

# ¿Cómo funcionan los routers?

Area de Ingeniería Telemática  
<http://www.tlm.unavarra.es>

Arquitectura de Redes, Sistemas y Servicios  
3º Ingeniería de Telecomunicación

# Temario

1. Introducción
2. Arquitecturas, protocolos y estándares
3. Conmutación de paquetes
4. Conmutación de circuitos
5. Tecnologías
6. Control de acceso al medio en redes de área local
7. Servicios de Internet

# Temario

1. Introducción
2. Arquitecturas, protocolos y estándares
3. **Conmutación de paquetes**
  - Principios
  - **Problemas básicos**
    - **Como funcionan los routers (Nivel de red)**
    - Encaminamiento (Nivel de red)
    - Transporte fiable (Nivel de transporte en TCP/IP)
    - Control de flujo (Nivel de transporte en TCP/IP)
    - Control de congestión (Nivel de transoporte en TCP/IP)
4. Conmutación de circuitos
5. Tecnologías
6. Control de acceso al medio en redes de área local
7. Servicios de Internet

# En próximas clases...

- Enrutamiento
- Algoritmos y técnicas de enrutamiento
  - ¿Cómo construyo la tabla de rutas?

## Hoy

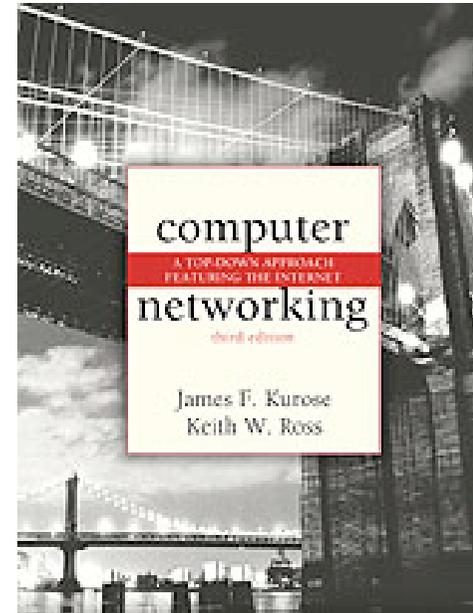
- Como funciona un router
- ¿Qué otros problemas hay que resolver para hacer un router aparte de construir la tabla de rutas?

# Material

- Capítulo 4 de Kurose & Ross, “Computer Networking a top-down approach featuring the Internet” Addison Wesley

- Capítulo 9 de Keshav (poco)

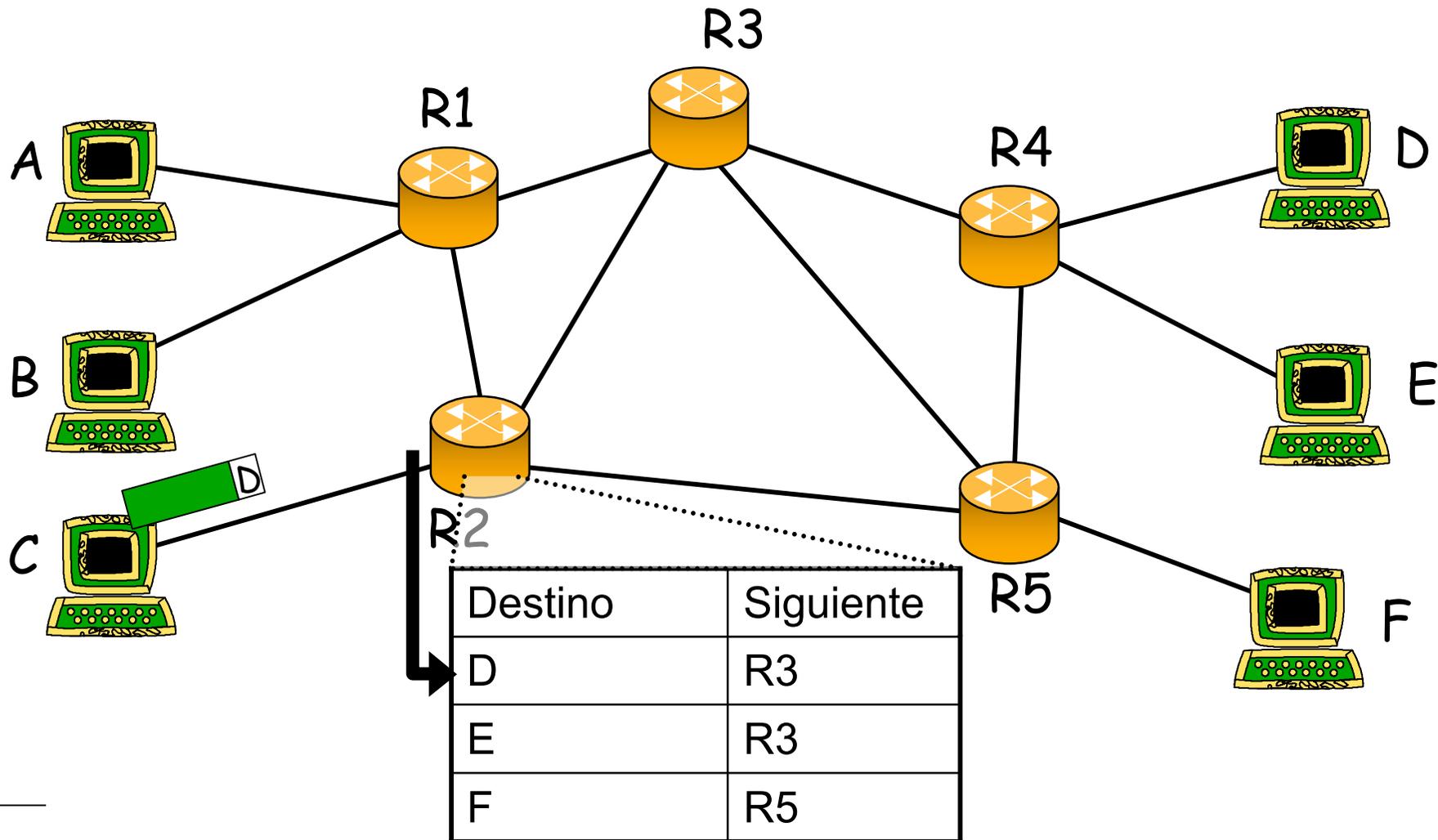
- Presentación basada en transparencias de AsstProf. Bhichate Chiewthanakul basadas a su vez en las de Profs. Nick McKeown and Balaji Prabahakar (Stanford) además de transparencias basadas en el libro de Kurose



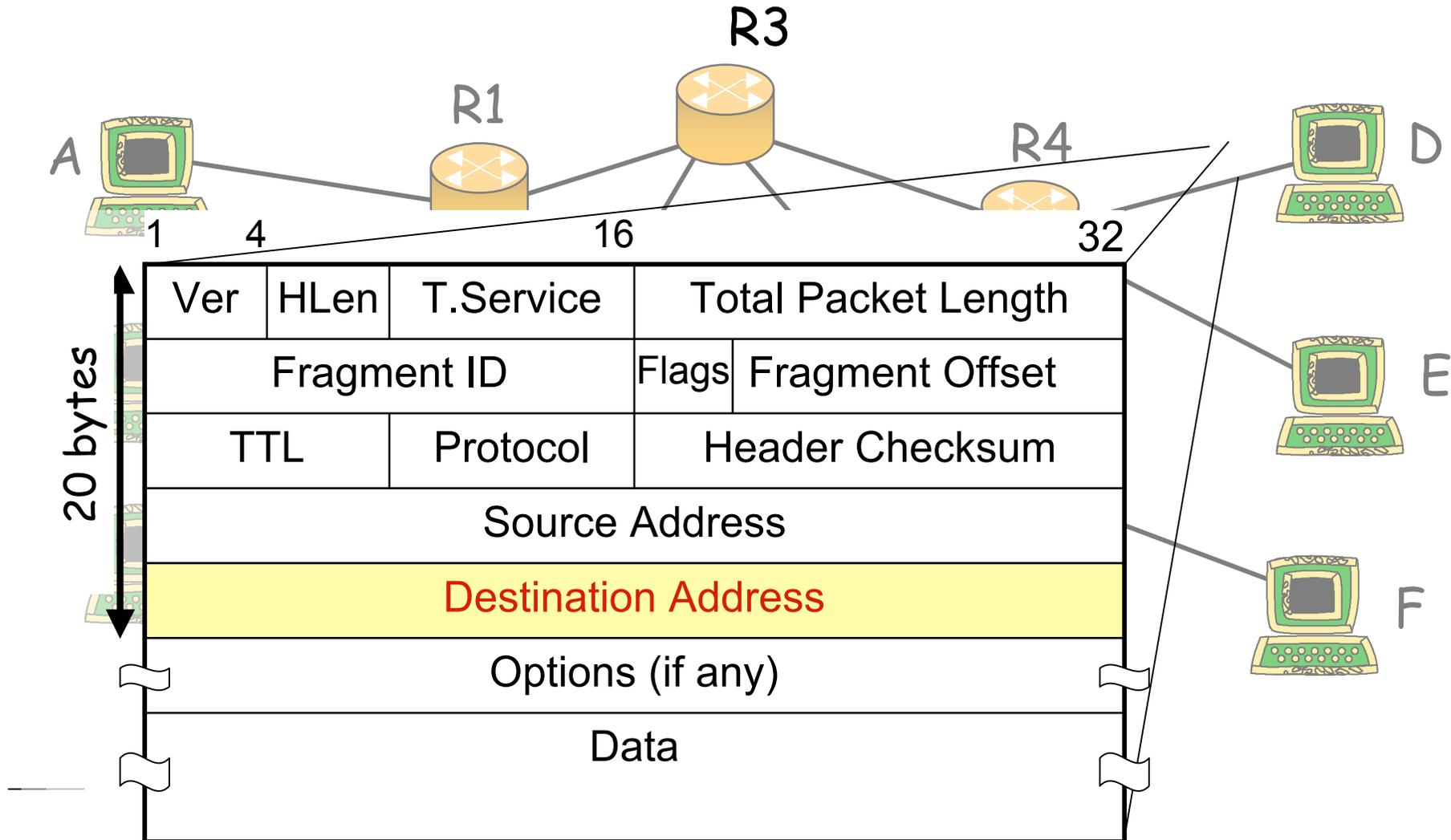
# Contenido

- Background
  - ❖ **Qué es un router?**
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ Búsqueda de dirección IP.
  - ❖ Almacenamiento de paquetes.
  - ❖ Conmutación.
  - ❖ Planificación.

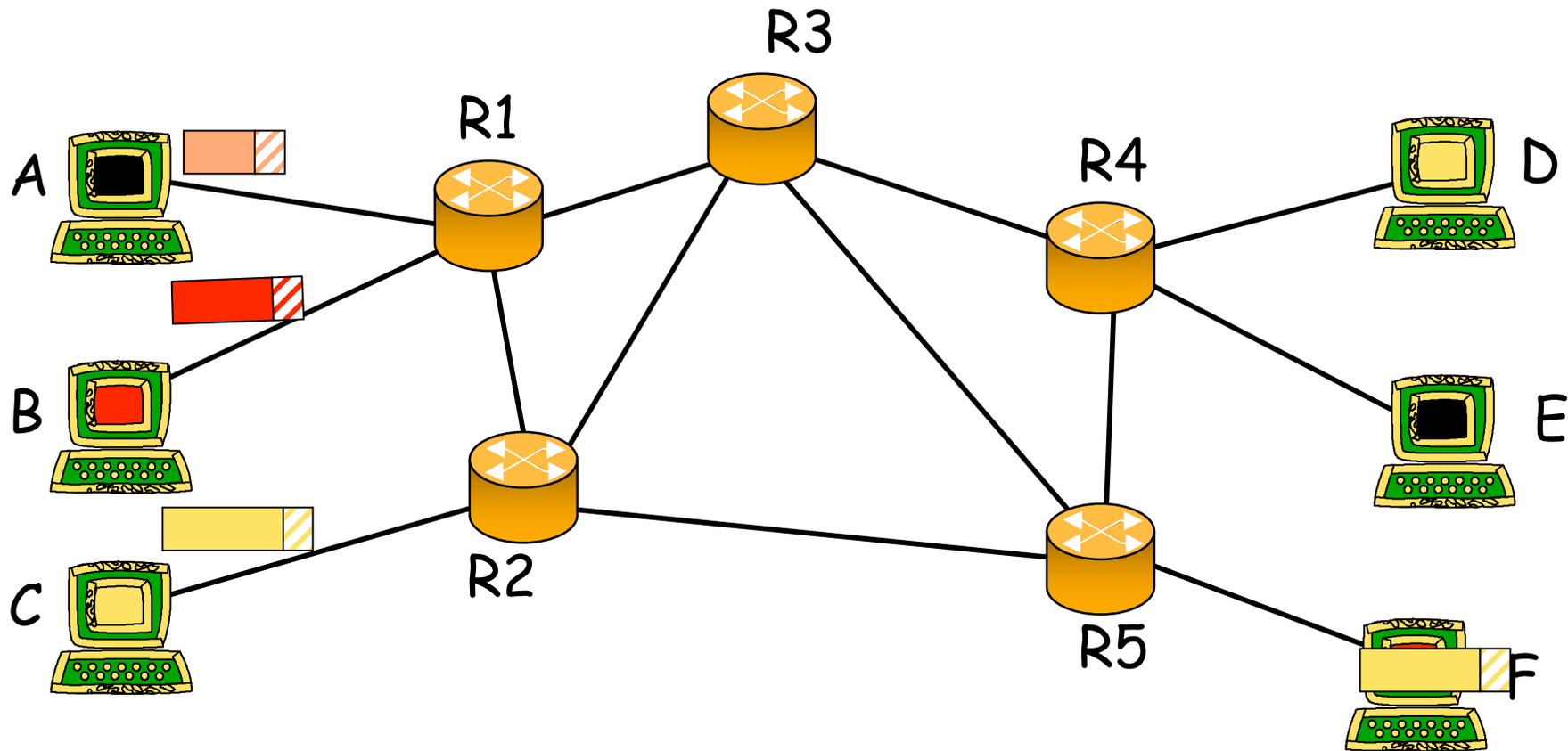
# ¿Qué es un router?



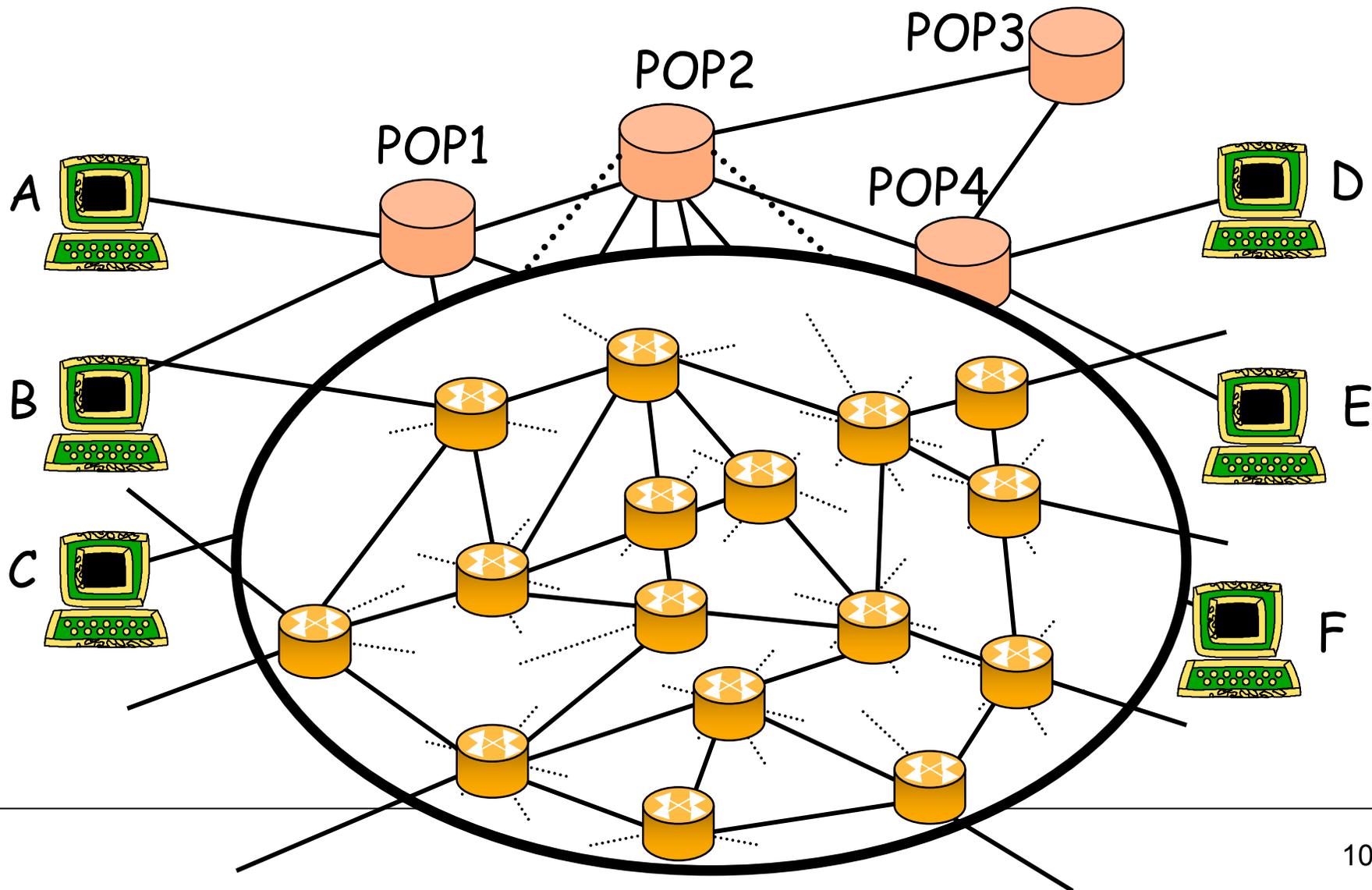
# ¿Qué es un router?



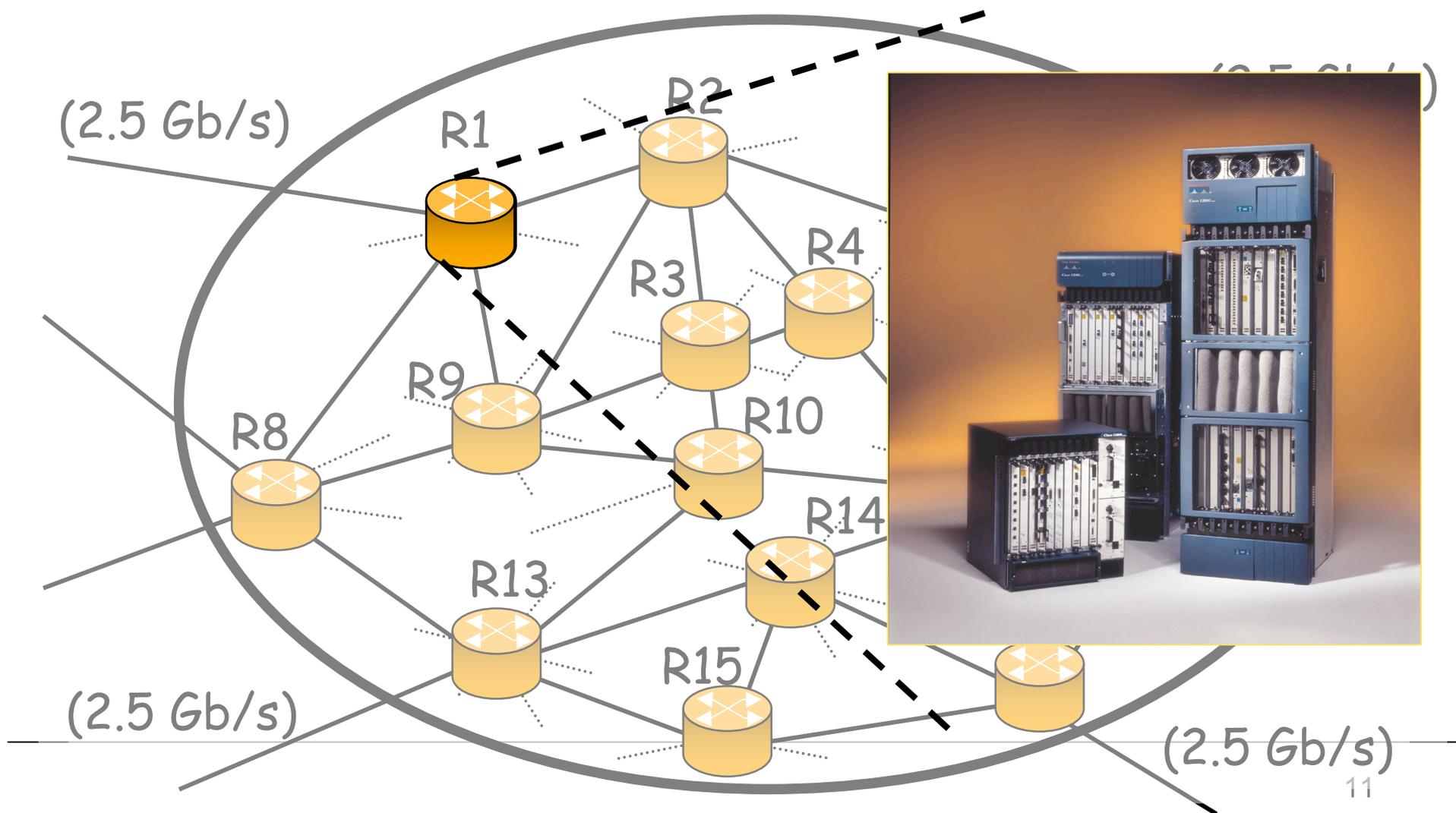
# ¿Qué es un router?



# Puntos de presencia (POPs Points of presence)

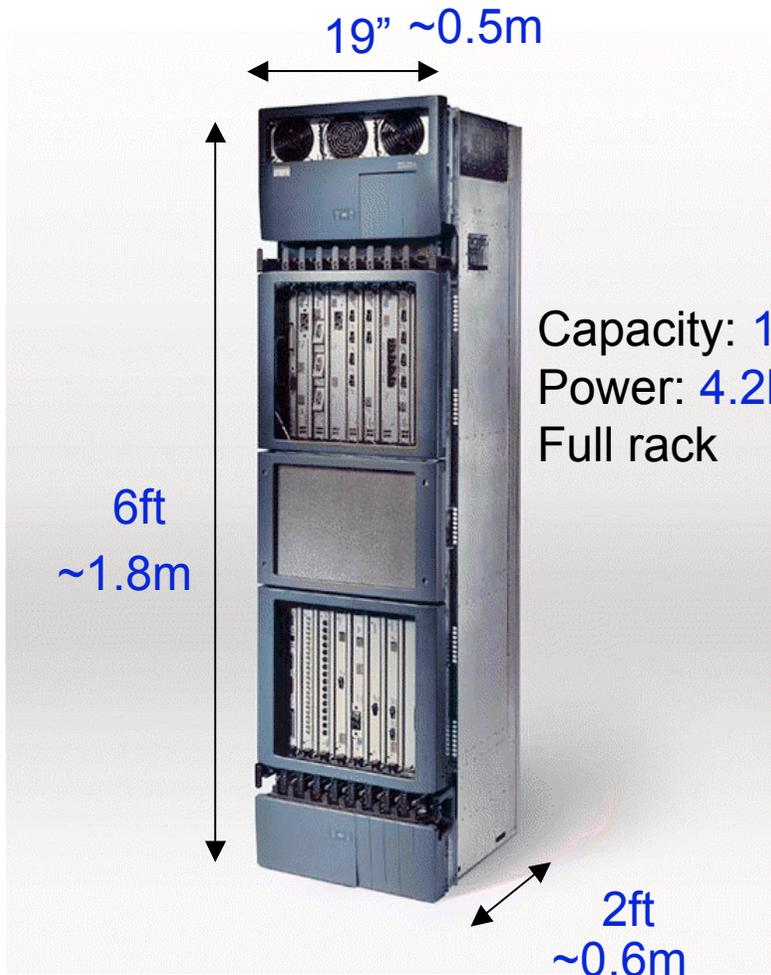


# Donde son necesarios los routers de altas prestaciones

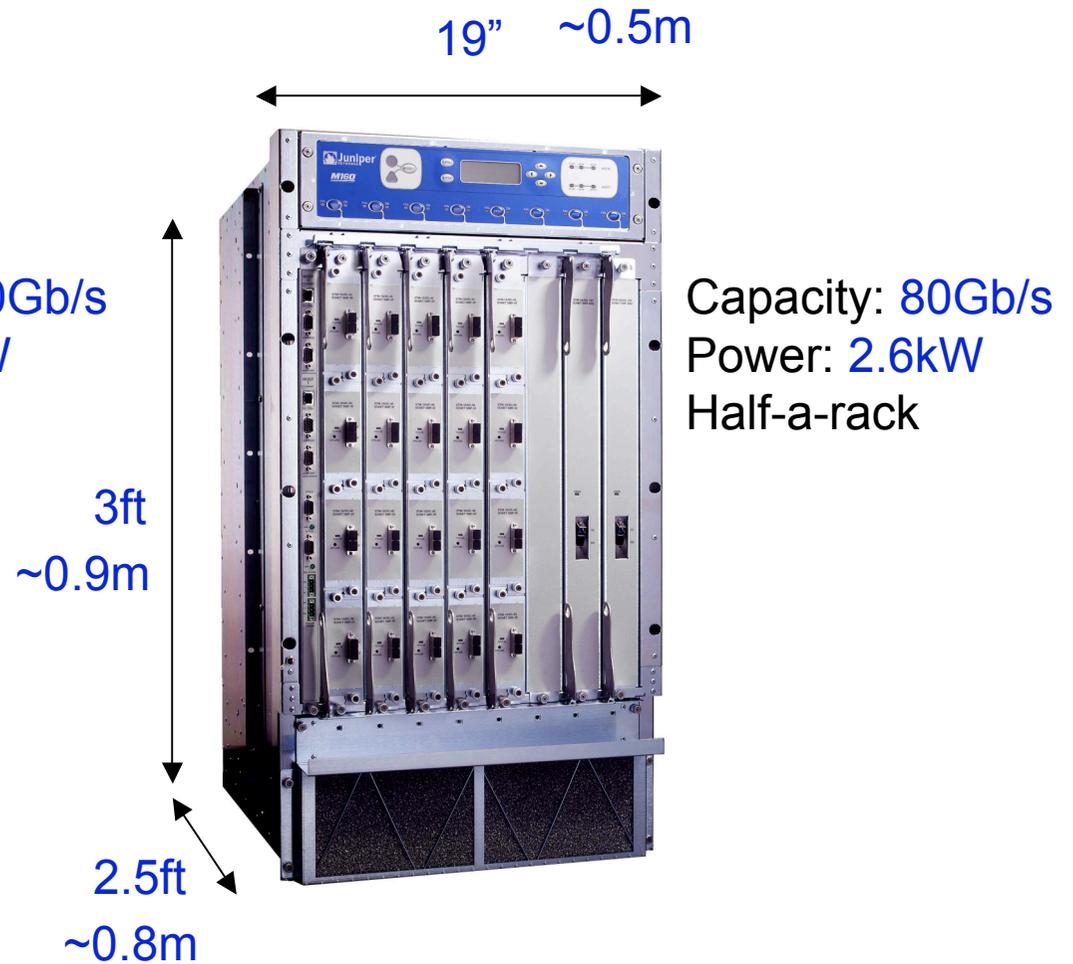


# ¿Qué pinta tiene un router?

Cisco GSR 12416



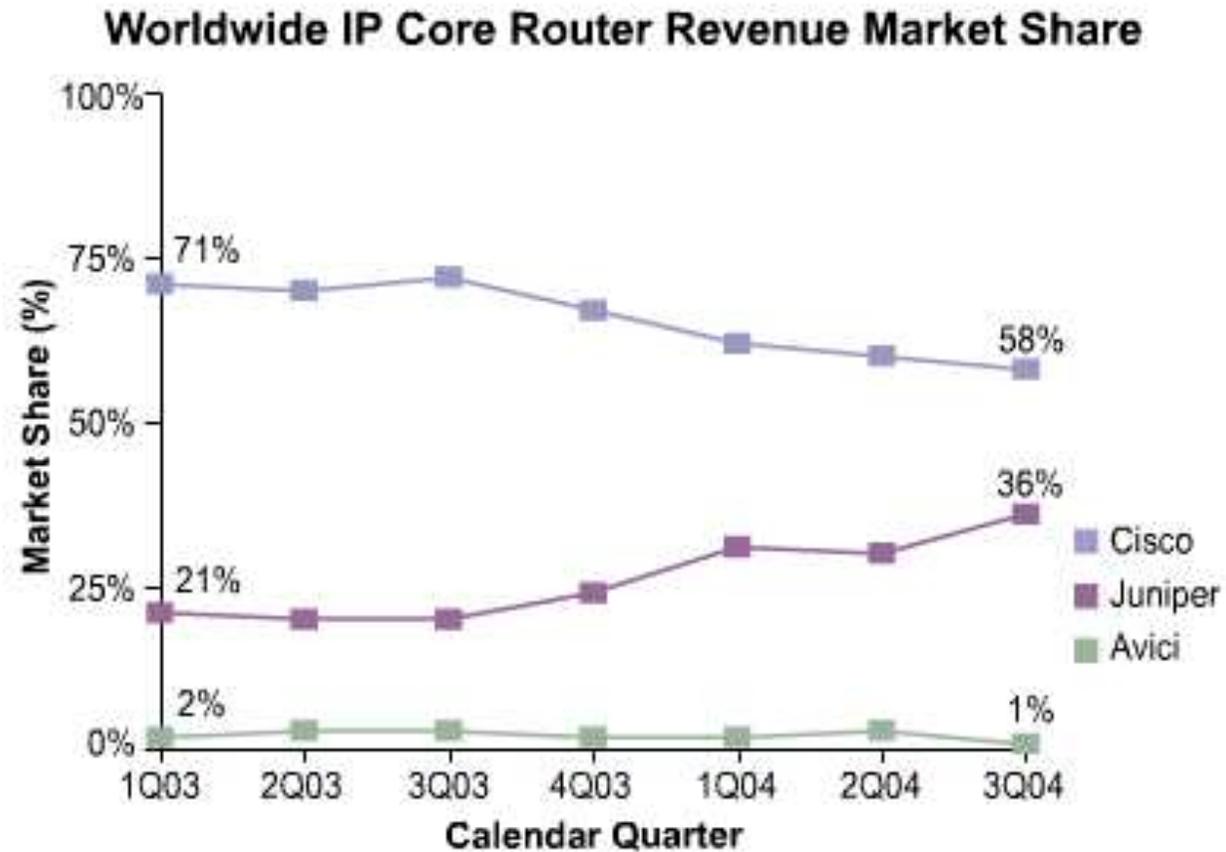
Juniper M160



# El mercado de routers

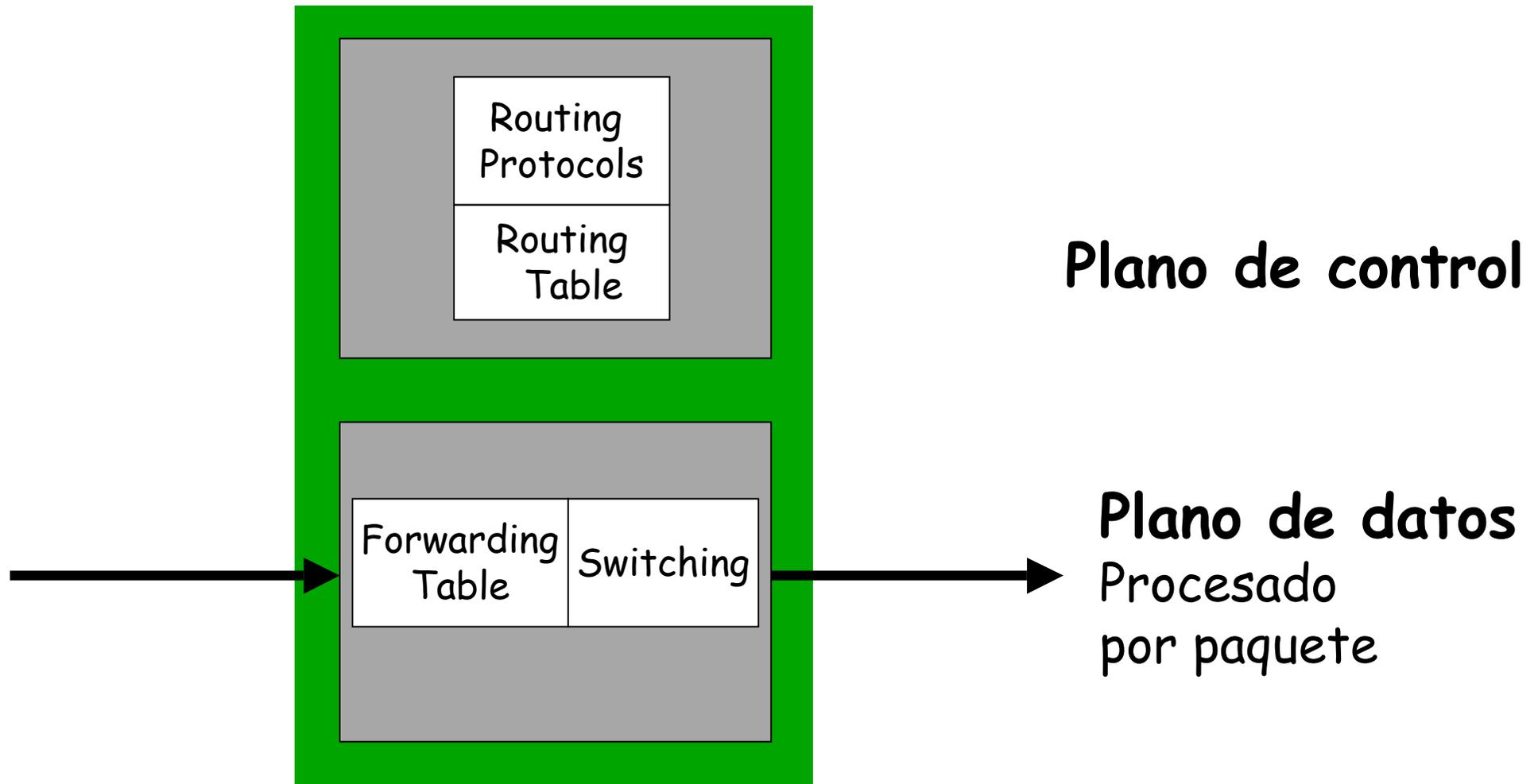
- Según Dell'Oro (Feb. 17, 2005):
  - \$1.2 billion in 2004 (up 66%)
  - Includes high-end (10Gbps) router market
- Según Infonetics (March 2, 2005):
  - \$6.1 billion in 2004 (up 26%)
  - Includes IP core/edge routers and multiservice core/edge switches
- Mercado de routers de núcleo
  - Relativamente pequeño (en comparación con el mercado de routers para empresas), pero da experiencia a las compañías
  - Fabricantes: Cisco, Juniper, Avici, Nortel, Lucent, Alcatel, Chiaro, Huawei, etc.

# Router Market



Source: Infonetics (Nov. 2004)

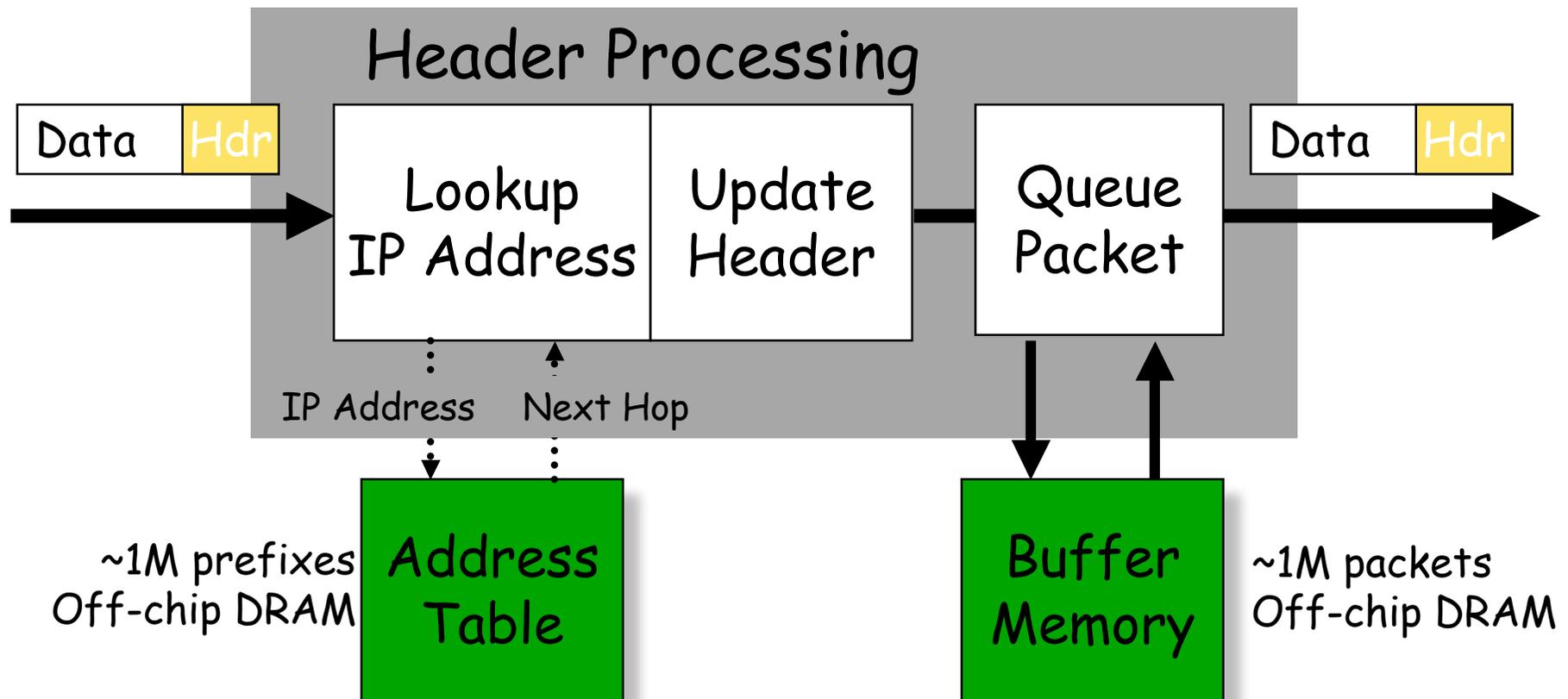
# Arquitectura básica de un router IP



# Proceso por paquete en un Router IP

-  Aceptar paquetes por las líneas de entrada
-  **Lookup:** búsqueda de la dirección de destino del paquete en la tabla de reenvío (para identificar puerto de salida).
-  **Header Processing:** Manipulación de la cabecera IP: decrementar TTL, recalcular checksum.
-  **Switching:** Enviar el paquete al puerto de destino correspondiente.
-  **Buffering:** Almacenar paquete en la cola.
-  Transmitir paquete en la línea de salida.

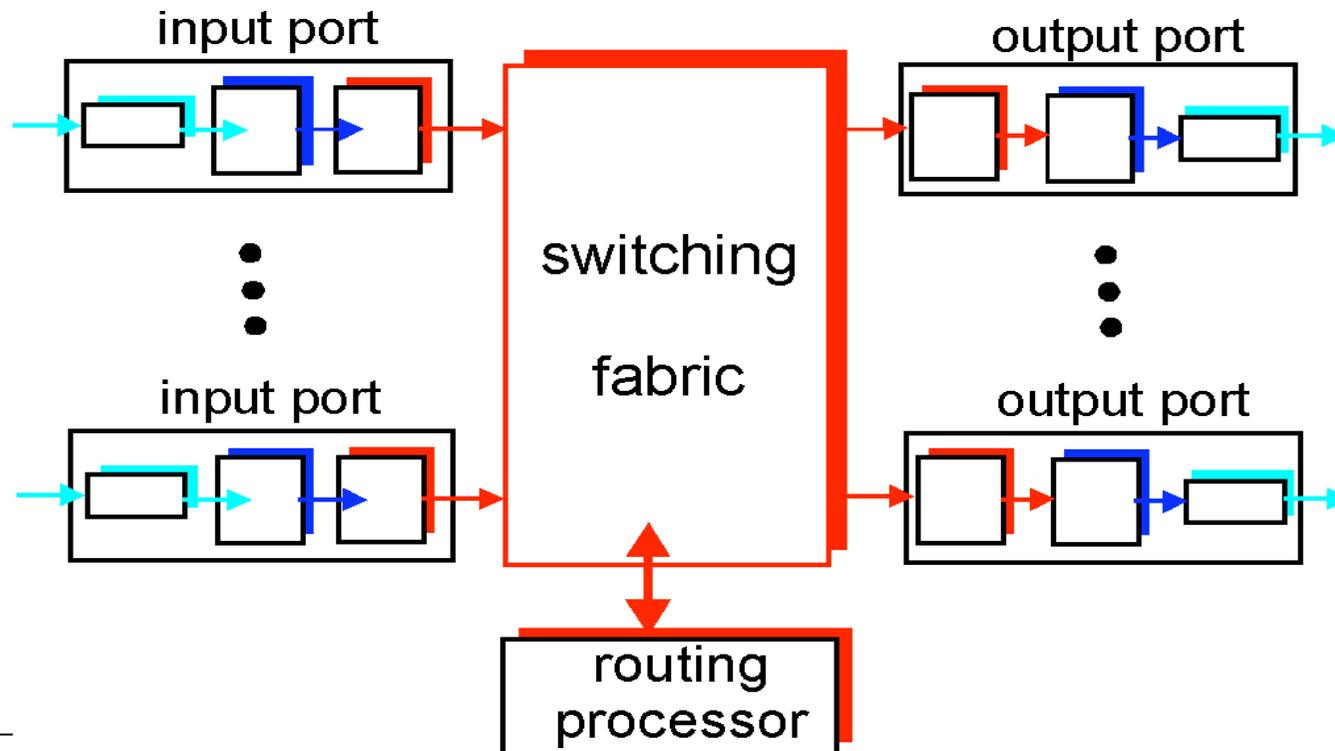
# Arquitectura de un router genérico



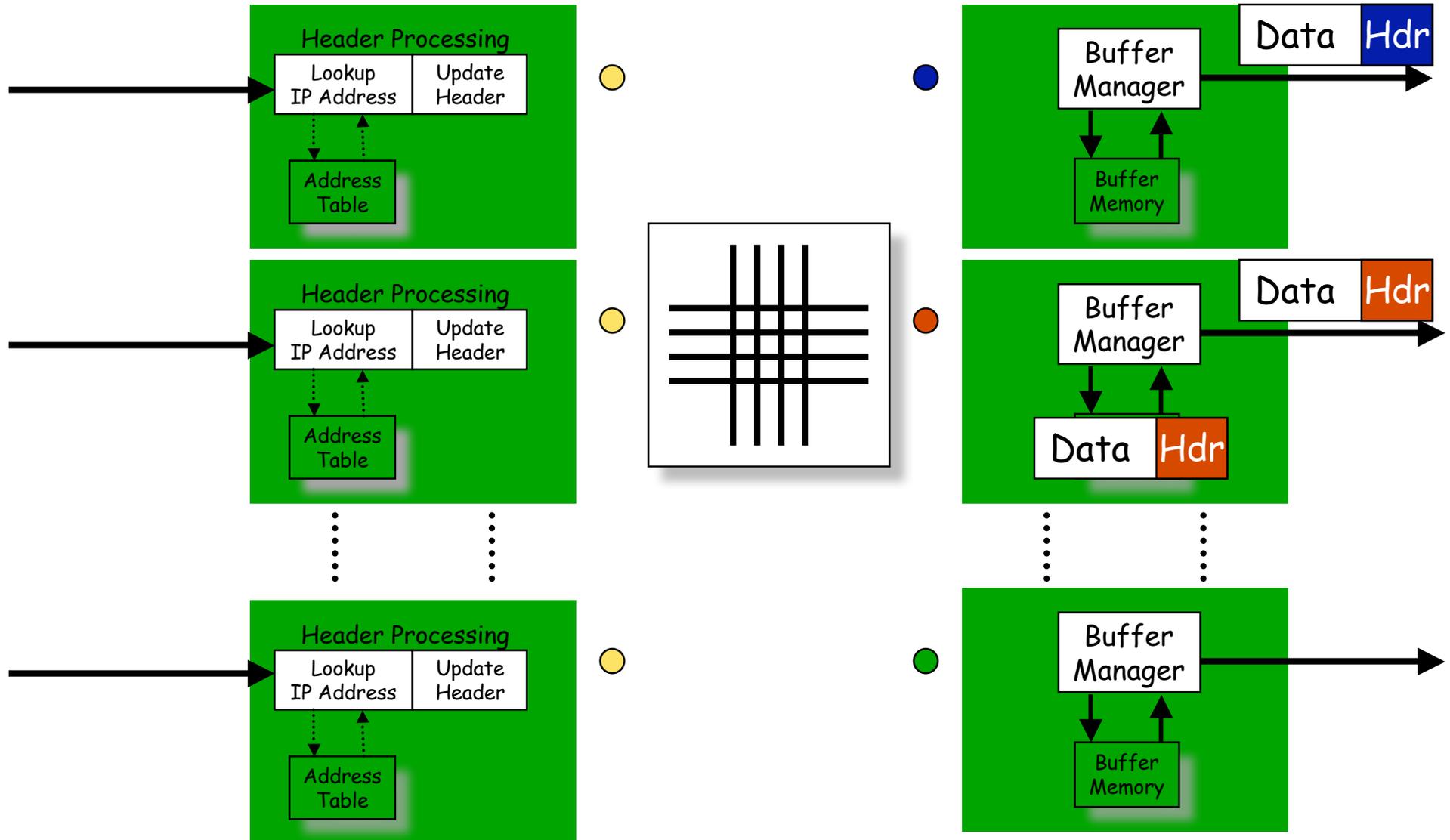
# Arquitectura de routers

Dos funciones básicas de los routers:

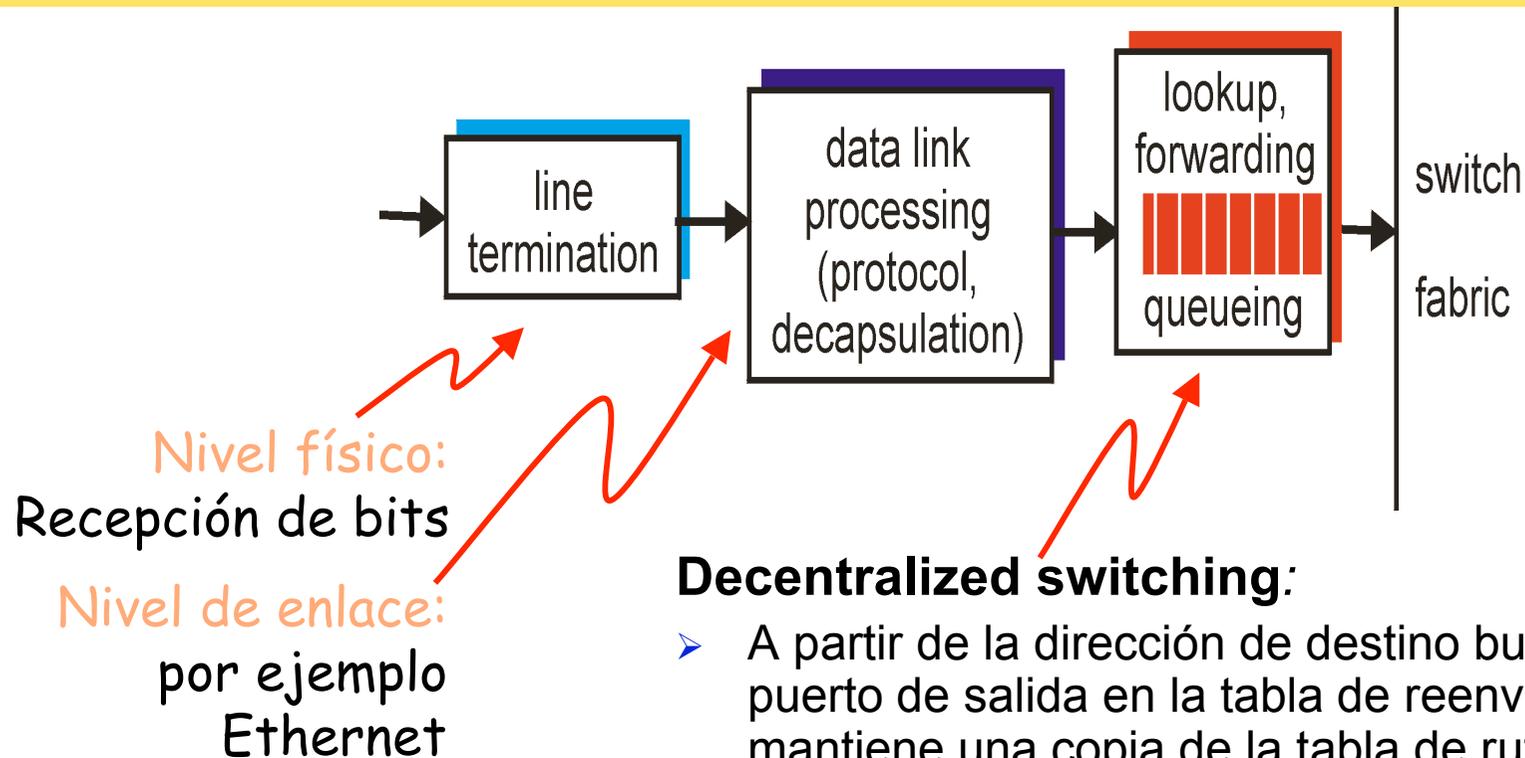
- Correr protocolos/algoritmos de enrutamiento (RIP, OSPF, BGP)
- *Dirigir paquetes de los puertos de entrada a los puertos de salida (forwarding)*



# Arquitectura de routers



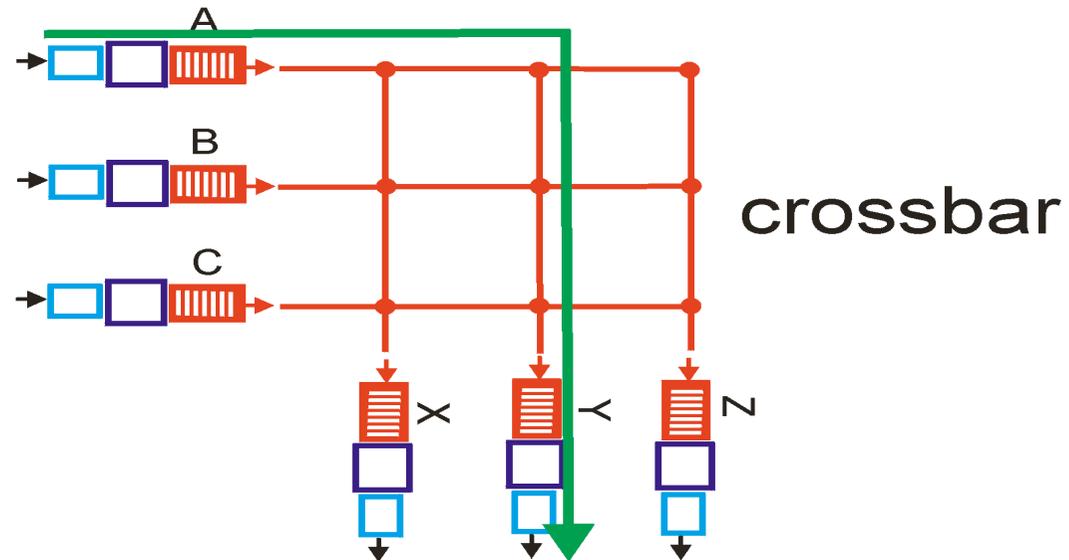
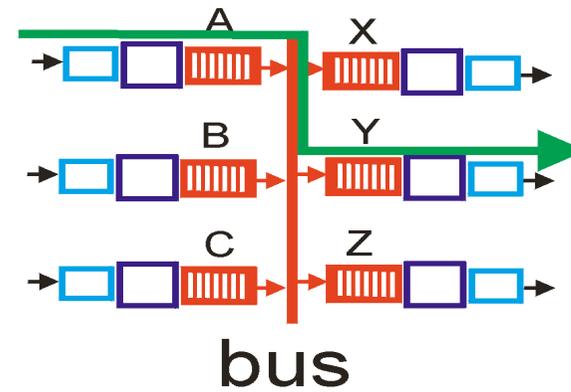
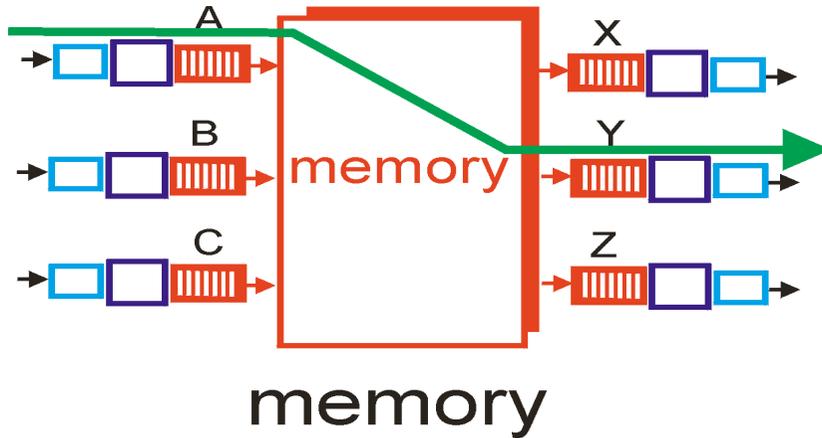
# Puertos de entrada



## Decentralized switching:

- A partir de la dirección de destino buscar el puerto de salida en la tabla de reenvíos (se mantiene una copia de la tabla de rutas en la memoria del puerto de entrada)
- objetivo: procesar paquetes entrantes a velocidad de línea
- Cola de entrada: si los paquetes llegan a mas velocidad que la velocidad de entrada a matriz de conmutación

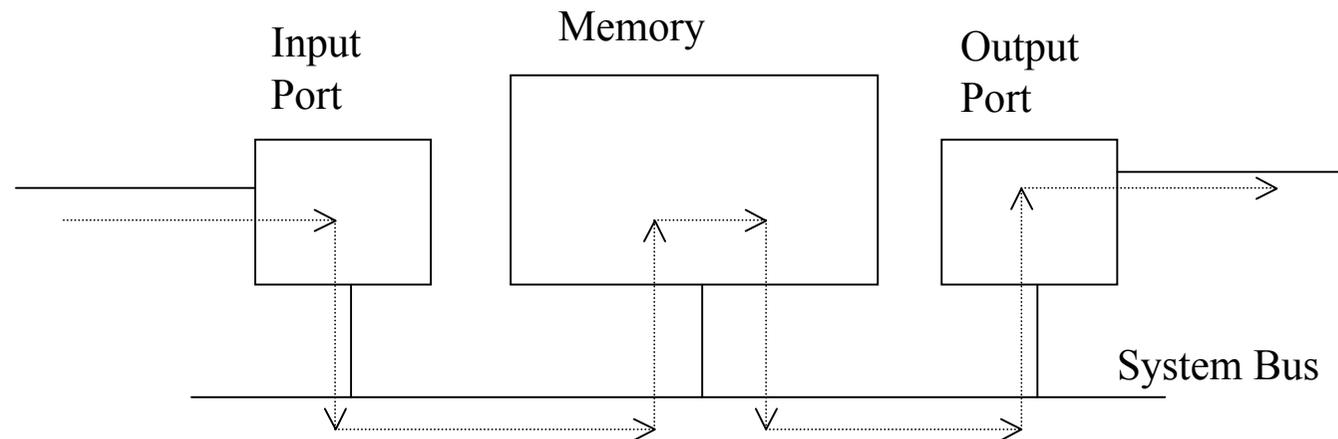
# Tipos de conmutación



# Conmutación por memoria

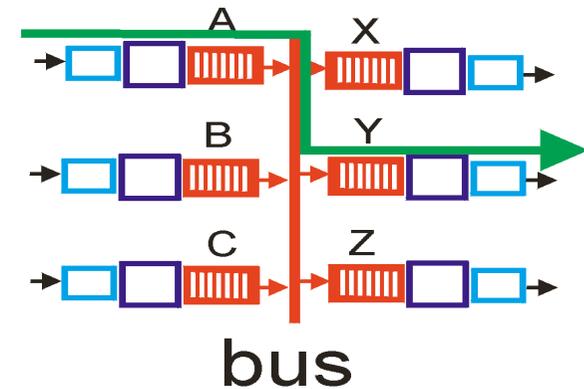
## Primera generación de routers:

- ordenador tradicional con conmutación bajo control de la CPU
- el paquete se copia a la memoria del sistema al recibirlo
- velocidad limitada por el ancho de banda a memoria (2 accesos al bus por paquete)



# Conmutación via Bus

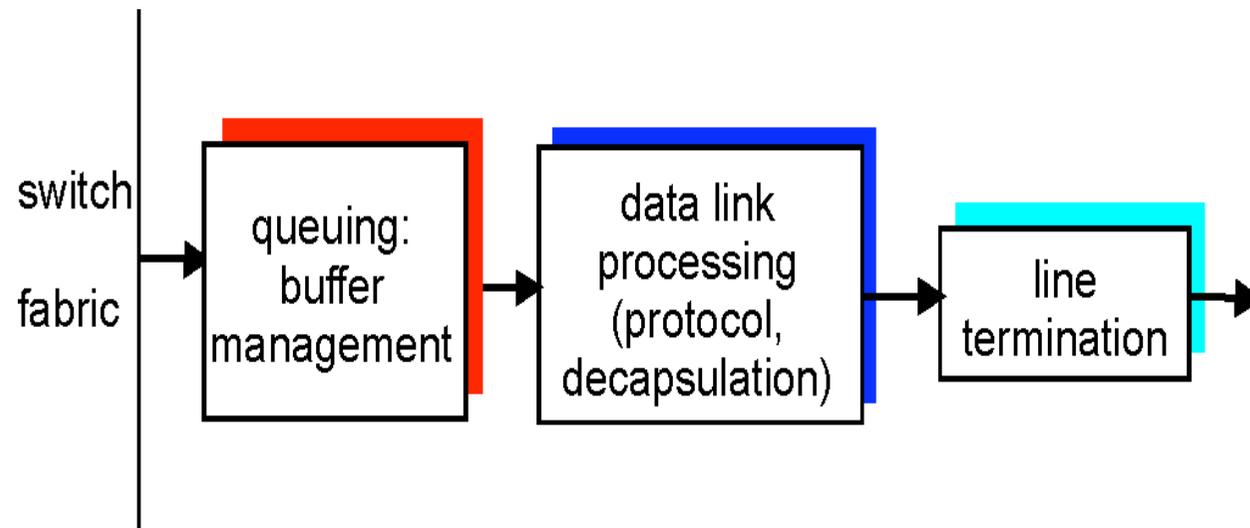
- El paquete se copia de la memoria del puerto de entrada a la memoria del puerto de salida a través de un bus compartido
- **Contención en el bus:**
  - Si el bus está ocupado el paquete espera en la memoria del puerto de entrada (colas a la entrada)
  - Velocidad de conmutación limitada por el bus
- Con bus de 1 Gbps, (Cisco 1900): es suficiente velocidad para accesos y routers de empresa (pero no para regionales o backbone)



# Conmutación via red de interconexión

- Superar las limitaciones de un bus
- Redes Banyan y otras redes de interconexión pensadas en principio para conectar procesadores en maquinas multiprocesador
- Diseño avanzado: fragmentar los paquetes en celdas de longitud fija, y conmutar las celdas por la matriz de conmutación.
- Cisco serie 12000: conmuta a Gbps a traves de la red de interconexión

# Puertos de salida



- *Buffer (cola de salida)* necesario si los paquetes pueden llegar a mas velocidad que la del puerto de salida desde la matriz de conmutación
- *Planificación (Scheduling)* elige entre los paquetes disponible para transmision (FIFO, u otras?)

# Contenido

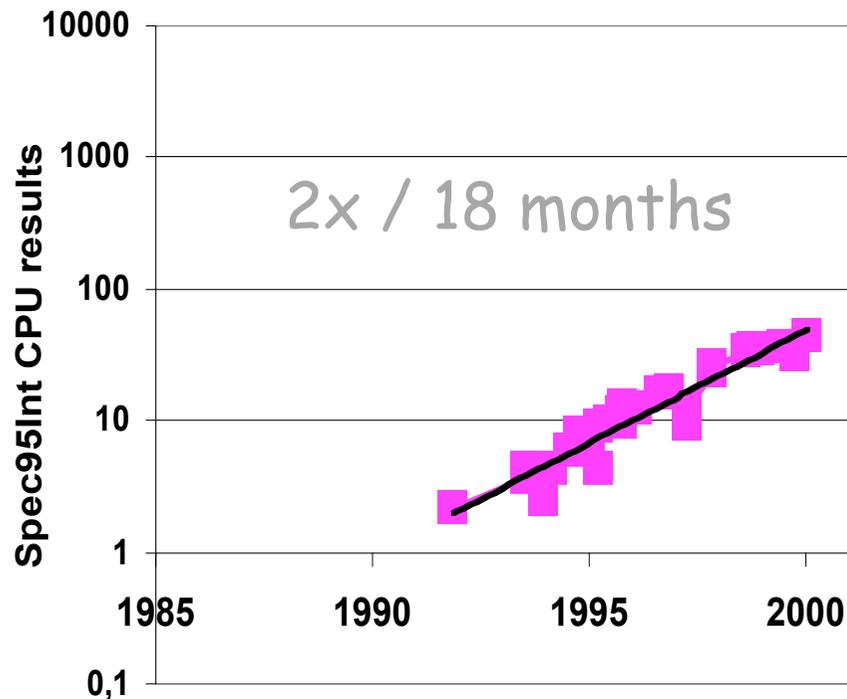
- Background
  - ❖ Qué es un router?
  - ❖ **Por qué necesitamos routers más rápidos?**
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ Búsqueda de dirección IP.
  - ❖ Almacenamiento de paquetes.
  - ❖ Conmutación.
  - ❖ Planificación.

# Por qué necesitamos routers más rápidos?

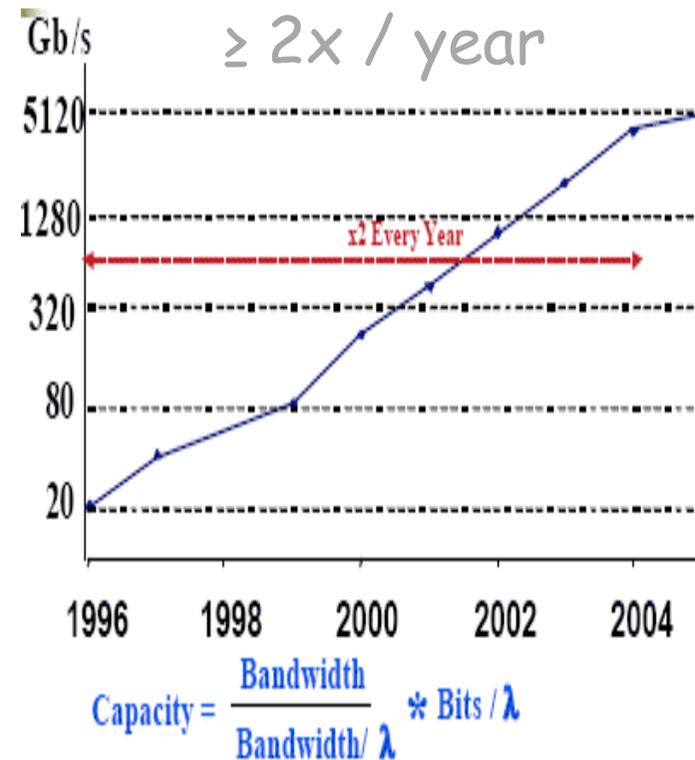
1. Para evitar que los routers se conviertan en el cuello de botella de Internet
2. Para aumentar la capacidad de los puntos de presencia (POP) y reducir su coste, tamaño y consumo

# Por qué necesitamos routers más rápidos? Evitar que los routers sean el cuello de botella

## Packet Processing Power



## Single Fiber Capacity (commercial)

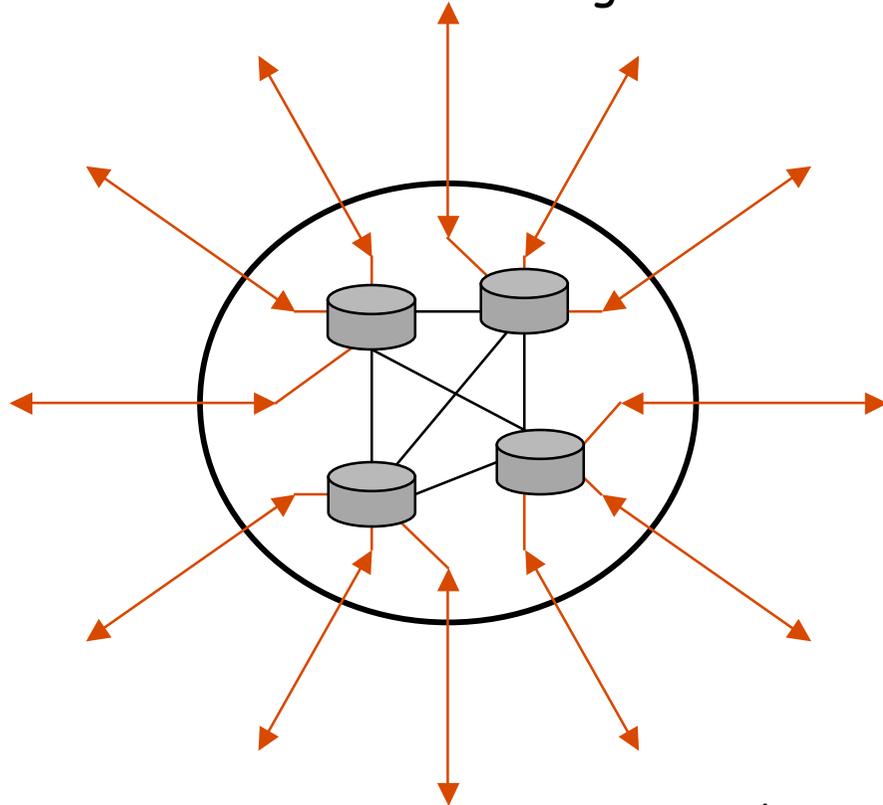


Source: SPEC95Int & Coffman and Odlyzko.

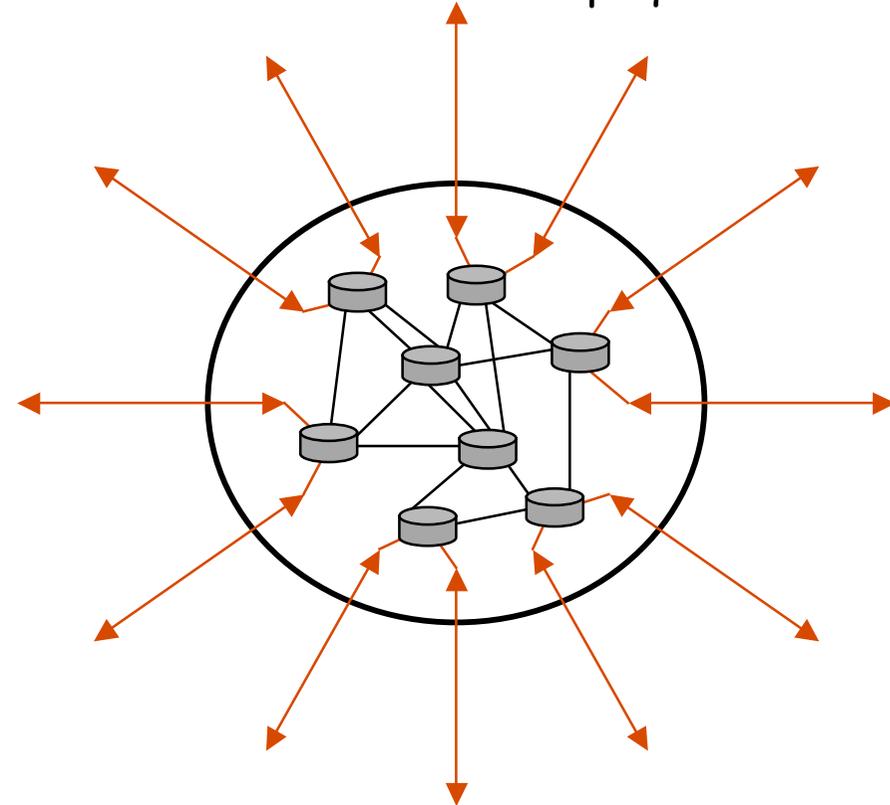
# Por qué necesitamos routers más rápidos?

## 2: Reducir coste, consumo y complejidad de POPs

POP con routers grandes



POP con routers pequeños



- ❖ Puertos: precio >\$50k, consumo > 400W.
- ❖ Alrededor de 50-60% de los puertos es para interconexión.

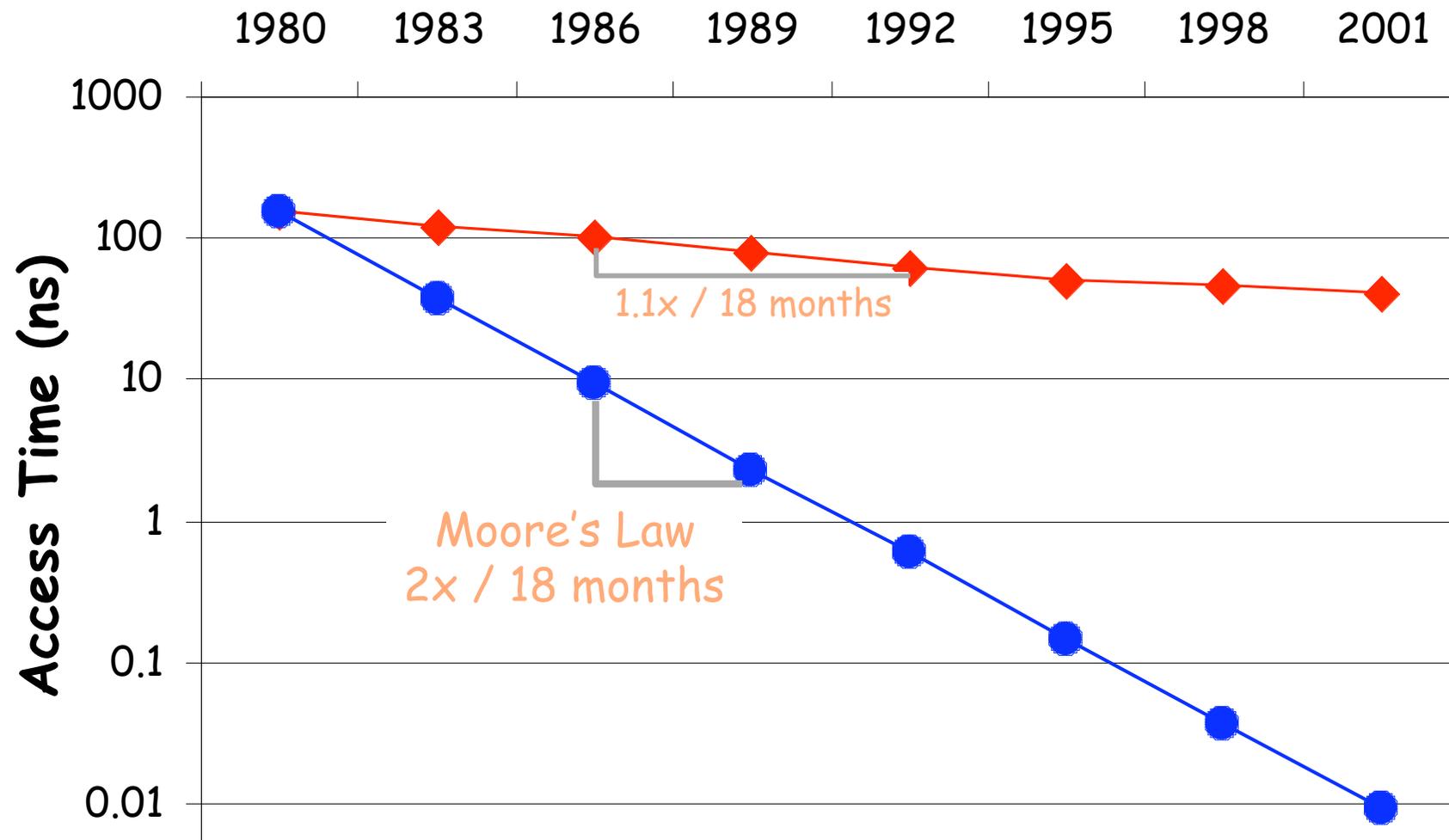
# ¿Por qué es difícil hacer routers rápidos?

## 1. **Es difícil seguir la ley de Moore:**

- El cuello de botella es la velocidad de la memoria
- La velocidad de la memoria no sigue la ley de Moore

# ¿Por qué es difícil hacer routers rápidos?

## *Velocidad de la DRAM comercial*



# ¿Por qué es difícil hacer routers rápidos?

## 1. Es difícil seguir la ley de Moore:

- El cuello de botella es la velocidad de la memoria
- La velocidad de la memoria no sigue la ley de Moore

## 2. La ley de Moore no es suficiente:

- Los routers necesitan mejorar más rápido que la ley de Moore

# Prestaciones de routers y la ley de Moore

Crecimiento de la capacidad de routers comerciales:

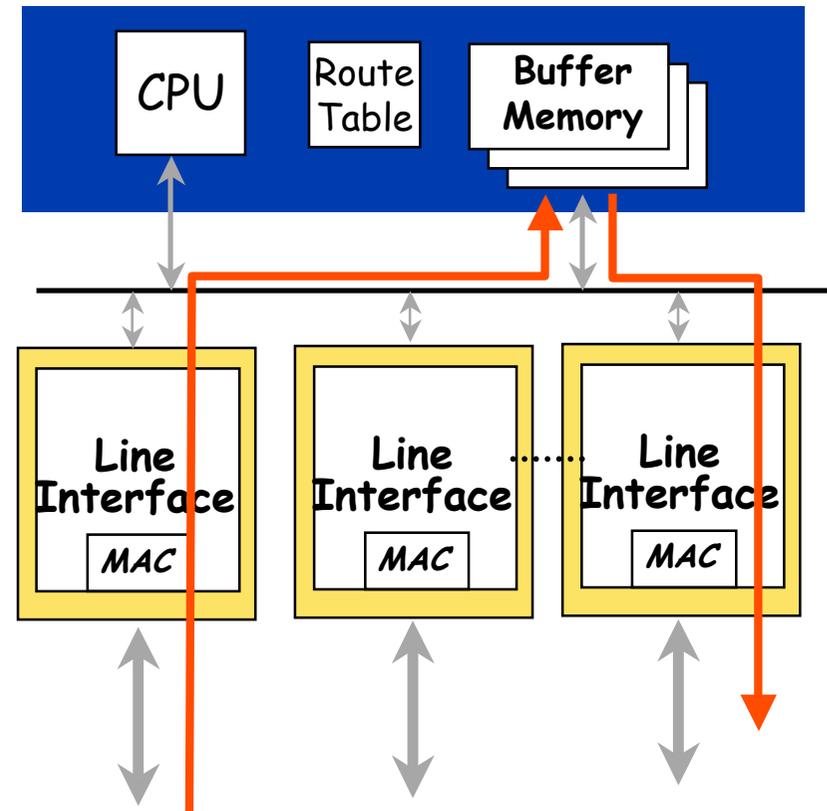
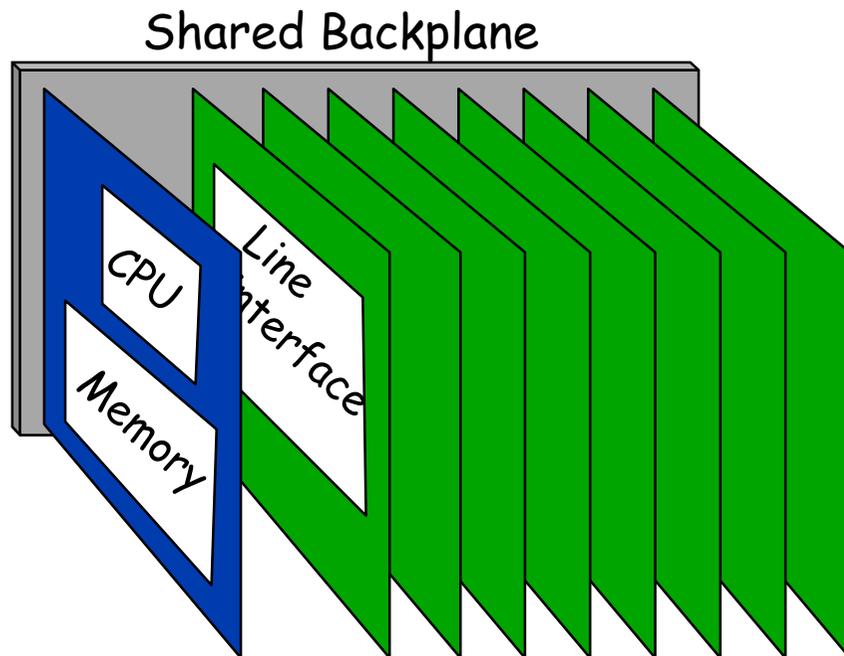
- Capacidad 1992 ~ 2Gb/s
- Capacidad 1995 ~ 10Gb/s
- Capacidad 1998 ~ 40Gb/s
- Capacidad 2001 ~ 160Gb/s
- Capacidad 2003 ~ 640Gb/s

Crecimiento medio: 2.2x / 18 months.

# Contenido

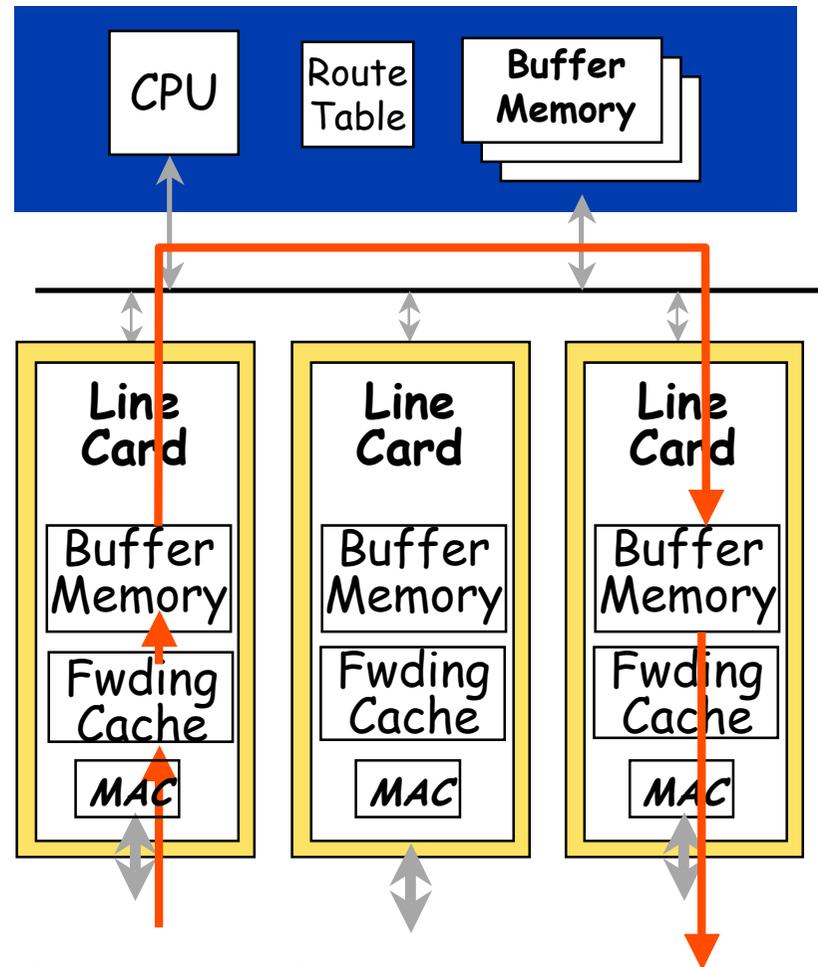
- Background
  - ❖ Qué es un router?
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ **Evolución de arquitecturas de routers.**
  - ❖ Búsqueda de dirección IP.
  - ❖ Almacenamiento de paquetes.
  - ❖ Conmutación.
  - ❖ Planificación.

# Routers de primera generación



Típico  $<0.5\text{Gb/s}$  capacidad agregada

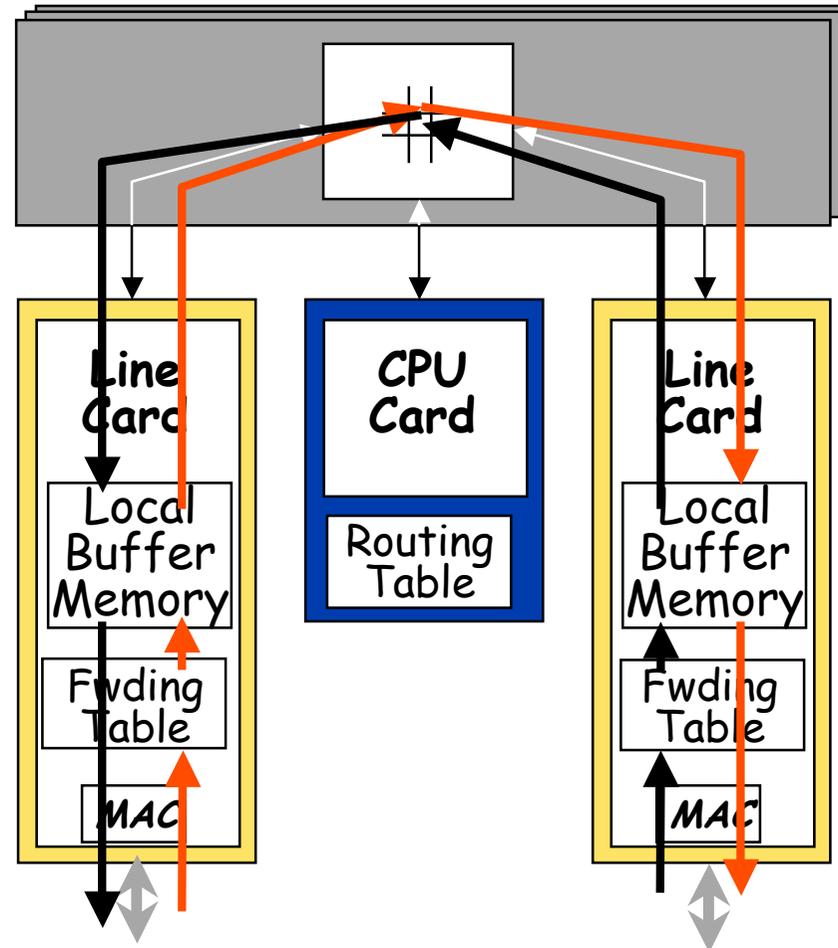
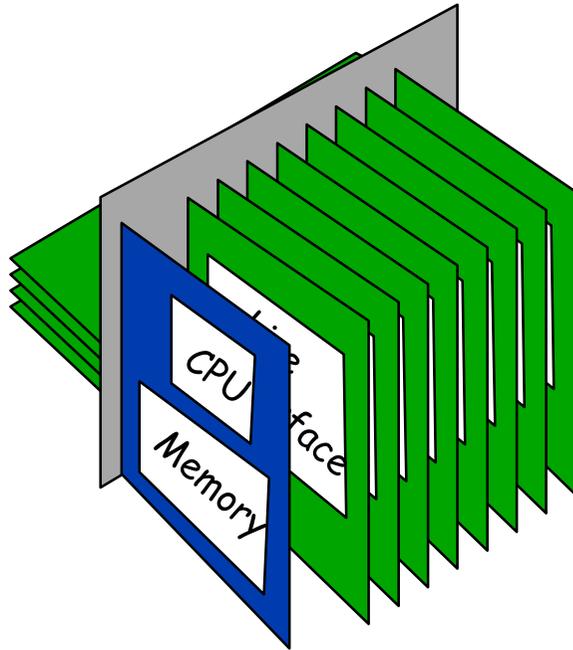
# Routers de segunda generación



Típico <5Gb/s capacidad agregada

# Routers de tercera generación

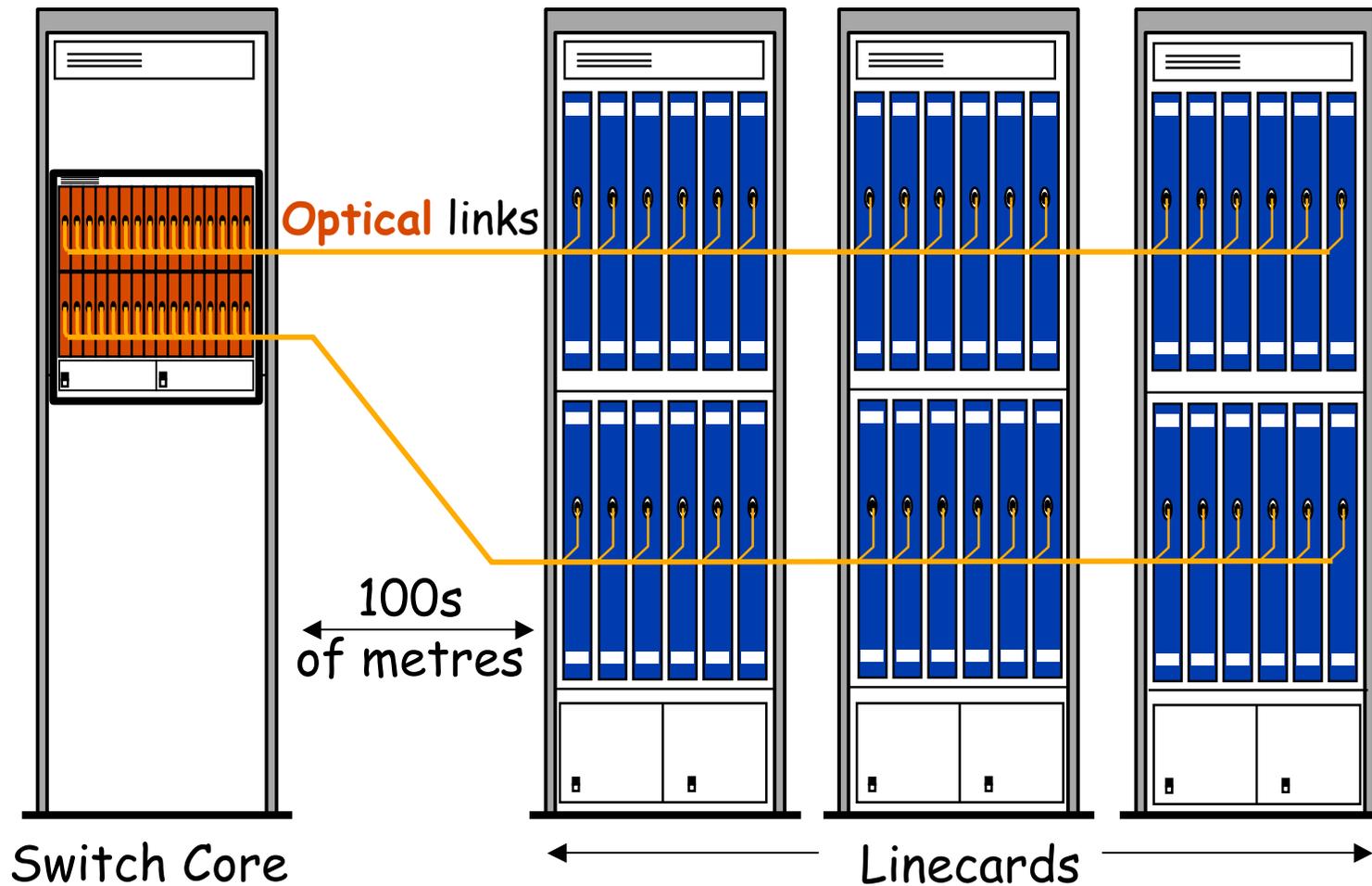
Backplane conmutado  
(Switched Backplane)



Típico <50Gb/s capacidad agregada

# Routers de cuarta generación

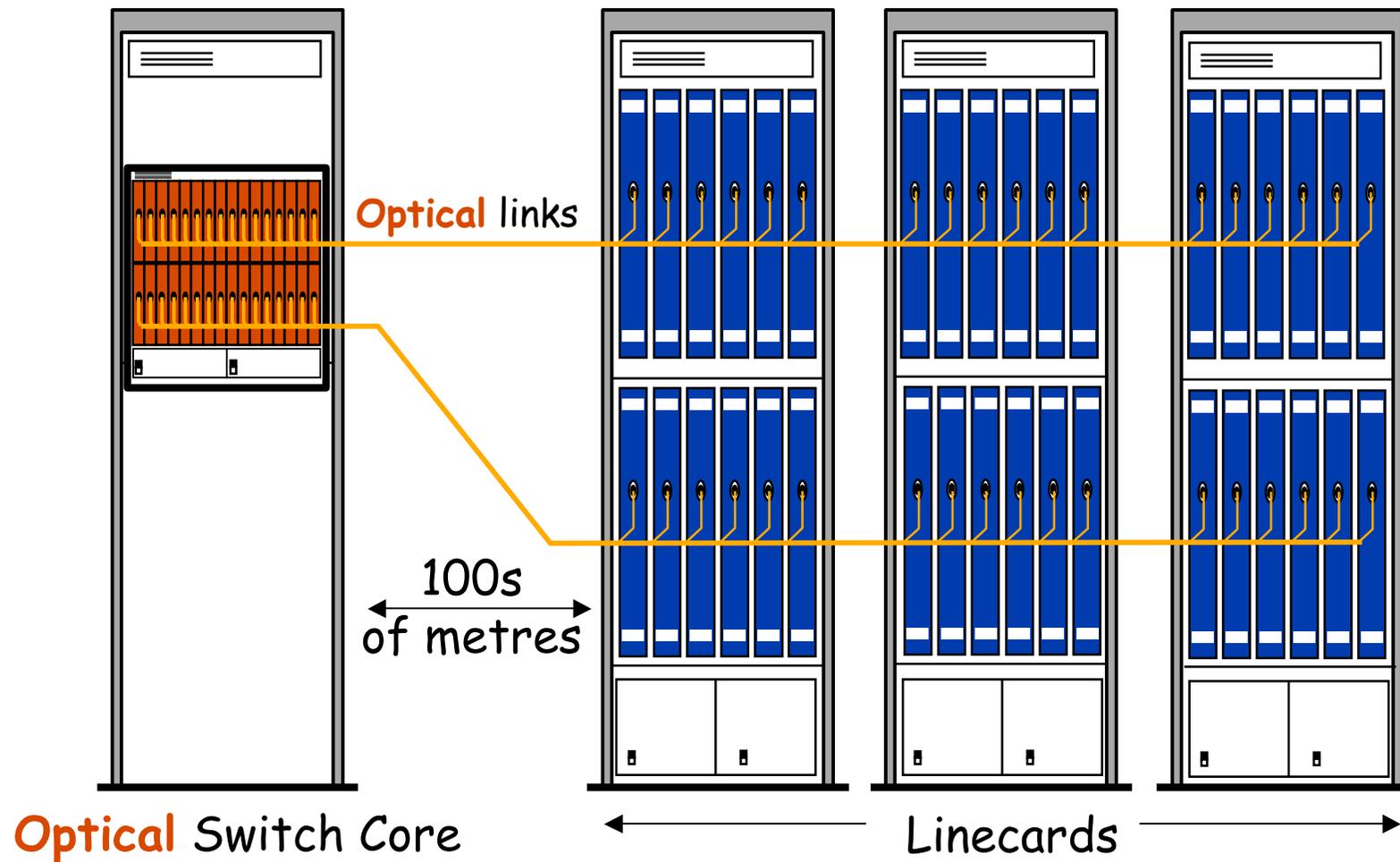
## *Multi-Racks, Optical Links*



Capacidades 0.3 - 10Tb/s

# (Futuros) Routers de quinta generación

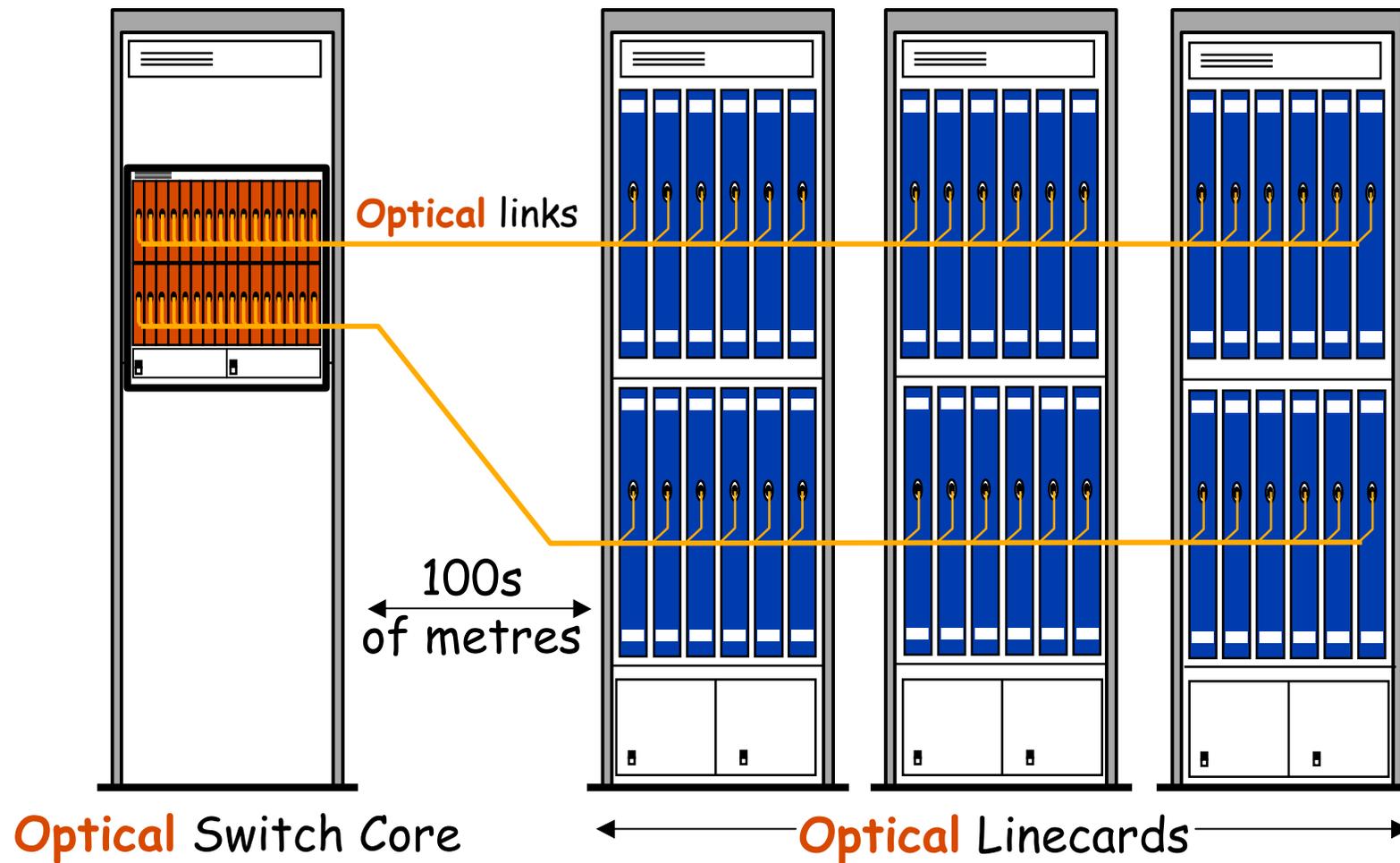
## *Optical Switch Core*



Routers de 10-100Tb/s en proyecto

# (Muy futuros) Routers de sexta generación

## *All-Optical Routers*

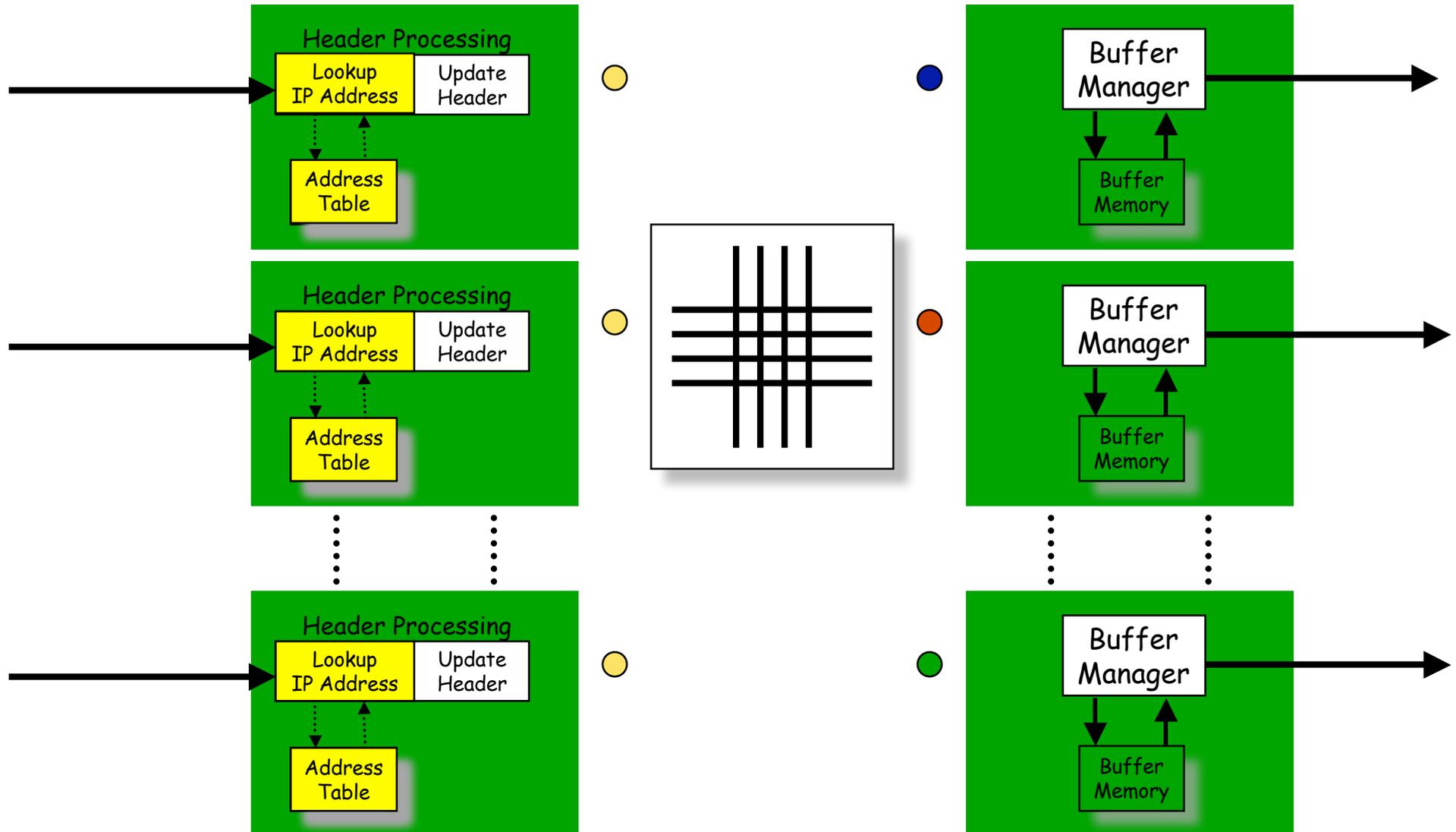


Routers de 100-1000Tb/s en un futuro lejano

# Contenido

- Background
  - ❖ Qué es un router?
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ **Búsqueda de dirección IP.**
  - ❖ Almacenamiento de paquetes.
  - ❖ Conmutación.
  - ❖ Planificación.

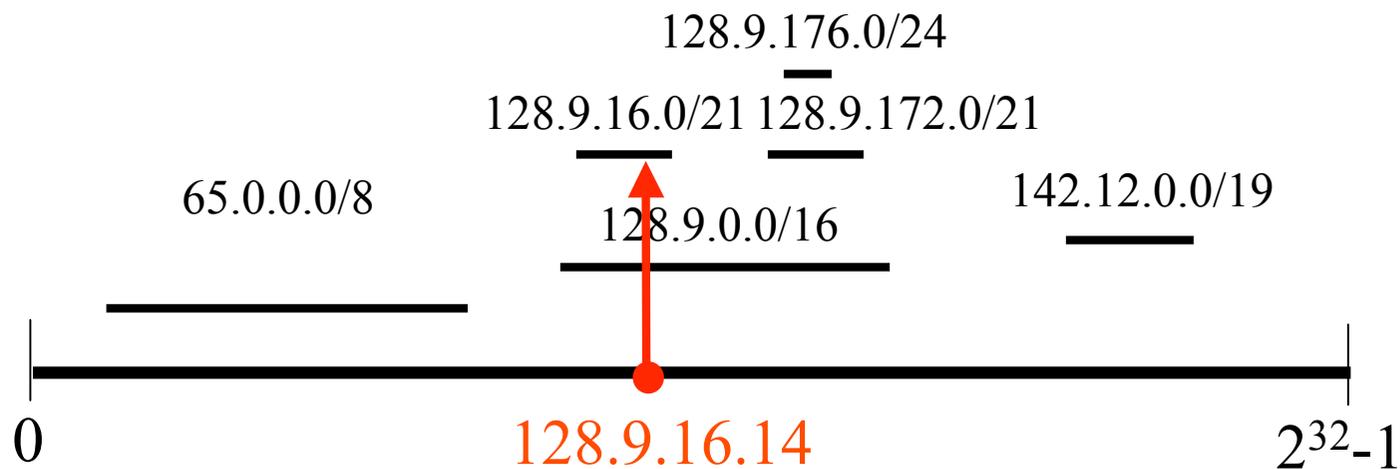
# Arquitectura de router



# Búsqueda de la dirección (Address lookup)

- ¿Por qué es difícil?:
  - ❖ No es una búsqueda exacta sino la búsqueda del prefijo más largo (longest prefix match).
  - ❖ La tabla es muy grande: unas 150,000 entradas y creciendo.
  - ❖ La búsqueda debe ser muy rápida: unos 30ns para una línea de 10Gb/s.

# IP Lookup con Longest Prefix Match

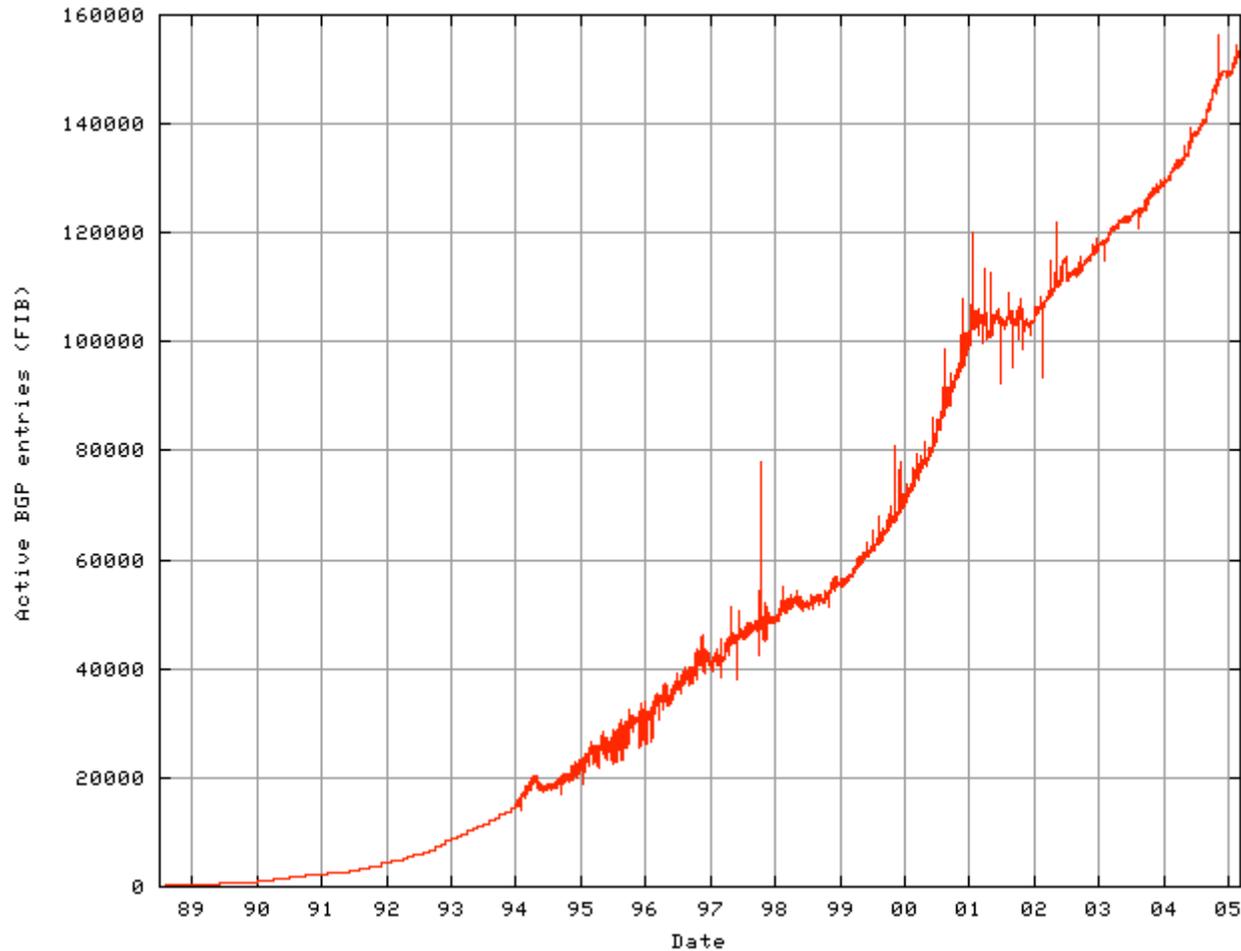


**Routing lookup:** Encuentra el prefijo más largo que coincida (o sea la ruta más específica) entre todos los prefijos que verifique esta dirección IP de destino

# Búsqueda de la dirección (Address lookup)

- ¿Por qué es difícil?:
  - ❖ No es una búsqueda exacta sino la búsqueda del prefijo más largo (longest prefix match).
  - ❖ La tabla es muy grande: unas 150,000 entradas y creciendo.
  - ❖ La búsqueda debe ser muy rápida: unos 30ns para una línea de 10Gb/s.

# Las tablas de rutas son grandes



Source: <http://www.cidr-report.org/>

# Búsqueda de la dirección (Address lookup)

- ¿Por qué es difícil?:
  - ❖ No es una búsqueda exacta sino la búsqueda del prefijo más largo (longest prefix match).
  - ❖ La tabla es muy grande: unas 150,000 entradas y creciendo.
  - ❖ La búsqueda debe ser muy rápida: unos 30ns para una línea de 10Gb/s.

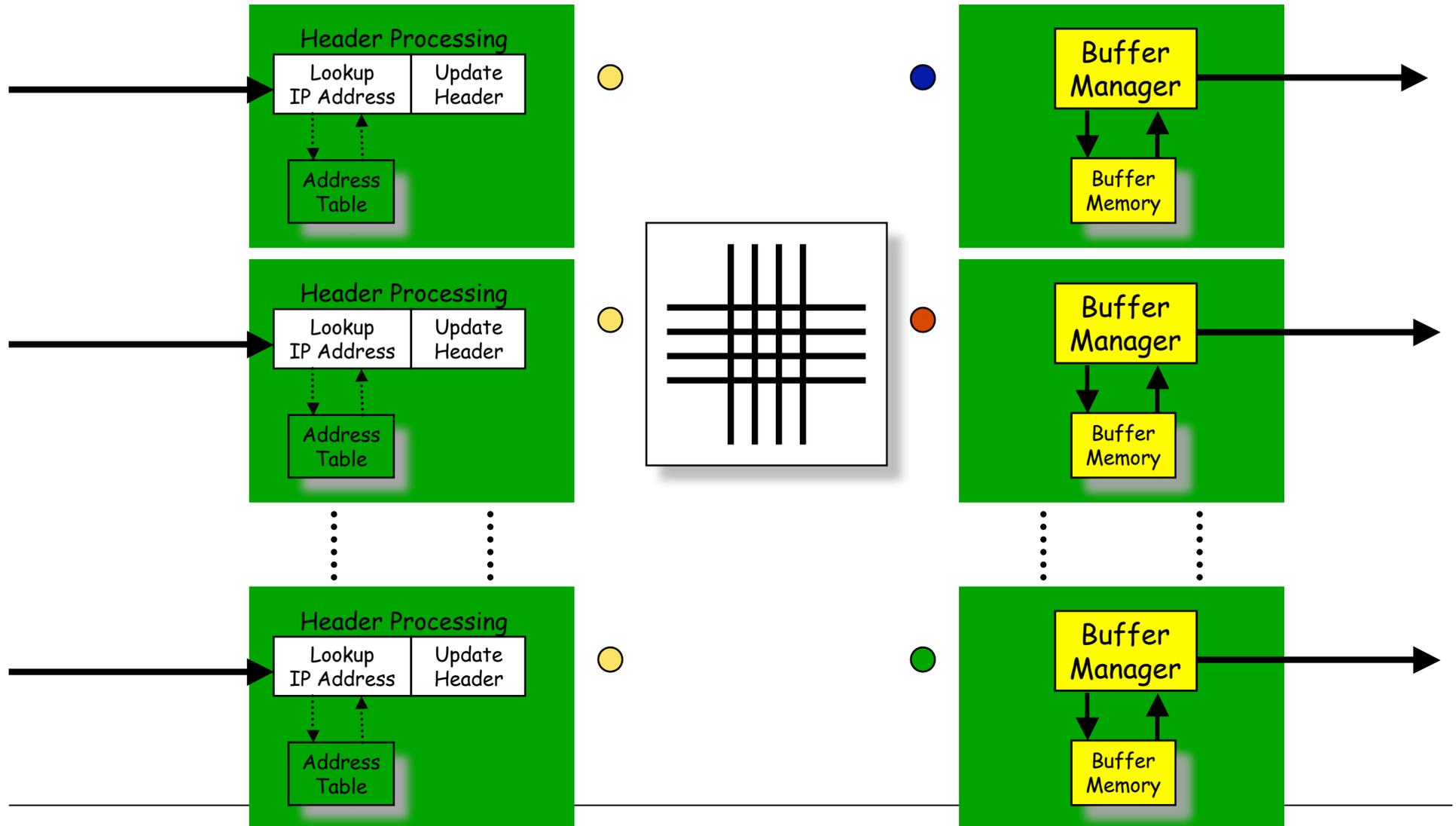
# La búsqueda debe ser muy rápida

Año	Línea típica	Paquetes de 40B (Mpkt/s)
1997	622Mb/s	1.94
1999	2.5Gb/s	7.81
2001	10Gb/s	31.25
2003	40Gb/s	125
2005?	160Gb/s	500

# Contenido

- Background
  - ❖ Qué es un router?
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ Búsqueda de dirección IP.
  - ❖ **Almacenamiento de paquetes.**
  - ❖ Conmutación.
  - ❖ Planificación.

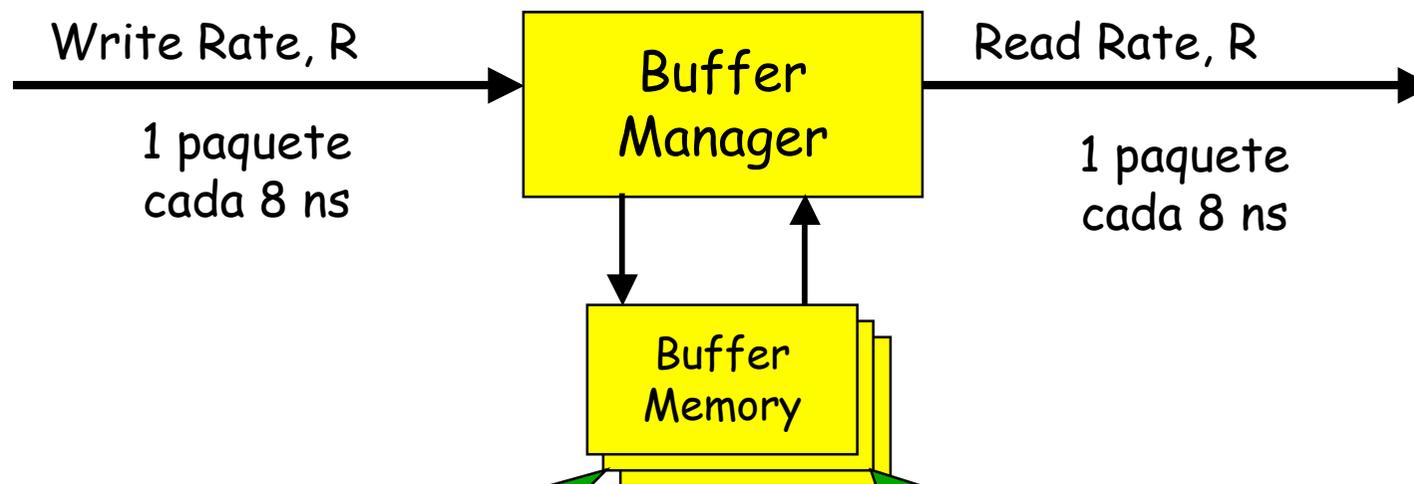
# Arquitectura de router



# Buffers de paquetes muy rápidos

Ejemplo: buffer para puertos de 40Gb/s

Tamaño =  $RTT * BW = 10Gb$ ; 40 byte packets



Usar SRAM?

- + acceso suficientemente rápido, pero...
- poca densidad para conseguir 10Gb.

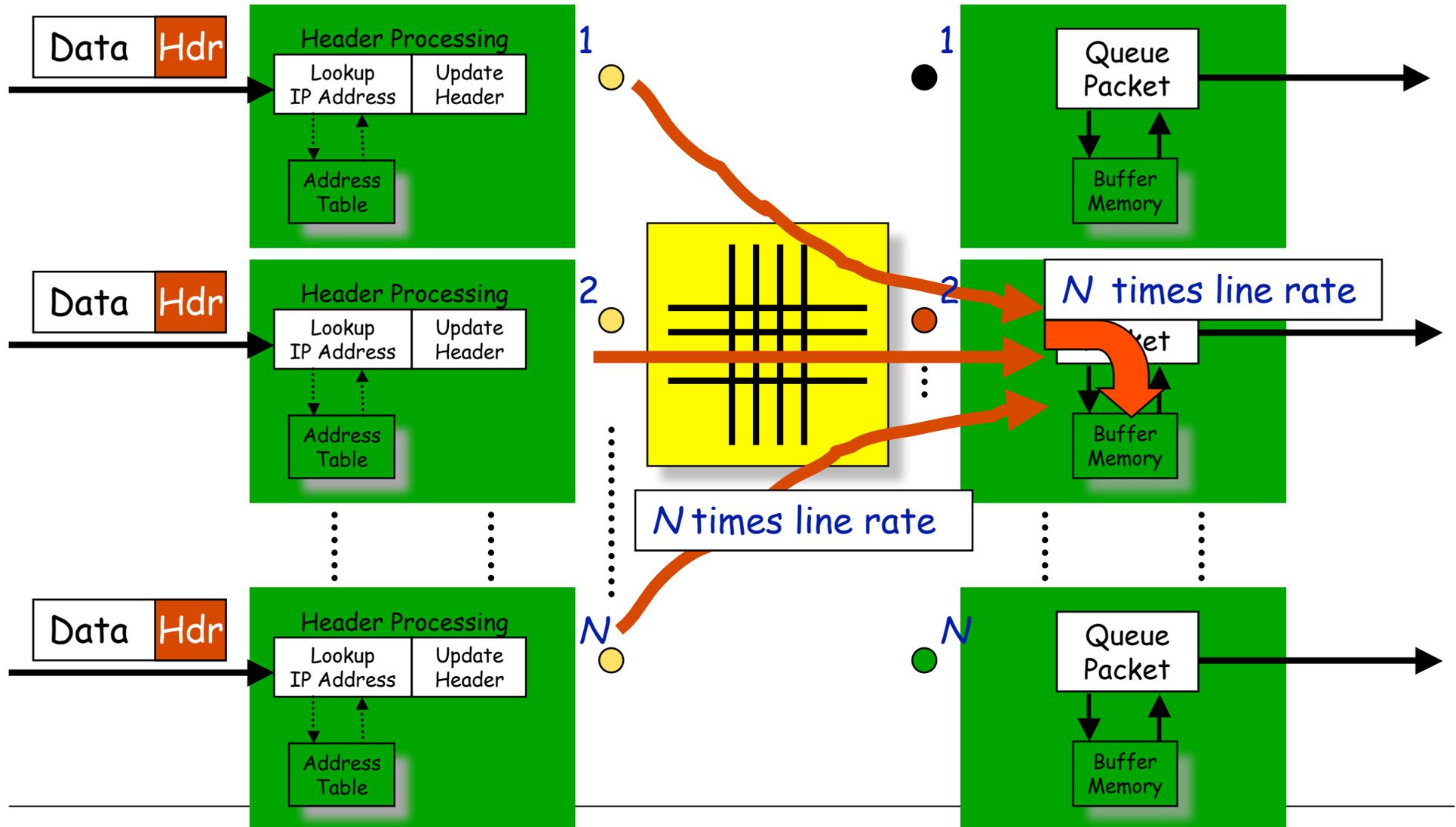
Usar DRAM?

- + alta densidad, suficiente memoria, pero...
- demasiado lenta (50ns tiempo acceso).

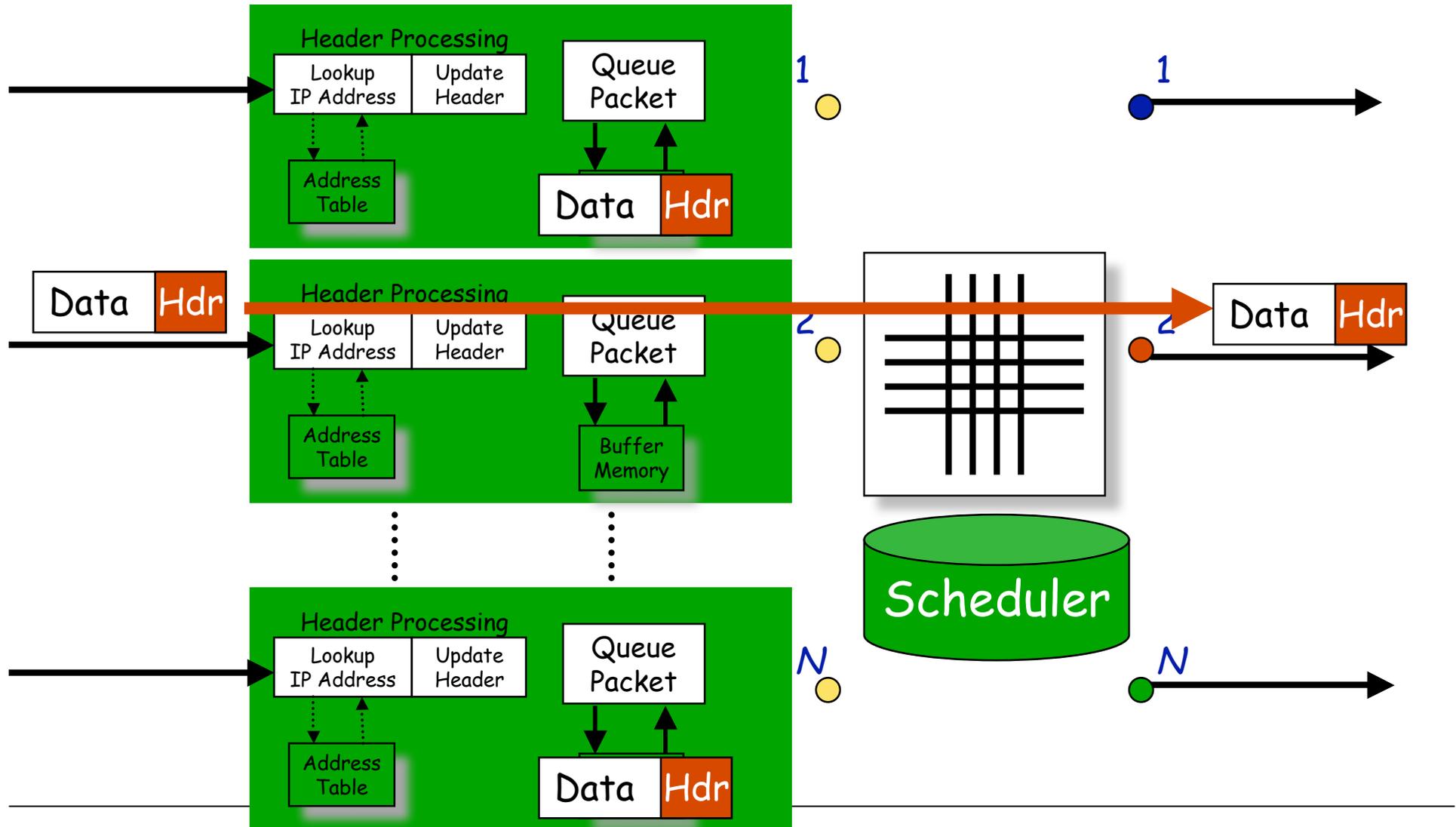
# Contenido

- Background
  - ❖ Qué es un router?
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ Búsqueda de dirección IP.
  - ❖ Almacenamiento de paquetes.
  - ❖ **Conmutación.**
  - ❖ Planificación.

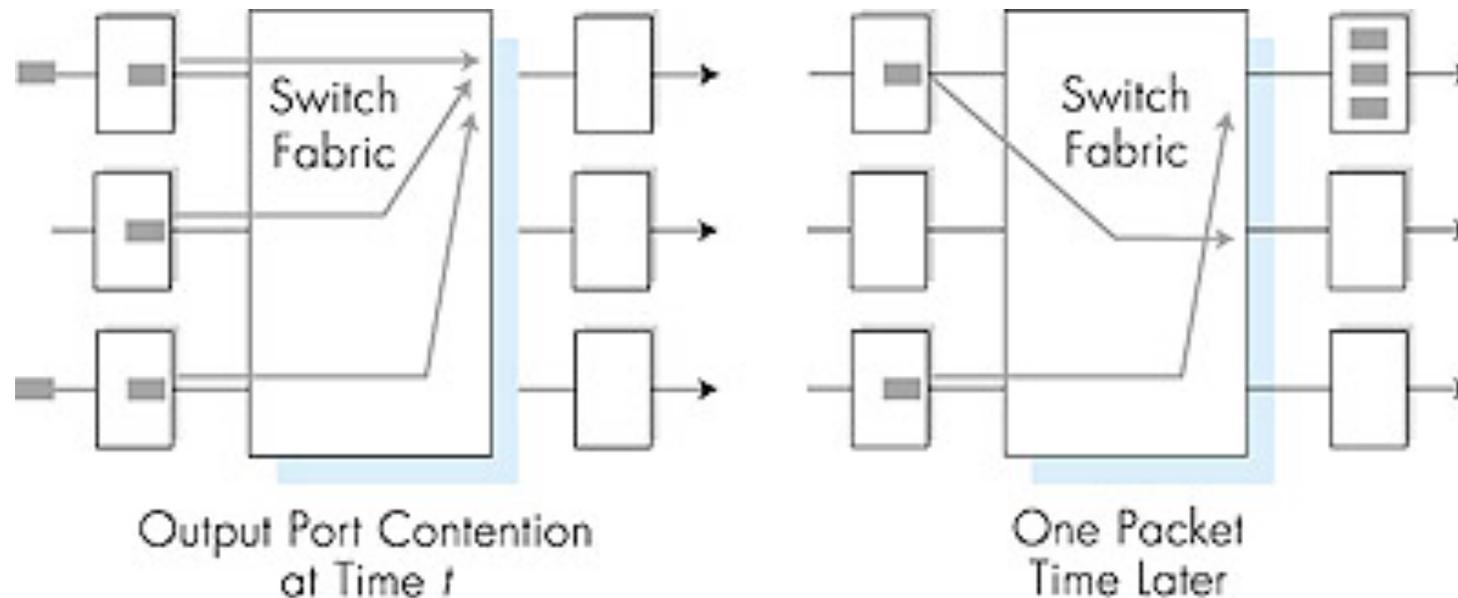
# Arquitectura de router



# Arquitectura de router



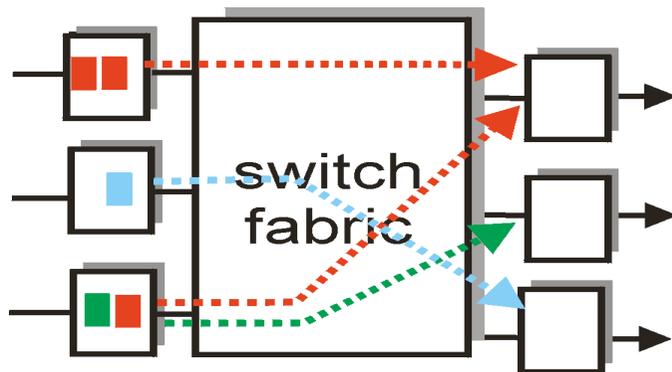
# Usando colas en los puertos de salida



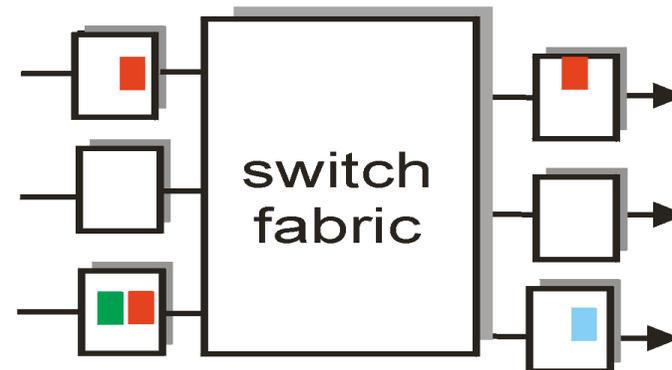
- Cuando la velocidad de llegada a través de la matriz de conmutación es superior a la velocidad de salida se almacenan en buffer (cola de salida)
- *almacenamiento (retraso) y pérdidas debidas al desbordamiento del buffer del puerto de salida*
- Problema: *matriz de conmutación debe ser  $N$  veces mas rápida que el puerto de entrada/salida*

# Usando colas en los puertos de entrada

- Matriz de conmutación más lenta que la suma de los puertos de entrada -> hay que almacenar en el puerto de entrada
- **Bloqueo Head-of-the-Line (HOL):** un paquete que no puede ir a un puerto bloquea al resto de paquetes en el puerto de entrada, aunque podrían ser servidos
- *Retraso y pérdidas debidas a desbordamiento de puerto de entrada*
- Combinación de las dos técnicas para conseguir poco HOL con matrices de conmutación rápidas pero no tanto como la suma de los puertos



output port contention  
at time t - only one red  
packet can be transferred



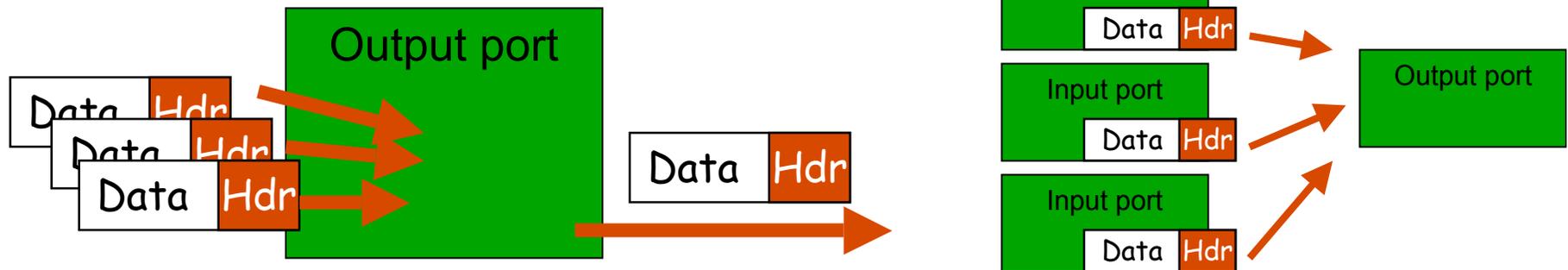
green packet  
experiences HOL blocking

# Contenido

- Background
  - ❖ Qué es un router?
  - ❖ Por qué necesitamos routers más rápidos?
  - ❖ Por qué son difíciles de construir?
- Arquitecturas y técnicas
  - ❖ Evolución de arquitecturas de routers.
  - ❖ Búsqueda de dirección IP.
  - ❖ Almacenamiento de paquetes.
  - ❖ Conmutación.
  - ❖ **Planificación.**

# Planificación (scheduling)

- Tráfico variable. Paquetes llegan al puerto de salida más rápido de lo que pueden ser enviados (o hay paquetes en varios puertos de entrada con colas a la entrada)



- No podemos mantener esta situación por tiempo ilimitado
- Pero podemos amortiguar ráfagas de cierta duración (colas)
- Tenemos varios candidatos de paquetes a enviar para un recurso limitado (sólo podemos enviar uno)

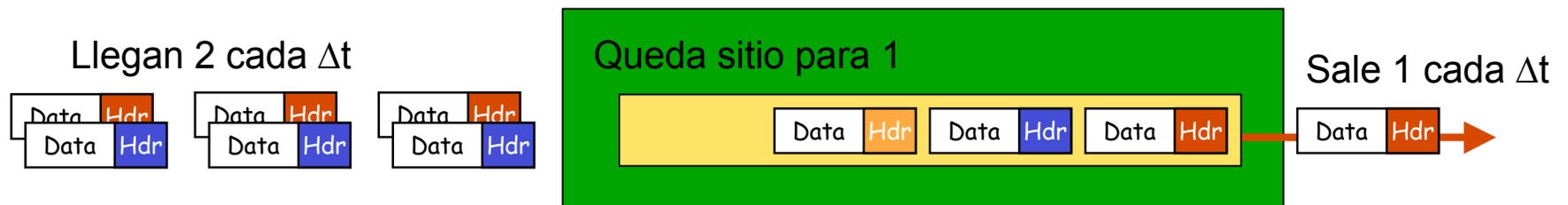
¿Cómo decidimos cual se envía y cual espera?

¿Qué hacemos con los paquetes que llegan cuando el almacenamiento es escaso?

Problema de **planificación** (o **scheduling**)

# Planificación (scheduling)

- Aproximación más simple: por orden de llegada (FCFS First come first served)  
Si un paquete llega y no hay sitio se elimina (DropTail)
- Problemas típicos
  - El orden de llegada puede ser arbitrario (llegar por el primer puerto de entrada?)
  - **Problemas de injusticia entre flujos**



- **A algunas aplicaciones les importa más el retardo que a otras**

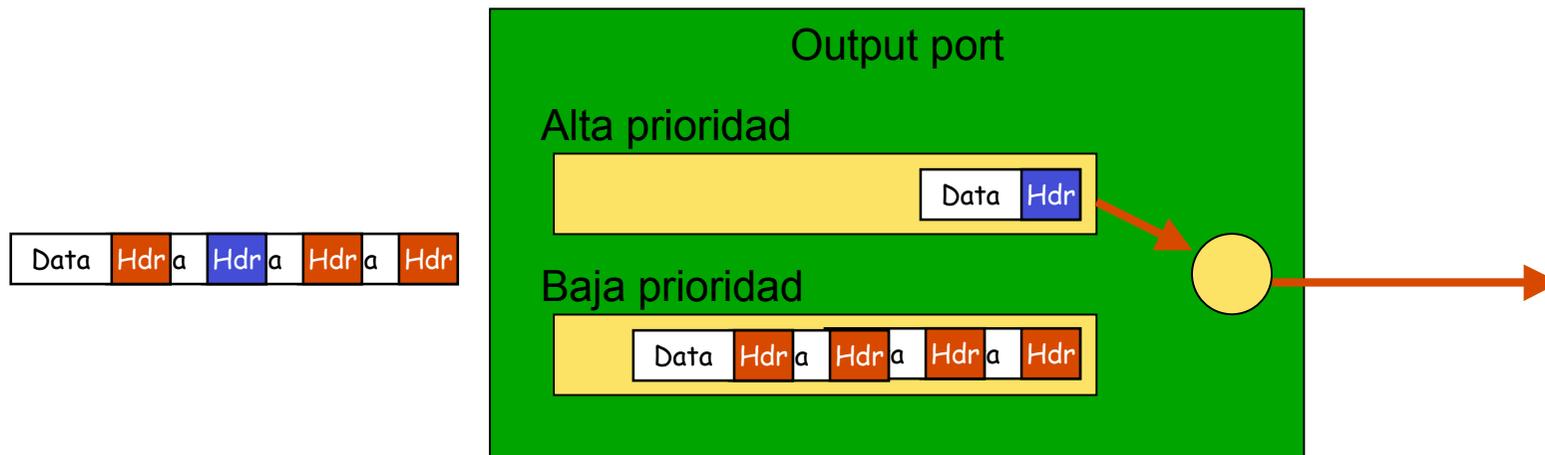


Data Hdr Voz sobre IP pocos paquetes pero muy importante que lleguen rápido (>150ms inutil)

Data Hdr Descarga de fichero muchos paquetes pero no importa que se retrasen muchos milisegundos

# Planificadores

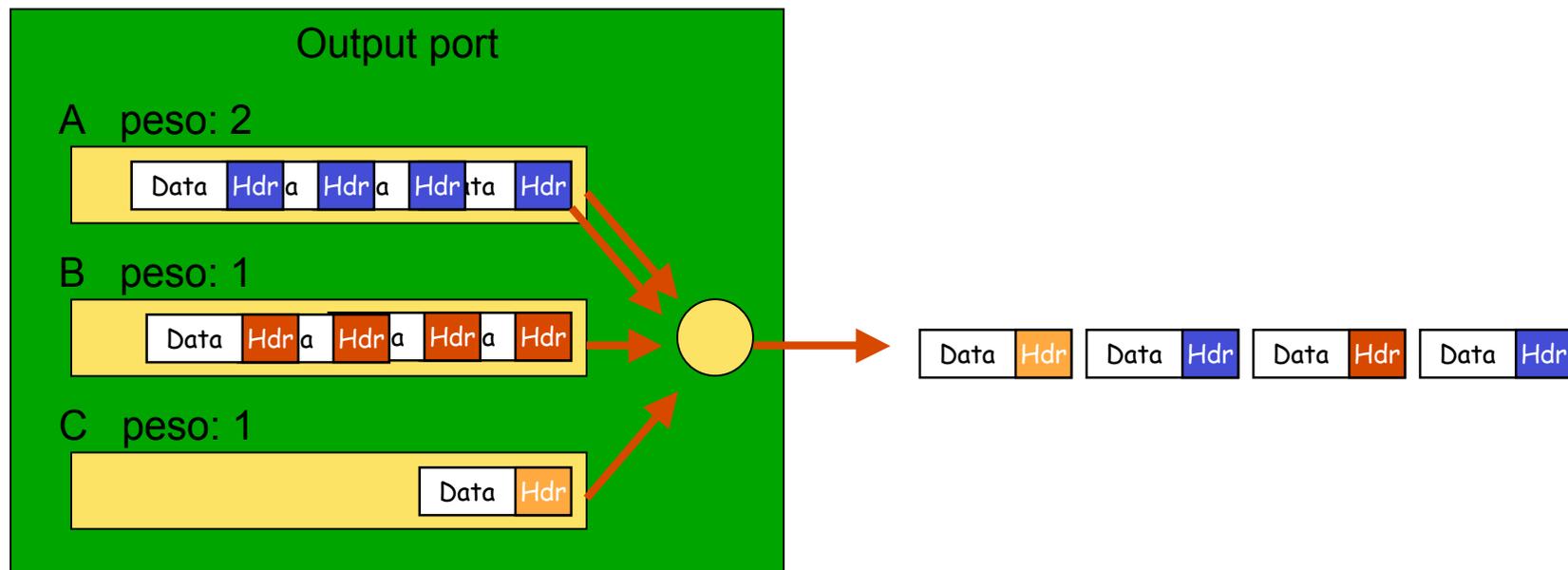
- Colas con prioridad (PQ Priority Queueing) y clases de aplicaciones
- Solo envía de las colas con menos prioridad si no hay nada en las de mas prioridad (con 2 o más niveles)



- Normalmente aplicaciones con requisitos de retardo vs aplicaciones normales. Y algún otro nivel más pero no muchos
  - Limitando clasificación del tráfico
  - Limitando estado interno (mayor complejidad con más colas)

# Planificadores

- Planificadores justos, reparten el envío por igual entre varias clases de paquetes: GPS (Generalized Processor Sharing), WFQ (Weighed Fair Queueing)
- Incluyendo reparto a partes proporcionales (por ejemplo: que la clase A tenga el doble de procesador que la B)



# Conclusiones

- Arquitectura básica de routers y evolución
  - Conmutación via memoria
  - Conmutación via bus
  - Conmutación via red de conmutación
- Problemas importantes en arquitectura de routers
  - Búsqueda rápida de direcciones
  - Almacenamiento de paquetes
  - Conmutación y planificación
- Próxima clase:
  - Problemas
  - y enrutamiento