An IPv6 Deployment Guide

Editor: Martin Dunmore

An IPv6 Deployment Guide

6NET was a three-year European IST project to demonstrate that continued growth of the Internet can be met using new IPv6 technology. The project built and operated a pan-European native IPv6 network connecting sixteen countries in order to gain experience of IPv6 deployment and the migration from existing IPv4-based networks.

6NET involved thirty-five partners from the commercial, research and academic sectors and represented a total investment of €18 million; €7 million of which came from the project partners themselves, and €11 million from the Information Society Technologies Programme of the European Commission. The project commenced on 1st January 2002 and officially finished on 30th June 2005. The network itself was decommissioned in January 2005, handing over the reigns of pan-European native IPv6 connectivity to GÉANT.

When we began 6NET, IPv6 code was in the form of early beta releases from most commercial companies. The 6BONE had been built but was only using tunnels; there were very few native IPv6 networks and none of these ran production traffic. One thing we strived for in the early days of 6NET was developing a pan-European testbed that had as much native IPv6 connectivity as was affordable. This gave everyone involved the chance to really exercise the IPv6 protocol developments we planned without the added complexity of tunnels that might detract from the real work. Soon we were able to peer with other IPv6 networks in the US (Abilene, 6TAP), Japan (NTT) and S.Korea (KOREN) to provide global IPv6 connectivity. The final stages of the project moved into exploiting the protocol and providing demonstrations that IPv6 was ready for full production service.

The information contained in this book is taken from the project's deployment cookbooks and other deliverables. Since each cookbook/deliverable generally concentrates only on specific IPv6 features or deployment scenarios (e.g. site transition, multicast, mobility, DHCP, routing etc.), we believe that providing all the important information in a single reference book is much more preferable to the reader than negotiating our multitude of project deliverables.

Chapter 12

IPv6 in the Backbone

In this chapter we present case studies of IPv6 in the backbone. First we look at the core 6NET backbone and the NRENs that were connected to it before the core network was decommissioned in January 2005.

Next, we detail case studies of IPv6 deployment by the NRENs themselves inside their own country backbone networks. The most common method to introduce IPv6 services into IPv4 networks will be through dual-stack networking. This complements the backbone transition and pushes the issue of deployment to the edge, e.g. to the universities.

In the timeframe of 6NET, many NRENs migrated to dual stack; the specific experiences of SURFnet, Funet and Renater are reported here.

12.1 6NET Backbone Case Study

A backbone IPv6 network connecting sixteen countries and running at 155 Mbps was established in 2002. This ran IPv6 over dedicated links, although for cost reasons, four links (to Greece, Hungary, Poland and Portugal) were provided by POS (Packet-over-SONET/SDH) over a Layer 2 VPN infrastructure.

Local access was provided through national IPv6 testbeds operated by partner NRENs (National Research and Education Networks) such as JANET (UK), RENATER (France) and SWITCH (Switzerland). Connectivity to the non-European 6NET partners in Japan and South Korea was provided via connections to London and RENATER respectively, and there are were connections to Abilene in the US (via SURFnet), Euro6IX (via the JANET- UK6X, GARR-TILab and SWITCH-Swisscom exchange points) and to the 6Bone.

The 6NET backbone, and interconnected national testbeds, collectively formed the largest native IPv6 network in the world. This provided plenty of scope for trialling the new technology, testing interoperability with existing networks, and demonstrating services and applications. In fact, it demonstrated that the IS-IS and BGP4+ routing protocols, IPv6 over IPv4 tunnelling, and DNS support were stable and usable. In addition, a multicast overlay network (M6Bone) was established and has been utilised for conferencing and radio broadcasting (e.g., Trondheim Underground Radio).

12.1.1 Network Topology

This section explains the topology of the network, Figure 12-1 presents the topology as it was in February 2003 for the 6NET core and NREN PoPs.

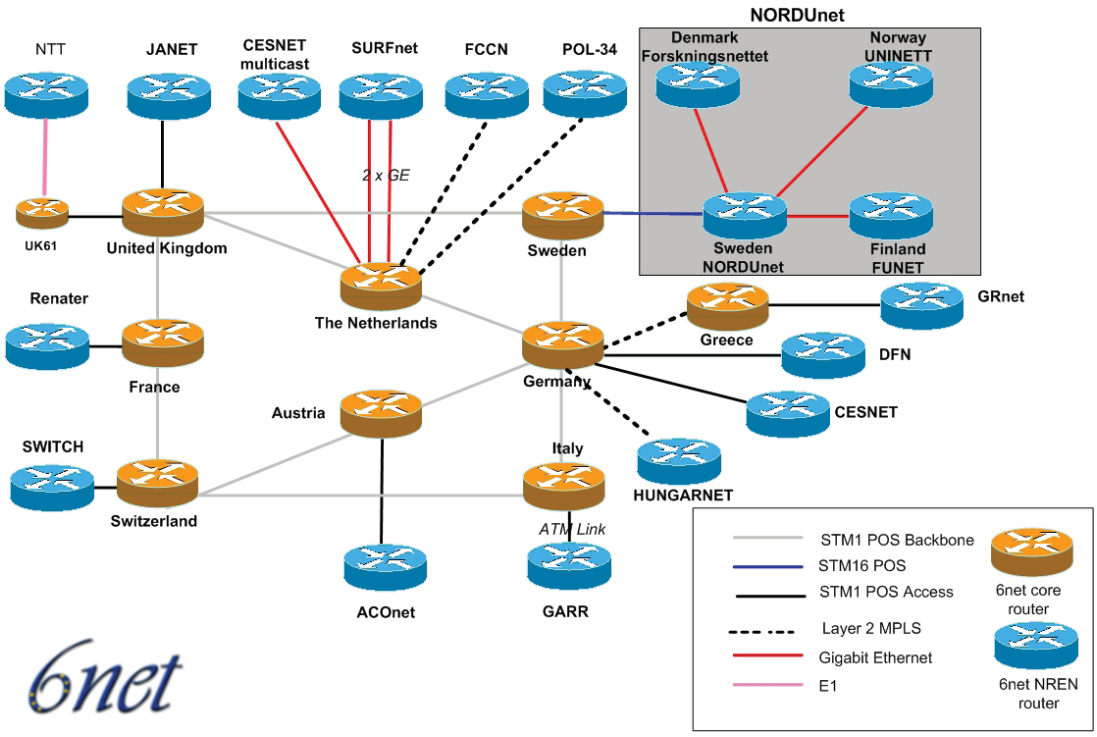

Figure 12-1 The 6NET Core and NREN PoPs

The 6NET network consisted of a set of international POS STM1/OC3 backbone circuits and a set of access circuits which mostly were native connections plus some tunnelled connections to 6NET partners. The 6NET core is represented by orange routers whilst blue routers represents the access routers of the NRENs.

In total, there were 16 partners directly connected to the 6NET core. POL34 was connected via a manual IPv4-to-IPv6 tunnel to the Swedish router. It also had a backup manual tunnel connection to the German Router. GRnet was connected to 6NET via a CCC/MPLS connection to the German router.

Previously, manual IPv4-to-IPv6 tunnels were configured for connecting HUNGARNET and CESNET. From January 2003 they were changed to STM1/OC3-155Mbps native connections. The rest of the access links were STM1/OC3-155Mbps, except NORDUnet with STM16/OC48-2.5Gbps, SURFnet with a 2xGE (2x1Gbps) access links and NTT with an E1 link to the UK router (2Mbps).

12.1.2 Addressing Scheme

Each NRENs used their IPv6 subTLA allocation. DANTE obtained an IPv6 subTLA for the pan-European IPv6 research backbone - divided into 6NET and GÉANT and other IPv6 purposes.

The IPv6 subTLA assigned to Dante is 2001:0798::/35. From this IPv6 address space a sub-section was used for 6NET. The allocated 6NET address space was ‘2001:0798::/40’.

The 6NET available address space '2001:0798:0::/40' was divided into logical sub portions to facilitate future expansions and enable simpler summarisation rules when required.

The address range can be seen as:

Table 12-1 6NET Prefix

35 bits	5 bits	8 bits	16 bits
DANTE subTLA	project	PoPs	SLA
	/35	/40 /48	/64

For 6NET, the 5 bit <project> part consists of all zeroes. Other <project> allocations will be made by DANTE.

To summarize, the assigned IPv6 address consists of the following parts:

Structure = <xyz>.<PoP>.<SLA>

Where:

<xyz> = <40 bit provider prefix> = 2001:0798::/40

<PoP> = assigned address range per PoP

<SLA> = segmentation of use within the PoP

SLA =Site Level Aggregate

12.1.2.1 PoP Addressing

The 6NET locations that had a 6NET core router got their own prefixes. The 8 bits for the PoP part are used as follows:

00. core links
01. reserved core links
…
10. AT
11. BE
12. CH
13. CZ
14. DE
15. ES
16. FR
17. GR
18. HU
19. IE
20. IT
21. LU

22. NL
23. PL
24. PT
25. SE/NORDUNET
26. SI
27. SK
28. UK

Note: The start of the country PoP areas started at 10 for visual reasons when reading the IPv6 prefix addresses. The hexadecimal addresses A-F were not used. This gave the following prefixes for the PoPs:

Table 12-2 PoP Addressing

PoP Location	IPv6 PoP addressing: 2001:0798:<PoP>::/48
Core	2001:0798:00::/48
Sweden	2001:0798:25::/48
Netherlands	2001:0798:22::/48
Germany	2001:0798:14::/48
Austria	2001:0798:10::/48
Italy	2001:0798:20::/48
Switzerland	2001:0798:12::/48
France	2001:0798:16::/48
UK	2001:0798:28::/48
Greece	2001:0798:17::/48

The only 'special' address range, which is not really bound to a geographical location, is the 'core' address range. This address range is used for the connections between PoPs.

The SLA (Site Level Aggregate) is used for various prefixes, (mostly) within a PoP:

Table 12-3 SLA Usage

Range	Use
0000 - 00FF	Loopback addresses
0100 - 01FF	Intra-PoP point-to-points
0200 - 02FF	connections to NREN PoPs
0300 - 03FF	external 6NET connectivity
0400 - 04FF	PoP LANs

Note: Using this convention there is room for 256 prefixes with /64 address space for each point-to-point link or broadcast media.

The prefix length selected for point-to-point connections was /64. A /64 was chosen because it seems to be the best current practice. It makes the number plan easy because every interface gets a /64. There is no need for a /126, /127 or /128 because address conservation is not a deciding factor with IPv6.

The potential future use of a /127 address space (or other address space) which initially seems to use only minimal address range will not work in the long run. More study material can be found in RFC 3627 [RFC3627].

It is also possible to use /128 on point-to-point links. The drawback here is that in Cisco IOS you need to add a static route to the remote side.

For Switzerland with <PoP>=12 this would give the following prefixes:

Table 12-4 Switzerland Prefixes

Use	Prefix
Loopback addresses:	2001:0798:0012:0000::/56
Intra-PoP Point-to-points:	2001:0798:0012:0100::/56
Connections to NREN PoP:	2001:0798:0012:0200::/56
External 6NET connection:	2001:0798:0012:0300::/56
PoP LANs:	2001:0798:0012:0400::/56

12.1.2.2 Loopback Addresses

In networking environments, it is seen as good practice to give each device an IPv6 loopback address. This is an IPv6 address that is not directly assigned to any physical interface and will typically be reachable when the networking appliance (in this case a router) is up and running.

This loopback address is also used for operational and management actions on the equipment and for routing protocols like eBGP, which use these addresses for terminating the peering sessions.

Loopback addresses typically have a prefix mask of /128. This avoids unnecessary unused addresses although address conservation is not really an issue in IPv6.

In the initial 6NET core network, the 6NET PoP routers were the only pieces of equipment that needed a loopback address. Below is a list which details the addresses used.

The <SLA> for the loopback addresses is '0000'.

Table 12-5 Loopback Addresses

6NET PoP location	6NET PoP address space 2001:0798:<PoP>:<SLA= 0>::/64	IPv6 loopback address
Sweden	2001:0798:25::/64	2001:0798:25::1/128
Netherlands	2001:0798:22::/64	2001:0798:22::1/128
Germany	2001:0798:14::/64	2001:0798:14::1/128
Austria	2001:0798:10::/64	2001:0798:10::1/128
Italy	2001:0798:20::/64	2001:0798:20::1/128
Switzerland	2001:0798:12::/64	2001:0798:12::1/128
France	2001:0798:16::/64	2001:0798:16::1/128
United Kingdom	2001:0798:28::/64	2001:0798:28::1/128
Greece	2001:0798:17::/64	2001:0798:17::1/128

12.1.2.3 Point-to-Point Addressing

Intra-PoP Point-to-Point Links

Intra-PoP point-to-point links were numbered from the PoP prefix with a <SLA> = '01xx', where 'xx' is a sequence number. This allowed for 256 point-to-point links per PoP with a /64 prefix.

An example of intra-PoP point-to-point prefixes which have a <SLA>=01xx for the Germany PoP with <PoP>=14 would be:

2001:0798:0014:0100::/64

2001:0798:0014:0101::/64

2001:0798:0014:0102::/64

2001:0798:0014:0103::/64

6NET PoP to NREN PoP Point-to-Point Links

The 6NET core routers were connected to the NREN routers as previously described. A special address range was selected for this type of point-to-point connectivity:

2001:0798:<PoP>:<SLA=02xx>::/56, where 'xx' is a sequence number.

The addresses utilised on the point-to-point links between the 6NET core and the NREN PoP had a prefix-length /64. Initially there was only one NREN PoP router attached to a 6NET PoP router.

The host part of the address was '::1' for the 6NET core PoP side, and '::2' for the NREN PoP side.

Table 12-6 6NET PoP to NREN PoP Point-toPoint-Links

6NET PoP location	6NET PoP address space	1st Point-to-Point prefix 2001:0798:<PoP>:<SLA=0200>::/64
Sweden	2001:0798:0025:0200::/56	2001:0798:0025:0200::/64
Netherlands	2001:0798:0022:0200::/56	2001:0798:0022:0200::/64
Germany	2001:0798:0014:0200::/56	2001:0798:0014:0200::/64
Austria	2001:0798:0010:0200::/56	2001:0798:0010:0200::/64
Italy	2001:0798:0020:0200::/56	2001:0798:0020:0200::/64
Switzerland	2001:0798:0012:0200::/56	2001:0798:0012:0200::/64
France	2001:0798:0016:0200::/56	2001:0798:0016:0200::/64
United Kingdom	2001:0798:0028:0200::/56	2001:0798:0028:0200::/64
Greece	2001:0798:0017:0200::/56	2001:0798:0017:0200::/64

Inter-PoP point-to-point links were numbered from the Core prefix (2001:0798:00::/48) with a <SLA> = '00xx' , where 'xx' is a sequence number. For the initial rollout of 6NET, this resulted in the following table:

Table 12-7 Point-to-point links between 6NET PoPs

Connectivity Between	IPv6 prefix	IPv6 address side (1)	IPv6 address side (2)
UK (1) - FR (2)	2001:0798:0:1::/64	2001:0798:0:1::1/64	2001:0798:0:1::2/64
FR (1) - CH (2)	2001:0798:0:2::/64	2001:0798:0:2::1/64	2001:0798:0:2::2/64
CH (1) - IT (2)	2001:0798:0:3::/64	2001:0798:0:3::1/64	2001:0798:0:4::2/64
IT (1) - DE (2)	2001:0798:0:4::/64	2001:0798:0:4::1/64	2001:0798:0:4::2/64
DE (1) - NL (2)	2001:0798:0:5::/64	2001:0798:0:5::1/64	2001:0798:0:5::2/64
NL (1) - UK (2)	2001:0798:0:6::/64	2001:0798:0:6::1/64	2001:0798:0:6::2/64
UK (1) - SE (2)	2001:0798:0:7::/64	2001:0798:0:7::1/64	2001:0798:0:7::2/64
SE (1) - DE (2)	2001:0798:0:8::/64	2001:0798:0:8::1/64	2001:0798:0:8::2/64
DE (1) - AT (2)	2001:0798:0:9::/64	2001:0798:0:9::1/64	2001:0798:0:9::2/64
AT (1) - CH (2)	2001:0798:0:a::/64	2001:0798:0:a::1/64	2001:0798:0:a::2/64
DE (1) - GR (2)	2001:0798:0:b::/64	2001:0798:0:b::1/64	2001:0798:0:b::2/64

12.1.3 Naming Scheme

<cc>.6net.org

The domain name chosen for the 6NET backbone network was short and easy to remember. It was applied to all equipment that was considered part of the native 6NET core environment.

12.1.3.1 PoP Naming Convention

Every PoP had its own subdomain within 6NET. The subdomain name corresponded to the top level domain name of the country where the PoP was located. The following names were defined domains:

at – Austria

be – Belgium

ch – Switzerland

cz – Czech Republic

de – Germany

es – Spain

fr – France

gr – Greece

hu – Hungary

ie – Ireland

it – Italy

lu – Luxemburg

nl – The Netherlands

pl – Poland

pt – Portugal

se – Sweden

si – Slovenia

sk – Slovakia

uk – United Kingdom

All domain names in the 6NET backbone belonged to equipment located in the above PoPs, thus all the names had a suffix from the list and were prepended to the common suffix '6net.org'.

12.1.3.2 Router Naming Convention

The Router name was comprised of the 'domain name', the 'PoP name' and a sequential ordinal number. The equipment ordinal numbers started with 6. The possible extra routers would be called uk61, uk62 and so on.

Example:

The 6NET core router in the United Kingdom: <uk6.uk.6net.org>

12.1.3.3 Interface Naming Convention

Every IP interface (either physical, virtual or ATM-subinterface) had its own name, with the suffix corresponding to the equipment the interface belonged to, and a distinguished name describing the interface. The interface ordinal numbers corresponded to the ones in the equipment configuration and usually started with 0.

Overview:

Loopback (2 or more):	lo0 and lo1
Fast-Ethernet in some VLANs (1 or 2):	fe0, fe1
Giga-Ethernet in some VLANs:	ge0

POS or ATM interfaces for connections to other pops, called by the PoP name on the opposite side of the network link: the interface in the UK router connecting to the FR PoP is called fr.uk6.uk.net.sixnet.org.

Example:

Loopback interface on the United Kingdom router: <lo0.uk6.uk.6net.org>

12.1.3.4 Equipment Naming

The naming convention as seen above needs to be applied to the initial equipment utilised within 6NET. The application of the convention provides the following initial equipment names:

Core PoP routers:

Sweden:	se6.se.6net.org
The Netherlands:	nl6.nl.6net.org
Germany:	de6.de.6net.org
Austria:	at6.at.6net.org
Italy:	it6.it.6net.org
Switzerland:	ch6.ch.6net.org
France:	fr6.fr.6net.org
UK:	uk6.uk.6net.org
Greece:	gr6.gr.6net.org

12.1.3.5 Interface Naming

Each interface as seen in the naming convention had its own name.

Example: For the Austria 6NET core router

Fast Ethernet:	fe0 fe0.at6.at.6net.org
Gig Ethernet:	ge0 ge0.at6.at.6net.org
Loopback:	lo0 lo0.at6.at.6net.org
POS from Austria to Germany:	de.at6.at.6net.org

12.1.4 DNS

The DNS for the 6NET core network was operated by DANTE serving the following zones:

Forward zones:

- se.6net.org
- nl.6net.org
- de.6net.org
- at.6net.org
- it.6net.org
- ch.6net.org
- fr.6net.org
- uk.6net.org
- gr.6net.org

Reverse zones:

- 0.8.9.7.0.1.0.0.2.ip6.int
- 1.8.9.7.0.1.0.0.2.ip6.int
- 0.8.9.7.0.1.0.0.2.ip6.arpa
- 1.8.9.7.0.1.0.0.2.ip6.arpa

For the core network SWITCH and SURFnet are operated the slave DNS servers. SURFnet operated the central DNS server for the 6net.org zone.

12.1.5 IGP Routing

The options available for an IGP routing protocol within 6NET were: static routing, RIPv6 or IS-IS. The option ‘static routing’ would have been very difficult to manage in practice, and would not have been very scalable. The dynamic routing options available were ‘RIPv6’ and ‘IS-IS’.

RIPv6 is a distance vector routing protocol while ‘IS-IS’ is a link-state protocol. Although a distance vector routing protocol is easier to troubleshoot and the operation simpler to understand, it was preferred to utilise a link-state protocol due to its advantages in convergence, tuning and additional

features (like opaque information, enhanced TLV (Type/Length, Value) information for Traffic Engineering, etc.).

Integrated IS-IS was used only to distribute the core router reachability. For everything else BGP4+ was used. No route exchange was used between IGP and EGP.

The 6NET ISIS topology is depicted in Figure 12-2.

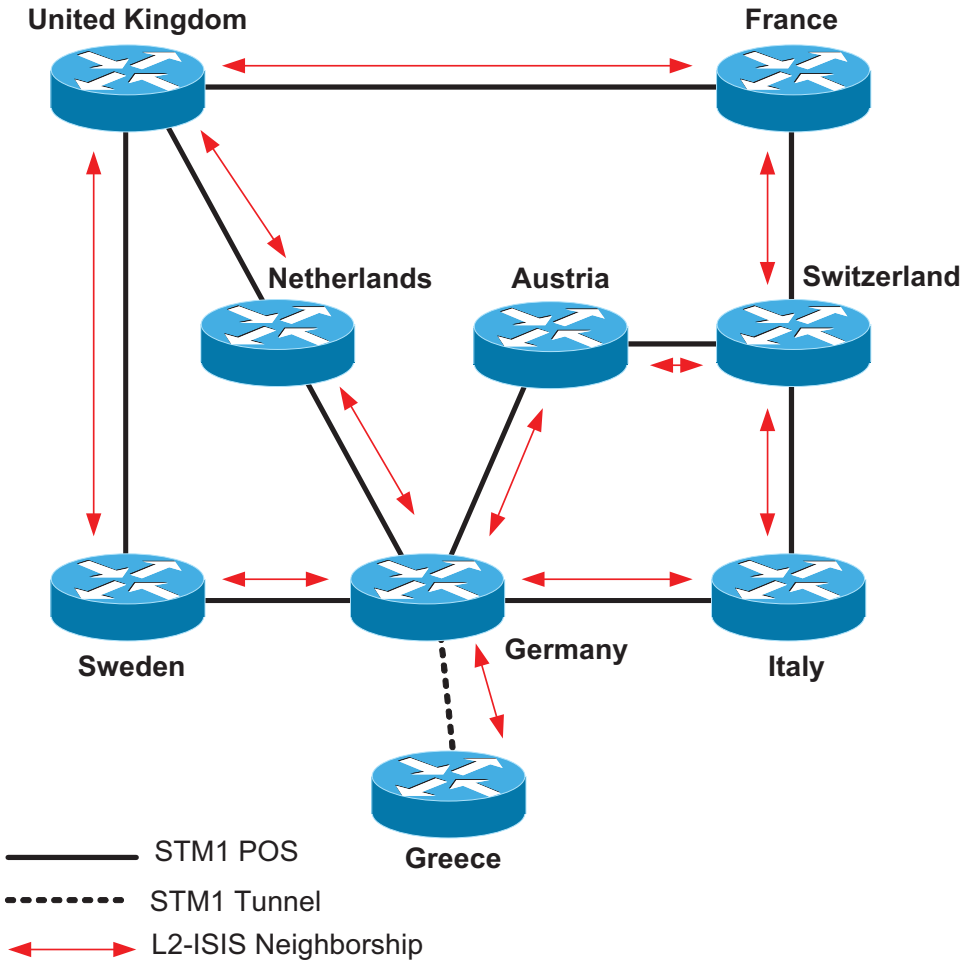

Figure 12-2 6NET IS-IS Topology

12.1.5.1 IS-IS structure

The hierarchical structure of ISIS enabled us to utilise two different levels of routing, allowing the routing protocol to scale through summarisation. The 6NET Core backbone initially consisted of a relatively small amount of routers and prefixes therefore, within 6NET a pure Level-2 routing domain was implemented with single area.

12.1.5.2 IS-IS authentication

Within the IS-IS routing protocol there are three basic password authentication mechanisms possible. The only authentication used was a password to authenticate the IIH packets for preventing accidentally forming wrong IS-IS adjacency.

12.1.5.3 IS-IS interface metrics

In order for a routing protocol to decide the best path from a one network to another it is important to assign a cost or metric to each interface. Typically, the metric is a direct correlation between value and the speed of the link. Other correlation can be delay, reliability, etc. Primarily, the link speed was used for the 6NET core according to the following table:

Table 12-8 Link Speed IS-IS Metric

Link Speed (Mbps)	IS-IS Metric
1	10000
10	1000
100	500
155	400
1000	300
2500	200
10000	100

12.1.5.4 Passive Interfaces

When a router interface is placed in the routing process it will send out IS-IS Hello packets in order to form an IS-IS neighbourship with peering IS-IS routers over that interface.

There are instances where it is desirable to include an interface in the IS-IS routing process to distribute the IPv6 prefix among the Level-2 area, but to never form an adjacency over that interface. To disable the sending of IIH packets the interface can be made passive with the command <passive-interface> under the IS-IS routing process. It is recommended to do this on stub LAN network interfaces.

In 6NET, all loopback interfaces and interfaces connecting to the NREN PoP routers and LAN stub interfaces were made passive.

12.1.5.5 6NET CLNS Addresses

The 6NET Core consisted of 9 routers that formed the International-backbone.

The following CLNS addresses were placed on the 9 routers:

Table 12-9 CLNS Addresses

Country	<domain>	<System ID>	<NSEL>
Sweden	49.0001	0025.0000.0001	00
Netherlands	49.0001	0022.0000.0001	00
Germany	49.0001	0014.0000.0001	00
Austria	49.0001	0010.0000.0001	00
Italy	49.0001	0020.0000.0001	00
Switzerland	49.0001	0012.0000.0001	00
France	49.0001	0016.0000.0001	00
United Kingdom	49.0001	0028.0000.0001	00
Greece	49.0001	0017.0000.0001	00

12.1.6 EGP Routing

At the time of deployment, BGP4+ was the only EGP supported for IPv6 when using Cisco IOS therefore the EGP used in 6NET was BGP4+.

12.1.6.1 Router-ID

When starting up a BGP process in an IPv6 only environment it is required to define a router-id on the BGP4+ router. The router-id has the same format as an IPv4 IP address.

For the 6NET environment the BGP router-id was identical to the initial IPv4 address placed on the Ethernet port for the initial management of the equipment.

12.1.6.2 eBGP

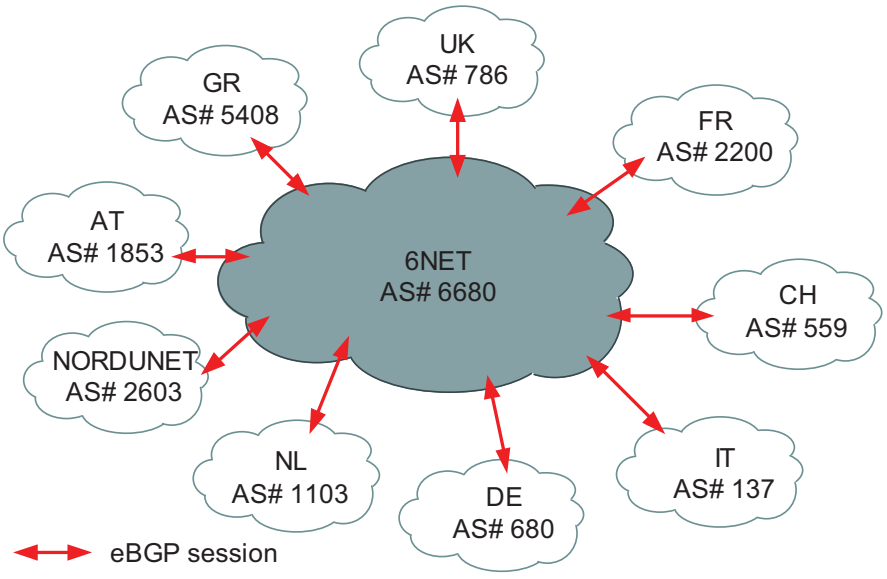

Figure 12-3 BGP peering with 6NET participants

BGP4+ was used between 6NET core and participant NRENs, 6NET and other organisations. Each NREN used their AS numbers. The 6NET core had a separate AS number from GÉANT as can be seen in Figure 12-3:

Neighbours with the same update policies can be grouped into peer groups to simplify configuration and to make updating more efficient.

Two initial peer groups were defined:

- Peer group for peerings between 6NET core PoP and NREN PoP:
  - 6NET_EXTERNAL_NREN_PEER
- Peer group for peerings between 6NET core PoPs:
  - 6NET_INTERNAL_6NET_PEER

12.1.6.3 iBGP

No Route-reflector was used due to the relatively small number of core routers. Instead full mesh BGP peering was configured inside the core. BGP peerings were secured with TCP MD5 authentication from the beginning to prevent the wrong peering by accident.

The external peers were grouped together to have same routing policy internally, with the groupname: 6NET_INTERNAL_NREN_PEER

The Next-hop-self functionality was not used in the 6NET rollout. The initial next-hop attribute checking of the NLRI was based on the IPv6 address from the NREN to which 6NET was peering.

12.1.6.4 BGP filtering

The complete 6NET IPv6 address range could be aggregated to ‘2001:0798::/40’. However, due to the aggressive filtering policies on the IPv6 Internet (see RFC 2772 [RFC2772]), initially the address range was aggregated to ‘2001:0798::/35’. To do this type of aggregation in a controllable way, some NLRI filtering was required. To enable the aggregation, 3 basic steps were required:

1. Configure static IPv6 route for ‘2001:0798::/35’ to the null0 interface
2. Redistribute the static summary route to the IPv6 BGP table (this required NLRI filtering in order to redistribute only the summary prefix)
3. Use BGP NLRI filtering to the iBGP peers to exclude the summary prefix from being exchanged.

12.2 SURFnet Case Study (Netherlands)

SURFnet is the national computer network for higher education and research in the Netherlands. SURFnet connects the networks of universities, colleges, research centres, academic hospitals and scientific libraries to one another and to other networks in Europe and the rest of the world. The SURFnet network enables its users to communicate with other network users and to consult the Internet from their office or their home PC.

In the summer of 2001 the national research infrastructure of SURFnet, called SURFnet5, was constructed as part of the GigaPort project and in partnership with BT Nederland and Cisco Systems. After a public procurement process these partners were awarded a six year contract at the very end of 1999, with BT Nederland supplying the infrastructure and Cisco Systems providing the router equipment.

The building of the successor to SURFnet5, called SURFnet6, started at the end of 2004 and is performed in partnership with Nortel, Avici Systems, and Telindus. SURFnet6 will be a hybrid optical and packet switching infrastructure delivering IPv4, IPv6, and lambda services to SURFnet's constituency. The IPv6 implementation for unicast and multicast on SURFnet6 will be native and truly dual-stack, i.e. not using MPLS.

12.2.1 The SURFnet5 Dual Stack network

The SURFnet5 network consists of a core that is situated at two locations in Amsterdam, at the Points of Presence (PoPs) called "Amsterdam1"[1] and "Amsterdam2"[2]. Two Cisco 12416 routers are installed at each location implementing the core functionality. Fifteen concentrator PoPs are each connected to both core locations over two separate unprotected SONET/SDH framed lambdas running at 10 Gbit/s. The engineering of the lambdas is such that it ensures that one connection is always maintained in case of a single transmission or core router failure. At each of the fifteen PoPs a router of type Cisco 124xx is installed together with a Cisco 7507 router for the low speed, legacy connections at speeds up to 155 Mbit/s. The backbone of SURFnet5 makes use of IP-over-DWDM, using POS framing. BT owns and operates the DWDM transmission infrastructure. Figure 12-4 shows the logical topology of SURFnet5.

From the start of SURFnet5 all routers running Cisco's IOS have IPv6 implemented. During the early days of the network, IOS versions in the 12.0ST train were used in the GSR routers. During 2002 the transition to 12.0S was made. The current version, in March 2004, in the GSR routers is "IOS Version 12.0(26)S1" and in the 7500 routers is "IOS Version 12.2(14)S2".

SURFnet5 started as a dual-stack network and has been running as such until March 2004. During March 2004 IPv6 routing within the core of the network was migrated from dual-stack to 6PE. The reason for this migration was to achieve line rate IPv6 forwarding in the SURFnet5 network. SURFnet5 is a 10 Gbit/s network largely build on Engine4 line cards and these cards handle IPv4 on the fast path while IPv6 is handled by the processor of the line card. After a large replacement in 2003 of Engine2 line cards with Engine3 line cards, the edges of SURFnet5 became capable of handling IPv6 traffic at line rate. By implementing 6PE in the core of the network it was ensured that IPv6 traffic was also handled at line rate in the core and potential bottlenecks were removed.

[1] This PoP is located at SARA in the eastern part of Amsterdam.

[2] This PoP is located at Hempoint, a BT Nederland co-lo facility, in the western part of Amsterdam.

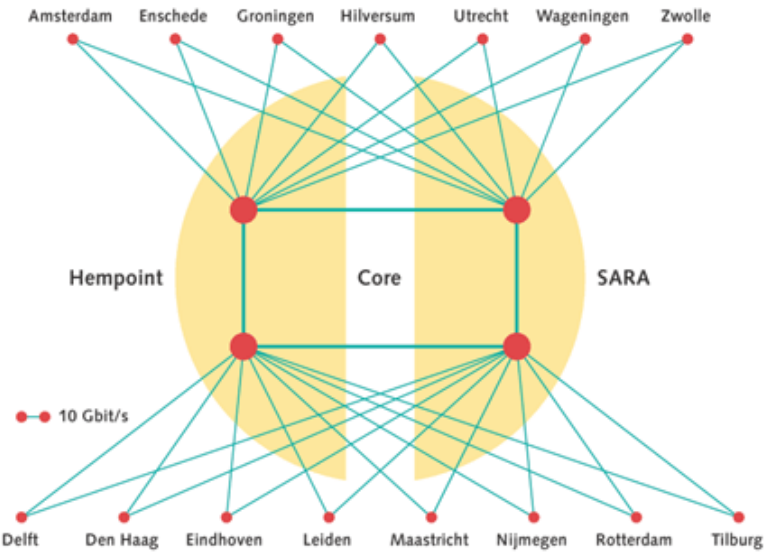

Figure 12-4 Logical topology for SURFnet5

The routing setup as implemented and running today runs without problems. A number of features are on SURFnet's requirements list, of which IPv6 multicast is high on the list. SURFnet already has hands-on experience with IPv6 multicast in the 6NET context using external MBGP for IPv6.

12.2.2 Customer Connections

Today, the vast majority of SURFnet's customer connections are at 1 Gbit/s using Ethernet technology. Three customers are connected at 10GE, connected to line cards in the backbone that do not support IPv6 at wire speed. In SURFnet6, the number of connected organizations using 10GE is expected to rapidly increase. Wire speed 10 Gbit/s IPv6 will be one of the features implemented in SURFnet6.

SURFnet5 supports customer connections on IPv4 as well as IPv6. For IPv4 unicast and multicast are supported, while for IPv6 this is currently unicast only. IPv6 can be delivered towards SURFnet's 180 customers in the field of research and higher education using either a native or tunnelled approach.

Native IPv6 connections are implemented either "dual stack" or "on a dedicated port". Dual stack implies that the customer connection runs IPv4 as well as IPv6 on the same local loop. A few customers, however, prefer IPv6 on a dedicated port, next to their IPv4 port for the simple reason that they have a dedicated router for IPv6 on their LAN.

Tunnelled connections to customers are only allowed in case the customer is not able to implement IPv6 in dual stack mode or on a dedicated port. In case we implement a tunnel connection with a customer, we always ask the customer to make plans for going native.

12.2.3 Addressing plan

SURFnet received an official IPv6 prefix from the RIPE NCC in August 1999 of length /35, which was enlarged to a /32 during the summer of 2002:

2001:0610::/32[1]

[1] For registration details on this prefix, see: http://www.ripe.net/cgi-bin/whois?2001:0610::/32.

This prefix is split up as follows:

Table 12-10 SURFnet Prefix

3 bits	13 bits	16 bits	3 bits	5 bits	8 bits
FP	TLA	sub-TLA	slow start	PoP (NLA 1)	site (NLA2)
/3	/16	/29	/35	/40	/48

Only the first /35 is currently used. The other 7 /35s are for future use right now. For each PoP a /40 prefix (NLA 1) is reserved and each /40 is again split into 256 /48 prefixes (NLA 2) to number the backbone links and customer networks. The breakdown of SURFnet's prefix on a per-PoP basis is shown in detail below.

The prefix 2001:0610::/40 is in use as the carrier network for SURFnet5, in which all links reside. For the links SURFnet uses /64 prefixes, in the format:

2001:0610:00xx:xyyy::zzz

In this format, xxx denotes the third octet of the link's or LAN's IPv4 prefix and yyy denotes the fourth octet of the link's or LAN's IPv4 prefix. The zzz denotes the actual address. An example of this is shown below in Table 12-11.

Table 12-11 SURFnet prefixes per POP

Prefix	NLA 1	PoP
2001:0610:0000::/40	0	(not used yet)
2001:0610:0100::/40	1	Amsterdam1
2001:0610:0200::/40	2	Amsterdam2
2001:0610:0300::/40	3	Hilversum
2001:0610:0400::/40	4	Leiden1
2001:0610:0500::/40	5	Utrecht
2001:0610:0600::/40	6	Chicago, USA
2001:0610:0700::/40	7	(not used yet)
2001:0610:0800::/40	8	(not used yet)
2001:0610:0900::/40	9	Delft
2001:0610:0A00::/40	10	(not used yet)
2001:0610:0B00::/40	11	Den Haag
2001:0610:0C00::/40	12	Rotterdam
2001:0610:0D00::/40	13	(not used yet)
2001:0610:0E00::/40	14	(not used yet)
2001:0610:0F00::/40	15	(not used yet)
2001:0610:1000::/40	16	(not used yet)

Prefix	NLA 1	PoP
2001:0610:1100::/40	17	Eindhoven
2001:0610:1200::/40	28	Maastricht
2001:0610:1300::/40	19	Nijmegen
2001:0610:1400::/40	20	Tilburg
2001:0610:1500::/40	21	(not used yet)
2001:0610:1600::/40	22	(not used yet)
2001:0610:1700::/40	23	(not used yet)
2001:0610:1800::/40	24	(not used yet)
2001:0610:1900::/40	25	Enschede
2001:0610:1A00::/40	26	Groningen
2001:0610:1B00::/40	27	(not used yet)
2001:0610:1C00::/40	28	(not used yet)
2001:0610:1D00::/40	29	Wageningen
2001:0610:1E00::/40	30	Zolle
2001:0610:1F00::/40	31	(not used yet)

12.2.4 Routing

During the build-up of SURFnet5, in the summer of 2001, IS-IS was used for IPv4 and IPv6. Since the transmission of SURFnet5, the 10G lambdas, is unprotected we used MPLS Traffic Engineering's extension Fast Re-Route as the protocol to protect the network against failures by making available an alternative path well within the 50 msec time frame. However, MPLS-TE was only supporting IPv4, yielding a situation in which the IPv4 topology was different from the IPv6 topology. At that time, IS-IS was not able to handle different topologies for the two protocols, and we had to move away from IS-IS for IPv6. We temporarily transitioned to RIPng for IPv6, while we stayed at IS-IS in combination with MPLS-TE FRR for IPv4.

In the mean time Cisco developed multi topology IS-IS, which enabled SURFnet to move away from RIPng back to the original plan. Before we went back to IS-IS for IPv6, we decided to abandon the MPLS-TE FRR ship for network management reasons, as we found the operations and maintaining of FRR too complex. At the same time, enhancements to IS-IS made it possible to fully rely on the routing protocols at the IP layer for the resilience in SURFnet5. During March 2004 the dual-stack approach was migrated to a 6PE implementation. The four core routers act as BGP route-reflectors in both the IPv4 as IPv6/6PE routing. All fifteen concentrator routers and border routers are BGP route-reflector clients hanging of these BGP route-reflectors.

12.2.5 Network Management and Monitoring

As a ground rule, SURFnet treats IPv6 on the network level equal to IPv4, as much as possible. This implies that procedures between SURFnet's NOC and SURFnet Network Services are streamlined to support this. Also, the external peering policy towards other non-European NRENs such as Abilene and CA*net 4 and towards the commodity Internet through SURFnet's upstream providers and through peerings at the Amsterdam Internet Exchange (AMS-IX) is the same for IPv4 and IPv6. At the

AMS-IX it means that IPv6 peering requests are handled exactly like IPv4 peering requests, as SURFnet's peering policy on the Exchange is identical for both protocols. In December 2004, SURFnet had approx. 65 IPv6 peering sessions enabled over the AMS-IX.

While SURFnet strives to be able to perform network management over IPv6 as well, today this is done only partly at this moment. SURFnet monitors the availability of the IPv6 customer connections as well as external connections. Also where software allows it (like SSH, Nagios and Rancid) IPv6 is used to manage the router equipment in SURFnet5.

12.2.6 Other Services

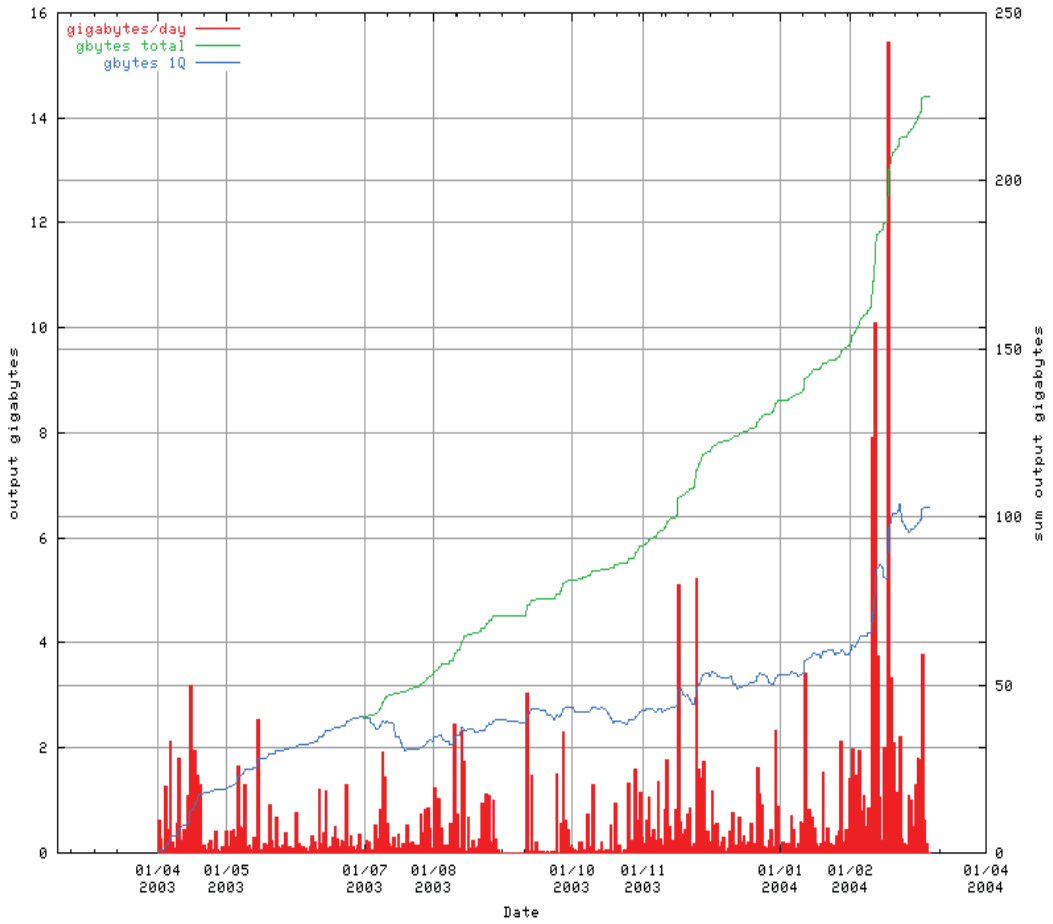

Figure 12-5 Anonymous-FTP over IPv6 volume

The following actions on applications and the like in the area of IPv6 are being or have been undertaken:

- The anonymous-FTP archive of SURFnet and NLUUG was enabled for IPv6 in the second quarter of 2002. See Figure 12-5 for the IPv6 volume transferred by this server.
- Four of SURFnet's Stratum 1 Network Time Protocol (NTP) servers are up and running and serving time over IPv6.

- All of SURFnet's three DNS server and three DNS resolvers are IPv6 aware as well as reachable over IPv6.
- Two Net News Transfer Protocol (NNTP) production feeders and three test feeders run IPv6 and have external peerings with Switch and Funet and other smaller peerings. In the test bed two news reader machines are used with IPv6 connectivity..
- The main web site of SURFnet[1] is enabled for IPv6.
- Several experimental services like video streaming (unicast, multicast) and internet telephony (SIP) run over IPv6 today.
- All SURFnet services are available over IPv6, either native or through a tunnel.

For new application services that are being developed and for which equipment or software is procured, we asked the potential suppliers on their readiness for IPv6.

[1] The English version is available at: http://www.surfnet.nl/en/.

12.3 FUNET Case Study (Finland)

In 2001, Funet's core network was upgraded to 2.5Gbit/s PoS links. Six Juniper routers (M20) were added to the previously all-Cisco network. There was not so much user demand for native IPv6 at this stage, but by the end of 2001, some plans for enabling dual-stack on these Juniper routers had been made.

During Q1/2002 various tests and trials were run.

After fixing all the issues, the Funet core had transitioned to complete dual-stack deployment in the core networks by Q2/2002. As of Q1/2005, dual-stack has worked well without any problems or performance impact.

12.3.1 History

The Funet experience involved some bleeding edge IPv6 deployment, leading to a subsequently stable service.

In Q2/2002, the minimum IPv6 working version 5.2R2 was installed on the Juniper routers. Later, versions 5.3R2, 5.3R3, and 5.4R3 were also used. Version 5.6R3 fixed a problem with IPv6 loopback access list breaking Neighbor Discovery, but since then (as of Q1/2005) no IPv6 issues have come up.

The IPv4 network used OSPFv2 and BGP for routing, but OSPFv3 for IPv6 was not ready. Funet didn't want to change the routing protocol, and for clarity, they wanted clearly separate IGP's for IPv4 and IPv6: OSPF and IS-IS. In addition, if IS-IS had been deployed for IPv4 and IPv6, multiple topologies would have to have been supported as IPv6 routing would have been different (in parts) from IPv4. So, IPv6-only IS-IS was clearly the best (and only, discounting RIPng) choice at that time.

Unfortunately, there were problems in the Juniper and Cisco devices with IPv6-only IS-IS. For example, the Juniper platform would not support IS-IS only for IPv6. The first, and worst, problem was that the Juniper routers would always also advertise IPv4 addresses used in loopbacks and point-to-point links. Cisco's IOS, unless you enabled IS-IS for IPv4 too (which was a non-starter for Funet), would discard all such attempts to form adjacencies: a total inoperability problem. Cisco's IS-IS implementation has an option 'no adjacency-check' to override this; however, an undocumented fact was that it would only work (at least in this case) when using level-2-only IS-IS circuit-type (which was not the default). A first step in interoperability was gained when these were enabled in IOS.

Some problems continued. IS-IS route advertisements from Ciscos to Junipers were accepted in the route database, but not put to the Junipers' routing table: this was caused by the above mentioned problem with adjacencies; this was reported, and fixed, in 5.2R2; a minimum workable version for Funet to use. The issue with Juniper always advertising IPv4 addresses in IS-IS was fixed, as (then undocumented) feature 'no-ipv4-routing' in 5.4R1. Also, one could not redistribute static discard routes to IS-IS (to generate a default route) until this was fixed in 5.3R3. You also could not set a metric when advertising a default route other than by redistributing a static discard route and applying a route-map in Cisco.

Fortunately, the rigorous tests in the lab network were enough to expose all of the above problems, which were fixed.

Also in 2002, most of Funet's Cisco routers were replaced by Junipers in an upgrade of the network.

In early 2003, Funet noticed significant bandwidth bottlenecks (in the order of dozens of megabits/second) with their remaining Cisco equipment - 7200's, 7500's, GSR's - regarding IPv6 forwarding capabilities, but the situation improved tremendously when software supporting CEFv6 come out later in 2003. This is very important for Funet as the IPv6 traffic level is at least 30-40 Mbit/s.

In early 2005, all IPv6 routing is done on production equipment; there are no longer any special "IPv6 routers", or any special IPv6-only links or connections. Almost all of these are Junipers (M10, M20, T320), about twenty, and two CSC's access routers are Ciscos (VXR with 12.2S software).

As of early 2005, Funet has not yet deployed IPv6 multicast support though the software capability has existed for a long time. Funet has waited for Embedded RP support (available in Junipers since 7.0R1 on Q4/2004), and availability of IPv6 multicast service from the upstream. It is expected that IPv6 multicast support is rolled out soon when the software becomes stable, within half a year or so.

Beyond that, no further IPv6 features have been identified missing in the core network.

The Funet IPv6 network is shown in Figure 12-6 (geography) and Figure 12-7 (topology).

Figure 12-6 Funet Network by Geography

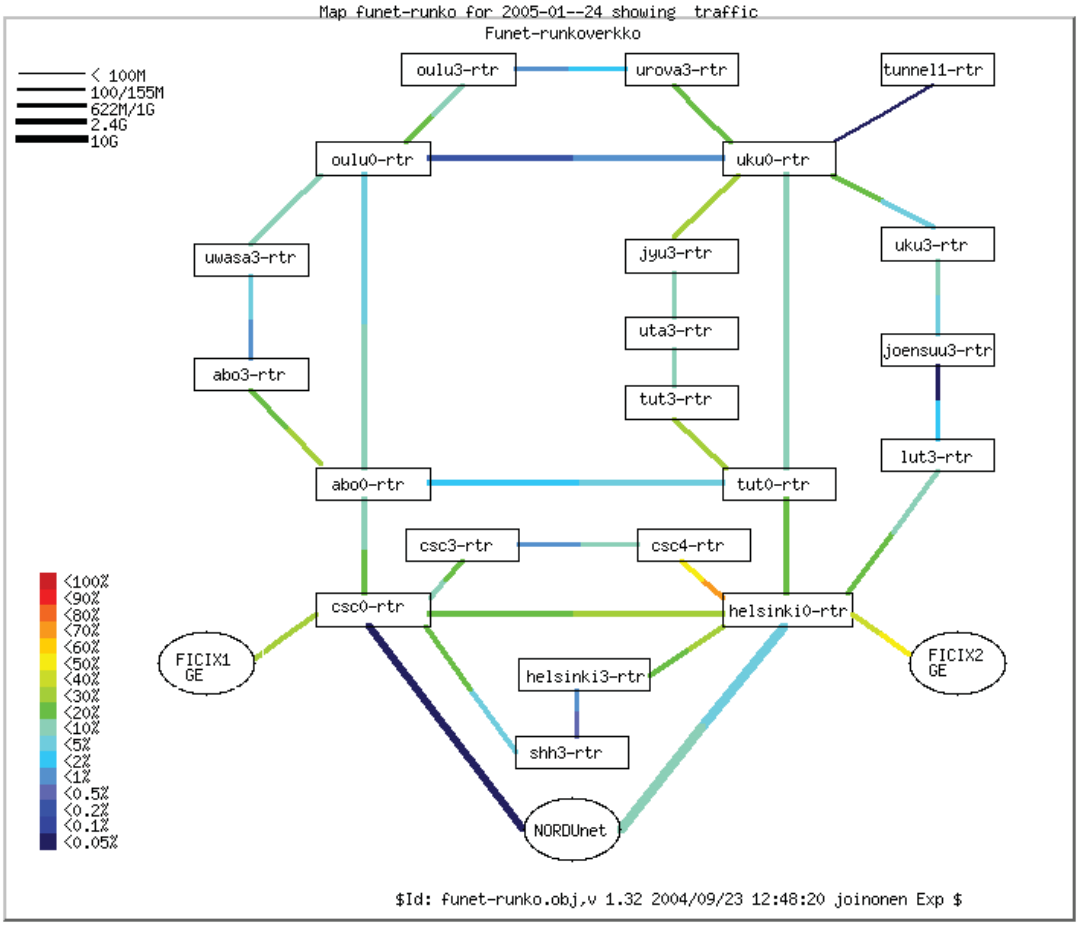

Figure 12-7 Funet Network by Topology

12.3.2 Addressing Plan

Funet has an IPv6 SubTLA address block of 2001:708::/32.

There are two different aspects to addressing: the customers and the network infrastructure.

12.3.2.1 Customers

There are six “SuperPoP” core routers; these are used when giving out /48 prefixes to customers, e.g.:

2001:708:0KLM::/48

“K” here is the number of the regional SuperPoP. This is not meant to be used for aggregation purposes, rather than to just create some mild hierarchy for the addresses, rather than allocate them in a sequential fashion.

“L” is the sequential number of customers within the SuperPoP. “0” is reserved.

“M” is reserved for expanding the number of customers in the SuperPoP and/or increasing the size of the assignment to the customers. To maintain future flexibility, this is still zero.

When assigning the prefix to the customers, Funet recommends they keep the first four bits (the first nibble) zero for now.

An example of a customer assignment is 2001:708:510::/48.

12.3.2.2 Network Infrastructure

Here, "KLM" has a slightly different meaning. "K" still means the SuperPoP, "L" means the PoP acting under the SuperPoP in an access network, and "M" is an identifier for the router in the PoP.

Loopback addresses are taken from the prefix:

2001:708:0:10:KLM::/112

So, a few loopback addresses will be:

tut0-rtr.funet.fi: 2001:708:0:10:300::1

tut1-rtr.funet.fi: 2001:708:0:10:301::1

uta3-rtr.funet.fi: 2001:708:0:10:313::1

These are configured using /128 prefix length. Note that the identifier of the router is also reflected in the naming ("uta_3_-rtr").

Point-to-point addresses in the core and the access networks are taken from one block - all addresses come from under a single /64:

2001:708:0:F000::/64

In particular:

2001:0708:0:F000::klmn:KLMm:z/112

klm and KLM identify the routers at the end of the point-to-point links, taken from the loopback addresses; in above, these would have been "300", "301", and "313". SuperPoP's or the smallest number goes first as klm. "n" and "m" are sequential numbers, used when necessary - for example if there are multiple links between routers which need to be numbered - defaulting to zero. "z" is the end-point of the point-to-point link: always "1" or "2". The same SuperPoP or smallest first rule applies here too. So, in consequence, the addressing becomes like:

uku0-jyu3: 2001:708:0:F000::4000:3230:[12]/112

uku0-oulu0: 2001:708:0:F000::4000:5000:[12]/112

uta3-jyu3: 2001:708:0:F000::3130:3230:[12]/112

The point-to-point links toward customers are always numbered from the customer's addresses, due to simplicity and policy reasons.

For peerings and miscellaneous use, a block of:

2001:708:0:F001::/64

is reserved.

In addition, some special use addresses are used inside 2001:708::/48, for example 2001:708::{1,2} (for a few routers), 2001:708::123 (NTP), 2001:708::53 (DNS) etc.

12.3.3 Routing

As was already noted, Funet network uses (as of Q1/2005) OSPFv2 for IPv4 and IS-IS (for IPv6 only) for IPv6. IPv4 infrastructure uses BGP with multiple (private) autonomous systems. Due to the

(relatively) low number of IPv6 customers and traffic, the same kind of BGP topology has not been built for IPv6. Instead, IPv6 BGP has only been set up at the border routers, and the rest use IS-IS.

A longer term idea has been to switch from OSPFv2 to using IS-IS for both IPv4 and IPv6, but this has been a low priority task, and as of this writing, has not been done yet. At the same time, the same kind of BGP topology would probably be built.

12.3.4 Configuration Details

In this section, core and customer (edge) configuration examples are listed.

12.3.4.1 Configuring the Core

The configuration of the Juniper routers is given below.

```
interfaces {
    lo0 {
        unit 0 {
            family iso {
                address 49.0001.1931.6600.5180.00;    "IS-IS address from IPv4"
            }
            family inet6 {
                address 2001:708:0:BB:eeee:ffff:0000:1111/128; "Loopback"
            }
            family inet {
                [...]
            }
        }
    }

    so-X/X/X {
    unit 0 {
            [...]
        family iso;                                "For IS-IS"
            family inet6 {
                address 2001:708:0:BB:aaaa:bbbb:cccc:dddd/112; "Core"
            }
            family inet {
                [...]
            }
        }
    }
}

protocols {
```

```
    isis {
        no-ipv4-routing;
        export ipv6-to-isis;
        level 1 disable;
        interface so-X/X/X.0 {       "Core connections"
        level 2 metric 2;
        }
    interface lo0.0 {
        passive;
    }
    }
}

policy-options {
    policy-statement ipv6-to-isis {
        from {
           protocol [ direct local static isis ];
           family inet6;
        }
        then {
            accept;
        }
    }
}
```

And respectively on Cisco:

```
interface POSX/Y
[...]
 ipv6 address 2001:708:0:BB:aaaa:bbbb:cccc:dddd/112
 ipv6 router isis
 isis circuit-type level-2-only
 isis metric 2
!
router isis
 passive-interface Loopback0
 !
 address-family ipv6
 redistribute static
 redistribute connected
 no adjacency-check
 exit-address-family
```

```
 is-type level-2-only
 net 49.0001.1931.6600.5181.00
 metric-style transition
 log-adjacency-changes
!
```

As can be seen, the model is such that all the routes of the router are redistributed in the IS-IS. An alternative approach would be to include all the interfaces in the IS-IS as passive interfaces.

This is not considered to have serious drawbacks, as none of these redistributed routes are advertised outside the autonomous system: the advertisement includes the aggregates only.

12.3.4.2 Connecting the Customers (edge)

Customers are connected using static routes. The configuration is very simple, like the below on Juniper:

```
interfaces {
    fe-X/X/X {
    unit Y {
        [...]
        family inet6 {
                rpf-check fail-filter RPF_FAIL_IPV6;
           address 2001:708:KLM:xxxx::2/64; # from the customer
        }
    }
    }
}
routing-options {
    rib inet6.0 {
        static {
            route 2001:708:KLM::/48 next-hop 2001:708:KLM:xxxx::1;
        }
    }
}

firewall {
    family inet6 {
        filter RPF_FAIL_IPV6 {
            term DEFAULT {
                then {
                    count count-rpf-fail-ipv6;
                    log;
                    discard;
                }
```

```
            }
         }
      }
}
```

And on Cisco this could look like:

```
interface Fa0/0
 [...]
 ipv6 address 2001:708:KLM:xxxx::2/64
!
ipv6 route 2001:708:KLM::/48  2001:708:KLM:xxxx::1
```

12.3.5 Monitoring

The addresses of all loopback and point-to-point addresses are entered into DNS in a special format. A script configured to allow zone-transfer of "ipv6.funet.fi" zone fetches this information and digs out the IP addresses which should be in use. The pinger periodically (once in five minutes) checks that the links (including the links to customers and the peers) are up and responding; if not, it sends an alert.

BGP and IS-IS adjacencies are also monitored using a tool which collects syslog warnings sent from routers to a central syslog server. If adjacencies or sessions flap, this can be noted in the monitoring page.

All routers and links are collected to a custom network map/monitoring tool, where the traffic levels and similar can be monitored easily.

A challenge in the dual-stack infrastructure is getting a feel how much traffic on the links is IPv4 and how much IPv6. As of this writing, there are no good mechanisms to get that. When IPv6 MIBs are complete and are implemented, getting such measurements may be easier.

The performance has not been rigorously tested, as the Junipers include a hardware IPv6 forwarding capability. When Cisco's performance issues were noticed, and CEFv6 was tested, Funet also briefly tested the backbone network's IPv6 forwarding capabilities; a PC with Gigabit Ethernet interface could not send enough traffic to cause any impact on the network, which was indeed what was expected.

12.3.6 Other Services

Funet has been using and advertising a 6to4 relay to everywhere (openly) since late 2001. This includes the advertisement of 2002::/16 and 192.88.99.0/24. The router in question is a FreeBSD system running zebra OSPF and BGP routing protocols.

IPv6 multicast has also been tested using separate infrastructure and tunnels, but this has been shut down as of Q1/2005, pending the upgrade of the core network to support IPv6 multicast.

A TCP/UDP relay (faith in FreeBSD) is used on server-side to experimentally enable IPv6 access to a few IPv4-only services.

An IPv6 newsfeed service is IPv6-enabled, and is generating 30+ Mbit/s of steady IPv6 traffic. Also, ftp.ipv6.funet.fi is also IPv6-operational.

12.3.7 Lessons Learned

It was noticed that especially if the network is built with Junipers, there is no need to worry about performance impacts, and building a dual-stack infrastructure in the core network is very easy and simple to maintain.

The more difficult part comes from getting the universities to use IPv6, either through tunnels or natively. This requires some education, but the main bottleneck is probably at the IT management at those organizations. The IT staff is typically overloaded with work, or otherwise reticent to start testing IPv6 or to provide that as a service to the university departments or other facilities.

As such, it seems that the researchers and departments, if they want to try out IPv6, should be more insistent in asking for it from their university's IT staff.

12.4 RENATER Case Study (France)

Within the framework of a pilot project by GIP RENATER, carried out by G6, the infrastructure of Renater2bis has been used to set up an IPv6 pilot network. The aim for the GIP RENATER was to begin to establish the means and mechanisms required to allocate the necessary resources to connect the test sites to the IPv6 pilot (NLA-ID allocation, reverse DNS delegation, IPv4 - IPv6 transition mechanisms, etc.)

The 2001:660::/35 prefix was used for addressing the pilot and connected academic sites. All industrial partners were addressed in the 3ffe:300::/24 address space. Dedicated ATM PVCs were used to transport IPv6 between the different IPv6 POPs. The original /35 prefix has been expanded to a /32 by the RIPE NCC (as part of common RIR policy) since the prefix was originally allocated.

12.4.1 Native Support

Thanks to this pilot experience, IPv6 is now offered as a native service on Renater3's backbone.

All Renater3's points of presence offer global IP connectivity to the regional networks and to the sites which contain both IPv4 unicast, IPv4 multicast and IPv6 unicast.

Both traffic types (IPv4 and IPv6) are carried in the backbone without any distinction, offering equal performance, availability, supervision and support levels. The Renater NOC was IPv6 trained to be able to achieve the same service for IPv6 and IPv4.

12.4.2 Addressing and Naming

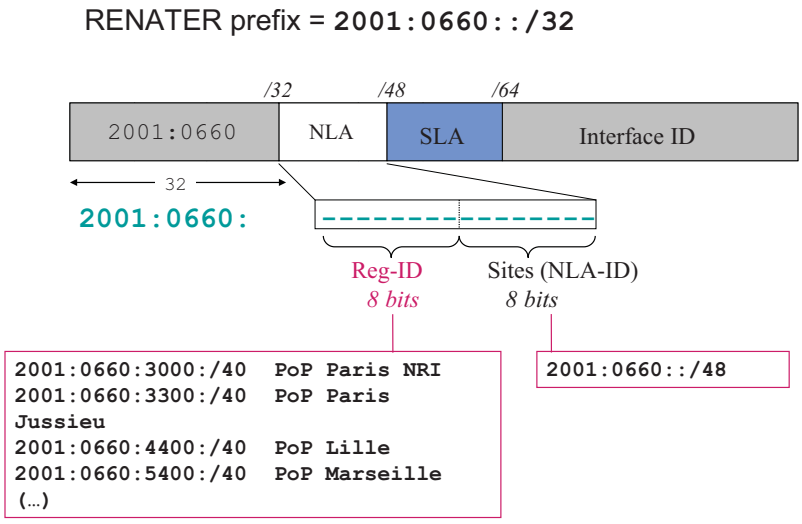

Figure 12-8 The Renater3 PoP Addressing Scheme

The addressing of the whole network is designed in a hierarchic way: this makes it possible to aggregate all routes, so reducing the routing table's size. Each point of presence of Renater3's backbone is allocated a /40 IPv6 prefix. Sites connected to a POP receive a /48 prefix derived from the /40 of the POP. For monitoring purpose, the /48 prefixes are not aggregated at the POP's in /40's. This way it is possible to have a view of sites routing announcements in the routing table. The addressing scheme is illustrated in Figure 12-8.

The GIP Renater manages the delegation of reverse zones of Renater3's SubTLA prefix (2001:0660::/32). It delegates to each site the reverse zone of the NLA-ID (/48) allocated to the site.

AFNIC, the French Network Information Center, is managing the .fr top-level domain name. They are connected to Renater3's Internet exchange point (SFINX) that supports IPv6.

12.4.3 Connecting to Renater 3

Using the experience gained with the IPv6 pilot of Renater2, procedures for connecting to Renater3 were designed and the teams were trained to be aware of the new processes. All the sites connected to the pilot have to be moved to Renater3, and be allocated a new prefix in the new address space. There was no D-day between the pilot and Renater3 as connectivity was not shut down for people connected through the pilot, to let them have time to do the procedures to connect to Renater3. Now all the sites have migrated to Renater3 and the pilot prefix is not used.

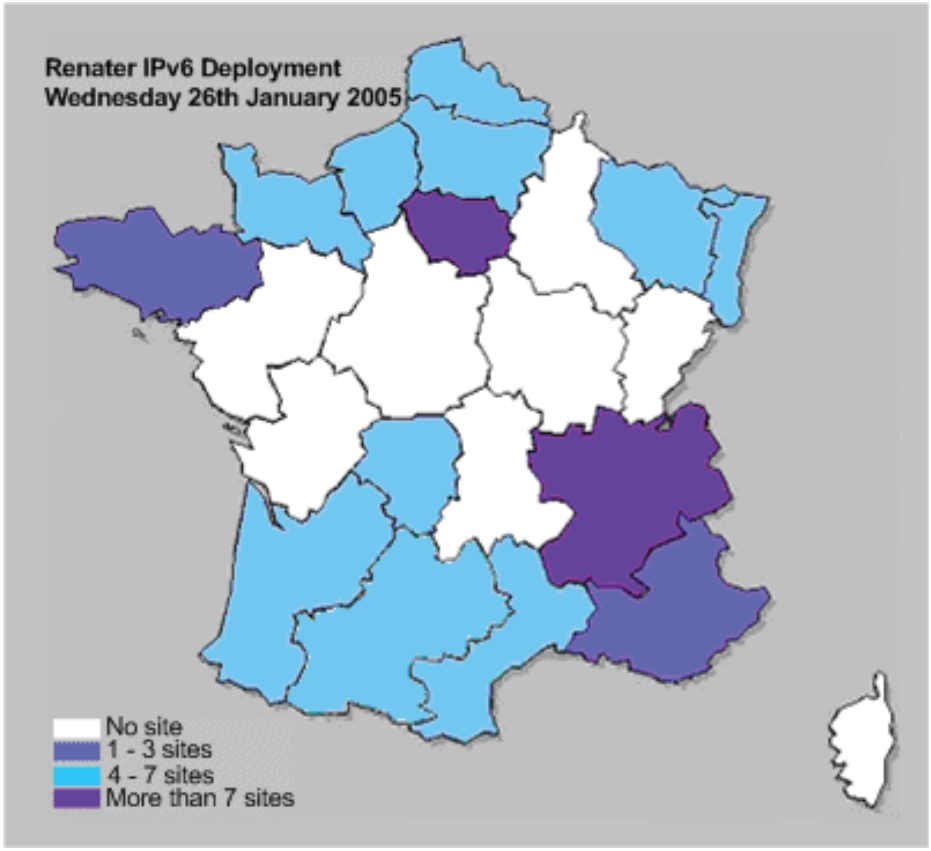

Figure 12-9 Density of IPv6 Connected Sites on Renater3

The procedures defined for IPv6 are very close to the ones defined for IPv4. This implies that a site can connect only if the network administrator fills some forms. An issue is that all these administrators are not IPv6 aware, and that some people in the site they manage need IPv6 connectivity. The prefix given to the site is aimed at addressing the whole network, and the administrator has to delegate a part of this prefix to the lab, which implies some IPv6 deployment forecast. This is not easy if the administrator is not IPv6 aware. This can lead to long delays to connect some sites to Renater3.

After 18 months, 50 sites are connected to the IPv6 service of Renater3 and 75 sites have received a prefix. The following map shows the density of IPv6 connected sites in the different French regions.

12.4.4 The Regional Networks

Renater3 is a national backbone with at least one POP in every French region. To connect to this POP, the sites use some access network (regional or metropolitan network). At the beginning of Renater3, none of these access networks were IPv6 enabled, meaning that the connection between the Renater3 POPs and the sites was done using tunnels or dedicated links (ATM PVCs, serial links, etc). As the core backbone is a Cisco GSR infrastructure, the choice was made not to set up tunnels on the core routers. Some dedicated equipment was deployed to concentrate the tunnels in the regions. Some of these routers are the ones used for the IPv6 pilot.

Eighteen months after the deployment of Renater3, there are 13 regional or metropolitan networks providing IPv6 connectivity to their customers, and 19 have received a prefix, so new networks should be connected soon. These networks used many different approaches for this IPv6 deployment: 6PE, VLANs, fully dual-stack, PVC ATM, tunnels, etc, and all these deployments are done in straight collaboration with Renater. It is clear that the deployment made in the core network is a real motivation for the other networks at the edges to follow.

12.4.5 International Connections

Renater3 offers IPv6 connectivity to national (RNRT) and Europeans (IST) projects. It is connected to GÉANT with a 10Gbps link where it exchanges traffic with the european NRENs. It has a connection to Asia via TEIN link, a 10Mbps ATM PVC is configured for IPv6, while the global link can offer 155Mbps. It also has an IPv6 connection to the transit network OpenTransit via one 2.5 Gbps link from the PoP of Lyon.

The IPv6 service is extended to the SFINX (Service for French Internet eXchange), which offers IP actors an interconnection point that carries both IPv4 and IPv6. IPv6 is exchanged in dedicated VLANs. This makes it possible to manage IPv6 traffic more easily. At this stage, 13 entities are connected to the SFINX using IPv6.

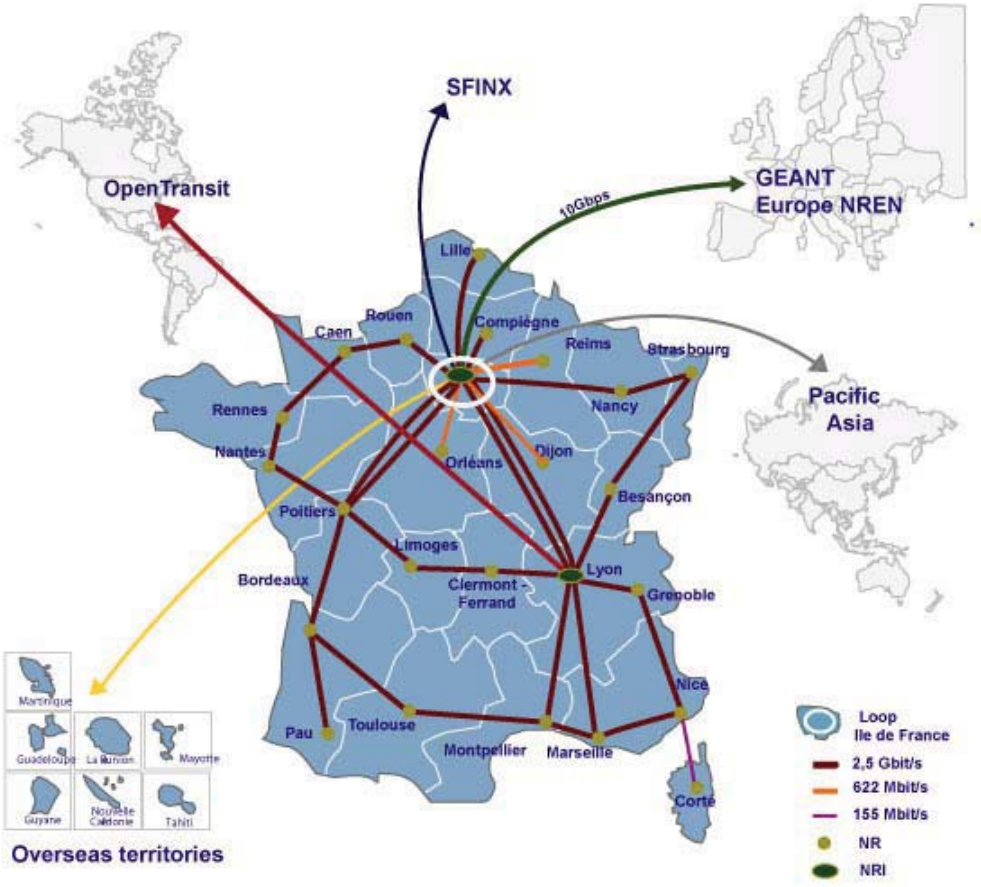

Figure 12-10 The Renater3 Network

12.4.6 Tunnel Broker Service Deployment

To overcome the lack of IPv6 in regional or metropolitan access networks, RENATER is deploying a tunnel broker service. The solution chosen is Hexago Migration Broker. Its main function is to connect sites or end-users who do not have IPv6 connectivity by creation of IPv6 dynamic tunnels. End-users can connect to the Tunnel Broker via a client software based on TSP (Tunnel Setup Protocol). This process will rely on MD5 authentication. The following scheme from Hexago shows the different use cases for the tunnel broker service.

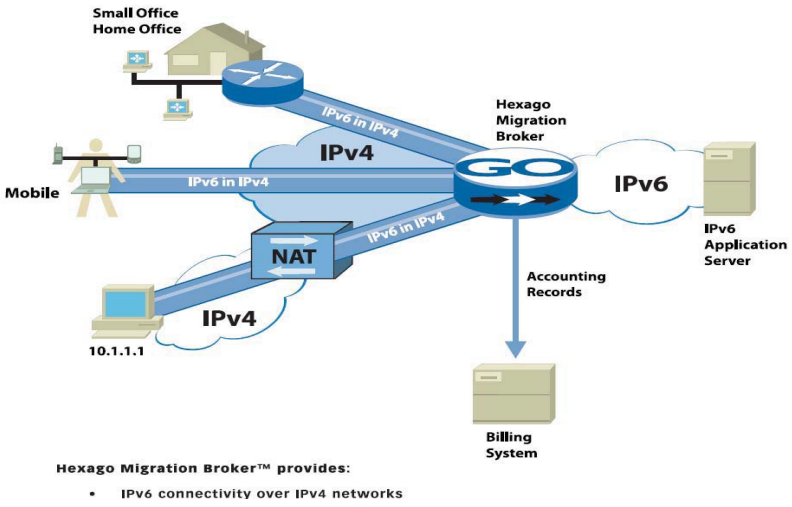

Figure 12-11 The RENATER Tunnel Broker

The tunnel broker will also be used for bringing IPv6 connectivity to meetings, conferences or any kind of event requiring IPv6 (IST, Launch events...).

12.4.7 Network Management

Being able to monitor the IPv6 service always stayed a priority since its deployment in Renater3 backbone. As it is offered now as a production service, same guarantees are required for IPv4 and IPv6.

12.4.7.1 Traffic monitoring

During the IPv6 pilot phase of Renater2, IPv6 was transported in dedicated ATM PVCs. The monitoring of IPv6 traffic was very easy as polling was made on separate interfaces. In Renater3, as IPv6 and IPv4 are transported on the same links and MIBs for monitoring IPv6 traffic are not yet implemented on Renater3 equipments, it is not possible to monitor IPv6 traffic the same way.

A script was developed to poll the routers using the CLI every 60 minutes. This feeds the Renater monitoring database and results can then be displayed using graphs or weathermaps.

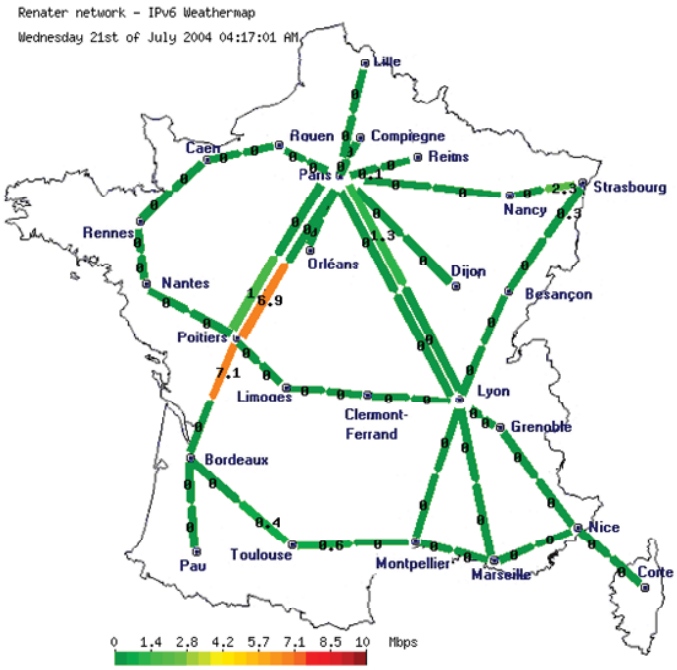

Figure 12-12 IPv6 Traffic Weathermap

12.4.7.2 Routing monitoring

Due to the lack of MIBs capable of monitoring IPv6 routing on Renater3 equipments, some scripts were deployed capable of checking the status of BGP4+ peerings. Alarms are sent when peering go down.

ASPath-tree was also deployed to give an overview of the routing policy in Renater. It cannot be used for operational routing monitoring but is useful for checking routing policies and make new routing plans.

12.4.7.3 Flow monitoring

At this stage, the routers on Renater3 are not capable of exporting IPv6 flows using Netflow v9. Nevertheless, a lot of work has been made to make the Netflow collector of RENATER capable of collecting these IPv6 flows. A first version of the collector is about to be submitted for tests to the interested 6NET partners.

12.4.8 IPv6 Multicast

An experimental IPv6 multicast network (M6Bone) is running on the Renater3 infrastructure. It allows the connection of lots of sites, all over the world. This network allows all the sites connected to test and develop IPv6 multicasting. It is connecting today over 80 sites and networks in Europe, Asia and Africa, making the network one of the most advanced multicast IPv6 network in the world.

12.5 SEEREN Case Study (GRNET)

The South East European Research and Education Networking (SEEREN) infrastructure interconnects the Research and Education Networks (NRENs) of Albania, Bosnia-Herzegovina, Bulgaria, FYR of Macedonia, Greece, Hungary, Romania and Serbia-Montenegro amongst themselves and to the European backbone network. In this respect, it constitutes the South-Eastern European segment of the multi-gigabit pan-European Research and Education network, GÉANT.

The SEEREN infrastructure was officially inaugurated in January 2004, with first IPv4 services provided in November 2003. Since then, the project has been developing several services and tools on top of the infrastructure, including a virtual network operations centre and a one-stop-shop for the management tools [SEEREN]. The deployment of IPv6 services has been planned since Spring 2004.

12.5.1 SEEREN Network

The SEEREN physical and logical network topologies are depicted in Figure 12-13 and Figure 12-14, respectively. The MPLS-enabled core network infrastructure, which is provided by a consortium of operators in SE Europe, has Points of Presence (PoPs) in all the SE European capital cities. Interconnection between the SEEREN NRENs is achieved with the Carrier supporting Carrier (CsC) technique (see Figure 12-15), which enables an MPLS VPN-based service provider, called the carrier provider, to allow other IP (or MPLS) service providers, called customers carriers (SEEREN NRENs[1]), to use a segment of its backbone network.

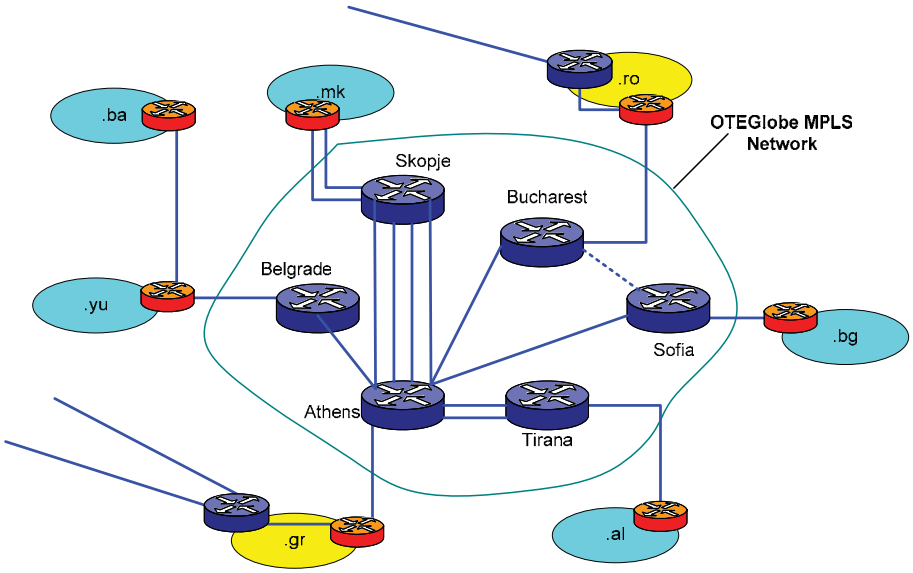

Figure 12-13 SEEREN Physical Network Topology

The access links of SEEREN NRENs range from 2Mbps up to 34Mbps, while connectivity is achieved by using diverse technologies from ATM to Ethernet over PPP (EoPPP). The SEEREN primary connection to GÉANT is a 95Mbps ATM PVC crossing GRNET. A secondary (backup) 34Mbps

[1] Each SEEREN NREN has deployed and manages one border router for its international connectivity. These routers consist of a "virtual" customer provider network.

ATM connection diverts traffic to RoduNET (Bucharest GÉANT PoP) in case of failure in the primary connection.

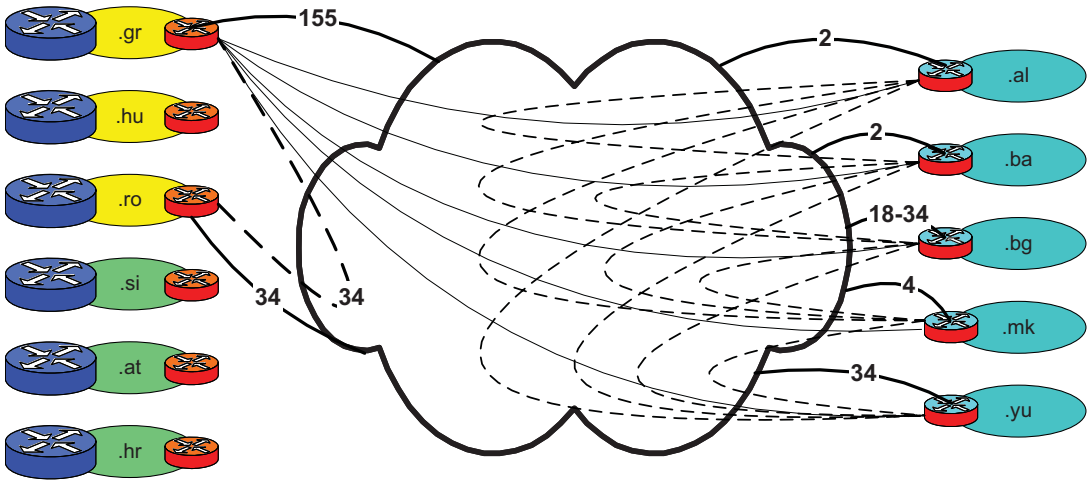

Figure 12-14 SEEREN Logical Topology

According to the CsC technique, in the data plane, the traffic exchanged between SEEREN NRENs border routers (CEs[1]) is encapsulated into MPLS frames and, then forwarded through the Carrier Providers MPLS network. The MPLS label is removed by the far-end along the path SEEREN CE router and forwarded as an ordinary IP packet. In the control plane, SEEREN CEs exchange with the Carrier Provider PEs only IGP routing information, i.e. CE loopback IP addresses and CE-PE interconnection subnets. The corresponding routing information populates a Carriers Provider virtual routing and forwarding (VRF) table. For each routing entry in the VRF a label is assigned by the carrier PE and advertised via eBGP or LDP to the directly connected SEEREN CE (see Figure 12-16).

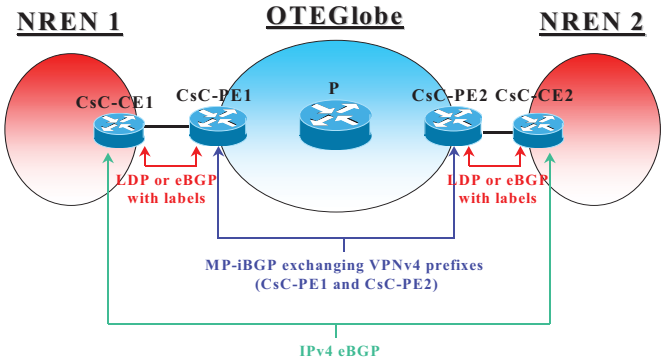

Figure 12-15 Label Exchange in the CsC Model

The CsC technique offers to the SEEREN NRENs multiple technical and economical advantages. SEEREN NRENs do not need to deploy, operate and maintain an international backbone infrastructure

[1] Each SEEREN NREN border router that is connected to the Carrier Provider's router is called customer edge (CE) router. The Carrier Provider's router connected to a Customer's router is called provider edge (PE) router.

but only need to exchange minimal routing information with its customers, e.g. only 15 routes are needed for the entire SEEREN network, even in cases where several instances of the entire Full Internet Routing Table is exchanged over the Provider's network. This allows SEEREN NRENs to define their routing policy (eBGP sessions) transparently over the carrier provider's network and the carrier provider to minimize the routing information stored in the VRF tables.

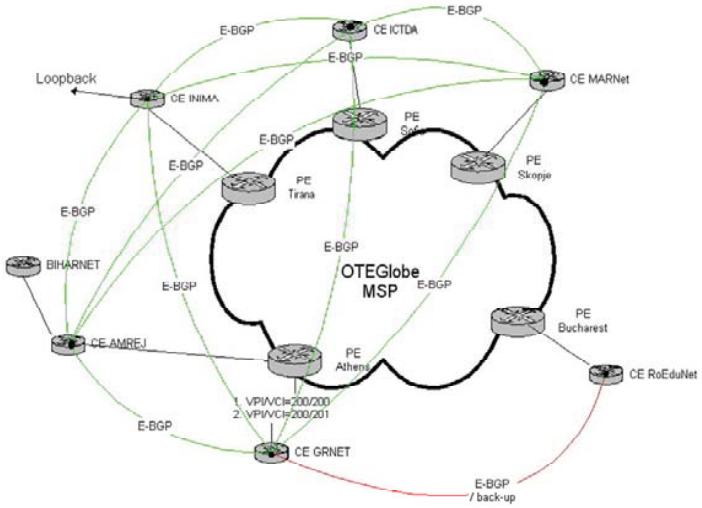

Figure 12-16 Routing Exchange in SEEREN

12.5.2 Implementation Details of CsC/6PE Deployment

The solution that was finally implemented in SEEREN infrastructure combined 6PE services at the NREN CE routers with transparent forwarding of IPv6 traffic over the Carrier Provider MPLS network. This solution did not require any software or hardware upgrades in the Carrier Provider network nor the creation of tunnels between the CE routers.

The 6PE deployment approach allows an ISP to support IPv6 services over an MPLS/IPv4 network. It has many technical implementation similarities to the MPLS VPN deployment solutions, where IP traffic is encapsulated into MPLS frames and sent over the MPLS core network.

At the SEEREN infrastructure, the NREN border routers act as 6PE routers and encapsulate IPv6 packets into MPLS frames. Concurrently, the NREN border routers act as CsC-CE routers and thus also encapsulate IPv4 (or IPv6) packets into MPLS frames. Consequently, the control and data plane of the CsC/6PE approach differs from the respective planes of either 6PE or CsC approaches.

In the CsC/6PE control plane, the 6PE routers (the NREN border routers) are dual stack, i.e. support IPv6 and IPv4 protocols, and communicate with the rest local NREN infrastructure with any common IPv6-enable routing protocol. The routing prefixes learned from the local networks are distributed among the 6PEs via multi-protocol MBGP (MP-iBGP). The BGP next-hop for each advertised IPv6 prefix derives from the IPv4 address of the connected 6PEs. Furthermore, the 6PE routers, now acting as CE routers in the CsC context, exchange routing and label information with the Service Provider (IPv4-only) PE routers. This process allows the 6PE routers to identify the labels for the IPv4 next-hop address for all destinations in the VPN, i.e. IPv4 addresses of CsC-CE routers. Finally, communication among CsC-PE routers is achieved via "pure" MPLS switching, where the (IPv4) IGP protocol is used for the distribution of appropriate MPLS labels for CsC-PE addresses.

In the data plane, IPv6 packets are sent from the local infrastructure routers to the NREN ingress 6PEs router. The latter imposes two labels on top of the IPv6 packet. The inner label identifies the IPv6 BGP next-hop and the outer label identifies the egress 6PE router. After the MPLS frame is received by the ingress CsC-PE router, the latter will replace the outer label in order to identify the VPN that the encapsulated packet belongs to. On top of the two labels, the ingress CsC – PE router will stack one additional label that identifies the MPLS-next hop along the path to the egress CsC – PE router. This label changes as the MPLS frame is switched inside the Service Provider MPLS core network. If MPLS penultimate hop poping operation is activated, the egress CsC-PE router receives an MPLS frame with two labels and removes the outer label as it forwards the frame to the 6PE router. Finally, the egress 6PE router removes the remaining label and forwards the IPv6 packet to the appropriate CE.