

IT-Operation (2)

OSPF

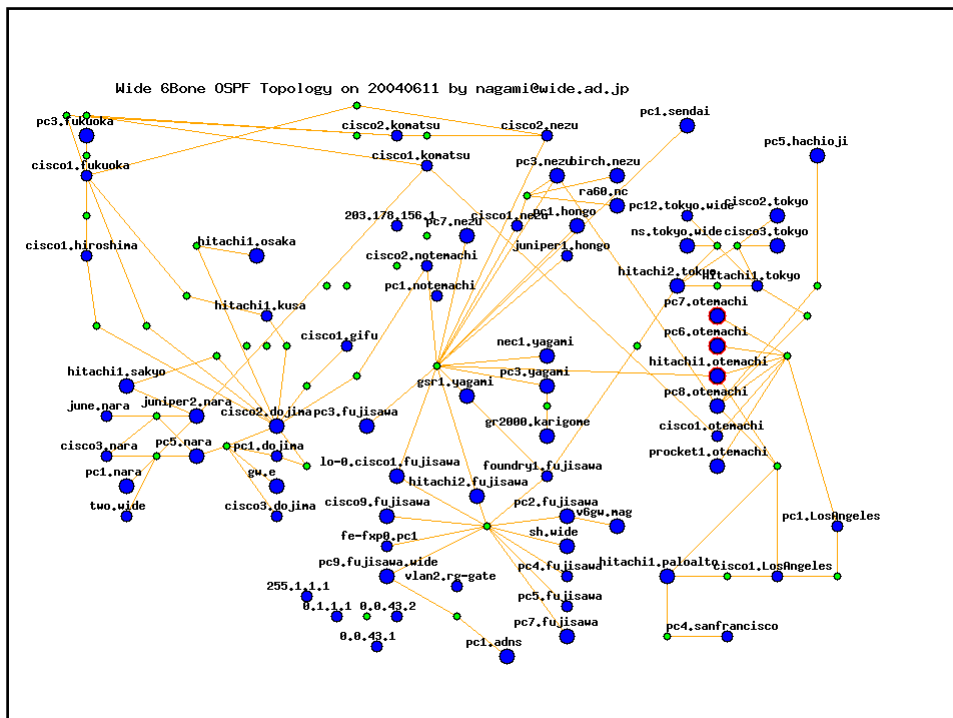
yasu@wide.ad.jp

Who am I

- Name: Yasuhiro Ohara
- 3rd grade of doctoral course in KEIO Univ.
- Author of Zebra ospf6d
- One of WIDE 6Bone operators
- developer or researcher
 - rather than operator

Zebra

- A famous routing software package
 - Free, open source software
 - Great ! the code is clean.
- Runs on PCs
 - Linux, FreeBSD, NetBSD, OpenBSD, ...
 - ripd, ripngd, ospfd, ospf6d, bgpd
 - Cisco-like user interface
- <http://www.zebra.org/>
- Various topics on Zebra ML archive

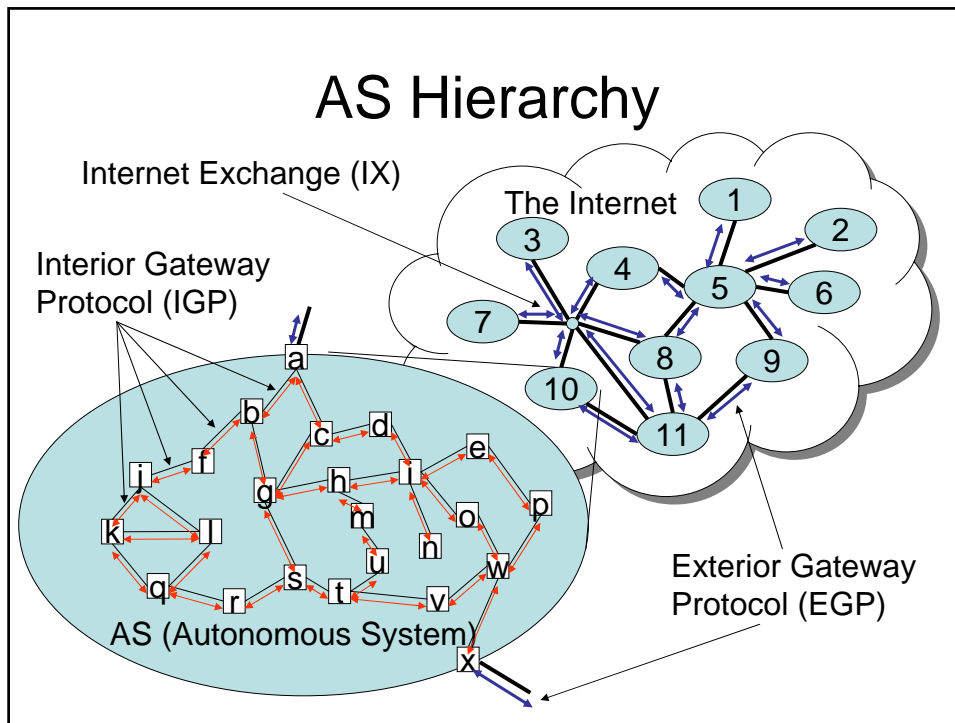


Objective of this lecture

- Understand Routing Concept/Principle
 - by introducing routing protocols briefly
- Know how OSPF works
 - so that you don't need to read specification
 - RFC2328 "OSPF" is approx. 240 pages !!
- Introduce some OSPF issues/problems

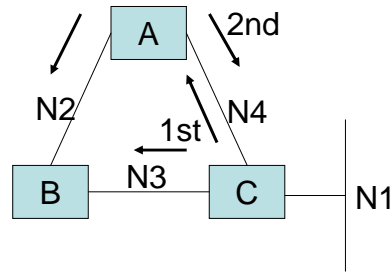
Routing Algorithms

- Distance Vector
 - RIP, EIGRP
- Link State
 - OSPF, IS-IS
- Path Vector (variant of DV)
 - BGP (only EGP)



- ## Why OSPF?
- RIP is dangerous
 - Counting to Infinity problem
 - Other restriction (e.g. 16 hops)
 - EIGRP is cisco proprietary protocol
 - other vendors router cannot co-exist
 - IS-IS may be a candidate
 - a few big U.S. ISPs use
 - simple (compared to OSPF)
 - OSPF
 - the most major (famous) IGP
 - complicated

RIP (Routing Information Protocol)



B's routing table

Dst	Nexthop	Metric
N2	-	0
N3	-	0
N4	C@N3	1
N1	C@N3	1

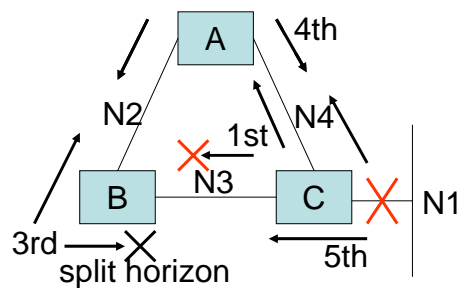
A's routing table

Dst	Nexthop	Metric
N2	-	0
N4	-	0
N1	C@N4	1
N3	C@N4	1

C's routing table

Dst	Nexthop	Metric
N4	-	0
N1	-	0
N3	-	0
N2	A@N4	1

DV's counting to infinity problem



B's routing table

Dst	Nexthop	Metric
N2	-	0
N3	-	0
N4	C@N3	1
N1	C@N3	4

A's routing table

Dst	Nexthop	Metric
N2	-	0
N4	-	0
N1	B@N2	2
N3	C@N4	1

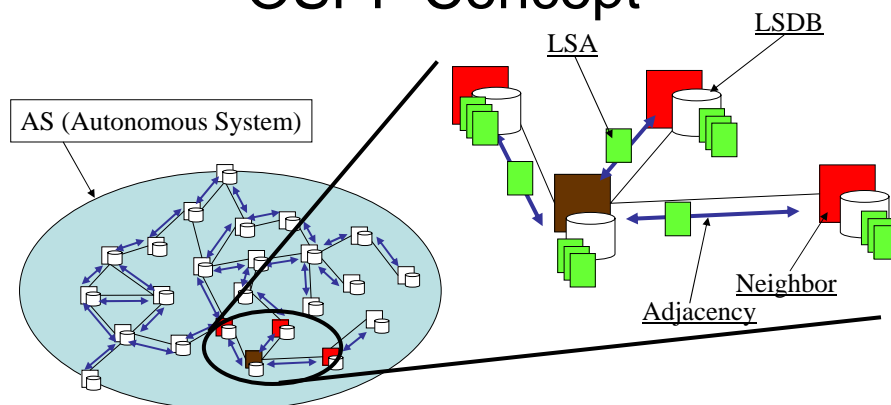
C's routing table

Dst	Nexthop	Metric
N4	-	0
N1	A@N4	3
N3	-	0
N2	A@N4	1

What is OSPF

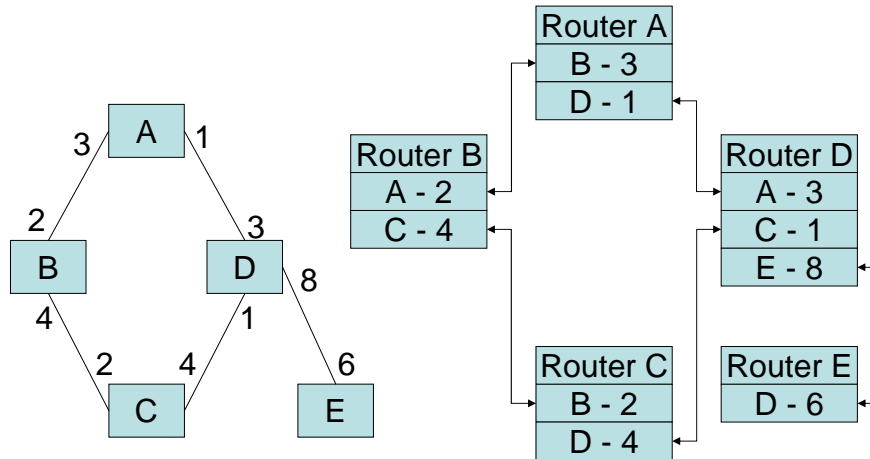
- Open Shortest Path First protocol
- the open routing protocol that employs SPF calculation (developed before 1991)
 - specification is OPEN
 - SPF calculation = Dijkstra algorithm
- Link state routing protocol
- Loop free
 - as long as LSDB is synchronized
 - as long as all routers employs the same calculation
- LSA (Link State Advertisement)
- LSDB (Link State DataBase)

OSPF Concept



Each router exchanges its LSAs with its neighbor router.
This leads to synchronization of all router's LSDB.
Then each router calculate routes to all destination independently (using Dijkstra algorithm).

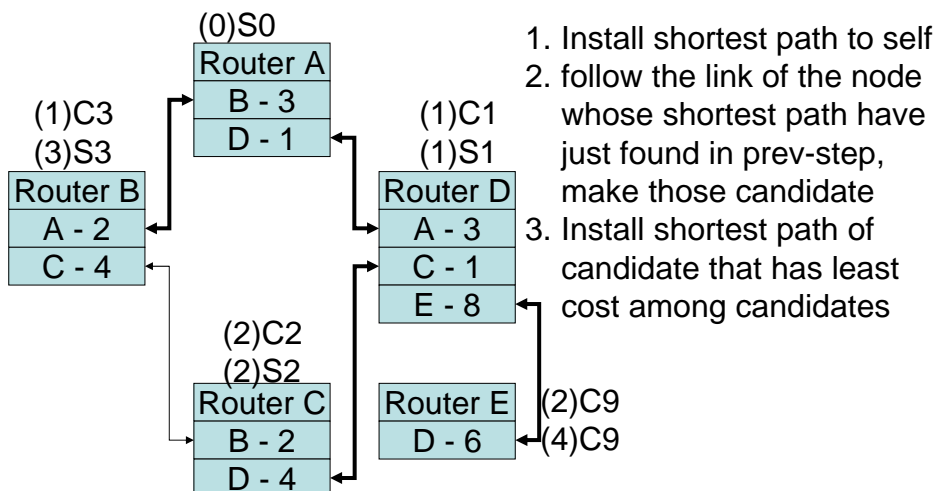
topology representation



Dijkstra algorithm

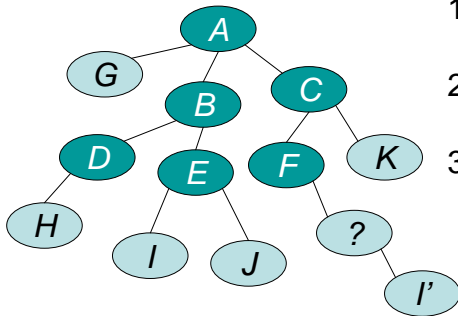
(N)Sc = Shortest path found as cost c at Nth step

(N)Cc = found as Candidate as cost c at Nth step



Dijkstra algorithm

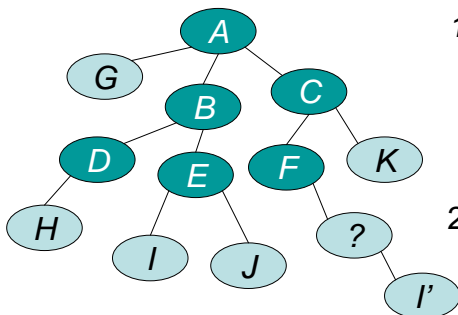
- the paths to the candidate vertex that is closest to the root are guaranteed to be shortest
- OSPF cost is defined to be a positive integer



1. Let $C(B-E)$ be the cost from B to E
2. Suppose $C(I)$ is the least cost among candidates
3. Can any I' that satisfies $C(I) > C(I')$ exist ?
 - $C(I) > C(?) + C(?-I')$ from 3
 - $C(I) < C(?)$ from 2 $C(?-I')$ satisfies above will be negative, so contradiction

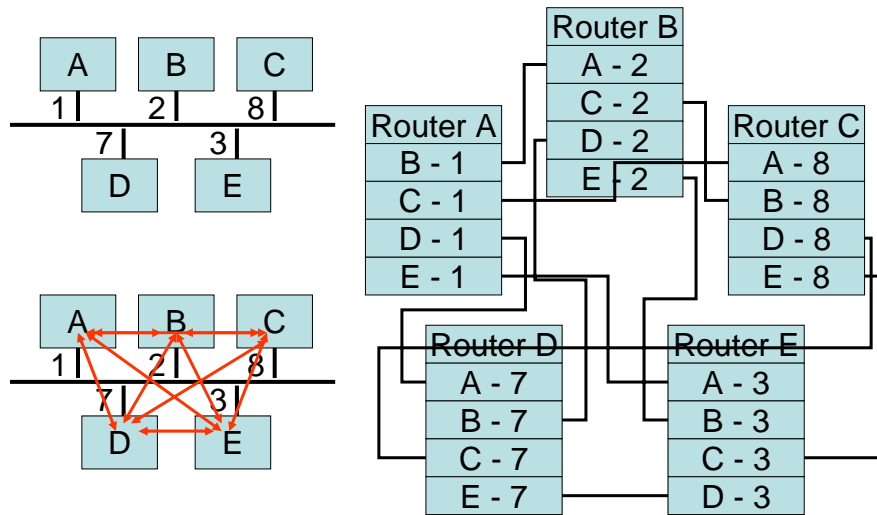
Dijkstra algorithm

- Is the path $A-B-E-I$ Loop free ?
- Yes, As long as LSDBs are synchronized

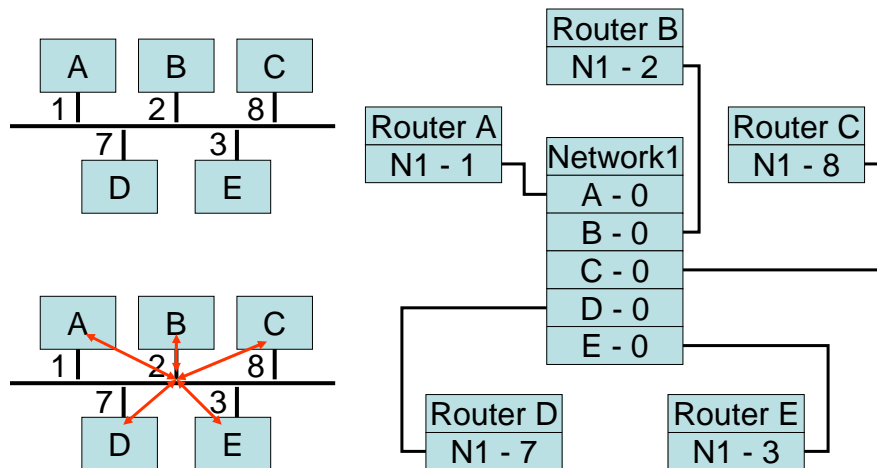


1. $C(A-C-F-?-I') > C(A-B-E-I)$ so $C(B-A-C-F-?-I') > C(B-E-I)$, no chance for B to get back to A (unless there's another path that we don't know)
2. the same goes for E and B, and for I and E

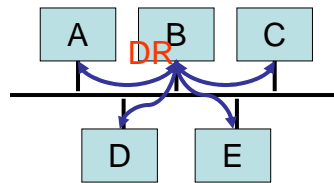
Without Designated Router and Network-LSA



With Designated Router and Network-LSA



Neighbor, Adjacency, DR/BDR



DR's role/task:

- LSDB synchronization proxy or relay point.
- Originate Network LSA on behalf of the network.

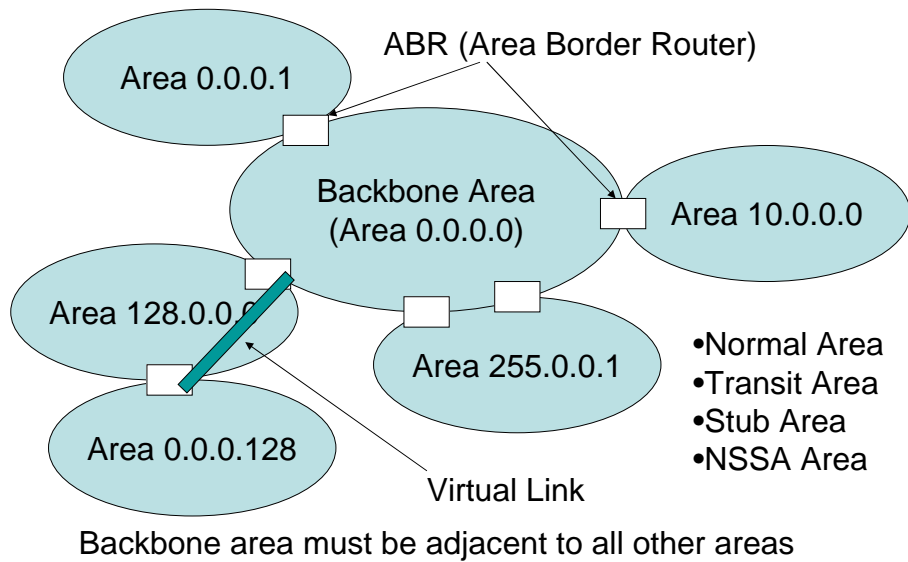
- Terminologies:

- **neighbor**: neighboring router
- **adjacency**: the relationship to synchronize LSDB
- e.g. Router B is a neighbor of router D, and they establish an adjacency between them (because B is the DR).

DR/BDR Election, neighbor states and Hello protocol

- DR is automatically elected on every multi-access network.
 - OSPF interface priority is used to elect.
 - DR is not preempted. It is sticky.
 - Changing DR may cause LSDB re-synchronization. In order to avoid this Backup DR (BDR) is also elected.
 - Every router on a multi-access network establishes adjacencies to both DR and BDR.
- Neighbor states:
 - to neighbors: Down-1Way-2Way.
 - to neighbors with an adjacency: Down-1Way-2Way-ExStart-Exchange-Loading-Full.
- OSPF Hello (sub)protocol
 - Discovers neighbors by means of multicast (like broadcast).
 - Makes sure that the router can communicate bi-directionally with the neighbor
 - Keeps watching the communication ability and detect loss of it.

OSPF Area Hierarchy



OSPF sample configuration

```

interface Loopback0
 ip address 203.178.136.2
 255.255.255.255
!
interface
 GigabitEthernet1/0.100
 ip address 203.178.137.89
 255.255.255.224
 ip ospf authentication-key
 <passwd>
 ip ospf cost 1
 ip ospf dead-interval 30
!

0.0.0.31=0.0.0.00011111(2)=28
203.178.137.64/28 is the range
203.178.137.64-95
so it includes 203.178.137.89

router ospf 2500
 router-id 203.178.136.2
 area 0.0.0.0 authentication
 area 203.178.141.224
 authentication
 network 203.178.136.2 0.0.0.0
 area 0.0.0.0
 network 203.178.136.96 0.0.0.3
 area 0.0.0.0
 network 203.178.137.64
 0.0.0.31 area 0.0.0.0
 network 203.178.138.96
 0.0.0.31 area 0.0.0.0
 network 203.178.138.224
 0.0.0.31 area 0.0.0.0
 network 203.178.141.224
 0.0.0.31 area
 203.178.141.224
    
```

2WAY or FULL

```
cisc01.fujisawa# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
lo-0.cisc01.not	1	2WAY/DROTHER	00:00:31	203.178.138.225	GigabitEthernet2/0.4
foundry2.otemac	0	2WAY/DROTHER	00:00:30	203.178.138.227	GigabitEthernet2/0.4
lo-1.foundry1.f	1	2WAY/DROTHER	00:00:29	203.178.138.253	GigabitEthernet2/0.4
lo-1.cisc01.nez	0	2WAY/DROTHER	00:00:31	203.178.138.231	GigabitEthernet2/0.4
lo-1.foundry4.o	1	2WAY/DROTHER	00:00:31	203.178.138.241	GigabitEthernet2/0.4
fe-0-7.hitachi2	0	2WAY/DROTHER	00:00:35	203.178.138.251	GigabitEthernet2/0.4
ge-0-1-0-v4.jun	1	2WAY/DROTHER	00:00:31	203.178.138.228	GigabitEthernet2/0.4
eth2.pcl.hongo.	0	2WAY/DROTHER	00:00:34	203.178.138.230	GigabitEthernet2/0.4
ge-0-0-0-v4.cis	0	2WAY/DROTHER	00:00:32	203.178.138.233	GigabitEthernet2/0.4
203.178.138.234	1	FULL/BDR	00:00:30	203.178.138.234	GigabitEthernet2/0.4
ve-4.foundry2.n	0	2WAY/DROTHER	00:00:37	203.178.138.237	GigabitEthernet2/0.4
ve-4.foundry1.y	0	2WAY/DROTHER	00:00:33	203.178.138.240	GigabitEthernet2/0.4
necl.yagami.wid	0	2WAY/DROTHER	00:00:32	203.178.138.242	GigabitEthernet2/0.4
ve-4.foundry3.n	1	FULL/DR	00:00:35	203.178.138.244	GigabitEthernet2/0.4
fe-fxp1.pc3.yag	0	2WAY/DROTHER	00:00:31	203.178.138.245	GigabitEthernet2/0.4
fe-fxp0.pc3.fuj	0	2WAY/DROTHER	00:00:29	203.178.138.254	GigabitEthernet2/0.4
lo-0.cisc011.fu	1	2WAY/DROTHER	00:00:25	203.178.137.78	GigabitEthernet1/0.100
lo-1.foundry1.f	1	FULL/BDR	00:00:19	203.178.137.91	GigabitEthernet1/0.100
fe-fxp0.pcl.fuj	0	2WAY/DROTHER	00:00:20	203.178.137.69	GigabitEthernet1/0.100
fe-0-7.hitachi2	0	2WAY/DROTHER	00:00:27	203.178.137.70	GigabitEthernet1/0.100
ve-100.foundry2	1	2WAY/DROTHER	00:00:20	203.178.137.74	GigabitEthernet1/0.100

show ip ospf database

```
OSPF Router with ID
(203.178.136.2) (Process ID 2500)
```

```
Router Link States
(Area 0.0.0.0)
```

```
Routing Bit Set on this LSA
```

```
LS age: 451
```

```
Options: (No TOS-capability, DC)
```

```
LS Type: Router Links
```

```
Link State ID: 203.178.136.0
```

```
Advertising Router: lo-
0.cisc01.otemachi.wide.ad.jp
```

```
LS Seq Number: 80004698
```

```
Checksum: 0x3CE6
```

```
Length: 96
```

```
Area Border Router
```

```
AS Boundary Router
```

```
Number of Links: 6
```

```
Link connected to: another Router (point-
to-point)
```

```
(Link ID) Neighboring Router ID:
203.178.136.33
```

```
(Link Data) Router Interface address:
203.178.136.161
```

```
Number of TOS metrics: 0
```

```
TOS 0 Metrics: 6000
```

```
Link connected to: a Stub Network
```

```
(Link ID) Network/subnet number:
203.178.136.160
```

```
(Link Data) Network Mask: 255.255.255.252
```

```
Number of TOS metrics: 0
```

```
TOS 0 Metrics: 6000
```

```
Link connected to: a Transit Network
```

```
(Link ID) Designated Router address:
203.178.140.221
```

```
(Link Data) Router Interface address:
203.178.140.221
```

```
Number of TOS metrics: 0
```

```
TOS 0 Metrics: 10
```

show ip ospf database

```
OSPF Router with ID (203.178.136.2) (Process ID 2500)

Net Link States (Area 0.0.0.0)

Routing Bit Set on this LSA
LS age: 1066
Options: (No TOS-capability, DC)
LS Type: Network Links
Link State ID: 203.178.140.221 (address of Designated Router)
Advertising Router: lo-0.cisco1.otemachi.wide.ad.jp
LS Seq Number: 80004593
Checksum: 0x75C
Length: 44
Network Mask: /27
  Attached Router: 203.178.136.0
  Attached Router: 203.178.136.15
  Attached Router: 203.178.136.19
  Attached Router: 203.178.136.34
  Attached Router: 203.178.145.177
```

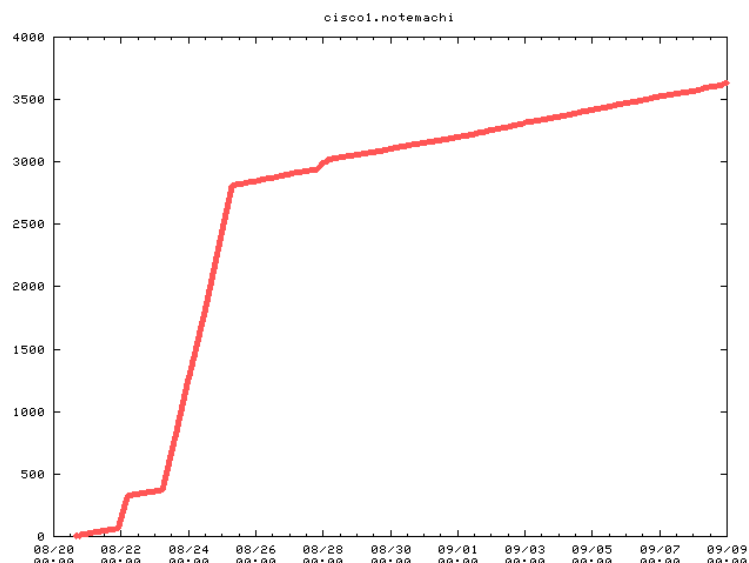
Common OSPF misconfigurations

- Interface netmask mismatch
- HelloInterval mismatch
- RouterDeadInterval mismatch
- Interface Area-ID or Area-Type mismatch
 - Result that the neighbor does not appear
 - “debug ip ospf hello” may help
- Interface MTU mismatch
 - Result that Neighbor state stops at ExStart
 - Neighbor relationship (Adjacency) never comes up (and so routes are not calculated)

Common OSPF misconfigurations

- Conflicting router-id
 - Routers conflicting router-id continue to refresh their LSA
 - The other's LSA seems as if it were an old LSA generated by the router itself
 - So those routers must override the other's LSA by generating a new LSA (LS seqnum is incremented)
 - Result in rapid growth of the LSA's LS Sequence Number
 - Persistent route oscillation at interval of about 5 seconds
 - routers in the area continue to execute SPF calculation again and again and again ...

LSA Seqnum sample growth



Open Issues

- LSA Refresh rate
 - periodic, rate cannot be changed (30 min.)
 - 140,000 LSAs result in refresh rate of 77 LSA/sec
 - IS-IS does not have this problem, but I don't recommend IS-IS
- Generating optimal cost is hard
- How to divide OSPF domain into Areas ?
- Area scalability
 - Some said less than 200 routers will be better
 - Other said 500 (or 1000?) has been experienced

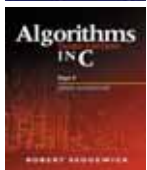
References



- *Christian Huitema, **Routing in the Internet**, Prentice Hall ; ISBN: 0130226475 ; 2nd edition (2000/01/15)*



- *John T. Moy, **Ospf: Anatomy of an Internet Routing Protocol**, Addison-Wesley ; ISBN: 0201634724 ; (1998/01/15)*



- *Robert Sedgewick, **Algorithms in C: Graph Algorithms (GRAPH ALGORITHMS)**, Addison-Wesley ; ISBN: 0201316633 ; 3rd edition, Graph Algorithms Pt.5 (2001/08/16)*

Today's Assignment

- Choose one of the assignments below.
 1. Consult yourself about OSPF's HelloInterval and RouterDeadInterval. Describe the effect of those values and relationship between them.
 2. Consult yourself about the functionality of AS-External-LSA (Type-5 LSA). Describe the functionality.
 3. Consult yourself about the types of OSPF area. Describe functionality for one of the area (select the one OSPF area type you would like to describe from Backbone, Normal, Transit, Stub and NSSA).