

## TOC: Switching & Forwarding

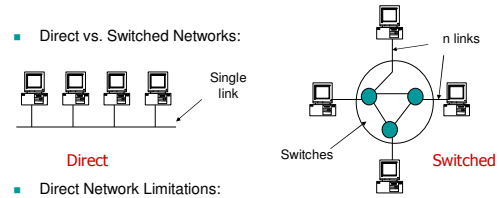
- [Why?](#)
- [Switching Techniques](#)
- [Switch Characteristics](#)
- [Switch Examples](#)
- [Switch Architectures](#)
- [Summary](#)

TOC - Switching

?

## Why?

- Direct vs. Switched Networks:



- Direct Network Limitations:
  - Distance (coordination delay; propagation limitation)
  - Number of hosts (collisions; shared bandwidth; address tables)
  - Single link technology (cannot mix optical, wireless, ...)
- Internetworking: Externality gain at low cost

TOC - Switching - Why?

?

## Techniques

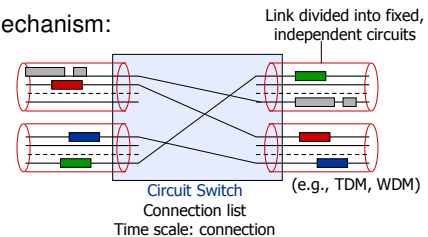
- [Circuit-Switching](#) (e.g., Telephone net.)
- [Packet-Switching](#)
  - [Datagram](#) (e.g., IP, Ethernet)
  - [Virtual Circuits](#) (e.g., MPLS, ATM)
  - [Source Routing](#)
- [Comparison](#)

TOC - Switching - Techniques

?

## Circuit-Switching

- Mechanism:



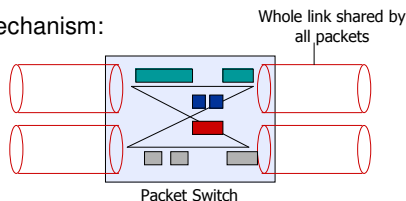
- Features:
  - Packets not switched independently (establish circuit before sending data)
  - Dedicated path and resources from source to destination
  - Setup time; low delays and guaranteed resources thereafter

TOC - Switching - Techniques - Circuit

?

## Packet-Switching

- Mechanism:



- Features:
  - Data separated into packets
  - Switching decision (**output port**) for each individual packet
  - Statistical multiplexing: Sum of peak rates may exceed link bandwidth (as long as mean does not)

TOC - Switching - Techniques - Packet

?

## PS - Datagram

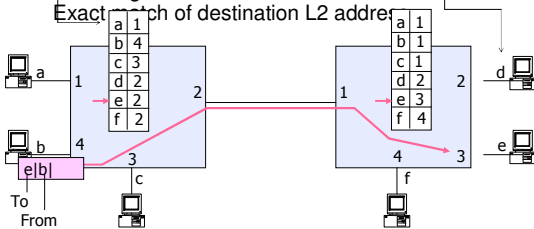
- General idea: no connection establishment, but each packet contains enough info to specify destination
- Switches contain forwarding tables (but no per-connection "state")
- Forwarding tables contain info on which **outgoing port** to use for each destination
- Two types of addressing:
  - [Layer 2](#) or [Layer 3](#)

TOC - Switching - Techniques - Packet - Datagram

?

## Layer 2 (e.g., Ethernet)

- Flat address space (no structure)
- Forwarding table:

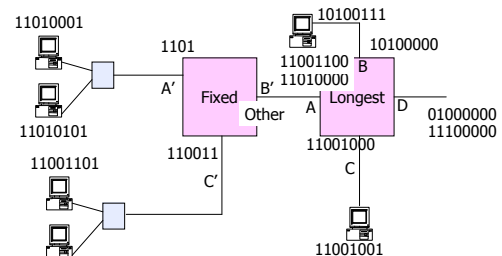


TOC - Switching - Techniques - Packet - Datagram - L2

## Layer 3 (e.g., IP)

### L3-network (e.g., IP)

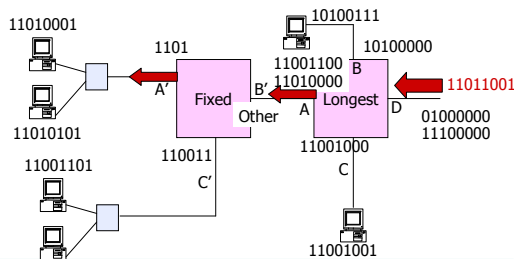
- Topological structure – match prefix
- Either **fixed prefix length** or **longest match**



TOC - Switching - Techniques - Packet - Datagram - L3

## Layer 3 (e.g., IP)

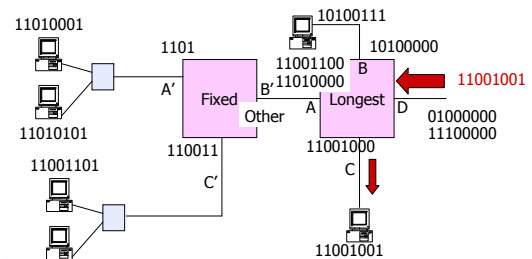
11011001 matches  
4 bits at A, 1 at B, 3 at C → A = LPM  
4 bits at A' → A' = EM



TOC - Switching - Techniques - Packet - Datagram - L3

## Layer 3 (e.g., IP)

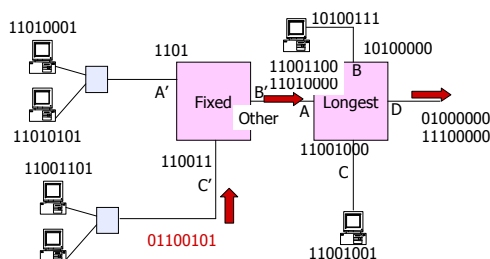
11001001 matches  
5 bits at A, 1 at B, 7 at C → C = LPM



TOC - Switching - Techniques - Packet - Datagram - L3

## Layer 3 (e.g., IP)

01100101 matches  
0 bit at A' → B'  
0 bit at B, 0 at C, 2 at D → D = LPM

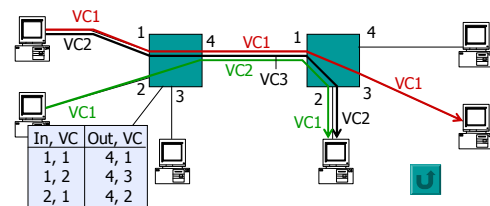


TOC - Switching - Techniques - Packet - Datagram - L3

## PS – Virtual Circuit

Connection setup establishes a path through switches

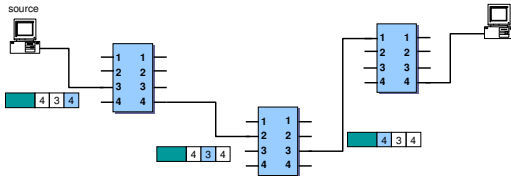
- A virtual circuit ID (VCI) identifies path
- Uses packet switching, with packets containing VCI
- VCIs are often indices into per-switch connection tables; change at each hop



TOC - Switching - Techniques - Packet - VC

## Source Routing

Each packet specifies the sequence of routers (or of output ports) from source to destination



TOC - Switching - Techniques - Source Routing

## Comparison

	Datagram	Virtual circuit switching	Circuit switching
Forwarding cost	high	low	none
Bandwidth utilization	high	flexible	low
Resource reservations	none	flexible	yes
Robustness*	high	low	low

\*The idea is that in case of failure, circuit and VC are lost; datagram routing can adapt after routing update ....

TOC - Switching - Techniques - Comparison

## Characteristics

- Ports
  - Fast Ethernet, OC-3, ATM, ...
- Protocols
  - ST, Link Agg., VLAN, OSPF, RIP, BGP, VPN, Load Balancing, WRED, WFQ
- Performance
  - Throughput, Reliability, Power, ...

TOC - Switching - Characteristics

## Examples

- [Juniper M160](#)
- [Cisco "GSR"](#)
- [Cisco "7600"](#)
- [Cisco "catalyst 6500"](#)
- [Extreme "Summit"](#)
- [Foundry "ServerIron"](#)

TOC - Switching - Examples

## Cisco GSR - 12416

- WAN Router – Large throughput; SONET links
- Up to 16 line cards at 10 Gbps each *Cisco GSR 12416*
- Crossbar Fabric
- Line Cards:
  - 1-port OC-192c
  - 4-port OC48c
  - Many others
  - (ATM, Ethernet, ...)



TOC - Switching - Examples - GSR

## Juniper M160

- WAN Router – Large throughput; SONET links
- Crossbar Fabric
- Line Cards:
  - 1-port OC-192c
  - 4-port OC48c
  - Many others
  - (ATM, Ethernet, ...)



TOC - Switching - Examples - M160

## Cisco 7600

- MAN-WAN Router
- Up to 128 Gbps with Crossbar Fabric
- 10Mbps – 10Gbps LAN Interfaces
- OC-3 to OC-48 SONET Interfaces
- MPLS, WFQ, LLQ, WRED, Traffic Shaping



TOC – Switching – Examples – 7600



## Cisco cat 6500

- From LAN to Access
- 48 to 576 10/100 Ethernet Interfaces
- 10 GE, OC-3, OC-12, OC-48, ATM
- QoS, ACL
- Load Balancing; VPN
- Up to 128Gbps (with crossbar)
- L4-7 Switching
- VLAN
- IP Telephony (E1, T1, inline-power Ethernet)
- SNMP, RMON



TOC – Switching – Examples – Cat6k



## Extreme - Summit

- 48 10/100 ports
- 2 GE (SX, LX, or LX-70)
- 17.5Gbps non-blocking
- 10.1 Mpps
- Wire speed L2
- Wire speed L3 static or RIP
- OSPF, DVRMP, PIM, ...



TOC – Switching – Examples – Summit



## Foundry - ServerIron

- Server Load Balancing
- Transparent Cache Switching
- Firewall Load Balancing
- Global Server Load Balancing
- Extended Layer 4-7 functionality including URL-, Cookie-, and SSL Session ID-based switching
- Secure Network Address Translation (NAT) and Port address



TOC – Switching – Examples – ServerIron



## Architectures

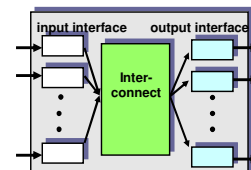
- [Generic Architecture](#)
- [First Generation](#)
- [Second Generation](#)
- [Third Generation](#)
- [Input Functions](#)
- [Output Functions](#)
- [Interconnection Designs](#)
  - [OUT](#)
  - [IN](#)
  - [VOB](#)
  - [Combined IN/OUT](#)

TOC – Switching – Architectures



## Generic

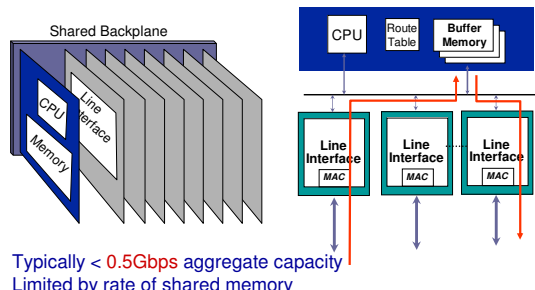
- Input and output interfaces are connected through an interconnect
- A interconnect can be implemented by
  - Shared memory
    - low capacity routers (e.g., PC-based routers)
  - Shared bus
    - Medium capacity routers
  - Point-to-point (switched) bus
    - High capacity routers



TOC – Switching – Architectures – Generic



### First Generation

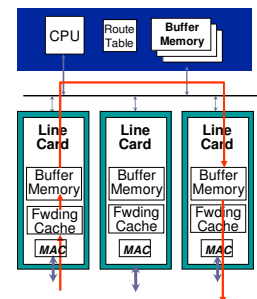


Slide by Nick McKeown

TOC - Switching - Architectures - First

### Second Generation

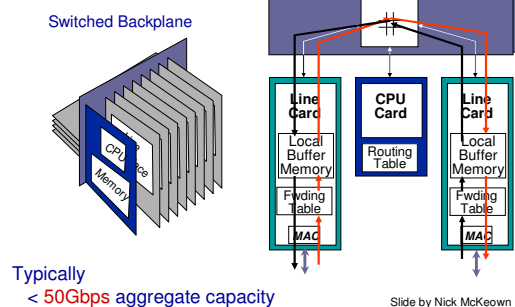
Typically  
< 5Gb/s aggregate capacity  
Limited by shared bus



Slide by Nick McKeown

TOC - Switching - Architectures - Second

### Third Generation



Slide by Nick McKeown

TOC - Switching - Architectures - Third

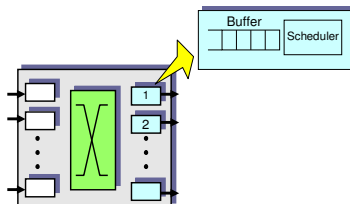
### Input Functions

- **Packet forwarding:** decide to which output interface to forward each packet based on the information in packet header
  - examine packet header
  - lookup in forwarding table
  - update packet header

TOC - Switching - Architectures - Input Functions

### Output Functions

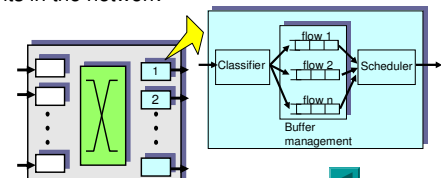
- **Buffer management:** decide when and which packet to drop
- **Scheduler:** decide when and which packet to transmit



TOC - Switching - Architectures - Output Functions

### Output Functions (ct)

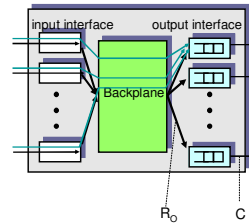
- **Packet classification:** map each packet to a predefined flow/connection (for datagram forwarding)
  - use to implement more sophisticated services (e.g., QoS)
- **Flow:** a subset of packets between any two endpoints in the network



TOC - Switching - Architectures - Output Functions

## Output Queued

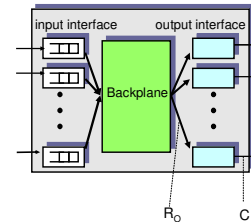
- Only output interfaces store packets
- Advantage
  - Easy to design algorithms: only one congestion point
- Disadvantage
  - Requires an output speedup  $R_0/C = N$ , where  $N$  is the number of interfaces  $\rightarrow$  not feasible for large  $N$



TOC - Switching - Architectures - Output Queued

## Input Queues

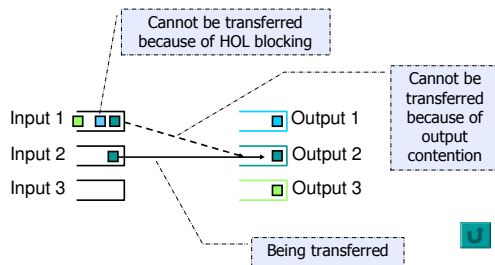
- Only input interfaces store packets
- Advantages
  - Easy to build
  - Simple algorithms
- Disadvantages
  - HOL Blocking
  - In practice: Speedup of 2 suffices



TOC - Switching - Architectures - Input Queued

## Note: Head-of-line Blocking

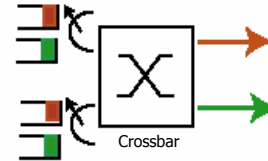
- The cell at the head of an input queue cannot be transferred, thus blocking the following cells



TOC - Switching - Architectures - Input Queued: HOL

## Virtual Output Buffers

- OUT buffers at each input port
  - Complexity: Matching Problem
  - Full throughput algorithm
  - Good

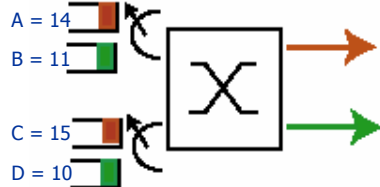


Note: Figure from Prof. Varaiya's notes for EE228b

TOC - Switching - Architectures - VOB

## VOB: Full Throughput

- Maxim

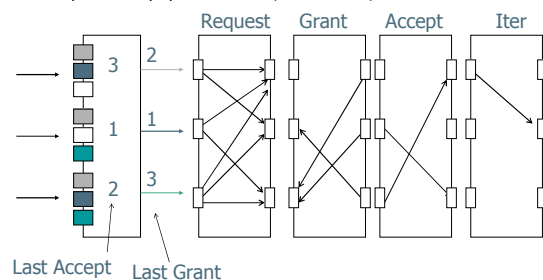


$$B + C > A + D \Rightarrow \text{Serve } (B, C)$$

TOC - Switching - Architectures - VOB: Full Throughput

## VOB: Good Heuristic - i-SLIP

- Inputs request permission to send from outputs
- Outputs grant permissions to inputs (round-robin)
- Inputs accept permissions (round-robin)

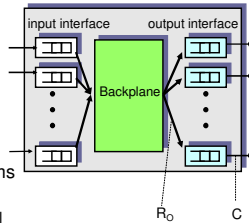


Last Accept Last Grant

TOC - Switching - Architectures - VOB: iSLIP

## Combined IN/OUT

- Both input and output interfaces store packets
- Advantages
  - Easy to built
    - Utilization 1 can be achieved with limited input/output speedup ( $\leq 2$ )
- Disadvantages
  - Harder to design algorithms
    - Two congestion points
    - Need to design flow control



TOC – Switching – Architectures – IN/OUT



## Summary

- Switching needed for big networks
- Internetworking externality
- Circuit
- Packet – VC: QoS possible
- Packet – Datagram
  - L2: Limited by flat address space
  - L3:
    - Exact Match: Easy lookup – less efficient
    - Longest Prefix Match
- Switch functions: control and data
- Different Architectures:
  - cost vs. performance

TOC – Switching – Summary

