

011691



Universidad Nacional Autónoma de México

*Análisis de Enlaces de Comunicación con
Conmutación Híbrida*

ZENG CHAOMING

T E S I S

*Presentada a la División de Estudios de
Posgrado de la*

FACULTAD DE INGENIERIA
de la

UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

*como requisito para obtener
el grado de*

MAESTRO EN INGENIERIA
DE CONTROL

Ciudad Universitaria

2002

TESIS CON
FALLA DE ORIGEN



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

ANALISIS DE ENLACES DE COMUNICACION CON CONMUTACION HIBRIDA

Créditos asignados a la tesis Doce (12)

APROBADA POR EL JURADO:

Presidente: Dr. Andrés Buzo de la Peña

~~Luis Andrés Buzo~~

Vocal: Dr. Federico Kuhlmann Rodriguez

F. Kuhlmann

Secretario: Dr. Guillermo Rebolledo Cortizo

G. Rebolledo

Suplente: Dr. Stanislaw Raczynski

S. R.

Suplente: Dr. Francisco Garcia Ugalde

F. Garcia Ugalde

獻
給
我
的
祖
國

一九五五年十月於墨

TESIS CON
FALLA DE ORIGEN

AGRADECIMIENTO

Deseo expresar mi sincero agradecimiento al Dr. Federico Kuhlmann, como el director de esta tesis, por su orientación, su ayuda y sus valiosos recomendaciones que me ha brindado durante todo el desarrollo de este trabajo.

También quiero agradecer al Dr. Andrés Buzo y otros profesores, los amigos de la DEEFI, por sus cooperaciones y apoyos.

Como el becario de la República Popular de China, agradezco a la Embajada de China en México, por el apoyo y el estímulo que me ha dado durante mis estudios en México.

TESIS CON
FALLA DE ORIGEN

RESUMEN

En esta tesis se presenta un panorama general de redes de comunicación híbrida, y se analiza particularmente un enlace de comunicación de redes híbridas. Se propone un esquema para multiplexaje con conmutación por circuitos y conmutación por paquetes mediante una frontera móvil bidireccional (FM2). Este sistema puede manejar dos diferentes tipos de tráfico en forma híbrida y adaptiva. Los dos tipos de tráfico se pueden caracterizar por voz que utiliza la conmutación por circuitos, y datos cortos o intermitentes que utilizan la conmutación por paquetes. Se modela el sistema por medio de una clase de cadenas de Markov bi-dimensional. Este modelo nos permite evaluar el desempeño del esquema propuesto, y comparar con algunos otros esquemas que han sido estudiados intensamente durante los últimos años. Los resultados muestran que el esquema FM2 es más flexible y ofrece algunas mejoras, cuando las cargas de tráfico varían en un rango amplio.

Otro trabajo que se presenta en esta tesis consiste en proponer un sistema integrado que permite aceptar tres tipos de tráfico, utilizando conmutación por paquetes y dos tipos de conmutación por circuitos (tipo pérdida y tipo espera). Este tipo de sistemas es una extensión de sistemas híbridos, añadiendo otro tipo de tráfico como transferencias de archivos, las cuales se caracterizan por sus grandes volúmenes de transacciones, y no requieren respuestas inmediatas. El desempeño del sistema se evalúa tanto para casos de buffer finito como para casos de buffer infinito, por medio de un modelo de cadena de Markov tridimensional, utilizando algunas técnicas numéricas o aproximaciones cuasi-estáticas. Finalmente, se sugiere una técnica de control de flujo de tráfico de paquetes para lograr un mejor desempeño, debido al aumento drástico de la cola de paquetes durante largas transacciones de los demás tipos de tráfico.

Los cálculos se efectuaron en Burroughs B0900 y Apple IIe.

Cd. Universitaria, México
Septiembre, 1985



ANALISIS DE ENLACES DE COMUNICACION CON CONMUTACION HIBRIDA

INDICE

Agradecimiento

Resumen

Indice

Capítulo 1 Introducción

- 1.1 Redes de computadoras
- 1.2 Comunicación entre computadoras
- 1.3 Una configuración de red híbrida-SENET
- 1.4 Trabajos anteriores

Capítulo 2 Herramientas para modelar sistemas de comunicación

- 2.1 Introducción a la teoría de colas
- 2.2 Procesos estocásticos
- 2.3 Algunos sistemas de colas importantes

Capítulo 3 Un multiplexor de conmutación híbrida con frontera móvil bidireccional (FM2)

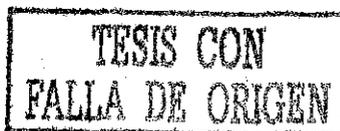
- 3.1 Las estrategias de multiplexaje
- 3.2 Protocolo del esquema FM2
- 3.3 Modelo
- 3.4 Solución en estados permanentes
- 3.5 Evaluación de desempeño
- 3.6 Análisis y comparación

Capítulo 4 Análisis de un multiplexor integrado para tres tipos de tráfico

- 4.1 Introducción
- 4.2 Modelo
- 4.3 Soluciones exactas aplicando técnicas numéricas
- 4.4 Aproximación cuasi-estática
- 4.5 Análisis de desempeño
- 4.6 Control de flujo

Capítulo 5 Conclusiones y trabajos futuros

Referencias



CAPITULO 1
INTRODUCCION

1.1 Redes de computadoras

La primera aparición de redes de computadoras (en los últimos 1960's) correspondió a la necesidad de distribuir servicios de computadora a un gran número de usuarios remotos: en tal red (usualmente llamadas redes "centralizadas") sólo se dispone una computadora central, conectada a un gran número de terminales pasivas por medio de una red de comunicación. Una representación esquemática de una red centralizada se muestra en la Fig.1.1.

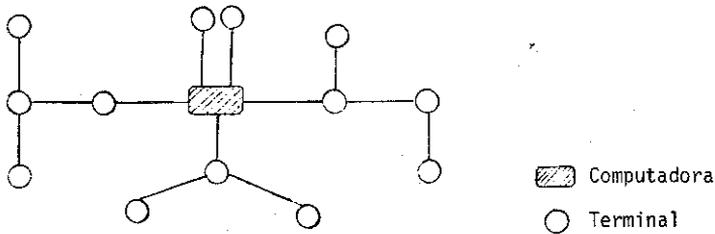


Fig.1.1 Red de computadora centralizada

Más adelante, en la búsqueda de una configuración de computadoras más eficiente, se sugería que los sistemas de computadoras existentes podrían ser utilizados mejor cuando se interconectan en una red, tal que los recursos pueden compartirse. Entre estos recursos mencionamos: capacidad de computadora, hardware y software especializados y bases de datos etc.. Este tipo de redes (llamadas redes "distribuidas") difiere del primero en que los recursos están distribuidos en los nodos, en lugar de acumularlos en un nodo central. Tal vez, el proyecto más grande y más ambicioso de las redes distribuidas fue la ARPA-NETWORK [1], Fig.1.2 es la configuración reciente de la ARPANET.

Algunos de los objetivos comunes de las redes de computadoras son [2]:

TESIS CON
FALLA DE ORIGEN

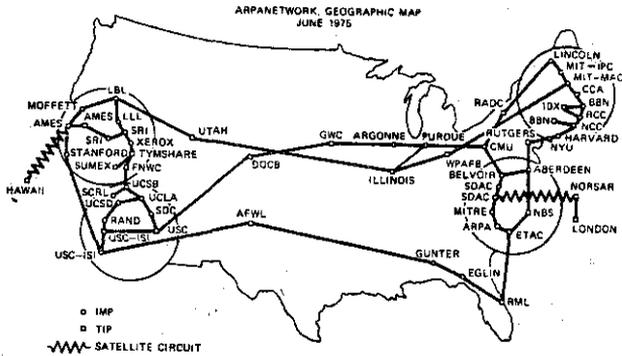


Fig.1.2 Configuración de ARPA-NET

- 1) Compartir los recursos distantes tales como información (bases de datos) o procesadores (CPU's).
- 2) Proporcionar comunicaciones entre usuarios, procesos y procesadores geográficamente distribuidos.
- 3) Proporcionar compatibilidad de equipos y software disímiles.
- 4) Aumentar la confiabilidad para los procesos si algún procesador falla.
- 5) Facilitar el control centralizado tanto para civil como para militar.

Es conveniente dividir redes de computadoras en dos subredes: la subred de usuario y la subred de comunicación.

La subred del usuario proporciona la capacidad de acceder a la red y contiene tres componentes: computadora anfitrión, controladores de terminal y terminales. La computadora anfitrión proporciona bases de datos y programas de aplicación para accesos de usuarios terminales. Por otro lado, la computadora de anfitrión está conectada con un nodo de la subred de comunicación, a través de un canal de multiplexor de alta velocidad, o directamente con un enlace de comunicación. El usuario terminal accede a la red generalmente a través de un controlador de terminales.

La subred de comunicación es una colección de nodos donde se dispone de los equipos (Modem's etc.) y procesos que permiten comunicar con otros nodos, por medio

de un conjunto de enlaces de comunicación. El nodo de la red tiene función como inter-
 face para la subred de usuario y otros nodos, recepción y transmisión de mensajes, ma-
 nejo de errores, control de flujo y evitar la congestión, enrutamiento etc..

Esta tesis se concentra en la subred de comunicación, y de aquí en adelante, la
 llamamos simplemente red de comunicación o red.

1.2 Comunicación entre computadoras

Muchos diferentes esquemas de comunicación pueden implementarse en una red de
 comunicación: línea dedicada, conmutación por circuitos (CC), conmutación por mensajes
 (CM), conmutación por paquetes (CP), radio-difusión, satélite o fibra óptica. También
 puede ser un esquema híbrido que incluye más de un esquema al mismo tiempo. Entre ellos,
 los más comunes son conmutación por circuitos, conmutación por mensajes o por paquetes.

a) Conmutación por circuitos:

Una red de CC transmite un mensaje al proporcionar una trayectoria completa de
 enlaces de comunicación desde el nodo origen al nodo destino. La trayectoria se esta-
 blece por un mensaje especial de señalización que se manda del origen al destino. Una
 respuesta al mensaje de señalización desde nodo destino informa al origen proceder con
 la transmisión de datos. Los enlaces de la trayectoria establecida permanecen asignados
 durante toda la duración de la comunicación. Fig. 1.3 es un proceso típico de CC.

Hay dos modos de CC: modo pérdida y modo espera (o cola).

El modo pérdida es similar a un sistema telefónico. En el si una llamada no
 encuentra una trayectoria disponible, la llamada se pierde, por lo que el usuario
 deberá intentar la comunicación en otra ocasión.

Al contrario del modo pérdida, el modo de espera permite que la llamada perma-
 nezca en el sistema (haciendo cola) hasta que se establezca una trayectoria. Se puede
 observar que el tiempo de retardo en CC tiene dos partes: tiempo para establecer la
 trayectoria y tiempo para la transmisión de datos.

Las medidas de desempeño del esquema CC pueden ser: probabilidad de bloqueo
 de las llamadas para el modo de pérdida, y tiempo de retardo total para el modo de
 espera.

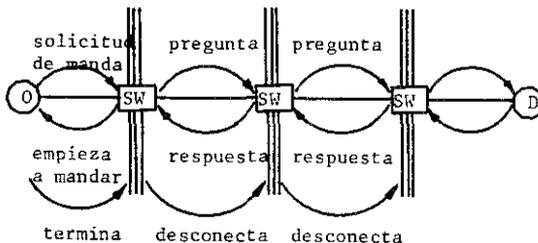


Fig.1.3 Esquema de CC

b) Conmutación por mensajes y por paquetes:

Una red de CM transmite un mensaje entre los nodos al mover el mensaje a través de varios enlaces de transmisión y buffers de mensaje (ver Fig. 1.4). Un mensaje se almacena primero en el buffer a esperar la transmisión al siguiente nodo; la transmisión no puede empezar hasta el buffer del siguiente nodo en la ruta todavía no ha asignado para él. El mensaje se tiene que retransmitir si el nodo siguiente no reconoce el mensaje transmitido (por error de transmisión o otras causas). Hay que mencionar que un mensaje para ser transmitido, necesita llevar algunas informaciones adicionales como la dirección, bit para checar error, la información de enrutamiento, el orden de mensaje etc..

El esquema de CP es una variación de CM. Los mensajes que van a transmitir se descomponen primero en varios mensajes pequeños de tamaño fijo (o variable) llamados paquetes, y estos paquetes se transmiten en forma independiente y con diferentes rutas para llegar a su destino (se conoce como efecto de pipelining). Por lo que el nodo destino se necesita un proceso de reordenamiento y reensamble de los paquetes.

Las medidas de desempeño de estos dos esquemas pueden ser: tiempo de retardo y probabilidad de que un buffer se llene (o overflow de buffer por lo tanto el mensaje transmitido se pierde).

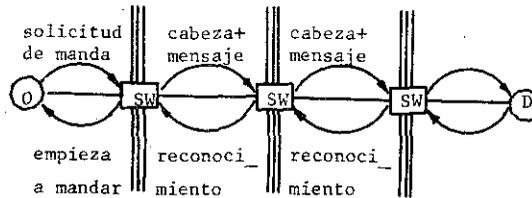


Fig.1.4 Esquema de conmutación por mensajes

c) Conmutación híbrida:

La tendencia de requerimiento futuro de sistemas de comunicación digital indica un incremento de diversidad de características de tráfico como [3]

- 1) Amplias disparidades entre rapidezces de tráfico, desde rapidez bajo como terminales TTY, que requieren cientos de bit/seg, hasta terminales de video y gráficas de banda amplia, que requieren cientos de Kbits/seg.
- 2) Amplias disparidades en tamaño de transacciones, empezando con mensajes interactivos de algunos cientos de bits, y continuando con las transferencias de archivos de datos de millones de bits.
- 3) Variación de tiempo de respuesta para acomodar voz y video, las cuales requieren respuesta continua y casi tiempo real; terminales interactivos

TESIS CON
FALLA DE ORIGEN

de datos, que requieren un tiempo de respuesta en orden de segundos;
y datos bultos como archivo, que puede esperar por unas horas.

De las descripciones anteriores, es evidente que una sola técnica de conmutación no puede satisfacer a los diferentes requerimientos de usuario. Por ejemplo, ni CP ni CM pueden manejar las comunicaciones de voz o señales de video de banda amplia (aunque ya existe técnicas para paquetizar voz). Esto es debido a que la CP o CM poseen muchas redundancias (por el mecanismo de detección de error, reconocimiento y reordenamiento), y las señales de voz o video no son sensibles a la tasa de error de canal, pues con el tiempo de respuesta de CP o CM no se puede mantener la integridad de información.

Por otro lado, la CC no es conveniente para los mensajes cortos intermitentes, porque el establecimiento de la trayectoria en CC toma más tiempo que la misma transmisión ocasionando un mayor retardo en CP. La característica fundamental de los tráficos entre computadoras es que generalmente se presenta en "ráfagas" de datos seguidos de largas pausas. Un simple cálculo por James Martin [4] muestra que sólo 17% de la capacidad total de comunicación se usa cuando un teletipo se conecta a una computadora a través de una línea telefónica pública.

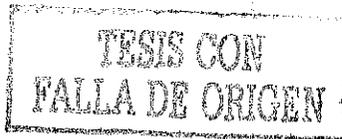
La Tabla I muestra las comparaciones de las características de CC, CM y CP en varios aspectos.

Tabla I Comparaciones de CC, CM y CP

Descripción	Código	CC	CM	CP
1. Permite enlaces de diferentes capacidades	+	no	sí	sí
2. Proporciona el efecto de pipelining	+	no	sí	sí
3. Puede seleccionar adaptivamente la ruta más rápida como la función de congestión de red	+	no	sí	sí
4. Puede transmitir mensajes cortos primero, debido a alguna causa	+	no	no	sí
5. Tiene control de error	+	algo	sí	sí
6. Requiere buffer en los nodos	-	no	sí	sí
7. Posible cola y congestión en la red	-	no*	sí	sí
8. Requiere proceso de enrutamiento y control de flujo	-	no*	sí	sí
9. Requiere reordenamiento y numeración	-	no	algo	sí
10. Requiere paquetización	-	no	no	sí
11. Aplicable a los usos de tiempo real	+	sí	bajo estudio	

*: No toma en cuenta el proceso de establecimiento de comunicación de CC.

Código: + para ventaja, - para desventaja.



La conclusión es que para la transmisión de datos de gran volumen o tiempo real entre localizaciones distantes, CC es más atractiva que CP y CM. Pero para la transmisión de datos de tamaño pequeño o intermitentes, la CP tiene mejor desempeño para el sistema (mayor utilización de canal) o para el usuario (menor tiempo de retardo y mayor seguridad).

Con el advenimiento de los nuevos avances de la tecnología de computadoras y comunicación, y para satisfacer los diferentes requerimientos futuros del usuario, es natural sugerir la mezcla de diferentes técnicas de conmutación en una red, con el fin de utilizar más eficientemente las facilidades de comunicación. Recientemente, se ha puesto un gran esfuerzo en la integración de CC y CP en una red híbrida, donde la capacidad de la red se comparte entre CC y CP, y los usuarios pueden escoger una manera apropiada para la transmisión de datos. Muchas investigaciones han mostrado que el sistema integrado es más eficiente y con menor costo que un sistema separado. Actualmente, una red híbrida comercial se ha lanzado al mercado [5,6], y con la tendencia de que el costo de procesamiento y de memoria sigue bajando más rápido que el costo de comunicación (ver Fig 1.5), un sistema de comunicación híbrida va a ser factible y de costo menor.

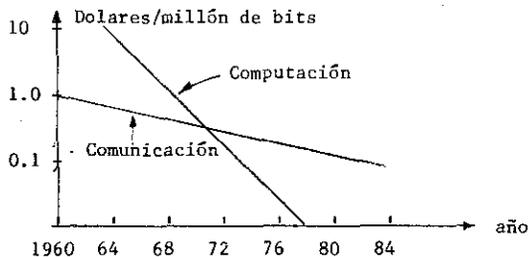
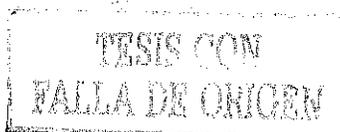


Fig. 1.5 Las tendencias de costo de computación y de comunicación

1.3 Una configuración de red híbrida-SENET

Muchos trabajos de investigación sobre redes híbridas están basados en el esquema SENET (Slotted Envelope Network). La idea fue originada por Coviello y Vena [7], y la implementación se discutió en [8] con más detalle. SENET es un esquema de multiplexaje en división de tiempo (TDM) en el cual el tiempo se divide en periodos fijos llamados marcos (frames), y cada marco se parte en dos regiones para transmisión de paquetes y voz. La región de voz es dividida en ranuras (slots) asignadas a cada conversación utilizando CC. Las ranuras reservadas a lo largo de la trayectoria en la red establecen una trayectoria virtual de CC entre los nodos de comunicación. La región



de paquetes del marco también se divide en varias ranuras para la transmisión de los paquetes de datos intermitentes utilizando CP. Los detalles de la estructura del marco se puede observar en Fig.1.6.

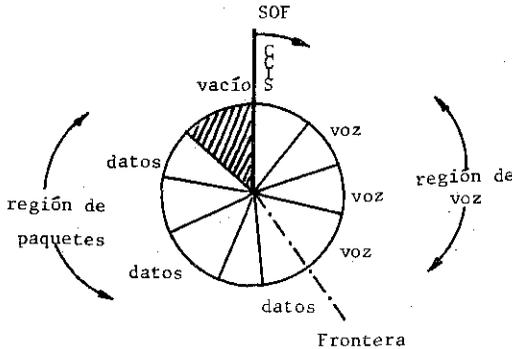


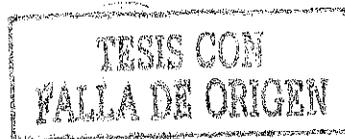
Fig.1.6 Estructura de marco de SENET

El funcionamiento de este esquema se puede explicar de la siguiente manera: empezando en la marca SOF (Start Of Frame), los tráficos de voz se asignan a las ranuras del marco en orden de sus llegadas. La asignación y la entrega de estas ranuras se realizan por medio de CCIS (Common Channel Interoffice Signaling). La marca de la frontera indica el fin de los tráficos de voz, si se requiere una nueva conexión, la región puede extenderse. De esta región, sigue la región de paquetes, y se asigna una o más ranuras para cada paquete de acuerdo de sus tamaños; si no hay más paquetes que se van a transmitir al siguiente nodo, la región final de este marco se encuentra vacía hasta que llegue otro marco. En caso de que no se pueda mandar todos los paquetes en un marco, se guarda en un buffer a esperar otro marco. La información de enrutamiento y de control del sistema pueden reservar una parte del marco, o también se puede transmitir como paquetes con mayor prioridad.

El buffer de marco puede aceptar una nueva entrada de marco después de copiar los paquetes al buffer del nodo siguiente, y copiar las ranuras de voz directamente al marco de la salida del nodo siguiente. Los paquetes pueden borrarse del buffer si es recibe un reconocimiento positivo.

1.4 Trabajos anteriores

Los importantes trabajos sobre redes de comunicación son bases de las redes híbridas. Las ideas de integrar los servicios de CC y CP en una red no son nuevas. Tal esquema, pero solamente para un enlace, fue propuesto y analizado por Otterman



en 1962 [9]. El interés en la investigación de redes híbridas había crecido desde el principio de la década 70, debido a las necesidades futuras y los avances tecnológicos. Los trabajos que han contribuido en este campo se pueden dividir básicamente en las siguientes ramas:

a) Comparación del costo y desempeño de CC y CP:

La comparación de CC y CP en los aspectos de costo y desempeño nos permite observar ventajas y desventajas de ellas, y la razón de la integración de ellas. Algunos detalles de este trabajo se presentarán en el Capítulo 4.

b) Realización del multiplexor y arquitectura de red híbrida:

En 1974, Zafiropulo propuso un esquema de multiplexor integrado y su realización utilizando el método de programas almacenados [10]. La idea del SENET fue originada por Coviello y Vena [7] y la implementación se discutió en [8]. En [11], se presentó una arquitectura de red híbrida, con algunos detalles de interface entre el sistema y los usuarios, y los protocolos de cada nivel.

c) Modelado y análisis de desempeño de los esquemas híbridos:

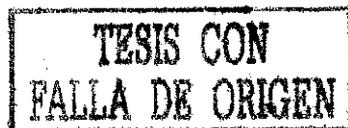
Esta rama es muy amplia, y depende de los esquemas propuestos. Para utilizar eficientemente los recursos de comunicación, se han propuesto muchos esquemas. El método de Frontera-Móvil (FM) apareció primeramente en [12]. El autor comparó el desempeño de este método con el de Frontera-Fija (FF) con algunas aproximaciones. Después de este, diferentes modelos se han presentado y analizado para FM, como el de Fisher y Harris [13], Alan [14] etc.. El modelo de Markov bidimensional fue desarrollado y analizado por Chang [3], Weinstein [15], Bhat [16] y Leon-Garcia [17]. Los resultados mostraron que FM es más eficiente que FF, pero también tiene algunos problemas, y estos problemas se pueden resolver de muchas maneras.

El esquema de Marco-Variable (MV) se analizó en varios trabajos como [18, 19, 20]. El desempeño de MV es mejor que FM, pero el sistema no es síncrono, y puede causar problemas en cuanto la realización.

En [21] se observó un fenómeno de que en alguna situación, la cola de paquetes aumenta fuertemente, y en este caso, las técnicas de aproximación con difusión y fluido son más adecuadas.

d) Enrutamiento en redes híbridas:

El problema de enrutamiento en redes híbridas es sumamente difícil. Mario Gerla propuso una medida de distancia entre dos nodos para CC y CP. Entonces, el problema se convierte en minimizar esta distancia con una tabla de enrutamiento [22].



e) Diseño óptimo de redes híbridas:

Como asignar las capacidades de cada enlace y para diferentes tráficos es un problema de diseño óptimo. En [23] se presentó un método para resolver este tipo de problemas. Dos criterios de optimalidad del sistema fueron discutidos en [24], y se apuntó que el sistema de FM es casi óptimo, bajo estos criterios, pero en un medio de tráficos previamente fijado.

Chang analizó las características de buffer en red híbrida, y sugirió usar una memoria secundaria para obtener un mejor desempeño de FM [3].

TESIS CON
FALLA DE ORIGEN

CAPITULO 2

HERRAMIENTAS PARA MODELAR SISTEMAS DE COMUNICACION

La teoría de colas y procesos estocásticos son las herramientas fundamentales para modelar y analizar sistemas de comunicación. Un modelo de sistema es una abstracción de la realidad que nos permite analizar esquemáticamente las características del sistema. Generalmente, un sistema de comunicación se puede modelar como un sistema de colas utilizando procesos estocásticos, pues la teoría de colas es campo bien estudiado y se puede resolver el problema con mayor facilidad. En muchos casos, se hacen algunas suposiciones para simplificar el modelo bajo la condición de que las suposiciones propuestas no afecten mucho al problema original. Un modelo de comunicación se puede resolver analíticamente o numéricamente, o también, si el sistema es bastante complicado, se puede usar el método de simulación.

2.1 Introducción a la teoría de colas

Un sistema de colas es un sistema de espera que consiste de servidores (o facilidades de servicio) y clientes. Para especificar completamente un sistema de colas, uno debe de identificar los procesos estocásticos que describen el flujo de llegadas de los clientes, así como la estructura y disciplina de las facilidades de servicio. Generalmente, se describe el proceso de llegadas del cliente en términos de la distribución de probabilidad de tiempo entre llegadas, y denotada $A(t)$, donde

$$A(t) = \Pr \{ \text{tiempo entre llegadas} \leq t \}$$

La suposición es que estos tiempos entre llegadas son variables aleatorias independientes, e idénticamente distribuidas (abreviada IID).

Otra cantidad estadística que se debe especificar es la demanda que se va a pedir al servidor, ésta se refiere al tiempo de servicio cuya distribución es $B(x)$, donde

$$B(x) = \Pr \{ \text{tiempo de servicio} \leq x \}$$

Aquí el tiempo de servicio se refiere a la duración de tiempo que gastará un cliente en la facilidad, y es conveniente suponer que $B(x)$ también es IID y además independiente de $A(t)$.

Respecto a la estructura de la facilidad, un factor importante es la capacidad de almacenamiento disponible para los clientes que esperan en la cola. Otra cantidad es el número de servidores o canales del sistema.



La disciplina de la cola describe la política o el orden en que el sistema atiende a los clientes de la cola. Las disciplinas pueden ser: FIFO (First-In First-Out), LIFO (Last-In First-Out), SIRO (Service In Random Orden), PRI (PRIority) o RR (Round-Robin) etc..

La descripción de un sistema de colas se puede abreviar con la siguiente notación:

A/B/m/K/F/d

Donde cada término tiene siguiente significado:

- A: distribución de tiempo entre llegadas.
- B: distribución de tiempo de servicio.
- m: número de servidores.
- K: capacidad de sistema (capacidad de almacenamiento + número de servidores).
- F: población de clientes.
- d: disciplina de servicio.

En caso de que $K=\infty$, $F=\infty$ y $d=\text{FIFO}$, se puede abreviar en forma A/B/m. Los términos de A y B pueden tomar las siguientes:

- M = exponencial (i.e. Markoviana)
- Er = Erlanguiana de r-etapas.
- Hr = Hiperexponencial de r-etapas
- D = Determinística
- G = distribución General

Un resultado muy importante en la teoría de colas, es la relación entre el número promedio de clientes en el sistema (N), la rapidez promedio de llegadas al sistema (λ) y el tiempo esperado que se tardan los clientes en el sistema (T). El resultado de Little [25] dice que:

$$N = \lambda T$$

La demostración de este resultado es muy simple y se puede ver [26].

2.2 Procesos estocásticos

Frecuentemente, se dice que la teoría de colas es una parte de la teoría de procesos estocásticos aplicada. En esta sección, introducimos algunos conceptos de procesos estocásticos, y analizamos algunos procesos importantes con el fin de modelar un sistema de comunicación híbrida.



2.2.1 Conceptos básicos

Considerese un sistema probabilístico (S, E, P) , que consiste de un espacio de ocurrencia S , un conjunto de eventos E (cada evento es un subconjunto de S), y las probabilidades que se asocian a estos eventos. A cada ocurrencia de un experimento $\zeta \in S$, se le asigna una función del tiempo $X(t, \zeta)$ real o compleja, de acuerdo a alguna regla. Entonces, la familia de funciones $X(t, \zeta)$, $\forall \zeta \in S$, forma un proceso estocástico.

Se puede observar que para un tiempo fijo t , $X(t, \zeta)$ es una variable aleatoria, y un proceso estocástico es una colección de variables aleatorias dependiendo de t . Usualmente $X(t)$ representa un proceso estocástico omitiendo ζ [27].

Con el propósito de especificar completamente un proceso estocástico $X(t)$, se define una función de distribución de probabilidad para cada t permisible:

$$F_x(x; t) \triangleq \Pr\{X(t) \leq x\}.$$

Además, se define una distribución conjunta para $\{t_1, t_2, \dots, t_n\}$ permisible:

$$\begin{aligned} F_x(x; t) &\triangleq F_{x_1 x_2 \dots x_n}(x_1, x_2, \dots, x_n; t_1, t_2, \dots, t_n) \\ &= \Pr\{X_1(t_1) \leq x_1, \dots, X_n(t_n) \leq x_n\}. \end{aligned}$$

La función de densidad de probabilidad está definida por:

$$f_x(x; t) = \frac{\partial F_x(x; t)}{\partial x}$$

El valor promedio de un proceso estocástico está dado por:

$$\overline{X(t)} = E\{X(t)\} = \int_{-\infty}^{\infty} x f_x(x; t) dx$$

La autocorrelación de $X(t)$ es:

$$R_{xx}(t_1, t_2) = E\{X(t_1)X(t_2)\} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} x_1 x_2 f_{x_1 x_2}(x_1, x_2; t_1, t_2) dx_1 dx_2.$$

La clasificación de los procesos estocásticos depende de tres parámetros: el espacio de estados, el tiempo (o índice) y las dependencias estadísticas entre las variables aleatorias $X(t)$ para diferentes tiempos.

a) Espacio de estados:

El conjunto de posibles valores (estados) que toma $X(t)$ forma un espacio de estados. Si los estados de $X(t)$ son finitos o contables, se dice que $X(t)$ es un proceso estocástico de estados-discretos o simplemente una cadena. Si $X(t)$ toma los valores en un intervalo continuo o infinito, se dice que $X(t)$ es un proceso



estocástico de estados-continuos.

b) Tiempo (o índice):

Si los tiempos permisibles de $X(t)$ son finitos o contables, se dice que $X(t)$ es de tiempo-discreto, y se escribe X_n en lugar de $X(t)$. Si los tiempos permisibles están dentro de un intervalo continuo ya sea finito o infinito, $X(t)$ es un proceso de tiempo-continuo.

c) Dependencia estadística:

La clasificación en base de las dependencias estadísticas depende del tipo y el grado de las dependencias, aquí no describimos en detalle.

Algunos procesos estocásticos que están muy relacionados con la teoría de colas se menciona en la siguiente:

1) Procesos estacionarios:

Se dice que un proceso $X(t)$ es estacionario si $F_x(x;t)$ es tal que

$$F_x(x;t+\Delta t) = F_x(x;t) \quad \text{donde } t+\Delta t = (t_1+\Delta t_1, \dots, t_n+\Delta t_n).$$

Además, si $E\{X(t)\}$ es independiente de t y $E\{X(t)X(t+\epsilon)\}$ depende únicamente de ϵ (es decir, $R_{xx}(t_1, t_2) = R_{xx}(t_1 - t_2)$), se dice que $X(t)$ es estacionario en sentido amplio.

2) Procesos independientes:

Uno de los más simples y triviales procesos estocásticos a considerar es la secuencia estocástica en el cual $\{X_n\}$ forma un conjunto de variables aleatoria independientes, o sea:

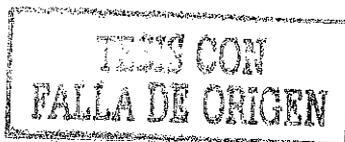
$$f_x(x;t) = f_{x_1 \dots x_n}(x_1 \dots x_n; t_1 \dots t_n) = f_{x_1}(x_1, t_1) \dots f_{x_n}(x_n, t_n).$$

Para el caso continuo, se puede definir un proceso estocástico independiente y eso se refiere comúnmente a "ruido blanco".

3) Procesos Markovianos:

Esta clase de procesos es la más importante y fundamental para la teoría de colas. Un proceso Markoviano de estados discretos también se llama cadena de Markov. Una cadena de Markov de tiempo-discreto es fácil de conceptualizar y entender.

Un conjunto de V.A. $\{X_n\}$ forma una cadena de Markov si la probabilidad de que el siguiente valor o estado X_{n+1} es x_{n+1} depende únicamente del estado presente X_n y no de cualquier estado anterior. Esto es, la única manera de que la historia pasada afecta al futuro del proceso está resumido en el estado actual del proceso.



Matemáticamente:

$$\Pr \{X_{n+1}=x_{n+1} \mid X_n=x_n, \dots, X_1=x_1\} = \Pr \{X_{n+1}=x_{n+1} \mid X_n=x_n\}.$$

En caso de tiempo-continuo, la cadena de Markov tiene la siguiente propiedad:

$$\Pr \{X(t_{n+1})=x_{n+1} \mid X(t_n)=x_n, \dots, X(t_1)=x_1\} = \Pr \{X(t_{n+1})=x_{n+1} \mid X(t_n)=x_n\}.$$

para $t_1 < t_2 < t_3 < \dots < t_n < t_{n+1}$

Se puede comprobar que para cada estado dado, si las distribuciones de tiempo de que el proceso permanece en estos estados son exponenciales, entonces el proceso es Markoviano [26].

4) Procesos de nacimiento-muerte:

Una clase importante y especial de las cadenas de Markov es el proceso de nacimiento-muerte. Ellos pueden ser de tiempo-discreto o continuo, y la condición es que las transiciones de estados solamente ocurran entre los estados adyacentes (vecinos inmediatos). En caso de tiempo-discreto, se tiene:

$$\Pr \{X_{n+1}=j \mid X_n=i\} = \begin{cases} P_{ij}^{(n+1, n)} & \text{si } j=i-1 \text{ o } i \text{ o } i+1 \\ 0 & \text{caso contrario} \end{cases}$$

Aquí se supone que los estados toman los números enteros $0, 1, 2, \dots, M$. El mayor detalle se presenta en las siguientes secciones.

2.2.2 Cadena de Markov discreta

En la sección anterior hemos definido la cadena de Markov discreta donde $P(X_n=j \mid X_{n-1}=i)$ es la probabilidad de transición del estado E_i al estado E_j en tiempo n . En esta sección presentamos algunos resultados importantes.

DEFINICION 1 Una cadena de Markov discreta es homogénea si las probabilidades de transición no depende del tiempo. Es decir,

$$P(X_n=j \mid X_{n-1}=i) = P_{ij}.$$

La P_{ij} es la probabilidad de transición del estado E_i al estado E_j de la cadena en una etapa. Se puede generalizar para m etapas:

$$P_{ij}^{(m)} = P(X_{n+m}=j \mid X_n=i).$$

Una propiedad de la cadena Markoviana es:

$$P_{ij}^{(m)} = \sum_k P_{ik}^{(m-1)} P_{kj}, \text{ para } m > 1.$$



DEFINICION 2 Una cadena de Markov discreta se dice que es irreducible si todos los estados son alcanzables de cualquier otro estado. Es decir, para cada par de estados E_i y E_j , existe un número entero m tal que:

$$P_{ij}^{(m)} > 0.$$

Antes de caracterizar los estados, es conveniente definir las siguientes cantidades probabilísticas:

$$f_j^{(n)} \triangleq \Pr\{\text{primera vez que regresa a } E_j \text{ después de } n \text{ etapas, partir de } E_j\}$$

y

$$f_j \triangleq \sum_{n=1}^{\infty} f_j^{(n)} = \Pr\{\text{regresa a } E_j \text{ en primera vez}\}.$$

DEFINICION 3 Un estado E_j de una cadena de Markov discreta se puede clasificar de las siguientes maneras:

- a) Recurrente, si $f_j = 1$.
- b) Transitorio, si $f_j < 1$.

DEFINICION 4 Un estado E_j de una cadena de Markov se dice que es periódico con periodo d_j si d_j es el máximo común divisor del conjunto

$$Q_j \triangleq \{n \mid P_{jj}^{(n)} > 0\}.$$

Si $d_j = 1$, se dice que E_j es aperiódico.

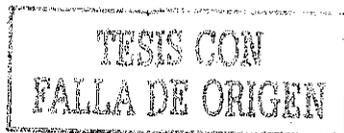
Ahora, considerando los estados recurrentes cuales $f_j = 1$, se define el tiempo promedio de recurrencia de E_j como:

$$M_j \triangleq \sum_{n=1}^{\infty} n f_j^{(n)}.$$

DEFINICION 5 Un estado recurrente de una cadena de Markov E_j se dice que es recurrente nulo si $M_j = \infty$, y es recurrente no-nulo si $M_j < \infty$.

Si definimos $\Pi_j^{(n)}$ como la probabilidad de que el sistema se encuentra en el estado E_j en n -ésima etapa, ésto es: $\Pi_j^{(n)} \triangleq P(X_n = j)$, tenemos el siguiente teorema.

TEOREMA 1 En una cadena de Markov homogénea, irreducible y aperiódica, los límites de las probabilidades



$$\pi_j = \lim_{n \rightarrow \infty} \pi_j^{(n)}, \text{ para todos } j$$

siempre existen y son independientes de la distribución de probabilidad del estado inicial. Además:

- a) Todos los estados son transitorios o todos son recurrentes nulos en tal caso $\pi_j = 0, \forall j$ y no existe una distribución estacionaria. 0
- b) Todos los estados son recurrentes no-nulos y que $\pi_j > 0, \forall j$, en tal caso el conjunto $\{\pi_j\}$ es una distribución estacionaria con: $\pi_j = \frac{1}{M_j}$.

En este caso la π_j se puede determinar con las siguientes ecuaciones:

$$\sum_i \pi_i = 1$$

$$\sum_i \pi_i P_{ij} = \pi_j.$$

DEFINICION 6 Una cadena de Markov es ergódica si todos los estados son aperiódicos y recurrentes no-nulos. O bien, $\pi_j = \lim_{n \rightarrow \infty} \pi_j^{(n)}, \forall j$ y no depende de $\pi(0)$.

Las probabilidades $\{\pi_j\}$ también son llamadas probabilidades de equilibrio. En caso de que tenga n estados finitos, se tiene la siguiente fórmula matricial por el teorema 1:

$$\Pi = \Pi P.$$

donde P es la matriz de transición, y Π es un vector de las probabilidades de equilibrio.

2.2.3 Cadena de Markov continua

DEFINICION 7 Un proceso estocástico X(t) forma una cadena de Markov de tiempo continuo si para cualquier n y cualquier secuencia t_1, t_2, \dots, t_{n+1} , tal que $t_1 < t_2 < \dots < t_{n+1}$, se tiene

$$\Pr \{ X(t_{n+1})=j | X(t_n)=i_n, \dots, X(t_1)=i_1 \} = \Pr \{ X(t_{n+1})=j | X(t_n)=i_n \}$$

Otra forma de interpretar esto es:

$$\Pr \{ X(t)=j | X(\tau) \text{ para } \tau_1 < \tau \leq \tau_2 < t \} = \Pr \{ X(t)=j | X(\tau_2) \}$$

TESIS CON
FALLA DE ORIGEN

Con la misma analogía que en el caso discreto, se define la probabilidad de transición para la cadena de Markov continua (no-homogénea):

$$P_{ij}(s,t) \stackrel{\Delta}{=} \Pr \{X(t)=j | X(s)=i\}, \quad t \geq s$$

Una propiedad importante (ecuación de Chapman-Kolmogorov) es:

$$P_{ij}(s,t) \stackrel{\Delta}{=} \sum_k P_{ik}(s,u) P_{kj}(u,t) \quad \text{para } s < u < t$$

La forma matricial de esta ecuación es:

$H(s,t) \stackrel{\Delta}{=} H(s,u)H(u,t)$ donde $s < u < t$, y $H(s,t) \stackrel{\Delta}{=} P_{ij}(s,t)$ con $H(t,t)=I$, la matriz de identidad.

Para analizar la cadena de Markov continua, se define la siguiente matriz:

$$P(t) \stackrel{\Delta}{=} P_{ij}(t,t+\Delta t), \quad \text{donde } \Delta t \text{ es suficientemente pequeña.}$$

Y observese:

$$\lim_{\Delta t \rightarrow 0} \frac{H(s,t+\Delta t) - H(s,t)}{\Delta t} = \lim_{\Delta t \rightarrow 0} \frac{H(s,t)H(t,t+\Delta t) - H(s,t)}{\Delta t} = H(s,t) \lim_{\Delta t \rightarrow 0} \left(\frac{P(t) - I}{\Delta t} \right)$$

Por lo tanto:

$$\frac{\partial H(s,t)}{\partial t} = H(s,t) Q(t), \quad t \geq s \quad \text{siendo } Q(t) = \lim_{\Delta t \rightarrow 0} \left(\frac{P(t) - I}{\Delta t} \right) \quad \dots (2.1)$$

La matriz $Q(t)$ se conoce como el generador infinitesimal de la función de $H(s,t)$, la matriz de transición. O también se llama la matriz de rapidez de transición. Los elementos de $Q(t)$ denotados $q_{ij}(t)$ son:

$$q_{ii}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{ii}(t,t+\Delta t) - 1}{\Delta t}$$

$$q_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{P_{ij}(t,t+\Delta t)}{\Delta t}, \quad i \neq j.$$

Similarmente, se puede calcular $\frac{\partial H(s,t)}{\partial s}$ por:

$$\begin{aligned} \frac{\partial H(s,t)}{\partial s} &= \lim_{\Delta s \rightarrow 0} \frac{H(s+\Delta s,t) - H(s,t)}{\Delta s} = \lim_{\Delta s \rightarrow 0} \frac{H(s+\Delta s,t) - P(s)H(s+\Delta s,t)}{\Delta s} \\ &= \lim_{\Delta s \rightarrow 0} \left(\frac{I - P(s)}{\Delta s} \right) H(s,t) = -Q(s)H(s,t) \quad \dots (2.2) \end{aligned}$$

Siendo $H(s,t) = P(s)H(s+\Delta s,t)$.

TESIS CON
FALLA DE ORIGEN

Es fácil de comprobar que $\sum_j q_{ij}(t) = 0, \forall i$.

La solución de las ecuaciones (2.1) y (2.2) es:

$$H(s,t) = \exp \left\{ \int_s^t Q(u) du \right\}, \text{ donde } e^{At} = I + At + \frac{A^2 t^2}{2!} + \dots$$

Ahora, definimos la probabilidad de equilibrio en tiempo t como

$$\Pi_j(t) \stackrel{\Delta}{=} \Pr \{X(t)=j\} \quad \text{o el vector de probabilidad de equilibrio}$$

$$\Pi(t) = [\Pi_0(t), \Pi_1(t), \Pi_2(t), \dots]$$

Si la distribución de probabilidad inicial de cada estado es dada como $\Pi(0)$, entonces se puede resolver $\Pi(t)$ con la siguiente fórmula:

$$\Pi(t) = \Pi(0) H(0,t) = \Pi(0) \exp \left\{ \int_0^t Q(u) du \right\} \quad \dots \dots \dots (2.3)$$

La ecuación diferencial de (2.3) es:

$$\frac{d\Pi(t)}{dt} = \Pi(t) Q(t), \quad \text{o bien}$$

$$\frac{d\Pi_j(t)}{dt} = q_{jj}(t) \Pi_j(t) + \sum_{k \neq j} q_{kj}(t) \Pi_k(t)$$

Los resultados anteriores se pueden aplicar al caso de una cadena de Markov homogénea, con las siguientes notaciones:

$$P_{ij}(t) \stackrel{\Delta}{=} P_{ij}(s, s+t)$$

$$q_{ij} \stackrel{\Delta}{=} q_{ij}(t)$$

$$H(t) \stackrel{\Delta}{=} H(s, s+t) = [P_{ij}(t)]$$

$$Q \stackrel{\Delta}{=} Q(t) = [q_{ij}]$$

Y resulta:

$$\frac{dH(t)}{dt} = H(t)Q = QH(t), \quad \text{con } H(0)=I$$

$$H(t) = e^{Qt}$$

$$\frac{d\Pi(t)}{dt} = \Pi(t) Q, \quad \text{o bien}$$

$$\frac{d\Pi_j(t)}{dt} = q_{jj}\Pi_j(t) + \sum_{k \neq j} q_{kj}\Pi_k(t).$$

Si además la cadena de Markov homogénea es irreducible, entonces podemos observar que los siguientes límites siempre existen y son independientes de los estados

iniciales de la cadena:

$$\lim_{t \rightarrow \infty} P_{ij}(t) = \Pi_j$$

El conjunto $\{\Pi_j\}$ forma el límite de las probabilidades de los estados. Para una cadena de Markov ergódica, se tiene los siguientes límites y son independientes de la distribución inicial:

$$\lim_{t \rightarrow \infty} \Pi_j(t) = \Pi_j$$

Estos límites de la distribución están dados únicamente como la solución de las siguientes ecuaciones lineales:

$$q_{jj} \Pi_j + \sum_{k \neq j} q_{kj} \Pi_k = 0, \quad \text{o bien} \quad \Pi Q = \underline{0}$$

donde $\Pi = \{\Pi_0, \Pi_1, \Pi_2, \dots\}$ y $\sum_i \Pi_i = 1$.

2.3 Algunos sistemas de colas importantes

Esta sección describe los sistemas de colas de tipo nacimiento-muerte en equilibrio, cual considera que cada tiempo que un cliente entra al sistema de colas, se puede considerar como un nacimiento en la población, donde la población consiste de todos los clientes presentes en el sistema. Cuando un cliente recibe un servicio, se sale del sistema, y eso se refiere a una muerte de un miembro de la población. La solución de equilibrio del sistema es lo que vamos a presentar, para diferentes tipos de colas.

2.3.1 M/M/1 : el sistema de colas clásico

M/M/1 es un sistema de colas de un servidor, con llegadas Poissonianas (tiempo entre llegadas exponencial), y tiempo de servicio exponencial. El diagrama de rapidez de transición entre los estados se muestra en siguiente:



donde $\lambda =$ rapidez de llegadas ($\lambda^{-1} =$ promedio de tiempo entre llegadas)

$\mu =$ rapidez de servicio ($\mu^{-1} =$ promedio de tiempo de servicio).

Con los conocimientos de la teoría de procesos estocásticos presentados en la sección anterior, se puede resolver fácilmente el problema. Las probabilidades de equilibrio de cada estado están dadas por :

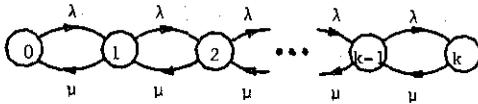


$$P_k = \frac{\rho^k}{\sum_{j=0}^{\infty} \rho^j} \quad \text{para } k=0,1,2,\dots$$

donde $\rho =$ carga de tráfico o intensidad de tráfico $= (\frac{\lambda}{\mu})$, y se supone $\rho < 1$.

2.3.2 M/M/1/K : capacidad finita

El mismo sistema de colas que M/M/1, excepto que la capacidad del sistema (tamaño K) es finita. Su diagrama de rapidez de transición es:



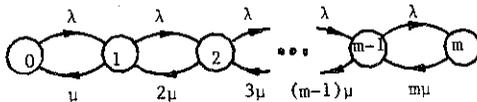
La solución es:

$$P_i = \frac{\rho^i}{\sum_{j=0}^k \rho^j}, \quad i=0,1,2,\dots,k$$

Este tipo de colas se puede utilizar para modelar un nodo de la red de comunicación con buffer finito.

2.3.3 M/M/m/m : sistema de pérdida con m-servidores

Este sistema es un modelo típico de un sistema telefónico, es decir, cuando los m líneas (o servidores) están ocupadas, una nueva llamada se bloquea. El diagrama de rapidez de transición es:



El resultado es fácil de obtener:

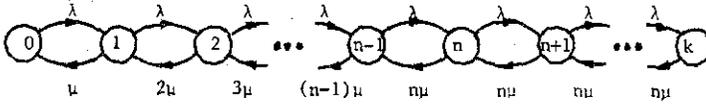
$$P_k = \frac{(\frac{\lambda}{\mu})^k / k!}{\sum_{j=0}^m (\frac{\lambda}{\mu})^j / j!} \quad k = 0,1,2,\dots,m$$

donde P_m es la probabilidad de que m servidores estén ocupados, o sea, la probabilidad de que entra una nueva llamada y se bloquea. Comúnmente, a P_m se le llama la probabilidad de bloqueo de Erlang.

2.3.4 M/M/n/K :

Este sistema de colas tiene n servidores disponibles para los clientes, los cuales se caracterizan por llegadas Poisson y tiempo de servicio requerido exponencial. La capacidad de espera es de tamaño K-n. El diagrama de rapidez de transición de este

sistema se muestra en la siguiente:



Resolviendo las siguientes ecuaciones de equilibrio:

$$P_0 \lambda = P_1 \mu$$

$$P_1 (\lambda + \mu) = P_0 \lambda + P_2 (2\mu)$$

⋮

$$P_{n-1} (\lambda + (n-1)\mu) = P_{n-2} \lambda + P_n (n\mu)$$

$$P_n (\lambda + n\mu) = P_{n-1} \lambda + P_{n+1} (n\mu)$$

⋮

$$P_{k-1} (\lambda + k\mu) = P_{k-2} \lambda + P_k (k\mu)$$

Se puede obtener las probabilidades de los estados permanentes:

$$P_j = \begin{cases} P_0 \frac{\rho^j}{j!} & j \leq n \\ P_0 \frac{\rho^j}{n! n^{j-n}} & j > n \end{cases}$$

donde $\rho = \lambda / \mu$

$$P_0 = \left\{ \sum_{j=0}^n \left(\frac{\rho^j}{j!} \right) + \sum_{j=n+1}^k \frac{\rho^j}{n! n^{j-n}} \right\}^{-1}$$

TESIS CON
FALLA DE ORIGEN

CAPITULO 3

UN MULTIPLEXOR DE CONMUTACION HIBRIDA CON FRONTERA MOVIL BIDIRECCIONAL (FM2)

La red híbrida (en especial el esquema de SENET), puede manejar dos tipos de tráfico: el tráfico como voz que usa CC (denotado CC-tráfico), y el tráfico de datos intermitentes que usa CP (denotado CP-tráfico). El recurso de comunicación es la capacidad de los enlaces, la cual se divide en marcos a lo largo del tiempo, y cada marco consiste de varias ranuras ya sean de mismo tamaño o distinto. Este esquema de transmisión es típicamente un multiplexaje en división de tiempo (TDM). Evidentemente, estos tráficos van a competir por los recursos limitados, y entonces, la pregunta es: ¿ como distribuir mejor los recursos para estos tráficos competitivos ? El problema asociado a la red híbrida es el diseño de la estrategia de multiplexaje 28 . Entre los trabajos de investigación dedicados a redes híbridas, muchos se han concentrado en la búsqueda de una estrategia de multiplexaje, para proporcionar una alta eficiencia de transmisión. Hasta ahora, algunas estrategias y sus variaciones se han propuesto, como la de Frontera-Móvil (FM) y Marcos-Variantes (MV).

El análisis de estos esquemas de multiplexaje generalmente se efectúa para un enlace de comunicación, debido a la complejidad de un enfoque global de la red. En este capítulo presentamos un nuevo esquema de multiplexaje, así como el modelo para evaluar el desempeño del sistema. Antes de esto, es necesario describir los esquemas de FM y MV, para dar un panorama de los trabajos hechos hasta ahora y para facilidad de comparación.

3.1 Las estrategias de multiplexaje

Basado en el importante concepto de SENET, han sido propuestas las estrategias de Frontera-Móvil y Frontera-Fija con marco fijo. Esto es: en un enlace de SENET, un marco de multiplexaje síncrono en división de tiempo se divide en N ranuras con dos regiones. La primera región (las primeras S ranuras) se dedican al CC-tráfico utilizando la conmutación clásica por circuitos, mientras la segunda región (las restantes (N-S) ranuras) son reservadas para el CP-tráfico utilizando la conmutación por paquetes convencional. La frontera entre las dos regiones se dice que es fija, si la primera región se puede usar únicamente por el CC-tráfico, y la segunda para CP-tráfico exclusivamente. En el caso de FM (o FM unidireccional para distinguir el esquema propuesto en este trabajo), el CP-tráfico puede ocupar toda la capacidad restante después de acomodar el CC-tráfico en la primera región (nótese que la frontera solo se mueve para un sentido). El CC-tráfico se maneja como un sistema de



espera con pérdida y se asigna a la primera región en base de First-come First-served en el inicio de cada periodo de marcos. Es decir, si las S ranuras de la primera región están ocupadas, las nuevas llegadas de CC-tráfico al sistema son bloqueadas. Pero, una vez se asigna una ranura para un CC-tráfico, permanece el uso hasta que termina la sesión. El CP-tráfico se trata de otra manera: en caso de que no hay ranuras disponibles para una nueva llegada de CP-tráfico, se guarda en el buffer del nodo de transmisión, y se espera hasta que se disponga de una ranura para su transmisión.

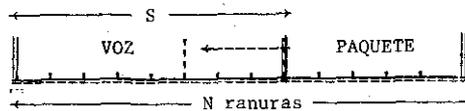


Fig. 3.1 Estructura de marco de Frontera-Móvil

Muchos resultados muestran que la estrategia de FM mejora el desempeño de CP-tráfico (el tiempo de retardo) con respecto a la estrategia de FF, pues la FM aprovecha las capacidades que no está usando el CC-tráfico (en lugar de la FF, que la desperdicia). Prácticamente, el sistema de FF es equivalente a dos sistemas separados que utilizan CC y CP respectivamente.

Otro esquema de multiplexaje es el de Marco-Variable, donde la longitud del marco puede variar de un mínimo de S ranuras a un máximo de N ranuras. Las primeras S ranuras también están reservadas para CC-tráfico, y el CP-tráfico puede usar toda la capacidad restante después de acomodar el CC-tráfico. En la situación en que el CP-tráfico es pequeño, podrían existir algunas ranuras desocupadas en la segunda región. El esquema de MV evita el desperdicio de la capacidad de marco en esta situación con un marco variable. Una condición simple para volver a iniciar un nuevo marco (Start of frame: SOP) después de cruzar la frontera, es que el buffer de paquetes de datos esté vacío. Obviamente este esquema es asíncrono, y tiene el problema de que es difícil sincronizar la transmisión de datos en la red.

En la siguiente sección, presentamos un nuevo esquema que es una combinación de las ideas de FM y MV. Es decir, este esquema puede evitar los desperdicios de capacidades en la situación de que el CC-tráfico o el CP-tráfico sea pequeños, con un sistema síncrono.

TESIS CON
FALLA DE ORIGEN

3.2 Protocolo del esquema FM

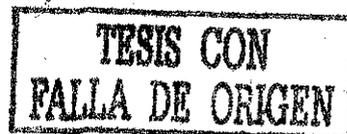
Por simplicidad y para permitir una comparación directa con los esquemas anteriores, se supone que el marco de STDM (Synchronous Time Division Multiplexing) es de b segundos, y se divide en N ranuras, todas del mismo tamaño. La rapidez de digitalización de voz (si el CC-tráfico consiste en voz digitalizada) y la capacidad de transmisión de enlaces están diseñadas tal que una simple conexión de voz conmutada por circuito puede acomodarse en una ranura.

El esquema de Frontera-Móvil bidireccional (FM2) consiste en una frontera virtual (de hecho, en la implementación no aparece) que divide el marco en dos regiones. La primera región (que contiene S ranuras) son reservadas inicialmente para CC-tráfico, mientras la segunda región (las restantes $N-S$ ranuras) son reservadas inicialmente para CP-tráfico. Una frontera fija está colocada en la segunda región en posición S_1 (donde $S < S_1 < N$), la cual es el límite hasta donde puede mover la frontera hacia la región de paquetes.

El esquema propuesto FM2 puede operar en uno de los dos modos en cualquier tiempo: modo-1 (la frontera mueve hacia la primera región) y modo-2 (la frontera mueve hacia la segunda región). La Fig. 3.2 muestra la estructura del marco de FM2 y sus dos modos de operación. Si, por ejemplo, las sesiones de CC-tráfico en tiempo t en el sistema $N_V(t) < S$, entonces el sistema opera en el modo-1, donde CP-tráfico puede usar toda la capacidad desocupada por CC-tráfico de la primera región. Se nota que el modo-1 es esencialmente idéntico que el esquema FM unidireccional. Si el número de CC sesiones en el sistema excede o igual S , entonces el FM2 multiplexor entra a operar en modo-2. En este modo, una nueva petición de CC conexión puede capturar una ranura desocupada en la segunda región, si el número total de CC-tráfico y CP-tráfico en el sistema en este momento no excede S_1 . Esto es, si $N_V(t) + N_D(t) < S_1$, la frontera puede moverse hacia la región de paquetes. Durante la operación en el modo-2, el sistema maneja los dos tipos de tráfico en una disciplina de FIFO (first in first out) y no-prevacio (nonpreemptive). Es decir, una vez asignada una CC sesión en el marco, se mantenga hasta que termine su comunicación.

Las $(N-S_1)$ ranuras del marco son reservadas exclusivamente para CP-tráfico, con el fin de garantizar el mínimo requerido del desempeño del CP-tráfico.

El CC-tráfico se maneja en base a pérdida. Es decir, si en algún tiempo una llamada no encuentra ninguna ranura disponible para ella ($N_V(t) \geq S$ y $N_V(t) + N_D(t) \geq S_1$), se pierde. Aquí suponemos que una llamada puede esperar hasta un máximo de una duración de marco. El CP-tráfico se maneja en base de espera (retardo), o sea, si algunos paquetes no se pueden transmitir en sus llegadas, se guarda en un buffer para esperar servi-



cios. El tamaño del buffer está seleccionado para una muy baja probabilidad de sobreflujo (overflow).

El protocolo del esquema FM2 se puede describir en un diagrama de flujo que se muestra en Fig. 3.3.

3.3 Modelo

Chang [1] y Weinstein et al [15] han mostrado que el sistema de conmutación híbrida puede ser modelado como una cadena de Markov ergódica bidimensional, si ignoramos el efecto de cuantización de tiempo introducido por la estructura de marco, y suponiendo que ambos tráficos tienen llegadas Poissonianas y servicios exponenciales.

El esquema de FM2, también se puede modelar como una cadena de Markov ergódica bidimensional, donde suponemos que las llegadas de CC-tráfico son en forma de un proceso Poissoniano, con promedio de rapidez de llegadas λ_v llamadas/seg., y las llegadas de CP-tráfico son en forma Poissoniana, con medio λ_d paquetes/seg.. También suponemos que ambos tráficos tienen servicios exponenciales, con medio μ_v^{-1} y μ_d^{-1} , respectivamente. Donde $\mu_d^{-1} = b/N$ seg, e ignoramos el efecto de cuantización de tiempo por la estructura del marco.

Las suposiciones son realistas en algún sentido, pues los tráficos que llegan a un nodo particular (superposición de los tráficos que llegan de otros nodos y los de este nodo) son de forma aleatoria y casi independientes. Entonces, el proceso que caracteriza mejor este tipo de llegadas es el proceso de Poisson. Generalmente, la rapidez de llegadas de paquetes es mucho mayor que la de voz, y el tiempo de servicio requerido por los paquetes es mucho menor que una conversación en voz. Por otra parte, si la duración del marco es muy pequeña, el efecto de cuantización de tiempo se puede despreciar.

El estado del modelo está dado por $\{V(t), D(t)\}$, $0 \leq V(t) \leq S_1$, y $0 \leq D(t) \leq m$, donde $V(t)$ es el número de CC sesiones en el sistema en tiempo t , y $D(t)$ es el número de paquetes en el sistema en t .

El diagrama de rapidez de transición de los estados se muestra en la Fig. 3.4. La rapidez de transición que sale del estado (i,j) es $(\lambda_d + K_4 \lambda_v + i \mu_v + K_2 \mu_d)$, y las rapideces que entran al estado (i,j) son $K_3 \lambda_v$, λ_d , $(i+1) \mu_v$, $K_1 \mu_d$ de los estados $(i-1,j)$, $(i,j-1)$, $(i+1,j)$, $(i,j+1)$, respectivamente, donde las K_i , $i=1,2,3,4$, están dadas por:

$$K_1 = \begin{cases} N-i & \text{si } j \geq N-i \\ j+1 & \text{si } j < N-i \end{cases}$$

$$K_2 = \begin{cases} N-i & \text{si } j \geq N-i \\ j & \text{si } j < N-i \end{cases}$$



$$K_3 = \begin{cases} 1 & \text{si } i \leq S \text{ o } i+j \leq S_1 \\ 0 & \text{caso contrario} \end{cases}$$

$$K_4 = \begin{cases} 1 & \text{si } i < S \text{ o } i+j < S_1 \\ 0 & \text{caso contrario} \end{cases}$$

La cadena de Markov tiene características de nacimiento-muerte, porque las transiciones entre los estados ocurren únicamente entre estados vecinos.

3.4 Solución en estado permanente

Por la ergodicidad de la cadena de Markov del modelo, se puede definir:

$$P_{ij} = \lim_{t \rightarrow \infty} \Pr\{V(t)=i, D(t)=j\}$$

como la probabilidad del estado permanente de que haya i llamadas y j paquetes en el sistema. Bajo las condiciones de los estados permanentes (estacionarios), se puede obtener una ecuación de equilibrio para cada estado de la cadena. Esto es:

$$P_{ij} (\lambda_d + K_4 \lambda_v + i \mu_v + K_2 \mu_d) - P_{i-1,j} (K_3 \lambda_v) - P_{i,j-1} (\lambda_d) - P_{i+1,j} (i+1) \mu_v - P_{i,j+1} (K_1 \mu_d) = 0 \quad \dots \dots \dots (3.1)$$

para $0 < i < S_1$ y $0 < j < m$.

Para los estados sobre fronteras, algunos términos de la ecuación anterior desaparecen, y tienen la siguiente forma:

$$\begin{aligned} P_{ij} (\lambda_d + K_4 \lambda_v) - P_{i+1,j} (i+1) \mu_v - P_{i,j+1} (K_1 \mu_d) &= 0 && \text{para } i=0, j=0 \\ P_{ij} (K_2 \mu_d + K_4 \lambda_v) - P_{i+1,j} (i+1) \mu_v - P_{i,j-1} (\lambda_d) &= 0 && \text{para } i=0, j=m \\ P_{ij} (\lambda_d + i \mu_v) - P_{i-1,j} (K_3 \lambda_v) - P_{i,j+1} (K_1 \mu_d) &= 0 && \text{para } i=S_1, j=0 \\ P_{ij} (K_2 \mu_d + i \mu_v) - P_{i-1,j} (K_3 \lambda_v) - P_{i,j-1} (\lambda_d) &= 0 && \text{para } i=S_1, j=m \\ P_{ij} (K_2 \mu_d + \lambda_d + K_4 \lambda_v) - P_{i,j-1} (\lambda_d) - P_{i+1,j} (i+1) \mu_v - P_{i,j+1} (K_1 \mu_d) &= 0 && \text{para } i=0, 0 < j < m \\ P_{ij} (\lambda_d + i \mu_v + K_2 \mu_d) - P_{i-1,j} (K_3 \lambda_v) - P_{i,j-1} (\lambda_d) - P_{i,j+1} (K_1 \mu_d) &= 0 && \text{para } i=S_1, 0 < j < m \\ P_{ij} (\lambda_d + K_4 \lambda_v + i \mu_v) - P_{i-1,j} (K_3 \lambda_v) - P_{i+1,j} (i+1) \mu_v - P_{i,j+1} (K_1 \mu_d) &= 0 && \text{para } 0 < i < S_1, j=0 \\ P_{ij} (K_4 \lambda_v + i \mu_v + K_2 \mu_d) - P_{i-1,j} (K_3 \lambda_v) - P_{i,j-1} (\lambda_d) - P_{i+1,j} (i+1) \mu_v &= 0 && \text{para } 0 < i < S_1, j=m. \end{aligned} \quad (3.2)$$

La forma matricial de las ecuaciones anteriores está dada por:



$$\Pi Q = \underline{0} \quad \dots (3.3)$$

donde $\Pi = [P_{k_{ij}}]$, el vector de renglón de las probabilidades de los estados

permanentes, y $k_{ij}=K(i,j)$ una transformación uno a uno ($K: (i,j) \rightarrow k$).

Q = la matriz de rapidez de transición entre los estados de la cadena.

$\underline{0}$ = vector de renglón de ceros.

Debido a la naturaleza del nacimiento-muerte de las transiciones de los estados, la matriz Q es rara. En otras palabras, solamente una banda (con una máxima anchura $2S_1$) de elementos sobre la diagonal de Q puede ser diferentes de cero. Entonces es suficiente almacenar estas rapidezces de transición cuando use la matriz Q para obtener la solución.

Existen algunos métodos para resolver eficientemente este tipo de ecuación. Un método es aproximación sucesiva que consiste en la siguiente fórmula recursiva [26]:

$$\Pi^{(k+1)} = \alpha \Pi^{(k)} Q + \Pi^{(k)}$$

donde $\Pi^{(k)}$ es la k -ésima iteración, y α es un escalar. Si $\Pi^{(k+1)} = \Pi^{(k)}$, entonces $\Pi^{(k)} Q = \underline{0}$ y $\Pi^{(k)}$ es una buena aproximación de las probabilidades de los estados permanentes.

La dificultad de este método es cómo seleccionar el escalar α , o específicamente, para nuestro modelo, con un α fijo no se garantiza la convergencia del método, o la velocidad de convergencia es muy lenta.

En el presente trabajo, utilizamos un método "directo" para resolver el sistema de ecuación (3.3). Este método está basado en la eliminación de Gauss: el programa implementado procesa la matriz Q columna por columna, y solamente almacena la columna resultante (con máximo S_1 elementos). claro que una columna de Q será dependiente de las demás, pues podemos quitar esta columna y utilizar la ecuación adicional de $\sum_k P_k = 1$. En la práctica, es mejor tomar inicialmente $\Pi' = [P'_1, P'_2, \dots, 1]$ y se evalúa las P'_k , $k=(S_1+1)(m+1), \dots, 1$ sucesivamente con la matriz resultante Q' después de la eliminación. Finalmente, las probabilidades de los estados permanentes del sistema se obtiene de la siguiente fórmula:

$$P_k = P'_k / \Pi' e, \quad \text{donde } e = [1, 1, \dots, 1]^T$$

Algunos pasos importantes de este método están ilustrados en la Fig. 3.5. El programa que resuelve esta ecuación fue computacionalmente eficiente. Por ejemplo, para un sistema de $N=20$, $S_1=15$, $S=5$ y $m=200$, sólo 6 segundos de tiempo de CPU se

**TESIS CON
FALLA DE ORIGEN**

requiere en la Burroughs B0900, usando Pascal.

3.5 Evaluación de desempeño

Primero, definamos las medidas de desempeño del sistema. Para el proceso de CC-tráfico, una posible medida de desempeño puede ser la probabilidad de bloqueo de las llamadas. Es decir, la razón entre el número de llamadas bloqueadas acumuladas en el sistema y el número total de llamadas intentadas, en el intervalo de tiempo $[0, t]$ cuando $t \rightarrow \infty$. En nuestro modelo, es exactamente la suma de las probabilidades de aquellos estados que ya no pueden aceptar llamadas, es decir:

$$PB_v = \sum_{i,j \in R} P_{k_{ij}} \quad \text{donde } R = \{(i,j) \mid i+j \leq S_i \text{ y } i \leq S_j\}$$

El tiempo de retardo de CC-tráfico también puede ser una medida de desempeño para el sistema. Pero, consideramos que el tiempo de retardo de CC-tráficos es la combinación de dos partes: tiempo para establecer la comunicación (es casi cero si utiliza la conmutación por circuito con pérdidas), y el retardo de propagación (es fijo si dada la duración de comunicación). Por la razón mencionada, no evaluamos el tiempo de retardo de CC-tráfico.

Para el proceso de CP-tráfico, una medida muy importante es el tiempo promedio de retardo de los paquetes (\bar{T}_d = tiempo de espera más el tiempo de servicio). Para calcular este tiempo promedio, es conveniente calcular primero el número promedio de paquetes en el sistema (\bar{N}_d), y utilizar la fórmula de Little. Es decir:

$$\bar{T}_d = \frac{\bar{N}_d}{\lambda_d}$$

donde λ_d es la tasa de llegadas de CP-tráfico, y \bar{N}_d se calcula de la siguiente forma:

$$\bar{N}_d = \sum_j (j \sum_i P_{k_{ij}})$$

La variancia del número de paquetes en el sistema también se puede calcular como:

$$\text{Var}\{N_d\} = \sum_j [(j - \bar{N}_d)^2 \sum_i P_{k_{ij}}]$$

Hay que mencionar que el tiempo de retardo \bar{T}_d no es el tiempo de retraso total, porque el tiempo de retraso total es igual a la suma de \bar{T}_d y la distancia entre la ranura en que ocupa este paquete y el inicio del siguiente marco (SOF).

Ahora, si la capacidad del sistema para el CP-tráfico es finita ($m < \infty$), entonces existe la posibilidad de que se pierdan los paquetes, debido a la insuficiencia



de la memoria del buffer. Esta probabilidad está dada simplemente por:

$$PFB = \sum_i P_{k_{im}}$$

Para el sistema, una medida de desempeño es el promedio de la utilización de la capacidad del canal, la cual es igual a la razón del número promedio de ranuras usadas por los tráficos en un intervalo de marco y el número total de ranuras del marco (N). O sea,

$$UC = \sum_{ij} C_{ij} P_{k_{ij}} / N, \quad \text{donde} \quad C_{ij} = \begin{cases} i+j & \text{si } i+j \leq N \\ N & \text{si } i+j > N \end{cases}$$

Como las probabilidades de los estados estacionarios $P_{k_{ij}}$ están dadas, estos desempeños del sistema se pueden calcular fácilmente.

3.6 Análisis y comparación

En esta sección, ponemos énfasis en la comparación de los sistemas que usan el esquema propuesto FM2 y el esquema de FM anterior. Los tráficos que vamos a considerar tienen una duración promedio de las llamadas $\mu_V^{-1} = 100$ seg., y los paquetes tiene un promedio de 1ms de longitud (exponencialmente distribuida).

Primero, suponemos que ambos esquemas coloquen sus fronteras en la misma posición (S) en el marco (ya que los dos esquemas pueden diseñar diferentes S en el marco). Además, el esquema de FM2 tiene una S_1 que permite a su frontera moverse hacia la región de CP-tráfico.

Para un sistema con los parámetros $N=20$, $S=10$, $S_1=17$, $m=100$, las probabilidades de bloqueo del CC-tráfico y las utilidades del canal están mostradas en la Fig. 3.6, con respecto a diferentes cargas de CC-tráfico (ρ_V), dado que la carga de CP-tráfico (ρ_d) es de 5 Erlangs. Los tiempos de retardo del CP-tráfico de ambos esquemas están en la Tabla 1. Se puede observar que para una carga grande de voz, el mejoramiento en la probabilidad de bloqueo de voz y la utilización del canal para el esquema FM2 es significativo, mientras el incremento de tiempo de retardo de paquetes es muy pequeño, comparado con FM.

Para el mismo parámetro del sistema anterior, el comportamiento de ambos sistemas con respecto a ρ_d , con una $\rho_V = 7$ Erlangs, se muestra en la Fig. 3.7 y la Tabla 2. La probabilidad de bloqueo para FM es constante (igual a la probabilidad de bloqueo de Erlang), debido a una carga de voz fija. Por otra parte, la probabilidad de bloqueo del sistema de FM2 varía con respecto a ρ_V , porque para una carga pequeña de paquetes, algunas llamadas pueden usar la capacidad reservada inicialmente para paquetes. Mien

tras ρ_d crece, el desempeño de ambos sistema tiende a ser mismo.

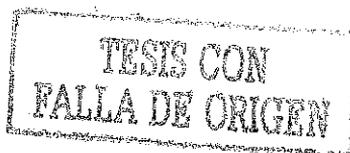
La Fig. 3.8 y la Tabla 3 muestran el efecto de S_1 sobre el desempeño de FM2, para $\rho_v=5$ Erlangs y dos diferentes ρ_d , en un sistema de $N=15$, $S=5$, $m=100$. Se nota que si S_1 tiende a N , la probabilidad de bloqueo y la utilización del canal mejora bastante, pero el incremento del tiempo de retardo de paquetes de FM2 también es mucho (ver Tabla 3). Una explicación es que al incrementar la zona accesible de los CC-tráficicos, lo cual algunas llamadas de voz son permitidas de usar las capacidades temporalmente desocupadas por el CP-tráfico, y en muchas ocasiones, los paquetes que llegan después tendrán que esperar toda la duración de estas conversaciones. Por lo tanto, una selección apropiada de S_1 es importante para un buen funcionamiento del multiplexor de FM2.

Un requerimiento crítico para un multiplexor de conmutación híbrida es la robustez a la variación de las cargas de los tráficos. Desafortunadamente, el esquema FM no es tan satisfactorio en este aspecto. En la Fig. 3.9, consideramos un caso más realista que consiste en una suma constante de las cargas de los dos tráficos (para $N=20$, $\rho_v+\rho_d=14$), y las cargas varían sujetas a esta condición. La comparación de probabilidad de bloqueo con respecto a S nos indica que, el esquema FM2 es más adaptivo que el FM en un medio de tráficos variables.

El efecto del tamaño del buffer para paquetes (m) sobre el desempeño del sistema de FM2 está mostrado en la Tabla 4. La probabilidad de sobreflujo se reducirá a menos de 10^{-4} con un tamaño de $m=50$. Prácticamente, el desempeño del sistema no es sensible al tamaño del buffer dentro de un rango razonable.

Hasta aquí, hemos analizado y comparado los desempeños de FM2 y FM unidireccional, considerando tomar la misma posición de frontera. Los resultados mostraron ser favorables para el esquema propuesto FM2, comparado con FM. Ahora, en el resto de esta sección, consideramos que los dos esquemas pueden tomar diferentes colocaciones de la frontera (claro que favorece a ellos mismos). Las comparaciones están basadas en un sistema de $N=20$, $m=100$ y $\rho_v+\rho_d=\text{constante}=13.785$ Erlangs.

Suponemos que las cargas de CC-tráfico son variables en un intervalo de tiempo (puede ser un día o una semana), pero con un valor promedio $\bar{\rho}_v=6$. Lo mismo para el tráfico de paquetes, cual también varía en el mismo periodo de tiempo y tiene un promedio de $\bar{\rho}_d=7.785$ Erlangs. Para que la probabilidad de bloqueo de voz de ambos sistemas sea igual y debaja de un nivel fijo (por ejemplo, para $PB_v < 4.3\%$), el esquema FM puede (y debe) colocar su frontera en $S=10$, mientras FM2 puede seleccionar una $S=9$ y $S_1=19$ para satisfacer el mismo requerimiento de la probabilidad de bloqueo predefinido. Ahora, como los tráficos de este ambiente son variables (por ejemplo, cada



hora del día o cada día de la semana, tienen un cambio notable), el desempeño del sistema que obtenemos por medio de los valores promedios de tráficos no es tan representativo de un caso real, lo más justo es ver globalmente el comportamiento del sistema con todos los posibles casos de tráficos.

En la Fig. 3.10 graficamos una curva de incremento (o decremento) de retardo de paquetes de FM2 sobre FM, y dos curvas de probabilidades de bloqueo de voz de ambos sistemas, para diferentes cargas de tráficos. Una observación importante es que cuando $\rho_v < 6$ (o $\rho_d > 7.785$), FM2 tiene menor tiempo de retardo de paquetes, y la probabilidad de bloqueo de voz es un poco mayor que el esquema FM; Cuando $\rho_v > 6$ (o $\rho_d < 7.785$), el sistema que usa FM2 tiene un poco aumento de retardos de paquetes, pero la probabilidad de bloqueo es mucho menor que FM. En promedio, desde el punto de vista global del sistema, FM2 puede operar en este ambiente con una probabilidad de bloqueo de voz mucho menor que FM (mantenga debajo de un nivel aceptable), mientras el tiempo de retardo de paquetes es casi igual que FM (quizá un poco mayor).

El mejoramiento del desempeño del sistema cuando usa FM2 en un medio de tráfico variable es debido a la adaptividad de este esquema, cual permite la frontera moverse en dos direcciones, dependiendo de las características de los tráficos. Se puede observar que el esquema FM2 no es tan sensible a la variación de los tráficos, comparado con el esquema de Frontera-Móvil unidireccional.

TESIS CON
FALLA DE ORIGEN

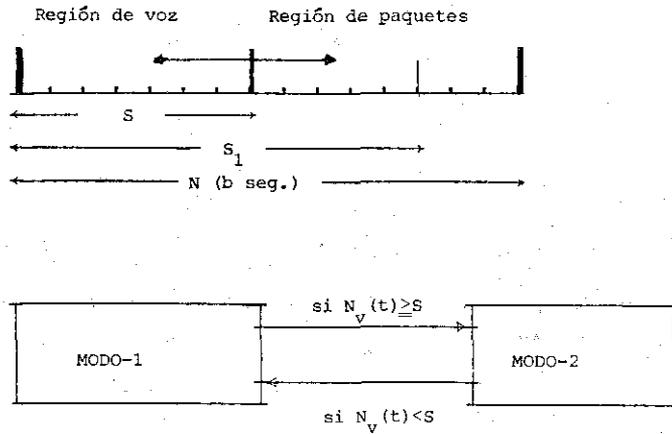


Fig. 3.2 La estructura del marco de Frontera-Móvil bidireccional y sus dos modos de operación.

TESIS CON
FALLA DE ORIGEN

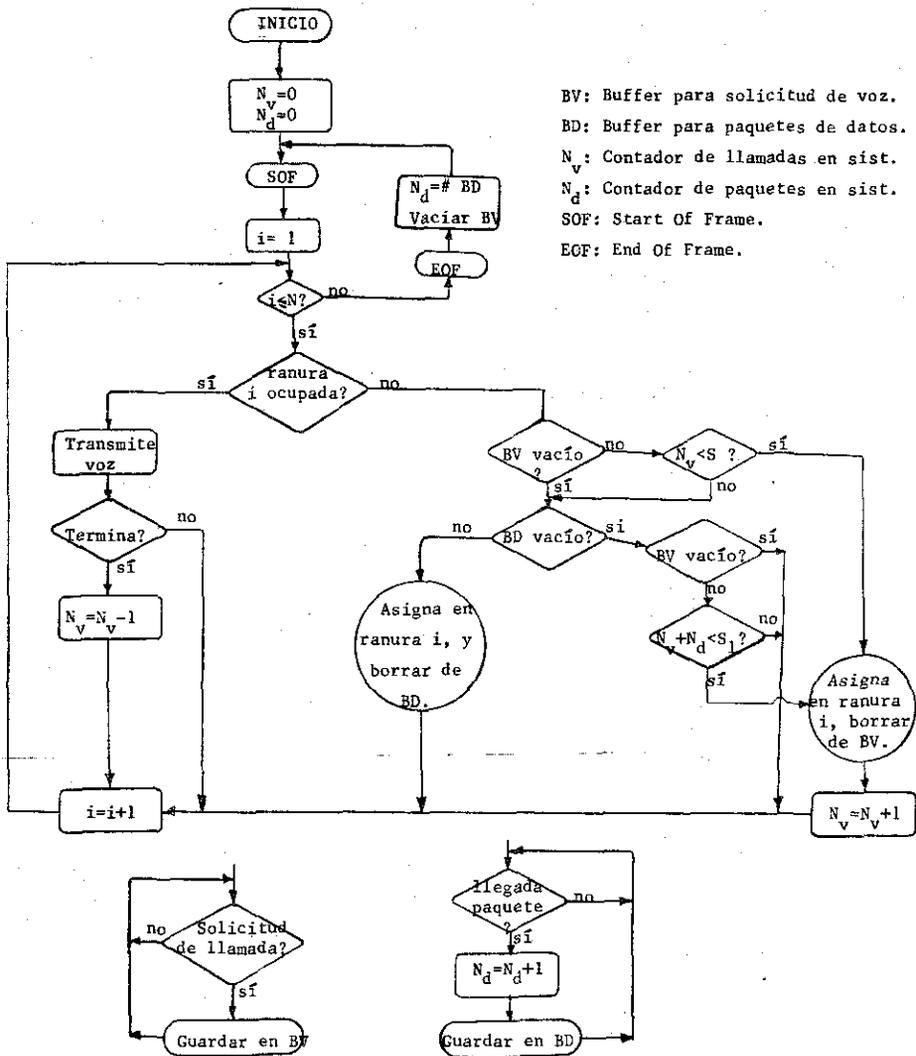


Fig. 5.3 Protocolo del esquema FN2 (Frontera Móvil Bidireccional).

TESIS CON
 FALLA DE ORIGEN

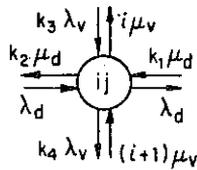
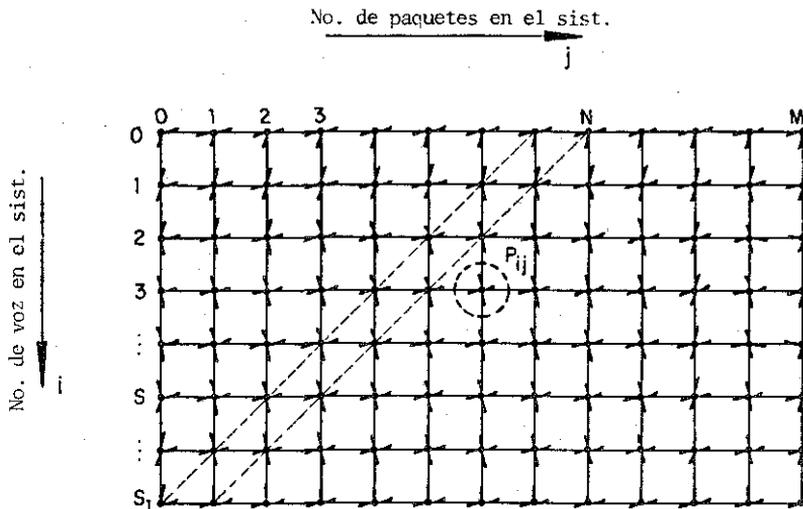


Fig. 3.4 Modelo de cadena de Markov bidimensional para FM2

TESIS CON
FALLA DE ORIGEN

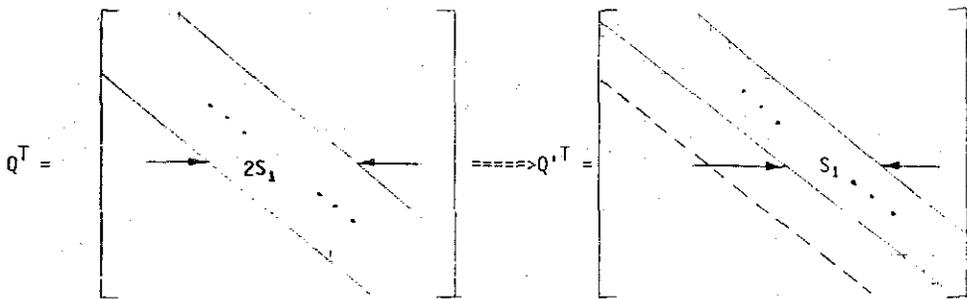


Fig.3.5 Las estructuras de Q y Q' (después de la eliminación Gaussiana), para resolver el sistema de ecuaciones $\pi Q = \underline{0}$.

TESIS CON
FALLA DE ORIGEN

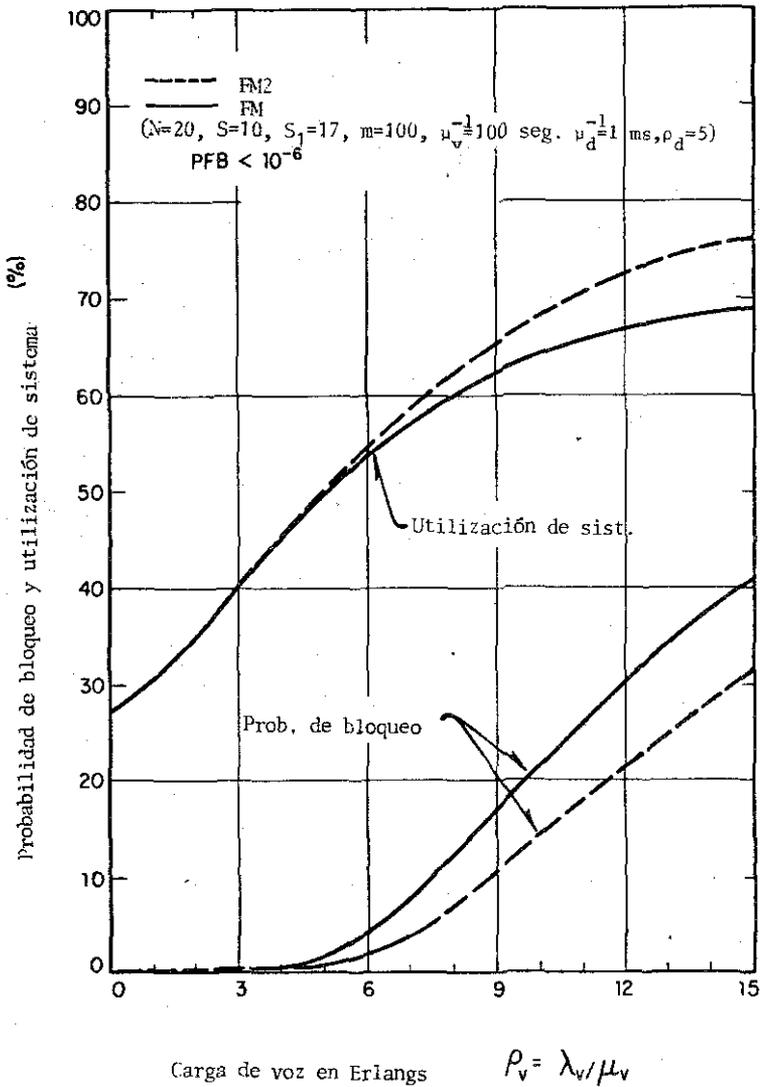


Fig. 3.6 Los comportamientos de FM y FM2, en función de las cargas de voz.

TESIS CON
 FALLA DE ORIGEN

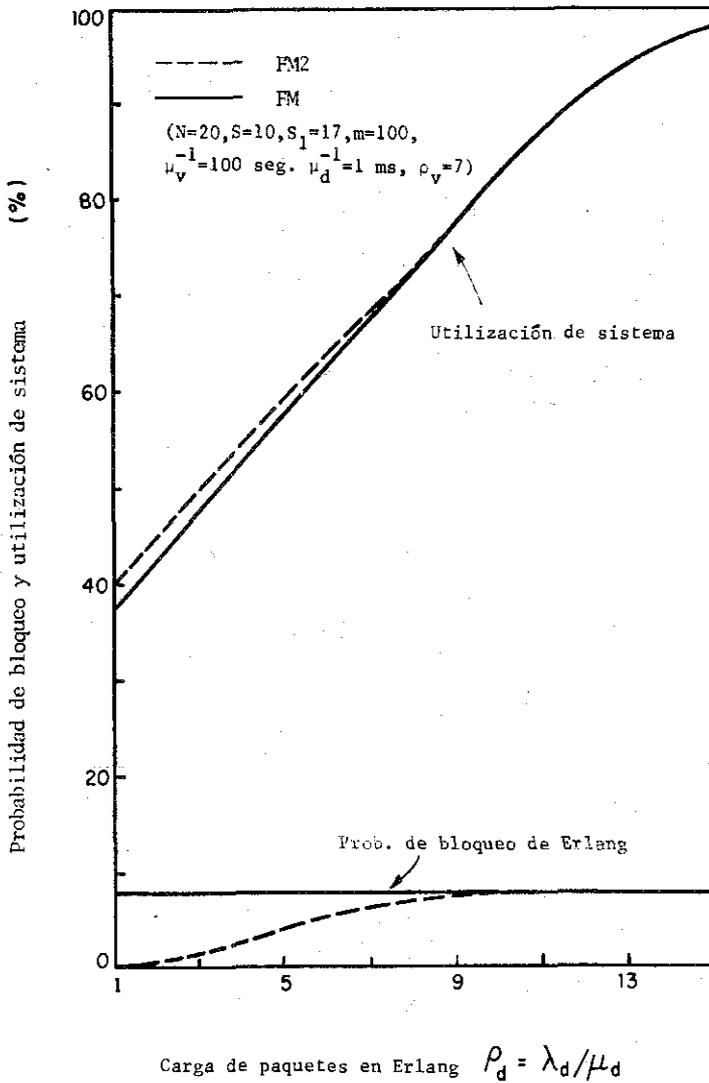


Fig. 3.7 Comparaciones de FM2 y FM, con respecto a la variación de la carga de paquetes.

TESIS CON
 FALLA DE ORIGEN

44

TESIS CON FALTA DE ORIGEN

Fig. 3.8 Variación de probabilidad de bloqueo y utilización de sistema, con respecto a S_1 , para $\rho_d=4$ y $\rho_d=6$.

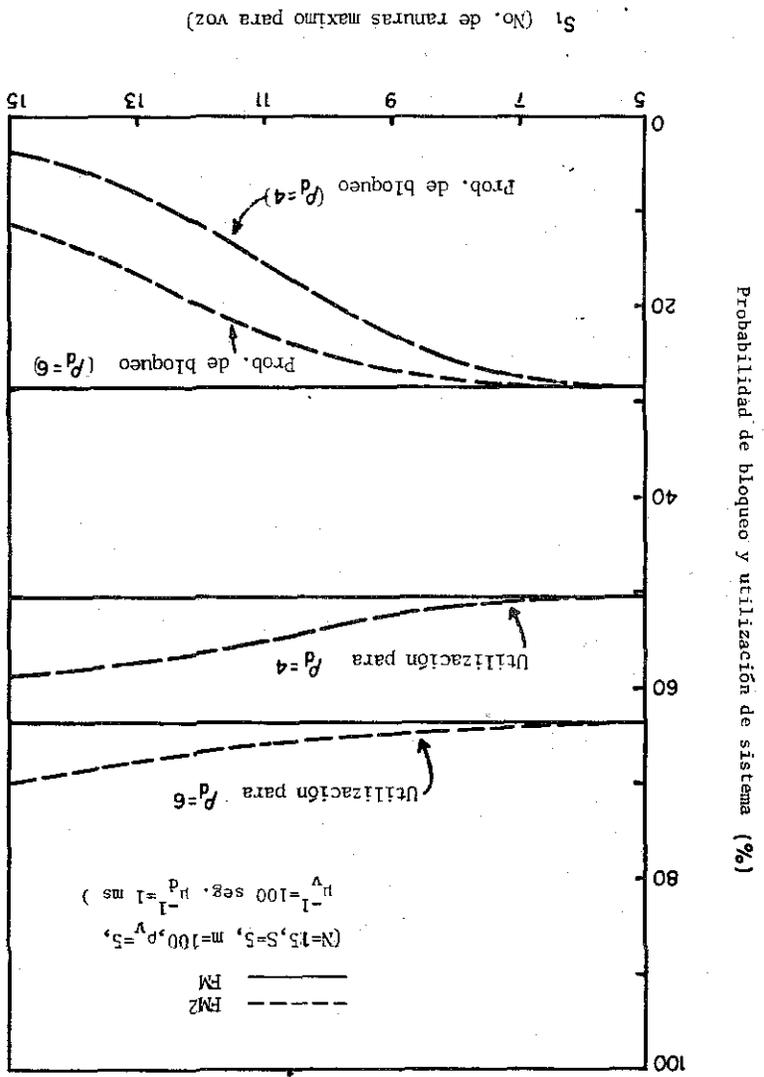


Tabla 1

Tiempos de retardo de paquetes para el sistema de Fig.3.6 (ms)

ρ_v	3	5	6	8	10	12	13	14	15
FM2	1.0000	1.0006	1.0015	1.0057	1.0140	1.0267	1.0347	1.0437	1.0538
FM	1.0000	1.0003	1.0006	1.0014	1.0023	1.0030	1.0033	1.0035	1.0038

Tabla 2

Tiempos de retardos y probabilidad de overflow del buffer de paquetes para el sistema considerado en Fig.3.7

ρ_d		1	3	5	7	9	11	13	15
FM2	T_d (ms)	0.9999	1.0004	1.0032	1.0189	1.1268	2.1624	3.7399	4.9544
	PFB	1.78×10^{-12}	3.8×10^{-9}	9.2×10^{-8}	1.4×10^{-6}	5.6×10^{-5}	0.0082	0.048	0.119
FM	T_d (ms)	1.0000	1.0000	1.0010	1.0124	1.1004	2.1624	3.7399	4.9544
	PFB	3.1×10^{-13}	9.1×10^{-12}	1.3×10^{-10}	10^{-9}	3.7×10^{-7}	0.0082	0.048	0.119

Tabla 3

Tiempo de retardo de paquetes para el sistema de Fig.3.8(ms)

S_1	6	7	8	9	10	11	12	13	14	15
$\rho_d=6$	1.0110	1.0113	1.0118	1.013	1.015	1.020	1.028	1.045	1.0736	1.109
$\rho_d=4$	1.0000	1.0004	1.0008	1.0012	1.0019	1.0032	1.0056	1.0111	1.022	1.037

TESIS CON
FALLA DE ORIGEN

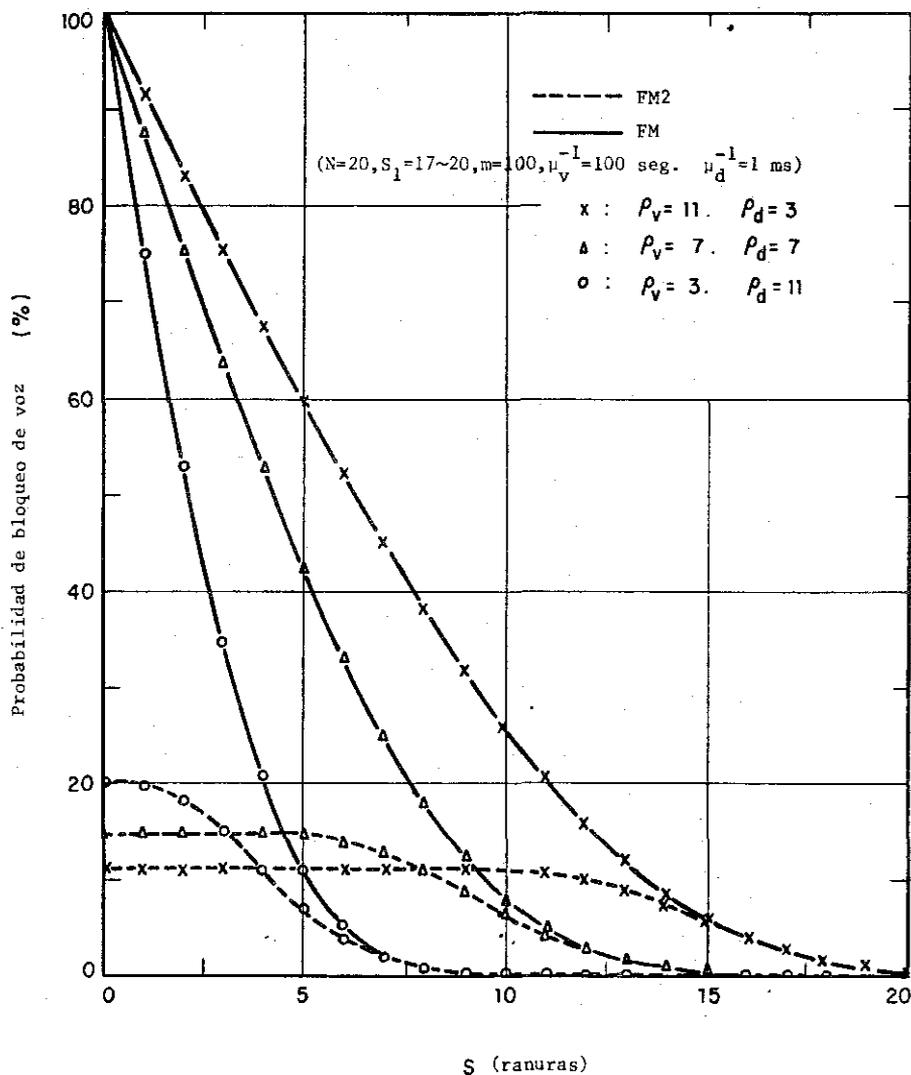


Fig. 3.9 Variaciones de probabilidad de bloqueo de voz, para diferentes S, en tres casos de tráficos ($\rho_v + \rho_d = 14$ Erlangs).

TESIS CON
 FALLA DE ORIGEN

Tabla 4
 Efecto de tamaño de buffer de paquetes sobre FM2
 ($N=15, S=8, S_1=14, \rho_v=0.04/0.01, \rho_d=6000/1000$ Erlangs)

m	PB _v (%)	\bar{N}_d	\bar{T}_d (ms)	PFB	\bar{UC} (%)
16	2.44	6.172	1.028	2.85×10^{-3}	65.90
20	2.49	6.231	1.039	9.70×10^{-4}	65.96
30	2.53	6.289	1.048	2.46×10^{-4}	65.98
50	2.54	6.339	1.057	9.30×10^{-5}	65.98
100	2.55	6.435	1.072	4.32×10^{-5}	65.99
200	2.55	6.613	1.102	2.05×10^{-5}	65.99
500	2.55	6.999	1.166	6.41×10^{-6}	65.99
1000	2.55	7.050	1.175	1.43×10^{-6}	66.00

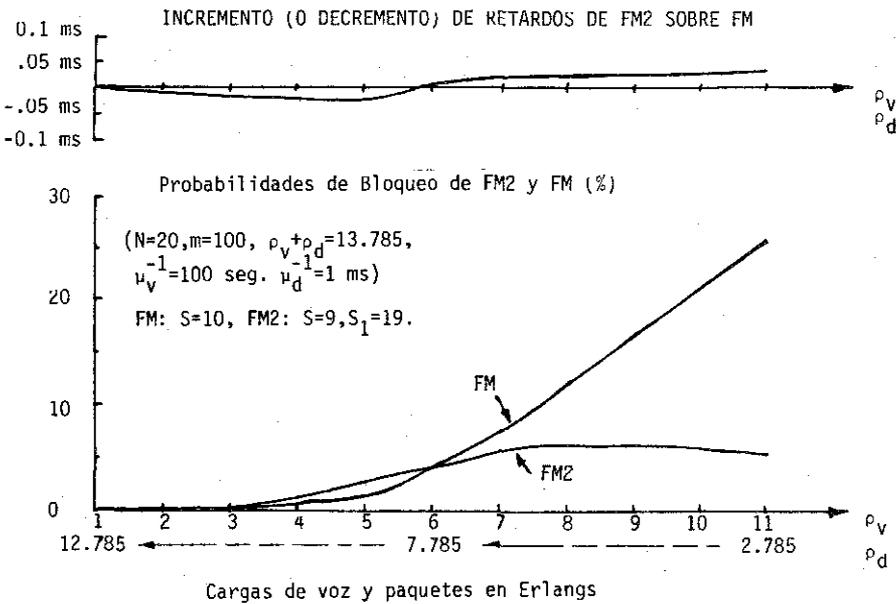


Fig.3.10 Comparación de FM2 y FM, en un medio de tráficos variables, donde los dos esquemas toman diferentes colocaciones de frontera en el marco.

TESIS CON
 FALLA DE ORIGEN

CAPITULO 4

ANALISIS DE UN MULTIPLEXOR INTEGRADO PARA TRES TIPOS DE TRAFICOS

4.1 Introducción

En el mundo en que vivimos, el humano tiene que procesar, transmitir, almacenar y mantener diversos tipos de información. Especialmente, en el campo de redes de computadora o de comunicación, la transmisión de datos es un problema de mayor interés. La información o los datos que circulan en una red de comunicación, dependiendo de sus características, se pueden clasificar en siguientes tipos:

Tipo I: caracterizado por transacciones largas que requieren una respuesta continua y en tiempo real (como la de voz, video y facsímil etc.).

Tipo II: caracterizado por transacciones cortas y discretas que requieren una respuesta casi en tiempo real (como los datos intermitentes o interactivos).

Tipo III: caracterizado por transacciones largas y continuas pero no necesariamente requieren una respuesta inmediata (por ejemplo, algunas transferencias de archivos grandes).

y otros que no consideramos aquí.

La red híbrida que hemos estudiado en los capítulos anteriores, está diseñada para los tráficos de tipo I y II, utilizando las técnicas de conmutación CC y CP respectivamente. Ahora, ¿que tipo de conmutación será recomendable para el tráfico del tipo III en una red "super-integrada"?

Conforme con las técnicas disponibles de CC o CP, no sería una solución eficiente y práctica. Por ejemplo, si usamos la CP para el tráfico del tipo III, las desventajas son:

- 1) Como el tamaño de un archivo (se supone que el tráfico del tipo III es transferencia de archivos grandes) generalmente es muy grande (hasta millones de bits), y como la CP lleva bastante información adicional (el encabezado de un paquete de tamaño máximo 967 bits en ARPANET es de 192 bits), la eficiencia de transmisión es baja.
- 2) El proceso de reensamble y reordenamiento para los paquetes de un archivo grande que llegan independientemente, a veces es un trabajo duro.
- 3) Existe posibilidad de que algunos paquetes se pierdan, debido a la insuficiencia de la memoria de buffer en el proceso de Store-and-Forward. Eso podría

TESIS CON
FALLA DE ORIGEN

costar mucho para la recuperación de la información perdida.

- 4) El costo de almacenamiento y mantenimiento de un gran número de paquetes es alto.

Itoh et al [29] sugirió que para los mensajes con más de 4000 bits es mejor manejar en CC, y para los mensajes cortos CP es preferible.

Por otro lado, si utilizamos CC de tipo pérdida, hay muchos problemas cuando la solicitud de circuito no se acepta inmediatamente y el usuario tiene que intentar nuevamente. Es decir, hay dificultad en encontrar inmediatamente todos los canales libres en la ruta seleccionada. Como y es la mayoría de los casos, las transferencias de archivos grandes no son tan urgentes, y pueden esperar mucho tiempo, la técnica de conmutación por circuitos de tipo espera será una opción adecuada.

La comparación de tiempo de retraso de CC tipo espera y de CP fue presentada por Kummerle y Rudin [30], basados de un modelo que consiste de una línea de tres nodos conectados. La Figura 4.1 es la variación de tiempo de retraso para ambas técnicas, con respecto a la utilización de información ρ_I (la fracción de la capacidad total utilizada por la información entre la capacidad total). Se puede observar que para un amplio rango de ρ_I (< 0.6) la CC con espera tiene menor tiempo de retraso. En la Fig.4.2, se compara el tiempo de retardo de ambas técnicas, en función de la longitud de los mensajes (para un $\rho_I = 0.5$ fija). Se puede observar que para una mayor longitud de mensajes, la CC es mucho mejor que CP.

El modelo de costo que se propuso en [30] (solamente el costo de utilización) resultó en que: CC es favorable cuando el mensaje es largo y continuo para una distancia corta.

Basado en las observaciones anteriores, proponemos un esquema de multiplexaje en un sistema integrado para tres tipos de tráfico, el cual consiste en utilizar las técnicas de conmutación de CC tipo espera, CC tipo pérdida y CP, para los tráficos que tienen las características de archivo, voz y paquetes de datos respectivamente. En la siguiente sección, describimos en detalle la operación de este tipo de multiplexor integrado para los tres tipos de tráficos, así como el modelo para analizar el sistema.

Evidentemente, en un sistema integrado con tres tipos de conmutación que mencionamos anteriormente, los usuarios de la red tendrán más opciones para escoger una forma adecuada para la transmisión de sus mensajes.

4.2 Modelo

Supongamos que el marco TDM de la Fig.4.3 es de duración b segundos, y cada



marco consiste en N ranuras de igual tamaño. El marco se divide en tres partes: la primera parte (que contiene N_v ranuras) se dedica a la comunicación de voz utilizando CC tipo pérdida, mientras la segunda parte (que contiene N_f ranuras) está dedicada a las transferencias de archivos utilizando CC tipo espera, con un buffer de tamaño m_f para las solicitudes de la transferencia de archivos. La tercera parte (las restantes $(N - N_v - N_f)$ ranuras) está reservada para los mensajes cortos utilizando CP, el tamaño de buffer para los paquetes es m_d , donde m_d puede ser finito o infinito. Los paquetes también pueden usar las ranuras desocupadas por voz y archivos (nótese que este esquema es similar al esquema de Frontera-Móvil, pero para tres tipos de tráfico).

Los tráfico de voz, archivo y paquetes se caracterizan por llegadas Poissonianas con tasas de llegadas λ_v , λ_f y λ_d respectivamente. Los tiempos de servicios requeridos de los tráfico de voz y archivos son de forma exponencial con medias μ_v^{-1} y μ_f^{-1} segundos respectivamente, y el tiempo de servicio de los paquetes también tiene una distribución exponencial con una rapidez de transmisión $\mu_d = 1/b$ en paquetes por ranuras disponibles por segundo (un servidor con capacidad variable).

Podemos modelar el sistema propuesto como una cadena de Markov ergódica tridimensional que se muestra en la Fig.4.4. El estado del sistema está dado por $\{V(t), F(t), D(t)\}$, $0 \leq V(t) \leq N_v$, $0 \leq F(t) \leq m_f$ y $0 \leq D(t) \leq m_d$, donde $V(t)$, $F(t)$ y $D(t)$ son los números de voz, archivo y paquetes en el sistema, respectivamente, al tiempo t .

4.3 Soluciones exactas aplicando técnicas numéricas

a) Solución parcial

Sea $P_{i,j,k} = \lim_{t \rightarrow \infty} \Pr\{V(t)=i, F(t)=j, D(t)=k\}$ la probabilidad de estado estacionario de que haya i llamadas, j archivos y k paquetes en el sistema. La probabilidad marginal del estado para voz y archivos $[P_{ij}]$ está dada por:

$$P_{ij} = \sum_{k=0}^{m_d} P_{ijk} \quad \dots \dots \dots (4.1)$$

Debido al esquema propuesto aquí, el proceso de comunicación de voz corresponde a un sistema de colas $M/M/N_v/N_v$, y el proceso de transferencia de archivos corresponde a un sistema de colas $M/M/N_f/m_f$. Entonces:

$$P_{ij} = \lim_{t \rightarrow \infty} \Pr\{V(t)=i\} \times \Pr\{F(t)=j\} = P_{v_i} \cdot P_{f_j} \quad \dots \dots \dots (4.2)$$



donde Pv_i y Pf_j se pueden calcular de la siguiente forma (ver Capítulo 2):

$$Pv_i = \frac{\rho_v^i / i!}{N_v \sum_{l=0}^{N_v-1} \rho_v^l / l!} \quad \rho_v = \lambda_v / \mu_v$$

$$Pf_j = \begin{cases} P_{f0}(\rho_f)^j / j! & \text{si } 0 \leq j \leq N_f \\ P_{f0}(\rho_f)^j / N_f! N_f^{j-N_f} & \text{si } m_f > j > N_f \end{cases} \quad \rho_f = \lambda_f / \mu_f$$

$$y \quad P_{f0} = \left\{ \sum_{j=0}^{N_f} \rho_f^j / j! + \sum_{j=N_f+1}^{m_f} \rho_f^j / N_f! N_f^{j-N_f} \right\}^{-1}$$

Los desempeño de los tráficos de voz y archivos se pueden evaluar mediante las siguientes fórmulas:

1) La probabilidad de bloqueo de voz es:

$$PB_v = Pv_{N_v}$$

2) La probabilidad de sobreflujo (overflow) del buffer de archivos es:

$$PB_f = Pf_{m_f}$$

3) El promedio de tiempo de retardo de los archivos es:

$$\bar{T}_f = \left(\sum_{i=0}^{m_f} i Pf_i \right) / \lambda_f$$

4) La utilización de la capacidad reservada para los archivos es:

$$E\{F\} / N_f = \left(\sum_{i=0}^{N_f} i Pf_i + \sum_{i=N_f+1}^{m_f} N_f Pf_i \right) / N_f$$

b) Técnicas recursivas

Definimos un mapeo uno a uno $L : (i, j) \rightarrow l_{ij}$, donde $0 \leq i \leq N_v$, $0 \leq j \leq m_f$ y $1 \leq l_{ij} \leq (N_v + 1)(m_f + 1)$.

Entonces, podemos definir un vector de probabilidades marginales para voz y archivos Π como:

$$\Pi \triangleq \{ \Pi_{l_{ij}} \} \triangleq \{ P_{ij} \}$$

y por lo tanto: $\Pi_{l_{ij}} = \sum_{k=0}^{m_f} P_{ijk}$ (4.3)



Por otro lado, las ecuaciones de equilibrio del modelo se puede escribir de la siguiente forma (ver Fig.4.4):

$$P_{ij(k+1)}^{(N-i-L_j)} \mu_d = P_{ij(k)} [K_i \lambda_v + i \mu_v + K_j \lambda_f + L_j \mu_f + \lambda_d + (N-i-L_j) \mu_d] \\ - P_{i-1,j(k)} (K_i^* \lambda_v) - P_{i+1,j(k)} K_i^{(i+1)} \mu_v - P_{i,j-1(k)} (K_j^* \lambda_f) \\ - P_{i,j+1(k)} K_j^{(L_{j+1} \mu_f)} - P_{ij(k-1)} (\lambda_d)$$

Para $1 \leq k \leq m_d - 1$ (4.4.a)

$$P_{ij(1)}^{(N-i-L_j)} \mu_d = P_{ij(0)} [K_i \lambda_v + i \mu_v + K_j \lambda_f + L_j \mu_f + \lambda_d] - P_{i-1,j(0)} (K_i^* \lambda_v) \\ - P_{i+1,j(0)} K_i^{(i+1)} \mu_v - P_{i,j-1(0)} (K_j^* \lambda_f) - P_{i,j+1(0)} (K_j^{L_{j+1} \mu_f})$$

Para $k=0$ (4.4.b)

donde:

$$K_i = \begin{cases} 1 & \text{si } i < N_v \\ 0 & \text{si } i = N_v \end{cases} \quad K_i^* = \begin{cases} 1 & \text{si } i > 0 \\ 0 & \text{si } i = 0 \end{cases}$$

$$K_j = \begin{cases} 1 & \text{si } j < m_f \\ 0 & \text{si } j = m_f \end{cases} \quad K_j^* = \begin{cases} 1 & \text{si } j > 0 \\ 0 & \text{si } j = 0 \end{cases}$$

$$L_j = \begin{cases} j & \text{si } j < N_f \\ N_f & \text{si } j \geq N_f \end{cases} \quad L_{j+1} = \begin{cases} j+1 & \text{si } j < N_f \\ N_f & \text{si } j \geq N_f \end{cases}$$

Utilizando la transformación L que definimos al principio de esta parte, la ecuación anterior (4.4) se puede escribir en una forma matricial como:

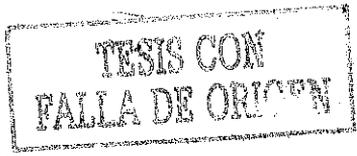
$$\begin{cases} \underline{P}_{k+1} = A_{k+1} \underline{P}_k + B_{k+1} \underline{P}_{k-1} & \text{para } 1 \leq k \leq m_d - 1 & \dots\dots (4.5.a) \\ \underline{P}_1 = A_1 \underline{P}_0 & \text{para } k=0 & \dots\dots (4.5.b) \end{cases}$$

donde los elementos de las matrices A_{k+1} , B_{k+1} y A_1 se pueden obtener facilmente por la ecuación (4.4). \underline{P}_k es un vector de dimensión $M=(N_v+1)(m_f+1)$.

Además, si las rapideces de transición no dependen de k (como es el caso de la Fig.4.4), la ecuación (4.5) se simplifica a:

$$\begin{cases} \underline{P}_{k+1} = A \underline{P}_k + B \underline{P}_{k-1} & \text{para } 1 \leq k \leq m_d - 1 & \dots\dots (4.6.a) \\ \underline{P}_1 = A_1 \underline{P}_0 & \text{para } k=0 & \dots\dots (4.6.b) \end{cases}$$

Si $m_d < \infty$ (buffer finito para paquetes), la ecuación anterior se puede resolver numéricamente por la recursión directa:



$$\begin{aligned} \Pi &= \underline{P}_0 + \underline{P}_1 + \underline{P}_2 + \dots + \underline{P}_{m_d} \\ &= \underline{P}_0 + A_1 \underline{P}_0 + A \underline{P}_1 + B \underline{P}_0 + \dots \\ &= [I + A_1 + (AA_1+B) + \dots] \underline{P}_0 \\ &= S_B \underline{P}_0 \end{aligned} \quad \dots \dots \dots (4.7)$$

$$\dots \underline{P}_0 = S_B^{-1} \quad \dots \dots \dots (4.8)$$

Como Π está dada por las ecuaciones (4.2) y (4.3), las probabilidades \underline{P}_k se pueden obtener con la ecuación (4.6), usando \underline{P}_0 de la ecuación (4.8).

Sin embargo, la técnica de recursión directa tiene dificultades numéricas. Para los parámetros típicos de voz, archivos y datos intermitentes, la ecuación (4.6) tendrá valores característicos mayores que uno, y como la solución involucra tomar potencias de estas matrices con estos valores característicos, la dificultad numérica crecerá si el tamaño del buffer de los paquetes (m_d) es muy grande [17].

En la parte siguiente, vamos a presentar algunas otras técnicas numéricas que resuelven el problema mencionado con mayor facilidad.

c) Casos de buffer finito ($m_d < \infty$)

Como nosotros sabemos, cualquiera ecuación diferencial de segundo orden se puede reducir a una ecuación de primer orden, si utilizamos la doble dimensionalidad. Ahora, si definimos un vector de dimensión $2M$ como:

$$X_k \stackrel{\Delta}{=} \begin{bmatrix} \underline{P}_{k+1} \\ \underline{P}_k \end{bmatrix} \quad \dots \dots \dots (4.9)$$

Entonces la ecuación (4.6.a) se convierte en:

$$X_{k+1} = K X_k, \quad \text{con la matriz } K = \begin{bmatrix} A & B \\ I & C \end{bmatrix}_{2M \times 2M} \quad \dots \dots (4.10)$$

La solución de esta ecuación tiene la siguiente forma:

$$X_k = K^k X_0 \quad \dots \dots \dots (4.11)$$

donde por la definición:

$$X_0 \stackrel{\Delta}{=} \begin{bmatrix} \underline{P}_1 \\ \underline{P}_0 \end{bmatrix} = \begin{bmatrix} A_1 \\ I \end{bmatrix} \underline{P}_0 \quad \dots \dots \dots (4.12)$$

Obsevamos las siguientes ecuaciones:

$$\sum_{k=0}^{m_d-1} X_k = \sum_{k=0}^{m_d-1} K^k X_0 \quad \dots \dots \dots (4.13)$$



$$\sum_{k=0}^{m_d-1} X_k = \sum_{k=0}^{m_d-1} \begin{bmatrix} P_{k+1} \\ P_k \end{bmatrix} = \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} \quad \dots\dots\dots (4.14)$$

Si igualamos (4.13) y (4.14), se tiene

$$\sum_{k=0}^{m_d-1} K^k X_0 = \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} \quad \dots\dots\dots (4.15)$$

y por lo que

$$X_0 = \begin{bmatrix} \sum_{k=0}^{m_d-1} K^k \end{bmatrix}^{-1} \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix}$$

Ahora, si definimos una matriz F de dimensión 2Mx2M como:

$$F \triangleq \begin{bmatrix} \sum_{k=0}^{m_d-1} K^k \end{bmatrix}^{-1} \triangleq \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix} \quad \dots\dots\dots (4.17)$$

donde las matrices F_1, F_2, F_3 y F_4 son de dimensión MxM. Entonces, la ecuación (4.16) se convierte en la siguiente ecuación, sustituyendo la ecuación (4.12):

$$\begin{bmatrix} A_1 P_0 \\ P_0 \end{bmatrix} = \begin{bmatrix} F_1 & F_2 \\ F_3 & F_4 \end{bmatrix} \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} \quad \dots\dots\dots (4.18)$$

Las soluciones de esta ecuación son:

$$P_0 = (C + F_4^{-1} - F_2^{-1} A_1)^{-1} C \Pi \quad \dots\dots\dots (4.19.a)$$

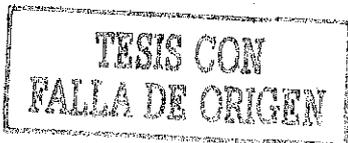
$$P_{m_d} = F_4^{-1} F_3 \Pi + \Pi - F_4^{-1} F_3 P_0 - F_4^{-1} P_0 \quad \dots\dots\dots (4.19.b)$$

$$\text{donde } C = F_4^{-1} F_3 - F_2^{-1} F_1$$

En el cálculo de la matriz F, uno puede diagonalizar la matriz K como[31]:

$$K = Q A_k Q^{-1}$$

donde



$$\sum_{k=0}^{m_d-1} k X_k = \sum_{k=0}^{m_d-1} k X^k X_0 = \left(\sum_{k=0}^{m_d-1} k K^k \right) F \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} \quad \dots (4.21)$$

Utilizando la diagonalización de la matriz K, la ecuación (4.21) tiene la forma:

$$\sum_{k=0}^{m_d-1} k X_k = Q \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_{2M} \end{bmatrix} Q^{-1} \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} \quad \dots (4.22)$$

donde:

$$J_i = \frac{\sum_{k=0}^{m_d-1} k \lambda_i^k}{\sum_{k=0}^{m_d-1} \lambda_i^k} = \begin{cases} \frac{\lambda_i}{1 - \lambda_i} - m_d \frac{\lambda_i^{m_d}}{1 - \lambda_i^{m_d}} & \text{si } \lambda_i \neq 1 \\ \frac{m_d - 1}{2} & \text{si } \lambda_i = 1 \end{cases} \quad \dots (4.23)$$

Si definimos una matriz G de dimensión 2Mx2M tal que:

$$G \triangleq \begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix} \triangleq Q \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_{2M} \end{bmatrix} Q^{-1}$$

Y por otro lado, observese que:

$$\sum_{k=0}^{m_d-1} k X_k = \sum_{k=0}^{m_d-1} k \begin{bmatrix} P_{k-1} \\ P_k \end{bmatrix} = \begin{bmatrix} \underline{g} - \Pi + P_0 \\ \underline{g} - m_d P_{m_d} \end{bmatrix} \quad \dots (4.24)$$

Igualemos las ecuaciones (4.22) y (4.24), y sustituyendo la matriz G, se tiene:

$$\begin{bmatrix} G_1 & G_2 \\ G_3 & G_4 \end{bmatrix} \begin{bmatrix} \Pi - P_0 \\ \Pi - P_{m_d} \end{bmatrix} = \begin{bmatrix} \underline{g} - \Pi + P_0 \\ \underline{g} - m_d P_{m_d} \end{bmatrix} \quad \dots (4.25)$$

La solución de la ecuación (4.25) para \underline{g} es :

$$\underline{g} = (G_1 + I)(\Pi - P_0) + G_2(\Pi - P_{m_d}) \quad \dots (4.26)$$



Nótese que en la ecuación (4.23), los valores de J_i se pueden obtener sin importar si los valores característicos λ_i son mayores o menores que uno, porque para un $\lambda_i > 1$, y m_d muy grande, J_i tiende a $\frac{\lambda_i}{1 - \lambda_i} + m_d$. Y para $\lambda_i \ll 1$, J_i tiende a $\frac{\lambda_i}{1 - \lambda_i}$.

La técnica que hemos presentado en esta parte, es computacionalmente eficiente, y evita las dificultades numéricas de la técnica de recursión directa. Es decir, en este método no importa si los valores característicos son mayores o iguales que uno, ni importa el tamaño del buffer de los paquetes.

d) Casos de buffer infinito

Consideramos la ecuación (4.6) cuando se supone que el buffer de los paquetes es infinito ($m_d \rightarrow \infty$). Neuts [32, 33] ha mostrado que una solución de este tipo de ecuaciones existe y tiene la siguiente forma geométrica matricial:

$$\underline{P}_k = R^k \underline{P}_0 \quad \dots\dots\dots (4.27)$$

donde la matriz R es de dimensión MxM con todos sus valores característicos menores que uno, y se satisface la ecuación:

$$R^2 = AR + B \quad \dots\dots\dots (4.28)$$

Neuts también ha mostrado que la recursión

$$R_{(n)} = A^{-1} (R_{(n-1)}^2 - B) \quad \text{con } R_{(0)} = 0 \quad \dots\dots\dots (4.29)$$

converge a una matriz de rapidez correcta R.

Williams y Leon-Garcia [15] observaron que la ecuación (4.29) convergerá muy lentamente para algunos parámetros típicos de un sistema integrado de comunicación. Ellos propusieron otro método para aproximar R, basándose en la siguiente manera cual vamos a describir brevemente a continuación.

Como la componente (i,j) de la ecuación (4.28) está dada por:

$$\sum_{k=1}^M r_{ik} r_{kj} = \sum_{k=1}^M a_{ik} r_{kj} + b_{ij} \quad \dots\dots\dots (4.30)$$

Entonces, para los elementos de la diagonal r_{ii} , se tiene:

$$r_{ii}^2 - a_{ii} r_{ii} - b_{ii} + \sum_{k \neq i, k=1}^M (r_{ik} - a_{ik}) r_{ki} = 0 \quad \dots\dots\dots (4.31)$$



La ecuación (4.31) se resuelve como una ecuación de segundo orden, y la fórmula recursiva para los r_{ii} será:

$$r_{ii}^{(\text{nuevo})} = 1/2 \{ a_{ii} \pm \sqrt{ a_{ii}^2 - 4 \left[\sum_{\substack{k=1 \\ k \neq i}}^M (r_{ik}^{(\text{viejo})} - a_{ik}) r_{ki}^{(\text{viejo})} - b_{ii} \right] } \} \dots (4.32)$$

Tomando las raíces negativas de esta ecuación, se puede lograr una R con valores característicos menores que uno.

Para los elementos fuera de la diagonal, la ecuación correspondiente es:

$$(r_{ii} + r_{jj} - a_{ii}) r_{ij} + \sum_{\substack{k=1 \\ k \neq i, j}}^M r_{ik} r_{kj} - \sum_{\substack{k=1 \\ k \neq i}}^M a_{ik} a_{kj} - b_{ij} = 0 \dots (4.33)$$

$$y \quad r_{ij}^{(\text{nuevo})} = \frac{\sum_{\substack{k=1 \\ k \neq i}}^M a_{ik} r_{kj}^{(\text{viejo})} + b_{ij} - \sum_{\substack{k=1 \\ k \neq i, j}}^M r_{ik}^{(\text{viejo})} r_{kj}^{(\text{viejo})}}{r_{ii}^{(\text{viejo})} + r_{jj}^{(\text{viejo})} - a_{ii}} \dots (4.34)$$

Este método es muy eficiente para el caso de que R es diagonalmente dominante (justamente es el caso de nuestro modelo).

Una vez obtenida R, el vector P_0 y los primeros dos momentos del tráfico de paquetes se pueden calcular mediante:

$$P_0 = (I - R) \Pi \dots (4.35)$$

$$m_1 = q = R(I - R)^{-1} \Pi \dots (4.36)$$

$$m_2 = [R(I - R)^{-2} + R^2(I - R)^{-2}] \Pi \\ = R(I + R)(I - R)^{-2} \Pi \dots (4.37)$$

Mediante las técnicas numéricas presentadas en esta sección, se puede evaluar el desempeño del sistema integrado para tres tipos de tráfico. El análisis del desempeño es lo que trataremos en la siguiente sección.

4.4 Aproximación cuasi-estática

Hemos presentado en la sección anterior las soluciones ("exactas") del modelo aplicando algunas técnicas numéricas. Pero, estas técnicas no son eficientes para resolver un sistema grande, debido a la necesidad de hacer varias operaciones computacionalmente difíciles (por ejemplo, encontrar valores y vectores característicos) sobre matrices de gran dimensión (en este caso, $2(N_v + 1)(m_f + 1)$). Algunos otros métodos, tanto exactos como aproximados, que evitan estas dificultades computacionales son



deseables, para evaluar el desempeño de los sistemas integrados de tres tráficos. A continuación, presentamos un método de aproximación del comportamiento cuasi-estático del tráfico de paquetes, o simplemente, aproximación cuasi-estática.

El proceso de colas de paquetes en un sistema integrado de tres tráficos, se comporta en forma diferente al proceso de un sistema en el cual la capacidad de transmisión es fija, porque cada cambio de número de ranuras ocupadas por voz y archivos en la primera y la segunda parte del marco, resulta en un cambio de capacidad de transmisión del sistema para paquetes. Por lo tanto, definimos

$$n_{vf}(t) \triangleq \text{número total de ranuras ocupadas por voz y archivos en el marco en tiempo } t. \quad (0 \leq n_{vf}(t) \leq N_v + N_f)$$

Entonces, la rapidez de transmisión de paquetes en un sistema de uniservidor en tiempo t es $(N - n_{vf}(t))\mu_d$.

Nosotros diremos que el sistema está en modo k si $n_{vf}(t) = k$. Por lo cual, la utilización (o carga) instantánea del sistema de paquetes en modo k está dada por:

$$\rho_d(k) = \frac{\lambda_d}{\mu_d(N-k)} < 1$$

La aplicabilidad de la aproximación cuasi-estática en un sistema integrado depende de la duración de cada modo y el tiempo de relajamiento del sistema de colas de paquetes en estos modos [34]. Si la duración de cada modo es mucho más larga que el tiempo de relajamiento del proceso de paquetes cuando está en estos modos, estos modos son estacionarios. O en otras palabras, si el sistema para paquetes alcanza a estabilizarse antes de que se cambie el modo, nosotros podemos analizar el comportamiento del proceso de paquetes en este modo sin importar otros modos.

Cuando todos los modos del sistema son estacionarios, el desempeño del proceso de paquetes se puede aproximar simplemente como [21]:

$$E\{D\} = E\{\text{número de paquetes en el sistema}\} = \sum_{k=0}^{N_v + N_f} E\{D | n_{vf} = k\} \Pr\{n_{vf} = k\}.$$

En caso de un buffer finito ($m_d < \infty$):

{PFB_d} = E{probabilidad de sobreflujo de buffer}

$$\cong \sum_{k=0}^{N_v + N_f} \{PFB_d\}_{n_{vf}=k} \Pr\{n_{vf}=k\}.$$

**TESIS CON
FALLA DE ORIGEN**

Donde $E\{D|n_{vf}=k\}$ es el número promedio de paquetes en el sistema cuando el modo es k , y $\{PFB_d\}_{n_{vf}=k}$ es la probabilidad de que el buffer se llene en modo k . Estas cantidades se pueden calcular mediante un sistema $M/M/1$ o $M/M/1/m_d$, debido a los modelos originales que presentamos en las secciones anteriores. O sea

$$E\{D|n_{vf}=k\} = \frac{\rho_d(k)}{1-\rho_d(k)} \quad (\rho_d(k) < 1) \quad \text{para } M/M/1$$

$$E\{D|n_{vf}=k\} = \frac{1 - \binom{m_d+1}{m_d} \rho_d^{m_d}(k) + m_d \rho_d^{m_d}(k)}{(1-\rho_d(k))(1-\rho_d^{m_d}(k))} \quad (\rho_d(k) \neq 1) \quad \text{para } M/M/1/m_d$$

$$E\{D|n_{vf}=k\} = \frac{m_d}{2} \quad (\rho_d(k) = 1) \quad \text{para } M/M/1/m_d$$

$$\{PFB_d\}_{n_{vf}=k} = \frac{\rho_d^{m_d}(k) - \rho_d^{m_d+1}(k)}{1 - \rho_d^{m_d+1}(k)} \quad (\rho_d(k) \neq 1) \quad \text{para } M/M/1/m_d$$

$$\{PFB_d\}_{n_{vf}=k} = \frac{1}{1 + m_d} \quad (\rho_d(k) = 1) \quad \text{para } M/M/1/m_d$$

La probabilidad de que esté en modo k está dada por:

$$\Pr\{n_{vf}=k\} = \sum_{i,j \in R_k} P_v(i)P_f(j), \quad R_k = \{(i,j) | i+j=k\} \quad \text{para } j < N_f$$

$$\Pr\{n_{vf}=k\} = \sum_{i,j \in R_k} P_v(i)P_f(j), \quad R_k = \{(i,j) | i+N_f=k\} \quad \text{para } m_f \geq N_f$$

La exactitud de la aproximación cuasi-estática se puede ver en la siguiente sección, cuando analizamos el desempeño del sistema, utilizando tanto la solución "exacta" como la aproximada.

4.5 Análisis del desempeño

En esta sección, analizamos el desempeño del sistema integrado para tres tráfico. La evaluación de desempeño se realiza utilizando tanto las técnicas numéricas como la aproximación cuasi-estática.

Como mencionamos previamente, los procesos de comunicación de voz y la transferencia de archivos son independientes, el proceso de voz corresponde a un sistema de pérdidas $M/M/N_v/N_v$, mientras que el proceso de archivos corresponde a un sistema de $M/M/N_f/m_f$. Entonces, para una carga de tráfico de voz ρ_v dada, se selecciona un número

TESIS CON
FALLA DE ORIGEN

de ranuras N_v para satisfacer una probabilidad de bloqueo específica PB_v . Si PB_v es muy pequeña, como es el caso típico, se debe seleccionar un N_v a satisfacer el requerimiento de flujo pico. Como el resultado, el número promedio de ranuras usadas por el tráfico de voz puede ser mucho menor que la capacidad reservada (N_v). Sea $V(t)$ el número de llamadas en el tiempo t en el enlace, el número promedio de ranuras usadas por el tráfico de voz es:

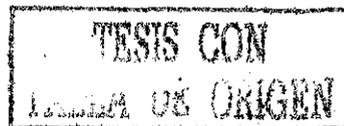
$$E\{V\} = \sum_i P_v(i) = \rho_v (1 - PB_v).$$

La Figura 4.5.a es una gráfica de $E\{V\}/N_v$, para una $PB_v = 2\%$. En general, una gran parte de la capacidad reservada para voz está desperdiciada. Particularmente, en la región de la carga $1 \leq \rho_v \leq 10$, que es el caso más realista, el número de ranuras requeridas es $4 \leq N_v \leq 17$, para una $PB_v = 2\%$, el desperdicio de la capacidad es enorme (ver Fig. 4.5.b).

Lo mismo sucede en el caso de transferencias de archivos. La Fig. 4.5.a también muestra la utilización de la capacidad de tráfico de archivos, para una probabilidad de sobreflujo del buffer de archivos $PB_f = 10^{-4}$, en lo cual suponemos que el tamaño del buffer m_f es el doble de N_f .

El desempeño de tráfico de archivos se muestra en la Fig. 4.6, donde la capacidad reservada para él es de 10 ranuras, y el número máximo de solicitudes de transferencia de archivos que se puede quedar en el sistema es de 50. Para diferentes cargas de tráfico de archivos, la variación de PB_f y el tiempo de retardo se presenta en la gráfica. Se puede observar que para las cargas $\rho_f < 8$, la PB_f es prácticamente aceptable (10^{-5}). El tiempo de retraso de archivos crece mucho cuando la carga ρ_f cruza 10 Erlangs. Un fenómeno "raro" es que cuando la carga pasa a 13 Erlangs, el tiempo de retraso baja un poco. Una explicación es que en estos momentos, como la PB_f aumenta, pues muchas solicitudes se pierden, el promedio de tiempo de retardo se reduce.

Las observaciones anteriores muestran que muchas capacidades reservadas para los tráficos de voz y archivos se van a desperdiciar. Por ejemplo, consideramos un sistema de $N=20$, y suponemos que $N_v=9$, $N_f=6$. Para una $PB_v=2\%$, se permite un tráfico de voz con intensidad $\rho_v=4.35$ Erlangs; y para una $PB_f=10^{-4}$, la carga de tráfico de archivos debe ser $\rho_f=2.42$ Erlangs, con un tamaño de buffer $m_f=12$. Con un esquema de frontera-fija para tres tráficos, sólo 5 ranuras/marco están disponibles para el tráfico de datos, pero si utilizamos un esquema de frontera-móvil, 13.32 ranuras/marco están disponibles para los paquetes, en promedio. La capacidad para el tráfico de paquetes se aumentó casi al doble, por lo cual el enlace puede manejar un tráfico de paquetes de mucho mayor intensidad o carga.



El esquema que propusimos en este trabajo para tres tipos de tráficos, es una versión de frontera-móvil para dos tráficos que se presentaron en trabajos previos. En la Fig. 4.7, consideramos un sistema integrado de $N=12$, $N_v=4$, $N_f=3$, $m_f=5$ y una duración de marco $b=1$ ms. Para una $PB_v=1.5\%$ fija y una $PB_f=1.5\%$ fija, las cargas correspondientes de voz y archivos son $\rho_v=1$ y $\rho_f=1.25$ Erlangs. En caso de un buffer de datos infinito, la técnica de aproximación iterativa de la matriz de forma geométrica R fue aplicada (ver ecuación (4.28)). El error de la iteración fue del orden de menor que 10^{-8} . Hemos observado que para las cargas de paquetes $\rho_d < 5$ Erlangs, el número promedio de paquetes en el sistema ($E\{D\}$) es relativamente pequeño (curva A), y para las cargas grandes, $E\{D\}$ crece drásticamente. Esto implica que para utilizar las capacidades desperdiciadas por los tráficos de voz y archivos, los paquetes tienen que aguantar una larga cola de buffer. O sea, cuando ρ_d excede la capacidad garantizada para el tráfico de paquetes ($N - N_v - N_f$), la cola de los paquetes aumenta por esperar las largas duraciones de conversaciones en voz y transferencias de archivos.

Como lo indicamos anterior, la suposición de buffer infinito no es realista para un sistema integrado de tres tipos de tráficos, porque una cola larga de paquetes ocurrirá tal que podría exceder cualquier capacidad de buffer razonable. El análisis de un sistema integrado con buffer finito se puede realizar utilizando la técnica de diagonalización de matrices, la cual hemos presentado en la sección 4.3. En la curva B de la Fig. 4.7, se muestra la variación de $E\{D\}$ con respecto a la carga de paquetes, para los mismos parámetros que el caso del buffer infinito, excepto que el tamaño del buffer es igual a 100. Se nota que para $\rho_d < (N - N_v - N_f)$, el desempeño es similar que en el caso del buffer infinito, mientras ρ_d crece, $E\{D\}$ se aumentará, pero no tan drásticamente como en el otro caso. Esto es debido a que la probabilidad de sobreflujo de buffer de paquetes aumenta (curva C).

El tiempo de retardo de los paquetes se puede calcular utilizando la fórmula de Little.

La aproximación cuasi-estática del desempeño de paquetes en el mismo sistema considerado anteriormente (en caso de buffer finito) se muestra en la curva D y la curva E, donde se puede apreciar una muy buena aproximación de los valores calculados con la solución "exacta".

Utilizando la aproximación cuasi-estática, la Fig. 4.8 muestra el comportamiento del tráfico de paquetes con diferentes tamaños de buffer para archivos (m_f), en un sistema grande $N=20$, $N_v=8$, $N_f=7$, $m_d=1000$, y $\rho_v=4$, $\rho_f=4$, $\mu_v^{-1}=\mu_f^{-1}=100$ seg., $\rho_d=7$. Se puede observar que para $m_f \geq 20$, $E\{D\}$ se tiende a un valor fijo (29.6).

En el mismo sistema, con $m_f=20$, se grafica el comportamiento del tráfico de paquetes, en función de tamaño de buffer de paquetes m_d , en la Fig. 4.9. El resultado



muestra que el promedio de paquetes en el sistema crece al aumentar el tamaño de buffer, esto es debido a que la intensidad de paquetes ($\rho_d=7$) en el sistema es mucho mayor que la capacidad reservada para ellos, al aprovechar las capacidades desocupadas por voz y archivos, los paquetes tienen que esperar mucho tiempo por las largas duraciones de transacciones, cuando las cargas de voz y archivos son altas.

Este fenómeno se puede apreciar claramente en la Fig. 4.10, donde variamos la carga de paquetes (ρ_d), para el sistema considerado anteriormente, excepto $m_d=10000$. Para $\rho_d > 5$, el número de paquetes en el sistema aumenta fuertemente conforme se incrementa ρ_d . Cuando $\rho_d=10$, $E\{D\}$ llega a 2186 y la probabilidad de rebalzo del buffer es 3.13%.

Una de las soluciones posibles del problema mencionado anteriormente es utilizar la técnica de control de flujo, la cual presentaremos a continuación.

4.6 Control de flujo

Control de flujo es un mecanismo para prevenir congestión de tráfico en la red o enlaces. La congestión ocurre cuando los nodos que mandan mensajes a un nodo particular exceden la capacidad de este nodo para procesar estos mensajes. Por lo cual el problema se reduce al que cada nodo proporciona un mecanismo para controlar la rapidez con que se reciben los mensajes de otros nodos. Se puede decir que el control de flujo es un proceso de regularización de rapidez de llegadas de un nodo específico.

Como indicamos previamente, al utilizar las capacidades reservadas para voz y archivos, los paquetes tienen que esperar mucho tiempo en el buffer para sus transmisiones, cuando temporalmente hay mucho tráfico de voz y archivos. En caso de que ocurra esto (a veces el tiempo de retardo de paquetes llega a más de una hora), algunas técnicas de control de flujo serán necesarias. En este trabajo, proponemos una forma para controlar flujos de llegadas de los paquetes, la cual consiste de observar la intensidad instantánea de paquetes que se define como:

$$\rho(t) = \frac{\lambda_d(t)}{\mu_d(N - n_{vf}(t))}$$

donde $\lambda_d(t)$ es la rapidez de llegada de paquetes al tiempo t , y $n_{vf}(t)$ es el número total de ranuras ocupadas por voz y archivos, en tiempo t .

Para cada cambio de $n_{vf}(t)$, el sistema tiene que controlar (regularizar) $\lambda_d(t)$ para que la intensidad de paquetes sea razonable (i.e. menor que uno).



Como ejemplo consideramos un sistema en el cual existen tres clases de tráfico de paquetes : 20% de clase de alta prioridad, 40% de media prioridad, y 40% de la clase de baja prioridad. El control de flujo de tráfico de paquetes se realiza de la siguiente manera: cada momento de incremento de $n_{vf}(t)$ (obviamente se aumenta $\rho(t)$), se compara la utilización instantánea $\rho(t)$ con un valor límite ρ_{max} ; si se excede, el flujo de paquetes debe de reducir $\lambda_d(t)$ (igualmente $\rho(t)$) hasta que la utilización instantánea sea menor que ρ_{max} . Esto es, si $\rho(t) > \rho_{max}$, el sistema tiene que negar algunos accesos de los paquetes, empezando con los paquetes de baja prioridad. Los paquetes que no se los aceptó el sistema, no quiere decir que se perdieron, sino se desviaron a otros enlaces de la red. En el momento de que se disminuye $n_{vf}(t)$, por la finalización de algunas comunicaciones en voz o archivos, el sistema puede manejar más paquetes para mantener $\rho(t)$ lo más cerca de ρ_{max} posible. Es decir, cuando se desocupan algunas ranuras por voz y archivos, el sistema aumenta la entrada de paquetes (seguramente se disminuyó el acceso de paquetes antes), para tener más utilizations de la capacidad de transmisión.

El desempeño del sistema utilizando este esquema de control de flujo se puede evaluar, mediante el mismo diagrama de rapidez de transición de la Fig. 4.4, cambiando las rapideces de llegadas de paquetes entre cualquier estados (i,j,k) y $(i,j,k+1)$ a:

$$\lambda_d: \quad \text{para} \quad 0 \leq n_{vf}(i,j) < N - \frac{\lambda_d/\mu_d}{\rho_{max}}$$

$$0.6\lambda_d: \quad \text{para} \quad N - \frac{\lambda_d/\mu_d}{\rho_{max}} \leq n_{vf}(i,j) < N - \frac{0.6\lambda_d/\mu_d}{\rho_{max}}$$

$$0.2\lambda_d: \quad \text{para} \quad N - \frac{0.6\lambda_d/\mu_d}{\rho_{max}} \leq n_{vf}(i,j) < N - \frac{0.2\lambda_d/\mu_d}{\rho_{max}}$$

$$0: \quad \text{para} \quad N - \frac{0.2\lambda_d/\mu_d}{\rho_{max}} \leq n_{vf}(i,j)$$

donde $n_{vf}(i,j) = \begin{cases} i+j & \text{si } j < N_f \\ i+N_f & \text{si } m_f \geq j \geq N_f \end{cases}$

Fig. 4.11 muestra el desempeño del sistema utilizando control de flujo, para el caso de buffer infinito del sistema considerado en la Fig. 4.7, con $\rho_{max}=1$. Es



claro que aplicando control de flujo, el desempeño del tráfico de paquetes mejora significativamente. Las discontinuidades en las curvas ocurren en los puntos de $\rho_{\max} = (N - n_{vf}(i,j))/k$, donde $n_{vf}(i,j) = 0, 1, \dots, N_v + N_f$ y $k = 1, 0.6, 0.2$.

Debido al esquema de control de flujo propuesto, el flujo de paquetes de datos entra al sistema con una rapidez de llegada variable ($\lambda_d, 0.6\lambda_d$ o $0.2\lambda_d$), dependiendo de la situación actual del sistema. Si definimos :

$PB_0 \triangleq$ El porcentaje del tráfico de paquetes de baja prioridad bloqueado

$PB_1 \triangleq$ El porcentaje del tráfico de paquetes de media prioridad bloqueado

$PB_2 \triangleq$ El porcentaje del tráfico de paquetes de alta prioridad bloqueado.

Entonces estas cantidades se pueden calcular como:

PB_0 = suma de las probabilidades de los estados que usan $0.6\lambda_d, 0.2\lambda_d$, o cero como su rapidez de llegadas de paquetes.

PB_1 = suma de las probabilidades de los estados que usan $0.2\lambda_d$ o cero como su rapidez de llegadas de paquetes.

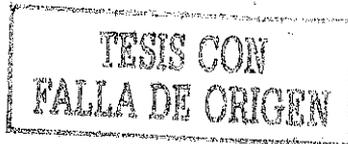
PB_2 = suma de las probabilidades de los estados que usan cero como su rapidez de llegadas de paquetes.

Por lo cual, la proporción del tráfico total de paquetes bloqueado PB_T es:

$$PB_T = 0.4 PB_0 + 0.4 PB_1 + 0.2 PB_2$$

En el sistema considerado anteriormente (Fig. 4.11), para una carga de paquetes $\rho_d = 9$ Erlangs y $\rho_{\max} = 1$, 16% del tráfico total de paquetes son bloqueados, de los cuales, 39.8% son de clase de baja prioridad y 0.2% son de media prioridad.

El resultado de control de flujo para el sistema considerado en la Fig. 4.10 muestra también la eficiencia de esta técnica.



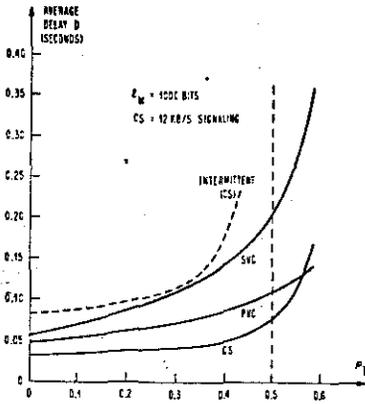


Fig.4.1 Promedio de retardo como una función de utilización de información. Longitud de mensaje es de un kbits. (CS: Circuit Switching, SVC: Packet Switching (in Switched Virtual Circuit mode), PVC: Packet Switching (in Permanent Virtual Circuit mode)).

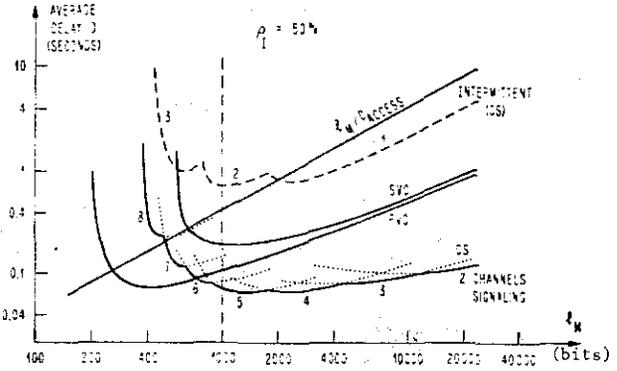


Fig.4.2 Promedio de retardo como una función de promedio de longitud de mensaje, para una utilización de información = 50%.

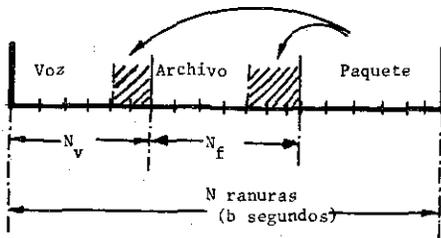
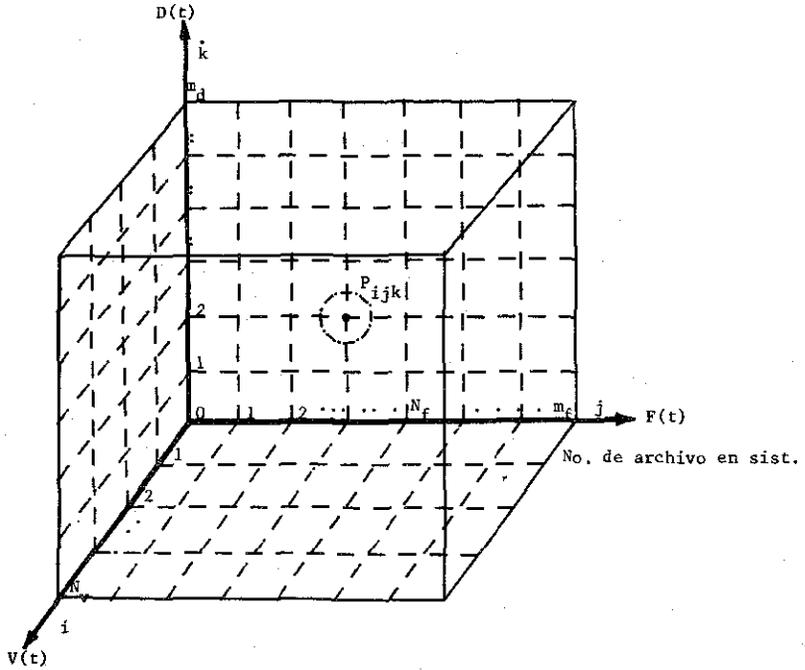


Fig.4.3 Estructura de marco del sistema integrado para tres tráfico.

TESIS CON
FALLA DE ORIGEN

No. de paquete en sistema



No. de voz en sistema

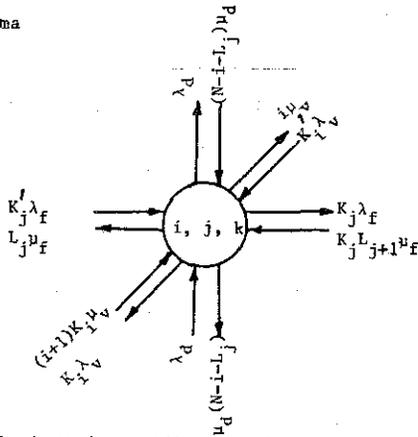


Fig. 4.4 Modelo de Markov tridimensional para sistema integrado de tres tráficos.

TESIS CON
FALLA DE ORIGEN

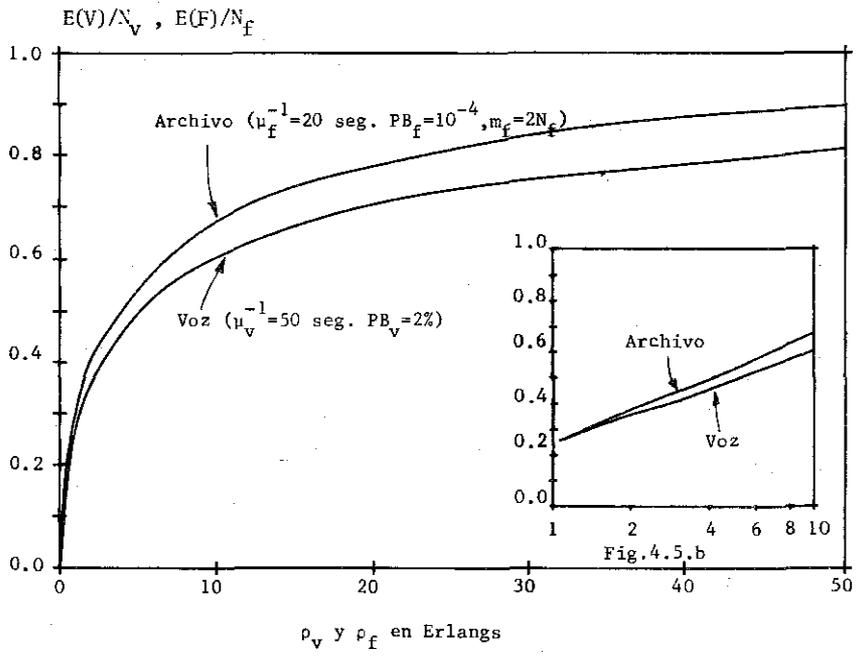


Fig. 4.5 Utilizaciones de capacidades reservadas para voz y archivo, con respecto a sus cargas de tráfico.

TESIS CON
FALLA DE ORIGEN

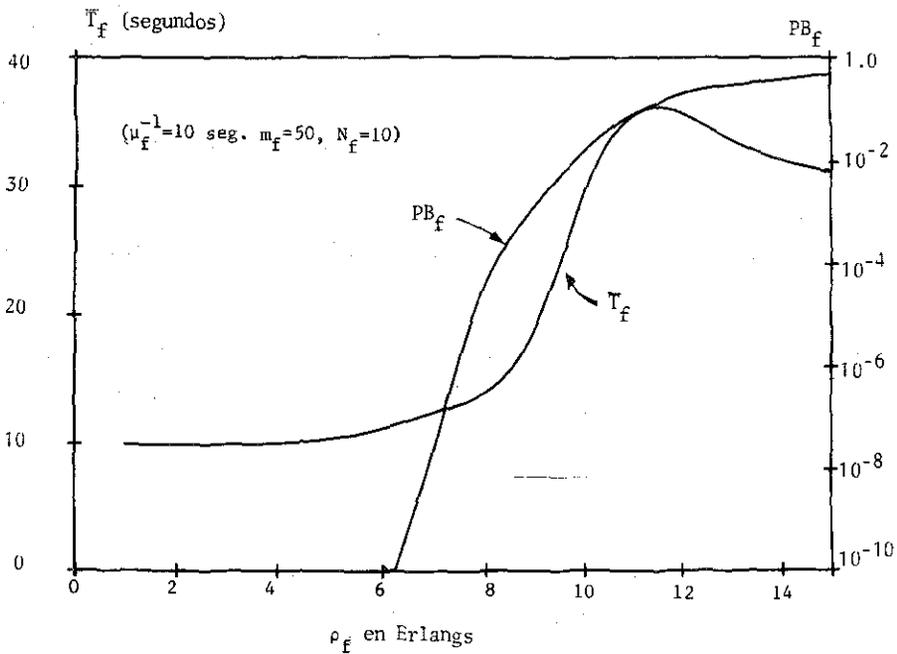


Fig. 4.6 Variación de tiempo de retardo y probabilidad de overflow del buffer para solicitud de archivos, como funciones de la carga de archivo.

TESIS CON
 FALLA DE ORIGEN

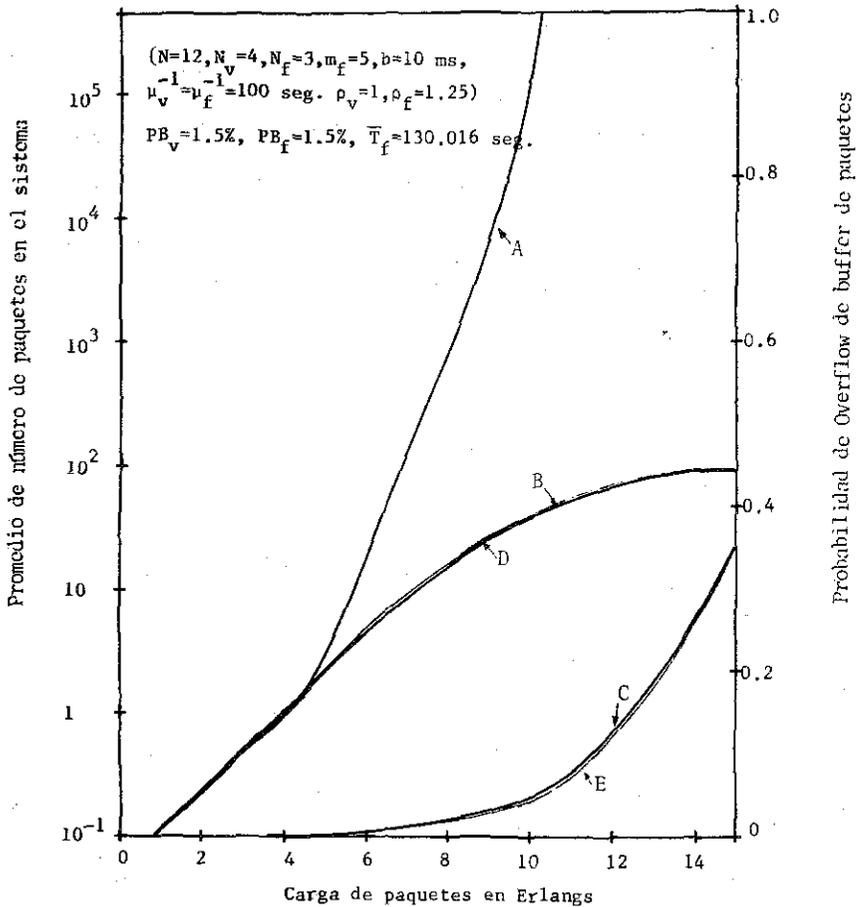


Fig. 4.7 El comportamiento del sistema de tres tráficos, con diferentes cargas de tráfico de paquetes. (Curva A: E(D) caso de buffer infinito, B: E(D) caso de buffer finito con $m_d=100$, C: PB_d del caso de buffer finito, D: Aprox. cuasi-estática de B, E: Aprox. cuasi-estática de C).

TESIS CON
 FALLA DE ORIGEN

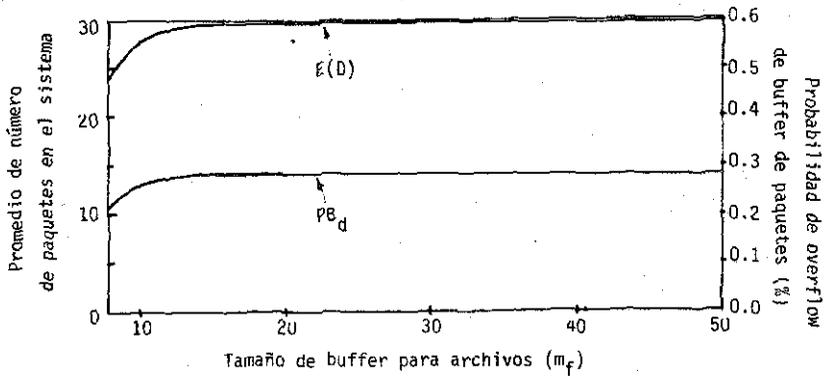


Fig.4.8 Los efectos del tamaño de buffer de archivos(m_f) sobre el tráfico de paquetes($N=20, N_v=8, N_f=7, m_d=1000, \rho_v=\rho_f=0.04/0.01, \rho_d=700/100$), utilizando aproximación cuasi-estática.

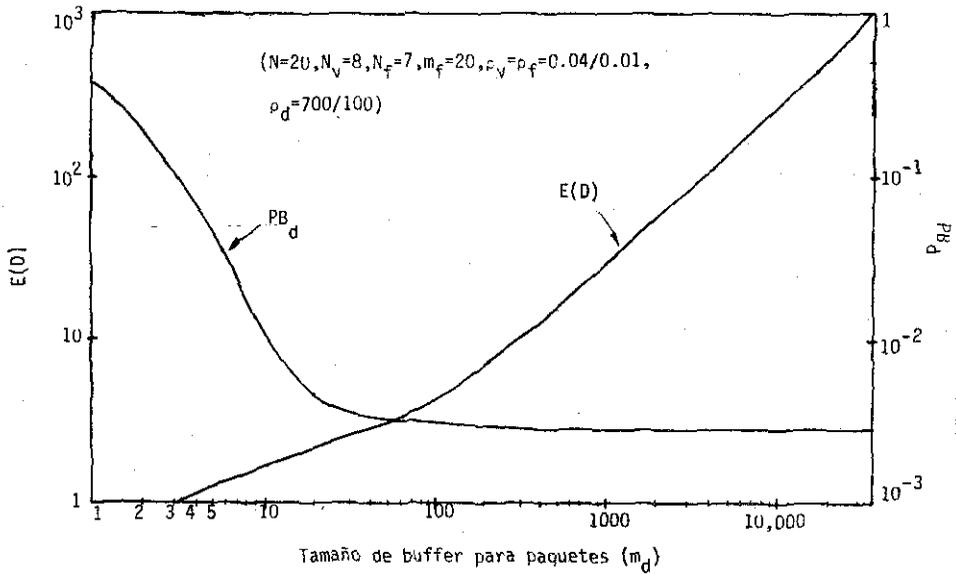


Fig. 4.9 Efecto de tamaño de buffer de paquetes sobre el desempeño del tráfico de paquetes, utilizando aproximación cuasi-estática.

TESIS CON FALLA DE ORIGEN

ESTA TESIS NO SALE DE LA BIBLIOTECA

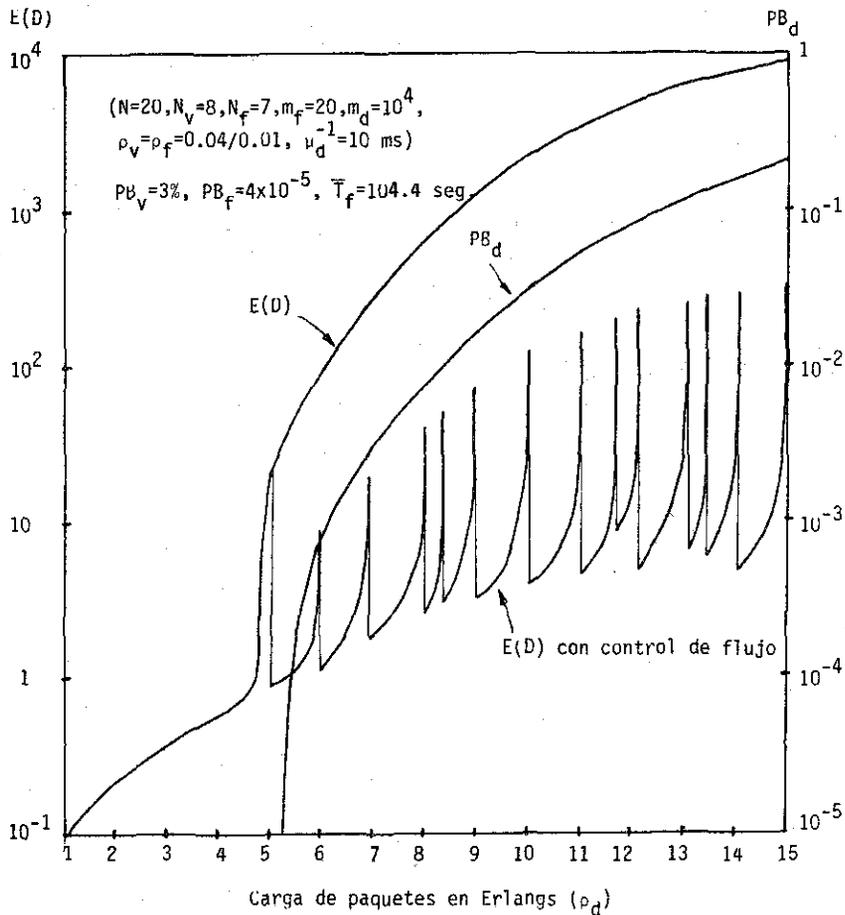


Fig. 4.10 El comportamiento del tráfico de paquetes, como una función de la carga de paquetes, utilizando aprox. cuasi-estática.

TESIS CON
 FALLA DE ORIGEN

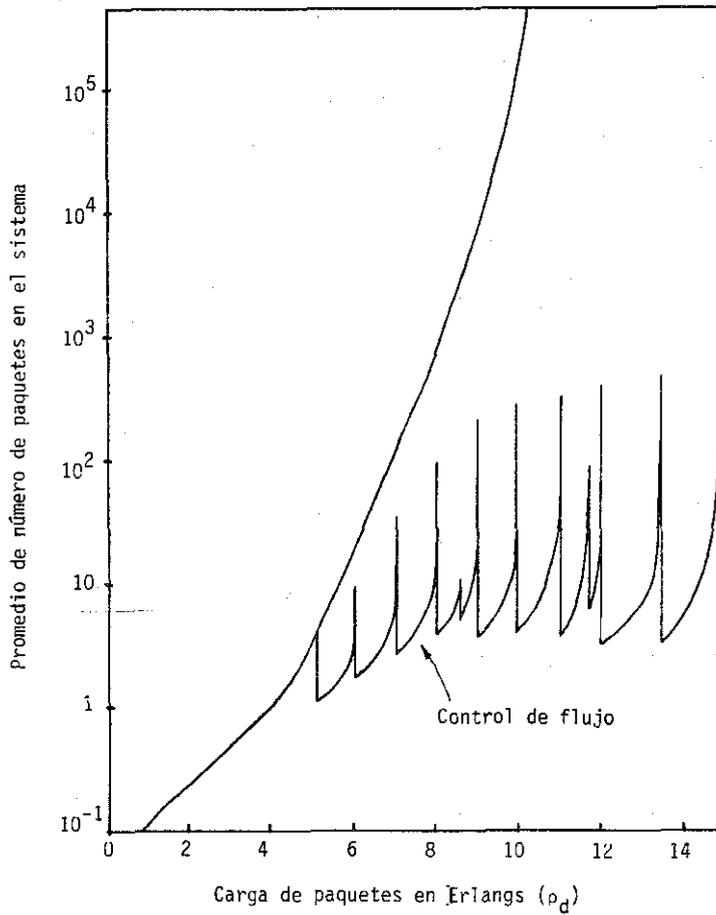


Fig. 4.11 Comparación de número promedio de paquetes en el sistema con y sin control de flujo, para caso de un buffer infinito del mismo sistema de Fig.4.7.

TESIS CON
FALLA DE ORIGEN

CAPITULO 5

CONCLUSION Y TRABAJOS FUTUROS

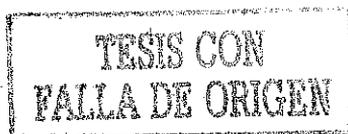
Esta tesis está concentrada en el análisis del desempeño de enlaces de comunicación con conmutación híbrida. La integración de los servicios de voz, datos y video en un ISDN (Integrated Service Data Network) es un campo de investigación de mucho interés. Este tipo de sistemas integrados, ofrece mejor servicio a los usuarios, así como mayor eficiencia de transmisión y menor costo, comparando con un sistema convencional separado.

Una estrategia adecuada de multiplexaje es muy importante para obtener un mejor desempeño del sistema. En el capítulo 3, propusimos un esquema de multiplexaje, el cual consiste en una frontera móvil bidireccional en el marco, en lugar de una frontera móvil uni-direccional de FM anterior. Un modelo de cadena de Markov bidimensional, nos permite evaluar el desempeño del sistema. La comparación del esquema propuesto FM2 y el esquema anterior FM se efectuaron en varios aspectos. Los resultados muestran que FM2 es más flexible que FM, y puede lograr algunos mejoramientos del desempeño, en un medio de tráficos variables.

En un medio de tráfico fijo (las cargas de voz y paquetes son constantes), el esquema FM puede diseñarse para lograr un desempeño óptimo o cerca del óptimo, utilizando los criterios de optimalidad de [24]. Pero, en un medio de tráficos variables, como modelamos en el capítulo 3, FM es muy difícil llegar a los puntos óptimos. Creemos que el esquema propuesto FM2 podría ser una opción.

De hecho, el esquema FM es el caso especial de FM2, porque un esquema de FM2 puede convertirse en FM, si ponemos $S_1=S$ en el esquema de FM2. La implementación de multiplexores de FM2 es un poco más complicada que la implementación de FM, pero su método es similar que FM.

En el capítulo 4, hemos ampliado el concepto de redes híbridas a un sistema que maneja tres tipos de tráficos utilizando CC tipo pérdida, CC tipo espera y CP. El esquema de multiplexaje que utilizamos en este tipo de sistemas, es similar que el de Frontera-Móvil, debido a la sencillez del análisis de desempeño, algunas otras estrategias también podrían implementarse. El modelo de cadena de Markov tri-dimensional, cual se puede reducir a la de dimensión dos, nos permite evaluar el desempeño del sistema, usando algunas técnicas numéricas para la solución, o aproximación cuasi-estática. La técnica de diagonalización de matrices puede evitar las dificultades computacionales en cuanto los valores característicos de matrices son mayores o iguales

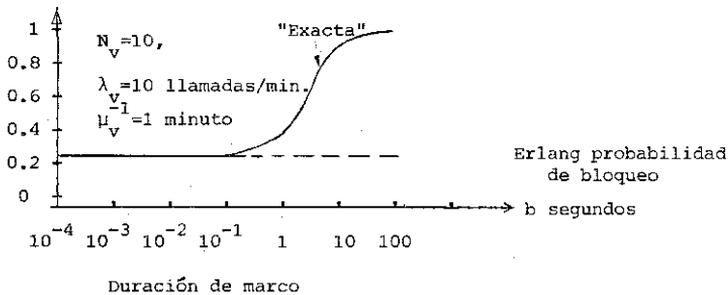


709

que uno, y el tamaño de buffer es muy grande. Sin embargo, esta técnica todavía no es muy eficiente debido a que tiene obtener los valores y vectores característicos de matrices. Por otra parte, la aproximación cuasi-estática es mucho más fácil de aplicar, comparado con la técnica de diagonalización de matrices. La única limitación de la aproximación cuasi-estática es que este método sólo se puede usarse cuando todos los modos (ver sección 4.4) son estacionarios.

La técnica de control de flujos es una manera para evitar el congestionamiento del sistema. El esquema de control de flujo que aplicamos en las redes híbridas de tres tráficos, no es una manera óptima, pero el resultado son satisfactorios. Los flujos de paquetes que no se los aceptaron para acceder al enlace, no implica que se bloquearon realmente, sino se desviaron a otros enlaces de la red que tienen menor tráfico.

En los dos modelos que presentamos para analizar los desempeños de los sistemas, no se considera el efecto de la cuantización de tiempo introducido por la estructura de marcos (esencialmente el efecto de la duración del marco). En la realidad, el desempeño del sistema no se altera mucho, si la duración de marco es relativamente pequeña. Por ejemplo, la probabilidad "exacta" de bloqueo de las llamadas es dependiente de la duración de marco (b segundos) [35], pero con la aproximación utilizando la fórmula de pérdida de Erlang, los dos resultados son casi iguales, si $b < 0.1$ segundos, la siguiente gráfica muestra este efecto.



Todavía hay muchos trabajos para futuro en este campo tan amplio, como el de la red de comunicación. Algunos posibles estudios en el futuro o sugerencias podrían ser:

- 1) Control de flujo y enrutamiento en redes híbridas todavía es una área que se necesita estudiar más, especialmente el enrutamiento, que es muy importante para lograr un buen sistema.
- 2) En base de los hardwares disponibles para la TASI (Time Assignment Speech Interpolat_

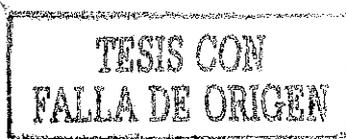
tion) [17], podemos aprovechar algunas en los periodos de silencio de las conversaciones en voz. El procedimiento y el efecto sobre la integridad de voz sería interesante.

- 3) Una selección óptima de la duración de marco y números de ranuras del marco, así como el diseño óptimo del esquema de multiplexaje, en función de los requerimientos, son importantes en cuanto la implementación de sistemas.
- 4) El concepto de redes híbridas de tres tráficos es una idea nueva, se requiere muchos estudios.
- 5) La técnica de diagonalización de matrices que usamos para resolver el modelo del sistema de tres tráficos, fue originalmente desarrollada para un sistema de dos tráficos. En caso de tres tráficos, la matriz K (ver ecuación (4.10)) es de dimensión $2(N_v+1)(m_f+1)$. En la práctica, para los m_f o N_v grandes, esta técnica tiene serios problemas en cuanto la computación. Encontrar una nueva técnica sencilla será importante.
- 6) Análisis de desempeño de una red híbrida, con un enfoque global, es un problema que todavía no ha resuelto. Para modelar este tipo de sistemas, una sugerencia podría ser la aplicación de la teoría de redes de colas [36], o también es práctico usar simulación.
- 7) Análisis de tráficos de voz y paquetes de datos en un canal de multiacceso (o acceso aleatorio) [37], es una área de investigación que apenas está empezando, las aplicaciones en las comunicaciones de satélites y radios, especialmente en redes locales, son muy amplias.
- 8) Sobre todo, la implementación de una red híbrida comercial es el objetivo principal de estas investigaciones. La estandarización y el diseño a nivel de protocolos, utilizando algunos esquemas nuevos y eficientes, es un trabajo muy práctico.

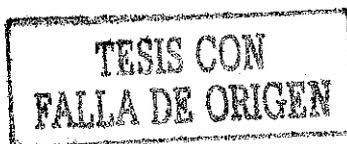
TESIS CON
FALLA DE ORIGEN

REFERENCIAS

1. Gerla, M. "The design of store-and-forward (S/F) networks for computer communications", U.C.L.A. Report, AD 758704., 1973.
2. Vijay Ahuja, Design and analysis of computer communication networks , McGraw-Hill, 1982.
3. Chang, L. H. "Analysis of integrated voice and data communication network", Ph. D. Dissertation, Dep. Elec. Eng., Carnegie-Mellon Univ., Pittsburgh, PA, 1977.
4. Martin, J. T. System analysis for data transmission , Englewood Cliffs, N. J. Prentice-Hall, 1972.
5. Gerla, M. y G. de Stasio, "Integration of circuit and PACUIT transport protocols in the TRAN data network", Proc. of the Comput. Net. Protocol Symposium, Liege, Belgium, Feb. 1978.
6. Vander Mey, J. E. "The architecture of a transparent, intelligent network", Proc. of the NTC'76, Dallas, Texas, Nov. 1976.
7. Coviello, G. J. y Vena, P. A. "Integration of circuit/packet switching in a SENET (Slotted Envelope NETWORK) concept", NTC'75, New Orleans, Dic. 1975.
8. Barbacci, M. R. y Oakley, J. D. "The integration circuit and packet switching networks toward a SENET implementation", The 15th NBS-ACM Annual Technique Symposium, 1976.
9. Otterman, J. "Grade of service of direct traffic mixed with store-and-forward traffic", IRE Trans. Commun. Syst., vol. CS-10, Mar. 1962, pp.130-138.
10. Zafiropulo, P. "Flexible multiplexing for networks supporting line-switched and packet-switching data traffic", Proc. of the 2nd ICCS, Stockholm, Sweden, Ago. 1974, pp.517-523.
11. Port, E., Kummerle, K., Rudin, H., Jenny, C. and Zafiropulo, P. "A network architecture for the integration of circuit and packet switching", Proc. of ICCS, Toronto, Canada, Ago. 1976, pp.505-514.
12. Kummerle, K. "Multiplexor performance for integrated line- and packet-switched traffic", Proc. of the 2nd ICCS, Stockholm, Sweden, Ago. 1974, pp.507-515.
13. Fisher, M. J. y Harris, T. C. " A model for evaluating the performance of an integrated circuit- and packet-switched multiplex structure", IEEE Trans. on



- Commun., vol. COM-24, Feb. 1976, pp. 195-202.
14. Alan, G. K. y Raymond, L. P. "Analysis of integrated voice/data multiplexing", IEEE Trans. on Commun., vol. COM-32, Feb. 1984, pp.140-147.
 15. Weinstein, C. J., Malpass, M. L. y Fisher, M. J. "Data traffic performance of an integrated circuit- and packet-switched multiplex structure", IEEE Trans. on Commun., vol. COM-28, Jun. 1980, pp.873-878.
 16. Bhat, U. N. y Fisher, M. J. "Multichannel queueing systems with heterogeneous classes of arrival", Naval Res. Logist Quart., Jun. 1976.
 17. Williams, G. F. y Leon-Garcia, A. "Performance analysis of integrated voice and data hybrid-switched links", IEEE Trans. on Commun., vol. COM-32, Ene. 1984, pp.695-706.
 18. Janakiraman, N., Paturek, B. y Neilson, J. E "Performance analysis of an integrated switch with fixed or variable frame rate and movable voice/data boundary", IEEE Trans. on Commun., vol. COM-32, Ene. 1984, pp.34-39.
 19. Maglaris, B. y Schwartz, M. "Performance evaluation of a variable frame multiplexor for integrated switched networks", IEEE Trans. on Commun., vol. COM-29, Jun. 1981, pp.800-807.
 20. Miyahara, H. y Hasegawa, T. "Performance evaluation of modified multiplexing technique with two types of packet for circuit and packet switched traffic", Int. Conf. Commun. Rec., Boston, MA, Jun. 1979.
 21. Leon-Garcia, A., Kwong, R. H. y Williams, G. F. "Performance evaluation methods for an integrated voice/data link", IEEE Trans. on Commun., vol. COM-30, Ago. 1982, pp.1840-1858.
 22. Gerla, M. y Mason, D. "Distributed routing in hybrid packet and circuit data network", Proc. Comput. Commun., Washington, D.C., Sep. 1978.
 23. Gerla, M. y Pazos-Rangel, P. A. "Bandwidth allocation and routing in ISDN's", IEEE Commun. Magazine, vol. 22, Feb. 1984, pp.16-26.
 24. Maglaris, B. y Schwartz, M. "Optimal bandwidth allocation in integrated line- and packet-switched channels", Inter. Conf. on Commun., Boston, Jun. 1979.
 25. Little, J. D. C. "A proof of $L=\lambda W$ ", Operations Research, 9, 1961, pp.383-387.
 26. Kleinrock, L. Queueing systems, vol.1, New York: Wiley, 1975.
 27. Papoulis, A. Probability, random variable, and stochastic processes, McGraw-Hill, 1984.
 28. Zeng, C.-M., Kuhlmann, F. y Buzo, A. "A model for a queueing system with



- priority and two different arrival processes", Proc. of the IASTED Inter. Conf.: ASM'85, Montreal, Canada, Jun. 1985.
29. Itoh, K., et al. "An analysis of traffic-handling capacity of packet-switched and circuit-switched networks", Proc. of Datacomm'73, Third Data Commun. Symposium, St. Petersburg, Florida, Nov. 1973, pp.29-37.
 30. Kummerle, K. y Rudin, H. "Packet and circuit switching: cost/performance boundaries", Computer Networks, vol. 2, 1978, pp.3-17.
 31. Gantmacher, F. R. The theory of matrices , vol. 1, New York: Chelsea, 1977.
 32. Neuts, M. F. "The M/M/1 queue with randomly varying arrival and service rates", Operation Research, vol. 15, No.4, 1978, pp.139-157.
 33. Neuts, M. F. "Further results on the M/M/1 queue with randomly varying rates", Operation Research, vol. 15, No. 4, 1978, pp.158-168.
 34. Kleinrock, L. Queueing systems , vol. 2, New York: Wiley, 1976.
 35. Occhiogrosso, B., Gitman, I., Hsieh, W. y Frank, H. "Performance analysis of integrated switching communications systems", Proc. of the NTC'77, L.A., 1977, pp. 12:4.1-12:4.13.
 36. Baskett, F., Chandy, K. M., Muntz, R. R. y Palacios, F. "Open, closed, and mixed networks of queues with different classes of customers", Journal of the ACM, vol. 22, No. 2, Abr. 1975, pp. 248-260.
 37. Gallager, R. G. "A perspective on multiaccess channels", IEEE Trans. on Inf. Theory, vol. IT-31, No. 2, Mar. 1985, pp. 124-142.

TESIS CON
FALLA DE ORIGEN