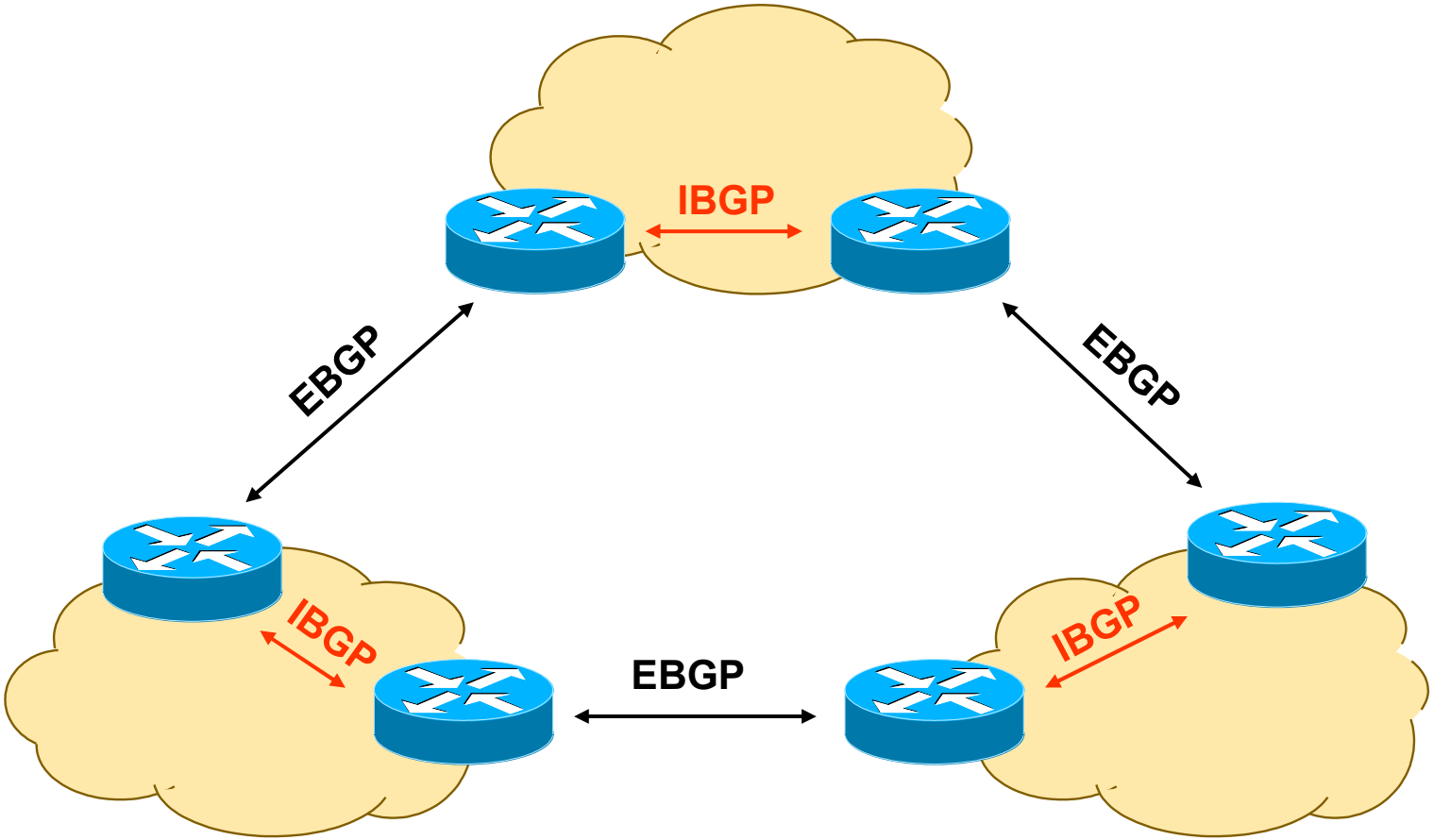
BGP

Internal and External BGP

EBGP and IBGP



Internal and External BGP

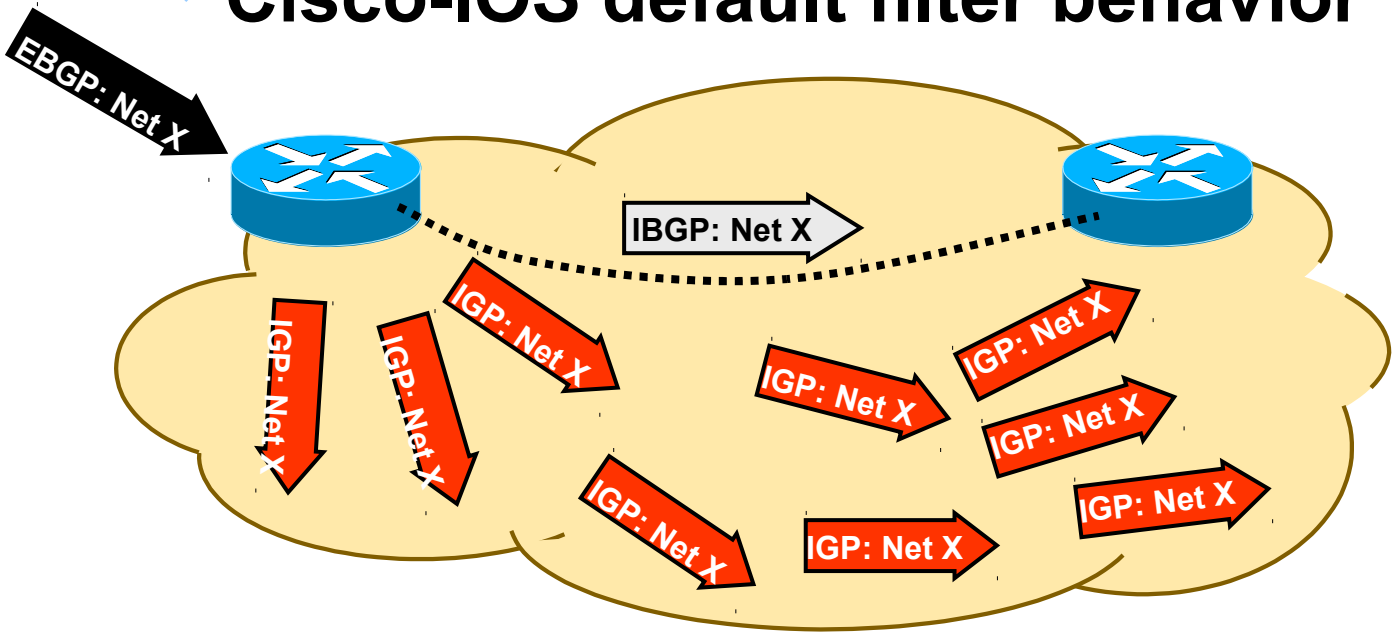
- EBGP messages are exchanged between peers of different ASs
  - EBGP peers should be directly connected
- Inside an AS this information is forwarded via IBGP to the next BGP router
  - IBGP messages have same structure like EBGP messages
- Administrative Distance
  - IBGP: 200
  - EBGP: 20 (preferred over all IGPs)

Loop Detection

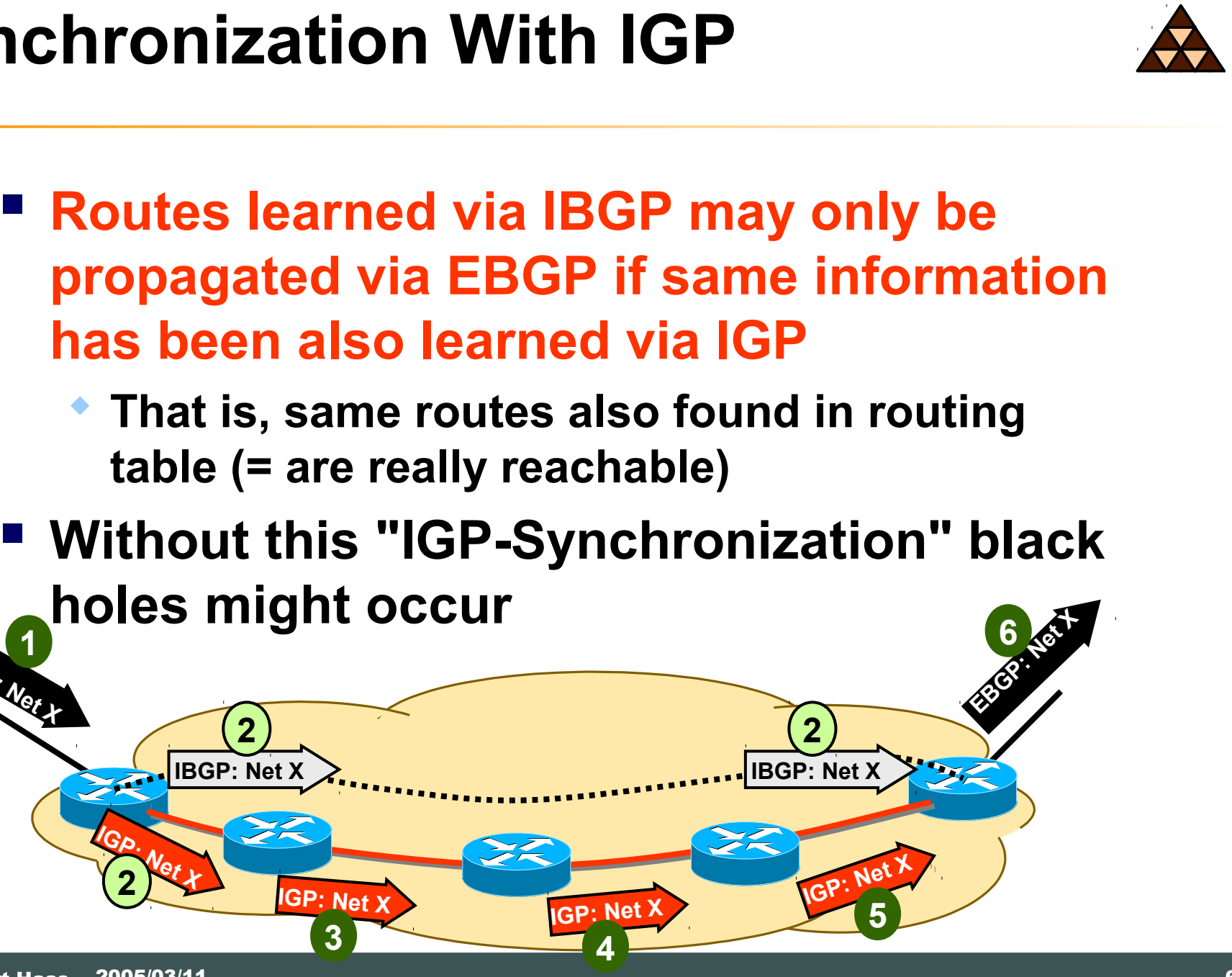
- Update is only forwarded if own AS number is not already contained in AS_Path
- Thus, routing loops are avoided easily
- But this principle doesn't work with IBGP updates (!)
- Therefore IBGP speaking routers must be fully meshed !!!

BGP ➔ IGP Redistribution

- Only routes learned via EBGP are redistributed into IGP
  - To assure optimal load distribution
  - Cisco-IOS default filter behavior



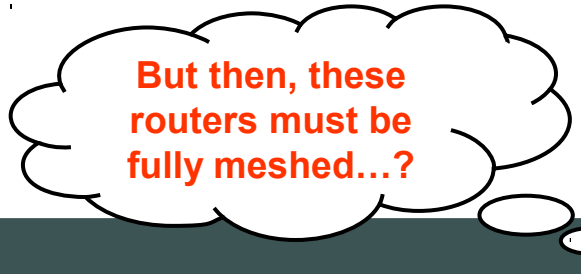
Synchronization With IGP



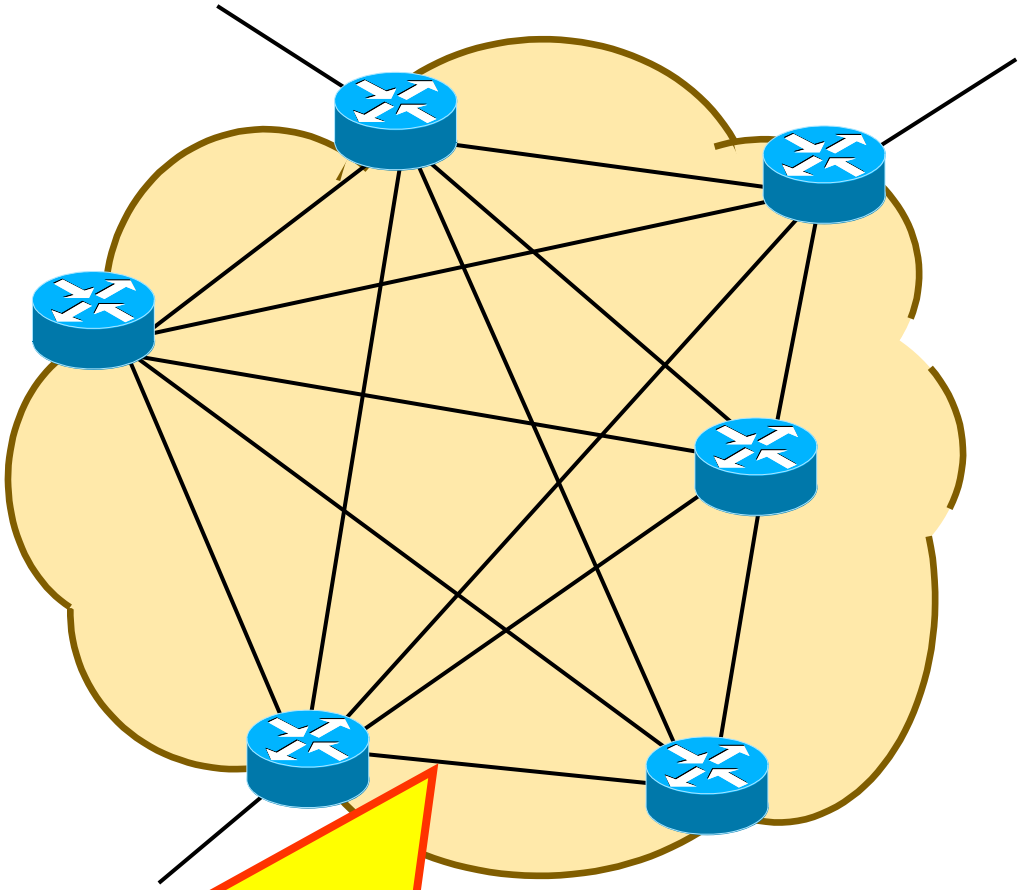
- **Routes learned via IBGP may only be propagated via EBGP if same information has been also learned via IGP**
  - That is, same routes also found in routing table (= are really reachable)
- **Without this "IGP-Synchronization" black holes might occur**

Avoid Synchronization

- Synchronization with IGP means injecting thousands of routes into IGP
  - IGP might get overloaded
  - Synchronization dramatically affects BGP's convergence time
- Alternatives
  - Set default routes leading to BGP routers (might lead to suboptimal routing)
  - Use only BGP-routers inside the AS !



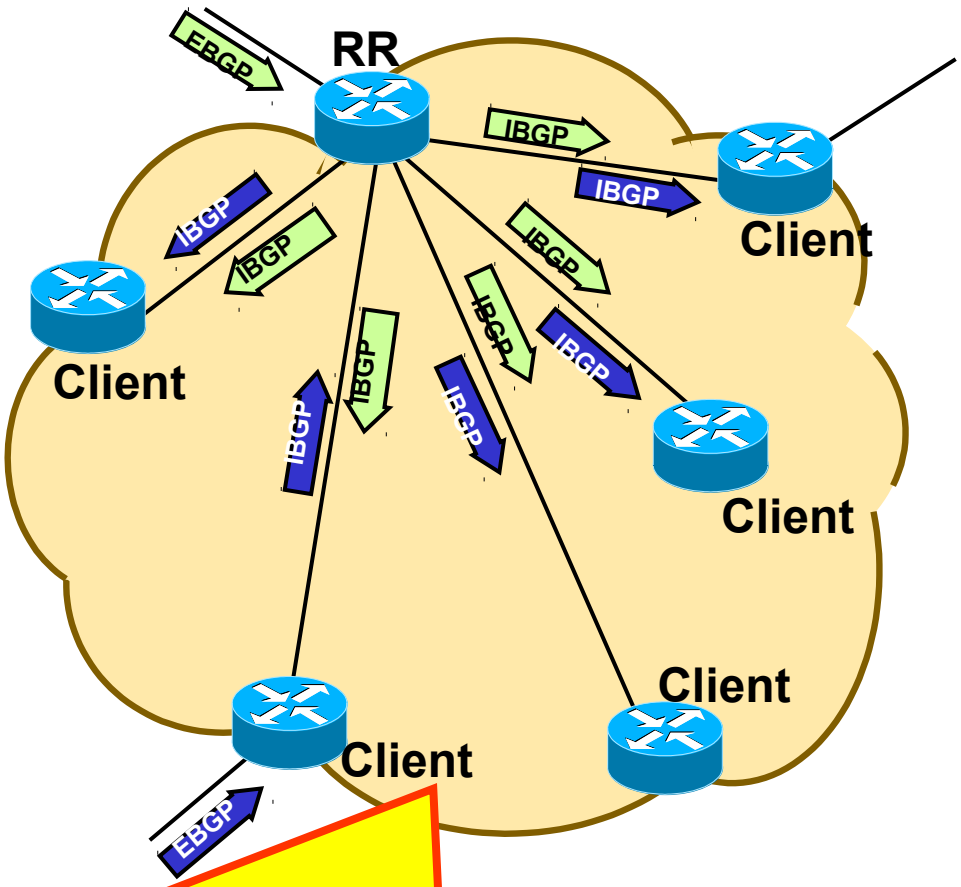
Fully Meshed IBGP Routers



Note: These are logical IBGP connections!
The physical topology might look different!

- Does not scale
  - n(n-1)/2 links
- Resource and configuration challenge
- Solutions:
  - Route Reflectors
  - Confederations

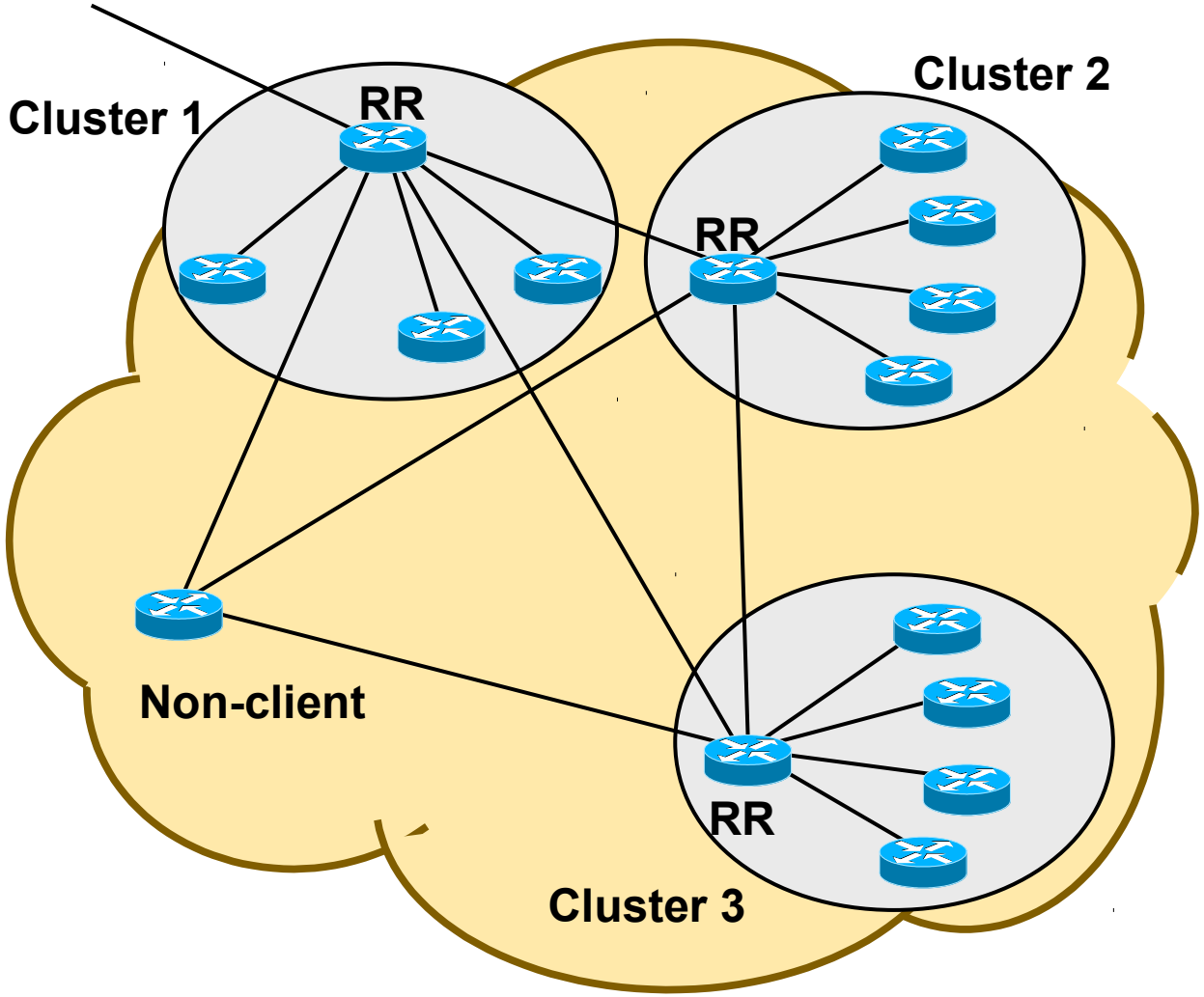
Route Reflector



Note: Although these are logical IBGP connections, the physical topology should be the main indicator for an efficient cluster design (which router becomes RR)

- RR mirrors BGP messages for "clients"
- RR and clients belong to a "cluster"
- Only RR must be configured
  - Clients are not aware of the RR

RR Clusters



- Only RRs are fully meshed
- Special Attributes care for loop-avoidance
- "Non-clients" must be fully meshed with RRs
  - And with other non-clients

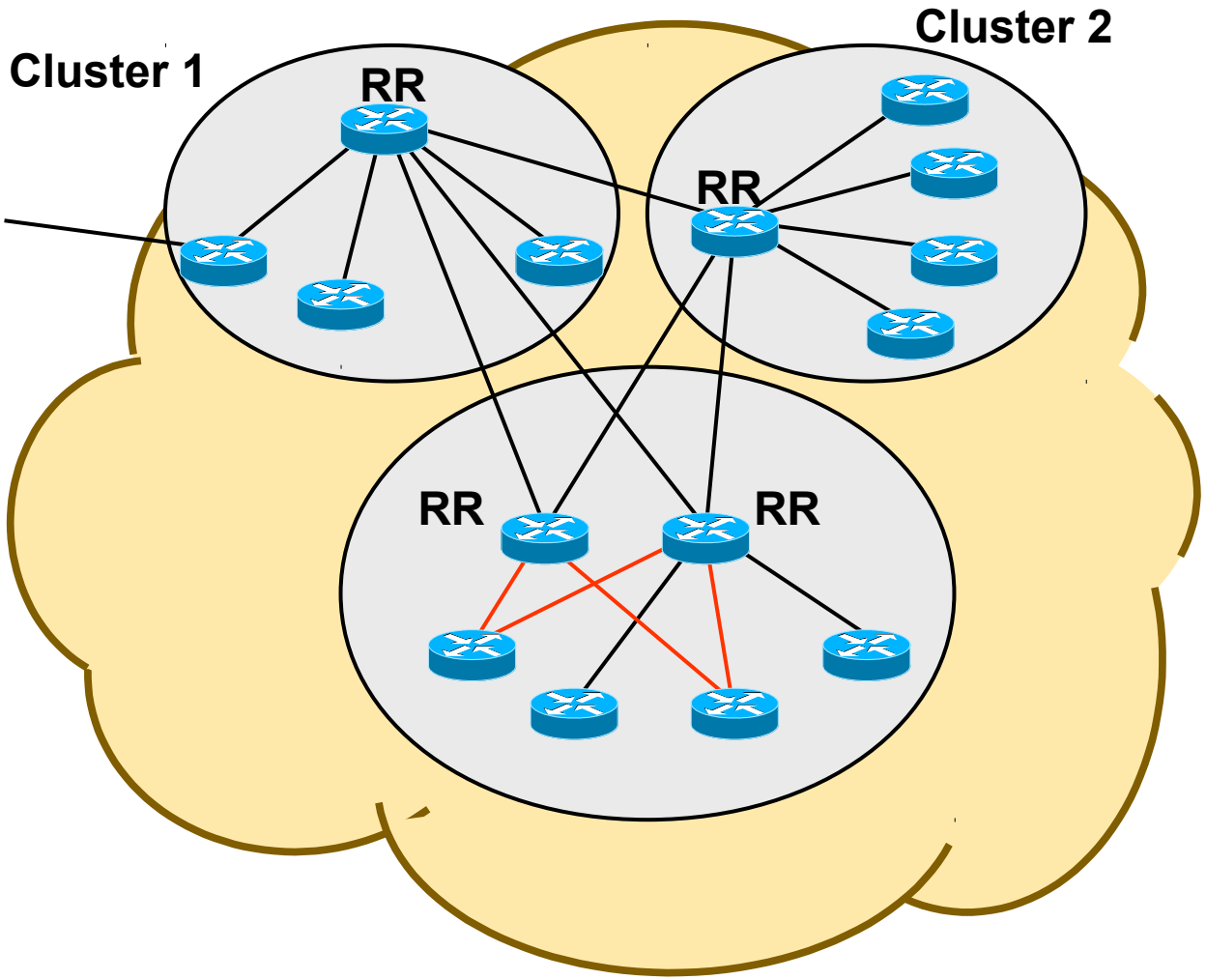
RR Issues

- RRs do not change IBGP behavior or attributes
- RRs only propagate best routes
- Special attributes to avoid routing updates reentering the cluster (routing loops)
  - ORIGINATOR_ID

    Contains router-id of the route's originator in the local AS; attached by RR (Optional, Non-Trans.)
  - CLUSTER_LIST

    Sequence of cluster-ids; RR appends own cluster-id when route is sent to non-clients outside the cluster (Optional, Non-Transitive)

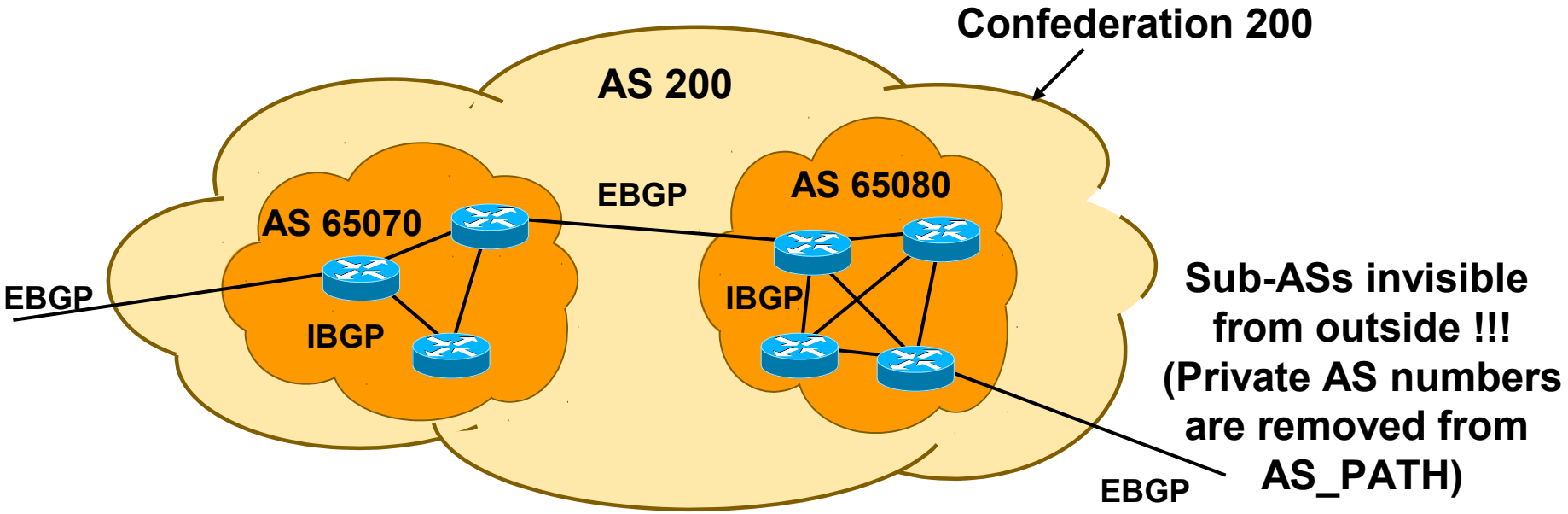
Redundant RRs



- RR is single point of failure
  - Other than fully meshed approach
- Redundant RRs can be configured
  - Clients attached to several RRs

Confederations

- Alternative to route reflectors
- Idea: AS can be broken into multiple sub-ASs
- Loop-avoidance based on AS_Path
- All BGP routers inside a sub-AS must be fully meshed
- EBGP is used between sub-ASs



RRs versus Confederations

- RRs are more popular
  - Simple migration (only RRs needs to be configured accordingly)
  - Best scalability
- Confederations drawbacks
  - Introducing confederations require complete AS-renumbering inside an AS
  - Major change in logical topology
  - Suboptimal routing (Sub-ASs do not influence external AS_PATH length)
- Confederations benefits
  - Can be used with RRs
  - Policies could be applied to route traffic between sub-ASs