Router Design

Lecture 21

http://www.cs.rutgers.edu/~sn624/352-F24

Srinivas Narayana

RUTGERS
UNIVERSITY | NEW BRUNSWICK

Review of concepts

Network layer’s main function: moving data from one endpoint to another

Analogy: postal system

endpoint

Network layer

endpoint

Addressing (IPv4)

Locate, not identify

10000000 11000011 00000001 01010000

128 . 195 . 1 . 80

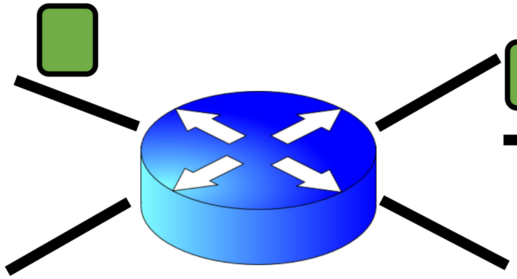
Classless (CIDR)

IP prefixes
==
zip code

128.195.0.0/20

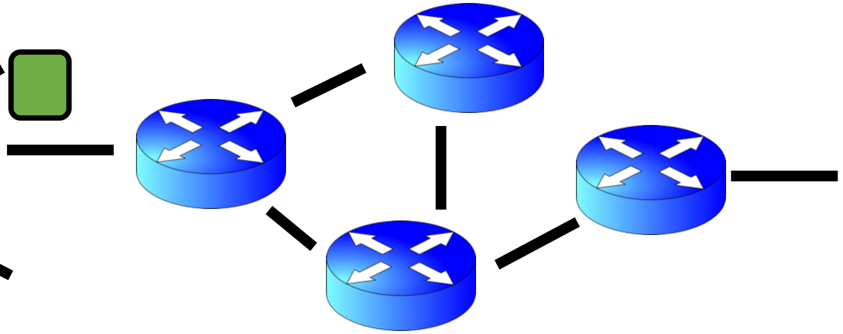
Forwarding

Data plane



Routing

Control plane



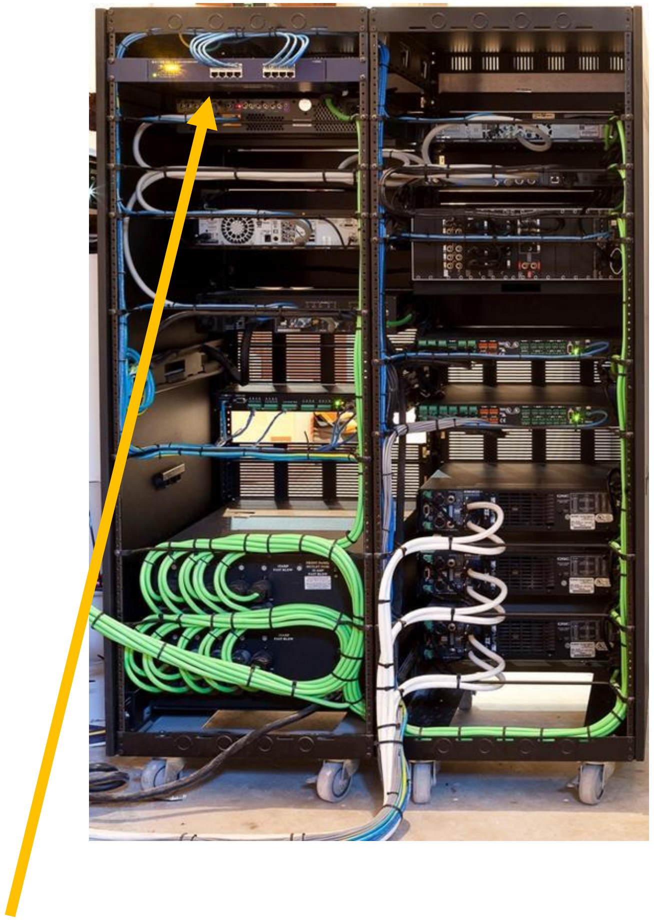
Advantages of prefix-based IP organization

- Aggregate information across endpoints for forwarding & routing
  - Don’t reason about individual addresses; do prefixes instead
  - Reduce the sizes of information exchanged and router data structures
- Prefixes (not individual IPs) are allocated to organizations by Internet registries
  - Each organization is delegated the work of assigning individual IPs
- Facilitates movement of entire groups of hosts between organizations
- (CIDR) IP address is decoupled from an explicit prefix length
  - Different routers can interpret an address with different prefix lengths
  - E.g., further away: more aggregated (shorter prefix); closer to destination: more granular (longer prefix)

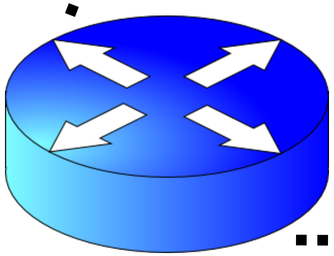
Next we'll talk about routers

Access routers

Internet core router

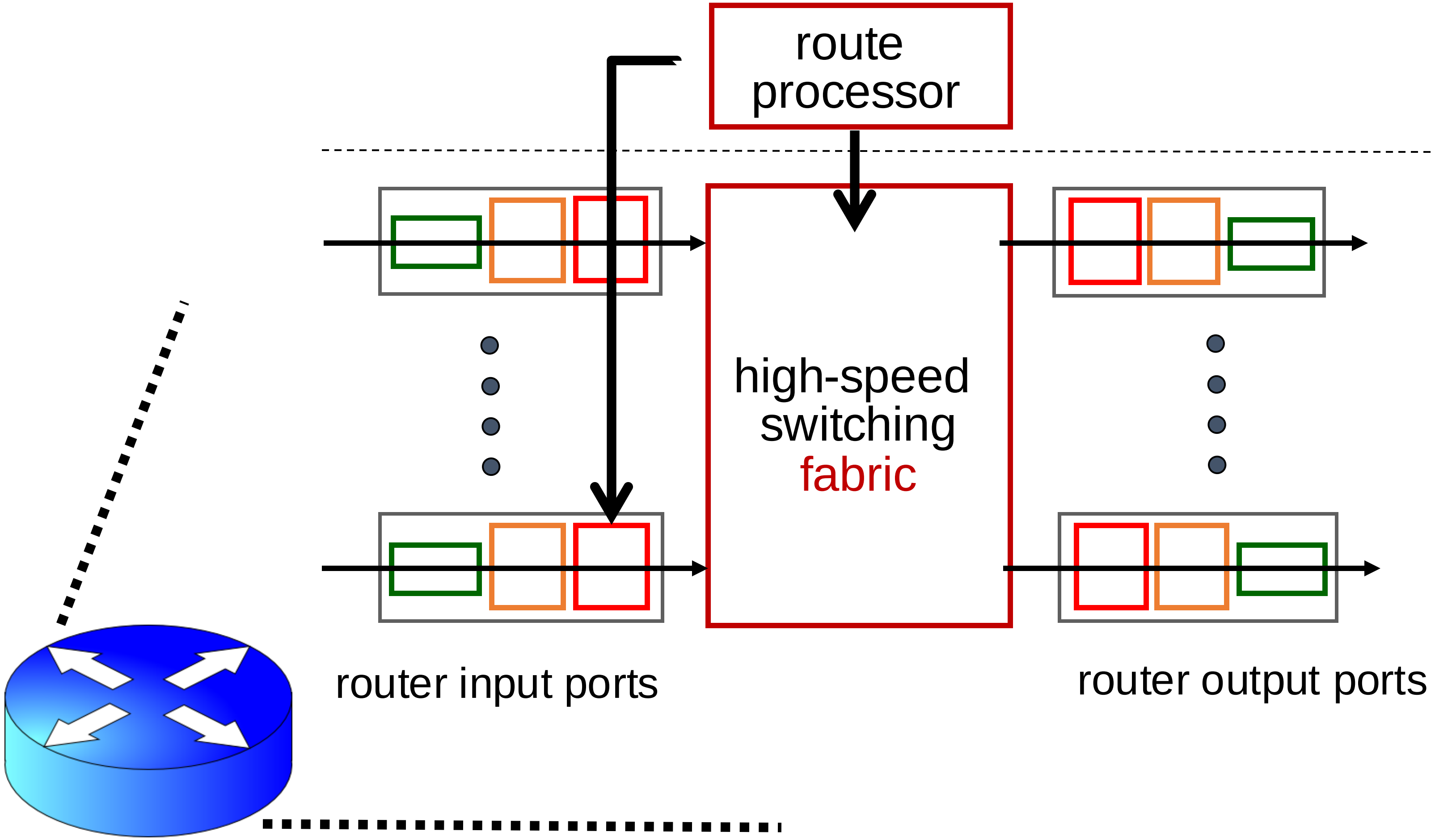


Data center top-of-rack switch



What’s inside a router?

Router architecture overview



Review: assuming distributed routing, routing function: decide which ports packets need to exit

Control plane

Data plane

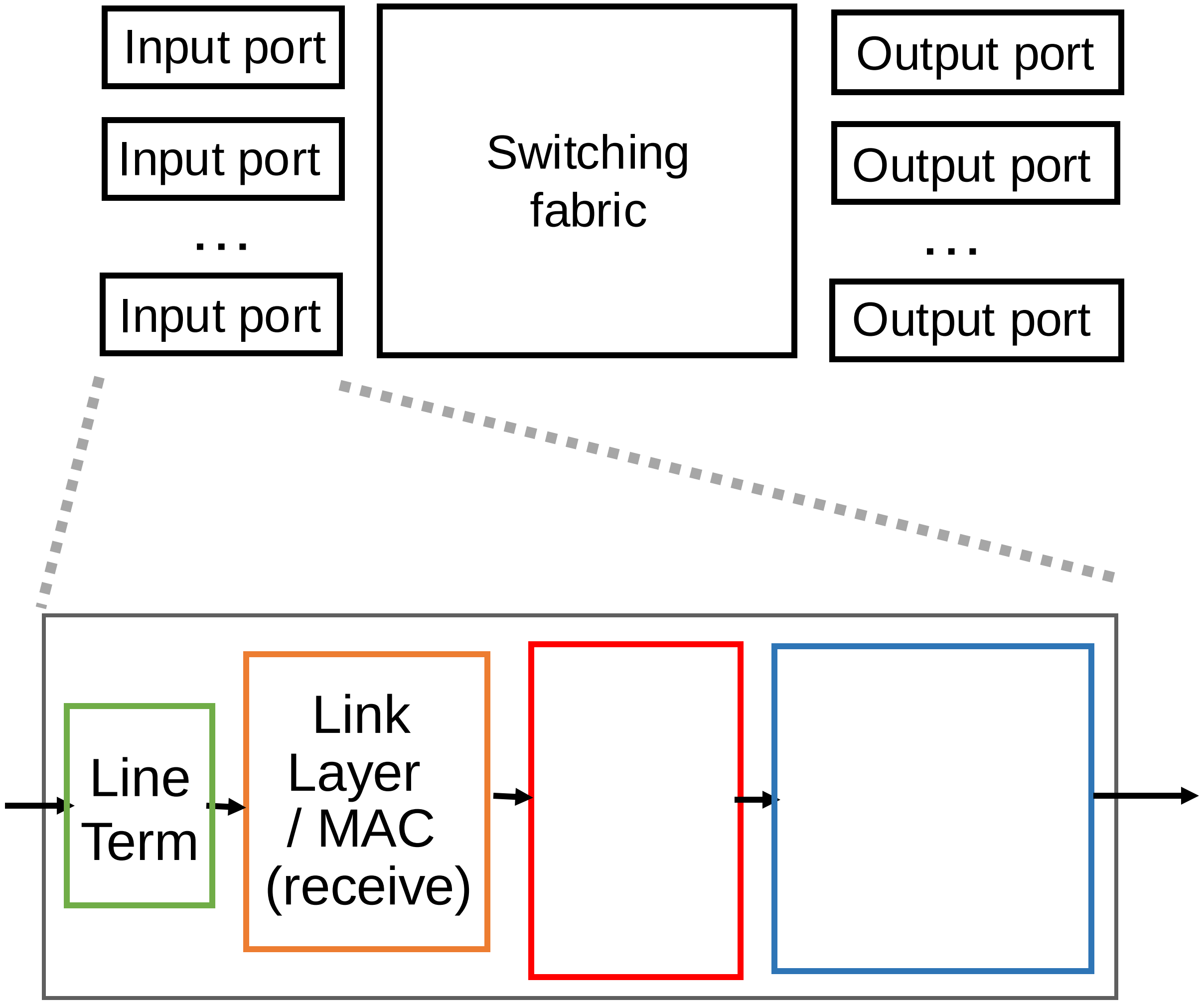
Review: Forwarding function: move packets from one input port to another.

Different and evolving designs

- There are different kinds of routers, with their own designs
  - Access routers (e.g., home WiFi), chassis/core routers, top-of-rack switches
- Router designs have also evolved significantly over time
- For simplicity and concreteness, we will learn about one high-speed router design from the early 2000s.
- Called the MGR (multi-gigabit router). It could support an aggregate rate of 50 Gbit/s (1 G = 10^9)
  - Today's single-chip routers can support aggregate rates of ~10 Tbit/s (1 T = 10^{12})

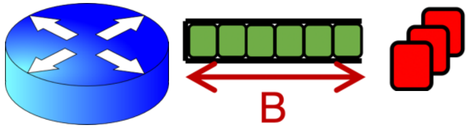
Input port functions

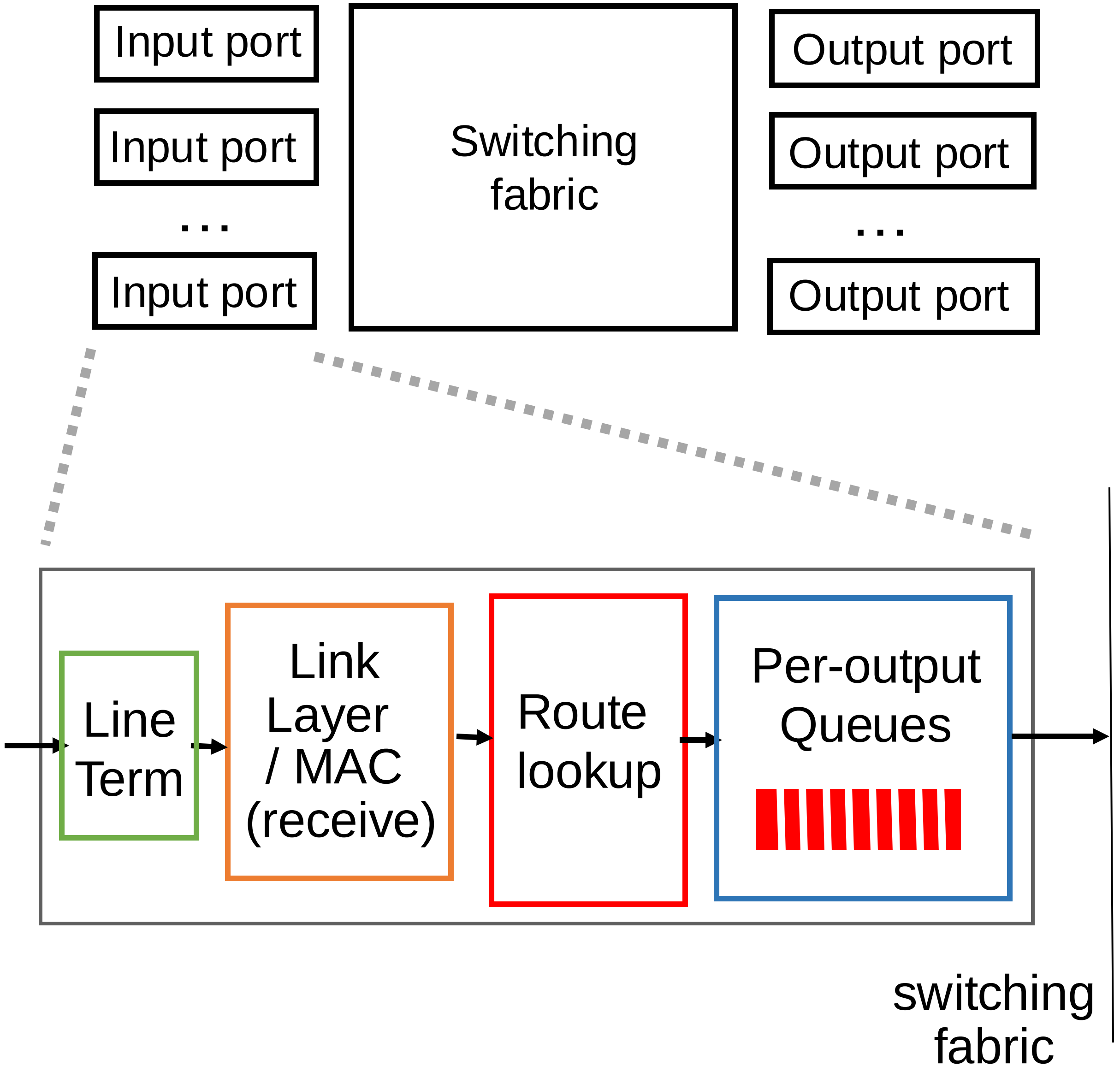
- Line termination: receives physical (analog) signals and turns them into digital signals (physical layer)
- Rate of link connecting to a single port termed line speed or line rate (modern routers: 100+ Gbit/s)
- Link layer: performs medium access control functions (e.g., Ethernet)



Input port functions

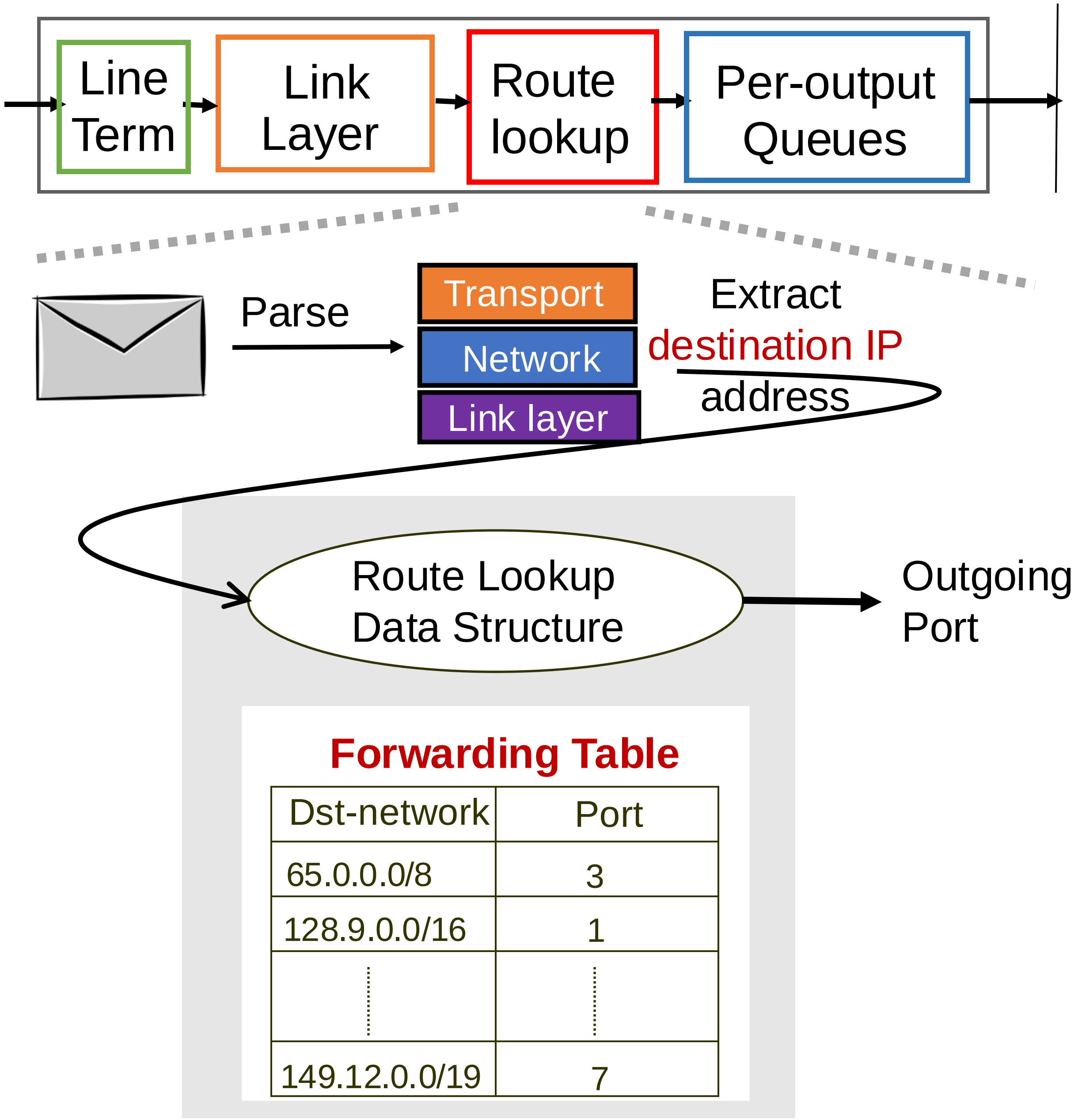
- Route lookup: high-speed lookup of which output port the packet is destined to
- Goal: must complete this processing at the line rate
- Queueing: packets may wait in per-output-port queues if packets are arriving too fast for the switching fabric to send them to the output port





Route lookups

Packet forwarding in the Internet is based on the destination IP address on the packet.



Example: if dst IP on packet is 65.45.145.34, it matches the prefix 65.0.0.0/8 (netmask 255.0.0.0) in table

`(IP & netmask == prefix)`

The packet is forwarded out port 3.

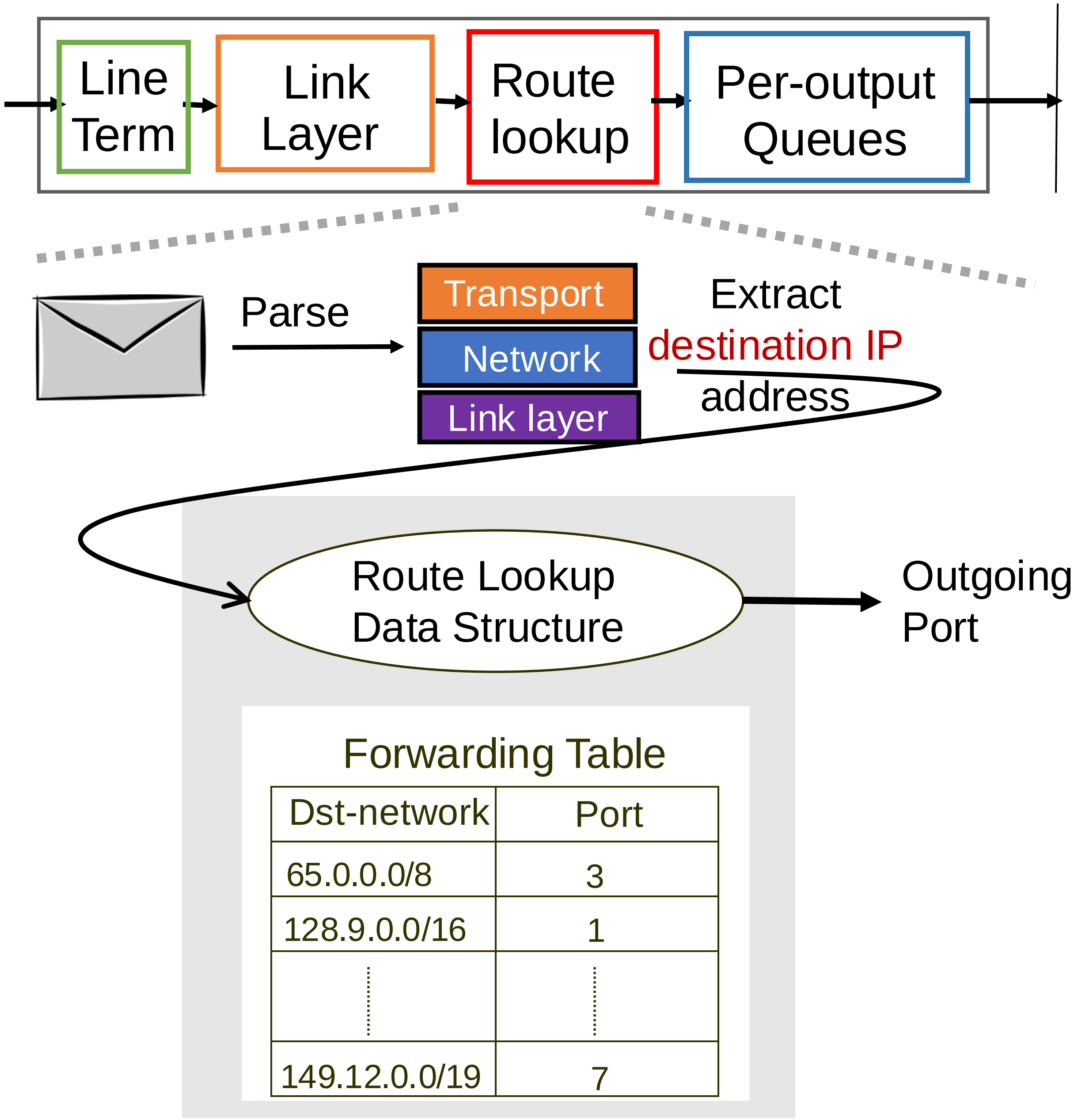
Example 2: what about dst IP 128.9.5.6?

Route lookups

Number of entries in the forwarding table matters.

Fitting into router memory

Designing hardware and software for fast lookups

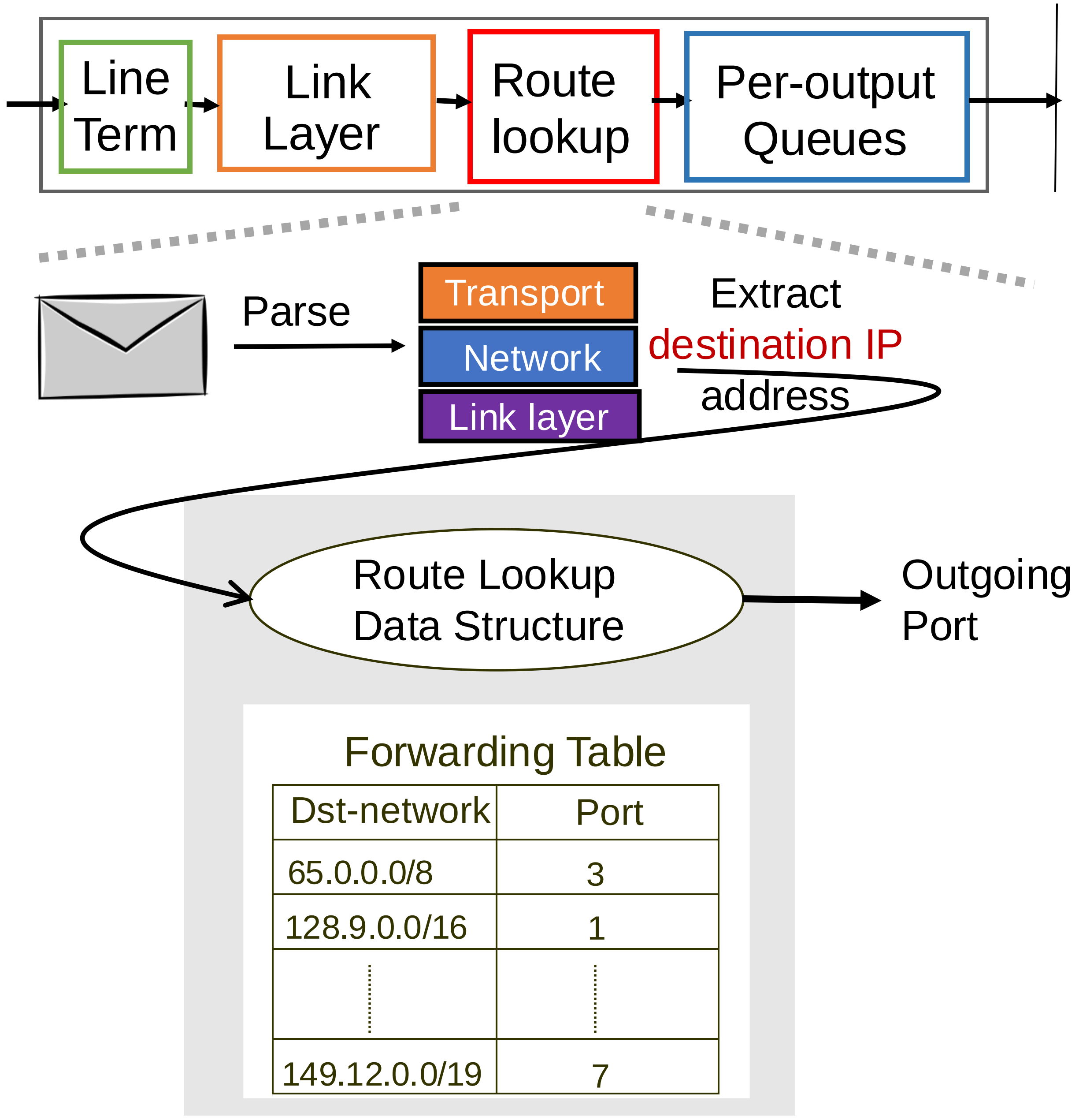


Route lookups

Recall: IP addresses can be aggregated based on shared prefixes.

The number of table entries in a router is proportional to the number of prefixes, NOT the number of endpoints.

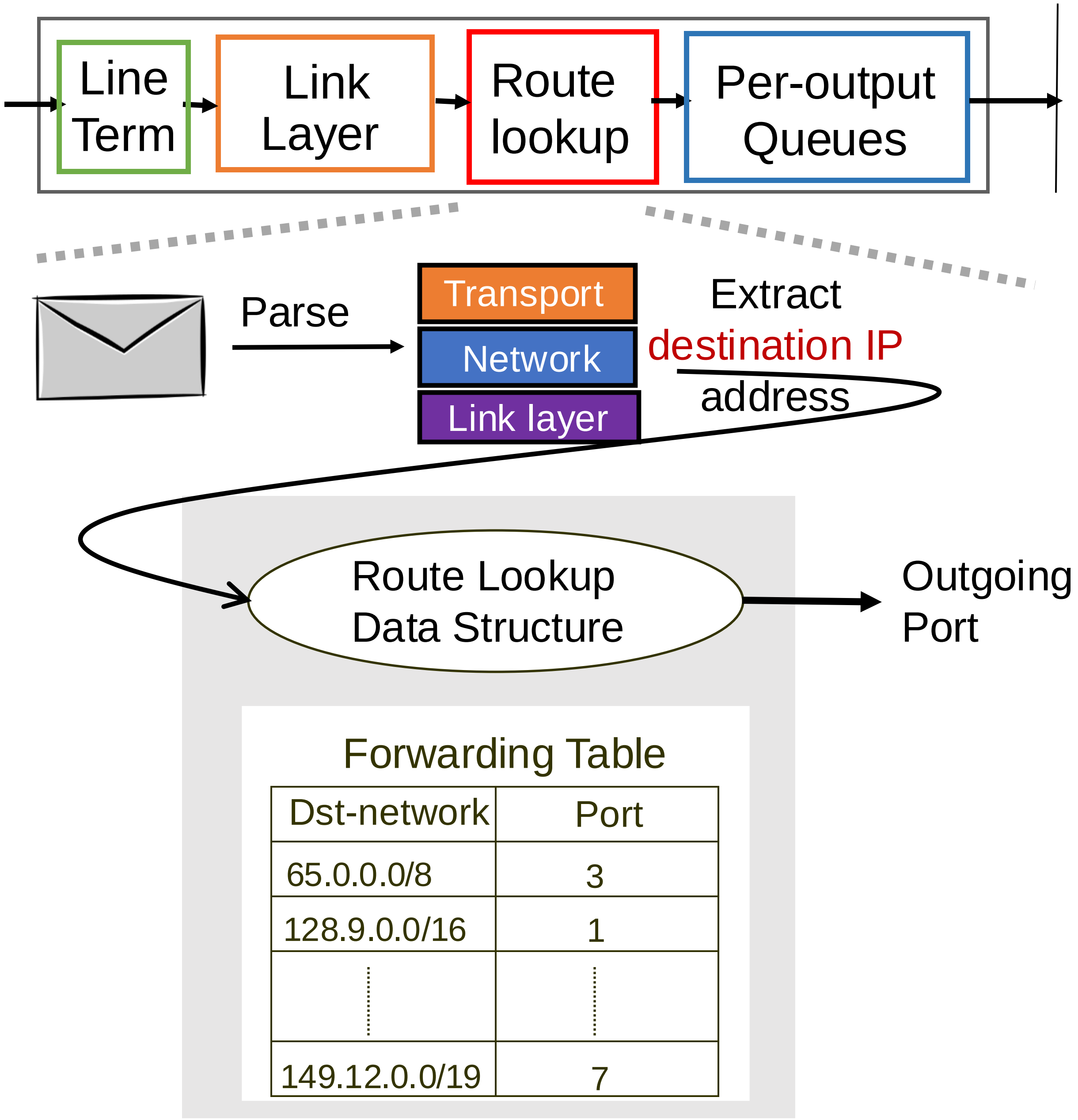
Today: ~ 1 million prefixes.



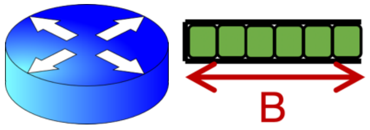
Route lookups

Destination-IP-based forwarding has consequences.

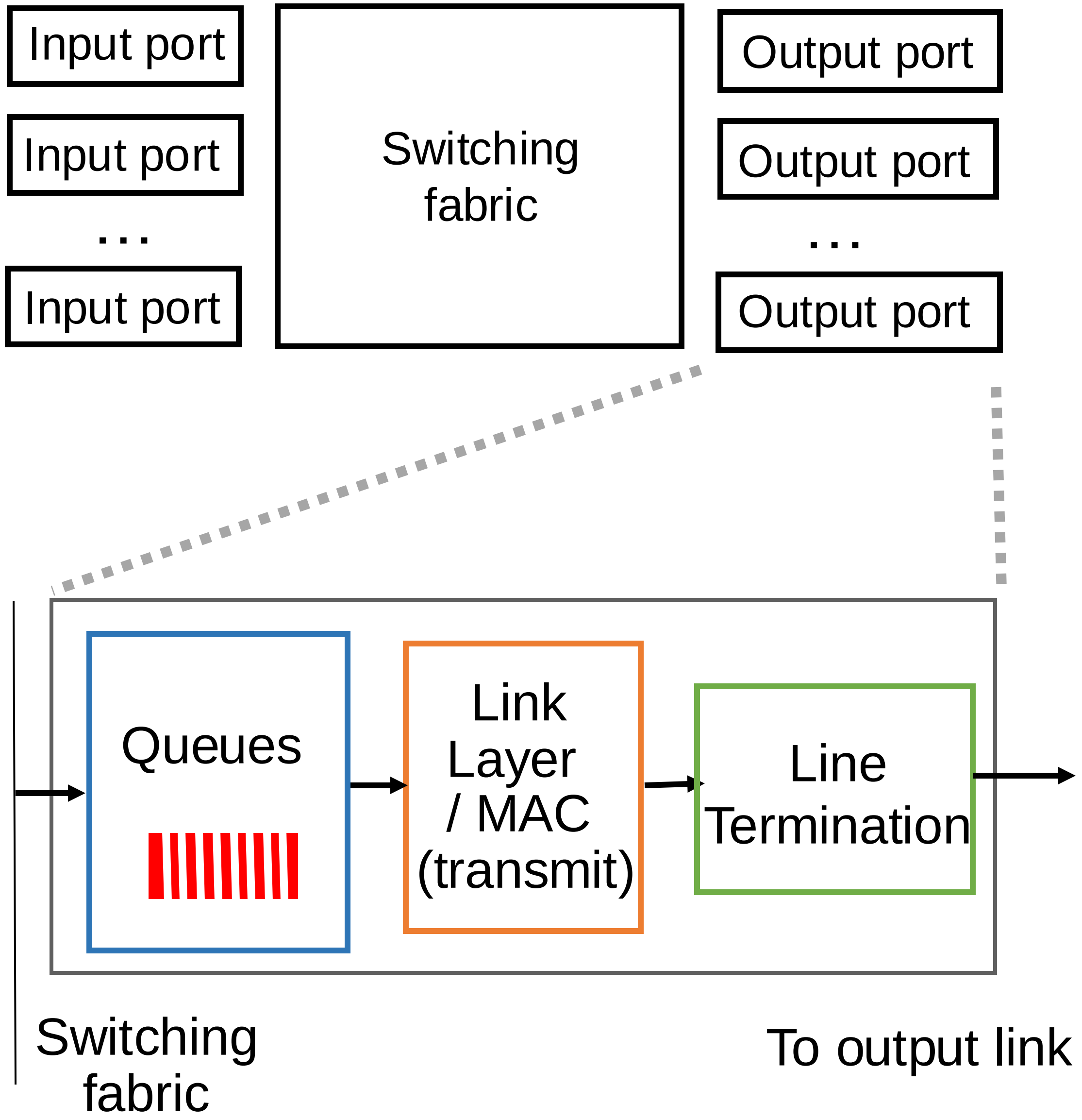
- Forwarding behavior is independent of the source: legitimate source vs. malicious attack traffic
- Forwarding behavior is independent of the application: web traffic vs. file download vs. video
- IP-based packet processing is “baked into” router hardware: evolving the IP protocol faces tall deployment hurdles



Output port functions

- Components in reverse order of those in the input port 

- This is where most routers have the bulk of their packet buffers
  - Recall discussions regarding router buffers from transport
- MGR uses per-port output buffers, but modern routers have shared memory buffers
  - More efficient use of memory under varying demands



Output port functions

- Two important policy decisions
- Scheduling: which among the waiting packets gets to be transmitted out the link?
  - Ex: First-In-First-Out (FIFO)
- Buffer management: which among the packets arriving from the fabric get space in the packet buffer?
  - Ex: Tail drop: later packets dropped first