Deploying OSPF for ISPs
Agenda

• OSPF Design in SP Networks
• Adding Networks in OSPF
• OSPF in Cisco’s IOS
OSPF Design

As applicable to Service Provider Networks
Service Providers

- SP networks are divided into PoPs
- PoPs are linked by the backbone
- Transit routing information is carried via iBGP
- IGP is only used to carry the next hop for BGP
- Optimal path to the next hop is critical
SP Architecture

- Major routing information is ~390K prefixes via BGP
- Largest known IGP routing table is ~9–10K
- Total of 400K
- 10K/400K is 2½% of IGP routes in an ISP network
- A very small factor but has a huge impact on network convergence!
SP Architecture

- You can reduce the IGP size from 10K to approx the number of routers in your network
- This will bring really fast convergence
- Optimise where you must and summarise where you can
- Stops unnecessary flapping
OSPF Design: Addressing

• OSPF Design and Addressing go together
  – Objective is to keep the Link State Database lean
  – Create an address hierarchy to match the topology
  – Use separate Address Blocks for loopbacks, network infrastructure, customer interfaces & customers
OSPF Design: Addressing

• Minimising the number of prefixes in OSPF:
  – **Number loopbacks out of a contiguous address block**
    • But do not summarise these across area boundaries: iBGP peer addresses need to be in the IGP
  – Use contiguous address blocks per area for infrastructure point-to-point links
    • Use *area range* command on ABR to summarise

• With these guidelines:
  – Number of prefixes in area 0 will then be very close to the number of routers in the network
  – It is critically important that the number of prefixes and LSAs in area 0 is kept to the absolute minimum
OSPF Design: Areas

• Examine physical topology
  – Is it meshed or hub-and-spoke?

• Use areas and summarisation
  – This reduces overhead and LSA counts
  – (but watch next-hop for iBGP when summarising)

• Don’t bother with the various stub areas
  – No benefits for ISPs, causes problems for iBGP

• Push the creation of a backbone
  – Reduces mesh and promotes hierarchy
OSPF Design: Areas

• One SPF per area, flooding done per area
  – Watch out for overloading ABRs

• Avoid externals in OSPF
  – **DO NOT REDISTRIBUTE** into OSPF
  – External LSAs flood through entire network

• Different types of areas do different flooding
  – Normal areas
  – Stub areas
  – Totally stubby (stub no-summary)
  – Not so stubby areas (NSSA)
OSPF Design: Areas

- Area 0 must be contiguous
  - Do NOT use virtual links to join two Area 0 islands
- Traffic between two non-zero areas always goes via Area 0
  - There is no benefit in joining two non-zero areas together
  - Avoid designs which have two non-zero areas touching each other
  - (Typical design is an area per PoP, with core routers being ABR to the backbone area 0)
OSPF Design: Summary

• Think Redundancy
  – Dual Links out of each area – using metrics (cost) for traffic engineering

• Too much redundancy...
  – Dual links to backbone in stub areas must be the same cost – otherwise sub-optimal routing will result
  – Too Much Redundancy in the backbone area without good summarisation will effect convergence in the Area 0
OSPF Areas: Migration

• Where to place OSPF Areas?
  – Follow the physical topology!
  – Remember the earlier design advice

• Configure area at a time!
  – Start at the outermost edge of the network
  – Log into routers at either end of a link and change the link from Area 0 to the chosen Area
  – Wait for OSPF to re-establish adjacencies
  – And then move onto the next link, etc
  – Important to ensure that there is never an Area 0 island anywhere in the migrating network
OSPF Areas: Migration

- Migrate small parts of the network, one area at a time
  - Remember to introduce summarisation where feasible
- With careful planning, the migration can be done with minimal network downtime
OSPF for Service Providers

Configuring OSPF & Adding Networks
OSPF: Configuration

• Starting OSPF in Cisco’s IOS

  `router ospf 100`
  – Where “100” is the process ID

• OSPF process ID is unique to the router
  – Gives possibility of running multiple instances of OSPF on one router
  – Process ID is not passed between routers in an AS
  – Many ISPs configure the process ID to be the same as their BGP Autonomous System Number
OSPF: Establishing Adjacencies

• Cisco IOS OSPFv2 automatically tries to establish adjacencies on all defined interfaces (or subnets)
• Best practice is to disable this
  – Potential security risk: sending OSPF Hellos outside of the autonomous system, and risking forming adjacencies with external networks
  – Example: Only POS4/0 interface will attempt to form an OSPF adjacency

```
router ospf 100
  passive-interface default
  no passive-interface POS4/0
```
OSPF: Adding Networks
Option One

• Redistribution:
  – Applies to all connected interfaces on the router but sends networks as external type-2s – which are not summarised
    
    ```
    router ospf 100
    redistribute connected subnets
    ```

• **Do NOT do this!** Because:
  – Type-2 LSAs flood through entire network
  – These LSAs are not all useful for determining paths through backbone; they simply take up valuable space
OSPF: Adding Networks
Option Two

• Per link configuration – from IOS 12.4 onwards
  – OSPF is configured on each interface (same as ISIS)
  – Useful for multiple subnets per interface

```plaintext
interface POS 4/0
  ip address 192.168.1.0 255.255.255.0
  ip address 172.16.1.0 255.255.255.224 secondary
  ip ospf 100 area 0
!
router ospf 100
  passive-interface default
  no passive-interface POS 4/0
```
OSPF: Adding Networks
Option Three

• Specific network statements
  – Every active interface with a configured IP address needs an OSPF network statement
  – Interfaces that will have no OSPF neighbours need passive-interface to disable OSPF Hello’s
    • That is: all interfaces connecting to devices outside the ISP backbone (i.e. customers, peers, etc)

```
router ospf 100
  network 192.168.1.0 0.0.0.3 area 51
  network 192.168.1.4 0.0.0.3 area 51
  passive-interface Serial 1/0
```
OSPF: Adding Networks
Option Four

• Network statements – wildcard mask
  – Every active interface with configured IP address covered by wildcard mask used in OSPF network statement
  – Interfaces covered by wildcard mask but having no OSPF neighbours need passive-interface (or use passive-interface default and then activate the interfaces which will have OSPF neighbours)

```
router ospf 100
  network 192.168.1.0 0.0.0.255 area 51
  passive-interface default
  no passive interface POS 4/0
```
OSPF: Adding Networks Recommendations

• Don’t ever use Option 1
• Use Option 2 if supported; otherwise:
  • Option 3 is fine for core/infrastructure routers
    – Doesn’t scale too well when router has a large number of interfaces but only a few with OSPF neighbours
    – → solution is to use Option 3 with “no passive” on interfaces with OSPF neighbours
  • Option 4 is preferred for aggregation routers
    – Or use iBGP next-hop-self
    – Or even ip unnumbered on external point-to-point links
OSPF: Adding Networks
Example One (Cisco IOS ≥ 12.4)

• Aggregation router with large number of leased line customers and just two links to the core network:

```cpp
interface loopback 0
  ip address 192.168.255.1 255.255.255.255
  ip ospf 100 area 0
interface POS 0/0
  ip address 192.168.10.1 255.255.255.252
  ip ospf 100 area 0
interface POS 1/0
  ip address 192.168.10.5 255.255.255.252
  ip ospf 100 area 0
interface serial 2/0:0 ...
  ip unnumbered loopback 0
! Customers connect here ^^^^^^^
routing ospf 100
  passive-interface default
  no passive interface POS 0/0
  no passive interface POS 1/0
```
OSPF: Adding Networks
Example One (Cisco IOS < 12.4)

• Aggregation router with large number of leased line customers and just two links to the core network:

```plaintext
interface loopback 0
  ip address 192.168.255.1 255.255.255.255
interface POS 0/0
  ip address 192.168.10.1 255.255.255.252
interface POS 1/0
  ip address 192.168.10.5 255.255.255.252
interface serial 2/0:0 ...
  ip unnumbered loopback 0
! Customers connect here ^^^^^^^^
router ospf 100
  network 192.168.255.1 0.0.0.0 area 51
  network 192.168.10.0 0.0.0.3 area 51
  network 192.168.10.4 0.0.0.3 area 51
  passive-interface default
  no passive interface POS 0/0
  no passive interface POS 1/0
```
OSPF: Adding Networks
Example Two (Cisco IOS ≥ 12.4)

- Core router with only links to other core routers:

```plaintext
interface loopback 0
  ip address 192.168.255.1 255.255.255.255
  ip ospf 100 area 0
interface POS 0/0
  ip address 192.168.10.129 255.255.255.252
  ip ospf 100 area 0
interface POS 1/0
  ip address 192.168.10.133 255.255.255.252
  ip ospf 100 area 0
interface POS 2/0
  ip address 192.168.10.137 255.255.255.252
  ip ospf 100 area 0
interface POS 2/1
  ip address 192.168.10.141 255.255.255.252
  ip ospf 100 area 0
router ospf 100
  passive interface loopback 0
```
OSPF: Adding Networks
Example Two (Cisco IOS < 12.4)

• Core router with only links to other core routers:

```plaintext
interface loopback 0
  ip address 192.168.255.1 255.255.255.255
interface POS 0/0
  ip address 192.168.10.129 255.255.255.252
interface POS 1/0
  ip address 192.168.10.133 255.255.255.252
interface POS 2/0
  ip address 192.168.10.137 255.255.255.252
interface POS 2/1
  ip address 192.168.10.141 255.255.255.252
router ospf 100
  network 192.168.255.1 0.0.0.0 area 0
  network 192.168.10.128 0.0.0.3 area 0
  network 192.168.10.132 0.0.0.3 area 0
  network 192.168.10.136 0.0.0.3 area 0
  network 192.168.10.140 0.0.0.3 area 0
  passive interface loopback 0
```
OSPF: Adding Networks
Summary

• Key Theme when selecting a technique: *Keep the Link State Database Lean*
  – Increases Stability
  – Reduces the amount of information in the Link State Advertisements (LSAs)
  – Speeds Convergence Time
OSPF in Cisco IOS

Useful features for ISPs
Areas

• An area is stored as a 32-bit field:
  – Defined in IPv4 address format (i.e. Area 0.0.0.0)
  – Can also be defined using single decimal value (i.e. Area 0)

• 0.0.0.0 reserved for the backbone area
Logging Adjacency Changes

• The router will generate a log message whenever an OSPF neighbour changes state

• Syntax:
  – [no] [ospf] log-adjacency-changes
  – (OSPF keyword is optional, depending on IOS version)

• Example of a typical log message:
  – %OSPF-5-ADJCHG: Process 1, Nbr 223.127.255.223 on Ethernet0 from LOADING to FULL, Loading Done
Number of State Changes

• The number of state transitions is available via SNMP (ospfNbrEvents) and the CLI:
  – `show ip ospf neighbor [type number] [neighbor-id] [detail]`
  – Detail—(Optional) Displays all neighbours given in detail (list all neighbours). When specified, neighbour state transition counters are displayed per interface or neighbour ID
State Changes (Continued)

• To reset OSPF-related statistics, use the `clear ip ospf counters` command
  – This will reset neighbour state transition counters per interface or neighbour id
  – `clear ip ospf counters [neighbor [<type number>] [neighbor-id]]`
Router ID

- If the loopback interface exists and has an IP address, that is used as the router ID in routing protocols – **stability**!
- If the loopback interface does not exist, or has no IP address, the router ID is the highest IP address configured – **danger**!
- OSPF sub command to manually set the Router ID:
  
  `– router-id <ip address>`
Cost & Reference Bandwidth

• Bandwidth used in Metric calculation
  – Cost = $10^8$/bandwidth
  – Not useful for interface bandwidths > 100 Mbps
• Syntax:
  – `ospf auto-cost reference-bandwidth <reference-bw>`
• Default reference bandwidth still 100 Mbps for backward compatibility
• Most ISPs simply choose to develop their own cost strategy and apply to each interface type
## Cost: Example Strategy

<table>
<thead>
<tr>
<th>Bandwidth Type</th>
<th>Bandwidth (Mbps)</th>
<th>Cost (units)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100GE</td>
<td>100Gbps</td>
<td>1</td>
</tr>
<tr>
<td>40GE/OC768</td>
<td>40Gbps</td>
<td>2</td>
</tr>
<tr>
<td>10GE/OC192</td>
<td>10Gbps</td>
<td>5</td>
</tr>
<tr>
<td>OC48</td>
<td>2.5Gbps</td>
<td>10</td>
</tr>
<tr>
<td>GigEthernet</td>
<td>1Gbps</td>
<td>20</td>
</tr>
<tr>
<td>OC12</td>
<td>622Mbps</td>
<td>50</td>
</tr>
<tr>
<td>OC3</td>
<td>155Mbps</td>
<td>100</td>
</tr>
<tr>
<td>FastEthernet</td>
<td>100Mbps</td>
<td>200</td>
</tr>
<tr>
<td>Ethernet</td>
<td>10Mbps</td>
<td>500</td>
</tr>
<tr>
<td>E1</td>
<td>2Mbps</td>
<td>1000</td>
</tr>
</tbody>
</table>
Default routes

• Originating a default route into OSPF
  – `default-information originate metric <n>`
  – Will originate a default route into OSPF if there is a matching default route in the Routing Table (RIB)
  – The optional `always` keyword will always originate a default route, even if there is no existing entry in the RIB
Clear/Restart

- **OSPF clear commands**
  - If no process ID is given, all OSPF processes on the router are assumed

- **clear ip ospf [pid] redistribution**
  - This command clears redistribution based on OSPF routing process ID

- **clear ip ospf [pid] counters**
  - This command clears counters based on OSPF routing process ID

- **clear ip ospf [pid] process**
  - This command will restart the specified OSPF process. It attempts to keep the old router-id, except in cases where a new router-id was configured or an old user configured router-id was removed. Since this command can potentially cause a network churn, a user confirmation is required before performing any action
Use OSPF Authentication

• Use authentication
  – Too many operators overlook this basic requirement
• When using authentication, use the MD5 feature
  – Under the global OSPF configuration, specify:
    area <area-id> authentication message-digest
  – Under the interface configuration, specify:
    ip ospf message-digest-key 1 md5 <key>
• Authentication can be selectively disabled per interface with:
  ip ospf authentication null
Point to Point Ethernet Links

- For any broadcast media (like Ethernet), OSPF will attempt to elect a designated and backup designated router when it forms an adjacency
  - If the interface is running as a point-to-point WAN link, with only 2 routers on the wire, configuring OSPF to operate in "point-to-point mode" scales the protocol by reducing the link failure detection times
  - Point-to-point mode improves convergence times on Ethernet networks because it:
    - Prevents the election of a DR/BDR on the link,
    - Simplifies the SPF computations and reduces the router's memory footprint due to a smaller topology database.

```bash
interface fastethernet0/2
  ip ospf network point-to-point
```
Tuning OSPF (1)

• DR/BDR Selection
  – `ip ospf priority 100` (default 1)
  – This feature should be in use in your OSPF network
  – Forcibly set your DR and BDR per segment so that they are known
  – Choose your most powerful, or most idle routers, so that OSPF converges as fast as possible under maximum network load conditions
  – Try to keep the DR/BDR limited to one segment each
Tuning OSPF (2)

• OSPF startup
  – `max-metric router-lsa on-startup wait-for-bgp`
  – Avoids blackholing traffic on router restart
  – Causes OSPF to announce its prefixes with highest possible metric until iBGP is up and running
  – When iBGP is running, OSPF metrics return to normal, make the path valid

• ISIS equivalent:
  – `set-overload-bit on-startup wait-for-bgp`
Tuning OSPF (3)

• Hello/Dead Timers
  – `ip ospf hello-interval 3` (default 10)
  – `ip ospf dead-interval 15` (default is 4x hello)
  – This allows for faster network awareness of a failure, and can result in faster reconvergence, but requires more router CPU and generates more overhead

• LSA Pacing
  – `timers lsa-group-pacing 300` (default 240)
  – Allows grouping and pacing of LSA updates at configured interval
  – Reduces overall network and router impact
Tuning OSPF (4)

• OSPF Internal Timers
  – `timers spf 2 8` (default is 5 and 10)
  – Allows you to adjust SPF characteristics
  – The first number sets wait time from topology change to SPF run
  – The second is hold-down between SPF runs
  – BE CAREFUL WITH THIS COMMAND; if you’re not sure when to use it, it means you don’t need it; default is sufficient 95% of the time
Tuning OSPF (5)

• LSA filtering/interface blocking
  – Per interface:
    • `ip ospf database-filter all out` (no options)
  – Per neighbor:
    • `neighbor 1.1.1.1 database-filter all out` (no options)
  – OSPF's router will flood an LSA out all interfaces except the receiving one; LSA filtering can be useful in cases where such flooding unnecessary (i.e., NBMA networks), where the DR/BDR can handle flooding chores
  – `area <area-id> filter-list <acl>`
  – Filters out specific Type 3 LSAs at ABRs

• Improper use can result in routing loops and black-holes that can be very difficult to troubleshoot
Summary

- OSPF has a bewildering number of features and options
- Observe ISP best practices
- Keep design and configuration simple
- Investigate tuning options and suitability for your own network
  - Don’t just turn them on!
Acknowledgement and Attribution

This presentation contains content and information originally developed and maintained by the following organisation(s)/individual(s) and provided for the African Union AXIS Project

Cisco ISP/IXP Workshops

Philip Smith: - pfsinoz@gmail.com

www.apnic.net
Deploying OSPF for ISPs

End