CPSC 317 COMPUTER NETWORKING

Module 5: Network Layer – Day 6 – Inter-domain Routing

ADMINISTRATION

- Quiz 3 grades will be released soon (this week)
- In my other role as a researcher, I will be attending a seminar on "Web Application Security" in Japan (Mar 18-24)
- Lectures for next week will be covered by Prof. Ivan Beschastnikh
  - Live and on zoom in section 202 (12 noon)
  - Recording in section 201 (3pm)
- ICAs as usual
- PA 4 will be due as usual on Mar 24
- Quiz 4 will be as usual next week

LEARNING GOALS

Inter-domain Routing

- Explain how routing decisions are made from the perspective of the AS
- Understand the terminology related to iBGP, eBGP, peering, transit, border
- List the types of information exchanged by eBGP.
- Given multiple routes to a destination, enumerate the factors that go into the router's decision to route a particular way.
- Explain, using hot potato routing, how a packet is forwarded from a router in one AS to its destination in another AS.
- Explain the issues with BGP and mitigations.

Software Defined Networking

- Explain where control plane is run in SDN

READING

- Reading: 5.4 (5.4.2)

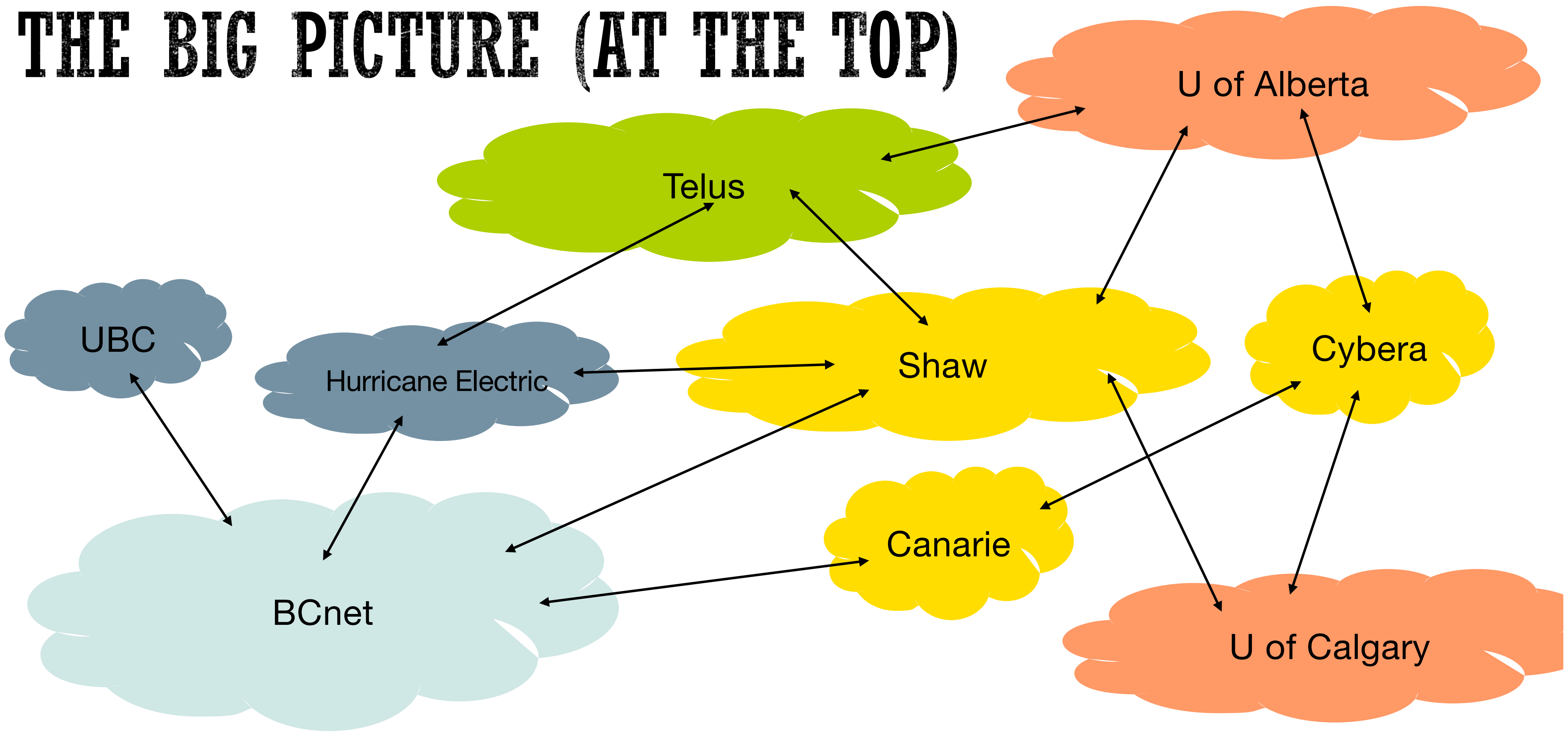
INTERIOR ROUTING PROTOCOLS

- Do not scale
- Do not account for administrative differences (administrative autonomy)
    - Political
    - Company
    - International boundaries
- Don’t allow policy to play a role

INTER-DOMAIN ROUTING

- The Internet is organized into ASes (autonomous systems)
- Each AS is responsible for some collection of IP addresses
- Each AS must “tell” other ASes
  - which addresses it “owns”
  - which addresses it is willing to route to

THE BIG PICTURE (AT THE TOP)



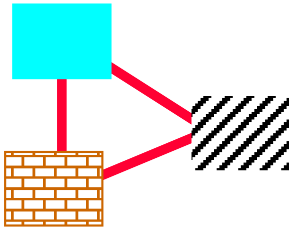
PEERING AND TRANSIT

- Peering
  - Two ISPs pass traffic between each other for their “customers”
- Transit
  - Passing traffic across an AS to get to a different AS
- Stub ASes
  - Have a single provider (or two) as their Internet Gateway

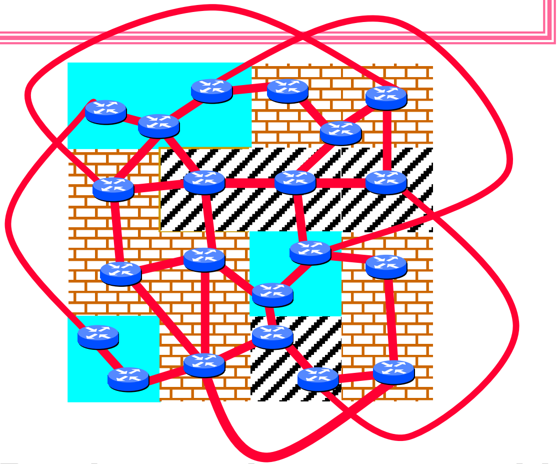
AS VIEW

… the administration of an AS appears to other ASes to have a single coherent interior routing plan and presents a consistent picture of what networks are reachable through it.

RFC 1930: Guidelines for creation, selection, and registration of an Autonomous System



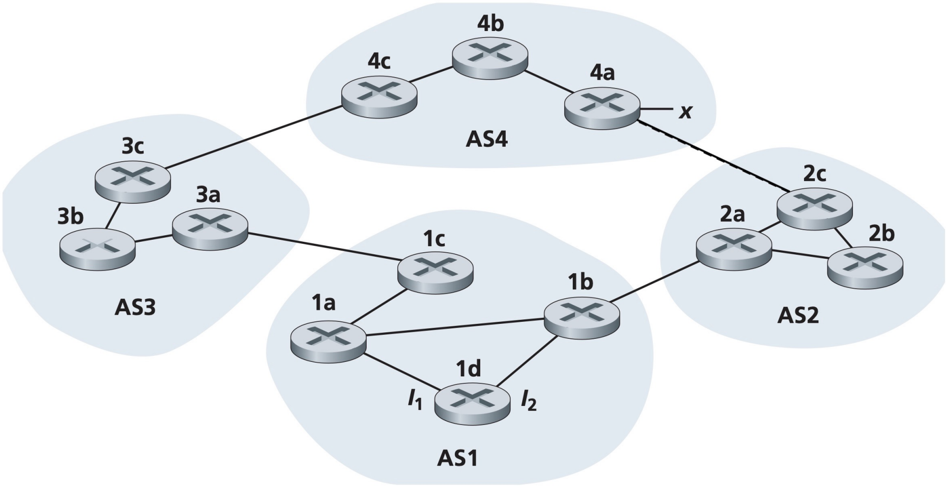
The AS graph may look like this.



Reality may be closer to this…

THE (MINI) INTERNET

- ASes are interconnected with routers
- Internal routers: to route traffic in its own network (IGP)
- Border routers: to route traffic at the edge of ASes
- Edge routers: connect a home or company network to the rest of the AS (and the world)



INTER-AS ROUTING REQUIREMENTS

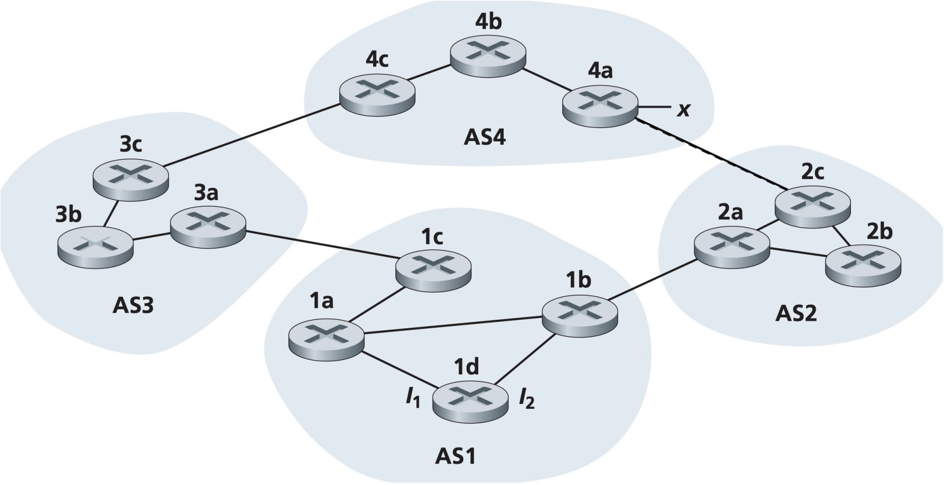
- Discover reachable subnets (network prefixes) from neighboring ASes
- Determine best routes to the reachable subnets (prefixes)

EXTERIOR GATEWAY PROTOCOL (EGP)

- BGP (Border Gateway Protocol)
- Executed by border routers (routers at the border between ASes)
- Forms the backbone of the Internet, along with IP
- iBGP and eBGP
  - Both are application layer protocols, built on top of TCP!
  - BGP connection or session between routers running BGP
- Between ASes it uses aggregation or summarization to reduce table sizes.

OBTAINING PREFIX REACHABILITY INFORMATION

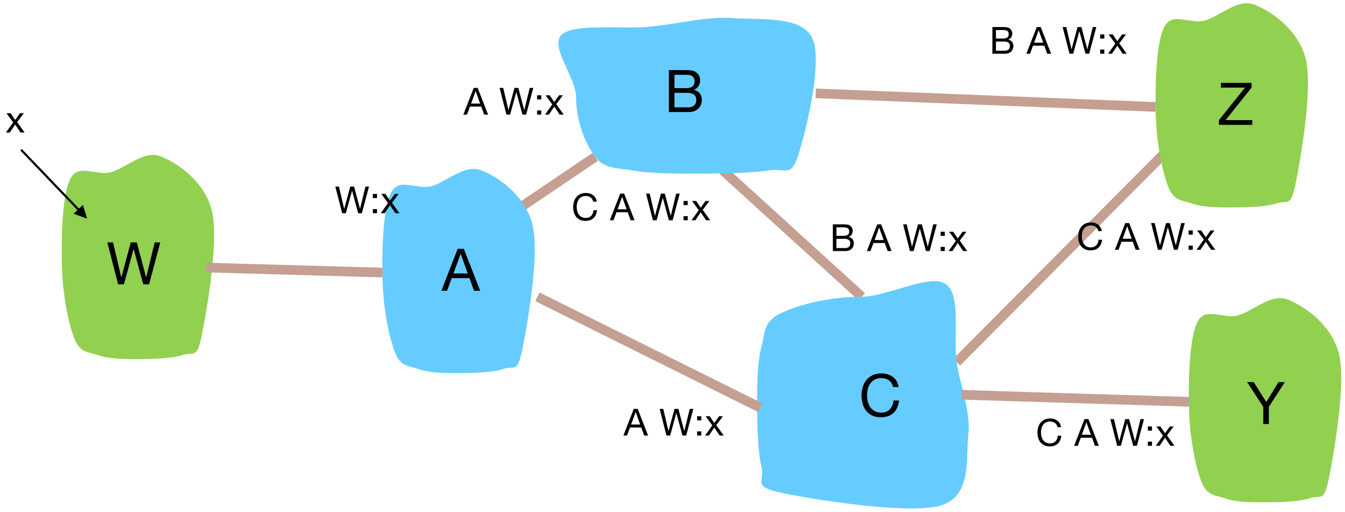
- eBGP: obtain subnet reachability information from neighboring ASes
- iBGP: propagate reachability information to all AS-internal routers



EBGP – PATH VECTOR ROUTING

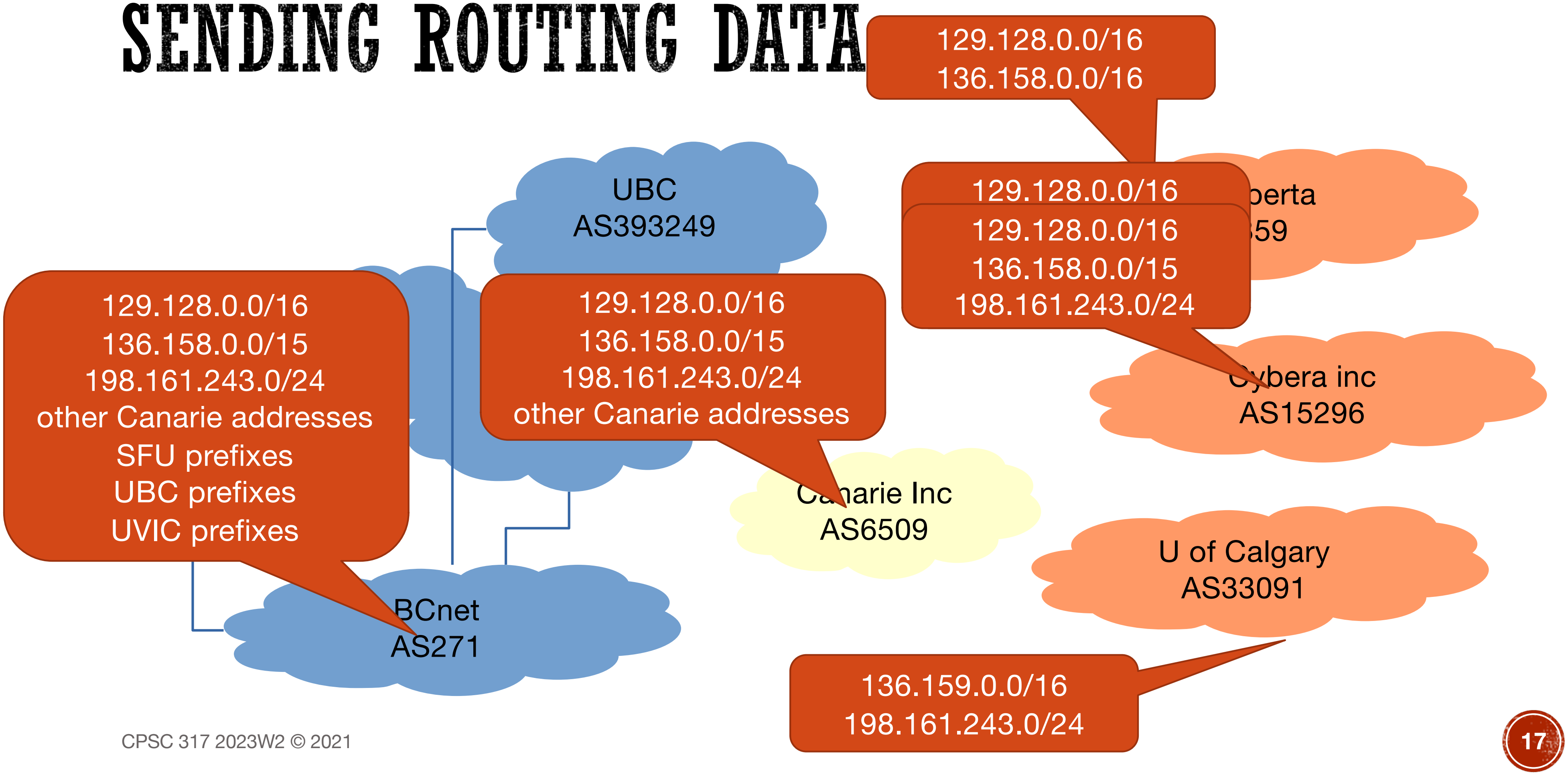
- Advertise reachability information of a prefix **x** to other ASes
- Path vector – variant of distance vector
- Advertise prefixes, which include:
  - AS number
  - AS path (prevents looping by allowing receiver to look for itself)
  - Next hop to take
- Receiver of advertisement builds forwarding table based on:
  - AS management decisions (policy)
  - Shortest route
  - Closest connecting router
  - Other factors

EBGP – PATH VECTOR ROUTING



How to find out about x?

SENDING ROUTING DATA



129.128.0.0/16
136.158.0.0/16

129.128.0.0/16

129.128.0.0/16
136.158.0.0/15
198.161.243.0/24

berta
59

Cybera inc
AS15296

UBC
AS393249

129.128.0.0/16
136.158.0.0/15
198.161.243.0/24
other Canarie addresses
SFU prefixes
UBC prefixes
UVIC prefixes

129.128.0.0/16
136.158.0.0/15
198.161.243.0/24
other Canarie addresses

Canarie Inc
AS6509

U of Calgary
AS33091

BCnet
AS271

136.159.0.0/16
198.161.243.0/24

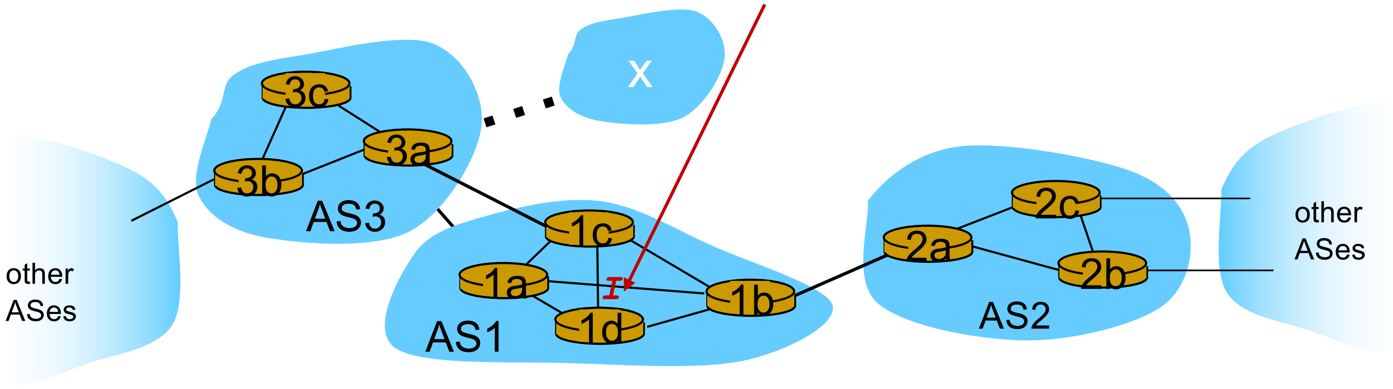
ROUTE AGGREGATION

- Route summarization
- Reduces the size of the routing table
- Reduces the number of advertisements

- Inside an AS it is called supernetting (200.23.26.0/21 is a supernet)

IBGP

- Suppose AS1 learns (via 1c, eBGP) that subnet ***x*** is reachable via AS3 (next hop 3a), but not via AS2
  - Border routers use iBGP to distribute info to all routers
- Router 1d determines from iBGP that the interface connected to 1c (call it ***I***) is on the least cost path to **x** via 3a
  - installs forwarding table entry ***(x,I)***



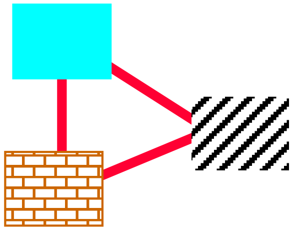
CLICKER QUESTION

- A Telus customer in Vancouver sends a message to a Shaw customer in Calgary. Ignoring other factors, and assuming all of these options are possible, which of the following options is better ***for Telus as an AS***?

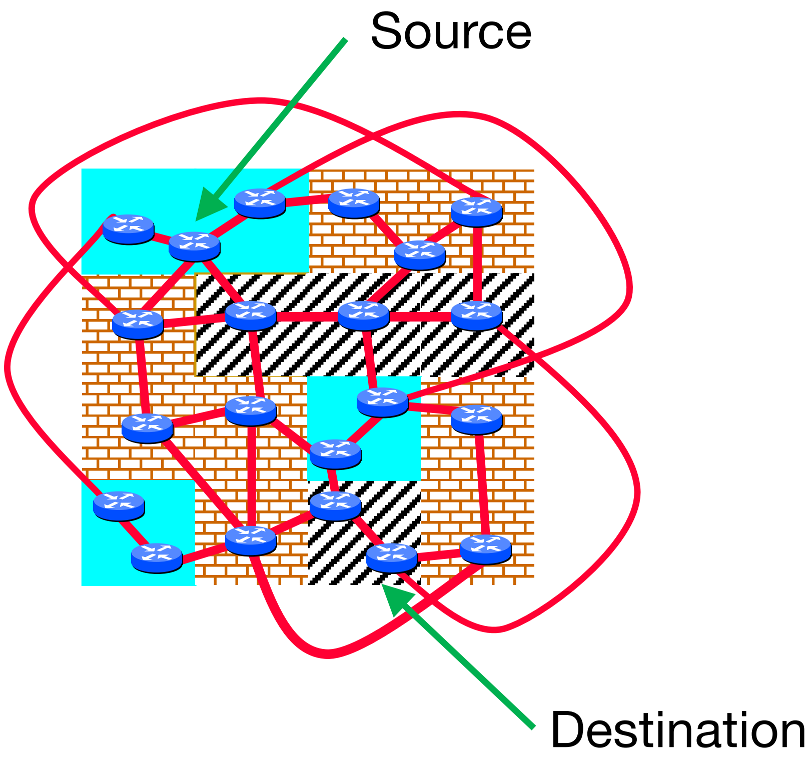
A. Transfer the data to Shaw in a Vancouver Exchange
B. Transfer the data to Shaw in a Calgary Exchange
C. Use a backbone AS (e.g., Hurricane Electric) to send the data to Calgary

DETERMINING ROUTES TO SUBNETS

Hot Potato Routing
Hand the packet off to the other AS as soon as possible

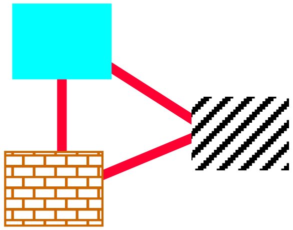


The AS graph

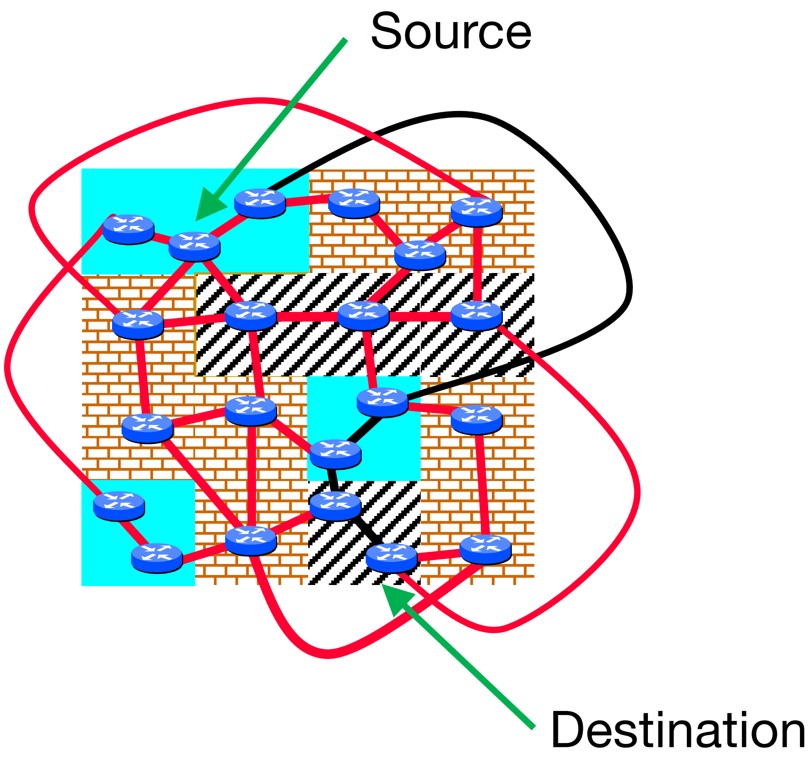


DETERMINING ROUTES TO SUBNETS

Hot Potato Routing
Hand the packet off to the other AS as soon as possible

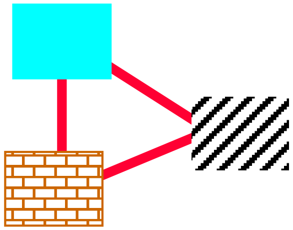


The AS graph

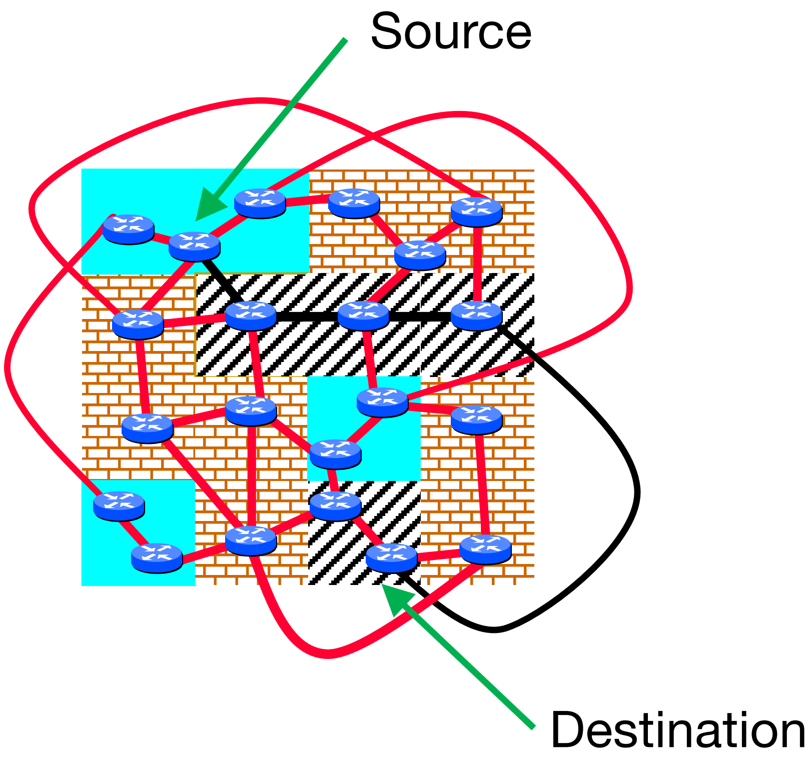


DETERMINING ROUTES TO SUBNETS

Hot Potato Routing
Hand the packet off to the other AS as soon as possible



The AS graph



CHALLENGES WITH BGP

- BGP is simple and globally consistent, but
- BGP policy is complex and locally determined
  - Routes accepted, rejected, trusted, propagated based on local decisions
- Pro: greater flexibility for organizations
- Con: vulnerable to bogus route propagations
  - BGP route leaks, route hijacking
  - MITM attacks

Slide based on https://web.stanford.edu/class/ee380/Abstracts/150211-slides.pdf

BGP ROUTE HIJACKING

- Advertise a prefix as part of a different AS than it actually is
  - Inadvertent – leak
  - Malicious – hijack
- If unused prefix – mildly bad behavior
- If prefix used in another AS – cause re-routing of traffic of the victim subnet and potential denial-of-service

Slide based on https://web.stanford.edu/class/ee380/Abstracts/150211-slides.pdf

MITM ATTACK

- Hijack a route from a source to a victim
- Re-route traffic via routers under adversarial control before forwarding it to the victim

Slide based on https://web.stanford.edu/class/ee380/Abstracts/150211-slides.pdf

BGP SECURITY INCIDENTS IN HISTORY

- 1997 – First BGP route “leak” (accidental)
- 2004 – Turk Telecom leak (target DNS)
- 2008 – Pakistan telecom hijack (target Youtube)
  - MITM explained by researchers
- 2010 – China Telecom leak (routed Verizon’s traffic)
- 2013 – first documented MITM case originating in Belarus (target credit card companies)
- 2022 – Attempts to block Twitter in Ukraine, Russia

Source: https://www.kentik.com/blog/a-brief-history-of-the-internets-biggest-bgp-incidents/

Slide based on https://web.stanford.edu/class/ee380/Abstracts/150211-slides.pdf

MITIGATIONS

- Transit AS providers need to check for false advertisements and filter them
- Custom filtering is laborious and error-prone
- Need to build filters from various routing registeries

Slide based on https://web.stanford.edu/class/ee380/Abstracts/150211-slides.pdf

SOFTWARE DEFINED NETWORK (SDN)

- An alternative to running routing algorithms in the routers
- A centralized service makes all the routing decisions, then tells each router what to do
- OpenFlow
- Used by a number of companies
  - Google
  - China Mobile
- Centralized solutions are better than distributed ones!

IN-CLASS ACTIVITY

- None