www.euro6ix.net

Title:	Document Version:
Deliverable D2.3 **Network Architecture Handbook**	0.8

Project Number:	Project Acronym:	Project Title:
IST-2001-32161	Euro6IX	European IPv6 Internet Exchanges Backbone

Contractual Delivery Date:	Actual Delivery Date:	Deliverable Type* - Security**:
31/07/2003	31/08/2003	R – PU

* Type: P - Prototype, R - Report, D - Demonstrator, O - Other
** Security Class: PU- Public, PP – Restricted to other programme participants (including the Commission), RE – Restricted to a group defined by the consortium (including the Commission), CO – Confidential, only for members of the consortium (including the Commission)

Responsible and Editor/Author:	Organization:	Contributing WP:
Sathya Rao	Telscom	WP2

Authors (organizations):

Alain Baudot (FRT&D), Riccardo De Luca (Telscom), Vladimir Ksinant (6WIND), Marcin Michalak (Telscom), Gabriel Lopez Millan (UMU), Mario Morelli (TILAB), Jordi Palet (Consulintel), Carlos Parada (PTIN).

Abstract:

This deliverable identifies key features to be considered in designing the correct networks with important features of high bandwidth, QoS, security and other key concepts researched in the Euro6IX project.

Keywords:

6to4, AAA, Addressing, Backbone, DNS, DSTM, Euro6IX, IPv6, IX, Monitoring, NAT-PT, Network Architecture, RADIUS, Routing, Routing Policies.

Revision History

The following table describes the main changes done in the document since its creation.

Revision	Date	Description	Author (Organization)
v0.1	04/04/2003	Document creation	Sathya Rao (Telscom)
v0.2	05/04/2003	ToC creation	Sathya Rao (Telscom)
v0.3	07/04/2003	Editorial updates	Marcin Michalak (Telscom)
v0.4	06/06/2003	Added 6WIND, TILAB, UMU, FRT&D, TELSCOM, PTIN contributions	Riccardo De Luca (Telscom)
v0.5	30/06/2003	Added 6WIND contribution and other add-ons	Riccardo De Luca (Telscom)
v0.6	09/07/2003	Several add-ons and corrections	Riccardo De Luca (Telscom) Marcin Michalak (Telscom)
v0.7	11/07/2003	Introduction, Conclusions	Riccardo De Luca (Telscom) Sathya Rao (Telscom) Marcin Michalak (Telscom)
v0.8	31/08/2003	Final Review	Jordi Palet (Consulintel)

Executive Summary

This deliverable represents the result of the work done in the context of the A2.3 activity inside the WP2 package of the Euro6IX project.

The main objective of this deliverable is to identify the key features to be considered in designing the IPv6 networks, taking into account the important features of high bandwidth, QoS, security and other key concepts researched.

In the first section (Chapter 2) there is an overall description of the Euro6IX backbone, dividing the connectivity issues into two main parts: The internal connectivity, meaning the interconnection of the different IXs and the external connectivity, meaning the interconnection with other IPv6 networks, such as 6NET, 6Bone, etc.

The second section (Chapter 3) contains the key network features for the implementation of the Quality of Service, Security, Multicast, Autoconfiguration and Multihoming in the IPv6 networks.

The third section (Chapter 4) provides a summary of the deliverable and draws up some conclusions.

Table of Contents

Table of Figures

n	16-n bits	64 bits
SLA1	Subnet	Interface ID

m	16-n-m bits	64 bits
SLA2	Subnet	Interface ID

1. INTRODUCTION

One of the objectives of the Euro6IX project is to research an appropriate architecture in order to design and deploy the first Pan-European non-commercial IPv6 Internet Exchange (IX) Network, connecting several regional neutral IPv6 Internet Exchange points across Europe. The network design will be based on the design of a native IPv6 network, which follows the hierarchical architecture of the global Internet by including:

- A set of regional native IPv6 Internet Exchanges.
- A core network to interconnect the exchanges.
- A second/access level of nodes, for ISPs.

The involvement of the major/incumbent European Telcos in this project, covering also the bigger areas that have a higher Internet user growth rate, reflects the strategic importance of the fast and timely introduction of IPv6 in Europe.

WP2 was related with the design of both backbone and internal networks following the design rules for commercial deployment: DNS management, addressing and routing, QoS, Bandwidth management, monitoring and operation.

D2.1 and D2.2 provided specification of internal network and backbone networks respectively. This deliverable, based on the Euro6IX network, has been developed as a result of these two deliverables and research carried out addressing general network design issues from network and service providers point of view.

The title of this deliverable is Network Architecture Handbook and is considered as a Network cookbook identifying key features to be addressed in designing the correct networks with important features of high bandwidth, QoS, security and other key concepts researched in the project.

The deliverable is organized in 2 major sections addressing internal connectivity and external connectivity.

After having described the details of such a network, we defined the main network features such as Quality of Service, Security, Multicast, Autoconfiguration and Multihoming.

2. THE EURO6IX BACKBONE

At the very highest level there are two types of geographically large networks:

- Homogeneous networks owned and operated by a single entity; Worldcom's global network being an example.
- Separate independently owned and operated networks that cooperate; the Internet being an example.

The Euro6IX network falls in the second category with a number of Internet Exchanges being owned and operated by separate organizations but which have agreed to cooperate to form a geographically larger network i.e. the Euro6IX network.

As such the Euro6IX project does not have a routed backbone network but rather a number of links that interconnect the cooperating Internet Exchanges. The Euro6IX network is therefore a consortium of cooperating Internet Exchanges interconnected by a number of links. Like any consortium there needs to be a set of guidelines that the partners adhere to. In the case of Euro6IX these guidelines need to cover:

- Acceptable usage policy – to define the type of traffic to be exchanged.
- Peering policies – to control the route traffic takes.
- Interface guidelines – QoS, multicast, etc.
- Experiments – provider independent addressing, etc.

In the following section we describe the network architecture dividing it in Internal and External connectivity. With "internal connectivity" we mean the interconnection of different IPv6 Internet Exchanges (IXs), hosted by the Telcos and links among IXs, sponsored by the Telcos and forming a ring topology as shown in the figure 2-1. With "external connectivity" we mean the links between the Euro6IX Network and other networks like 6NET, 6Bone and so on.

2.1 Internal Connectivity

2.1.1 Network Topology

The internal network architecture of Euro6IX is basically made by Internet IPv6 Exchanges (6IX or IX from now on). Several European Telcos of Euro6IX consortium provide the necessary infrastructure to deploy and place the IXs.

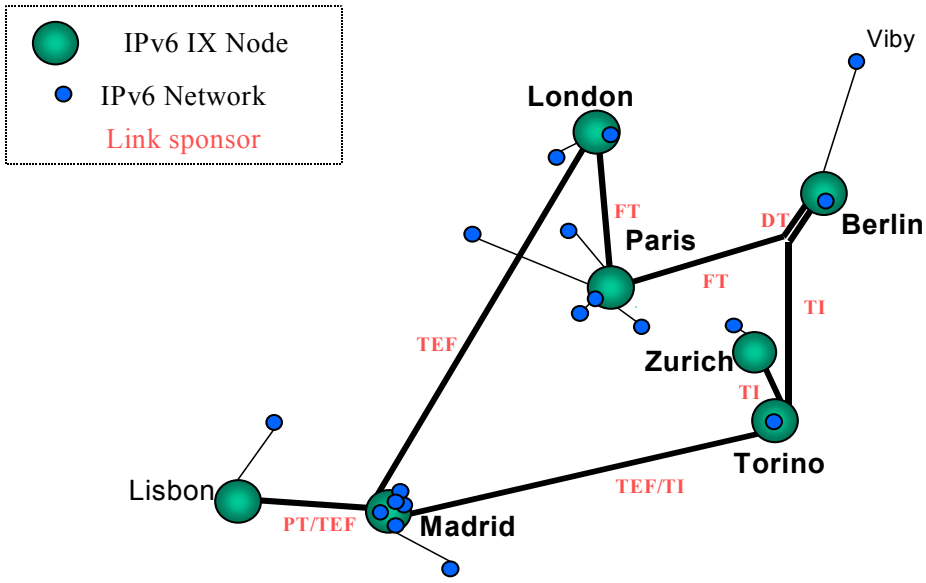

Figure 2-1: Euro6IX Internal Connectivity

2.1.2 Layer-2 Issues

This chapter describes the Layer 2 (OSI Data Link Layer) architecture that should be implemented in each IX node. The Layer 2 provides basic connectivity between the equipments of an IX while allowing logical separation between distinguishable functional blocks.

Each IX will consist of a Layer 2 Switched Network providing connectivity between the main IX blocks: IX Backbone Routers, Traditional IX LAN, Transitions/Access Facilities, Services and Monitoring/Management network. Traditional IX Customers can connect to the IX using a logical separated segment in the Euro6IX IX node.

The following picture describes the L2 main blocks.

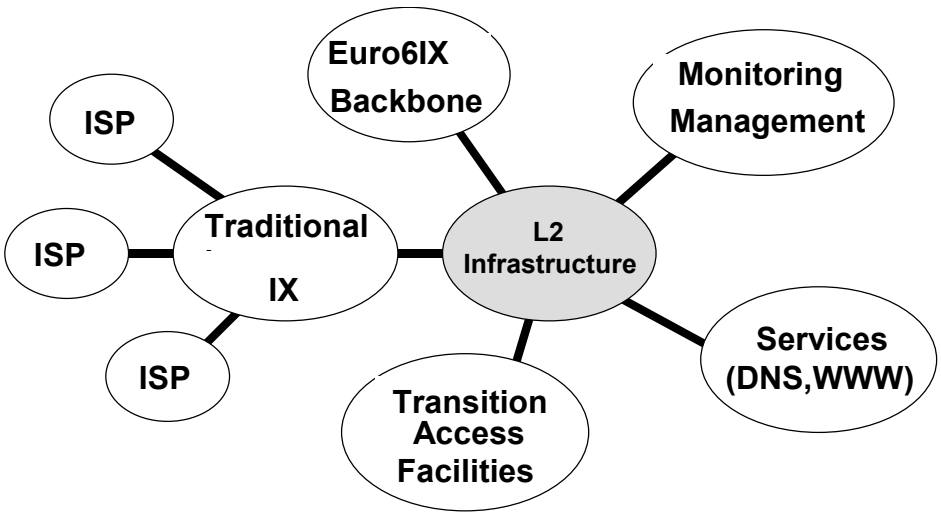

Figure 2-2: L2 Main Blocks

According to each IX node particularities, a Euro6IX node can have some or all the components defined. Anyway the basic and common components will be the Euro6IX Backbone Access and the Traditional IX or the Access Facilities Block depending on if the IX has traditional IX facilities or only allows Euro6IX backbone access.

The Switched Network can be built based on two main technologies: Ethernet or ATM. Ethernet is a broadcast, mature, cheap and well-understood technology. ATM is a point-to-point technology and holds the promise of scaling well beyond 100 Mbps to Gigabit rates, while still delivering QoS guarantees. The development of Layer 2 switching in hardware with technologies such as virtual LANs (VLANs) allows logical separation between segments and will provide increased scalability being the most recommended technology for L2 in an IX node.

In the following section, it will be provided a step-by-step description that allows the implementation of an IX node defining the technologies and the architecture for L2 connection of the main blocks previously described.

2.1.2.1 Basic Configuration- Layer 2 Infrastructure

Euro6IX IX is supposed to provide services collocated in the IX facilities: Either Traditional IX Service or Euro6IX Access Services using either Native IPv6 Dial-Up, IPv6 over IPv4 Tunneling or using IPv6/IPv4 Transition mechanisms. Typically, dedicated equipments will provide these services so connectivity between them is needed. The easiest and cheapest way to establish local connectivity is using Ethernet technology.

Each Euro6IX IX node should consist of at least two Ethernet switches supporting local LAN segments. The IX Backbone Router will have local connectivity using two high-speed interfaces (Fast-Ethernet/Gigabit-Ethernet). For redundancy, each IX Backbone Router should have a link for each one of layer2 Ethernet switches.

Other equipments installed in the IX can also be connected directly to these switches. For equipments installed in other rooms/buildings, the local LAN is extended installing a hub/switch in each room and connecting them to the backbone switches.

The following picture shows the general connection model for the IX Local LAN.

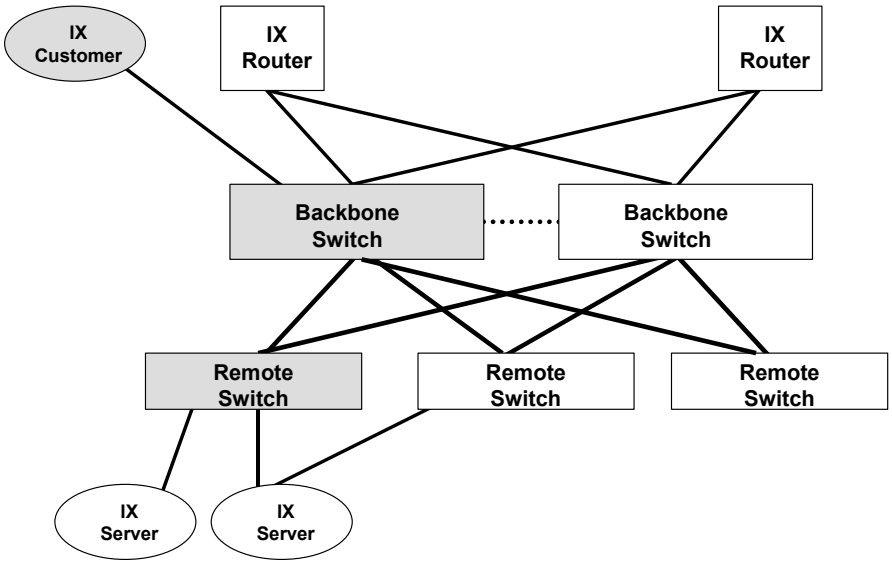

Figure 2-3: General Model for IX Local LAN

For higher availability Remote switches should have dual links to the backbone Switches. The main advantage of this design is that each remote switch maintains two equal-cost paths to every destination network, so recovery from any link failure is fast. This design also provides double the trunking capacity into the IX switched backbone.

2.1.2.2 Basic Configuration – Traditional IPv6 IX Customers

A simple Euro6IX IX will work as traditional IXs do, allowing ISP customers to install a router in IX facilities, establishing a local interface to the IX LAN segment and allowing peering traffic with other ISPs.

Traditional customers are those who use the IX in the traditional way (i.e. just to set-up their own peering with the participating ISPs). Next generation customers are those who use the Layer 3 mediation function of the IPv6 IX.

Customers can have one or multiple links for redundancy. Each IX should establish rules for establishing connectivity defining user interfaces allowed: Half-duplex 10BaseT, Full-duplex 100BaseTX, Full-duplex 1000BaseSX, etc.

One possibility to establish connectivity between ISPs that could be tested during the Euro6IX project lifetime is the usage of VLANs to control the connectivity between customers.

2.1.2.3 Basic Configuration – Next Generation Customers

Euro6IX IX nodes are supposed to provide a more flexible architecture to connect customers to the IX. According to this, one IX node could have an L3-mediation router that receives customers' traffic and forwards it to the local ISPs and to long-haul providers. This router must be attached to the IX LAN. The following picture represents this router in the IX architecture.

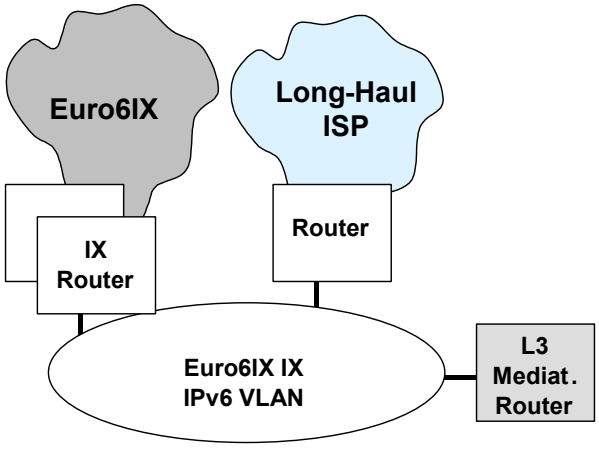

Figure 2-4: Layer 3 Mediation Router

The connection to the long-haul provider could use the local LAN or a Remote connection.

2.1.2.4 Basic Configuration – Connection between Regional IXs

Some IX could want to interconnect existent Regional IXs to the Euro6IX IX. The establishment of an Ethernet extension of the local bridge to the Remote switches can do this.

2.1.2.5 IEEE 802.1P – QoS Prioritization

The 802.1P is applied both for the *Ethernet* (IEEE 802.3) and the *Fast Ethernet* (IEEE 802.3u) and the *Gigabit Ethernet* (IEEE 802.3ab). 802.1P traffic is simply classified and sent to the destination; no bandwidth reservations are established. 802.1P is a spin-off of the 802.1Q (VLANs) standard. The 802.1Q standard specifies a tag to be appended to the MAC frame. The VLAN tag carries VLAN information.

Priority	Traffic Type
1	Background
2	Spare
0 (default)	Best Effort
3	Excellent Effort
4	Controlled Load
5	Video
6	Voice
7	Network Control

Figure 2-5: IEEE 802.1P Priority Levels Table

The VLAN tag has two parts: The VLAN ID (12-bit) and Prioritization (3-bit). The Prioritization field was never defined in the VLAN standard. The 802.1P implementation defines this prioritization field. 802.1P establishes eight levels of priority similar to IP Precedence (see Figure 2-5). Network adapters and switches route traffic based on the priority level. Using Layer 3 switches allows you to map 802.1P Prioritization field to IP Precedence field before forwarding to the routers.

2.1.2.6 Basic Configuration – Management Network

The basic IX will also have a LAN segment where to connect Network Management Systems (Monitoring/Management). Access to these systems will be protected by a firewall. The rules applied on the firewall should be defined according to each IX implementation. Firewall could also control the access to Services Area (DNS/WWW). The following picture represents the basic model for implementation of the Management Network.

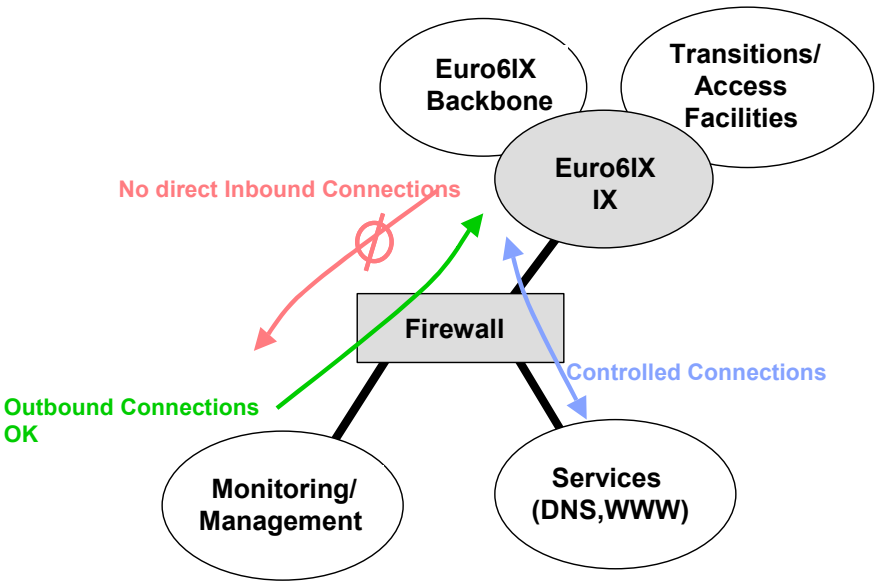

Figure 2-6: Management Network with Single Firewall

The firewall filtering should be implemented using dedicated equipments. In more critical IX firewalls, redundancy should be applied and in this case should be implemented based on two distinct boxes.

Typically each firewall has three Ethernet interfaces connected to Management LAN, IX Backbone LAN, and to the other firewall for synchronization. Traffic directed to the firewall uses virtual interfaces that represent the whole firewall system, regardless of which box forwards the packets. There should be two virtual interfaces, one for the Management LAN and the other for IX LAN.

In a dual-firewall environment the typical configuration should be that one described in the following scenario.

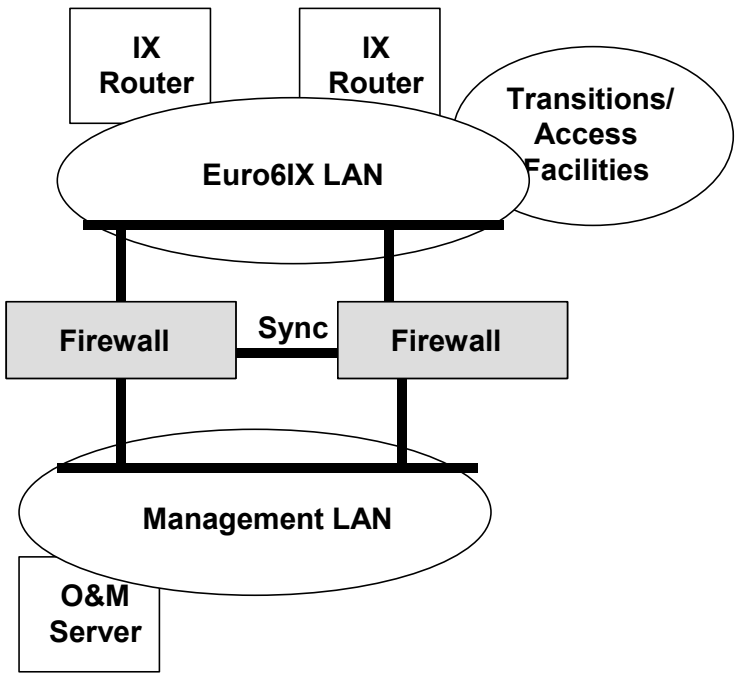

Figure 2-7: Management Network with Dual Firewall

Preferably Management LAN should be implemented in a separated VLAN secured by the firewall. In simple IX the firewall could be implemented using a router and defining simple ACL (Access Control Lists).

2.1.2.7 Advanced Configuration – Transition Facilities

A more complex IX will have mechanisms to translate packets from IPv4 to IPv6. These mechanisms will need fast access to the public IPv6 network and, as they route customer traffic no advanced security mechanisms are needed. The best way to connect these equipments is to directly connect them to the Euro6IX default VLAN allowing them to send/receive traffic from IPv6 customers. Layer 3 routing mechanisms should be applied in order to route traffic to these equipments.

Traditional IPv4 IX could evolve to IPv6 IX. To separate IPv6 and IPv4 traffic a special VLAN could be created in order to transport IPv4 traffic and a bridge between IPv6 and IPv4 world would route traffic among the VLANs using different layer-2 logic interfaces.

The following picture represents the architecture for IPv4 and IPv6 traffic separation in the IX.

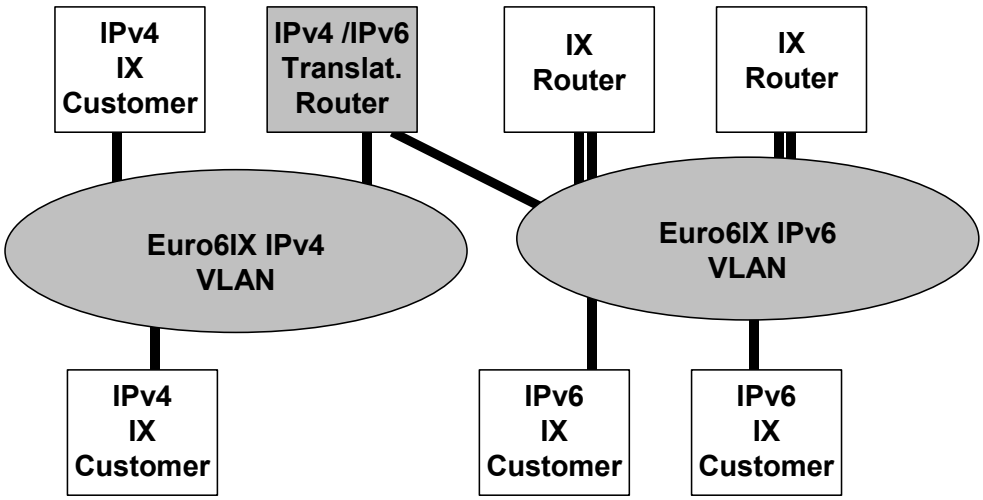

Figure 2-8: IPv4/IPv6 Transition

2.1.2.8 Advanced Configuration – Remote Access Customers

The IX will also provide remote access to customer either connected though leased lines or dial-up. The customers access will typically be done either by a direct connection to an IX backbone Router or though an Access router.

Access Routers will belong to IX manager and should have connectivity to the IX long-haul Backbone.

The following picture describes the standard architecture for connectivity between IX equipments.

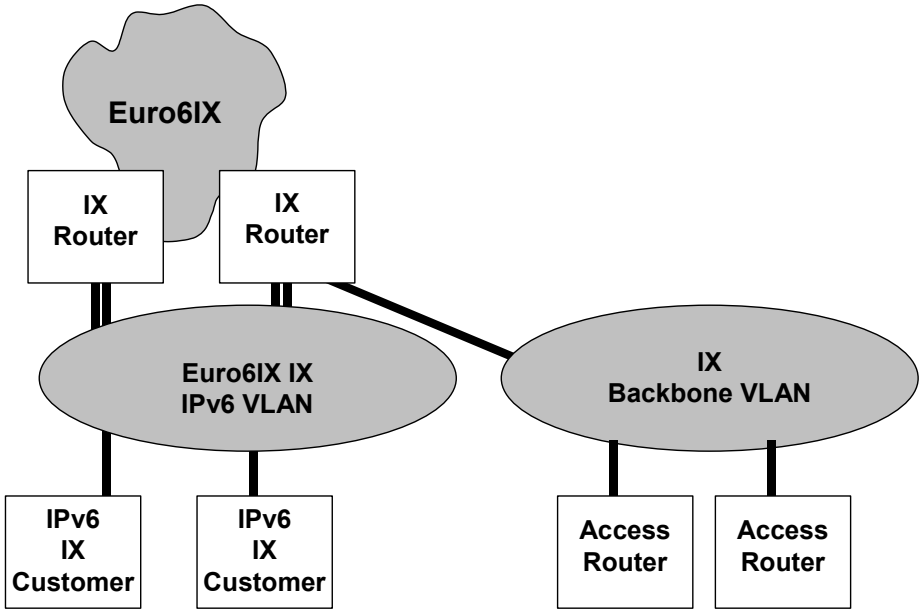

Figure 2-9: Inter-IX Connectivity

2.1.2.9 Advanced Configuration – Servers Farm

The IX will also support servers offering either Basic Internet Services like (DNS, IRC, NTP) or Content Delivery Services (WWW, FTP). These equipments will be connected to a specific network segment that should have high-speed links to the backbone segment and secured access from IPv6 customers.

The following picture represents the generic diagram for connecting a server farm to the IX backbone segment.

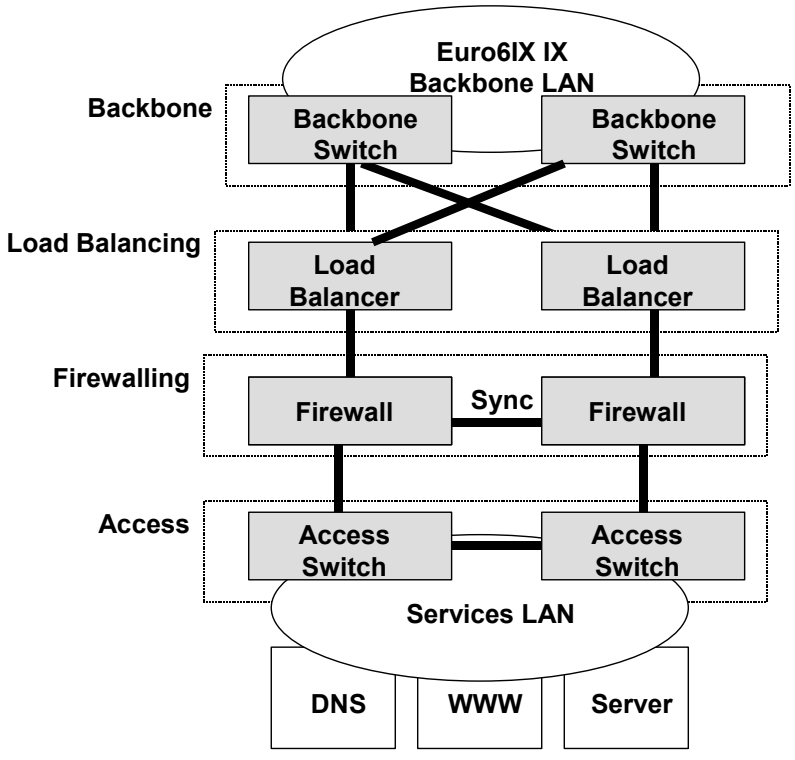

Figure 2-10: Server Farm

2.1.3 Layer-3 Issues

In this section some general guidelines about routing and addressing inside the IX are provided and two different scenarios will be shown. These scenarios refer to the different architectural models considered for two different phases in the evolution of the IX model. In any case, topic related to this section is currently under discussion among the partners.

2.1.3.1 Routing

Here we consider a new concept of the IX based on the so-called "layer 3 mediation function". This new role is based on the possibility, from the IX to assign IPv6 prefixes independent of the provider. In this case, each customer accessing the IX chooses the provider (one or more) and the IX assigns the IPv6 prefixes. If a customer decides to change the provider, it does not have to change IPv6 prefixes, because they are provider independent and are assigned by that particular IX. In this scenario, the router linking the IXs between each other can belong to a different administrative domain than that managed by Telco owning the IX. So an inter-domain routing protocol could run on this router to exchange routing information with the IX routers belonging to other administrative domains.

The elements that compose the network and the links among them identify the traffic flows to be exchanged. All these flows must be identified and controlled, so that a Network Routing Policy definition is needed.

Network Routing Policy is mandatory to identify, classify and manage network traffic flows. This work must consider the following Routing Levels and policies associated with those levels.

For the Internal Euro6IX backbone routing, the routing policy is described in the context of traffic exchange among IXs. The policy defined for this type of routing is:

- Every IX have to reach another IX using as much as possible the Euro6IX network infrastructure. For this reason Euro6IX prefixes received from Euro6IX links have to be always preferred respect to the same prefixes received from another link (e.g. 6Bone).
- Routes will be aggregated as much as possible and non-IANA allocated prefixes will be filtered out.

Regarding the Internal IX Routing, the routing policy is described in terms of exchanging traffic between one IX and other national networks, such as NRENs or other National Telcos, etc.

- This policy is freely implemented by Internet Exchanges Administrators.
- Since some IX can have already defined their own routing policies, this point is open to the election of the administrators of each IX.

For more details about the configuration issues refer to the deliverable D2.2.

2.1.3.2 Addressing

The addressing plan for the Exchanges and for the whole Euro6IX network is in full conformance with the RFC2374 "An IPv6 Aggregatable Global Unicast Address Format".

In the commercial phase, it is expected that every IX will have a commercial prefix delegated from 2001::/16. The aggregatable address format is designed to support long-haul providers, Exchanges, multiple levels of providers and subscribers, as shown in the following figure.

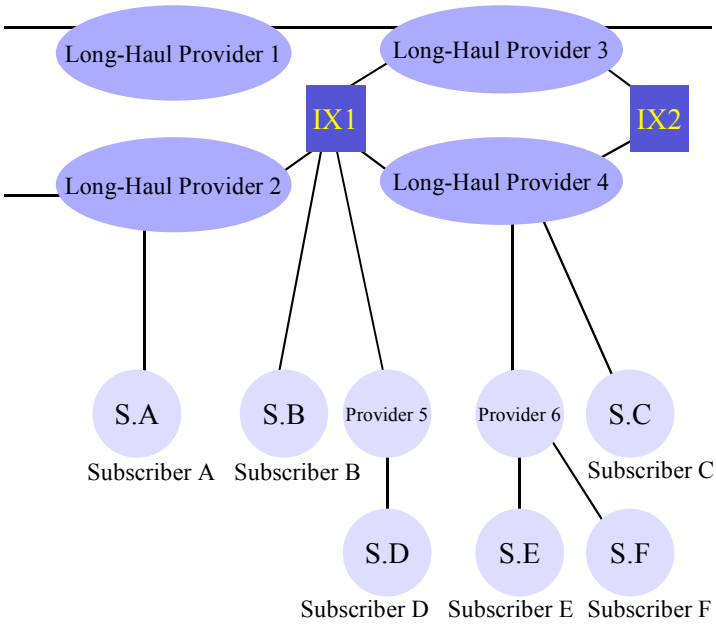

Figure 2-11: Public Topology Hierarchy

Exchanges will allocate IPv6 addresses. This condition is fulfilled since the L3 IX is owned by one of the providers. This ISP assigns addresses to its customers delegated from its own 2001:xxxx::/32 prefix(es).

Based on RFC2374 statements a new concept to study and research on is that, considering that the delegated addresses belong to the ISP in charge of the IX, organizations that connect to these IXs will achieve addressing independence from long-haul providers. Then, they will be able to change long-haul providers without renumbering their organization. They can also be multihomed via the IX to more than one long-haul provider.

The next figure shows the Aggregatable Global Unicast Address Structure.

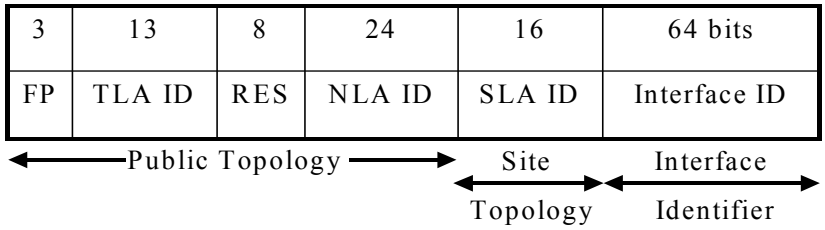

Figure 2-12: Aggregatable Global Unicast Structure

The different fields of the previous figure are:

- FP: Format Prefix (001).
- TLA ID: Top-Level Aggregation Identifier.
- RES: Reserved for future use.
- NLA ID: Next-Level Aggregation Identifier.
- SLA ID: Site-Level Aggregation Identifier.

- INTERFACE ID: Interface Identifier.

The following sections specify each part of the IPv6 Aggregatable Global Unicast address format.

- Top-Level Aggregation ID. Top-Level Aggregation Identifiers (TLA ID) are the top level in the routing hierarchy. The routing topology at all levels must be designed to minimize the number of routes into the routing tables.
- Reserved. The Reserved field is reserved for future use and must be set to zero.
- Next-Level Aggregation Identifier. Next-Level Aggregation Identifiers are used by organizations assigned a TLA ID to create an addressing hierarchy and to identify sites. The organization can assign the top part of the NLA ID in a manner to create an addressing hierarchy appropriate to its network. Each organization assigned a TLA ID receives 24 bits of NLA ID space. This space can be delegated to approximately as many organizations as the current IPv4 Internet.

Organizations assigned a TLA ID can provide service to organizations that provide public transit service and to organizations that do not provide public transit service. The organizations receiving an NLA ID may also choose to delegate their space to another NLA ID's. This is shown in the following picture.

n	24-n bits	16	64 bits
NLA1	Site ID	SLA ID	Interface ID

m	24-n-m bits	16	64 bits
NLA2	Site ID	SLA ID	Interface ID

o	24-n-m-o bits	16	64 bits
NLA3	Site ID	SLA ID	Interface ID

Figure 2-13: Next-Level Aggregation Identifier

- Site-Level Aggregation Identifier. The SLA ID field is used by an individual organization to create its own local addressing hierarchy and to identify subnets. It is a 16 bit field, so it supports 65.535 subnets. The approach chosen for structuring an SLA ID field is the responsibility of the individual organization. This is show in the next figure.
- Interface ID. Interface identifiers are used to identify interfaces on a link. They are required to be unique on that link.

n	16-n bits	64 bits
SLA1	Subnet	Interface ID

m	16-n-m bits	64 bits
SLA2	Subnet	Interface ID

Figure 2-14: Site-Level Aggregation Identifier

2.1.4 Transition Mechanisms

It is expected that there will be a long transition period during which IPv4 and IPv6 nodes coexist and communicate. A strong, flexible set of IPv4-to-IPv6 transition and coexistence mechanisms will be required during this transition period. A wide range of transition techniques have been identified and implemented, basically falling into three categories:

- Dual Stack Transition Mechanism that allow IPv4 and IPv6 to co-exist in the same devices and networks;
- Tunneling techniques, which encapsulate IPv6 packets inside IPv4 packets or MPLS frames to get through IPv6-ignorant routers/switches;
- Translation techniques, which use IPv6-IPv4 protocol translation to ensure IPv4-only devices to communicate with IPv6-only devices.

Different techniques may have different advantages and disadvantages and may suit different transition scenarios.

This section lists the different transition facilities that an IPv6 IX may offer. These facilities are based on the usage of transition tools currently defined today. The goal is to provide temporary solution enabling access to the IPv6 IX and/or enabling some means of heterogeneous communication between IPv4 and IPv6 nodes. Since it is assumed that legacy IPv4 world will coexist with the new IPv6 world for a long period, an IPv6 IX may implement such facilities.

2.1.4.1 Dual Stack Transition Mechanism

The Dual Stack Transition Mechanism (DSTM) is a technology developed by ENST Bretagne that enables a dual-stack host connected to an IPv6-only network to get some temporary IPv4 connectivity. In some way, DSTM provides the facility complementary to that one provided by the Tunnel Broker tool.

When a DSTM client wants to communicate with an IPv4-only node, it gets an IPv4 temporary address from the DSTM server (the IPv4 address is temporary allocated using a protocol that is implementation dependant). The IPv4 traffic is tunneled to the tunnel end point (TEP) through the dynamic tunnel interface (DTI), and then forwarded within the legacy IPv4 network to its destination.

This transition mechanism, which relies both on the DSTM server and on the TEP, may be easily managed and provided by an IPv6 IX, in order to offer native IPv4 connectivity and it may be needed in the medium or later phase of the transition, when IPv4 global addresses are part of scarce resource. This facility assumes in fact that most of the traffic is based upon IPv6 and that only some remaining communications need IPv4 connectivity.

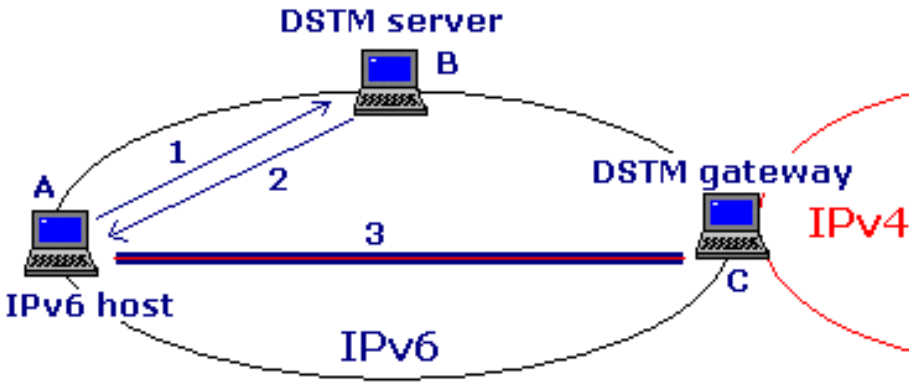

Figure 2-15: Dual Stack Transition Mechanism

The DSTM allows support for both IPv4 and IPv6 applications and services during the period that IPv4 services are gradually replaced by IPv6 versions, and new IPv6 services emerge. The main limitation is that an IPv4 address must be retained and thus as more devices become IP enabled and IPv4 addresses become more scarce, this approach will become less viable.

2.1.4.2 Tunneling Mechanism

A way to use the IPv6 protocol on the pre-existent network, based on IPv4, is to create a tunnel. A tunnel is a mechanism to let any packet be forwarded across different networks without being modified. Technically, the packet, entering inside another network (typologically different from the original one), is encapsulated inside the frame of the new network, forming its payload. So, the endpoints of a tunnel encapsulate a data packet into another one. Tunneling is a powerful technique used in many different areas.

Within Euro6IX, some applications have been tested using this technique.

2.1.4.2.1 6to4 Relay

6to4 tool defines a way to assign an interim unique IPv6 prefix, and a mechanism to encapsulate IPv6 prefix over IPv4 without explicit tunnel configuration. The main characteristics of 6to4 tool are reminded below:

- A 6to4 router is placed at the border of the unique 6to4 domain it serves, and the legacy IPv4 Internet.
- The interim unique IPv6 prefix is formed as follow: 2002:V4ADDR::/48, where V4ADDR is a global IPv4 address.
- The global IPv4 address V4ADDR is used by the encapsulating mechanism, since it ends every non-explicit tunnel to concerned 6to4 router.

Thus, 6to4 domains get an easy and transparent IPv6 connectivity with each other.

RFC3056 identifies mixed scenarios, which enables a so called 6to4 relay router to provide, either IPv6 connectivity to 6to4 domains, or IPv6 connectivity to isolated regular IPv6 domains using normal 2001:: TLA.

An IPv6 IX may deploy such a 6to4 relay router as alternative access facility, in order to provide IPv6 connectivity, either to isolated 6to4 domains, or to regular IPv6 domain. The IX's 6to4 relay router, located at the border of the IPv4 legacy Internet and the IPv6 new world, has two configuration options, with or without BGP4+:

- 6to4 relay without BGP4+: The 6to4 relay router may apply source address filtering, in order to accept traffic only from chosen domains. This is the only routing policy that can be deployed in this case.
- 6to4 relay router with BGP4+: This option enables to deploy a real, more complex, routing policy, choosing the routes to advertise and the traffic to accept.

Such a facility may be needed during the earlier transition phase, while only little ISPs offer IPv6 access services. This solution aims to meet short-term transition needs.

2.1.4.2.2 Protocol-41 Forwarding

A machine or network behind a NAT box can obtain IPv6 connectivity if the router supports the feature of forwarding the protocol 41.

This behavior provides a big opportunity to rapidly deploy a huge number of IPv6 nodes and networks, w/o the need of new transition mechanism. So exploring this option is very important to facilitate the IPv6 deployment.

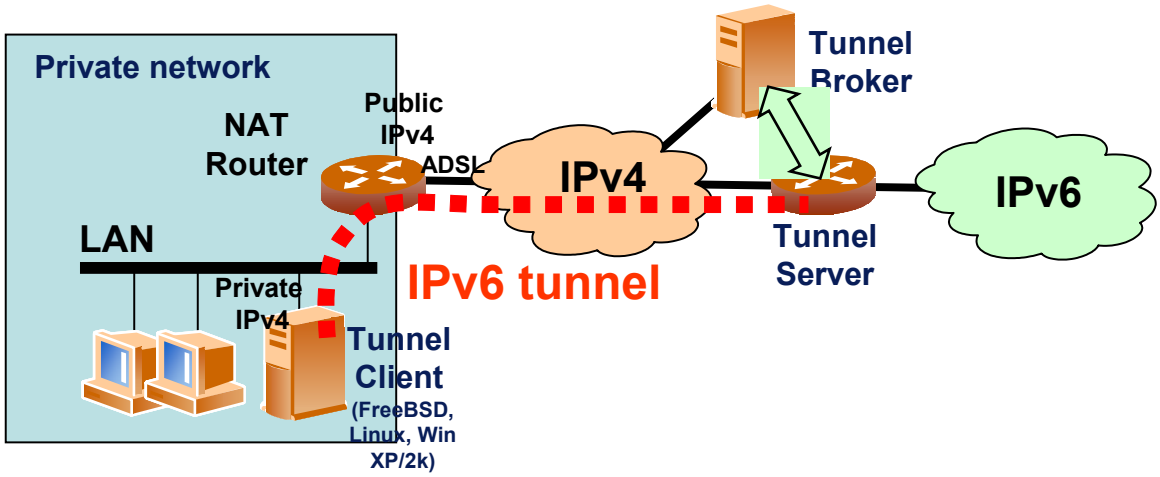

Figure 2-16: IPv6 Tunnel through a NAT-ADSL Router

An Internet Draft [draft-palet-v6ops-proto41-nat-02.txt] provides more details about this mechanism.

Such a facility may be needed during the beginning of the transition phase, when only a few number of routers provide IPv6 functionalities.

2.1.4.2.3 IPv6-over-UDP

It is also possible to have IPv6 connectivity behind IPv4 NAT using UDP tunneling.

IPv6 packets are encapsulated into UDP/IPv4 and sent over the network. This allows for passing through the NAT with UDP traffic, and works also with home ADSL modems/routers.

In the following picture it is shown a Linux PC, a NAT box/router providing IPv4 connection, a dual stack PC with a public IPv4 and IPv6 address connected to an IPv6 network. The latter PC is used as Tunnel Server.

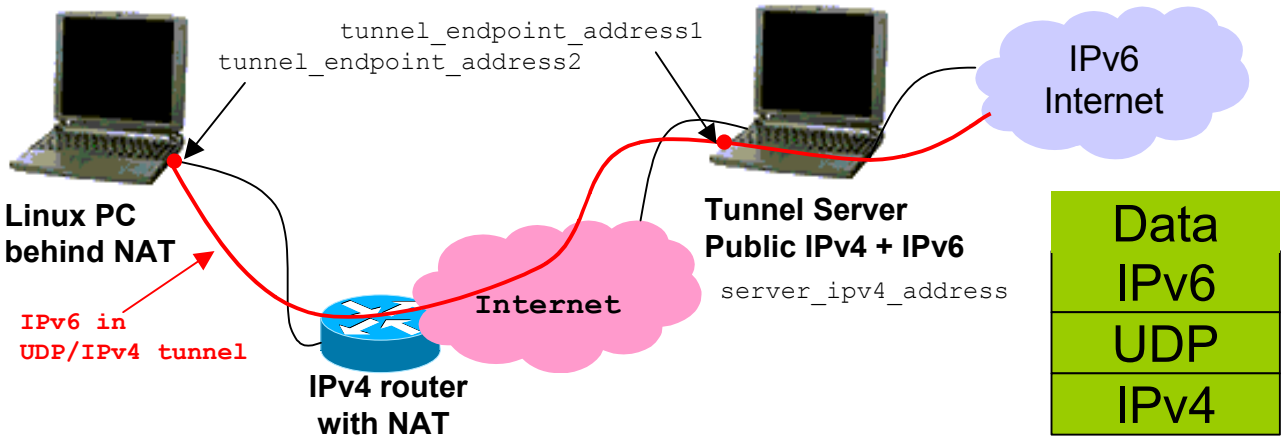

Figure 2-17: IPv6 Connectivity

With this type of configuration it is possible to provide to a PC machine, located behind a router with a Network Address Translation the access to an IPv6 backbone.

2.1.4.2.4 Tunnel Broker

The Tunnel Broker tool enables an isolated dual-stack host or an isolated IPv6 domain behind a dual-stack router, within the legacy IPv4 Internet, to get IPv6 connectivity through an IPv6-in-IPv4 tunnel. After a negotiation phase between the IPv6 host/router and the tunnel broker (TB), the tunnel server (TS) will set up the tunnel part from the IPv6 world to the dual-stack host/domain, while the TB returns some means (e.g. a script) to set up the reverse part of the tunnel. The Tunnel Broker will also allocate a permanent global IPv6 address or prefix, and permanent DNS names as well.

The Tunnel Broker tool is an efficient way to automate tunnel set up, and then to provide IPv6 connectivity over a legacy IPv4 network, IPv6 global address and DNS names as well.

Since the control of such a service relies on the Tunnel Broker that can be easily managed, an IPv6 IX may offer services based on this access technique. Basically, two levels of service can be defined:

- Simple Access Service: In this case, an IPv6-in-IPv4 tunnel replaces the classical L2 link. Tunnel Broker tool provides to the customer a basic IPv6 connectivity, typically allocating to the client remote node a /127 address, and setting up a tunnel between the former node and the IX. Then, a classical BGP session can be set up.
- Complete Access Service: In this case, where no BGP session needs to be set up, the IX provides IPv6 connectivity and prefix allocation (e.g. /64 or /48 prefix). The service may include also permanent DNS names management.

The main constraint relies on the user itself, which must have a fixed global IPv4 address that is used for the tunnel configuration with the Tunnel Server. Such a constraint may not be met by users (e.g. dial up users), which get access to the Internet from an ISP that shares a pool of IPv4 addresses.

2.1.4.3 Translation Mechanism

The representative of the third category of the translation techniques is Network Address Translation and Protocol Translation (NAT-PT). It is described in RFC2766 and SIIT proposal from IETF.

NAT-PT translates IPv6 packets into IPv4 and vice-versa. NAT-PT is a gateway located at the boundary between an IPv6 and an IPv4 network. It is normally a dual stack box with at least one interface connecting IPv4 Internet and another interface connecting to the IPv6 network.

NAT-PT can be implemented in a router platform or Linux/FreeBSD/Windows computer platform. There are several available implementations for NAT-PT. For example, BT Ultima is an experimental interworking device developed within BT. Ultima supports the NAT-PT mechanism including NAPT-PT and DNS/FTP Application Level Gateways (ALGs). NAPT-PT (Network Address Port Translation-Protocol Translation) takes the address translation a stage further by enabling the translation of port numbers as well. This makes it possible to re-use one IPv4 address and map this one IPv4 address to many IPv6 hosts. Cisco has implemented NAT-PT on its wide range of router models. Ericsson Telebit has developed NAT-PT on the AX11460 and the RX1820 research platforms. The protocol translator implementation distributed as part of the Microsoft IPv6 stack is another implementation of NAT-PT.

Two types of NAT-PT are identified: basic NAT-PT that enables only communication initiated by IPv6 nodes, and bi-directional NAT-PT that enables communication initiated from both sides. The former case assumes that the IPv6 domains do not need to deploy servers that need to be visible from the legacy IPv4 Internet.

In both cases, and in order to avoid any specific configuration on the IPv6 domains side, the NAT-PT facility needs to support the defined DNS-ALG. This is the major constraint, among the various restrictions of the NAT-PT tool, since it assumes that all the traffic has to go through the NAT-PT and its DNS-ALG, creating at least a single point of failure. Some other issues have been identified in [draft-durand-natpt-dns-alg-issues-00.txt], and a distributed solution (NAT-PT is a centralized solution) has been proposed in [draft-durand-ngtrans-nat64-nat46-00.txt].

2.1.5 Network Management

Network management in Euro6IX is based on Magalia tool developed at TID labs for the project.

The concept and tool are described in details in the deliverable D3.2.

2.2 External Connectivity

2.2.1 Network Topology

The network is interconnected with other IPv6 networks, in general, from different points as depicted in the following figure. Each external network connecting to Euro6IX, i.e. 6NET, must have a BGP4+ peering with one or more Euro6IX IXs to establish that connectivity. There will be also national networks connected to any IX. Euro6IX connectivity to these networks can be achieved establishing dynamic routing (BGP4+ peering) or just with static routing.

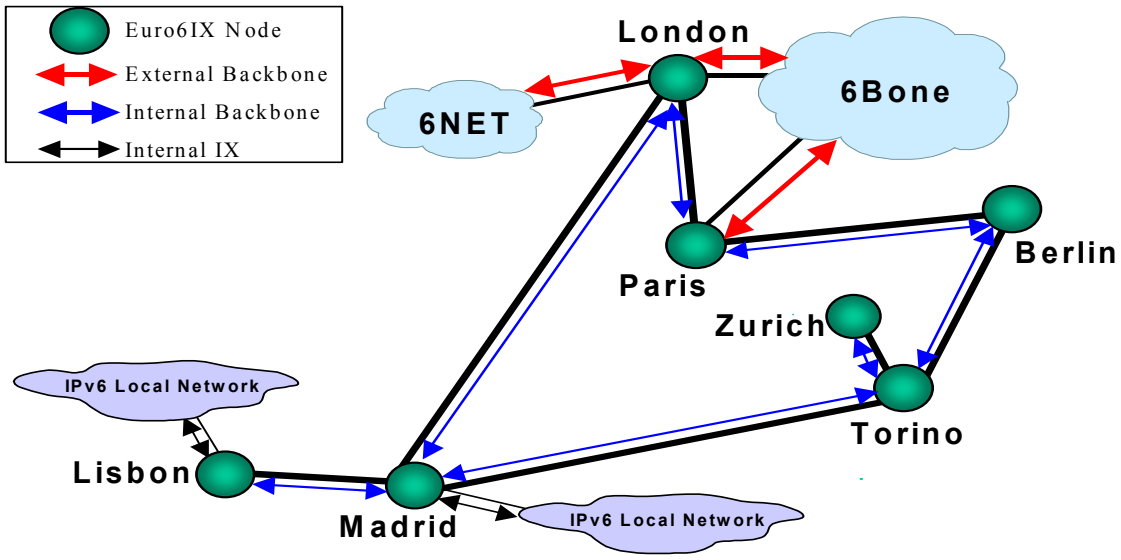

Figure 2-18: External Euro6IX Connectivity

2.2.1.1 6Bone

The 6Bone is an IPv6 test-bed that is an outgrowth of the IETF IPng project that created the IPv6 protocols intended to eventually replace the current Internet network layer protocols (IPv4).

The 6Bone is currently a world wide informal collaborative project, informally operated, initially, with oversight from the "NGtrans" (IPv6 Transition) Working Group of the IETF.

The 6Bone started as a virtual network (using IPv6 over IPv4 tunneling/encapsulation) operating over the IPv4-based Internet to support IPv6 transport, and is slowly migrating to native links for IPv6 transport. The initial 6Bone focus was on testing of standards and implementations, while the current focus is more on testing of transition and operational procedures.

A commercial situation for Euro6IX is that all IX have a native connection and peering with the 6Bone.

An intermediate solution is that some IX peer with the 6Bone natively, others peer through a tunnel and others use temporarily another IX connection and peering.

2.2.1.2 6NET

6NET is a three-year European project to demonstrate that continued growth of the Internet can be met using new IPv6 technology. It also aims to help European research and industry play a leading role in defining and developing the next generation of networking technologies.

The project will build a native IPv6-based network with both static and mobile components in order to gain experience of IPv6 deployment and migration from existing IPv4-based networks. This will be used to extensively test a variety of new IPv6 services and applications, as well as interoperability with legacy applications.

2.2.1.3 Other Networks

The LON6IX peers with an ever-increasing range of networks, transit to these networks is provided via the full routing table being made available to the Euro6IX partners that the LON6IX natively peers with (France Telecom and Telefónica). Of course it depends on the routing policies of these partners if they forward these routes to other Euro6IX partners.

For the publicly available list of LON6IX members please look at: http://www.uk6x.com under Operational Info, Public Members List.

The LIS6IX has another important peering with FCCN, the Portuguese NREN and the owner of the actual Portuguese IXv4 (GigaPIX). They have already started a pilot of the public PIXv6 a few months ago.

2.2.2 Layer-2 Issues

2.2.2.1 Basic Configuration – External Connectivity

The switched Infrastructure technology is dependent on the services that each IX is willing to provide and in particular if the Euro6IX IXs node will switch traffic as normally traditional IX do (using a common shared medium where ISP establish peering sessions to flow traffic between them).

A common building block to all the Euro6IX IXs is the international connectivity to other Euro6IX nodes. In the Euro6IX project, Internet exchanges will be interconnected with each other through dedicated trunks. These dedicated trunks terminate on an IX and can be: Dedicated lines (E3, STM1/OC3) or logical connections provided by a public ATM network.

Each IX will also provide long-haul connectivity to regional customers by creating logical connections between customer's equipments and the IX Router. For flexibility and scaling reasons this connections will normally use ATM or Frame-Relay technology.

So, normally IX external connectivity will be established using ATM or Frame-Relay connections passing though to a commercial ATM/FR network, exception being made for links to other IXs.

The following picture represents a basic model for interconnection with IX external entities.

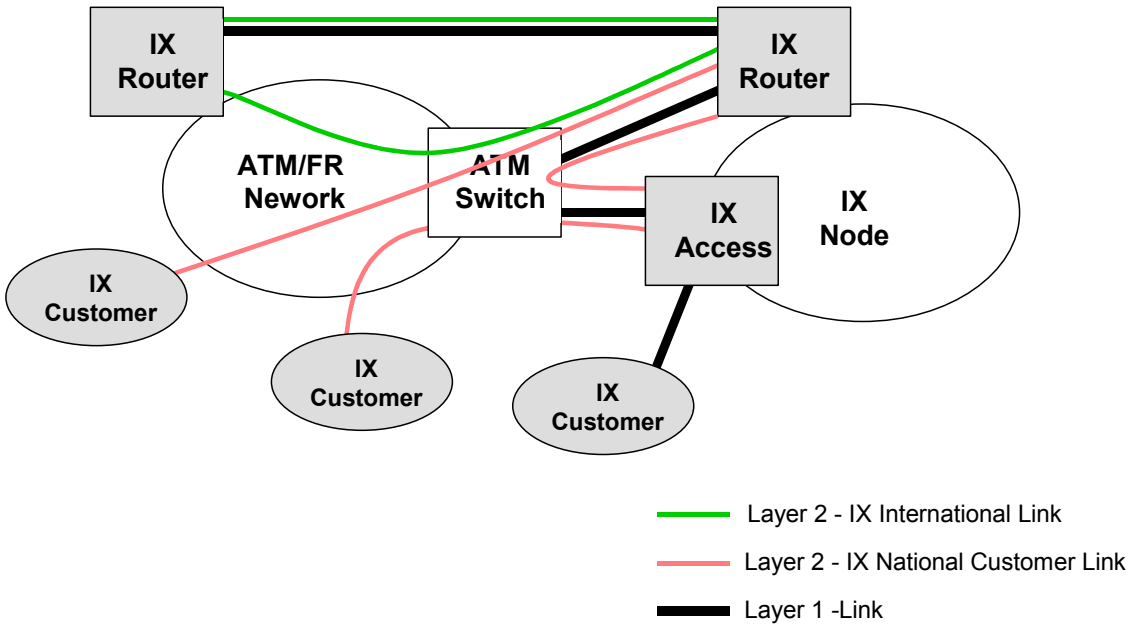

Figure 2-19: External Connections

Basically IX equipments installed in the IX can have connections using: Physical direct links either for the international link to other IXs or for leased lines access if the IX provides Access Services; logical links passing though a ATM Network. Typically, IX equipments will connect to an L2 switch port of the ATM Network Switches.

An IX node could also have a local ATM Switch (L2 Switch) allowing ATM local connectivity between equipments installed in the IX. ATM/FR public Networks will policy traffic sent in each logical connection to assure that the contracted traffic parameters are respected. So to establish multiple logical connections inside a single logical connection created on the ATM Network (VP Tunneling) is not a possibility.

This architecture could be used for an IX providing long-haul connectivity.

2.2.2.2 Basic Configuration – International Connectivity

Euro6IX IX can also function as NAPs (Network Access Points) to an IPv6 international network. In this case Customers will peer with a long-haul IPv6 Operator that could also be connected to the IX node. In this case peering agreements are more complex and long-haul providers could: Either charge for international connectivity based on measurement of the total traffic transferred per Operator or to limit the total bandwidth provided per Operator.

The long-haul router should do bandwidth limitation by defining: Layer-3 traffic shaping on the egress of the long-haul provider router interface; ATM policing through an ATM switch.

When a long-haul provider offers international connectivity through the local Euro6IX LAN, the following Layer 2 configuration applies:

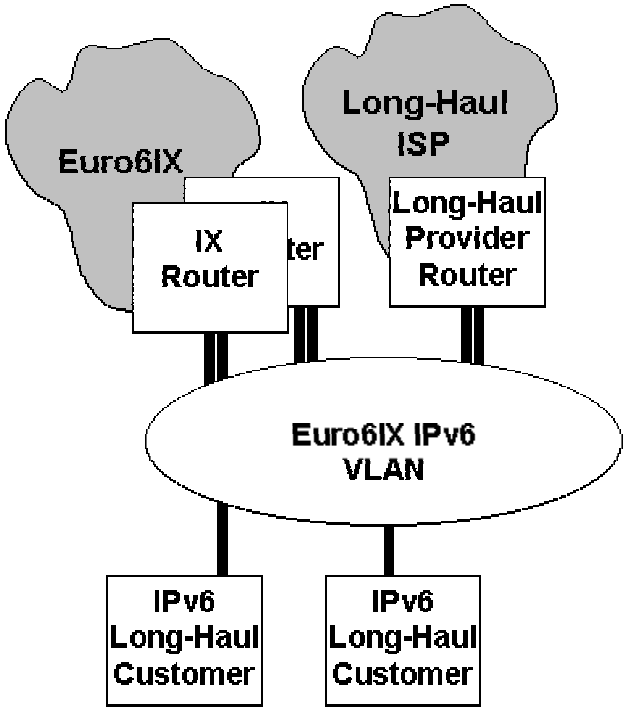

Figure 2-20: International Connectivity with VLAN

When a long-haul provider offers international connectivity though an ATM connection the following configuration applies:

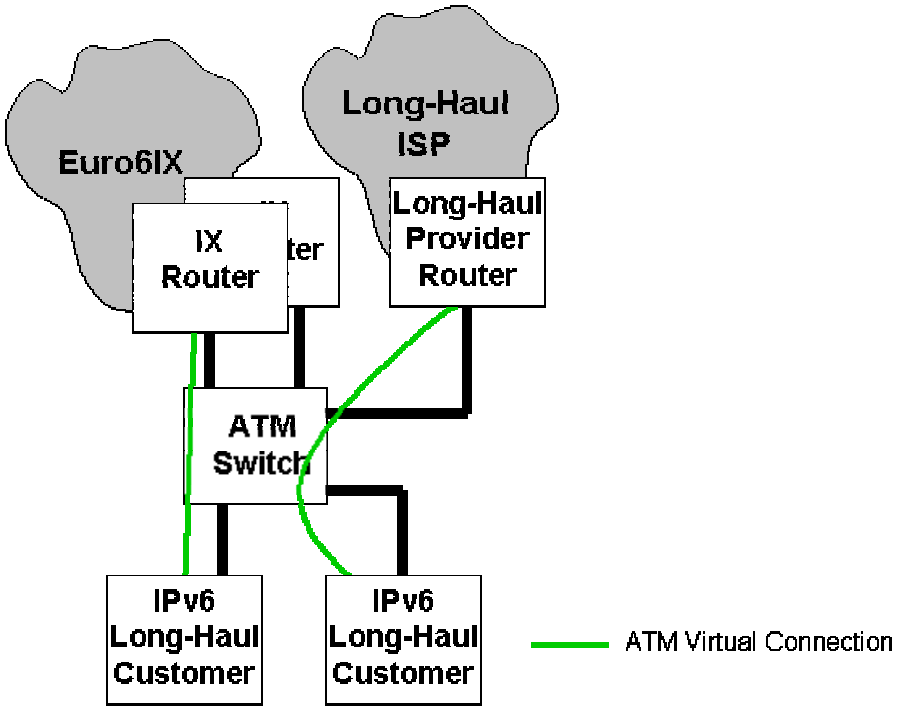

Figure 2-21: International Connectivity with ATM

The use of ATM technology for customer-connections provides guaranteed, adjustable, dedicated bandwidth for exchanges of traffic among interconnected networks.

Thanks to ATM PVC operating features this connections can provide:

- Inter-peer interconnection throughput guarantee (configurable speed).

- Impenetrability of inter-peer flows enabling control of throughputs allocated.
- Differentiated Qualities of Service (QoS), favoring purchase or sale of connectivity among providers with quality of service configurable per PVC.
- Granularity and flexibility for each ATM PVC bandwidth.
- Scalability by simple reconfiguration of PVCs corresponding to inter-peer throughputs.

ATM interconnection is implemented by setting up ATM PVCs between interconnection routers. IX managers would assign a PVC number to customers for their interconnections with long-haul provider router. IX customers that want to have ATM international connectivity must negotiate the ATM PVC with de the long-haul provider. This PVC will be dedicated to peering or transit traffic between the two ISPs.

The class of ATM service that is available on the switch should be: Variable Bit Rate Non Real Time (VBR-NRT). The VBR-nrt class corresponds best to support IP data flows from the Internet by offering a guarantee on an average cell rate. The switch should have a UPC (User Parameter Control) that can be activated by VC. This control operates PCR (Peak Cell Rate), SCR (Sustainable Cell Rate) and MBS (Maximum Burst Size) parameters and uses up to two levels of GCRA (leaky bucket) algorithms. For VBR-nrt traffic, a first level of the GCRA algorithm will police on the PCR(0) and a second level on the SCR(0) while marking outsize cells. Some rules should be defined like the PCRs (Peak Cell Rate) and SCRs (Sustainable Cell Rate) may have values of 1Mbps to 34Mbps or 155 Mbps by increments of 2Mbps. The SCR is ½ of the PCR selected.

2.2.3 Layer-3 Issues

2.2.3.1 Routing

About the External Euro6IX Backbone routing, the routing policy has to be described in order to exchange traffic between IXs and other networks like 6Bone, 6NET, Asia Pacific Networks, etc. Respect to these networks, the routing policy defined is:

- In general IX doesn't provide transit between two external networks (prefixes coming from an external network aren't announced to another external network), except for some special networks.
- Every IX have to reach external networks exiting from the interconnection point between Euro6IX network and the External network. If more than one interconnection point is present between Euro6IX network and external network the routers will make a decision based on the nearest link (ASPath based or IGP based).
- Special procedures will be specified for special routes to/from particular experimental networks like 6Bone, 6NET:
    - **6Bone:** Many of the Euro6IX IXs peer with parts of the 6Bone. IXs will inject 6Bone routes onto the Euro6IX backbone. Euro6IX IX will prefer 6Bone routes received from 6Bone peers. IX will not provide transit to 6Bone network. Euro6IX prefixes will be announced to 6Bone peers so that the 6Bone PoPs will choose (in the most cases) 6Bone path instead of Euro6IX path. In the future, additional policies (e.g. preferring 6Bone native links respect to tunneled) may be specified.

- **6NET:** Some of the Euro6IX IXs peer either directly with 6NET or via a transit provider (both cases will be treated the same). 6NET routes received from an IX will be announced Euro6IX backbone only to Euro6IX partners: Routes will be preferred using normal metric (e.g. BGP AS path length). Euro6IX prefixes will be announced to 6NET using no metric change. This is a requirement of 6NET that Euro6IX only advertise their routes to companies in the Euro6IX consortium**.**

For more details about the configuration issues refer to the deliverable D2.2.

2.2.4 DNS

The Internet Domain Name System (DNS) is a distributed hierarchical database that permits to establish a correspondence between names and IP addresses. Clients look up information in the DNS by calling a resolver library, which sends queries to one or more name servers and interprets the responses.

The Berkeley Internet Name Domain (BIND) implements a domain name server for a number of operating systems:

- IBM AIX 4.3.
- Compaq Digital/Tru64 UNIX 4.0D.
- Compaq Digital/Tru64 UNIX 5 (with IPv6 EAK).
- HP HP-UX 11.
- IRIX64 6.5.
- Sun Solaris 2.6, 7, 8.
- NetBSD 1.5 (with unproven-pthreads 0.17).
- FreeBSD 3.4-STABLE, 3.5, 4.0, 4.1.
- Red Hat Linux 6.0, 6.1, 6.2, 7.0.

BIND 9 fully supports all currently defined forms of IPv6 name-to-address and address-to-name lookups. It will also use IPv6 addresses to make queries when running on an IPv6 capable system.

For forward lookups, BIND 9 supports both AAAA and A6 records. A6 was moved to experimental in IETF environment, but it is still useful for hosts to have both A6 and AAAA records to maintain backward compatibility with installations where A6 records are used. In fact, the stub resolvers currently shipped with most operating system support only AAAA lookups, because following A6 chains is much harder than doing A or AAAA lookups.

For IPv6 reverse lookups, BIND 9 supports the new "bitstring" format used in the ip6.arpa domain, as well as the older, deprecated, "nibble" format used in the ip6.int domain.

DNS Server placed in the several Euro6IX IXs will have mainly the function of cache DNS server that permits routers and other servers to resolve direct (AAAA) Reverse Records and inverse (PTR) queries made by DNS clients.

Nowadays, there is an open discussion about the possibility that the IX assigns its own addresses: In this case the IX should manage the reverse delegation of its addresses.

2.2.5 AAA

Authentication, Authorization, and Accounting (AAA) concept is used in IP based network management and policy administration. There are several AAA protocols, some of them in standardization process and some of them that need an IPv6 implementation. Within the scope of Euro6IX the most interesting AAA protocols for the IXs are:

2.2.5.1 RADIUS

The Remote Authentication Dial In User Service (RADIUS) is a protocol for carrying authentication, authorization and configuration information between a Network Access Server, which desires to authenticate its links and a shared Authentication Server.

RADIUS can run over IPv6 and RADIUS attributes can be used to support IPv6 network access. With this tool the IXs could offer dial-up and broadband network access including managed VPN and wireless.

RADIUS Authentication Servers can be set up in a variety of ways, depending upon the security scheme of the network they are serving, but the basic process for authenticating a user is essentially the same.

Using a modem, a remote dial-in user connects to a network access server (NAS) with a built-in analog or digital modem. The NAS then prompts the user for a name and password. For protection against eavesdropping by hackers, the NAS, acting as the RADIUS client, encrypts the password before it sends it to the authentication server. If the primary security server cannot be reached, the security client or NAS device can route the request to an alternate server. When an authentication request is received, the authentication server validates the request, and if the user name and password are correct it enables the necessary procedure to allow the user the access rights to network services and resources. If at any point in this login process all necessary authentication conditions are not met, the security database server sends an authentication reject instruction to the NAS device and the user cannot access the network.

2.2.5.2 DIAMETER

The DIAMETER base protocol is intended to provide an AAA framework for applications such as network access or IP mobility. DIAMETER is also intended to work both with local AAA and with roaming situations. The IETF AAA working group is specifying the DIAMETER protocol for communication between servers where RADIUS is currently being used. The basic concept behind DIAMETER is to provide a base protocol that can be extended in order to provide AAA services to new access technologies.

A DIAMETER implementation could be used to support IPv6 network access to Euro6IX. For sure both RADIUS and DIAMETER test will be carry out.

2.2.6 Network Monitoring

In Euro6IX backbone all the machines inside the network (i.e. routers, switches, servers) have to be monitored. Section seven of the deliverable D2.1 describes monitoring at different levels, from the infrastructure, through route server, server farm, reachability to traffic monitoring, taking into account also the implementation options.

2.2.7 Redundancy

2.2.7.1 L2-Redundancy

In the following, there are some technologies that support L2 redundancy.

2.2.7.1.1 Spanning Tree Protocol (STP)

The Spanning-Tree Protocol (STP) is a Layer 2 protocol designed to run on bridges and switches. The specification for STP is called 802.1d. The main purpose of STP is to ensure that we do not run into a loop situation when we have redundant paths in our network.

To provide a path redundancy and avoid a loop condition, STP defines a tree that spans all switches in an extended network. STP forces certain redundant data paths into a standby (blocked) state, while leaving others in a forwarding state. If a link in a forwarding state becomes unavailable, STP reconfigures the network and re-routes data paths by activating the appropriate standby path.

With STP, the switches in the network elect a root bridge that becomes the focal point in the network. All other decisions in the network, such as which port is blocked and which port is in the forwarding mode, are made from the perspective of this root bridge. A switched environment, which is different from that of a bridge, most likely deals with multiple VLANs. When implemented in a switching network, the root bridge is usually referred to as the root switch. Each VLAN (because it is a separate broadcast domain) must have its own root bridge. All the root bridges of the different VLANs can reside in a single switch, or they can reside in various switches.

All the switches exchange information to use in the selection of the root switch, as well as for subsequent configuration of the network. This information is carried in Bridge Protocol Data Units (BPDUs). The BPDU contains parameters that the switches use in the selection process. Each switch compares the parameters in the BPDU that they are sending to their neighbors with those one that they are receiving from their neighbors.

If the root ID (that a Switch A is advertising) is smaller than the *root ID* that its neighbor (Switch B) is advertising, Switch A's information is better. Consequently, Switch B stops advertising its *root ID*, and instead accepts that of Switch A.

According to IX L2 Topology, Backbone Switches should be root switches. These switches should have a bridge priority that is lower than all remote switches so that the remote switches will automatically select one of them as the root switch. Typically the bridge priority default value is 32768.

2.2.7.1.2 VLAN Trunking Protocol (VTP)

VTP is a Layer 2 messaging protocol that maintains VLAN configuration consistency by managing the addition, deletion, and renaming of VLANs on a network-wide basis. VTP minimizes miss-configurations and configuration inconsistencies that can result in a number of problems, such as duplicate VLAN names, incorrect VLAN-type specifications, and security violations.

VTP capable devices can be configured to operate in the following three modes:

- **Server:** In VTP server mode, we can create, modify, and delete VLANs and specify other configuration parameters (such as VTP version and VTP pruning) for the entire VTP domain. VTP servers advertise their VLAN configuration to other switches in the same VTP domain and synchronize their VLAN configuration with other switches based on advertisements received over trunk links. VTP server is the default mode.
- **Client:** VTP clients behave the same way as VTP servers, but we cannot create, change, or delete VLANs on a VTP client.
- **Transparent:** VTP transparent switches do not participate in VTP. A VTP transparent switch does not advertise its VLAN configuration and does not synchronize its VLAN configuration based on received advertisements. However, in VTP version 2, transparent switches do forward VTP advertisements that they receive out their trunk ports.
  VTP pruning increases available network bandwidth by reducing unnecessary flooded traffic, such as broadcast, multicast, unknown, and flooded unicast packets. VTP pruning increases available bandwidth by restricting flooded traffic to those trunk links that the traffic must use to access the appropriate network devices. By default, VTP pruning is disabled.

2.2.7.1.3 VLAN Routing

Four different protocols are available for routing between VLANs. All these technologies are based on OSI Layer 2 bridge multiplexing mechanisms.

- **Inter-Switch Link Protocol:** The Inter-Switch Link (ISL) protocol is used to interconnect two VLAN-capable Ethernet, Fast Ethernet, or Gigabit Ethernet devices. The ISL protocol is a packet-tagging protocol that contains a standard Ethernet frame and the VLAN information associated with that frame. The packets on the ISL link contain a standard Ethernet, FDDI, or Token Ring frame and the VLAN information associated with that frame. ISL is currently supported only over Fast Ethernet links, but a single ISL link, or trunk, can carry different protocols from multiple VLANs.

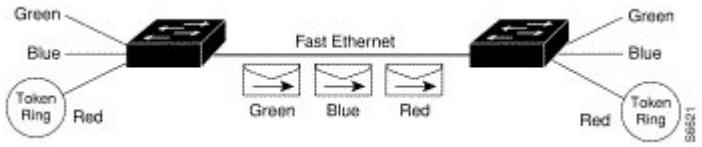

Figure 2-22: Inter-Switch Link Protocol

- **IEEE 802.10 Protocol:** The IEEE 802.10 protocol provides connectivity between VLANs. Originally developed to address the growing need for security within shared LAN/MAN environments, it incorporates authentication and encryption techniques to ensure data confidentiality and integrity throughout the network. Additionally, by functioning at Layer 2, it is well suited to high-throughput, low-latency switching environments. The IEEE 802.10 protocol can run over any LAN or HDLC serial interface.
- **IEEE 802.1Q Protocol:** The IEEE 802.1Q protocol is used to interconnect multiple switches and routers, and for defining VLAN topologies.
- **Layer 3 Routing:** If L2 have routing capacity they can route traffic across VLANs. In this case IP addressing in each VLAN cannot be overlapped.

2.2.7.2 L3-Redundancy

In this section a brief overview about two redundancy protocols is given. These mechanisms are implemented on the routers, but it is important that Layer infrastructure 2 is configured to support them.

2.2.7.2.1 Layer 3 VLAN Redundancy: HSRP

The Hot Standby Router Protocol (HSRP) is a Cisco innovation, which provides excellent fault tolerance and enhanced routing performance for IP networks. HSRP allows Cisco routers to monitor each other's operational status and very quickly assume packet-forwarding responsibility if the current forwarder in the HSRP group fails or goes down for maintenance. This mechanism remains transparent to the attached hosts and can be deployed on any LAN type. With Multi-Group Hot Standby, routers can simultaneously provide redundant backup and perform load sharing across different IP subnets.

2.2.7.2.2 Layer 3 VLAN Redundancy: VRRP

The Virtual Router Redundancy Protocol (VRRP) is a standard protocol that provides a function very similar to that one provided by the Cisco proprietary protocol HSRP. A group of routers individuate the so-called VRRP group where a router is elected the “Master virtual router” and the others are the “Back-up virtual routers”. These routers are seen as a single virtual router that is configured as default gateway for the clients in a LAN. If the Master virtual router goes down, then the back-up virtual router becomes the Master virtual router (according to a VRRP priority number) forwarding the packets in place of the Master virtual router gone down.

3. NETWORK FEATURES

3.1 QoS

This section intends to give some ideas regarding the deployment of Quality of Service (QoS) in the Euro6IX network.

In the last years, we have assisted to the increase of applications with special QoS requirements over the Internet. Services like voice, video and audio streaming, audio-conference, e-learning or tele-medicine, among many others. The future networks must be able to differentiate such kind of traffic, providing to them better conditions than the others in order allow their correct operation. Other traditional applications like www or ftp can deal better with rate variation or loss and their resources should be demoted if necessary.

There are two approaches for QoS provisioning, IntServ and DiffServ. Both can provide QoS for the networks. However, only DiffServ is able to work properly at a large scale, therefore, it should be the one to be deployed on the Euro6IX core.

3.1.1 QoS Appliance to Euro6IX

The Euro6IX is a network composed by several IPv6 Internet Exchangers (6IXs). To every 6IX, we will have several Service Providers (SP) connected, including the 6IX itself, which will act also as a commercial operator providing direct connectivity services to the customers as defined on D2.2 (L3 6IXs). The following figure shows the actual deployment of the 6IXs over Europe.

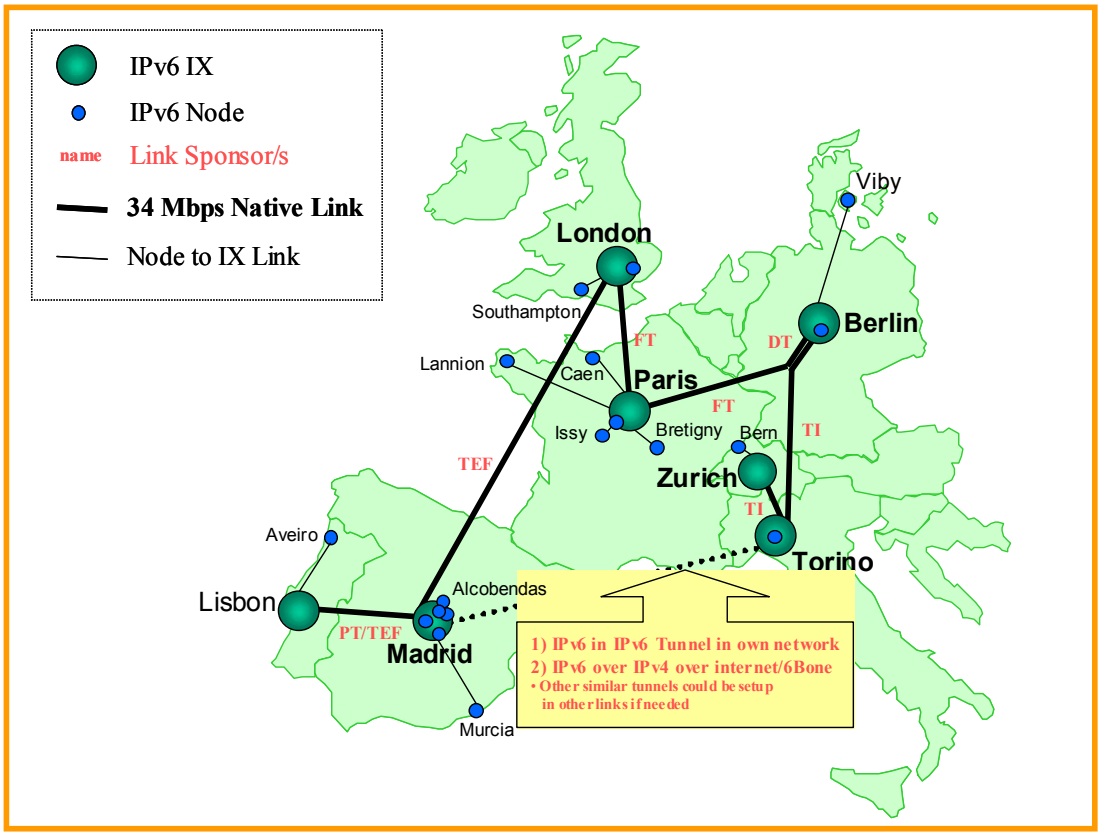

Figure 3-1: Euro6IX Connectivity

3.1.1.1 Complexity of Euro6IX

However, although we call it "Euro6IX network", actually this is something more complex, because there are a lot of administrative domains on this network and hence different administrators and policies. For this reason, in such scenario, the difficulty to provisioning QoS over the entire Euro6IX network become quit more complex.

The following figure depicts a typical model for QoS implementation over an administrative domain:

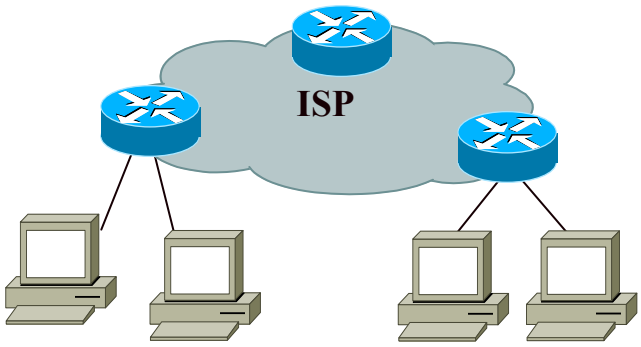

Figure 3-2: Typical Network Model for QoS Provisioning

This network is composed by core and edge (usually PoPs) routers. Following the DiffServ model in its simplest way, the traffic coming from the customers is received by the edge routers, marked on DSCP, and policed or shaped according to the SLAs. The EF and the twelve AFs classes could be used as well as the default BE. The core routers will perform queuing and congestion avoidance according to the marking performed by edge routers. Eventually, the core routers could remark packet as specified for AF DPs.

In this simple model, with a single Service Provider all the tasks are performed within the same administrative domain, and it can be insured that the marking, policing and queuing tasks applied to the packets along the network are coherent. However, the real world, and, in particular the Euro6IX, is not a single Service Provider, but a lot of them, as shown in the following figure:

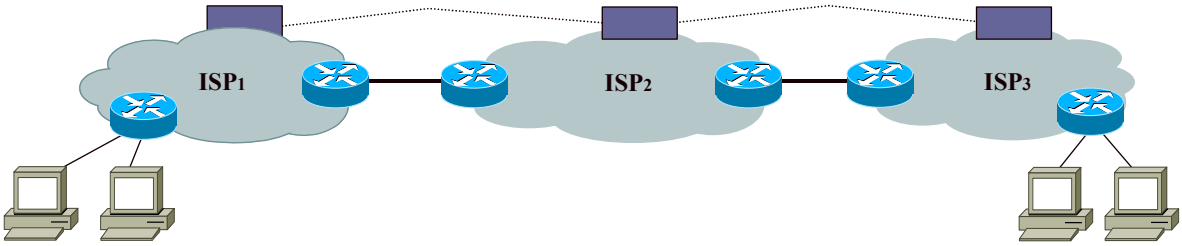

Figure 3-3: Euro6IX QoS Model

To implement a global QoS in such environments, we need to define agreements of SLSs between the partners, regarding the PHB that a given class receives in the different administrative domains. In some cases the remarking could be performed. Also, the definition of SLAs is needed in order to define how much traffic is allowed between the providers.

In the case where the Providers, to manage the QoS provisioning, use some external entities, interoperability is already needed between these entities in order to negotiate the SLAs dynamically.

3.1.1.2 The 6IX Model

One of the important issues to be taken into account is the Euro6IX model and how it could influence the QoS deployment model. From our point of view, the conceptual model is the same as the one shown in the following figure.

Administrative domains are connected ones each other's the same way, either by Ethernet technologies (within the 6IX) or WAN technologies like FR or ATM between IXs. The figure depicts this vision.

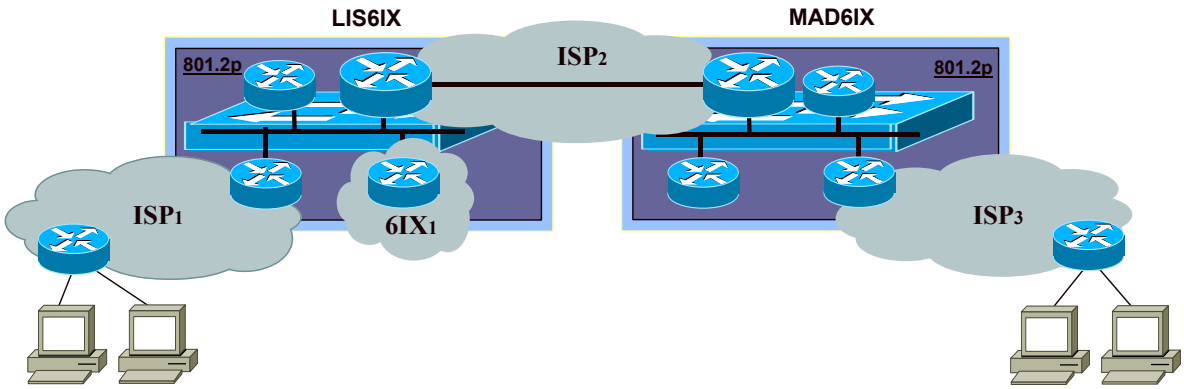

Figure 3-4: Conceptual Scenario for Euro6IX in terms of QoS Deployment

The Euro6IX models also brings the new L3 6IX innovative model, allowing the 6IX to provide connectivity services to the final customers as well as other ISP traditional services. This situation doesn't changes anything the model. The IX network ($6IX_1$ in the Figure) should be considered as one Service Provider more. This provider could also provide connectivity to other 6IXs as ISP2 does.

The only thing new required from the 6IX is the provisioning of L2 QoS, using 801.2P technology. This is needed to support in a coherent way the L3 QoS. But not just the Ethernet-based should do it. All the other connections either ATM or FR should do the same, if possible.

3.1.2 Classes Of Service

One of the main issues to discuss between the partners is the CoS definition. How much they would be, and to what services and users it would be applied.

Our proposal goes to the use the EF class for all the multimedia traffic. Applications like Isabel, VoIP, video, audio and TV streaming and also Management should have the maximum priority. This should be consensual for all the partners. Of course, in this class is assumed that is just to be used, when traffic from and to Euro6IX partners are present.

Next levels of priority can be divided in two. Classes AF1 and AF2, and Classes AF3 and AF4 (here I assume AF1 and AF2 are most priority classes. RFCs are not clear about this). The first two (AF1 and AF2) should be used between Euro6IX partners, differentiating between the non-multimedia (or non-real-time) applications, the ones who could be more and less important, or require more or less resources. The other classes should be used by special Euro6IX users (not partners), defining the same two levels of traffic, with the same philosophy. Finally, the BE should be applied to all the rest of the traffic, insuring that external traffic can transit the Euro6IX, but not affect the activities of the partners and special users.

This is just a generic proposal. These aspects should be discussed and defined more carefully and in a more detail, of course with the contribution of all the partners.

Also the SLAs must be defined between the partners, but this is perhaps a latter issue.

3.1.3 Equipment Support

It seems that all the equipment deployed in the Euro6IX network supports that DiffServ model. At least, this is what the activities performed during the first year shown.

In DiffServ, the compatibility problems don't use to arise, because the interoperation between the equipment is not so much. In any case we must check if any problem exist.

3.1.4 Liaison with WP4

The deployment of a DiffServ model does not invalidate the test of other more innovative models like IntServ over DiffServ proposed in WP4 whose architecture is shown in the following picture.

With slight changes these experiences could be set up between some partners, while the rest of the partners could have a more traditional QoS services.

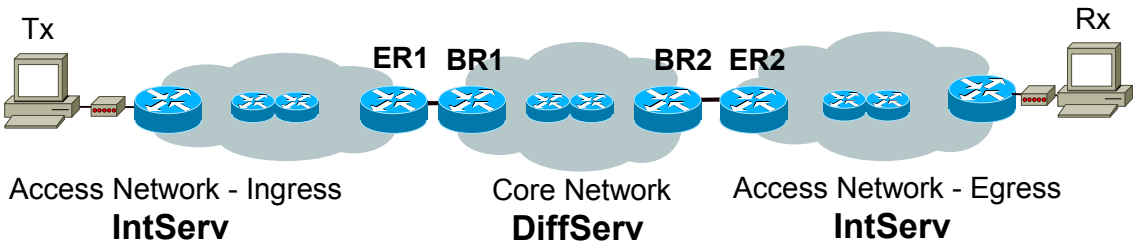

Figure 3-5: General Architecture for IntServ over DiffServ Model

3.2 Security

3.2.1 Global Approach

Security is a concept that must be addressed through a global approach. In fact, the security level of a system is equal to the level of its weakest component. Building security solutions in the Information Technology domain consists in designing the solution with respect of one or more of these three main aspects:

- Confidentiality (C), information should be available only to those who rightfully have access to it.
- Integrity (I), information should be modified only by those who are authorized to do so.
- Availability (A), information should be accessible to those who need it when they need it.

Analyzing the use of Internet according to each aspect, we can deduce the various risks like Eavesdropping (C), Information alteration (I), and Denial of Service (A).

On the other side, each component of a system may have security weaknesses, resulting of bugs, bad configurations and so on. In fact, weaknesses may appear at each step of the development of a system: Specification, Implementation and Configuration.

Security problems may occur at the intersection of those risks and those weaknesses.

The objective of a security solution is to limit (or avoid in the best case) the impact of security risks. It can only be achieved by a global solution addressing each layer of the architecture. Thus adding new services means that the new security requirements, listed in the following sections, will have to be addressed. However, though it is necessary, it will not be enough. Obviously, a network protocol alone does not provide a solution when applications have bugs or security weaknesses but it heavily contributes to enable the global solution.

3.2.2 Current Security Solutions with IPv4

Several well-known solutions were designed for IP networks:

- The IPsec protocol allows building **VPN** (Virtual Private Network). With this concept, the network is private since a security policy restricts the communication service to a community of end-points that trust each other. It is virtual because it actually makes use of a public network. The VPN is established using authentication mechanisms and protocols implemented at the end points. It covers mainly the Confidentiality and Integrity aspects.
- The **Firewall** is the most appropriate mechanism to protect a private network against insecure traffic from outside networks while allowing the advantages and benefits of interconnectivity. It completes the previous solution with an Availability aspect. It is advised to use dynamic ("stateful") packet filtering which keeps a state on opened sessions. This technology is now mature and more secure than static filtering.
- **Access control** mechanisms such as Radius

Combining these solutions brings a very good coverage of security aspects. Next figure gives an example of a network including:

- Nomadic accesses protected by an IPsec tunnel.
- IPsec VPN securing communications with a "Remote Office" over a public network.
- A front firewall combined with an IPsec gateway. The IP filters control accesses service by service.
- An internal firewall creating a more secure zone, "Back Office", and allowing accesses only from the internal zone. This higher security level is recommended for sensitive data, like financial or medical data, customer accounts and so on.

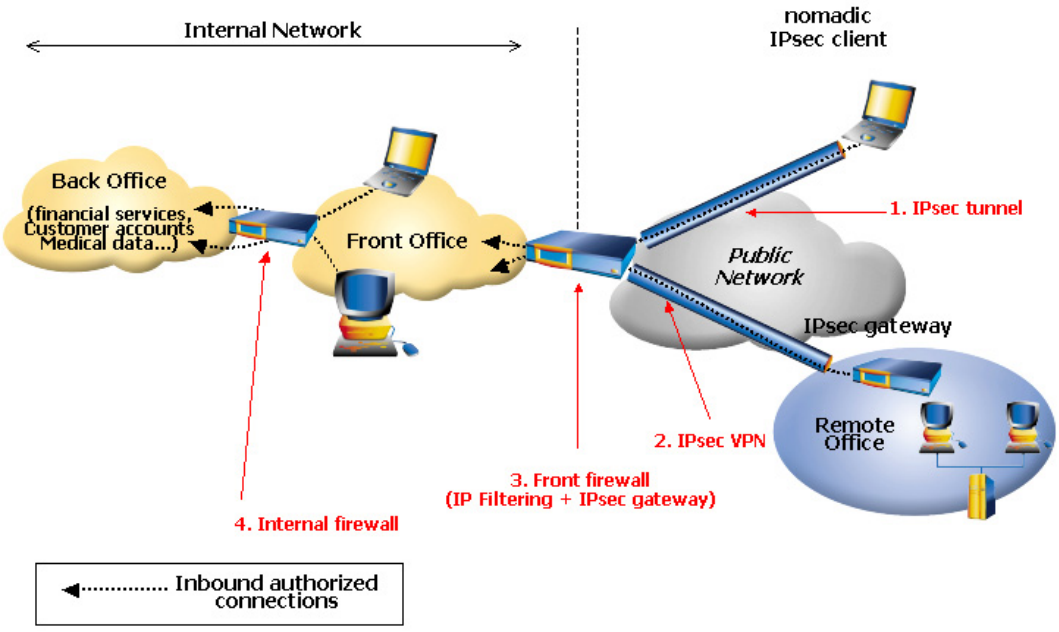

Figure 3-6: Security Network Scenario

Despite the existence of security solutions, IPv4 has some limitations that are essentially associated with the usage of NAT:

- **Complex network architecture**: NAT implies more complexity in network architecture. Little by little, while the use of IP networks was growing, architectures became a real "patchwork". This complexity leads to a rather fragile structure.
- **IPsec not supported**: IPsec has a limited coverage due to NAT: the large deployment of IPsec is difficult in IPv4. The problem is that IPsec is not compatible with NAT. NAT replaces IP addresses, which prevents IPsec from verifying authenticity and integrity of packets.

3.2.3 Security Solutions with IPv6

IPv6 is compatible with existing architectures and support the same security features as IPv4 described in the previous section. Therefore IPv6 provides the same solutions as IPv4 does for standard services such as mail, web browsing, file transfers, etc.

One of the big advantages of IPv6 regarding security is that IPv6 restores network simplicity. The complexity induced by the usage of NAT will disappear with IPv6 and simplicity will bring better security.

Security is a global issue and each component must have a coherent security level. Obviously, IPv6 alone does not provide solution when applications or implementations have bugs or security weaknesses. Nevertheless, the conjunction of features including IPv6 addressing accessibility combined with largely deployed end-to-end IPsec tunnels or more secure hosts internal architecture will allow to keep a security level that will be equivalent and probably higher than within IPv4 networks.

3.2.4 Issues and On-going Research Efforts

Still, IPv6 implementations are not always as complete as IPv4 ones from a security point of view. For instance, IPSEC for IPv6 or stateful filtering are not always supported, even in very well known implementations. And even when they are, their impact on performance should be measured.

3.2.5 Impact on IPv6 IX

In WP4, TID has issued a report on "Security Guidelines for Euro6IX Networks". Most of the elements below are inspired by this report.

The D2.1 document "Specification of the Internal Network Architecture of each IX point" specifies the architecture of each Euro6IX Internet Exchange:

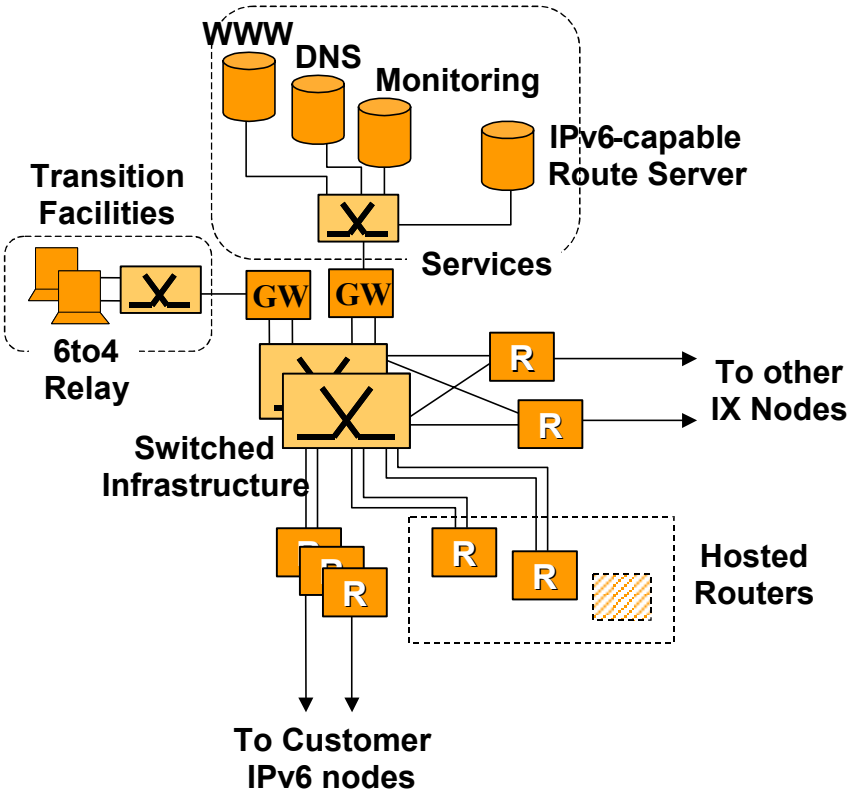

Figure 3-7: Euro6IX Architecture

Taken from a security point of view, this architecture includes several zones:

- The "IX" zone which includes routers switches and two other sub-zones:
  - A "services" zone containing resources that must be accessible from the outside world such as DNS servers, web servers, etc… Transition facilities may be placed here too.
  - A "Network Management" zone that includes management stations. This zone must be very secure.

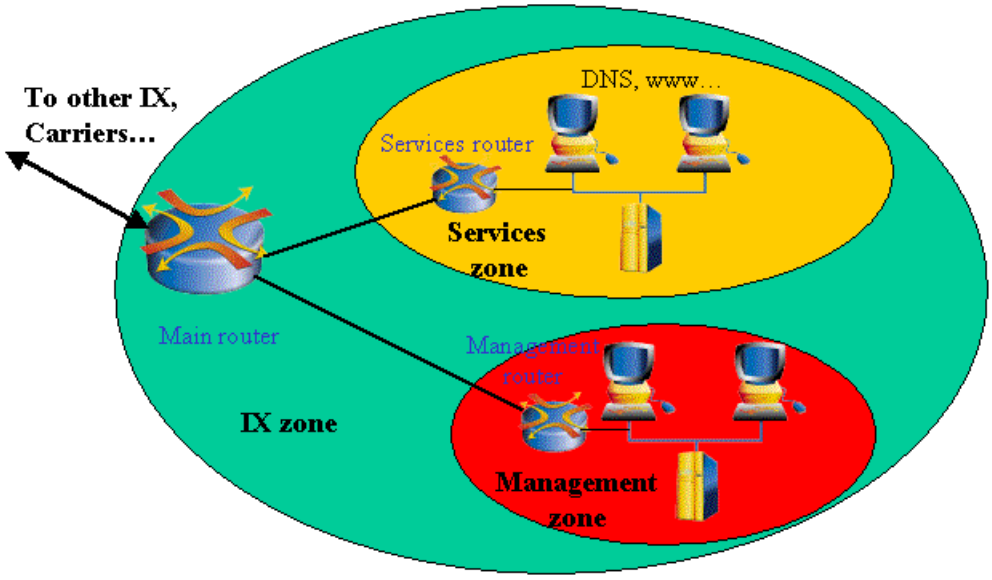

Figure 3-8: Security Zones

3.2.5.1 Protection of the IX Zone

This zone is the less secure one. Security in it must be enforced on elements, which provide external connectivity. It is mainly based on the use of filters and return routability checks. Filters used here should be based on IP source and destination addresses. They must be kept simple because they must not impact the throughput of the router. In fact, stateless filtering is here sufficient.

3.2.5.2 Protection of the "Services" Zone

This zone already benefits from the protection of the IX zone. However, security should be stricter here and port numbers should also be filtered. Stateful filtering is here a must.

3.2.5.3 Protection of the "Management" Zone

Securing this zone is essential. If ever, it is compromised, then the whole IX will be compromised: an intruder could then take control of any element of the IX.

Strict stateful filters are needed here and logs should be kept carefully.

3.3 Multicast

Multicast technology is based mainly on the capability of a network to accept a single packet from a given application (resident on a particular host) and to forward it to different destinations inside the network. In the following picture, we see the difference between a unicast network and a multicast network.

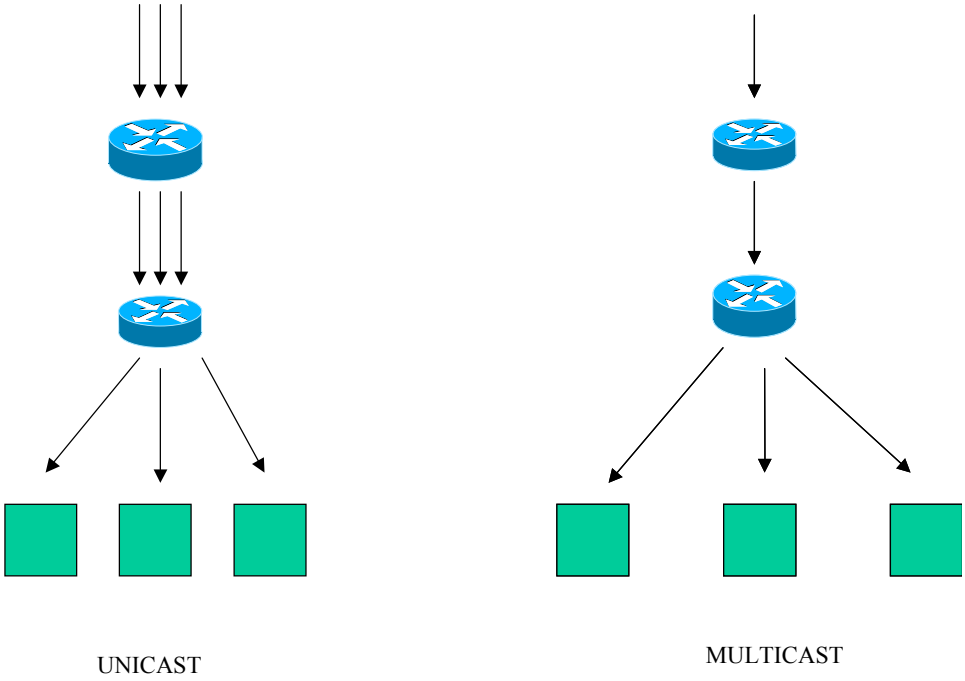

Figure 3-9: Unicast and Multicast Logic

In both of cases, there are three destinations where the packets have to be sent. In the first case (unicast network), the source sends as many packets as destinations, whereas in the second case the sends only a single packet and the network, multicast capable, will make the copies of the packets (as many copies as many destination).

This approach has the following basic advantages:

- Network bandwidth optimization (e.g. multicast network makes a copy of the packets only when needed).
- Source optimization (e.g. no need to transmit a packet per destination).
- Well scalable technology (e.g. when a receiver is added to a network, no addiction cost is required).

In a multicast capable network foresees the multicast routing protocol runs over the network in order to generate the multicast routing table. Particularly, the core routers of the network will run the multicast routing protocol, while the access routers (connected to the customer site) will run both the multicast routing protocol and the protocol for managing the groups inside the network.

Moreover, users are grouped in the so-called “multicast groups”. All the hosts inside a multicast group have the same multicast address and when a host wants to add or to leave a group, it has only to communicate this to the border router of the network where it is connected.

So summarizing, inside the backbone there will be:

- Core routers running quite typically PIM-Sparse Mode.

- Access routers running PIM-SM and additionally MLD (for group managing).
- Hosts implementing the MLD protocol.

3.3.1 Multicast Protocols

All the multicast routing protocols are based on a particular algorithm that is used in order to build the distribution tree. This tree is built searching for optimizing the path for delivering the packets from source to multiple destinations.

Main algorithms used are: Flooding, Spanning Tree, Reverse Path Broadcasting, Truncated Reverse Path Broadcasting, Reverse Path Multicasting, Core Based Tree and Steiner Tree. The protocols based on these algorithms are: DVMRP, MOSPF, CBT, PIM.

It is possible to make several classifications of these multicast routing protocols. First classification can be made based on the unicast routing protocol (in fact multicast protocols need inside the network also a unicast protocol for building the unicast routing table):

- PIM, CBT (Core Based Tree) are independent on the unicast protocol inside the network.
- DVMRP (Distance Vector Multicast Routing Protocol), MOSPF (Multicast OSPF) dependent on the particular type of routing protocol (RIP or OSPF).

Another classification can be made based on the distribution tree. So it is possible to distinguish among:

- Source Based Tree Protocols, where there are as many distributions tree as the sources.
- Shared Tree Based Protocols where there is only one distribution tree per group shared by all the sources that send data to that group.

Another classification can be made on the basis of the distribution tree of the hosts inside the group. In this case, it is possible to distinguish:

- Dense Mode Protocols (this protocol is called “dense” because the number of hosts receiving the multicast packets inside a given group is high).
- Sparse Mode Protocols (“sparse” when the multicast receivers density is low, inside a group).

At the time of writing, main operating systems can be used for implementing IPv6 multicast by host side (e.g. *BSD, Linux, Windows XP, Windows 2000). On the other hand, by router side, main router vendors already support IPv6 Multicast and moreover it is available the IPv6 multicast KAME software for *BSD equipments, with implementation of PIM-SM (and Dense Mode) and MLD protocols.

3.3.2 “m6bone” Initiative

M6Bone network was developed inside a project started basically in France in 2001 with the support of groups like Aristote, G6 and Renater, involved in IPv6 research and technology fields.

The aim of this project was to offer IPv6 multicast connectivity and, in the same time, to set-up a test bed to test some multimedia IPv6 multicast applications giving also the possibility to the participants to share their IPv6 knowledge.

This network is based on tunnel architecture due that in the actual IPv6 backbone network the routers and intermediate systems do not enable multicast traffic. If a local network wants to access to the M6bone backbone it only must configure a tunnel between a local router and a M6bone central point. This tunnel will route the multicast traffic to the local site.

At the being time, more than 35 sites all over the world are connected to M6Bone. Euro6IX network is connected to M6bone through several links, from UMU and UPM, PTIN and Consulintel networks but in a near future Euro6IX IXs, like MAD6IX, will offer M6Bone traffic inside the Euro6IX core network and future clients.

3.3.2.1 M6Bone World Map

The M6Bone network is a backbone of IPv6 Multicast nodes, which allows clients to use multicast IPv6 characteristics. Figures 2-1 to 2-4 show the actual network status.

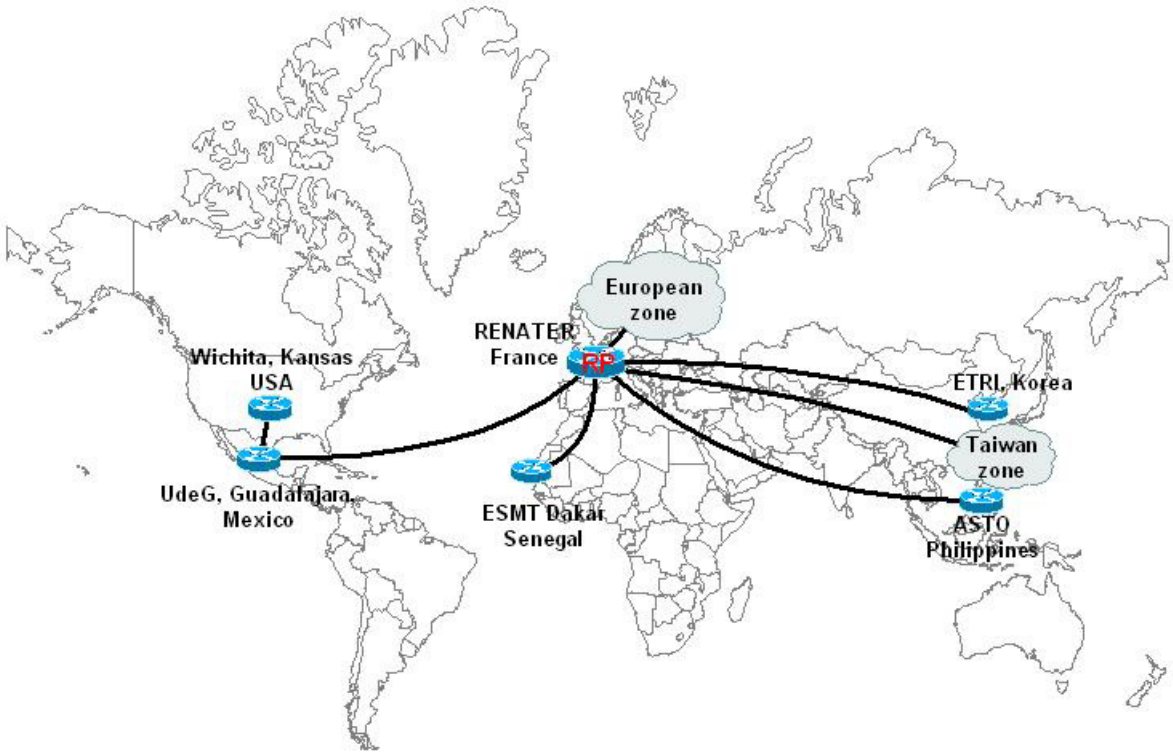

Figure 3-10: M6bone World Connections

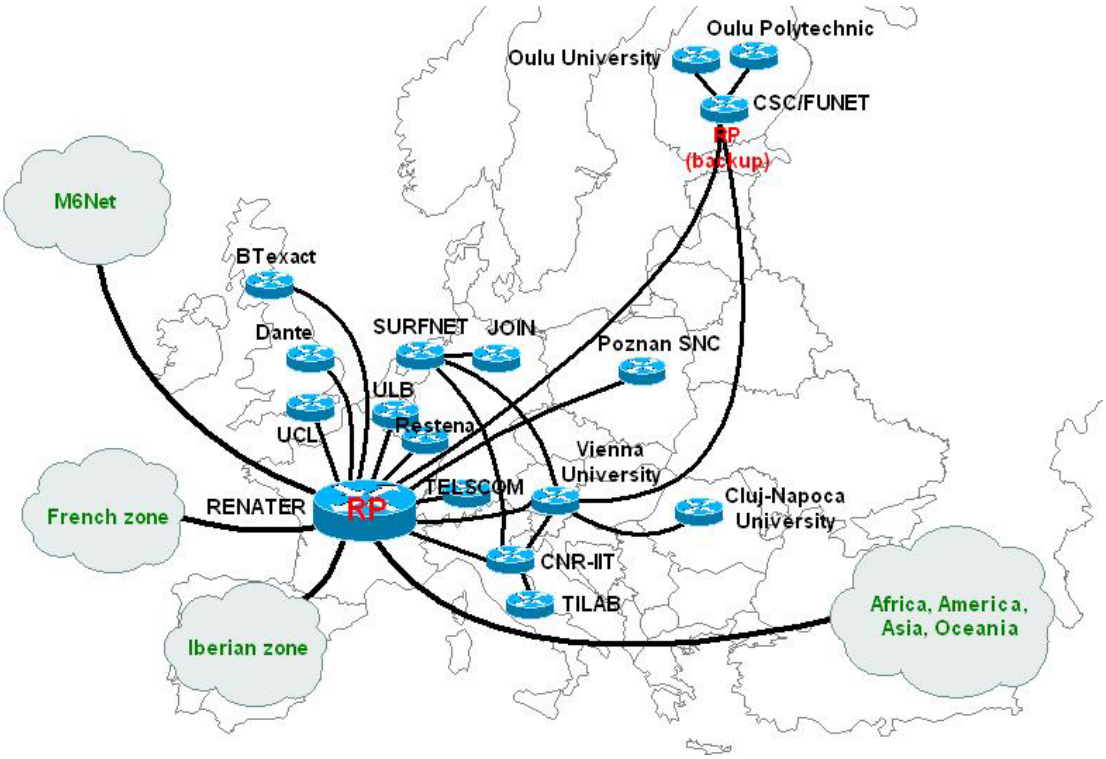

Figure 3-11: M6bone European Network

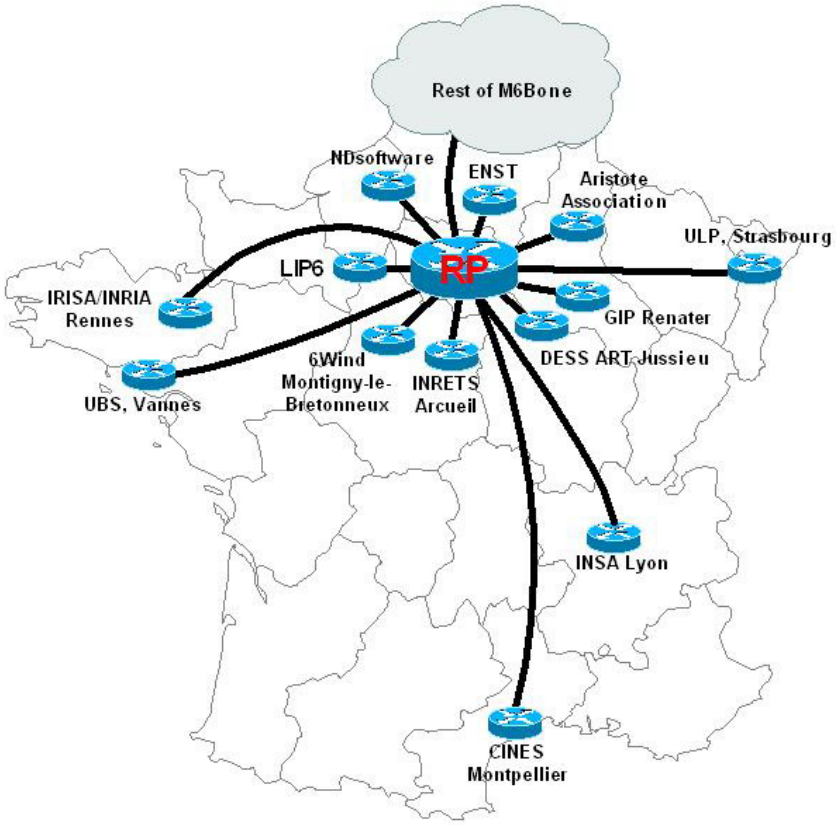

Figure 3-12: M6bone French Connection

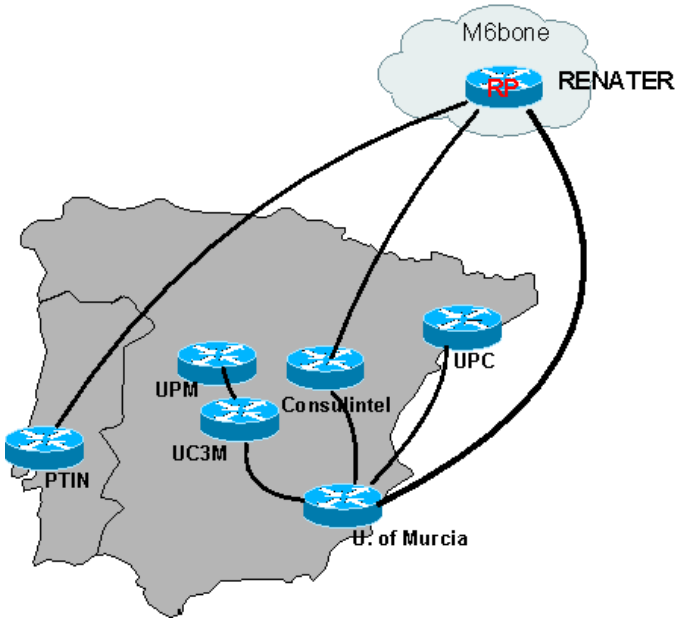

Figure 3-13: M6bone Iberian Connection

Actual 6Bone [6BONE] network is a heterogeneous system formed by several types of IPv6 routers. All these routers do not know Multicast IPv6, so the M6Bone is a Multicast network over the 6Bone formed by multicast IPv6 routers.

[M6BONE] is based on the interconnection of IPv6 sites through tunnels that can be:

- IPv6 (multicast) in IPv6 (unicast) tunnels for those sites that have already IPv6 connectivity.
- IPv6 (multicast) in IPv4 (unicast) tunnels for those sites with only IPv4 connectivity.

As shown in the picture above, inside the M6Bone there are two Rendezvous Points (RP). First one is the basic one managed by Renater (that works also as Bootstrap router) and second one, acting as backup RP, managed by the Finnish entity CSC/Funet.

Inside the network, the protocol used is PIM-SM (for building the multicast routing table) and RIPng (for the unicast routing table). Actual M6Bone policies, foresee that each site entering the network has to indicate the prefixes it will announce inside the network. Every prefix unknown will be filtered out by the RP. Moreover new sites will configure both RIPng and PIM-SM between their own router and the router of the nearest M6bone PoP.

3.3.2.2 M6bone Multicast Applications

Main multicast IPv6 tools used inside the M6bone are:

VIC (Video Conferencing Tool)

This tool was developed for videoconference by LBL (Lawrence Berkeley National Laboratory) in collaboration with Berkeley University. It can be used for point-to-point and multicast transmissions with bit rate from 10 Kbps to 3 Mbps. It is supported by different operating systems like Windows, *BSD and Linux.

RAT (Robust Audio Tool)

This tool was developed by UCL and, likewise VIC, can be used both for point-to-point and multicast transmissions. It is based on the RTPv2 protocol and it permits to reach a high quality of the audio thanks to the mechanisms of packet redundancy, interleaving and error correction on the receiver side.

SDR (Session Directory)

It is a tool used to announce the beginning of a multicast event and to let the other users (on a multicast network) to register to the event without knowing IPv6 multicast address or UDP port where the event is transmitted. This tool can be used on Unix and Windows platforms.

WBD (Shared Whiteboard Application)

This tool can be used as shared whiteboard to share ASCII texts or Adobe Post Script documents. The user that set-up the session can write on the whiteboard while the other users can only read.

NTE (Network Text Editor)

This tool has been developed by UCL. More users can edit a text in the same time, but, if two different users edit a line text in the same time, only one of the modifications is accepted.

3.3.2.3 M6Bone Services

The main service M6Bone offered is the use of videoconferencing between the involved nodes connected to the multicast network. Over this network the classical multicast tools with IPv6 support are used like VIC, RAT, NTE, etc.

Of course, any IPv6 multicast tool can be used over the network.

Another service offered is a IPv6-IPv4 gateway for multicast. With this gateway IPv4 and IPv6 sites can use the same session to transmit multicast traffic. It is a powerful transition mechanism with allows having session in the M6Bone and MBone at the same time.

3.3.2.4 Euro6IX-M6Bone Connection

Euro6IX network is connected to the M6Bone network through several partners. The following diagram shows the actual connection:

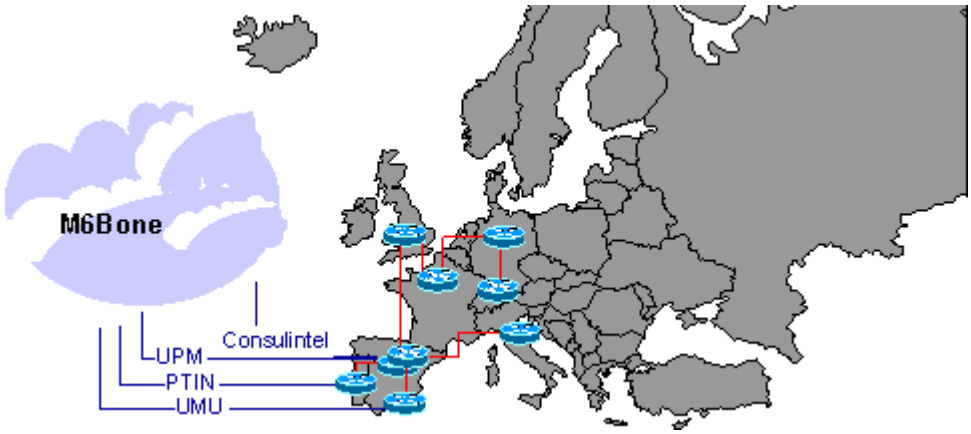

Figure 3-14: Euro6IX-M6Bone Connection

This is a transition scene until the Euro6IX IXs can be connected to the M6Bone network. Due that multicast IPv6 is a demanded and very used service, it is important to allow Euro6IX partners and clients to have a good multicast connectivity. The ideal sites from where offer this service is an IX.

Multicast technology is one of the most suitable to introduce inside an IX-based network like Euro6IX. Particularly, in a scenario where PIM-SM protocol is taken into account, it could be very interesting to put a Rendezvous-Point inside one of the IXs, in order to take advantage of the user aggregation that an IX permits to obtain.

Moreover note that most of the partners already have the connectivity with M6Bone network and this could be used, at least in a first phase, to make easier the interconnection process.

The interconnection between Euro6IX and M6Bone network could be hence realized in this way:

First scenario

Euro6IX network uses the connections already up between its partners and the M6Bone network to interconnect to the M6Bone RP managed by Renater. This scenario requires basically defining the multicast connection to M6Bone network and the suitable routing policies to be applied. This solution seems to be easier to realize than the second one described later and could be adopted in the first phase to make sure the interconnection in the short time among the two networks and to provide the multicast service to the IX users.

Second scenario

A RP is introduced inside one of the partners' IXs. In this case, Euro6IX network would have an its own RP independent on that one managed by Renater. This scenario is more complex to manage for the presence of two RPs (one in the M6Bone and another one in Euro6IX) but it would permit to analyze some interesting research issues related to the multi-domain multicast routing.

3.4 Autoconfiguration

3.4.1 Goal of Autoconfiguration Mechanisms

The goal of autoconfiguration mechanisms is to allow a network to configure itself in an automatic manner.

Usually, these mechanisms are used in the parts of the networks where the configuration tasks are expensive. This is for instance the case in LANs: up to several thousands hosts may have to be configured here. It is also the case in access networks for the configuration of customer premises equipments (CPEs). In fact, these mechanisms are mostly used at the edge of the network and not in the core. In the core, the number of devices is small and the additional complexity added by autoconfiguration mechanisms are not worth the benefits they bring.

The parameters that usually need to be configured in a network device are:

- The network prefixes or IP addresses to be used in the device.
- The address of the Domain Name System server (DNS) to use.
- The address of the default gateway (if required).

Some other parameters may also be needed sometimes (address of an NTP server, for example).

3.4.2 IPv4 Autoconfiguration Mechanism

The most common IPv4 autoconfiguration mechanism is the Dynamic Host Configuration Protocol (DHCP). It provides configuration parameters to Internet hosts.

DHCP consists of two components: a protocol for delivering host-specific configuration parameters from a DHCP server to a host and a mechanism for allocation of network addresses to hosts. It is built on a client-server model, where designated DHCP servers allocate network addresses and deliver configuration parameters to dynamically configured hosts.

DHCP is designed to supply DHCP clients with the configuration parameters required. After obtaining parameters via DHCP, a DHCP client should be able to exchange packets with any other host in the Internet.

DHCP allows but does not require the configuration of client parameters not directly related to the IP protocol. For instance, DHCP allows the configuration of the address of the DNS server to be used. DHCP also does not address registration of newly configured clients with the Domain Name System (DNS).

Last, DHCPv4 is not intended for use in configuring routers.

3.4.3 IPv6 Autoconfiguration Mechanisms

IPv6 specifications describe several types of configuration processes:

- **Manual configuration**: Parameters are written in a local configuration file in every host. This method does not require the implementation of a dedicated configuration protocol but is expensive in case of reconfiguration.

- **Stateless configuration**: Stateless auto-configuration requires no manual configuration of hosts, minimal configuration of routers, and no additional servers. It allows a host to generate its own addresses using a combination of locally available information and information advertised by routers.
- **Stateful address configuration**: This method relies on a specific protocol such as DHCP. A host that wants to obtain a parameter has to request it from a remote server.
- **Prefix delegation**: The prefix delegation mechanism provides the automated delegation of prefixes from a delegating router to another IPv6 node.

Stateless, stateful mechanisms and prefix delegation are auto-configuration mechanisms. These mechanisms are shortly described hereafter.

3.4.3.1 IPv6 Stateless Autoconfiguration

The stateless autoconfiguration mechanism mostly relies on the Neighbor Discovery protocol and on the IPv6 address structure. IPv6 addresses are made of a network prefix and of an interface identifier. Network prefixes are advertised by routers on the links while the interface identifier is built locally in the host either from the MAC address of the network card, or from a random token (for privacy purposes). From these elements, any host can build their own IPv6 addresses. The manual configuration is limited to prefix configuration in routers, hosts build automatically their IPv6 addresses and the address of the default gateway.

The role of the router is important in this method since it has to periodically advertise the prefixes to be used on the medium through the Neighbor Discovery protocol. However, the use of a router is not compulsory: if no router is used, host stations still may use link-local addresses.

3.4.3.2 IPv6 Stateful Autoconfiguration

In the stateful autoconfiguration model, hosts obtain configuration information and parameters from a server. Servers maintain a database that keeps track of which addresses have been assigned to which hosts.

DHCPv6 is the stateful autoconfiguration protocol defined by the IETF. DHCPv6 is a client/server protocol that provides managed configuration of devices. Although it relies on the same concepts, it is not an extension to the current DHCPv4 protocol, but a new protocol.

This stateful autoconfiguration protocol allows hosts to obtain addresses and other configuration information from a server.

3.4.3.3 Comparison of Stateful and Stateless Autoconfiguration

The stateless approach is useful when a site is not particularly concerned with the exact addresses hosts use, so long as they are unique and properly routable. This is usually the case in LANs.

The stateful approach is used when a site requires tighter control over exact address assignments. This may be the case in an ISP environment.

Both stateful and stateless auto-configuration may be used simultaneously. For example, a host can use stateless auto-configuration to configure its own addresses, but use stateful auto-configuration to obtain other information such as the address of the DNS server.

Another example of this coexistence is the use of the prefix delegation mechanism described below.

3.4.3.4 Prefix Delegation

DHCPv6 may be used in order to automatically delegate a prefix to an IPv6 node (most of the time, a router). This is done through the use of an IPv6 prefix option in DHCPv6. The mechanism can be used across an administrative boundary. This prefix option is currently defined in draft-ietf-dhc-dhcpv6-opt-prefix-delegation-03.txt

The prefix delegation mechanism has been first defined for automated delegation of prefixes from an ISP to its customers. Then, the delegating router is typically present at the point of presence of the ISP. It does not require knowledge about the topology of the links attached to the CPEs. A basic use case is the assignment of a /48 prefix to a CPE that itself assigns a /64 subnet from this delegated space to each of its LAN interfaces, and begins transmitting router advertisements.

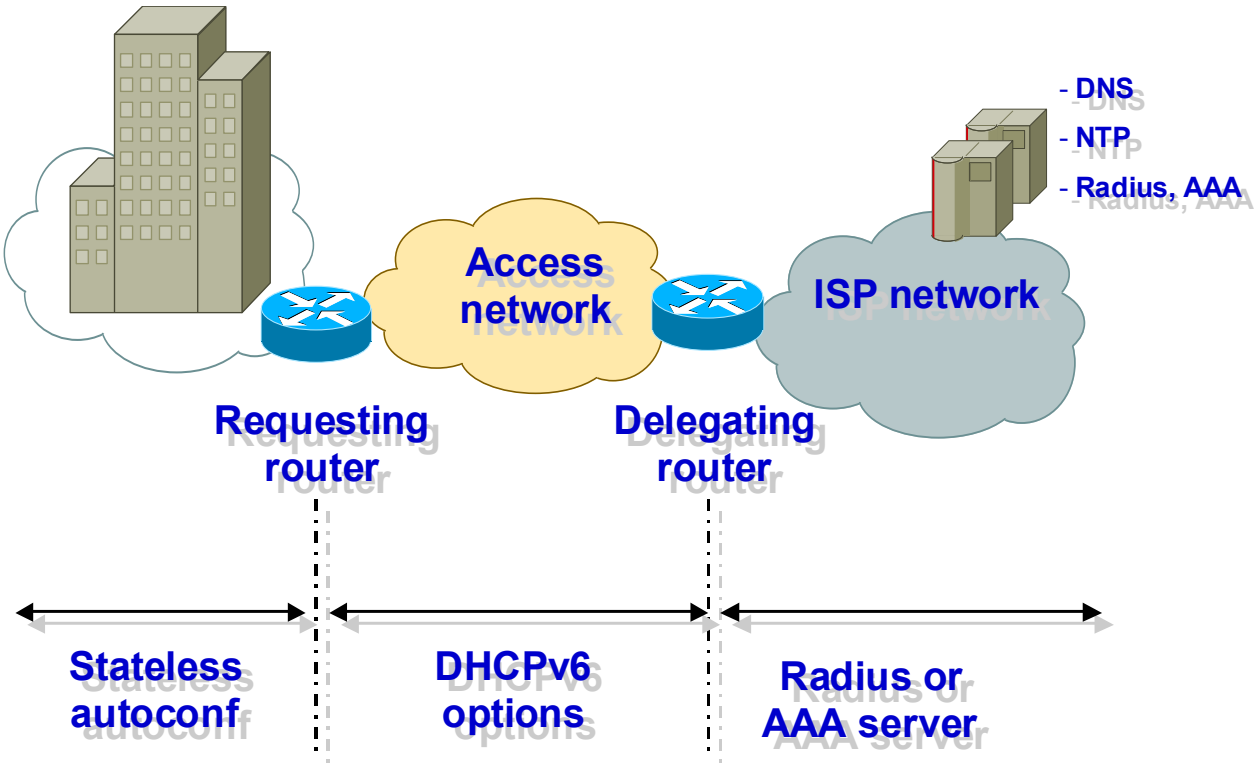

Figure 3-15: Prefix Delegation

One of the big advantages of this prefix delegation method is that it may be coupled easily with Radius or AAA servers. The process is as follow:

- When the DHCPv6 server receives a solicitation from a client, it checks if he has a configuration for this client, watching to the configuration file parsing result. At this point it uses the client duid for identification.
    - If it finds one, it sends a reply to the client, containing the prefix it found in the configuration file.

- If it cannot find any, it tries to authenticate the client using the radius (or AAA) server. It checks among client duid, client MAC address, or client link-local address if a binding exists in the hosts file.

- When a binding is matched, the DHCPv6 server asks for the associated login and password, and tries to authenticate it to the radius (or AAA) server. If the authentication is valid, a new client configuration is created, and added to the hosts configuration list (to avoid another authentication request, this client configuration is kept in memory). A new binding is added, and the server replies to the client, including the radius-acquired prefix.

3.4.3.5 DNS Dynamic Update

Once a device has been auto-configured, it has IPv6 addresses attached to its interfaces. The device is then able to initiate an IP exchange, but it has not been recorded yet in the DNS. Other correspondents have no means to obtain the IPv6 addresses of the device and cannot initiate IPv6 exchanges with it.

This is not a problem when the device only hosts client applications. But when server applications are hosted, the device must be automatically registered in the DNS. Using DNS Dynamic Updates as defined in RFC 3007 can do this.

This mechanism allows under certain conditions the addition, the modification or the deletion of a DNS record without stopping the server. This can be done zone by zone, but of course it has to be secure. To secure Dynamic Updates, one can use a simple authentication based on TSIG (Transaction Signatures) or a more complex security scheme.

When using stateful auto-configuration, the DHCPv6 server can perform the Dynamic Update.

3.4.4 Issues and On-going Research Efforts

Currently, there are no standardized methods for stateless automatic DNS discovery. Stateful mechanisms can perform this configuration thanks to the DHCPv6 DNS Configuration option (draft-ietf-dhc-dhcpv6-opt-dnsconfig-02).

This configuration option can be combined with the prefix delegation ones. Then, the delegating router can provide IPv6 addresses of name servers. The problem is that the CPE has no standard way to redistribute these addresses to IPv6 hosts located on the customer networks. Of course, it could use DHCPv6 to do that, but it would then why not using it for address configuration and default gateway configuration?

Some proposals have been made at the IETF in order to extend the current stateless auto-configuration mechanism, but no consensus was reached.

3.4.5 Impact on IPv6 IX

3.4.5.1 IX Architecture Overview

The D2.1 document "Specification of the Internal Network Architecture of each IX point" specifies the architecture of each Euro6IX Internet Exchange:

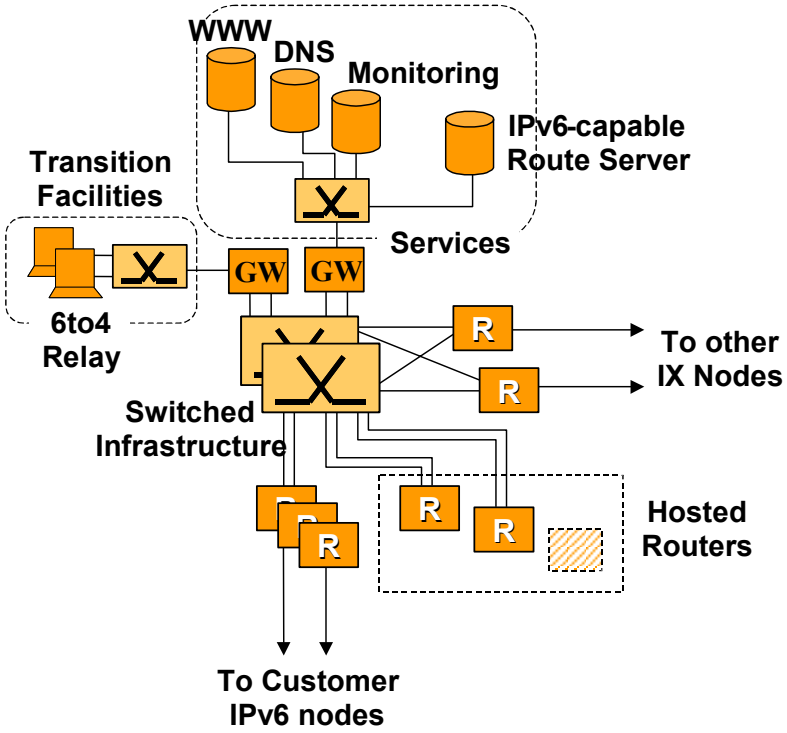

Figure 3-16: Euro6IX Architecture

This architecture includes:

- Service-Transition infrastructure (including e.g. Tunnel Broker, 6to4 relay, NAT-PT).
- Service-Farm infrastructure (including e.g. WWW and SMTP servers).
- Hosted routers and some of them may be connected to access networks.

The D2.1 specification also defines two scenarios for the Layer 3 infrastructure:

- The first scenario consists in connecting directly the customer routers to the layer 2 infrastructure. In this scenario, the router linking the IXs between each other belongs to the domain administratively managed by the Telco owning that IX.
- The second phase scenario considers on a new conception of the IX based on the so-called "layer 3 mediation function". This new role is based on the possibility, from the IX to assign IPv6 prefixes independent of the provider. In this case, each customer accessing the IX chooses the provider (one or more) and the IX assigns the IPv6 prefixes. If a customer decides to change the provider, it does not have to change IPv6 prefixes, because they are provider independent and are assigned by that particular IX.

In both cases, a /48 prefix should be enough for the IX devices.

3.4.5.2 Autoconfiguration Requirements

The first option in the IX consists in using manual configuration.

However, the IX architecture includes several hosts that could take benefits from auto-configuration mechanisms. Most of them are servers. Depending on the level of control required, one could use either stateless auto-configuration or DHCPv6. In any case for servers, DNS dynamic updates are needed and they should be performed in a secure manner.

If stateless auto-configuration is selected, a DNS discovery mechanism should be used. The most standard one for the moment is DHCPv6. Other stateless extensions could be explored too.

If stateful auto-configuration is selected, a DHCPv6 server must be deployed in the IX. It will provide configuration parameters to the hosts in the IX. The DHCPv6 server could be associated to a RADIUS or AAA server.

If an access network is directly connected to the IX, then prefix delegation should be deployed. In fact, a DHCPv6 server is then required (in the PE of the access Network) and a Radius or AAA server would also be a plus.

3.5 Multihoming

The goal of the section is to analyze issues related to multihoming, and to identify the way they may be improved or even solved within IPv6 network environment.

After a brief definition of multihoming, the current IPv4 multihoming practices are listed. Then the IPv6 multihoming requirements are detailed according to the ongoing work at IETF. Last the impact on the Euro6IX IPv6 architecture is identifying, as well as potential solutions.

3.5.1 Definition

The definition of multihoming is simple and clear: a site that is connected to more than one IP provider is a multihomed site.

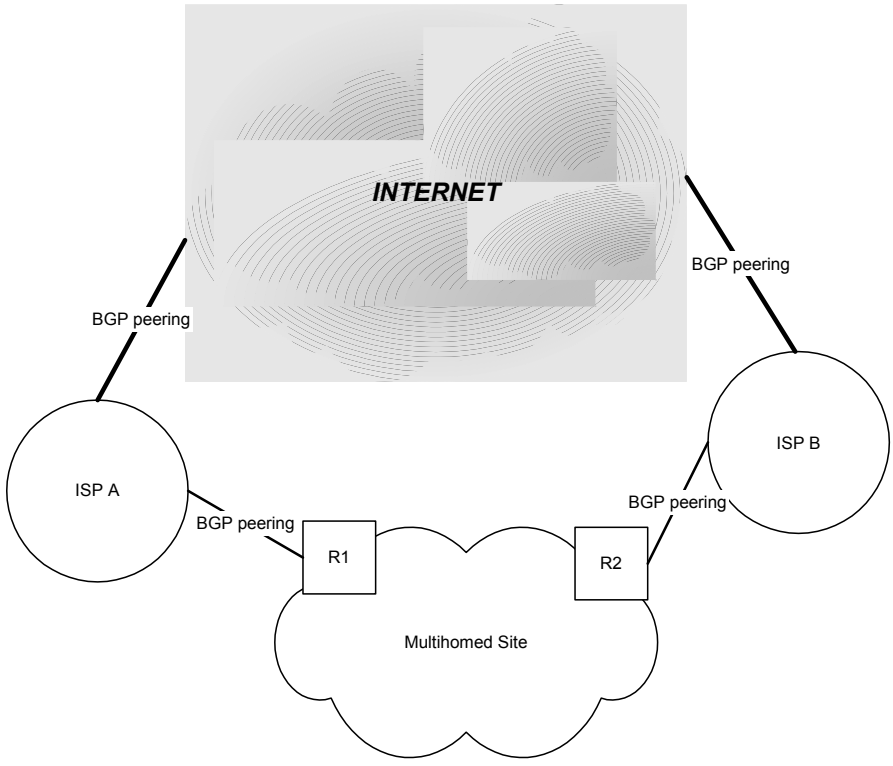

Figure 3-17: Typical Site Multihoming Architecture

The Figure 3-17 shows a typical multihoming architecture of site using dedicated routers for the connection to the ISPs.

The multihomed site announces the same route to ISP A and ISP B through its direct BGP peering. ISP A and ISP B are exchanging full routing with the global Internet. Hence, ISP A has two separate entries in its BGP table matching the multihomed site AS with different path. ISP B has two separate entries in its BGP table matching the multihomed site AS with different path, as well.

3.5.2 Multihoming Issues

This section describes the multihoming main issues:

- Routing table growth.
- DFZ convergence time increase.
- IPv6 aggregation principle break.

3.5.3 IPv4 Multihoming Practices

Since IPv6 is not as widely deployed as IPv4, a close look at IPv4 multihoming current practices must help to consider and optimise IPv6 multihoming during future deployment.

The current practices and the different benefits IPv4 multihoming intent to get are detailed below:

- Redundancy.
- Load sharing.
- Performance.
- Policy.

3.5.4 IPv6 Requirements

This section describes the requirement IPv6 multihoming has to meet.

3.5.5 Impact on IPv6 IX

This section analyses the impact and the consequences of multihoming on IPv6 IXs, and will discuss the applicability to Euro6IX test-bed.

4. SUMMARY AND CONCLUSIONS

This document has presented issues related to designing and maintaining the IPv6 network. Its main focus includes IXs, their interconnection and network features. All the concepts and solutions have been treated related to the operational Euro6IX network. Most of the features described are already implemented and tested, which means that the solutions presented are reliable and may be used by the other operators.

Starting from the Internal Connectivity, through External Connectivity finishing with Network Features like QoS, Security, Multicast, Autoconfiguration, Multihoming this deliverable covers topics, which any network operator willing to implement IPv6 network may come across. It also provides references to other documents - both project - internal and external - which might be of help.

With this information, it should be possible to design an operational and reliable IPv6 network.

5. REFERENCES

[L2, L3] Euro6IX Deliverable D2.1

[6to4] Euro6IX Deliverable D2.1

[DNS] Euro6IX Deliverable D2.1

[AAA] Euro6IX Deliverable D2.1

[Routing] Euro6IX Deliverable D2.2

[network topology] Euro6IX Deliverable D2.2

[Addressing] Euro6IX Deliverable D2.2

[M6BONE] M6bone Home Page [www]. www.m6bone.net. December 2002.

[Tunnel] http://vtun.sourceforge.net/

[UCL] UCL Networked Multimedia Research Group [www]. http://www-mice.cs.ucl.ac.uk/multimedia/software/. December 2002.

[6BONE] 6Bone Home Page [www]. www.6bone.net. December 2002.

[FBSD] FreeBSD Home Page [www]. www.freebsd.org. December 2002.