Implementation

Major Internet routing protocols:

- RIP (v1 and v2): intra-domain, Bellman-Ford

  → also called "distance vector"

  → metric: hop count

  → UDP

  → nearest neighbor advertisement

  → popular in small intra-domain networks

- OSPF (v1 and v2): intra-domain, Dijkstra

  → also called "link state"

  → metric: average delay

  → directly over IP: protocol number 89

  → broadcasting via flooding

  → popular in larger intra-domain networks

- IS-IS: intra-domain, Dijkstra

  → "link state"

  → directly over link layer (e.g., Ethernet)

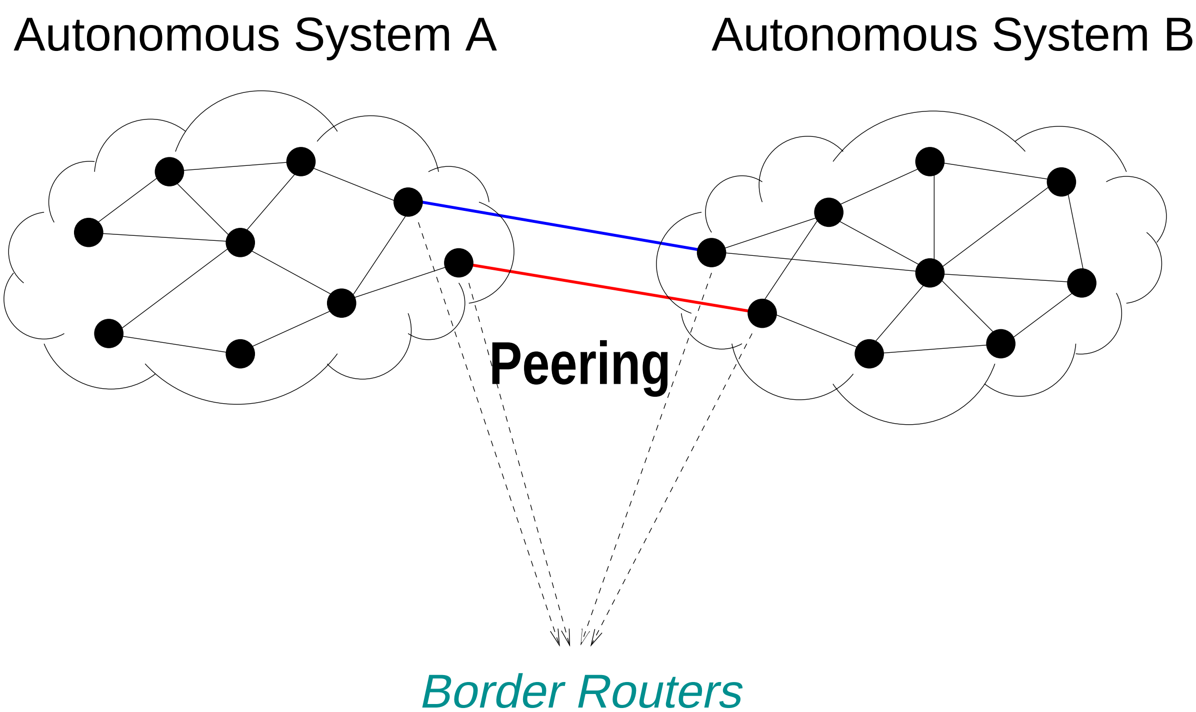
  → more recently: also available over IP

  → flooding

  → popular in larger intra-domain networks

BGP (Border Gateway Protocol):

$\longrightarrow$ inter-domain routing



$\longrightarrow$ "peering" between two domains

$\longrightarrow$ includes customer-provider relationship

$\longrightarrow$ Internet exchanges: multiple domains

- CIDR addressing

  → i.e., $a.b.c.d/x$

  → Purdue: 128.10.0.0/16, 128.210.0.0/16, 204.52.32.0/20

  → check at `www.iana.org` (e.g., ARIN for US)

- Metric: policy

  → e.g., shortest-path, trust, pricing

  → meaning of "shortest": delay, router hop, AS hop

  → mechanism: path vector routing

  → BPG update message

BGP route update:

$\longrightarrow$ BGP update message propagation

BGP update message:

$$\mathrm{ASN}A_k \to \cdots \to \mathrm{ASN}A_2 \to \mathrm{ASN}A_1; \mathrm{a.b.c.d/x}$$

Meaning: ASN A_1 (with CIDR address a.b.c.d/x) can be reached through indicated path

$\longrightarrow$ "path vector"

$\longrightarrow$ called AS-PATH

Some AS numbers:

- Purdue: 17
- BBN: 1
- UUNET: 701
- Level3: 3356
- Abilene (aka "Internet2"): 11537

Policy:

- if multiple AS-PATHs to target AS are known, choose one based on policy
  - → e.g., shortest AS path length, cheapest, least worrisome
- advertise to neighbors target AS's reachability
  - → also subject to policy
  - → no obligation to advertise
  - → specifics depend on bilateral contract (SLA)

SLA (service level agreement):

⟶ bandwidth (e.g., 1 Gbps, OC-3, DS3

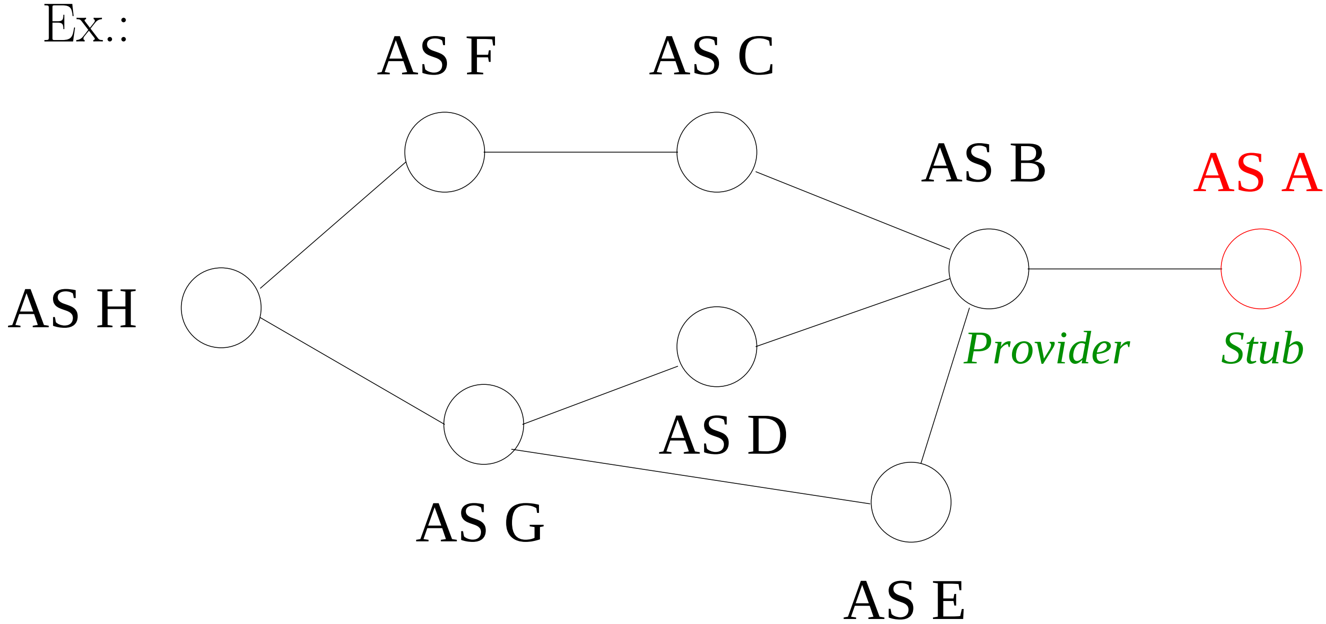
⟶ delay (e.g., avrg. 25ms US), loss (e.g., 0.05%)

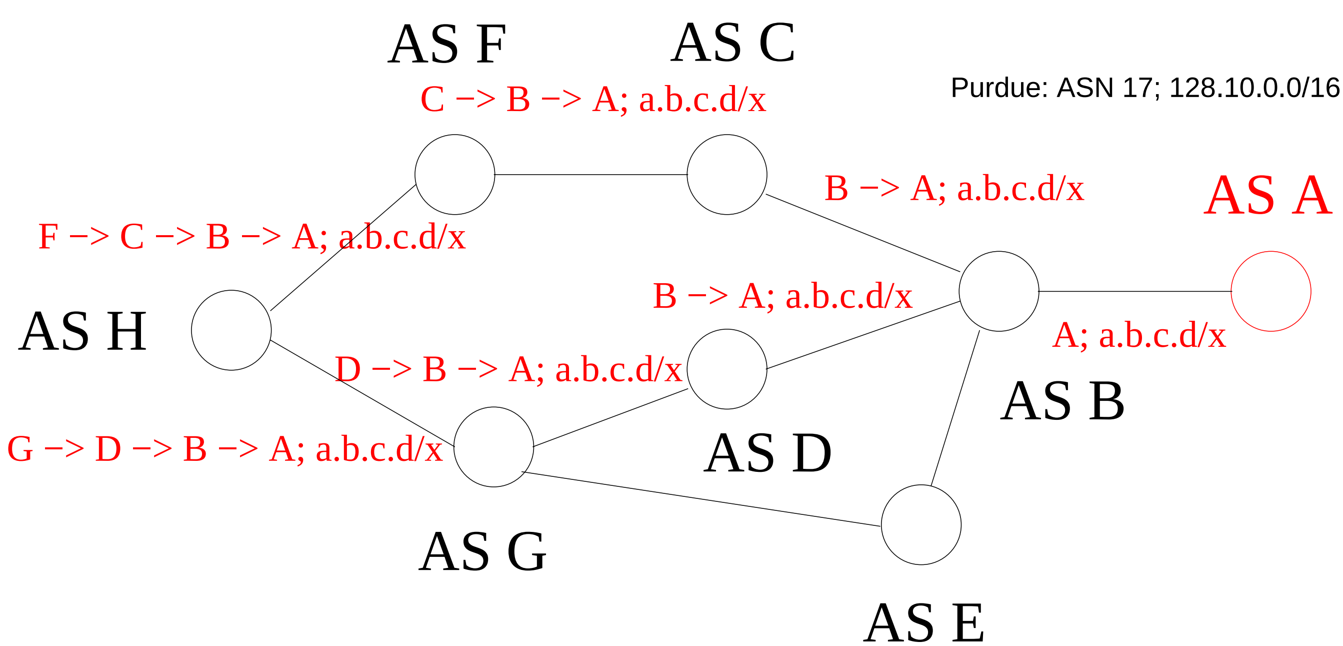
⟶ pricing (e.g., 1 Mbps: below $100)

⟶ availability (e.g., 99.999%)

⟶ etc.

Ex.:


AS F
AS C
AS B
AS A
AS H
Provider
Stub
AS D
AS G
AS E


AS F
AS C
Purdue: ASN 17; 128.10.0.0/16
C –> B –> A; a.b.c.d/x
B –> A; a.b.c.d/x
AS A
F –> C –> B –> A; a.b.c.d/x
B –> A; a.b.c.d/x
AS H
A; a.b.c.d/x
D –> B –> A; a.b.c.d/x
AS B
G –> D –> B –> A; a.b.c.d/x
AS D
AS G
AS E

Performance

Route update frequency:

$\longrightarrow$ routing table stability vs. responsiveness

$\longrightarrow$ rule: not too frequently

$\longrightarrow$ 30 seconds

$\longrightarrow$ stability wins

$\longrightarrow$ hard lesson learned from the past (sub-second)

$\longrightarrow$ legacy: TTL

Other factors for route instability:

$\longrightarrow$ selfishness (e.g., fluttering)

$\longrightarrow$ BGP's vector path routing: inherently unstable

$\longrightarrow$ more common: slow convergence

$\longrightarrow$ target of denial-of-service (DoS) attack

Route amplification:

- $\longrightarrow$ shortest AS path $\neq$ shortest router path
- $\longrightarrow$ e.g., may be several router hops longer
- $\longrightarrow$ AS graph vs. router graph
- $\longrightarrow$ inter- vs. intra-domain routing: separate subsystems
- $\longrightarrow$ policy: company in Denmark

Route asymmetry:

- $\longrightarrow$ routes are not symmetric
- $\longrightarrow$ estimate: $> 50\%$
- $\longrightarrow$ mainly artifact of inter-domain policy routing
- $\longrightarrow$ various performance implications
- $\longrightarrow$ source traceback

Black holes:

- $\longrightarrow$ persistent unreachable destination prefixes
- $\longrightarrow$ BGP routing problems
- $\longrightarrow$ further aggrevated by DNS
- $\longrightarrow$ purely application layer: end system problem