BGP Routing

The BGP Routing Principles
and Route Decisions based on AS-Path

Agenda

- **BGP Internals**
- **BGP Session Topologies**
- **Synchronization with IGP**
- **Fail-over Handling**

BGP Routing and BGP Policy

- **in a simple topology of AS´s**
  - routing policy is reduced to a minimal function
    - demonstrated in example 1 and 2 (BGP-4 Fundamentals module)
    - a BGP router can decide only which networks within the own AS should be announced to external BGP neighbors and which learned networks should be advertised into the own AS
  - no route decision must be taken
- **in a complex topology of AS´s**
  - routing policy is necessary to decide which routes should be propagated to other peers
    - BGP policy based on agreements between AS´s
  - in case of several paths to same destination
    - route decision (= selecting the best path) is necessary

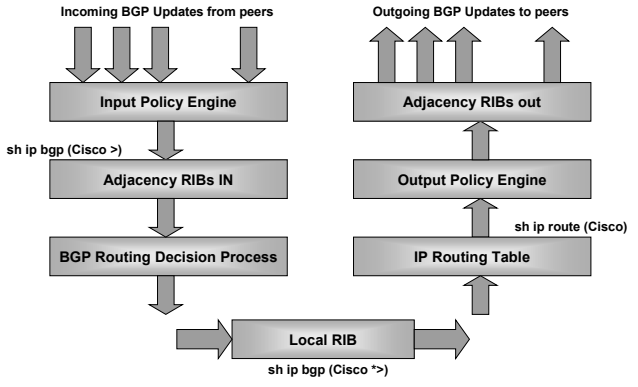
BGP Routing Information Base

- **in a meshed topology of AS´s routing decisions are necessary**
  - to maintain network connectivity across AS´s in case of topology changes
    - e.g. physical link between two AS´s is broken
  - to select the best path in case of several paths to same destination
- **in order to handle routing policy and route decision**
  - BGP routes are stored in three conceptual Routing Information Bases (RIBs) within a BGP router
  - Adj-RIBs-In, Loc-RIB, Adj-RIBs-Out

BGP Routing Information Bases

- Adj-RIBs-In
  - store routing information that has been learned by inbound Update messages; there contents represent routes that are available as an input to the decision process
  - an optional Input Policy Engine can filter routes or manipulate their attributes (policy decision) before this routing information is given to the route decision process
- Loc-RIB
  - contains the local routing information that the BGP router has selected after applying its local input policies and route decisions
  - an optional Output Policy Engine can filter routes or manipulate their attributes before this information is given to peers
- Adj-RIBs-Out
  - contains the routing information that the BGP router has selected for advertisement to its peers

Model of BGP Process



BGP Routing - Path Vector Protocol

- **with the lack of an Input and Output Policy Engine**
  - routing decisions alone will control propagation of BGP routes to peers
- **without special assumptions about use of attributes**
  - only AS_Path, Next_Hop and Origin (the mandatory attributes) are used for routing decisions only
- **path vector protocol**
  - every routing update (BGP route) contains full list of transit networks (AS_Path)
  - handling very similar to distance vector algorithm
    - e.g. split horizon, hop metrics

Path Vector Protocol Details 1

- **basic algorithm of path vector routing**
  - when receiving an update with own AS number contained in the AS_Path, the BGP route will be refused
    - loop protection
  - when an update is received, the path will be compared to the current best path to this destination
  - if the new path is better (e.g. shorter) than the old path
    - the BGP routing table is modified to include the new path
    - corresponding updates are sent to BGP neighbors
  - implementations might include extra features

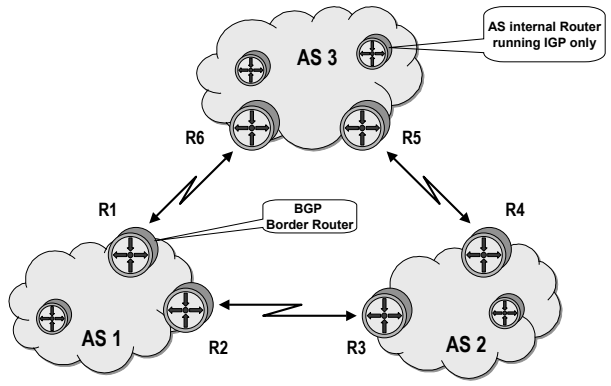
Path Vector Protocol Details 2

- **only the best paths are forwarded to the neighboring ASes**
  - same concept as in distance vector protocols
  - however, if this path is broken the next best path should be advertised
- **the BGP routing table is also modified**
  - if a new path is received from the currently selected neighbor for that destination
  - if a the currently selected neighbors withdraw a route
    - in this case, the last advertisements from all other neighbors are compared and the best path is selected

Agenda

- **BGP Internals**
- **BGP Session Topologies**
- **Synchronization with IGP**
- **Fail-over Handling**

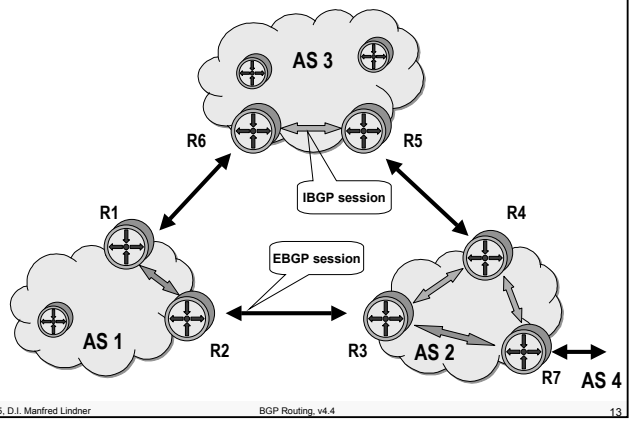
BGP Meshed Topology (example 3)



External <–> Internal BGP Sessions

- **path information between different AS´s is transported by BGP Update message**
  - external BGP (EBGP) sessions
  - direct physical link is necessary
    - for simplicity only, not a strict requirement
      - to eliminate the need for yet another DMZ routing protocol
    - remark: some implementations support non-directly connected EBGP peers (e.g. Cisco´s ebgp multi-hop feature)
- **in order to pass path information within the same AS to other BGP routers**
  - internal BGP (IBGP) sessions are necessary
    - remark: IGP protocols might not be able to pass AS numbers or any other path attributes

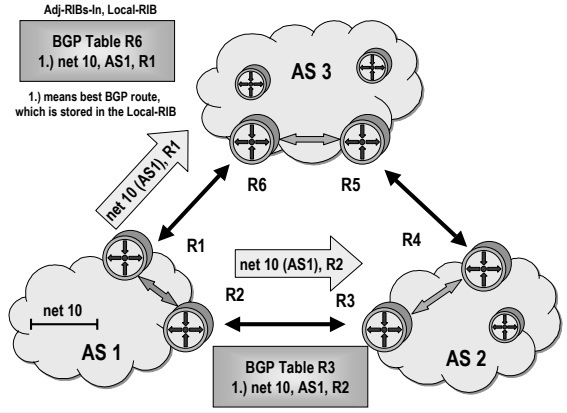
BGP Internal - External Sessions



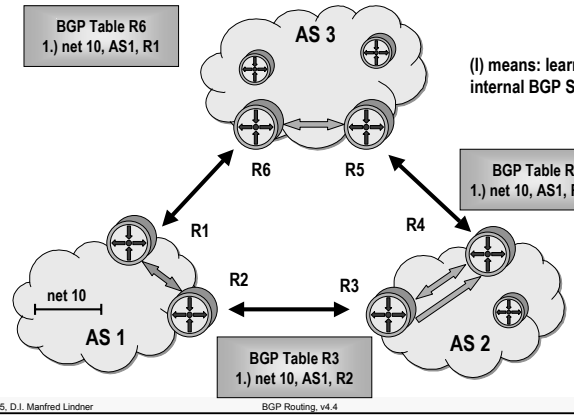
External <–> Internal BGP Sessions

- **different behavior concerning routing updates**
  - update received on an internal connection will not be propagated to other BGP routers of same AS
    - basic BGP loop avoidance does not work when AS number is the same through multiple hops!
  - therefore internal BGP routers must be fully meshed
    - IBGP sessions to every other BGP router of same AS
  - if an external update is received and propagated via an internal BGP connection the Next Hop will be that of the BGP router which originates this external update
    - external router (IP address of this router) must be reachable by IGP or some other means from that internal router
    - recursive lookup of routing table is necessary to find the real next hop in any way, if next hop is not directly connected interface

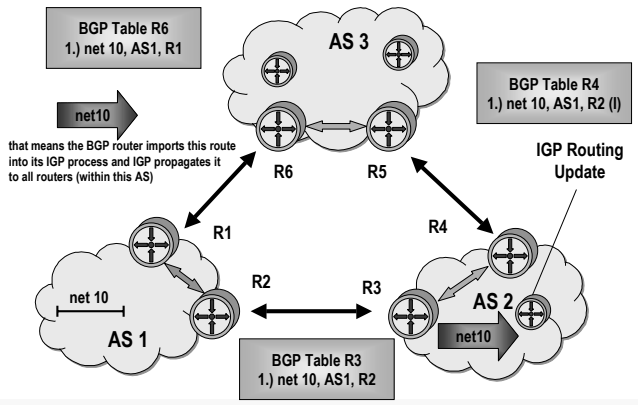
BGP Update (1)



BGP Update (2)



IGP Routing Update into AS2 (3)



Agenda

- **BGP Internals**
- **BGP Session Topologies**
- **Synchronization with IGP**
- **Fail-over Handling**

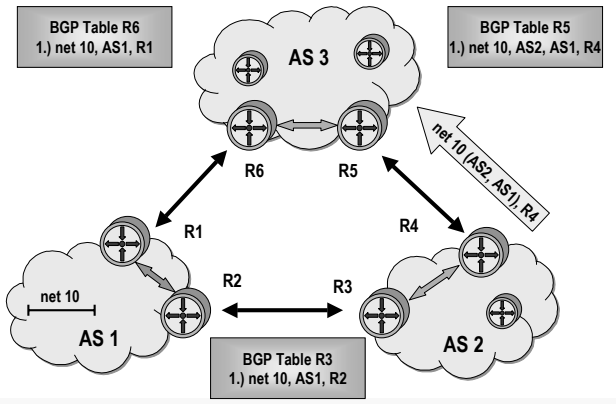
BGP and IGP

- **for an internal IGP-only router in order to achieve reachability of Net-ID´s of other AS´s**
  - default routing within an AS towards BGP Border
    - or
  - route redistribution from BGP into IGP
    - maybe a problem in case of carrying all Internet routes
- **rules for redistribution**
  - only the best BGP route will be installed in the BGP-routers IP-forwarding (routing) table
  - only routes learned via BGP external sessions are redistributed into IGP
    - remark: Cisco-IOS default filter behavior

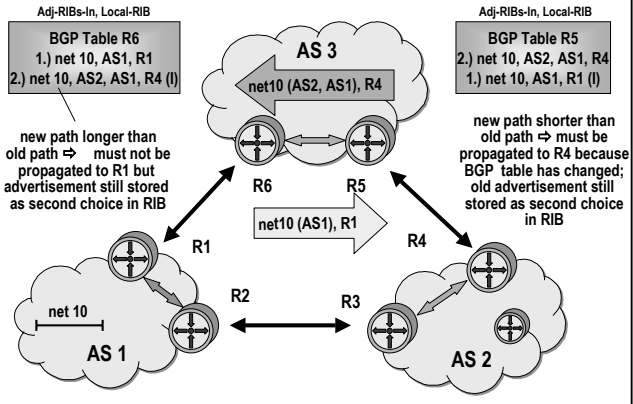
BGP and IGP Synchronization

- **if an BGP router receives**
  - a route via internal BGP, it must wait until this route is reachable via IGP too before this route could be announced to an external BGP peer
  - this is called BGP synchronization
- **reason:**
  - if router would propagate this route earlier, the AS would get traffic for that destination but this traffic could not be passed through the own AS

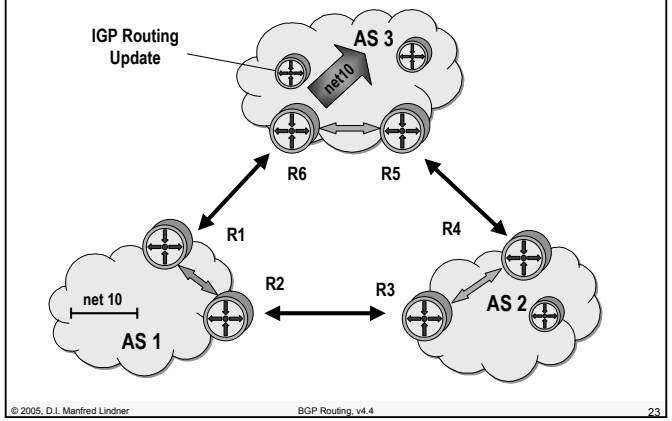
BGP Update after BGP Synchronization (4)



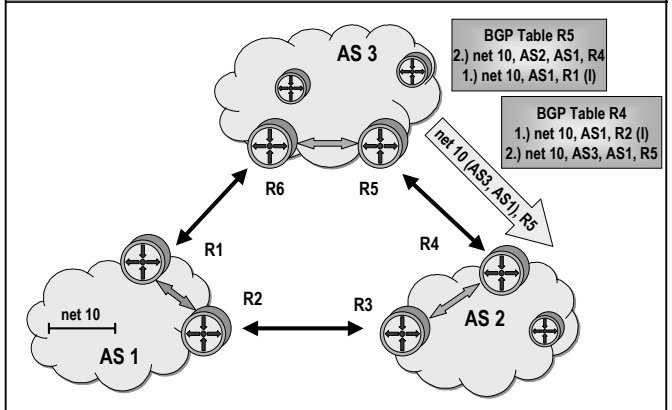
BGP Update (5)



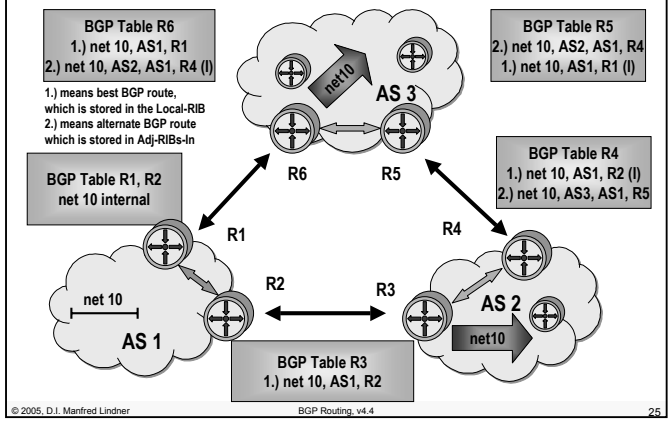
IGP Routing Update AS3 (6)



BGP Update Meshed (7)



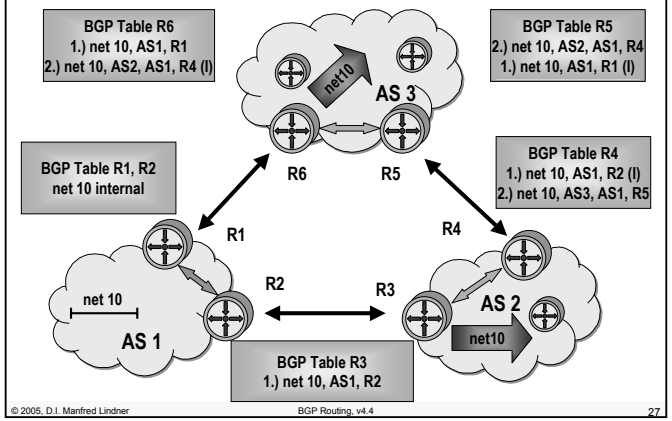
BGP Final Picture (8)



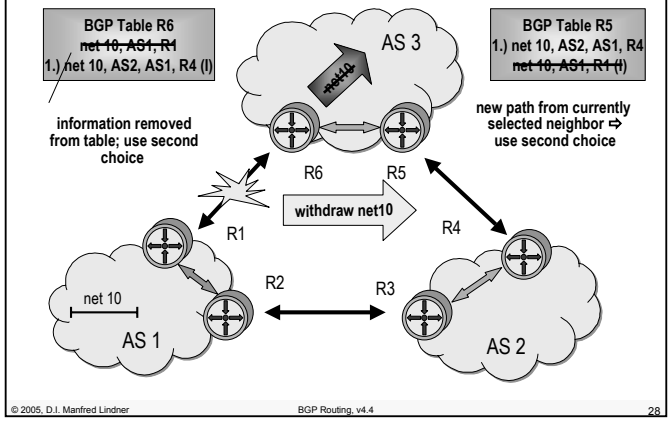
Agenda

- BGP Internals
- BGP Session Topologies
- Synchronization with IGP
- Fail-over Handling

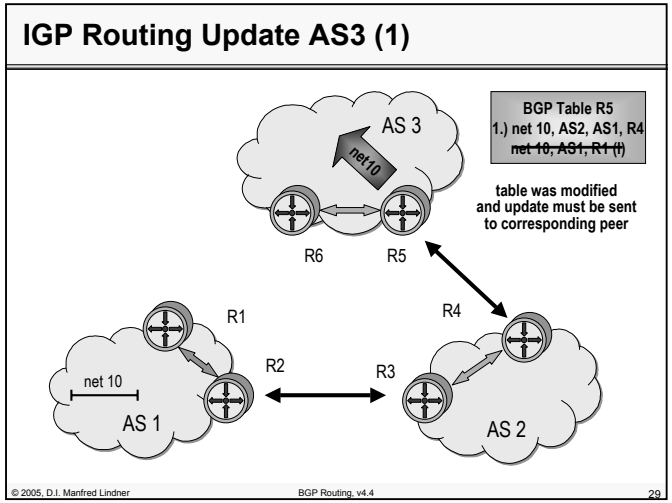
BGP Start Picture (Scenario 1)



Session Loss R1 – R6 (Scenario 1)

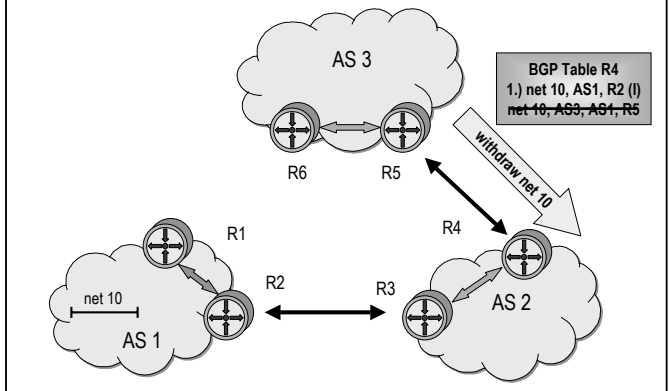


IGP Routing Update AS3 (1)



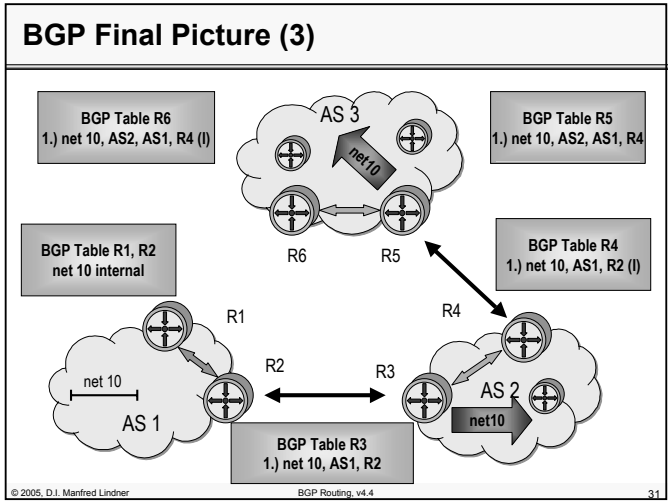
 BGP Routing, v4.4 29

BGP Incremental Update (2)



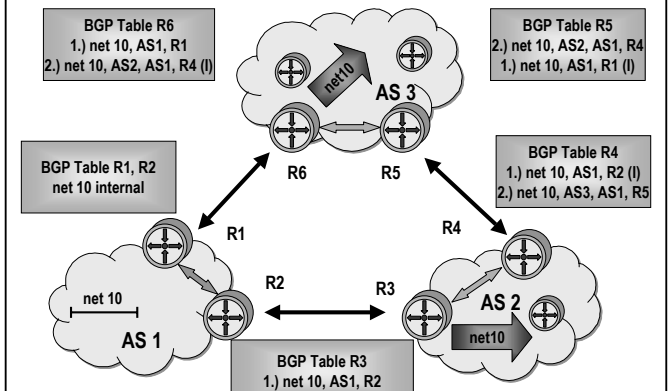
 BGP Routing, v4.4 30

BGP Final Picture (3)



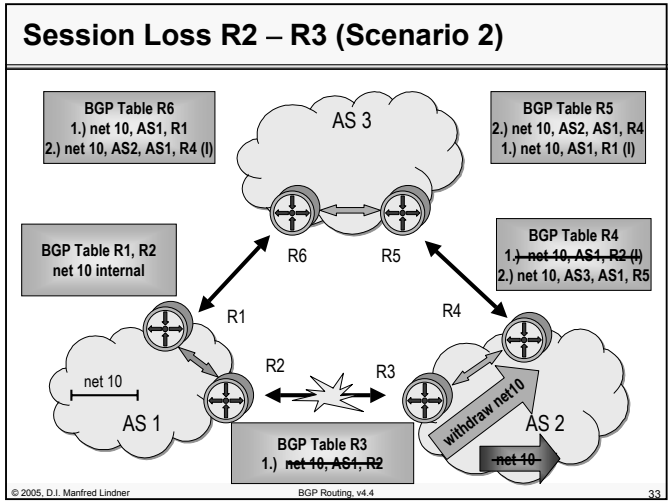
 BGP Routing, v4.4 31

BGP Start Picture (Scenario 2)

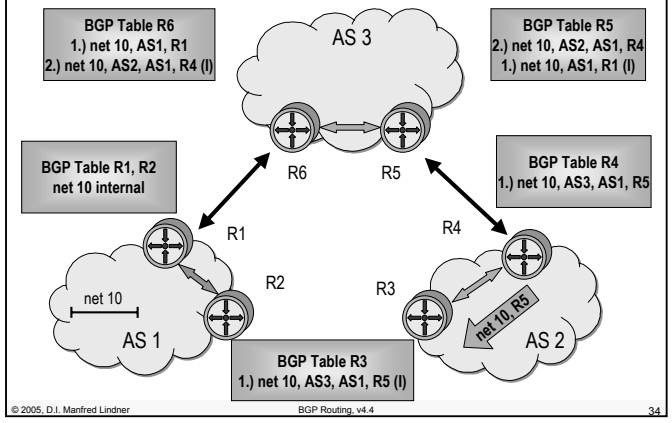


 BGP Routing, v4.4 32

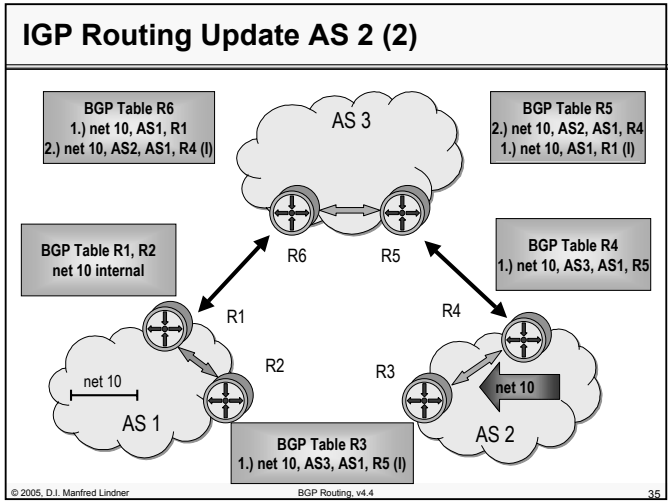
Session Loss R2 – R3 (Scenario 2)



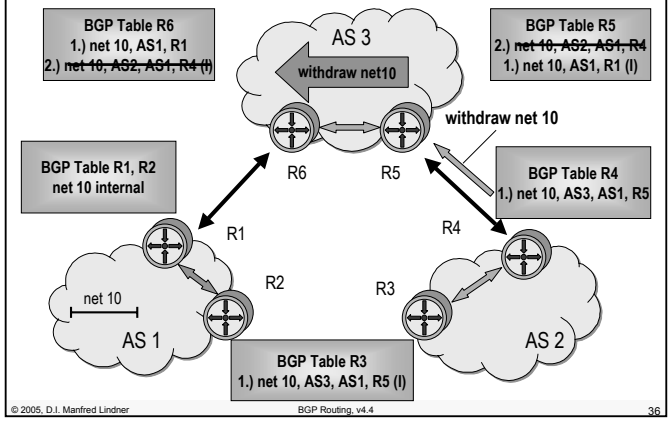
BGP Incremental Update (1)



IGP Routing Update AS 2 (2)



BGP Incremental Update (3)



BGP Final Picture (4)