OSPF Fundamentals

Open Shortest Path First Routing Protocol
Internet´s Second IGP

Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

OSPF - Open Shortest Path First

- **distance vector protocols like RIP have several dramatic disadvantages:**
  - slow adaptation of network modifications !!
  - size of routing update is proportional to network size !!

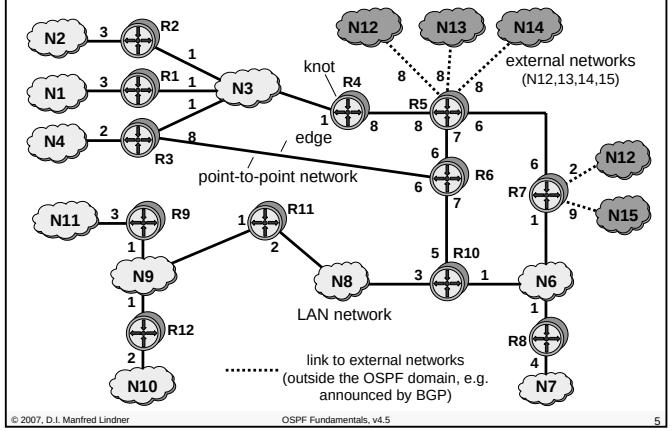
  this led to

  link-state protocols
- **OSPF is an important implementation for IP**
- **basic idea: every router knows topology of the whole network, including subnets and other routers**
- **see RFC 2328 (2178, 1583 are obsolete)**

OSPF Topology Database

- **every router maintains a topology database**
  - is like a "network roadmap"
  - describes the whole network !!
  - note: RIP provides only "signposts"
- **database is based on a graph**
  - where each knot stands for a router
  - where each edge stands for a subnet
    - connecting the routers
    - path-costs are assigned to the edges
  - where the actual router uses the graph as root
    - to calculate shortest paths to all subnets

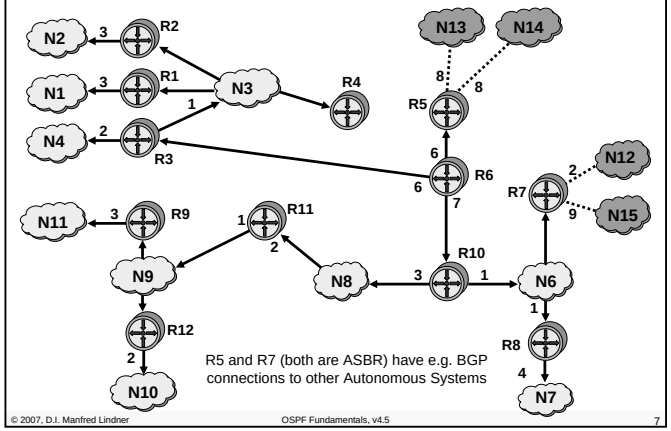
OSPF Domain



OSPF Principles

- **with this topology-database a router can calculate the best path to the desired destination-network**
  - applying Dijkstra's SPF (Shortest Path First) algorithms
- **the topology-database describes all other possible paths too**
  - in critical situations (failures) the router can independently calculate an alternative path
- **no waiting on rumors of other routers anymore**
  - which was reason for several RIP problems !!!
  - because other routers are also relying on rumors

Shortest Paths regarding Router R6



Routing Table

- **after calculating the shortest path, this path is entered in the routing table**
- **OSPF is able to differentiate between internal and external net-IDs**
  - Autonomous System Boundary Router ASBR imports external net-IDs
  - paths to ASBRs are considered as special internal destinations
  - path-costs to external networks can be calculated (E1 metric) by adding:
    - internal costs to ASBR
    - and external costs (from ASBR to the external networks)

Routing Table Router 6

NET-ID	NEXT HOP	DISTANCE
N1	R3	10
N2	R3	10
N3	R3	7
N4	R3	8
N6	R10	8
N7	R10	12
N8	R10	10
N9	R10	11
N10	R10	13
N11	R10	14
R5	R5	6
R7	R10	8

internal destinations of router 6;
R5 and R7 are ASBRs

Routing Table Router 6

NET-ID	NEXT HOP	DISTANCE
N12	R10	10
N13	R5	14
N14	R5	14
N15	R10	17

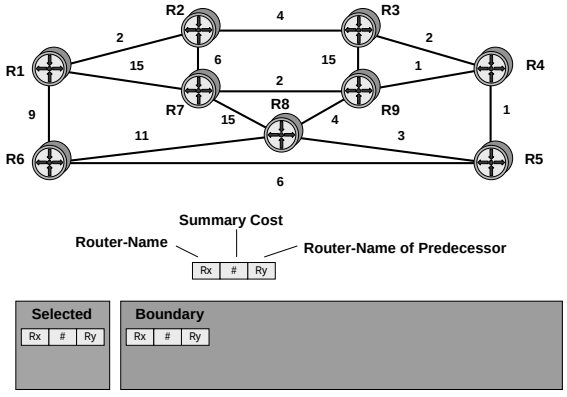
external destinations of router 6

assumption: ASBR R5 and R7 announce external networks (which have been noticed using e.g. BGP-4) via type 1 external link advertisements (E1 metric)

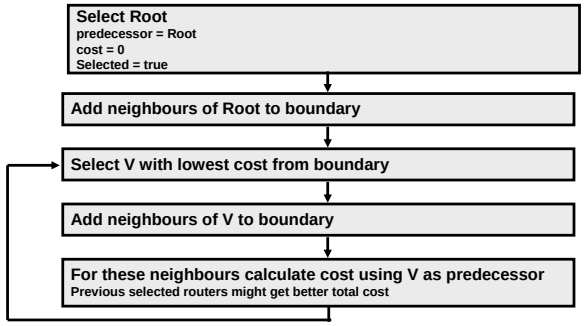
Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

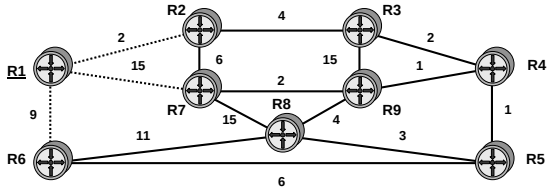
Example Topology for Dijkstra Algorithm



The Principle



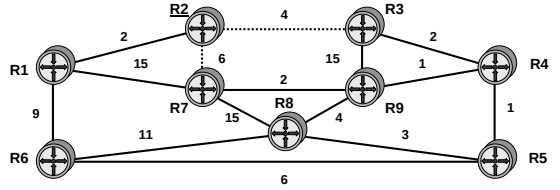
Select root (R1)



Selected		
R1	0	R1

Boundary								
R2	2	R1	R6	9	R1	R7	15	R1

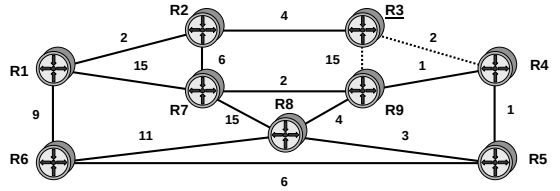
Select router with lowest cost in boundary (R2), calculate cost for neighbours R3, R7



Selected		
R1	0	R1
R2	2	R1

Boundary								
R2	2	R1	R6	9	R1	R7	15	R1
R6	9	R1	R7	8	R2	R3	6	R2

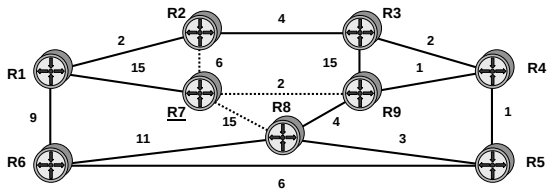
Select router with lowest cost in boundary (R3), calculate cost for neighbours R9, R4



Selected		
R1	0	R1
R2	2	R1
R3	6	R2

Boundary												
R2	2	R1	R6	9	R1	R7	15	R1				
R6	9	R1	R7	8	R2	R3	6	R2				
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3	

Select one router with lowest cost in boundary (R7), calculate cost for neighbours R8, R9



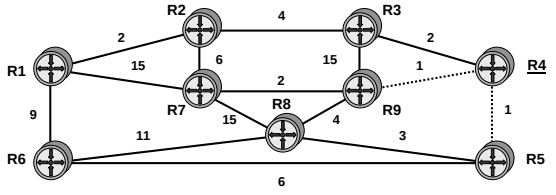
Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2

Boundary

R2	2	R1	R6	9	R1	R7	15	R1						
R6	9	R1	R7	8	R2	R3	6	R2						
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3			
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7			

Select router with lowest cost in boundary (R4), calculate cost for neighbours R9, R5



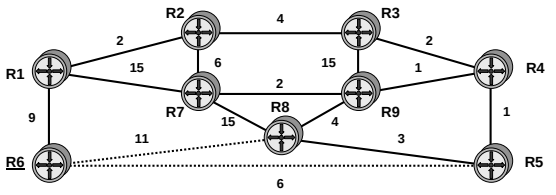
Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2
R4	8	R3

Boundary

R2	2	R1	R6	9	R1	R7	15	R1						
R6	9	R1	R7	8	R2	R3	6	R2						
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3			
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7			
R6	9	R1	R8	23	R7	R9	9	R4	R5	9	R4			

Select one router with lowest cost in boundary (R6), calculate cost for neighbours R5 and R8



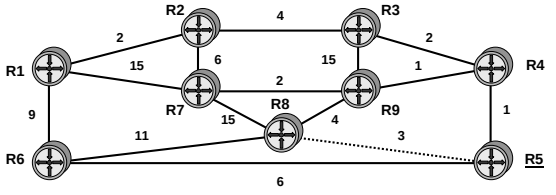
Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2
R4	8	R3
R6	9	R1

Boundary

R2	2	R1	R6	9	R1	R7	15	R1						
R6	9	R1	R7	8	R2	R3	6	R2						
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3			
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7			
R6	9	R1	R8	23	R7	R9	9	R4	R5	9	R4			
R9	9	R4	R8	20	R6	R5	9	R4						

Select one neighbour with lowest cost in boundary (R5), calculate cost for neighbour R8



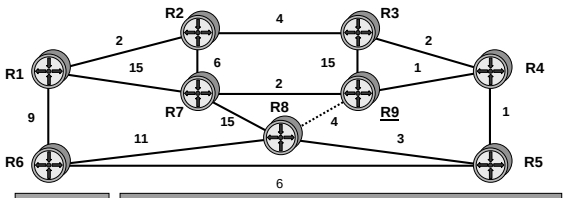
Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2
R4	8	R3
R6	9	R1
R5	9	R4

Boundary

R2	2	R1	R6	9	R1	R7	15	R1						
R6	9	R1	R7	8	R2	R3	6	R2						
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3			
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7			
R6	9	R1	R8	23	R7	R9	9	R4	R5	9	R4			
R9	9	R4	R8	20	R6	R5	9	R4						
R9	9	R4	R8	12	R5									

Select router with lowest cost in boundary (R9), calculate cost for neighbours R8



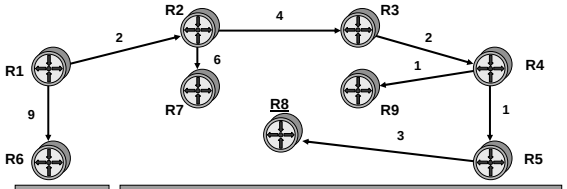
Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2
R4	8	R3
R6	9	R1
R5	9	R4
R9	9	R4

Boundary

R2	2	R1	R6	9	R1	R7	15	R1									
R6	9	R1	R7	8	R2	R3	6	R2									
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3						
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7						
R6	9	R1	R8	23	R7	R9	9	R4	R5	9	R4						
R9	9	R4	R8	20	R6	R5	9	R4									
R9	9	R4	R8	12	R5												
R8	12	R5															

Select last router in boundary (R8), algorithm terminated, all shortest paths found



Selected

R1	0	R1
R2	2	R1
R3	6	R2
R7	8	R2
R4	8	R3
R6	9	R1
R5	9	R4
R9	9	R4
R8	12	R5

Boundary

R2	2	R1	R6	9	R1	R7	15	R1									
R6	9	R1	R7	8	R2	R3	6	R2									
R6	9	R1	R7	8	R2	R9	21	R3	R4	8	R3						
R6	9	R1	R4	8	R3	R9	10	R7	R8	23	R7						
R6	9	R1	R8	23	R7	R9	9	R4	R5	9	R4						
R9	9	R4	R8	20	R6	R5	9	R4									
R9	9	R4	R8	12	R5												
R8	12	R5															

Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

Creating the Database

- **until now an a priori existing consistent database in every router has been assumed**
- **in fact, the basic means for creating and maintaining the database are the so-called**

  link states
- **a link state stands for a local neighbourhood between two routers**
  - the link state is created by these two routers
  - other routers are notified about this link state via a broadcast-mechanism ("traffic-news")
  - link states are verified continuously

How are Link States used?

- **adjacent routers declare themselves as neighbours by setting the link state up (or down otherwise)**
  - the link-state can be checked with hello messages
- **every link state change is published to all routers of the OSPF domain using Link State Advertisements (LSAs)**
  - is a broadcast mechanism
  - LSAs are much shorter than routing tables
    - because LSAs contain only the actual changes
    - that's why distance vector protocols are much slower
  - whole topology map relies on LSAs

OSPF Communication Principle 1

- **OSPF messages are transported by IP**
  - ip protocol number 89
- **during initialization a router sends hello-messages to all directly reachable routers**
  - to determine its neighbourhood
  - can be done automatically in broadcast networks and point-to-point connections by using the IP multicast-address 224.0.0.5 (all OSPF routers)
  - non-broadcast networks: configuration of the neighbourhood-routers is required (e.g. X25)
- **this router also receives hello-messages from other routers**

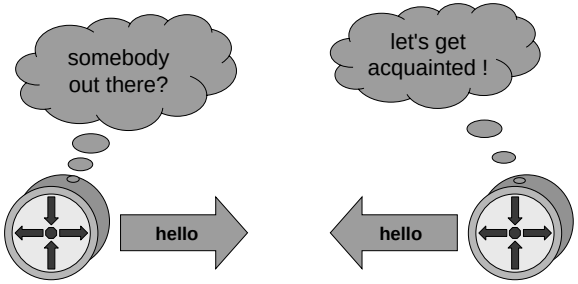
OSPF Communication Principle 2

- **each two acquainted routers send database description messages to each other, in order to publish their topology database**
- **unknown or old entries are updated via link state request and link state update messages**
  - which synchronizes the topology databases
- **after successful synchronization both routers declare their neighbourhood (adjacency) via router LSAs (using link state update messages)**
  - distributed across the whole network

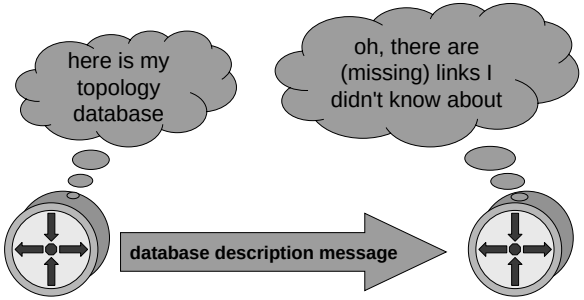
OSPF Communication Principle 3

- **periodically, every router verifies its link state to its adjacent neighbours using hello messages**
- **from now only changes of link states are distributed**
  - using link state update messages (LSA broadcast-mechanism)
- **if neighbourhood situation remains unchanged, the periodic hello messages represents the only routing overhead**
  - note: additionally all Link States are refreshed every 30 minutes with LSA broadcast mechansim

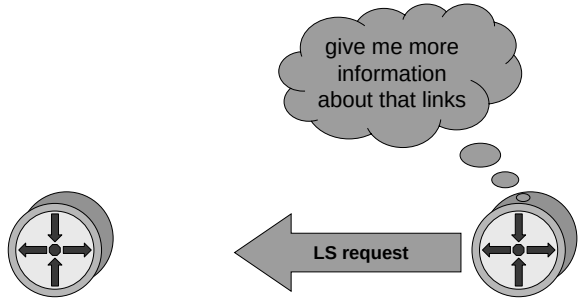
OSPF Communications Summary 1



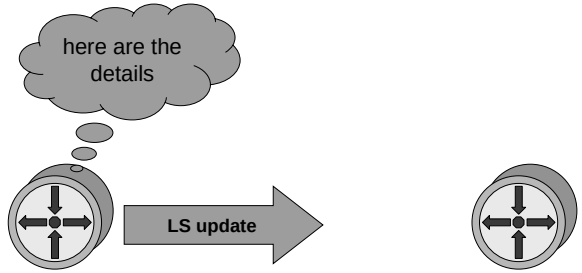
OSPF Communications Summary 2



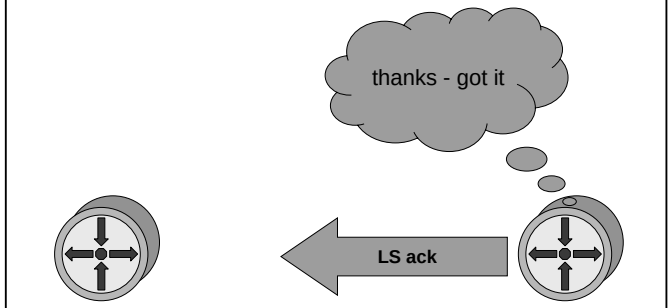
OSPF Communications Summary 3



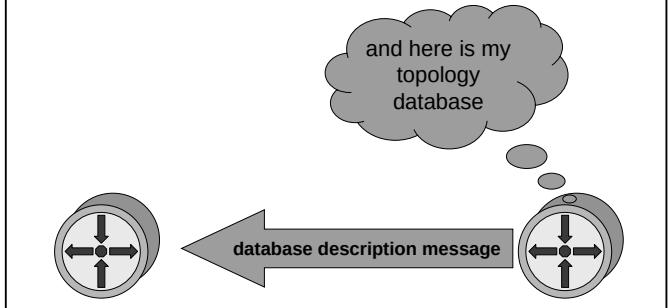
OSPF Communications Summary 4



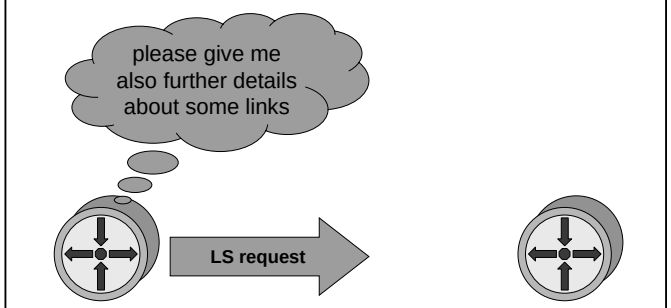
OSPF Communications Summary 5



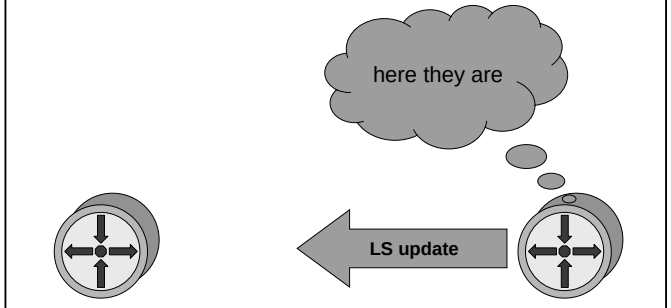
OSPF Communications Summary 6



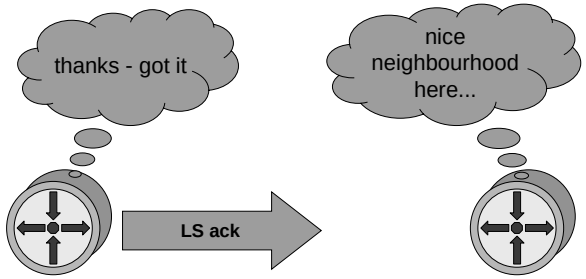
OSPF Communications Summary 7



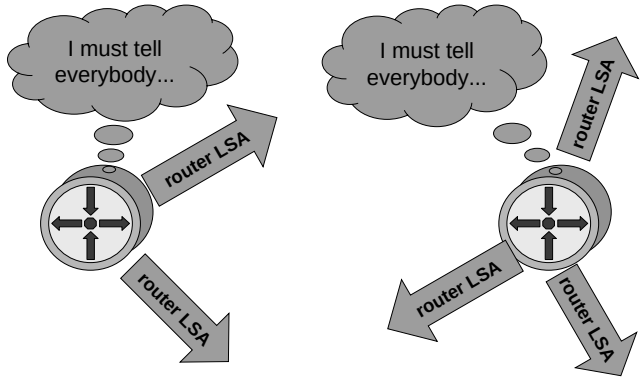
OSPF Communications Summary 8



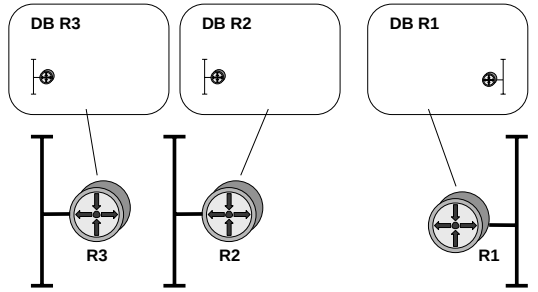
OSPF Communications Summary 9



OSPF Communications Summary 10

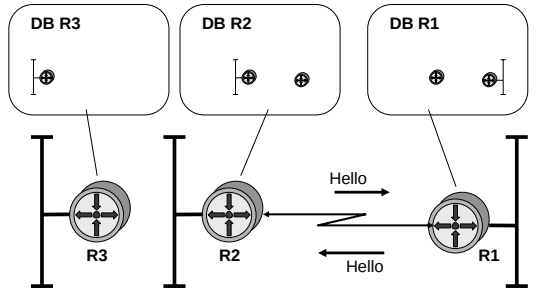


OSPF Start-up



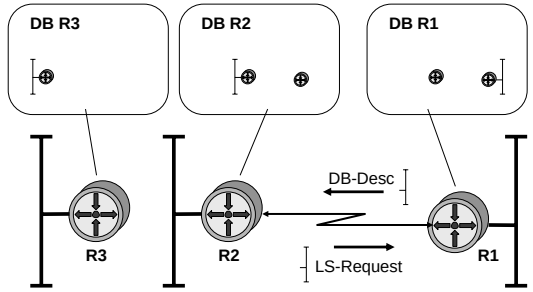
starting position: all routers initialized, no connection between R1-R2 or R2-R3

OSPF Hello R1 - R2



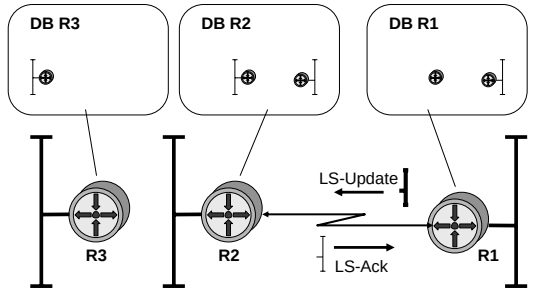
link between R1-R2 activated: get acquainted using hello messages

OSPF Data Base Description R1 -> R2



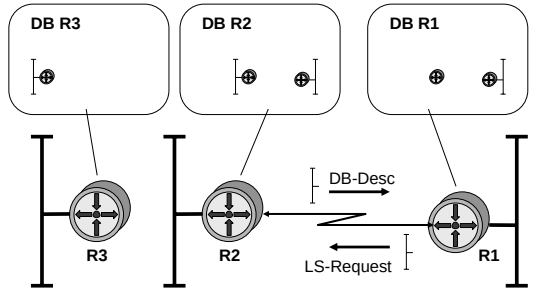
database synchronization: R1 master sends Database-Description, R2 slave sends Link-State Request

OSPF Data Base Update R1 -> R2



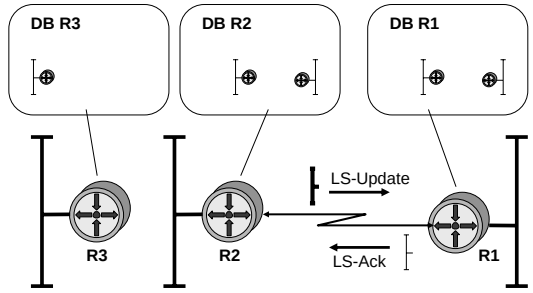
database synchronization: R1 master sends Link-State Update, R2 slave sends Link-State Acknowledgement

OSPF Data Base Description R2 -> R1



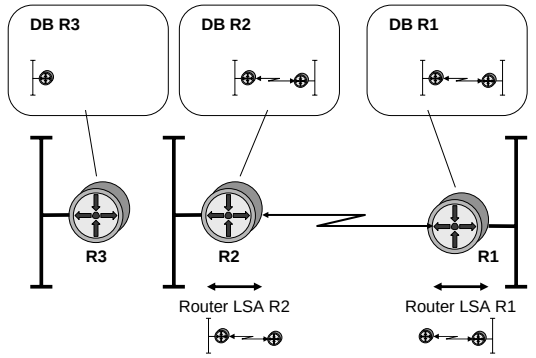
database synchronization: R2 master sends Database-Description, R1 slave sends Link-State Request

OSPF Data Base Update R2 -> R1



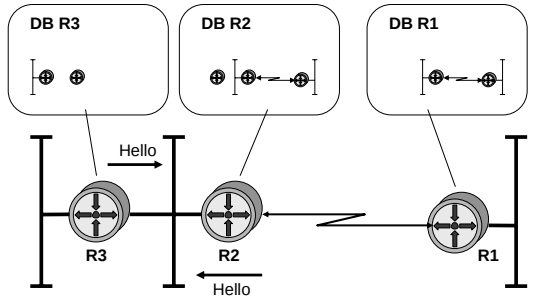
database synchronization: R2 master sends Link-State Update, R1 slave sends Link-State Acknowledgement

OSPF Router LSA Emission



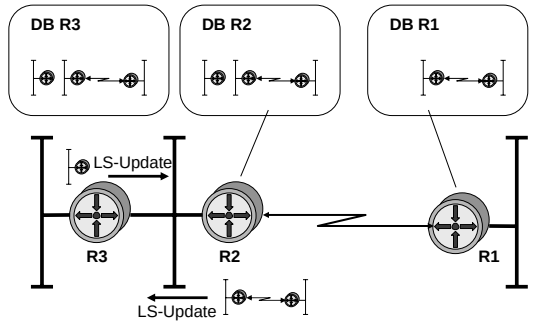
R1 and R2 have synchronized their database completely and notify other nodes about their links

OSPF Hello R2 - R3



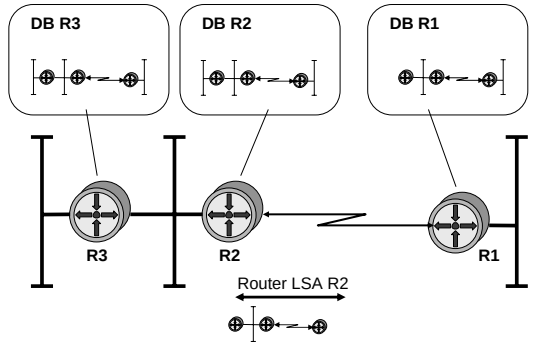
link between R2-R3 activated: get acquainted using Hello, determination of designated router

OSPF Database Update



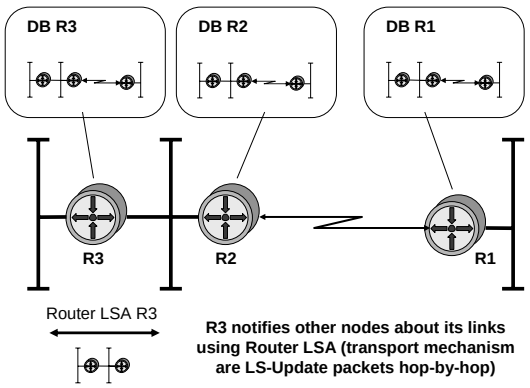
R2 and R3 synchronize their databases (DB-Des., LS-Req.,LS-Upd., LS-Ack.)

OSPF Router LSA Emission R2



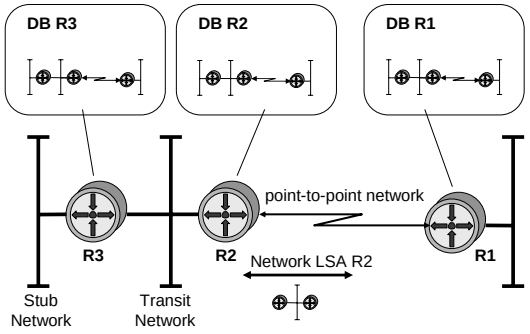
R2 notifies other nodes about its links using Router LSA, (transport mechanism are LS-Update packets hop-by-hop)

OSPF Router LSA Emission R3



R3 notifies other nodes about its links using Router LSA (transport mechanism are LS-Update packets hop-by-hop)

OSPF Network LSA R2



Designated Router R2 notifies other nodes about the multi-access network using Network-LSA (transport mechanism are LS-Update packets hop-by-hop

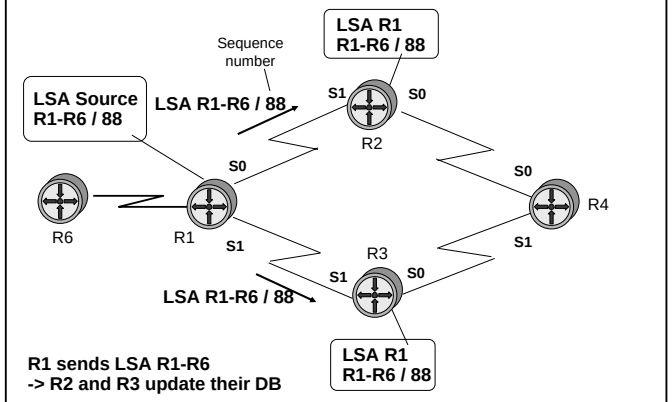
Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

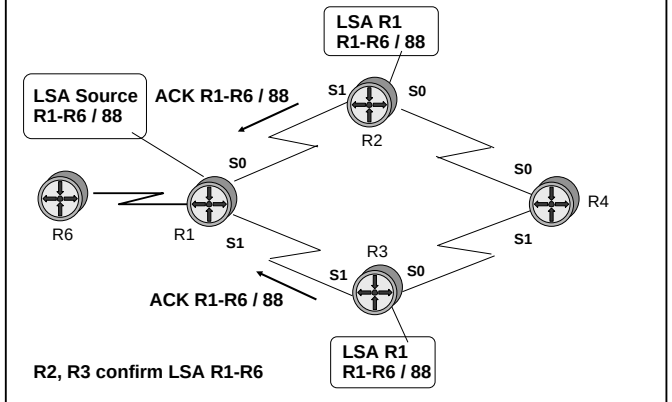
LSA Broadcast Mechanism

- **LSA must be distributed safely to all routers within an area (domain)**
  - consistency of the topology-database depends on it
  - every LS-update is acknowledged explicitly (using LS-ack) by the neighbour-router
  - if a LS-ack stays out, the LS-update is repeated (timeout)
  - if the LS-ack fails after several trials, the adjacency-relation is cleared
  - this method assures a reliable distribution of LSAs

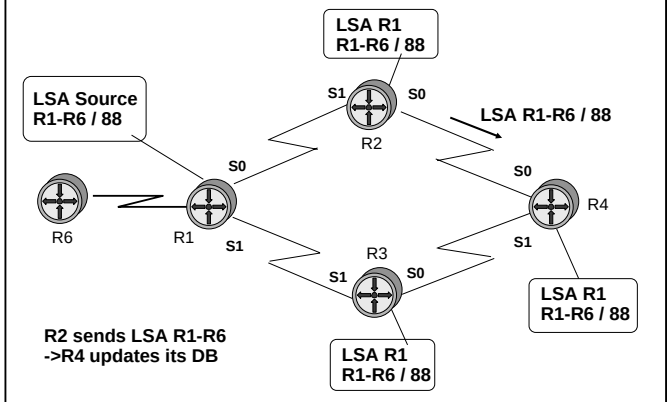
LSA Broad 1



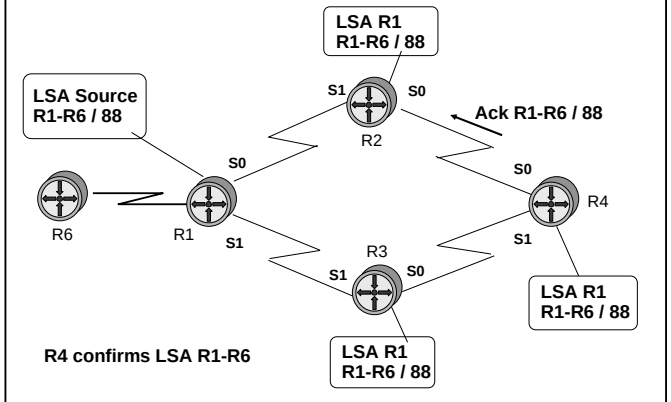
LSA Broad 2



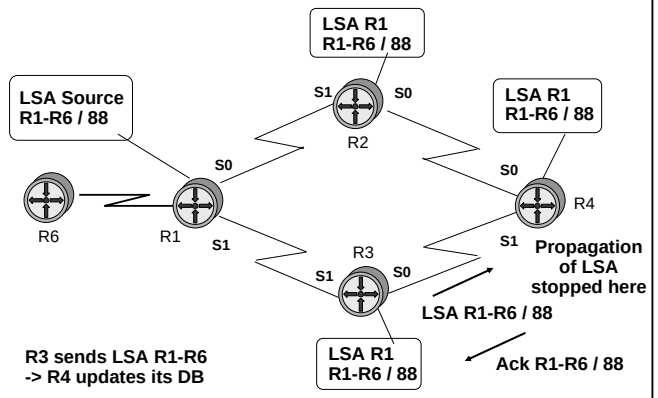
LSA Broad 3



LSA Broad 4



LSA Broad 5



LSA Usage

- **additionally, link states are repeated every 30 minutes to refresh the databases**
  - link states become obsolete after 60 minutes and are removed from the databases
- **reasons:**
  - automatic correction of unnoticed topology-mistakes (e.g. happened during distribution or some router internal failures in the memory)
  - combining two separated parts of an OSPF area (here OSPF also assures database consistency without intervention of an administrator)

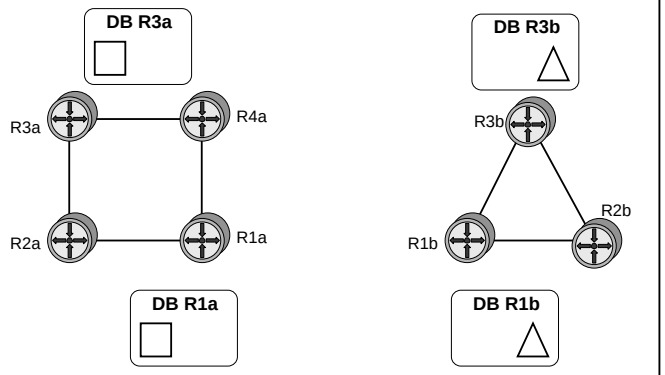
How are LSA unique?

- **each router as a node in the graph (link state topology database)**
  - is identified by a unique Router-ID
  - note: automatically selected on Cisco routers
    - either numerically highest IP address of all loopback interfaces
    - or if no loopback interfaces then highest IP address of physical interfaces
- **every link and hence LS between two routers**
  - can be identified by the combination of the corresponding Router-IDs
  - note:
    - if there are several parallel physical links between two routers the Port-ID will act as tie-breaker

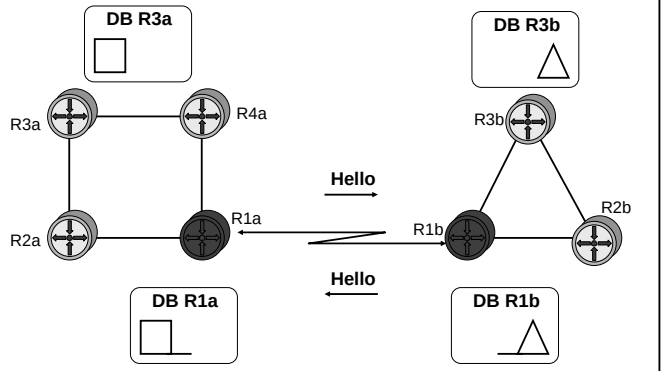
Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

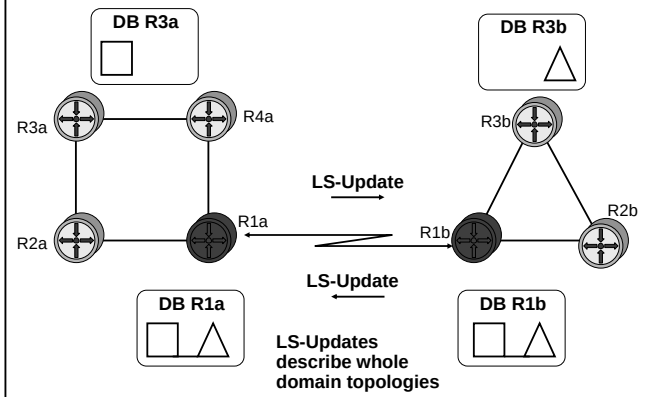
Initially: Two Separated Parts



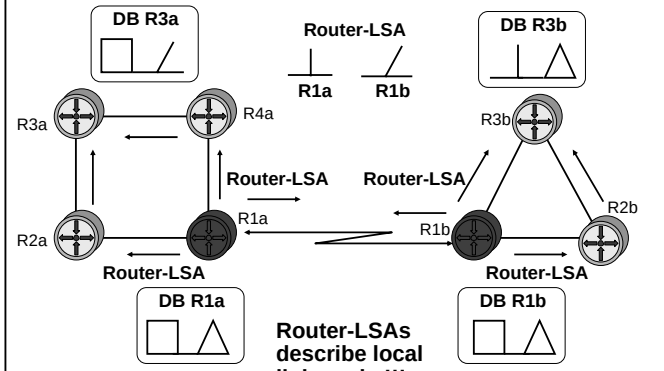
Hello R1a <-> R1b



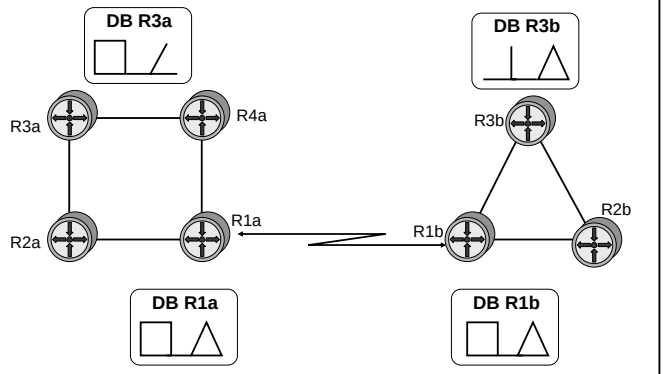
Database Synchronization R1a <-> R1b



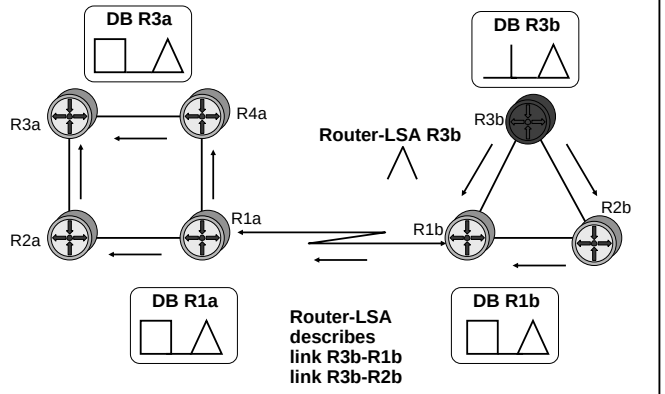
Router LSA Broadcast



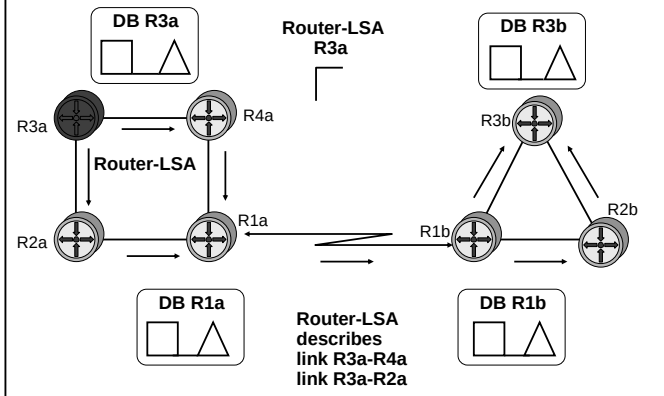
Inconsistent Databases !



R3b: Router LSA Broadcast (Refresh)



R3a: Router LSA Broadcast (Refresh)



Improvement

- **until now it seems**
  - that combining a splitted area of a OSPF domain will take a lot of time in the worst case (30 min refresh)
- **the situation can improved in such a case**
  - by triggering database synchronization between any two routers in the network
  - so whenever a router is informed by a Router-LSA about some changes in the network this router additionally will do a database synchronization with the router from which the Router-LSA was received
  - database description packets will help to reduce traffic to the necessary minimum
    - design rule: avoid splitting of an area in a OSPF environment by avoiding single point of failures

Agenda

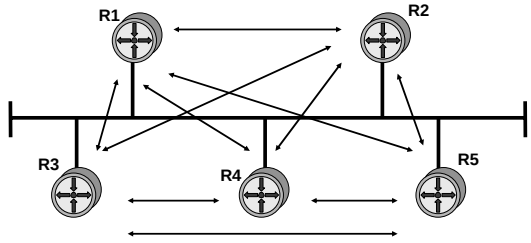
- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

OSPF Broadcast Networks

- **basic concept of link state**
  - point-to-point relationships
- **that fits best for**
  - point-to-point networks like serial lines
- **that causes a problem with shared media multi-access networks**
  - e.g. with LANs or with networks running in NBMA-mode (Non Broadcast Multi Access) like X.25, Frame Relay, ATM
  - hello, database description and LSA updates between each of these routers can cause huge network traffic and CPU load

Broadcast Network

Any-to-Any Relation



Router-Router Neighbourhood-Relation

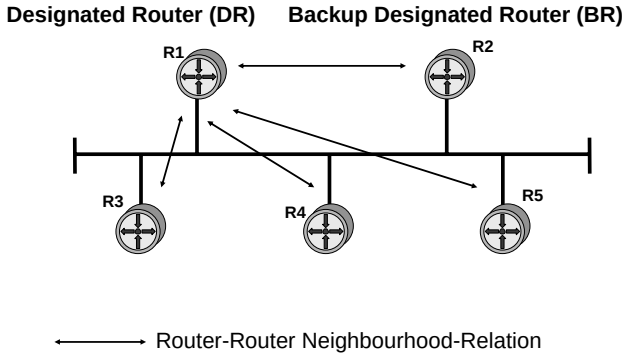
OSPF Broadcast Networks

- **if several routers share a multi-access network**
  - any-to-any doesn't scale well -> N*(N-1)/2 problem
  - information about all possible neighbourhood-relations seems to be redundant
  - concept of virtual (network) node (or virtual router) is introduced to solve the problem
- **only the virtual node needs**
  - to maintain N-1 point-to-point relationship to the other nodes
  - any-to-any is not necessary
- **in OSPF the virtual node is called**
  - Designated Router (DR)

OSPF Broadcast Networks

- **in case of a failure the Designated Router**
  - would be single point of failure
- **therefore**
  - a Backup Designated Router (BR) is used
- **DR and BR**
  - are elected by exchanging hello-messages at start-up
- **Attention !!!**
  - this concept influences only how routing information is exchanged among those routers
  - no influence on actual IP forwarding which is based on routing tables

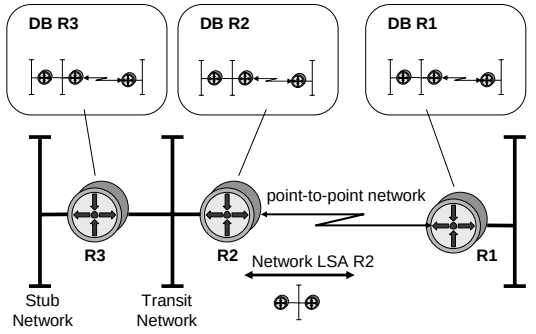
Designated and Backup Router



Designated Router and Network LSA

- **Designated Router (DR) is responsible**
  - for maintaining neighbourhood relationship via virtual point-to-point links using the already known mechanism
    - DB-Description, LS-Request LS-Update, LS-Acknowledgement, Hello, etc.
- **Router-LSA´s implicitly describe**
  - these virtual point-to-point links by specifying such a network as transit-network
  - remark: Stub-network is a LAN network where no OSPF router is behind
- **To inform all other routers of domain about such a situation**
  - DR is additionally responsible for emitting Network LSAs
- **a Network LSA describes**
  - which routers are members of the corresponding broadcast network

OSPF Network LSA R2



Designated Router R2 notifies other nodes about the multi-access network using Network-LSA (transport mechanism are LS-Update packets hop-by-hop

DR and Backup Router

- **Designated and Backup router are determined using the router-priority field of the Hello message**
- **on DR failure, a Backup Router (BR) continues the service**
- **BR listens to the traffic on the virtual point-to-point link**
  - multicast addresses are used
- **BR recognizes a DR failure through missing acknowledge messages**
  - note: every LS-Update message demands for an LS-Acknowledgement message

OSPF Multicast Addresses

- **OSPF uses dedicated IP multicast addresses for exchanging routing messages**
  - 224.0.0.5 (“All OSPF Routers")
  - 224.0.0.6 (“All Designated Routers")
- **224.0.0.5 is used as destination address**
  - by all routers for Hello-messages
    - DR and BR determination at start-up
    - link state supervision
  - by DR router for messages towards all non-DR routers
    - LS-Update, LS-Acknowledgement
- **224.0.0.6 is used as destination address**
  - by all non-DR routers for messages towards the DR
    - LS-Update, LS-Request, LS-Acknowledgement and database description messages

Agenda

- **OSPF Principles**
  - Introduction
  - The Dijkstra Algorithm
  - Communication Procedures
  - LSA Broadcast Handling
  - Splitted Area
  - Broadcast Networks
  - Summary of Benefits

Distance-Vector **versus** Link-State

- Distance-Vector:
  - every router notifies directly connected routers about all reachable routes
  - using broadcast messages
  - maintains its routing table according to information from neighbor routers
- Link-State:
  - every router notifies all routers about the state of his directly connected links
  - using flooding mechanism (LSA)
  - calculates optimal paths whenever a new LSA is received

OSPF Benefits 1

- **network load is significantly smaller than that of distance vector protocols**
  - short hello messages between adjacent routers versus periodical emission of the whole routing table
- **even update messages after topology modifications are smaller than the routing table of distance vector protocols**
  - LSAs only describe the local links for which a router is responsible -> incremental updates !!!
- **massive network load**
  - occurs only on combining large splitted network parts of an OSPF domain (many database synchronizations)

OSPF Benefits 2

- **SPF-techniques take advantages from several features:**
  - every router maintains a complete topology-map of the entire network and calculates independently its desired paths (actually based on the original LSA message)
  - this local ability for route calculation grants a fast convergence
  - LSA is not modified by intermediate routers across the network
  - the size of LSAs depends on the number of direct links of a router to other routers and not on the number of subnets!

OSPF Benefits 3

- **during router configuration, every physical port is assigned a cost value**
  - depends on ToS (Type of Service)
  - each ToS can be assigned a separate topology map (8 possible combinations)
  - IP's ToS field is examined for packet forwarding
  - note: OSPF ToS support disappeared in RFC 2328
- **determination of the best path with a specific ToS is based on the costs of the particular segments (RIP uses hop count only)**
- **equal costs automatically enables load balancing between these paths**

OSPF Benefits 4

- **subnet masks of variable length can be attached to routes (in contrast to RIPv1)**
- **external routes are marked (tagged) explicitly to be differentiated from internal routes**
- **OSPF messages can be authenticated to grant secure update information**
- **OSPF routing messages use IP-multicast addresses: lower processing effort**
- **point-to-point connections do not need own IP-address: economic use of address space**

OSPF in Large Networks

- **OSPF area concept can be used**
  - a two level hierarchy is used to decrease
    - CPU time for SPF calculations
    - memory requirement for storing topology database
  - one backbone area
  - several non-backbone areas
    - non-backbone area can be connected by area border router to backbone area only
  - summarization possible at area border routers
    - route aggregation to reduce size of routing tables
    - summarization means that some net-IDs can be summarized as one net-ID only
- **For a detailed description of OSPF areas see the corresponding advanced OPSF chapter**

