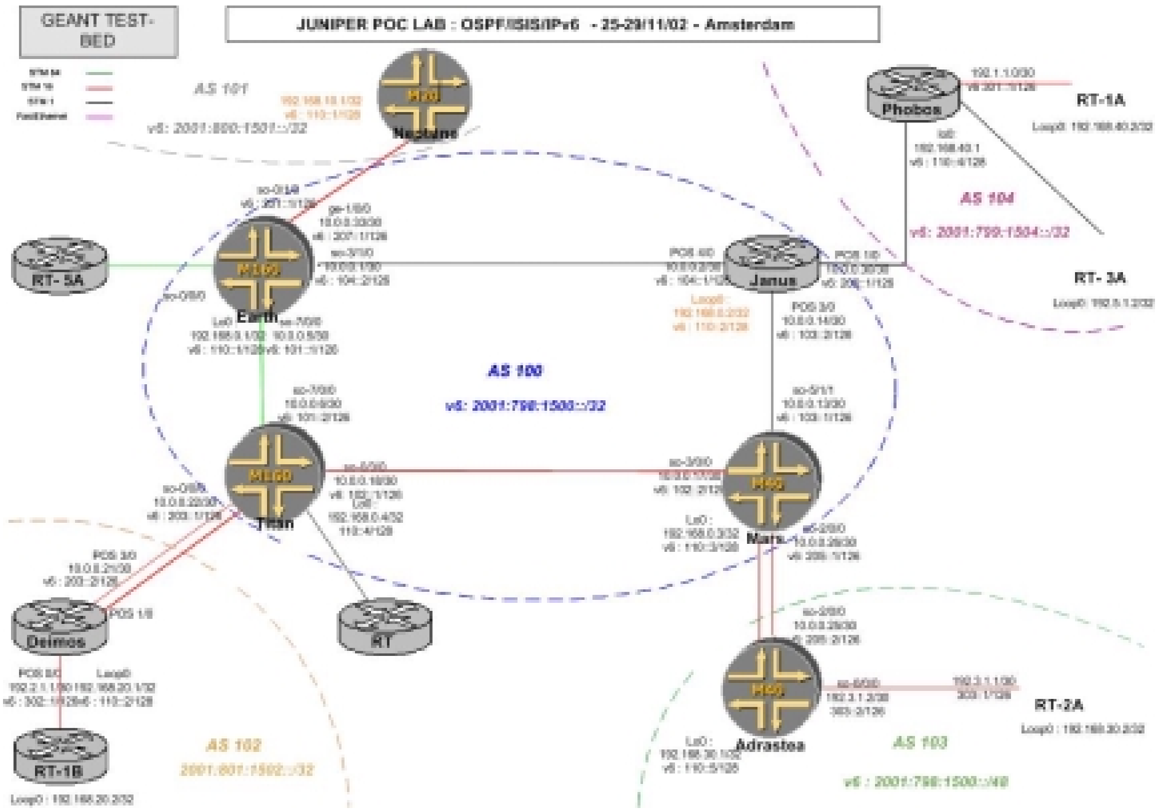
POC lab Test – Dante: November 25-29
IS-IS Migration and IPv6 Integration
Summary report

Editors	
	Agnes Pouele
	Marian Vidondo Garcia
	Ioannis Kappas
	Aurelia Peltier

References:

NEP-02-081 Draft IPV6 lab test plan
Jean-marc Uze's summary

GEANT Test-bed



This drawing represents the test-bed that we have set-up.
During the week, Phobos has moved from AS104 to AS100, we've renumbered slightly the IP addresses too.

1. PART 1 OSPF-ISIS migration

1.1. Test 1: Performance tests on 7k and M40 with IS-IS and OSPF in cohabitation + change of preference/distance of protocols

Results: test skipped due to Agilent interoperability issue when injecting OSPF routing

Conclusion: no impact as test integrated in Test 2

1.2. Test 2: Performance tests on GEANT test-bed with IS-IS and OSPF in cohabitation + change of preference/distance of protocols

Results:

- configuration of IS-IS successful (+ preference set-up to 169 to be the less preferred IGP, OSPF has its default of 10)
- no issue with adjacency between Cisco and Juniper
- redistribution of 400 BGP routes into OSPF and IS-IS
- write of the Python shell script “Sigkrisis”to compare the OSPF and IS-IS routes into Inet0
- comment: on a local router, we observed in inet0 OSPF direct connected networks announced by the peers (interconnection network routes between the local router and the peers). Is-is ignores such routes. No operational impact but must be taken in account to monitor the number of routes and compare between OSPF and IS-IS.

Conclusion: test successful

1.3. Test 3: Performance tests on GEANT test-bed with IS-IS and OSPF in cohabitation + GEANT services activated

Results:

- IS-IS configuration for Multicast successful (IS-IS routes are redistributed into inet2)
- Configuration of 3 MPLS LSPs using CSPF successful
- Traffic generation of Multicast traffic and unicast traffic
    - 10 Mb/s Multicast between Netptune and Deimos
    - 10 Mb/s Multicast between Netptune and Phobos
    - 5 Gb/s Unicast between Earth and Titan

  - 2 Gb/s Unicast between Adrastea and Deimos
  - 10 Mb/s Unicast between Adrastea and Phobos
- Note: 30% CPU utilization on Phobos (7500)

Conclusion: set-up ready. Test successful.

Comments:

The set-up tried is validated and implemented in GEANT. Multicast and MPLS services are running with this set-up.

1.4. Test 4: Following of test number 3 => migration

Results:

- Migration router one by one starting from Janus (7200), Mars (M40), Titan (M160), Earth (M160), by deactivating OSPF.
- No interruption of traffic (unicast and multicast, no LSP failure)
- Use of the script to check activation of IS-IS routes in inet0 (all routes previously enabled by OSPF)

Conclusion: test successful. Migration process qualified.

Comments:

The difficulty consist in counting the routes in INET.0
This test permitted us to confirm IS-IS configuration and to establish the procedures of the migration. The following documents have been updated or edited

  - Update of NEP 2002 87 GEANT ISIS for IS-IS

The detailed configuration and check list for the migration can be found under:

  - AP01G009_CHK0100_1-GeantIsisMigrationChecklist and document
  - AP01G009_MOP0099_1-GeantIsisMigrationConfiguration

2. PART II: IPV6 design

2.1. Test 5: Configuration of IPV6 on GEANT test-bed

- generation of v4&v6 traffic,
- injection of 400 routes of Ipv4 and Ipv6 into IS-IS to simulate GEANT internal routing.

Results:

a) Limitation with current IS-IS implementation is lack of multi-topology support. That means that the IPv4 and IPv6 topology must be the same. So all routers of the GEANT backbone must be dual stack and be part of the IS-IS topology for both IPv4 and IPv6.

b) Some differences have to be taken in account with IS-IS design. By using wide metrics, traffic engineering is possible but the notion of external routes is IS-IS is not possible any more, which has some incidence to the selection of best internal route to a destination. The design of GEANT with OSPF supported the notion of external routes, so the change has to be taken in account.

c) BGP design has 2 possibilities (for both I-BGP and E-BGP):

One BGP TCP session over v4 supporting both IPv4 AFI and IPv6 AFI. The advantage is to integrate both v4 and v6 in one unique BGP/TCP session for each peer. As the TCP source/destination addresses are IPv4, the next-hop chosen by IPv6 NLRIs is an IPv6 address automatically derived from the IPv4 source address, to build an IPv4 compatible address (::IPv4).
Unfortunately this NH is not part of the IPv6 addressing scheme, so it is not a real IPv6 addresses known by the IGP.
A workaround would consist of configuring a static route for each peering E-BGP to tell the router that the IPv4 compatible next-hop address is reachable via the physical interface of the e-BGP peering on that router.
This static route would need to be redistributed into the IGP for the other routers in the core backbone to learn about the BGP next-hop.
For I-BGP the BGP peering is made over IPv4 loopback @, the e-BGP nexthop is not modified for routes announced via the I-BGP therefore there is no problem.

We haven't tested the routing policy "Next-hop self" applied towards the internal peers in case of one BGP TCP connections announcing V4 and V6 routes.

We haven't been able to configure one BGP peering announcing V4 and V6 routes on the 7500 with the IOS 12.2.(13)T, the commands where not available.

We've decided to design GEANT with two BGP TCP sessions. One over IPv4 for IPv4 routing (already existing in GEANT configuration), and in parallel one new over IPv6 for IPv6 routing.
All BGP sessions are duplicated to support IPv6 peering, both for the E-BGP sessions and I-BGP sessions. In this case the Next-hop for the IPv6 addresses will be automatically the real addresses part of the IPv6 addressing scheme and exchanged in the IGP.
The dis-advantage of this solution is the duplication of the I-BGP mesh by two. GEANT doesn't have Route Reflectors neither for Ipv4 or Ipv6. However due to the small amount of routers (20), the network should be able to support a dual full mesh.

d) Generation of traffic

	Ipv4 traffic	Ipv6 traffic	Multicast Ipv4	results
On Juniper routers M160&M40 2385.40 Mbps of Ipv4 and Ipv6	700 Mb/s 64 byte pkts 200 Mb/s 576 byte pkts 190 Mb/s 1550 byte pkts	700 Mb/s 64 byte pkts 200 Mb/s 576 byte pkts 190 Mb/s 1550 byte pkts	10 Mbps	No packet lost Average latency 62 micro seconds
On 7500	100 Mbps	1 Mbps	From 500kbps to 3M	100 % CPU load

Comment:

For a dual stack core backbone, which can forward natively Ipv4 and Ipv6 traffic, it doesn't make sense to use translated addresses from one protocol to another one.
We can't design a solution where V4 and V6 would be linked. Both designs must be independent and evolve separately.

Conclusion: test successful. All routers supported the traffic load. CPU on Cisco 7500 100% but no packet loss.

2.2. Test 6: Test performance on BGPv4+, announce of 100000 Ipv4 routes and 2000 routes for Ipv6 to GEANT test-bed and forwarding of traffic

Result:

7500 not stable. Loss of control of the router. IPv6 packet drops observed.
No problem observed on Juniper. Although the slow convergence of the 7500 impacted the RE CPU of the M40.

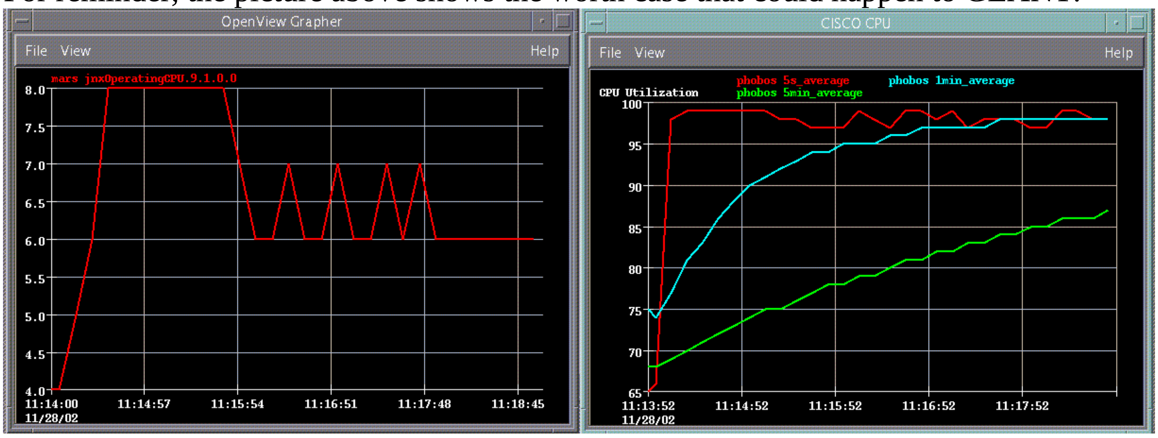
Conclusion:

The Cisco 7500 cannot be included in the IPv6 forwarding path and can’t deliver Ipv6 services. But based on Test 5 results, it has to be part of the control plane (IS-IS + BGP) and will be a potential Ipv6 transit router for the core backbone.
So two solutions are envisaged

a. Metric of IS-IS can be defined in order to avoid any transit traffic on the 7500 (it will concerns both v6 and v4 as we can’t separate them). So except if there are major failures that impact all other potential paths (e.g. Italy looses all links except the one to Israel, or Belgium looses all links except the one to Luxembourg), the 7500 (in Israel or Luxembourg) will never to forward transit traffic.
b. The other solution consists of applying IPv6 firewall on Juniper interfaces facing the 7500 routers, to limit the output traffic (for example 300 Kb/s). By lack of time this solution was not tested.

Comment:

GEANT will implement solution “a” until these routers will be replaced by more recent platforms.
For reminder, the picture above shows the worth case that could happen to GEANT.



The left view is the M40, Mars, and the “CPU” usage of the routing engine.
The right view is the 7500 and the CPU usage.
The Agilent stopped to announce in the network 100000 Ipv4 prefixes.
The M40 converged very quickly for the withdrawing of these 100.000 while the 7500 could not face such demand.
The whole network became unstable (red picks on the left figure). The Juniper platforms continue to learn from the 7500 prefixes that they have withdrawn and re-announce them until the 7500 has finished to converge.

2.3. Test 7: Test on BGPv4+ routing policy + traffic + GEANT services

Results:

a) We tested Local Pref, MED, and BGP communities applied for V4 and V6 routing policy. We can use for both Ipv4 and ipv6 the same community tagging.
Local pref and MED can be indifferently applied to V4 and V6 peering with different values.

b) Prefix list filters have been tested with v4 and v6

c) No traffic was finally sent in the testbed, as forwarding performance were already qualified in the previous tests.

Conclusion: test successful

Remark:

We have to configure two external and internal BGP peering groups: for “V4” and “V6” Therefore it’s preferable that the type of BGP peering “family inet any” or “family inet6 any” is specified at the group level.

Routing policy have to be written carefully.
One mis-place “then accept” or “then reject” can avoid the whole set of routing policy to be applied to the peer.

Here an example of correct syntax that we should use for the configuration of Ipv6 BGP routing policy:

```
policy-statement ipv4_export {
    term med {
        from family inet;
        then next term;
    }
    term check_community_nrn {
        from {
            community GEANT-NRN;
            family inet;
        }
        then next policy;
    }
    term check_community_dws {
        from {
            community GEANT-DWS;
            family inet;
        }
        then next policy;
    }
    term check_community_others {
        from {
            protocol bgp;
            family inet;
        }
```

```
        then reject;
    }
}
policy-statement ipv6_export {
    term med {
        from family inet6;
        then next term;
    }
    term check_community_nrn {
        from {
            community GEANT-NRN;
            family inet6;
        }
        then next policy;
    }
    term check_community_dws {
        from {
            community GEANT-DWS;
            family inet6;
        }
        then next policy;
    }
    term check_community_others {
        from {
            protocol bgp;
            family inet6;
        }
        then reject;
    }
}
```

2.4. Test 8: GEANT Ipv4&Ipv6 routing policy

Test not realized because GEANT routing policy not yet defined.

2.5. Test 9bis: Test Vlan access described in CS's document and tunnel access

Results:

a) If the IPv4 service and IPv6 service are provided through the same access port, it is recommended to separate the traffic in two different channels. That will simplify the monitoring (stat per logical interface provides automatically IPv6 stats or IPv4 stats). These channels can be VLAN on a GE interface or FR DLCI on a POS interface.

b) Test was performed with 2 routers connected with GE, 300 Mb/s IPv4 on one VLAN and 500 Mb/s IPv6 on a second VLAN. The routers were Neptune and Earth. Line rate observed. No drop.

c) Results with tunnels

To be added

Conclusion: For Monitoring purpose we'll use virtual channels on the primary access of the NRENs with DLCIs dedicated to the Ipv6 traffic or Vlans.

2.6. Test 10,11 and 12 are missing

They correspond to performance forwarding of Ipv4 and Ipv6 on Juniper routers and test of access lists.

They will be done later on.