Inter-domain Routing

- Autonomous Systems and the AS graph
- Routing at the edge
- Routing in the core
- BGP

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Inter-domain Routing

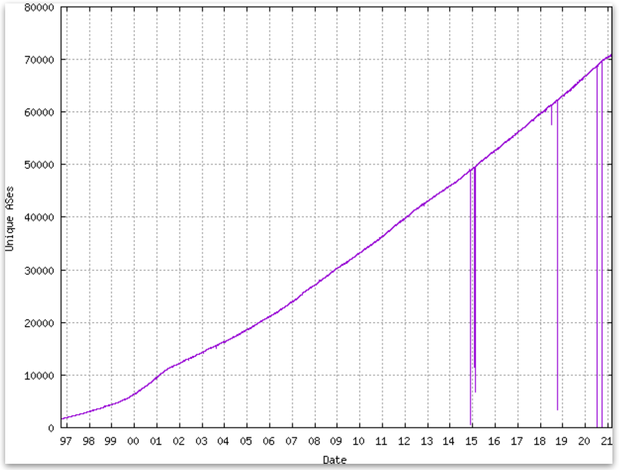
- The Internet is a network of networks
  - Each network is an **Autonomous System (AS)**
  - Each network is a separate routing domain

- **Inter-domain routing** is the problem of finding the best path from the source network to the destination network
  - Treat each network as a node on the routing graph (the “AS topology graph”)
  - Treat connections between networks as edges in the graph
  - What is the best set of networks to choose to get from source to destination

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Autonomous Systems

- An **Autonomous System (AS)** is an independently operated network
  - An Internet service provider, or other organisation, that operates a network and wants to participate in interdomain routing
  - Some organisations operate more than one AS
    - For ease of administration
    - Due to company mergers
    - …
  - Each AS is identified by a unique number, allocated by the RIR
    - 114685 AS numbers allocated as of 6 March 2021; of these 70,926 are advertised in BGP
    - http://www.potaroo.net/tools/asn32/ – current state of allocation of AS numbers
    - http://bgp.potaroo.net/cidr/autnums.html – complete list AS numbers and names, with links to details
    - http://www.cidr-report.org/as2.0/ – classless interdomain routing (CIDR) report

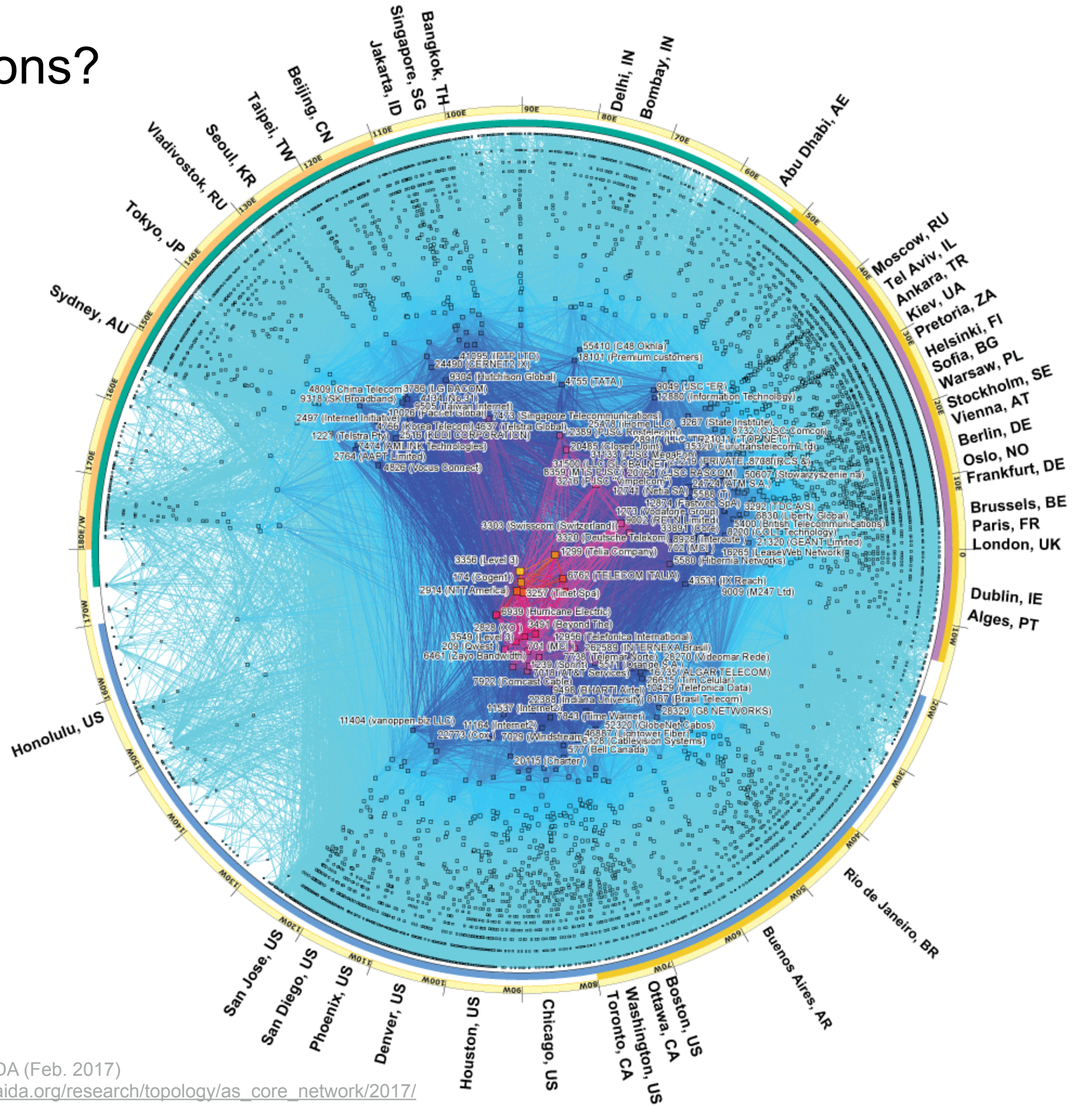


Source: http://www.cidr-report.org/as2.0/

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

AS-level Internet Topology Graph (IPv4)

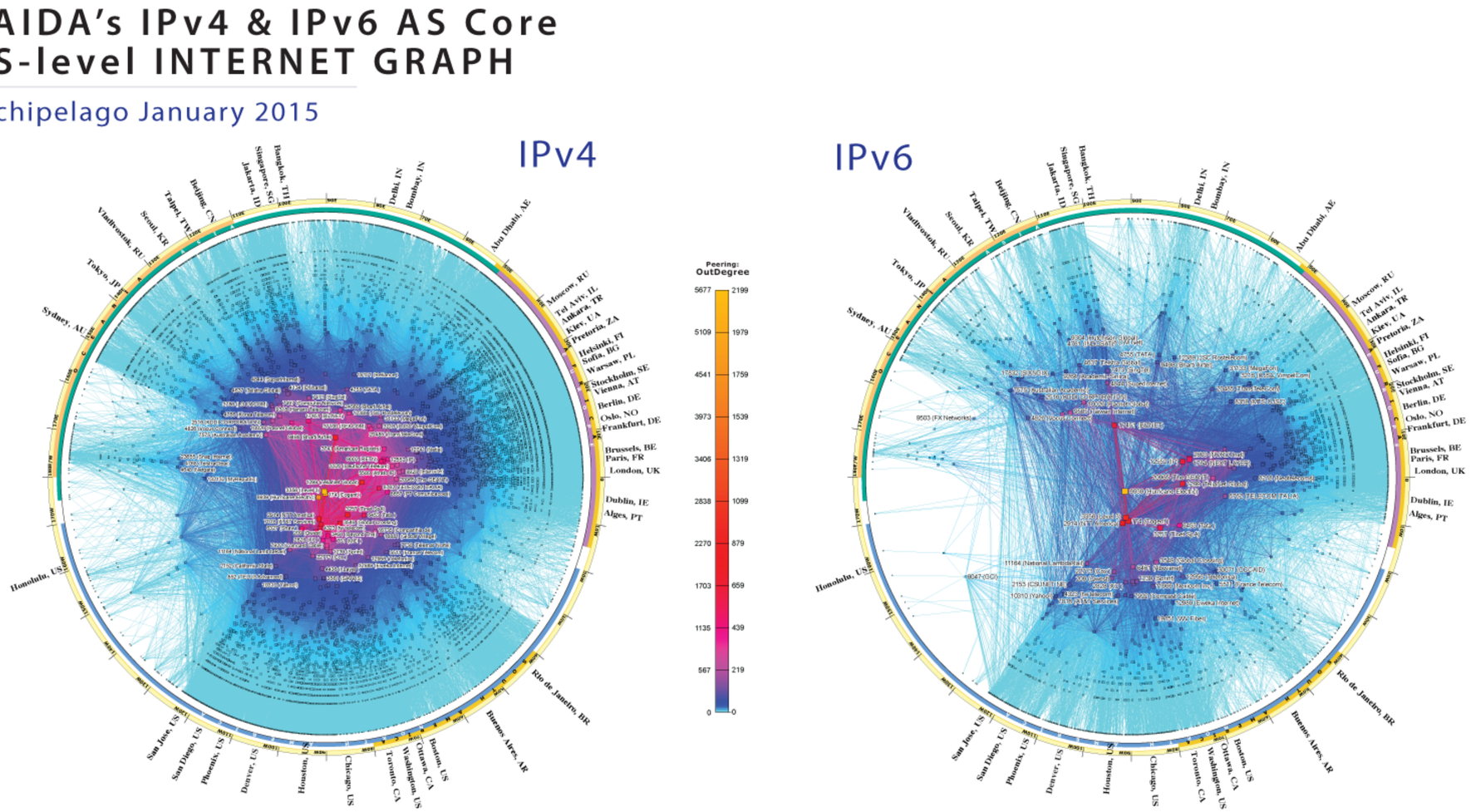
- AS-level topology – what are the inter-AS connections?
  - Each node is an autonomous system
    - Position around circle based on geographic location
    - Distance from centre based on number of links to other networks → more links, closer to the centre
  - Edges show links between autonomous systems
    - A **potential route** traffic can flow
    - Not all possible routes are in use



Source: CAIDA (Feb. 2017)
http://www.caida.org/research/topology/as_core_network/2017/

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

AS-level Internet Topology Graph (IPv4 and IPv6)

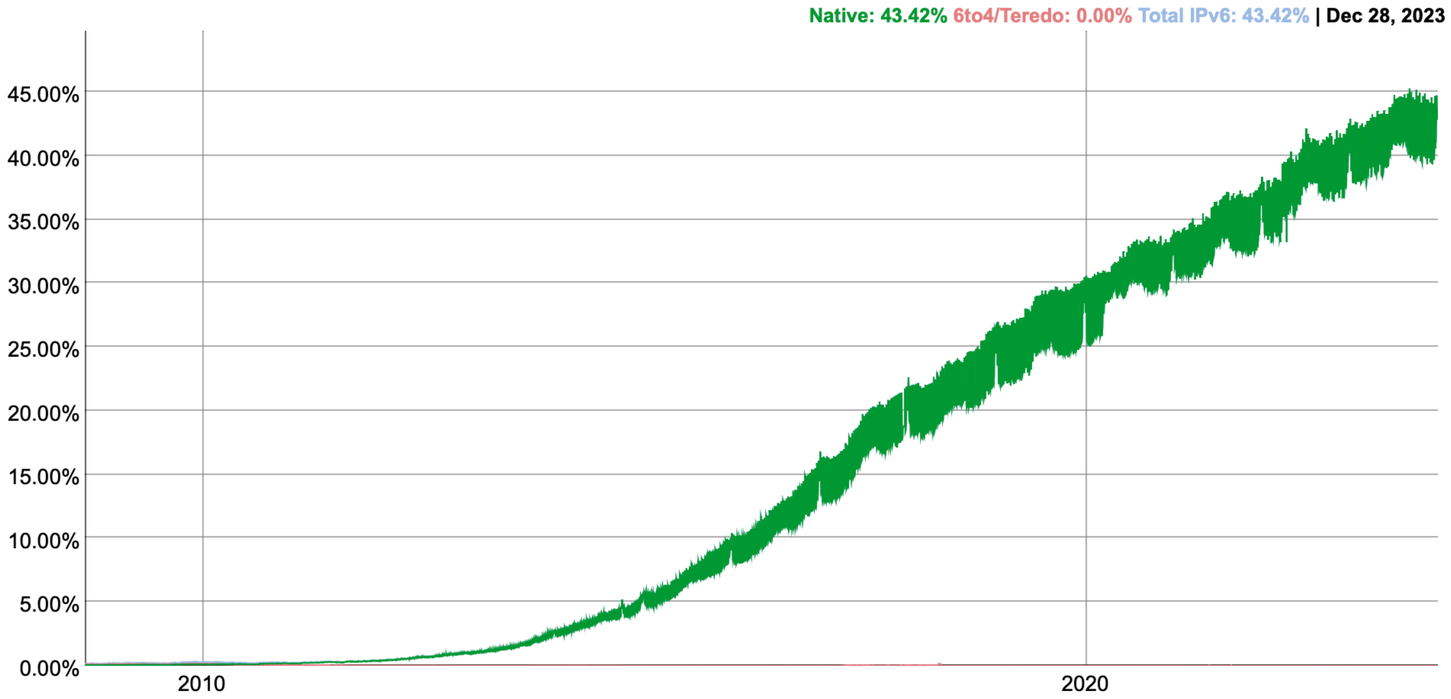


Copyright © 2015 UC Regents. All rights reserved.

Source: http://www.caida.org/research/topology/as_core_network/2015/

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Differences Between IPv4 & IPv6 AS-level Topologies



Source: https://www.google.com/intl/en/ipv6/statistics.html

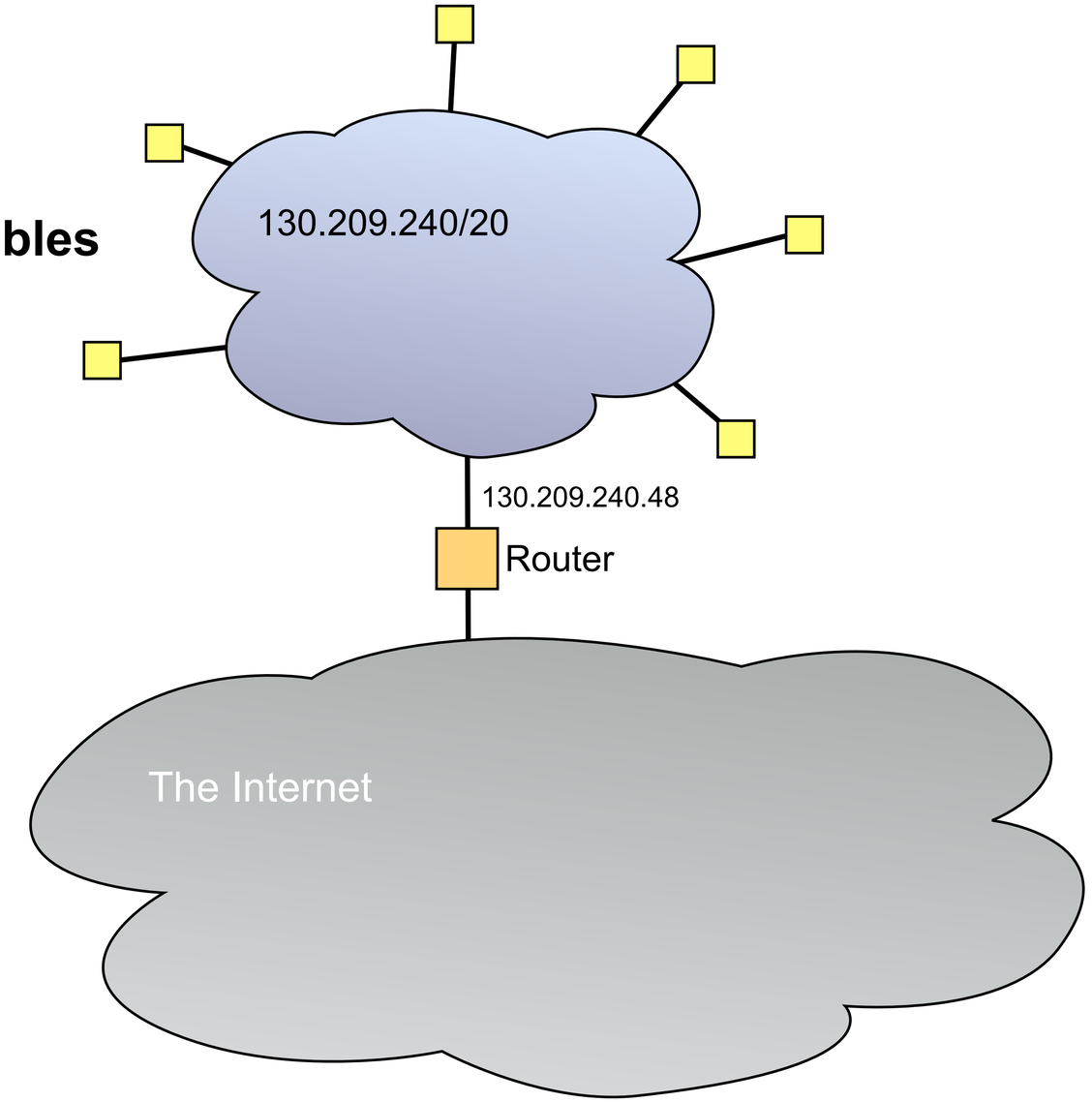
- IPv6 AS-level topology graph is sparse compared to IPv4, but there is significant IPv6 usage → IPv4 has 30 years head start on deployment; the topology can be expected to be more densely interconnected

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Routing at the Edge

- Devices at the edge tend to have simple **routing tables**
  - The devices on the local network
  - A **default route** for everything else
  - e.g., routing table for hosts in Glasgow SoCS:

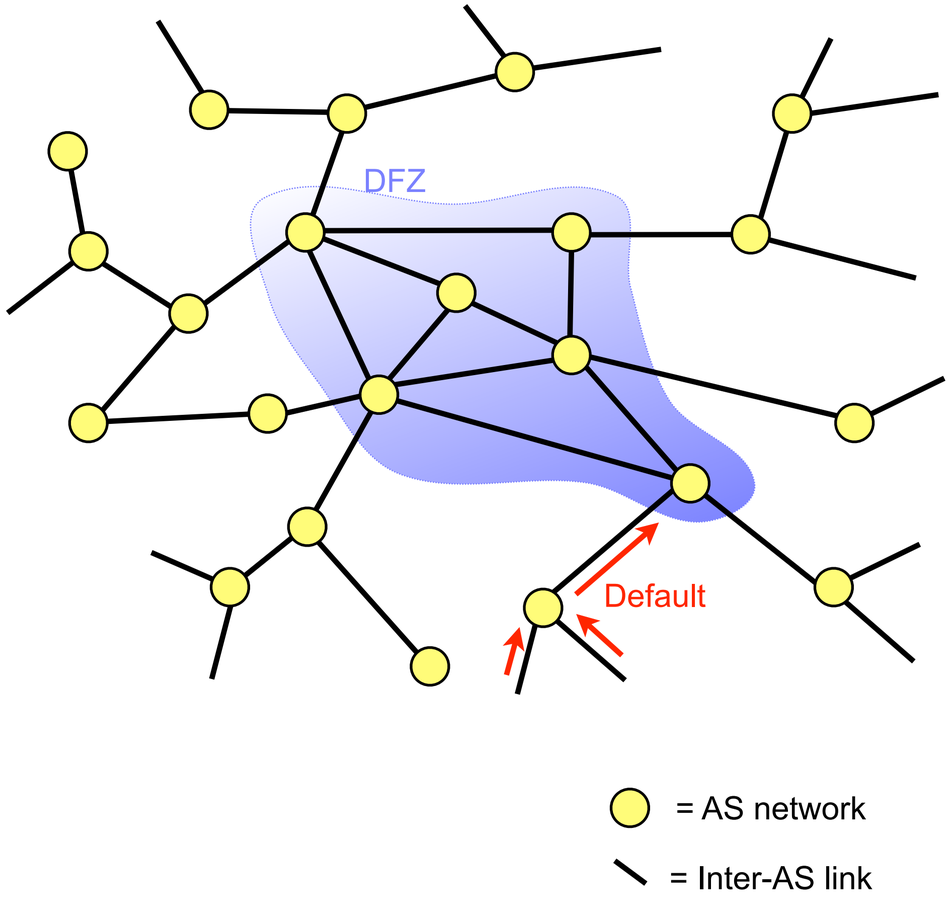
```
Network:        Netmask:        Gateway:
130.209.240.0   255.255.240.0   eth0
default         0.0.0.0         130.209.240.48
```



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

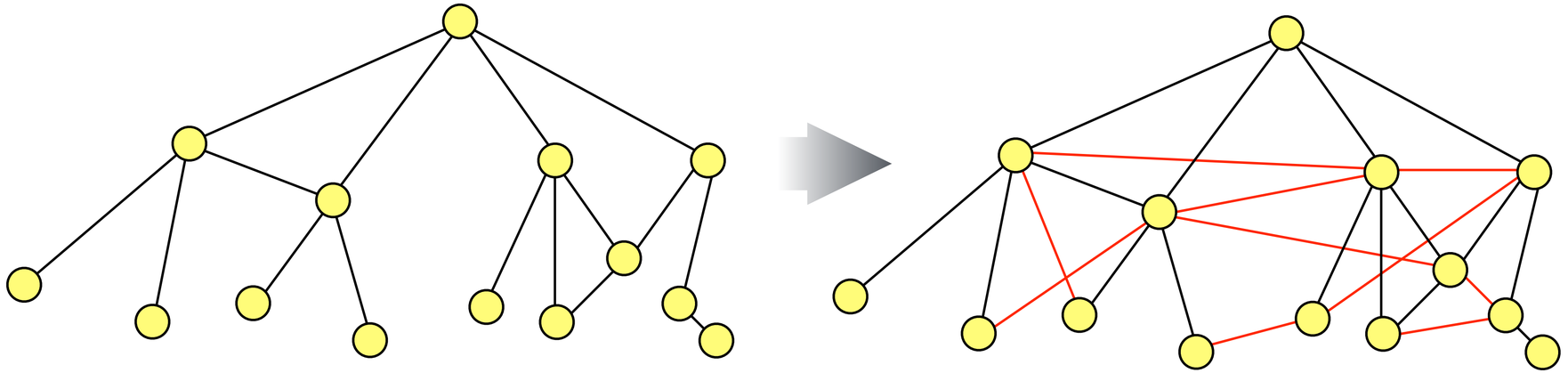
Routing in the Core of the Internet (1/2)

- Networks near the edge can often use a **default route** for much of their traffic
- Core networks need full routing table
  - The **default free zone** (DFZ)
  - With enough interconnections, they must store routing information about the broad network topology
  - Many possible paths – need information to choose the most appropriate



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Routing in the Core of the Internet (2/2)



- AS-level topology flattening → increasingly rich interconnections
- **Internet Exchange Points (IXPs)** are becoming commonplace
  - Example: London Internet Exchange
  - >850 different networks interconnected
  - >4.3 Tbps of traffic



Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Routing Policy

- Interdomain routing is between competitors
  - Each AS is independently operated and may compete for customers
  - Business, political, and economic relationships influence routing
- Routing **must** consider policy
  - Policy restrictions on who can determine the network topology
  - Policy restrictions on which route traffic should follow between a particular source and destination
  - Policy restrictions might prioritise non-shortest path routes
    - Do you prioritise cost, bandwidth, or latency when choosing a route?
    - Traffic between certain networks may be prohibiting from passing through other networks
    - Traffic between certain regions should avoid other regions

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Border Gateway Protocol (BGP)

- Internet routing uses the **Border Gateway Protocol (BGP)**
  - External BGP (**eBGP**) used to exchange routing information between ASes
    - An eBGP session is used by an AS to exchange information about the AS-level topology with a neighbouring AS
    - Runs over a TCP connection between two routers, one in each AS
    - Exchanges knowledge of the AS-level topology of the network, filtered according to policy
    - Used to compute appropriate **interdomain routes** to each destination AS
  - Internal BGP (**iBGP**) propagates this information to routers within an AS
    - An **intra-domain routing protocol** determines routes within the AS
    - iBGP distributes information to internal routers on how to reach external destinations

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Routing Information Exchanged in eBGP

- eBGP routers advertise lists of IP address ranges (“prefixes”) and their associated AS-level paths
- Combined to form a routing table

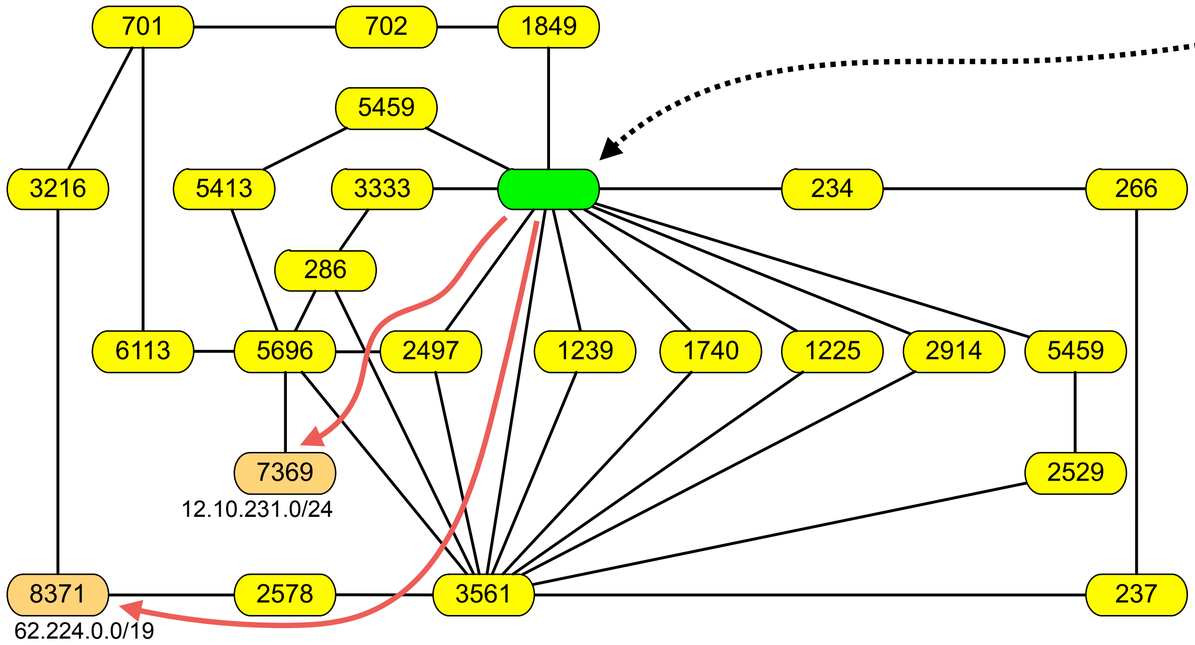
	Prefix	Next Hop	AS Path
…			
*	12.10.231.0/24	194.68.130.254	5459 5413 5696 7369 i
*		158.43.133.48	1849 702 701 6113 5696 7369 i
*		193.0.0.242	3333 286 5696 7369 i
*		204.212.44.128	234 266 237 3561 5696 7369 i
*>		202.232.1.8	2497 5696 7369 i
*		204.70.4.89	3561 5696 7369 i
*		131.103.20.49	1225 3561 5696 7369 i
*	62.224.0.0/19	134.24.127.3	1740 3561 2578 8371 i
*		194.68.130.254	5459 2529 3561 2578 8371 i
*		158.43.133.48	1849 702 701 3216 3216 3216 8371 8371 i
*		193.0.0.242	3333 286 3561 2578 8371 i
*		144.228.240.93	1239 3561 2578 8371 i
*		204.212.44.128	234 266 237 3561 2578 8371 i
*		202.232.1.8	2497 3561 2578 8371 i
*		205.238.48.3	2914 3561 2578 8371 i
*>		204.70.4.89	3561 2578 8371 i
*		131.103.20.49	1225 3561 2578 8371 i
…			

- Hosts with IP addresses in the range 12.10.231.0 - 12.10.231.255 are in AS 7369.
- That AS is best reached via AS 2497, then AS 5696.
- Packets destined for those addresses should be sent to address 202.232.1.8 next, from where they will be forwarded.

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

AS Topology Graph

An example fragment of the AS topology graph:

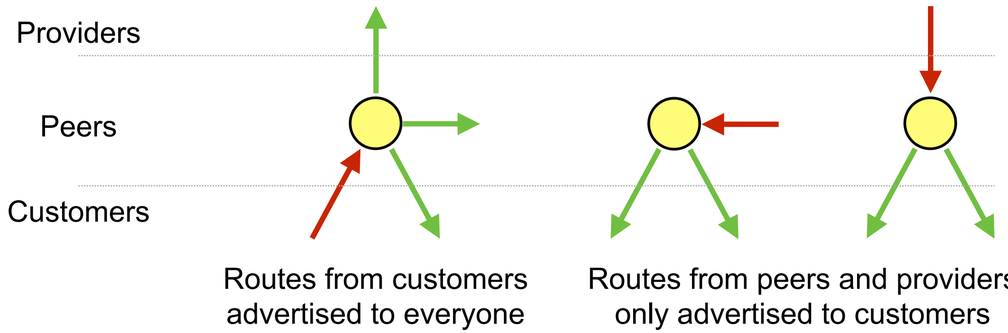


	Prefix	Next Hop	AS Path
…			
*	12.10.231.0/24	194.68.130.254	5459 5413 5696 7369 i
*		158.43.133.48	1849 702 701 6113 5696 7369 i
*		193.0.0.242	3333 286 5696 7369 i
*		204.212.44.128	234 266 237 3561 5696 7369 i
*>		202.232.1.8	2497 5696 7369 i
*		204.70.4.89	3561 5696 7369 i
*		131.103.20.49	1225 3561 5696 7369 i
*	62.224.0.0/19	134.24.127.3	1740 3561 2578 8371 i
*		194.68.130.254	5459 2529 3561 2578 8371 i
*		158.43.133.48	1849 702 701 3216 3216 3216 8371 8371 i
*		193.0.0.242	3333 286 3561 2578 8371 i
*		144.228.240.93	1239 3561 2578 8371 i
*		204.212.44.128	234 266 237 3561 2578 8371 i
*		202.232.1.8	2497 3561 2578 8371 i
*		205.238.48.3	2914 3561 2578 8371 i
*>		204.70.4.89	3561 2578 8371 i
*		131.103.20.49	1225 3561 2578 8371 i
…			

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Routing Policy in eBGP

- Each AS chooses what routes to advertise to its neighbours
- Doesn't need to advertise everything it receives
  - Usual to drop some routes from the advertisement – depends on the chosen routing policy
  - Common approach: the Gao-Rexford rules:




Lixin Gao


Jennifer Rexford

- Gao-Rexford rules specify **what routes are advertised**, not how traffic is forwarded – difference between **control plane** and **data plane**
- Resulting graph is valley-free DAG: traffic flows up to common provider, then down
- Designed to optimise **profit**, not routing efficiency – don't advertise expensive routes

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

BGP Routing Decision Process (1/2)

- Each AS exchanges routing information with neighbours
  - AS-level topology information, filtered based on policies applied by each AS
  - Gives a partial view of AS-level network topology
- Each AS applies policy to choose what routes to use:
  - Choose what route to install in forwarding table for each destination prefix
  - BGP may not find a route, even if one exists, since may be prohibited by policy
  - Routes are often not the shortest AS path – shortest policy-compliant path
- Each AS applies policy to determine what routes to advertise to neighbours
  - Filter out routes the neighbour should not use (e.g., Gao-Rexford rules, and other reasons)
  - De-prioritise certain other routes
  - Tag routes for special processing where business agreements exist
  - Mapping business goals to BGP policies is poorly documented → many operational secrets

Table 2: Simplified BGP decision process [6, 24]. This table was also provided with the survey.

#	Criteria
1	Highest LocalPref
2	Lowest AS Path Length
3	Lowest origin type
4	Lowest MED
5	eBGP-learned over iBGP-learned
6	Lowest IGP cost to border router (hot-potato routing)
7	If both paths are external, prefer the path that was received first (i.e., the oldest path) [6]
8	Lowest router ID (to break ties)

Source: Phillipa Gill, Michael Schapira, and Sharon Goldberg, "A Survey of Interdomain Routing Policies", ACM CCR, V44, N1, January 2014, p29-34

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

BGP Routing Decision Process (2/2)

- BGP exchanges routes between autonomous systems
  - Policy, economic, and political concerns outweigh shortest path in many cases
  - Routing is between competitors – there is little trust and little public data
  - The BGP decision process is public – the input data is not, making it difficult to evaluate how routing decisions are made

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow

Inter-domain Routing

- Autonomous Systems and the AS graph
- Routing at the edge
- Routing in the core
- BGP

Colin Perkins | https://csperkins.org/ | Copyright © 2020 University of Glasgow