$, el final de ese marco chocará con el comienzo del marco sombreado. De hecho, el destino del marco sombreado se selló aún antes de enviar el primer bit, pero dado que en Aloha puro una estación no escucha el canal antes del transmitir, no tiene manera de saber que otro marco ya está en camino. De manera parecida, cualquier otro marco que salga entre $t_0 + t$ y $t_0 + 2t$ chocará con el final del marco sombreado.

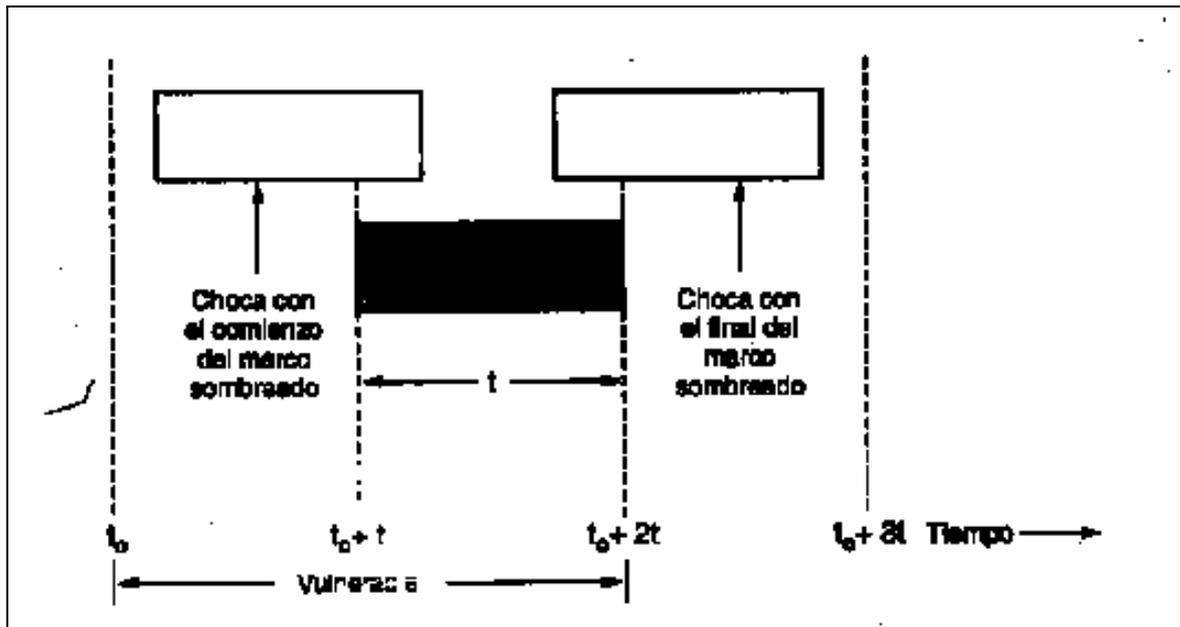


figura n° 104 Período vulnerable para el marco sombreado

Para una transmisión exitosa de un paquete es necesario que ningún conflicto ocurra durante $2t$. Esto significa que durante el periodo vulnerable $2t$ ningún otro nodo comience una transmisión. La probabilidad de que x marcos sean generados durante un tiempo de marco dado está dado y se asume que es una variable aleatoria con distribución de Poisson:

$$P(X = x) = (G^x / x!). e^{-G} \quad \text{para todo } x=0,1,\dots$$

Y la probabilidad de cero marcos es simplemente e^{-G} . En una ranura de dos tiempos de marco de longitud, el número medio de marcos es de $2G$. La probabilidad de que no se inicie otro tráfico durante todo el periodo vulnerable está dada entonces por

$$P(X = 0) = P_0 = e^{-2G}$$

Si $S = GP_0$, se obtiene que: $S = Ge^{-2G}$

La transmisión exitosa de un paquete significa que había exactamente un usuario transmitiendo en el canal durante el intervalo $2t$. El rendimiento del canal también puede ser visto como la probabilidad de que hay exactamente un usuario transmitiendo en el canal durante el “periodo vulnerable $2t$ ”

$$S = P(X = 1) = (G^1 / 1!). e^{-2G} = G.e^{-2G}$$

La relación entre el tráfico ofrecido y el rendimiento de acuerdo a esta expresión analítica de S , se muestra en la figura n° 105 generada con el software Mathematica (comparar con la figura n° 81 obtenida usando el simulador STPNPlay sobre el modelo de Aloha puro), y el cuadro de valores para G desde 0 hasta 10. El rendimiento máximo ocurre $G = 0,5$, con $S = 1/2e$, que es aproximadamente 0.184. En otras palabras, lo más que podemos esperar es un uso del canal del 18%.

En la figura n° 106 se muestra la tabla correspondiente a los valores analíticos y del simulador (con un cierto nivel de confiabilidad) para las curvas de las figuras n° 105 y n° 81, respectivamente. La primera columna es el tráfico ofrecido G (variable independiente de la

expresión matemática de S, eje de abscisas de la gráfica); la segunda columna es el rendimiento S (variable dependiente o función de la expresión, eje de ordenadas de la gráfica); la tercera columna son los resultados del rendimiento S obtenidos del simulador STPNPlay (Apéndice A.1); y la cuarta columna es el error relativo entre los resultados analíticos y del simulador. Se observa que el error relativo es despreciable, aún cuando el simulador se ejecuto en el modo standard de precisión que es configurable.

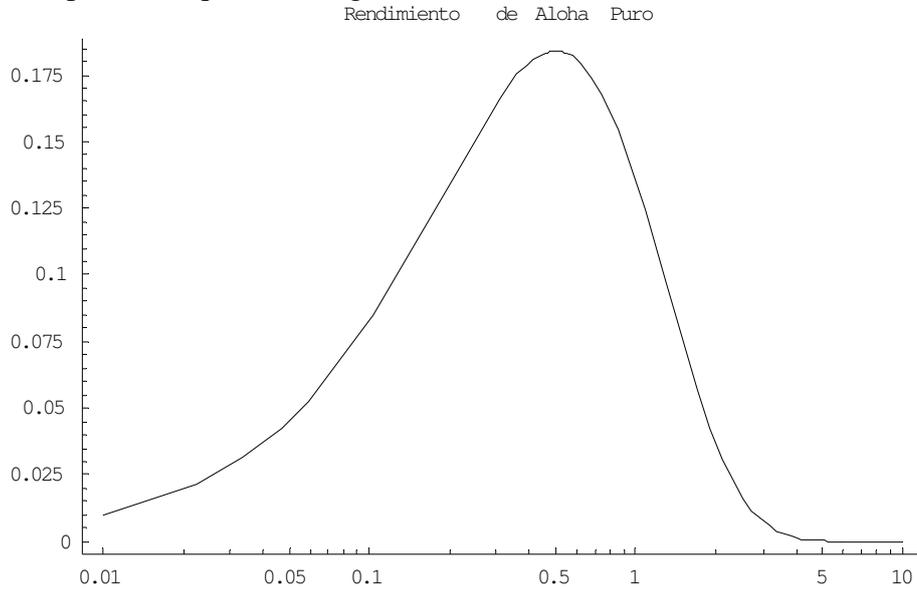


figura nº 105 Rendimiento S en función del tráfico ofrecido G Aloha puro

TABLA COMPARATIVA DE RENDIMIENTO EN ALOHA PURO

G	ANALITICO	SIMULADOR	ERROR REL
0,00457	0,00453	0,00453	-0,00035
0,00637	0,00629	0,00623	0,00944
0,00925	0,00908	0,00913	-0,00546
0,01387	0,01349	0,01345	0,00300
0,01939	0,01865	0,01877	-0,00630
0,02835	0,02679	0,02651	0,01035
0,03900	0,03607	0,03632	-0,00683
0,05835	0,05192	0,05180	0,00237
0,08382	0,07088	0,07092	-0,00052
0,11977	0,09426	0,09409	0,00178
0,17164	0,12177	0,12229	-0,00429
0,25208	0,15226	0,15463	-0,01557
0,35987	0,17521	0,17481	0,00230
0,52176	0,18377	0,18410	-0,00179
0,74221	0,16821	0,16775	0,00273
1,07631	0,12505	0,12391	0,00908
1,55106	0,06973	0,06937	0,00511
2,22561	0,02596	0,02598	-0,00076
3,21021	0,00523	0,00489	0,06432
4,63816	0,00043	0,00035	0,19392
5,52049	0,00009	0,00005	0,43503

figura nº 106 Comparación entre los valores de los métodos analítico y del simulador

IV.7.2.2 Aloha ranurado

En Aloha ranurado, el tiempo es ranurado en segmentos, cuya duración es exactamente al tiempo de transmisión de paquete t . La transmisión de paquete comienza siempre al inicio de un tiempo de slot. Al slotear el tiempo en segmentos causa que la probabilidad de ocurrencia de una colisión se reduzca y una colisión siempre afecte un intervalo de tiempo de longitud t . En caso de una colisión, dos o más paquetes se solapan completamente. El rendimiento del canal es la probabilidad que haya solo un usuario sobre el canal durante el “intervalo t ”.

$$S = P(X = 1) = (G^1 / 1!) \cdot e^{(-G)} = G \cdot e^{(-G)}$$

La relación entre el tráfico ofrecido y el rendimiento de acuerdo a esta expresión analítica de S , se muestra en la figura nº 107; generada con el software Mathematica, el Aloha ranurado tiene su pico en $G = 1$, con un rendimiento de $S = 1/e$, o aproximadamente 0.368, el doble de Aloha puro (comparar con la figura nº 85 obtenida usando el simulador STPNPlay sobre el modelo de Aloha ranurado). Lo mejor que podemos esperar usando Aloha ranurado es 37% de ranuras vacías, 37% de éxitos y 26% de colisiones.

En la figura nº 108 se muestra la tabla correspondiente a los valores analíticos y del simulador (con un cierto nivel de confiabilidad) para las curvas de las figuras nº 107 y nº 85, respectivamente. La primera columna es el tráfico ofrecido G (variable independiente de la expresión matemática de S , eje de abscisas de la gráfica); la segunda columna es el rendimiento S (variable dependiente o función de la expresión, eje de ordenadas de la gráfica); la tercera columna son los resultados del rendimiento S obtenidos del simulador STPNPlay (Apéndice A.2); y la cuarta columna es el error relativo entre los resultados analíticos y del simulador. Se observa que el error relativo es despreciable, aún cuando el simulador se ejecuto en el modo standard de precisión que es configurable.

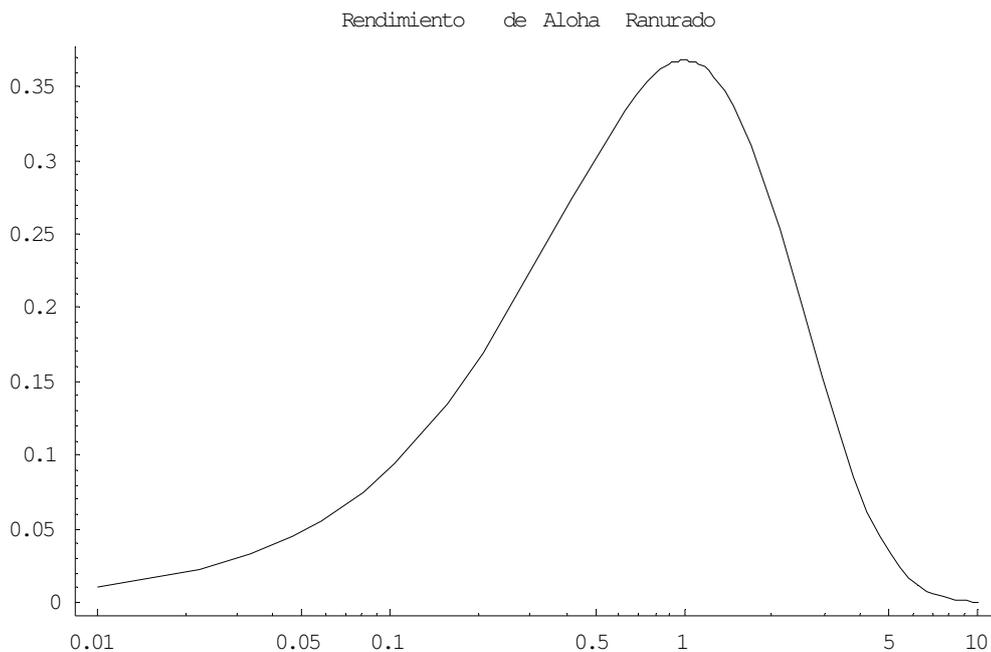


figura nº 107 Rendimiento S en función del tráfico ofrecido G Aloha ranurado

TABLA COMPARATIVA DE RENDIMIENTO EN ALOHA RANURADO

G	ANALITICO	SIMULADOR	ERROR REL
0,00473	0,00471	0,00467	0,00800
0,00663	0,00659	0,00659	-0,00058
0,00984	0,00974	0,00976	-0,00168
0,01348	0,01330	0,01332	-0,00154
0,01908	0,01872	0,01866	0,00317
0,02853	0,02773	0,02755	0,00640
0,04120	0,03954	0,03970	-0,00412
0,05775	0,05451	0,05460	-0,00166
0,08337	0,07670	0,07682	-0,00155
0,12193	0,10793	0,10942	-0,01377
0,16997	0,14340	0,14318	0,00155
0,25075	0,19514	0,19471	0,00219
0,35553	0,24916	0,24940	-0,00098
0,51888	0,30883	0,31077	-0,00628
0,74990	0,35426	0,35124	0,00853
1,07026	0,36701	0,36576	0,00341
1,54031	0,33011	0,32790	0,00670
2,24086	0,23835	0,23820	0,00064
3,21146	0,12941	0,12902	0,00305
4,62533	0,04533	0,04671	-0,03044
5,57317	0,02117	0,02113	0,00185

figura n° 108 Comparación entre los valores de los métodos analítico y del simulador

IV.7.3 CSMA

El esquema CSMA fue propuesto al principio de los setentas para ser usado en las comunicaciones vía canales de radio. Dos tipos básicos de métodos CSMA se presentan en la literatura: El CSMA persistente (también llamado CSMA I-persistente), que produce la transmisión de un paquete inmediatamente cuando el canal está libre, y el CSMA no persistente, que difiere la transmisión del paquete por un retardo aleatorio, que comienza inmediatamente después que el canal está libre.

IV.7.3.1 CSMA No persistente

En el CSMA no persistente, la idea para reducir la interferencia entre los paquetes es replanificar siempre un paquete que encuentra en canal estado ocupado. Un nodo que atiene un paquete para transmitir, opera de la siguiente manera:

- Si el canal es sensado libre, el nodo transmite el paquete inmediatamente,
- Cuando el nodo sensa que canal ocupado, el planifica la retransmisión del paquete de acuerdo a la distribución del retardo de retransmisión. El procedimiento MAC repite este nuevo punto en el tiempo.

El CSMA no persistente difiere del CSMA persistente porque replanifica un paquete después que el canal es sentido ocupado. Mientras que CSMA persistente transmite el paquete inmediatamente después que el canal se vuelve libre, el CSMA no persistente replanifica la transmisión del paquete. Para pequeños G , este procedimiento extiende el tiempo medio que un paquete gasta en el sistema. Para grandes G , por el contrario, reduce la probabilidad de colisión después que el canal se vuelve libre, de modo que el rendimiento de canal máximo es grande comparado c CSMA persistente.

Recordemos que G es la frecuencia de arribo de los paquetes nuevos y replanificados, los cuales constituyen la carga del canal ofrecida, pero en algunos casos como en los modelos sin recuperación de CSMA no persistente, sólo una fracción de G constituye el tráfico del canal a causa de los paquetes que encuentra el canal ocupado son replanificados sin ser transmitidos y el mecanismo de replanificación no es parte del modelo (hablamos de un sistema con clientes perdidos).

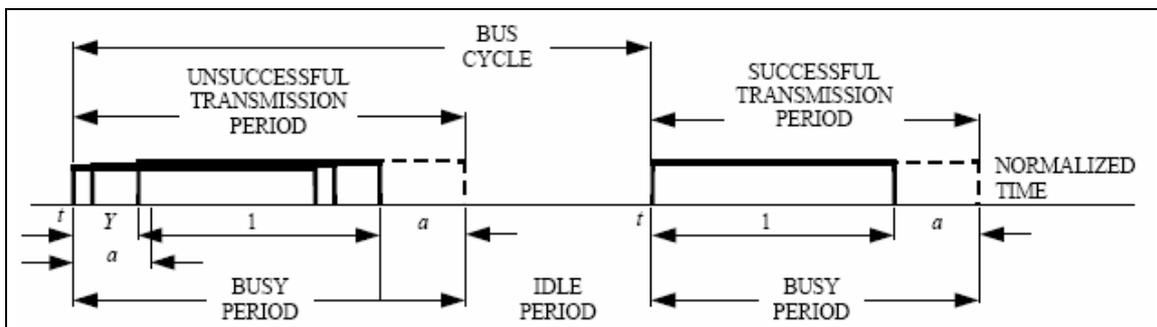


figura n° 109 CSMA no persistente: periodos ocupados y libres

Consideremos el eje x en la figura n° 109, y nombramos a t el tiempo del arribo del paquete que sensa el canal libre, y tal que ningún otro paquete está en el proceso de transmisión. Cualquier otro paquete, que arribe entre t y $t+a$, encontrará el canal en estado libre. Tal paquete será transmitido y por lo tanto, causará un conflicto. Si ningún otro paquete es transmitido al canal durante el “periodo vulnerable” a , entonces el primer paquete será transmitido exitosamente.

Nombramos a $t+Y$ como el tiempo de la ocurrencia del último paquete arribando entre t y $t+a$. Sólo a segundos luego el canal será sentido libre por todos los nodos. Ahora, cualquier nodo que estaba listo entre $(t+a$ y $t+Y+1+a)$ sensará el canal ocupado, y por lo tanto, replanificará su paquete. El intervalo entre $(t$ y $t+Y+1+a)$ es llamado el periodo de transmisión TP, que en este caso también es el periodo ocupado. Se nota que puede haber al menos una transmisión exitosa durante TP. Se define al periodo libre al tiempo entre dos TP consecutivos. Un periodo ocupado más el siguiente periodo libre constituyen el ciclo.

Llamemos B a la duración esperada del periodo ocupado, I a la duración esperada del periodo libre, y $B + I$ a la longitud esperada de un ciclo. Llamemos U el tiempo durante el ciclo que el canal está usado sin conflictos. Usando los argumentos, la utilización del canal promedio está dado simplemente por

$$S = U / B+I$$

La duración promedio del periodo libre es simplemente $1/G$. La duración promedio del intervalo ocupado es $1+Y+a$, donde Y es el valor esperado de Y . La probabilidad

que un TP sea exitoso es simplemente la probabilidad que ninguna terminal transmita durante los primeros a segundos del periodo. Por lo tanto,

$$U = e^{-aG}$$

Se deduce que la ecuación del rendimiento S de CSMA no persistente, tanto en su versión pura y sloteada son:

$$S = (G \cdot e^{-aG}) / (G(1+2a) + e^{-aG})$$

$$S = (a \cdot G \cdot e^{-aG}) / (1 - e^{-aG} + a)$$

La relación entre el tráfico ofrecido y el rendimiento de acuerdo a la segunda expresión analítica de S, para $a = 0,01$, se muestra e en la figura n° 110, obtenida con el software Mathematica, el CSMA no persistente puro tiene su pico en $G = 10$, con un rendimiento de aproximadamente 0.86 (comparar con la figura n° 103 obtenida usando el simulador STPNPlay sobre el modelo de Aloha CSMA no persistente).

En la figura n° 111 se muestra la tabla correspondiente a los valores analíticos y del simulador (con incierto nivel de confiabilidad) para las curvas de las figuras n° 110 y n° 103, respectivamente. La primera columna es el tráfico ofrecido G (variable independiente de la expresión matemática de S, eje de abscisas de la gráfica); la segunda columna es el rendimiento S (variable dependiente o función de la expresión, eje de ordenadas de la gráfica); la tercera columna son los resultados del rendimiento S obtenidos del simulador STPNPlay (Apéndice A.3); y la cuarta columna es el error relativo entre los resultados analíticos y del simulador. Se observa que el error relativo es despreciable, aún cuando el simulador se ejecuto en el modo standard de precisión que es configurable.

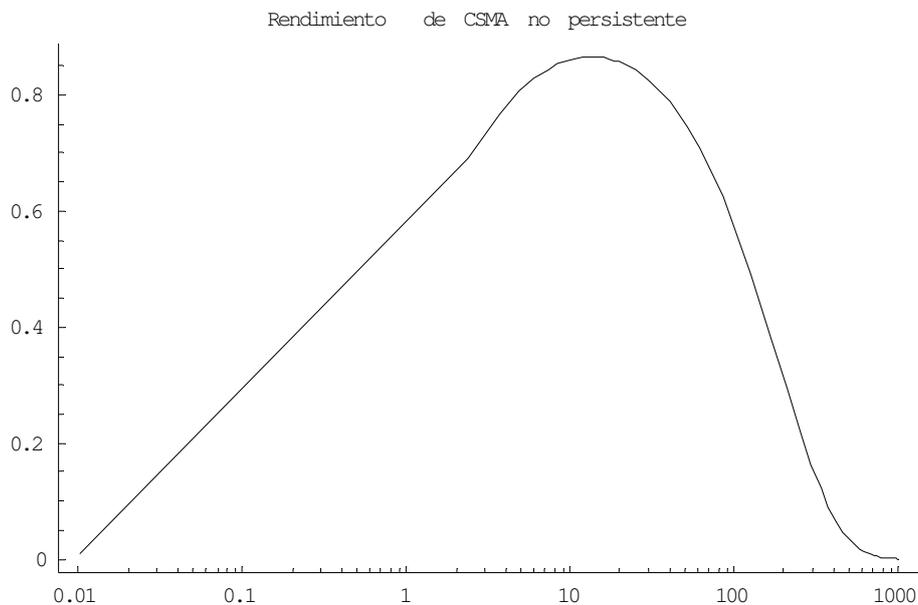


figura n° 110 Rendimiento S en función del tráfico ofrecido G CSMA no persistente

TABLA COMPARATIVA DE RENDIMIENTO EN CSMA NO PERSISTENTE

G	ANALITICO	SIMULADOR	ERROR REL
0,00145	0,00145	0,00144	0,00198
0,00241	0,00240	0,00240	0,00129
0,00414	0,00412	0,00412	0,00136
0,00676	0,00672	0,00672	-0,00035
0,01150	0,01136	0,01136	-0,00002
0,01923	0,01886	0,01887	-0,00073
0,03237	0,03134	0,03158	-0,00757
0,05478	0,05188	0,05192	-0,00073
0,09209	0,08417	0,08411	0,00077
0,15547	0,13416	0,13423	-0,00052
0,26131	0,20621	0,20643	-0,00111
0,43984	0,30321	0,30324	-0,00012
0,74369	0,42154	0,42119	0,00082
1,24896	0,54541	0,54549	-0,00014
2,10516	0,65932	0,65943	-0,00017
3,54466	0,74687	0,74659	0,00037
5,96670	0,79980	0,79991	-0,00014
10,04130	0,81478	0,81454	0,00030
16,90870	0,78924	0,78896	0,00035
28,47090	0,71886	0,71903	-0,00023
47,93580	0,59945	0,59913	0,00053
80,72440	0,43498	0,43575	-0,00177
135,90800	0,25139	0,25154	-0,00056
228,83600	0,09940	0,09908	0,00327
385,34200	0,02079	0,02093	-0,00653

figura nº 111 Comparación entre los valores de los métodos analítico y del simulador

IV.7.3.2 CSMA Persistente

El CSMA Persistente obtiene el rendimiento razonablemente máximo del canal al ocupar el canal libre cuando cualquier nodo tiene datos para transmitir. Un nodo, que tiene datos para la transmisión, opera de la siguiente forma:

- Cuando el canal es sentido libre, el transmite el paquete,
- Cuando el canal es sentido ocupado, el nodo espera hasta que el canal se vuelve libre, y luego transmite el paquete.

El diagrama de tiempo del CSMA persistente se muestra en la figura nº 112.

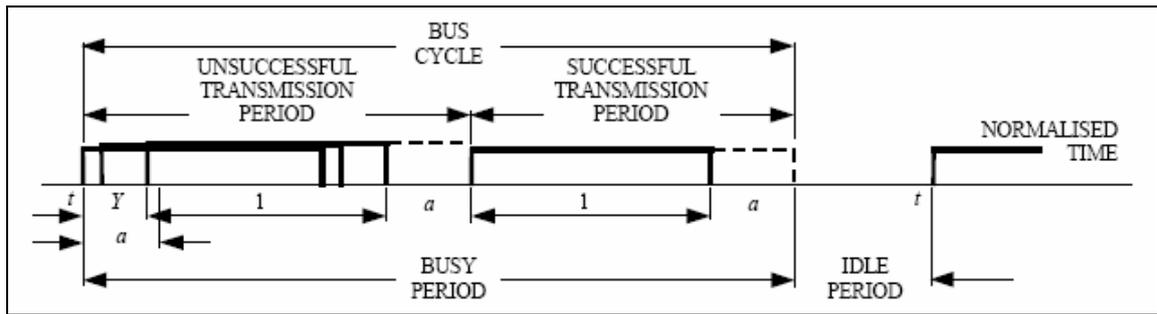


figura nº 112 CSMA persistente: periodos ocupados y libres

En este protocolo, cualquier paquete que arriba en el intervalo de tiempo $[t+a, t+Y+1+a]$ sentirá el canal ocupado, y por lo tanto, esperará hasta que el canal sea sentido libre (tiempo $t+Y+1+a$). En ese momento, el paquete es transmitido. El número de paquetes acumulados en el extremo de TP es el número de los arribos durante $(1+Y)$ segundos. Si este total iguala a 0 o es mayor que 2, entonces un conflicto seguramente ocurrirá en el siguiente TP.

El rendimiento del canal obtenido para CSMA persistente en sus versiones puras y ranuradas, están dadas por las siguientes fórmulas:

$$S = G[1 + G + aG.(1 + G + aG/2)]. e^{-G(1+2a)} / G.(1+2a) - (1 - e^{-aG}) + (1+aG). e^{-G(1+a)}$$

$$S = G. e^{-G(1+2a)} (1 + a. e^{-aG}) / (1 + a).(1 - e^{-aG}) + a e^{-G(1+a)}$$

La relación entre el tráfico ofrecido y el rendimiento de acuerdo a la segunda expresión analítica de S , para $a = 0,01$, se muestra en la figura nº 113, generada con el software Mathematica, el CSMA persistente puro tiene su pico en $G = 1$, con un rendimiento de aproximadamente 0.55 (comparar con la figura nº 87 obtenida usando el simulador STPNPlay sobre el modelo de CSMA persistente).

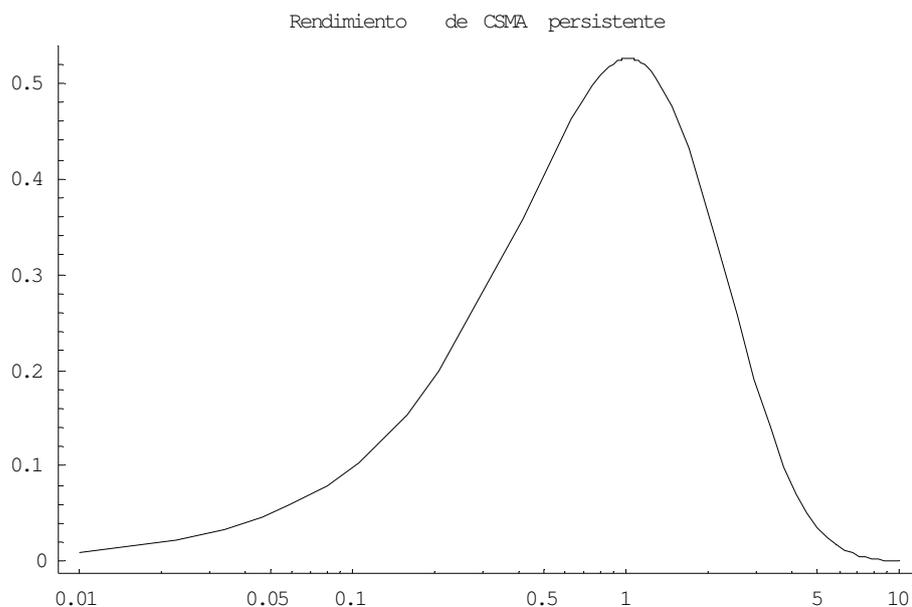


figura nº 113 Rendimiento S en función del tráfico ofrecido G CSMA persistente

En la figura n° 114 se muestra la tabla correspondiente a los valores analíticos y del simulador para las curvas de las figuras n° 113 y n° 87, respectivamente. La primera columna es el tráfico ofrecido G (variable independiente de la expresión matemática de S, eje de abscisas de la gráfica); la segunda columna es el rendimiento S (variable dependiente o función de la expresión, eje de ordenadas de la gráfica); la tercera columna son los resultados del rendimiento S obtenidos del simulador STPNPlay (Apéndice A.4); y la cuarta columna es el error relativo entre los resultados analíticos y del simulador. Se observa que el error relativo es despreciable, aún cuando el simulador se ejecuto en el modo standard de precisión que es configurable.

TABLA COMPARATIVA DE RENDIMIENTO EN CSMA PERSISTENTE

G	ANALITICO	SIMULADOR	ERROR REL
0,00134	0,00134	0,00134	-0,00003
0,00244	0,00244	0,00244	-0,00005
0,00401	0,00401	0,00401	-0,00010
0,00681	0,00681	0,00681	-0,00018
0,01128	0,01128	0,01128	-0,00035
0,01954	0,01952	0,01954	-0,00077
0,03310	0,03304	0,03303	0,00038
0,05359	0,05338	0,05336	0,00043
0,07272	0,07225	0,07214	0,00145
0,09219	0,09128	0,09134	-0,00070
0,15508	0,15120	0,15116	0,00025
0,26024	0,24394	0,24424	-0,00125
0,44106	0,37390	0,37431	-0,00110
0,74330	0,49877	0,49750	0,00254
1,24972	0,51373	0,51445	-0,00140
2,10122	0,34762	0,34744	0,00051
3,52262	0,12649	0,12622	0,00212
5,96554	0,01663	0,01675	-0,00702
10,04750	0,00043	0,00046	-0,07922
16,90400	0,00000	0,00000	1,00000
28,48910	0,00000	0,00000	1,00000
47,91460	0,00000	0,00000	1,00000
80,74340	0,00000	0,00000	1,00000
135,96500	0,00000	0,00000	1,00000
228,86200	0,00000	0,00000	1,00000

figura n° 114 Comparación entre los valores de los métodos analítico y del simulador

IV.7.3.3 CSMA P-Persistente

El CSMA p-persistente generaliza el CSMA persistente en el sentido que el nodo transmite con probabilidad p, después de sensar el canal libre. Un nodo que tiene datos para la transmisión opera de la siguiente forma:

- Cuando el canal es sensado libre, transmite el paquete con una probabilidad p. Con la probabilidad (1-p), el nodo retarda la transmisión del paquete por un slot de tiempo a. En este nuevo punto en el tiempo, si el canal es aún sensado libre, el mismo procedimiento

se repite nuevamente, el nodo planifica la retransmisión del paquete de acuerdo a la distribución de retardo de retransmisión.

- Cuando el canal es sentido ocupado, el nodo espera hasta que el canal se vuelve libre, y luego opera como antes.

Ambos métodos, CSMA persistente y no persistente existen en sus modificaciones puras y ranuradas. Obviamente, solo la versión ranurada de CSMA p-persistente existe.

IV.7.3.4 CSMA Mp-persistente

CSMA Mp-persistente incorpora las ventajas de los métodos CSMA persistente y no-persistente. Es posible decir, que CSMA persistente y no persistente con casos especiales de CSMA Mp-persistente. Un nodo que tiene un dato para transmitir, opera de la siguiente manera:

- Si el canal es sentido libre, el transmite el paquete,
- Si el canal es sentido ocupado, entonces el nodo planifica la retransmisión del paquete con probabilidad (1-p) de acuerdo a la distribución de retardo de retransmisión, y con probabilidad p, y espera hasta que el canal se vuelve libre.

El rendimiento del canal de CSMA Mp-persistente ranurado está dado por la siguiente fórmula. Se observa que no se analizan cualesquier otros métodos CSMA que las versiones persistentes y no-persistente, pero mencionamos algunos métodos CSMA de interés educacional. Por ejemplo, la fórmula de rendimiento Mp-persistente es interesante en simplicidad.

$$S = pG + aG - pG \cdot e^{-aG} / a + (1 + a) \cdot (e^{(a+p)G}) - e^{-pG}$$

IV.7.3.5 CSMA Mp-persistente Optimo

El CSMA p-persistente Óptimo también incorpora los métodos CSMA persistente y no persistente. El optimiza el rendimiento del canal al variar el parámetro p usado en el método CSMA Mp-persistente. En el caso óptimo el método provee un rendimiento que iguala a o es mejor que el máximo del rendimiento CSMA persistente y no persistente. El rendimiento del canal de CSMA p-persistente óptimo está dado por la siguiente fórmula. Un nodo que tiene un dato para transmitir opera de la siguiente forma:

- Si el canal se sentido libre, transmite el paquete,
- Si el canal es sentido ocupado, entonces el nodo planifica la retransmisión del paquete, y entonces con probabilidad (1-p) de acuerdo a la distribución de retardo de retransmisión, y con probabilidad p, el nodo espera hasta que el canal se vuelva libre.
- El parámetro p se varía con respecto a la utilización del canal.

$$S = \max [pG + aG - pG \cdot e^{-aG} / a + (1 + a) \cdot (e^{(a+p)G}) - e^{-pG}], p=[0,..1]$$

Las curvas de rendimiento de la figura número n° 115, fueron graficadas usando las fórmulas de Aloha ranurado, de CSMA MP-persistente y CSMA Mp-persisten óptimo. El rendimiento del canal CSMA depende del retardo de propagación del canal *a*. Por ejemplo, cuando el retardo de propagación del canal iguala aproximadamente al tiempo de transmisión del paquete, los métodos CSMA dan peores resultados que Aloha ranurado.

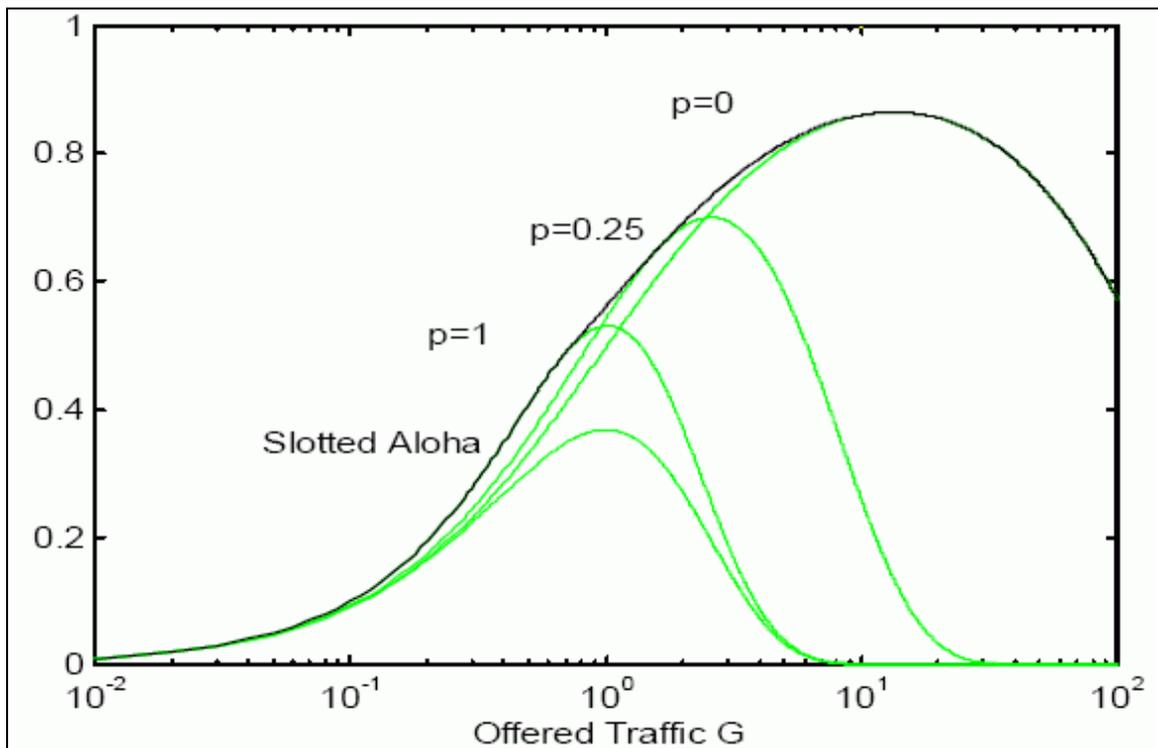


figura nº 115 Rendimiento del Canal de Aloha ranurado y métodos CSMA para $a = 0.01$

IV.8) CONCLUSION

Dentro de los objetivos y la metodología a seguir en el trabajo de tesis, se han presentado los principios de evaluación de performance, se ha seleccionado una herramienta para la evaluación del performance, y efectuado la comparación entre los resultados de métodos analíticos y los resultados del simulador seleccionado.

En este Capítulo se han presentado los aspectos teóricos sobre Principios de Evaluación de Performance, y específicamente de los protocolos de acceso al medio. Luego, se ha seleccionado la herramienta de simulación para evaluación de performance STPNPlay, que se corresponde con la propuesta de los modelos FPQSPN. Es una herramienta de simulación para modelación interactiva que funciona en el entorno Windows. El énfasis en su desarrollo ha sido la exactitud y la velocidad del algoritmo de simulación para calcular performance, y la portabilidad de los modelos del sistema y los resultados a otras herramientas y plataformas. El simulador puede accederse gratuitamente para fines educativos y de investigación desde el sitio WEB: <http://dce.felk.cvut.cz/capekj/StpnPlay/>. Esta herramienta no permite la animación del protocolo, por lo que es imposible ver la dinámica del protocolo como en HPSim.

Luego, se han incluido las gráficas de los resultados de la evaluación de performance para los protocolos modelizados, que podría verse en el simulador STPNPlay en funcionamiento, para distintos casos, como sería la cantidad de nodos ó el retardo de propagación.

Además, se ha puesto de manifiesto que a partir de las propiedades o características de los protocolos, analizadas a partir de los modelos FPQSPN y STPNPlay, que este enfoque modelando con Redes de Petri es complementario con las de origen analítico. Además, que pone en evidencia aspectos sobre los protocolos que los métodos analíticos muestran parcialmente o son de imposible estudio.

Finalmente, se ha efectuado la comparación que asegure la validez funcional de los modelos FPQSPN y del simulador STPNPlay. Con dicho objeto, en este apartado se han efectuado las comparaciones de los resultados obtenidos con el simulador STPNPlay (rendimiento en función de la carga) y los resultados de origen analítico. La calidad de la comparación asegura la confiabilidad del simulador STPNPlay y la verificación de los modelos de los protocolos estudiados.

CAPITULO V

CONCLUSIONES

V.1 Conclusiones

En la tesis se han presentado las Redes de Petri FPQSPN como la herramienta matemática de modelación para establecer de una manera uniforme, las características, similitudes y diferencias cualitativas y cuantitativas, entre algunos de los protocolos de red de canal compartido utilizados en el ámbito de las redes LAN y WAN. Además, se han seleccionado herramientas software de simulación que han demostrado su utilidad para complementar los estudios analíticos sobre los protocolos propuestos, y favorecer las interpretaciones de las características MAC a partir de sus animaciones con HPSIM, y de la mejora incremental del rendimiento con el uso de un analizador de performance STPNPlay.

Se realizaron las siguientes actividades de tesis:

- 1) Se compilaron las novedades teórico-prácticas de las Redes de Petri, y se seleccionó el tipo más apto de Red de Petri para la modelación de los protocolos de comunicaciones. Se presentó los aspectos teóricos sobre Redes de Colas, las Redes de Petri No Temporizadas y Temporizadas. Se presentaron tablas comparativas de las Redes de Petri y las Redes de Colas, y de las Redes de Petri entre sí.
- 2) A los efectos de los objetivos del trabajo de tesis, se seleccionó las FPQSPN (Finite Population Queuing Systems Petri Nets). FPQSPN es una clase de Redes de Petri que es apropiada para modelación simplificada de sistemas de cola con población finita. La ventaja principal para su adopción es que FPQSPN permite modelación eficiente de los sistemas de cola de población finita, y mantiene la transparencia de las Redes de Petri de Bajo Nivel. Además, a la propuesta de FPQSPN se asocia un algoritmo eficiente para el cálculo computacional de rendimiento y graficación de funciones de algunas variables, usando representaciones matriciales.
- 3) Se han presentado los aspectos teóricos sobre los Protocolos de Red, sus medidas de performance y características de tráfico de red. Ha quedado en evidencia que las FPQSPN son apropiadas para mantener la transparencia de las Redes de Petri de Bajo Nivel, siendo relativamente fácil observar la creciente complejidad de los modelos desde Aloha puro. Efectivamente, desde la representación gráfica del modelo FPQSPN de Aloha puro puede observarse que la modelación es del canal y que los estados (lugares) básicos del mismo son: libre, ocupado, y colisionado. Y a este modelo básico se han agregado nuevos estados (lugares) que corresponden al funcionamiento ranurado, detección de portadora o detección de colisiones, para los restantes protocolos superiores, cuyo objeto es el incremento de la performance.
- 4) Se ha seleccionado una herramienta para la simulación de la animación de protocolos, para observar la dinámica, investigar propiedades o características. Se presentaron los aspectos teóricos sobre Principios de Simulación y una compilación de simuladores de Redes de Petri comerciales o de libre disponibilidad, definiendo sus características, funcionalidades y áreas de aplicación. Se adoptó el simulador HPSim que provee animación, y

aunque no provee análisis de performance avanzada puede usarse para un rudimentario estudio de la cuenta y movimiento de las marcas en los diferentes estados.

- 5) Se han presentado los principios de evaluación de performance, se ha seleccionado una herramienta para la evaluación del performance, y efectuado la comparación entre los resultados de métodos analíticos y los resultados del simulador seleccionado. Se seleccionó la herramienta de simulación para evaluación de performance STPNPlay, que se corresponde con la propuesta de los modelos FPQSPN. El énfasis en su desarrollo ha sido la exactitud y la velocidad del algoritmo de simulación para calcular performance.. Esta herramienta no permite la animación del protocolo, por lo que es imposible ver la dinámica del protocolo como en HPSim.
- 6) Se han incluido las gráficas de los resultados de la evaluación de performance para los protocolos modelizados, que podría verse en el simulador STPNPlay en funcionamiento, para distintos casos, como sería la cantidad de nodos ó el retardo de propagación. Ha quedado expuesto que a partir de las propiedades o características de los protocolos, analizadas a partir de los modelos FPQSPN y STPNPlay, que este enfoque modelando con Redes de Petri es complementario con las de origen analítico. Además, que pone en evidencia aspectos sobre los protocolos que los métodos analíticos muestran parcialmente o son de imposible estudio.
- 7) Se ha efectuado la comparación que asegure la validez funcional de los modelos FPQSPN y del simulador STPNPlay. Con dicho objeto, se han efectuado las comparaciones de los resultados obtenidos con el simulador STPNPlay (rendimiento en función de la carga) y los resultados de origen analítico. La calidad de la comparación asegura la confiabilidad del simulador STPNPlay y la verificación de los modelos de los protocolos estudiados.

V.2 Sumario de Contribuciones

Se han alcanzado los siguientes resultados:

- 1) Se demostró la validez de las Redes de Petri FPQSPN, y del simulador STPNPlay para la modelación y su utilidad para estudiar y evaluar aspectos que en otros tipos de estudios analíticos no son de fácil obtención, o que permitan demostrar que ambos enfoques son complementarios, tales como la evaluación de la dependencia del número de nodos y del retardo de propagación.
- 2) Se demostró una de ventajas comparativas de las Redes de Petri para el estudio del comportamiento de sistemas, como es su naturaleza gráfica, que facilita la interpretación, comunicación y transferencia entre los interesados que usan y discuten sobre los protocolos de red,
- 3) Se puso de manifiesto que dado los fundamentos matemáticos del formalismo es posible efectuar un enfoque integrado de validación de los parámetros de eficiencia para la comparación entre los protocolos, y que los protocolos superiores han seguido un curso de perfeccionamiento del protocolo Aloha, como modelo base,
- 4) Se efectuó una comparación entre los resultados de origen analíticos y los obtenidos entre los modelos de Red de Petri FPQSPN para los protocolos seleccionados (carga ofrecida y rendimiento). Los resultados de la comparación aseguran la confiabilidad del

simulador STPNPlay y la verificación de los modelos de los protocolos estudiados, para su utilización académica.

- 5) Se ha planteado un enfoque alternativo para la enseñanza-aprendizaje de los protocolos de red de canal compartido, desarrollando al efecto una secuencia enseñanza aprendizaje complementaria de los protocolos de red de canal compartido, usando este instrumento matemático, y las herramientas de software seleccionadas (Anexo C).

APENDICE A: TABLAS CON LOS VALORES NUMERICOS DE SIMULACION

A.1 Tabla de Aloha puro exportada desde el simulador STPNPlay

	Tráfico Ofrecido G	Disparo de T2	Disparo de T3	Disparo de T4	Rendimiento S	Disparo de T6	
220,00000	0,004570	0,004550	0,000020	0,000000	0,004530	0,000020	0,000000
183,34300	0,005640	0,005580	0,000060	0,000000	0,005520	0,000060	0,000000
152,79400	0,006370	0,006300	0,000070	0,000000	0,006230	0,000070	0,000000
127,33600	0,007840	0,007790	0,000050	0,000000	0,007740	0,000050	0,000000
106,11900	0,009250	0,009190	0,000060	0,000000	0,009130	0,000060	0,000000
88,43700	0,011120	0,011010	0,000110	0,000000	0,010900	0,000110	0,000000
73,70150	0,013870	0,013660	0,000210	0,000000	0,013450	0,000210	0,000000
61,42130	0,016160	0,016000	0,000160	0,000000	0,015840	0,000160	0,000000
51,18720	0,019390	0,019080	0,000310	0,000000	0,018770	0,000310	0,000000
42,65830	0,024110	0,023390	0,000690	0,000030	0,022690	0,000690	0,000030
35,55050	0,028350	0,027420	0,000910	0,000020	0,026510	0,000910	0,000020
29,62700	0,034120	0,032930	0,001170	0,000020	0,031760	0,001170	0,000020
24,69050	0,039000	0,037630	0,001300	0,000070	0,036320	0,001300	0,000070
20,57660	0,050310	0,047640	0,002530	0,000140	0,045110	0,002530	0,000140
17,14810	0,058350	0,054990	0,003190	0,000170	0,051800	0,003190	0,000170
14,29080	0,069960	0,065340	0,004310	0,000310	0,061030	0,004310	0,000310
11,90970	0,083820	0,077080	0,006160	0,000580	0,070920	0,006160	0,000580
9,92526	0,101590	0,091790	0,008800	0,001000	0,082980	0,008800	0,001000
8,27150	0,119770	0,106200	0,012100	0,001470	0,094090	0,012100	0,001470
6,89329	0,144500	0,124890	0,016560	0,003050	0,108330	0,016560	0,003050
5,74472	0,171640	0,144780	0,022490	0,004370	0,122290	0,022490	0,004370
4,78752	0,208000	0,169340	0,031750	0,006910	0,137590	0,031750	0,006910
3,98982	0,252080	0,197530	0,042900	0,011650	0,154630	0,042890	0,011650
3,32503	0,300020	0,222050	0,057450	0,020520	0,164590	0,057450	0,020520
2,77101	0,359870	0,250820	0,076010	0,033040	0,174810	0,076010	0,033040
2,30930	0,432160	0,280040	0,099220	0,052900	0,180820	0,099220	0,052900
1,92452	0,521760	0,309950	0,125850	0,085960	0,184100	0,125840	0,085960
1,60385	0,623240	0,332190	0,153730	0,137320	0,178460	0,153730	0,137320
1,33662	0,742210	0,353710	0,185960	0,202540	0,167750	0,185960	0,202540
1,11391	0,899160	0,366570	0,216890	0,315700	0,149680	0,216890	0,315700
0,92831	1,076310	0,366600	0,242690	0,467020	0,123910	0,242690	0,467020
0,77363	1,290500	0,355290	0,257940	0,677270	0,097350	0,257930	0,677270
0,64473	1,551060	0,329420	0,260050	0,961590	0,069370	0,260040	0,961590
0,53730	1,858590	0,290220	0,245160	1,323210	0,045060	0,245150	1,323210
0,44778	2,225610	0,241030	0,215050	1,769530	0,025980	0,215050	1,769530
0,37317	2,664100	0,184850	0,172520	2,306730	0,012330	0,172510	2,306730
0,31099	3,210210	0,129530	0,124640	2,956040	0,004890	0,124630	2,956040
0,25917	3,852880	0,082070	0,080080	3,690730	0,001990	0,080070	3,690730
0,21599	4,638160	0,044540	0,044190	4,549430	0,000350	0,044180	4,549430
0,18000	5,564140	0,021860	0,021790	5,520490	0,000070	0,021780	5,520490

A.2 Tabla de Aloha ranurado exportada desde el simulador STPNPlay

	Tráfico Ofrecido G	Disparo de T2	Disparo de T3	Disparo de T4	Rendimiento S	Disparo de T6					
220,00000	0,00473	0,00470	0,00003	0,00000	0,00467	0,00003	0,00000	0,00473	0,99990	0,99991	0,00473
183,34300	0,00524	0,00522	0,00002	0,00000	0,00520	0,00002	0,00000	0,00524	0,99990	0,99991	0,00524
152,79400	0,00663	0,00661	0,00002	0,00000	0,00659	0,00002	0,00000	0,00663	0,99990	0,99991	0,00663
127,33600	0,00755	0,00751	0,00004	0,00000	0,00747	0,00004	0,00000	0,00755	0,99990	0,99991	0,00755
106,11900	0,00984	0,00980	0,00004	0,00000	0,00976	0,00004	0,00000	0,00984	0,99990	0,99991	0,00984
88,43700	0,01131	0,01125	0,00006	0,00000	0,01119	0,00006	0,00000	0,01131	0,99990	0,99991	0,01131
73,70150	0,01348	0,01340	0,00008	0,00000	0,01332	0,00008	0,00000	0,01348	0,99990	0,99991	0,01348
61,42130	0,01655	0,01636	0,00019	0,00000	0,01617	0,00019	0,00000	0,01655	0,99990	0,99991	0,01655
51,18720	0,01908	0,01887	0,00021	0,00000	0,01866	0,00021	0,00000	0,01908	0,99990	0,99991	0,01908
42,65830	0,02350	0,02325	0,00024	0,00001	0,02301	0,00024	0,00001	0,02350	0,99990	0,99991	0,02350
35,55050	0,02853	0,02804	0,00049	0,00000	0,02755	0,00049	0,00000	0,02853	0,99990	0,99991	0,02853
29,62700	0,03394	0,03335	0,00058	0,00001	0,03277	0,00058	0,00001	0,03394	0,99990	0,99991	0,03394
24,69050	0,04120	0,04044	0,00074	0,00002	0,03970	0,00074	0,00002	0,04120	0,99990	0,99991	0,04120
20,57660	0,04870	0,04741	0,00129	0,00000	0,04612	0,00129	0,00000	0,04870	0,99990	0,99991	0,04870
17,14810	0,05775	0,05615	0,00155	0,00005	0,05460	0,00155	0,00005	0,05775	0,99990	0,99991	0,05775
14,29080	0,07007	0,06767	0,00231	0,00009	0,06536	0,00231	0,00009	0,07007	0,99990	0,99991	0,07007
11,90970	0,08337	0,08003	0,00321	0,00013	0,07682	0,00321	0,00013	0,08337	0,99990	0,99991	0,08337
9,92526	0,09945	0,09498	0,00436	0,00011	0,09062	0,00436	0,00011	0,09945	0,99990	0,99991	0,09945
8,27150	0,12193	0,11556	0,00614	0,00023	0,10942	0,00614	0,00023	0,12193	0,99990	0,99991	0,12193
6,89329	0,14544	0,13512	0,00979	0,00053	0,12533	0,00979	0,00053	0,14544	0,99990	0,99991	0,14544
5,74472	0,16997	0,15616	0,01297	0,00084	0,14318	0,01297	0,00084	0,16997	0,99990	0,99991	0,16997
4,78752	0,20975	0,18914	0,01909	0,00152	0,17004	0,01909	0,00152	0,20975	0,99990	0,99991	0,20975
3,98982	0,25075	0,22159	0,02687	0,00229	0,19471	0,02687	0,00229	0,25075	0,99990	0,99991	0,25075
3,32503	0,30166	0,26033	0,03748	0,00385	0,22285	0,03748	0,00385	0,30166	0,99990	0,99991	0,30166
2,77101	0,35553	0,29948	0,05008	0,00597	0,24940	0,05008	0,00597	0,35553	0,99990	0,99991	0,35553
2,30930	0,43491	0,35245	0,07180	0,01066	0,28065	0,07180	0,01066	0,43491	0,99990	0,99991	0,43491
1,92452	0,51888	0,40606	0,09528	0,01754	0,31077	0,09528	0,01754	0,51888	0,99990	0,99991	0,51888
1,60385	0,62549	0,46543	0,12995	0,03011	0,33547	0,12995	0,03011	0,62549	0,99990	0,99991	0,62549
1,33662	0,74990	0,52590	0,17466	0,04934	0,35124	0,17466	0,04934	0,74990	0,99990	0,99991	0,74990
1,11391	0,89267	0,59003	0,22581	0,07683	0,36421	0,22581	0,07683	0,89267	0,99990	0,99991	0,89267
0,92831	1,07026	0,65688	0,29112	0,12226	0,36576	0,29112	0,12226	1,07026	0,99990	0,99991	1,07026
0,77363	1,28801	0,72364	0,36935	0,19502	0,35429	0,36935	0,19502	1,28801	0,99990	0,99991	1,28801
0,64473	1,54051	0,78538	0,45748	0,29765	0,32790	0,45747	0,29765	1,54051	0,99990	0,99991	1,54051
0,53730	1,86105	0,84473	0,55525	0,46106	0,28948	0,55525	0,46106	1,86104	0,99989	0,99990	1,86104
0,44778	2,24086	0,89311	0,65591	0,69182	0,23720	0,65590	0,69182	2,24084	0,99989	0,99990	2,24084
0,37317	2,67868	0,93025	0,74605	1,00235	0,18419	0,74605	1,00235	2,67865	0,99989	0,99990	2,67865
0,31099	3,21146	0,95876	0,82914	1,42351	0,12962	0,82913	1,42351	3,21141	0,99989	0,99990	3,21141
0,25917	3,85680	0,97830	0,89753	1,98097	0,08077	0,89752	1,98097	3,85680	0,99989	0,99990	3,85680
0,21599	4,62533	0,99019	0,94348	2,69158	0,04671	0,94347	2,69158	4,62525	0,99989	0,99990	4,62525
0,18000	5,57317	0,99642	0,97529	3,60142	0,02113	0,97528	3,60142	5,57313	0,99989	0,99990	5,57313

A.3 Tabla de CSMA no persistente exportada desde el simulador STPNPlay

	Tráfico Ofrecido G	Disparo de T2	Disparo de T3		Rendimiento S	Disparo de T6					
700,0000	0,00145	0,00144	0,00000	0,00000	0,00144	0,00000	0,00000	0,00144	0,00144	0,00144	0,00001
539,45300	0,00182	0,00182	0,00000	0,00000	0,00182	0,00000	0,00000	0,00182	0,00182	0,00182	0,00000
415,72800	0,00241	0,00240	0,00000	0,00000	0,00240	0,00000	0,00000	0,00240	0,00240	0,00240	0,00001
320,38000	0,00310	0,00310	0,00000	0,00000	0,00310	0,00000	0,00000	0,00310	0,00310	0,00310	0,00000
246,90000	0,00414	0,00412	0,00000	0,00000	0,00412	0,00000	0,00000	0,00412	0,00412	0,00412	0,00002
190,27300	0,00528	0,00525	0,00000	0,00000	0,00525	0,00000	0,00000	0,00525	0,00525	0,00525	0,00003
146,63300	0,00676	0,00672	0,00000	0,00000	0,00672	0,00000	0,00000	0,00672	0,00672	0,00672	0,00004
113,00300	0,00894	0,00887	0,00000	0,00000	0,00887	0,00000	0,00000	0,00887	0,00887	0,00887	0,00008
87,08510	0,01150	0,01136	0,00000	0,00000	0,01136	0,00000	0,00000	0,01136	0,01136	0,01136	0,00013
67,11190	0,01490	0,01468	0,00000	0,00000	0,01468	0,00000	0,00000	0,01468	0,01468	0,01469	0,00022
51,71970	0,01923	0,01887	0,00000	0,00000	0,01887	0,00000	0,00000	0,01887	0,01887	0,01888	0,00035
39,85760	0,02481	0,02424	0,00001	0,00000	0,02423	0,00001	0,00000	0,02424	0,02424	0,02425	0,00057
30,71620	0,03257	0,03159	0,00001	0,00000	0,03158	0,00001	0,00000	0,03159	0,03159	0,03159	0,00098
23,67130	0,04227	0,04050	0,00002	0,00000	0,04048	0,00002	0,00000	0,04050	0,04050	0,04051	0,00176
18,24230	0,05478	0,05195	0,00003	0,00000	0,05192	0,00003	0,00000	0,05195	0,05195	0,05198	0,00281
14,05840	0,07108	0,06619	0,00005	0,00000	0,06614	0,00005	0,00000	0,06619	0,06619	0,06624	0,00484
10,83400	0,09209	0,08418	0,00007	0,00000	0,08411	0,00007	0,00000	0,08418	0,08418	0,08425	0,00783
8,34922	0,11940	0,10659	0,00012	0,00000	0,10647	0,00012	0,00000	0,10659	0,10659	0,10671	0,01269
6,43430	0,15547	0,13443	0,00019	0,00000	0,13423	0,00019	0,00000	0,13443	0,13443	0,13462	0,02085
4,95858	0,20152	0,16725	0,00035	0,00000	0,16691	0,00035	0,00000	0,16725	0,16725	0,16760	0,03391
3,82132	0,26132	0,20695	0,00052	0,00000	0,20644	0,00052	0,00000	0,20695	0,20695	0,20748	0,05384
2,94489	0,33775	0,25157	0,00084	0,00000	0,25073	0,00084	0,00000	0,25157	0,25157	0,25241	0,08534
2,26947	0,43984	0,30451	0,00127	0,00000	0,30324	0,00127	0,00000	0,30451	0,30451	0,30578	0,13406
1,74896	0,57118	0,36244	0,00200	0,00000	0,36044	0,00200	0,00000	0,36244	0,36244	0,36444	0,20674
1,34783	0,74369	0,42434	0,00315	0,00001	0,42119	0,00315	0,00001	0,42434	0,42434	0,42750	0,31619
1,03870	0,96335	0,48797	0,00467	0,00002	0,48330	0,00467	0,00002	0,48797	0,48797	0,49266	0,47069
0,80048	1,24896	0,55230	0,00681	0,00005	0,54549	0,00681	0,00005	0,55230	0,55230	0,55915	0,68981
0,61688	1,62078	0,61510	0,00981	0,00010	0,60529	0,00981	0,00010	0,61510	0,61510	0,62501	0,99577
0,47540	2,10516	0,67334	0,01391	0,00014	0,65943	0,01391	0,00014	0,67334	0,67334	0,68740	1,41777
0,36637	2,73117	0,72643	0,01984	0,00027	0,70660	0,01984	0,00027	0,72643	0,72643	0,74654	1,98463
0,28234	3,54466	0,77338	0,02678	0,00047	0,74659	0,02678	0,00047	0,77338	0,77338	0,80063	2,74402
0,21758	4,59486	0,81439	0,03652	0,00078	0,77787	0,03652	0,00078	0,81439	0,81439	0,85169	3,74318
0,16768	5,96670	0,84886	0,04895	0,00139	0,79991	0,04895	0,00139	0,84886	0,84886	0,89920	5,06750
0,12922	7,74188	0,87741	0,06504	0,00262	0,81237	0,06504	0,00262	0,87741	0,87741	0,94506	6,79682
0,09958	10,04130	0,90089	0,08636	0,00442	0,81454	0,08636	0,00442	0,90089	0,90089	0,99166	9,04964
0,07674	13,03350	0,91965	0,11231	0,00754	0,80734	0,11231	0,00754	0,91965	0,91965	1,03950	11,99400
0,05914	16,90870	0,93466	0,14570	0,01276	0,78896	0,14570	0,01276	0,93466	0,93466	1,09312	15,81560
0,04558	21,94220	0,94647	0,18737	0,02137	0,75910	0,18737	0,02137	0,94647	0,94647	1,15521	20,78700
0,03513	28,47090	0,95563	0,23660	0,03551	0,71903	0,23660	0,03551	0,95563	0,95563	1,22774	27,24310
0,02707	36,94080	0,96277	0,29691	0,05779	0,66586	0,29691	0,05779	0,96277	0,96277	1,31747	35,62340
0,02086	47,93580	0,96809	0,36896	0,09547	0,59913	0,36896	0,09547	0,96809	0,96809	1,43253	46,50330
0,01608	62,21770	0,97217	0,45013	0,15373	0,52205	0,45013	0,15373	0,97217	0,97217	1,57602	60,64160
0,01239	80,72440	0,97511	0,53936	0,24668	0,43575	0,53936	0,24668	0,97511	0,97511	1,76115	78,96320
0,00955	104,74400	0,97715	0,63484	0,38840	0,34232	0,63484	0,38840	0,97715	0,97715	2,00039	102,74300
0,00736	135,90800	0,97857	0,72703	0,60383	0,25154	0,72703	0,60383	0,97857	0,97857	2,30943	133,59900
0,00567	176,33800	0,97945	0,81099	0,91483	0,16845	0,81099	0,91483	0,97945	0,97945	2,70527	173,63200
0,00437	228,83600	0,97996	0,88089	1,36138	0,09908	0,88089	1,36138	0,97996	0,97996	3,22223	225,61400
0,00337	296,94300	0,98022	0,92969	1,98015	0,05054	0,92969	1,98015	0,98022	0,98022	3,89006	293,05300
0,00260	385,34200	0,98033	0,95940	2,81799	0,02093	0,95940	2,81799	0,98033	0,98033	4,75772	380,58500
0,00200	499,98600	0,98036	0,97369	3,93040	0,00667	0,97369	3,93040	0,98036	0,98036	5,88445	494,10100

A.4 Tabla de CSMA persistente exportada desde el simulador STPNPlay

	Tráfico Ofrecido G	Disparo de T2	Disparo de T3	Disparo de T4	Rendimiento S	Disparo de T6				
700,00000	0,00134	0,00134	0,00000	0,00000	0,00134	0,00000	0,00000	0,00134	0,00134	0,00134
539,45300	0,00198	0,00198	0,00000	0,00000	0,00198	0,00000	0,00000	0,00198	0,00198	0,00198
415,72800	0,00244	0,00244	0,00000	0,00000	0,00244	0,00000	0,00000	0,00244	0,00244	0,00244
320,38000	0,00328	0,00328	0,00000	0,00000	0,00328	0,00000	0,00000	0,00328	0,00328	0,00328
246,90000	0,00401	0,00401	0,00000	0,00000	0,00401	0,00000	0,00000	0,00401	0,00401	0,00401
190,27300	0,00506	0,00506	0,00000	0,00000	0,00506	0,00000	0,00000	0,00506	0,00506	0,00506
146,63300	0,00681	0,00681	0,00000	0,00000	0,00681	0,00000	0,00000	0,00681	0,00681	0,00681
113,00300	0,00869	0,00869	0,00000	0,00000	0,00869	0,00000	0,00000	0,00869	0,00869	0,00869
87,08510	0,01128	0,01128	0,00000	0,00000	0,01128	0,00000	0,00000	0,01128	0,01128	0,01128
67,11190	0,01518	0,01518	0,00000	0,00000	0,01518	0,00000	0,00000	0,01518	0,01518	0,01518
51,71970	0,01954	0,01954	0,00000	0,00000	0,01954	0,00000	0,00000	0,01954	0,01954	0,01954
39,85760	0,02508	0,02508	0,00000	0,00000	0,02508	0,00000	0,00000	0,02508	0,02508	0,02508
30,71620	0,03310	0,03307	0,00003	0,00000	0,03303	0,00003	0,00000	0,03307	0,03306	0,03310
23,67130	0,04255	0,04251	0,00004	0,00000	0,04247	0,00004	0,00000	0,04251	0,04251	0,04255
18,24230	0,05359	0,05347	0,00011	0,00001	0,05336	0,00011	0,00001	0,05347	0,05347	0,05359
14,05840	0,07272	0,07243	0,00029	0,00000	0,07214	0,00029	0,00000	0,07243	0,07243	0,07272
10,83400	0,09219	0,09186	0,00032	0,00001	0,09154	0,00032	0,00001	0,09186	0,09186	0,09219
8,34922	0,12020	0,11915	0,00103	0,00002	0,11812	0,00103	0,00002	0,11915	0,11915	0,12020
6,43430	0,15508	0,15308	0,00191	0,00009	0,15116	0,00191	0,00009	0,15308	0,15307	0,15508
4,95858	0,19962	0,19551	0,00383	0,00028	0,19168	0,00383	0,00028	0,19551	0,19551	0,19962
3,82132	0,26029	0,25197	0,00773	0,00059	0,24424	0,00773	0,00059	0,25197	0,25197	0,26029
2,94489	0,33954	0,32270	0,01520	0,00164	0,30750	0,01520	0,00164	0,32270	0,32270	0,33954
2,26947	0,44106	0,40509	0,03077	0,00520	0,37431	0,03077	0,00520	0,40509	0,40508	0,44106
1,74896	0,56953	0,49911	0,05839	0,01203	0,44071	0,05839	0,01203	0,49911	0,49910	0,56953
1,34783	0,74330	0,60540	0,10789	0,03001	0,49750	0,10789	0,03001	0,60540	0,60539	0,74330
1,03870	0,95746	0,70844	0,18103	0,06799	0,52740	0,18103	0,06799	0,70844	0,70843	0,95746
0,80048	1,24972	0,80777	0,29332	0,14862	0,51445	0,29331	0,14862	0,80777	0,80776	1,24971
0,61688	1,62226	0,88350	0,43307	0,30569	0,45042	0,43307	0,30569	0,88350	0,88349	1,62226
0,47540	2,10122	0,93662	0,58917	0,57542	0,34744	0,58917	0,57542	0,93662	0,93661	2,10121
0,36637	2,72397	0,96722	0,73735	1,01939	0,22987	0,73734	1,01939	0,96722	0,96721	2,72396
0,28234	3,52262	0,98222	0,85600	1,68440	0,12622	0,85599	1,68440	0,98222	0,98221	3,52262
0,21758	4,58392	0,98779	0,93531	2,66081	0,05248	0,93530	2,66081	0,98779	0,98778	4,58391
0,16768	5,96556	0,98933	0,97258	4,00363	0,01675	0,97257	4,00363	0,98933	0,98932	5,96554
0,12922	7,75408	0,98961	0,98597	5,77849	0,00364	0,98596	5,77849	0,98961	0,98960	7,75407
0,09958	10,04570	0,98952	0,98906	8,06699	0,00046	0,98905	8,06699	0,98952	0,98951	10,04560
0,07674	13,03190	0,98939	0,98937	11,05310	0,00002	0,98936	11,05310	0,98939	0,98938	13,03190
0,05914	16,90400	0,98919	0,98918	14,92560	0,00001	0,98917	14,92560	0,98919	0,98918	16,90400
0,04558	21,96760	0,98893	0,98893	19,98960	0,00000	0,98892	19,98960	0,98893	0,98892	21,96750
0,03513	28,48910	0,98862	0,98861	26,51170	0,00001	0,98860	26,51170	0,98862	0,98861	28,48890
0,02707	36,96310	0,98824	0,98823	34,98640	0,00001	0,98822	34,98640	0,98824	0,98823	36,96280
0,02086	47,91470	0,98782	0,98782	45,93900	0,00000	0,98781	45,93900	0,98782	0,98781	47,91460
0,01608	62,25140	0,98726	0,98726	60,27660	0,00000	0,98725	60,27660	0,98726	0,98725	62,25120
0,01239	80,74380	0,98668	0,98668	78,77000	0,00000	0,98667	78,77000	0,98668	0,98667	80,74340
0,00955	104,73700	0,98602	0,98602	102,76500	0,00000	0,98601	102,76500	0,98602	0,98601	104,73700
0,00736	135,96600	0,98530	0,98530	133,99500	0,00000	0,98529	133,99500	0,98530	0,98529	135,96500
0,00567	176,32800	0,98457	0,98457	174,35800	0,00000	0,98456	174,35800	0,98457	0,98456	176,32700
0,00437	228,86200	0,98388	0,98387	226,89200	0,00001	0,98386	226,89200	0,98388	0,98387	228,86000
0,00337	296,95200	0,98324	0,98323	294,98500	0,00001	0,98322	294,98500	0,98324	0,98323	296,95100
0,00260	385,27600	0,98265	0,98265	383,30900	0,00000	0,98264	383,30900	0,98265	0,98264	385,27400
0,00200	500,03300	0,98215	0,98215	498,06600	0,00000	0,98214	498,06600	0,98215	0,98214	500,03000

APENDICE B

ESPECIFICACION DE SERVIDORES, CONFLICTOS Y POLITICAS

B.1 Resolución de conflictos, y Conflictos y Precedencia

Resolución de conflictos: Carrera vs Preselección. Consideremos el sistema de red con prioridad de la figura n° B1. Ella contiene tres tipos diferentes de situaciones conflictivas:

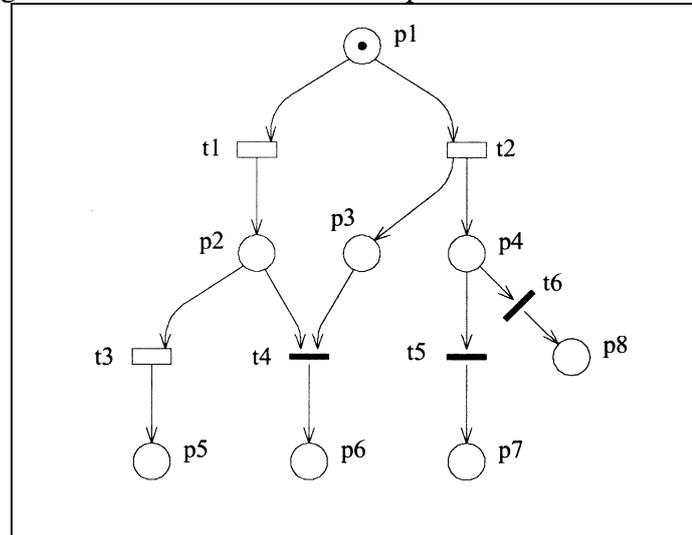


figura n° B1 Algunos tipos de conflictos en una sistema de red con prioridad

- conflicto de elección libre entre las transiciones temporizadas t1 y t2;
- conflicto de elección no libre entre la transición temporizada t3 y la transición inmediata t4;
- conflicto de elección libre entre las transiciones inmediatas t5 y t6.

En el caso a, en principio podemos imaginar dos diferentes formas de resolver el conflicto, ambas perfectamente razonables desde el punto de vista de la modelación: o bien tomar en cuenta la definición de tiempo (política de competencia), o asumir una elección independiente del tiempo (política de preselección). De acuerdo a la política de competencia, cada transición decide independientemente sus propios retardos de disparo.; las dos habilitaciones comienzan en el mismo tiempo, de modo que la transición que está asociada con el retardo de disparo más corto “gana la competencia” y se dispara; la otra transición “pierde la competencia”, y su habilitación es malograda por el conflicto. De acuerdo a la política de preselección, en cambio, primero un acuerdo es alcanzado (idealmente en tiempo cero) entre las transiciones conflictivas, acordando un protocolo o algoritmo que no tiene necesariamente que tomar en cuenta sus retardos de disparo; entonces, la transición que es seleccionada para disparar “declara” su retardo de disparo, y espera esa cantidad de tiempo antes de disparar; por el contrario, la otra transición (que perdió la resolución de conflicto) se comporta como si ella no estuviera habilitada.

En el caso b, primero de todo que no tiene sentido considerar elección libre en el caso de conflictos entre transiciones con diferente nivel de prioridad (de otro modo las transiciones de prioridad más baja estaría muerta). Por lo tanto, si queremos dar a la transición temporizada una chance de disparo, tenemos que considerar los casos en el cual t3 está

habilitado y t4 no. La única política de resolución de conflicto significativa es entonces la política de competencia en este caso, donde t3 comienza contando su retardo de disparo, y “gana la carrera” si y solo si t4 no se vuelve habilitada (debido al disparo de t2 en nuestro ejemplo) antes de que el retardo de disparo de t3 termine.

En el caso c, solo un política de preselección tiene sentido, dado que el retardo de disparo para ambas transiciones es idénticamente igual a 0 (de modo que les política de competencia debería siempre resultar en un empate, y no podría decidir cual va a dispararse).

Podemos así resaltar que las políticas de competencia y preselección tienen sentido, aún mezcladas en el mismo modelo para resolver conflictos diferentes. La política de preselección refuerza una separación clara entre la resolución del conflicto y la temporización de las transiciones, de modo que en algún sentido está teóricamente claramente delimitado. Por el contrario, la política de competencia se necesita para modelar un conjunto de fenómenos interesantes que son usualmente encontrados en sistemas distribuidos (condiciones de competencia, timeouts, etc.), de modo que su introducción en los modelos de Redes de Petri temporizadas es invaluable.

Finalmente, se observa que la preselección entre transiciones temporizadas pueden ser simuladas por un modelo estructuralmente diferente, en el que el caso a está reemplazado por el caso c, y las transiciones inmediatas son seguidas por transiciones temporizadas persistentes. Por ello, en un modelo mezclado conteniendo transiciones temporizadas e inmediatas, uno podría siempre asumir la política de competencia entre las transiciones conflictivas sin mayor pérdida de potencia de modelación con respecto al caso en el cual las políticas de competencia y preselección están permitidas para las transiciones temporizadas conflictivas.

Conflictos y precedencia: políticas de memoria. En el caso de transiciones temporizadas no persistentes, la elección de la política de la competencia no es suficiente para determinar la conducta del modelo. Tenemos aún que decidir que se va a hacer con los tiempos de temporización de las transiciones que pierden la carrera. Notamos que esta distinción no tiene sentido en el caso de la política de preselección, dado que solo la instancia de transición seleccionada realmente comienza, de modo que el problema de las instancias interrumpidas no aparece. Al menos dos alternativas principales pueden ser identificadas que ven a una consistente conducta temporizada con la conducta cualitativa del modelo de red subyacente no interpretado para las transiciones temporizadas conflictivas con política de competencia:

- Memoria de paso del tiempo: una instancia de transición que pierde la competencia, que se vuelve deshabilitada antes de su disparo, es considerada estar congelada para la “foto final” de la competencia. Tan pronto como la misma transición se habilita nuevamente, la instancia congelada es reasumida y continúa su actividad hasta el disparo, o para ser congelada nuevamente por una instancia de transición conflictiva más rápida.

- Memoria de habilitación: una instancia de transición que pierde la competencia, es considerada perdida. Cuando la misma transición es habilitada nuevamente, la instancia interrumpida previamente es olvidada, y una nueva instancia de transición comienza su actividad desde el principio.

Las dos alternativas ven conductas completamente diferentes desde el punto de vista de la modelación: la primera permite una representación trivial de la disciplina de planificación “round-robin” y hace imposible representar un mecanismo de time-out es una forma natural.: la última tiene características opuestas en estos aspectos. Notamos que estas políticas son ortogonales con respecto a la forma en que el no determinismo es reducido.

Ambas políticas de memoria deben ser ulteriormente especificadas en el caso de transiciones múltiple server: cuando más que una instancia de transición está habilitada, y el grado de habilitación de la transición es reducido a un valor menor mayor que cero, un algoritmo

debe ser especificado para seleccionar que instancia de transición tiene que continuar y cual tiene que ser congelada o descartada. Diferentes disciplinas de deshabilitación podrían ser especificadas, exactamente como en el caso de las disciplinas de cola para modelos de colas, tales como: FIFO, LIFO, elección aleatoria, valor más grande, valor más pequeño, etc. En el caso de transición de memoria de paso de tiempo múltiple server, también una política de replanificación debe ser especificada es una forma similar para decidir cual entre las instancias congeladas ha de reasumir cuando el grado de habilitación de transición se incrementa de un valor menor que el número total de instancias congeladas corrientemente.

Por otro lado, en el caso que el tiempo de disparo está sujeto a variabilidad (o bien estocásticas o no determinística) al menos una alternativa “intermedia” podría ser considerada:

- Memoria de reset: el trabajo realizado por una instancia de transición que pierde la competencia es considerada pérdida. Sin embargo, un seguimiento se mantiene del trabajo que fue descartado. Cuando la misma transición es habilitada nuevamente, la actividad interrumpida previamente debe ser recomenzada desde el comienzo, de modo que el mismo valor será usado como el tiempo de disparo de transición. Esta política es algunas veces llamada PRI (preempt-repeat-identical) en las colas estocásticas, como opuesto a PRD (preempt-repeat-different) que es nuestro concepto de memoria de habilitación.

Las diferencias entre memoria de habilitación y de reset son tenues: si el tiempo de disparo está sujeto a variabilidad (o bien estocástica o no determinística), la política de memoria de habilitación no puede garantizar que una instancia de transición recomience desde cero la misma actividad que fue interrumpida por deshabilitar una instancia de habilitación previa de la misma transición, mientras la política de memoria de reset garantiza esta condición. La memoria de reset y de habilitación colapsan a la misma política cuando el tiempo de disparo de la transición es determinística.

APENDICE C

CASOS DE MARKOV Y SEMI-MARKOV

C.1 Casos de Markov y semi-Markov

Hay algunos casos particulares para los cuales los asuntos de memoria, las políticas de desplanificación y replanificación se simplifican drásticamente con respecto al caso general. Este es el caso llamado “sin memoria”, en el cual todas las elecciones diferentes ven la misma conducta debido a una propiedad particular de la temporización estocástica. Por ejemplo, esta propiedad particular se podría referir a la suposición de tiempos de disparo distribuidos exponencialmente negativos, independientes y aleatorios para todas las instancias de disparo de transición.

Simplificaciones de la política de competencia. Debido a la propiedad sin memoria de la distribución exponencial, algunas simplificaciones pueden ser introducidas en la especificación de un sistema de red markoviana temporizada con respecto al caso general. Por ejemplo, en caso de competencia entre transiciones temporizadas de memoria de paso del tiempo y de habilitación, se especifica la misma conducta. Este no es el caso, sin embargo, para la política de memoria de reset, para la cual la distribución de un tiempo de disparo reasumido no tiene la misma distribución como una nueva instancia de tiempo de disparo (dado que ya conocemos que la primera forma tiene una duración mayor que el intervalo de tiempo más grande para la cual la instancia de transición estaba habilitada sin disparo). También la elección de disciplinas de desplanificación y replanificación en el caso de competencia con políticas de memoria de paso del tiempo o competencia con habilitación, se vuelve irrelevante bajo la hipótesis de tiempo de disparo de transición markoviana.

Cargas de trabajo sin memoria. En algunos casos, la suposición sin memoria para algunos de los parámetros del modelo pueden inducir simplificaciones sustanciales no solo en el análisis de los modelos de red temporizados, sino también en la estructura de la representación del sistema de red subyacente.

Dependencia del mercado. La propiedad sin memoria puede sugerir una forma inteligente de abstraer algunos de los detalles de una representación del modelo del sistema, mientras se mantiene un modelo de performance exacto. Esta es la idea de servidor equivalente tomada de los modelos de cola. Si una cola de clase simple ordena la propiedad sin memoria, entonces toda la información acerca de su estado presente puede ser sumariada en un simple número que cuenta el número total de clientes internos. Una caracterización detallada de su conducta puede ser evitada dado que para cada valor que el parámetro de carga puede asumir, especificamos la frecuencia de servicio esperada, por ejemplo, la velocidad de partida del cliente primero. La partida de un cliente cambiará el parámetro de carga, por lo tanto cambiando también el ritmo de partida del cliente siguiente. También el arribo de un nuevo cliente cambiará el parámetro de carga induciendo la selección de una frecuencia de partida diferente.

Una forma de explotar esta característica de colas sin memoria es realizar análisis jerárquico, por computar primero las frecuencias de partida dependientes de la carga basados con un modelo detallado de una simple cola compleja. Entonces la cola es sustituida por su “servidor equivalente” en un modelo más abstracto, que tiene en cuenta la interacción de la cola cuya conducta está siendo abstraída con otras colas (o aún los servidores equivalentes de otras colas).

En la estructura de una Red de Petri, si la propiedad sin memoria mantiene el estado completo del modelo, puede ser sumariado al especificar el mercado en cada lugar. Siguiendo el mismo ejemplo, se puede así sustituir la representación explícita detallada de un

submodelo sin memoria con retardos de disparo de transiciones distribuidas exponencialmente, que son una función del mercado corriente.

En algunos casos, sin embargo, la adopción de frecuencias de disparo dependientes del mercado para las transiciones en una Red de Petri markoviana puede inducir reducción substancial en la complejidad del modelo, con respecto a una representación detallada de los mecanismos del sistema a ser modelado. Usualmente tal explotación inteligente de las frecuencias dependientes del mercado dependen del conocimiento del modelador de alguna regularidad y/o simetría intrínseca en la conducta del modelo. Tal característica, útil en algunos casos particulares, debería explotarse con gran cuidado y autocontrol por usuarios experimentados del formalismo. En general, las frecuencias de disparo dependientes del mercado para las transiciones crea dependencias entre eventos que no estarían relacionados de acuerdo al análisis de la estructura del sistema de red subyacente, así impidiendo explotar uno de las características fundamentales de Redes de Petri temporizadas bien definidas, llamada la ortogonalidad de la interpretación temporizada con respecto a las propiedades causales del modelo de Red de Petri.

APENDICE D

REDES DE PETRI ESTOCASTICAS EXPONENCIALES (ESPN EXPONENTIAL STOCHASTIC PETRI NETS) [17][18]

D.1 Introducción

Los sistemas de Redes de Petri Estocásticas pueden ser convertidos en cadenas de Markov y como sus análisis pueden ser realizados para computar interesantes índices de performance. La construcción de la cadena de Markov asociada con un sistema de Red de Petri Estocástica es la base para la derivación subsecuente del modelo probabilístico asociado con una Red de Petri Estocásticas Generalizada. Sólo las Redes de Petri Estocásticas con espacio de estado finito son consideradas en este sentido, por la relación con las cadenas de Markov, que son resolubles con técnicas numéricas estándar. Métodos de solución más avanzados se estructuran sobre la teoría de matriz geométrica. Finalmente, pueden adicionarse técnicas de reducción que pueden ser empleadas para obtener grafos de alcanzabilidad más pequeños.

D.2 El Proceso Estocástico Asociado con una Red de Petri Estocástica

Debido a la propiedad sin memoria de la distribución exponencial de los retardos de disparo, es relativamente fácil mostrar que los sistemas de Redes de Petri Estocásticas son isomórficas para cadenas de Markov de tiempo continuo (CTMC). En particular, un sistema de Red de Petri k-limitado puede mostrarse como isomórfico en una CTMC finita.

D.2.1 Máquina de Estado Finito y Grafo Marcado de las Redes de Petri Estocástica

Esto puede ser fácilmente visto cuando la estructura de la Red de Petri Estocástica es una máquina de estado finito (ninguna transición tiene más que un lugar de entrada y uno de salida) y de un grafo marcado (ningún lugar tiene más que una transición de entrada y una de salida) con solo una marca en su marcado inicial. En este caso, cada lugar de la red corresponde unívocamente a un estado de la red y la posición de la marca en un instante dado de tiempo, identifica el estado que la red tiene en ese mismo momento. Cada lugar de la red mapea en un estado de la correspondiente CTMC y cada transición mapea en un arco anotado por la frecuencia de la distribución de tiempo de disparo correspondiente. Por otra parte, si los tiempos de disparo de las transiciones tienen distribuciones exponenciales negativas y si la estructura de la red es de un grafo marcado, el tiempo gastado en cada lugar por la red está completamente identificada por las características de la única transición que puede retirar la marca desde ese lugar. Cuando la red tiene la estructura de una máquina de estado finito, los conflictos entre transiciones habilitadas simultáneamente aparecen, dado que algunas transiciones pueden compartir el mismo lugar de entrada. En esta situación, algunas posibilidades existen para la selección de la transición a dispararse luego. Una política de preselección puede ser aplicada en la que una de las transiciones es inicialmente seleccionada, y luego una cantidad de tiempo exponencialmente distribuida transcurre, antes de que la transición realmente se dispare. Si esta es la forma en la que la red opera, es necesario una especificación probabilística del mecanismo de preselección, para construir el modelo probabilístico correspondiente. Otra posibilidad es asumir que una competencia toma lugar entre las transiciones habilitadas, y que la transición asociada con la actividad que finaliza primero es seleccionada para el disparo. Como la

competencia es ganada por una de las transiciones, las otras son interrumpidas, y una elección debe hacerse sobre como tratar estas actividades parcialmente completadas. Aunque algunas posibilidades existen, la forma más simple es descartar el trabajo asociado con las transiciones interrumpidas, de modo que ellas comiencen nuevamente la próxima vez que ellas sean habilitadas. Si todas las actividades tienen duraciones distribuidas exponencialmente negativas, la CTMC correspondiente a la Red de Petri se obtiene desde la red de una manera directa. Nuevamente, cada lugar de la Red de Petri mapea en un estado de la CTMC correspondiente, y cada transición de la Red de Petri mapea en un arco de la CTMC anotada con la frecuencia de la distribución de tiempo de disparo correspondiente. También en este caso, el tiempo gastado por la red en cada lugar tiene una distribución exponencial negativa, pero su frecuencia está dada por la suma de sus frecuencias de disparo de todas las transiciones que retiraran las marcas desde ese lugar.

Situaciones más complejas aparecen cuando algunas marcas se permiten en el marcado inicial de las Redes de Petri, de los tipos mencionados previamente. Estas son debido al hecho que los lugares no corresponden a estados de los modelos probabilísticos asociados. Por lo tanto, en estos casos las conductas de los modelos probabilísticos son también afectados por las políticas de selección de servicio desde los lugares de entrada, como así también por las políticas de selección de marcas adoptadas en los momentos de disparo de transición.

Aun el caso muy simple de dos marcas esperando en el único de lugar de entrada de una transición genera un conjunto de cuestiones interesantes que deben ser dirigidas en desarrollar el modelo probabilístico correspondiente. La primera tiene que ver con la velocidad en la cual la transición deja las marcas desde su lugar de entrada en esta situación. Existen algunas posibilidades. Una es que de considerar el lugar de entrada como una sala esperando: la transición habilitada “sirve” solo una marca en el tiempo, y su velocidad de servicio no es afectada por el número de marca “esperando” en la sala de espera. Otra es que de asumir un “servicio múltiple” en la cual cada marca comienza siendo servida tan pronto ella arriba al lugar de entrada. Esto corresponde a algunas instancias que están simultáneamente activas: tantas como marcas hay en el lugar de entrada. Finalmente, una última es que de asumir una forma de dependencia de carga en la cual la frecuencia de disparo de la transición es una función de su grado de habilitación. En este caso, una especificación adicional debe ser introducida en el modelo para definir la función de dependencia de carga asociada con cada transición. La segunda cuestión se refiere a la selección de la marca que es removida desde el lugar de entrada sobre el disparo de la transición. Desde un punto de vista de la Red de Petri “clásica”, esta política de selección es innecesaria dado que las marcas no llevan ninguna identidad (excepto en el caso de redes coloreadas) y ninguna marca es “movida” a través de la red durante su operación. En varias aplicaciones, sin embargo, es conveniente asociar un significado físico con las marcas (por ejemplo, clientes), de modo que las preguntas sobre su flujo a través de la red puedan ser respondidas. En estas situaciones, cuando algunas marcas están simultáneamente presentes en el lugar de entrada de una transición, si este se asume opera con una política de simple server, un interrogante sobre la política de cola aplicada a estas marcas se vuelve interesante. La política más natural (desde el punto de vista de las Redes de Petri) que se asume en estas situaciones es una de orden aleatorio. Cuando los tiempos de disparo están distribuidos exponencialmente, y cuando los número de performance de interés están solo relacionados a los momentos del número de marcas en el lugar de entrada de una transición, es posible mostrar que varias políticas de colas ven los mismos resultados (por ejemplo, random, FIFO, LCFS). Sin embargo, debe observarse que en otros casos, la elección de la política de selección de marcas debe ser importante y que políticas diferentes desde la aleatoria, deben estar explícitamente implementadas a través de construcciones de red apropiadas.

D.2.2 Redes de Petri Estocástica con estructura general

En general, la CTMC asociada con un sistema de Red de Petri Estocástica dada se obtiene aplicando las siguientes simples reglas:

1. El espacio de estado CTMC $S = \{s_i\}$ corresponde al conjunto de alcanzabilidad $RS(m_0)$ de la Red de Petri asociada con la Red de Petri Estocástica ($m_i \leftrightarrow s_i$).
2. La frecuencia de transición desde el estado s_i (correspondiente al marcado m_i) al estado $s_j(m_j)$ se obtiene como la suma de las frecuencias de disparo de las transiciones que están habilitadas en m_i y cuyos disparos generan marcado m_j .

Basado en estas simples reglas, es posible deducir algoritmos para la construcción automática del generador infinitesimal (también llamado la matriz de frecuencia de estado de transición) de la CTMC isomórfica, comenzando desde la descripción de la Red de Petri Estocástica.

Una Red de Petri Estocástica (SPN) es una Red de Petri con una interpretación temporal que reduce el no-determinismo inherente en el modelo autónomo y le hace adecuado para definir y evaluar índices de prestaciones. La interpretación temporal debe incluir la asociación de una duración temporal a las actividades (transiciones) y también una política de resolución de conflictos. Con respecto a la duración de actividades, se suele asociar una variable aleatoria a cada transición, que modela su tiempo de servicio. Denotamos con $s[t]$ el tiempo medio de servicio de la transición t . Con respecto a la resolución de conflictos, se pueden asociar tasas de encaminamiento a cada subconjunto de transiciones en conflicto, de manera que quede definida una distribución discreta de probabilidad para elegir la transición que debe dispararse al presentarse el conflicto. Las tasas de encaminamiento pueden definirse explícitamente mediante una anotación asociada a cada transición inmediata (transición cuyo tiempo de servicio es nulo y sirve para modelar decisiones), o bien pueden calcularse para el caso de transiciones temporizadas (que tienen un tiempo de servicio no nulo) usando una política de competición (la primera transición que termina su actividad es la elegida para ser disparada).

Una red de Petri Markoviana (MPN) es un caso particular de Res de Petri Estocástica en el que los tiempos de servicio son variables aleatorias independientes y con distribución de probabilidad exponencial (con tasa $\lambda_i = 1/s[t_i]$ para cada transición t_i), y las tasas de encaminamiento se definen mediante la política de competición.

Las Redes de Petri Estocásticas Generalizadas (GSPN) son una extensión de las Redes de Petri Markoviana que incluyen, entre otros aspectos, transiciones inmediatas que pueden usarse para modelar conflictos con unas tasas de encaminamiento determinadas (separando así la política de resolución de conflictos de la duración de las actividades).

Si nos centramos en técnicas de análisis, se conocen algoritmos eficientes para el cálculo de la distribución del estado estacionario del modelo únicamente para algunas subclases muy particulares de RdPE que poseen la denominada solución con forma producto. Por tanto, en general, para obtener los índices exactos de prestaciones debe emplearse un método numérico para resolver una cadena de Markov en tiempo continuo (CMT) isomorfa a la red. Si bien es esta la técnica más utilizada y más fácilmente implementable en herramientas software, el problema de la explosión de estados (crecimiento asintótico exponencial del espacio de estados en relación con el número de lugares y transiciones de la red) puede limitar su utilidad

para el análisis de modelos grandes, tanto por el costo de almacenar la matriz generadora infinitesimal de la CMTC como por la complejidad temporal de los algoritmos de resolución.

Las técnicas dirigidas por la estructura de la red tratan de reducir la complejidad en espacio y en tiempo de los algoritmos de solución, mediante la utilización de información extraída de la estructura del modelo. Hay múltiples ejemplos de técnicas dirigidas por la estructura, como las de eliminación de transiciones inmediatas, las de utilización de simetrías en GSPN coloreadas, o las basadas en programación lineal para el cálculo de cotas de índices de rendimiento. La estructura de la red se ha utilizado también en técnicas aproximadas de divide y vencerás para limitar el problema de la explosión de estados. Igualmente, la estructura de la red resulta fundamental en las técnicas de álgebra tensorial que permiten expresar la matriz generadora infinitesimal de la CMTC en términos de matrices más pequeñas obtenidas para las componentes (subredes con espacio de estados más pequeño), y así calcular la distribución del estado estacionario sin calcular ni almacenar la matriz completa.

Para el caso de una MPN, los ciclos de actividad en la red se entrelazan y pueden describirse perfectamente con el grafo de alcanzabilidad de la red. Si en el grafo de alcanzabilidad se etiquetan los arcos con las tasas de las distribuciones exponenciales en lugar de con los nombres de las transiciones correspondientes, el grafo se denomina diagrama de transición de estados de la cadena de Markov en tiempo continuo (CMTC) asociada a la MPN

Si la MPN es limitada, todas las transiciones de la red operan con una semántica simple servidor, velocidad independientes del marcado (denotado como matriz Q) y tiene estado hogar, la CMTC asociada a ella es ergódica (el marcado inicial es estado hogar), en cuyo caso es posible calcular la distribución de probabilidad de estado estacionario de los marcados, resolviendo el siguiente sistema de ecuaciones lineales:

$$\pi \cdot Q = 0$$

$$\pi \cdot 1 = 1$$

en el que π es el vector de la distribución estacionaria (es decir, tiene tantos elementos como marcados alcanzables), 0 es un vector de igual tamaño que π y con todos los elementos nulos, 1 es un vector (de la misma dimensión) con todas sus componentes igual a 1, y la matriz Q , llamada generador infinitesimal de la CMTC, es la siguiente:

$$q_{ij} = \sum_{m_i} t_{k_{mi}} \lambda_k \quad \text{para } i \neq j$$

$$q_{ii} = - \sum_{j \neq i} q_{ij}$$

Es decir, el elemento q_{ij} de la matriz Q es la suma de las tasas de las transiciones que pueden llevar al modelo desde el estado m_i al estado m_j ("velocidad" de salto de m_i a m_j). A partir de la distribución en estado estacionario es posible el cálculo de los índices de rendimiento del modelo. Existe una aproximación general para la definición de índices de rendimiento basada en el concepto de función de ganancia. Las funciones de ganancia se definen sobre los marcados alcanzables, de manera que la ganancia media (calculada para la distribución estacionaria de marcado) constituya el índice de rendimiento deseado. Es decir, si se considera la función de ganancia $r(m)$, la ganancia media es la siguiente:

$$R = \sum_{m_i \in RS(m_0)} r(m_i) \pi_i$$

Por ejemplo, si el índice que se desea calcular es el marcado medio μ_i de un lugar p_i , (valor medio del número de marcas), que puede aportar información sobre la utilización media de un recurso, o el grado medio de ocupación de un almacén, basta con definir la ganancia r_{μ_i} como el número de marcas en ese lugar p_i :

$$r_{\mu_i}(m) = m[p_i]$$

Si el índice de interés es el throughput X_i de una transición t_i (de tasa λ_i), es decir, el número medio de disparos de la transición por unidad de tiempo (índice de productividad), basta con definir la ganancia r_{x_i} como:

$$r_{x_i}(m) = \begin{cases} \lambda_i, & \text{si } t_i \text{ está sensibilizada en } m \\ 0_i, & \text{en otro caso} \end{cases}$$

En el caso en que se decida separar la decisión de conflictos de la duración de las tareas, pueden utilizarse las transiciones inmediatas para el modelado del conflicto. En otras situaciones, es posible que existan conflictos entre varias transiciones inmediatas, tales conflictos se resuelven asociando tasas de encaminamiento a las transiciones (se realiza un sorteo entre las transiciones en conflicto para decidir cual es la que se dispara). Puede suceder que un diagrama no represente exactamente una CMTC, ya que en ese tipo de procesos estocásticos no se incluye la posibilidad de tasas infinitas. Es posible transformar ese diagrama eliminando los estados fugaces del modelo (estados en los que hay una tasa de salida de velocidad infinita), para obtener el diagrama de transición de estados reducido (figuras n° B1 y B2), en el que únicamente aparecen los estados tangibles del modelo, es decir, aquellos en los que no hay transiciones inmediatas sensibilizadas y, por tanto, en los que el sistema permanece un tiempo no nulo. La solución en estado estacionario de este último modelo se obtiene resolviendo el sistema lineal de ecuaciones (figura n° B3).

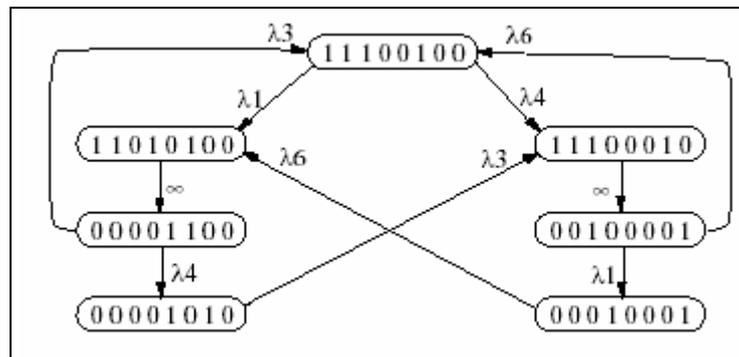


figura n° D1 Diagrama de transición de estados de una Red de Petri Generalizada

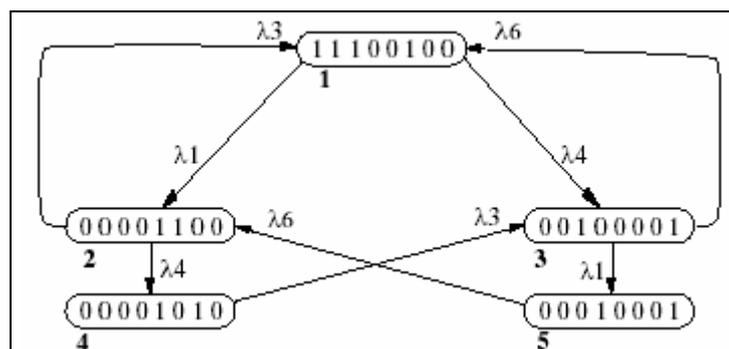


figura n° D2 Diagrama de transición es de estados reducido

$$\begin{array}{c} (\pi_1, \pi_2, \pi_3, \pi_4, \pi_5) \cdot \begin{bmatrix} -(\lambda_1 + \lambda_4) & \lambda_1 & \lambda_4 & 0 & 0 \\ \lambda_3 & -(\lambda_3 + \lambda_4) & 0 & \lambda_4 & 0 \\ \lambda_6 & 0 & -(\lambda_1 + \lambda_6) & 0 & \lambda_1 \\ 0 & 0 & \lambda_3 & -\lambda_3 & 0 \\ 0 & \lambda_6 & 0 & 0 & -\lambda_6 \end{bmatrix} = 0 \\ \pi_1 + \pi_2 + \pi_3 + \pi_4 + \pi_5 = 1 \end{array}$$

figura n° D3 Sistema de ecuaciones

Resolviendo el sistema se puede calcular, por ejemplo, la tasa de utilización de un estado, que puede coincidir con la probabilidad en estado estacionario del único estado en el que ese lugar está marcado:

$$\mu[p_2] = \pi_i$$

ANEXO A: BIBLIOGRAFIA

- [1] Y. Dallery "Queueing Networks"- Performance Models for Discrete Event Systems with synchronisations: Formalism and Analysis Techniques" - Capítulo 1- set/98.
- [2] L. Kleinrock "Queueing Systems"- John Wiley, New York, 1975.
- [3] F. Baskett, K. Chandy, R. Muntz, F. Palacios "Open, Closed and Mixed Networks of Queues with Different Classes of Customers" Journal of the ACM, 22, 2, pp. 248-260, 1975.
- [4] Allen, A.O. "Probability, Statistics and Queueing Theory". Academic Press. 1990
- [5] E. Teruel, G. Franceschinis, M Silva "Untimed Petri Nets" - Performance Models for Discrete Event Systems with synchronisations: Formalism and Analysis Techniques" - Capitulo 2- set/98.
- [6] M. Silva "Las Redes de Petri en la Automática y la Informática" - AC. 1985.
- [7] M. Silva "Introducing Petri nets" – In Practice of Petri Nets in Manufacturing, paginas 1-62 – Chapman & Hall, 1993.
- [8] G. Chiola, C. Dutheillet, G. Franceschinis, S. Haddad "Stochastic well-formed coloured nets form symmetric modeling applications" – IEEE Transactions on Computers, 42 (11), November 1993.
- [9] G. Chiola "Timed Petri Nets"- Performance Models for Discrete Event Systems with synchronisations: Formalism and Analysis Techniques" - Capitulo 3- set/98.
- [10] J. Campos, G. Chiola, J. Colom, M. Silva "Properties and performance bounds for timed marked graphs" – IEEE Transactions on Circuits and Systems: Fundamental Theory and Applications, 39(5) paginas 386-401, Mayo 1992
- [11] M. Marsan, G. Chiola, A. Fumagalli "Timed Petri net model for accurate performance analysis of CSMA/CD bus LAN" – Computer Communications, 10(6) paginas 304-312, Diciembre 1987.
- [12] M. Moloy "Performance analysis using stochastic Petri nets" – IEEE Transaction on Computers, 31(9) paginas 913-917, Septiembre 1982.
- [13] G. Chiola, M Marsan, G. Balbo, G. Conte "Generalized stochastic Petri nets: A definition at the net level and its implications" – IEEE Transactions on Software Engineering, 19(2) páginas 89-107, Febrero 1993.
- [14] J.M. Colom, E. Teruel, M. Silva "Logical Properties of P/T systems and their analysis"- Performance Models for Discrete Event Systems with synchronisations: Formalism and Analysis Techniques" - Capitulo 6- set/98.
- [15] J. Campos, G. Chiola, M. Silva "Properties and performance bounds for closed free choice synchronized monoclase queueing networks" – IEEE Transactions on Automatic Control, 36(12), Diciembre 1991.
- [16] M. Silva "Logical Controllers" – In Proceedings of IFAC Symposium on Low Cost Automation, paginas 157-166, 1989.
- [17] G. Balbo "Exponential stochastic Petri nets"- Performance Models for Discrete Event Systems with synchronizations: Formalism and Analysis Techniques" - Capitulo 9- set/98.
- [18] J. Campos "Evaluación de Prestaciones de Sistemas Concurrentes Modelados con Redes de Petri"- Departamento de Informática e Ingeniería de Sistemas – Universidad de Zaragoza ago/04.
- [19] K. Jensen "Coloured Petri Nets – Basic Concepts, Analysis Methods and Practical Use Volume 1 and 2 – Springer-Verlag Berlin Heidelberg New York - 1997
- [20] K. Jensen, G. Rozenberg "High Level Petri Nets" – Springer Verlag, 1995
- [21] J. Capek "Petri Net Simulation of Non-deterministic MAC Layers of Computer Communications Network" - Department of Control Engineering – Faculty of Electrical Engineering, Czech Technical University -2001

- [22] G. Chiola "Petri Nets versus Queueing Networks: Similarities and Differences" - Performance Models for Discrete Event Systems with synchronizations: Formalism and Analysis Techniques" - Capitulo 4- set/98.
- [23] K. Jensen "Una introducción al uso práctico de las Redes de Petri Coloreadas" – Teoret. Computation, SCI, volumen 13, pp 109-135, 1981
- [24] Zerguini "Performance Evaluation of Finite-Population Retrial Queueing Systems Using Generalized Stochastic Petri Nets", Institut de Mathematiques U.S.T.H.B., BP. 32
- [25] L. Kleinrock "On the Modeling and Analysis of Computer Networks" – Special Issue of IEEE Proceedings, página 1179-1191, Agosto 1993.
- [26] V. Frost , B. Melamed "Traffic Modeling for Telecommunications Networks" – IEEE Communications Magazine, 32(3) páginas 70-80, March 1994.
- [27] L. Kleinrock, F. Tobagi "Packet Switching in radio Channels: Part I, II y III. IEEE Transactions on Communications, Corn-23, pp 1400-1433, Dic/1975 and pp 832-844, Dic/1976
- [28] A. Baiocchi, N. Blefari-Melazzi "Steady-State Analysis of the MMPP/G/I/K Queue", IEEE Trans. Commun. Vol. COM-41, pp 531-534, Abril/1993
- [29] P Tran-Gia "A Renewal Approximation for the Generalized Poisson Process", Proc. Int. Workshop on App. Math. And Performance/Reliability Model in Computer/Communication Systems, Pisa, Italy, Sept/1983
- [30] R. German, A. Heindl "Performance Evaluation of IEEE 802.11 Wireless LANs with Stochastic Petri Nets" – Technische Universitat Berlin 1999
- [31] G. Bianchi "Throughput evaluation of the IEEE 802.11 distributed coordination function" – In Proc. 5th Int. Wsho. On Mobile Multimedia Communications, pagina 307-318, Berlin, 1998.
- [32] L. Kleinrock "Queueing Systems" – Volumen II: Computer Applications, John Wiley & Sons, 1976.
- [33] A. Tanenbaum "Redes de Computadoras" – Cuarta Edición Prentice Hall Hispanoamericana - 2003
- [34] A. Ferscha "Simulations Principles" - Performance Models for Discrete Event Systems with synchronizations: Formalism and Analysis Techniques" - Capitulo 24- set/98.
- [35] Henryk Anschuetz - Simulador HPSim 1999-2002
http://www.winpesim.de/petrinet/e/hpsim_e.htm
- [36] W. Stallings "Comunicaciones y Redes de Computadores" – Sexta Edición Printice Hall 2000
- [37] Jose Capek – Simulador STPNPlay -2000
<http://dce.felk.cvut.cz/capekj/StpnPlay/>

ANEXO B: GLOSARIO

Acrónimo en Ingles y Castellano

ACK	Acknowledge-Reconocimiento
ARPA	Advance Research Projects Agency-Agencia de Proyectos de Avanzada
BEB	Binary Exponential Backoff-Retroceso Exponencial Binario
CPNs	Coloured Petri Nets-Redes de Petri Coloreadas
CSMA	Carrier Sense Multiple Acces-Acceso Multiple con Sensado de Portadora
CSMA/CD	CSMA with Collision Detection-CSMA con Detección de Colisiones
CSMA/CA	CSMA with Collision Avoidance- CSMA con Prevención de Colisiones
CTMC	Continuous Time Markov Chains-Cadenas de Markov de Tiempo Continuo
CTS	Clear to Send-Listo para Enviar
DCF	Distributed Control Function-Función de Control Distribuída
DEDS	Discrete Event Dinamic Systems-Sistemas Dinámicos de Eventos Discretos
DIFS	DCF Inter Frame Space-Espacio Intertrama DCF
DSPNs	Deterministic and Stochastic Petri Nets-Redes de Petri Determinística y Estocástica
EIFS	Extended Inter Frame Space-Espacio Interframe extendido
FCFS	First Come First Served- Primero en Llegar Primero en ser Atendido
FDM	Frecuency Division Multiplexing-Multiplexación por división de frecuencia
FIFO	First Input First Ouput-Primero en Entrar Primero en Salir
FTP	File Tranfer Protocol-Protocolo de tranferencia de ficheros
GSPNs	Generalized Stochastic Petri Nets-Redes de Petri Estocásticas Generalizadas
GSPP	Generalizad Switched Poisson Process-Procesos de Poisson Comnutados General
HLPNs	High Level Petri Nets-Redes de Petri de Alto Nivel
IEEE	Institute of Electrical and Electronics Engineers-Instituto de ingenieros eléctricos y electrónicos
IPP	Interrupted Poisson Process-Procesos de Poisson Interrumpidos
ISO	Internacional Organization for Standardization-Organización Internacional para normalizaciones
LAN	Local Area Network-Red de Area Local
LCFS	Last Come First Served-Último en Llegar Primero en ser Atendido
LIFO	Last Input First Ouput-Último en Entrar Primero en Salir
LLC	Logical Link Control-Control de enlace lógico
LLPNs	Low Level Petri Nets-Redes de Petri de Bajo Nivel
MAC	Medium Access Control-Control de Acceso al Medio
MACA	Multiple Access with Collision Avoidance-Acceso Múltiple con Prevención de Colisiones
MACAW	Wireless MACA-MACA Inalambrica
MMPP	Markov Modulated Poisson Process-Procesos de Poisson Modulado de Markov
MPNs	Markov Petri Nets-Redes de Petri Markoviana
NNTF	Network News Transfer Protocolo-Protocolo de Transferencia de Nuevas Redes
OSI	Open Systems Interconexión-Interconexión de sistemas abiertos
PCF	Puntual Coordination Function-Función de Coordinación Puntual
PLC	Programmable Logical Computer-Computador Lógico Programable
PRI	Preempt Repeat Identical-Preferido Repetir Idéntico
PRD	Preempt Repeat Different-Preferido Repetir Diferente
RG	Reachability Graph-Grafo de Alcanzabilidad
RTS	Request to Send-Petición de Envío

SIFS	Short Inter Frame Space-Espacio Intertrama corto
SMTP	Simple Mail Transfer Protocol-Protocolo Simple de Transferencia de correo
SIRO	Service in random order-Servicio en Orden Aleatorio
SWNs	Stochastic Well-formed Nets-Redes Estocásticas bien formadas
SPNs	Stochastic Petri Nets-Redes de Petri Estocásticas
TDM	Time Division Multiplexing-Multiplexación por División de Tiempo
TSF	Time Synchronization Function-Function de Sincronización de tiempo
WAN	Wide Area Network-Red de Area Amplia
WLANs	Wireless LANs-LANs Inalambricas
WNs	Well-formed Nets-Redes bien formadas

ANEXO C: HERRAMIENTAS ACADÉMICAS DE ENSEÑANZA

C.1 Introducción

Atento al desarrollo en los sucesivos Capítulos del trabajo de tesis, y a los conceptos de modelación sobre sistemas de estado finito, nos planteamos el uso de las Redes de Petri, como herramienta formal de enseñanza para la transferencia de conocimientos y contenidos, cuestión que ha sido eventualmente incluida en algunas bibliografías sobre la temática de protocolos..

Aunque existe siempre la alternativa de enseñanza de dichos contenidos en base a métodos descriptivos en forma narrada, o el uso de cualquiera de los formalismos de representación y análisis, tales como las redes de colas, grafos dirigidos, diagramas de fase, en este caso nos plantearemos la preparación de una clase de enseñanza usando los conceptos y herramientas de simulación disponibles para Redes de Petri sobre protocolos de comunicaciones, con la siguiente secuencia:

- Introducción de los conceptos elementales de las Redes de Petri (10 minutos)
- Presentación con conceptos elementales de las Redes de Petri FPQSPN (10 minutos)
- Presentación y análisis de los distintos modelos disponibles de Redes de Petri FPQSPN para los protocolos de comunicaciones seleccionados, como por ejemplo Aloha puro, Aloha ranurado y CSMA persistente. Debería hacerse hincapié en la similitud de los modelos y la complejidad creciente, con el objeto de perfeccionar el protocolo y reducir la cantidad de colisiones, y por lo tanto, la mejora en el rendimiento general (30 minutos).
- Reconocimiento y uso del Simulador HPSim como herramienta de animación del modelo de uno de los protocolos seleccionados, por ejemplo Aloha puro. Se considera que se analiza y discute la evolución del canal por los distintos estados, y si es posible, hacer una breve introducción a los conceptos de rendimiento, paquetes colisionados, con éxito, etc. en la medida del desarrollo de la actividad. El simulador es una herramienta muy útil en este sentido dado que lleva una cuenta de la cantidad de veces que el canal se encuentra en cada estado posible. HPSim no es un simulador de análisis de rendimiento pero puede introducir la noción conceptual necesaria (20 minutos).
- Reconocimiento y uso del Simulador STPNPlay como herramienta de análisis de rendimiento de algunos de los modelos de los protocolos seleccionados, por ejemplo Aloha puro y Aloha ranurado. Se considera que se analiza y discuten las curvas de rendimiento en función del tráfico ofrecido G , para distintas situaciones cambiando los parámetros del simulador. Además, de poner en evidencia las curvas sobre la evolución de los distintos niveles de colisión (20 minutos).
- Presentación y comparación con las curvas de rendimiento en función del tráfico ofrecido obtenidas desde el punto de vista analítico con el software Mathematica (10 minutos).
- Realización de actividades prácticas complementarias con los simuladores. Se entiende que se provee a los alumnos de una Guía Práctica con ejercitación y de los simuladores para la evaluación de los protocolos seleccionados en distintas situaciones operativas a definir (30 minutos).

Todos los tiempos asignados a las actividades son orientativos, y dependerían de las disponibilidades horarias para la temática dentro de la planificación de la materia, o curso.

VI.2 Láminas de Presentación en multimedia

A continuación se han seleccionado las figuras mínimas para el desarrollo de las actividades, de acuerdo a la secuencia mencionada en el párrafo anterior. Además, se supone que los simuladores están disponibles para mostrarlos en el momento oportuno vía un equipo multimedia.

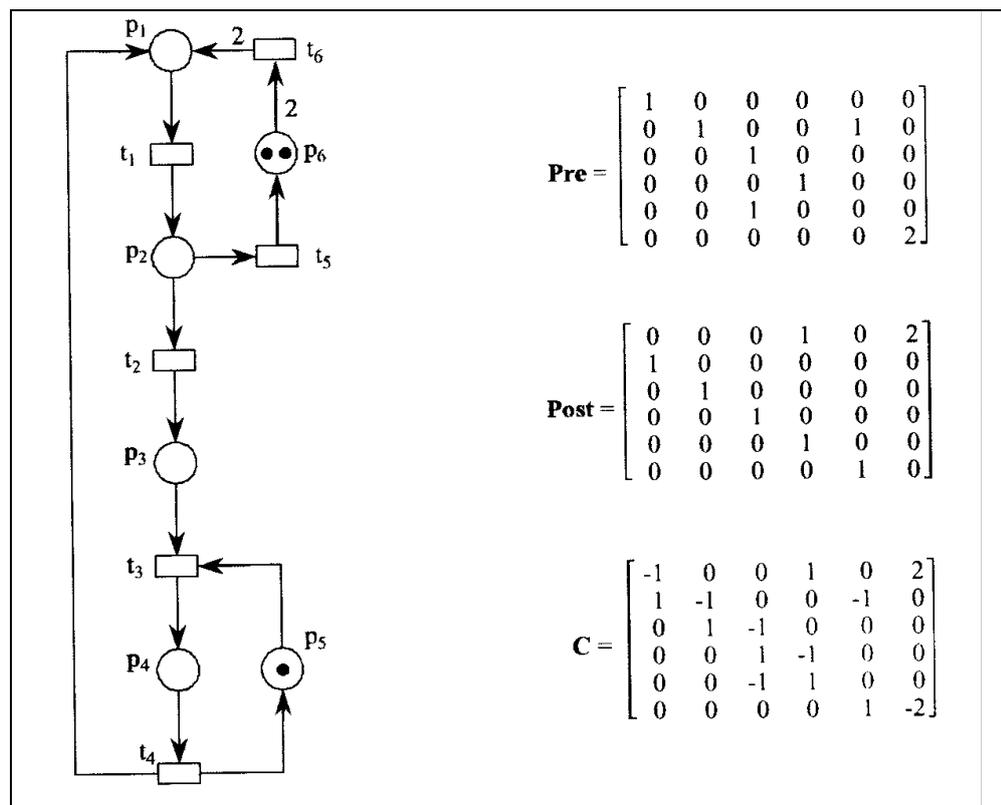


Lámina n° 1 Un sistema de red Lugar/Transición (P/T) y sus matrices incidentes

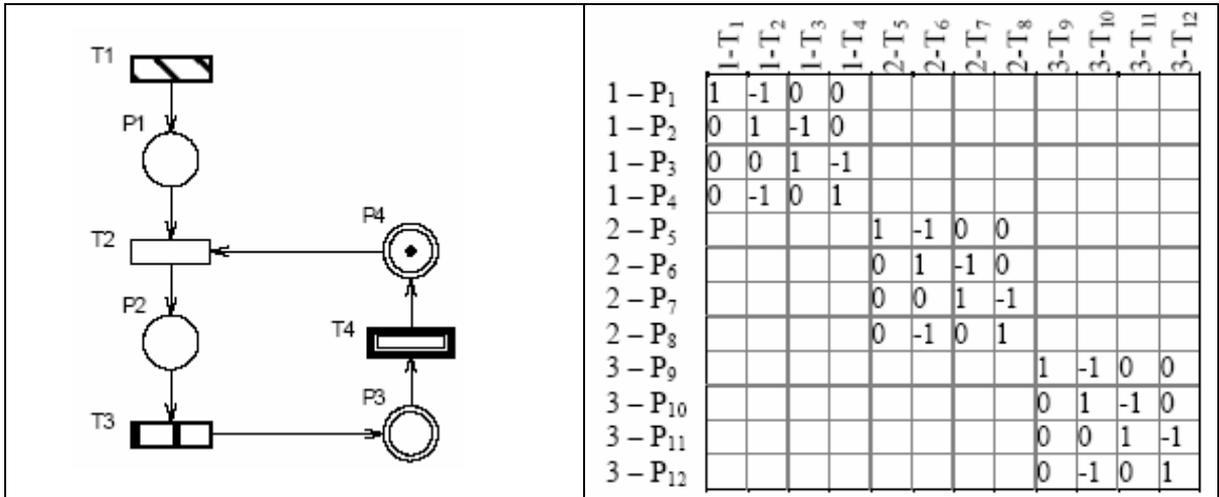


Lámina nº 2 Modelo FPQSPN de una sección crítica

Lámina nº 3 Tres modelos independientes

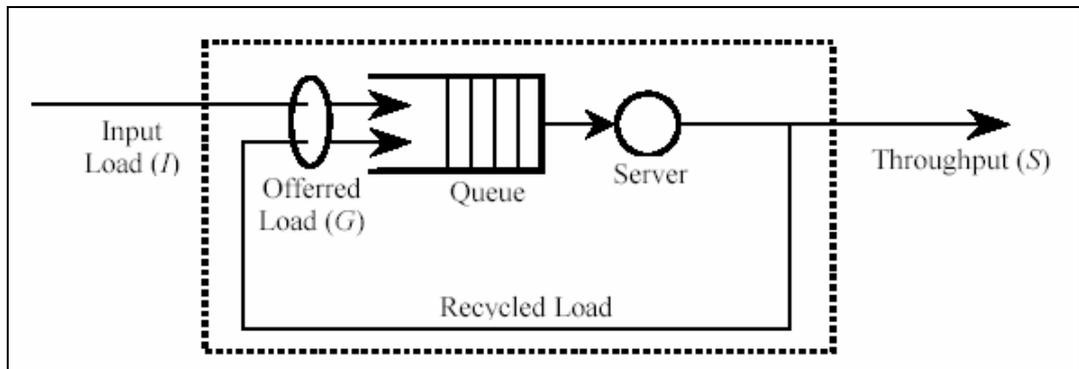


Lámina nº 4 Un modelo de cola general para una red o dispositivo de red

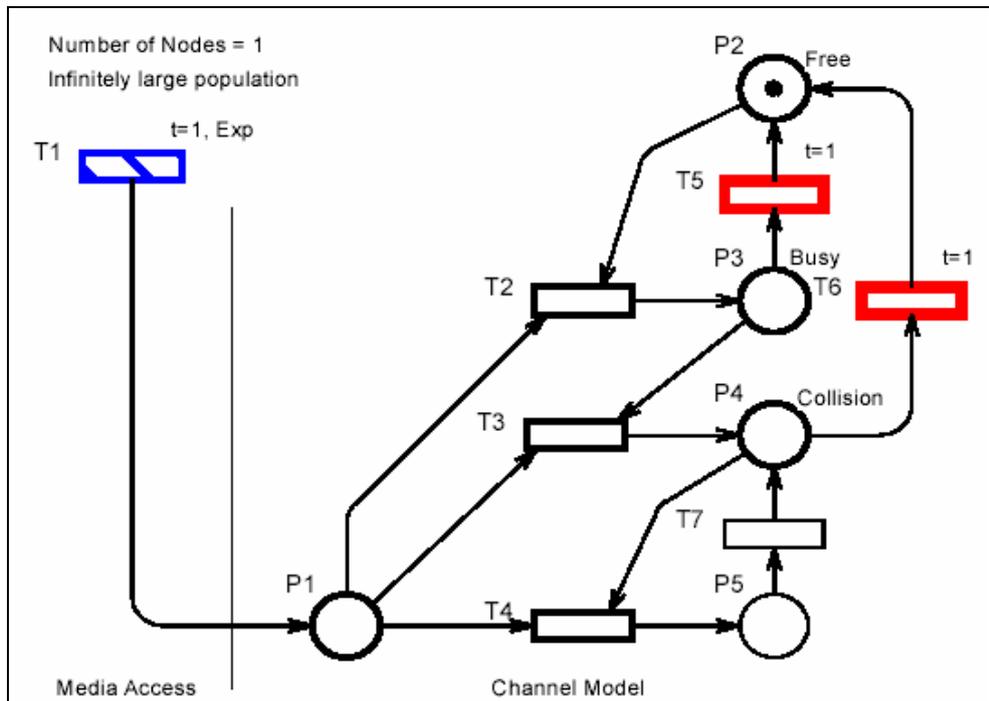


Lámina nº 5 Modelo FPQSPN del protocolo Aloha Puro

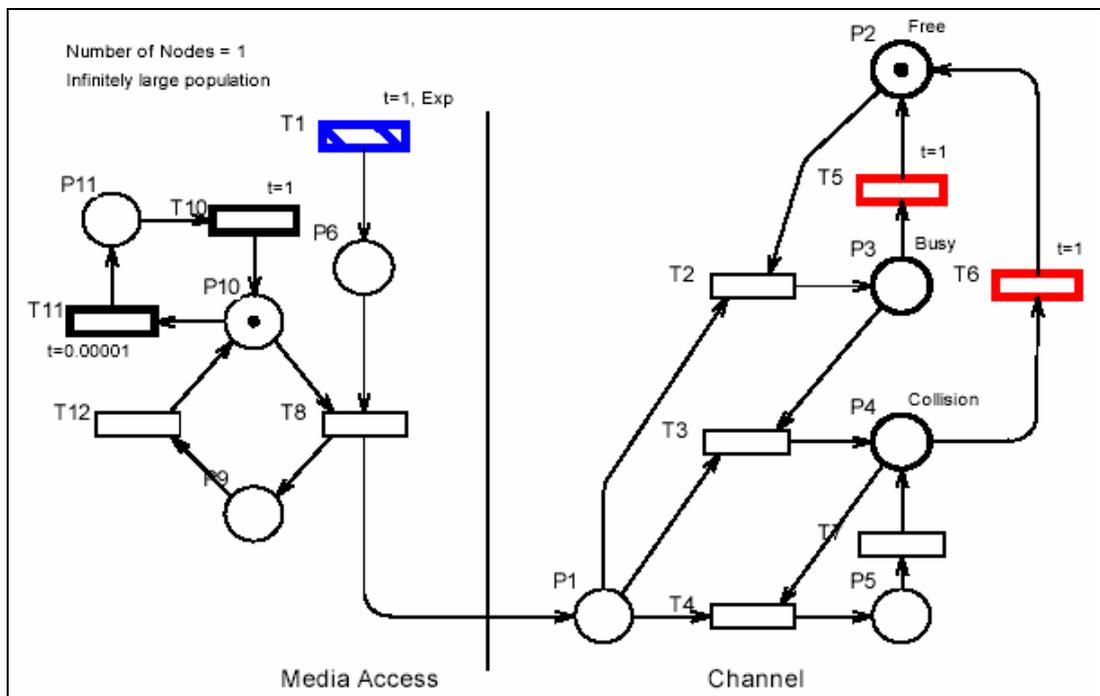


Lámina nº 6 Modelo FPQSPN del protocolo Aloha Ranurado

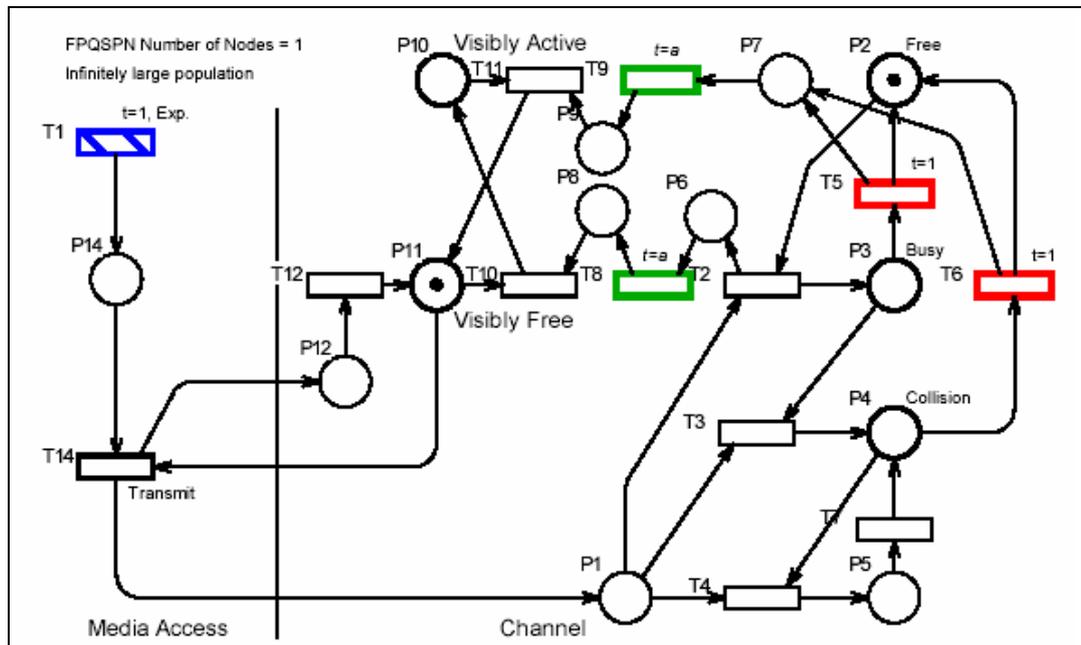


Lámina nº 7 Modelo FPQSPN del protocolo CSMA Persistente

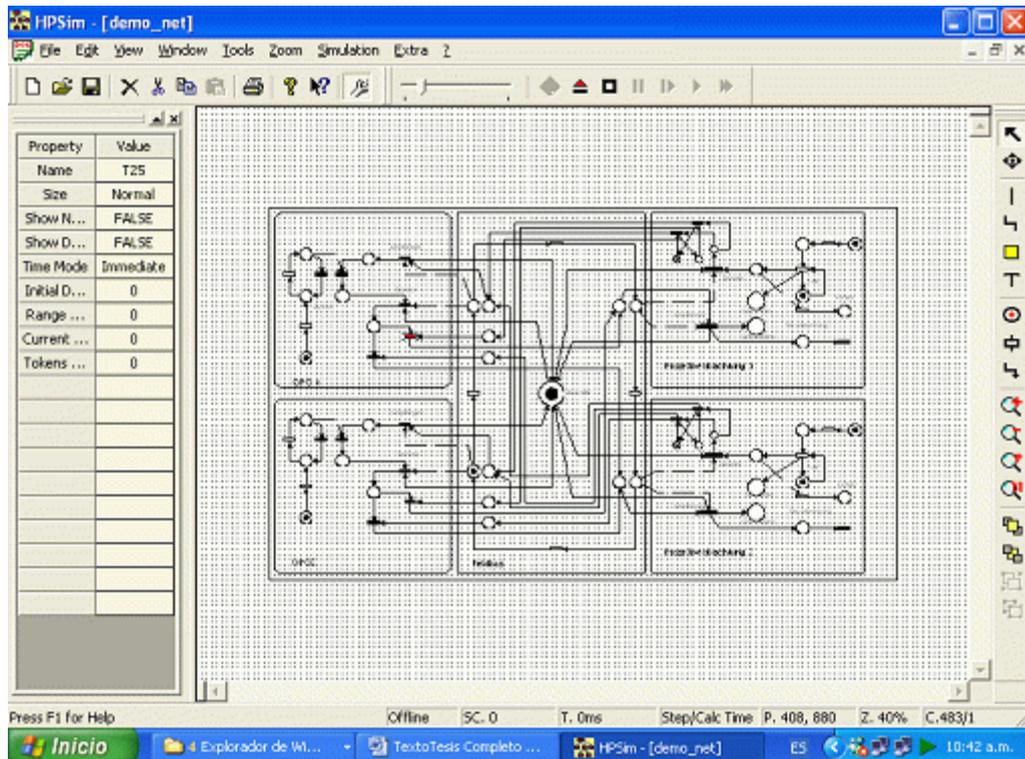


Lámina nº 8 Ambiente de trabajo del simulador HPSim

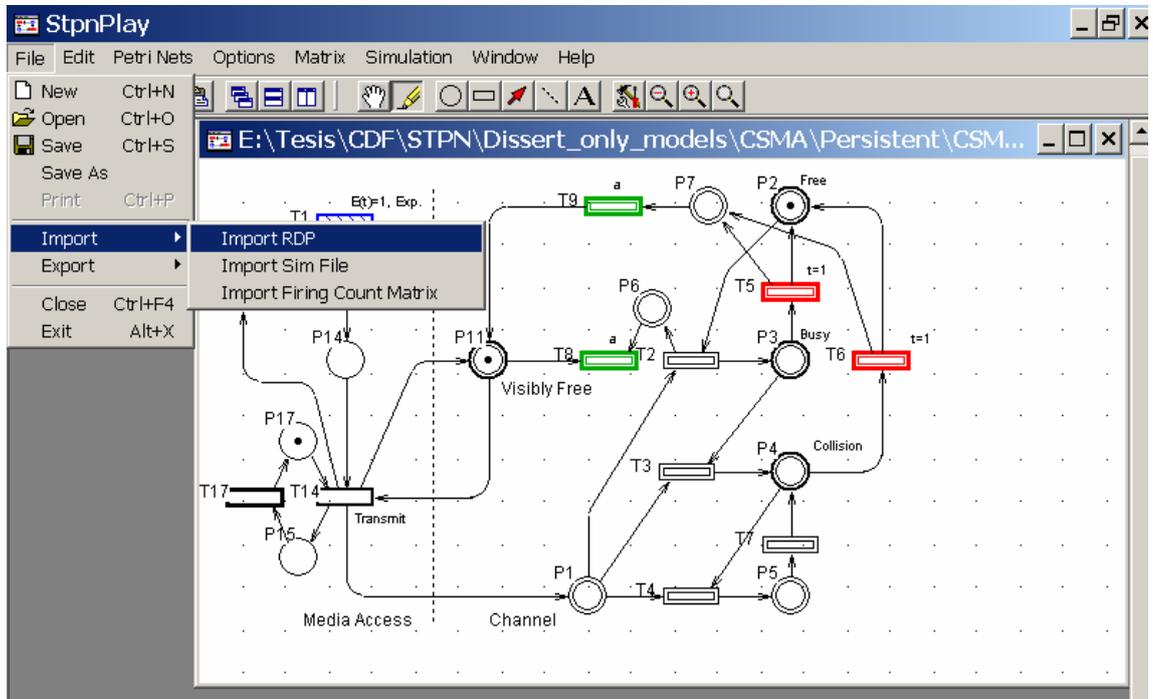


Lámina nº 9 STPNPlay: Menú de File

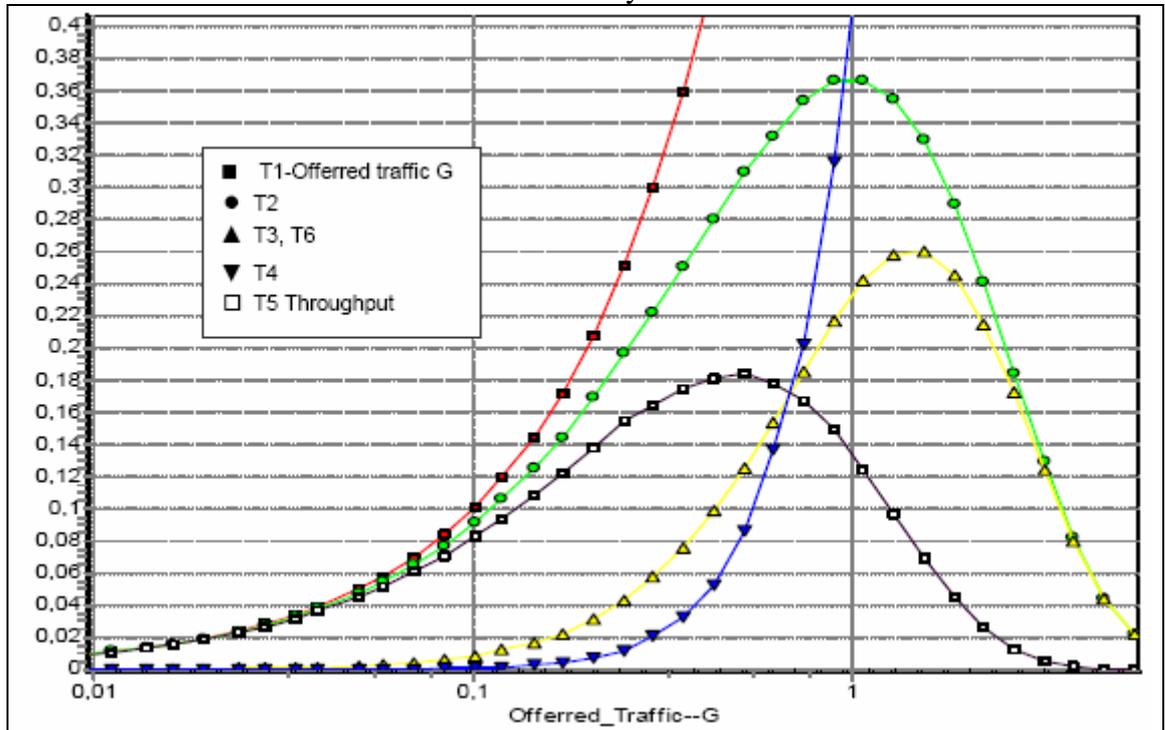


Lámina nº 10 Rendimiento de Aloha puro – Frecuencias de disparo de transición

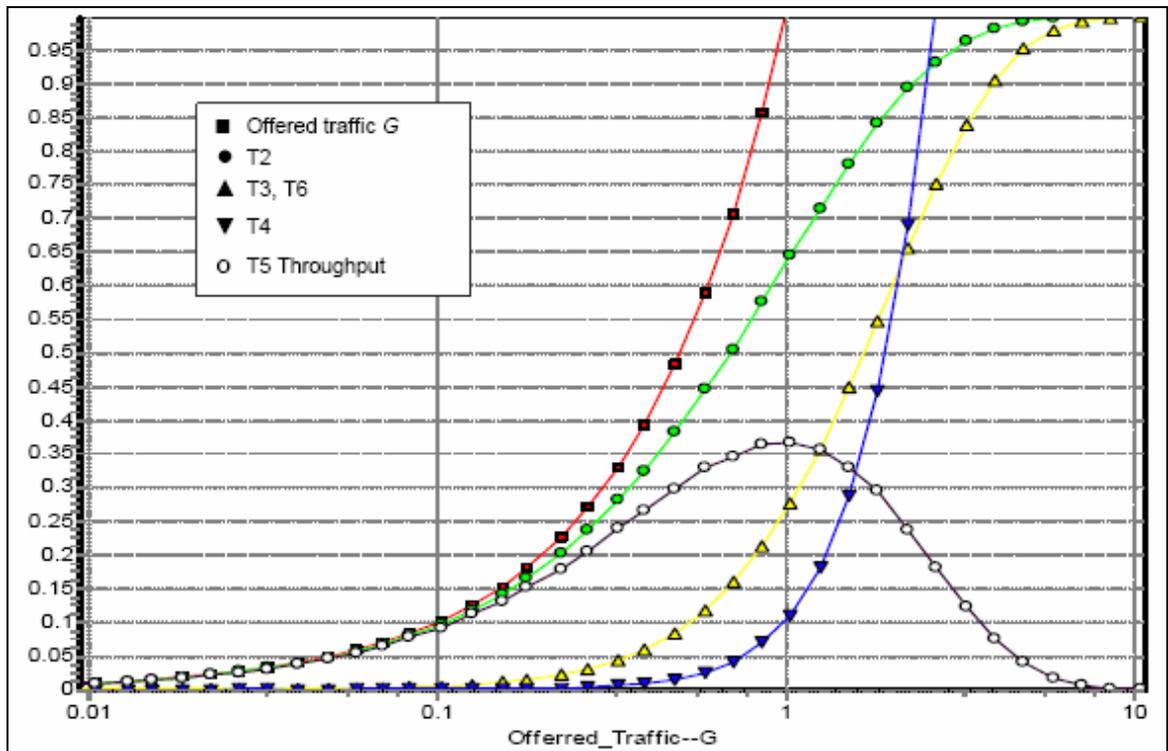


Lámina nº 11 Rendimiento de Aloha ranurado – Frecuencias de disparo de transición

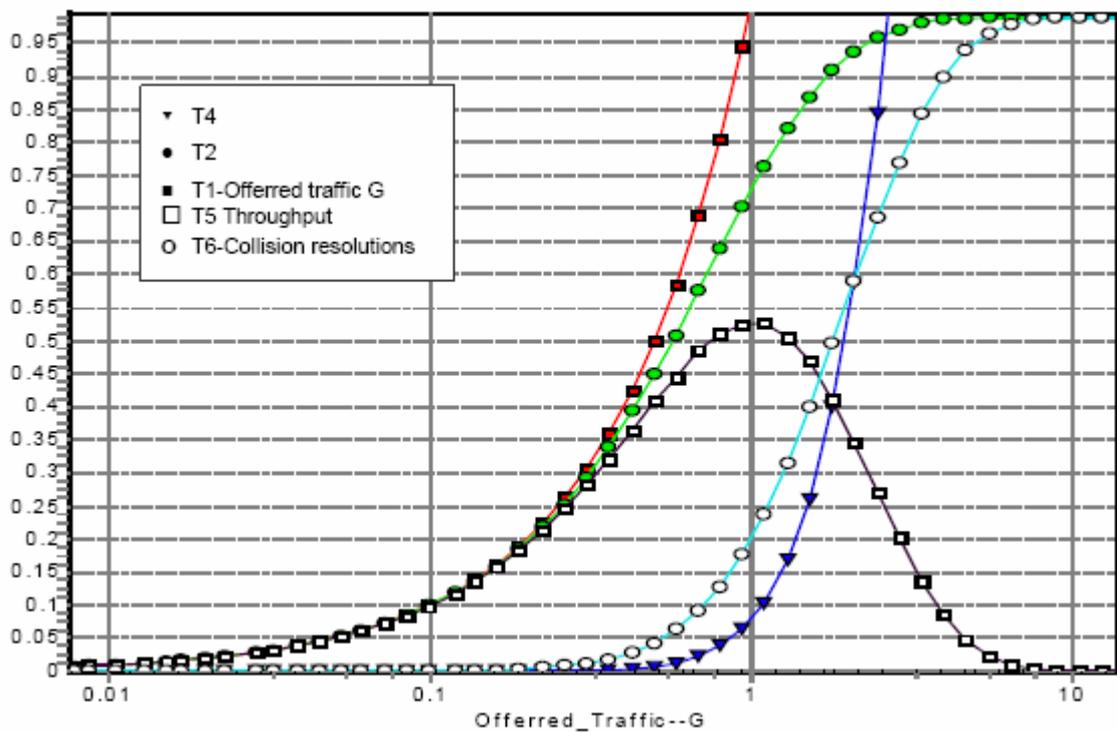


Lámina nº 12 Rendimiento de CSMA Persistente – Frecuencias de disparo de transición para $a = 0.01$

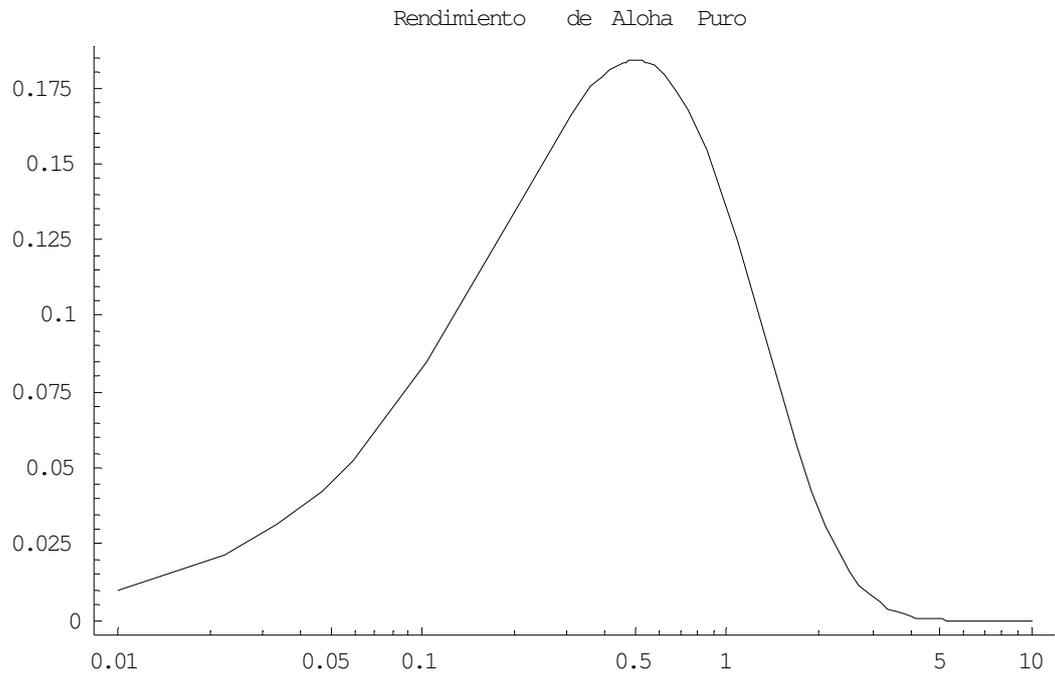


Lámina nº 13 Rendimiento S en función del tráfico ofrecido G Aloha puro

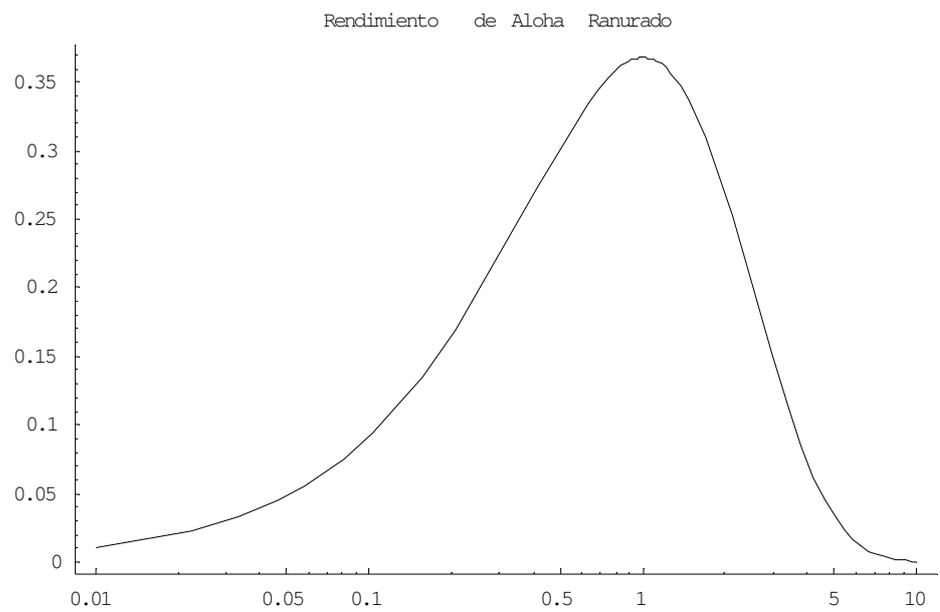


Lámina nº 14 Rendimiento S en función del tráfico ofrecido G Aloha ranurado

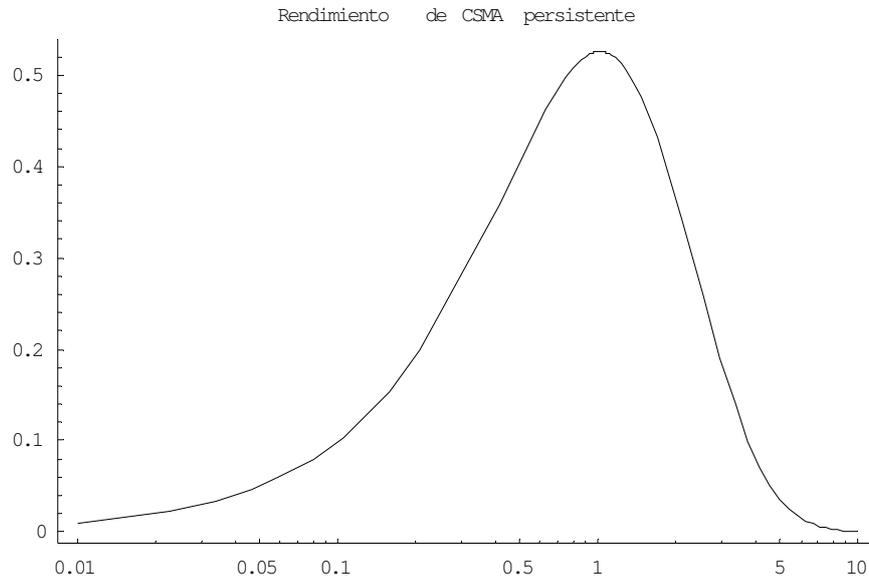


Lámina nº 15 Rendimiento S en función del tráfico ofrecido G CSMA persistente

Ejercitación Tentativa

Con el uso del simulador STPNPlay, y los modelos de Aloha, Aloha ranurado y CSMA persistente, resuelva los siguientes ejercicios:

- **Ejercicio nº 1:** Indique (aproximadamente), con que probabilidad, en el punto de máximo rendimiento de Aloha, un nuevo paquete se encontrará con el canal ocupado, y con el canal en estado de colisión.
- **Ejercicio nº 2:** Reitere el ejercicio anterior para $G=1$ y $G=10$.
- **Ejercicio nº 3:** Obtenga las curvas características de rendimiento para 1, 2 y 10 nodos, y explique los resultados para Aloha puro.
- **Ejercicio nº 4:** Reitere la ejercitación anterior para Aloha ranurado y CSMA persistente.