Enabling IPv6 in the Internet Exchanges

Jordi Palet (jordi.palet@consulintel.es)
CEO/CTO - Consulintel

NAPLA 2004

October/04, Sao Paulo

Agenda

1. Introduction & Background
2. Header Formats
3. Addressing & Routing
4. Security
5. Quality of Service
6. Mobility
7. Autoconfiguration
8. Multicast
9. IPv4-v6 Coexistence & Transition
10. Euro6IX
11. Conclusions

IPv6 Tutorial

1. Introduction & Background

Why a New IP?

Only *compelling* reason: more addresses!

- for billions of new devices,
  e.g., cell phones, PDAs, appliances, cars, etc.
- for billions of new users,
  e.g., in China, India, etc.
- for “always-on” access technologies,
  e.g., xDSL, cable, ethernet-to-the-home, etc.

But Isn't There Still Lots of IPv4 Address Space Left?

- ~ Half the IPv4 space is unallocated
  - if size of Internet is doubling each year, does this mean only one year's worth?!
- No, because today we deny unique IPv4 addresses to most new hosts
  - we make them use methods like NAT, PPP, etc. to share addresses
- But new types of applications and new types of access need unique addresses!

Why Are NAT's Not Adequate?

- They won't work for large numbers of "servers", i.e., devices that are "called" by others (e.g., IP phones)
- They inhibit deployment of new applications and services
- They compromise the performance, robustness, security, and manageability of the Internet

Incidental Benefits of Bigger Addresses

- Easy address auto-configuration
- Easier address management/delegation
- Room for more levels of hierarchy, for route aggregation
- Ability to do end-to-end IPsec (because NATs not needed)

Incidental Benefits of New Deployment

- Chance to eliminate some complexity, e.g., in IP header
- Chance to upgrade functionality, e.g., multicast, QoS, mobility
- Chance to include new enabling features, e.g., binding updates

Summary of Main IPv6 Benefits

- Expanded addressing capabilities
- Server-less autoconfiguration (“plug-n-play”) and reconfiguration
- More efficient and robust mobility mechanisms
- Built-in, strong IP-layer encryption and authentication
- Streamlined header format and flow identification
- Improved support for options / extensions

Why Was 128 Bits Chosen as the IPv6 Address Size?

- Some wanted fixed-length, 64-bit addresses
  - easily good for 10^{12} sites, 10^{15} nodes, at .0001 allocation efficiency (3 orders of mag. more than IPng requirement)
  - minimizes growth of per-packet header overhead
  - efficient for software processing
- Some wanted variable-length, up to 160 bits
  - compatible with OSI NSAP addressing plans
  - big enough for autoconfiguration using IEEE 802 addresses
  - could start with addresses shorter than 64 bits & grow later
- Settled on fixed-length, 128-bit addresses
  - (340,282,366,920,938,463,463,374,607,431,768,211,456 in all!)

What Ever Happened to IPv5?

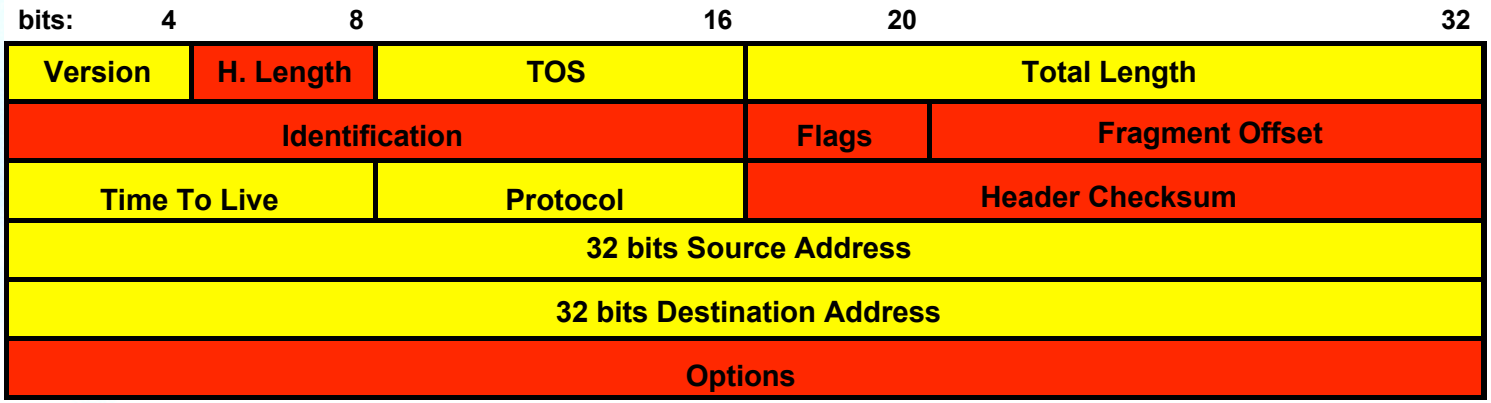
0–3		unassigned
4	IPv4	(today's widespread version of IP)
5	ST	(Stream Protocol, not a new IP)
6	IPv6	(formerly SIP, SIPP)
7	CATNIP	(formerly IPv7, TP/IX; deprecated)
8	PIP	(deprecated)
9	TUBA	(deprecated)
10-15		unassigned

IPv6 Tutorial

2. Header Formats

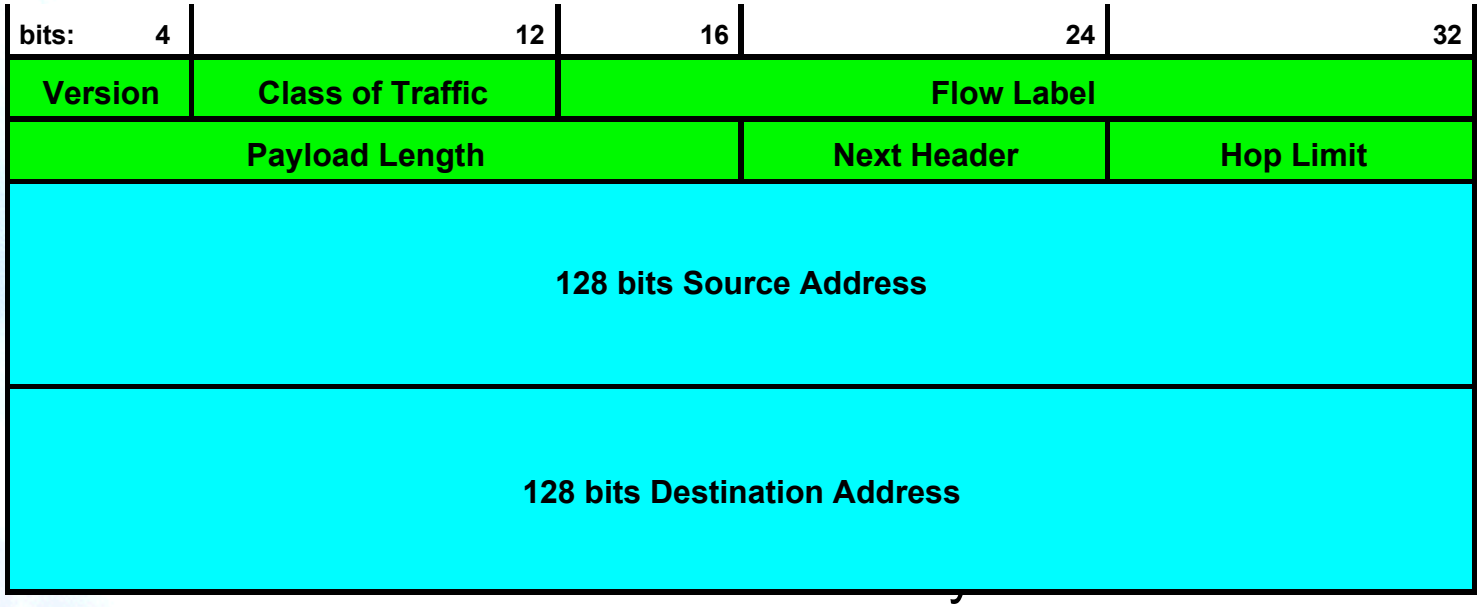
IPv4 Header Format

- 20 Bytes + Options



IPv6 Header Format

- From 12 to 8 Fields (40 bytes)



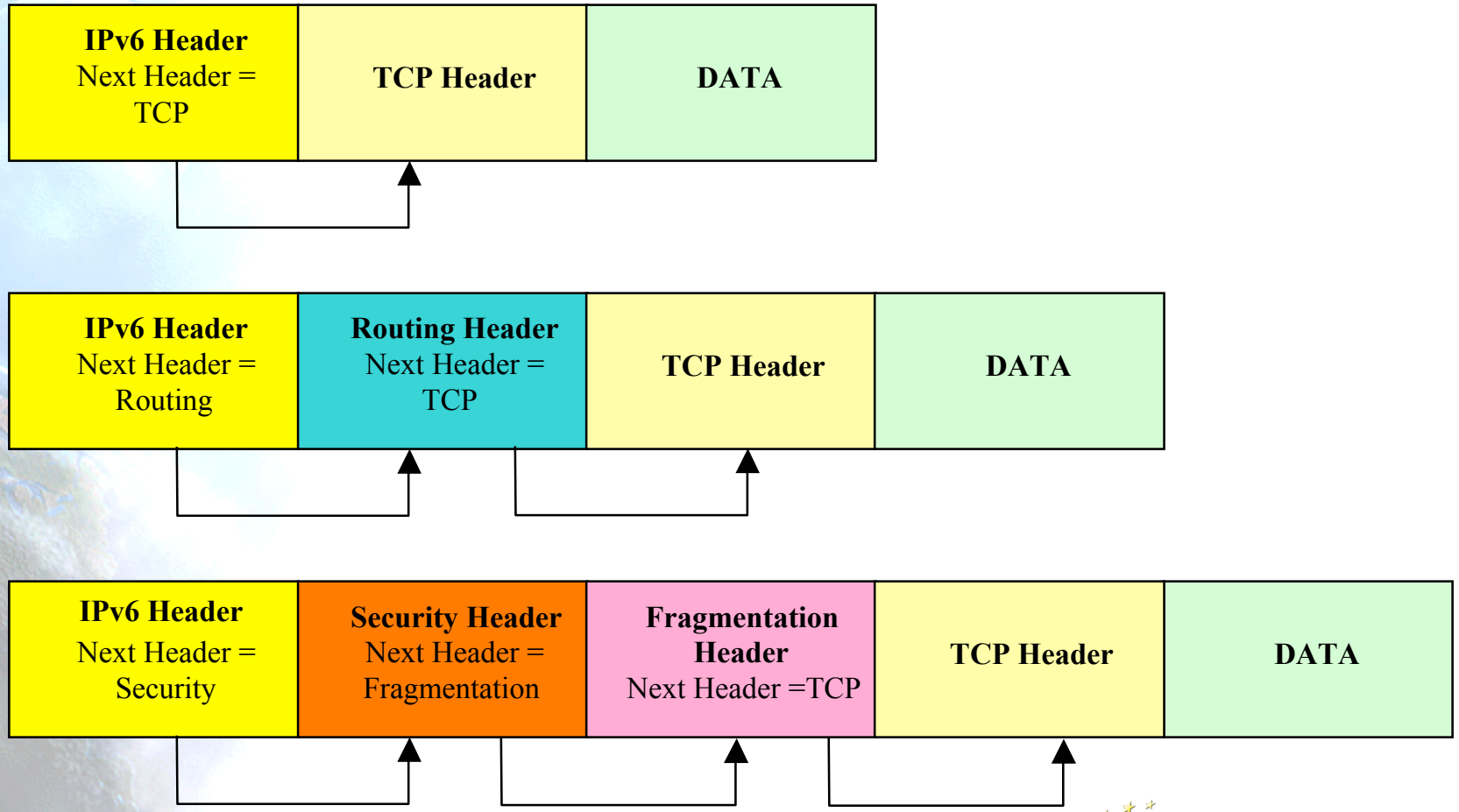
  - Fragmentation end to end

Summary of Header Changes

- 40 bytes
- Address increased from 32 to 128 bits
- Fragmentation and options fields removed from base header
- Header checksum removed
- Header length is only payload (because fixed length header)
- New Flow Label field
- TOS -> Traffic Class
- Protocol -> Next Header (extension headers)
- Time To Live -> Hop Limit
- Alignment changed to 64 bits

Extension Headers

- "Next Header" Field



Extension Headers Goodies

- Processed Only by Destination Node
  - Exception: Hop-by-Hop Options Header
- No more “40 byte limit” on options (IPv4)
- Extension Headers defined currently:
  - Hop-by-Hop Options
  - Routing
  - Fragment
  - Authentication (RFC 2402, next header = 51)
  - Encapsulating Security Payload (RFC 2406, next header = 50)
  - Destination Options

IPv6 Tutorial

3. Addressing and Routing

Text Representation of Addresses

“Preferred” form:	1080:0:FF:0:8:800:200C:417A
Compressed form:	FF01:0:0:0:0:0:0:43
	becomes FF01::43
IPv4-compatible:	0:0:0:0:0:0:13.1.68.3
	or ::13.1.68.3
URL:	http://[FF01::43]/index.html

Address Types

- Unicast (one-to-one)
  - global
  - link-local
  - site-local (deprecated)
  - IPv4-compatible
- Multicast (one-to-many)
- Anycast (one-to-nearest)
- Reserved

Interface IDs

The lowest-order 64-bit field of unicast addresses may be assigned in several different ways:

- auto-configured from a 48-bit MAC address (e.g., Ethernet address), expanded into a 64-bit EUI-64
- assigned via DHCP
- manually configured
- auto-generated pseudo-random number (to counter some privacy concerns)
- possibly other methods in the future

Some Special-Purpose Unicast Addresses

- The unspecified address, used as a placeholder when no address is available:

  0:0:0:0:0:0:0:0

- The loopback address, for sending packets to self:

  0:0:0:0:0:0:0:1

Routing

- Uses same “longest-prefix match” routing as IPv4 CIDR
- Straightforward changes to existing IPv4 routing protocols to handle bigger addresses
  - –unicast: OSPF, RIP-II, IS-IS, BGP4+, …
  - –multicast: MOSPF, PIM, …
- Can use Routing header with anycast addresses to route packets through particular regions
  - –e.g., for provider selection, policy, performance, etc.

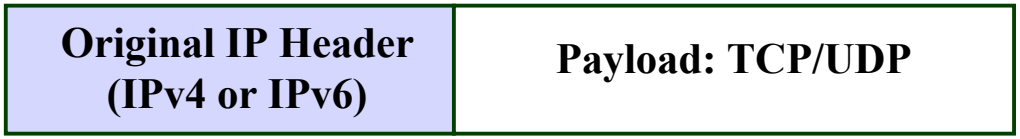
IPv6 Tutorial

4. Security

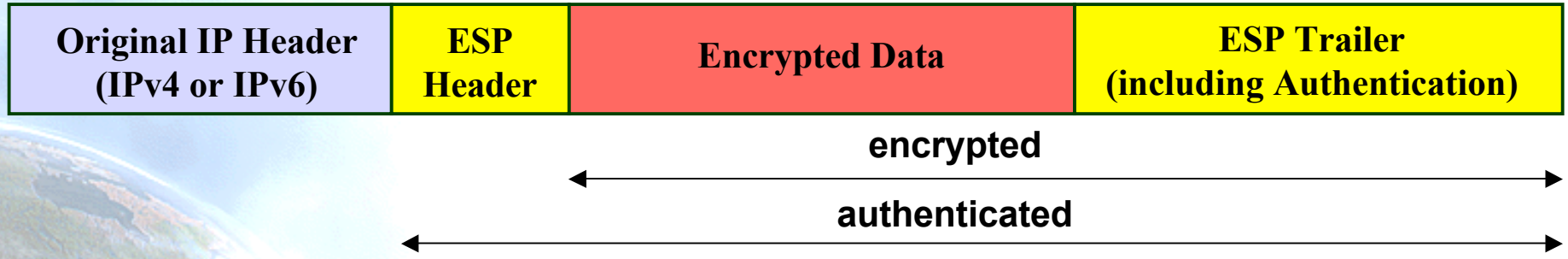
IPv6 Security

- IPsec is part of the IPv6 "core" specs:
  - All implementations expected to support authentication and encryption headers ("IPsec")
- Authentication separate from encryption for use in situations where encryption is prohibited or prohibitively expensive
- Key distribution protocols are under development (independent of IP v4/v6)
- Support for manual key configuration required

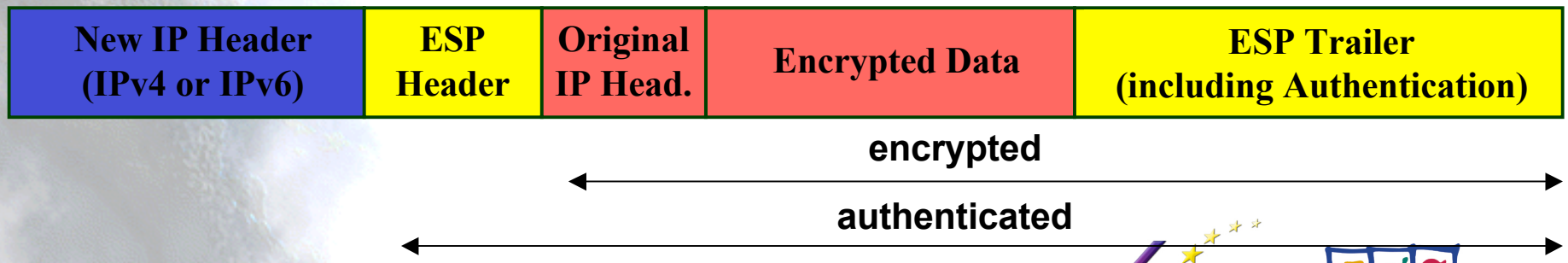
Transport and Tunnel Modes



Transport Mode



Tunnel Mode



IPv6 Tutorial

5. Quality of Service

Concept of QoS

- Quality: Reliable delivery of data ("better than normal")
  - Data loss
  - Latency
  - Jittering
  - Bandwidth
- Service: Anything offered to the user
  - Communication
  - Transport
  - Application

Abstract

- “Quality of Service is a measurement of the network behavior with respect to certain characteristics of defined services” !!!!!

- Common concepts to all definitions of QoS:
  - Traffic and type of service differentiation
  - Users may be able to treat one or more traffic classes differently

IP Quality of Service Approaches

Two basic approaches developed by IETF:

- "Integrated Service" (int-serv)
  - fine-grain (per-flow), quantitative promises (e.g., x bits per second), uses RSVP signalling
- "Differentiated Service" (diff-serv)
  - coarse-grain (per-class), qualitative promises (e.g., higher priority), no explicit signalling

IPv6 Support for Int-Serv

20-bit Flow Label field to identify specific flows needing special QoS

- each source chooses its own Flow Label values; routers use Source Addr + Flow Label to identify distinct flows
- Flow Label value of 0 used when no special QoS requested (the common case today)
- this part of IPv6 is not standardized yet, and may well change semantics in the future

IPv6 Support for Diff-Serv

8-bit Traffic Class field to identify specific classes of packets needing special QoS

- same as new definition of IPv4 Type-of-Service byte
- may be initialized by source or by router enroute; may be rewritten by routers enroute
- traffic Class value of 0 used when no special QoS requested (the common case today)

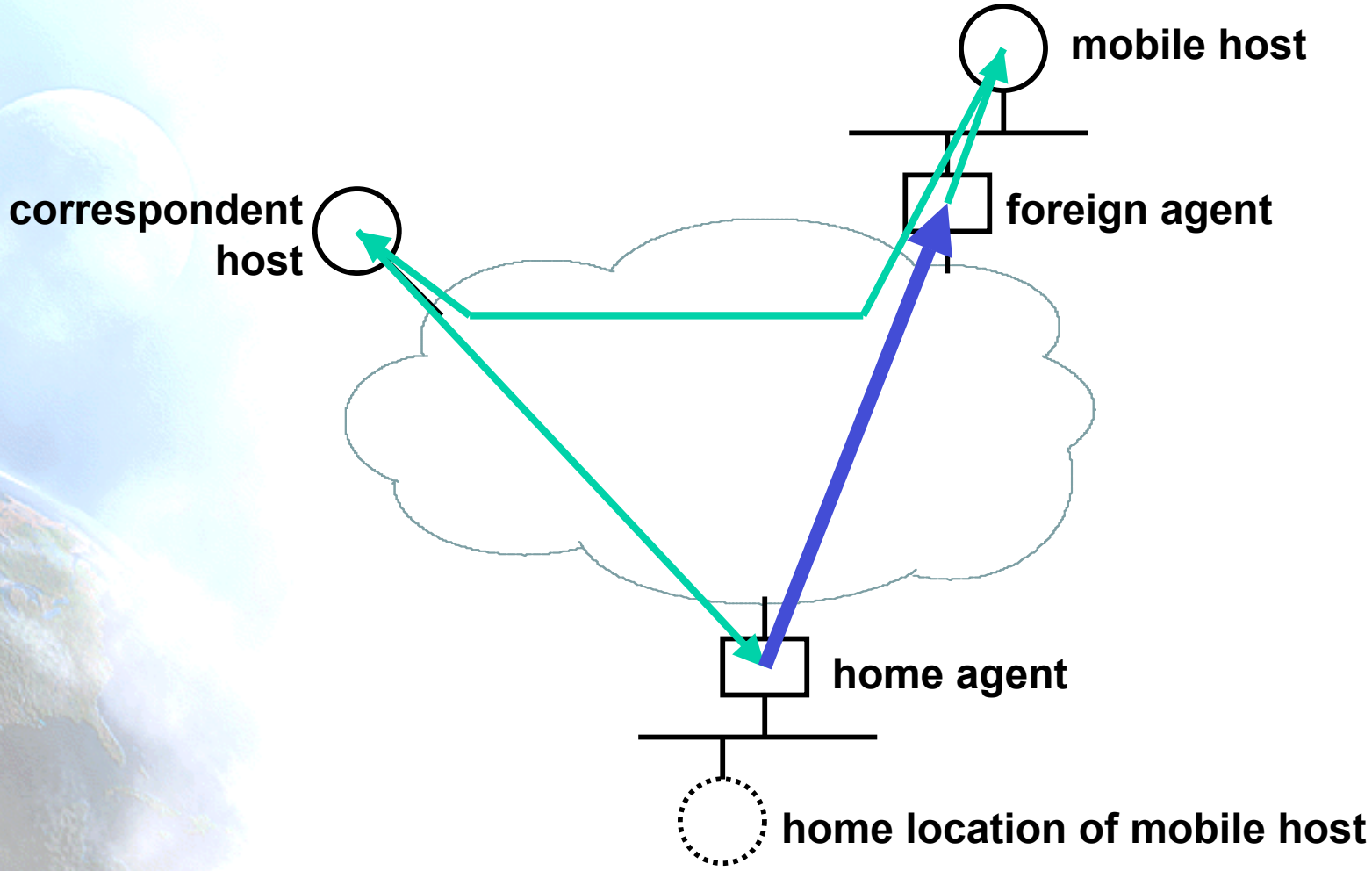
IPv6 Tutorial

6. Mobility

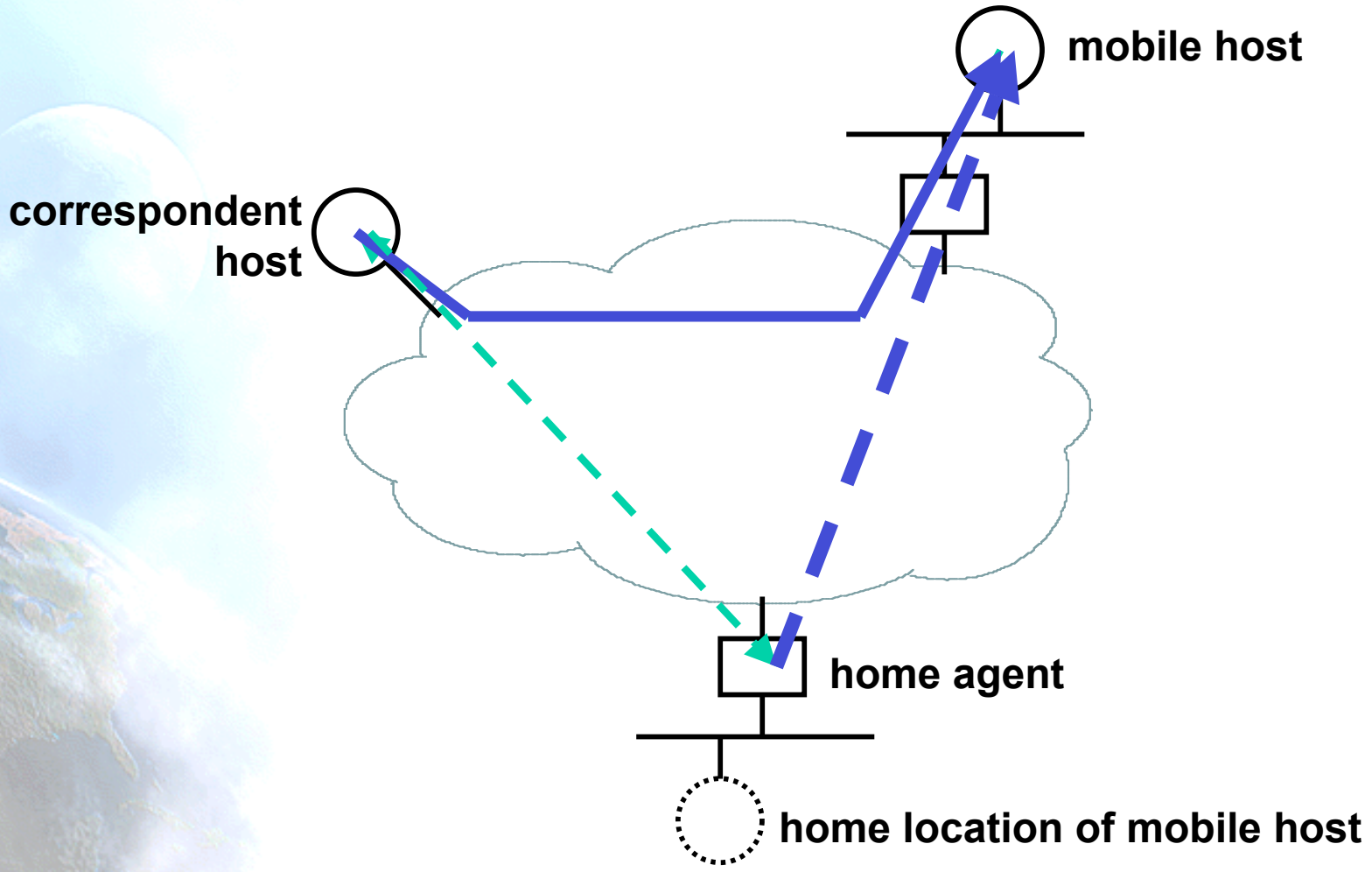
IPv6 Mobility

- A mobile host has one or more home address(es)
  - relatively stable; associated with host name in DNS
- When it discovers it is in a foreign subnet (i.e., not its home subnet), it acquires a foreign address
  - uses auto-configuration to get the address
  - registers the foreign address with a home agent, i.e, a router on its home subnet
- Packets sent to the mobile's home address(es) are intercepted by home agent and forwarded to the foreign address, using encapsulation

Mobile IP (v4 version)



Mobile IP (v6 version)



Standards

- Mobility Support in IPv6
  - RFC3775 – June 2004
- Using IPsec to Protect Mobile IPv6 Signaling Between Mobile Nodes and Home Agents
  - RFC3776 – June 2004

IPv6 Tutorial

7. Autoconfiguration

Stateless or Serverless Autoconfiguration

- Stateless mechanism allows a host to generate its own addresses using a combination of locally available information and information advertised by routers.
- Routers advertise prefixes that identify the subnet(s) associated with a link.
- Hosts generate an "interface identifier" that uniquely identifies an interface on a subnet, locally generated, e.g., using MAC address.
- An address is formed by combining the both.
- In the absence of routers, a host can only generate link-local addresses.
- Link-local addresses are sufficient for allowing communication among nodes attached to the same link.

Stateful Autoconfiguration

- Hosts obtain interface addresses and/or configuration information and parameters from a server.
- Servers maintain a database that keeps track of which addresses have been assigned to which hosts.
- Stateless and stateful autoconfiguration complement each other.
- Both stateful and stateless address autoconfiguration may be used simultaneously.
- The site administrator specifies which type of autoconfiguration to use through the setting of appropriate fields in Router Advertisement messages.

Address Life Time

- IPv6 addresses are leased to an interface for a fixed (possibly infinite) length of time, that indicates how long the address is bound to an interface.
- When a lifetime expires, the binding (and address) become invalid and the address may be reassigned to another interface elsewhere in the Internet.
- To handle the expiration of address bindings gracefully, an address goes through two distinct phases while assigned to an interface.
  - Initially, an address is "preferred", meaning that its use in arbitrary communication is unrestricted.
  - Later, an address becomes "deprecated" in anticipation that its current interface binding will become invalid.

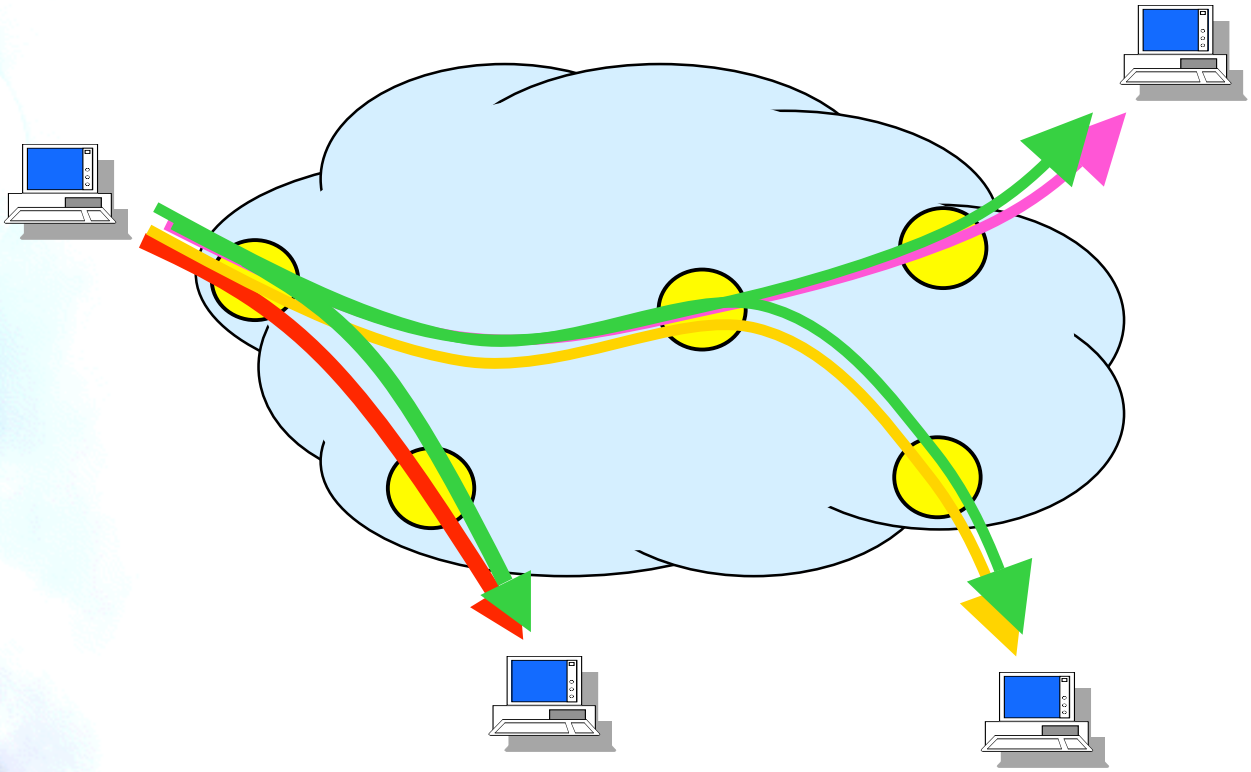
Duplicate Address Detection

- To insure that all configured addresses are likely to be unique on a given link, nodes run a "duplicate address detection" algorithm on addresses before assigning them to an interface.
- The Duplicate Address Detection algorithm is performed on all addresses, independent of whether they are obtained via stateless or stateful autoconfiguration.
- The procedure for detecting duplicate addresses uses Neighbor Solicitation and Advertisement messages.
- Since host autoconfiguration uses information advertised by routers, routers will need to be configured by some other means. However, it is expected that routers will generate link-local addresses using the same mechanism.
- Routers are expected to successfully pass the Duplicate Address Detection procedure on all addresses prior to assigning them to an interface.

IPv6 Tutorial

8. Multicast

What's Multicast?



Applications

- Distributed systems
- Video on Demand (VoD)
- Radio/TV Diffusion
- Multipoint Conferencing (voice/video)
- Network Gaming
- Network level functions

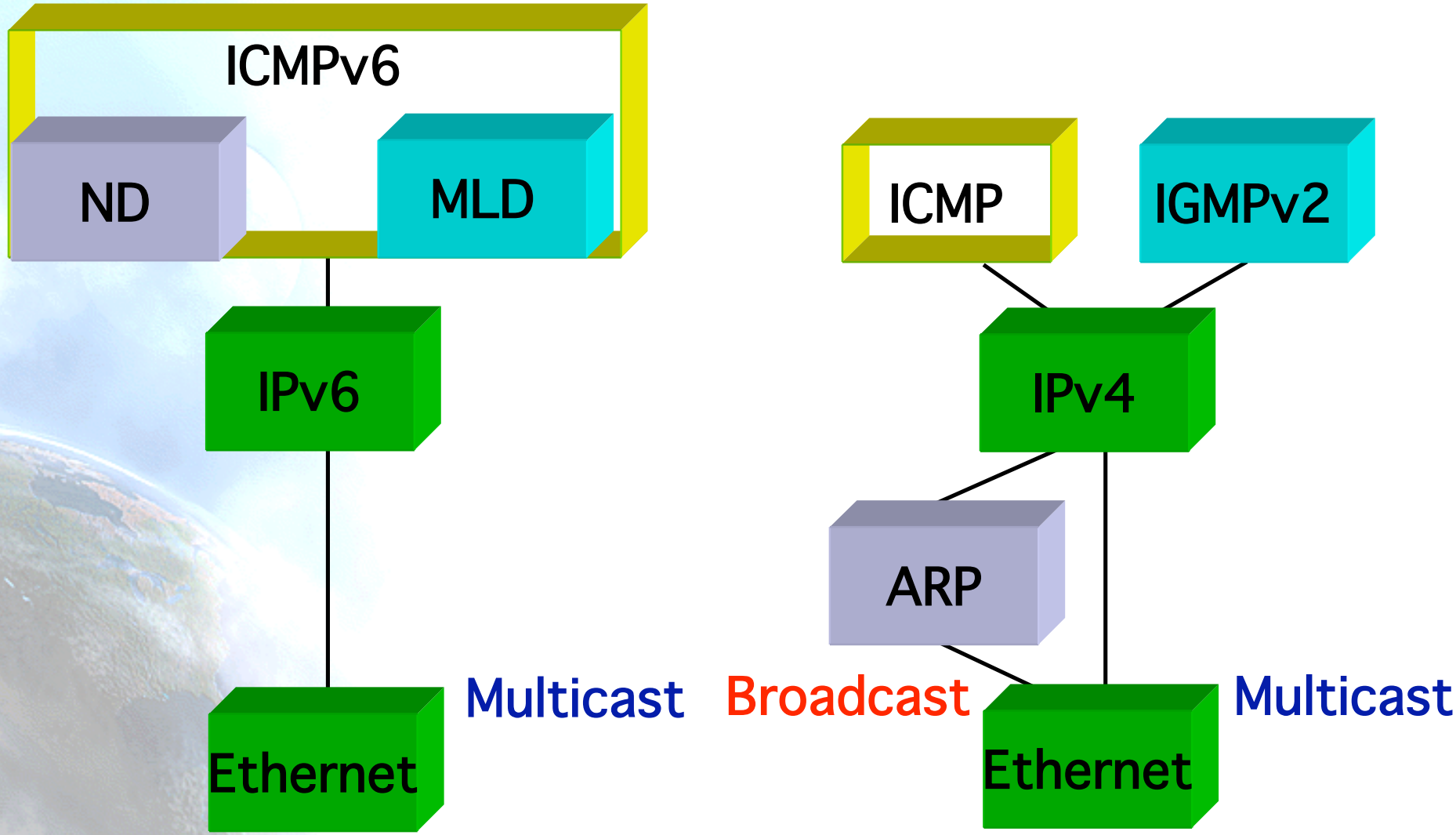
How it Works ?

- The host joins/signoff the multicast group
- No restriction about number of groups or members per group
- Sending to the group don't means belonging to it
- The destination address is a group address (multicast address)
- Connection-Less service

IPv4 vs. IPv6

- IPv4
  - Broadcast
    - Limited: 255.255.255.255
    - Directed: <network>11..1
  - Multicast
    - D Class:
      224.0.0.0 - 239.255.255.255
- IPv6
  - Multicast

Control Plane IPv4 vs. IPv6



IPv6 Tutorial

9. IPv4-IPv6 Coexistence & Transition

Transition / Co-Existence Techniques

A wide range of techniques have been identified and implemented, basically falling into three categories:

(1) dual-stack techniques, to allow IPv4 and IPv6 to co-exist in the same devices and networks

(2) tunneling techniques, to avoid order dependencies when upgrading hosts, routers, or regions

(3) translation techniques, to allow IPv6-only devices to communicate with IPv4-only devices

Expect all of these to be used, in combination

Dual-Stack Approach

- When adding IPv6 to a system, do not delete IPv4
  - this multi-protocol approach is familiar and well-understood (e.g., for AppleTalk, IPX, etc.)
  - note: in most cases, IPv6 will be bundled with new OS releases, not an extra-cost add-on
- Applications (or libraries) choose IP version to use
  - when initiating, based on DNS response:
    - if (dest has AAAA or A6 record) use IPv6, else use IPv4
  - when responding, based on version of initiating packet
- This allows indefinite co-existence of IPv4 and IPv6, and gradual app-by-app upgrades to IPv6 usage

Tunnels to Get Through IPv6-Ignorant Routers

- Encapsulate IPv6 packets inside IPv4 packets (or MPLS frames)
- Many methods exist for establishing tunnels:
  - manual configuration
  - "tunnel brokers" (using web-based service to create a tunnel)
  - "6-over-4" (intra-domain, using IPv4 multicast as virtual LAN)
  - "6-to-4" (inter-domain, using IPv4 addr as IPv6 site prefix)
- Can view this as:
  - IPv6 using IPv4 as a virtual link-layer, or
  - an IPv6 VPN (virtual public network), over the IPv4 Internet (becoming "less virtual" over time, we hope)

Translation

- May prefer to use IPv6-IPv4 protocol translation for:
  - new kinds of Internet devices (e.g., cell phones, cars, appliances)
  - benefits of shedding IPv4 stack (e.g., serverless autoconfig)
- This is a simple extension to NAT techniques, to translate header format as well as addresses
  - IPv6 nodes behind a translator get full IPv6 functionality when talking to other IPv6 nodes located anywhere
  - they get the normal (i.e., degraded) NAT functionality when talking to IPv4 devices
  - methods used to improve NAT functionality (e.g, RSIP) can be used equally to improve IPv6-IPv4 functionality

Enabling IPv6 in the IX

10. Euro6IX

Euro6IX: The Concept

- How to pronounce it: forget IX and read 6 (“SIX”)
- Build a large, scalable and native IPv6 Backbone of Traffic Exchanges, with connectivity across Europe and other IPv4/v6 Exchangers
- In order to promote and allow other players to trial v6 and port/develop key applications and services
- In order to break the chicken and egg issue !
- Gain REAL IPv6 experience, in a real world with not just research users, involving Telcos/ISPs/ASPs, among others: Allow new players into our trials
- Bring IPv6 into a production transit service

Euro6IX Goal

- Support the fast introduction of IPv6 in Europe.
- Main Steps:
  - Network design & deployment
  - Research on network advanced services
  - Development of applications validated by user groups & international trials
  - Active dissemination:
    - participation in events/conferences/papers
    - contributions to standards
    - project web site

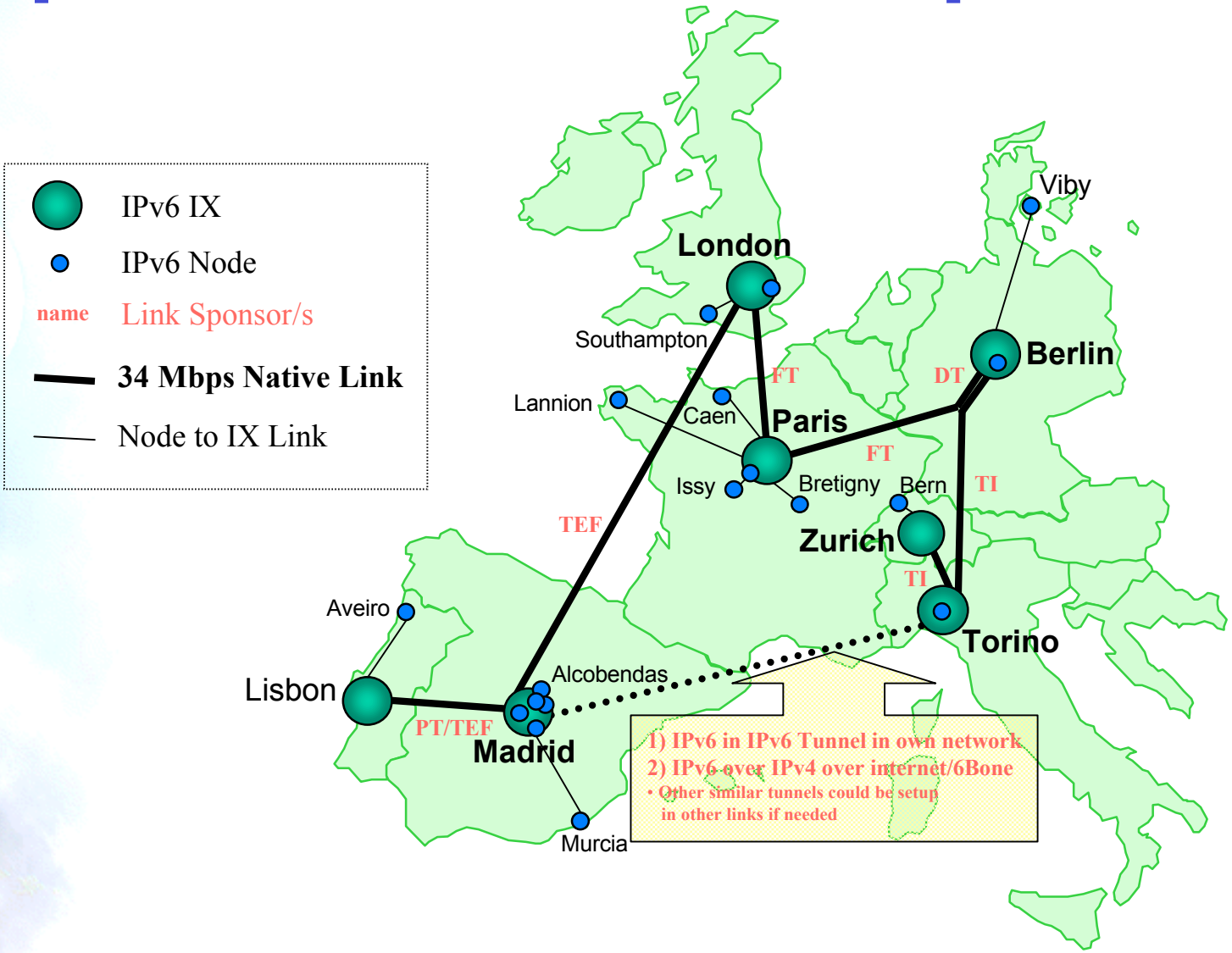
Objectives

1. Research an appropriate architecture, to design and deploy the first Pan-European non-commercial IPv6 Internet Exchange Network.
2. Use this infrastructure to research, test and validate IPv6-based applications & services.
3. Open the network to specific User Groups for its validation in trials.
4. Dissemination, liaison and coordination with clusters, fora, standards organizations (e.g. IETF, RIPE) and third parties.

Consortium Members (17)

- Telcos/ISPs (7):
  - Telecom Italia LAB (WP2 leader), Telefónica I+D (WP3 leader and project coordinator), Airtel-Vodafone, British Telecom Exact, T-Nova (Deutsche Telecom), France Telecom RD, Portugal Telecom Inovação
- Industrial (2):
  - 6Wind, Ericsson Telebit
- Universities (3):
  - Technical University of Madrid (WP4 leader), University of Southampton, University of Murcia
- Research, System Integrators and Consultancy (3):
  - Consulintel (WP1 leader and project coordinator), Telscom (WP5 leader), novaGnet systems
- Others (2):
  - Écija & Asociados Abogados, Eurocontrol

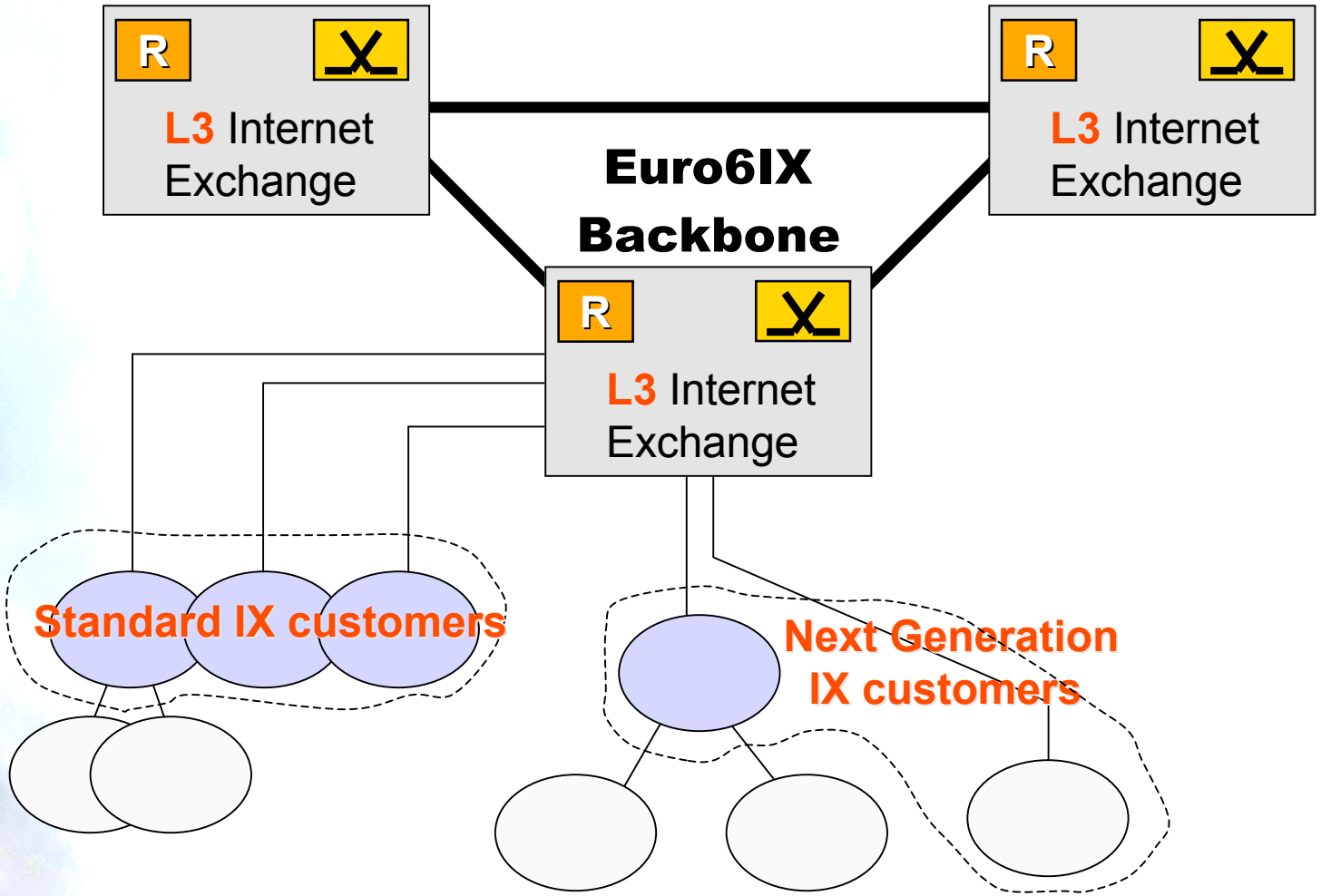
Updated Network Map



Layer 3 IX

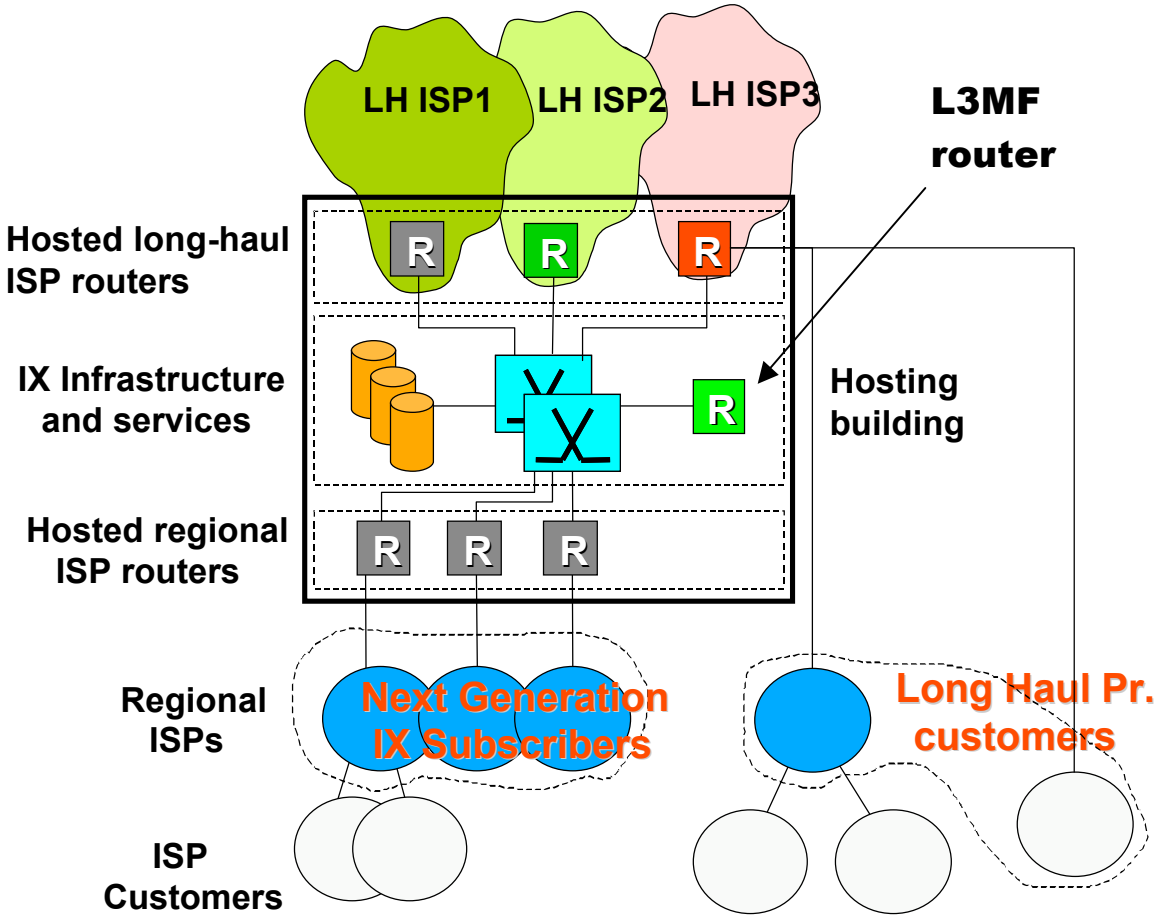
- Infrastructure providing both layer 2 and layer 3 interconnection service.
- Several IXs can make direct peering offering also Wide Area Layer 3 transport as an Internet Service Provider. Every IXs will use an assigned xTLA prefix (x=p or s) to assign NLA prefixes to ISPs or customers connecting to the IX.
- Project partners will use their xTLA prefix to assign NAL to customers and regional ISP connecting to IX.

Layer 3 IXs Network Architecture



IX Model C

- L2 infrastructure (fully redundant) where the IX services are placed
- Routers infrastructure (long-haul providers and customers)
- Layer 3 mediation function router (L3MF) is the real new element of this model

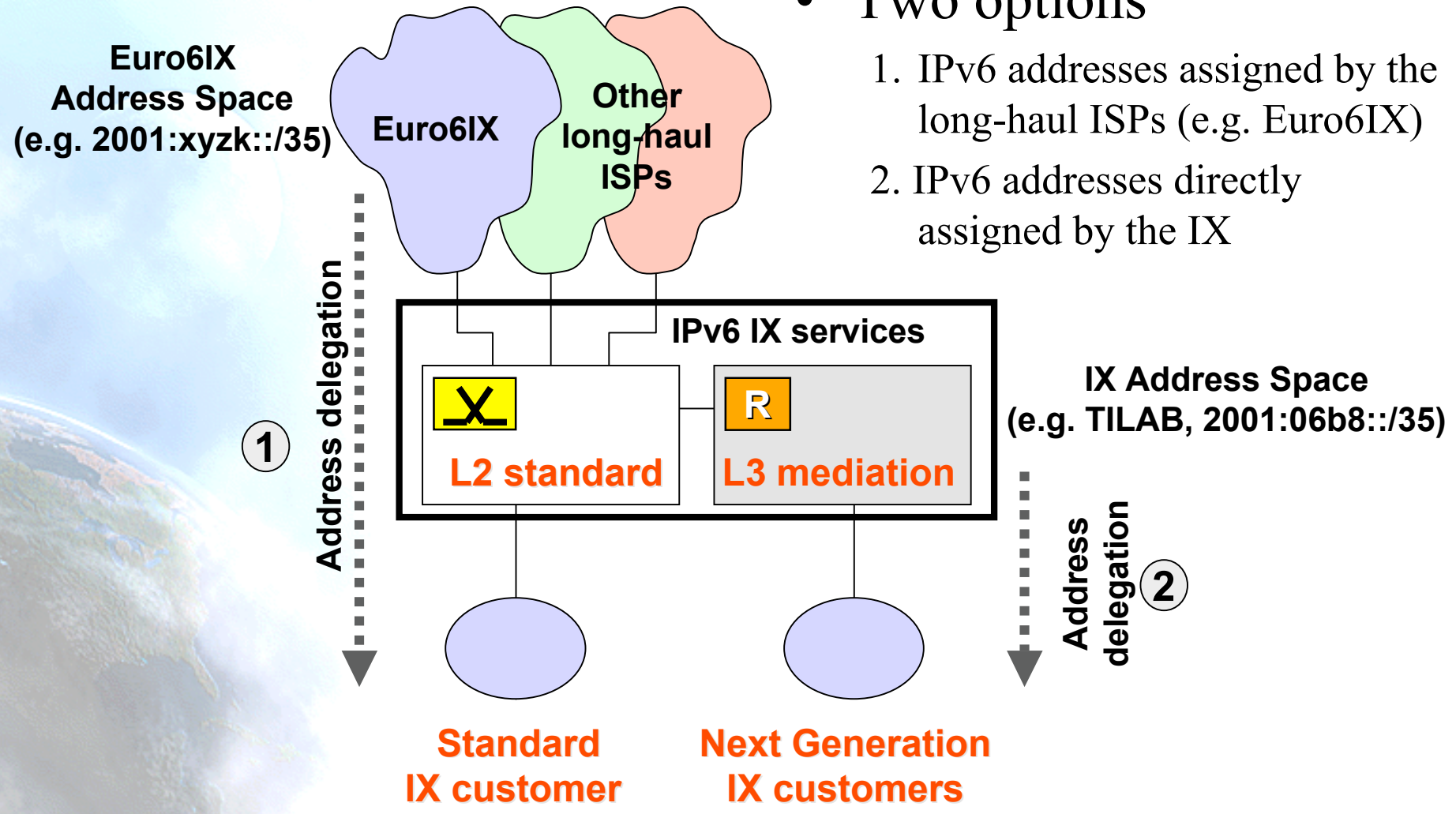


RFC2374 Benefits

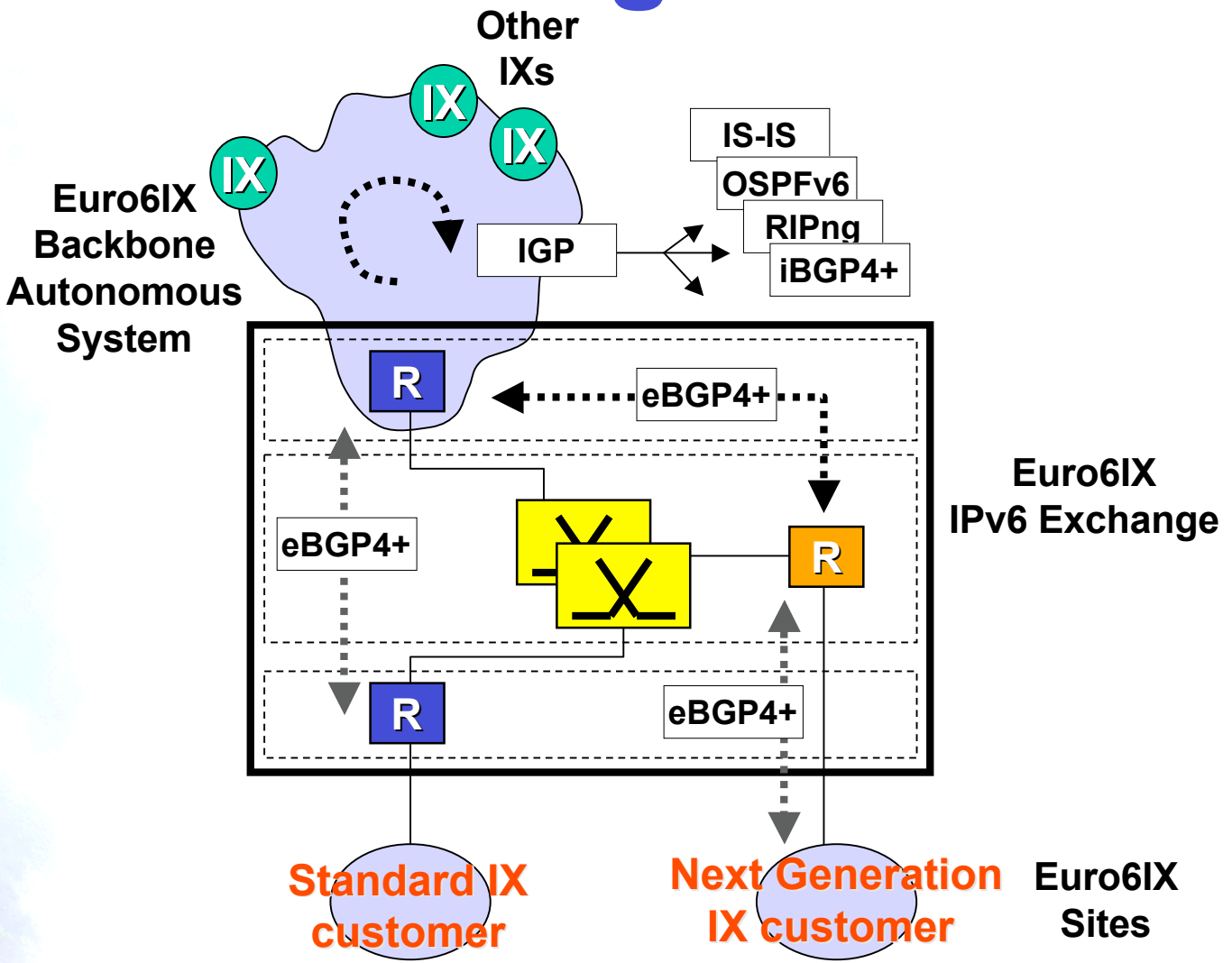
- This model is based on the RFC 2374 to verify that:
  - a customer could change its service provider without changing its addressing space
  - the renumbering functionality could be realized more easily (no renumbering in the better case)
  - the multihoming functionality could be realized more easily
- IX plays an intermediation role between the ISP and the customers (Layer 3 mediation function router)
- Routing:
  - iBGP+IGP: inside the Long Haul Provider
  - Euro6IX is the collection of the routers inside the IX emulating the LHP (single AS)
  - eBGP4+: between the customers and the IX
  - eBGP4+: between the IX and the LHPs

Address Assignment

- Two options
  1. IPv6 addresses assigned by the long-haul ISPs (e.g. Euro6IX)
  2. IPv6 addresses directly assigned by the IX



Routing


Other
IXs
IX
IX
IX
Euro6IX
Backbone
Autonomous
System
IS-IS
OSPFv6
RIPng
iBGP4+
IGP
R
eBGP4+
eBGP4+
R
R
eBGP4+
Euro6IX
IPv6 Exchange
Standard IX
customer
Next Generation
IX customer
Euro6IX
Sites

Mobility

- Definition of mobility scenarios for IPv6
- Identification of macro-mobility technologies to be used in the test-beds
- First Identification and evaluation of available implementations for macro-mobility for a common platform
- Selection of access technologies to be used in the test-beds
- Every participant will design their own access network based on the available implementations identified before.

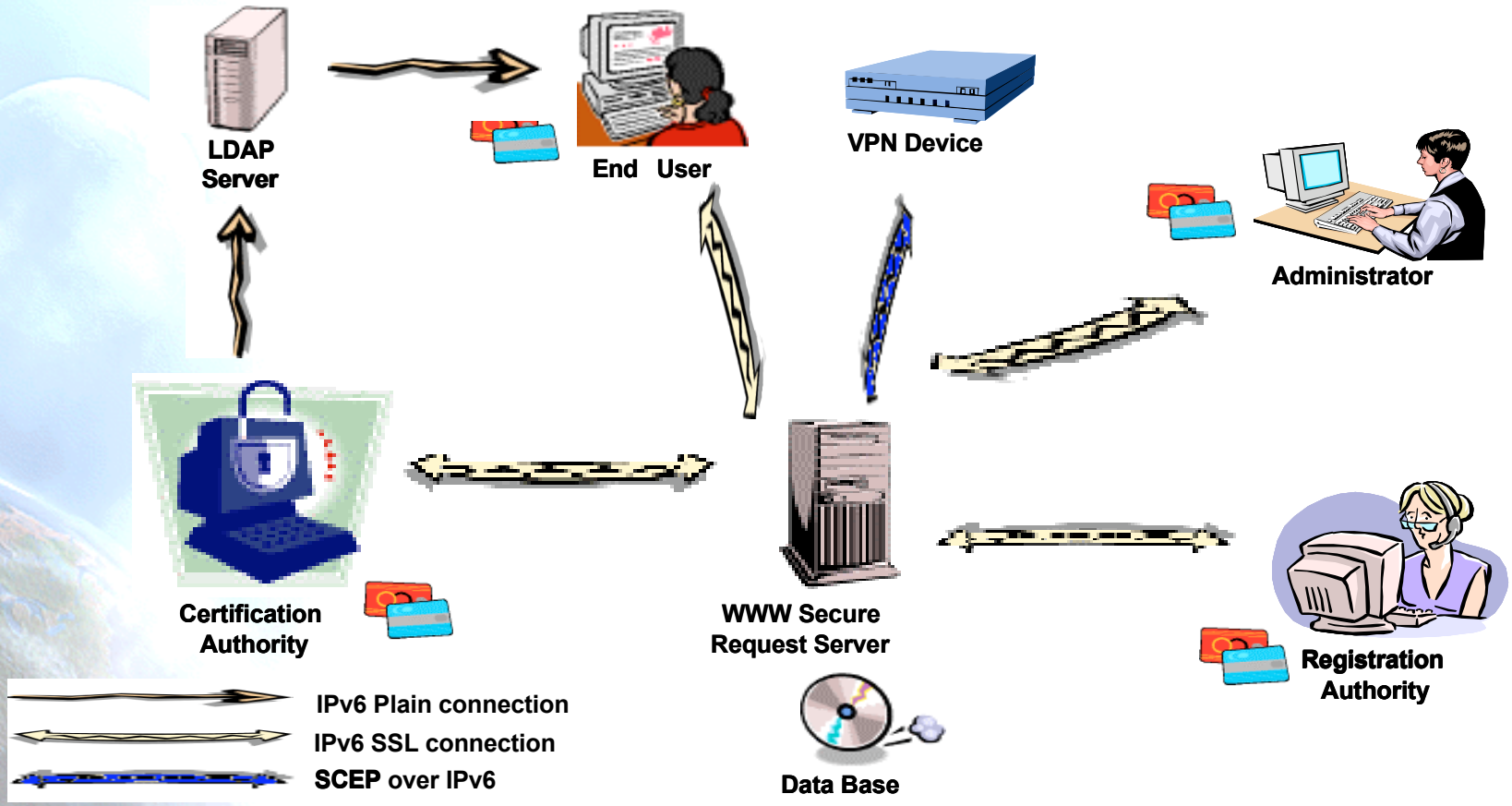
Static VPNs with IPv6

- To evaluate the current status of the main open source IPsec/IKE implementations and some commercial IPsec/IKE solutions
- To deploy of a static VPN service in the Euro6IX test-bed
- Configuration and installations guides for IPsec/IKE
- Test reports of interoperability and conformance

UMU – PKIv6 Description

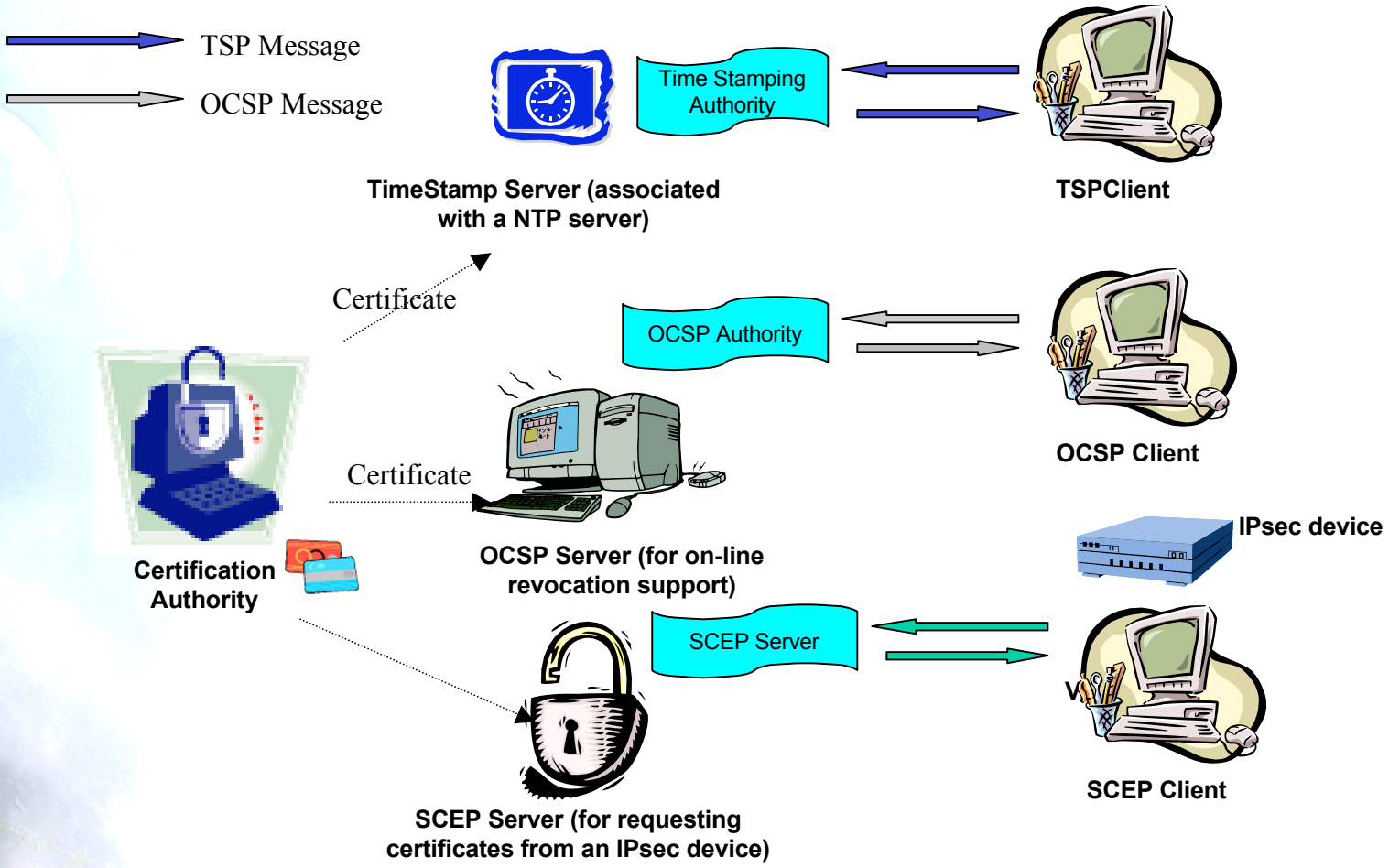
- Main Objective: Establish a high security infrastructure for distributed systems
- Main Features:
  - PKI supporting IPv6
  - Developed in Java ➔ Multiplatform
  - Issue, renew and revoke certificates
  - Final users can use either RAS or Web
  - LDAPv6 directory support
  - Use of smart cards (file system, RSA or Java Cards) ... allowing user mobility and increasing security
  - PKI Certification Policy support
  - VPN devices certification support (using the SCEP protocol)
  - Support for the OCSP protocol and Time Stamp
  - Web administration

UMU – PKIv6 Architecture

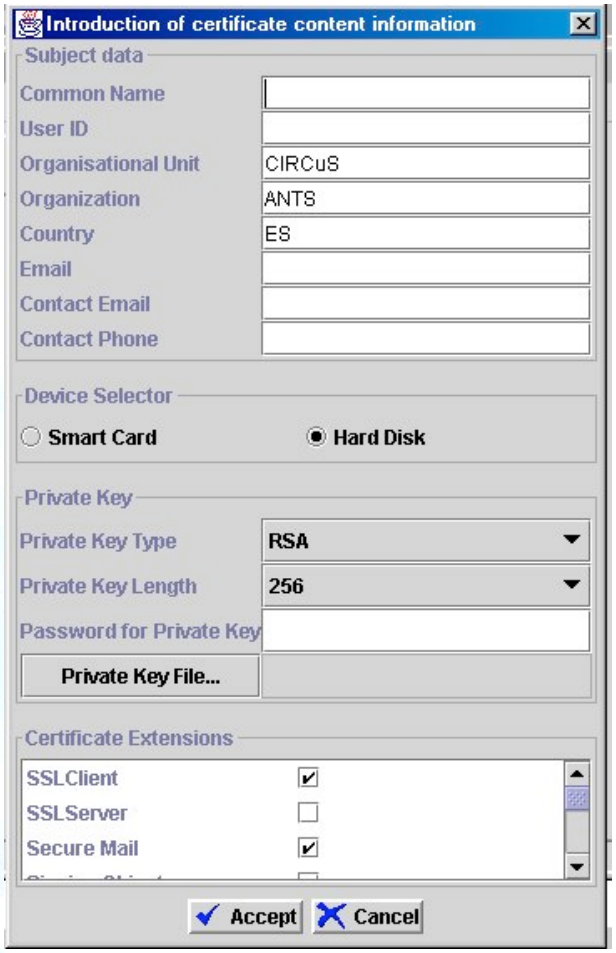


https://pki.ipv6.um.es

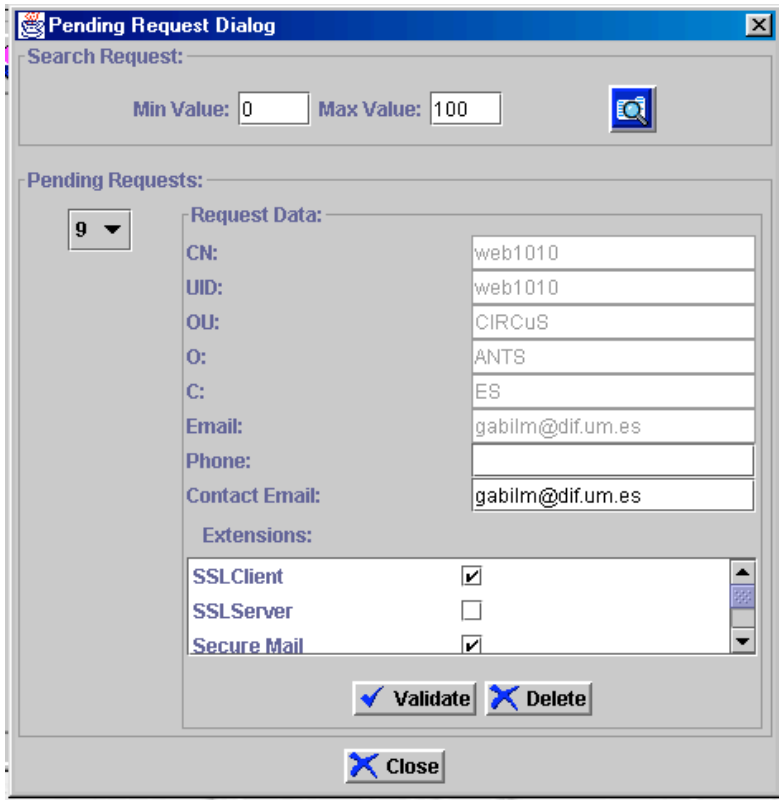
UMU – PKIv6 Advanced Services



UMU – PKIv6 RA Snapshot

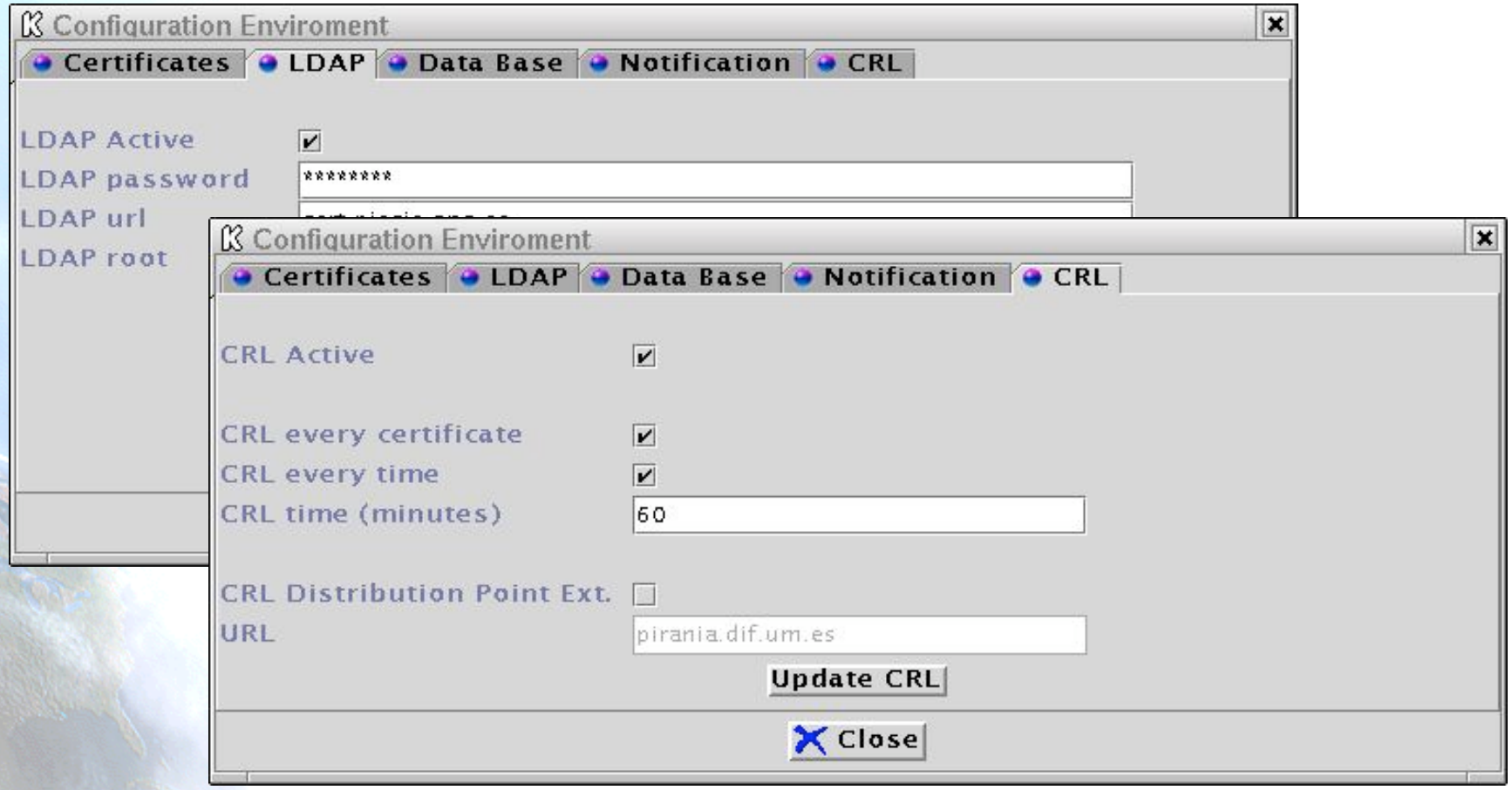


Requesting a certificate



Validating a certificate

UMU – PKIv6 CA Snapshot



CA Internal Management Process

Other Applications

- Messaging Systems:
  - Peer-to-peer
- Audio and video-conferencing:
  - Include multi-conference and collaboration
- Web mail tools
- VNC over IPv6
- Network Management, Analysis, test & diag:
  - IPv6 Network Management Tool (Magalia)
  - Intrusion Detection System
  - Route Server

IX Based Services

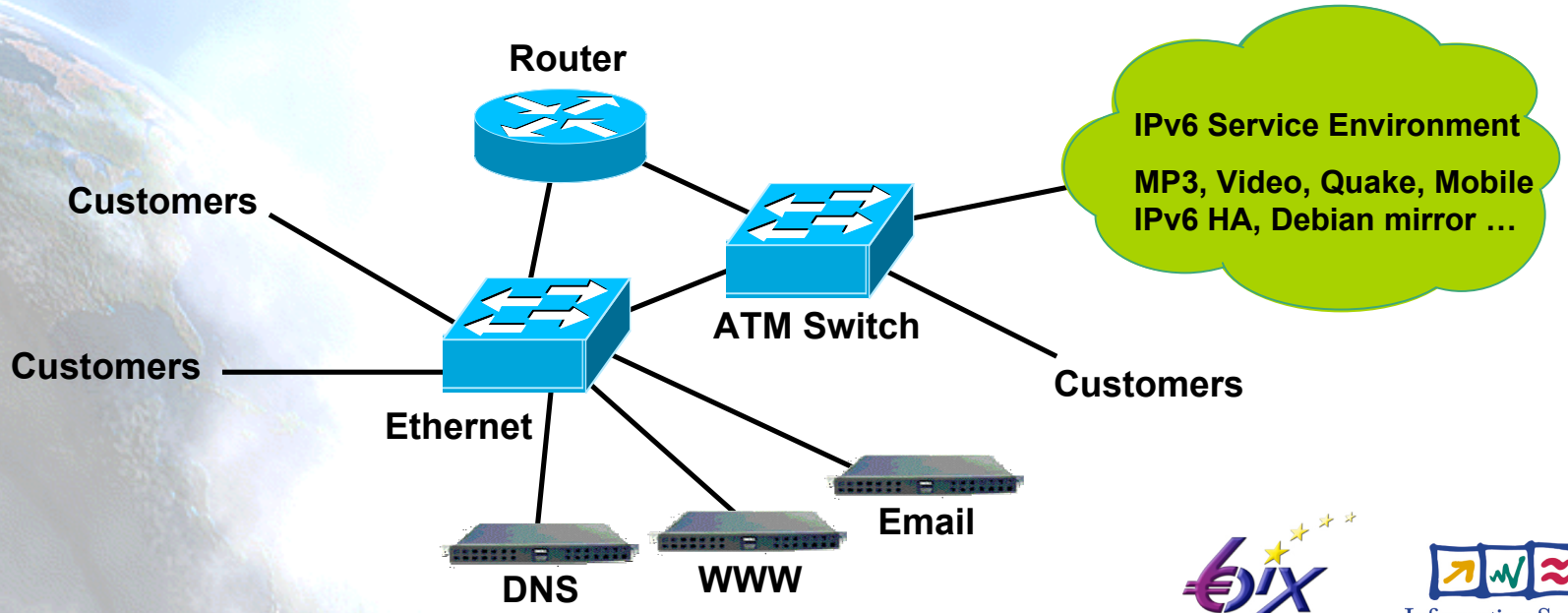
- IX becomes a place where new services are offered to the users.
- IX is an aggregation point, so it can provide those services who can benefit by this “user aggregation” (e.g. in a based multicast network, the RP could be located inside the IX, because a lot of users connect to it).
  - Network Services
    - Multicast, AAA, QoS, DNSSec
    - Transition Mechanisms: NAT-PT, Tunnel Broker, 6to4
    - Route Server mechanism
  - Application Services
    - HTTP, FTP, SMTP
    - VideoConference/e-learning services
    - P2P applications
  - Monitoring Services
    - Routing/Traffic/Reachability Monitoring (Magalia, AS-Path tree, Looking Glass)

The UK6x (LON6IX)

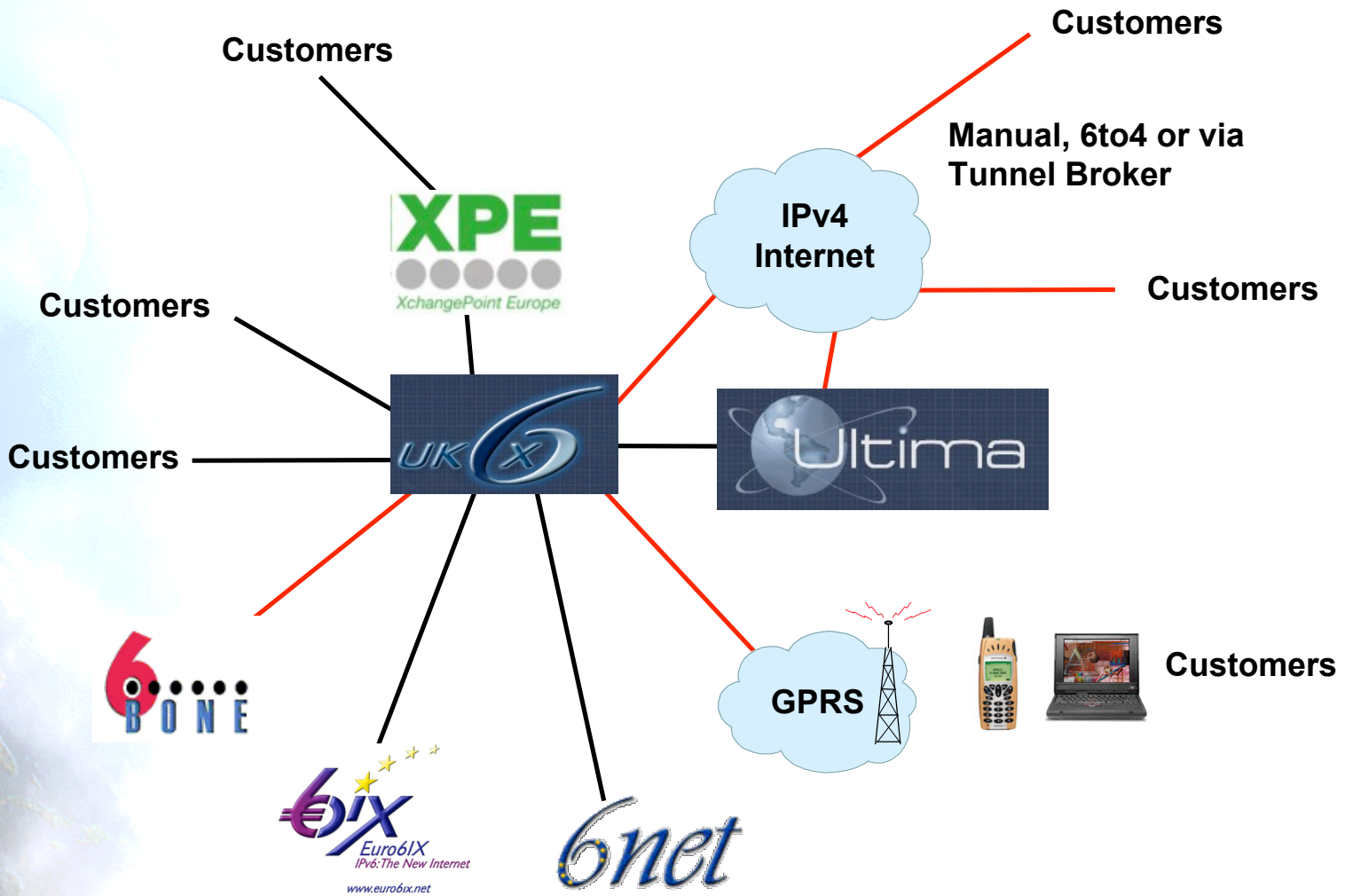
- Layer 2 & 3 IPv6 Internet exchange
- First in the UK
- Uses commercial IPv6 addresses
- Located at the heart of the UK Internet – Telehouse
- Open to all
- Primary aims are:
  - to stimulate the IPv6 environment in the UK, Europe and the World
  - to further the understanding of IPv6

UK6x Core Architecture

- Ethernet switch for Layer 2 peering
- ATM switch for additional customer access mechanisms
- Router for Layer 3 functionality
- 2001:618::/32 used for address allocation
- 2001:7F8:2::/48 used for infrastructure
- Maintenance via Looking Glass, ASpath-tree etc.



UK6x Connectivity

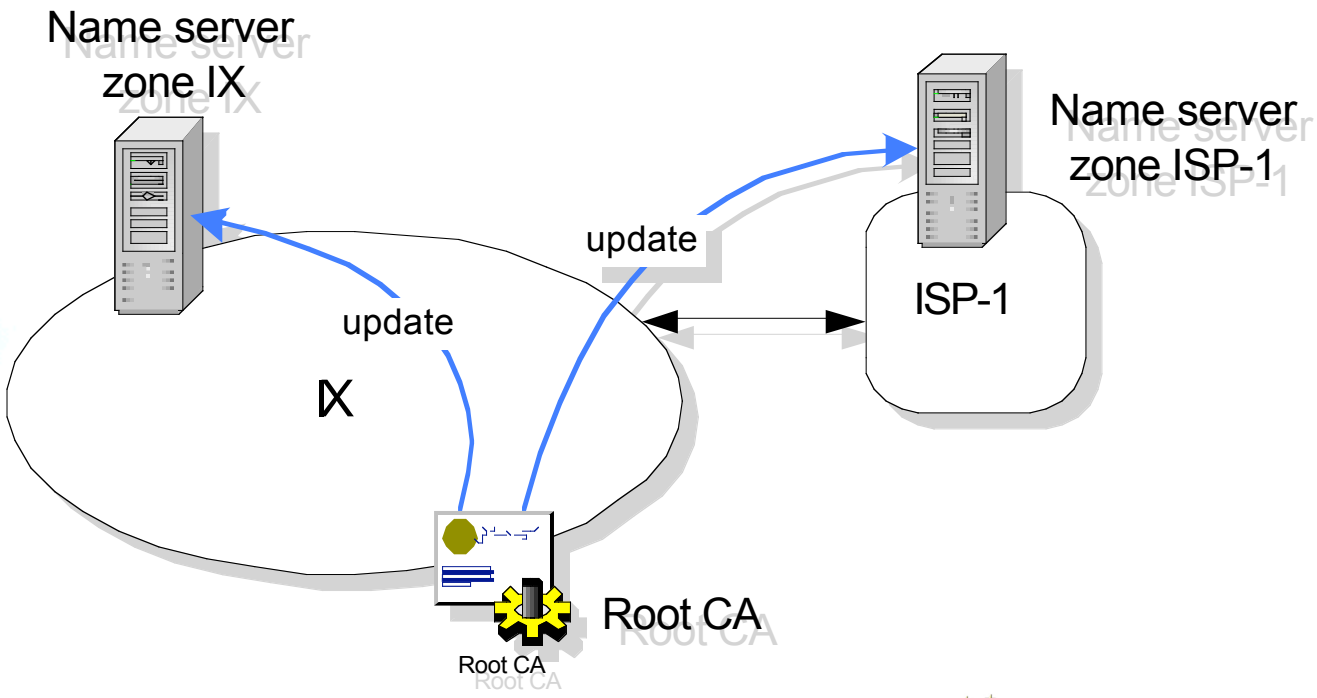


DNSsec Services

- UPM is completing the DNS emulation environment
- Developing a complete set of DNSSEC example configurations using the emulation environment
- DNSSEC pilot work on setting-up and maintaining experiment between UMU, Consulintel and UPM
- Publishing certificates using DNSsec
  - Models analyzed to publish certificates:
    - TSIG Model: symmetric keys.
    - SIG Model: asymmetric keys.
  - Support in PKIv6:
    - PKIv6 supports TSIG Model
      - BIND 9.2.0 or newer for TSIG
    - PKIv6 will support SIG Model
      - BIND 9.3.0 (snapshot) for SIG(0)

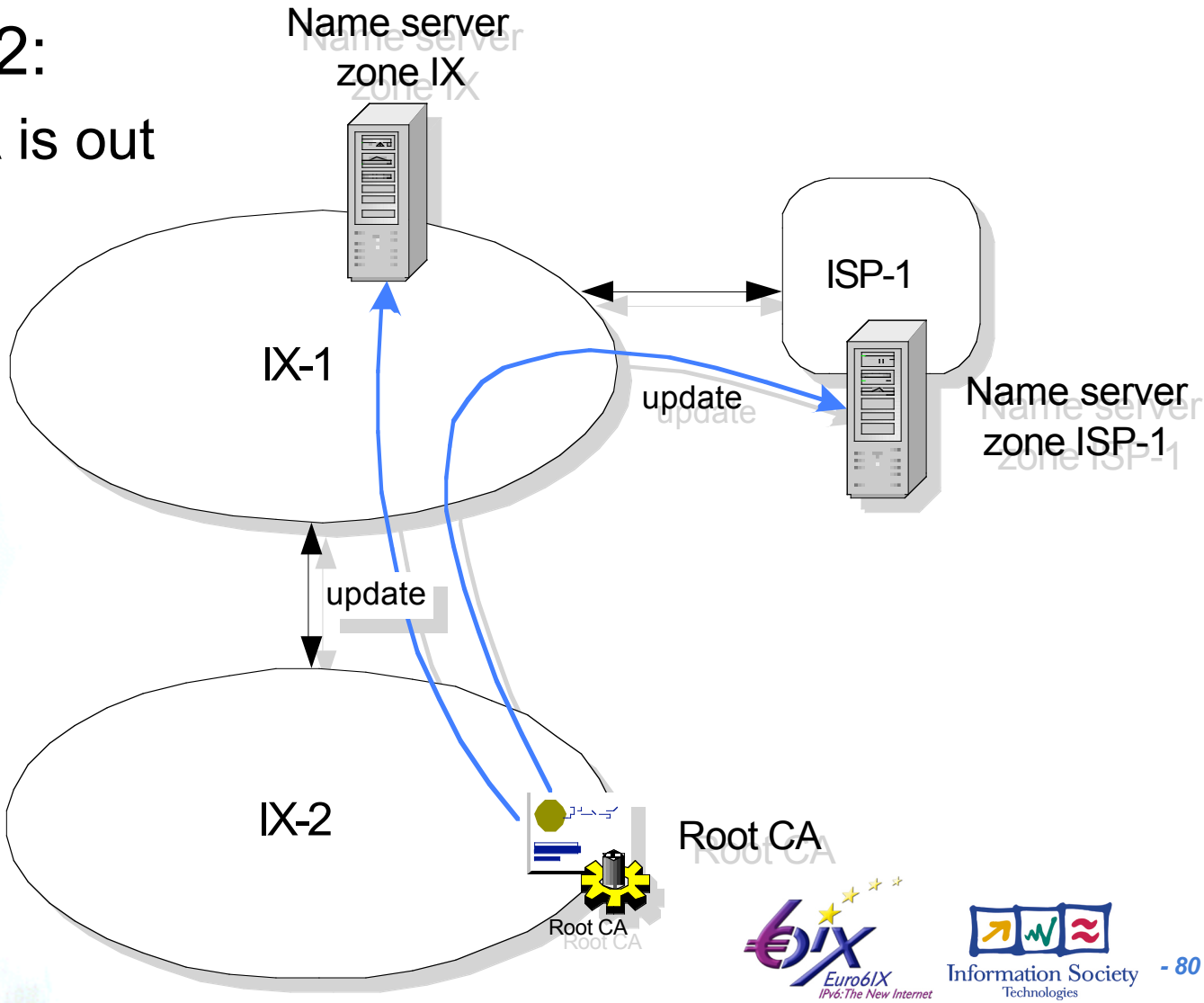
IX service PKIv6 to publish certificates using DNSSEC

- Scenario 1:
  - Root CA and Name Server are together in the IX



IX service PKIv6 to publish certificates using DNSSEC

- Scenario 2:
  - Root CA is out



Security Framework

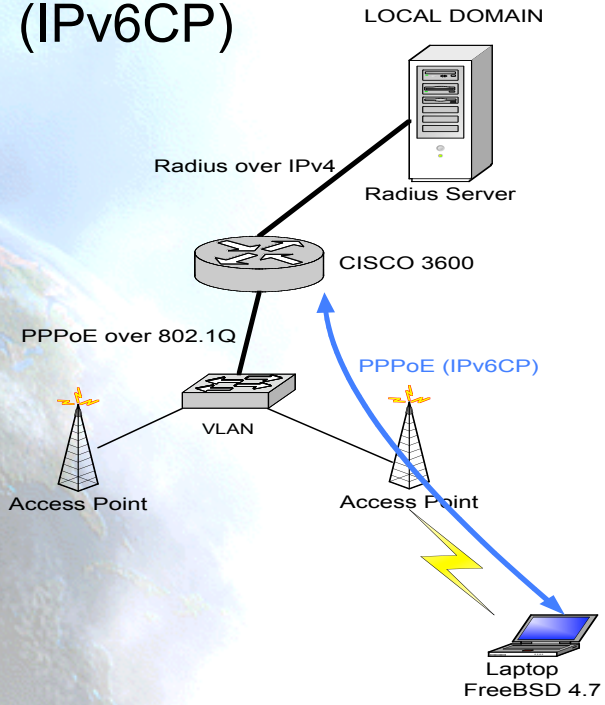
- General VPN Policy Definition. Tools VPNEtool
- Tested with UCL in 6NET-Euro6IX collaboration
- 6WIND VPN Enforcement element working, and being tested by 6WIND
- CISCO: Waiting CISCO IOS version that could be accessible with support for IPsec for IPv6. Actually working with IPv4

VOCAL

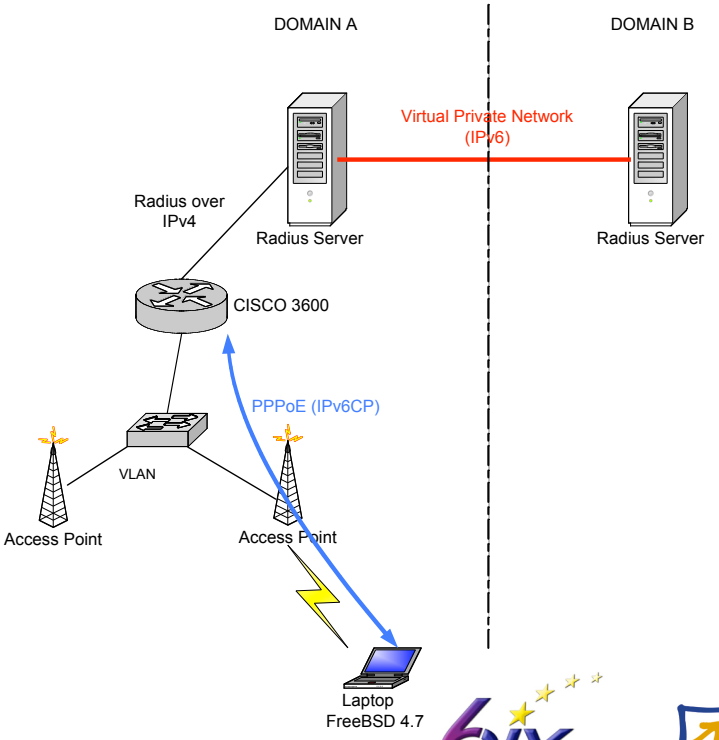
- Porting was undertaken within the Euro6IX project (www.euro6ix.org)
  - But also in conjunction with 6NET (www.6net.org)
  - Work done by a researcher between degree and PhD
  - Being used in 6NET, 6WINIT and Euro6IX
  - Quality of VoIP depends largely on latencies in hardware
- Now moving to VOCAL+ENUM integration
  - A lot of issues to be sorted out

User Auth. DSL, PPP connections based on IPv6

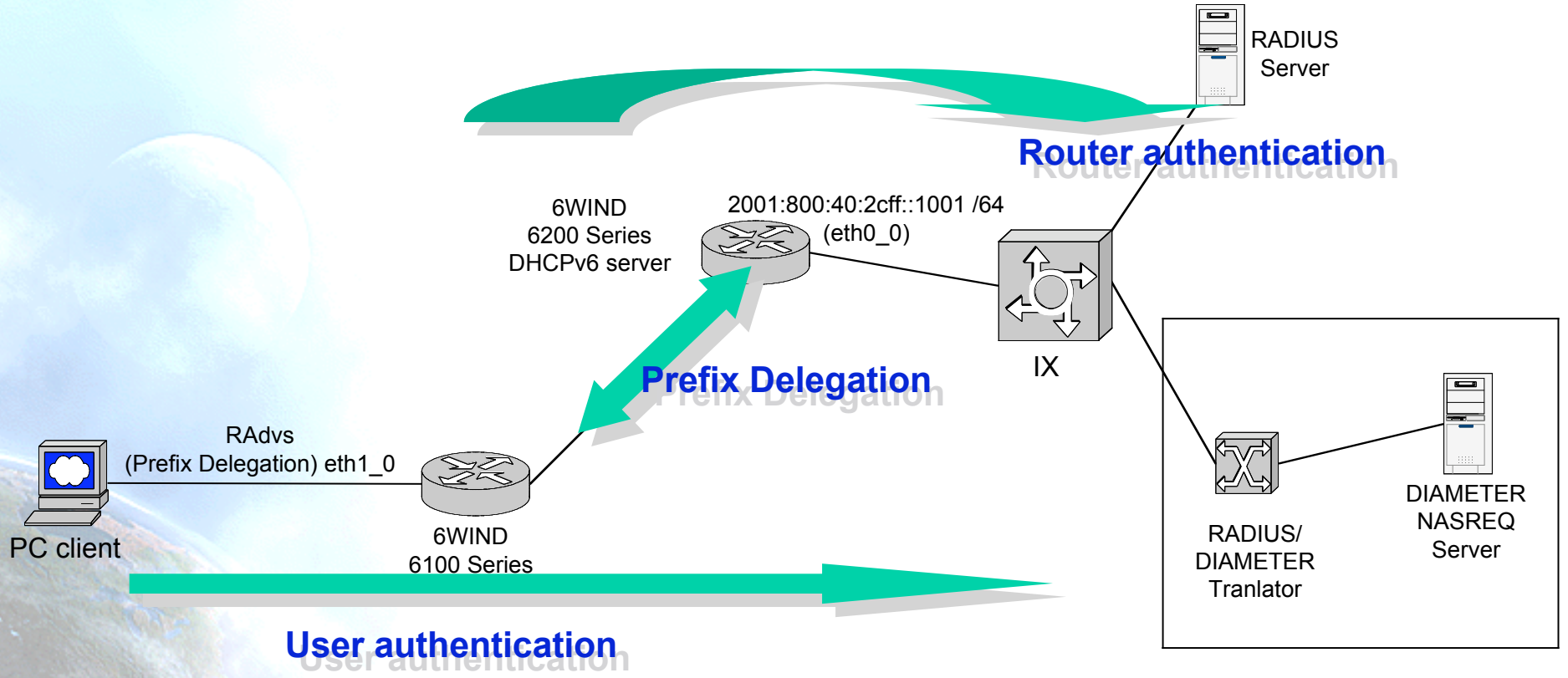
- First scenario:
  - Unique domain
  - End-user is authenticated
  - End-user obtains a prefix (IPv6CP)



- Second scenario:
  - several domains
  - Security between Radius servers is a concern => VPN



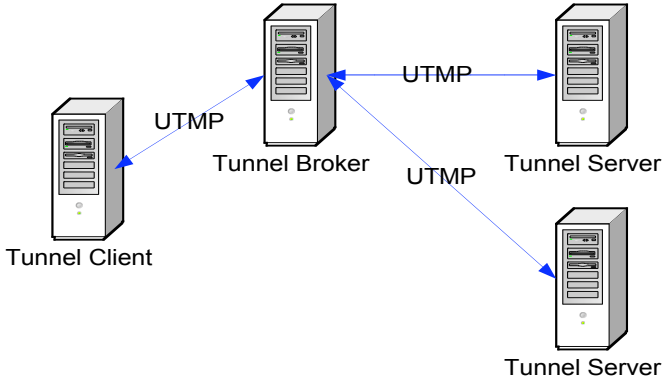
RADIUS/DIAMETER Translator

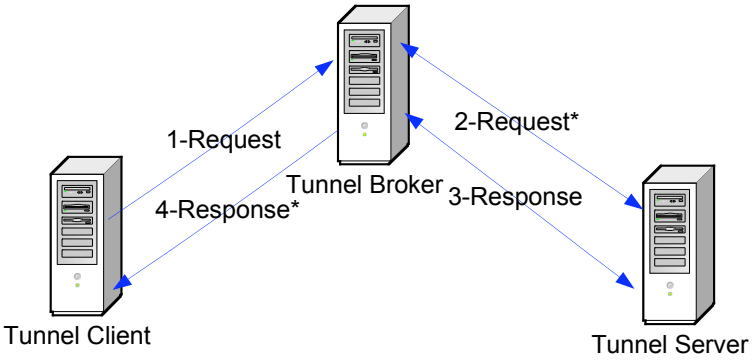


- **Future**: PANA Protocol for carrying Authentication for Network Access (PANA) and DIAMETER Protocol that allows clients to authenticate themselves to the access network using IP protocols
- *Collaboration with PANA-developers for integration with DIAMETER pure scenario.*

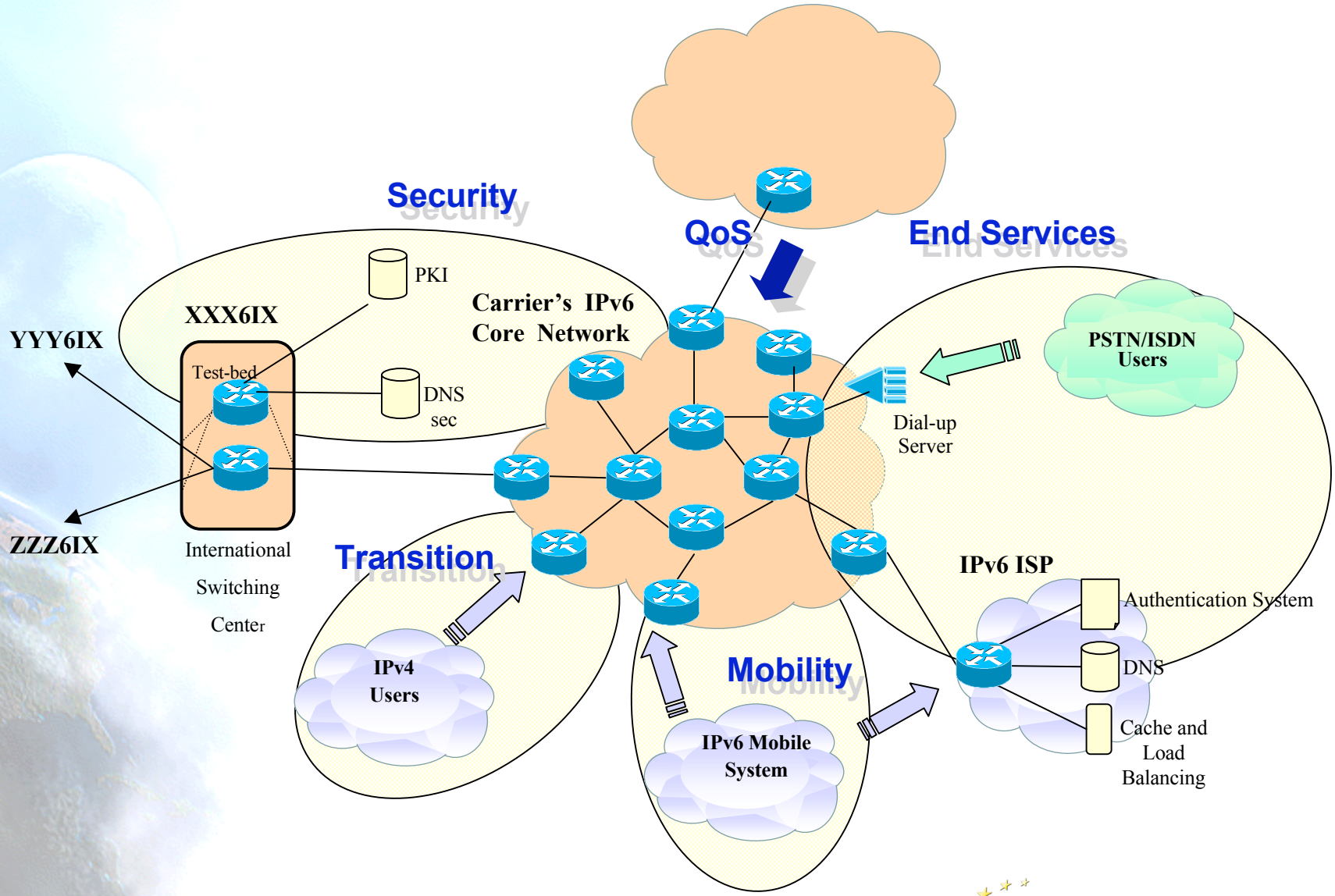
Extended TB architecture

- Integrate new functionality over TB RFC
- Supports entities authentication (Integration with PKIv6)
- UMTP Universal Tunnel Management Protocol
  - used between all devices
  - messages can be “secured” using signs
  - supports several tunnel types (IPv6 in IPv4, IPv6 over UDP, IPSECv6 tunnels)





Advanced Services Vision



Enabling IPv6 in the IX

11. Conclusions

IPv6 is Ready. Do you ?

- Can be deployed now
  - … is being deployed already
- Deployment could take time for ISPs
  - Expertise and knowledge needed
- Brings innovation, not just more addresses
  - New Services = New Revenues
- Restoring Internet paradigms

- Internet Exchanges are the ideal place to start
  - Local peering helps to make it easier and reduce cost
- Are you ready ?
- How we can help ? Let me know !

Thanks !

Contact:

- Jordi Palet (Consulintel): jordi.palet@consulintel.es
- Madrid 2005 IPv6 Summit, soon more info at:
  http://www.ipv6-es.com
- Euro6IX Project Coordinators (coordinators@euro6ix.org):
  - Jordi Palet Martínez (Consulintel): jordi.palet@consulintel.es
  - Carlos Ralli Ucendo (Telefónica I+D): ralli@tid.es