BGP

1. Internet Routing

Internet Routing

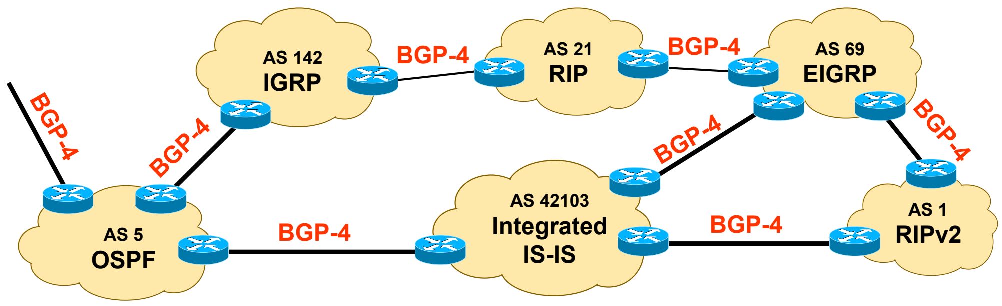
- Interior Gateway Protocols (IGPs) not suitable for Inter-ISP routing
  - Technical metrics only
  - No policy features
- Inter-ISP routing is a question of contracts, money, services, policy
- Lead to development of Exterior Gateway Protocols (EGP)

It is impossible to use IGPs for inter-ISP routing because of the metric they use and only basic policy which could be implemented. For inter-ISP routing there is a need for some other protocols - EGPs

Routing Protocol Realms

- **Autonomous System (AS) = domain of one routing policy**
  - **16 bit AS number (maintained by Internet Registries)**
  - **Private range: 64512-65534 (RFC 1930)**

- **IGPs are used inside an AS**
- **EGPs are used between ASs**



 3

An Autonomous System (AS) is a group of IP networks run by one or more network operators that has a single and clearly defined routing policy.

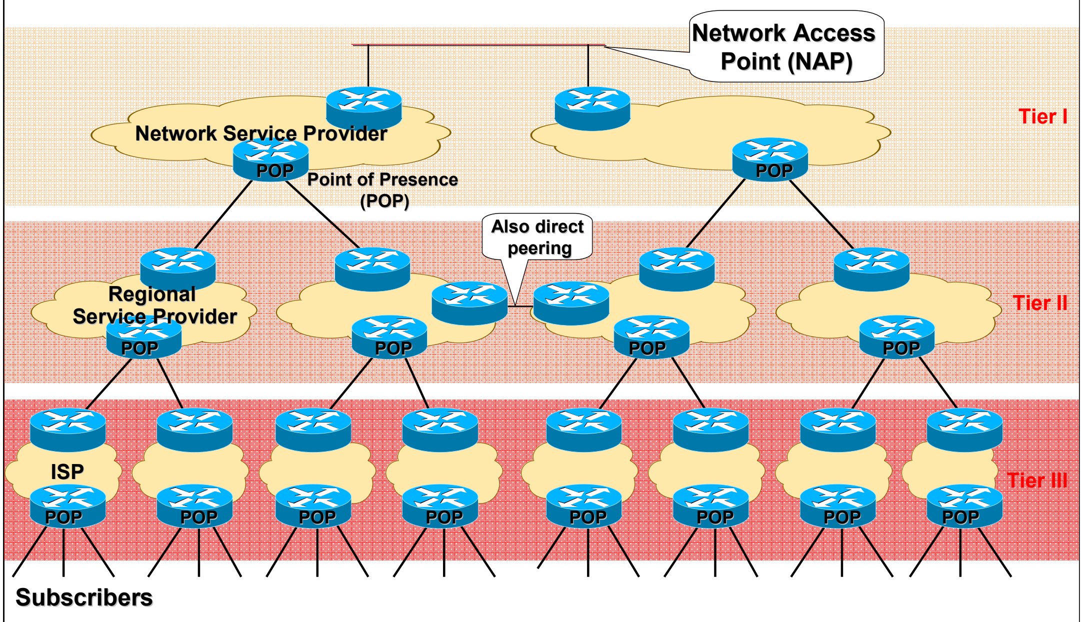
Different types of routing protocols are used between Autonomous Systems (ASs) and inside an AS. Between AS the used routing protocols are called Exterior Gateway Protocols (EGPs) while Interior Gateway Protocols (IGP) are used inside an AS.

Policy Routing

- ISPs exchange routes with each other
  - Policies affect paths and reachabilities
- For example, policies define
  - Which routes to accept
  - Which routes to advertise
  - Which path to use
- Each advertized route has several attributes assigned to it
  - Metric is only one of them

The routing policy is easy to implement when using BGP, but impossible to implement with any other routing protocol. EIGRP, for example, can do route filtering only on individual IP subnets, not on all prefixes belonging to an autonomous system. Link-state protocols, such as OSPF, cannot do powerful route filtering at all. BGP, however, can do this, based on AS numbers, which makes it possible to scale over the Internet.

Internet Architecture



Today the Internet Service Providers (ISPs) are divided into a three layer structure. Those ISPs who are directly connected to the Internet (i. e. who constitute the Internet!) are called "Tier 1" or Network Service Providers.

Tier 2 or Regional Service Providers are often interconnected and connect Tier 3 ISPs to the backbone of the Internet. Eventually, the subscribers are connected via "Point of References" (POPs) to the Internet.

In real world even higher-order providers connect subscribers directly to the Internet.

Peering

- **NAPs are typically multipoint LANs such as Ethernet, FDDI, ATM**
- **In real world, peering not only done at NAPs between Network Service Providers**
  - **Also between Tier-II and Tier-III**
  - **Also directly point-to-point**
  - **Also without *Route Servers***

What the hell are Route Servers?

A peering agreement established between two service providers is a question of traffic patterns. If the traffic is heavier in one direction, then the originating SP will have to pay for the peering. Peering is free if both sides benefit from it.

Routing Arbiter (RA)

- Routing Arbiter Database (RADB)
  - NSF-funded *Routing Arbiter Project* proposed centralized database of routes and policies
  - To support manageability and stability
  - RIPE-181 as database language
- RADB is maintained by Route Server
  - UNIX workstation running BGP
  - Instead of n(n-1)/2 peerings, every service provider only peers with route server

Using a Routing Arbiter DataBase (RADB) significantly reduces the number of peerings. The basic idea is to replace a any-to-any peering topology by a star topology.

Internet Registries

- Internet Registry (IR) care for
  - Network numbers
  - AS numbers
  - Domain Names
  - Domain Name Servers
- IR passes responsibility to Regional Internet Registries (RIRs)
  - E. g. RIPE NCC, APNIC, ARIN

The Internet Registry hierarchy consists of the following levels of hierarchy as seen from the top down: IANA, Regional IRs, Local IRs.

IANA - The Internet Assigned Numbers Authority has authority over all number spaces used in the Internet. This includes Internet Address Space. IANA allocates parts of the Internet address space to regional IRs according to its established needs.

Regional IRs operate in large geopolitical regions such as continents. Currently there are three regional IRs established; InterNIC serving North America, RIPE NCC serving Europe, and AP- NIC serving the Asian Pacific region. Since this does not cover all areas, regional IRs also serve areas around its core service areas. It is expected that the number of regional IRs will remain relatively small. Service areas will be of continental dimensions. Regional IRs are established under the authority of the IANA. This requires consensus within the Internet community of the region. A consensus of Internet Service Providers in that region may be necessary to fulfill that role. The specific duties of the regional IRs include coordination and representation of all local IRs in its respective regions.

Local IRs are established under the authority of the regional IR and IANA. These local registries have the same role and responsibility as the regional registries within its designated geographical areas. These areas are usually of national dimensions.

Routing Registries

- To avoid conflicting policies of different ISPs when interconnecting them
- Required when peering at public peering points
- A must when connecting to major ISPs

Routing Registries are important when connecting to a major ISP.

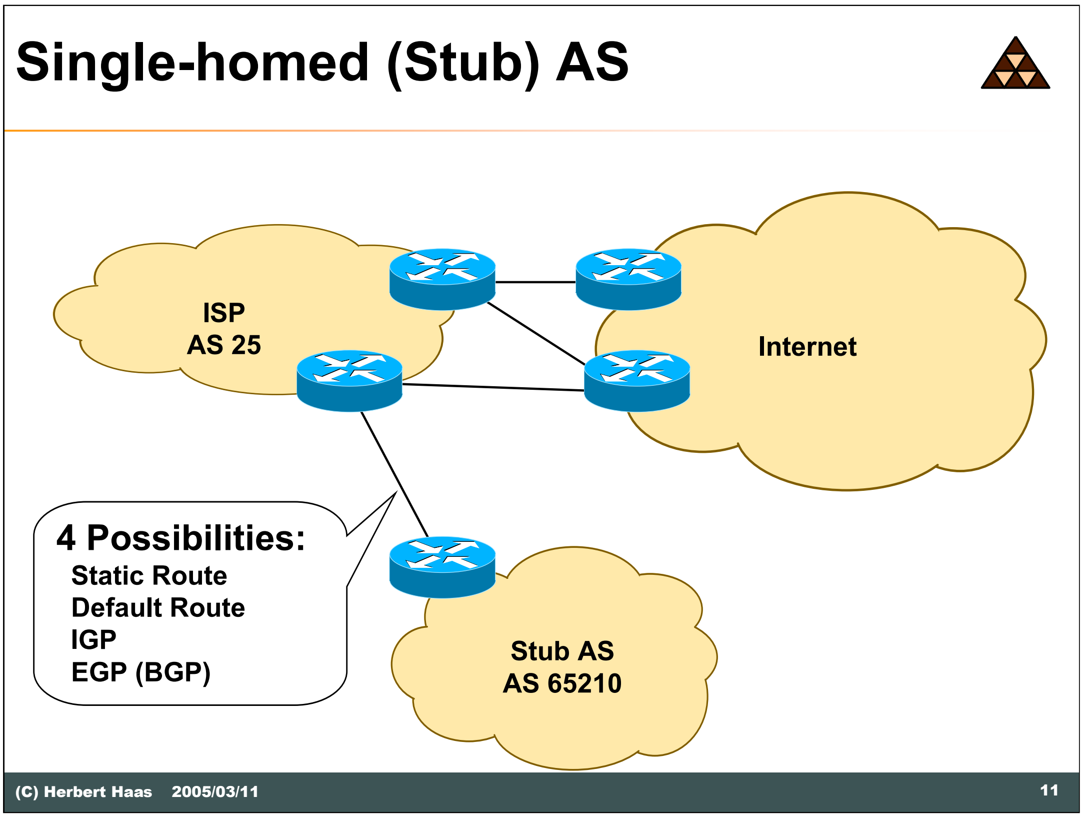
Autonomous Systems

- Stub or Single-Homed AS
- AS with multiple upstream transit provider: Multi-homed AS
  - Multi-homed Non-Transit
  - Multi-homed Transit

See RFC 1930 **Guidelines for creation, selection, and registration of an Autonomous System (AS)**

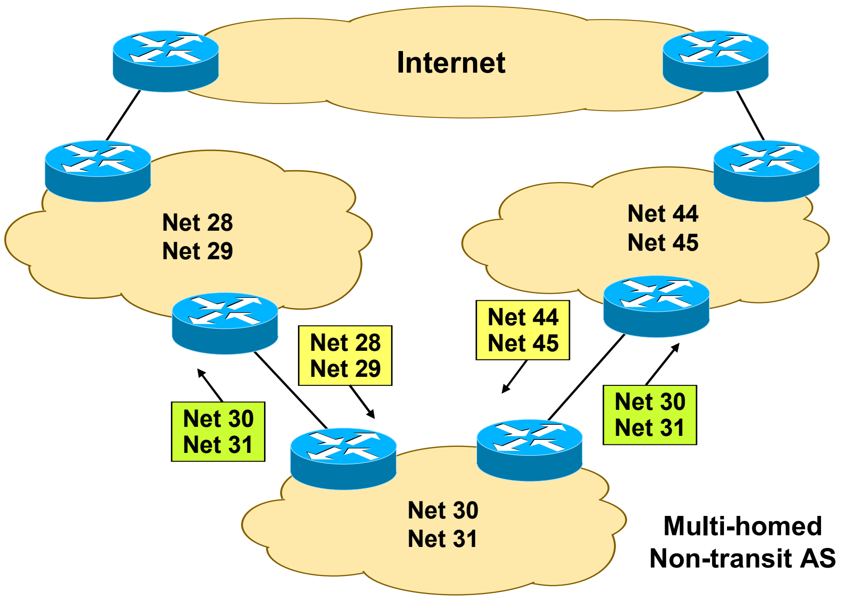
There are three types of autonomous systems:

1. Stub or Single-homed
2. Multi-homed Non-Transit
3. Multi-homed Transit



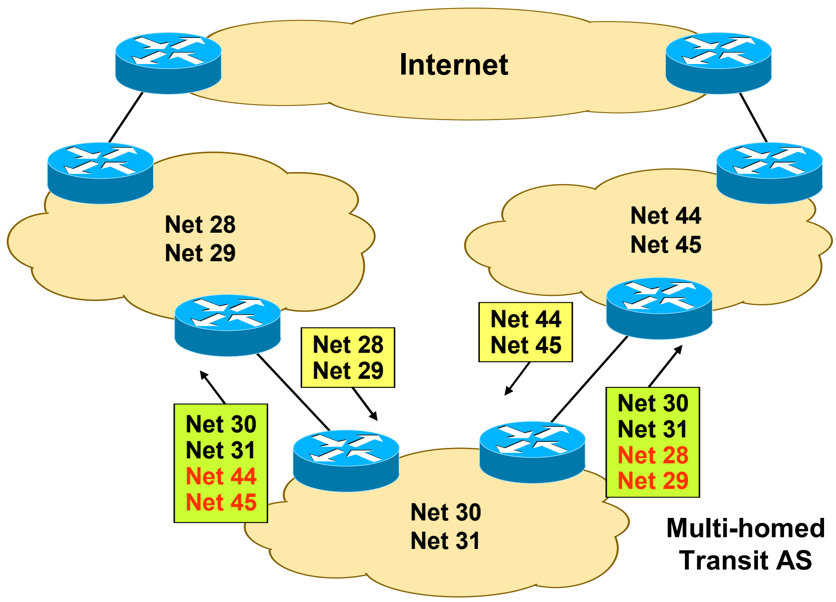
The above scenario where a stub AS is connected to the Internet using a single Internet Service Provider is generally not a case where BGP is used. A normal Internet access to a single ISP does not require BGP, static routes are more commonly used to handle this situation. Small ISPs buying Internet connectivity from other ISPs use this type of connectivity more often, especially if they want to start their business the proper way—by using their own AS number and having their own address space.

Multi-homed Non-transit AS



Although there are designs where BGP could be avoided, most multi-homed customers use BGP with their service providers. The multi-homed customers must have their own AS number and it is recommended to use a public AS number. Multi-homed customers should use a provider-independent address space which is allocated to them directly by an Internet registry.

Multi-homed Transit AS



BGP is used most commonly in service provider networks that ensure connectivity between their customers and the rest of the Internet. An ISP might exchange BGP updates with the customers or use static routing toward them. It also connects to other ISPs and is required to forward the routes received from the customers to the rest of the Internet, as well as in the other direction. As a result, user data traffic starts to flow between the customers and the rest of the Internet. Such a network, providing transit services to traffic originated in other networks, is thus called a *transit autonomous system*. A transit AS is an AS that exchanges BGP routing information with other ASes and forwards information received from one AS to other ASes.

EGP Requirements (1)

- Multiple expressive policy attributes for each route
- Provide a consistent routing state
  - No forwarding loops
  - No black holes
- Upon physical or policy changes
  - Compute new distributed routing state

(C) Herbert Haas 2005/03/11 14

For EGP it is absolutely nessesary to have a reach metric and a possibility to implement a particular policy.

EGP Requirements (2)

Very Important!

- **Large scale stability**
  - Damp dynamic changes rather than amplify
- **Bandwidth and CPU efficiency**
  - Data volume is very large
- **QoS Routing**
  - "Best-effort" path
  - "Low-latency" path
  - "High-bandwidth" path

EGPs must scale to a huge network like an Internet.

EGPs

- EGP-1
- EGP-2
- BGP-3
- BGP-4

Classful !

Several EGPs were defined, wherethere BGP-4 is the most important today.

Quick EGP-2 Overview

FYI

- Distance Vector Protocol
  - Maximum Hop Count 255
- All ASs must be attached to a Core-AS
  - To avoid count to infinity
  - Does not scale
- Mapping of EGP-Metric to IGP-Metric not defined
  - Vendor/Provider specific
  - Difficult to define global policy rules
- Classful

EGP-2 was the only standard protocol until the early 1990s. It had several disadvantages and would have hampered the evolution of the Internet. The main drawback was its demand to attach all ASs to a single core-AS. Obviously this would not scale.

Furthermore it was a classful protocol, that is, EGP-2 did not support supernets and CIDR.