

# THE TERASTREAM IPv6 NATIVE NETWORK ARCHITECTURE

How to build a modern service provider using IPv6 and Optics

Ian Farrer, Deutsche Telekom AG



LIFE IS FOR SHARING.

# TeraStream

## Packet Cloud Architecture Fundamentals

**One truly converged network - de-layered, IP and Optical are one, bits over wavelengths, digital over analog**

**The same technology for LAN and WAN - for LAN: IP packets in Home, Office, Data Center; for WAN: IP packets in Metro, Country, Continent**

**Digital services for consumers and businesses - communication, information, cloud-compute and -storage**

**Real-time OSS - instantaneous service provisioning, guaranteed good user experience**

**Cloud-era economics - service flexibility, fast-paced innovation, agile implementation, reduced system complexity, lower cost**

**Stay with the IP Internet Architecture using IPv6 for all functions and services**

**Have a standardized toolbox, “services in the data center”**

**Standard computers instead of specialized appliances**

**Look forward, what can we do, not backwards what we used to do**

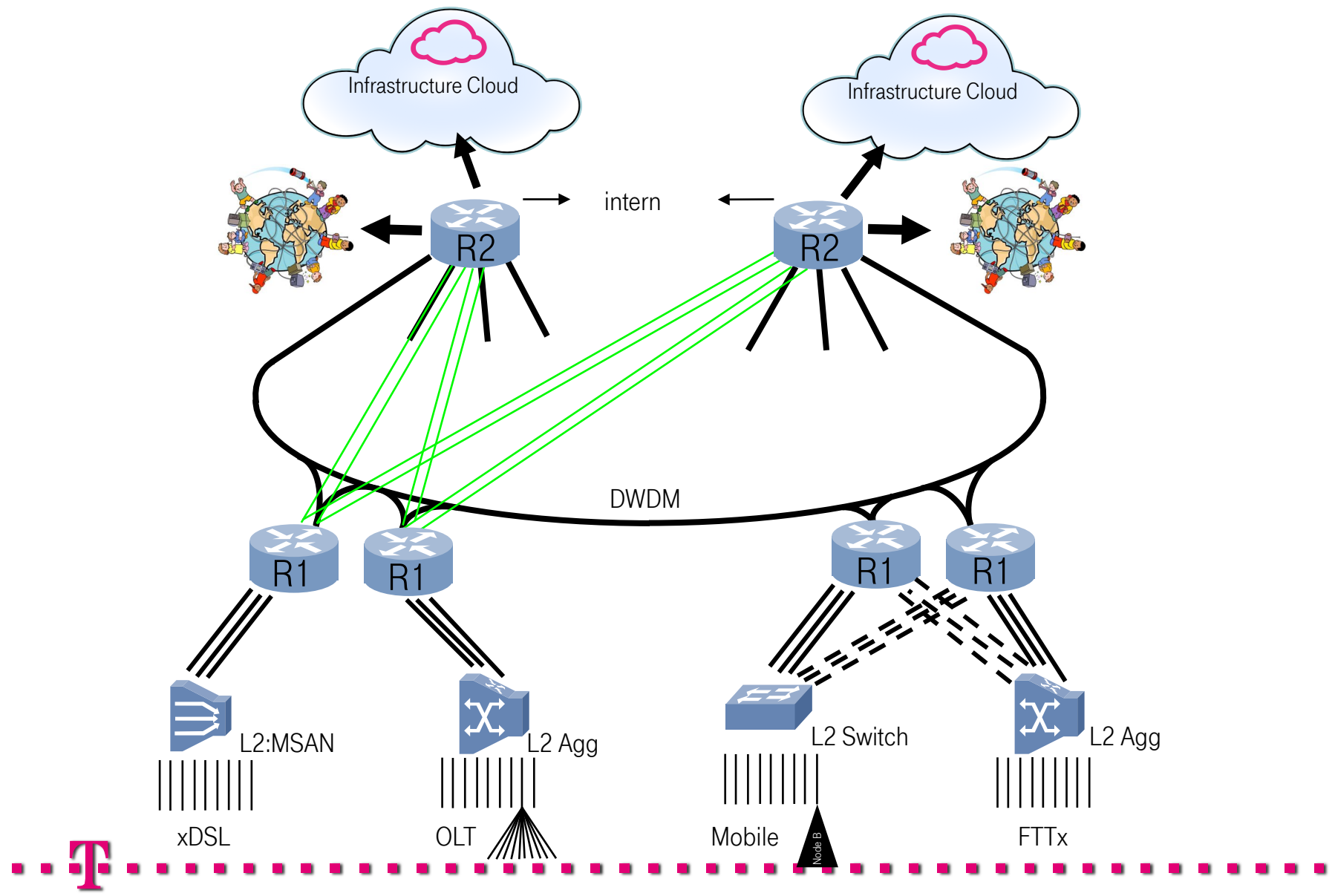
**No L1, L2, L3 dependencies**

**No “best effort” you get what you pay for (Best effort is just a service class..)**

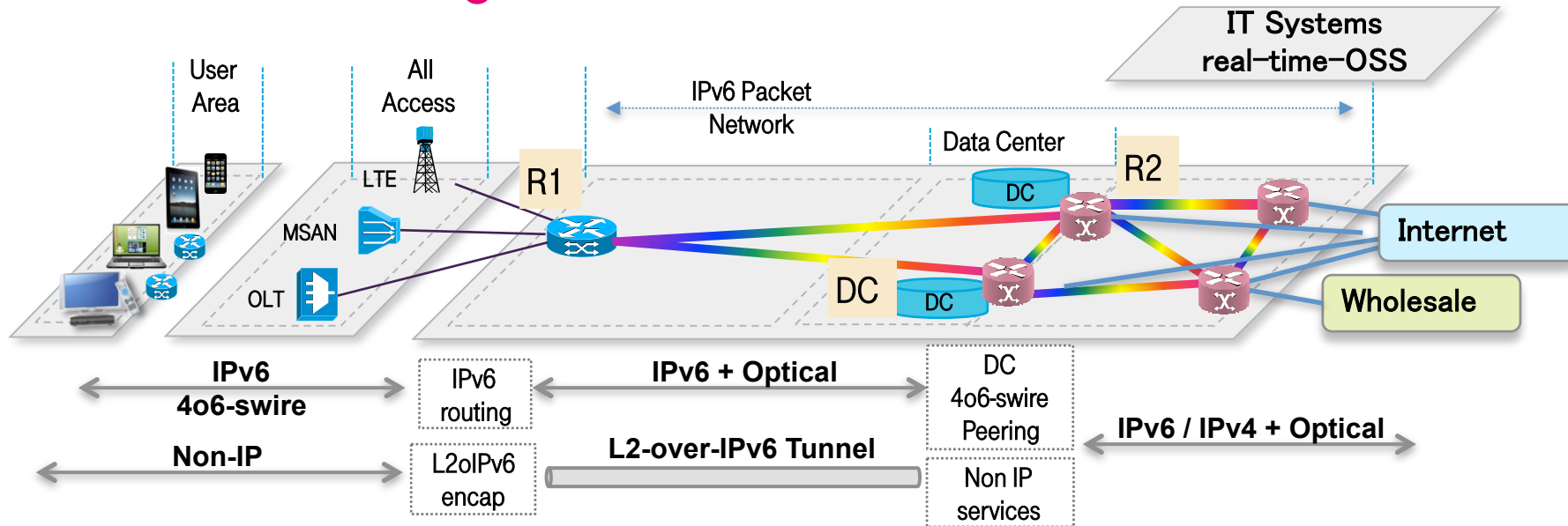
# TeraStream Design Principles

Principle	Applied to TeraStream design
Reduce the amount of technologies used	Use IP and optical transmission only No OTN, L2, MPLS switching
Use IPv6 for all internal functions and services	No native IPv4 support in the network IPv4 is a service IPv6 based “carrier Ethernet service”
Avoid internal interfaces	Minimize non-customer, non-peering facing interfaces Distribute Internet peerings, offload external traffic ASAP
Size the network to handle all IP traffic without IP packets losses	Dimension the network for peak hour IP traffic, no oversubscription, packet loss is extreme exception
Integrate optical networks and IP networks as much as possible	Integrate IP and optical layers into routers to simplify the network, avoid redundant mechanisms e.g. failure handling, reduce total cost
Use one network for all services – Internet, IP TV, business, ...	Single converged packet network Note: <b><u>Dominant traffic drives the design!</u></b>
Deterministic and short routing path for all on-net traffic	Network distance between R1 access routers is at most two R2 backbone routers away and R1 is multi homed to two R2
Service policy for packets are outside the payload	Encode service type, traffic class, direction etc in the IPv6 address
Data Centers are directly connected to backbone routers	DCs connect directly to R2s to avoid building internal IP interfaces for very large amount of traffic

# THE TERASTREAM ARCHITECTURE



# TeraStream – Design in a Nutshell



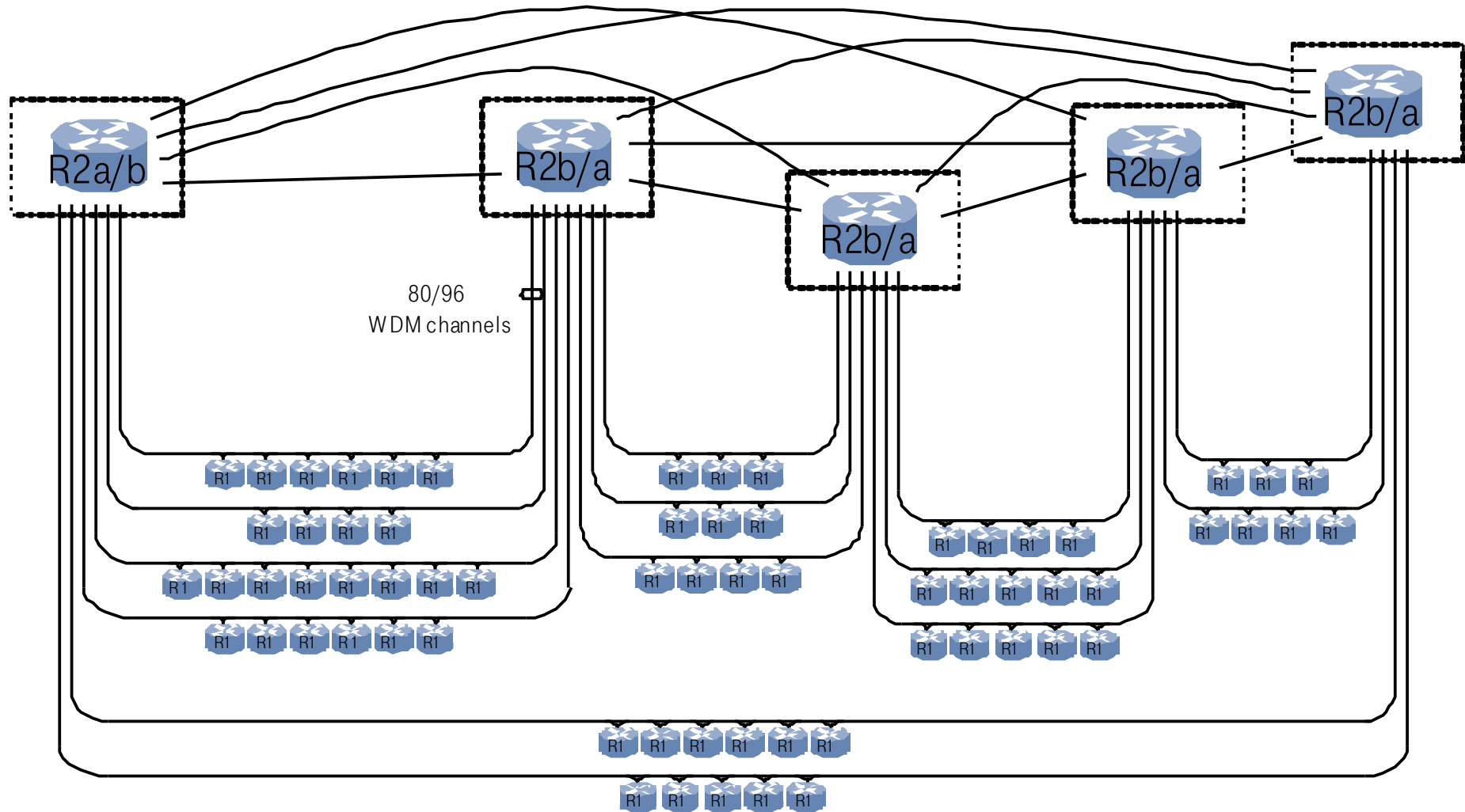
## TeraStream key functional elements

R1	R2	Data Center / Services
<ul style="list-style-type: none"> <li>▪ Terminate access interfaces</li> <li>▪ Runs IPv6 routing only, integrates optical</li> <li>▪ Access services                             <ul style="list-style-type: none"> <li>▪ IPv6 - dealt with natively</li> <li>▪ IPv4 – IPv4 over IPv6 softwire between HGW / CPE and DC, R1 not involved</li> <li>▪ non-IP - L2-over-IPv6 encapsulation</li> </ul> </li> <li>▪ User configuration                             <ul style="list-style-type: none"> <li>▪ using Netconf / Yang</li> <li>▪ Driven by real-time OSS i.e. self-service portal</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>• Connects R1s, Data Centers and Internet peerings</li> <li>• Runs IPv6 and IPv4 routing, integrates optical</li> <li>• Closely integrated with Data Centers                             <ul style="list-style-type: none"> <li>Optimized handling of locally sourced services</li> </ul> </li> <li>• High scale IP bandwidth</li> </ul>	<ul style="list-style-type: none"> <li>• Distributed design                             <ul style="list-style-type: none"> <li>fully virtualized x86 compute and storage environment</li> </ul> </li> <li>• Network support functions - DNS, DHCP, NMS</li> <li>• Real-time OSS incl. user self-service portal</li> <li>• Cloud DC applications, XaaS services</li> <li>• Complex network services e.g. high-touch subscriber handling</li> </ul>

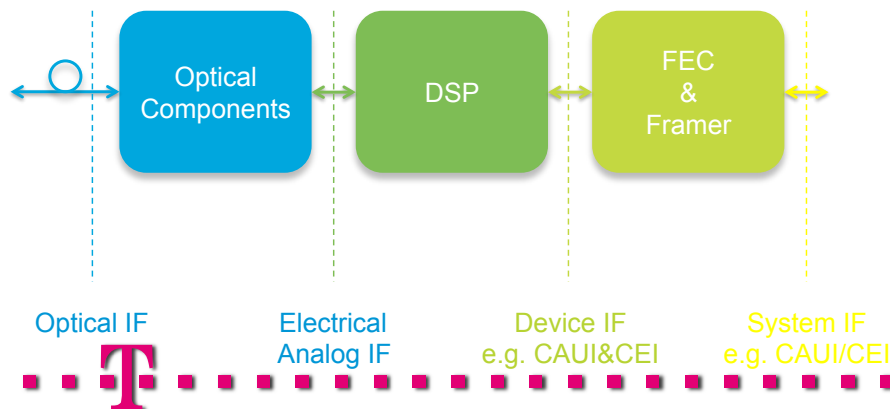
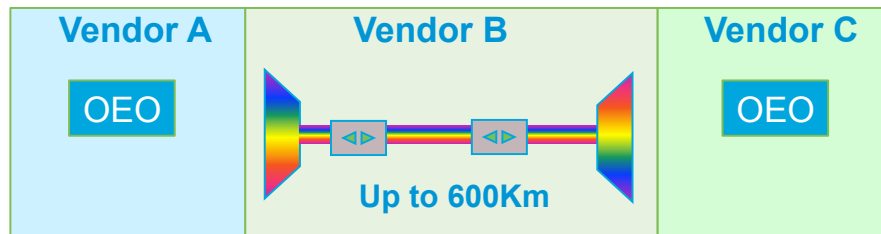
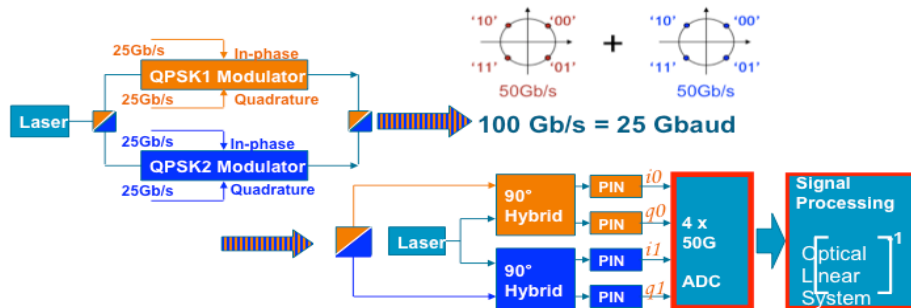
# TRANSPORT



# R1 ↔ R2: OPTICAL FIBER LINKS



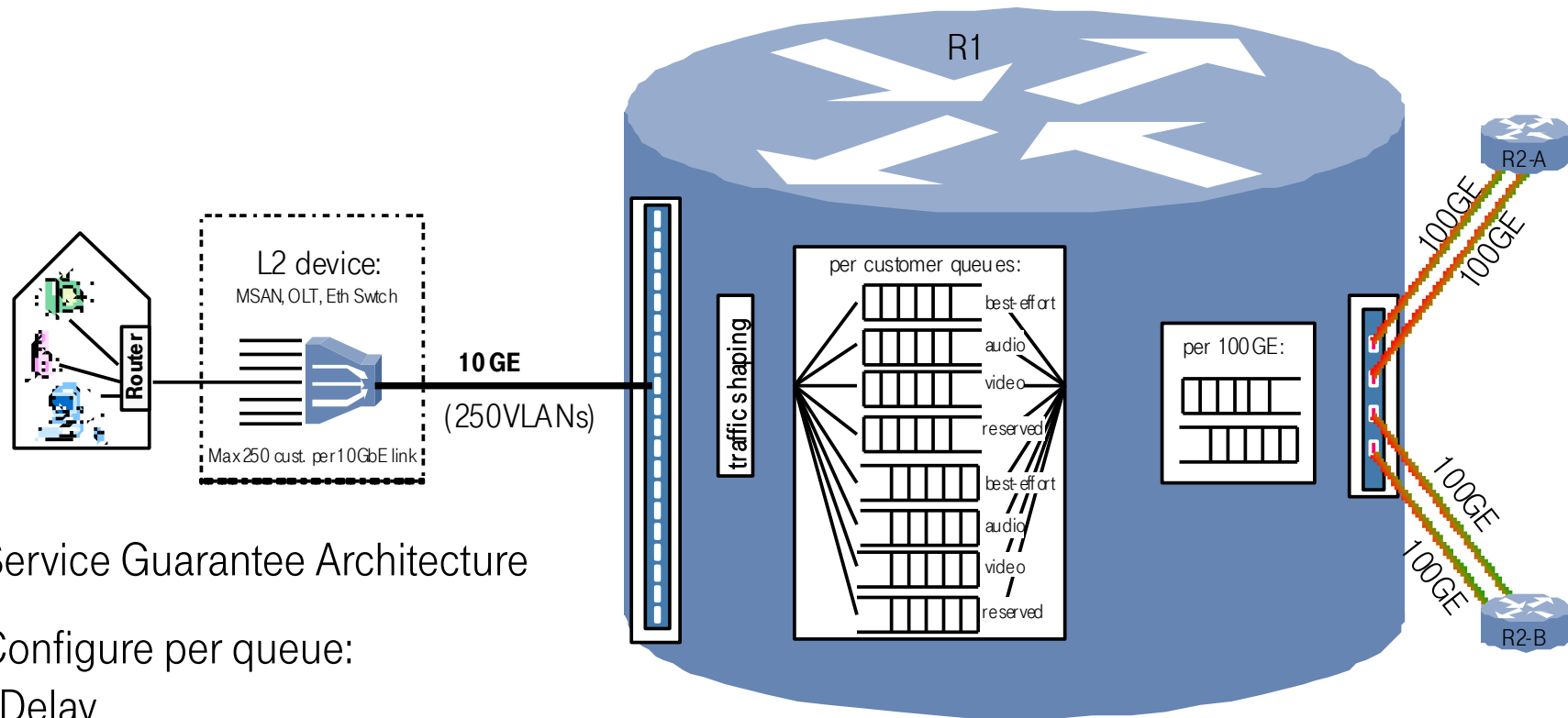
# 100G Coherent DWDM Interop and Pluggable Technology



- Agree on a common set of parameters for the 100G line side
- Enable innovation by many players in the silicon optics arena
- Hard FEC, typ 800km
- If price is right, use in data center
- Coding
- Carrier Recovery
- Acquisition (blind)
- Reach
- Framing (works with both OTU4.4 and OTU4.10)
- Forward Error Correction (Hard FEC Staircase)



# TeraStream user facing router R1



## Service Guarantee Architecture

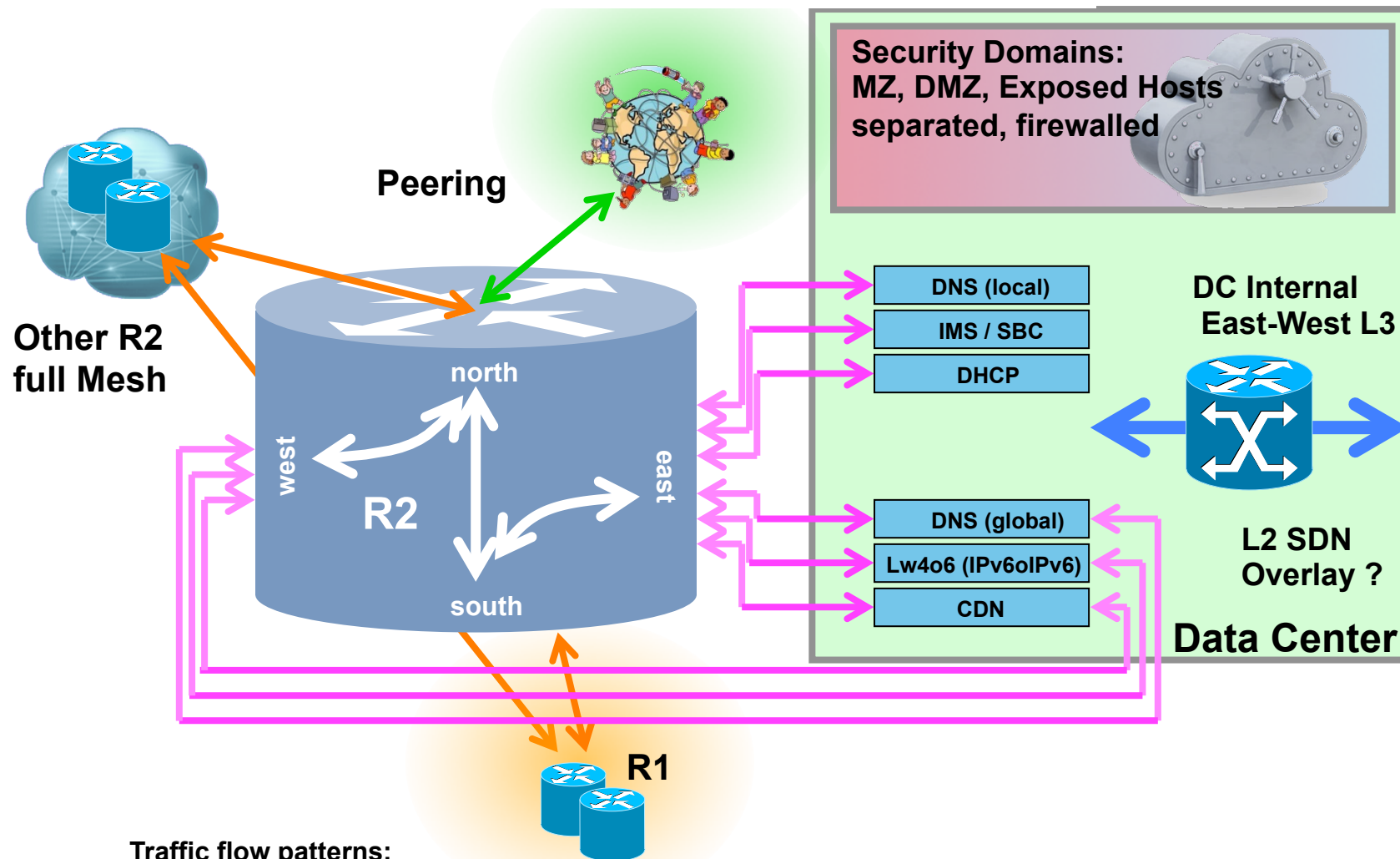
Configure per queue:

- Delay
- Drop
- Bandwidth
- Reorder
- Etc...

- IP traffic shaped to capabilities of L2 device
- 5000 customers connections per R1
- 20 \* 10GE port for L2 device
- 4 \* 100GE for R2 link



# R2 router and traffic patterns



## Traffic flow patterns:

- R1 ⇔ Peers and Other R2 going north ⇔ south (example: IPv6 Internet traffic)
- R1 ⇔ Data Center services going south ⇔ east (example: DHCP)
- R1 ⇔ Data Center ⇔ Peers going south ⇔ east ⇔ west ⇔ north (example: IPv4 Internet traffic)



# TERASTREAM HOMEGATEWAY PLATFORM

Existing CPE procurement models are far too slow to keep up with the necessary rate of development

The current state of IPv6 support on commercial platforms leaves ... room for improvement!

We believe that we are not the only carrier with this problem

**So ...**

DT are currently developing a carrier grade' CPE software platform

Based on OpenWRT, extended with IPv6 and SP functionality

Plan to register this with Linux Foundation

A fully open source initiative

**We are looking for other interested carriers to build a development community around the platform**



# IPv6



# IPv6 NATIVE NETWORKING

IPv6 is fundamental, not an Afterthought

**A new network - designed from the ground up specifically for IPv6**

**IPv6 for ALL internal interfaces (incl. management, control plane etc) – with just a bit of IS-IS for IGP**

**Treat IPv6 as a new protocol – It can be much more than just ‘IPv4, but bigger’**

**Don't let native IPv4 into the network**

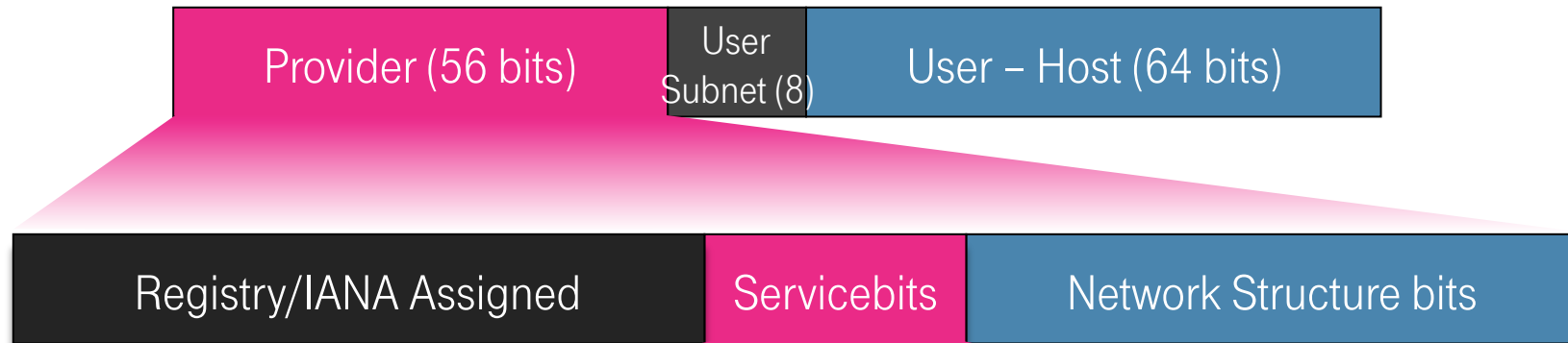
**No Dual-Stack**

**Treat IPv4 as a long-tail overlay service – (more on this later)**

**Once you let ANY IPv4 into the design, you'll never get rid of it!**

# SERVICE DIFFERENTIATION BASED ON ADDRESSES

## USING IPv6 ADDRESS SPACE AS LABELS



```

P Public          0=SP-intern, 1=extern
I Infrastructure  0=end user, 1=infrastructure packet
E Endpoint/Service 0=endpoint, 1=service
SSS Service Type 0=res, 1=internet, 4=video, 5=L2, 6=voice, 7=mgmt
M                0=fixed, 1=mobile endpoint

Examples:
Source           Destination
PIESSS          PIESSS
-----
User -> IMS      000110          011110
IMS -> User      011110          000110
User -> User (best effort) X00001          X00001
User -> Internet (best effort) 100001          XXXXXX
Internet -> User (best effort) XXXXXX          100001
Lan-Lan service  010101          010101
    
```



# MULTIPLE PREFIXES ON THE CLIENT...

Brings benefits, but also a new (old!) problem

How do you ensure that the client selects the right source address for each different service?

Current source address mechanisms are based on variants of longest source/destination prefix matching policy

- This places constraints on your addressing architecture
- Often require additional policy to be provisioned to the client
- Doesn't give users or applications information about the 'properties' or 'suitability' of a prefix for use

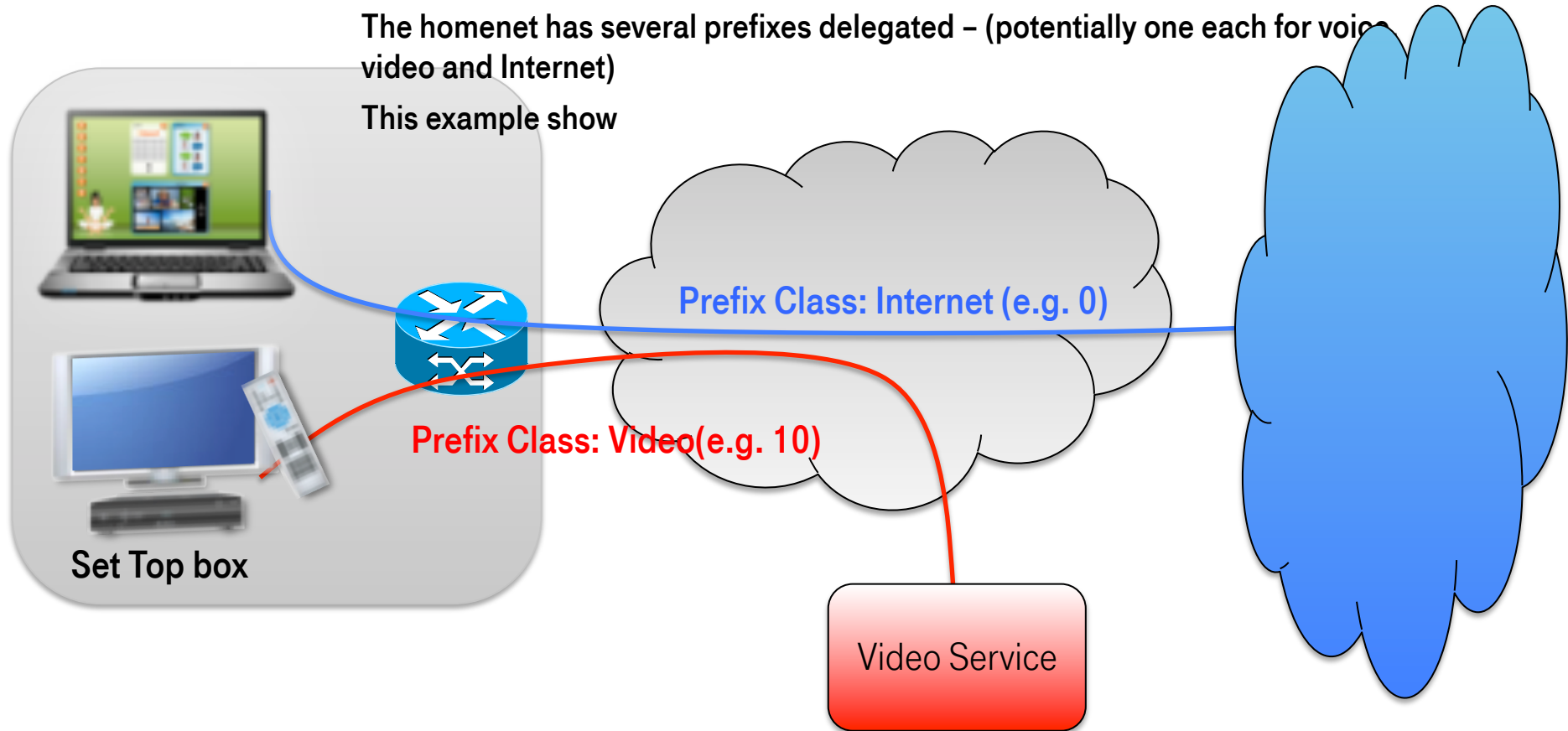
## The Solution? - "Prefix Colouring"

- Adds additional metadata to DHCP prefix allocations
- Allows applications and users to select a source prefix based on this metadata
- Source address selection is decoupled from the destination address prefix matching
- Can also help with source address selection in multi-homing deployments
- Described in 'draft-lepape-6man-prefix-metadata'



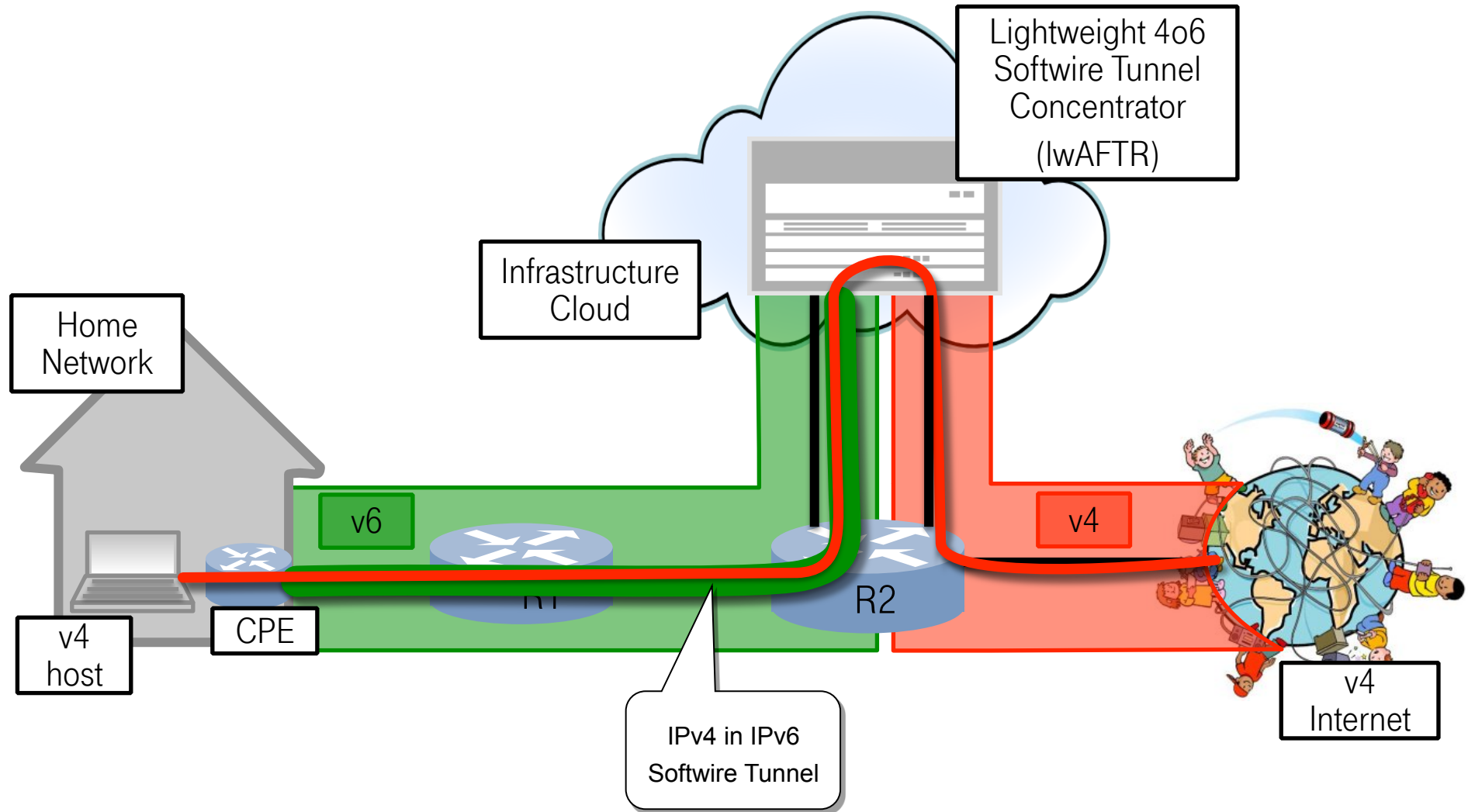
# Home network with video class service

In this example, two different services are being run on the same network.  
The service provider wishes that traffic is sourced from different prefixes by the home network clients for Video on demand service as against general Internet access  
The homenet has several prefixes delegated – (potentially one each for voice, video and Internet)  
This example show





# IPv4 AS A SERVICE - LIGHTWEIGHT 4o6 SOFTWARES





# DEPLOYING v6 ONLY – SOME LESSONS LEARNT SO FAR...

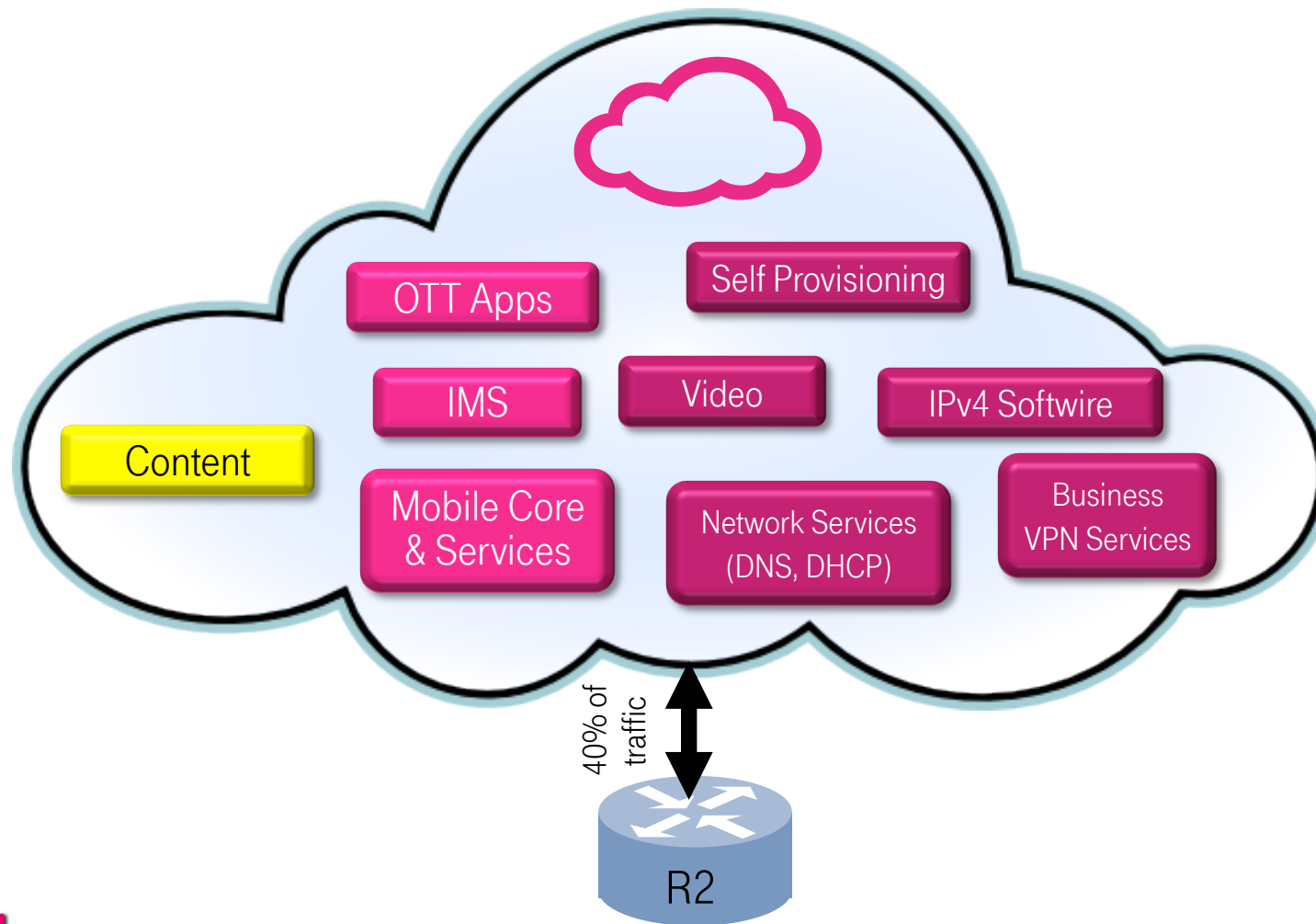
- There are still mainstream products that do not have complete IPv6 implementations:
  - Transport interfaces tend to have more complete implementations
  - Management and control plane functionality may not be so good
- The level of IPv6 testing in shipping products is not on a par with IPv4 – we've found some pretty hairy bugs!
- Vendor's need constant 'encouragement' to resolve these problems
- If you are planning any kind of similar rollout, get your requirements fixed and test well in advance!

# NETWORK FUNCTION VIRTUALISATION

“THE INFRASTRUCTURE  
CLOUD”



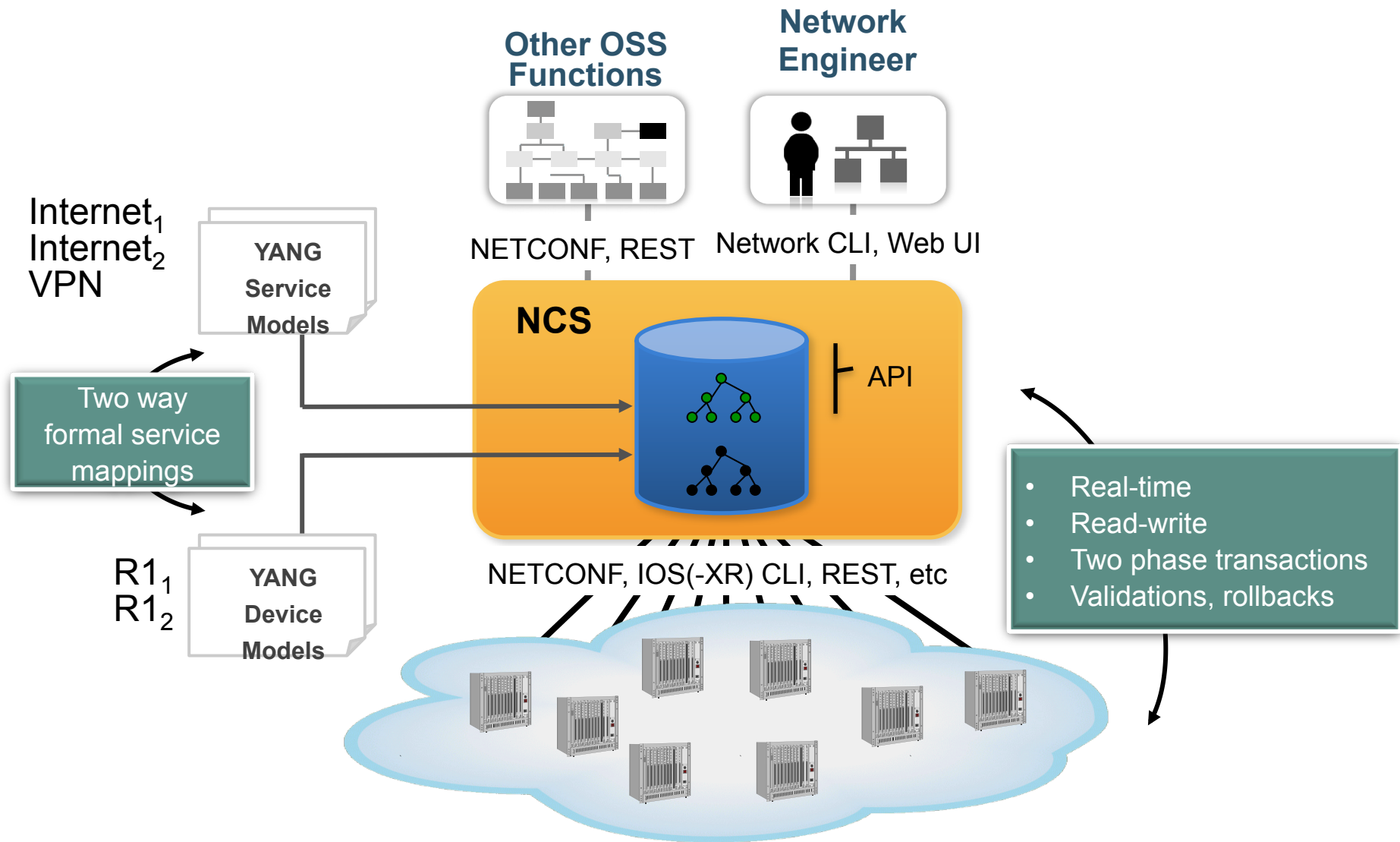
# INFRASTRUCTURE CLOUD NETWORK FUNCTION VIRTUALIZATION



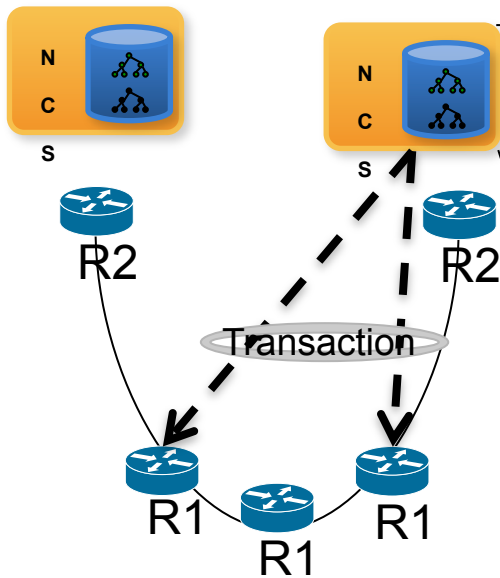
# REAL-TIME OSS



# Model-Driven Network Abstraction Layer



# Define “Services” in a data-model language (Yang)



## NCS CLI : EVPL Service over R1 and R2

```

ncs# services service evpl2 type evpl bandwidth 100MB qos-template
gold port-1 r1-sto0-ge0/0/1.65 port-2 r1-gbg1-ge0/0/5.65
ncs(config-service-evpl2)# show
ncs(config-service-evpl2)# show conf
services service evpl2
type evpl qos-template gold
type evpl bandwidth 100MB
type evpl port-1 r1-sto0-ge0/0/1.65
type evpl port-2 r1-gbg1-ge0/0/5.65
!
ncs(config-service-evpl2)# commit
    
```

## NCS WebUI



## NCS REST

```

$ curl -v -X POST -T evpl.xml -u admin:admin \
http://localhost:8080/api/running/services/service
    
```

## NCS NETCONF

```

$ netconf-console --edit-config=evpl.xml
    
```



# TERASTREAM CROATIA TRIAL



# TERASTREAM PILOT SUCCESSFULLY LAUNCHED @ HRVATSKI TELEKOM ON DEC 10, 2012

100 Gb/s network using IP and  
Optical integration

Full integration of Network  
and Cloud technologies for  
service production

Native IPv6 network delivering  
consumer service

- ▶ Built in a record time - decision in September, launch on Dec 10th
- ▶ 500 customers with up to Gigabit access speeds
- ▶ Agile execution – small cross-functional teams (DT, HT, Cisco, Combis)
- ▶ Continued development and iterative improvement
  - Technology refinements
  - New vendors being integrated
  - More customers connected





# TERASTREAM

## SUMMARY

### TeraStream

- Radical simplification of the IP network architecture
- Removing the legacy from the core (IPv4, MPLS), improving services
- Optical transmission is integrated into the IP routers using 100G coherent technologies
- Combining network and cloud for scalable service production
- Control using a SDN paradigm – Realtime OSS

### Benefits

- Improve user experience, real Internet services to more users
- Use just enough complexity to do the job and no more
- Get the revenue and cost balance right



# Questions?

Now you can bring out your tar and feathers and start throwing things at me..

# THANKS!

