Link-State Routing Algorithms and Protocols

Petr Grygárek

Link state algorithms

Why “link state” ?

1. Every router continually checks links to it’s neighboring routers (using Hello protocol)
2. If a link changes state (i.e. fails, goes up or Hello response time changes considerably), every router in routing domain is informed.

Basic principle of LS algorithm

- Hello protocol continually checks links to neighbors
  - Neighborship states maintained in Adjacency database
- If state of some link changes, link state update is sent to other routers
  - Sometimes called Link State Advertisements (LSA)
- Link state advertisements flooded to whole routing domain (reliably)
- All routers have (the same) topology database, i.e. graph of network topology
  - Sometimes called Link-State Database
  - Every router is uniquely identified
    - link is a connection between two uniquely identified routers
  - It is also necessary to be able to represent multiaccess network segments
- Every router computes tree of shortest paths to other networks with itself as root
  - Uses Dijkstra algorithm see http://students.ceid.upatras.gr/~papagel/project/kef5_7_1.htm
  - See http://www.cs.uwa.edu.au/undergraduate/courses/230.300/readings/graphapplet/dijkstra.html for Dijkstra algorithm operation example.
- From shortest path tree, routing table is created

Link state change detection

Link state change mean that

- Neighbor router is now reachable/unreachable
  - Link or neighbor router goes up/down (including addition of a new link)
- Metric of a link changed
  - Change made by administrator
  - Delay of Hellos passed between neighbors averaged over time changes considerably (implemented rarely)

- Detected by hardware: interface state changes
- Detected by IGP protocol (Hello):
  - interface still up, but neighbor lost
  - new neighbor found

Note: On multiaccess network, interface can be up, but some neighbor down

Link State Advertisement

- Sent every time link state changes
- Contains
  - Sending router identity
  - List of neighboring routers
    - together with current metric for link to each of them
  - Sequence number
    - Needed to distinguish between older and newer LSAs
    - Incremented every time new LSA is sent
  - MaxAge
    - To defend against outdated information, LSAs have limited lifetime
    - Need of periodic refresh of LSAs generated by every router
      - Typical period of 20 – 30 mins, randomized to avoid congestion

LSA Flooding

- LSA has to be processed on each router
- Need for reliable and loop-free flooding

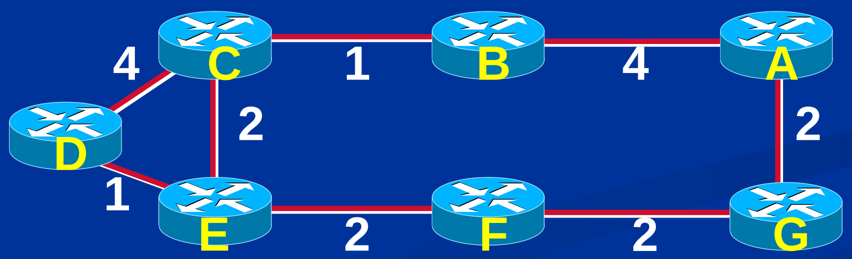
On LSA arrival

- If router already has this version (i.e. seq. number originated by the same router) of LSA, router ignores it
- If LSA carries any changes, router schedules SPF tree calculation (full/partial)
- Router installs LSA in the database (or reset aging timer)
- Router resends LSA to adjacent neighbors except that from which LSA came

Link State Database (topology database)

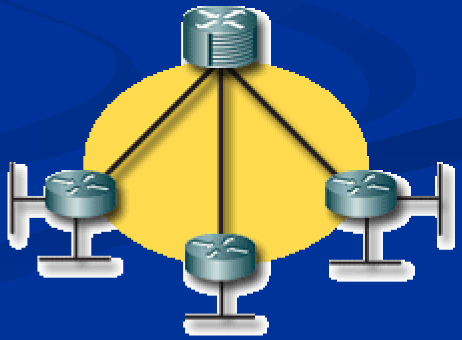
Topology database is a collection of LSAs originated by individual routers

A	B	C	D	E	F	G
B/4 G/2	A/4 C/1	B/1 D/4 E/2	C/4 E/1	C/2 D/1 F/2	E/2 G/2	A/2 F/2



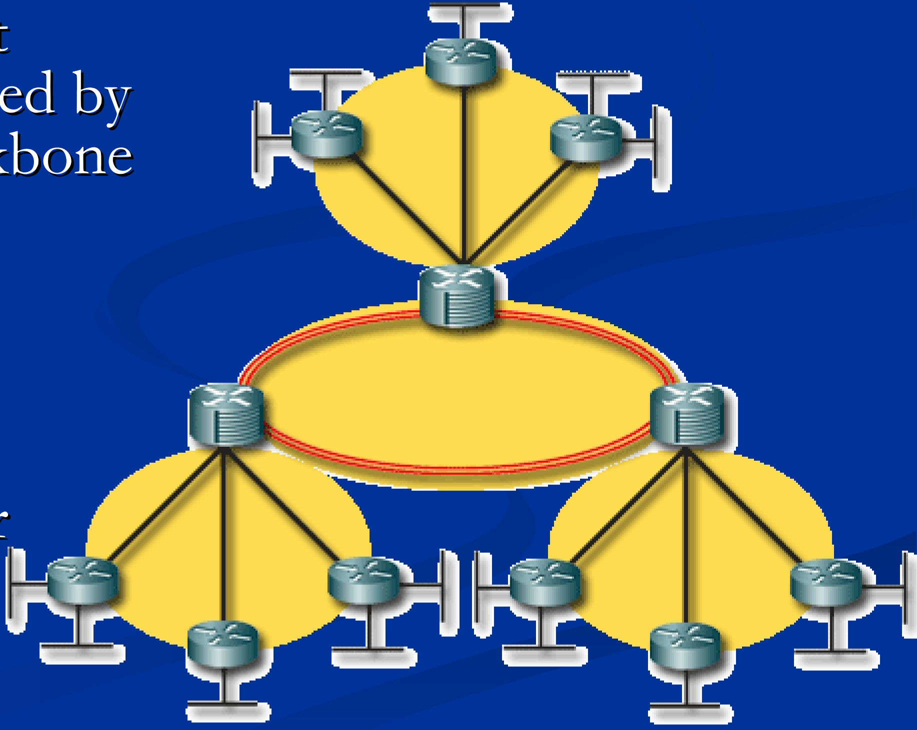
LS Algorithms Scalability (1)

- Hierarchical routing principle (2-level hierarchy)
- Autonomous system can be divided into areas
  - Area – a set of routers sharing the complete topology information for that area
- Routers in the same area have the same topology database
- Routers at area boundary maintain topology databases for of all adjacent areas
- Topology of the area not known outside area
  - only summary information about networks in that area is exported



LS Algorithms Scalability (2)

- Router has separate topological database for every connected area
- LSA flooding bounded by area boundary
- Only summary information about networks in the area are propagated by the area border router to the backbone
- SPF calculation is performed independently for each area
- Distance vector routing principle between areas
- Hierarchical addressing allows for summarization



LS - Advanced issues

- Route flapping – holddown timer on link state update generation and shortest path tree recalculation
- Partial SPF calculation
- LS algorithms on dial-on-demand circuits
  - Hellos not sent over link
  - "don't age" flag in LSA (no need for periodic refresh)
- Designated router – represents multiaccess network
  - propagates network pseudonode
  - Limits number of adjacencies on multiaccess network and thus number of hellos transmitted
- LSA refresh – periodic reflood of LSA + limited lifetime of LSAs
- Passive interfaces
  - adjacencies not created

Route filtering in LS routing

- IN filtering
  - Filtering of routes accepted to routing tables
  - Does not affect flooding of LSAs
  - Done on every router independently
- Out filtering
  - Only on area or AS boundary
  - Filtering of (summary) LSAs generated by ABR and ASBR – inter-area filtering (out of area, to area)

Advantages of LS routing algoritms

- Fast convergence
- Minimal overhead after reaching convergence
- Event-driven - updates sent only on link state change
  - Incremental updates
- All routers have the same and complete information about network topology
  - less chance of calculating wrong routing tables
- Better metric than hop count
  - Cost defined by administrator for each link
  - No hop-count limit on network topology

Disdvantages of LS routing algoritms

- Higher processor utilization
- Higher memory requirements
- More planning before implementation
- Route filtering not so transparent as with DV algoritms

OSPF

Open Shortest Path First (OSPF)

- SPF – Shortest paths are calculated first, i.e. before routing tables are created and routing of packets starts
- “Open” = Open standard
- Development began 1987 by IETF
- OSPFv2 on 1991, latest revision is RFC 2328 (1998)
- Full support for VLSM and route summarization
- Metric: administratively defined cost
  - (default 100 Mbps/ bandwidth)
- Uses IP for transport, IP protocol 89
- Uses multicast addresses for neighbor maintenance and flooding of LSAs
  - 224.0.0.5 – All OSPF Routers
  - 224.0.0.6 – All DRouters

Data structures of OSPF router

- Adjacency database
- Topological database
- Routing table

Neighbor and Adjacent routers

Neighboring routers – connected to the same link (exchange Hello packets)

Adjacent routers – exchange routing information (LSAs)

- On point-to-point links, routers are always adjacent
- On multiaccess networks, star topology of adjacencies is created (not full mesh)
  - The center of the star is called "Designated router"

Adjacency database - keeps track of all directly connected routers (neighbors) and states of the respective adjacencies.

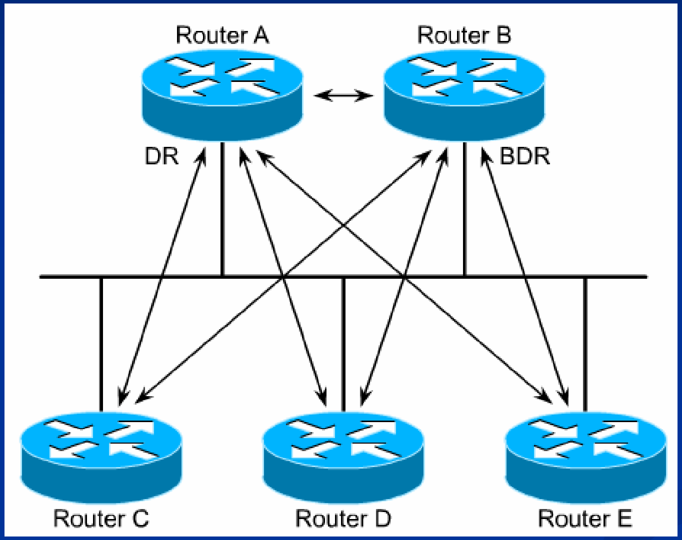
OSPF Hello protocol

- Neigbor discovery and continuous testing of neighbor reachability
- Multicast (224.0.0.5) on all router interfaces
  - Hello interval: 10 sec. LAN, 30 sec. NBMA
  - Dead time interval: 4 x Hello interval
- Used to form adjacencies between routers
  - Checks for match of IP subnet, Hello interval and Dead interval
- Hello packet carries a list of neighbors known to the sender

Designated router and Backup Designater Router

- DR = router that is elected on multiaccess networks to be the focal point for routing updates
- The goal is to lower routing traffic and make LSA propagation more organized
  - Eliminate many copies of the same LSA on the network if passed recursively along full-mesh to adjacency graph edges
- BDR behaves like normal router, but takes over DR if DR fails
- Every router (including BDR) forms adjacency with DR
- On multiaccess network, every router sends its updates to DR. DR sends it back to other routers on that network
  - DR serves as collection point for Link State Advertisements (LSAs)
- Multicast addresses reserved for DR/BDR group and All-SPF-Routers group
  - Link State Updates are sent to DR/BDR multicast group and then back to AllSPFRouters group

Adjacencies on multiaccess network



- Routers are adjacent with DR and BDR, but only DR resends LSA sent to it to other routers on multiaccess networks
  - BDR sits quietly and only listens whether DR responds

DR and BDR election

- Election occurs if routers see Hellos with no DR/BDR address
- The router with the highest priority configured on the interface connected to multiaccess network becomes DR (/BDR) on that network
  - In case of equal priorities, router with highest IP address configured on any interface wins
  - Priority 0 means that router cannot be elected
- The BDR is chosen the same way
- The election is non-preemptive
  - If a router with higher priority comes up and DR/BDR is already elected, nothing changes
  - If a DR fails, the BDR becomes DR even if a new router with a higher priority comes up during the election process
  - In fact, the router which goes up first becomes DR commonly, disregarding priorities (commonly criticized OSPF behavior)

OSPF adjacency states (1)

Every adjacency relaltion passes through a sequence of states

- Init
  - Entered when a router receives first Hello packet
  - Router periodically sends Hello packets
- 2Way
  - Entered when a router receives Hello packet containing it's own Router ID (meaning that neighboring router mentioned it's presence)
  - In 2Way state, router decides who to establish a full adjacency with, depending upon the type of network that the particular interfaces resides on.
    - point-to-point link: the sole link partner
    - multi-access link: DR/BDR if elected
      - If DR/BDR not elected, otherwise election process takes place here
  - On multiaccess networks, adjacencies with non-DR neihgbors ends up in 2Way state

OSPF adjacency states (2)

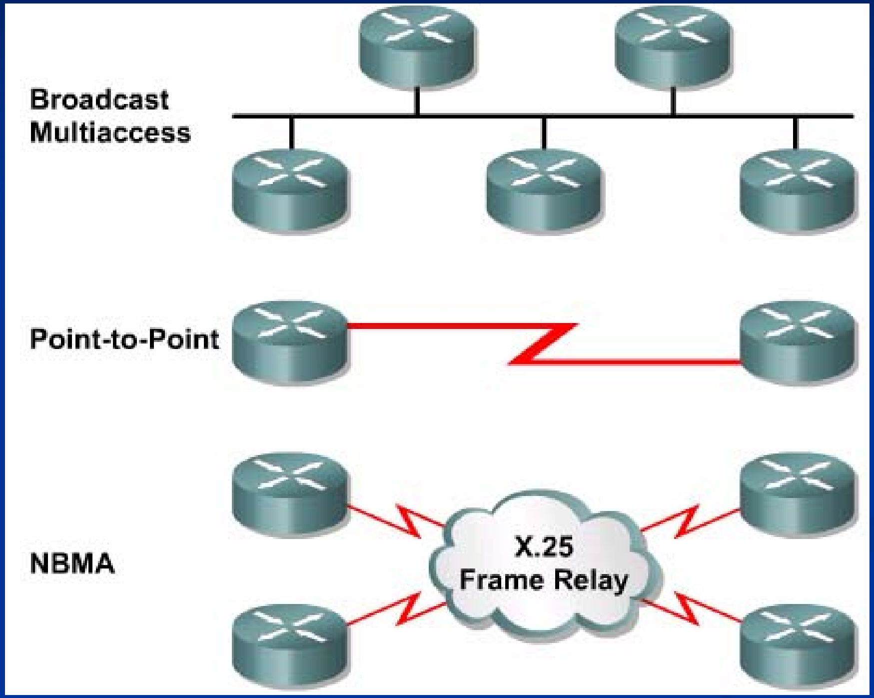
- ExStart
  - The router which will start the link state database description exchange is determined here
- Exchange
  - Exchange of Database Description Packets (DBDs) – reliable process
  - DBDs contains LSA headers of the topology database
  - Topology databases are exchanged mutually
    - (consider a link of two separate set of routers going on)
- Loading
  - Link state Request (LSR) for specific part of topology database (LSA)
  - Only Link State Updates (LSU) with required LSAs sent on response
- Full
  - link state databases of adjacent routers are now identical

OSPF timers

- Hello and Dead timers
  - Contained in Hello packets –must match to form adjacency
- Hello interval setting consideration
  - Hellos must be able to pass even in case of congested network
  - Router must be able to send hellos fast enough even if busy with another tasks
- Timers increasing stability
  - LSA generation delay after event detection – 500ms
    - (the purpose is to collect changes implied by that event into one LSA)
  - Maximum frequency of LSA origination (once per 5s)
    - (defends against flapping links)
  - Maximum frequency of LSA reception (1s, LSA received at higher frequency are discarded)
    - SPF calculations are delayed by 5 seconds after receiving an LSU (Link State Update)
  - Cisco: Delay between consecutive SPF calculations is 10 seconds by default

Stability vs. fast convergence

OSPF Network Types



NBMA=Non-Broadcast MultiAccess

- more than one neighbor can be reached by one interface (using virtual circuits), but broadcasting to all neighbors is not supported
  - broadcast packet has to be replicated into every virtual circuit
- NBMA cloud may be use single IP subnet

Operation on various network types

- Slight difference of OSPF operation on various network types
  - Topology of adjacencies
  - Neighbor discovery (multicast/preconfigured neighbors)
  - Different default timers ((hello, holdtime) for different network types
- Network types introduced because of need to represent network node in topology graph
- Problems with DR election on NBMA if the topology is not full mesh

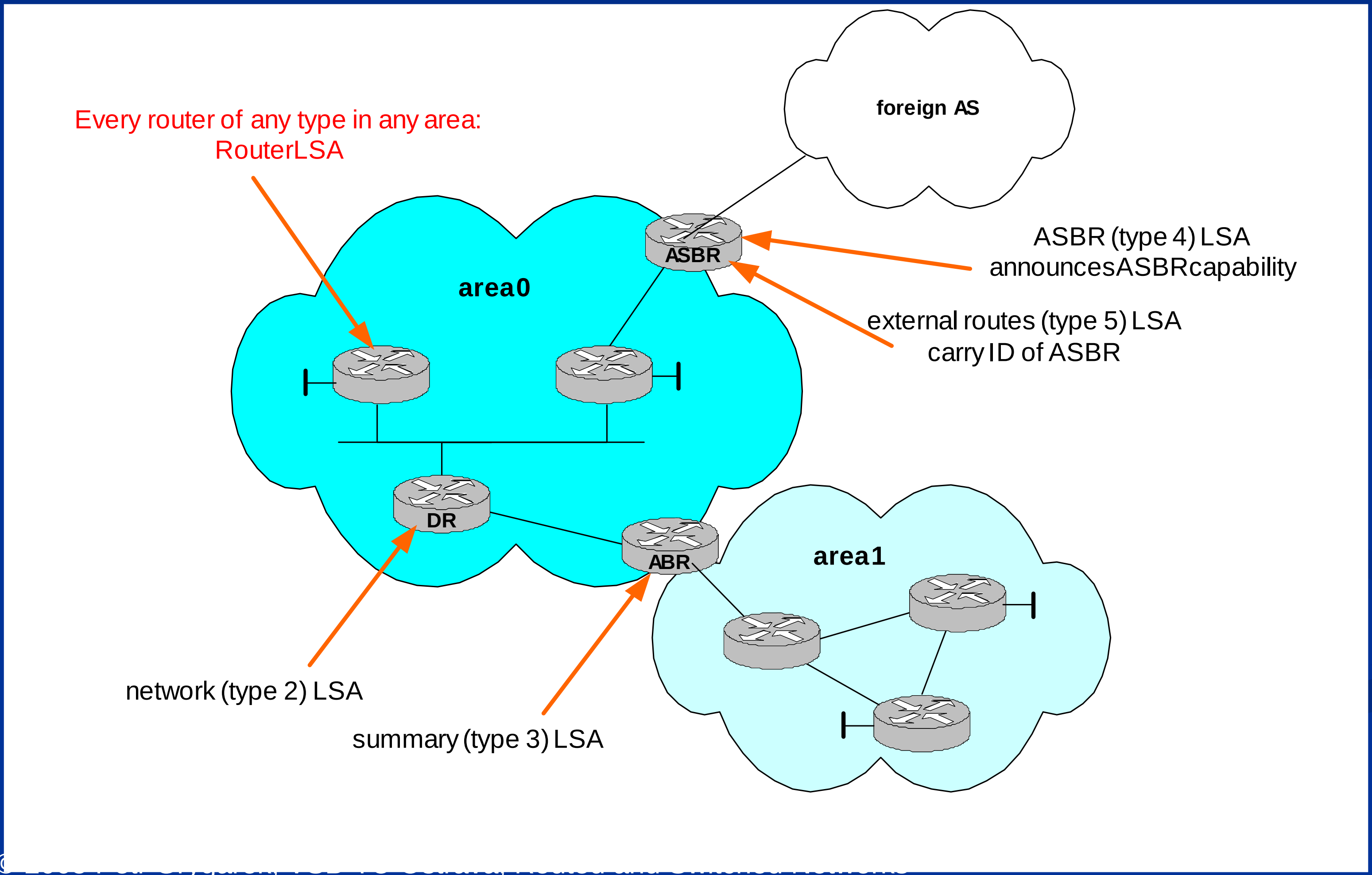
OSPF packet types

1	2	4
Version #	Type	Packet Length
Router ID		
Area ID		
Checksum		Autype
Authentication		
Authentication		

Type	Description
1	Hello (establishes and maintains adjacency relationships with neighbors)
2	Database description packet (describes the contents of an OSPF router's link-state database)
3	Link-state request (requests specific pieces of a neighbor router's link-state database)
4	Link-state update (transports link-state advertisements (LSAs) to neighbor routers)
5	Link-state acknowledgement (Neighbor routers acknowledge receipt of the LSAs)

OSPF LSA types

LSAs carried in Link State Update (LSU) packets



LSA Types – Router LSA

- Type 1 – Router LSA
  - Describes states (including costs) of links in the same area
    - Network number and subnet mask for every link also included
    - Link type (point-to-point, multiaccess) indicated
    - Stub network indicator
  - Indicators that router is ABR, ASBR or virtual link endpoint

LSA Types - Network LSA

- Type 2 – Network LSA
  - Generated for every (transit) broadcast and NBMA network
  - Describes all routers connected to that network
  - Generated by DR
  - Contains network IP address and subnet mask

LSA Types – Summary LSA

- Type 3 – Summary LSA
  - Generated by ABR
  - Describes networks inside/outside area
  - Also used to distribute default route

- Type 4 – information about ASBR (router ID)

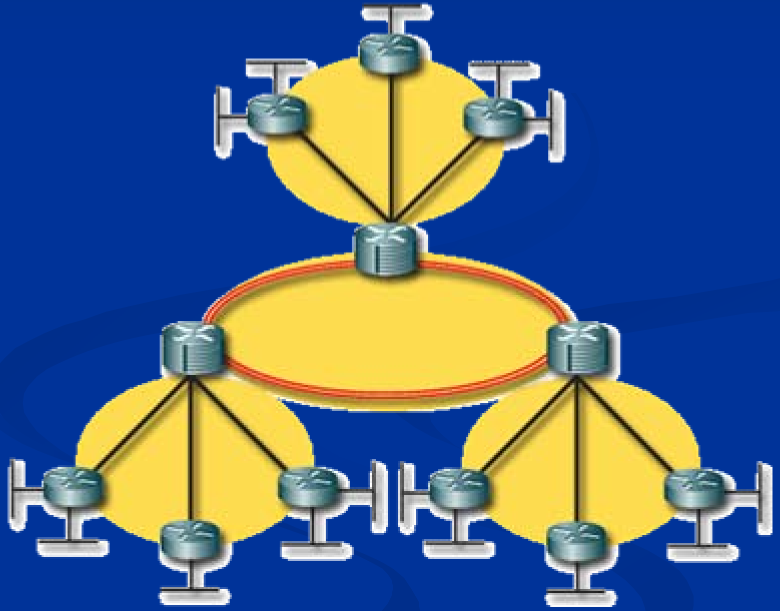
LSA Types – External LSA

- Type 5 – External LSA
  - Describes routes to networks external to AS
  - Also used to distribute default route
  - Two external route types
    - E2 – works with external metric only
    - E1 – adds internal metric to the external metric
  - Contains advertising router ID, network IP address, subnet mask, metric, metric type

To be continued on the next lecture…

OSPF Scalability

- Uses hierarchical routing concept
  - 2-level hierarchy (multiple areas connected to single backbone – area 0)
  - Every area must be connected (only) with backbone area

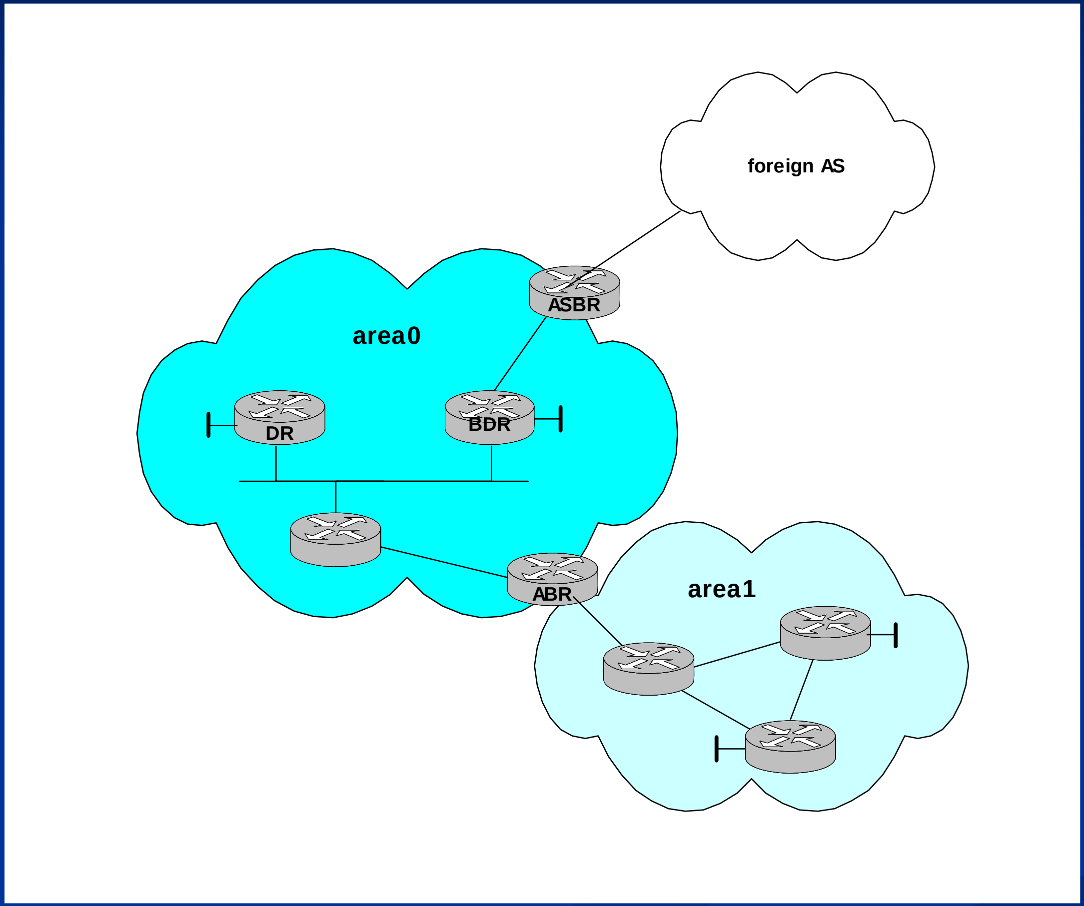


OSPF scalability - areas

- Areas labeled with a 32-bit number
  - single decimal or IP address format
  - Backbone area always denoted as 0.0.0.0

OSPF router types

In OSPF, boundary between areas goes across routers, i.e. individual interfaces of the same router may belong to different areas



OSPF Router Types

- Internal routers – all interfaces in the same area
- Backbone routers – at least one interface in backbone area
- Area Border Routers (ABR) – backbone router connected to at least one non-backbone area
  - area can be connected to backbone by more than one router
- Autonomous System Border Router (ASBR) – imports external routing information into OSPF AS

OSPF Area Classification

- Areas differ in what information they have about rest of AS
  - i.e what they hold in topology database and how routers inside area reach networks outside the area
- In particular area type, all or only some of LSA types are flooded
- The purpose is to limit topology database and routing tables and LSA flooding in some areas

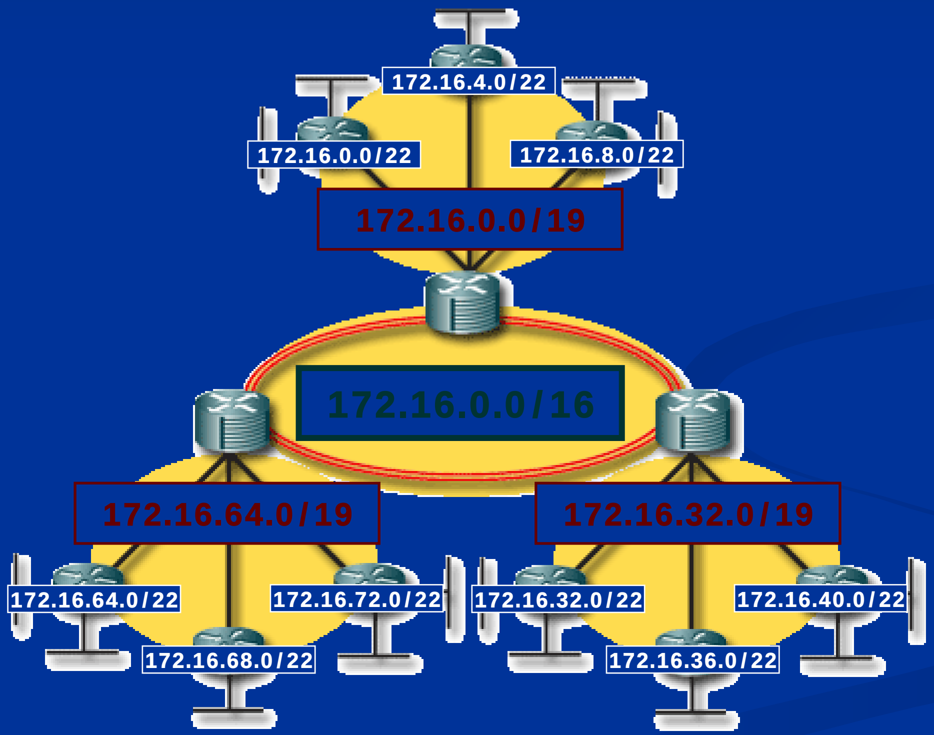
OSPF Area Types

- Backbone
- Normal
- Stub
  - does not propagate Type 5 LSA
  - networks external to AS reached using default
- Totally Stubby (Cisco proprietary, but recommended)
  - does not propagate Type 5 and 3 LSA
  - networks external to area reached using default
- Not-So-Stubby-Area (NSSA)
  - Stub area where ASBR is connected
    - Instead of Type 5 LSA, new Type 7 LSA is used in NSSA
    - Type 7 converted to Type 5 on ABR when injected into backbone

Summarization in OSPF

- Support for route summarization
  - (on ABR/ASBR)
- Summarization is possible if IP addressing properly designed
- Summarization helps tp defend against instability
  - (route flapping not propagated outside area)
- Possibly results to suboptimal routing

Summarization example



Other OSPF features

- Support for authentication
- Equal-cost load balancing
- Support for IP unnumbered
- Virtual links
- Authentication: per neighbor/per area
- Default propagated as 0.0.0.0 network
- OSPFv3
  - Multi-topology routing (QoS-oriented routing)
  - Support for multicast distribution

IS-IS

IS-IS

- Designated to be routing protocol for ISO ConnectionLess Network Protocol (CLNP)
  - IS=Intermediate System (router in ISO terminology)
  - ES=End System
- Works with ISO addresses
  - router identified by Network Entity Tittle (NET) : Area ID + System ID (MAC address)
- Support for multiple network layer protocols (multiprotocol routing)
  - Integrated IS-IS: addresses of foreign address families (including IP) treated as additional attributes

IS-IS: Basic Concepts

- Principles and operation similar to OSPF
- Two-level hierarchy: backbone (level 2) + areas (level 1)
- L1, L2, L1/L2 routers
- Adjacency created between
  - L2-L2, L1/L2-L2
  - L1/L2-L1, L1-L1 – if both in the same area (area ID in Hello packets)
- Backbone is a contiguous set of routers working at Level 2
  - does not have backbone area
- Boundary between areas goes across links – router belongs to exactly one area
- Strict hierarchical routing
  - Only area IDs propagated at level 2
    - advantage of strictly hierarchical ISO addresses
  - L1/L2 routers advertise themselves as candidates to default for L1 routers
  - L1 routers reach networks outside their area using default
    - no concept of summary LSAs sent between areas as with OSPF
    - may result to suboptimal routing
- Broadcast networks: Concept of DR
  - DR called Designated Intermediate System (DIS) in ISO terminology

IS-IS: Comparison with OSPF (1)

- Developed in parallel with OSPF – mutual influence
- Supports areas with more routers (thousands)
  - Less LSA exchange overhead than OSPF
- Sometimes criticized for small interface metric (6b)
- Used as IGP of some large Internet providers
  - Not very popular
- Not bound to TCP/IP – multiprotocol support, routing info exchanged in 802.2 frames
- Broadcast and Point-to-Point network types, no NBMAs

IS-IS: Comparison with OSPF (2)

- Slightly different operation of designated router (Designeted Intermediate System-DIS)
  - full mesh of adjacencies
  - Topology database synchronization based on periodical multicasts of database description from DIS
  - Adjacent router may request particular LSA if it does not have one or has older version than LSA announced by DIS
  - Router can send DIS a LSA it does not hear it in database description or if it has newer version
  - Acknowledgements replaced by waiting for next database description multicast
    - LSA to DIS: DIS announces acceptance implicitly by presence of the accepted LSA in next database description
    - LSA from DIS: if adjacent router misses some information from DIS, it will mention that when next database description is sent and requests the particular information again

Labs

- OSPF single area - observing adjacencies
  - Debug ip ospf adjacency
- OSPF multiple area – various area types (difference of routing table contents)
- Virtual link