



# Why OSPF paths aren't always shortest

David Applegate

Carsten Lund

Aman Shaikh

**AT&T Labs (Research)**

**NANOG 54**

**February 06, 2012**



# Introduction

OSPF is widely used for intra-AS routing

OSPF routes packets along shortest paths

- In terms of weights configured on links

For scalability, OSPF divides domains into areas

- Areas are widely used as well

## Areas make OSPF routing complicated

- Paths are no longer shortest
- Other tweaks to areas have increased complexity
  - Multi-Area Adjacencies (RFC 5185)
  - Multi-Area Routers (RFC 3509)

# Outline

## Introduction to OSPF

- Basics of OSPF routing

## Areas

- How routes are computed with areas

## Tweaks to areas

- Multi-area adjacency (RFC 5185)
- Multi-area routers (RFC 3509)

## AS Border Routers (ASBRs)

## Quiz

# OSPF as a Link-State Protocol

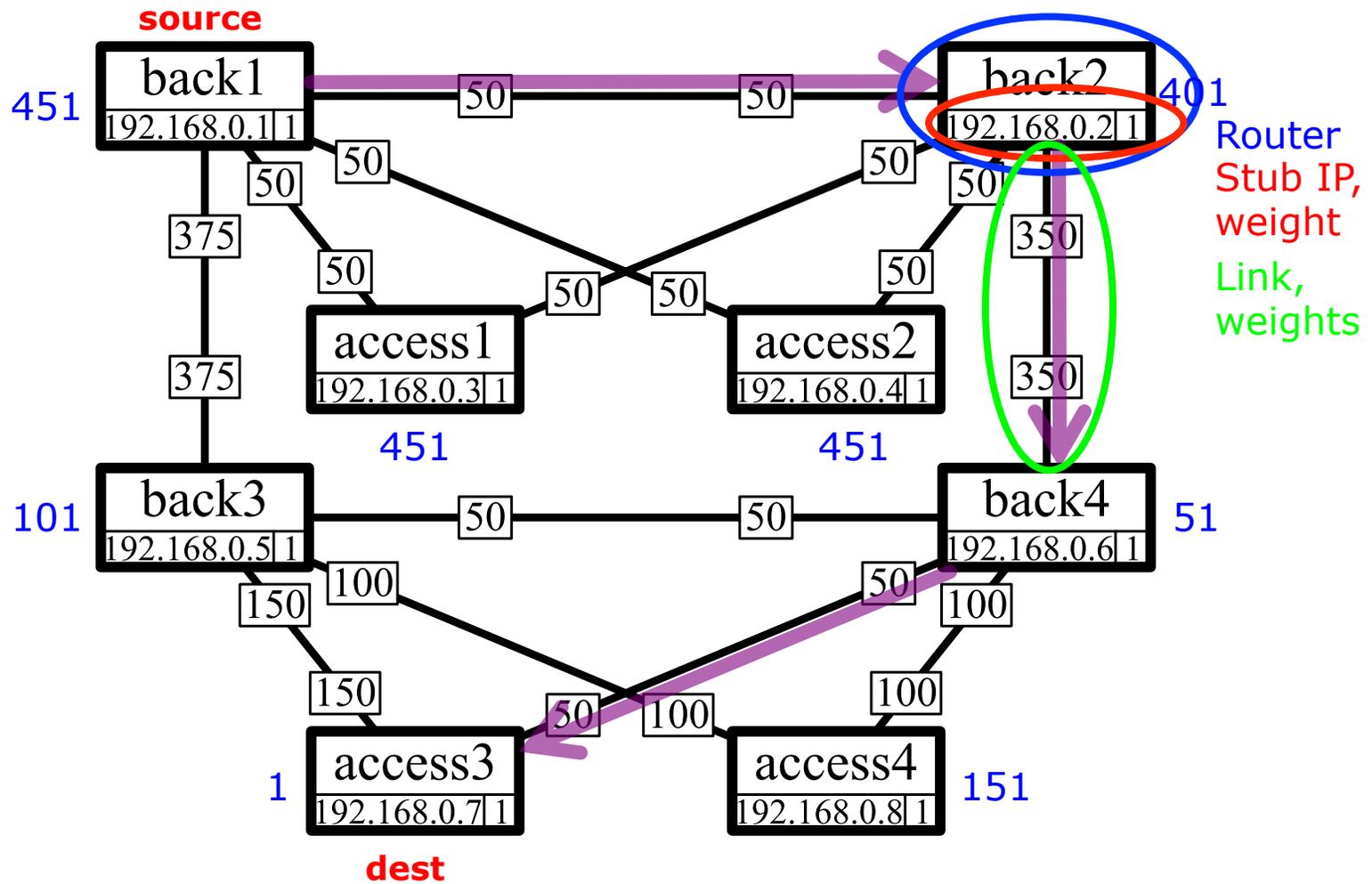
With a link-state protocol, every router ...

- learns entire network topology
  - represents topology as a weighted graph
- computes Shortest Path Tree rooted at itself

OSPF follows this with some tweaks

- Nodes are of two types:
  - Transit: routers, subnets
  - Stub: prefixes advertized by routers
    - Example: /32 for loopbacks
- **Path is computed from a source router to a stub**
  - Via one or more transit nodes

# Example



# Areas

For scalability, OSPF domain is divided in areas

- Areas are numbered 0, 1, 2, ...
- Conceptually a hub-and-spoke
  - Area 0 is hub, non-zero areas are spokes

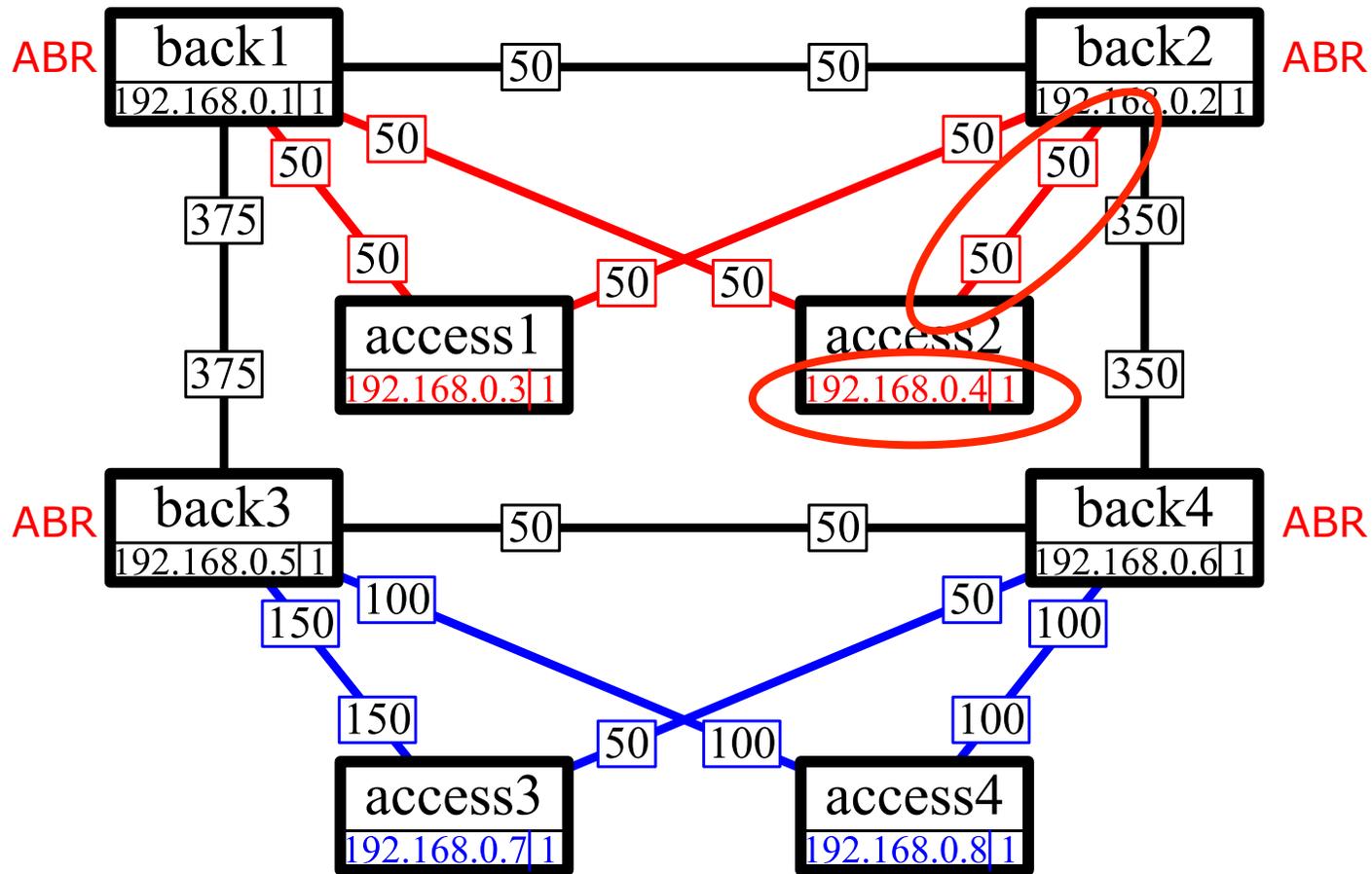
Each link and stub is assigned to a single area

- A router can have links in multiple areas
  - Such a router is called an **area border router (ABR)**
  - An ABR must have a link in area 0 (RFC 2328)

With areas, every router learns ...

- entire topology of areas it has links to
- distance from ABRs to stubs in remote areas

# Example of OSPF Areas



Color indicates area  
(Links and Stubs)

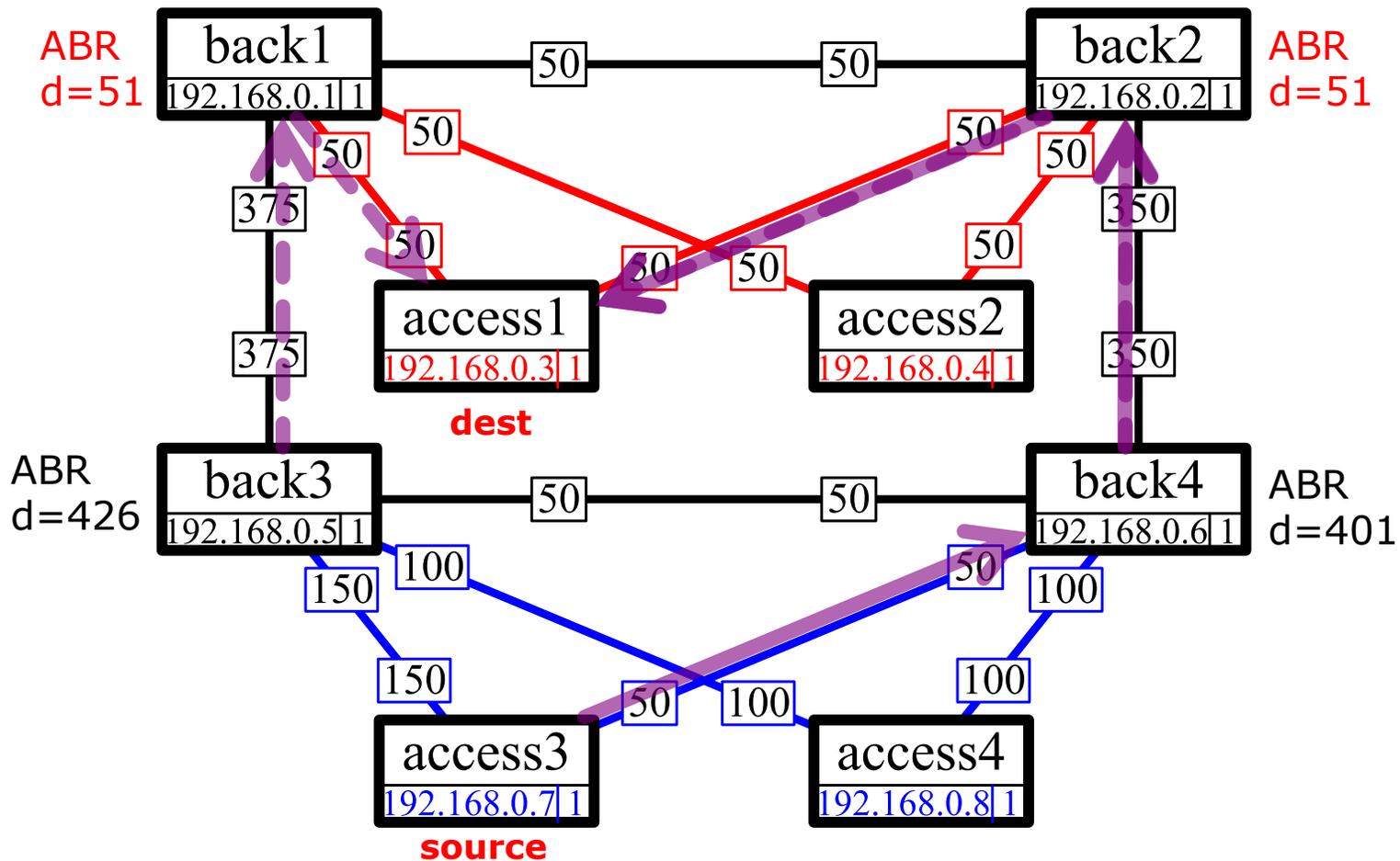
Black is area 0

# Path Calculation with Areas

A router ....

- Calculates SPTs for all attached areas
  - Leads to **intra-area paths** to stubs
- Calculates paths to all remote stubs
  - Minimize the total distance from itself to the stub
    - Total distance =  $\text{dist}(\text{router}, \text{ABR})$  in an attached area + advertized  $\text{dist}(\text{ABR}, \text{stub})$
  - Leads to **inter-area paths** to (remote) stubs

# Example: Path Calculation with Areas



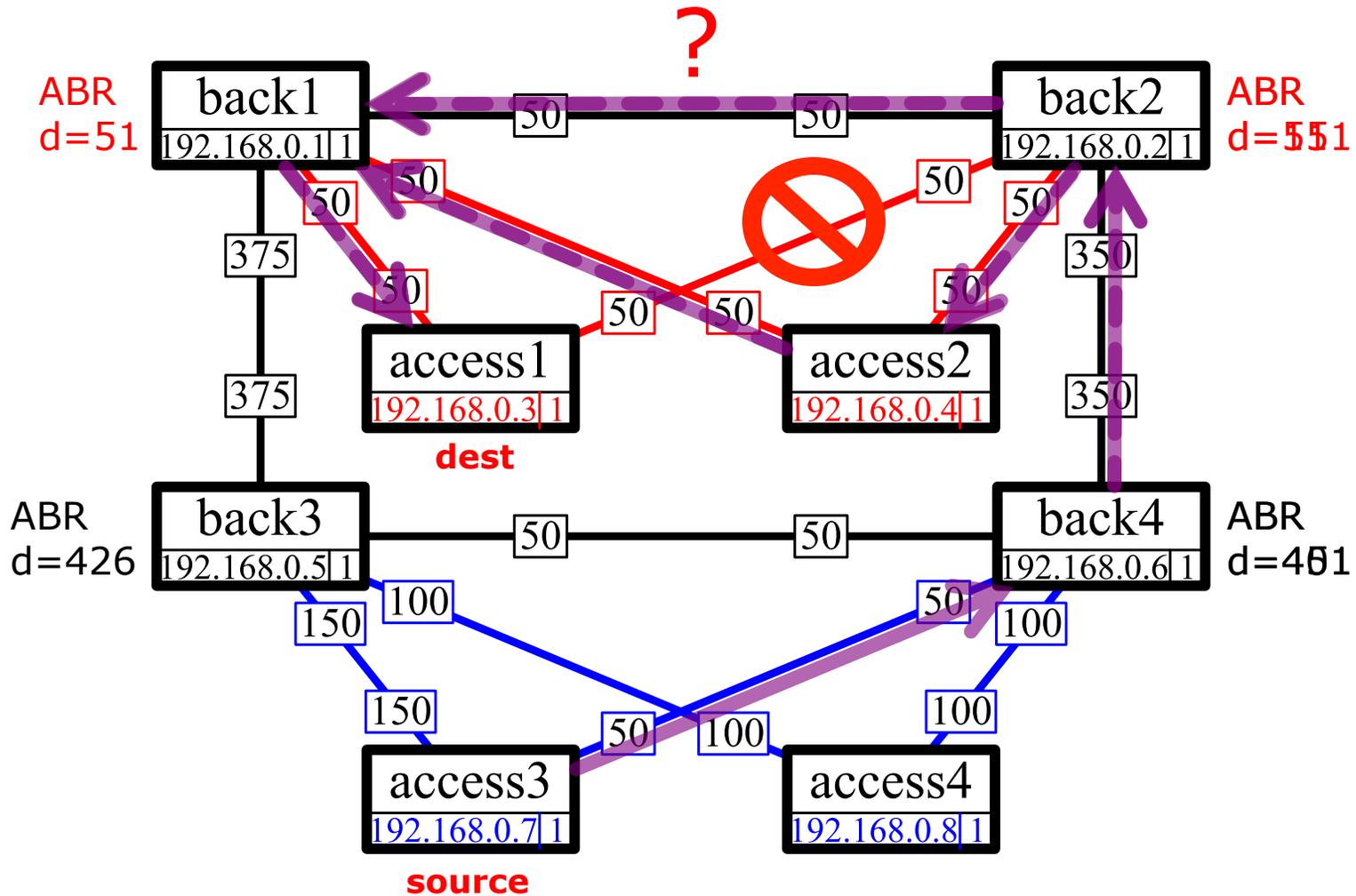
# Intra-area v/s Inter-area Paths

When a router has to choose between intra-area and inter-area paths, it always chooses intra-area path

Leads to ...

- Sub-optimal paths (within an area)
  - Packet takes a longer intra-area path over a shorter inter-area path
- **Area hijacking** at ABRs
  - Actual path (and the distance) differs from the path (and distance) calculated by the source router

# Area Hijacking Example



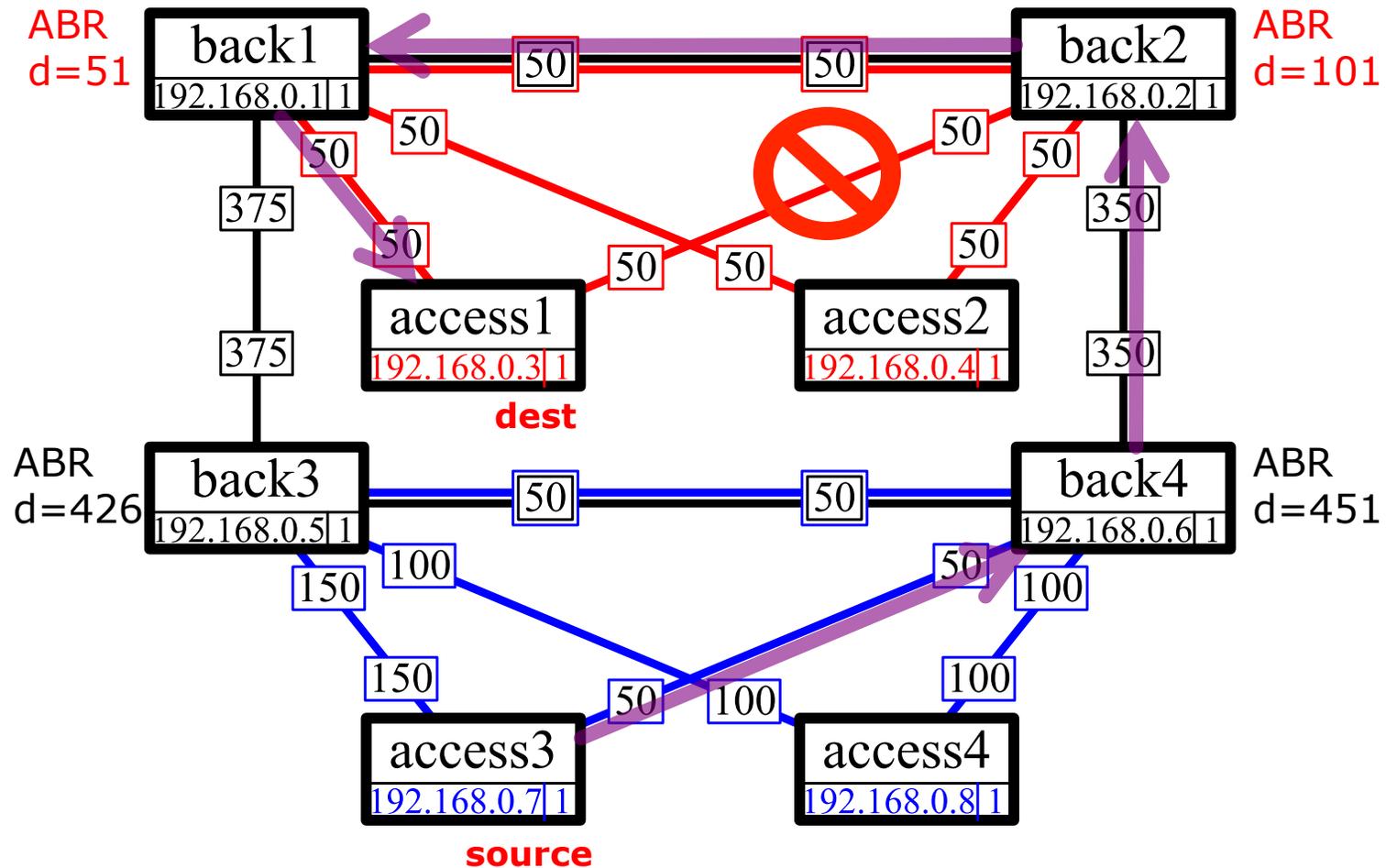
# Multi Area Adjacencies (MADJ)

RFC 5185 allows a link to be in multiple areas

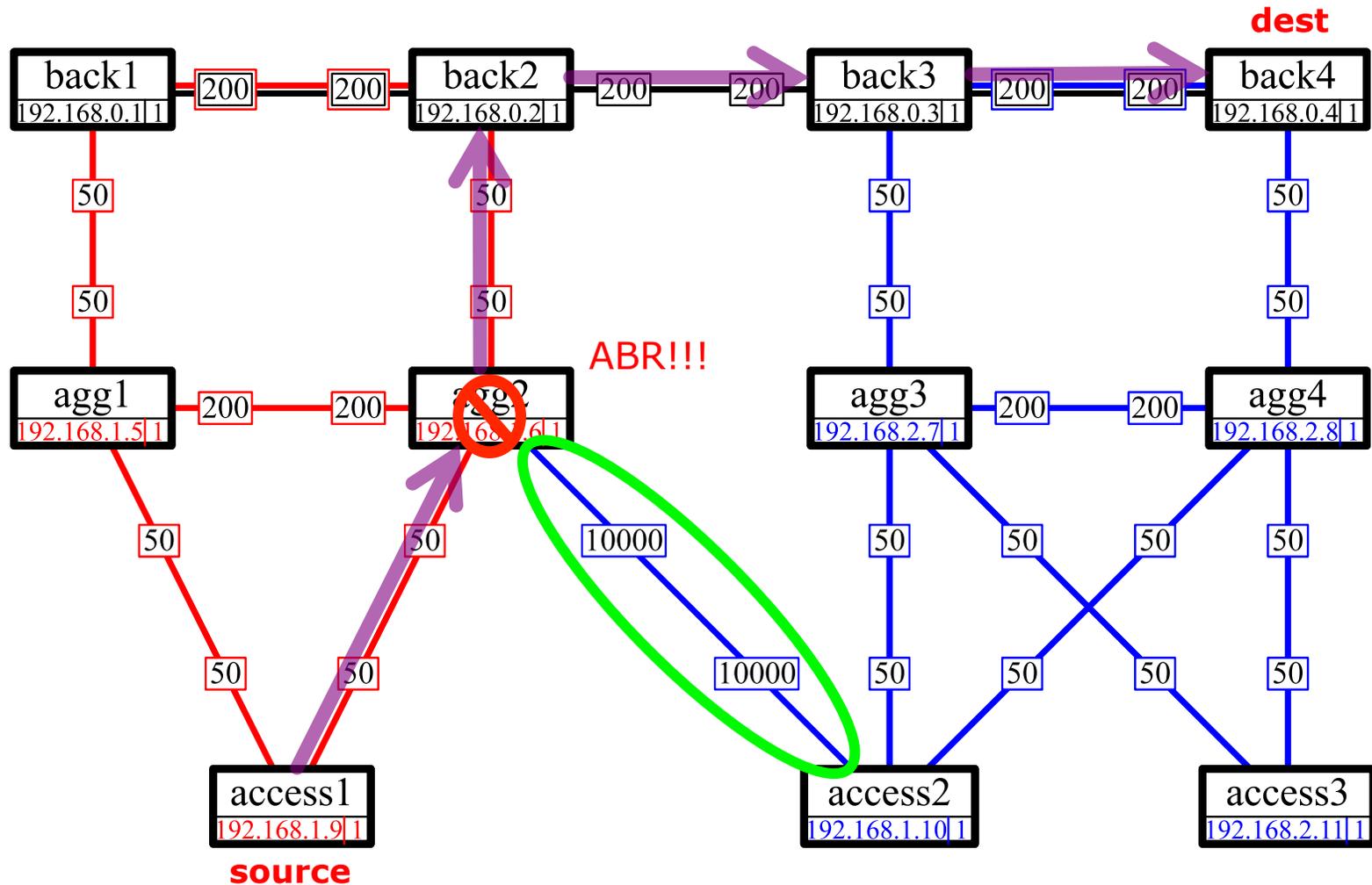
- Stubs can only be in a single area
- Links have a primary area for their interface stubs

Protects against some hijacking cases, but not all

# Area Hijacking Example with MADJ



# Multi Area Routers



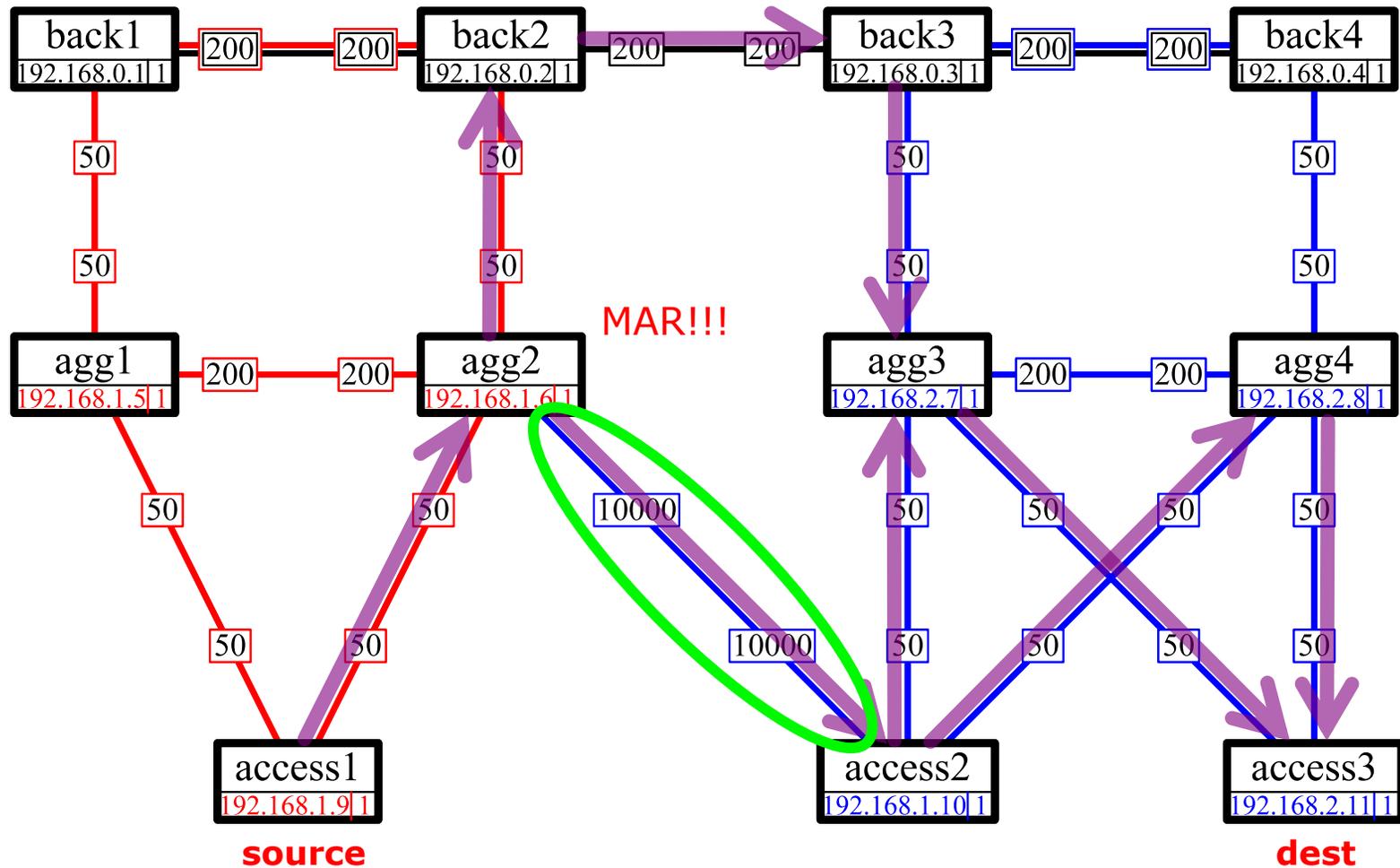
# Multi Area Routers (MAR)

RFC 3509 allows a router to be in multiple non-zero areas without being in area 0

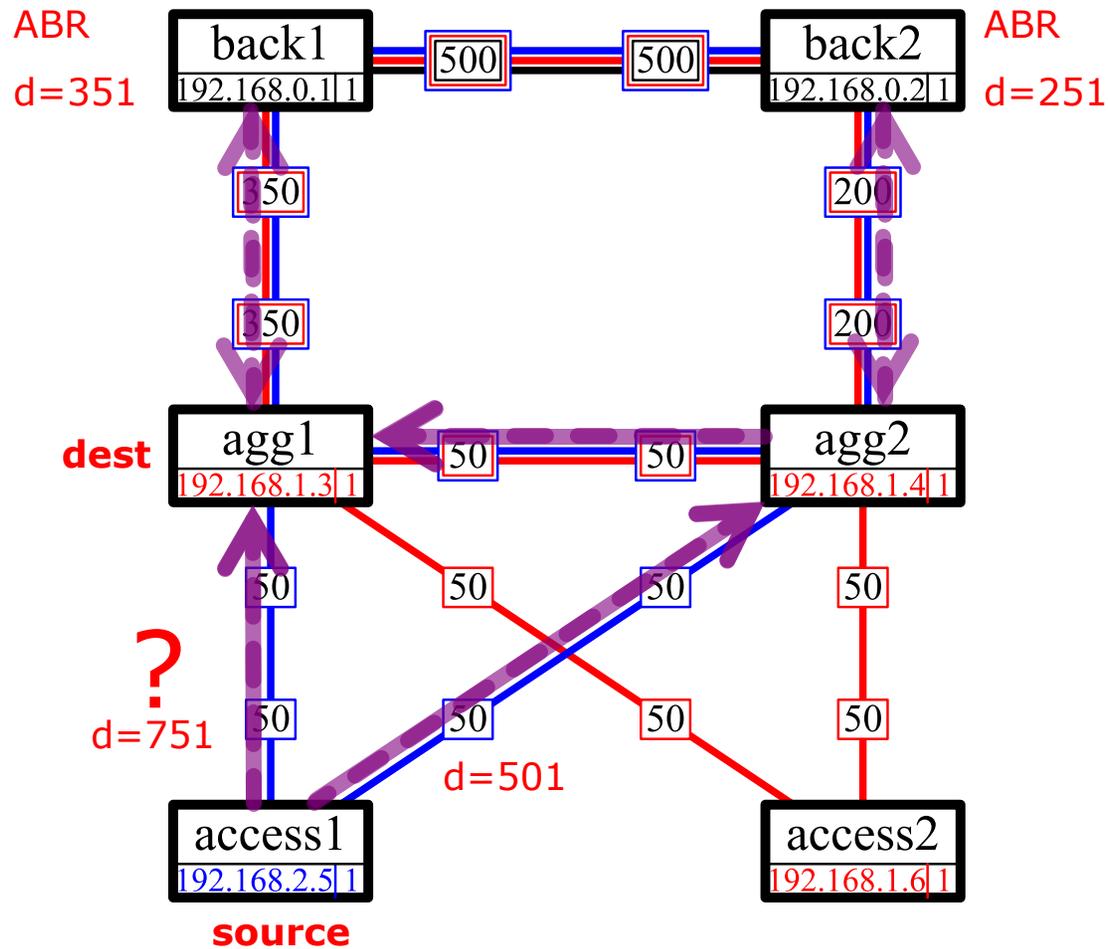
- Protects against dropping traffic
- Specifies subtly different behavior of Cisco and IBM routers (only)

Leads to more opportunities for area hijacking

# Area Hijacking Example with MAR



# More Area Hijacking (Stub Area Matters)



# Importing External Information

External routes can be imported into OSPF

- Example: static routes

Router importing external routes is called an **ASBR** (AS Border Router)

- Route contains distance from ASBR to prefix

A router ....

- Calculates SPTs for all the attached areas
- Calculates paths to all remote stubs
- Calculates paths to all external stubs
  - Calculates path (and distance) to ASBR and combines that with  $\text{dist}(\text{ASBR}, \text{stub})$

# Routing to an ASBR

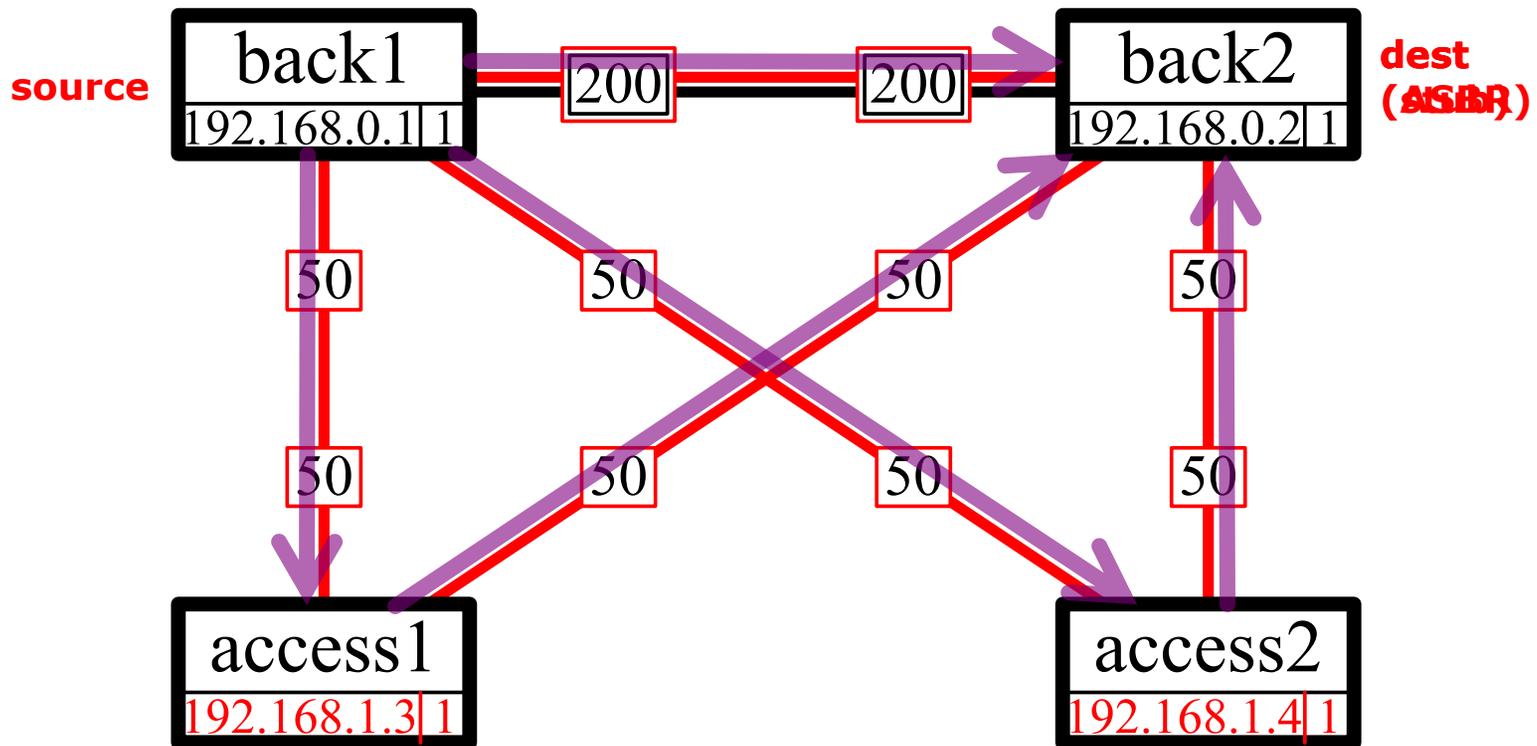
Unusual, because ASBR is not in an area

- ASBR could be reachable in multiple areas

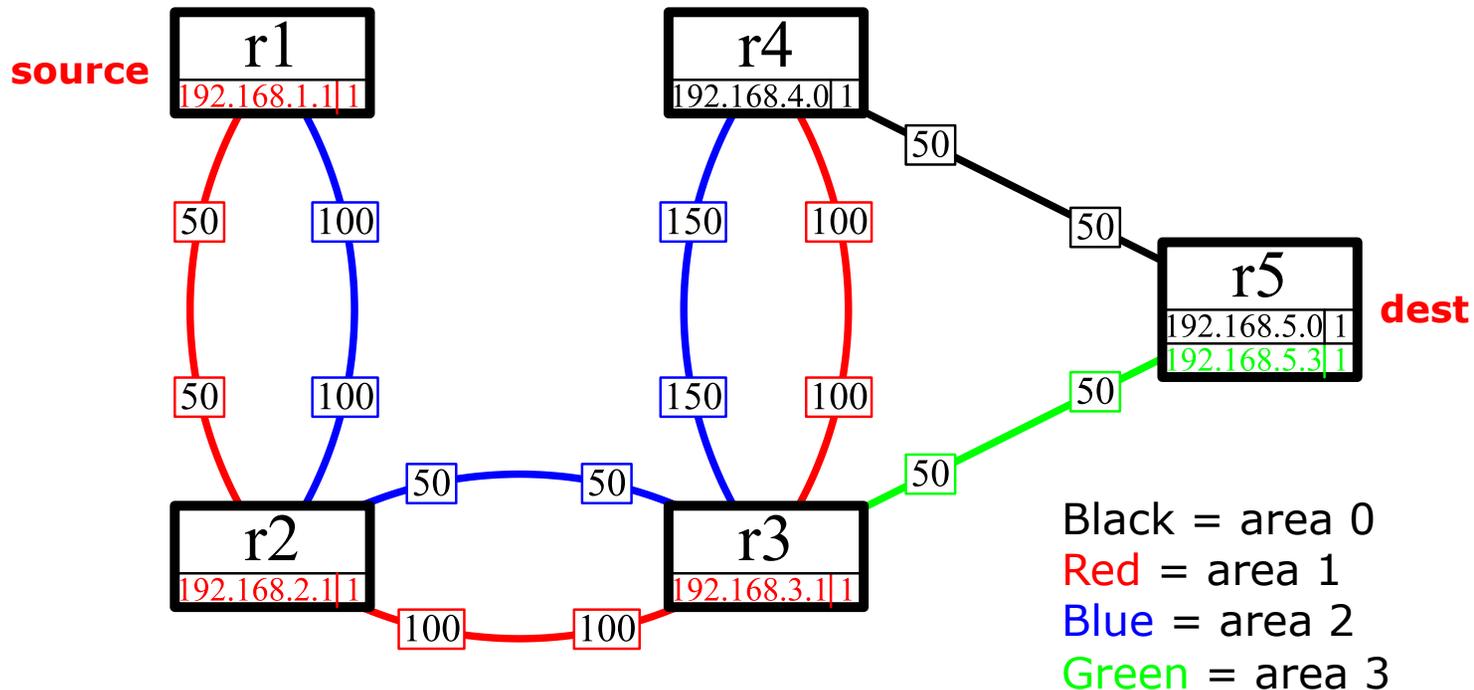
So a router has to calculate per-area path to an ASBR, and then choose the best path

- Tie-breaking rules depend on whether RFC1583compatibility is set to disabled
  - RFC 1583 is the older OSPF RFC
- Ties broken by
  - Least cost
  - Highest area number of the link

# Example of ASBR routing



# Quiz: What's the path from source to dest?

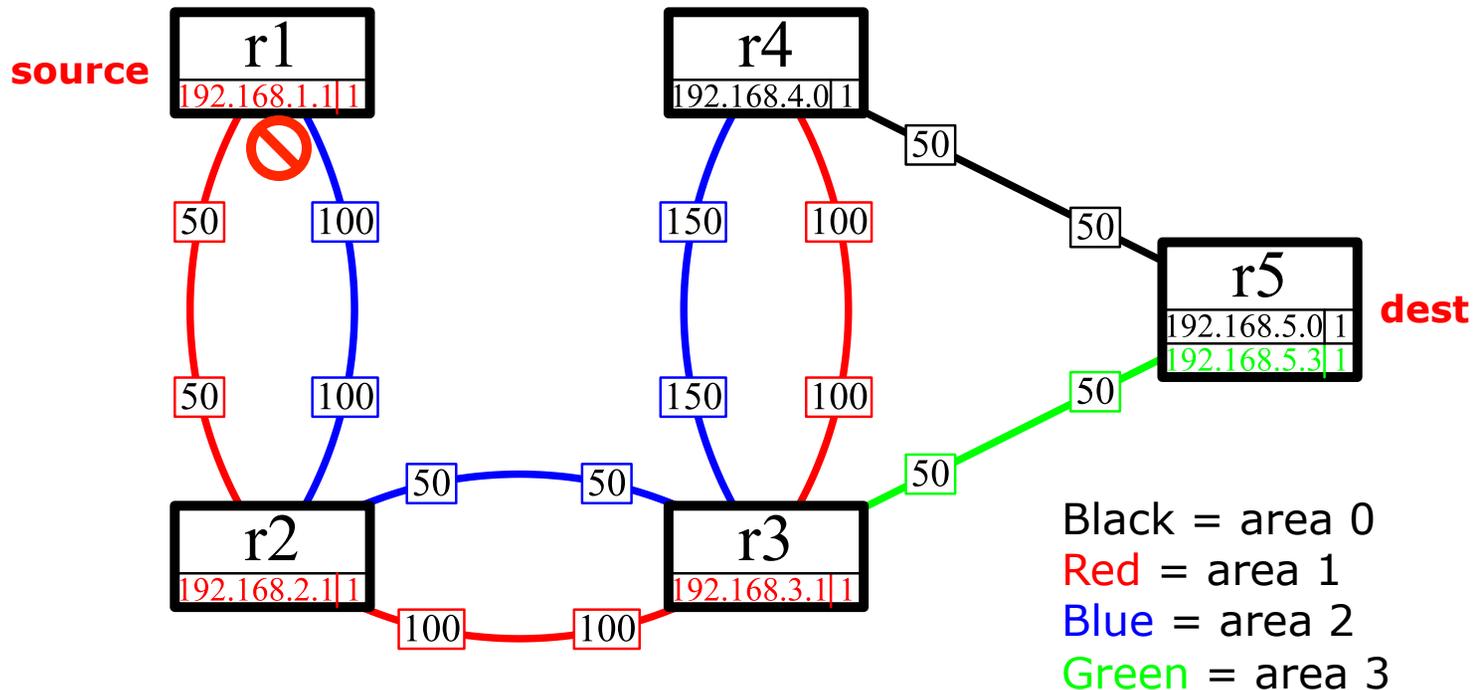


# Answer

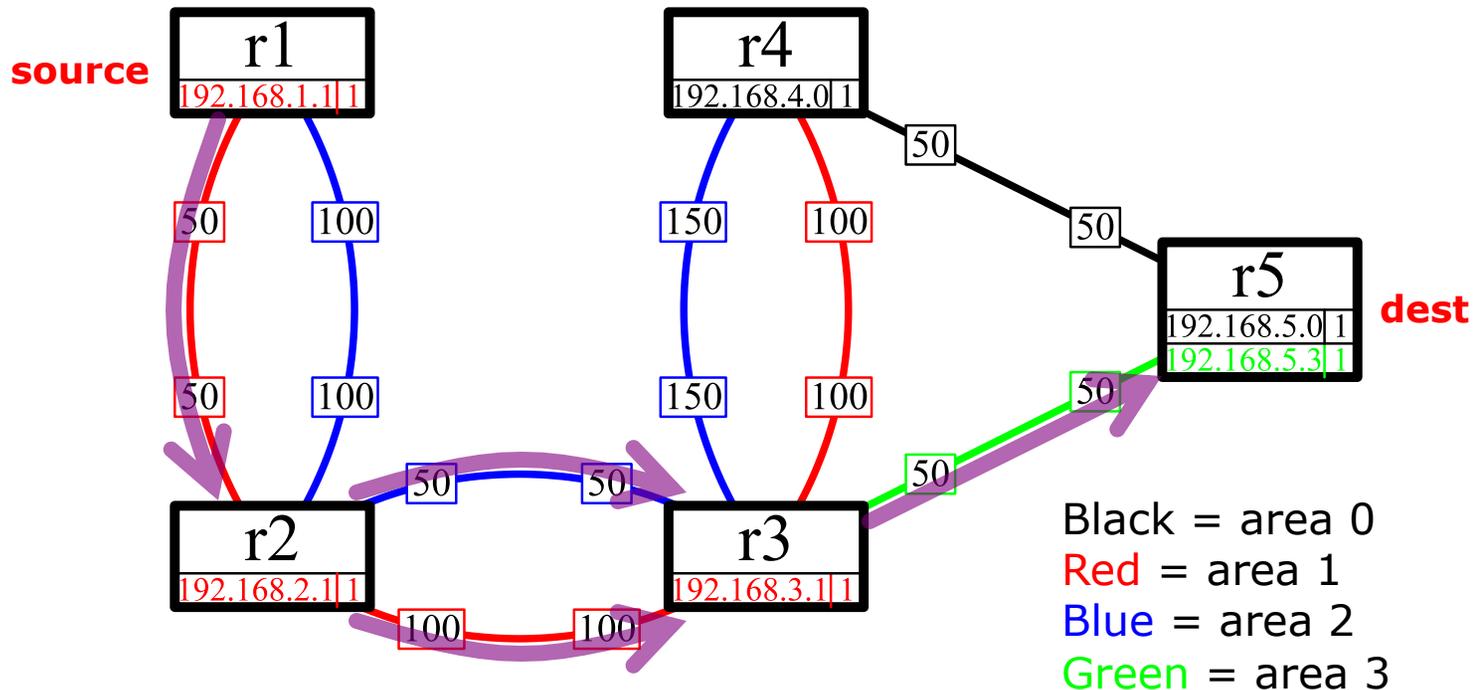
Depends on ...

- What 'dest' refers to...
  - Choices: r5:192.168.5.0, r5:192.168.5.3 or r5:ASBR
- Whether routers are RFC-3509 compliant or not
  - i.e., can r1, r2, and r3 act as true MARs or not
- When RFC-3509 compliant:
  - Whether r3 advertises routes learned in green area into red and blue areas or not
    - Vendor dependent
      - Cisco and Junipers behave differently

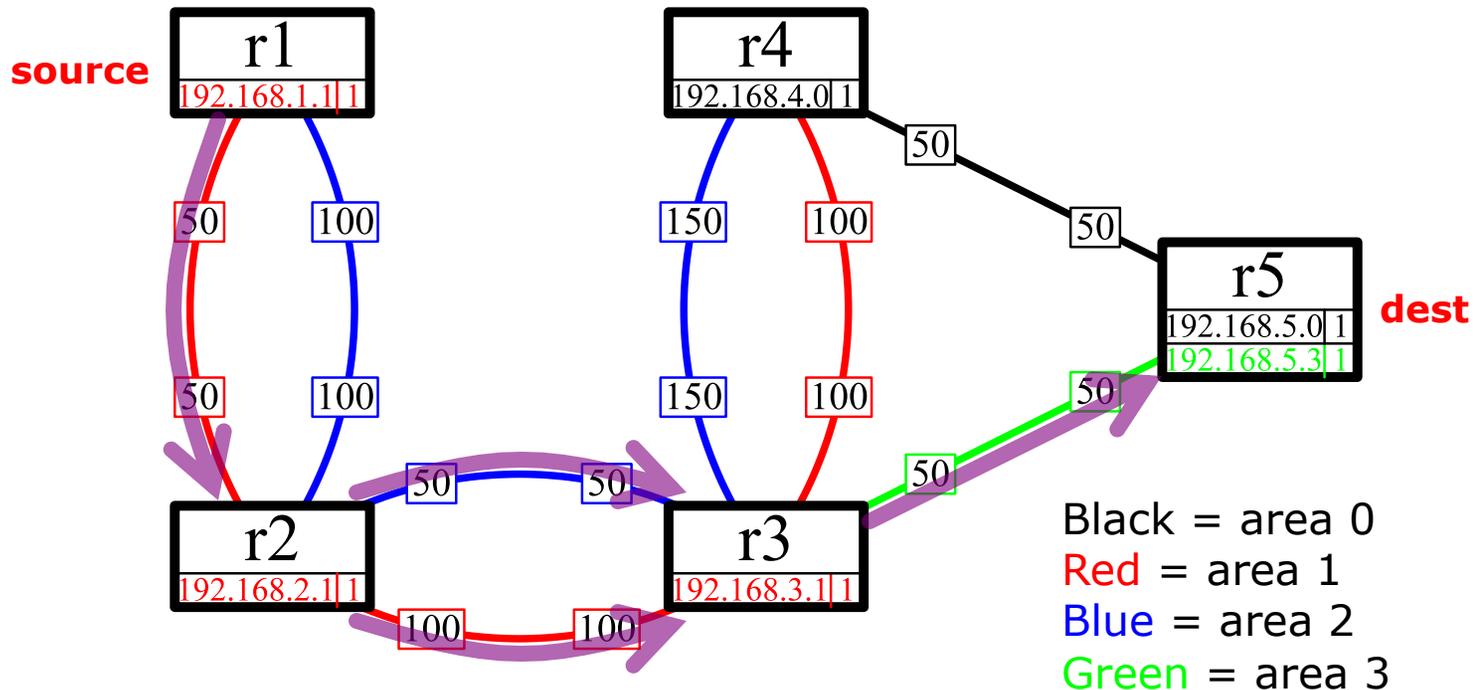
# Quiz: What's the path from source to dest? dest 192.168.5.0, r1 not RFC3509



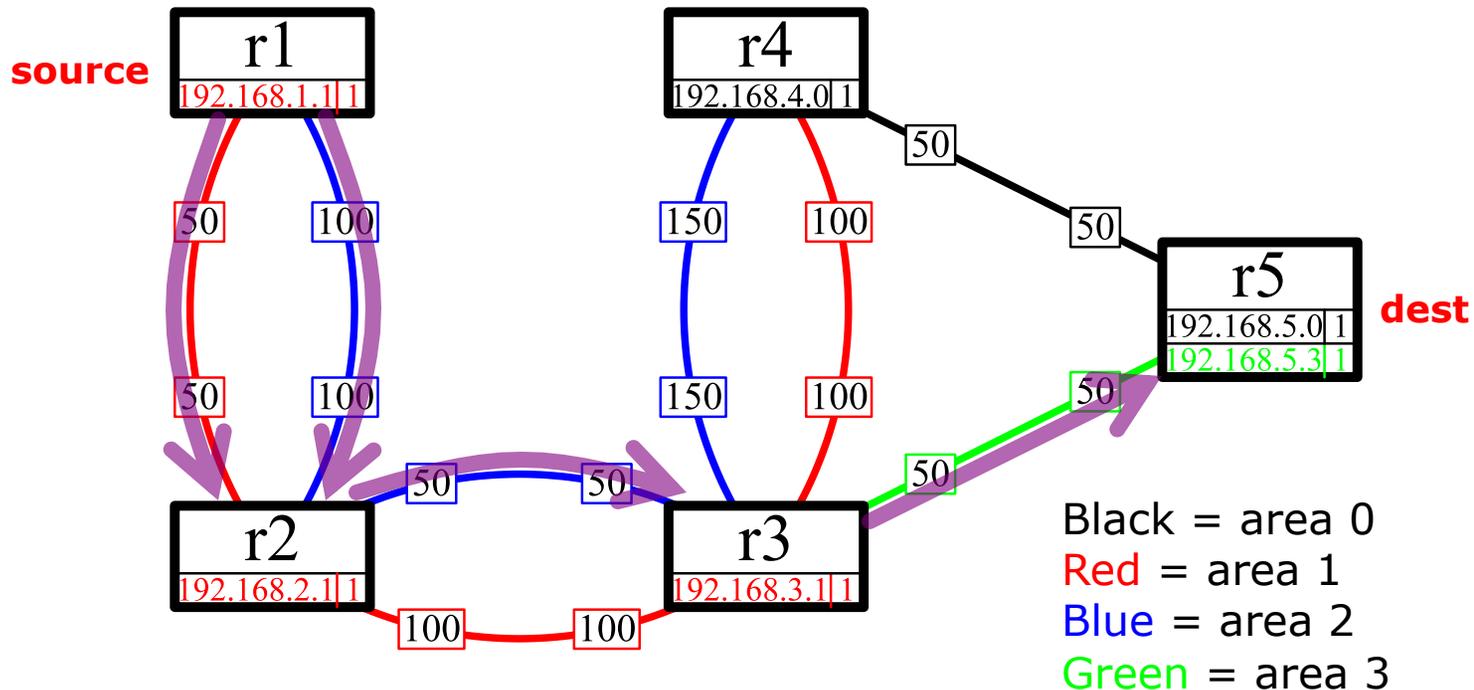
# Quiz: What's the path from source to dest? dest 192.168.5.0, RFC3509



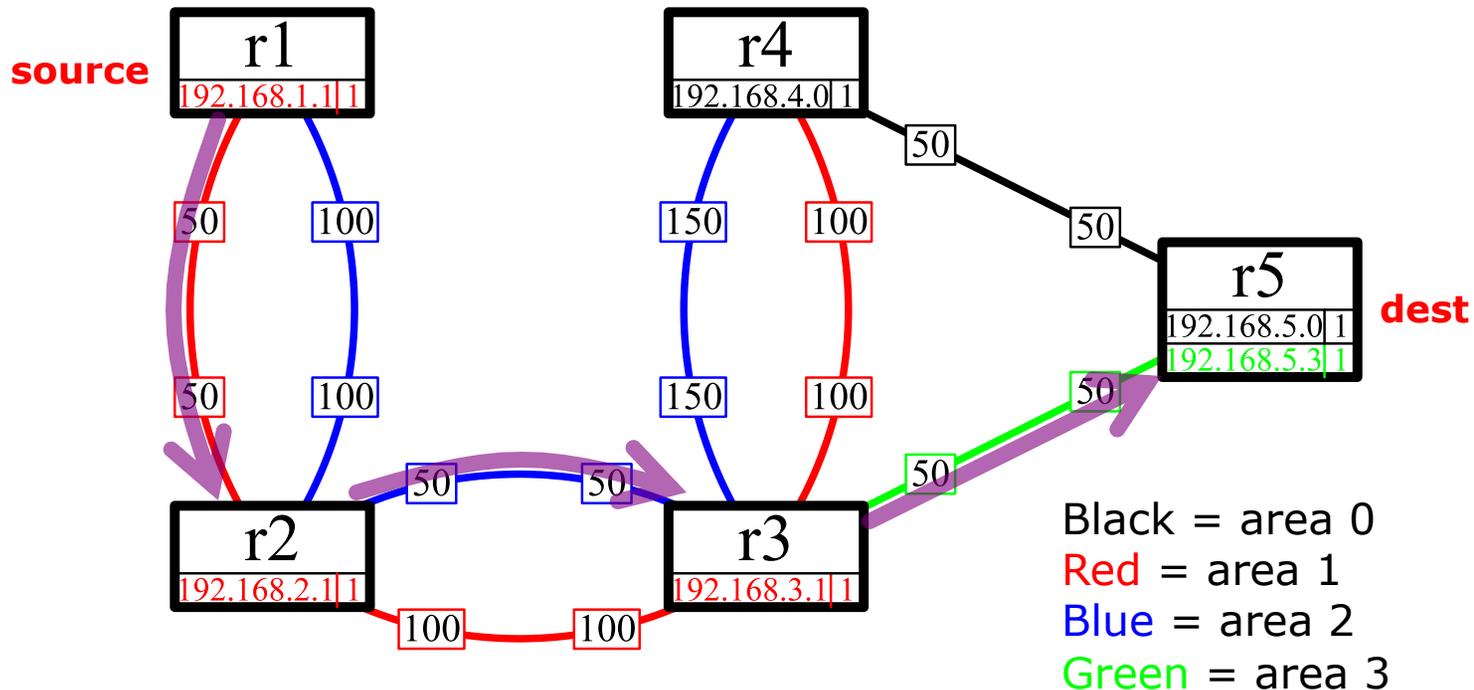
# Quiz: What's the path from source to dest? dest 192.168.5.3, RFC3509, Cisco



# Quiz: What's the path from source to dest? dest 192.168.5.3, RFC3509, Juniper



# Quiz: What's the path from source to dest? dest r5:ASBR, RFC3509, Cisco



# Quiz: What's the path from **source** to **dest**?

**dest** r5:ASBR, RFC3509, Juniper

