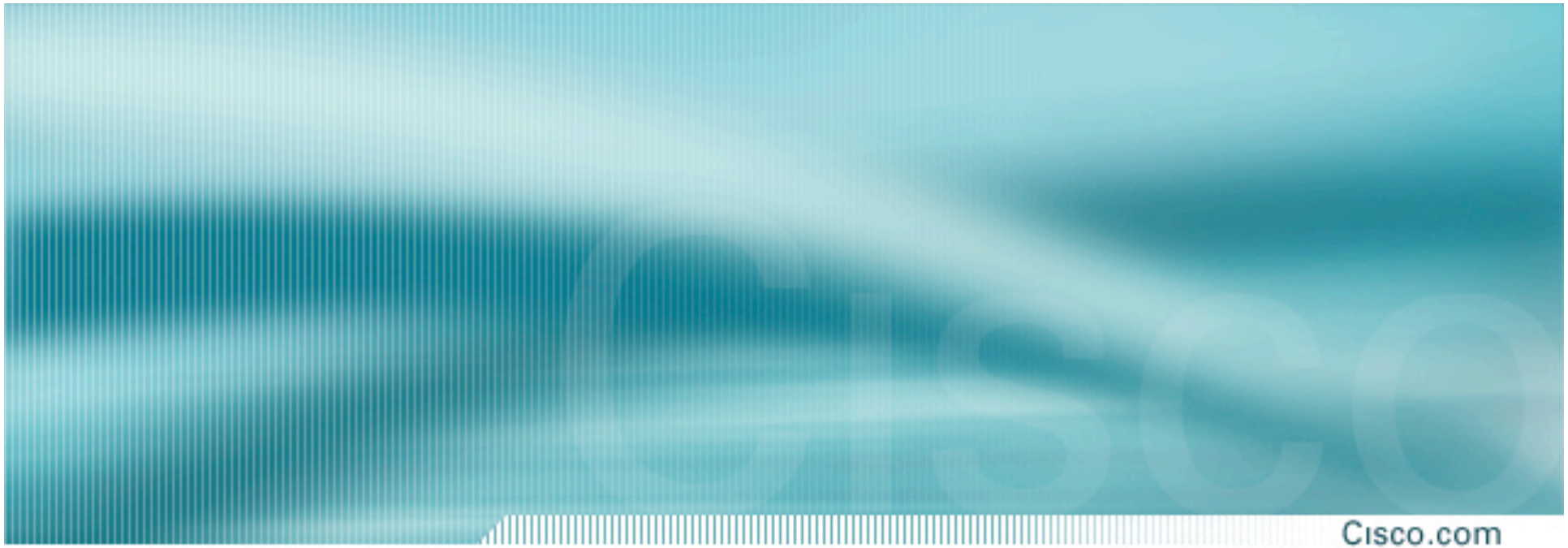


Multiprotocol BGP (MBGP)

Agenda

- MBGP Overview
- MBGP Update Messages
- MBGP Capability Negotiation
- MBGP NLRI Information
- Advanced MBGP Features

MBGP Overview

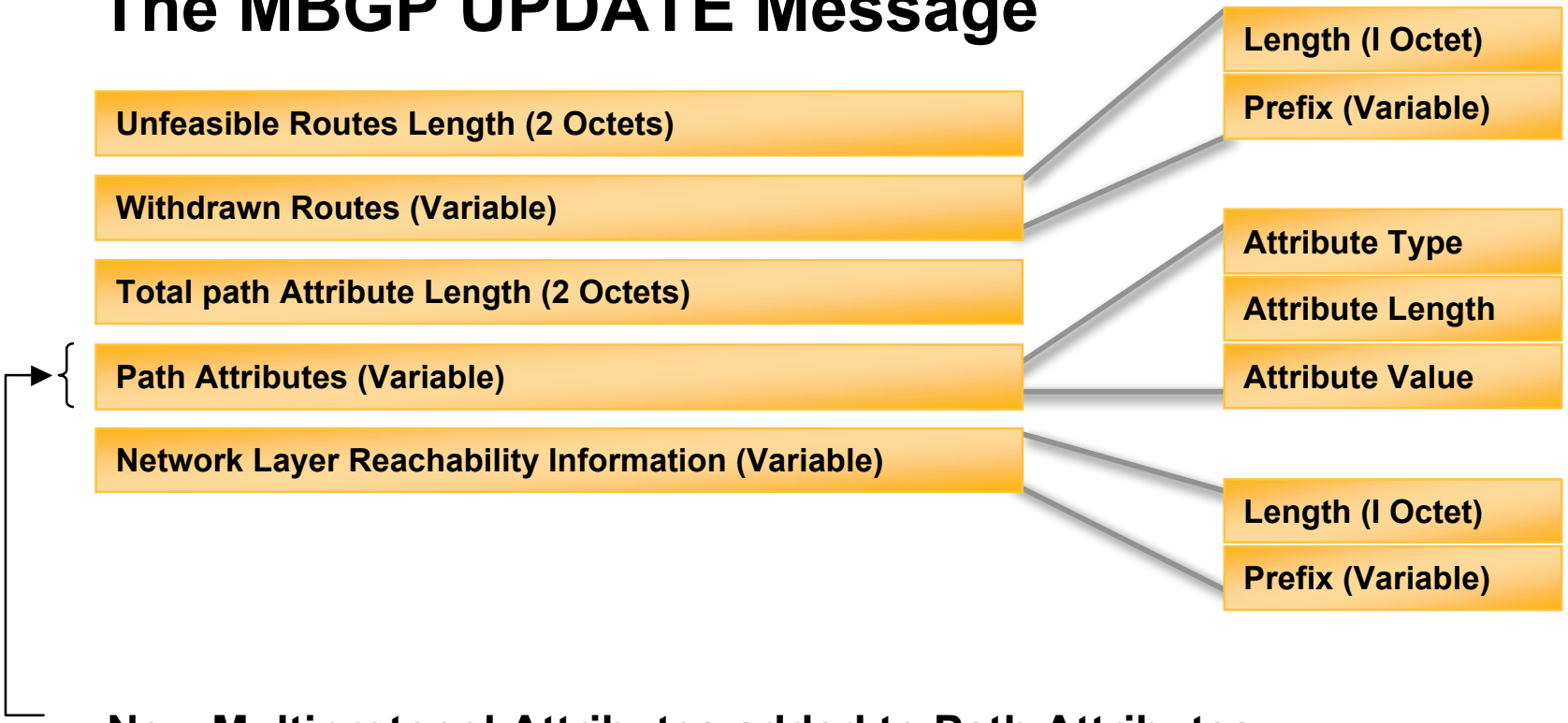
- MBGP: Multiprotocol BGP
  - Defined in RFC 2283 (extensions to BGP)
  - Can carry different types of routes
    - IPv4 Unicast
    - IPv4 Multicast
    - IPv6 Unicast
  - May be carried in same BGP session
  - Does not propagate multicast state info
    - Still need PIM to build Distribution Trees
  - Same path selection and validation rules
    - AS-Path, LocalPref, MED, …

MBGP Overview

- Separate BGP tables maintained
  - Unicast BGP Table (U-Table)
  - Multicast BGP Table (M-Table)
  - New BGP ‘nlri’ keyword specifies which BGP Table
  - Allows different unicast/multicast topologies or policies
- Unicast BGP Table (U-Table)
  - Contains unicast prefixes for unicast forwarding
  - Populated with BGP unicast NLRI
- Multicast BGP Table (M-Table)
  - Contains unicast prefixes for RPF checking
  - Populated with BGP multicast NLRI

MBGP Update Message

The MBGP UPDATE Message



- New Multiprotocol Attributes added to Path Attributes:
  - MP_REACH_NLRI
  - MP_UNREACH_NLRI

MBGP Update Message

- **Address Family Information (AFI)**
  - **Identifies Address Type (see RFC1700)**
    - **AFI = 1 (IPv4)**
    - **AFI = 2 (IPv6)**
- **Sub-Address Family Information (Sub-AFI)**
  - **Sub category for AFI Field**
  - **Address Family Information (AFI) = 1 (IPv4)**
    - **Sub-AFI = 1 (NLRI is used for unicast)**
    - **Sub-AFI = 2 (NLRI is used for multicast RPF check)**
    - **Sub-AFI = 3 (NLRI is used for both unicast and multicast RPF check)**

MBGP—Capability Negotiation

- BGP routers establish BGP sessions through the OPEN message
- OPEN message contains optional parameters
- BGP session is terminated if OPEN parameters are not recognised
- New parameter: CAPABILITIES
  - Multiprotocol extension
  - Multiple routes for same destination

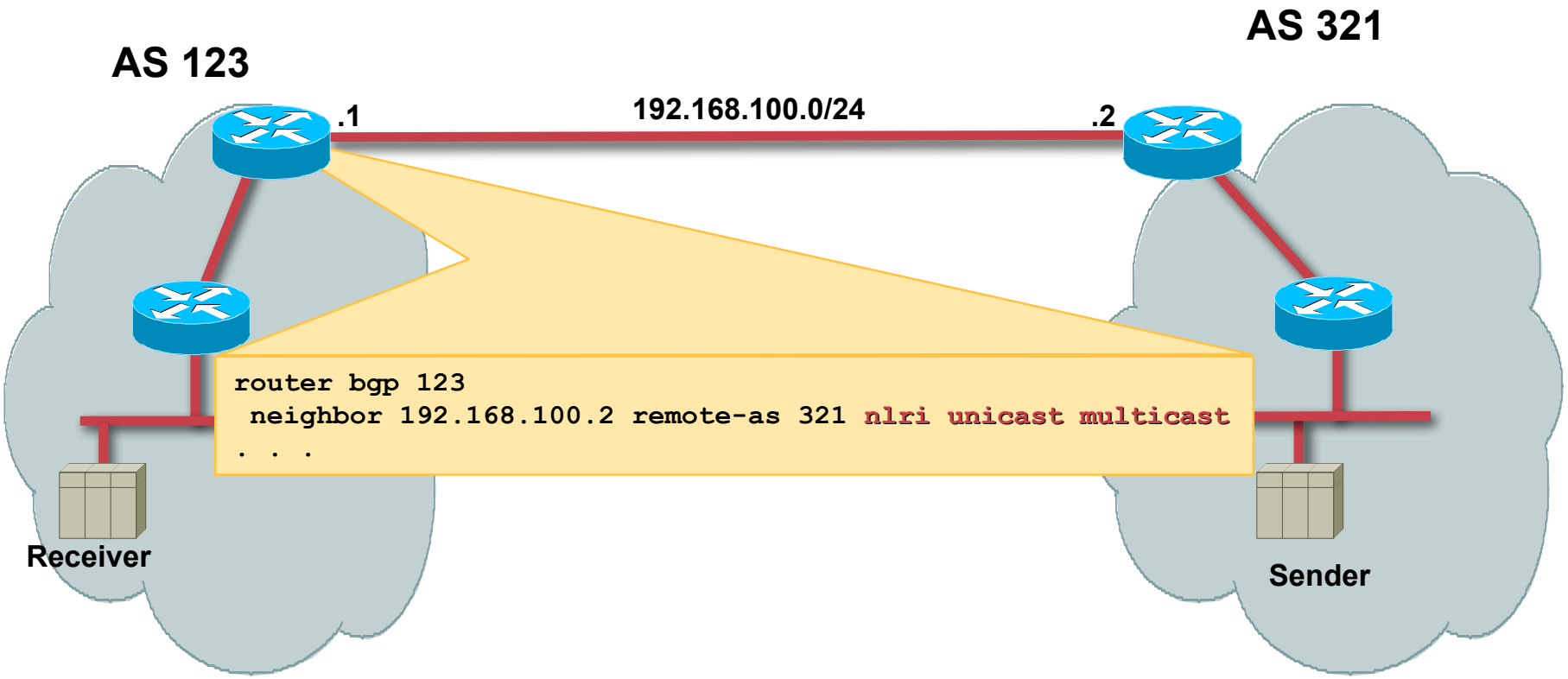
MBGP—Capability Negotiation

- New “nlri” keyword on “neighbor” command

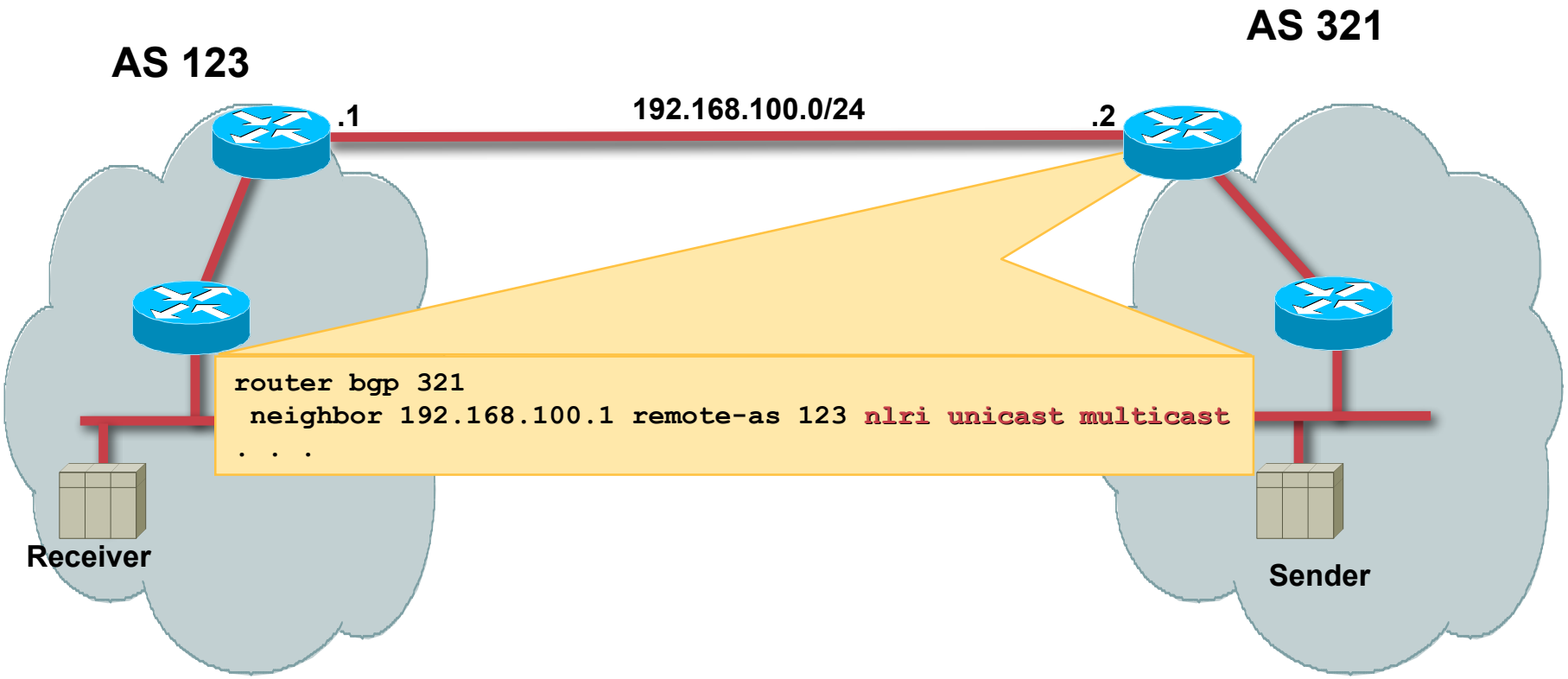
```
neighbor <foo> remote-as <asn> nlri multicast unicast
```

- Configures router to negotiate either or both types of NLRI
- If neighbor configures both or subset, common NRLI is used in both directions
- If there is no match, notification is sent and peering doesn’t come up

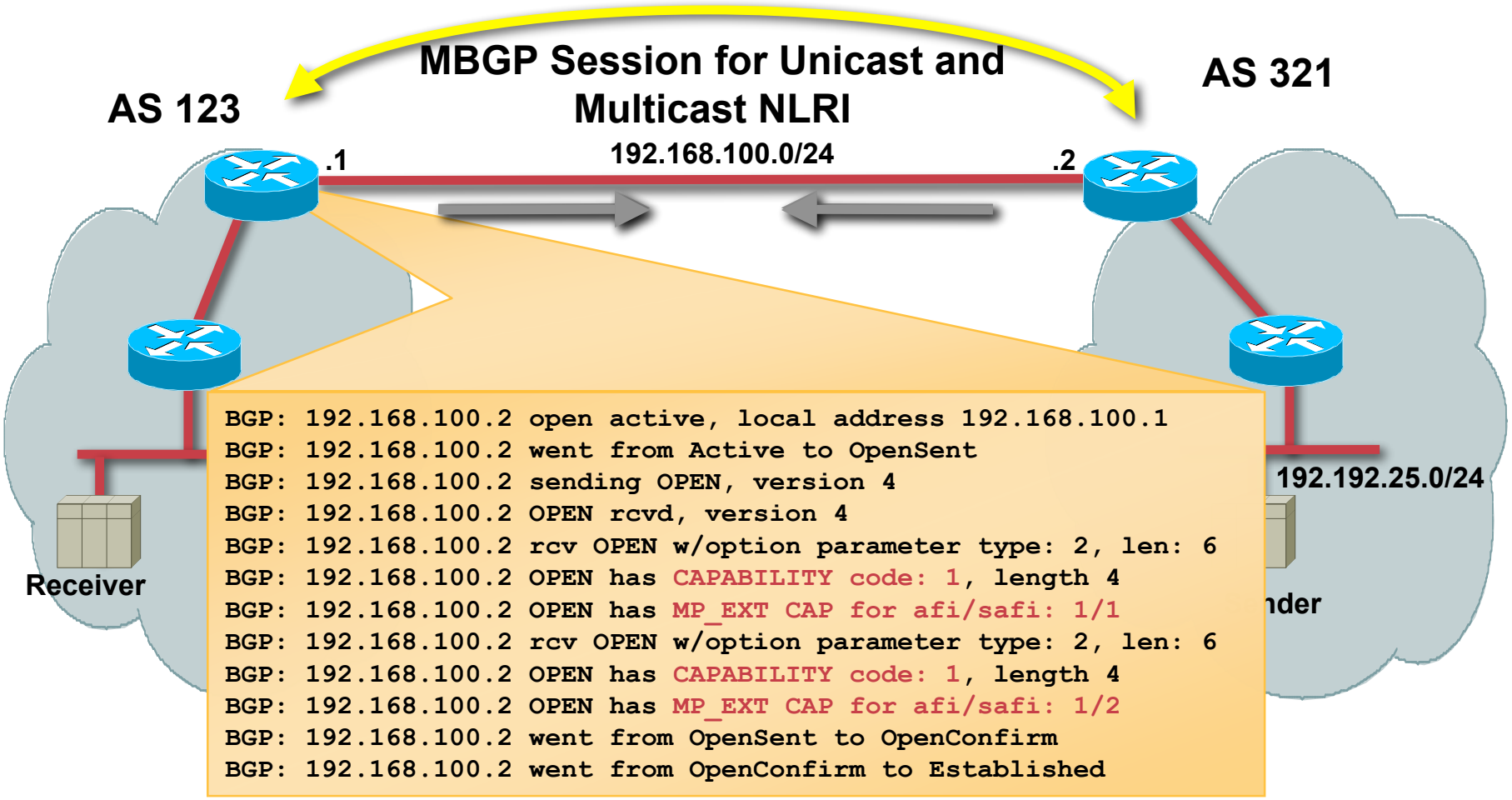
MBGP — Capability Negotiation



MBGP — Capability Negotiation



MBGP — Capability Negotiation



MBGP — Capability Negotiation

- If neighbor doesn't include the CAPABILITY parameters in open, Cisco backs off and reopens with no capability parameters
- Peering comes up in unicast-only mode
- Hidden command

```
neighbor <foo> dont-capability-negotiate
```

MBGP — NLRI Information

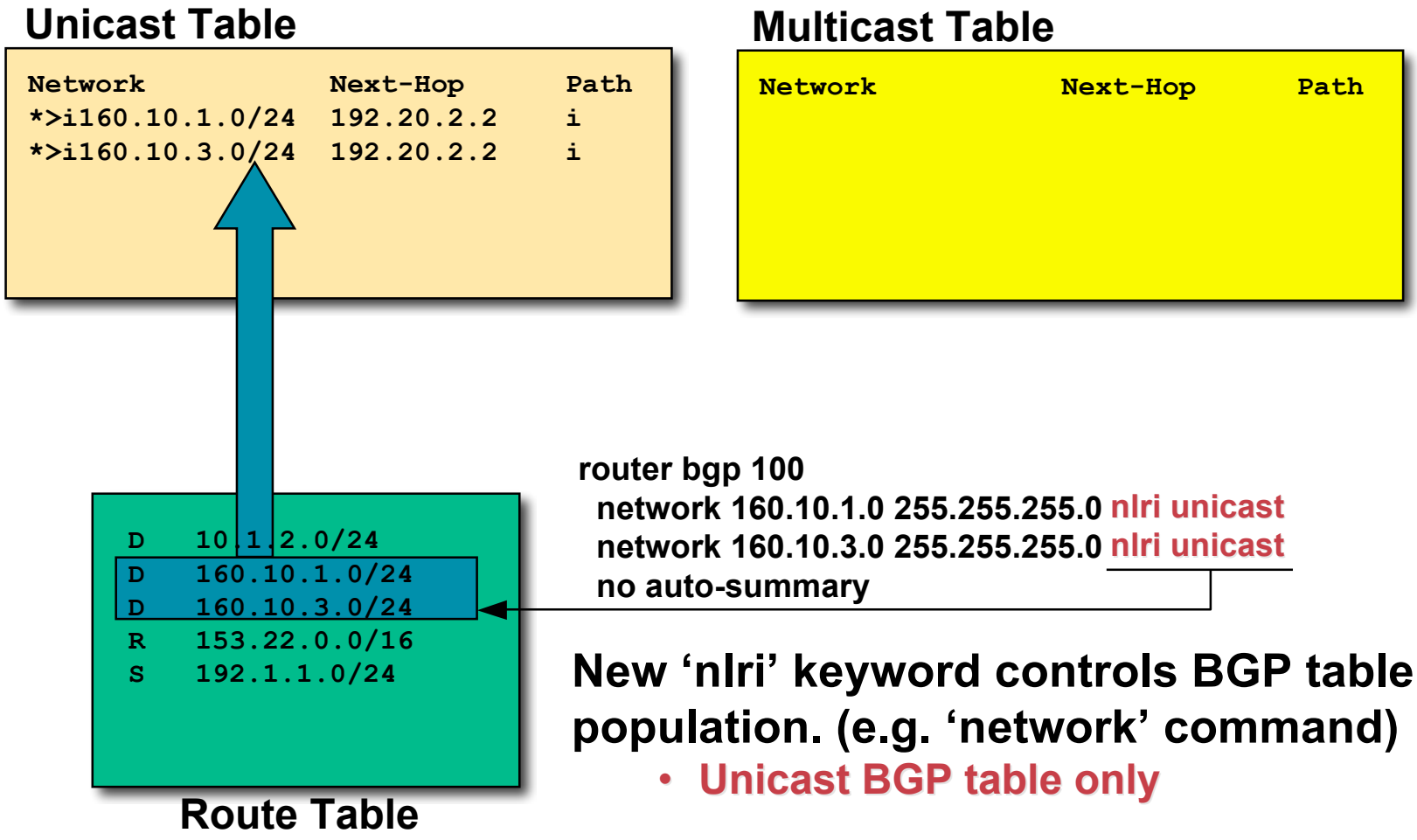
BGP Tables may be populated by:

- **Network commands**

```
network <foo> <foo-mask> [nlri multicast unicast]
```

  - **New "nlri" keyword controls in which BGP table the matching route(s) is(are) stored**
    - **M-Table if "multicast" keyword specified**
    - **U-Table if "unicast" keyword specified (or if nlri clause omitted)**
    - **Both BGP Tables if both keywords specified**

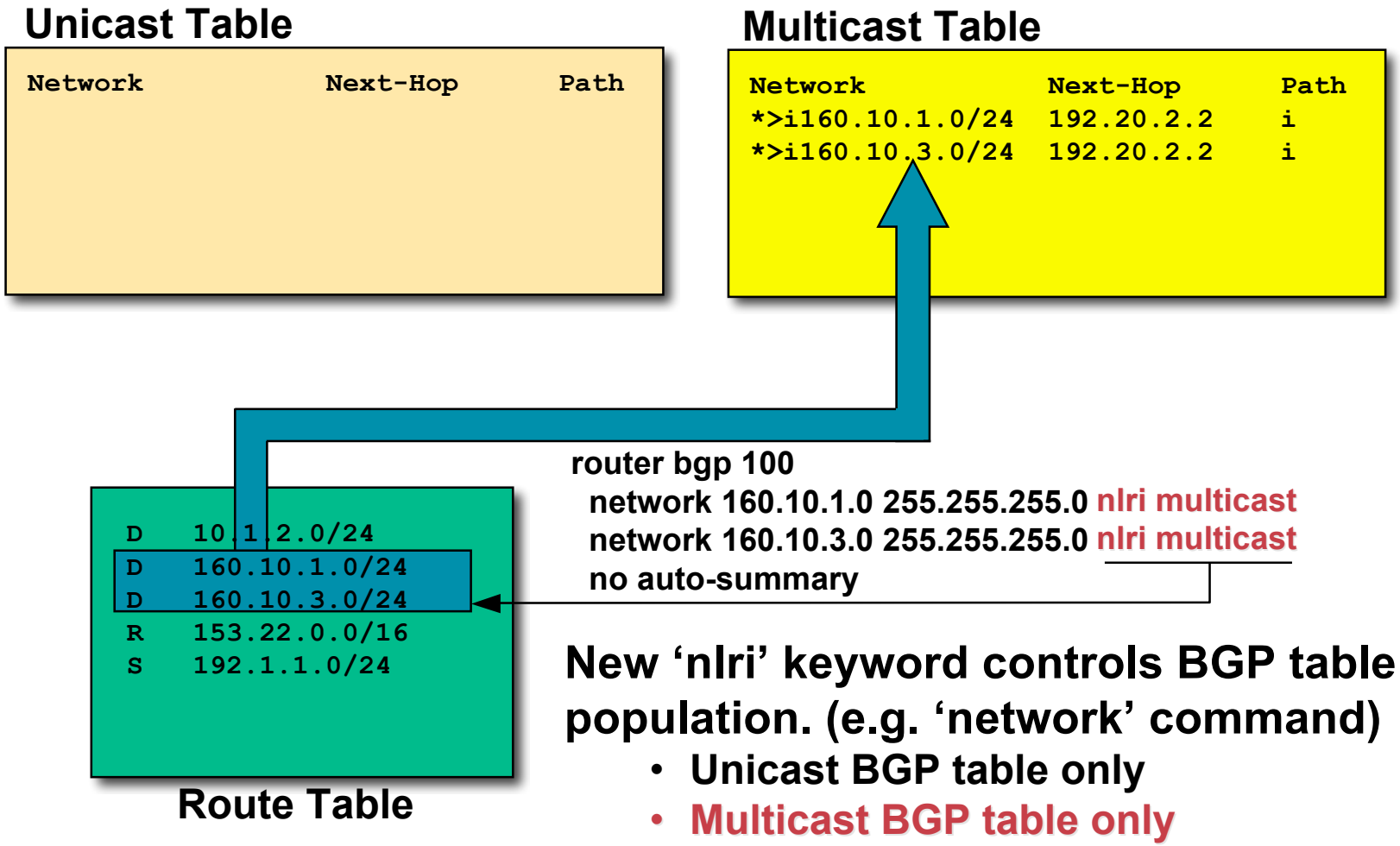
MBGP — NLRI Information



New ‘nlri’ keyword controls BGP table population. (e.g. ‘network’ command)

- **Unicast BGP table only**

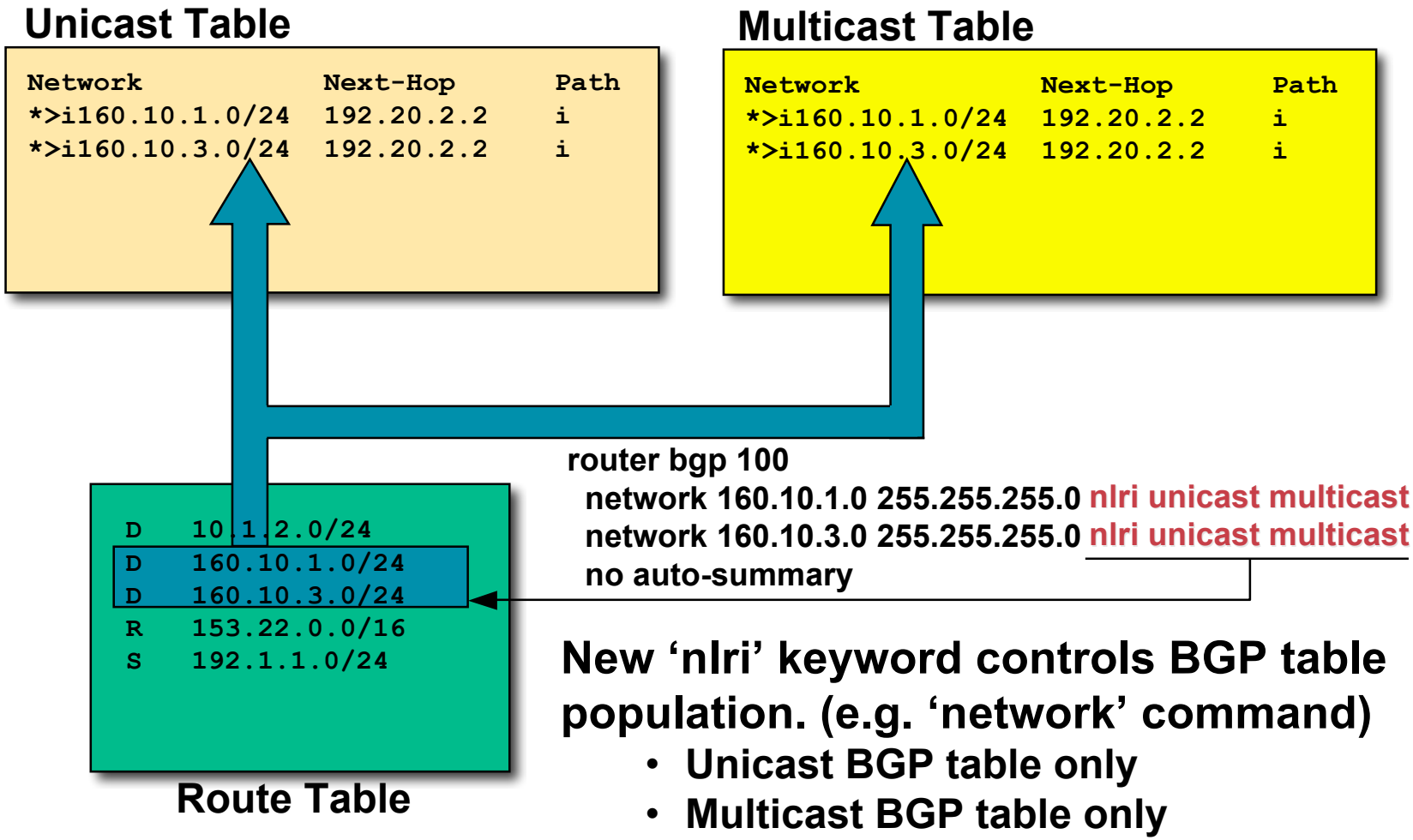
MBGP — NLRI Information



New 'nlri' keyword controls BGP table population. (e.g. 'network' command)

- **Unicast BGP table only**
- **Multicast BGP table only**

MBGP — NLRI Information



New ‘nlri’ keyword controls BGP table population. (e.g. ‘network’ command)

- **Unicast BGP table only**
- **Multicast BGP table only**
- **Both BGP tables**

MBGP — NLRI Information

- **Other "nlri" keyword commands**
  - **Aggregation**

```
aggregate-address <foo> <foo-mask> [nlri multicast unicast]
```

    - **Generates an aggregate route for network <foo>**
  - **In Route Maps**

```
match nlri multicast unicast
```

    - **Matches on the NLRI type**

```
set nlri multicast unicast
```

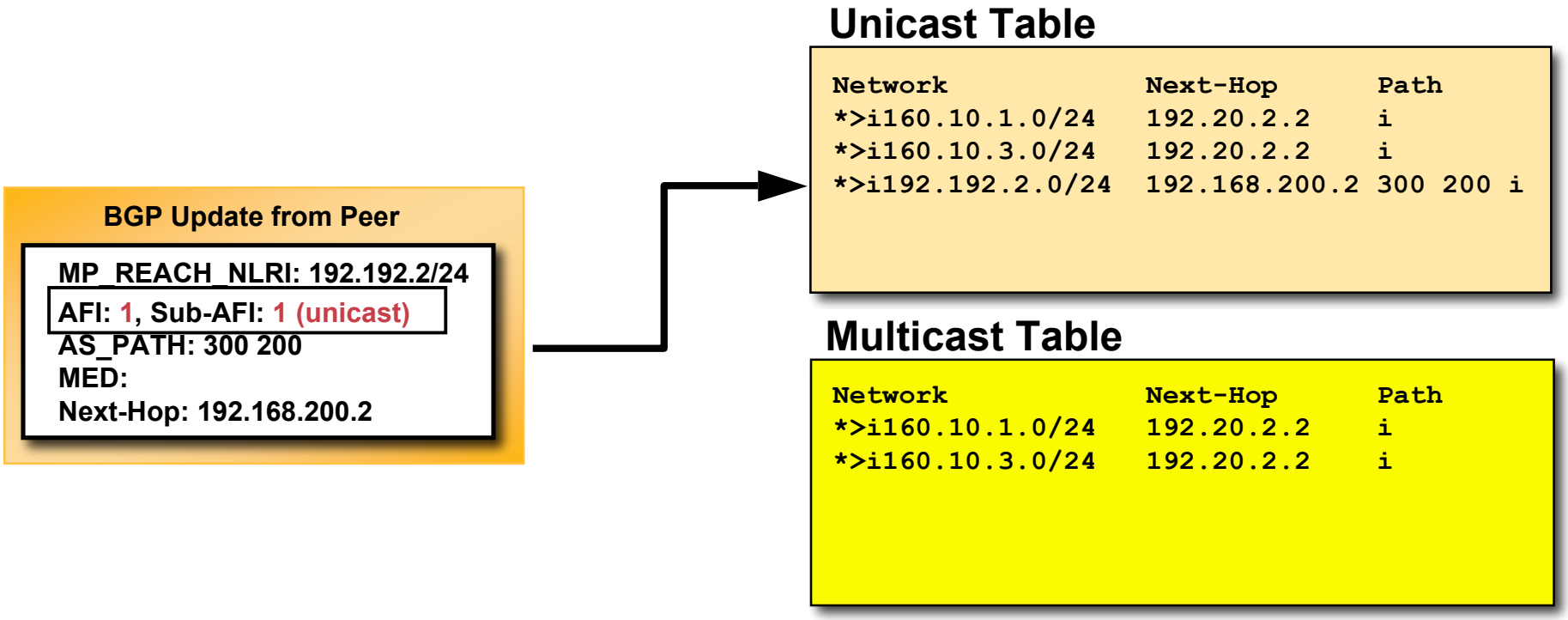
    - **Injects the matched route into the specified unicast or multicast BGP table**

MBGP — NLRI Information

BGP tables are also populated by:

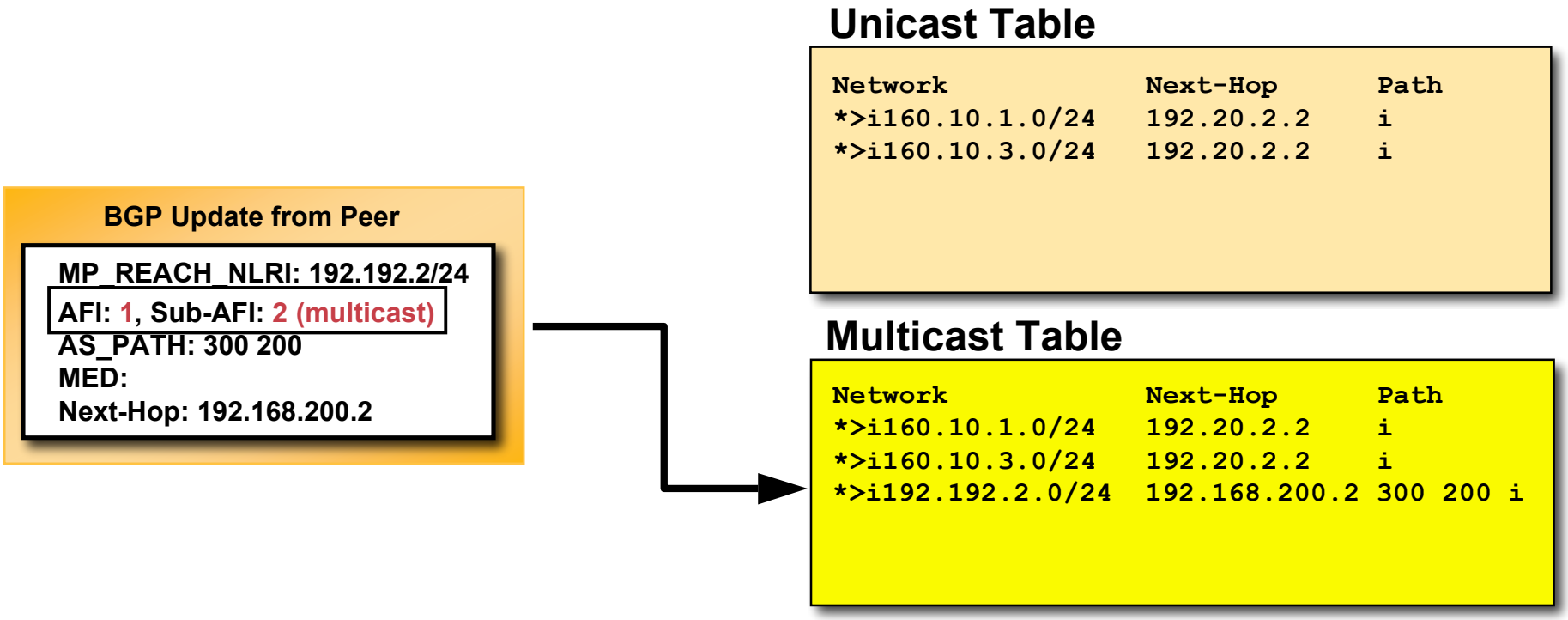
- Receiving MP_REACH_NLRI from Peers
  - Storage controlled by AFI/SAFI value:
    - AFI/SAFI = 1/1 (IPv4 / Unicast) : U-Table only
    - AFI/SAFI = 1/2 (IPv4 / Multicast) : M-Table only
    - AFI/SAFI = 1/3 (IPv4 / Unicast-Multicast) : Both Tables
- Receiving NLRI (old style) from Peers
  - Stored in U-Table only

MBGP — NLRI Information



- Storage of arriving NLRI information depends on AFI/SAFI fields in the Update message
  - Unicast Table only (AFI=1/SAFI=1 or old style NLRI)

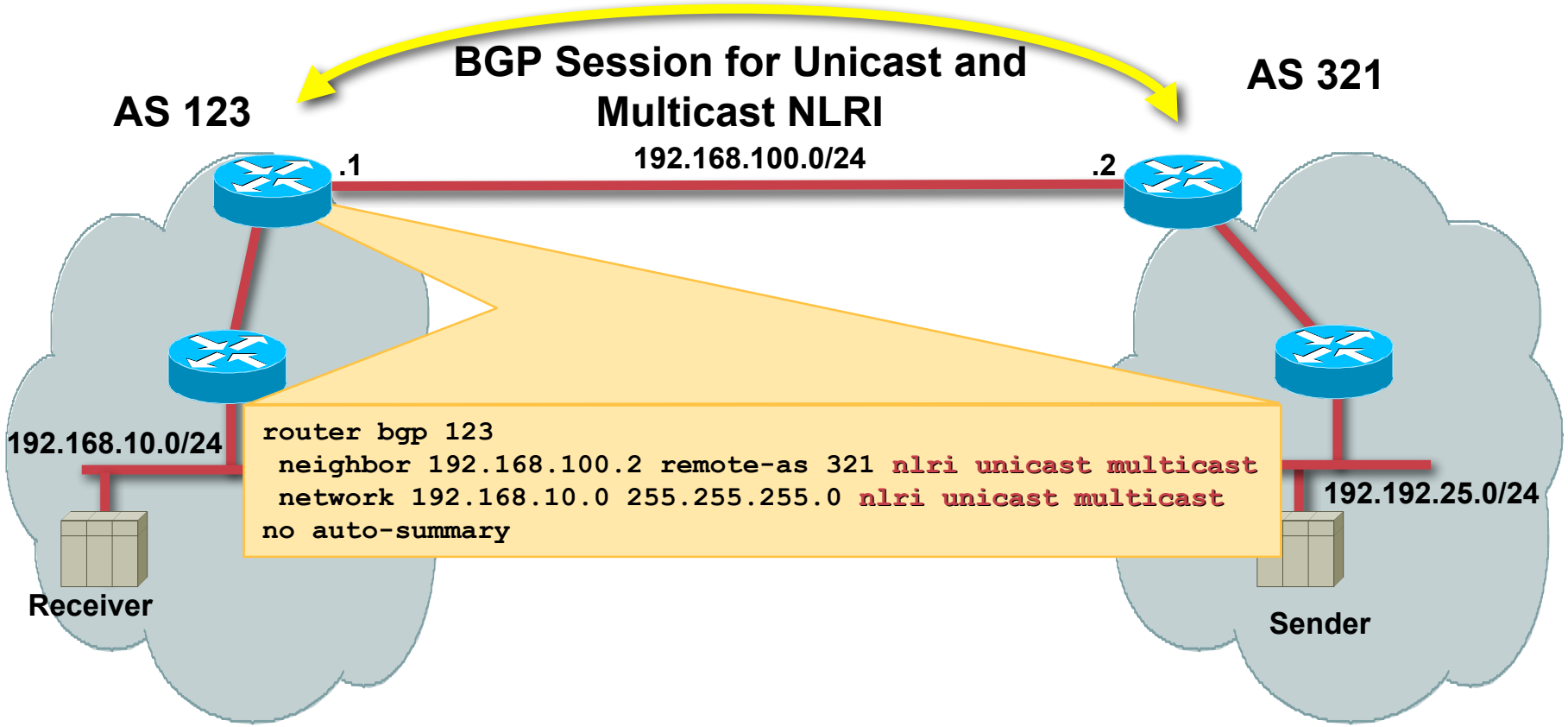
MBGP — NLRI Information



- Storage of arriving NLRI information depends on AFI/SAFI fields in the Update message
  - Unicast Table only (AFI=1/SAFI=1 or old style NLRI)
  - Multicast Table only (AFI=1/SAFI=2)

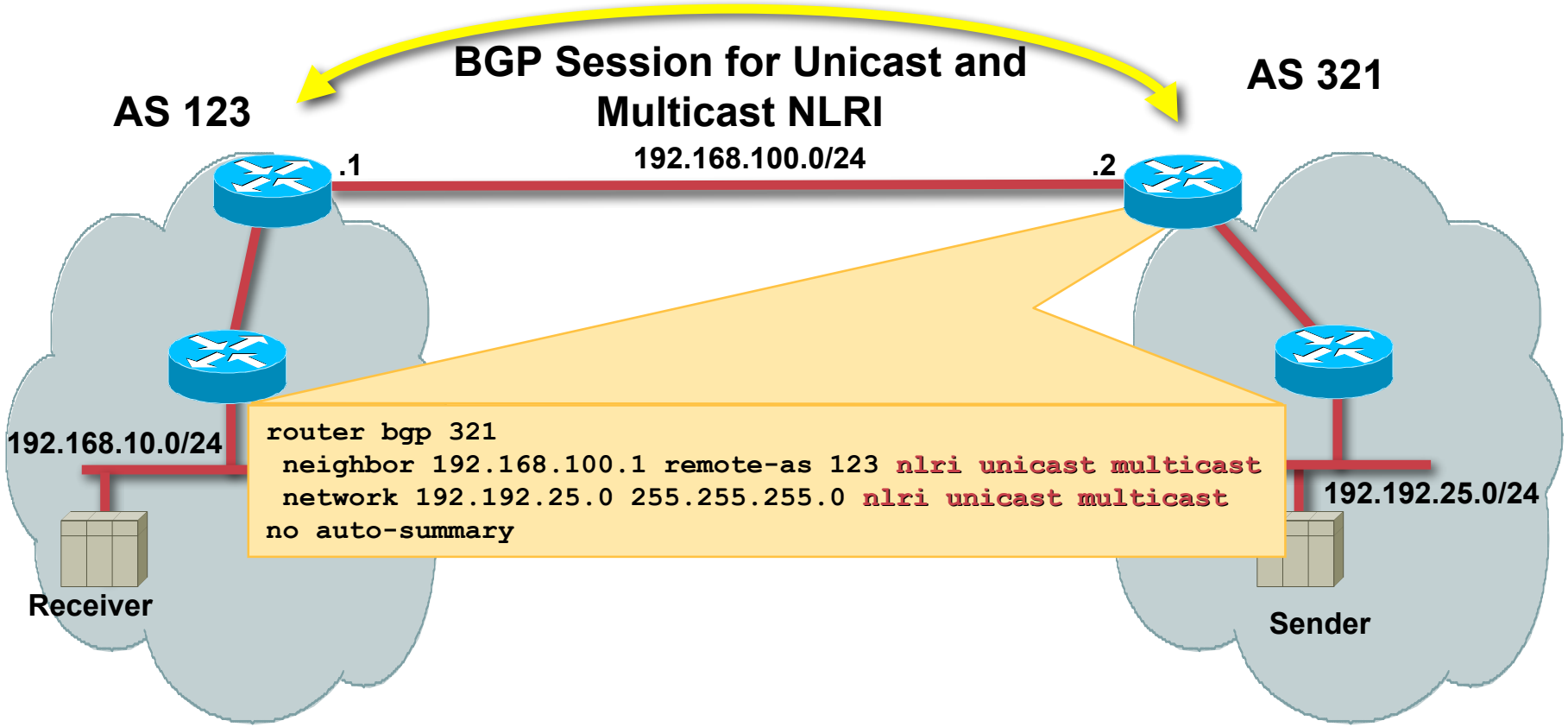
MBGP — NLRI Information

Congruent Topologies



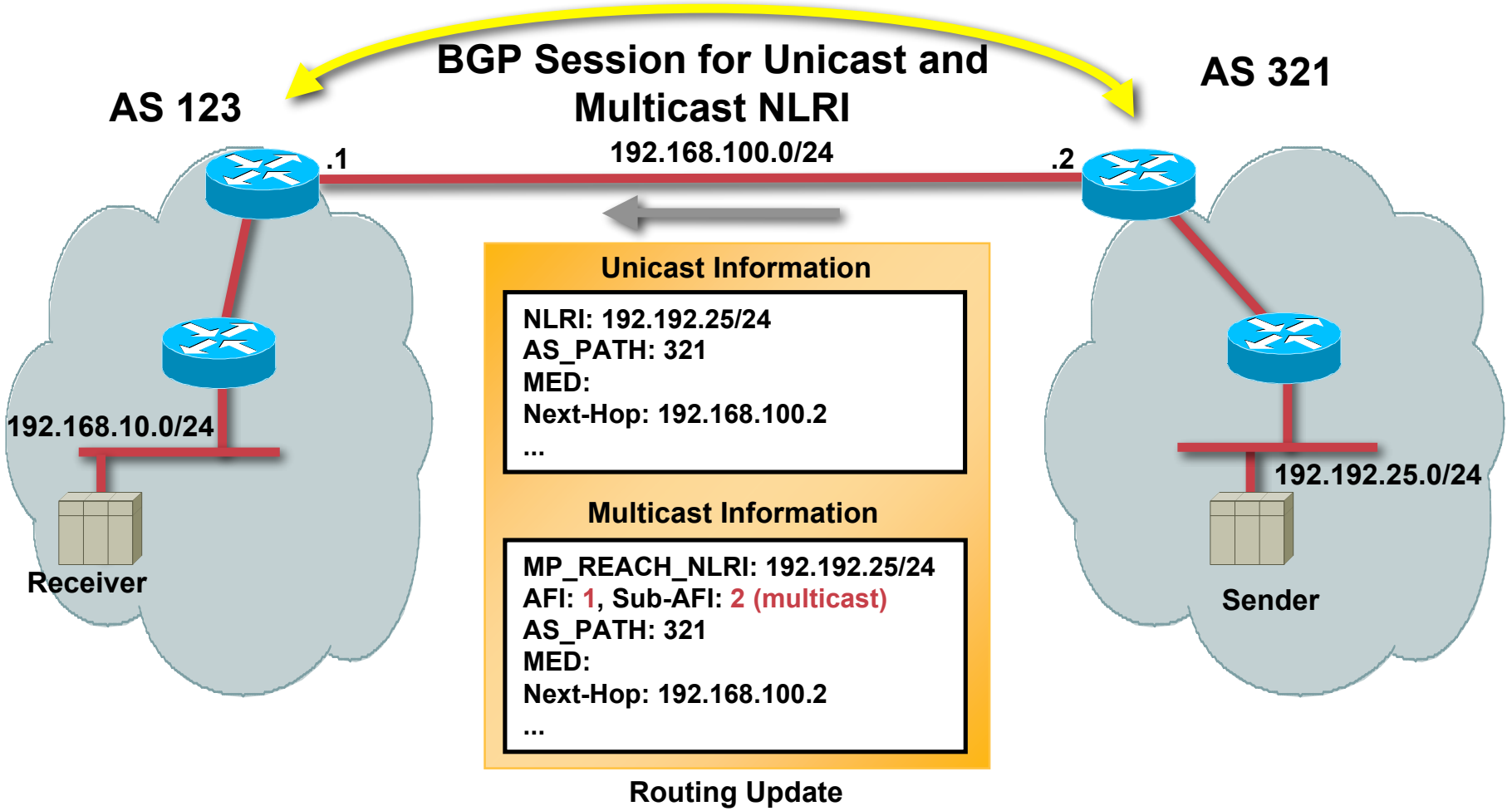
MBGP — NLRI Information

Congruent Topologies



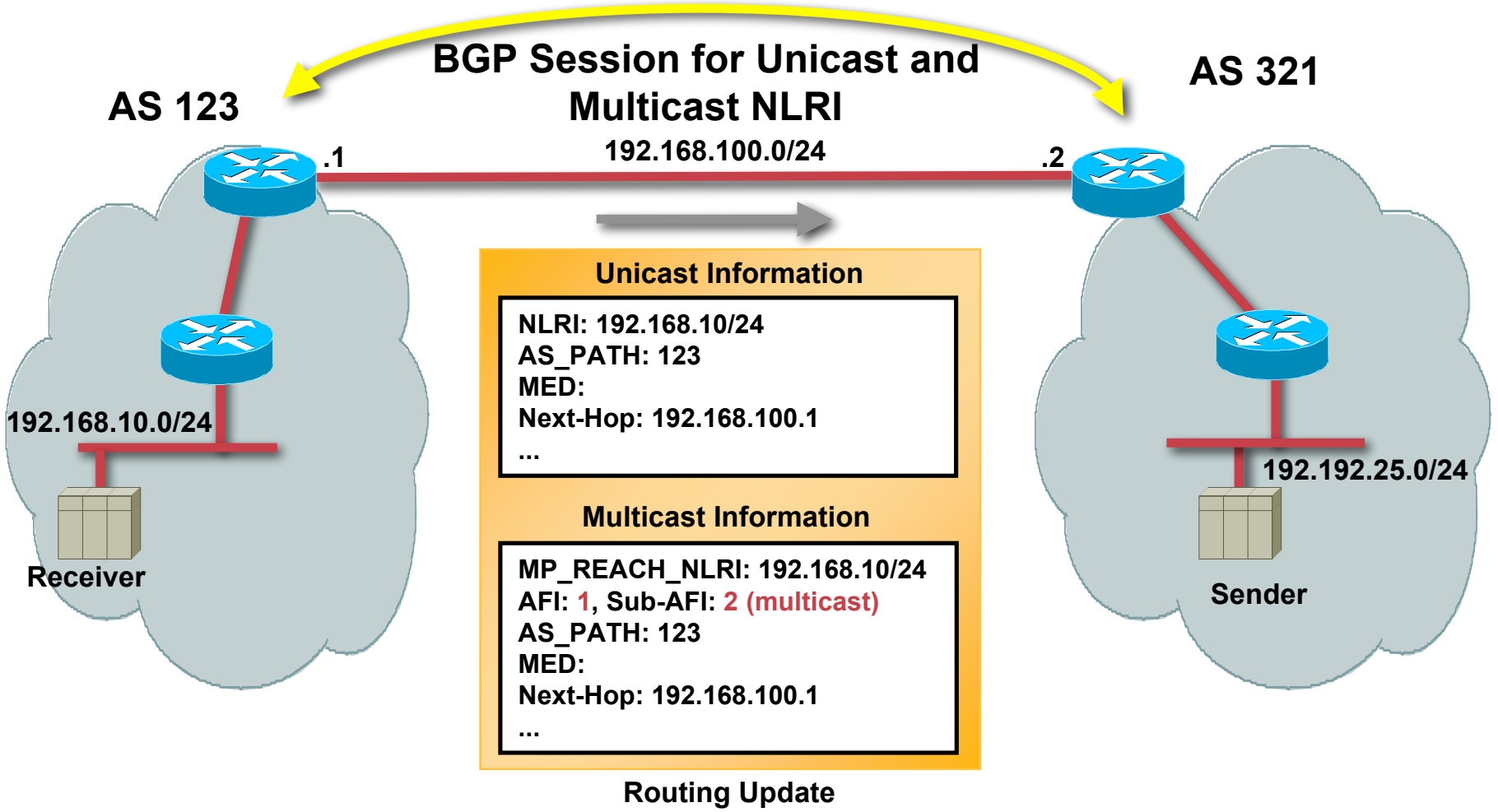
MBGP — NLRI Information

Congruent Topologies



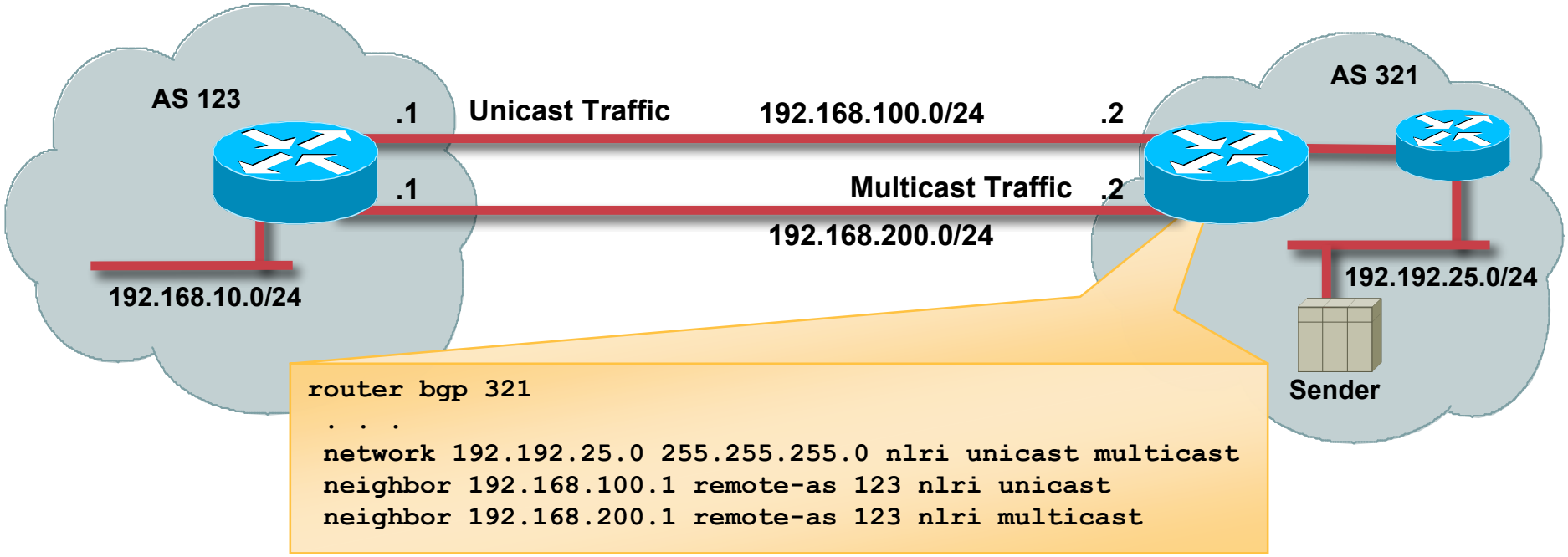
MBGP — NLRI Information

Congruent Topologies



MBGP — NLRI Information

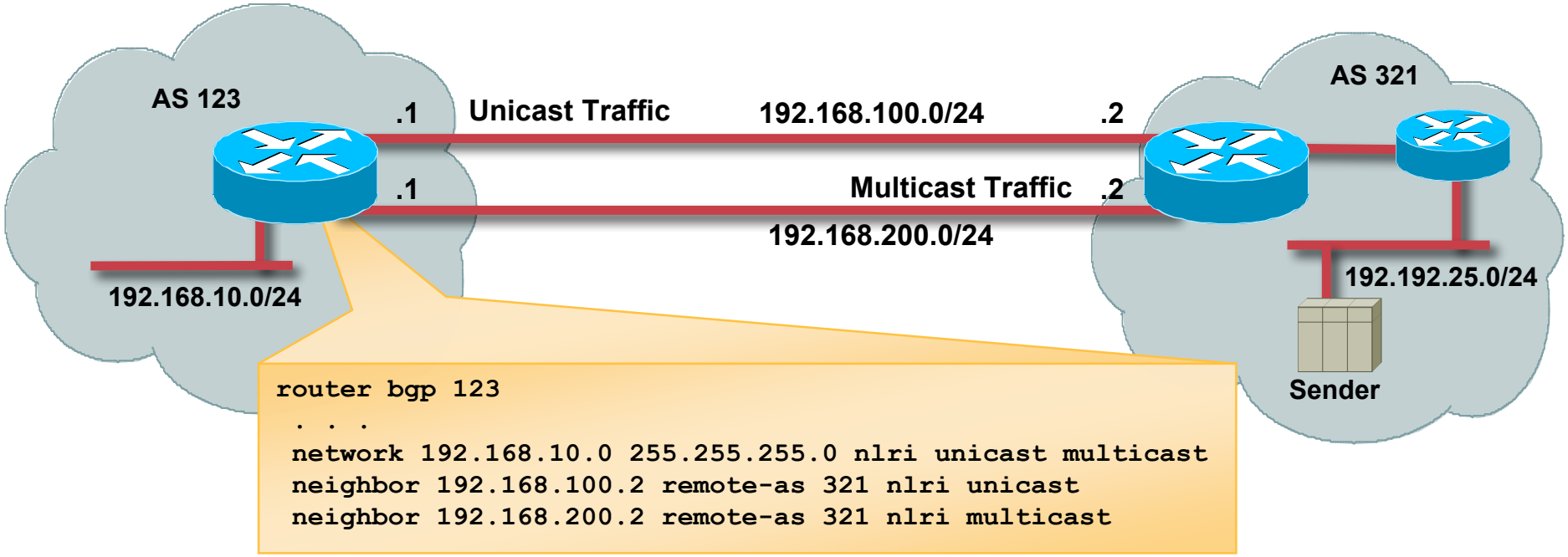
Incongruent Topologies



MBGP — NLRI Information

Cisco.com

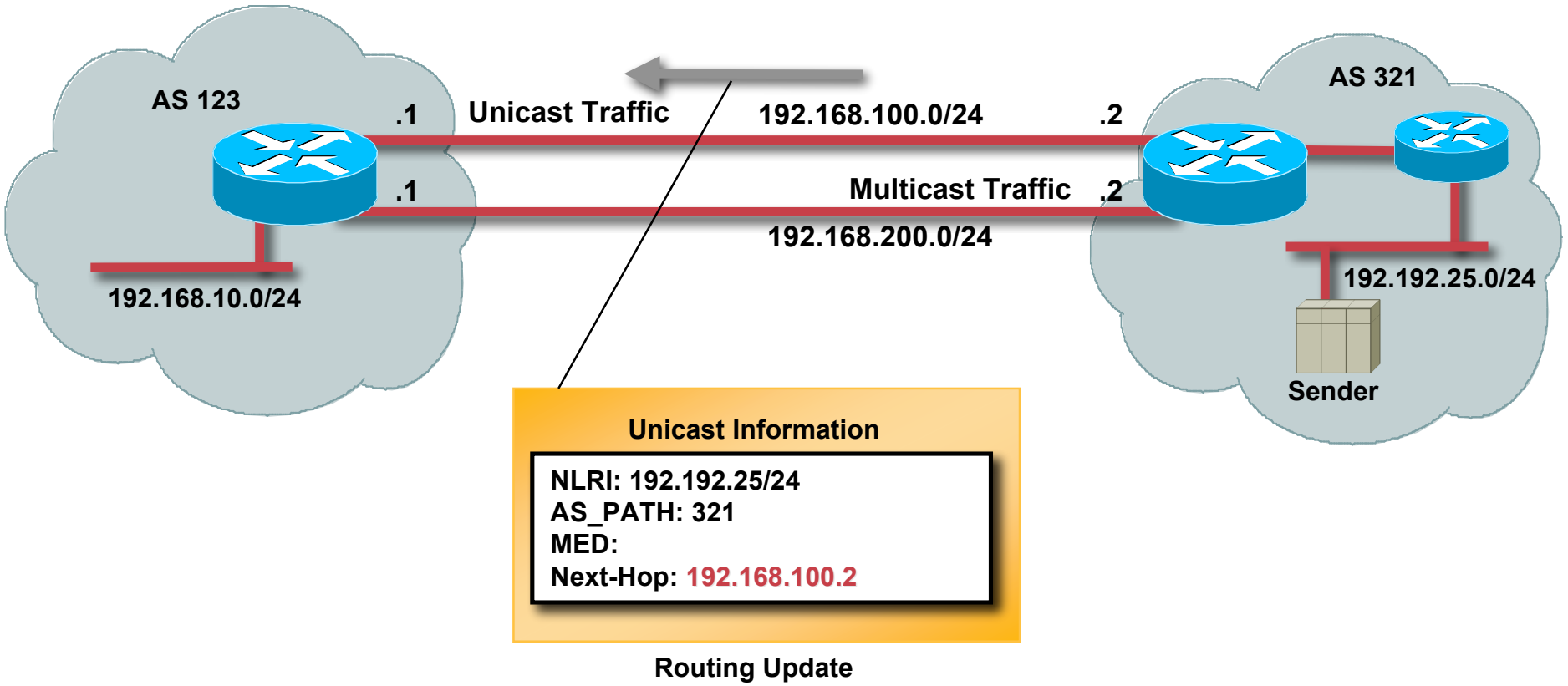
Incongruent Topologies



MBGP — NLRI Information

Cisco.com

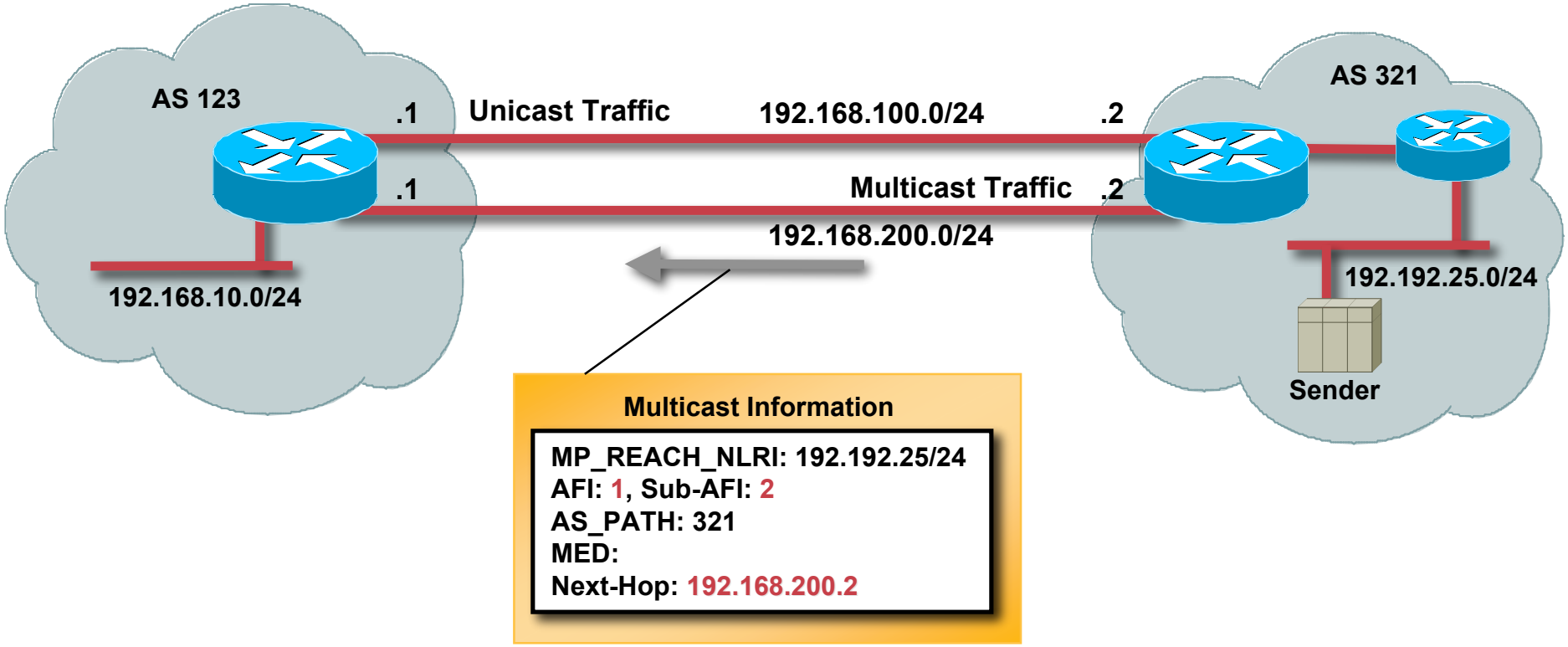
Incongruent Topologies



MBGP — NLRI Information

Cisco.com

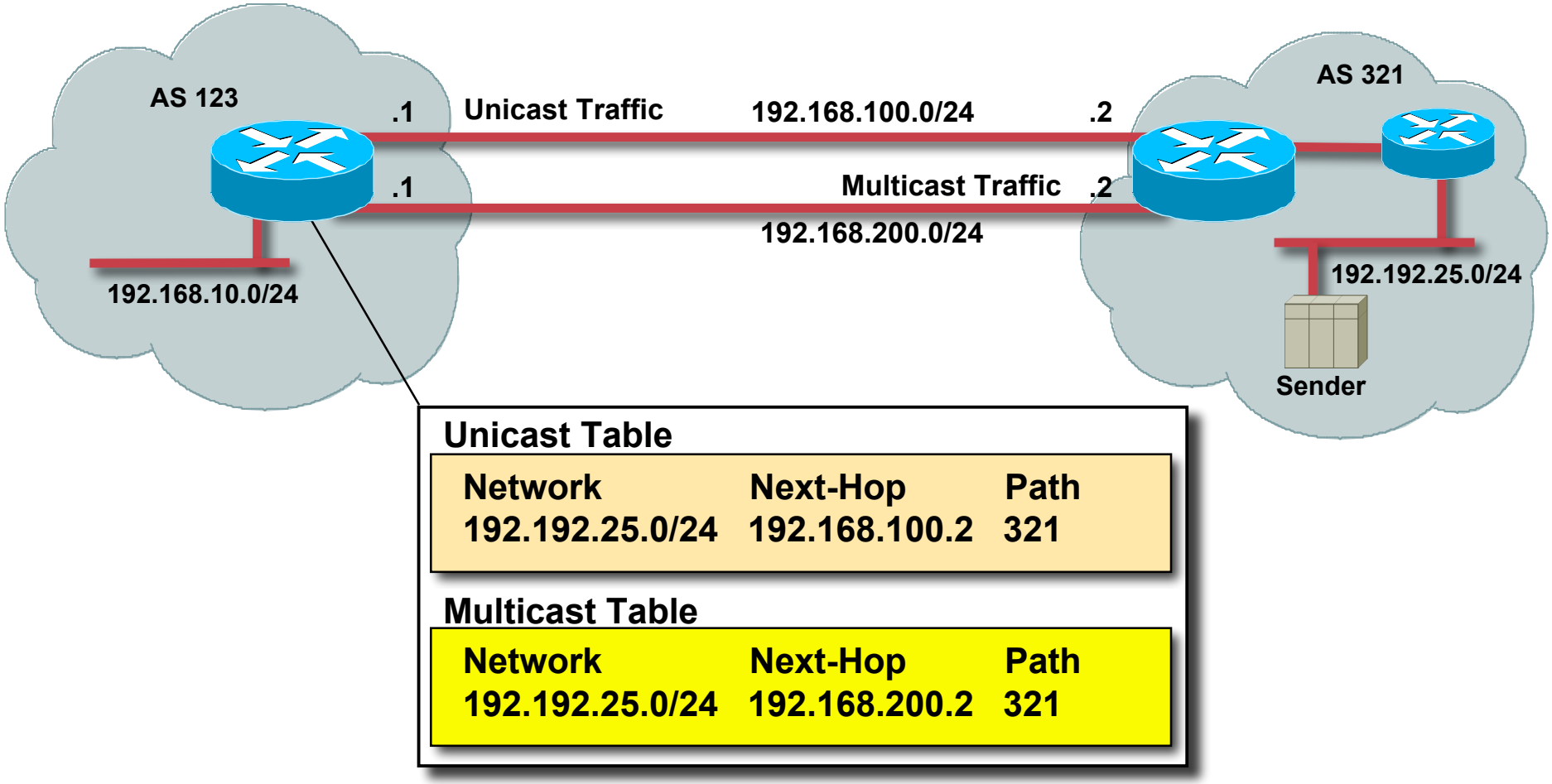
Incongruent Topologies



MBGP — NLRI Information

Cisco.com

Incongruent Topologies



Unicast-Multicast NLRI Translation

- BGP stubs that don’t have MBGP support need to get their prefixes into the Multicast backbone
- They get external routes via MBGP default or static default

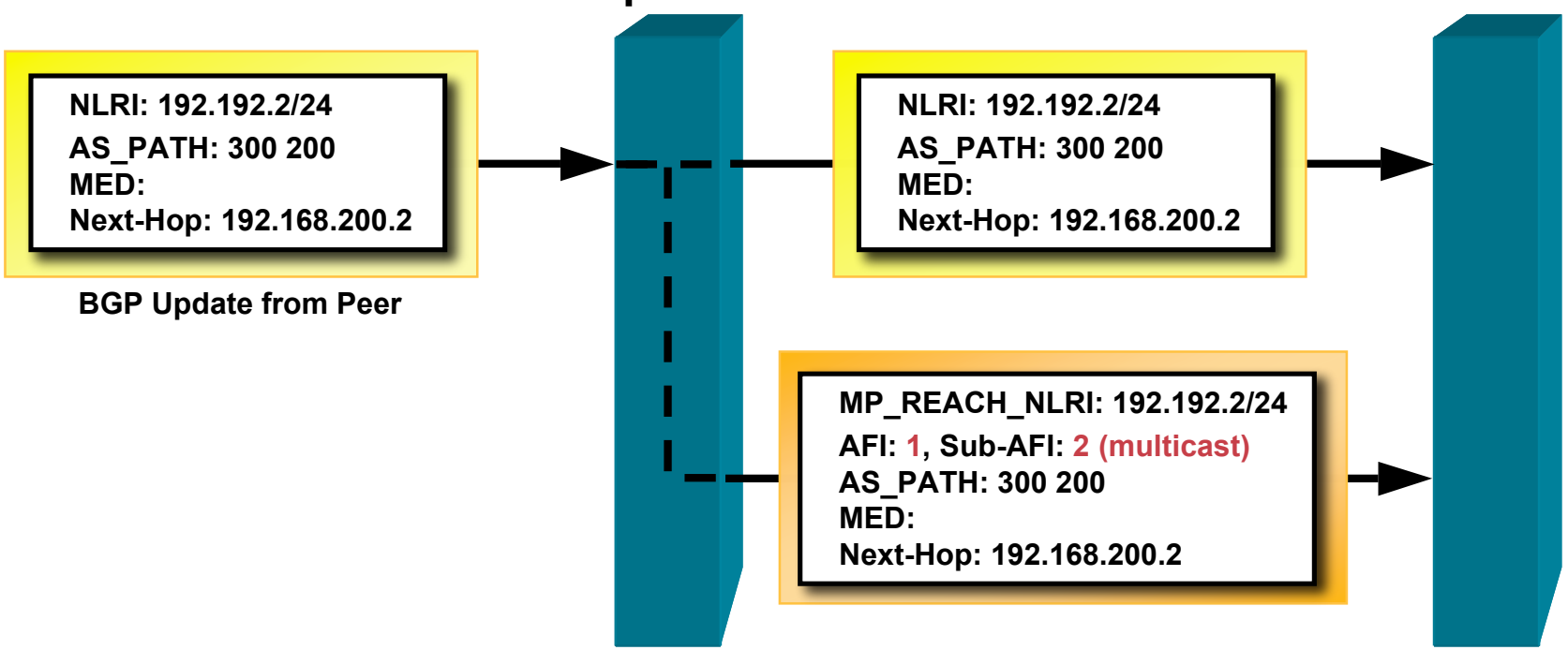
Unicast-Multicast NLRI Translation

- **Use command**

```
neighbor <foo> translate-update [nlri multicast]
```

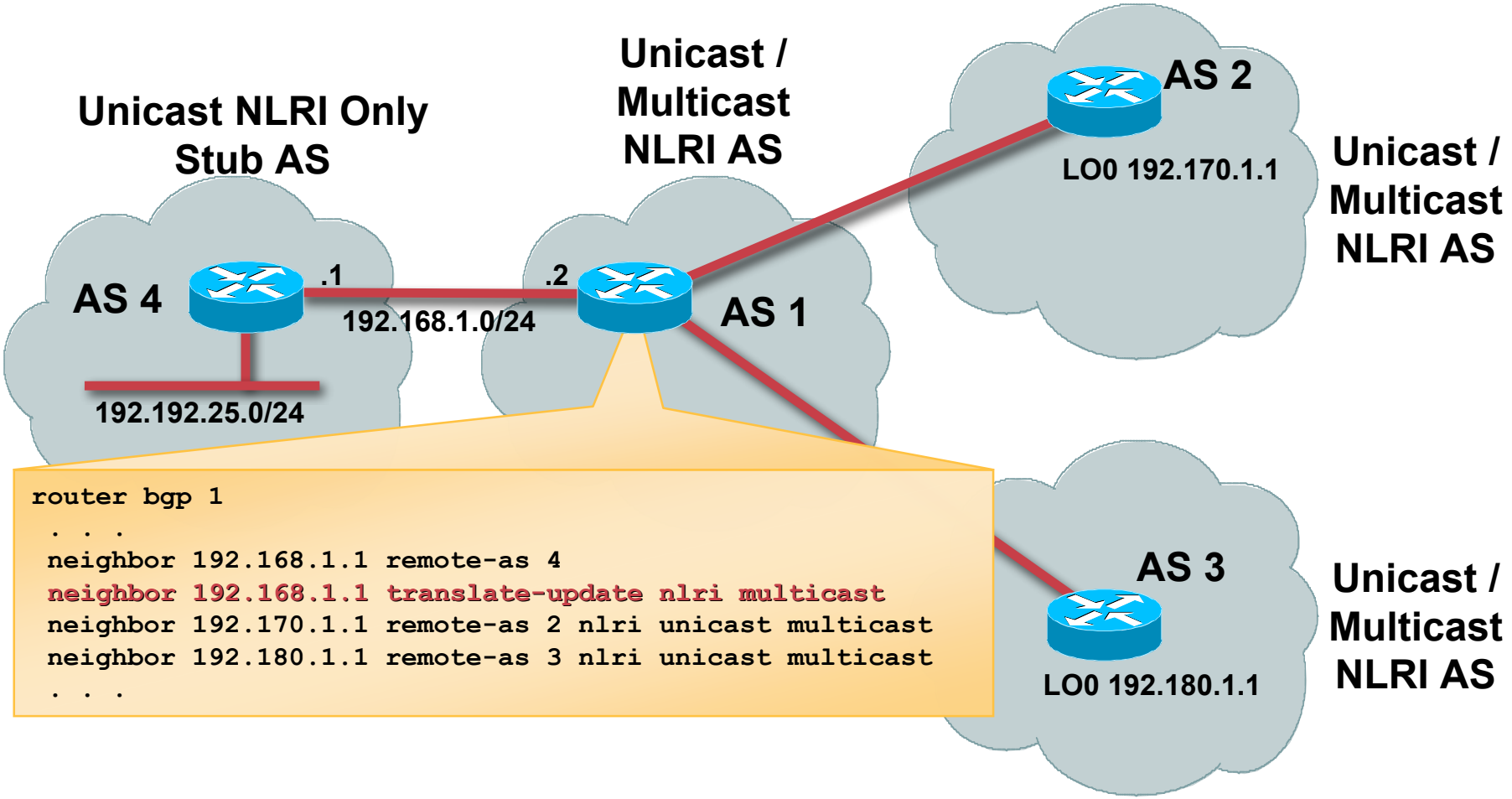
- **Arriving BGP Updates are translated into an MP_REACH_NLRI attribute**
  - **As if the neighbor sent AFI 1/SAFI 2 routes**
  - **Results written into the Multicast BGP Table**
- **Original BGP Update processed as normal**
  - **Results written into the Unicast BGP Table**

Unicast-Multicast NLRI Translation

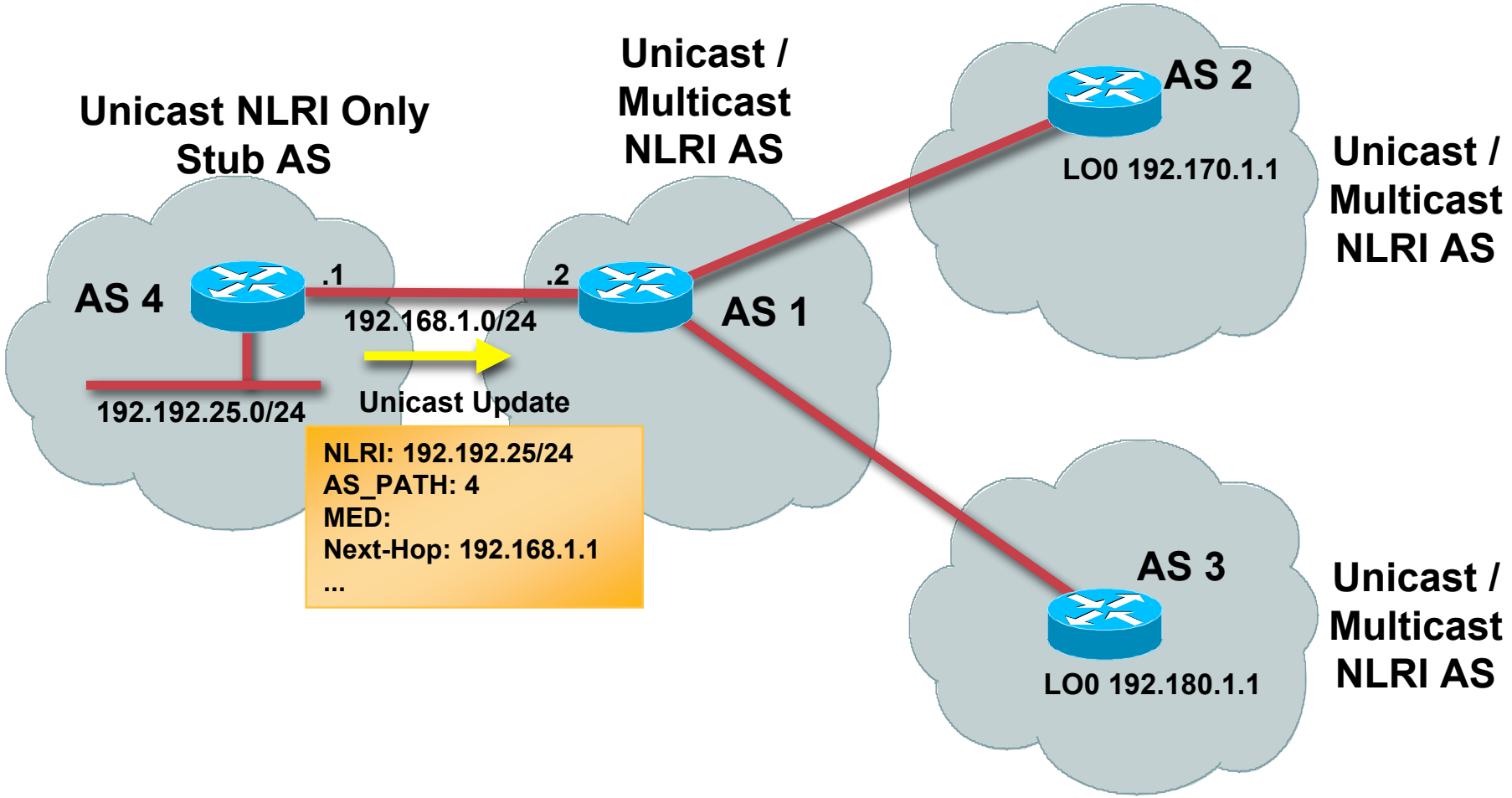


- Arriving Unicast update intercepted by 'translate-update' Front-end
- A translated Multicast update is created & passed to the IN Process
- Original Unicast update is passed on to the IN Process
- Both updates processed normally by the IN Process

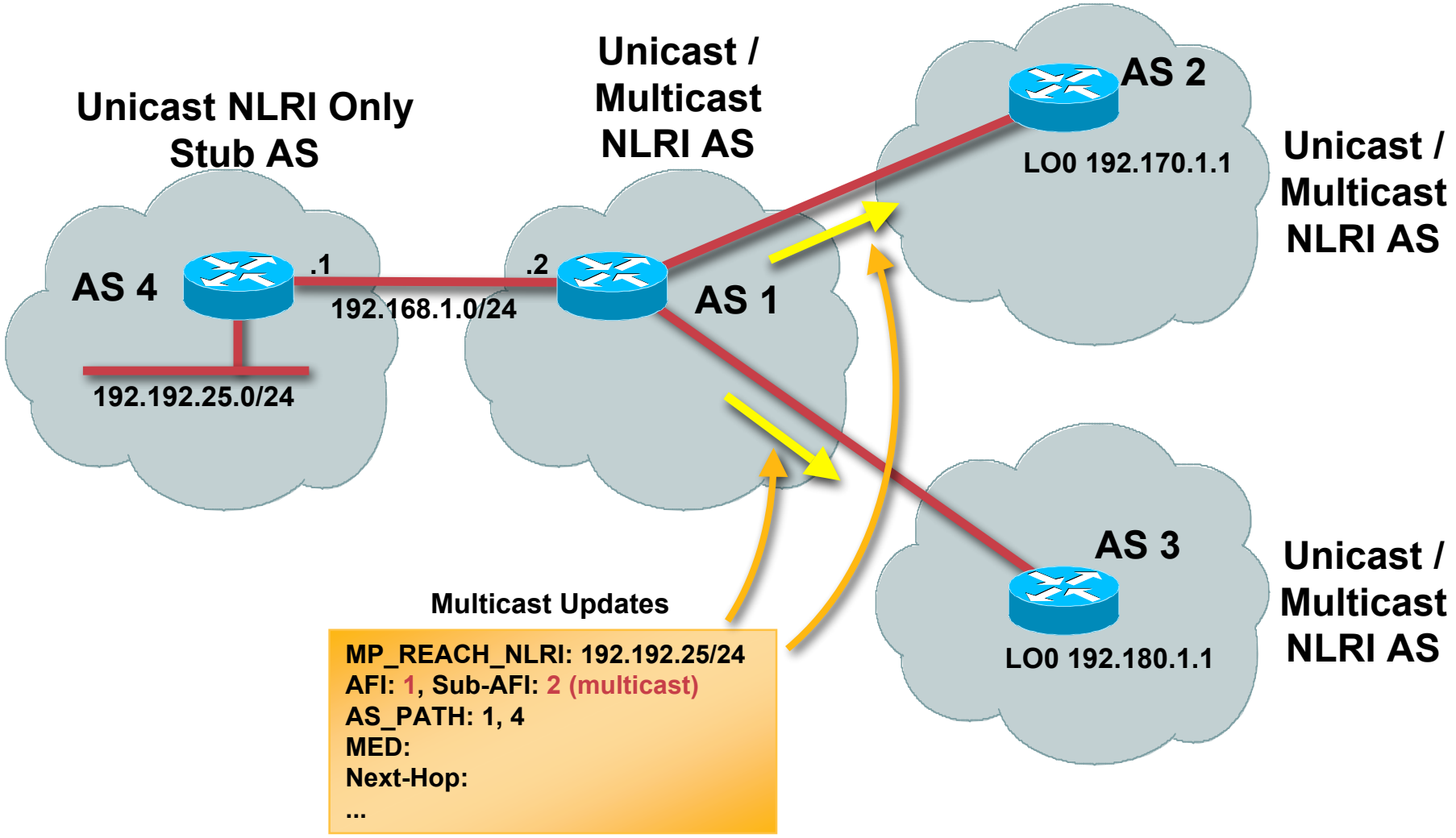
Unicast-Multicast NLRI Translation



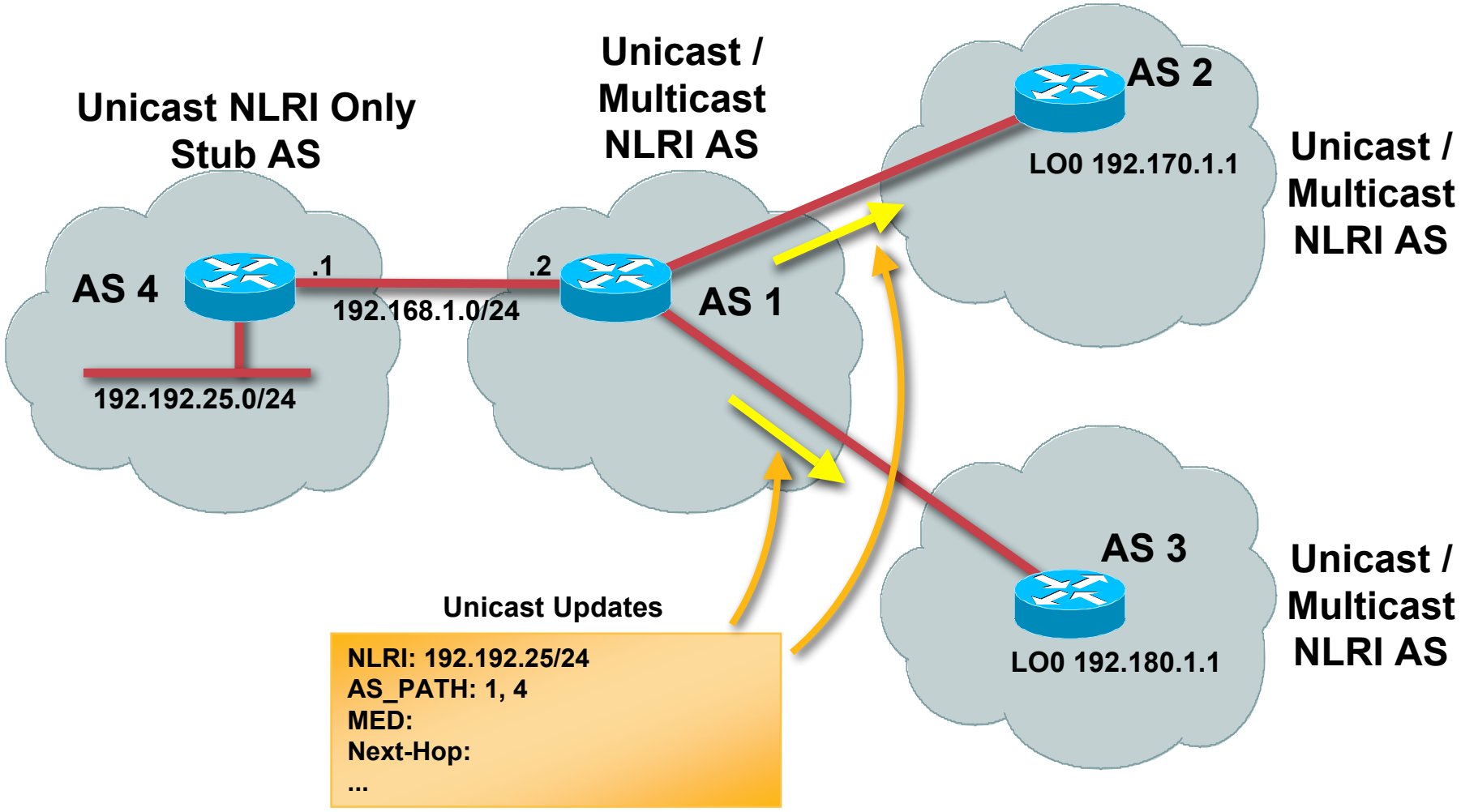
Unicast-Multicast NLRI Translation



Unicast-Multicast NLRI Translation



Unicast-Multicast NLRI Translation



MBGP—Summary

- Solves part of inter-domain problem
  - Can exchange multicast RPF information
  - Uses standard BGP configuration knobs
  - Permits separate unicast and multicast topologies if desired
- Still must use PIM to:
  - Build multicast distribution trees
  - Actually forward multicast traffic
  - PIM-SM recommended