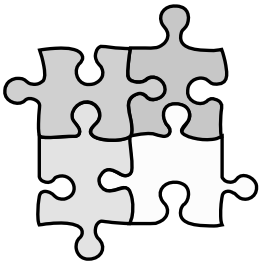


OSPF – Introduction

The IETF Routing Master
Part 1

“Dijkstra probably hates me”

Linus Torvalds in kernel/sched.c

"Open Shortest Path First"

- **Official (IETF) successor of RIP**
  - **RIP is slow**
  - **RIP is unreliable**
  - **RIP produces too much routing traffic**
  - **RIP only allows 15 hop routes**

- **OSPF is a link-state routing protocol**
  - **Inherently fast convergence**
  - **Designed for large networks**
  - **Designed to be reliable**

OSPF was developed by IETF to replace RIP. In general link-state routing protocols have some advantages over distance vector, like faster convergence, support for lager networks.

Some other features of OSPF include the usage of areas, which makes possible a hierarchical network topologies classless behavior,there are no such a problem like in RIP with discontiguous subnets. OSPF also supports VLSM and authentication.

OSPF Background

- **OSPF is the IGP recommended by the IETF**
- **"Open" means "not proprietary"**
- **Dijkstra's Shortest Path First algorithm is used to find the best path**
- **OSPF's father: John Moy**
  - **Version 1: RFC 1131**
  - **Version 2: RFC 2328 (244 pages !!!)**
  - **And a lot of additional OSPF related RFCs available...**

The Internet Engineering Task Force (IETF) strictly recommends to use OSPF for Interior Gateway routing (i. e. within an AS) instead of RIP or other protocols. Integrated IS-IS is an alternative routing protocol but not explicitly recommended by the IETF. Note that IS-IS has been standardized by the ISO world.

Both (Integrated) IS-IS and OSPF use Dijkstra's famous Shortest Path First (SPF) algorithm to determine all best paths for a given topology.

OSPF version 2 has been specified in RFC 2328. Note that there are a lots of additional RFCs around OSPF. Use http://www.rfc-editor.org/rfcsearch.html to find them all.

Dijkstra's SPF Algorithm

- **Used in graph theory**
- **Very efficient**
- **Calculates all paths to all destinations at once**
- **Creates a (loop-free) tree with local router as source**
- **See SPF section for more details**

Edsger W. Dijkstra (1930-2002)

The Dijkstra's SPF algorithm is generally used in graph theory and was not invented especially for IP routing. The most interesting point on the SPF algorithm is its efficiency. SPF is capable to calculate all paths to all destinations at once. The result of the SPF algorithm is a loop-less tree with the local router as source.

OSPF Ideas

- **Metric: "Cost" = 10^8/BW (in bit/s)**
  - **Therefore easily configurable per interface**
- **OSPF Routers exchange real topology information**
  - **Stored in dedicated topology databases**
- **Now Routers have a "roadmap"**
  - **Instead of signposts (RIP)**
- **Incremental updates**
  - **NO updates when there is NO topology change**

In the Cisco IOS implementation starting with 11.2, the cost is calculated automatically by the simple formula 10,000,000/BW.

Here the bandwidth parameters on a routers interface are used, thus it is especially important to configure it on the serial interfaces.

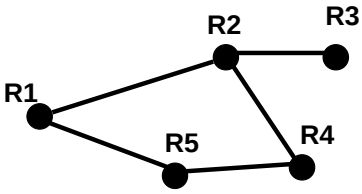
In other OSPF implementations cost must be configured manually for each of the interfaces.

OSPF—and other link state protocols—exchange true topology information which is stored in a dedicated database by each router. This database acts like a "roadmap" and allows a router to determine all best routes.

Note that once OSPF got the toplogy database there is no need to exchange further routing traffic—unless the topology changes. In this case only incremental updates are made.

What is Topology Information?

- **The smallest topological unit is simply the information element ROUTER-LINK-ROUTER**
- **So the question is: Which router is linked to which other routers?**



=

Link Database:

R1– R2
R1– R5
R2– R3
R2– R4
R4– R5

The Link Database exactly describes the roadmap

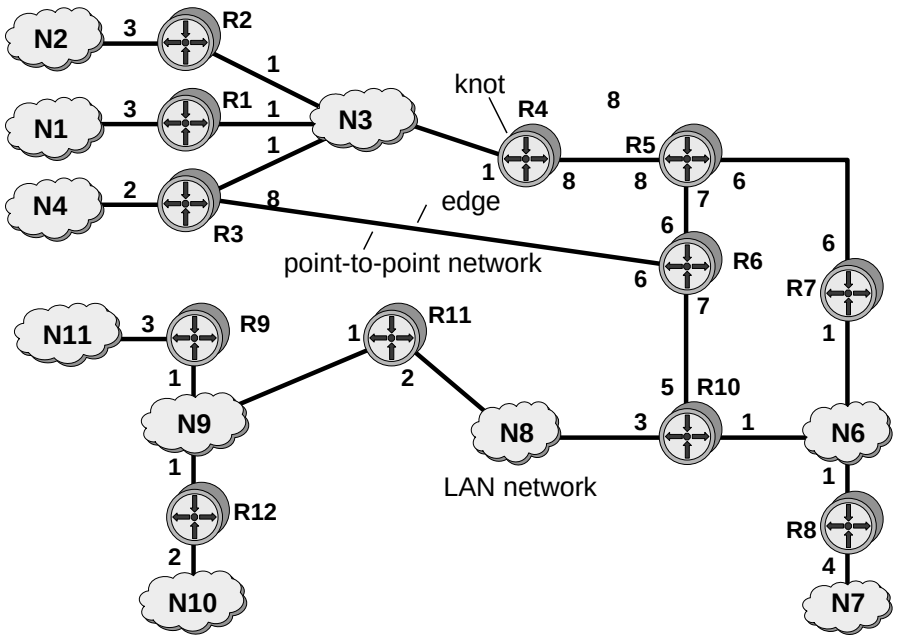
Obviously the dots are routers and the links between the routers are actually networks. The basic idea of OSPF and the topology table is that simple.

OSPF is actually much more complicated. There are 5 types of networks defined in OSPF: point-to-point networks, broadcast networks, non-broadcast multi-access networks, point-to-multipoint networks, and virtual links. Furthermore it is reasonable to divide the topology into multiple "areas" to increase performance ("divide and conquer"). These are the reasons why OSPF is a rather complex protocol. This is explained later.

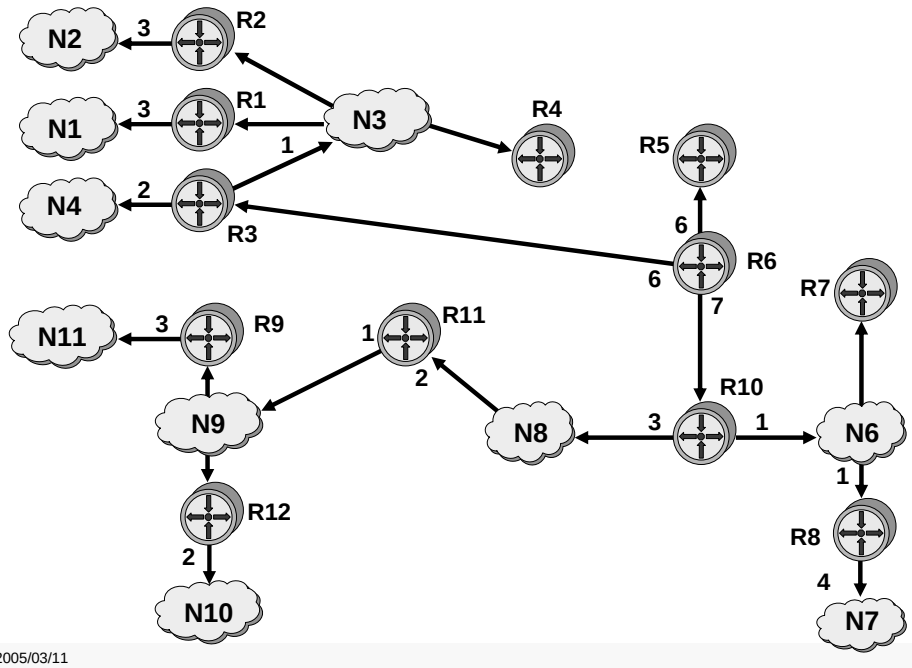
OSPF Topology Database

- **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


N2
R2
N1
R1
N3
knot
R4
R5
N4
R3
edge
point-to-point network
R6
R7
N11
R9
R11
N9
N8
R10
N6
LAN network
R12
R8
N10
N7

Shortest Paths regarding Router R6



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 Routing Updates

- The routing updates are actually link state updates
  - Parts of link state database are exchanged
  - Instead of parts of routing table (RIP)
- Applying the SPF algorithm on the link state database, each router can create routing table entries by its own

The Links State Updates LSUs are sent in a special packets – Link State Advertisments LSAs. There are several types of LSAs, depending on what kind of information is sent and which router originated it.

OSPF Protocol

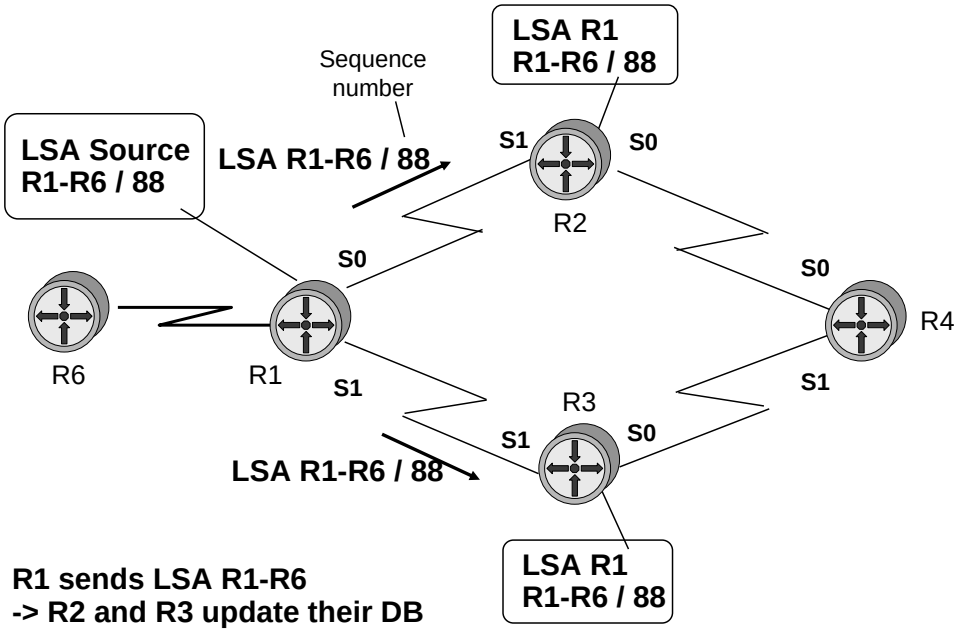
- All OSPF messages are carried within the IP payload ("raw IP")
  - Protocol number 89
- Error recovery and session management is covered by OSPF itself
- Multicast address 224.0.0.5
  - "All OSPF routers"

LSUs are encapsulated in IP packet directly, unlike RIP where we have an additional UDP overhead. IP is not reliable by itself, but OSPF updates are transmited reliable using Link State Acknolegements LSAck. There are 2 multicast addresses which are reserved for OSPF, 224.0.0.5 – for all OSPF routers and 224.0.0.6 for designated and back designated OSPF routers.

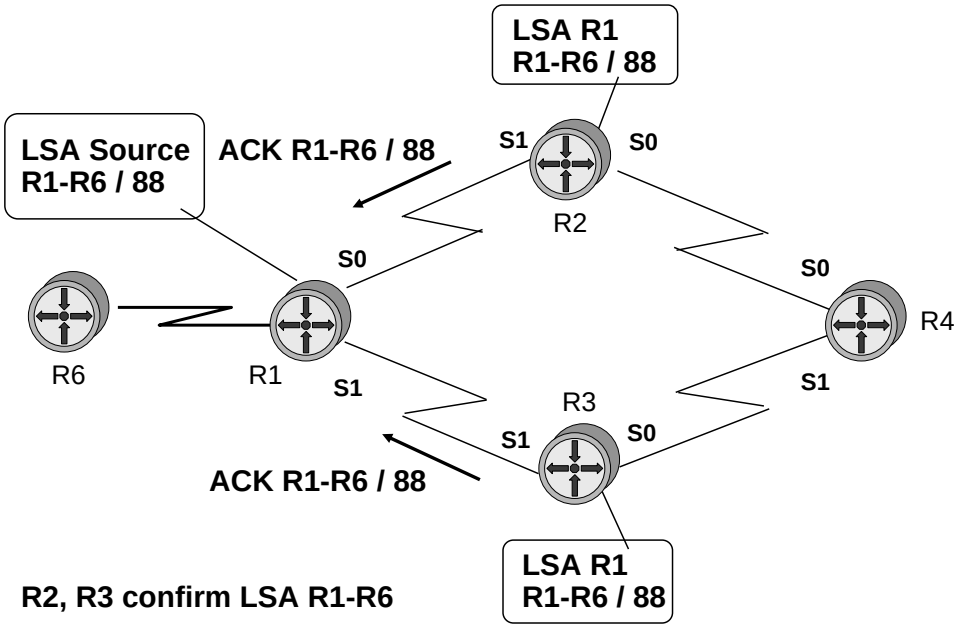
LSA Broadcast Mechanism

- **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

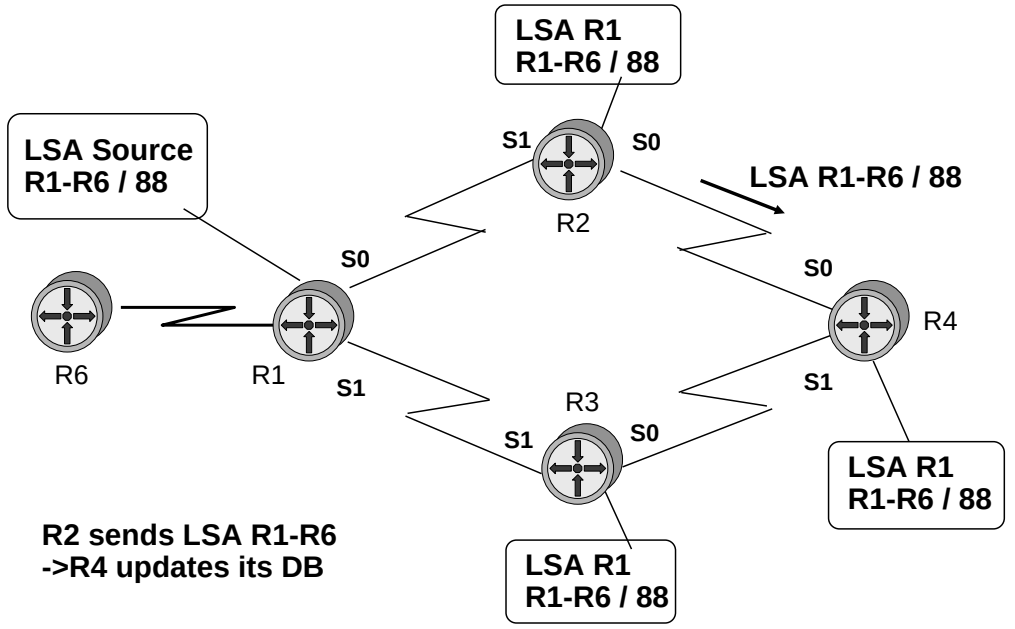
LSA Broad 1


LSA R1
R1-R6 / 88
Sequence
number
LSA Source
R1-R6 / 88
LSA R1-R6 / 88
S1
S0
R2
S0
S0
R6
R1
R4
S1
S1
R3
S1
S0
LSA R1-R6 / 88
LSA R1
R1-R6 / 88
R1 sends LSA R1-R6
-> R2 and R3 update their DB

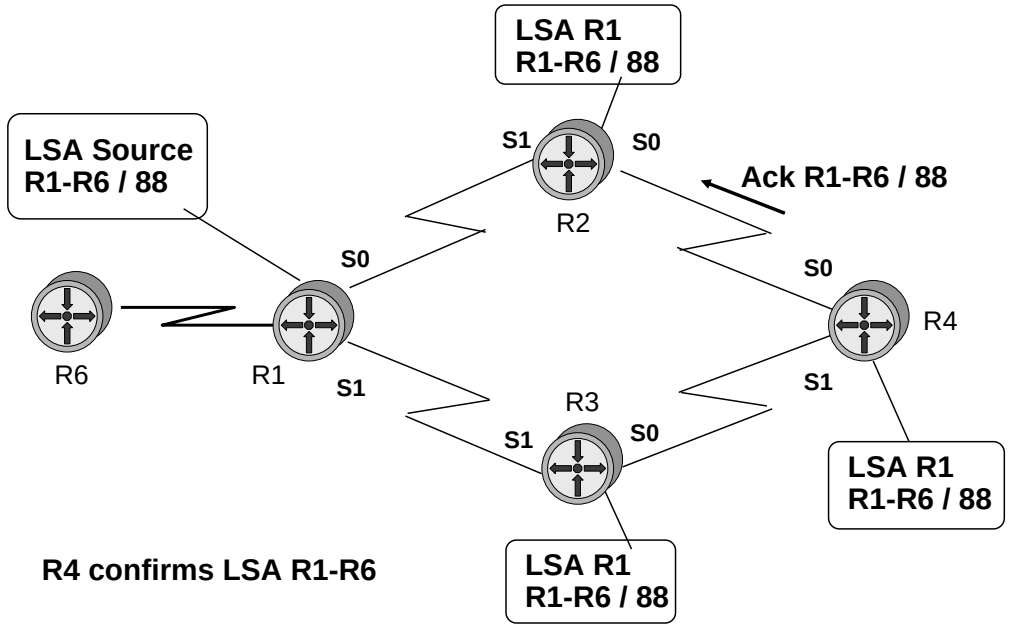
LSA Broad 2


LSA R1
R1-R6 / 88
LSA Source
R1-R6 / 88
ACK R1-R6 / 88
S1
S0
R2
S0
S0
R6
R1
R4
S1
S1
R3
S1
S0
ACK R1-R6 / 88
LSA R1
R1-R6 / 88
R2, R3 confirm LSA R1-R6

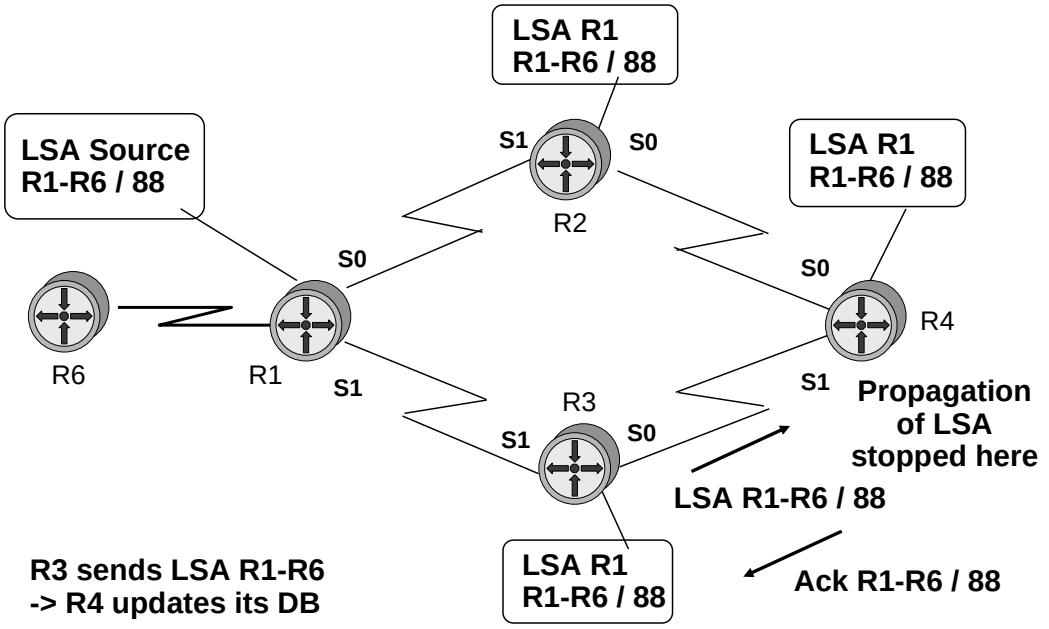
LSA Broad 3


LSA R1
R1-R6 / 88
LSA Source
R1-R6 / 88
S1
S0
LSA R1-R6 / 88
R2
S0
S0
R4
R6
R1
S1
S1
R3
S1
S0
LSA R1
R1-R6 / 88
R2 sends LSA R1-R6
->R4 updates its DB
LSA R1
R1-R6 / 88

LSA Broad 4


LSA R1
R1-R6 / 88
LSA Source
R1-R6 / 88
S1
S0
Ack R1-R6 / 88
R2
S0
S0
R4
R6
R1
S1
S1
R3
S1
S0
LSA R1
R1-R6 / 88
R4 confirms LSA R1-R6
LSA R1
R1-R6 / 88

LSA Broad 5


LSA R1
R1-R6 / 88
LSA Source
R1-R6 / 88
S1
S0
LSA R1
R1-R6 / 88
R2
S0
S0
R4
R6
R1
S1
S1
R3
S1
S0
Propagation
of LSA
stopped here
LSA R1-R6 / 88
Ack R1-R6 / 88
R3 sends LSA R1-R6
-> R4 updates its DB
LSA R1
R1-R6 / 88

LSA Flooding

- **LSA's are small packets, forwarded by each router without much modifications through the whole OSPF area (!)**
- **Much faster than RIP updates**
  - **RIP must receive, examine, create, and send**
- **Convergence time**
  - **Detection time + LSA flooding + 5 seconds before computing the topology table = "a few seconds"**

When the router gets a new information in its link state database it should send this information to all adjacent routers – flood. The packets are small, only the changes are sent and not the whole database. All other routers do the same, receive new information, update link state database, flood changes to others.

OSPF Overview

- **Large networks: "Divide and conquer" into areas**
  - **LSA-procedures inside each area**
  - **But *distance-vector updates between areas***
- **Additional complexity because of performance optimizations**
  - **Limit number of adjacencies in a multi-access network OSPF**
  - **Limit scope of flooding through "Areas"**
  - **Deal with stub areas efficiently**
  - **Learn external routes efficiently**
  - **Realized through different LSA types**
- **Fast convergence, almost no routing traffic in absence of topology changes**

Perfomance is very important with OSPF, to run SPF algorithm a CPU resources are required, to store a link state database an additional memory, compared to RIP we need much more router‘s resources. Some additional improvements were made to OSPF in order to improve performance. Areas were introduced to limit the flooding of LSAs, Stub Areas to minimize a link state database and routing tables.

Several types of LSAs were implemented:
Type 1 – Router LSA
Type 2 – Network LSA
Type 3 – Network Summary LSA
Type 4 – ASBR Summary LAS
Type 5 – AS External LSA
Type 6 – Group Membership LSA
Type 7 – NSSA External LSA
and others