Management and Operations of Networks, Services, and Systems

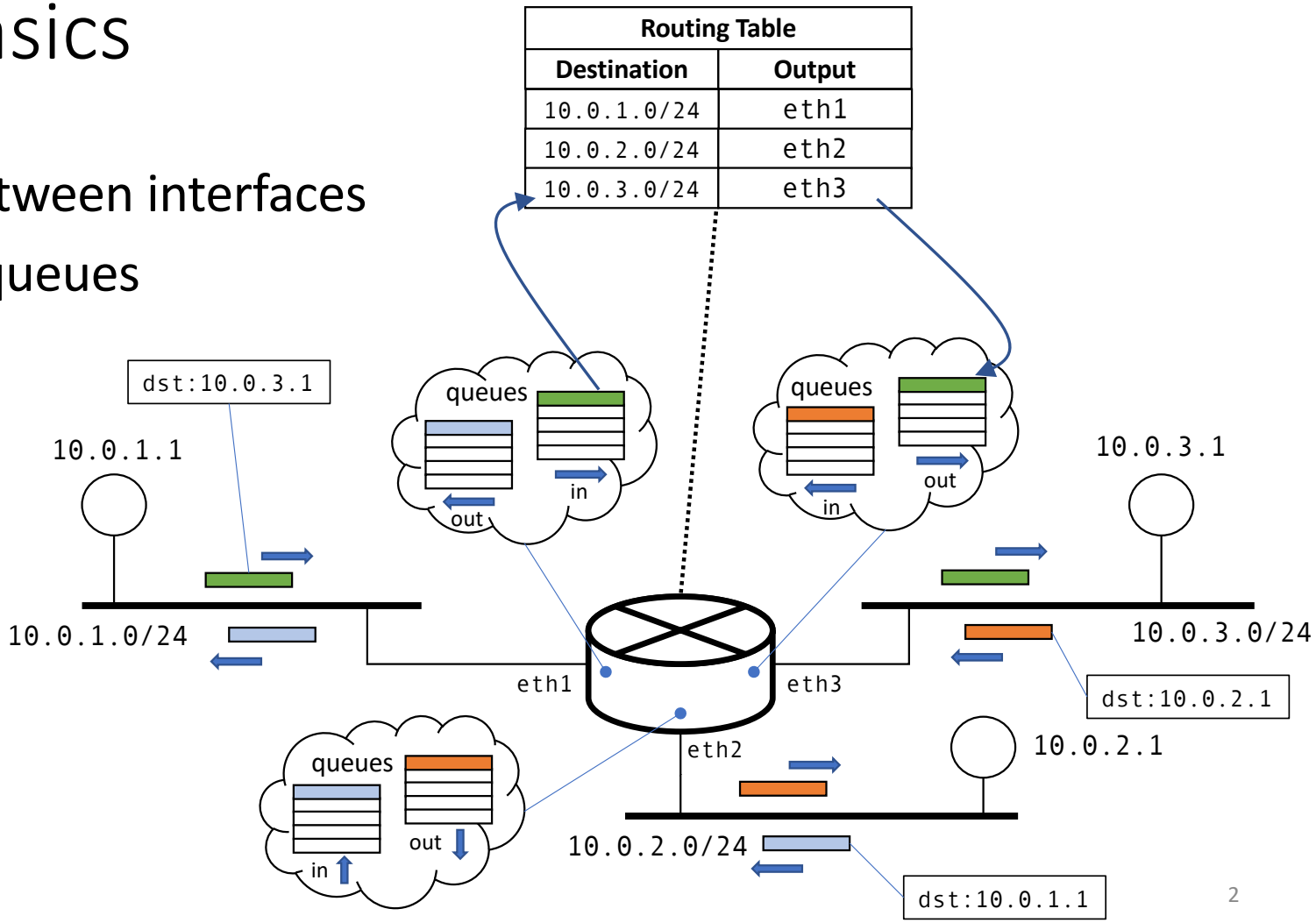
Routing

Ricardo Morla

FEUP – GORS/M.EEC, GRS/M.EIC

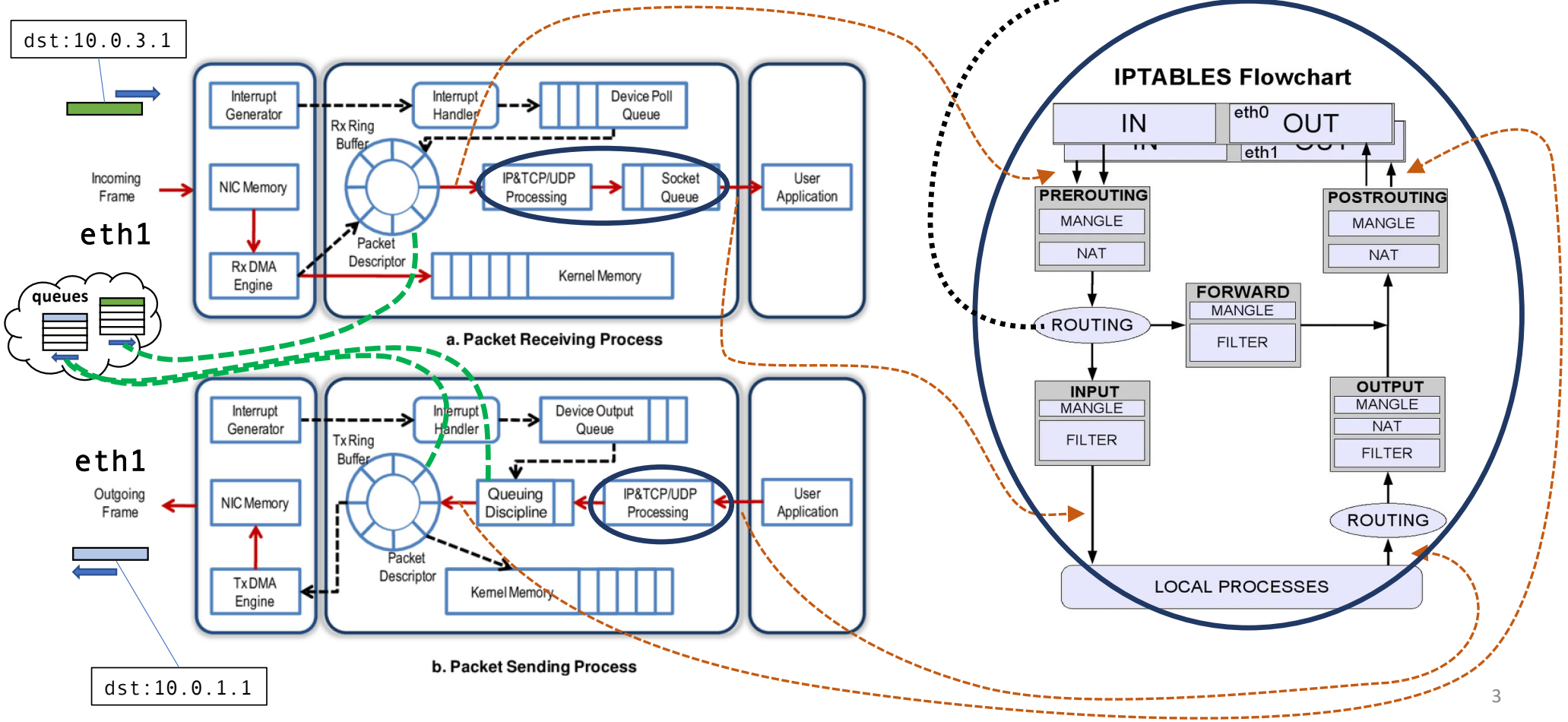
Routing basics

- Forwarding between interfaces
- Input/output queues
- Routing table



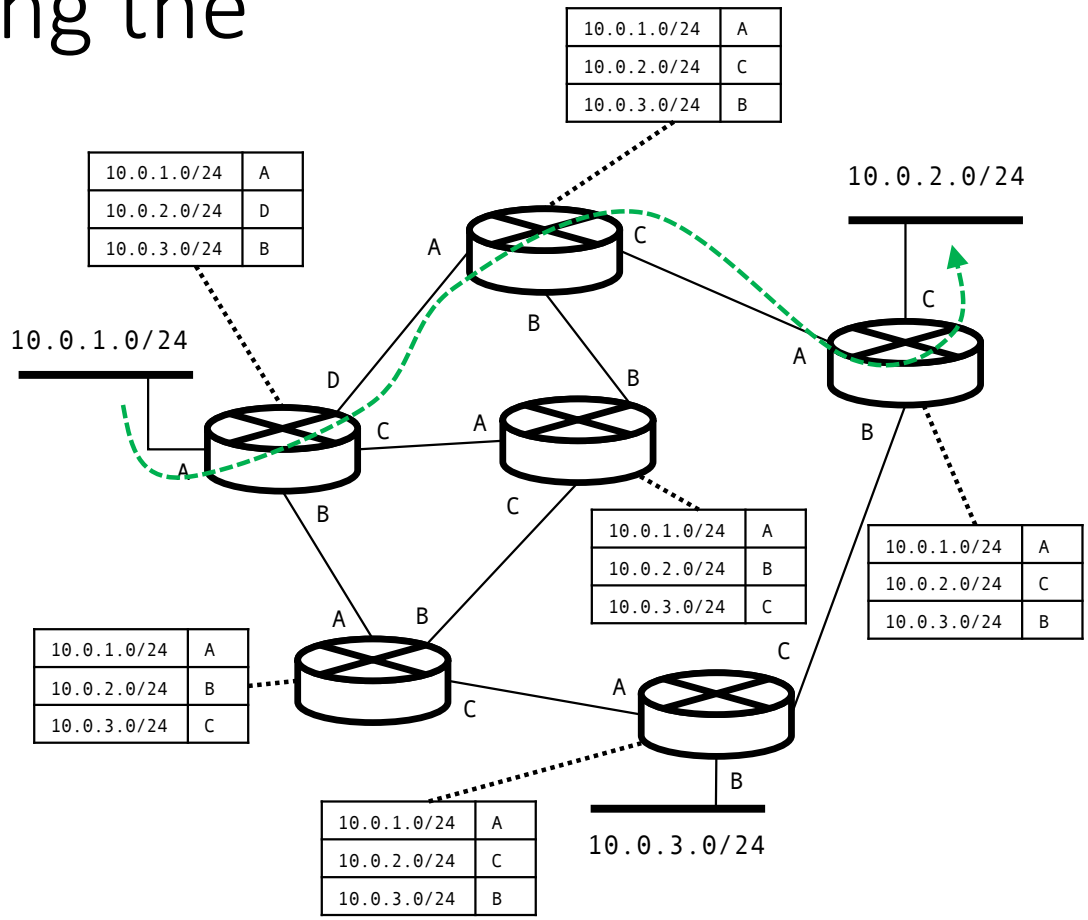
Inside the router

Routing Table	
Destination	**Output**
10.0.1.0/24	eth1
10.0.2.0/24	eth2
10.0.3.0/24	eth3



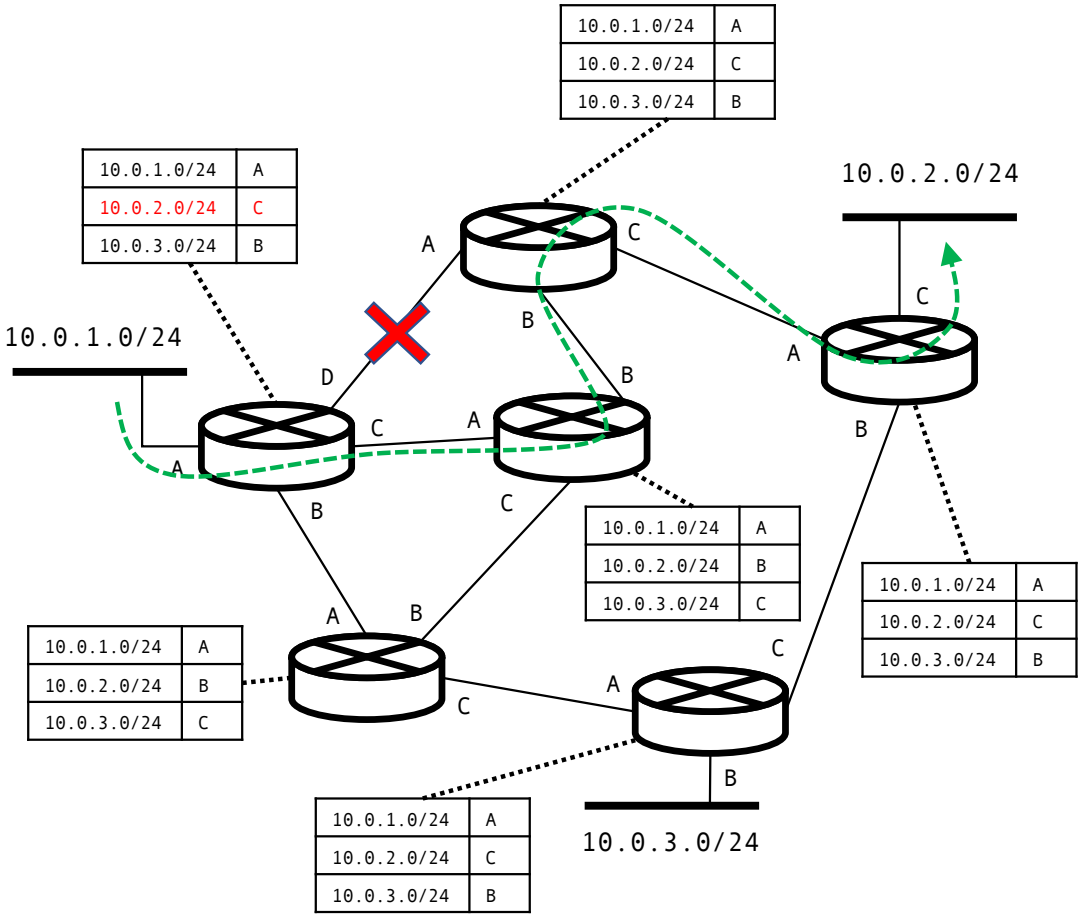
Routing – as in "choosing the paths for packets"

- Each router has its own routing table
- If routing tables are coherent, packets can go from any node in the network to any other node
- If they're not:
  - Loops
  - Blackholes



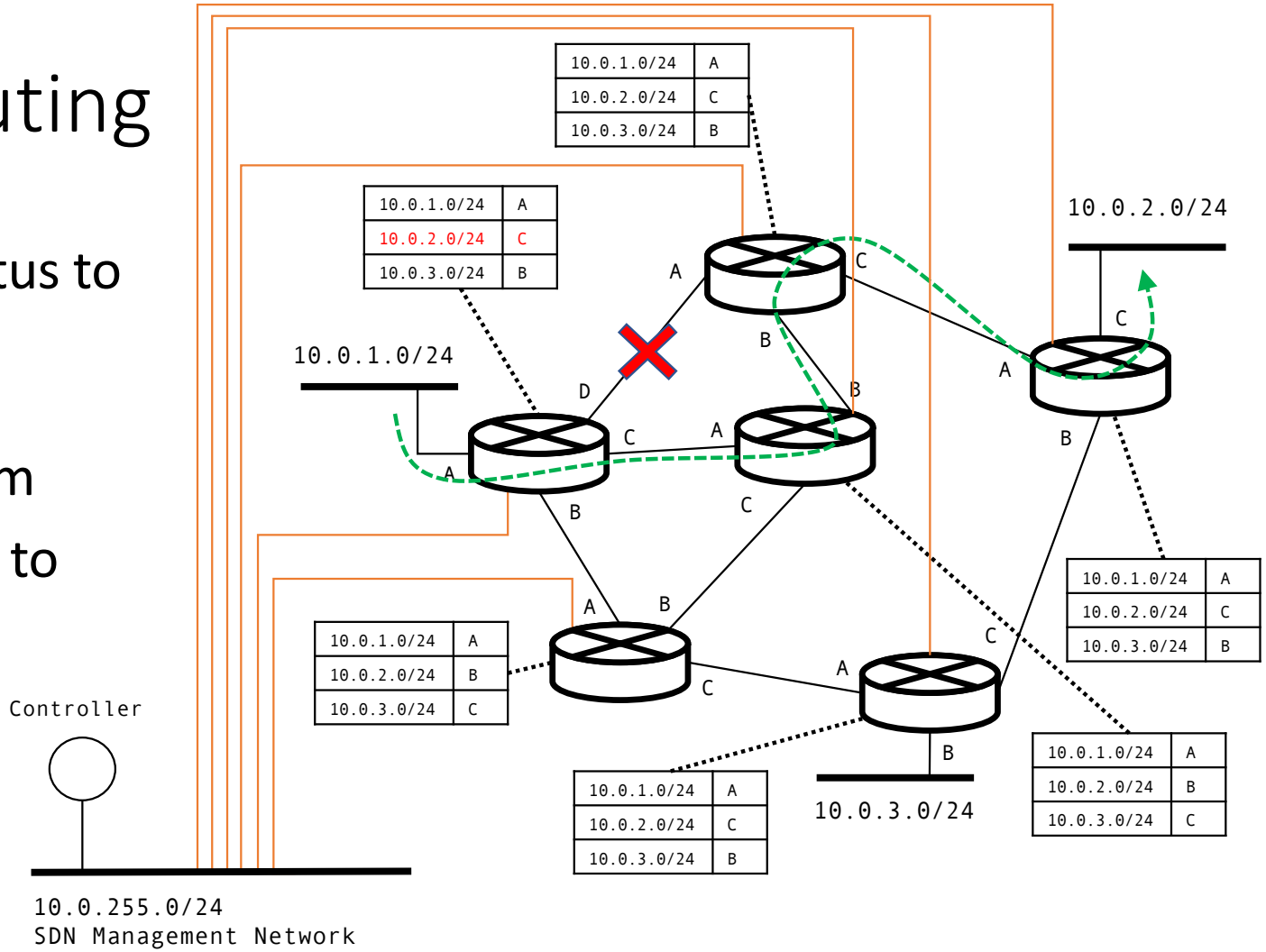
Reacting to changes in topology

- Each router can sense if their links are up or down
  - Or if new networks are added
- Automatically update routing tables
  - New path bypasses down link
- Based on some objective to optimize:
    - Smallest number of hops
    - Smallest delay
    - Largest throughput
    - Balance load on links
    - Administrative distance (…)



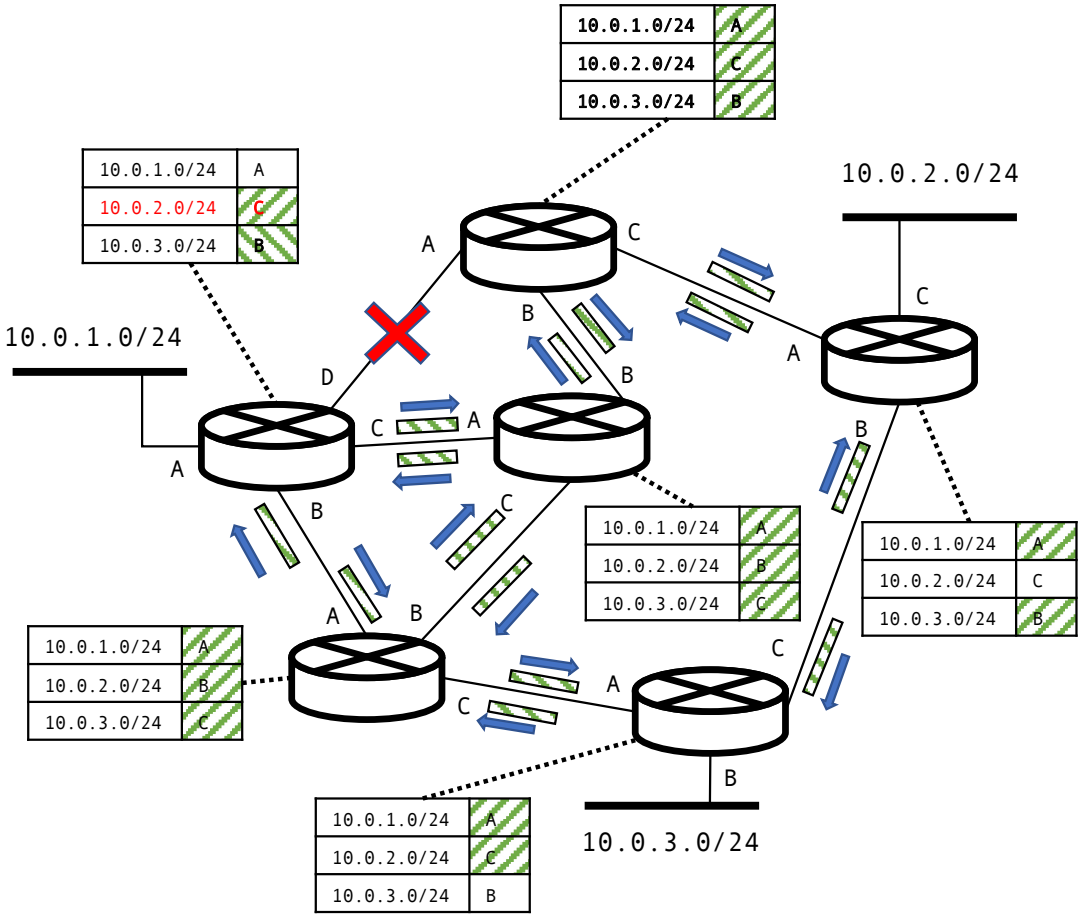
Centralized routing

- Nodes report link status to central location
- Central location runs optimization algorithm
- Returns routing table to each router
- SDN-like solution
  - Data centers



Distributed routing

- Update routing tables dynamically
- According to information from other nodes
  - Each router knows the networks it is directly connected to
  - Routers send **control packets** to other routers
- Geographically dispersed network



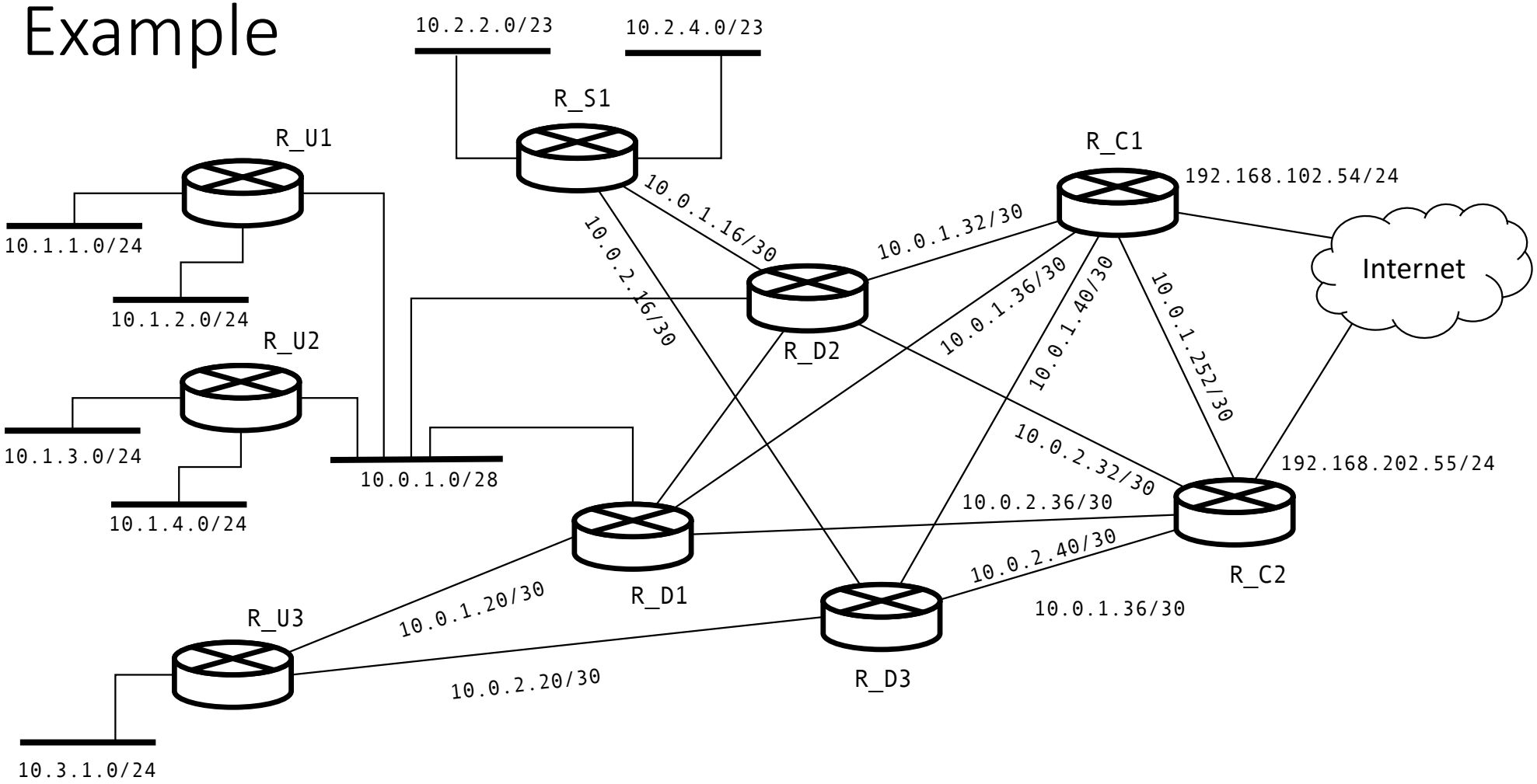
Routing algorithms

- Link state
  - Full topology of the network, then Dijkstra
  - Distributed: flood link state announcements – OSPF
- Distance vector
  - Distance to destination node, Bellman-Ford
  - Distributed: update distances, resend to neighbors
- Path vector
  - List of nodes to destination node, local preferences
  - Distributed: update path vectors, resend to neighbors

Interior routing

- Same administrative domain
  - All networks and devices administred by same entity
  - Typically the internal network of an organization
- Protocols
  - RIP (outdated) – distance vector
  - EIGRP (Cisco) – distance vector, optimized, hybrid
  - OSPF – link state

Example

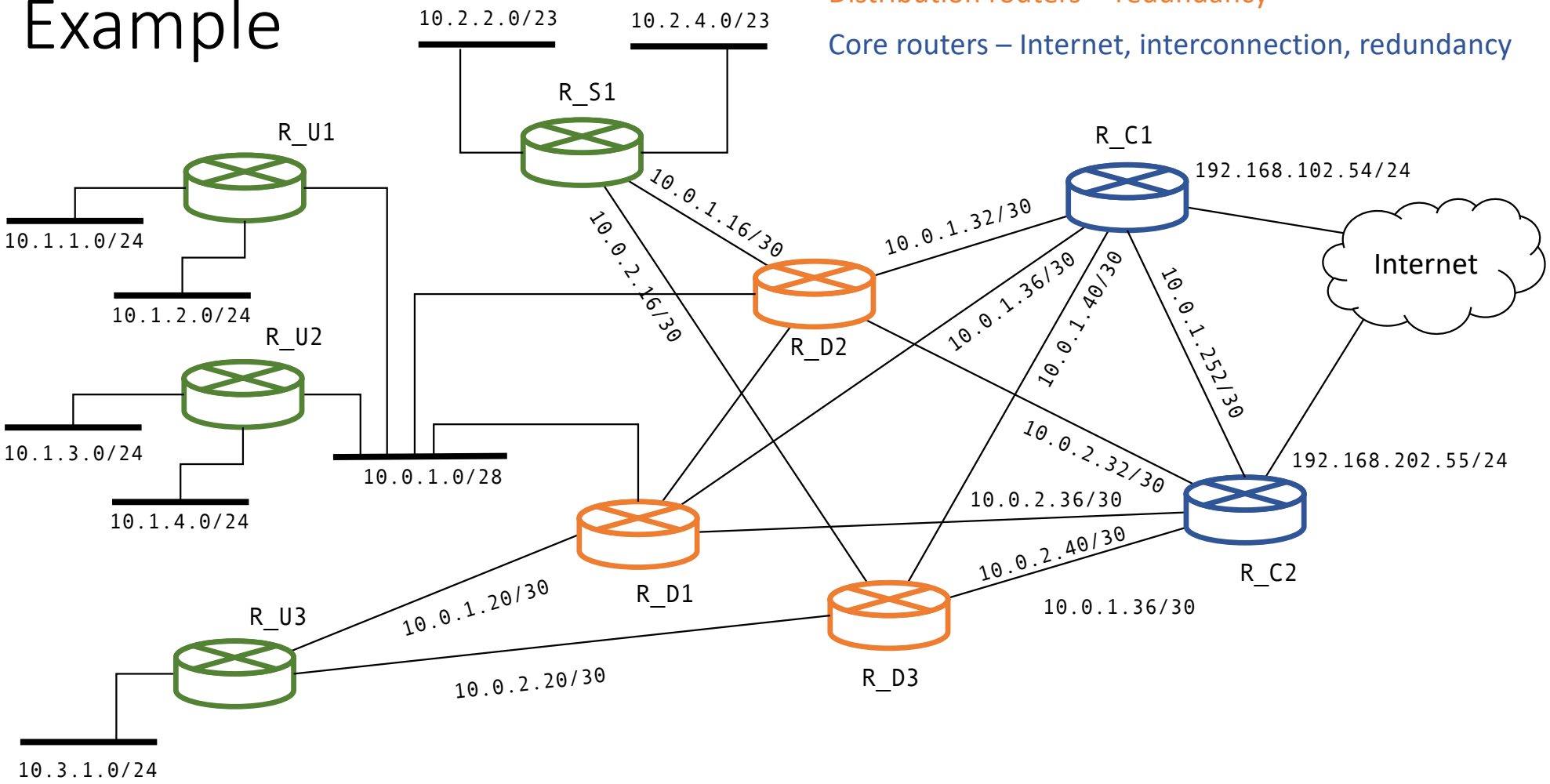

10.2.2.0/23
10.2.4.0/23
R_S1
R_U1
R_C1
192.168.102.54/24
10.1.1.0/24
10.1.2.0/24
10.0.1.16/30
10.0.1.32/30
10.0.2.16/30
Internet
R_D2
R_U2
10.0.1.36/30
10.0.1.40/30
10.0.1.252/30
10.1.3.0/24
10.1.4.0/24
10.0.1.0/28
10.0.2.32/30
192.168.202.55/24
10.0.2.36/30
R_D1
R_C2
10.0.2.40/30
R_U3
10.0.1.20/30
10.0.1.36/30
R_D3
10.0.2.20/30
10.3.1.0/24

Example

Access routers – end user traffic to Internet and servers

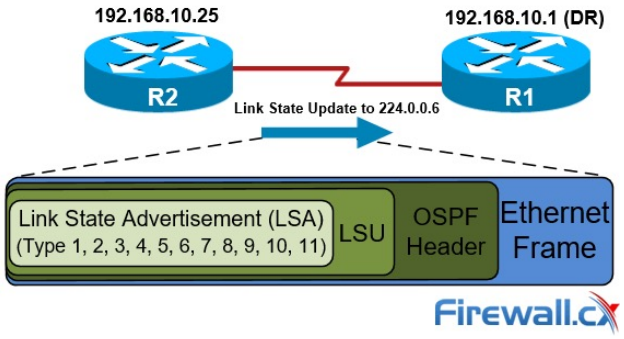
Distribution routers -- redundancy

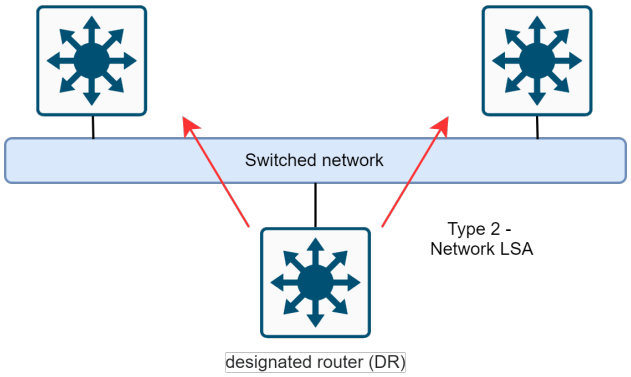
Core routers – Internet, interconnection, redundancy

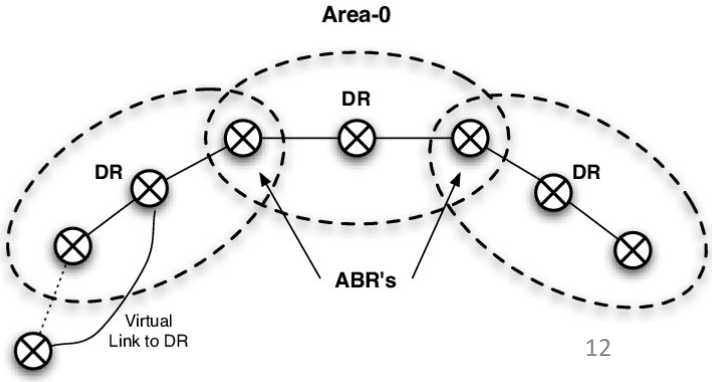


OSPF

- Messages
  - Hello packets between routers to establish adjacency
  - Link state advertisement, link-state database
- Designated routers, network adjacency vs. point-to-point link
- Cost metric manually defined, or multiple of reference bandwidth
- Backbone area, separate link-state database, Area Border Router



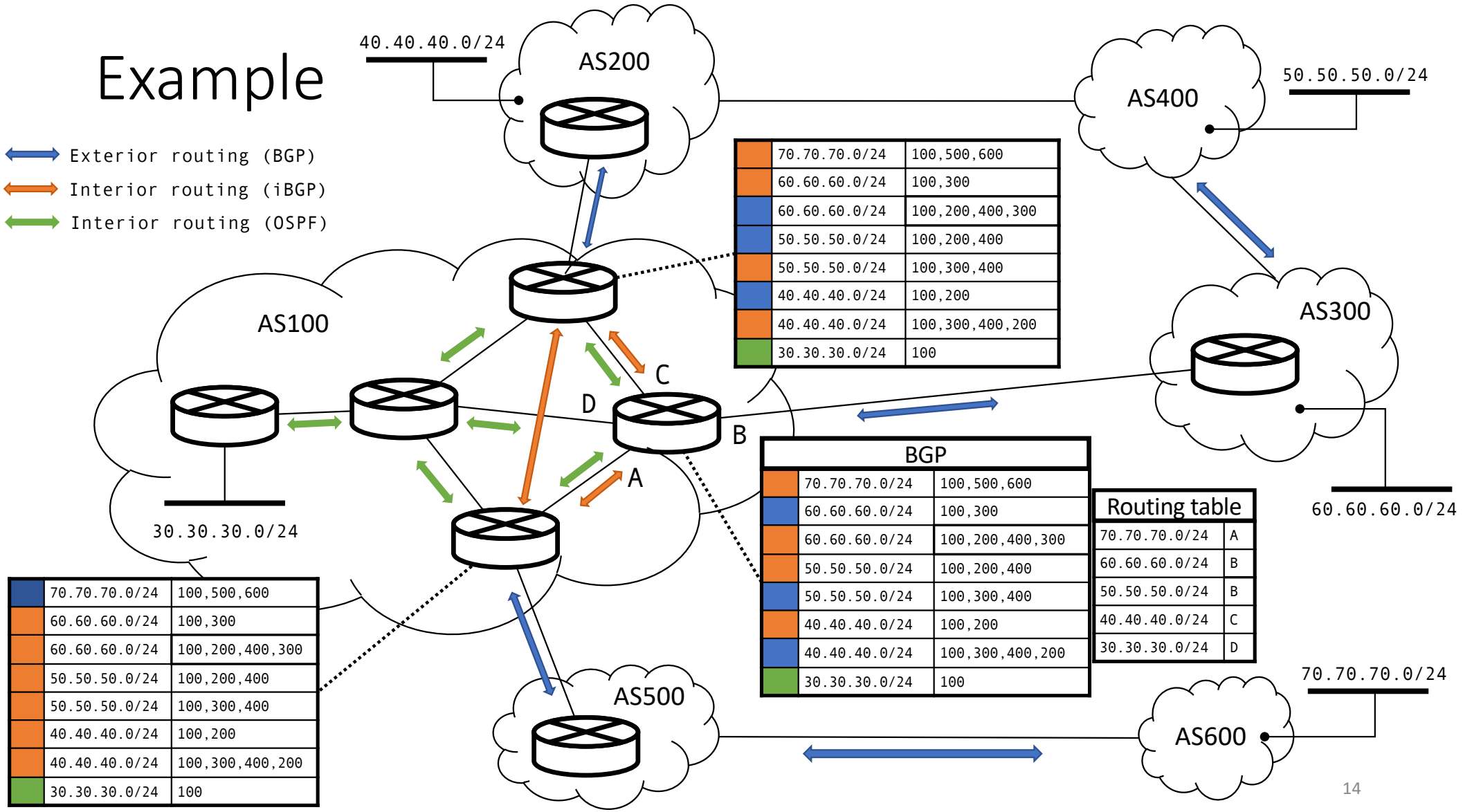




Exterior routing

- Different administrative domains
  - Each node in this network has a different administrator
  - The Internet
  - Each node is an Autonomous System – and a network by itself
- Protocols
  - BGP – path vector

Example


Exterior routing (BGP)
Interior routing (iBGP)
Interior routing (OSPF)
40.40.40.0/24
AS200
AS400
50.50.50.0/24
AS100
AS300
AS500
AS600
30.30.30.0/24
60.60.60.0/24
70.70.70.0/24
A
B
C
D
70.70.70.0/24 100,500,600
60.60.60.0/24 100,300
60.60.60.0/24 100,200,400,300
50.50.50.0/24 100,200,400
50.50.50.0/24 100,300,400
40.40.40.0/24 100,200
40.40.40.0/24 100,300,400,200
30.30.30.0/24 100
BGP
70.70.70.0/24 100,500,600
60.60.60.0/24 100,300
60.60.60.0/24 100,200,400,300
50.50.50.0/24 100,200,400
50.50.50.0/24 100,300,400
40.40.40.0/24 100,200
40.40.40.0/24 100,300,400,200
30.30.30.0/24 100
Routing table
70.70.70.0/24 A
60.60.60.0/24 B
50.50.50.0/24 B
40.40.40.0/24 C
30.30.30.0/24 D

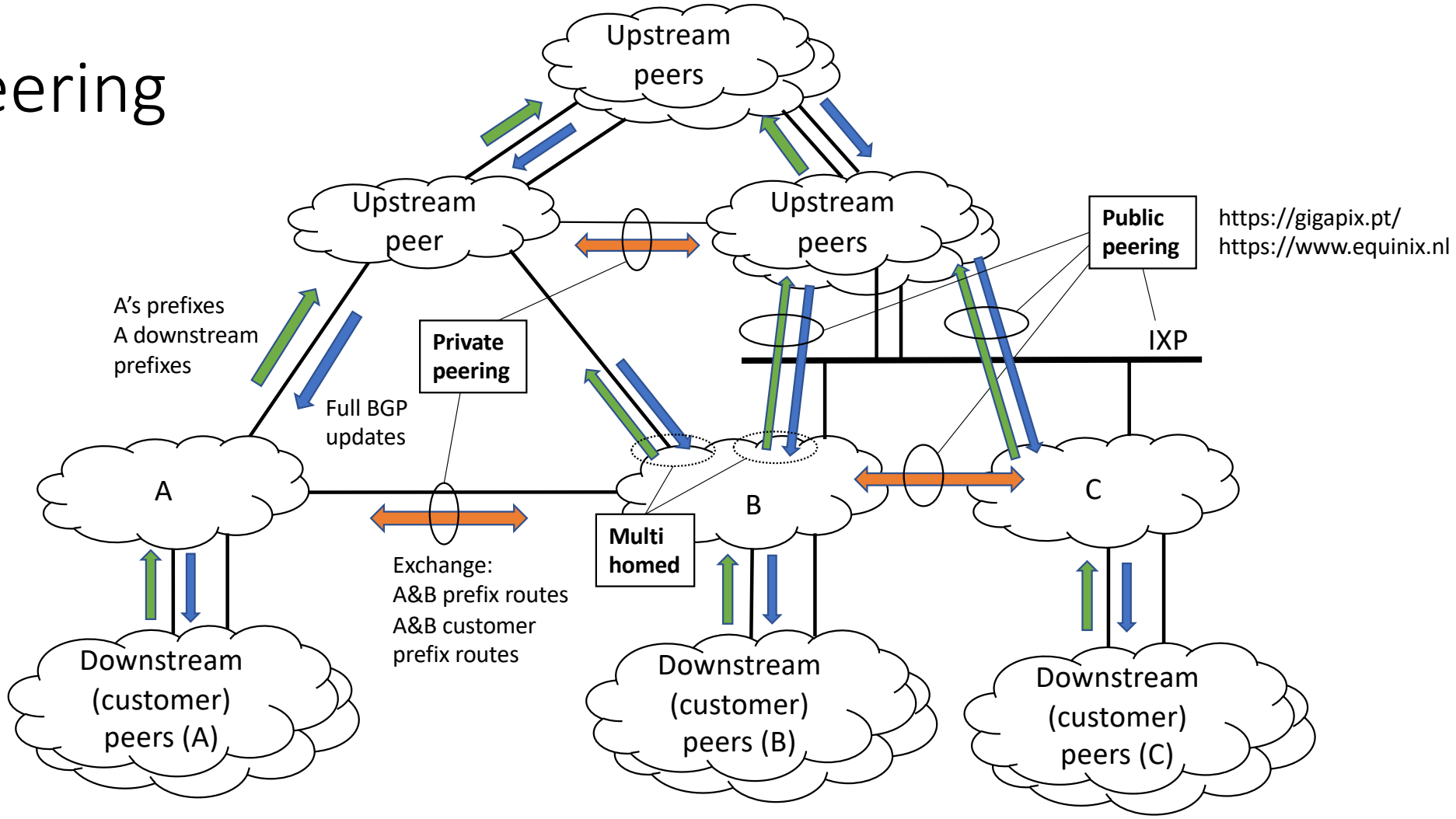
BGP protocol

- RFC 4271, BGP4
- Keep-alive messages, BGP update messages
- Exchange path-vector routes to other BGP routers
  - For each network aka network prefix
  - Select one route, announce it to neighbors, insert it in routing table
- Route map: set of rules to check what to do with route
  - Drop route, modify route and add it to routing table
- Route selection criteria
  - Weight (local to each router); AS-local preference for route (local to each AS), local route; length of AS path; MED multiple exit hint to other AS's, …

BGP protocol

- TCP port 179
- Authentication, RFC 2385
  - MD5 hash of password and BGP data
- IP header TTL set to 1
  - Or more, if BGP connection over multiple IP hops
- Route aggregation
  - 32.32.32.0/24, 32.32.33.0/24 announced as 32.32.32.0/23
- eBGP vs. iBGP
  - e: between routers of different AS
  - i: between routers of same AS in multiple peer AS's, full mesh, route reflectors

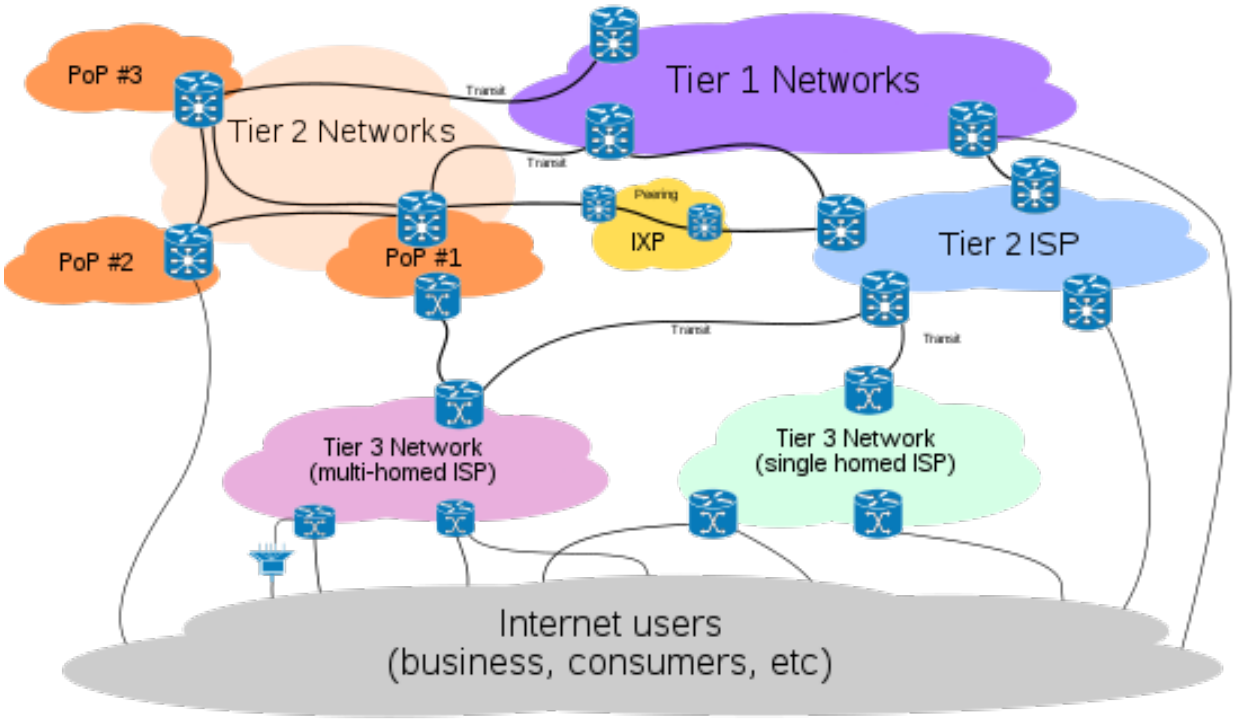
Peering



http://www.bgphelp.com/2017/01/05/bgp-types-of-peering/

Internet Tiers

- Tier 1
  - Settlement-free peering only
- Tier 2
  - Some free, other paying
- Tier 3
  - Paying only peering



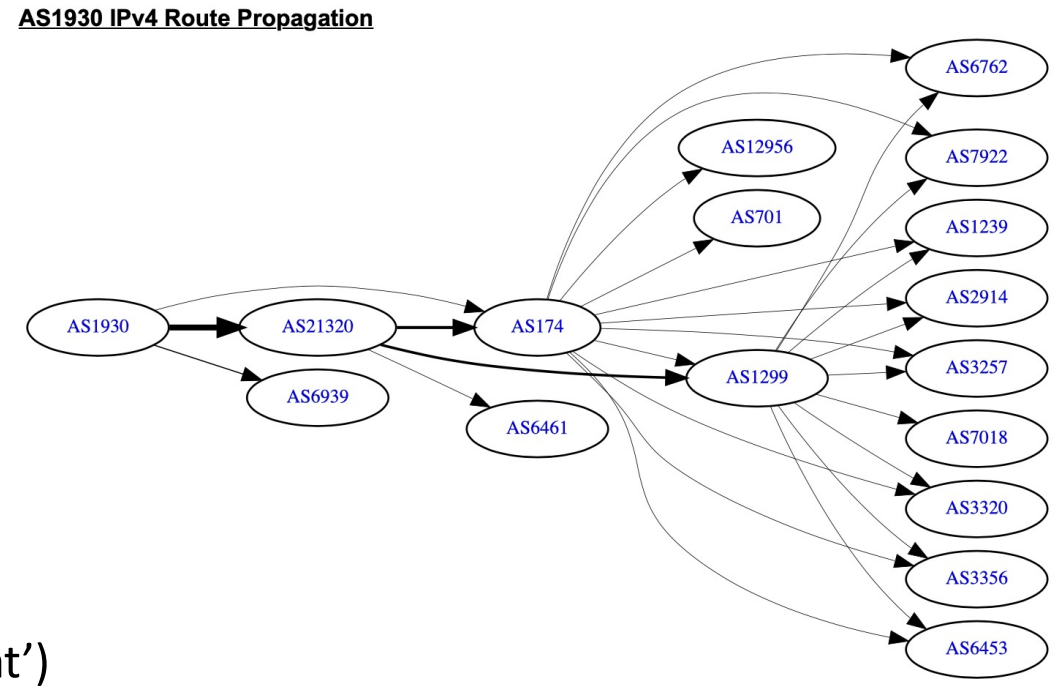
https://en.wikipedia.org/wiki/Tier_1_network

Looking glass

- http://www.routeviews.org
- https://bgp.he.net
- RCCN
  - AS 1930
  - https://bgp.he.net/AS1930
- Archive
  - RIBS and UPDATES
- pyasn

```
import pyasn
asndb = pyasn.pyasn('ipasn_20140513.dat’)
asndb.lookup('8.8.8.8')
```

https://bgp.he.net/AS1930#_graph4



Management and Operations of Networks, Services, and Systems

Routing

Ricardo Morla

FEUP – GORS/M.EEC, GRS/M.EIC