

99
2ef.



UNIVERSIDAD NACIONAL AUTONOMA
DE MEXICO

FACULTAD DE INGENIERIA

DESEMPEÑO DE LOS METODOS DE ACCESO
MÚLTIPLE ALEATORIO Y POR DIVISION DE
TIEMPO VIA SATELITE

T E S I S
QUE PARA OBTENER EL TITULO DE
INGENIERO MECANICO
ELECTRICISTA

P R E S E N T A N
JOAQUIN ERNESTO LEYVA MACIAS
RAUL STAVOLI PATRON

TESIS CON
FALLA DE ORIGEN

MEXICO, D.F.

1997



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.

*A mis Padres
A mis Hermanos
A mis Amigos*

Joaquín Ernesto Leyva Macías

A mis Padres
A mis Abuelos

Raúl Stavoli Patrón

Índice

Introducción	iii
1 Esquemas de Acceso Múltiple	1
Técnicas de Multiplexaje y Conmutación, 1	
Conmutación de Circuitos, 4	
Conmutación de Paquetes, 5	
Protocolos de Acceso Múltiple, 8	
Análisis de Métodos de Acceso Vía Satélite, 17	
2 Técnicas de Acceso Aleatorio I	19
Aloha Puro, 19	
Aloha Ranurado, 25	
Capacidad Excedida, 31	
Apéndice A, 33	
3 Técnicas de Acceso Aleatorio II	37
Protocolos de Reservación, 37	
Aloha de Retransmisión Selectiva, 48	
Apéndice B, 57	
4 Teoría de Colas	62
Líneas de Espera M/M/1, 64	
Líneas de Espera M/G/1, 72	
5 Canal TDMA	78
Tiempo de Espera con Memoria Finita, 80	
Resultados Numéricos, W_q y P_B , 84	
Tiempo de Retraso, Canal TDMA, 98	
Resultados Numéricos, Canal TDMA, 104	

Conclusión

110

Referencias

113

INTRODUCCION

El tema principal de este trabajo de tesis es analizar los métodos de acceso por división de tiempo. Este tipo de protocolos se encuentran en una importante nueva etapa de desarrollo especialmente por dos motivos: la evolución de los sistemas de comunicaciones personales (PCS) y de los protocolos de conmutación de paquetes como Frame Relay e IP.

El objetivo de esta tesis es evaluar el desempeño de los diversos esquemas de acceso múltiple, específicamente en términos del tráfico efectivo en el canal (throughput) y el tiempo de respuesta. El análisis cubre todo el espectro de esquemas, desde los aleatorios más simples hasta los de asignación fija.

Los esquemas de acceso múltiple se han diseñado para asignar el uso del canal entre un grupo de usuarios. Estos usuarios tienen requerimientos de tráfico aleatorio y restricciones en el tiempo de respuesta. El problema central es la resolución del "conflicto" entre usuarios para acceder el canal. Estos esquemas pueden diseñarse en forma independiente de funciones de comunicación tales como modulación, conversión de protocolo y calidad de la señal en términos de relación señal-ruido.

El diseño de estos sistemas tiene como objetivo establecer una capacidad óptima del canal que resulte en una alta eficiencia sin congestión excesiva. El diseño de red busca lograr un balance entre alto grado de utilización y un aceptable nivel de congestión bajo condiciones pico.

El criterio de selección de un protocolo satelital se basa especialmente en el tipo de tráfico y el desempeño requerido por los usuarios. Cuando el tráfico de la red es altamente predecible, en cuanto al ancho de banda requerido y el tiempo de utilización, es más factible coordinar el uso eficiente del canal mediante una técnica relativamente simple.

El criterio de selección de un protocolo satelital se basa especialmente en el tipo de tráfico y el desempeño requerido por los usuarios. Cuando el tráfico de la red es altamente predecible, en cuanto al ancho de banda requerido y el tiempo de utilización, es más factible coordinar el uso eficiente del canal mediante una técnica relativamente simple.

El precio a pagar por esta eficiencia es el procesamiento adicional tales como una forma de identificación y de formado de paquetes de datos, así como el requerimiento de capacidad de memoria temporal en cada nodo para almacenar los datos que se acumulen en la línea.

La formación de paquetes se realiza por un protocolo de nivel de red en el equipo banda base del nodo de comunicación considerando las condiciones de acceso al canal satelital.

En el caso opuesto, un bajo nivel de predictibilidad implica ya sea un uso menos eficiente del ancho de banda o técnicas más sofisticadas de asignación de tiempos de transmisión.

El capítulo 1 está dedicado a la clasificación de esquemas de acceso en general, una breve descripción y su relación con técnicas de multiplexaje y de conmutación. Se introducen además los aspectos básicos de acceso aleatorio.

Los sistemas de Acceso Aleatorio son analizados con detalle en los capítulos 2 y 3. El capítulo 2 está dedicado a los esquemas Aloha en su versión simple y en el caso ranurado. Se le da una particular atención al factor de retransmisión de paquetes que han sufrido colisión en el canal y su efecto en el cálculo de tráfico efectivo y tiempo de respuesta.

En el capítulo 3 se estudian los accesos que siendo aleatorios tienen un alto nivel de complejidad. Estos incluyen los protocolos SR Aloha, de Retransmisión Selectiva, orientados a optimizar la eficiencia del canal y el Aloha de Reservación, una forma híbrida de acceso aleatorio y determinístico de tipo sincrónica y que fue el primer diseño de división inteligente y variable de tiempos de transmisión en el canal.

Uno de los aspectos analizados con detalle son los diferentes enfoques que se le han dado en la caracterización del tráfico en el canal, tales como Teoría de Colas y las condiciones de equilibrio de estado estable, además de la suposición general de la estadística de Poisson del tráfico generado.

En el capítulo 5 se analiza el extremo determinístico de la asignación de tiempos de transmisión. El tema principal de este capítulo es la derivación de la fórmula del cálculo del tiempo de respuesta de un canal TDMA. El cálculo está basado en el trabajo original de S. Lam que es la más notable aportación en esta disciplina. En este capítulo se analiza además un aspecto que normalmente es considerado en sistemas de conmutación de circuitos. Este parámetro es la probabilidad de bloqueo de un mensaje en el sistema en función de la capacidad de memoria en el nodo de comunicación.

Capítulo 1

Esquemas de Acceso Múltiple

Técnicas de Multiplexaje y Conmutación

Las redes de comunicación de datos han experimentado una acelerada evolución desde la aparición de las primeras aplicaciones de transmisión de datos que comunicaban terminales remotas con una computadora central.

La función básica de una red de comunicación es la de proporcionar una línea de acceso entre usuarios separados geográficamente. Esta línea de acceso significa un grupo de funciones que permite a un usuario estar no solo físicamente conectado a una red sino tener comunicación total superando desde interferencias en el canal físico hasta diferencias en formatos de velocidad de transmisión y de protocolos.

Una importante función desempeñada por la red de comunicación es la verificación de la comunicación en los extremos de un canal. Esta es una de las funciones de los protocolos de "nivel de datos" (DLC), que confirman que la señal de datos en la recepción es una réplica exacta de la señal original transmitida. Los bits son agrupados en grupos (tramas) para su transmisión.

La función que es precisamente objeto de este estudio es la distribución del canal de comunicación entre varios usuarios con necesidades de comunicación variables e intermitentes. Estos esquemas se llaman de acceso múltiple y han sido diseñados para asignar el uso del canal entre un grupo de usuarios que comparten el medio de comunicación.

Adicionalmente se encuentran las funciones de enrutamiento, corrección de error, almacenamiento temporal y control de flujo.

El rápido incremento del tráfico de comunicación digital motivó el interés en la optimización de recursos de comunicación de datos y el desarrollo de métodos más eficientes de cursar este tipo de señales.

Uno de los más importantes esfuerzos se ha enfocado en los métodos de conmutación de y enrutamiento de paquetes como un método efectivo de manejo del tráfico de datos en un ambiente que originalmente estuvo dedicado a la conmutación de circuitos.

La eficiencia de una técnica de conmutación en particular esta en función de la estadística del volumen y la frecuencia de los mensajes de datos a transmitir. En forma simplista, es más eficiente utilizar conmutación de paquetes para mensajes cortos y conmutación de circuitos para mensajes largos, o grupos de mensajes cortos pero transmitidos a una frecuencia tal que justifique conservar la conexión.

El punto de equilibrio entre mensajes cortos y largos depende de las características de cada red en particular y del tiempo que toma establecer un circuito.

El canal de transmisión constituye el recurso primario de una red de comunicaciones. Este recurso puede consistir en un canal de comunicación punto a punto, una red de cierta topología de enlaces punto a punto, un canal de difusión (broadcast) o una combinación de los mismos. Entre los aspectos más importantes en el diseño de una red es la selección de la técnica de conmutación.

Tecnología de multiplexaje. Todo canal de comunicación puede ser multiplexado en un cierto número de subcanales. Existen dos tipos básicos de multiplexaje: en frecuencia y en tiempo.

El multiplexaje en frecuencia (FDM) divide el canal de comunicación en subcanales a los que corresponde un segmento del ancho de banda del canal todo el tiempo.

En el multiplexaje en tiempo (TDM) el tiempo de transmisión es dividido en segmentos de tiempo (slots), de esta forma el ancho de banda total del canal es disponible para uno de los subcanales durante el "slot" asignado. Estos "slots" son organizados en periodos de tiempo llamados "tramas".

Asignación de canales.

Asignación fija. Esta es la forma más simple de asignación de recurso de canal. El canal de comunicación a ser compartido es multiplexado en un cierto número de subcanales, suficiente para asignar uno a cada pareja de transmisor-receptor en cada extremo del enlace. Asignación fija no significa que los subcanales sean del mismo "tamaño", ya sea en tiempo o en frecuencia.

En comunicación de datos la asignación fija es propensa a uso ineficiente, especialmente porque los usuarios de cada subcanal pueden no tener información a transmitir durante una considerable mayoría del tiempo, sino en forma interactiva, con lapsos de inactividad en los que el subcanal no es utilizado. El tráfico interactivo tiene una importante aplicación de comunicación de datos.

Esta ineficiencia puede ser minimizada por el método de asignación de canal por demanda. En este tipo de asignación se dispone de un número de subcanales menor al de parejas de usuarios transmisor-receptor, asumiendo que únicamente un subconjunto de ellos utiliza el subcanal en un momento determinado. De esta forma los subcanales se asignan únicamente cuando hay una solicitud de inicio de transmisión y son desconectados a la terminación de dicha transmisión.

Los subcanales de asignación por demanda tienen una mayor utilización que los canales de asignación fija, para este efecto se requiere una capacidad adicional de procesamiento en los nodos de comunicación. Se debe considerar además un retraso (o rechazo como en el caso de telefonía) en el servicio de comunicación cuando una solicitud de transmisión no pueda ser atendida por la indisponibilidad de subcanales.

Multiplexaje estadístico

Esta técnica está orientada a la asignación de tiempo de transmisión únicamente cuando tiene datos a transmitir. El nodo de conmutación es referido como concentrador de datos y el modelo consiste en una línea de espera de un solo servidor. El canal de comunicación es el servidor y la línea es formada por los datos generados por los usuarios que comparten el canal. Este método permite utilizar en forma eficiente el canal de comunicación aun cuando cada usuario genera tráfico intermitente en forma individual.

El precio a pagar por esta eficiencia es el procesamiento adicional necesario. Como las unidades de datos correspondientes a cada pareja de transmisor-receptor son mezcladas en la línea de espera del canal, deben contener alguna forma de identificación. Además, se requiere contar con memoria temporal en cada nodo para almacenar los datos que se acumulen en la línea.

La superioridad del multiplexaje estadístico se basa en el fenómeno de escalamiento en teoría de colas. El desempeño retraso-tráfico efectivo ("throughput") de un canal de alta velocidad es mejor que el de varios canales de baja velocidad, aun cuando el ancho de banda total en ambos casos es el mismo.

Conmutación de Circuitos.

Esta técnica de conmutación para datos establece un circuito directo a través de la red y es ampliamente utilizada en redes de tráfico telefónico.

En general, un enlace de circuito conmutado está comprendido por una serie de sub-canales cuya secuencia se determina al inicio de la conexión. Estos canales pueden ser de asignación fija en forma indefinida o por demanda cuando se presenta una solicitud de diálogo entre dos usuarios.

Tiene la obvia ventaja de ser compatible con los conceptos de las redes de comunicación de voz e incluso puede utilizar los mismos equipos y canales.

El procedimiento de establecer el enlace es común al de un canal de voz. Esta técnica requiere compatibilidad absoluta (código, tasa y formato) entre los usuarios terminales ya que la red no efectúa procesamiento alguno y esta sujeta a bloqueo, ya sea internamente o en la conexión terminal, cuando el equipo está en comunicación a través de otro circuito.

La conmutación de circuitos es apropiada cuando los usuarios tienden a comunicarse a una tasa aproximadamente constante por un tiempo relativamente largo, comparado con el tiempo que toma establecer el enlace.

El desempeño de una red de circuito conmutado se cuantifica por el porcentaje de llamadas que exceden la capacidad de la red. El modelo más empleado en el diseño de estas redes es el de Erlang, que relaciona la capacidad de la red en número de canales, la intensidad de tráfico de los usuarios y el porcentaje de bloqueo de llamadas.

Los parámetros que definen la intensidad promedio de tráfico son la tasa de llamadas generadas (cuya distribución de probabilidad se considera de Poisson) y el tiempo de duración de las llamadas (con densidad exponencial de probabilidad).

Conmutación de Paquetes.

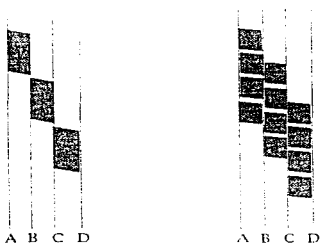
La técnica de conmutación de paquetes consiste en dedicar parcial o totalmente la capacidad de transmisión de un sistema a un usuario a la vez por un período corto de tiempo.

La implementación de una red de paquetes requiere de funciones adicionales como el formato de paquetes a partir de los datos de usuario y la conversión inversa, de paquetes a unidades de datos originales. En este esquema los nodos efectúan el procesamiento de los paquetes de cada mensaje.

Cada paquete contiene información sobre el nodo al que los datos están dirigidos. los paquetes atraviesan una secuencia de canales y nodos. Los paquetes pueden ser retenidos temporalmente en un nodo de la red hasta que el canal se encuentre disponible.

Estas unidades de datos experimentan un cierto retraso durante la transmisión a través de la red debido al tiempo de procesamiento y de conmutación.

Tanto la capacidad de procesamiento de paquetes en los nodos (conmutación y memoria) y la capacidad de los canales que los unen son multiplexados estadísticamente por los diversos usuarios.



CONMUTACION DE MENSAJES
Y CONMUTACION DE PAQUETES
A TRAVES DE LOS NODOS DE LA RED

En las redes de comunicación por división de tiempo hay dos subcategorías: de conmutación de mensajes y de conmutación de paquetes. Una red es de mensaje conmutado cuando el mensaje completo constituye la unidad de datos transmitida a través de la red.

Una red de conmutación de paquetes es la que maneja un tamaño máximo de paquete como unidad de transmisión. Los mensajes cuyo tamaño excede el del paquete son fragmentados. Esta partición de mensajes en paquetes es efectuada por un protocolo de nivel red o de mayor nivel.

La técnica de conmutación de paquetes presenta ventajas sobre la de mensajes, requiriendo funciones adicionales de procesamiento:

La conmutación de paquetes reduce el retraso entre puntos terminales. Como se ilustra en el diagrama, los diferentes paquetes de un mensaje pueden ser procesados en paralelo por distintos nodos. La reducción del tamaño de las unidades de datos facilitan el control de capacidad de memoria de los nodos de la red.

El diseño de estas redes tiene como objetivo establecer una capacidad óptima del canal que resulte en una alta eficiencia sin congestión excesiva.

Los sistemas de conmutación de paquetes son diseñados para establecer un recurso que es suficiente para cubrir la demanda promedio de tráfico, no así la demanda pico. Es razonable considerar que esta última se presentará solamente en forma ocasional. El diseño de red busca lograr un balance entre alto grado de utilización y un aceptable nivel de congestión bajo condiciones pico.

Cuando los usuarios de una red generan tráfico intermitente, un canal de conmutación de paquetes es utilizado más eficientemente que uno de conmutación de circuitos.

Los sistemas de conmutación de paquetes y de circuitos se consideran los extremos del espectro de técnicas de comunicación de datos.

Protocolos de Acceso Múltiple

Las redes de difusión de paquetes se definen como las redes de conmutación de paquetes en la que el canal que proporciona conectividad a los usuarios está constituido por un canal de difusión ("Broadcast"). Como ejemplos de medios de difusión están los canales vía satélite y los canales de radio.

El concepto de difusión de paquetes inicio su aplicación en forma extensa a partir del proyecto Aloha, que ejercio una influencia decisiva en el diseño de dos sistemas que utilizan esquemas de acceso múltiple en un canal de banda amplia: Comunicaciones de datos via satélite y el de las redes de área local.

El objeto de este estudio son los protocolos de acceso de un canal de difusión por un grupo de usuarios. Estos usuarios tienen requerimientos de tráfico aleatorio y restricciones en el tiempo de respuesta. El problema central es la resolución del "conflicto" entre usuarios para acceder el canal. Los metodos de resolución de esos conflictos son conocidos como protocolos de acceso múltiple. En esta sección se presenta una clasificación de esos protocolos, una breve descripción y una comparación de su desempeño ante diversas condiciones de tráfico y del canal mismo.

La medida principal del desempeño de un protocolo de acceso múltiple es su característica de tráfico efectivo ("throughput") contra retraso promedio.

Técnicas Tradicionales

El problema de acceso múltiple en el diseño de sistemas vía satélite ha sido resuelto con especial interés en la comunicación de voz. El objetivo del diseño es maximizar el número de canales para una potencia y ancho de banda determinados.

Las técnicas comunes de acceso múltiple vía satélite son de división de frecuencia (FDMA), de división de tiempo (TDMA), de división de código (CDMA), también llamada de Espectro Disperso ("Spread Spectrum").

TDMA es generalmente más eficiente que el resto de los métodos de acceso, el costo de esta eficiencia está en la complejidad del equipo requerido para su implementación.

Los protocolos de acceso múltiple han sido normalmente orientados al concepto de canal. La capacidad de transmisión disponible se subdivide en canales independientes, que en conmutación de paquetes son mapeados a circuitos virtuales. La forma de asignación de recursos del canal se clasifica a su vez en:

- a) Asignación Fija
- b) Asignación por Demanda.

Con asignación por demanda, se requiere establecer un canal para señalización entre usuarios. La asignación por demanda es realizable mediante un controlador central o con control distribuido.

Los protocolos de acceso múltiple orientados al canal (uso continuo del canal) son apropiados para comunicación de voz y para un cierto tipo de tráfico de datos y video.

La comunicación de datos en general tiene un amplio espectro de requerimientos de capacidad de transmisión en función de la aplicación, desde sistemas interactivos de consulta-respuesta con tráfico intermitente hasta transferencia de archivos de grandes volúmenes de datos.

Las aplicaciones de comunicación de datos han evolucionado a partir del tráfico interactivo de uso transaccional, en el que los nodos remotos cuentan con terminales que se comunican con un procesador central. Este tipo de tráfico resultó compatible con el concepto de conmutación de paquetes de tipo X.25 y con los métodos originales de acceso al canal en forma aleatoria. Las aplicaciones subsecuentes como de transferencia de archivos han requerido de un uso más continuo del canal, lo que ha influenciado el uso de esquemas de asignación determinística de tiempos de transmisión. Las aplicaciones de comunicación de voz y de nuevos

protocolos como Frame Relay requieren de tiempos de transmisión también determinísticos pero flexibles, lo que se ha dado en llamar "ancho de banda por demanda".

El presente estudio está orientado a los protocolos de conmutación de datos de usuarios que comparten el canal de difusión con un esquema de división de tiempo. El análisis de las técnicas de acceso múltiple tiene como medida de desempeño el tráfico efectivo y el retraso.

Modelo de Tráfico

En la descripción del modelo de tráfico, el término usuario incluye terminales de computadora, programas de computadora de interacción con operador u otro programa. Usuario se puede definir como elemento generador de tráfico cuyos tiempos de interacción se pueden modelar como una variable aleatoria. Se define además como mensaje a un bloque de datos con un tiempo de respuesta sujeto a restricción. En realidad, en una red de conmutación de paquetes, un mensaje puede consistir de varios paquetes.

Un cierto tipo de tráfico, especialmente el interactivo de consulta-respuesta se describe como intermitente. La intermitencia del tráfico es una importante característica considerada en el diseño de un sistema de comunicación de paquetes.

Características de la Red

En esta sección se describen las características y parámetros de operación de una red de canal de difusión y la nomenclatura utilizada. Se considera un grupo de N usuarios compartiendo el canal. El usuario "i" tiene una tasa de generación de mensajes λ_i [mensajes/segundo].

Cada paquete contiene una dirección de destino tal que cuando es transmitido exitosamente a través del canal sea recibido por el destinatario. Los errores debido al ruido del canal no son considerados.

Los requerimientos de conectividad de los usuarios son relevantes solamente en forma indirecta en función del conjunto de tasas de mensajes.

Para un conjunto de tasas $\{r_i\}$ dado, es irrelevante si los usuarios se comunican entre ellos o si lo quieren hacer a una estación central.

El aspecto de interés es el acceso al canal de difusión. Cada usuario tiene la posibilidad de enviar y recibir datos a la tasa de transmisión del canal, llamada C [bps].

En algunos de los protocolos TDMA los usuarios están sincronizados de manera que el canal se puede visualizar como una secuencia de ranuras de tiempo. La duración de las ranuras es suficiente para contener un paquete de datos. Se considera la utilización de ranuras de menor duración que se mezclen con las ranuras de datos para transmitir paquetes de menor tamaño para información de control.

Un usuario puede enviar datos a cualquier otro usuario cuando se le concede el acceso al canal. El protocolo de acceso es esencialmente un algoritmo para determinar el derecho de los usuarios para acceder el canal. En algunos protocolos, el derecho de acceso no está absolutamente determinado y son probables las colisiones entre paquetes de diferentes usuarios en el canal. Se asume que las colisiones impiden la recepción correcta de los paquetes implicados en la colisión.

El tiempo R [segundos] que toma detectar si un paquete transmitido ha sufrido una colisión se le denomina tiempo de detección de colisión.

En varios tipos de sistemas el usuario detecta el resultado de su propia transmisión monitoreando el canal, una de las características particulares de un canal de difusión. En este caso R es aproximadamente el tiempo de propagación de la señal en el canal. Cuando este tipo de detección no es utilizado, R corresponde al periodo de expiración de espera de respuesta positiva de recepción.

Clasificación de Protocolos

Los protocolos de acceso múltiple se pueden clasificar de la siguiente manera.

Protocolos de Acceso Múltiple

- **Usuarios Pasivos**
 - Encuesta
 - Encuesta Adaptiva

- **Usuarios Activos**
 - Contención
 - Reservación

Protocolos para Usuarios Pasivos

Este tipo de protocolos es conocido como protocolos de encuesta ("polling"). Se requiere de la presencia de un controlador central. Los usuarios son pasivos en el sentido de que pueden acceder el canal únicamente cuando les es específicamente solicitado por el procesador central.

El controlador tiene dos funciones, la primera es identificar usuarios con datos a enviar, la segunda es programar el uso del canal por esos usuarios.

Protocolos de encuesta convencional. En este protocolo el controlador central las dos funciones mencionadas encuestando el grupo de usuarios uno por uno. A la llegada de un mensaje de encuesta, el usuario asignado transmite todos los mensajes acumulados en su memoria temporal. Al tiempo necesario para encuestar todos y cada uno de los usuarios una vez y transmitir sus datos acumulados se le llama ciclo de encuesta tc. Se puede medir como el tiempo entre dos encuestas sucesivas a un mismo usuario.

Hay variantes adicionales de estos protocolos que incluyen prioridad en el orden de encuesta, frecuencia de encuesta diferente entre los usuarios y limitaciones en el volumen de datos a transmitir por ciclo de encuesta.

Protocolos para usuarios activos

Los protocolos a ser descritos en esta sección son los empleados por usuarios que buscan el acceso al canal en lugar de esperar una solicitud de transmisión.

Estos protocolos son llamados de contención y de reservación. En los primeros, no hay coordinación entre usuarios para evitar una colisión, cada usuario decide cuando acceder al canal. Hay sin embargo mecanismos de transmisión para minimizar interferencia entre usuarios.

Protocolos de Contención

Estos protocolos son adecuados para un número grande de usuarios de tráfico intermitente. Para una red de tráfico reducido, el desempeño de un protocolo de contención en términos de retraso llega a ser muy superior al de los protocolos de encuesta.

El protocolo Aloha.

En el protocolo Aloha los usuarios no cuentan con sincronización alguna. Cada usuario transmite un paquete de datos en cualquier momento que este se genere. Si ninguna otra estación transmite durante ese tiempo, el paquete será recibido sin problema en la estación destino.

En el caso de que dos o más paquetes accedan el canal y sufran una colisión, cada usuario involucrado se entera después de un cierto tiempo de detección R y retransmite su paquete después de un tiempo aleatorio de espera. Este tiempo de espera tiene especial importancia en cuestión de estabilidad y en la relación de desempeño tráfico efectivo-retraso de todos los protocolos de contención.

El diseñador del protocolo Aloha, Norman Abramson, determinó el tráfico efectivo máximo de un canal en el caso límite de un número muy grande de usuarios de tráfico muy bajo. Los mensajes consisten de un solo paquete. El proceso de generación de paquetes es un proceso de Poisson a una tasa de "S" paquetes por tiempo de transmisión.

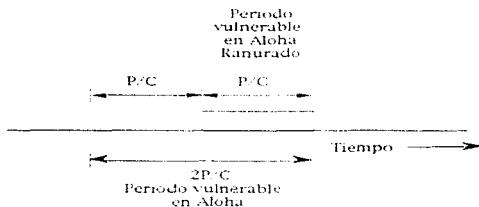
La otra suposición interesante es que la suma de nuevas transmisiones y retransmisiones en el canal se modelan como un proceso de Poisson

también a una tasa de "G" paquetes por tiempo de transmisión. Adicionalmente, se considera equilibrio estadístico.

La probabilidad de que un paquete sea transmitido sin colisión es:

$$S \cdot G = e^{-2G}$$

Como se ve en la figura, cada paquete transmitido tiene un periodo vulnerable que es el doble del tiempo de duración de un paquete. La transmisión de un paquete será exitosa únicamente si ningún otro paquete es transmitido dentro del periodo vulnerable.



El máximo tráfico efectivo posible en un canal Aloha se obtiene cuando $G=0.5$ y, bajo las condiciones anteriormente mencionadas:

$$C_A = 1/(2e) \approx 0.18$$

La máxima eficiencia de canal alcanzable para el acceso Aloha es un poco menor de $1/(2e)$ cuando se transmiten paquetes de tamaño variable. Por otro lado, los slots de tamaño fijo de Aloha ranurado resultan en degradaciones en la relación retraso - eficiencia.

En el caso de tráfico de paquetes de tamaño constante, la eficiencia se puede duplicar sincronizando los tiempos de inicio de ráfagas de transmisión de todas las estaciones.

Una desventaja del sistema de acceso Aloha es la necesidad de estabilizar el canal. Esta técnica de acceso es inestable si los intervalos de retransmisión son muy cortos en relación a la tasa de tráfico. En el otro extremo, intervalos demasiado largos pueden resultar en retrasos excesivos.

A diferencia de un sistema de asignación fijo, el desempeño de una estación en particular en una canal Aloha depende directamente del tráfico transmitido por el resto de las estaciones.

Protocolo Aloha Ranurado

Este protocolo tiene la característica adicional de que el tiempo de uso del canal está dividido en periodos de tiempo de aproximadamente la duración del paquete, se dice que está "ranurado".

Los usuarios deben sincronizar la transmisión de sus paquetes en intervalos uniformes de tiempo. Dada la sincronía de tiempos de inicio de paquete, las colisiones por interferencia parcial son evitadas y el periodo vulnerable de un paquete transmitido es solamente la duración de una ranura de tiempo.

Tomando las mismas consideraciones se tiene:

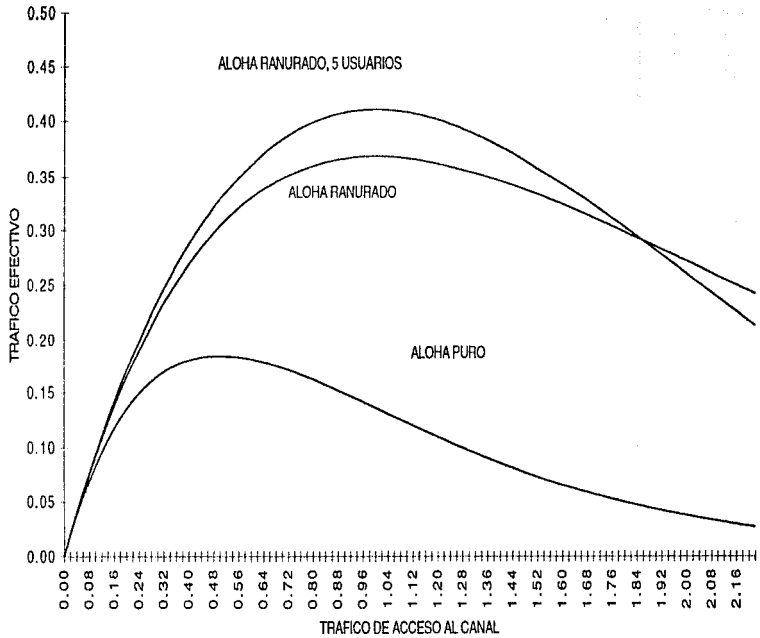
$$S/G = e^{-G}$$

donde S se maximiza cuando $G = 1$. El máximo tráfico efectivo es del doble del caso de Aloha puro:

$$C_{SA} = 1/e \approx 0.37$$

Aloha

EFICIENCIA DEL CANAL DE DIFUSION



La parte lineal de las curvas representa el desempeño realizable en operación estable. La sección de pendiente negativa de las curvas indica la región de operación en condiciones de inestabilidad, cuando la eficiencia decrece y el retraso se incrementa por el aumento de retransmisiones en el canal.

La selección entre los dos esquemas Aloha, en su forma simple o la versión ranurada depende de la relación entre el encabezado de la rafaga y el tamaño del mensaje, de la distribución de probabilidad del tamaño mismo y el punto de operación de la eficiencia seleccionado para satisfacer las consideraciones de retraso y estabilidad.

Una de las importantes funciones del esquema Aloha es que una estación en particular cuenta con la posibilidad de utilizar toda la capacidad del canal por un periodo arbitrario de tiempo, dependiendo de los requerimientos de transmisión de las otras estaciones durante ese tiempo.

En esos casos la eficiencia promedio del canal puede exceder $1/(2e)$ o $1/e$, dependiendo del esquema, sin presentar inestabilidad, cuando las tasas de transmisión de las estaciones no son iguales y el tráfico tiende a concentrarse en unas estaciones.

Esta "capacidad excedida" de un canal Aloha puede dar una eficiencia que se aproxima al 100% en el caso de un usuario con una alta tasa de transmisión de paquetes y el resto de los usuarios con una tasa baja. La desventaja la constituye el retraso que encuentran las estaciones de bajo volumen de tráfico, mayor que en el caso de estaciones de tasas uniformes.

Consideraciones de Desempeño.

El análisis de los sistemas Aloha puro y Aloha ranurado se basa en tres consideraciones básicas:

- La independencia estadística del tráfico del canal.
- El número de usuarios es muy grande.
- Se supone equilibrio estadístico.

Análisis de Metodos de Acceso Via Satelite.

Los protocolos analizados en este estudio consideran el uso de un sistema de satélites geostacionario.

El canal de comunicación que es el transpondedor del satélite es accesado por las estaciones terrenas de una manera determinada por el propio protocolo. Las señales transmitidas desde las estaciones terrenas hacia el satélite son amplificadas y re-transmitidas hacia el segmento terrestre de la red.

El canal satelital es compartido por las estaciones empleando un metodo de acceso por división de tiempo, sincrono o asincrono. Cada estación accesa un tiempo de utilizacion del canal en un conjunto de frecuencias común a la red tanto para transmision como para recepcion.

El canal satelital es compartido por varios usuarios. Cada nodo cuenta con un subsistema receptor por el que recibe realimentacion de su propio acceso y de comunicacion de retorno desde la estación central.

Estos canales de nodo central a nodos remotos y viceversa operan a diferente frecuencia que conforman un canal bidireccional "full duplex".

El criterio de seleccion de un protocolo satelital se basa especialmente en el tipo de tráfico y el desempeño requerido por los usuarios. Si el tráfico de cada estación es altamente predecible, en cuanto el ancho de banda requerido y el tiempo de utilizacion, es mas factible coordinar el uso eficiente del canal mediante una técnica relativamente simple.

En el caso opuesto, un bajo nivel de predictibilidad implica ya sea un uso menos eficiente del ancho de banda o técnicas más sofisticadas de asignación de tiempos de transmisión.

Entre los aspectos que influyen en el desempeño de un protocolo está la intermitencia del tráfico y el tamaño de los mensajes. Otros factores de desempeño son el equilibrio uso-necesidad de los usuarios en el canal y la flexibilidad y capacidad de crecimiento del tráfico de los usuarios.

Este estudio considera protocolos de comunicación de datos en paquetes. La formación de paquetes se realiza por un protocolo de nivel de red en el equipo banda base del nodo de comunicación considerando las condiciones de acceso al canal satelital.

Cada paquete contiene dirección e información de control. Puede tener restricciones de longitud máxima, seleccionada de acuerdo al protocolo utilizado.

Capítulo 2

Técnicas de Acceso Aleatorio I

En comunicaciones de datos de transmisión intermitente, de bajo "ciclo de trabajo", una técnica de acceso apropiada es el acceso aleatorio. Por la naturaleza estadística de los procesos de generación de datos y la característica intermitente de los mensajes, hay condiciones en las que la probabilidad de que dos usuarios transmitan al mismo tiempo es baja.

En la forma más simple de este tipo de acceso, cuando un paquete está listo para transmitir, simplemente se transmite sin usar sincronía en la red e independientemente del resto de los usuarios. Si ninguna otra estación transmite durante ese tiempo, el paquete será recibido sin problema en la estación receptora.

En el caso de que más de una estación sea la que transmite, ocurre una colisión y muy probablemente los paquetes no son detectados correctamente en la recepción.

La característica de canal de difusión del satélite permite la detección de la colisión por la estación transmisora, la que procede a la retransmisión del paquete.

Aloha Puro

Se considera que k usuarios están cada uno transmitiendo en forma independiente con probabilidad de inicio de mensaje de tipo Poisson a una tasa λ [mensajes/segundo].

Los mensajes se suponen de longitud fija, de τ [segundos] de duración. El máximo tráfico efectivo del canal, alcanzable únicamente en el caso extremo de que a un usuario se le concediera el acceso al canal y transmitiera paquetes sin interrupciones, sería obviamente $1/\tau$ [paquetes/segundo], lo que representa la capacidad del canal.

Quando varios usuarios tienen acceso aleatorio al canal, surge la posibilidad de transmisiones simultaneas. El trafico efectivo real en esas condiciones es considerablemente menor que el maximo posible.

El sistema de acceso Aloha se basa en la caracteristica ocasional de la transmision de mensajes de cualquier usuario. Considera la relación:

$$\lambda \tau \ll 1$$

que significa que la tasa de generacion de paquetes es muy inferior a la capacidad del canal.

Como en el multiplexaje estadístico, un grupo de usuarios pueden compartir el canal y utilizarlo en forma mucho más eficiente que un solo usuario ocasional. El precio a pagar es que el máximo tráfico efectivo obtenible con este esquema es mucho menor que la capacidad del canal. Un usuario podría utilizar el canal más eficientemente que el grupo de usuarios, si su transmisión es continua.

Con k usuarios, la utilización del canal (referida a la tasa máxima de transmisión de $1/\tau$) es:

$$S = k\lambda\tau < 1$$

S es el tráfico efectivo ("throughput").

Para el sistema Aloha puro, con un número muy grande de usuarios del mismo tráfico, el máximo tráfico efectivo es:

$$S_{max} = 1/2e$$

Esto significa que este esquema tiene aproximadamente un 18% de utilización de la capacidad del canal.

El esquema de transmisión se describe a continuación:

Un usuario espera un mensaje de reconocimiento ("acknowledgment") desde el nodo central (transmitido en modo "broadcast" también en un canal independiente, a todos los usuarios).

Si después de un tiempo de espera no recibe respuesta (este tiempo de espera debe ser por lo menos el tiempo de propagación del sistema en ambos sentidos), procede a repetir el mensaje original.

En un sistema de acceso aleatorio, la no recepción del reconocimiento se considera debida a que dos o mas usuarios accedieron el canal al mismo tiempo, resultando una colisión.

Una forma de evitar el mismo problema al retransmitir es programar tales retransmisiones en forma escalonada, la tecnica mas simple es agregar un lapso aleatorio a la expiración del tiempo de espera.

El efecto de las colisiones se refleja en la reducción del tráfico efectivo en el canal, que consiste de dos tipos de paquetes:

- Los generados originalmente (a un tasa promedio de S paquetes por intervalo de tiempo τ), y
- Los paquetes retransmitidos.

A medida que la generación de paquetes originales se incrementa (ya sea incrementando el número de usuarios k o la tasa de generación de mensajes λ), la probabilidad de colisión aumenta. Esto a su vez produce un aumento en el número de retransmisiones y por lo tanto el número de colisiones, esto produce un efecto de inestabilidad en el canal. Es precisamente este efecto el que limita el tráfico efectivo del canal.

La clasificación del tráfico en sus dos formas es útil en la derivación del máximo tráfico efectivo.

G se define al tráfico (original) promedio accedando el canal por intervalo τ

$$G = S + \text{promedio de retransmisiones en } \tau$$

Una retransmisión ocurre debido a una colisión. Ya que los mensajes consisten de un paquete de τ segundos de duración, una colisión ocurre si otro usuario decide transmitir en un intervalo $\pm \tau$ alrededor del tiempo de iniciación del mensaje.

La siguiente suposición es que el tráfico total en el canal tiene una función de probabilidad de Poisson, con parámetro G . (Esto no es generalmente válido en las retransmisiones, pero es aplicable en muchos casos).

La probabilidad de que ocurra por lo menos una colisión (debido a otros paquetes generados dentro del intervalo de 2τ) es entonces: $1 - e^{-2G}$. El número promedio de retransmisiones es $G(1 - e^{-2G})$, por lo tanto:

$$G = S + G(1 - e^{-2G})$$

Lo que resulta en un tráfico promedio de:

$$S = Ge^{-2G}$$

Esta es la ecuación del tráfico efectivo para un sistema Aloha puro operando en condiciones de estabilidad. Una forma alternativa de obtenerla es la siguiente, de acuerdo a la siguiente definición:

G como el promedio del tráfico total en el canal y S como la fracción de G correspondiente al tráfico original (paquetes nuevos) entonces S/G es la fracción de paquetes que corresponden a una transmisión exitosa, que es igual a la probabilidad de transmisión sin colisión en el primer intento. Es exactamente la probabilidad de que una colisión no ocurra, que es precisamente e^{-2G} en estadística de Poisson.

El otro aspecto de marcado interés es el retraso en el sistema, adicional al tiempo de propagación y despreciando el tiempo de procesamiento. Denominando E al número promedio de retransmisiones por paquete, el promedio de intentos de retransmisión por paquete generado es:

$$1 + E = G/S = e^{2G}$$

Esta es una medida del retraso en el sistema. A medida que el tráfico efectivo S se aproxima al valor máximo de 0.18, el retraso se incrementa proporcionalmente. Como una cota mínima en el retraso adicional al de propagación, en unidades de intervalos de τ segundos es:

$$D_{\min} = e^{2G}$$

Incluyendo los tiempos de retraso adicionales:

- El de propagación en ambas direcciones R (en intervalos de τ segundos), requerido para recibir un mensaje positivo de recepción.
- El de espera antes de retransmisión, suponiéndolo una variable aleatoria uniformemente distribuida de 1 a k intervalos. El promedio de este tiempo de espera escalonado es $(1+k)/2$ intervalos.

El tiempo promedio de retraso es:

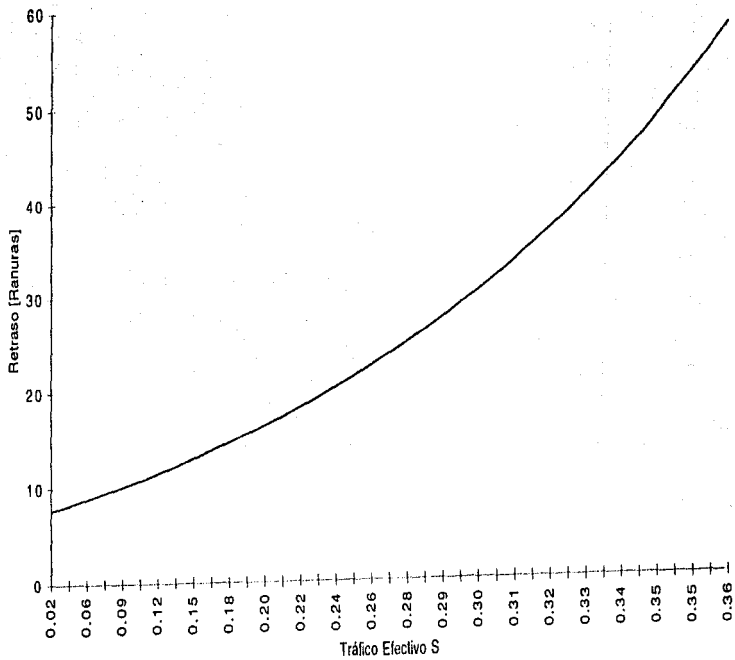
$$D = R + 1 + E \left(R + \frac{k+1}{2} \right)$$

Donde $E = e^{2G} - 1$

La siguiente gráfica muestra la curva que relaciona el tiempo de retraso contra tráfico efectivo (throughput). Los parámetros empleados en la elaboración de la curva son $R=6$ y $k=5$. La escala del eje x es no lineal.

Retraso en Aloha

Retraso en Transmisión



Es de hacer notar que esta curva constituye una aproximación en el sentido que asume una espera de retransmisión de varios intervalos después de la colisión. Se ha encontrado que existe un óptimo tiempo de espera de retransmisión en el sentido de minimizar el retraso total D para un tráfico efectivo S dado. Si el intervalo de retransmisión es muy corto es mayor la probabilidad de una nueva colisión. Por otro lado, si es muy largo, el retraso total se extiende proporcionalmente. Se ha demostrado también que ese valor óptimo tiene un amplio rango y que la selección del parámetro k no es crítica.

Aloha Ranurado

El esquema de Aloha puro tiene una utilización máxima de $1/(2e)$. El parámetro crucial es el intervalo (2τ) en torno a un paquete susceptible a colisiones. El Aloha Ranurado es un sistema en el que todos los usuarios están sincronizados a una referencia común de tiempo. El tiempo se divide en intervalos de duración del paquete o ranuras de τ segundos. Los usuarios transmiten únicamente en los instantes de inicio de ranura. De esta manera las colisiones ocurren cuando dos o más paquetes se transmiten en la misma ranura. Ya que no hay interferencia de paquetes de intervalos vecinos, el intervalo vulnerable se reduce a la mitad y la capacidad del protocolo es exactamente del doble del sistema Aloha puro. Esta vez la expresión del tráfico efectivo es:

$$S = Ge^{-G}$$

y su máximo valor se duplica también.

$$S_{\max} = 1/e$$

En este caso se considera también procesos de función de probabilidad de Poisson para el tráfico original y los paquetes retransmitidos.

Aloha Ranurado aporta una ventaja considerable en eficiencia de utilización del canal sobre la técnica de Aloha Puro a un costo de complejidad del sistema. Para redes cubriendo una pequeña área esto no representa ningún problema especial. Para redes dispersas en áreas grandes o en sistemas vía satélite, la utilización de técnica Aloha Ranurado presenta requerimientos adicionales de implementación.

La característica del retraso depende del método de retransmitir los paquetes que sufren una colisión. Para este análisis que el tiempo de espera de retransmisión (asumiendo que el mensaje de recepción positivo no se recibe y en adición al tiempo de propagación) es de j ranuras, siendo j de un valor de 1 a k como se describe anteriormente. Es el mismo proceso de aleatorización antes mencionado, expresado en una escala discreta de tiempo.

El tiempo promedio de retraso se denomina D nuevamente, medido en ranuras y con el ajuste para considerar los paquetes originados antes de que el intervalo de la ranura se inicie:

$$D = R + 1.5 + E \left(R + \frac{1}{2} + \frac{k+1}{2} \right)$$

La evaluación de los parámetros E , G , S y D está en función de k , el número máximo de ranuras que transcurren hasta el próximo intento de retransmisión.

p_i es la probabilidad de que un paquete requiera exactamente " i " retransmisiones para ser recibido correctamente.

$$p_i = (1 - q_n) (1 - q_t)^{i-1} q_t \quad i \geq 1$$

y el número promedio de retransmisiones es:

$$E = \sum_{i=1}^{\infty} i p_i = (1 - q_n) / q_t$$

La derivación de q_n y q_t se encuentra en el apéndice A.

Para un k grande (el retraso de retransmisión aleatorio es del orden de varias ranuras) se tienen los límites:

$$E \rightarrow e^G - 1$$

y

$$S/G \rightarrow e^{-G}$$

Para un k de orden finito, en la referencia [6] Lam realizó el análisis encontrando que existe un k óptimo que minimiza el retraso para un tráfico efectivo S dado.

Lam demostró utilizando un ejemplo de transmisión vía satélite que el k óptimo es del orden de 6 para un tráfico efectivo S pequeño, haciendo notar que el valor es cercano al óptimo para casi cualquier valor de S .

En este estudio se expresa tanto el promedio de retransmisiones como el retraso de la retransmisión (promedio) en función de k .

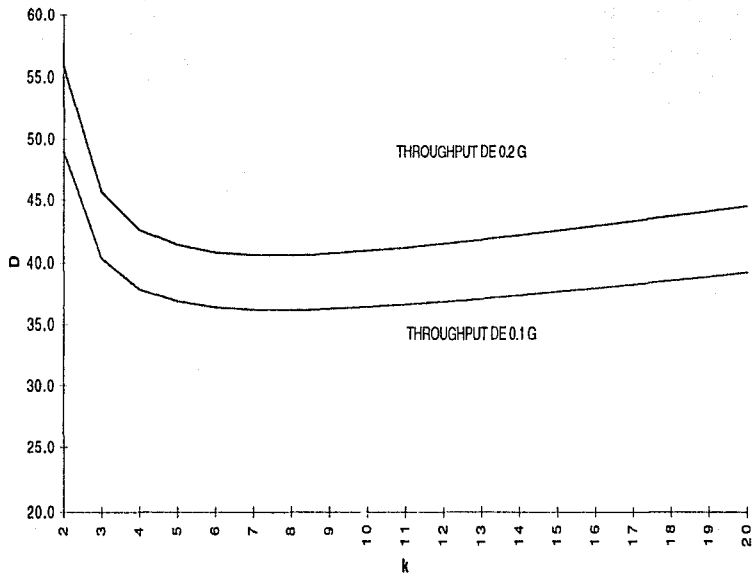
El valor de k que minimiza el retraso promedio en el canal es efectivamente del orden de 6 en el caso de transmisión vía satélite (k es una variable discreta, con los valores empleados en este análisis el óptimo es 7, con una mínima diferencia en el desempeño de $k=6$).

Sin embargo, como se puede apreciar en la gráfica, el retraso no varía mucho en función de k para un S dado. Para valores de k de 4 a 10 el retraso es sumamente uniforme.

Para un tráfico en el canal más allá del punto óptimo de S normalizado de $1/e$ se presenta una reducción en el tráfico efectivo y un incremento en el retraso. Esto es explicable por el incremento en colisiones y por lo tanto en la retransmisión de paquetes.

La selección entre Aloha en su forma simple y ranurada depende de la relación entre el tamaño del encabezado de la ráfaga y el tamaño del mensaje, la distribución de probabilidad del tamaño del mensaje mismo y el punto de operación del tráfico entrante y saliente seleccionado para satisfacer las condiciones de retraso y estabilidad.

RETRASO EN RETRANSMISION



El sistema Aloha presenta inestabilidad en la región de mayor eficiencia, que se puede compensar reduciendo el número de usuarios accediendo al canal o incrementando el retraso de retransmisión k . Una efectiva y elegante solución para mantener estable un canal potencialmente inestable es la implementación de procedimientos de control dinámico de tráfico.

El criterio de selección de un protocolo satelital se basa especialmente en el tipo de tráfico y el desempeño requerido por los usuarios. Si el tráfico de cada estación es altamente predecible, es más factible coordinar el uso eficiente del canal mediante una técnica de asignación de ancho de banda y tiempo de transmisión.

En el caso opuesto, un bajo nivel de predictibilidad implica ya sea un uso menos eficiente del ancho de banda o técnicas más sofisticadas de asignación de tiempos de utilización del canal.

Se ha dedicado una considerable atención a la optimización del tráfico efectivo de un canal Aloha.

Se han propuesto una variedad de esquemas de reservación en los que los usuarios solicitan tiempos de transmisión por anticipado. Otra posibilidad es la de monitorear el canal y efectuar la transmisión cuando se prevé que no haya colisión, como en los protocolos de contención en banda base utilizado en las redes locales (CSMA).

El método de reservar tiempo de transmisión logra un mayor tráfico efectivo en el canal a cambio de un mayor retraso.

El esquema de sensar la portadora tiene una eficiencia teórica de hasta más del 80% en el caso de redes de radioenlaces. Una importante condición para implementar este esquema es que las estaciones se encuentren lo suficientemente cerca, de manera que el tiempo que toma detectar las transmisiones de los demás sea tan solo una pequeña fracción del intervalo de la ranura. A medida que el tiempo de propagación se incrementa, la eficiencia del canal decrece.

Capacidad Excedida

Existe un caso particular en la aplicación del esquema de Aloha ranurado en el que la eficiencia es mayor que el 36% establecido. Estas configuraciones son llamadas "de capacidad excedida".

Se encuentra por supuesto el caso trivial de que un solo usuario transmita información a la capacidad del canal, tal que la eficiencia sea del 100%. Sin embargo, el motivo de aplicar técnicas de acceso aleatorio es el compartir el canal entre un grupo de usuarios que transmiten con baja frecuencia y densidad.

El análisis original del Aloha presupone una estadística de Poisson para un número grande de usuarios. Esta generalización parte de la observación en teoría de probabilidad de que la suma de variables independientes de probabilidad de Bernoulli se aproxima bajo ciertas condiciones a una variable de probabilidad de Poisson.

Cuando un sistema de acceso Aloha ranurado tiene un número limitado de usuarios puede tener una "capacidad excedida", especialmente cuando uno de los usuarios transmite con mayor frecuencia y el resto lo hace a muy baja frecuencia.

El hecho de que la "capacidad excedida" permite a un número reducido de estaciones utilizar el canal a altas tasas de transmisión presenta la posibilidad de asignar a esas estaciones un acceso determinístico al canal. Para ese objetivo se diseñó la variante del esquema Aloha llamado de Reservación.

Este protocolo consiste en la asignación fija de tiempos en un canal Aloha ranurado, además del requerimiento de control adicional en la estación. En este protocolo, si el tráfico es uniformemente distribuido en las estaciones y los "ciclos de trabajo" son altos, el sistema se comporta de forma semejante al método de acceso por división de tiempo. Si en cambio el ciclo de trabajo es bajo, se aproxima a la operación de un sistema Aloha ranurado simple.

La búsqueda de técnicas de un uso flexible del canal con una eficiencia comparable a la de TDMA de asignación fija ha conducido a la evolución de esquemas de reservación explícita.

APENDICE A

El tiempo promedio de retraso se denomina D , medido en términos del número de ranuras k que transcurren antes de la retransmisión.

$$D = R + 1.5 + E \left(R + \frac{1}{2} + \frac{k+1}{2} \right)$$

p_i es la probabilidad de que un paquete requiera exactamente "i" retransmisiones para ser recibido correctamente.

$$p_i = (1 - q_n) (1 - q_t)^{i-1} q_t \quad i \geq 1$$

y el número promedio de retransmisiones es:

$$E = \sum_{i=1}^{\infty} i p_i = (1 - q_n) \cdot q_t$$

q_n es la probabilidad de una transmisión exitosa, cuando un paquete nuevo se genera.

q_t es la probabilidad de transmisión exitosa cuando un paquete es retransmitido.

La derivación de q_t sigue el siguiente procedimiento:

Se considera que la colisión que produce esta retransmisión ocurrió en la j -ésima ranura anterior.

Para una re-transmisión exitosa se deben cumplir tres condiciones:

1.- Ningún otro paquete, desde la j -ésima ranura deberá ser retransmitido en la ranura presente (la probabilidad de este evento se llama q_c).

2.- Los paquetes que puedan retransmitirse desde las $(k-1)$ ranuras anteriores además de la j -ésima no se transmitirán en la ranura presente (la probabilidad es q_0^{k-1}).

3.- Ningún paquete nuevo se origina en la presente ranura. La probabilidad es e^{-S} .

Asumiendo que estos eventos son independientes, la probabilidad de una retransmisión exitosa q_c es:

$$q_t = e^{-S} q_c q_0^{k-1}$$

Considerando ahora la probabilidad q_c de que ningún otro paquete desde la j -ésima ranura fue retransmitida en la ranura presente, q_c se puede escribir:

$$q_c = \sum_{\gamma=1}^{\infty} q_c(\gamma)$$

$q_c(\gamma)$ viene a ser la probabilidad de que γ paquetes, además del paquete bajo consideración fueron generados en la ranura j , causando la colisión y que ninguno de ellos se traslape en la presente ranura.

$$q_c = (e^{-G \cdot k} - e^{-G}) \cdot (1 - e^{-G})$$

q_0 es la probabilidad de que ningún paquete de alguna de las otras $(k-1)$ ranuras, excluyendo la j -ésima que podría estar involucrada en una retransmisión en la ranura presente es realmente retransmitida en esa ranura.

$$q_0 = e^{-G \cdot k} + G \cdot k \cdot e^{-G}$$

mientras que q_n , la probabilidad de la transmisión exitosa de un paquete nuevo está dada por:

$$q_n = q_c^k e^{-S}$$

Para las gráficas de "promedio de Retransmisiones" y "Retraso en Retransmisión" en función de k se consideran dos casos de tráfico efectivo "throughput" normalizado: de $S = 0.1$ y 0.2 .

Los parámetros empleados para la construcción de estas gráficas son:

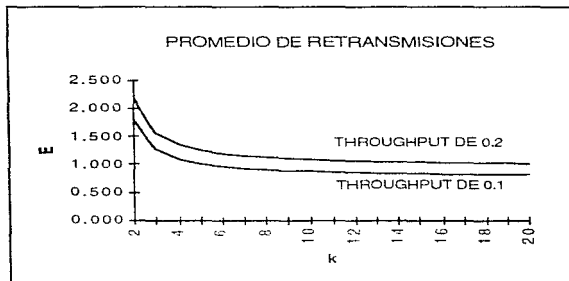
Tasa de canal = 64 KBPS

Paquete: 1000 Bits

R = 0.25 Seg. (retraso de propagación en un canal vía satélite, de estación a satélite y de satélite a estación)

El rango de k es de 2 a 20

La gráfica muestra una curva continua para fines ilustrativos, ya que k es una variable discreta.



Naturalmente, se presenta un mayor retraso y número promedio de retransmisiones en las condiciones de mayor tráfico en el canal, esto es, cuando $S=0.2$, por aumentar la probabilidad de colisión.

La tabla de valores de la gráfica es la siguiente:

k	E		k	D	
	S=0.2	S=0.1		S=0.2	S=0.1
2	2.193	1.795	2	55.9	-18.9
3	1.561	1.269	3	45.6	40.3
4	1.355	1.098	4	42.6	37.8
5	1.254	1.014	5	41.3	36.8
6	1.193	0.964	6	40.8	36.3
7	1.153	0.93	7	40.6	36.1
8	1.125	0.907	8	40.6	36.1
9	1.103	0.889	9	40.7	36.2
10	1.087	0.875	10	40.9	36.3
11	1.073	0.864	11	41.1	36.5
12	1.063	0.855	12	41.4	36.7
13	1.054	0.848	13	41.7	37.0
14	1.046	0.841	14	42.1	37.3
15	1.040	0.836	15	42.5	37.6
16	1.034	0.831	16	42.8	37.9
17	1.029	0.827	17	43.2	38.2
18	1.025	0.824	18	43.6	38.5
19	1.021	0.82	19	44.0	38.8
20	1.017	0.818	20	44.5	39.2

Capítulo 3

Técnicas de Acceso Aleatorio II

Protocolos de Reservación

Ya que el efecto de capacidad excedida permite a un grupo de estaciones de alto tráfico utilizar el canal, una manera viable de optimizar el uso del canal es dedicar un intervalo determinado a esas estaciones y no hacer necesario el acceso aleatorio al canal.

En la sección anterior se analizó la utilización de un canal para un grupo de usuarios de bajo tráfico. Cuando el tráfico es más denso el desempeño del canal declina por el efecto de las colisiones. Si el tráfico fuera muy denso y de tendencia determinística el mejor esquema de uso de utilización del canal sería uno de tipo de asignación fija TDMA. Este tipo de protocolos tendría un bajo desempeño en el caso de un grupo grande de usuarios intermitentes. Con un esquema de asignación rígida fija se presenta la posibilidad de un uso ineficiente del canal, por los tiempos asignados pero no utilizados.

El esquema Aloha de Reservación proporciona una solución a ambos tipos de tráfico, tiene la opción de operar en modo Aloha desde un extremo (tráfico muy ligero) y en asignación fija en el otro extremo (tráfico muy denso).

Los esquemas de Aloha de Reservación son técnicas que incrementan el tráfico efectivo en un canal a un nivel mayor del máximo posible con Aloha ranurado, al costo de un retraso adicional.

El concepto básico consiste en que el tiempo de uso del canal es organizado en tramas ("frames"). Cada trama consiste de dos sub-tramas: la de transmisión y la de reservación.

La sub-trama de reservación es a su vez dividida en ranuras de corta duración que son utilizadas para paquetes de reservación que accesan el canal en forma aleatoria dentro de la sub-trama en modo Aloha ranurado. Por la característica de difusión del canal "broadcast", un paquete de reservación transmitido exitosamente sin interferencia es recibido por todos los usuarios, quienes reciben aviso de la reservación.

Lo que los usuarios reservan es una posición en la línea de espera global para un grupo de paquetes a transmitir. Cada usuario mantiene información del status de la línea de espera. La información esencial es la longitud de la línea de espera así como la posición de su propio mensaje a transmitir.

Varios de estos tipos de esquemas han sido analizados y estudiados. En este análisis se revisan los resultados de Binder [2] y Roberts [10]. En el primero, las ranuras son organizadas en tramas de por lo menos R ranuras (R es el tiempo de propagación, en número de ranuras). La trama es una referencia de tiempo a las que todas las estaciones deben estar sincronizadas, cada ranura tiene una posición relativa definida en la trama.

Si un usuario ha tenido una transmisión exitosa de paquete en la trama anterior, tiene derecho a transmitir en la misma ranura en los marcos subsiguientes, hasta que ya no tiene paquetes a transmitir.

Las ranuras "desocupadas", incluyendo las que hayan tenido colisión en la trama anterior se consideran disponibles y pueden ser accedidas en modo aleatorio por cualquier usuario.

Este método de reservación tiene su punto inicial en la transmisión exitosa de un paquete. Este sistema proporciona una técnica particular de Aloha ranurado para los casos en los que las estaciones tienen mensajes largos a transmitir o bien transmisión continua.

Un aspecto interesante es que la reservación agrega una capacidad de conmutación de línea al modo normal de conmutación de paquetes al esquema Aloha. Este canal llega a alcanzar una velocidad de tráfico

efectivo que se puede aproximar a 1, a un costo de un mayor retraso en la transmisión de paquetes.

Para un tráfico efectivo dado, el retraso de transmisión de paquetes es proporcional al número de estaciones en la red. Este tipo de reservación se limita esencialmente en número de estaciones al orden de R , el número de ranuras que toma el tiempo de propagación.

Un segundo esquema de reservación propuesto por Binder [2] combina multiplexaje por división de tiempo con asignación dinámica de ranuras no utilizadas.

A cada usuario de la red se le asigna una ranura por trama. Todas las ranuras no utilizadas, así como las no asignadas se encuentran disponibles para la red en general en un orden de primera llegada-primer salida.

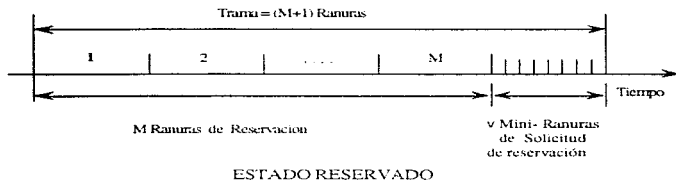
Un usuario puede recuperar su ranura asignada si se encuentra ocupada, simplemente transmitiendo un paquete y generando una colisión. En la siguiente trama toma posición normal de su ranura y transmite su paquete. Este método es útil en aplicaciones de tráfico interactivo, es capaz de disponer de casi toda la capacidad del canal para un solo usuario mientras el resto de ellos no transmite o lo hace con poca frecuencia.

Los resultados de pruebas de simulación han mostrado una característica retraso-tráfico efectivo superior al del multiplexaje por división de tiempo.

Un esquema que ha tenido una amplia aplicación es el diseñado por Roberts [10]. En este sistema se definen dos estados, de acuerdo al modo de operación:

- El estado Aloha
- El estado Reservado

En el estado Reservado la trama de tiempo consiste de $M+1$ ranuras.



Las primeras M ranuras son asignadas a usuarios con reservación; la última ranura de la trama se divide en v mini-ranuras, las que se utilizan para comunicación de paquetes cortos de solicitud de reservación.

El usuario con información a transmitir debe solicitar primeramente al sistema central la reservación de las ranuras requeridas. Los paquetes de solicitud son transmitidos en una de las v mini-ranuras con acceso Aloha ranurado.

Al recibir la autorización del nodo central, la información real a ser transmitida queda en línea de espera hasta que ocurre su ranura reservada de transmisión.

Los usuarios en general reciben la señal de autorización y actualizan continuamente la lista de reservaciones. De esta forma cada usuario sabe qué ranuras le han sido asignadas para transmitir su mensaje.

En el caso de que el paquete de solicitud sufra una colisión o error en la transmisión, se sigue el procedimiento normal de Aloha ranurado: de no recibirse reconocimiento de la estación central al término de un intervalo de espera ("timeout"), retransmite el paquete de solicitud en una de las v mini-ranuras de la trama posterior, con el mecanismo de aleatorización apropiado en la retransmisión para reducir la probabilidad de una nueva colisión.

Cuando la línea de espera en modo Reservado se agota tal que no hay mensajes en la red a ser transmitidos, el sistema se conmuta al estado Aloha, con todas las $M+1$ ranuras por trama divididas en v mini-ranuras cada una y disponibles para solicitud de reservación. Tan pronto como una solicitud es autorizada, el sistema regresa al modo Reservado.

El esquema de reservación de Roberts produce una notable mejora en la eficiencia del canal comparado con Aloha ranurado. La eficiencia tiene un valor máximo de $M/(M+1)$, a comparación del $1/e = 36.8\%$ del Aloha ranurado.

Como se ha puntualizado, mientras la eficiencia se optimiza el retraso se incrementa. Para un tráfico efectivo de baja densidad el tiempo de retraso mínimo es del doble del tiempo de propagación (el del paquete de solicitud más el del mensaje) más el tiempo de transmisión, contrastando con el retraso del tiempo de transmisión del sistema de Aloha ranurado.

A continuación se encuentra el análisis del desempeño de este esquema de reservación, el objetivo es encontrar una expresión del tiempo de retraso en la transmisión a través del canal en función de los parámetros que definen el sistema y el punto óptimo de operación como el balance entre número de ranuras, de mini-ranuras y el tráfico accediendo al canal.

Se consideran k usuarios, cada uno transmitiendo a una tasa de λ [mensajes/seg]. Z es la capacidad del canal en [paquetes/seg], (los paquetes largos, los mensajes) Z es igual a $1/\tau$, siendo τ la duración del paquete en [seg].

La primera parte del análisis es sobre el canal en estado Aloha. La utilización del canal dividido en sub-tramas de v mini-ranuras para paquetes de solicitud es:

$$S_1 = \frac{k\lambda}{Zv}$$

Cuando el canal está en estado Aloha se transmiten únicamente mensajes cortos, por lo que la tasa de mensajes se escala de $k\lambda$ a $k\lambda/v$.

La utilización del canal por las mini-ranuras en el estado de reservación es:

$$S_2 = k\lambda(M+1)/Zv$$

ya que la capacidad efectiva para el paquete de solicitud en el estado Reservado se reduce de Z a $Z/(M+1)$.

La utilización del canal por parte de los mensajes formados en línea de espera y después transmitidos utilizando las ranuras reservadas es:

$$S_3 = BK\lambda(M+1)/MZ$$

En este punto se considera la transmisión de mensajes de cualquier tamaño, a diferencia del cálculo de S_1 y S_2 , que consideran únicamente mensajes cortos.

B es el tamaño promedio del mensaje, en paquetes.

Ya que S_1 y S_2 están referidos a un canal Aloha ranurado, su valor máximo es $1/e$, mientras S_3 siempre es menor que 1.

S_3 representa en realidad la intensidad de tráfico para una línea de espera M/G/1 con capacidad efectiva en el canal de $MZ/(M+1)$.

La utilización real del canal o intensidad de tráfico para el sistema es:

$$\rho = BK\lambda(M+1)/MZ$$

que es la relación entre paquetes/seg introducidos al sistema a la capacidad Z , en ranuras o paquetes/seg.

ρ se puede expresar en términos de S_3 :

$$\rho = MS_3/(M+1)$$

y por lo tanto $\rho < M/(M+1)$

Desde este punto de vista, a mayor M , mayor eficiencia en el sistema. M es la más importante variable para alcanzar una mayor utilización del canal que en un esquema Aloha puro.

Sin embargo, hay un límite en el que M se puede incrementar ya que el número de reservaciones aceptadas deben estar dentro del límite que marca la tasa máxima de paquetes de solicitud que opera en modo Aloha ranurado.

Concretamente se tienen los límites:

$$S_3 = MS_3/Bv \leq 1. e$$

Para encontrar el punto óptimo de operación de la red hay varios enfoques, considerando la dependencia entre todas las variables involucradas. El procedimiento de Roberts es seleccionar primeramente v . Existe un valor óptimo de M para cada valor de ρ que minimiza el tiempo de retraso. Este resultado es similar al de optimizar el intervalo de retransmisión k que minimiza el retraso en Aloha ranurado.

Se ha encontrado que tampoco la selección de M es crítica y que la función de tiempo de retraso tiene un rango uniforme en función de M .

La función de tiempo de retraso es la suma del tiempo de propagación y de otras dos componentes:

- El tiempo transcurrido en hacer la reservación por el canal Aloha ranurado y
- El tiempo de espera de los mensajes a que ocurra su intervalo de transmisión ya reservado.

El retraso en una línea de espera de tipo M/G/1 se puede expresar como la suma del tiempo de transmisión $1/\mu$ más el tiempo promedio de espera:

$$E(w) = E(t^2) / (2(1-\rho))$$

siendo $E(t^2)$ el segundo momento de la duración del mensaje en segundos. λ es la tasa efectiva de llegada de mensajes en [mensajes/seg] y ρ es la intensidad de tráfico.

En el sistema analizado, la tasa efectiva de transmisión de mensajes de la red al canal es $k\lambda$. La intensidad de tráfico, en lo que al tiempo de espera del mensaje se refiere es S_2 y la capacidad efectiva del canal en ranuras por segundo es $MZ/(M+1)$.

El segundo momento de la duración del mensaje en unidades de "paquetes al cuadrado" es $E\{B^2\}$; para convertirlo a segundos al cuadrado se divide entre el cuadrado de $MZ/(M+1)$. El tiempo de transmisión en segundos es $B/(M+1)/MZ$.

El tiempo en la línea de espera D_q está dado por:

$$D_q = B/(M+1) / MZ + Y S_2 (M+1) / 2 MZ (1-S_2)$$

Donde $Y = E\{B^2\}/B$

El tiempo total de retraso es:

$$D_r = D_s + D_q + R$$

siendo D_s el retraso por reservación y R el retardo de propagación en intervalos de paquete.

Para obtener el retraso por reservación hay que considerar que el sistema puede estar en cualquiera de los dos estados, Reservado o Aloha. El retraso en modo Aloha se denominará D_1 y tiene una probabilidad de ocurrencia de $(1-S_2)$ ya que S_2 es por definición la fracción del tiempo en el que los mensajes son transmitidos o simplemente la probabilidad de que el sistema esté en modo Reservado.

Similarmente, el tiempo de retraso del paquete de solicitud en modo Reservado se le nombra D_2 y tiene una probabilidad de ocurrencia S_3 .

El tiempo promedio de retraso de un paquete de solicitud, considerando los dos estados es:

$$D_s = D_1(1-S_2) + D_2S_3$$

Donde D_1 y D_2 son los retrasos en cada uno de los dos estados. D_1 es el retraso en estado Aloha. Del resultado del retraso en Aloha ranurado, dado en unidades de paquetes de mensaje o "ranuras largas" es:

$$D_1 = R + \frac{1.5}{v} + E_1 \left(R + \frac{1}{2v} + \frac{k+1}{2v} \right) \text{ [ranuras]}$$

Esta ecuación es la misma del retraso de Aloha ranurado en términos de las "mini-ranuras". Para expresar D_1 en segundos hay que dividirlo entre Z [ranuras/seg]. El número promedio de retransmisiones E_1 es

$E_1 = G_1/S_1 - 1$ donde G_1 es el tráfico total (de solicitudes) tanto de transmisiones nuevas como de retransmisiones. Para efectos de construcción de gráficas se asumirá $G_1/S_1 = e^{G_1}$ por simplicidad.

El tiempo de retraso D_2 del paquete de solicitud en estado Reservado se calcula de modo semejante con la característica especial de que las ranuras para solicitud se encuentran únicamente hasta el final de cada trama. Esto agrega un tiempo de retraso adicional de $M/2$ ranuras de mensaje en promedio por paquete de solicitud. Para casos en los que $R > (M+1)$ y $k \leq v$, D_2 se puede expresar en unidades de ranuras de mensaje:

$$D_2 = R + M/2 + 1.5/v + E_2 [R + M/2 + (k+1)/2v] \text{ [ranuras]}$$

De manera semejante, E_2 es el mismo promedio de retransmisiones por paquete de solicitud generado.

Se considera que S_2 y G_2 , la intensidad y tráfico recién generado están relacionados por $G_2/S_2 = e^{G_2}$

Con el cálculo de D_T se puede realizar una evaluación de ρ en función de D_T . S_1 y S_2 son calculados a partir de la utilización del canal en cada estado. El aspecto más importante del cálculo es encontrar las variables G_1 y G_2 . Esto se hace a partir de S_1 y S_2 en forma recursiva.

Con las parejas de datos $\{G_1, S_1\}$ y $\{G_2, S_2\}$ son calculados E_1 y E_2 y posteriormente D_1 y D_2 , los retrasos para cada caso de retransmisión.

La variable independiente es siempre el tráfico de usuario accedando el canal k .

La utilización del canal S_2 para ranuras largas se calcula suponiendo que el canal está en modo reservado, ya que el estado Aloha es un resultado de la no utilización de las ranuras, no la causa.

La fracción promedio del tiempo en que el sistema esta en modo reservado es igual a S_2 , la utilización del canal por ranuras largas, ya que es la fracción del tiempo en que los paquetes reservados son transmitidos.

A continuación se presenta un ejemplo de calculo de retraso basado en el ejemplo de Roberts [10]. La diferencia mas importante con relacion al caso de Roberts es el cociente del segundo momento del tamaño de los paquetes dividido sobre el tamaño promedio del mismo. $Y = E\{B^2\}/B$.

Cuando se comparan los tiempos de respuesta de un sistema Aloha con un sistema de reservacion se observa una superioridad del sistema de acceso aleatorio para baja intensidad de trafico. El sistema de reservacion presenta el mejor equilibrio entre el reducido tiempo de respuesta del acceso aleatorio y la alta eficiencia del esquema TDMA de asignacion fija.

La ventaja del sistema del sistema de reservacion se presenta en la mejor capacidad de tráfico efectivo, que tiene un limite máximo de $M/(M+1)$.

CALCULO DEL RETRASO, ALOHA CON RESERVACION

Probabilidad de:

G1= Tráfico de solicitudes en el canal en modo Aloha

G2= Tráfico de solicitudes en el sub-canal en modo Reservado

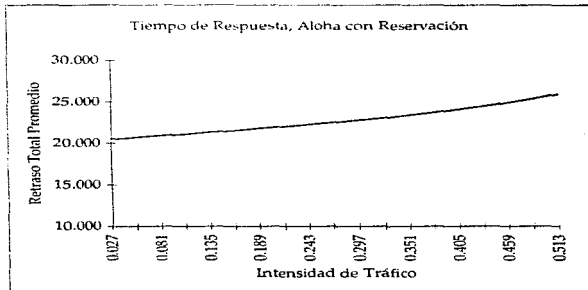
G3= Tráfico de informacion en el canal en modo Reservado

TASA = 50 KBPS
 PAQUETE: 1350 Bits (mensaje largo)
 R = 0.27 SEGUNDOS
 t = 27 mSeg
 R= 10 Ranuras en el tiempo de propagación
 k= 6 Ranuras de espera para retransmisión
 V= 6 Ranuras en sub-canal de reservación
 M= 2 Ranuras por trama para reservación
 B= 1 Tamaño promedio del mensaje [paquetes]
 Z= 37.0 Capacidad del canal [paquetes/seg]
 Y= 1.0 =(prom(B^2))/B

$$D_r = D_s + D_q + R/Z$$

$$D_s = D_1 \cdot (1 - S_3) + D_2 \cdot S_3$$

$$D_q = B \cdot (M+1) / M \cdot Z + Y \cdot S_3 \cdot (M+1) / [2 \cdot M \cdot Z \cdot (1-S_3)]$$



Aloha de Retransmisión Selectiva

El tema central de esta sección es la evaluación de sistemas de transmisión de mensajes conteniendo más de un paquete.

En sistemas de acceso aleatorio, la estrategia convencional es retransmitir el mensaje completo cuando ocurre una colisión. Este mecanismo reduce el tráfico efectivo "throughput" comparado con el obtenido con mensajes de un solo paquete.

Para compensar esa reducción se han diseñado esquemas de retransmisión de mensajes que resultan en una significativa mejora del desempeño del protocolo global.

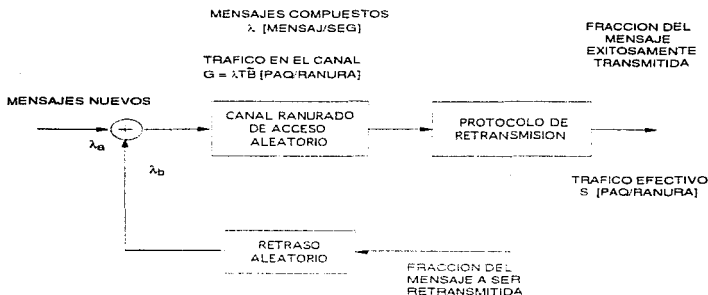
Hay dos casos de protocolos en particular: el sistema de "Retransmisión a partir de N" (GBN) que ofrece un modesto incremento en eficiencia y el protocolo de Retransmisión Selectiva, que alcanza la misma eficiencia del sistema Aloha ranurado, independientemente de la estadística del tamaño de los mensajes.

Los análisis tradicionales de los sistemas Aloha se han concentrado en los casos especiales de mensajes de un solo paquete y canales asincrónicos o ranurados en intervalos iguales a la duración del paquete.

En esta sección se analizan los sistemas Aloha con mensajes consistentes de varios paquetes con una arbitraria distribución de probabilidad de longitud del mensaje.

El principio básico de los sistemas de retransmisión de este análisis consiste en limitar la retransmisión del mensaje completo cuando éste sufre una colisión parcial. Este mecanismo se deriva de los protocolos de enlaces de datos (DLC).

El modelo esquemático del algoritmo se muestra a continuación.



Los mensajes transmitidos (los nuevos y los que son retransmitidos) son considerados procesos de Poisson con tasas de transmisión λ_a y λ_b respectivamente. El tiempo de transmisión en el canal es dividido en intervalos de T segundos y los paquetes se sincronizan para comenzar y terminar dentro de este intervalo.

Las distribuciones de probabilidad se denotan:

$a(i)$, $i = 1, 2, 3, \dots$ para los mensajes nuevos y

$b(i)$, $i = 1, 2, 3, \dots$ para las transmisiones compuestas

i es el número de paquetes en el mensaje. El tráfico normalizado en el canal es G y S es el tráfico efectivo que es comunicado exitosamente a través del canal. En condiciones de estado estable S debe ser $\lambda_a T \bar{a}$. Se han agregado las consideraciones de que el tiempo de espera para retransmitir es aleatorizado y que la generación de nuevos mensajes es interrumpida hasta que los paquetes del mensaje en transmisión han sido recibidos exitosamente.

Mensajes de múltiples paquetes en Aloha ranurado.

El primer caso a ser analizado es el convencional en el que los mensajes de uno o más paquetes son completamente retransmitidos al sufrir una colisión. Expresando el tráfico efectivo S en términos de probabilidad condicional de transmisión exitosa del mensaje:

$$S = \sum_{j=1}^{\infty} j \cdot [\text{Prob (TS exitosa del mensaje l long mensaje = j)} \\ \cdot \text{Prob (long del mensaje = j)}]$$

Esa probabilidad condicional es el producto de la probabilidad de tres eventos:

- P₁ De que el mensaje anterior no se extienda a la presente ranura
- P₂ De que haya exactamente una transmisión en la presente ranura
- P₃ De que no haya transmisiones en las próximas j-1 ranuras

De acuerdo a la consideración de distribución de Poisson de número de transmisiones en un intervalo T:

$$P_2 = \lambda \cdot T \cdot e^{-\lambda T}$$

$$P_3 = e^{-\lambda T(j-1)}$$

Para calcular P₁:

$$P_1 = \prod_{i=1}^{\infty} \text{Prob (el mensaje de la ranura anterior = i \\ no exceda de i paquetes)}$$

Definiendo B(•) como la distribución de probabilidad acumulada de los mensajes compuestos (los nuevos y los retransmitidos):

$$P_1 = \prod_{i=1}^{\infty} \sum_{k=i}^{\infty} [(\lambda \cdot T)^k / k!] \cdot e^{-\lambda T} \cdot (B(i))^k$$

Simplificando esta expresión, se tiene un resultado que depende únicamente del tamaño promedio de los mensajes compuestos.

$$P_1 = e^{-\lambda T} (\bar{B} - 1)$$

Que, sustituyendo en la fórmula del tráfico efectivo resulta:

$$S = \lambda T e^{-\lambda T} (\bar{B} - 1) \sum_{j=1}^{\infty} j e^{-\lambda T} b(j)$$

El tráfico efectivo S está dado en terminos del tamaño del mensaje retransmitido, concretamente su distribución $b(\bullet)$ y su valor medio \bar{B} . Estas variables son desconocidas pero se pueden definir en función del tráfico en el canal y la función de distribución del tráfico recién generado $\alpha(\bullet)$.

Para calcular $b(\bullet)$ en terminos de $\alpha(\bullet)$ se consideran condiciones de estado estable, una condición de equilibrio a un tamaño de paquete j en la que la tasa de los mensajes recién generados es igual a la tasa de a la que los mensajes salen del sistema.

$$\lambda \alpha(j) = \lambda e^{-\lambda T} (\bar{B} - 1) e^{-\lambda T j} b(j)$$

definiendo $M_2(x) = \sum_j e^{-\lambda T j} \alpha(j)$ como la función de generación de momentos de la distribución $\alpha(\bullet)$:

$$b(j) = \alpha(j) e^{-\lambda T j} / M_2(\lambda T)$$

y determinando \bar{B} como $M_2'(\lambda T) / M_2(\lambda T)$, se obtiene S en terminos de $\alpha(\bullet)$:

$$S = \lambda T \bar{B} M_2(\lambda T) \exp \{-\lambda T [(M_2'(\lambda T) / M_2(\lambda T)) - 1]\}$$

Se puede observar que S es función únicamente de la función de generación de momentos de la variable α , el tamaño de los mensajes nuevos.

Este resultado concuerda con el analisis clasico de Aloha puro con mensajes de tamaño variable.

En el caso particular de mensaje de tamaño constante K (mayor que 1, ocupando más de una ranura de tiempo), $S = G e^{-G(2-1/K)}$. Esta ecuación tiene los casos particulares de $k=1$ para Aloha ranurado y $k \rightarrow \infty$ para Aloha puro.

Para una distribución $o(\bullet)$ de tipo geométrica se ha encontrado que la distribución de la variable $b(\bullet)$ es geométrica también con media:
 $\bar{B} = \bar{A} / (\bar{A} - 1) e^{-\lambda T} \geq \bar{A}$.

El tráfico efectivo resultante en terminos del tráfico en el canal G para $\bar{A} / (\bar{A} - 1) \geq e^{-\lambda T}$ es:

$$S = (\bar{A} - (\bar{A} - 1)e^{-\lambda T})^2 G e^{-G}$$

Retransmisión en Aloha ranurado con mensajes de paquetes múltiples.

Uno de los recursos más utilizables para incrementar la eficiencia de un canal Aloha ranurado es evitar la retransmisión total de un paquete después de una colisión parcial. Una de las formas más simples es la retransmisión parcial del mensaje a partir del primer paquete erróneo. Con el uso de esta técnica es conocida como "Retransmisión a partir de N" (GBN) los paquetes son recibidos en el orden original y la información de dirección del mensaje se requiere únicamente en el primer paquete de la transmisión.

Un método de retransmisión que alcanza una mayor eficiencia en el canal es el de retransmisión selectiva. Este método consiste en retransmitir únicamente los paquetes con error en la recepción. La aplicación de este sistema implica que cada paquete debe contener información de numeración y dirección. Además, los receptores deben contar con funciones de memoria y reordenamiento de los paquetes. Esto significa la necesidad de un mayor encabezado ("overhead") en cada paquete y mayor procesamiento en los nodos. A cambio se obtiene una mayor eficiencia en el canal en términos de tráfico efectivo.

En la evaluación del tráfico S se considera nuevamente el grupo de ranuras en estado estable que tienen exactamente una transmisión. La probabilidad de estos eventos en estado estable se obtiene considerando la distribución del número de transmisiones en el sistema en la k -ésima ranura (N_k).

Prob{ $N_k = j$ } = $\sum_{i=0}^{\infty}$ Prob (i llegadas e (i-j) salidas en el intervalo de k ranuras)

Esta probabilidad puede definirse en términos de la distribución acumulada $B(\bullet)$:

$$\text{Prob} \{ N_k = j \} = \sum_{i=0}^{\infty} e^{-\lambda T k} \frac{(\lambda T k)^i}{i!} \binom{i}{j} \left[1/k \sum_{l=0}^{k-1} 1 - B(l) \right]^j \left[1/k \sum_{l=0}^{k-1} B(l) \right]^{i-j}$$

El siguiente paso es expandir $B(\bullet)$ en términos de $b(\bullet)$ y tomar el límite cuando $k \rightarrow \infty$. El resultado es:

$$\text{Prob} \{ N_k = j \} = \frac{(\lambda T B)^j}{j!} e^{-\lambda T B}$$

La distribución en estado estable resulta ser de tipo Poisson con media $\lambda T B$. El tráfico efectivo se obtiene por:

$$S = \lambda T B e^{-\lambda T B} = G e^{-G}$$

Este resultado también se puede obtener considerando las colisiones encontradas por un determinado paquete en el sistema. El tráfico en el canal consiste de mensajes con probabilidad de Poisson a una tasa λ y distribución de tamaño $b(\bullet)$.

La probabilidad de que no haya colisión en todas las transmisiones anteriores es $P_1 = e^{-\lambda T (1 - P_1)}$, que depende únicamente del tamaño promedio de los paquetes compuestos. La probabilidad de una transmisión exitosa del paquete es por lo tanto P_1 multiplicado por la probabilidad de que la transmisión de otro paquete se encuentre separada por lo menos una ranura, esto es $e^{-\lambda T}$. Como la tasa de transmisión de paquetes es $\lambda T B$, el tráfico efectivo es $\lambda T B e^{-\lambda T B}$ tal como se calculó previamente.

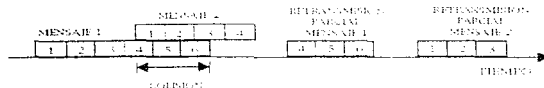
El aspecto más interesante de este resultado es que la eficiencia del protocolo SREJ Aloha es la misma que en el caso de mensajes de un solo paquete en Aloha ranurado, que es el mejor desempeño que cualquier sistema de retransmisión puede lograr.

El tráfico efectivo máximo de $1 - e \approx 0.368$ representa un notable aumento en eficiencia comparado con el Aloha convencional. Estrictamente hablando, esta eficiencia máxima es ligeramente menor de $1 - e$ por la adición de bits de dirección y numeración en cada paquete.

Aloha de Retransmisión Selectiva en canales no-ranurados.

Los resultados de SREJ Aloha en canales ranurados motivaron el interés en la aplicación del esquema en canales no-ranurados que no requieren sincronización.

El esquema consiste también en dividir los mensajes en paquetes de tamaño fijo conteniendo cada uno información de direccionamiento. El esquema de retransmisión se ilustra por la siguiente figura.



La eficiencia en el caso no-ranurado, sin considerar los bits del encabezado, se pueden aproximar a la del caso ranurado cuando el número promedio de paquetes en un mensaje es alto.

La medida en que la eficiencia se aproxime a su límite máximo depende de la estadística del tamaño de los mensajes. La optimización de este sistema consiste en encontrar el tamaño óptimo de paquete en un mensaje. Un paquete demasiado largo se aproxima al caso convencional de Aloha, un paquete muy corto contendrá un menor porcentaje de bits de información. La ecuación del tráfico efectivo de SREJ Aloha so-ranurado se puede obtener de la siguiente manera:

Considerando T segundos como la duración de los paquetes, en el caso no ranurado, la probabilidad de que un paquete no interfiera con transmisiones previas es $e^{-\lambda T B}$.

Como en el análisis del caso ranurado, la probabilidad de que la próxima transmisión se presente hasta después de un intervalo T es $e^{-\lambda T}$, la tasa total de transmisión es $\lambda T B$ y el tráfico efectivo es:

$$S = \lambda T B e^{-\lambda T B (1+B)} = G e^{-\lambda T B (1+B)}$$

Comparando el tráfico efectivo con el caso ranurado se observa que el factor $(1 + B)$ dentro del exponencial es proporcional a la degradación del tráfico efectivo debido al modo de operación asincrónica.

El siguiente paso es calcular B en términos de la estadística de los mensajes originales $\alpha(\bullet)$ y λT .

Mientras que B no fue calculada explícitamente en el caso ranurado, el procedimiento en el caso no ranurado sí lo requiere.

El enfoque de la solución a este problema es calcular los límites superior e inferior de B, lo que permite establecer a su vez los límites de S.

En el apéndice B se describen los procedimientos para calcular los límites de B en términos de A.

El límite superior: $B \leq (A+1 - e^{-\lambda A T}) / (2 - e^{-\lambda A T})$

Mientras el límite inferior: $B \geq A(1 + \lambda T) / (1 + A \lambda T)$

En tanto que el tráfico efectivo S tiene el siguiente límite inferior:

$$S(A, \lambda T) \geq \lambda T A (1 + \lambda T) / (1 + A \lambda T) \exp\{-\lambda T [(A+1 - e^{-\lambda A T}) / (2 - e^{-\lambda A T}) + 1]\}$$

Para el caso trivial de que el número de paquetes en el mensaje $A=1$ se tiene el resultado conocido de $\lambda T e^{-2\lambda T}$.

La evaluación total de la eficiencia del esquema de retransmisión debe incluir, además de la optimización de la relación paquetes/mensaje, el número de bits de información por paquete. Se define como L al número de bits en el mensaje original. Se denomina k al número de bits que se agregan como encabezado a cada paquete conteniendo información de direccionamiento y detección de error.

Cuando el mensaje es segmentado en paquetes, es necesario agregar el mismo encabezado en cada paquete, además de los bits de numeración de los paquetes. Para un nivel constante de protección de error se requieren menos bits de detección de error a medida que el tamaño del paquete disminuye. Para efectos de implementación es más realista considerar el número de bits de paridad como una constante.

La función de encabezado es $O(A) = k + \lceil \log_2 A \rceil$, donde $\lceil x \rceil$ significa el entero más próximo mayor de x . El tráfico efectivo S_N cuando los mensajes son segmentados en A paquetes es:

$$S_N(A, \lambda T) = (L + A \cdot O(A))^{-1} L \cdot S(A, \lambda T)$$

El tráfico efectivo que incluye encabezado puede calcularse como una función de A para determinados parámetros L y k . Las gráficas del tráfico efectivo en función de A se muestran en las páginas siguientes. La evaluación se hizo de acuerdo al método del autor del algoritmo. Dipankar Raychaudhuri [30] considerando B mínimo. Para el mismo análisis de cálculo de B_{max} se optimizó S_N de acuerdo a $\partial S_N / \partial G$ y calculando S en forma recursiva.

En ambos casos, el máximo tráfico efectivo se obtuvo con A en el rango de 4 a 10 paquetes por mensaje. Lógicamente, el valor óptimo de A y el tráfico efectivo máximo se reducen a medida que el tamaño del encabezado se incrementa.

En resumen, se demuestra que el máximo tráfico efectivo alcanzable con el esquema de retransmisión SREJ Aloha es idéntico al de Aloha ranurado con mensajes de un solo paquete, independientemente de la distribución de probabilidad del tamaño del mensaje.

Apéndice B

Cálculo de los límites de $\bar{B} = f(A)$.

Este análisis tiene como referencia el diagrama de estados de ARQ-Aloha (Solicitud de Retransmision Automática).

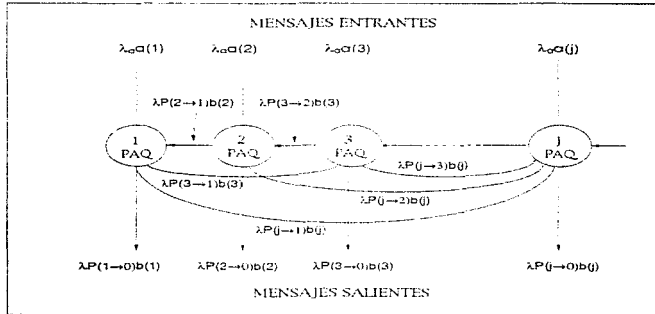


Diagrama de estados de mensajes Aloha con esquema de retransmision tipo ARQ.

En los esquemas de retransmision tipo ARQ, el tamaño de un mensaje con j paquetes es transformado secuencialmente en j_1, j_2, \dots, j_k en forma decreciente hasta que $j_k = 0$.

En el diagrama de estados se muestra la generación y transmisión de mensajes de tamaño j .

Para analizar un sistema definido por este diagrama de estado se debe determinar la probabilidad de que la transmisión de un mensaje de tamaño j se convierta en una de tamaño i ($i \leq j$). Esta probabilidad está denotada por $\text{Prob}(j \rightarrow i)$.

Del diagrama se observa que en el caso SREJ Aloha, las probabilidades de transición de la forma $\text{Prob}(j \rightarrow 0)$ pueden ser determinadas, a diferencia del caso general $\text{Prob}(j \rightarrow i)$ para $i \leq j$. El cálculo exacto de B requiere conocer todas las probabilidades de transición.

Una forma de resolver este problema es obtener las cotas inferior y superior del tamaño promedio B a partir de la probabilidad $\text{Prob}(j \rightarrow 0)$, únicamente si el análisis se restringe al caso de mensajes nuevos de intervalo fijo. Del diagrama de estados, la tasa a la que los mensajes entran y salen del sistema se pueden igualar de la siguiente forma:

$$\lambda_0 = \lambda \sum_{j=1}^{\infty} \text{Prob}(j \rightarrow 0) b(j)$$

La probabilidad de no tener interferencia de otros mensajes en SREJ no ranurado es $\text{Prob}(j \rightarrow 0) = e^{-\lambda T \bar{B}} - i \lambda T$ y el tráfico efectivo es $S = \lambda_0 T \bar{A}$, de esto se obtiene la igualdad:

$$\lambda \bar{B} e^{-\lambda T \bar{B}} - e^{-\lambda T} = \lambda \bar{A} e^{-\lambda T \bar{B}} \sum_{j=1}^{\infty} e^{-j \lambda T} b(j)$$

Que se puede simplificar en:

$$\bar{B} = \bar{A} \sum_{j=1}^{\infty} e^{-(j-1) \lambda T} b(j)$$

Considerando que los mensajes recién generados tienen un intervalo fijo de T_p segundos, si los paquetes son de duración T , se tienen $A = \bar{A} = T_p/T$ paquetes por mensaje tipo Aloha. Ya que el mensaje no tiene más de A paquetes, se tiene $b(i) = 0$ para $i > A$. Esto significa un límite superior finito a la sumatoria anterior, desde $j=1$ hasta A . Para el intervalo de k : $0 \leq k \leq A$ se tiene la desigualdad: $e^{-k\lambda T \bar{B}} \leq 1 - k(1 - e^{-(A\lambda T)})/A$ que, sustituyéndola en la expresión de B se tiene:

$$\bar{B} = A + 1 - e^{-A\lambda T} - \bar{B}(1 - e^{-A\lambda T})$$

Despejando \bar{B} se tiene:

$$\bar{B} \leq (A + 1 - e^{-A\lambda T}) / (2 - e^{-A\lambda T})$$

B max

$$B \leq (A+1-e^{-(A \cdot G)}) / (2-e^{-(A \cdot G)})$$

$$B \geq A(1+G) / (1+AG)$$

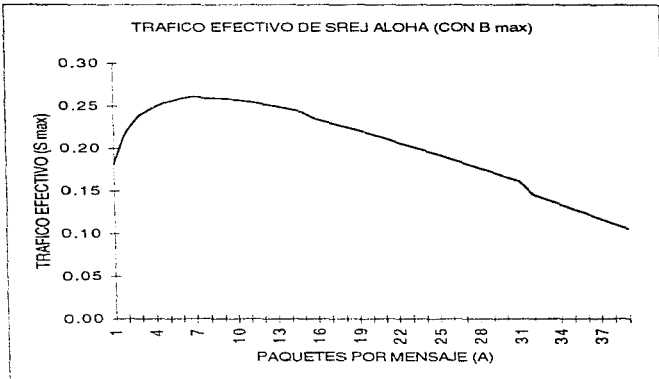
SREJ ALOHA

L= 1024

k= 12

A	Init B	Opt G	B max	S max	O(A)	S(N) max
1	1.00	0.50	1.00	0.18	13	0.18

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39



28	2.31	0.70	14.50	0.33	17	0.18
29	2.32	0.70	15.00	0.33	17	0.17
30	2.32	0.70	15.50	0.33	17	0.17
31	2.32	0.70	16.00	0.33	17	0.16
32	2.33	0.70	16.50	0.33	18	0.15
33	2.33	0.70	17.00	0.33	18	0.14
34	2.33	0.70	17.50	0.33	18	0.13
35	2.33	0.70	18.00	0.33	18	0.13
36	2.33	0.70	18.50	0.33	18	0.12
37	2.34	0.70	19.00	0.34	18	0.12
38	2.34	0.70	19.50	0.34	18	0.11
39	2.34	0.70	20.00	0.34	18	0.11

B min

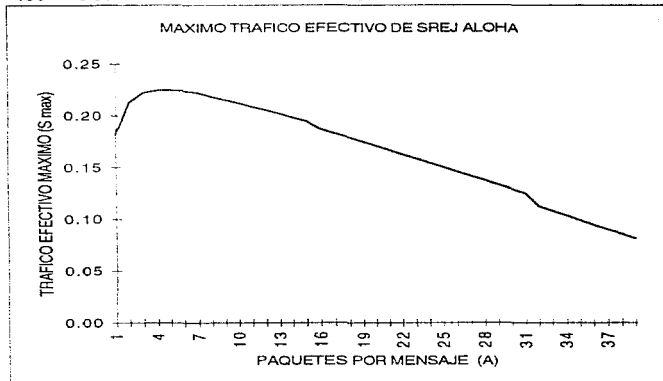
$$B \leq (A+1-e^{-(A \cdot G)}) / (2-e^{-(A \cdot G)})$$

$$B \geq A(1+G) / (1+AG)$$

SREJ ALOHA

L = 1024
k = 12

A	Init B	Opt G	B min	S max	O(A)	S(N) max
1	1.00	0.50	1.00	0.18	13	0.18



28	2.31	0.70	2.31	0.26	17	0.14
29	2.32	0.70	2.32	0.26	17	0.13
30	2.32	0.70	2.32	0.26	17	0.13
31	2.32	0.70	2.32	0.26	17	0.12
32	2.33	0.70	2.33	0.26	18	0.11
33	2.33	0.70	2.33	0.26	18	0.11
34	2.33	0.70	2.33	0.26	18	0.10
35	2.33	0.70	2.33	0.26	18	0.10
36	2.33	0.70	2.34	0.26	18	0.09
37	2.34	0.70	2.34	0.26	18	0.09
38	2.34	0.70	2.34	0.26	18	0.09
39	2.34	0.70	2.34	0.26	18	0.08

Capítulo 4

Teoría de Colas

La teoría de colas desempeña un papel muy importante en el análisis de redes de comunicación de datos, particularmente en redes de conmutación de paquetes. En el capítulo anterior se esbozó una expresión que relaciona dos variables muy importantes en el diseño de redes de computadoras: el tiempo de espera y la capacidad del canal. Esta relación, es el tema central de este capítulo que incluye otras variables tanto del tráfico de datos como de la capacidad de procesamiento de los nodos de una red.

Los análisis de sistemas de acceso aleatorio consideran el retraso en la transmisión de un mensaje como uno de los parámetros a ser minimizados en el sistema. Ese retraso tiene una componente representada por el tiempo que toma el sistema en darle al mensaje del usuario acceso al canal de comunicación. A este retraso en particular se le llama tiempo de espera.

El tiempo de espera es a su vez función de la estadística de los mensajes a ser transmitidos y al método de acceso al canal, que determina la forma en que los mensajes son procesados en cada nodo de comunicación.

Una red de conmutación de paquetes en general se puede considerar como una red interactiva de líneas de espera o colas. Un modelo simple de línea de espera se muestra a continuación:



El modelo original de las líneas de espera asume que estas se originan en forma aleatoria e independiente, siguiendo un procedimiento en el que los paquetes de datos ingresan a un nodo (o se generan en él), son almacenados, procesados y transferidos a un puerto de comunicación correspondiente para su transmisión.

Los paquetes se generan a una tasa λ en unidades de datos por unidad de tiempo. Son formados en una línea de espera en la memoria y son transmitidos a una tasa μ promedio.

El servicio de transmisión es realizado por un subsistema del nodo llamado "servidor" en la nomenclatura de la teoría de colas. En el caso particular de redes de datos, el servidor lo constituye el canal de comunicación de capacidad C .

El desempeño y la eficiencia de un canal de transmisión se define por variables aleatorias como los tiempos de espera, tiempo de servicio y tráfico (entrante y saliente) del nodo.

El tiempo total de retraso está comprendido por dos componentes:

- El tiempo de espera
- El tiempo de servicio de transmisión.

El tiempo de espera depende entre otras cosas, del "estado" en la memoria, esto es, el número de mensajes que se encuentran en la línea de espera.

Este tamaño de línea de espera es determinado estadísticamente en función de varios factores:

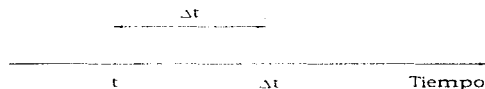
- 1.- La estadística de los mensajes entrantes, incluyendo tamaño del mensaje.
- 2.- La disciplina del servicio de transmisión:

- Estrategia de orden y prioridad de servicio
- Capacidad de memoria del nodo
- Número de puertos de transmisión.

Líneas de Espera M/M/1

Este modelo es el más empleado en teoría de líneas de espera y análisis de tráfico. El modelo M/M/1 asume distribución exponencial tanto de llegada de mensajes como de tiempo de servicio y la existencia de un canal de transmisión, esto es, la presencia de sólo un "servidor". El modelo analiza la distribución de mensajes en espera en el sistema, de los periodos de mayor tráfico y de los tiempos de espera.

Considerando un intervalo de tiempo Δt , con $\Delta t \rightarrow 0$. Las condiciones que definen a un proceso de Poisson son las siguientes:



1.- La probabilidad de que se genere un mensaje en Δt es $\lambda \Delta t \ll 1$, independientemente de la generación de otros mensajes en intervalos pasados o futuros. Esta condición implica que esa probabilidad es proporcional a la duración del intervalo de tiempo Δt . λ es por lo tanto la constante de proporcionalidad.

2.- La probabilidad de que no se generen mensajes durante Δt es $1 - \lambda \Delta t$. El proceso excluye la generación de más de un mensaje en Δt .

La estadística de llegadas de mensajes en un intervalo T tiene una distribución de Poisson. La probabilidad de que k mensajes sean generados en T segundos es:

$$P(k) = (\lambda T)^k e^{-\lambda T} / k! \quad k = 0, 1, 2$$

El número promedio de llegadas en T segundos es por lo tanto:

$$E[k] = \sum_{k=0}^{\infty} kp(k) = \lambda T$$

La constante λ representa por lo tanto la tasa promedio de llegadas de mensaje.

Hasta este punto se ha descrito la estadística de una variable discreta k en el dominio de una variable continua t . El parámetro a definir ahora es el del tiempo entre llegadas τ . τ es una variable aleatoria de distribución continua. Para una estadística de Poisson, τ es una variable de distribución exponencial cuya función de densidad de probabilidad es:

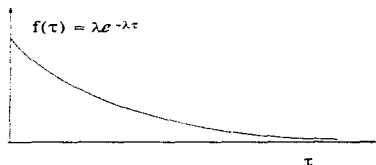
$$f(\tau) = \lambda e^{-\lambda\tau} \quad \tau \geq 0$$

Llegadas



Tiempo

Tiempo entre llegada de mensajes en un proceso de Poisson



Densidad Exponencial

El tiempo promedio entre llegadas es:

$$E(\tau) = \int_0^{\infty} \tau f(\tau) d\tau = 1/\lambda$$

En procesos de llegadas de Poisson, los tiempos entre llegadas tienen más probabilidad de ser cortos que largos. La probabilidad de un tiempo entre dos eventos sucesivos decrece exponencialmente con el tiempo τ entre ellos.

El siguiente parámetro a considerar es el tamaño del mensaje. El tiempo que dura la transmisión de un mensaje está directamente relacionado con el tamaño del mensaje.

En las líneas de espera del tipo M/M/1 se asume que los mensajes tienen un tamaño variable r , aleatorio y con una distribución exponencial con tamaño promedio $1/\mu$.

La función de densidad de probabilidad de r es:

$$f(r) = \mu e^{-\mu r}$$

El proceso del servicio de transmisión es análogo al proceso de llegadas de mensajes. Los mensajes cortos ($r < 1/\mu$) son más probables que los mensajes largos ($r > 1/\mu$).

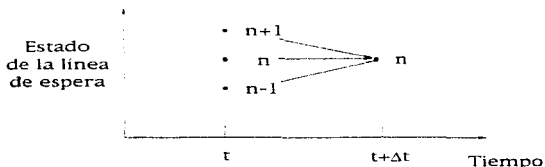
Es de hacer notar que siendo el tamaño de un mensaje una variable discreta, al ser el múltiplo entero de una unidad de datos, se le trata como una variable continua. Esta suposición permite un análisis más simplificado y se ha encontrado muy apropiada en el diseño de redes de datos.

Si la capacidad del canal de transmisión es C [unidades de datos/seg], los mensajes de tamaño r [unidades de datos] tomarán r/C [segundos], de transmisión.

Las propiedades estadísticas de la línea de espera M/M/1 se definen una vez calculadas las probabilidades de estado $p_n(t)$ en la línea, esto es, la probabilidad de que n mensajes se encuentren en línea de espera en el

tiempo t . estas probabilidades de estado son calculadas en estado estable en un sistema invariante en el tiempo.

Una forma de calcular las probabilidades de estado es analizando dos intervalos sucesivos de tiempo: t y $t+\Delta t$, donde $\Delta t \rightarrow 0$ como se muestra en la siguiente figura:



En un proceso de Poisson, considerando el estado n de la línea en el instante $t+\Delta t$, tiene tres posibles estados anteriores: $n+1$, n , $n-1$ en el instante anterior t . Esto significa también que durante Δt no ocurre más de una llegada y no más de una transmisión.

La probabilidad de que la fila se encuentre en el estado n en el instante $t+\Delta t$, denominada $p_n(t+\Delta t)$ es la suma de las siguientes probabilidades mutuamente exclusivas:

$$p_n(t+\Delta t) \text{ (Probabilidad de que la línea este en el estado } n \text{)}$$

- [Prob(no cambio de estado) = Prob(una llegada y una salida)]

$$p_{n+1}(t) \text{ (Probabilidad de que el estado sea } n+1 \text{)}$$

- [Prob(una salida, sin llegada)]

$$p_{n-1}(t) \text{ (Probabilidad de que el estado sea } n-1 \text{)}$$

- [Prob(una llegada, sin salida)]

El primer término es la probabilidad de que estando la línea en n , se conserve en el mismo estado.

El segundo término representa la probabilidad de que, estando la fila en el estado $n+1$, se desplace al estado n .

El tercer término es la probabilidad de estar en el estado $n-1$ y recorrerse al estado n . En notación resumida:

$$p_n(t+\Delta t) = p_n(t) + p_{n-1}(t) p(n+1 \rightarrow n) + p_{n+1}(t) p(n-1 \rightarrow n)$$

Esta ecuación se puede resolver para $p_n(t)$ asumiendo que las probabilidades de estado son continuas. Representando $p_n(t+\Delta t)$ por los dos primeros términos de su serie de Taylor:

$$p_n(t+\Delta t) = p_n(t) + dp_n(t) \cdot dt \cdot \Delta t$$

Tomando el límite cuando $\Delta t \rightarrow 0$ y representando las probabilidades de transición en términos de λ y μ :

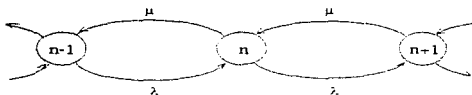
$$dp_n(t) \cdot dt = -(\lambda + \mu) p_n(t) + \lambda p_{n-1}(t) + \mu p_{n+1}(t)$$

que es la ecuación diferencial- en diferencias que hay que resolver para $p_n(t)$.

Como se estableció anteriormente, se considera un comportamiento estadísticamente estacionario, por lo que se asume que las probabilidades de estado son independientes del tiempo. Esta condición implica que $dp_n(t)/dt = 0$ y la expresión se simplifica a lo siguiente:

$$(\lambda + \mu) p_n = \lambda p_{n-1} + \mu p_{n+1} \quad n \geq 1$$

Esta ecuación se puede representar en el diagrama de la siguiente figura que representa las dos formas en las que la línea de espera puede encontrarse en el estado n a partir de los estados $(n-1)$ y $(n+1)$.



REPRESENTACION DE LOS ESTADOS DE UNA LINEA DE ESPERA

Las transiciones entre estados adyacentes se presentan a las tasas λ y μ mostradas, considerando que el proceso de entradas y salidas es de Poisson en estado estable.

El término $(\lambda + \mu) p_n$ representa la tasa a la que la línea de espera sale del estado n (dado que se encontraba en n con probabilidad p_n).

Los términos del lado derecho de la ecuación representan en cambio la tasa a la que el sistema entra al estado n , ya sea desde el estado $(n-1)$ o desde el estado $(n+1)$.

Para que las probabilidades de estado sean estacionarias, las dos tasas deben estar en equilibrio.

Una de las formas más aplicadas para resolver estas ecuaciones de equilibrio de tráfico tienen su origen en el análisis de flujo.

Obteniendo una ecuación de balance de tráfico entrante con saliente en el conjunto de estados de 0 a n :

$$\lambda p_n = \mu p_{n+1}$$

Esta ecuación es válida para cualquier conjunto de estados contiguos. La forma más sencilla es incluir el estado $n=0$, que solamente tiene dos términos de transición, hacia y desde el estado $n=1$:

$$\lambda p_0 = \mu p_1$$

Para cualquier estado n , se puede obtener p_n en forma recursiva a partir de p_0 :

$$p_n = (\mu/\lambda) p_{n-1} = (\mu/\lambda)^n p_0 = \rho^n p_0$$

Donde $\rho = \mu/\lambda$ es denominada intensidad de tráfico. Es en esta expresión donde se puede observar más claramente que la condición $\mu/\lambda < 1$ es un indicativo de la estabilidad del sistema y de que las probabilidades de estado decrecen a medida que n se incrementa. Esta característica concuerda con el concepto de que la tasa λ de llegadas a un nodo debe ser menor a la tasa μ de salidas, de otra manera el sistema tiende a la saturación y la fila crece indefinidamente, así como el tiempo de espera.

Para encontrar una relación que involucre la intensidad de tráfico con la probabilidad del estado inicial p_0 , se emplea la condición de probabilidad:

$$\sum p_n = 1$$

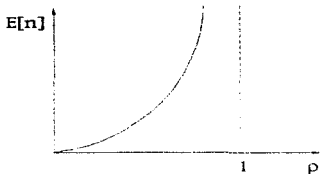
De donde $p_0 = 1 - \rho$, en función de las tasas de tráfico; $p_n = (1 - \rho)^n$

$\rho = 1 - p_0$, representa la probabilidad de que la línea de espera no se encuentre vacía. La distribución de probabilidad de estados en el caso M/M/1 es de tipo geométrica.

El tamaño promedio de la línea de espera (en mensajes) es:

$$E[n] = \rho / (1 - \rho)$$

Este tamaño promedio crece a medida que $\rho \rightarrow 1$. Este es el mismo fenómeno de tiempo de espera que se analizó para los sistemas de acceso aleatorio. A medida que la eficiencia del canal se aproxima a su límite máximo, el tiempo de respuesta se incrementa. La relación $E[n]$ vs. ρ se ilustra a continuación:



En condiciones de baja densidad de tráfico, el tráfico efectivo a través del canal es proporcional al tráfico de mensajes entrando al sistema. Cuando el tráfico se incrementa, el punto de operación se aproxima a $\rho = 1$ y el sistema se comienza a saturar, formando una cada vez más larga línea de espera y en consecuencia un mayor retraso.

Existe un compromiso en el diseño de redes de comunicación de datos entre un factor positivo que es el tráfico efectivo contra un factor negativo que es el retraso.

Obtenido el número promedio de mensajes en la línea de espera en función de los parámetros de ocupación del sistema, el siguiente cálculo es el del tiempo promedio de retraso.

Cuando un mensaje llega a la fila con $E[n]$ mensajes esperando ser transmitidos antes que él, el tiempo total de retraso de servicio es:

$$T = \text{tiempo de servicio} + \text{tiempo de espera}$$

La derivación más general de esta ecuación viene de la fórmula de Little, que establece que el tamaño promedio de una línea de espera es igual al retraso promedio en el sistema multiplicado por la tasa promedio de llegadas.

$$\lambda E[T] = E[n]$$

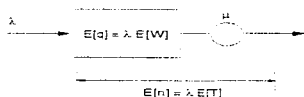
$$E[T] = E[n] / \lambda = 1 / (\mu(1-\rho))$$

Para una intensidad de tráfico muy baja $\rho \ll 1$, el tiempo promedio de espera es despreciable y el tiempo total de retraso es $1/\mu$, el tiempo de servicio de transmisión.

Para líneas de espera de un solo "servidor", la relación entre el tiempo promedio de espera $E[W]$ y el retraso promedio total $E[T]$ en todo el proceso es:

$$E[T] = E[W] + 1/\mu$$

Esta relación se muestra en la siguiente figura:



El teorema de Little establece una relación explícita para calcular el número promedio de mensajes $E[q]$ en la línea de espera, antes de acceder al servicio de transmisión.

$$\begin{aligned} E[q] &= \lambda E[W] = \lambda E[T] - \lambda / \mu \\ &= E[n] - \rho \end{aligned}$$

Líneas de Espera M/G/1

En esta sección se extiende el análisis de línea de espera al caso de una distribución general de tiempo de servicio. Los mensajes pueden tener una distribución arbitraria en la duración del servicio. El proceso de llegada de mensajes se considera de Poisson con un solo "servidor" y una capacidad ilimitada de memoria en el nodo.

El objetivo de este análisis es caracterizar la ocupación promedio de la línea de espera y, por medio de la fórmula de Little, la expresión del tiempo promedio de retraso.

La expresión resultante se conoce como fórmula de Pollaczek- Khinchine. Esta fórmula permite manejar modelos estadísticos de tamaño de mensaje que representen casos reales con mayor detalle. El modelo permite además visualizar como las diferentes distribuciones estadísticas determinan el tiempo de retraso.

Un caso típico es el de los mensajes de tamaño constante, un modelo muy común para redes interactivas de datos y redes de conmutación de paquetes, en las que la mayoría de los mensajes son de este tipo.

El principio de este análisis se enfoca en los intervalos de tiempo (de distribución aleatoria) en los que el servicio de transmisión de un mensaje es completado.

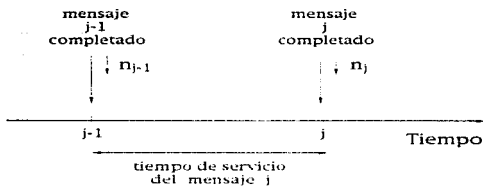
Se considera una secuencia de mensajes ingresando a la memoria del nodo, cada uno con un tiempo de secuencia aleatorio.

Definiendo el número variable de mensajes en la línea de espera: n_j es el tamaño de la línea en [mensajes] después de la transmisión del mensaje j . Una ecuación que relacione n_j con n_{j-1} después de la transmisión del mensaje $j-1$ será:

$$n_j = (n_{j-1} - 1) + v_j \quad n_{j-1} \geq 1$$
$$v_j \quad n_{j-1} = 0$$

v_j es el número de mensajes que ingresan al nodo durante el tiempo de servicio del mensaje j . v_j es en si una variable aleatoria.

La ecuación establece simplemente que el tamaño de la fila de mensajes reduce su tamaño en 1 al transmitirse un mensaje y lo aumenta de acuerdo a las llegadas al sistema.



DISTRIBUCIÓN GENERAL DE TIEMPO DE SERVICIO

Utilizando la notación de Schwartz [11] :

$$n_j = (n_{j-1} - 1) \dagger + v_j$$

El significado de la notación:

$$\begin{aligned} a \dagger &= a & \text{si } a \geq 0 \\ a \dagger &= 0 & \text{en caso diferente} \end{aligned}$$

Las variables del lado derecho de la ecuación son variables aleatorias independientes. La variable de estado n_j es la suma de dos variables aleatorias independientes. Su distribución de probabilidad es por lo tanto la convolución de dos distribuciones de probabilidad.

La herramienta matemática empleada para convertir la convolución de dos variables discretas en un dominio a la multiplicación de sus respectivas transformadas en otro dominio es la función de generación de momentos.

Para la variable n , su transformada es:

$$G_n[z] = E[z^n] = \sum_{k=0}^{\infty} p(n=k) z^k$$

Y la transformada de n_j en función de las transformadas de a y v_j :

$$G_{n_j}[z] = G_a[z] G_{v_j}[z]$$

Considerando que el sistema ha estado en operación en equilibrio estadístico, la expresión se puede establecer independientemente de j :

$$G_n[z] = G_a[z] G_v[z]$$

$G_v[z]$ es obtenible a partir de la estadística de v , la que se supone conocida y de la que se pueden obtener las probabilidades de estado $p(n=k)$ que en forma simplificada se denomina p_n .

$$G_a[z] = p_0 + p_1 z + p_2 z^2 + \dots$$

Las probabilidades de $(n-1)!$ se pueden denotar en función de las probabilidades de n .

$$G_a[z] = p_0 + (G_n[z] - p_0) z$$

que sustituyendo para $G_n[z]$ se tiene:

$$G_n[z] = (p_0(z+1) G_v[z]) / (z - G_v[z])$$

Las propiedades de estado de la línea de espera pueden ser derivadas a partir de $G_n[z]$. Schwartz demuestra en [11] que partiendo de las propiedades de la función de generación de momentos:

$$E[v] = d G_v[z] / dz |_{z=1}$$

y en el caso M/G 1:

$$G_n[z] = (1-\rho) / (1-\rho z)$$

Expandiendo $(1-\rho z)^{-1}$ en una serie infinita en z y asumiendo: $\rho < 1$ y $|z| \leq 1$ para la convergencia de la función de generación de momentos:

$$G_n[z] = (1-\rho) \sum_{k=0}^{\infty} (\rho z)^k$$

$$= \sum_{k=0}^{\infty} (1-\rho) \rho^k z^k$$

El coeficiente de z^k es entonces:

$$p_k = (1-\rho) \rho^k$$

que representa la probabilidad de que haya k mensajes en la línea de espera.

Para encontrar el número promedio de mensajes se recurre a la evaluación en $z=1$ de la derivada con respecto a z de la función de generación de momentos.

$$E[n] = d G_n[z] / dz |_{z=1}$$

$$= (1/(1-\rho)) ((E[v] - 2E^2[v] + E[v^2])/2)$$

$$= (1/2(1-\rho)) (\sigma_v^2 + \rho(1-\rho))$$

$\sigma_v^2 = E^2[v] + E[v^2]$ es la varianza del número de mensajes entrantes a la línea de espera durante un intervalo aleatorio de servicio.

De la segunda derivada con respecto a z de $G_n[z]$:

$$\sigma_v^2 = d^2 G_v[z] / dz^2 |_{z=1} + \rho - \rho^2$$

$$d^2 G_v[z] / dz^2 |_{z=1} = \lambda^2 (\sigma^2 + (1/\mu)^2)$$

En este caso σ^2 es la varianza de la distribución del tamaño del mensaje. El tamaño promedio de la línea en mensajes está relacionado directamente con el factor de intensidad de tráfico ρ , la duración promedio del servicio del mensaje $1/\mu$ y la varianza del tamaño de los mensajes.

$$E[n] = (1/(1-\rho)) (\rho - \rho^2 (1-\mu^2 \sigma^2)/2)$$

Finalmente, el retraso promedio es calculado con la fórmula de Little:

$$E[T] = E[n] / \lambda = (1 - 2\rho + (1 - \rho) / (2 - \rho) (1 - \rho^2 \sigma^2))$$

Esta expresión de retraso para sistemas M/G/1 es muy similar a la de los sistemas M/M/1. Ambos casos difieren en el segundo término, dependiendo únicamente de la diferencia $1 - \rho^2 \sigma^2$. Una distribución con alta varianza tiene una mayor probabilidad de mensajes largos y por lo tanto un mayor retraso promedio.

La fórmula de Pollaczek - Khinchine tiene como objetivo calcular el tiempo de retraso promedio en una línea M/G/1. De manera similar, para calcular el tiempo que el mensaje de espera, en el que el mensaje es almacenado antes de ser transmitido $E[W]$, este es simplemente el tiempo promedio de retraso menos el tiempo de servicio $E[T] = E[W] + 1/\mu$.

$$E[W] = \lambda E[\tau^2] / (2(1 - \rho))$$

El tiempo promedio de espera $E[W]$ depende del segundo momento $E[\tau^2]$ de la distribución del tamaño del mensaje.

Capítulo 5

Canal TDMA

El método de acceso TDMA ha demostrado tener una eficiencia superior en el uso del canal de comunicación, a un costo de mayor complejidad en el equipo que compone la red. El uso de este esquema de acceso requiere normalmente de un control centralizado, considerando especialmente los segmentos de tiempo que son asignados en forma determinística. Desde el análisis de métodos híbridos de acceso como el de reservación descrito en el capítulo 3, se asume que la administración del canal de retorno es realizado desde el nodo central.

En este capítulo se analiza el caso específico de la eficiencia y tiempo de respuesta de un sistema TDMA. Los sistemas de acceso aleatorio mostrados en los capítulos anteriores requieren de una complejidad menor y tienen un tiempo de respuesta más corto, con la desventaja de una baja eficiencia en la utilización del ancho de banda del canal.

El objetivo principal de los métodos de acceso es resolver los dos problemas principales de "resolución de conflicto", que son:

- Identificar a los usuarios con datos a enviar
- Asignar el recurso del canal a esos usuarios

Un postulado de la ley de la conservación de la teoría de colas es que el desempeño retraso-"throughput" de un sistema de acceso múltiple se determina esencialmente por el tiempo transcurrido en identificar al usuario listo para transmitir y asignarle acceso al canal.

En el caso de tráfico intermitente, que es de naturaleza más impredecible, la utilización del canal es estadística. Por otro lado, cuando el canal conduce transferencia de archivos en forma rutinaria, el canal se puede considerar dedicado al menos por un periodo de tiempo, durante el cual se le puede considerar determinístico.

Los orígenes del esquema TDMA encontraron especial aceptación en escenarios de tráfico interactivo de tipo terminal-computadora central. El uso de canales vía satélite para transmisión de datos tiene un número de ventajas significativas comparado con canales terrestres. Una de las principales ventajas es el potencial de máxima conectividad, de tipo punto - multipunto. Las redes satelitales utilizando el método TDM/TDMA fueron diseñadas originalmente para reemplazar las redes terrestres de computadoras de protocolo SNA/SDLC y X.25. Por la característica del tiempo de propagación del canal satelital, se diseñaron también técnicas de compensación de retraso, como emulación local de protocolos de tipo encuesta "polling".

Desde el punto de vista protocolo satelital, los paquetes de datos son procesados ya sea como unidades uniformes de datos o asignando prioridades de tiempo de respuesta, de control de flujo y de método de retransmisión.

Las técnicas tradicionales de acceso múltiple se basan en la división del canal principal en subcanales ya sea fijos o variables. En cuanto a la distribución de tiempos de transmisión en particular, los esquemas de tiempos fijos de transmisión han encontrado su sucesor en los sistemas de paquetes de tamaño fijo ("celdas") como los utilizados en sistemas de alta tasa de comunicación como ATM. Los esquemas de tiempos variables de distribución, en cambio, tienen como objetivo ajustar o adaptar el tiempo de acceso a un nodo determinado en función de sus requerimientos de tráfico.

En esquemas de acceso TDMA de asignación fija, se considera un equilibrio entre el tráfico que ingresa a la red con el tráfico efectivo "throughput": no hay bloqueos de tráfico, no hay colisiones y el desempeño del esquema es cuantificado en términos de tiempo de respuesta. Este desempeño se deteriora en la región en que el tráfico en la red se aproxima a la capacidad del canal, donde el sistema tiende a la inestabilidad.

Para los esquemas de asignación de tiempos de transmisión, se presentan variables de configuración ya sea del canal o en los nodos de la red, que no son aplicables en los casos de acceso aleatorio. Uno de estos parámetros es la capacidad finita de memoria en el nodo, parámetro que

ESTA
SALIR
DE LA
BIBLIOTECA
NO DEBE
TESIS

se supone lo suficientemente grande como para considerar el bloqueo de mensajes a ingresar al sistema de comunicaciones.

En este protocolo el tiempo de uso del canal es asignado a cada estación de acuerdo a su requerimiento anticipado de tráfico.

En su esquema más simple, cada una de las M estaciones tiene un tiempo preasignado de transmisión y el marco de tiempo T se divide en lapsos fijos que pueden ser de diferente duración para cada estación. A continuación se muestra el esquema básico de distribución de tiempos.

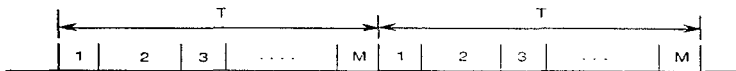


Diagrama de asignación fija de tiempos de transmisión.

En este caso el protocolo de transmisión consiste en que la estación espera la llegada de su correspondiente intervalo de tiempo y transmite los paquetes que se hayan acumulado durante el tiempo de espera.

Dentro del intervalo de transmisión de cada estación es posible hacer una subdivisión de tiempo entre los usuarios de dicha estación.

Esto permite la comunicación de varios paquetes de datos en una sola ráfaga. Cada paquete contiene una dirección de la estación destino, como si hubieran sido enviadas en distintas ráfagas.

Esta subdivisión de tiempo puede ser de asignación fija también o estar sujeta a distribución variable, de acuerdo al tráfico de la estación.

Tiempo de espera con memoria finita

En esta sección se analiza el desempeño de un sistema TDMA con asignación fija de ranuras y una capacidad finita de memoria "buffer" en el nodo. El modelo de línea de espera formulado por Hillier y Jabbari [3] tiene por objetivo derivar una expresión para el tiempo promedio de espera y la probabilidad de bloqueo del mensaje.

En este modelo se considera que los mensajes generados en las estaciones están compuestos de uno o más paquetes, mientras el tamaño de los mensajes originales depende de las aplicaciones de la comunicación de datos, el tamaño de los paquetes esta en funcion del sistema de comunicación en si. Los mensajes son segmentados en paquetes de tamaño fijo por el sistema procesador de comunicaciones a nivel banda base, normalmente un convertidor de protocolo.

Los paquetes son almacenados en un espacio de memoria temporal "buffer" antes de su transmisión por el canal. Como el espacio de memoria es finito, los mensajes generados despues de que la memoria está llena están sujetos a bloqueo. En este analisis se aplican modelos de lineas finitas de espera para representar el sistema. Este sistema considera que los mensajes pueden provenir de una o varias fuentes de información en el nodo, cuyas señales son multiplexadas para su transmisión por el canal. Estos mensajes son generados de acuerdo a un proceso de Poisson.

En cuanto al tiempo de respuesta a ser cuantificado, se define como el tiempo transcurrido desde que el mensaje se entrega en el nodo origen hasta que es recibido en el nodo destino. Esto significa básicamente el tiempo que dura el paquete almacenado mas el tiempo de propagación nodo origen - satélite - nodo destino. El tiempo de procesamiento de la señal en los nodos tales como ensamble y desensamble de paquetes se considera despreciable comparado con el tiempo de respuesta. El cálculo subsecuente se concentra en el retraso en el "buffer" y se asume como constante el tiempo de propagación por el canal.

El canal de comunicacion se supone exclusivamente satelital, de acceso por tiempo dividido en ranuras de duracion constante e igual a la duracion del paquete a transmitirse.

La red a ser analizada consiste de N estaciones compartiendo un canal con una tasa de transmision de C [bits-seg] y un tiempo de trama TDMA de T [segundos]. Cada estacion puede transmitir paquetes en un número S de ranuras asignadas por cada trama. Los mensajes generados en el nodo son almacenados en el "buffer" hasta que ocurre su tiempo de transmisión. El

análisis se puede concentrar en una sola estación y asumir que el tráfico del resto de la red no afecta el del nodo en estudio.

Sobre los mensajes generados en el nodo en particular se presumen las siguientes características de tráfico:

- Los mensajes se generan o concentran en el nodo con una probabilidad de Poisson a una tasa de λ [mensajes · seg].

- Cada mensaje se segmenta en uno o más paquetes de P [bits · paquete].

- La memoria del "buffer" tiene una capacidad de almacenar un máximo de M [paquetes]. Los mensajes, o parte de ellos serán aceptados en la memoria del nodo si hay capacidad disponible, de otra manera serán descartados.

- Los mensajes son transmitidos, un paquete por ranura, con un esquema FIFO, los primeros paquetes en llegar son los primeros en salir, sin prioridades establecidas. El número de paquetes transmitidos por trama es el mínimo de S y el número de paquetes esperando en el "buffer".

- El número S es pequeño en relación al número total de ranuras en la trama. Esto permite suponer que la transmisión de mensajes en la misma trama ocurre al mismo tiempo.

En términos de teoría de colas, se trata de un modelo con un solo servidor, una línea finita de espera, un proceso de llegada de Poisson y tiempos constantes de servicio. Los clientes son los paquetes, la línea de espera es el "buffer" y el servidor es el canal, que se considera ocupado tanto durante el tiempo de transmisión de los S paquetes como el "tiempo de recuperación", hasta que se presentan los S ranuras de la trama siguiente. Una diferencia importante con los modelos típicos de teoría de colas es que en este caso el servidor está siempre ocupado, en el sentido de que todas las ranuras de la trama son asignadas entre los nodos de la red, aún cuando no tengan tráfico a transmitir.

Modelo matemático del sistema.

El modelo a utilizar es el de cadenas de Markov. a_n se define como la probabilidad de que se generen n paquetes durante la trama T :

$$a_n = (\lambda T)^n e^{-\lambda T} / n! \quad n = 0, 1, 2$$

Definiendo como estado del sistema a el número de paquetes en la memoria "buffer", excluyendo los paquetes que acaban de ser transmitidos y con la condición de que $S \leq M$, que es el caso de interés, se tiene la matriz de probabilidad de transición P de orden $m \times m$.

Las probabilidades de estado estable $\{\pi_j, j = 0, 1, \dots, m\}$ se calculan por medio de:

$$\pi = \pi P$$

donde $\pi = [\pi_0, \pi_1, \pi_2, \dots, \pi_m]$

Un mensaje es bloqueado si la cadena de Markov estaba en estado "j" y hay al menos $(M - j)$ mensajes generados previamente a que comience la siguiente trama.

Sea R el número de mensajes rechazados durante una trama el número estimado de mensajes bloqueados por trama es:

$$E(R) = \sum_{j=0}^m \pi_j \sum_{n=j}^M a_{n,j}$$

La probabilidad de que un mensaje sea bloqueado está dada por:

$$P_b = E(R) \cdot \lambda T$$

El segundo factor de desempeño del sistema a ser calculado es el tiempo de espera de los paquetes en el buffer (W_q). En términos de teoría de colas este es el tiempo de espera en la línea. Se denomina w a la variable aleatoria del retraso que experimenta un mensaje en el buffer. Si $w > 0$ significa que el mensaje no es bloqueado, mientras que si $w=0$ significa que el mensaje es bloqueado con probabilidad 1. La probabilidad de que un mensaje se genere a exactamente el mismo tiempo de inicio de la trama se considera cero. Por lo tanto:

$$\begin{aligned} E(w) &= E(w|w > 0) (1 - P_b) + E(w|w=0) P_b \\ &= W_q (1 - P_b) \end{aligned}$$

Para el cálculo de $E(w)$, se define E_m como el evento de que la generación de un mensaje cumpla con las siguientes condiciones:

- Quedan i mensajes en el buffer.

- Es uno de los n mensajes generados durante la trama presente.

$P(E_m)$ es la probabilidad de que el evento E_m ocurra para la variable w :

$$E(w) = \sum_{i=0}^n \sum_{k=0}^i P(E_{ik}) E(w|E_{ik})$$

El siguiente paso es encontrar una expresión de $P(E_m)$ el cálculo de la probabilidad de que la generación de un mensaje dado es uno de los n en la trama requiere involucrar a_n con n , de manera que:

$$P(E_m) = \pi a_n \sum_{j=0}^n j! a_j$$

Ahora se procede al cálculo de $E(w|E_m)$, considerando el caso de que todos los mensajes generados durante el intervalo $[0, T]$ de la trama serán almacenados en el buffer, esto es: $i + n \leq M$.

El tiempo promedio de espera para cada uno de los mensajes generados, calculado en [3] está dado por:

$$E(w|E_m) = T \left[2 + \frac{1}{n} \sum_{k=1}^n \min \{n, (n+i-S_k)\} \right] \{ T$$

donde $(x)^+ = \max \{0, x\}$

Resultados numéricos. W_q y P_b

Para los casos de mensajes de un solo paquete se desarrollaron los cálculos de P_b y de W_q .

Las gráficas de probabilidad de bloqueo P_b con intensidad de tráfico λT como variable independiente se muestran en la siguiente sección. Una pequeña diferencia con la forma de expresar los resultados con los autores del análisis. Hillier y Jabbari [3] es la variable independiente que ellos utilizan de $\rho = \lambda T/S$. Para el eje de P_b se aplicó una escala lineal, a diferencia de los autores, que utilizaron una escala logarítmica.

El primer caso analizado consistió de los parámetros $S = M = 2$, con dos ranuras por trama por cada estación y una capacidad de memoria de dos paquetes. El segundo caso consideró $S=1$, $M=3$.

Los resultados para W_q con la misma variable de intensidad de tráfico se muestran asumiendo la duración de la trama T como la unidad de tiempo.

Como se anota en [3], a medida que se incrementa el tamaño del buffer M se obtiene el límite:

$$\lim_{M \rightarrow \infty} W_q = 0.5 T + \rho T / (2(1 - \rho))$$

que es el resultado del caso del buffer infinito cuando el mensaje consiste de un paquete. El primer término es debido a la espera promedio de que ocurra el tiempo asignado al nodo, el segundo término es el valor esperado del tiempo de espera en línea para un sistema $M/D/1$. Los resultados muestran que cuando un canal de capacidad C es compartido entre un número grande de usuarios, el desempeño del canal (probabilidad de bloqueo y tiempo de respuesta) asignando una ranura por estación es superior al caso de varias ranuras por trama.

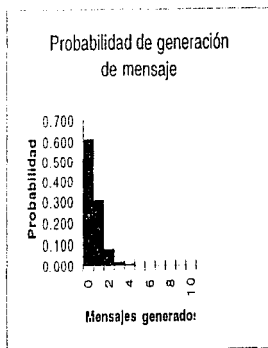
Ecuación 1, a_n es la probabilidad de que se generen n mensajes en el frame.

$$a_n = \frac{[(\lambda T)^n \cdot e^{-\lambda T}]}{n!}$$

$$n \geq 0$$

$\lambda = 0.5$ [packet/sec]
 $T = 1$ [sec/packet]
 $\lambda T = 0.5$

n	a_n	
0	0.607	0.000
1	0.303	0.303
2	0.076	0.152
3	0.013	0.038
4	0.002	0.006
5	0.000	0.001
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000



$$\Sigma = 1.000 \quad E[n] = 0.500$$

Ecuación 2, Matriz de transiciones

$$\pi = \pi P$$

$m = M - S = 0$ para este segundo ejemplo

M es el tamaño de la memoria buffer

S es el número de slots asignados en el frame para una determinada estación

m es el estado del buffer

Observación: Este análisis se aplica para

Ejemplo: con $S=2$, $m=0$, $M=2$ tráfico ligero, de manera que los mensajes no se bloqueen ya que el buffer es muy pequeño

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\pi_{0,0} = 1$$

$$v_2(j=0) = 0.0163$$

Ecuación 3, Promedio de mensajes rechazados

$$E(R) = 0.0163$$

Ecuación 5, Probabilidad de bloqueo de mensajes

$$P_B = E(R) / \lambda T = 0.032699$$

Ecuación 8, $P(E_{in})$ para calcular $P(w)$

$$\sum_j a_j = 0.500$$

	n=1	n=2
$P(E_{in})$:	0.607	0.303

Ecuación 10, $E(w|E_{in})$

Nota: i y n serán parte de una sumatoria, para el cálculo de $P(w)$
 los límites de i son $\{0,0\}$ (sólo un término) y los de n serán $\{1,\infty\}$,
 con la limitante de que, para que no haya bloqueo: $n+i \leq M$

i	n	$f(i,n)/T$	$E(w E_{in})$	$P(E_{in})$	
0	1	0.0	0.500	0.607	0.30327
0	2	0.5	1.000	0.303	0.30327

Estos son factores que relacionan el
 tamaño del buffer con la relación de mensajes
 generados/almacenados

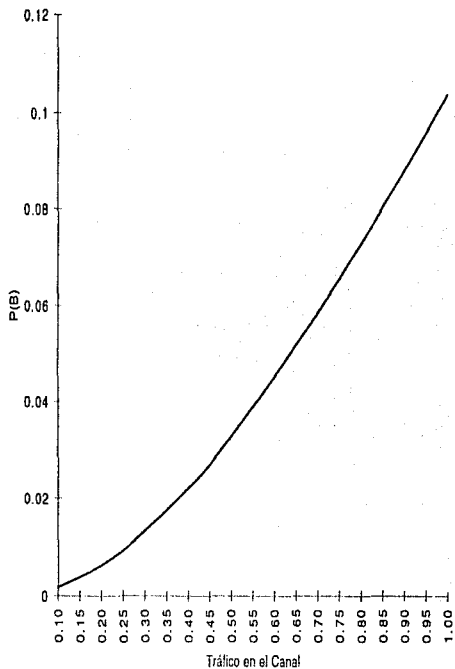
Ecuación 7, $E(w) = \sum \sum P(E_{in}) E(w|E_{in})$

$$E(w) = \sum \sum P(E_{in}) E(w|E_{in}) : 0.60653070$$

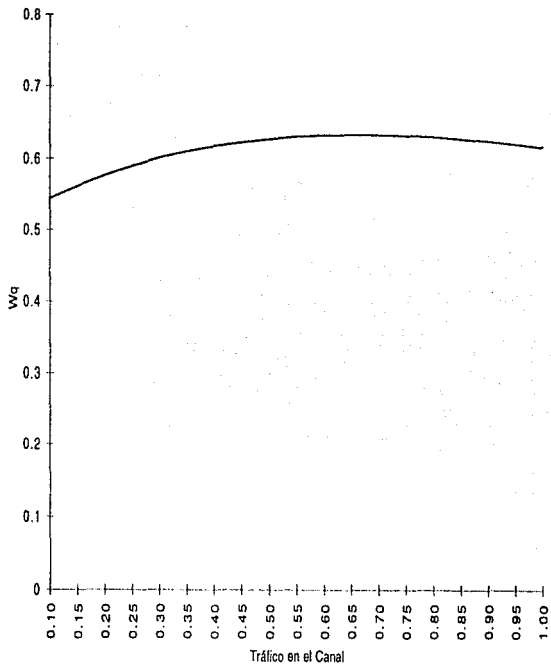
Ecuación 6, Tiempo de espera en línea en función del tráfico

	λT	P_B	W_q
$W_q = E(w)/(1-P_B)$			
$W_q = 0.6270$	0.10	0.001643	0.5438
	0.15	0.003537	0.5614
	0.20	0.006093	0.5766
	0.25	0.0093	0.5896
	0.30	0.01299	0.6005
	0.35	0.017241	0.6095
	0.40	0.0220	0.6168
	0.45	0.027134	0.6226
	0.50	0.0327	0.627
	0.55	0.03863	0.6301
	0.60	0.044893	0.6321
	0.65	0.051459	0.6329
	0.70	0.058298	0.6328
	0.75	0.065383	0.6318
	0.80	0.072689	0.6299
	0.85	0.080193	0.6273
	0.90	0.087871	0.624
	0.95	0.0957	0.6201
	1.00	0.10367	0.6157

Probabilidad de Bloqueo
Acceso TDMA



Tiempo de Espera
Acceso TDMA



Tiempo de Espera
Hillier/Jabbari

Ecuación 1, a_n es la probabilidad de que se generen n mensajes en el frame.

$$a_n = ((\lambda T)^n \cdot e^{-\lambda T}) / n!$$

$n \geq 0$

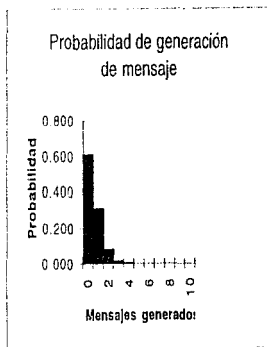
$\lambda = 0.5$ [packet/sec]

$T = 1$ [sec/packet]

$\lambda T = 0.5$

n	a_n	
0	0.607	0.000
1	0.303	0.303
2	0.076	0.152
3	0.013	0.038
4	0.002	0.006
5	0.000	0.001
6	0.000	0.000
7	0.000	0.000
8	0.000	0.000
9	0.000	0.000
10	0.000	0.000

$\Sigma = 1.000$ $E[n] = 0.500$



Ecuación 2, Matriz de transiciones

$$\pi = \pi P$$

$$m = M - S$$

M es el tamaño de la memoria buffer

S es el número de slots asignados en el frame para una determinada estación

m es el estado del buffer

Ejemplo: con $S=1$, $m=2$, $M=3$

P=	0.9998	0.0758	0.0144
	0.6065	0.3033	0.0902
	0.0000	0.6965	0.3935

$$\pi_0 = 0.7460 \quad V_2(j=0) = 0.0020$$

$$\pi_1 = 0.1270 \quad V_2(j=1) = 0.0163$$

$$\pi_2 = 0.0189 \quad V_2(j=2) = 0.1065$$

Observación: Este análisis se aplica para tráfico ligero, de manera que los mensajes no se bloqueen ya que el buffer es muy pequeño

Ecuación 3, Promedio de mensajes rechazados

$$E(R) = 0.0056$$

Ecuación 5, Probabilidad de bloqueo de mensajes

$$P_B = E(R) \cdot T = 0.0111$$

Ecuación 8, $P(E_{in})$ para calcular $P(w)$

$$\sum_j a_j = 0.500$$

$P(E_{0n})$:	0.452	0.226	0.057	0.009	0.001
$P(E_{1n})$:	0.077	0.039	0.010	0.002	0.000
$P(E_{2n})$:	0.011	0.006	0.001	0.000	0.000

Ecuación 10, $E(w|E_{in})$

Nota: i y n serán parte de una sumatoria, para el cálculo de $P(w)$

los límites de i son $\{0, m\} = \{0, 2\}$ y los de n serán $\{1, \infty\}$,

con la limitante de que, para que no haya bloqueo: $n+i \leq M$

i	n	$f(i, n)/T$	$E(w E_{in})$	$P(E_{in})$	
0	1	0.0	0.500	0.452	0.22623
1		1.0	1.500	0.077	0.11555
2		2.0	2.500	0.011	0.02864
0	2	0.5	1.000	0.226	0.22623
1		1.5	2.000	0.039	0.07703
0	3	1.0	1.500	0.057	0.08484

Estos son factores que relacionan el tamaño del buffer con la relación de mensajes generados/almacenados

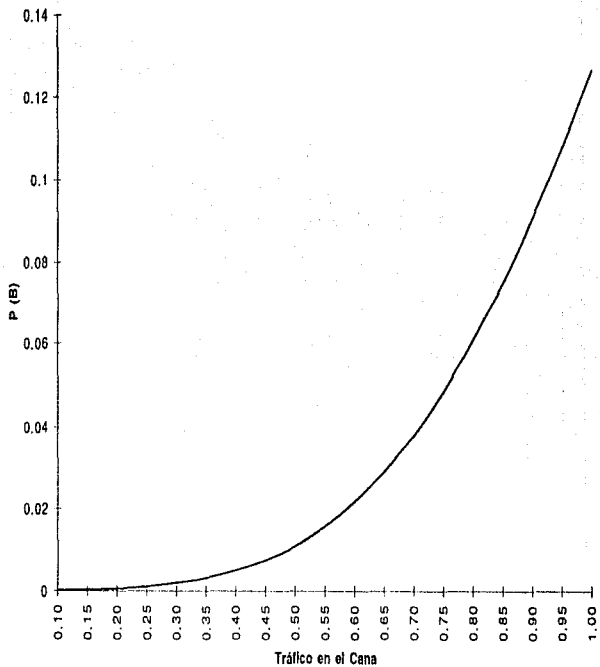
Ecuación 7, $E(w) = \sum \sum P(E_{in}) E(w|E_{in})$

$$E(w) = \sum P(E_{in}) E(w|E_{in}) = 0.75851989$$

Ecuación 6, Tiempo de espera en línea en función del tráfico

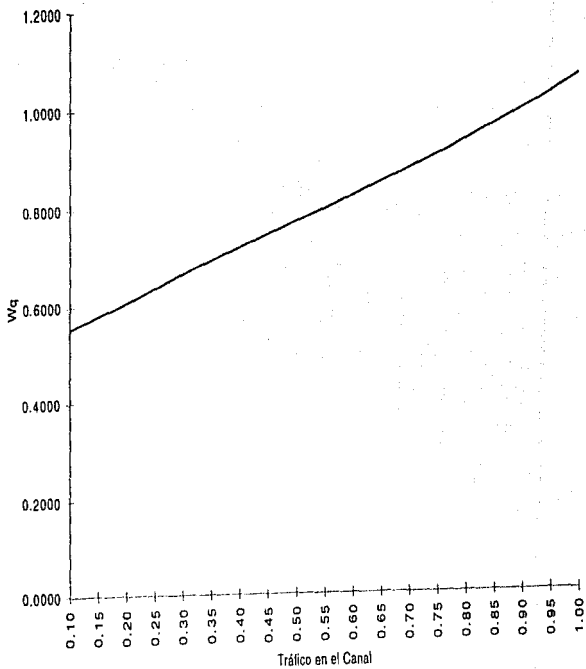
$W_q = E(w)/(1-P_B)$	λT	P_B	W_q
0.7671	0.10	0.0001	0.5521
	0.15	0.0003	0.5790
	0.20	0.0005	0.6060
	0.25	0.0010	0.6338
	0.30	0.0019	0.6610
	0.35	0.0031	0.6880
	0.40	0.0050	0.7147
	0.45	0.0076	0.7410
	0.50	0.0111	0.7671
	0.55	0.0157	0.7930
	0.60	0.0216	0.8190
	0.65	0.0289	0.8452
	0.70	0.0378	0.8720
	0.75	0.0484	0.8995
	0.80	0.0607	0.9280
	0.85	0.0747	0.9577
	0.90	0.0905	0.9887
	0.95	0.1078	1.0213
	1.00	0.1266	1.0555

Probabilidad de Bloque Acceso TDMA



Probabilidad de Bloqueo
Hillier/Jabbari

Tiempo de Espera
Acceso TDMA



Tiempo de Espera
Hillier/Jabbari

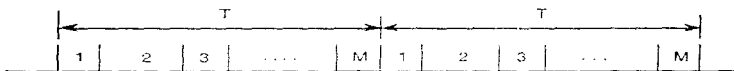
Cálculo del Tiempo de Retraso de un Canal TDMA

El análisis más completo de tiempo de respuesta de un canal TDMA lo realizó Lam en un artículo [5] que se convirtió en un clásico de la disciplina. Se considera un canal con ranuras de asignación fija, nodos sin restricción en el tamaño del "buffer" y un proceso de generación de mensajes de probabilidad de Poisson.

Cada mensaje consiste de uno o más paquetes de tamaño fijo que son transmitidos en ranuras conteniendo un paquete cada una. El número de paquetes que comprenden un mensaje es una variable aleatoria con una distribución de probabilidad general. Las condiciones estadísticas del tamaño de la línea de espera en el nodo a ser analizadas son de estado estable y la expresión a ser derivada calcula el tiempo promedio de retraso en la comunicación del mensaje, incluyendo en tiempo de espera y de transmisión.

El esquema de la señal TDMA es la típica, como se muestra en la figura. El tiempo es ranurado y las ranuras se organizan en marcos de M ranuras, numeradas del 1 al M .

Siendo la duración de la trama de T segundos, cada estación transmite un paquete en un intervalo de tiempo de T/M segundos y deja de transmitir por $(M-1)M/T$ segundos antes de restablecer la transmisión, a cada estación le corresponde una ranura por trama.



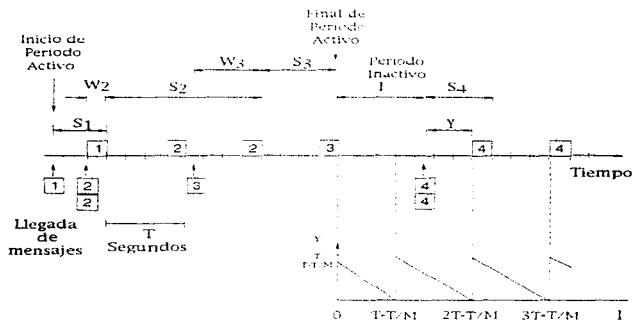
Para propósito de este análisis, se supone también que el desempeño de un canal TDMA es independiente para cada nodo de la red.

De manera similar al modelo de memoria finita en el nodo, el tamaño de los paquetes es función del sistema de comunicación y los mensajes son segmentados en paquetes de tamaño fijo por el sistema procesador de comunicaciones. Se trata también de una línea finita de espera de un solo servidor y tiempos constantes de servicio.

El tiempo de servicio de un paquete incluye por definición, el tiempo de espera en la línea y el tiempo de transmisión. Mientras el esquema de asignación de tiempo en un canal TDMA es determinístico, la relación entre el tiempo de generación de un mensaje con el tiempo que se le asigna para ser transmitido, es aleatoria.

Para un paquete en general, el tiempo de transmisión es T/M , el tiempo de espera se encuentra en el rango de 0 segundos (cuando el paquete ingresa al canal justo en el momento de transmisión) a T segundos (cuando el paquete espera toda una trama).

El tiempo de servicio es por lo tanto una variable aleatoria denominada X , distribuida entre T/M y $T+T/M$. El tiempo total de servicio de un mensaje es la suma de los tiempos de servicio de los paquetes que lo componen.



Estas definiciones de tiempo se ilustran en la figura. En ella se muestra la generación de cuatro paquetes y sus correspondientes tiempos de transmisión. S_i representa el tiempo de servicio y W_i el tiempo de espera del i -ésimo mensaje. Para una estadística de Poisson de llegada de mensajes, el período de canal desocupado I tiene distribución exponencial.

La variable aleatoria Y , definida como $X - T/M$, es una función de I . La relación entre estas dos variables se muestra en el diagrama.

El siguiente análisis tiene como objetivo encontrar la función de generación de probabilidad de estado estable del sistema. Sea N_t el número de mensajes en el sistema, tanto en la línea de espera como en servicio en el instante t . Se considera una línea de espera con un solo servidor y un proceso de generación de mensajes de probabilidad de Poisson de una tasa de λ [mensajes/segundo]. La distribución del tiempo de servicio de un mensaje que marca el inicio de un período activo del canal es $\hat{B}(x)$ con primeros y segundos momentos \hat{b}_1 y \hat{b}_2 . Todos los siguientes mensajes en el mismo período activo tiene tiempos independientes de servicio de distribución $B(x)$ con primeros y segundos momentos b_1 y b_2 .

Se tiene \hat{B}

$$P_n = \lim_{t \rightarrow \infty} \text{Prob} \{N_t = n\}$$

Y se definen las transformadas de P_n en z y de $\hat{B}(x)$ en s :

$$P(z) = \sum_{n=0}^{\infty} z^n P_n$$

$$B^*(s) = \int_0^{\infty} e^{-sx} dB(x)$$

$$\hat{B}^*(s) = \int_0^{\infty} e^{-sx} d\hat{B}(x)$$

Se emplea también el resultado del teorema invocado en el artículo de Lam [5] en el que para la condición de bajo tráfico $\lambda b_1 < 1$ se cumple:

$$P(z) = P_0 [z \hat{B}^*(\lambda - \lambda z) - B^*(\lambda - \lambda z)] / [z - B^*(\lambda - \lambda z)]$$

donde

$$P_0 = (1 - \lambda b_1) / [1 - \lambda(b_1 - \hat{b}_1)]$$

El teorema anterior se emplea para la obtención de las fórmulas del retraso promedio y del tamaño de la línea de espera en una estación utilizando un canal TDMA. Los momentos estadísticos de mayor orden de la variable del retraso del mensaje se puede obtener de manera similar. Para efectos de este análisis se consideran únicamente el promedio.

Denominando L como el número de paquetes en un mensaje, una variable aleatoria discreta con una densidad de probabilidad $\{g_i\}$ con primeros y segundos momentos L_1 y L_2 , donde

$$g_i = \text{Prob} \{L = i\}$$

Las distribuciones del tiempo de servicio se obtienen con las siguientes transformadas:

$$B^*(s) = \sum_{i=1}^{\infty} g_i [\beta^*(s)]^i$$

$$\tilde{B}^*(s) = \hat{\beta}^*(s) \sum_{i=1}^{\infty} g_i [\beta^*(s)]^{i-1}$$

donde:

$$\beta^*(s) = e^{-sT}$$

$$\hat{\beta}^*(s) = e^{-sL/M} \int_0^T e^{-sy} dF(y)$$

$$F(y) = \text{Prob} \{Y \leq y\}$$

La parte inferior de la figura muestra a la variable Y como una función del tiempo inactivo I que es exponencialmente distribuido:

$$F(y) = [e^{-\lambda(T-T/M)} [e^{\lambda y} - 1]] / (1 - e^{-\lambda T}) \quad 0 \leq y \leq T$$

$$= 1 \quad \text{para } y > T$$

donde $U(x)$ es la función escalón.

Los primeros y segundos momentos de Y están dados por:

$$\bar{Y} = (T - T/M - 1/\lambda) + T e^{-\lambda(T-T/M)} / (1 - e^{-\lambda T})$$

$$\overline{Y^2} = (T - T/M - 1/\lambda)^2 + (1/\lambda)^2 \\ + (T^2 - 2T/\lambda) e^{-\lambda(T-T/M)/(1-e^{-\lambda T})}$$

Y los primero y segundos momentos de X están dados por:

$$\overline{X} = \overline{Y} + T/M \\ \overline{X^2} = \overline{Y^2} + 2T\overline{Y}/M + (T/M)^2$$

Y en el caso de los tiempos de servicio de mensajes:

$$b_1 = L_1 T \\ \hat{b}_1 = (L_1 - 1) T + \overline{X} \\ b_2 = L_2 T^2 \\ \hat{b}_2 = \overline{X^2} + 2\overline{X}(L_1 - 1) T + (L_2 - 2L_1 + 1) T^2$$

La probabilidad P_0 de que el sistema no tenga tráfico:

$$P_0 = (1 - \lambda b_1) / [1 - \lambda(b_1 - \hat{b}_1)] \\ = (1 - \lambda L_1 T)(1 - e^{-\lambda T}) / [\lambda T e^{-\lambda(T-T/M)}]$$

\overline{N} es el número promedio (el valor esperado) de mensajes en el sistema y se obtiene evaluando la primera derivada de la función de generación de momentos $P(z)$ para $z=1$.

$$\overline{N} = \lambda \hat{b}_1 / [1 - \lambda(b_1 - \hat{b}_1)] \\ + \lambda^2 (\hat{b}_2 - b_2) / 2 [1 - \lambda(b_1 - \hat{b}_1)] \\ + \lambda^2 b_2 / 2(1 - \lambda b_1)$$

Una vez obtenido el número (esperado) de mensajes en el sistema se puede obtener, aplicando la fórmula de Little, el retraso promedio del sistema:

$$\overline{D} = \hat{b}_1 / [1 - \lambda(b_1 - \hat{b}_1)] \\ + \lambda(\hat{b}_2 - b_2) / 2 [1 - \lambda(b_1 - \hat{b}_1)] \\ + \lambda b_2 / 2(1 - \lambda b_1)$$

El valor esperado del tiempo de servicio de un mensaje es:

$$\begin{aligned}\bar{S} &= (1 - P_0)b_1 + P_0\hat{b}_1 = b_1 - P_0(b_1 - \hat{b}_1) \\ &= \hat{b}_1 / [1 - \lambda(b_1 - \hat{b}_1)]\end{aligned}$$

que es precisamente el primer término de los componentes de la expresión del retraso esperado en el sistema. Los dos términos restantes representan el valor esperado del tiempo de espera:

$$\begin{aligned}\bar{W} &= \lambda(\hat{b}_1 - b_2) / 2[1 - \lambda(b_1 - \hat{b}_1)] \\ &+ \lambda b_2 / 2(1 - \lambda b_1)\end{aligned}$$

Substituyendo en \bar{N} , \bar{D} , \bar{S} y \bar{W} los valores calculados de los momentos de los tiempos de servicio de los mensajes:

$$\begin{aligned}\bar{N} &= \lambda L_1 T - \lambda T / 2 + \lambda T^2 M + \lambda^2 L_2 T^2 / 2(1 - \lambda L_1 T) \\ \bar{D} &= L_1 T - T / 2 + T^2 M + \lambda L_2 T^2 / 2(1 - \lambda L_1 T) \\ \bar{S} &= 1/\lambda + (L_1 T - 1/\lambda) (1 - e^{-\lambda T}) / [\lambda T e^{-\lambda(T-\hat{b}_1)}] \\ \bar{W} &= L_1 T - T / 2 + T^2 M + \lambda L_2 T^2 / 2(1 - \lambda L_1 T) \\ &- 1/\lambda - (L_1 T - 1/\lambda) (1 - e^{-\lambda T}) / [\lambda T e^{-\lambda(T-\hat{b}_1)}]\end{aligned}$$

El cálculo del retraso promedio \bar{D} significa un resultado especialmente importante. Siendo C la tasa de transmisión del canal C [bits/seg], la tasa de transmisión de cada subcanal TDMA es entonces C/M [bits/seg].

Para fines de comparación con el tiempo de retraso de un canal FDMA, se considera un canal FDMA de tasa C/M [bits/seg]. Para un canal de este tipo el retraso se puede calcular por la fórmula de Pollaczek-Khinchin [4, Vol II]:

$$\bar{D}_{PK} = L_1 T + \lambda L_2 T^2 / 2(1 - \lambda L_1 T)$$

Este resultado es directamente proporcional al retraso del subcanal TDMA, de la siguiente forma:

$$\bar{D} = \bar{D}_{PK} - T/2 + T/M$$

Este resultado es interesante por su simplicidad. Es de hacer notar que si la duración del marco de tiempo T se comprime hasta cero, como por ejemplo, reduciendo el tamaño del paquete, a una distribución de tamaño de mensajes constante, se tiene:

$$\lim_{T \rightarrow 0} \bar{D} = \bar{D}_D$$

Resultados Numericos

En esta sección se ilustra el desempeño del sistema en términos de retraso promedio, la variable más importante para calificar el desempeño de sistemas TDMA de asignación fija. El número de paquetes de usuario en el sistema se puede calcular por la fórmula de Little. Se utilizan ejemplos numéricos con los siguientes parámetros:

Tasa de Transmisión del Canal: $C = 1.5$ MBPS

Número de Ranuras en la Trama: $M = 450$

Para la distribución de tráfico en el canal se consideraron tres casos, uno por tipo de mensaje: de 1, 2 y 3 paquetes por mensaje. La tasa de generación de mensajes es la misma variable para cada curva de la gráfica.

Las gráficas muestran el retraso promedio como variable dependiente, para diferentes intensidades de tráfico, que es la variable independiente. Para los diferentes casos, el retraso tiende a infinito a medida que el tráfico efectivo se aproxima al límite del canal.

La primera gráfica muestra el comportamiento del retraso para el caso de mensajes de un solo paquete. No incluye el retraso de propagación del canal, que es constante. La curva es monótonicamente creciente y la pendiente se acentúa al acercarse el límite de la capacidad del canal. En esta región, los datos en el canal sufren un retraso intolerable, en los niveles que usualmente los protocolos de nivel DLC y de red han calificado el envío del paquete como erróneo y los tiempos de espera máximo han expirado. Otra de los parámetros a saturarse es la memoria del nodo.

En la segunda gráfica se ilustran los tres casos mencionados de tamaño de paquete. La tasa de mensajes en el canal es la misma para los tres casos, por lo que los mensajes multi-paquetes saturaran el canal más rápido para una misma tasa de mensajes. Como se puede observar, el retraso es también mayor.

En la tercer gráfica se muestra el efecto de la variable duración de la trama T en el retraso promedio para un tráfico efectivo y un tamaño de mensaje fijo.

En cuanto a la importancia del tamaño del mensaje, si este es aleatorio, un paquete muy largo puede ser ineficiente por la fracción no utilizada en el último paquete del mensaje. Por otro lado, un tamaño muy pequeño de paquete significa un mayor porcentaje de encabezado ("overhead") por cada uno de ellos.

La regla aplicable es utilizar un tamaño mínimo natural en la distribución del tamaño de mensajes para dimensionar el paquete. Subsecuentemente, los mensajes largos son fragmentados linealmente en paquetes del tamaño mínimo seleccionado.

Análisis de Tiempo de Respuesta, Canal TDMA

Parámetros del tráfico en el canal

Perfil del tráfico:	Paq1	Paq2	Paq3	Paq4	
Tamaño en [paq] de los mensajes (l_i)	1				
Probabilidad del mensaje (p_i)	1	0	0		1

$$p_i \cdot l_i^2 \quad 1 \quad 0 \quad 0 \quad 0$$

$$L_2 = E[L^2] = 1$$

$$p_i \cdot l_i \quad 1 \quad 0 \quad 0 \quad 0$$

$$L_1 = 1$$

$$C = 1.5 \text{ [Mbps]}$$

$$M = 450 \text{ [ranuras/frame]}$$

$$T = 0.30 \text{ [seg/frame]}$$

$$L = 1000 \text{ [bits] (tamaño del paquete)}$$

$$\lambda_{q1} = 0.125 \text{ [mensaje/seg] (por nodo)}$$

$$E[N] = \lambda L_1 T - \lambda T/2 + \lambda T/M + \lambda^2 L_2 T^2 / [2(1 - \lambda L_1 T)]$$

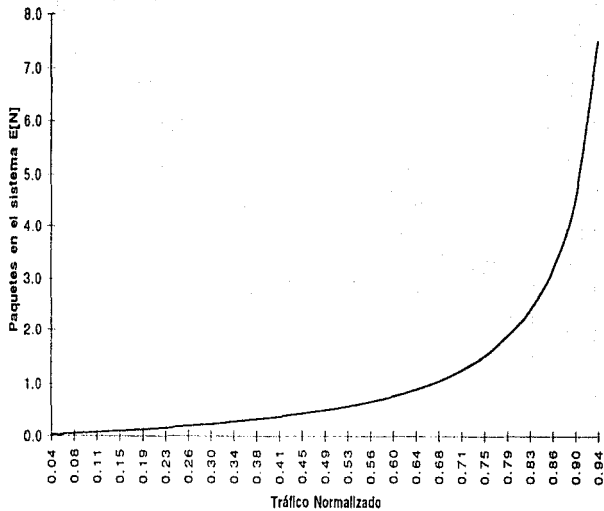
$$E[D] = E[N]/\lambda$$

$$E[S] = 1/\lambda + (L_1 T - 1/\lambda) (1 - e^{-\lambda T}) / [\lambda T e^{-\lambda T - T/M}]$$

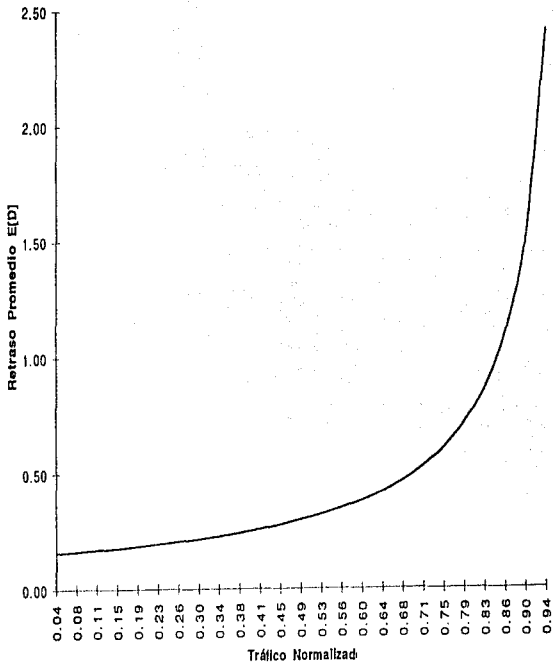
$$E[W] = E[D] - E[S]$$

ρ	λ	$E[N]$	$E[D]$	$E[S]$	$E[W]$
0.04	0.125	0.020	0.157	0.155	0.002
0.08	0.250	0.041	0.163	0.158	0.004
0.11	0.375	0.064	0.170	0.162	0.007
0.15	0.500	0.089	0.177	0.167	0.011
0.19	0.625	0.116	0.185	0.171	0.015
0.23	0.750	0.146	0.194	0.175	0.019

Paquetes en el Sistema
Canal TDMA

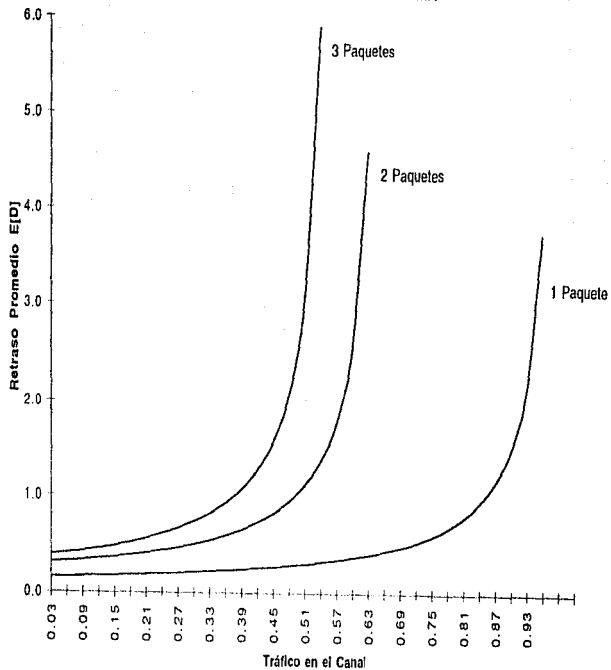


Retraso Promedio
Canal TDMA



Lam, Gráfica 1.1

Retraso Promedio Canal TDMA



Lam, Gráfica 2

CONCLUSION

La principal aportación de este trabajo fue reunir y analizar toda una familia de esquemas de acceso por división de tiempo.

Con los resultados obtenidos se pueden evaluar los parámetros promedio de desempeño de un canal. Para un tráfico determinado en el canal, se puede calcular el tiempo de respuesta del sistema de comunicación.

El esquema de acceso óptimo depende especialmente de la aplicación y está sujeto a condiciones adicionales al patrón de tráfico de la red, tales como costo y calidad de servicio. No hay un esquema que funcione en forma óptima en todos los escenarios.

La tecnología de las comunicaciones vía satélite ha dado la oportunidad de presenciar la evolución de la tecnología de la conmutación de paquetes utilizando sistemas inteligentes de acceso múltiple. Todos los esquemas analizados han sido observados y operados en redes instaladas.

Se puede afirmar que la motivación principal en la selección de este tema fue la integración de conceptos teóricos, bastante bien desarrollados en el mundo académico y de la investigación, con el mundo real.

Los esquemas de acceso aleatorio, por ejemplo, han sido utilizados para comunicación interactiva de datos y particularmente en el canal de señalización de redes de comunicación de voz tipo SCPC DAMA, de acceso por demanda. El caso Aloha de Retransmisión Selectiva fue analizado en redes en operación, en un sistema que utiliza conversión de protocolo de IP en mensajes TDM/TDMA.

Nos gustaría puntualizar dos áreas de estudio que tienen como objetivo optimizar estos sistemas de acceso, el aprendizaje del perfil de tráfico en el canal y el desarrollo de un esquema de control adaptivo de acceso al canal.

- **El análisis de las distribuciones de probabilidad de las variables que definen el tráfico que accesa el canal. Este primer aspecto está orientado a la coordinación de los mensajes de información desde su mismo origen, antes de acceder al canal.**

Se ha observado que el perfil de la distribución de probabilidad de generación de tráfico no siempre coincide con la de Poisson, consideración muy utilizada en el análisis de teoría de colas. Esta observación se ha efectuado en particular en el tráfico de redes locales (LANs) de tipo Ethernet y de redes extendidas WANs.

Los reportes sobre el monitoreo de tráfico masivo demuestran que este tipo de redes tienen una característica de tipo fractal. Esto significa que el tráfico tiene propiedades estadísticas similares para diferentes escalas de tiempo. Esto representa la necesidad de una redefinición de la relación entre parámetros de redes como capacidad de memoria de los nodos con la probabilidad de bloqueo de paquetes y su tiempo de respuesta.

Una solución propuesta en este aspecto es la implementación de esquemas de control dinámico de congestión, desarrollada en un protocolo de alto nivel, antes de que los paquetes soliciten acceso al canal de comunicación.

- **Predicción de tráfico y control adaptivo. El segundo nivel de control está dedicado al problema clásico de asignación de recurso de canal, dado un tráfico generado y solicitando acceso. Este tipo de conceptos se han propuesto implementar en esquemas de bajo nivel, a desempeñar funciones tanto de prevención como de adaptación ante el modelo de tráfico. La idea es esencialmente derivada del método de Roberts analizado en el capítulo 3.**

La asignación óptima del canal tendría una característica adaptiva de acuerdo a la dinámica del tráfico en la red y la variable a optimizar es el tiempo asignado de transmisión por nodo. El objetivo incluye definir esa variable por nodo en cada trama.

Entre los tipos de algoritmos más interesantes se encuentran las Redes Neuronales, la Lógica Difusa y los Algoritmos Genéticos.

Las Redes Neuronales encuentran una excelente aplicación en la implementación de controladores de contención por su capacidad de aprendizaje y de adaptación, además de su velocidad de procesamiento y de convergencia.

En cuanto a la Lógica Difusa, este tipo de algoritmos operan especialmente en los niveles de congestión de tráfico y de asignación de ancho de banda. En este tipo de sistemas de control adaptivo de tráfico, los parámetros de control no son establecidos mediante análisis previo, son aprendidos por la red y las decisiones son tomadas con base en información monitoreada directamente en la red en operación, por lo que están sujetos a un margen de incertidumbre.

Los algoritmos genéticos están siendo aplicados como procedimientos de búsqueda de soluciones a problemas de enrutamiento en redes de conmutación de paquetes basados en procesos de selección natural y de ingeniería genética.

Estos tres algoritmos han encontrado una especial aplicación especialmente en los protocolos de conmutación de paquetes de tamaño constante ("celdas"), como ATM y de conexión de redes locales.

REFERENCIAS

- [1] N. Abramson, " The Throughput of Packet Broadcasting Channels"
Vol COM-25, No 1, pp 117 - 128, Jan 1977.
- [2] R. Binder, "A dynamic packet switching system for satellite broadcast channels".
Proc. Int. Conf. Commun., San Francisco, CA, June 1975.
- [3] F. Hillier, B. Jabbari, "Analysis of the Fixed-Assigned TDMA Technique with Finite Buffer Capacity".
Vol COM-35, No 7, pp 724 - 729, July 1987.
- [4] L. Kleinrock, " Queuing Systems".
New York: Wiley, 1976
- [5] S.S. Lam, " Delay Analysis of a time division multiple (TDMA) channel".
IEEE Trans. Communications
Vol COM-25, pp 1489 - 1494, Dec 1977.
- [6] S.S. Lam, " Satellite packet communication- multiple access protocols and performance".
IEEE Trans. Communications
Vol COM-27, pp 1456 - 1466, Oct 1979.
- [7] A. Papoulis, " Probability, Random Variables, and Stochastic Processes".
McGraw Hill, 1965

[8] D. Raychaudhuri, "Aloha with Multipacket Messages and ARQ-Type Retransmission Protocols- Throughput Analysis"
IEEE Trans. Communications
Vol COM-32, pp 148 - 154, Feb 1984.

[9] D. Raychaudhuri, "Stability, Throughput and Delay of Asynchronous Selective Reject Aloha "
IEEE Trans. Communications
Vol COM-35, pp 767 - 772, July 1987.

[10] L. Roberts, " Dynamic Allocation of Satellite Capacity Through Packet Reservation"
Proc. Nat. Comp. Conf., NCC '73, pp 711 - 716

[11] M. Schwartz, " Computer Communication Network Design and Analysis",
Prentice-Hall, Englewood Cliffs, N.J., 1977