Computer Networks
COL 334/672

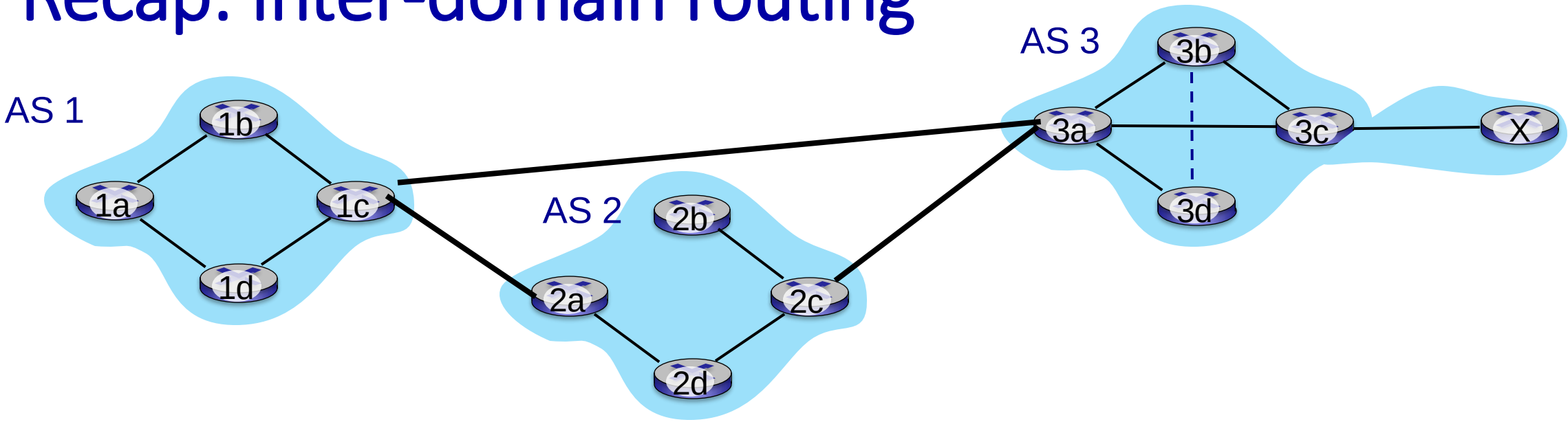
Inter-domain Routing

Tarun Mangla

Slides adapted from KR

Sem 1, 2024-25

Recap: Inter-domain routing

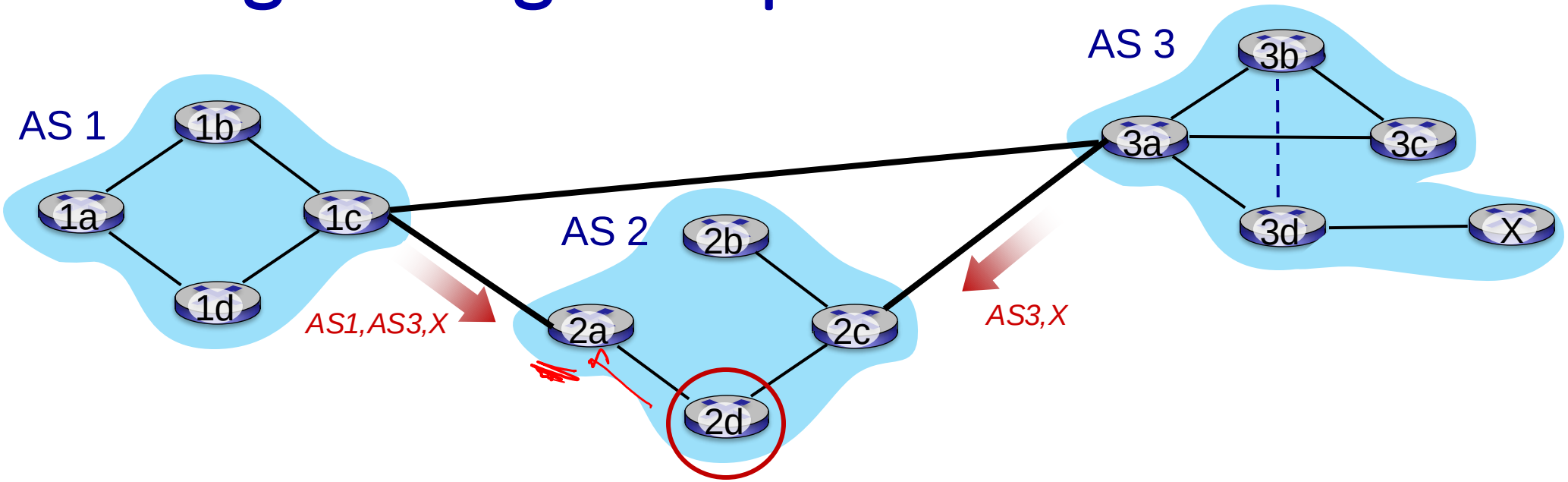


Border Gateway Router (BGP) is the de facto inter-domain routing protocol

- Consists of eBGP and iBGP connections [TCP Connections]
- ASes announce BGP advertisements to neighboring AS
- The advertisement consists of path attributes and IP prefix
- Path attributes include AS path and next hop

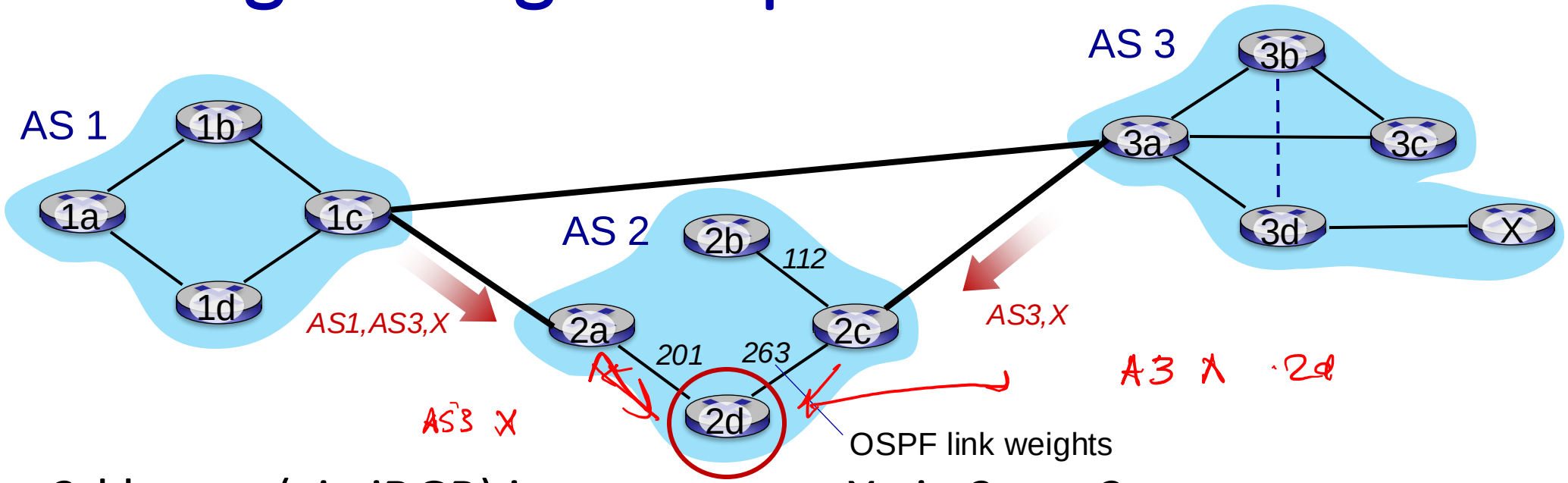
How does a BGP router select among the multiple announcements?

Selecting Among Multiple Announcements



- 2d learns (via iBGP) it can route to X via 2a or 2c
- shortest AS path: choose local gateway that has fewest number of AS hops to the destination

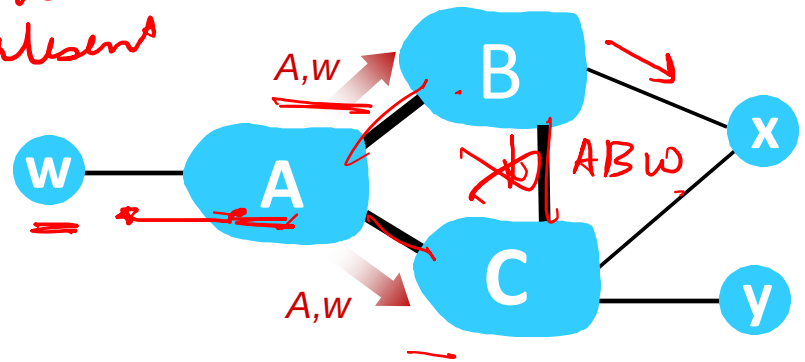
Selecting Among Multiple Announcements

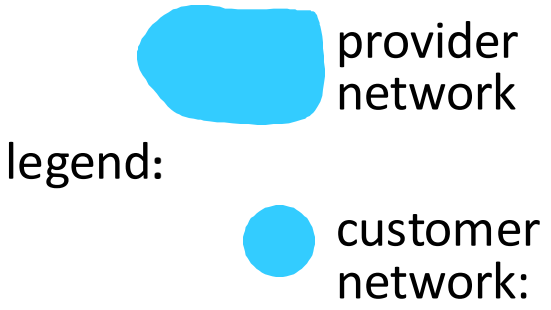


- 2d learns (via iBGP) it can route to X via 2a or 2c
- shortest AS path: choose local gateway that has fewest number of AS hops to the destination
- hot potato routing: choose local gateway that has least *intra-domain* cost (e.g., 2d chooses 2a, even though more AS hops to *X*): don't worry about inter-domain cost!
- network policy: for both path selection and announcements!

BGP: achieving policy via advertisements

Policy
① Route selection
② Route advertisement





ISP only wants to route traffic to/from its customer networks (does not want to carry transit traffic between other ISPs – a typical "real world" policy)

- A advertises path Aw to B and to C
- B *chooses not to advertise* BAw to C!
  - B gets no "revenue" for routing CBAw, since none of C, A, w are B's customers
  - C does *not* learn about CBAw path
- C will route CAw (not using B) to get to w

Inter-AS Routing: Policies

- *No transit traffic through certain ASes*
- *Never put China on a route starting from Ministry of Defence*
- *Do not use Singapore to get from India to Maldives*
- *Traffic starting or ending at Google should not transit through Microsoft*
- *Don't share information about this peering link to other routers*

1. local pref
2. AS path
3. hot potato routing

What is the mechanism in BGP?

Set local preference on inbound routes

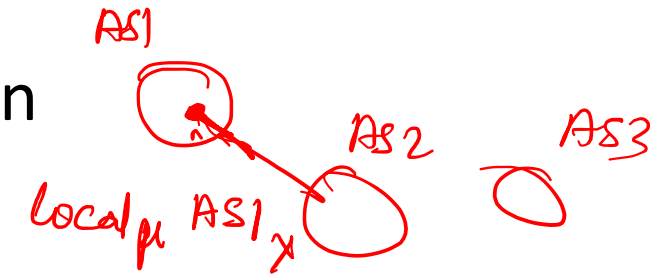
BGP route selection

NOC: Network Operation Center

AS2 ↓ 200

AS3 ↑ 100

- router may learn about more than one route to destination AS, selects route based on:
  1. local preference value attribute: policy decision
  2. shortest AS-PATH
  3. closest NEXT-HOP router: hot potato routing
  4. additional criteria ..



cold potato routing

Summary

- Routing Algorithms: Finding "shortest" path from sending host to receiving host
- Intra-domain routing and inter-domain routing
  - Intra: focus more on performance
  - Inter: focus more on policy
- Intra-domain routing
  - Distance vector (e.g., RIP, EIGRF)
  - Link state (e.g., OSPF)
- Inter-domain routing
  - Path vector routing
  - Border Gateway Protocol (BGP)
- All examples of per-router control plane or a distributed control plane

Centralize control plane → SDN
Software Defined Network

Two key network-layer functions

network-layer functions:

- *forwarding:* move packets from a router's input link to appropriate router output link
- *routing:* determine route taken by packets from source to destination
  - *routing algorithms*

analogy: taking a trip

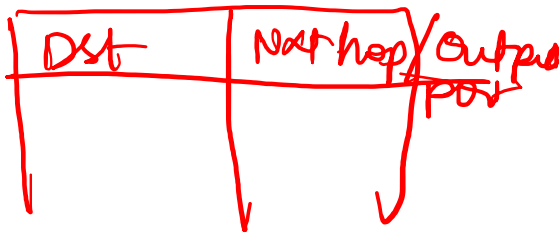
- *forwarding:* process of getting through single interchange
- *routing:* process of planning trip from source to destination

forwarding

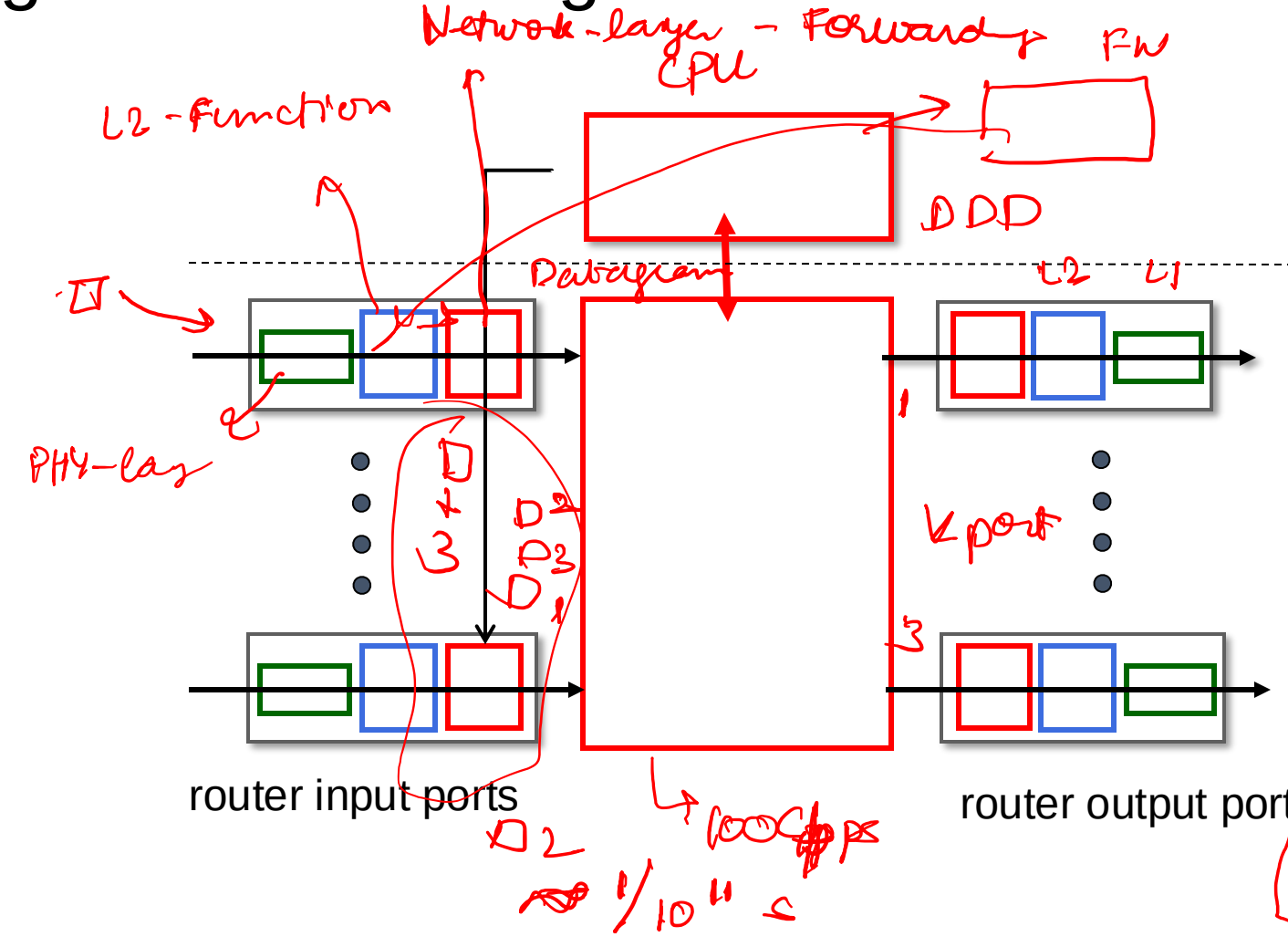
routing

SRAM & DRAM

Router architecture overview

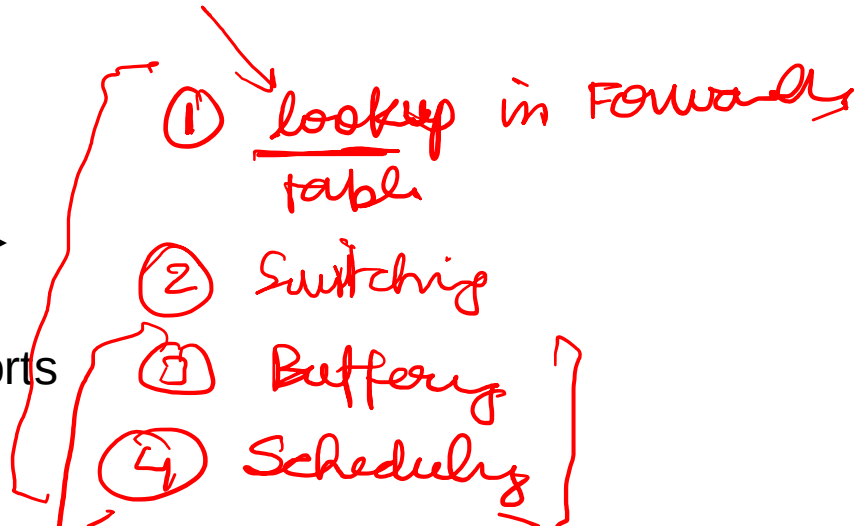


high-level view of generic router architecture:



routing, management control plane (software) operates in millisecond time frame

forwarding data plane (hardware) operates in nanosecond timeframe



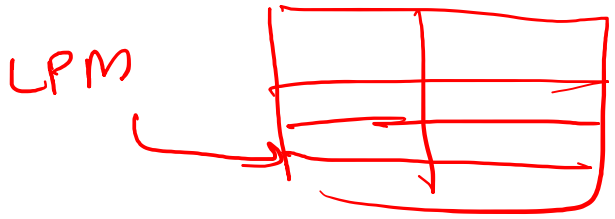
Lookup

forwarding table

Destination Address Range	Link Interface
11001000 00010111 00010000 00000000 through 11001000 00010111 00010000 00000100	0
11001000 00010111 00010000 00000101 through 11001000 00010111 00010000 00000111	1
11001000 00010111 00011000 11111111	2
11001000 00010111 00011001 00000000 through 11001000 00010111 00011111 11111111	2
otherwise	3

Q: but what happens if ranges don't divide up so nicely?

Longest prefix matching



longest prefix match

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** ********	0
11001000 00010111 00011000 ********	1
11001000 00010111 00011*** ********	2
otherwise	3

match!

examples:

11001000 00010111 00010110 10100001	which interface?
11001000 00010111 00011000 10101010	which interface?

Longest prefix matching

longest prefix match

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** ********	0
11001000 00010111 00011000 ********	1
11001000 00010111 00011*** ********	2
otherwise	3

match!

examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?

Longest prefix matching

longest prefix match

when looking for forwarding table entry for given destination address, use *longest* address prefix that matches destination address.

Destination Address Range	Link interface
11001000 00010111 00010*** ********	0
11001000 00010111 00011000 ********	1
11001000 00010111 00011*** ********	2
otherwise	3

match!

examples:

11001000 00010111 00010110 10100001 which interface?

11001000 00010111 00011000 10101010 which interface?