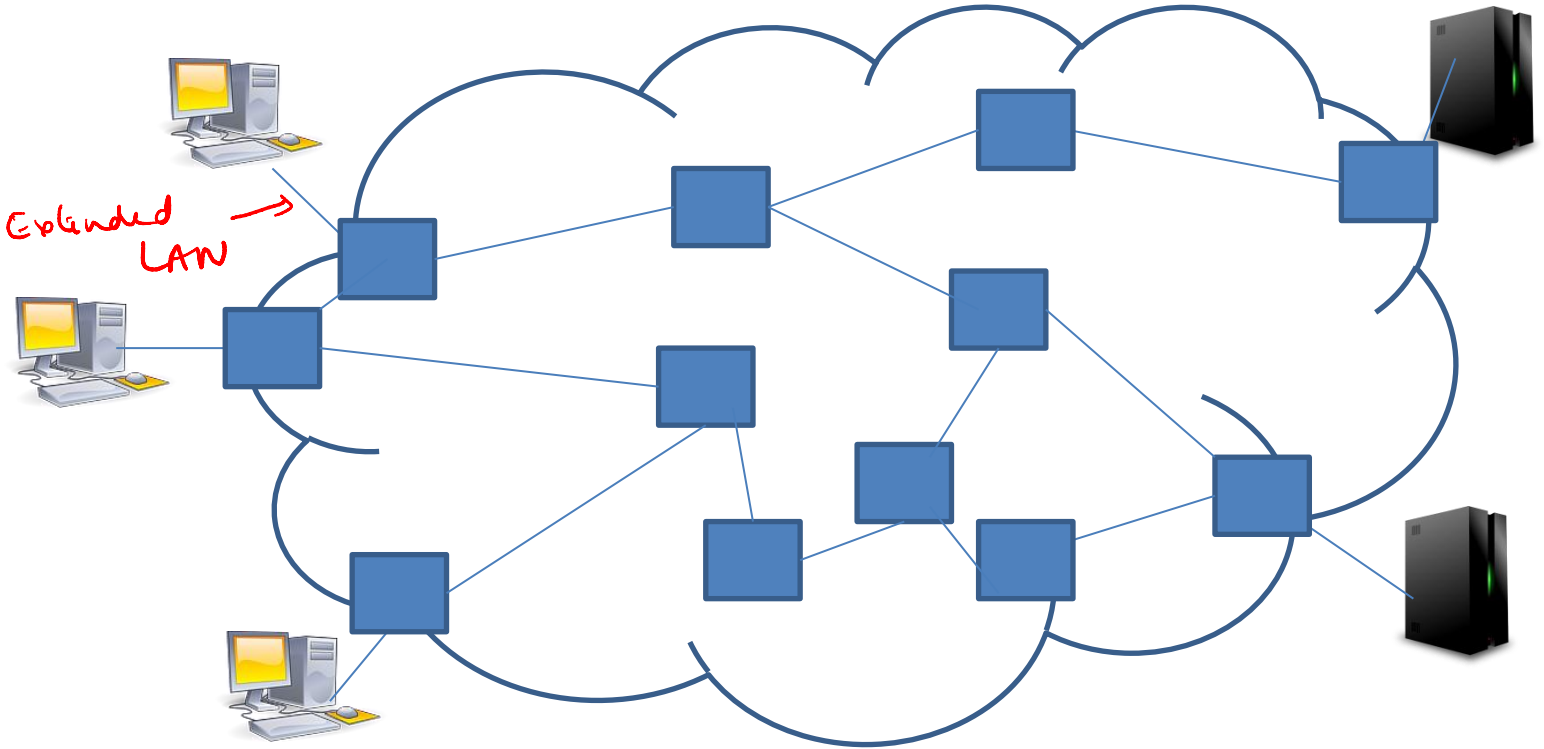
Interdomain Routing: Border Gateway Protocol (BGP)

Kameswari Chebrolu

Background

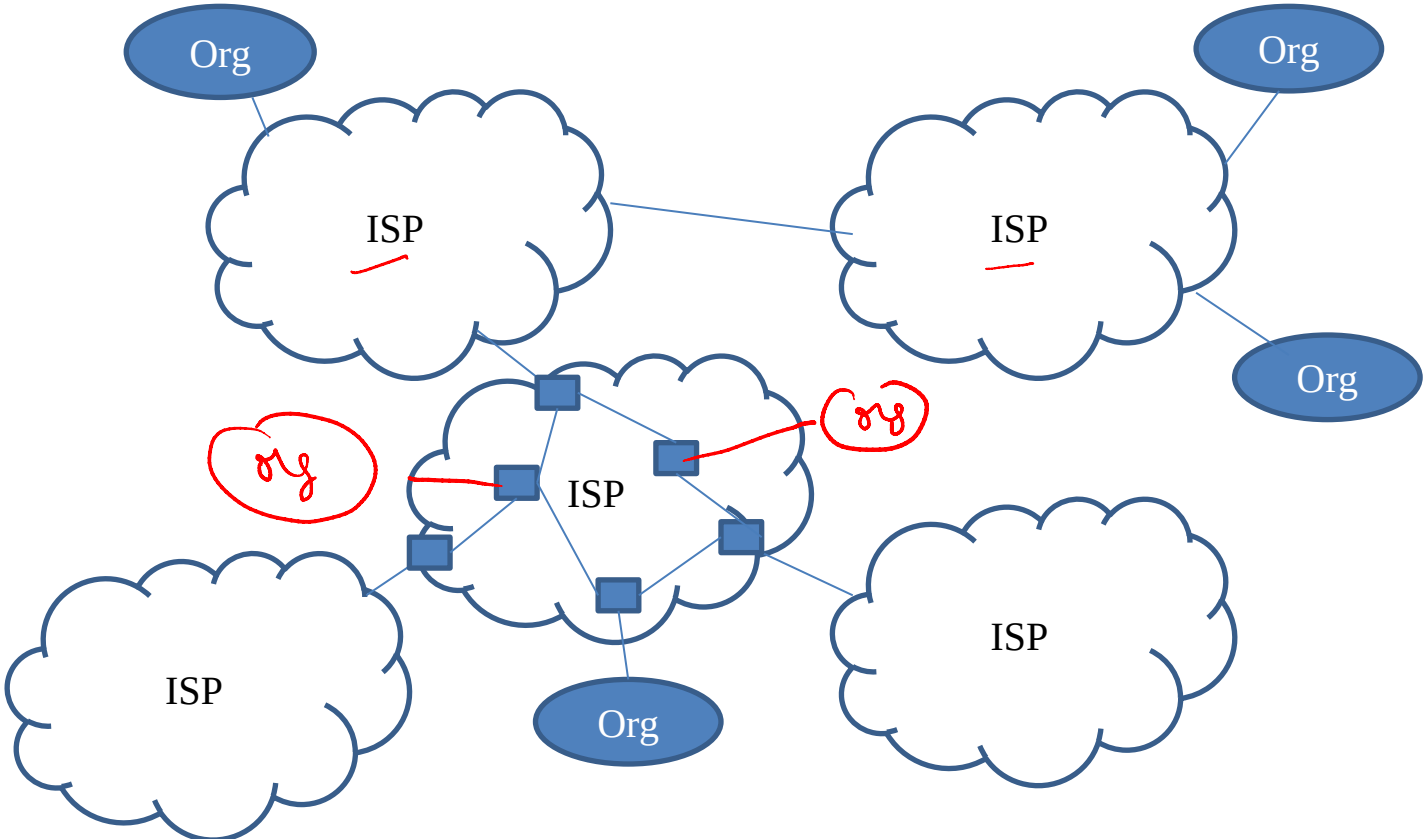
- Routing process builds forwarding tables at routers
- Two types of algorithms: DV and LS
- Routing in Internet is lot more complex: Need to handle policy, scale and performance
- BGP protocol is extremely complex
  - Many issues still not well understood, very few possess good knowledge of the protocol

Internet Architecture – v1

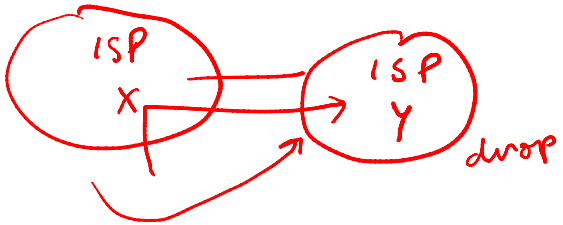


Wish it were so!

Internet Architecture – v2

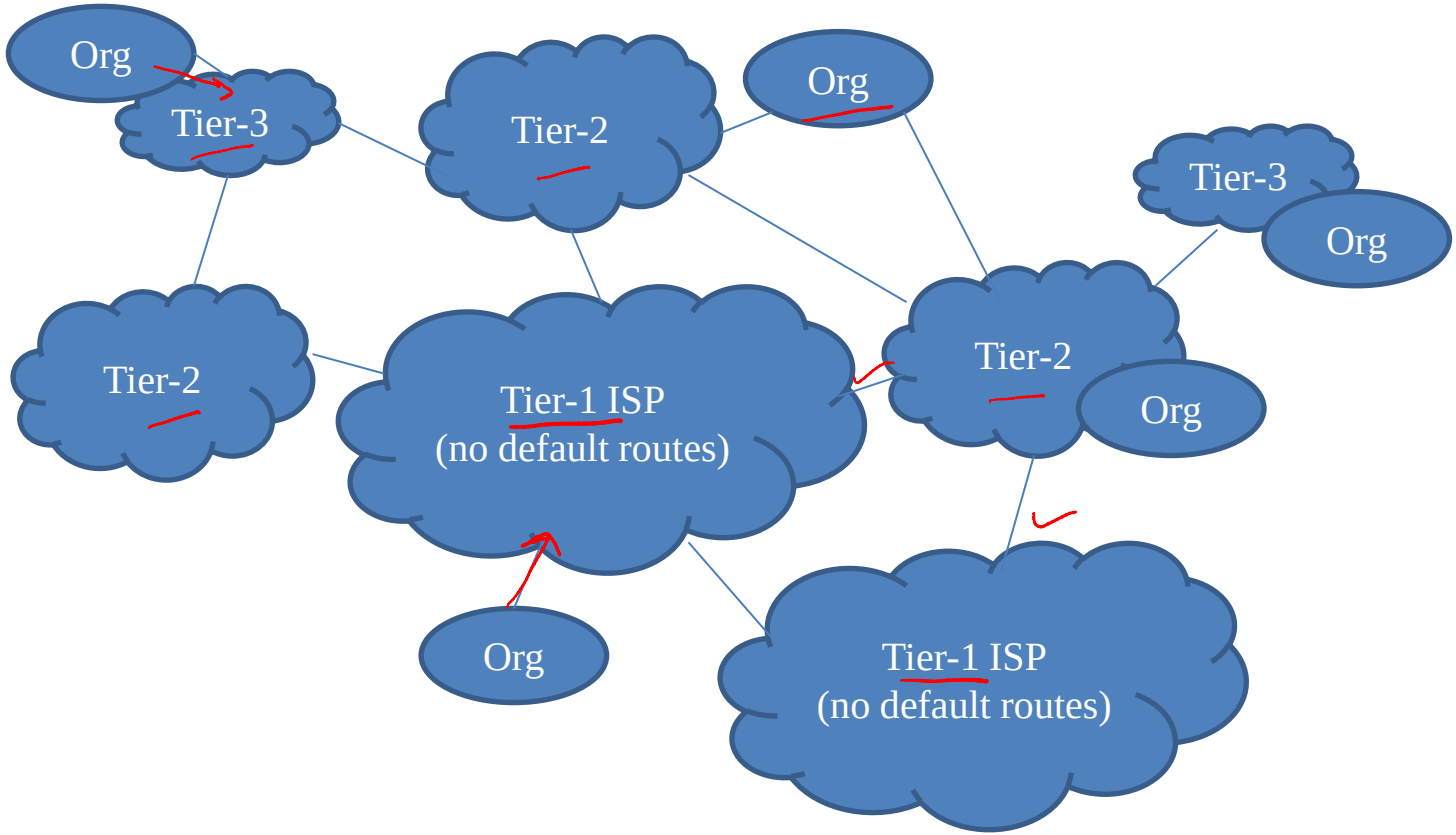


Reality



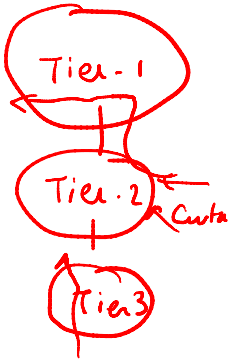
- Connectivity, Cooperation and Competition
- Many ISP's want complex policies
- Not all ISP's are equal
  - Tier1, Tier 2, Tier3
- Scalability is a big concern

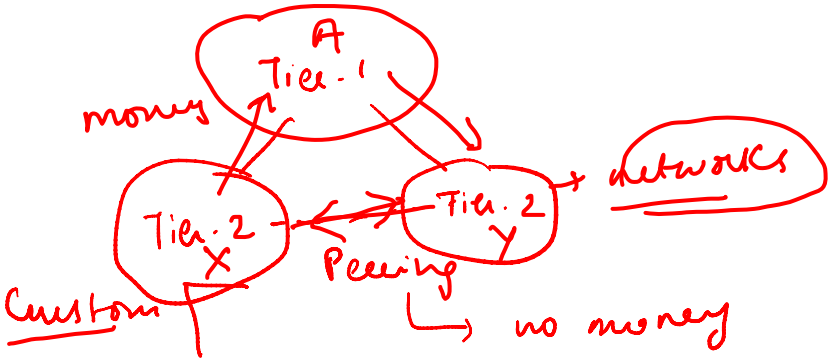
Internet Architecture – V3



Relations between ~~ASs~~ ISPs

- Transit: provider/customer relation; typically financial settlement is involved
- Peering: Mutual access to subset of routing tables; typically no financial settlement involved





Routing in the Internet

- Autonomous System == Routing Domain: Controlled by a single administrative entity
  - Network within an organization; network within an ISP
- Routing problem: Two aspects
  - Routing within an AS
  - Routing between ASs

Routing within an AS

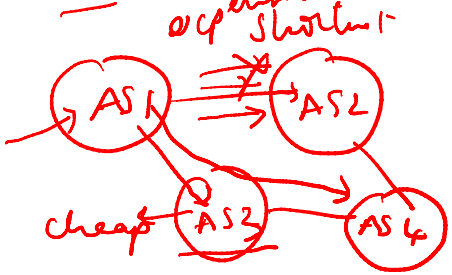
- Intradomain routing: Employs interior gateway protocol (IGP)
  - E.g. OSPF, RIP
- Focuses on finding 'optimal' paths within the domain
- Different AS can implement different IGPs within

Routing across ASs

- Interdomain routing: Employs exterior gateway protocol
  - Border Gateway Protocol (BGPv4)
- Focuses on Reachability, Policy and Scalability
- Needs to be common across ASs

Policies

- Dictated by political, security and economic consideration
- No transit through certain ASs
- Use a specific AS only if there are no other routes to destination
- Traffic from X country can't go through Y country
- Traffic starting or ending at Google can't go through Facebook

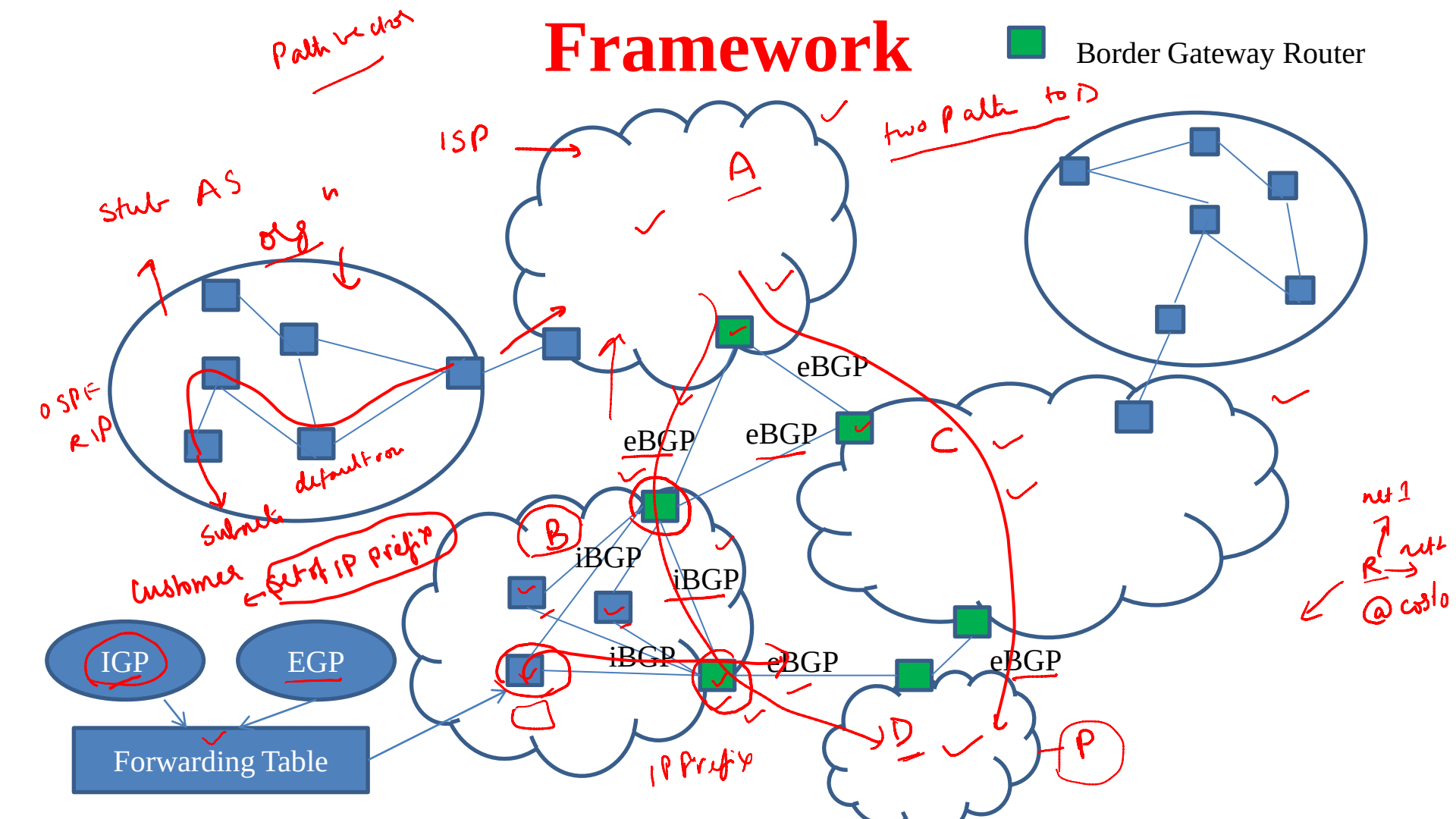


Break

Routing across ASs

- BGP messages exchanged using TCP, port 179.
- eBGP: Facilitate inter-AS communication
  - Routers are in different ASs, often directly connected
- iBGP: Facilitate communication
  - Routers within same AS, need not be directly connected
  - Install learned routes (via eBGP) within AS

Framework


Border Gateway Router
ISP
A
stub AS
eBGP
eBGP
eBGP
B
iBGP
iBGP
iBGP
eBGP
eBGP
C
D
P
IGP
EGP
Forwarding Table
OSPF
RIP
default
IP prefix
IP Prefix

Summary of Framework

- Border gateway routers employ eBGP to exchange IP prefix information
  - Underlying route determination (which next hop AS to take) is based on path vector
  - An AS need not export all the IP prefixes it has learnt (to be covered under exporting routes)
  - When there are multiple routes to a given destination, policy takes precedence over optimality (to be covered under importing routes)

Summary of Framework

- Learned information via eBGP is injected within AS via iBGP sessions
  - Border gateway routers form a mesh of iBGP sessions with all routers within AS
- A forwarding table at a router is dictated by both the IGP and EGP protocols