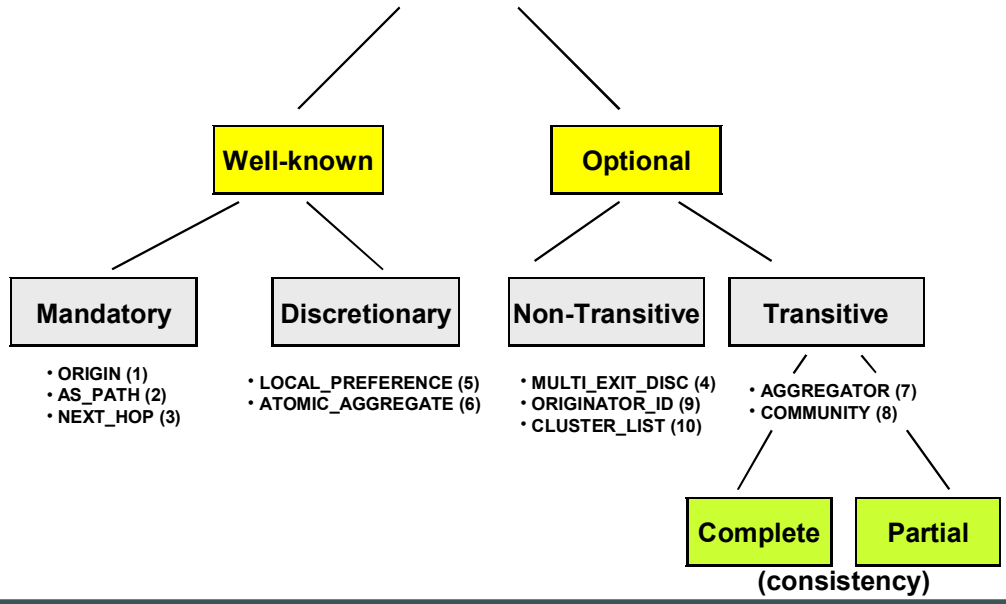
BGP

Attributes

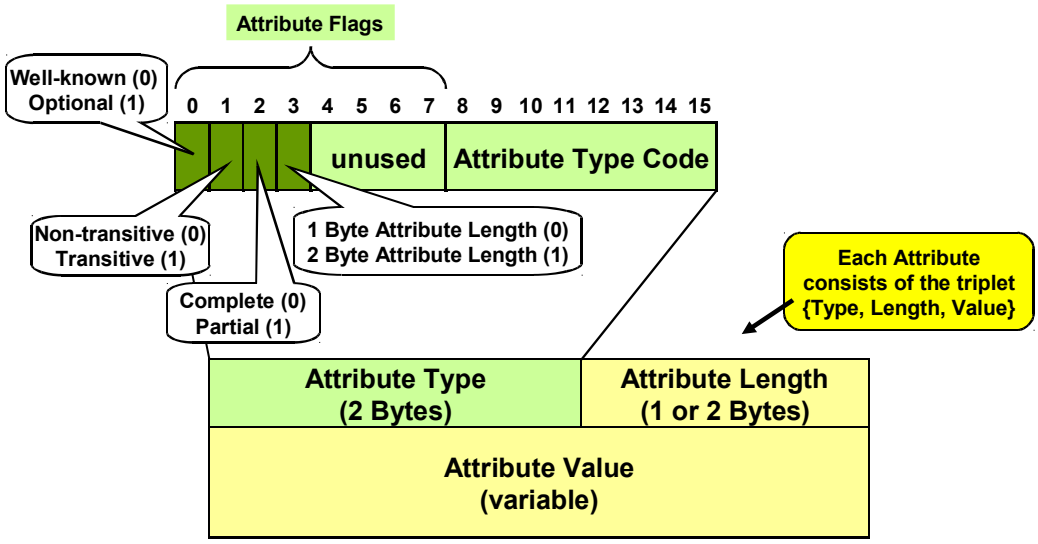
Attribute Types



Each BGP update consists of one or more IP subnets and a set of attributes attached to them. Some of the attributes are required to be recognized by all BGP implementations. Those attributes are called *well-known* BGP attributes.

Attributes that are not well known are called *optional*. These could be attributes specified in a later extension of the BGP protocol or even private vendor extensions not documented in a standard document.

Path Attributes



Each attribute consists of a so called TLVs – Type, Length, Value.

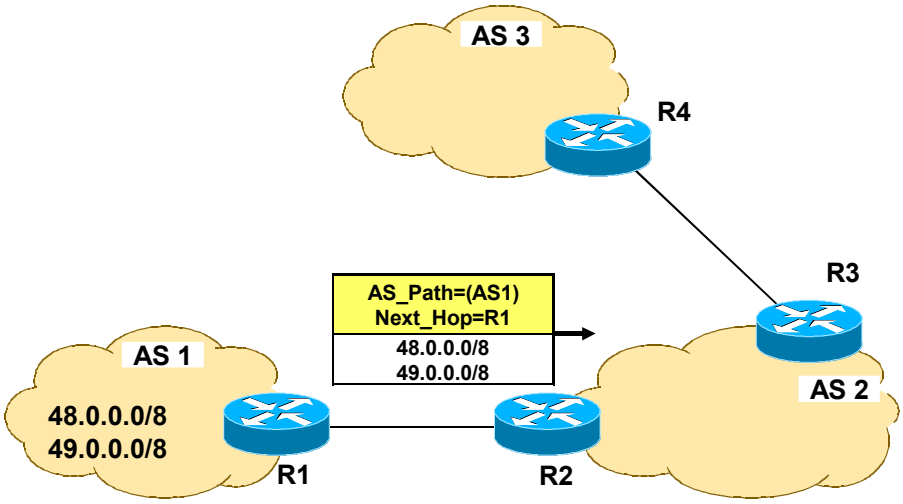
Well-known Mandatory

- AS_Path contains all ASs traversed for this route
- Next_Hop indicates the last EBGP router leading to this route
  - Not necessarily the physical next hop
- Origin indicates how this route was learned

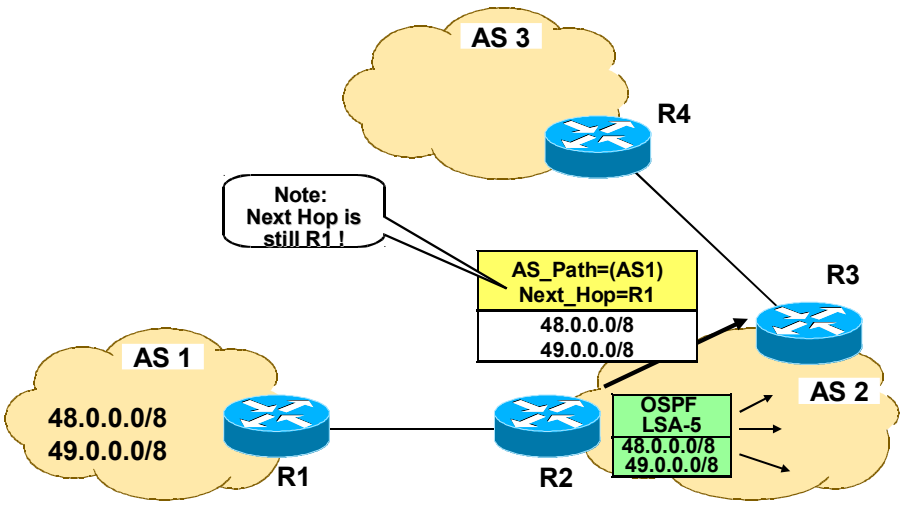
There is a small set of three specific well-known attributes that are required to be present on every update. These three are the **AS-path, next-hop** and **origin** attributes. They are referred to as *well-known mandatory* attributes.

Other well-known attributes may or may not be present depending on the circumstances under which the updates are sent and the desired routing policy. The well-known attributes that could be present, but are not required to be present, are called *well-known discretionary* attributes.

Path Vector Protocol (1)

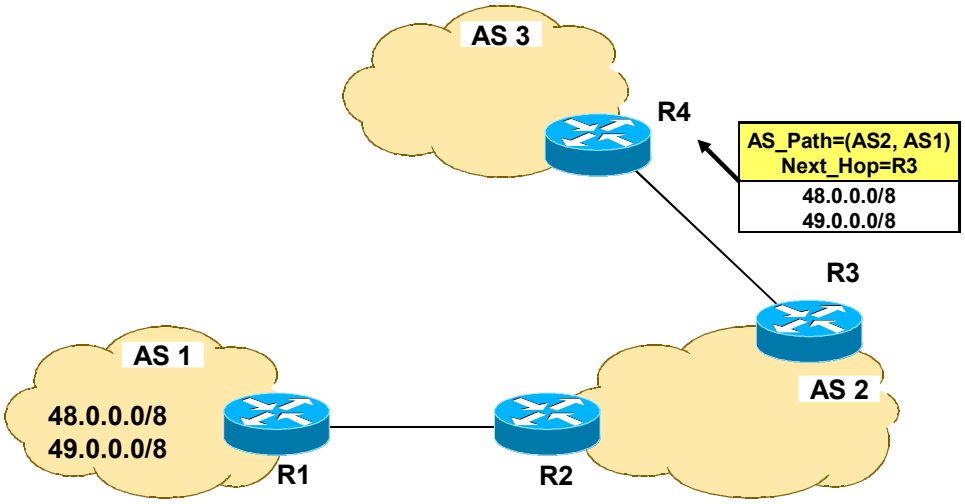


Path Vector Protocol (2)

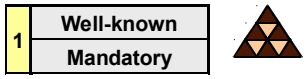


Redistribution into IGP
(e. g. OSPF)

Path Vector Protocol (3)



ORIGIN



- Value 0: IGP
  - Routes learned via network statement (NLRI is member of originating AS)
- Value 1: EGP
  - Learned via redistribution from EGP to BGP
- Value 2: INCOMPLETE
  - Learned via redistribution from IGP to BGP
  - Example: redistribute static (Cisco)

The origin attribute is set when the route is first injected into the BGP. If information about an IP subnet is injected using the **network** command or via aggregation (route-summarization within BGP) the origin attribute is set to *IGP*. If the IP subnet is injected using redistribution, the origin attribute is set to *unknown* or *incomplete* (these two words have the same meaning). The origin code, *EGP,* was used when the Internet was migrating from EGP to BGP and is now obsolete.

AS_PATH

2 | Well-known | Mandatory

- Composed of a sequence of AS path segments
- An AS path segment is represented by a triple
  - Path segment type (1 byte)
    - 1 = AS_Set (unordered set of ASs)
    - 2 = AS_Sequence (ordered set of ASs)
  - Path segment length (1 byte)
  - Path segment value (variable, 2 bytes per AS)

The *AS-path* attribute is modified each time the information about a particular IP subnet passes over an AS border. When the route is first injected into the BGP the AS-path is empty.

Each time the route crosses an AS boundary the transmitting AS prepends its own AS number to appear first in the AS-path. The sequence of ASes, through which the route has passed, can therefore be tracked using the AS-path attribute.

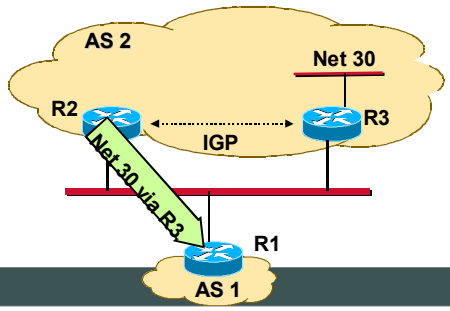
Who is NEXT_HOP?

3	Well-known
	Mandatory

- The boundary router that advertized the route in this AS is the next hop
  - Recursive routing table lookup might be necessary to determine the true physical next hop
- Exception:
  - On multi-access media (Ethernet, FDDI) always the physical next hop must be indicated

R1 and R2 have BGP session established, R3 speaks IGP only.

R2 advertises R3 as next hop to Net 30 because R3 is on the same physical media.

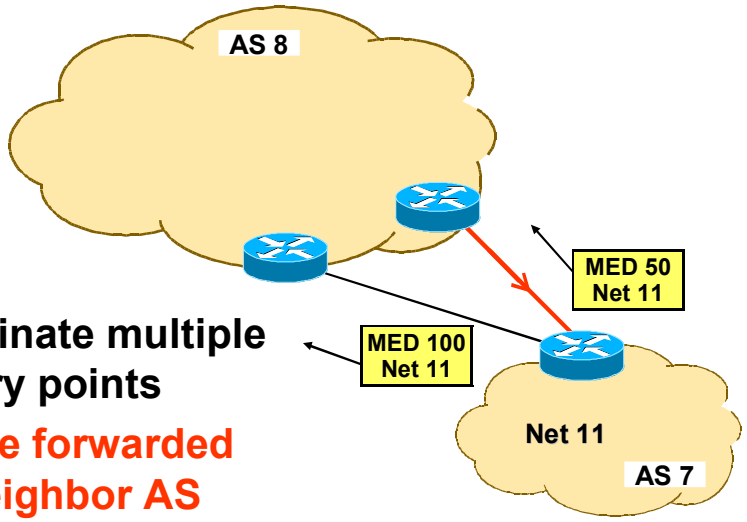


(C) Herbert Haas 2005/03/11 10

The next-hop attribute is also modified as the route passes through the network. It is used to indicate the IP address of the next-hop router—the router to which the receiving router should forward the IP packets toward the destination advertised in the routing update.

MULTI_EXIT_DISC

4 | Optional / Non-transitive



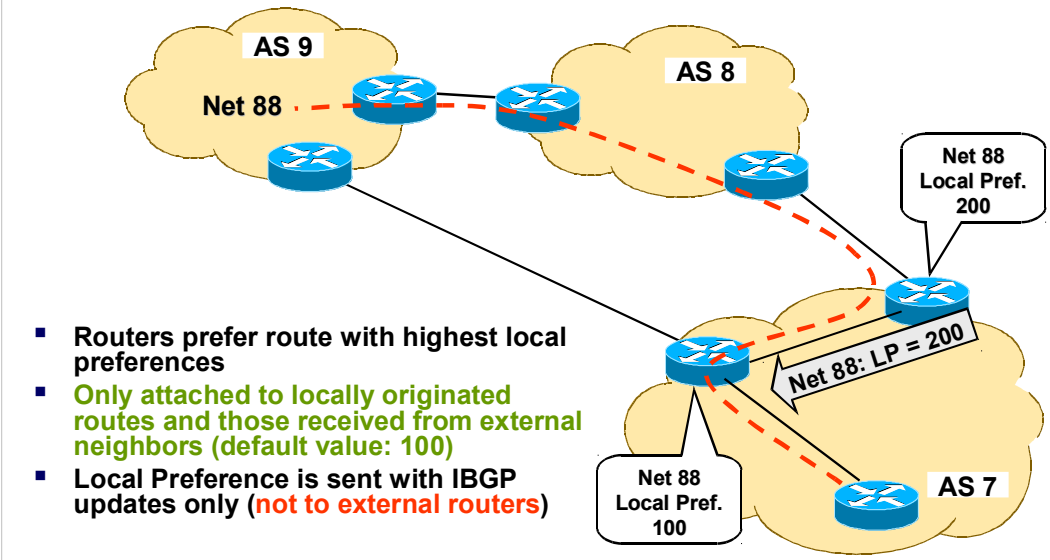
- To discriminate multiple exit or entry points
- Must not be forwarded to other neighbor AS

(C) Herbert Haas 2005/03/11 11

One of the non-transitive optional attributes is the *Multi-Exit-Discriminator* (*MED*) attribute which is also used in the route selection process. Whenever there are several links between two adjacent ASes, multi-exit-discriminator may be used by one AS to tell the other AS to prefer one of the links over the other for specific destinations.

LOCAL_PREF

5	Well-known
	Discretionary



- Routers prefer route with highest local preferences
- Only attached to locally originated routes and those received from external neighbors (default value: 100)
- Local Preference is sent with IBGP updates only (not to external routers)

(C) Herbert Haas 2005/03/11 12

Local Preference is used in the route selection process. The attribute is carried within an AS only. A route with a high local preference is preferred over a route with a low value. By default, routes received from peer AS are tagged with the local preference set to the value 100 before they are entered into the local AS. If this value is changed through BGP configuration, the BGP selection process is influenced. Since all routers within the AS get the attribute along with the route, a consistent routing decision is made throughout the AS.

ATOMIC_AGGREGATE

6	Well-known
	Discretionary

- Optionally the Atomic_Aggregate attribute indicates that some BGP router made an AS aggregation
  - When selecting the less specific route on overlapping routes (rejecting the more specific route)
- Length 0

The **Atomic Aggregate** attribute is attached to a route that is created as a result of route summarization (called aggregation in BGP). It signals that information that was present in the original routing updates may have been lost when the updates where summarized into a single entry.

AGGREGATOR

7	Optional
	Transitive

- Contains the AS number and IP address of the BGP speaker that formed the aggregate route
- Useful for troubleshooting

Aggregator identifies the AS and the router within that AS that created a route summarization, aggregate.

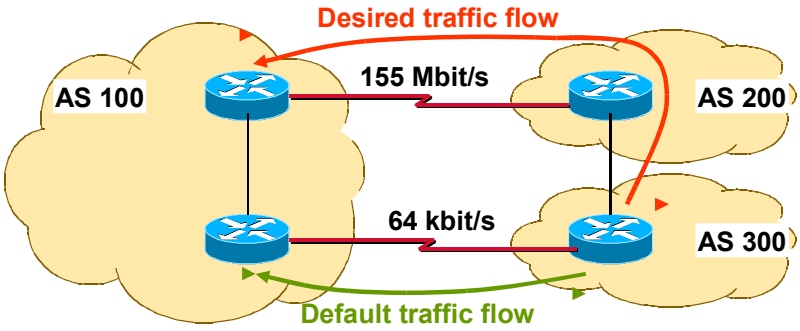
COMMUNITY

8 | Optional | Transitive

- Group of destinations that share a common policy
  - Each destination could be member of multiple communities
  - Carried across ASs
- Community strings are simple policy labels
  - Any BGP router can tag routes in incoming and outgoing routing updates or when doing redistribution
  - Any BGP router can filter routes in incoming or outgoing updates or select preferred routes based on communities

A **Community** is a numerical value that can be attached to certain routes as they pass a specific point in the network. The community value can then be checked at other points in the network for filtering or route selection purposes. BGP configuration may cause routes with a specific community value to be treated differently than others.

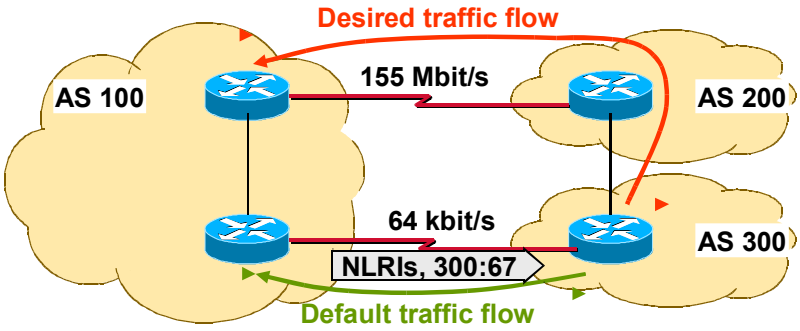
Community Example (1)



- Assume AS 100 wants AS 300 to use the 155 Mbit/s link to reach own networks
  - MED: not possible (non-transitive)
  - Local Preference: will admin of AS 300 set it?
- Best and easiest: Use community !

The picture above gives an example where the comunity could be implemented.

Community Example (2)



- Receiving a community string means "apply the predefined policy"
- In our example 300:67 means: "set local preference to 50"

The picture above gives an example where the comunity could be implemented (continued from previous slide).

Defining Communities

- More than one BGP community per route allowed
  - By default, communities are stripped in outgoing BGP updates
- Private range:
  0x00010000 - 0xFFFEFFFF
- Common practice
  - High order 16 bit: AS number
  - Low order 16 bit: Local significance

Well-known Communities

- **Reserved ranges:** 0x00000000 - 0x0000FFFF and 0xFFFF0000 - 0xFFFFFFFF
- 0xFFFFFF01 means: NO_EXPORT
  - Routes received carrying this value should not be advertised to EBGP peers, except ASs of a confederation
- 0xFFFFFF02 means: NO_ADVERTISE
  - Routes received carrying this value should not be advertised at all (both IBGP and EBGP peers)
- 0xFFFFFF03 means: NO_EXPORT_SUBCONFED
  - Routes received carrying this value should not be adverised to EBGP peers, including members of a confederation (Cisco: LOCAL_AS)

Easy to memorize: Values of all-zeroes and all-ones in high-order 16 bits are reserved.

Administrative Weight (Cisco)

- No attribute – just a local parameter
- Applies only to routes within an individual router
- Number between 0 and 65535
  - The higher the weight the more preferable the route
- Initially invented to translate public routing policies (EGP)

Note that the Administrative Weight is a Cisco specific atrribute.

Decision Hierarchy

1. Prefer highest weight (Cisco)
2. Prefer highest local preference
3. Prefer locally originated routes
4. Prefer shortest AS-Path
5. Prefer lowest origin code
6. Prefer lowest MED
7. Prefer EBGP path over IBGP path
8. Lowest IGP metric to next hop
9. Prefer oldest route for EBGP paths
10. Prefer path with lowest neighbor BGP router ID

If routes have same local preference the route that was locally originated will be preferred. At last the BGP router ID can be used as tie-breaker.