Active and passive measurements: networks

Prof. Anja Feldmann, Ph.D.
Dr. Steve Uhlig
Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
- Intra-domain routing
- Inter-domain routing
- Organization of the Internet
A map of the Internet in 2000
Organization of Internet Routing

- More than 30,000 autonomous routing domains:
  A domain is a set of routers, links, hosts and local area networks under the same administrative control
- Domains size: from one to millions of hosts
- Interconnections between domains are complex
Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
- Intra-domain routing
- Inter-domain routing
- Organization of the Internet
Types of domains: transit

- **Transit domains:**
  A *transit domain* allows external domains to use its own infrastructure to send packets to other domains

- Implicit hierarchy of transit domains according to “size”

- Examples: AT&T, UUNet, Level3, Opentransit, KPN,...

- 15% of all ASs
Tier-1: UUNET
Tier-2: GEANT
Types of domains: stub

- **Stub domains:**
  
  A *stub domain* does not allow external domains to use its infrastructure to send packets to other domains.

- A stub is connected to at least one transit domain.

- Content stub domains: Yahoo, Google, MSN, BBC,...

- Access stub domains: ISPs providing Internet access via CATV, DSL,...

- 85% of all ASs
Stub: SURFNET
BELNET
Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
- Intra-domain routing
- Inter-domain routing
- Organization of the Internet
Intra- and inter-domain routing

- **Interior Gateway Protocol (IGP):**
  - Routing of IP packets *inside each domain*
  - Only knows topology of its domain

- **Exterior Gateway Protocol (EGP):**
  - Routing of IP packets *between domains*
  - Each domain is considered as an atomic structure
Intra- and inter-domain routing
Advertizing a prefix
Traffic paths

Inter-domain link
Intra-domain link
BGP is not shortest-paths!

AS 1
AS 2
AS 3
AS 4
AS 5
AS 6

Effect of policy

AS path
Inter-AS edge
Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
  - Intra-domain routing
  - Inter-domain routing
- Organization of the Internet
Intra-domain routing

- **Goal:** allow routers to transmit IP packets along the best path towards their destination
  - *best* usually means the *shortest* path
  - Allow to find alternate routes in case of failures
- **Behavior:** all routers exchange routing information
  - Each domain router can obtain routing information for the whole domain
  - The network operator or the routing protocol selects the cost of each link
Types of IGPs

- **Static routing:** only useful in very small domains
- **Distance vector routing:**
  - Routing Information Protocol (RIP)
    - Still widely used in small domains despite its limitations
- **Link-state routing:**
  - Open Shortest Path First (OSPF): widely used in enterprise networks
  - Intermediate System- Intermediate-System (IS-IS): widely used by ISPs
Distance-vector routing

- Each router sends periodically a distance vector containing, for each known prefix:
  1. The IP prefix
  2. The distance between itself and the destination
     - The distance vector is a summary of the router's routing table
- Each router receives its neighbor's distance vectors and builds its routing table based on those vectors
Link-state routing

- Each router builds link state packets containing its local topology
  - Link state packets are created at regular intervals and when the local topology changes
- Link state packets are reliably flooded to all routers inside the domain
- Each router knows the complete domain topology by maintaining a LSP database
Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
- Intra-domain routing
- Inter-domain routing
- Organization of the Internet
Inter-domain routing

- **Goal:** allow to transmit IP packets along the **best path** towards their destination
  - From an interdomain viewpoint, **best path** often means **cheapest path**

- **Behavior:**
  - Each domain specifies inside its **routing policy** the domains for which it agrees to provide a transit service and the method it uses to select the best path to reach each destination
  - Each router of the domain chooses its best path according to the routing policies, and advertises them to its neighboring routers
Inter-domain routes redistribution

- **Between domains (eBGP):**
  - Goal: propagate external reachability to neighbors
  - Implementation: private peerings, public interconnection points

- **Inside a domain (iBGP):**
  - Goal: propagate the routes learned from neighbors to the routers inside the domain
  - Implementation: full-mesh between BGP routers, route-reflection, or confederations
Conceptual operation of a BGP router

BGP Routing Information Base
Contains all the acceptable routes learned from all Peers + internal routes
- BGP decision process selects the best route towards each destination

BGP Loc-RIB
All acceptable routes

Peer[1] Import filter Attribute manipulation
Determines which BGP Msgs are acceptable from Peer[1]

Peer[N] Import filter Attribute manipulation
BGP Msgs from Peer[N]

Peer[1] Export filter Attribute manipulation
Determines which routes can be sent to Peer[1]

Peer[N] Export filter Attribute manipulation
BGP Msgs to Peer[N]

BGP Adj-RIB-In
Peer[1] from Peer[N]

BGP Adj-RIB-Out
Peer[N] to Peer[1]
Conceptual operation of a BGP router

1. highest weight
2. highest LOCAL–PREF
3. shortest AS–PATH
4. lowest MED
5. eBGP over iBGP
6. nearest IGP neighbor
Path selection in a domain
Routing policies

- BGP allows each domain to define its own routing policy
- Some policies are common:
  - customer-provider peering
  - Customer C buys Internet connectivity from provider P
  - shared-cost peering
  - Domains x and y agree to exchange packets by using a direct link or through an interconnection point
Routing policies

- Routing policies implement business relationships between domains
- The routing policy of a domain is implemented via the route filtering mechanism on BGP routers:
  - **Inbound filtering**: Upon reception of a route from a peer, a BGP router decides whether the route is acceptable, and if so whether to change some of its attributes.
  - **Outbound filtering**: Before sending its best route towards a destination, a BGP router decides which peers should receive this route and whether to change some of its attributes before sending it.
Customer-provider peering

- Customer sends to its provider its internal routes and the routes learned from its own customers => Provider will advertise those routes to the entire Internet to allow anyone to reach the Customer
- Provider sends to its customers all known routes => Customer will be able to reach anyone on the Internet
Shared-cost peering

- Peer X sends to Peer Y its internal routes and the routes learned from its own customers
  - Peer Y will use shared link to reach Peer X and Peer X's customers
  - Peer X's providers are not reachable via the shared link
- Peer Y sends to Peer X its internal routes and the routes learned from its own customers
  - Peer X will use shared link to reach Peer Y and Peer Y's customers
  - Peer Y's providers are not reachable via the shared link
Business relationships and BGP

- Relationships between ASs stem from their mutual interest:
  - if interest is not balanced, customer-provider will typically arise
  - if balanced interest, peer-peer or other partial agreement will arise
- As long as business relationships remain so important, the interdomain routing protocol does not matter so much, i.e. technical changes in BGP should not affect interdomain routing so much
Classical iBGP

- iBGP full-mesh:
  - connect all routers
  - each router knows all the best routes of all other routers
- Redistribution rules:
  - Redistribute best route to all peers except the one from which the route was received
  - Do not redistribute a best route if it was learned from an iBGP peer (iBGP peers must have learned this route directly from the concerned peer)
iBGP full-mesh: path selection
Route-reflection

• When ASs become large (hundred of routers), full-mesh does not scale ⇒ route-reflection
• **2 types of route-reflector peers:** *client* and *non-client* [RFC2796]
• **Implicit hierarchy:**
  • clients are “down”
  • non-clients are either “peer” or “up”
• **Redistribution rules:**
  • Best route received from client or eBGP peer ⇒ redistribute to clients and non-client peers (down, peer and up)
  • Best route received from non-client peer ⇒ redistribute to clients only (down only)
• iBGP graph is generally not a forest ⇒ prevent looping of routes (cluster-ID)
Route-reflection: path selection
iBGP and loss of path diversity
iBGP: summary

- **iBGP full-mesh:**
  - Pro’s: full visibility of external routes, small convergence time
  - Con: $N*(N-1)/2$ iBGP sessions

- **Route-reflection:**
  - Pro: # iBGP sessions $\sim$ # physical links
  - Con’s: opaqueness of best route selection, slow convergence, route oscillations

For more details:


Outline

- Organization of Internet routing
- Types of domains
- Intra- and inter-domain routing
- Intra-domain routing
- Inter-domain routing
- Organization of the Internet
Internet hierarchical structure

- Tier-1 ISPs
  - About 20 large ISPs
  - Provide transit service
- Tier-2 ISPs
  - Regional or National ISPs
  - Customers of T1 ISP(s)
  - Providers of T3 ISP(s)
- Tier-3 ISPs
  - Smaller ISPs, Corporate Networks, Content providers
  - Customers of T2 or T1 ISPs