Multisession BGP

IDR WG, IETF-58, Minneapolis, MN
November 13, 2003

John Scudder <jgs@cisco.com>
Chandra Appanna <achandra@cisco.com>

What

- Run different AFI/SAFI over different BGP sessions
- Sessions still use port 179, supported AFI/SAFI determined during OPEN phase
- No changes to ESTABLISHED state machinery
- Can totally remove multiplexing or multiplex selected AFI/SAFI

Design Considerations

- No requirement for multiple loopbacks
- Minimal configuration (for default behavior)
- Support for multiplexing selected AFI/SAFI ("grouping")

Why

- Protocol robustness
  - Corrupt PDU on one session won't affect other sessions
- Software engineering
  - Demux in transport allows separation of code to handle different "applications"
    - (If indeed the code differs enough to make this worth while)

Proposal

- Passive peer waits for OPEN from active peer before sending own OPEN
- Passive peer's OPEN replies with AFI/SAFI a subset of those sent by active peer
- Repeat process for any non-ESTABLISHED AFI/SAFI
- "Passive" and "active" roles determined dynamically -- "passive" = port 179 end

Proposal

- After session establishes, normal BGP rules apply

Proposal

- Multisession capability

- G bit indicates support for "grouping" (multiplexing)

Errors

- Three new OPEN Message Error subcodes
  - No Supported AFI/SAFI
  - Grouping Conflict
  - Grouping Required

No Supported AFI/SAFI

- Active peer could propose session with no AFI/SAFI supported by passive peer
  - Possible now, but non-multiplexed sessions make it much less unlikely
- Close session with No Supported AFI/SAFI error

Grouping Conflict

- Active peer drives AFI/SAFI supported on each session
- Thus, if passive peer wants different AFI/SAFI groups, tough luck
- Grouping Conflict error communicates collision in hard grouping policies

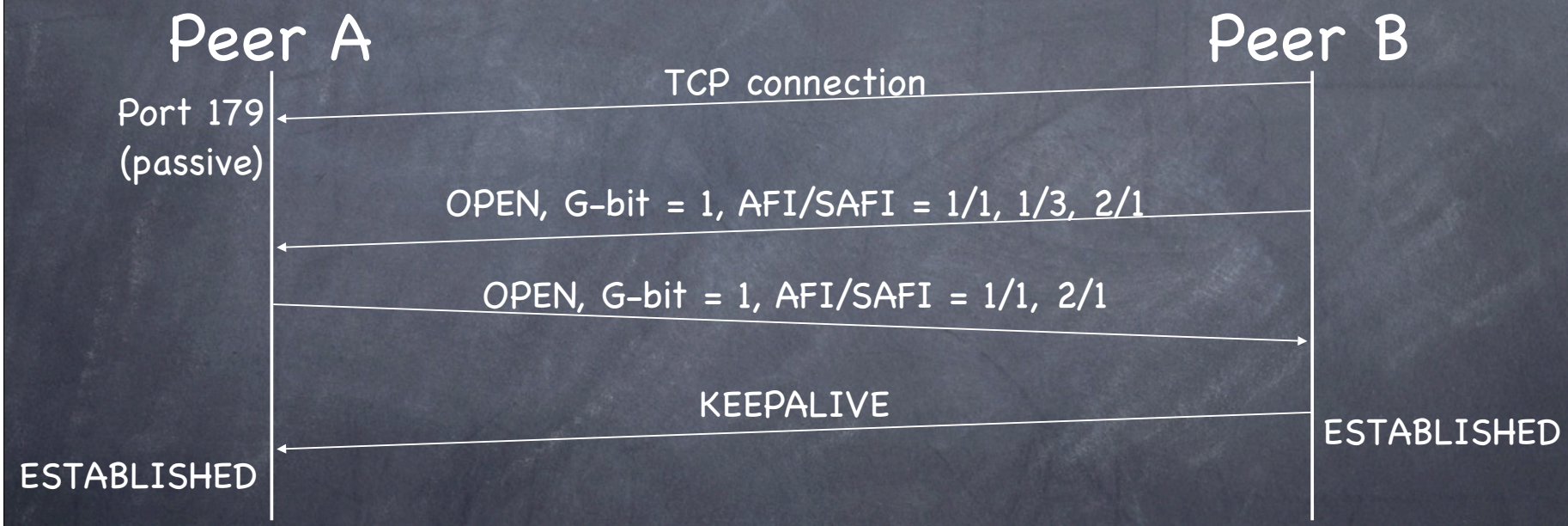
Grouping Conflict — Options

- Accept connection
- Close connection with Grouping Conflict error
  - Can subsequently attempt to re-establish another connection
- Arbitrary set operations were considered but rejected due to poor cost/benefit ratio

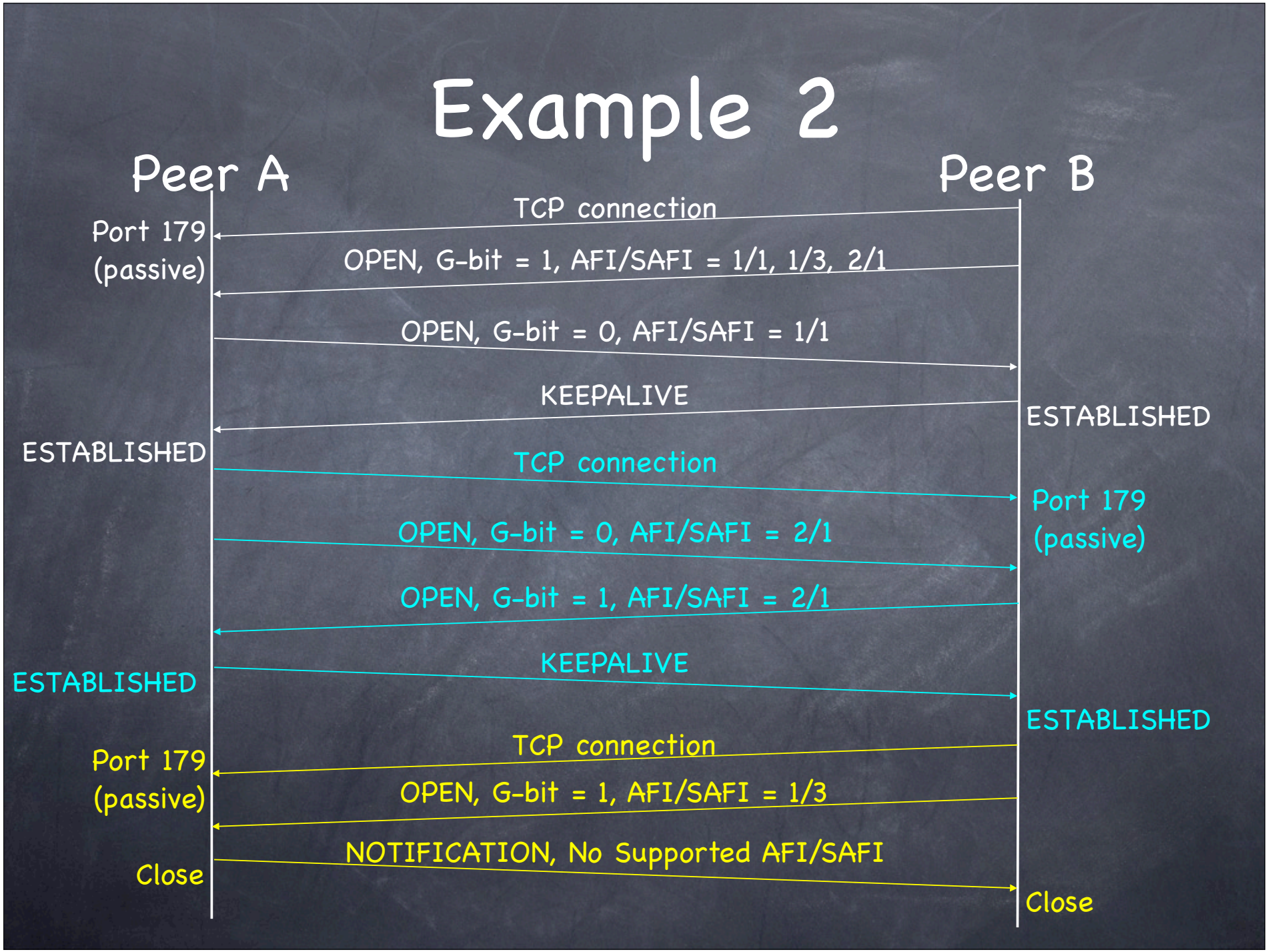
Grouping Required

- Active peer requires grouping and passive peer does not support grouping
  - Active peer MAY close connection with Grouping Required
  - MAY also accept non-grouped connection
- Special case of Grouping Conflict

Example 1


Peer A
Peer B
TCP connection
Port 179
(passive)
OPEN, G-bit = 1, AFI/SAFI = 1/1, 1/3, 2/1
OPEN, G-bit = 1, AFI/SAFI = 1/1, 2/1
KEEPALIVE
ESTABLISHED
ESTABLISHED

Example 2


Peer A
Peer B
TCP connection
Port 179
(passive)
OPEN, G-bit = 1, AFI/SAFI = 1/1, 1/3, 2/1
OPEN, G-bit = 0, AFI/SAFI = 1/1
KEEPALIVE
ESTABLISHED
ESTABLISHED
TCP connection
Port 179
(passive)
OPEN, G-bit = 0, AFI/SAFI = 2/1
OPEN, G-bit = 1, AFI/SAFI = 2/1
KEEPALIVE
ESTABLISHED
ESTABLISHED
TCP connection
Port 179
(passive)
OPEN, G-bit = 1, AFI/SAFI = 1/3
NOTIFICATION, No Supported AFI/SAFI
Close
Close

Other Approaches

- Configure multiple loopbacks and peer between them
  - Can do this today!

Other Approaches

- Well-known port per AFI/SAFI
  - No grouping (or inflexible grouping)
  - Need to assign port
  - Less backward compatible?

Other Approaches

- Manually configured listen ports
  - Defaults to multiplexed mode unless ports configured
  - More brittle (prone to configuration error)

Other Approaches

- Control session to chat about groupings and port numbers, then establish sessions based on outcome
  - Yuck

Draft

- ftp://ftpeng.cisco.com/jgs/draft-scudder-bgp-multisession-00.txt
- (or your favorite Internet Drafts repository in a day or three)
- Comments to authors {achandra,jgs}@cisco.com or IDR mailing list