

PIC
port d'informació
científica



Despliegue de la red de altas prestaciones para el LHC sobre RedIRIS-NOVA

Gerard.Bernabeu@pic.es
Alberto.Escolano@rediris.es

LHC will generate O(10) Petabytes of RAW data per year.

Several hundred thousands CPUs needed to produce scientific results.

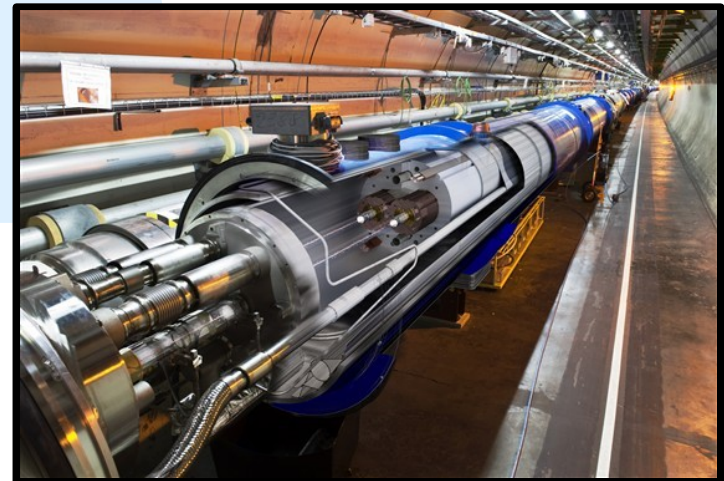
1000s of researchers in 10s of countries around the world.

Decided to build a distributed computing infrastructure: WLCG.

Started in 2002 as a project coordinated from CERN.

Infrastructure built, grown to nominal capacity and thoroughly tested in these last 8 years to have it ready for the LHC start.

Also, key point: Enable the funding agencies to spend the money inside the countries (and not at CERN).



The successful LHC data analysis has shown that a more intensive use of the network leads to a more efficient analysis.

To this end the experiments are re-evaluating their computing models and a **greater use of the network resources is foreseen** (Bos-Fisk paper).

Bos-Fisk paper categorizes Tier-2s depending on their resources as minimal (1Gbps), nominal (5Gbps) and Leadership (10Gbps+).

Generic IP networks should not saturate by the increase of LHC traffic.

The objective of LHCONE is to provide a collection of access locations that are effectively entry points into a network that is private to the LHC T1/2/3 sites. LHCONE is not intended to replace the LHCOPN but rather to complement it.

Primary goal is to get more TierX-TierX available bandwidth.

LHCONE is still in the prototyping phase. Currently deploying and evaluating a private layer 2 shared network.

Layer 2 network is, in general, cheaper for the carriers.



Use of available underused capacity (ie. backup links).

To ensure all sites can connect to LHCONE only standard multi-vendor technologies should be used.

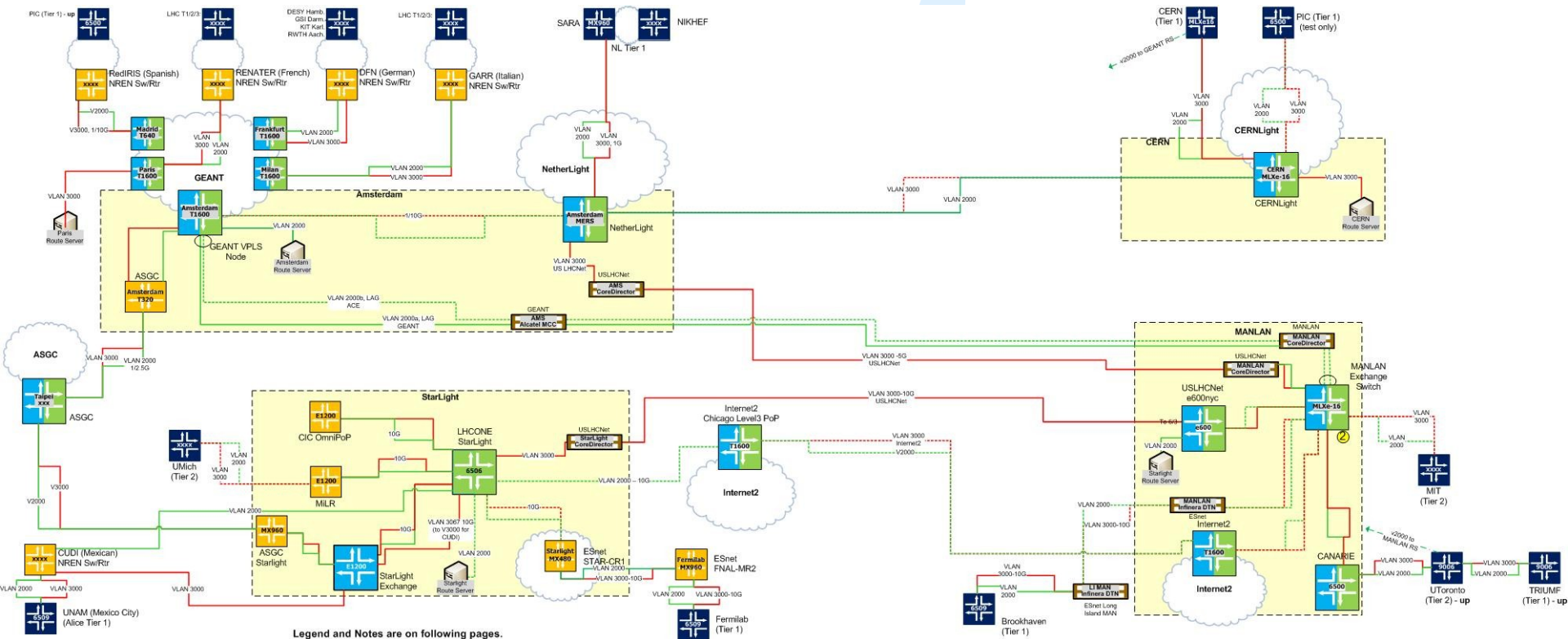
Specially during the pilot phase it has been agreed to avoid additional expenses for network carriers and sites.

LHCONE is designed for high throughput data transfers, JumboFrame (IP MTU=9000bytes) and Path MTU Discovery (RFC1191) are a must.

LHCONE layer 2 prototype - International

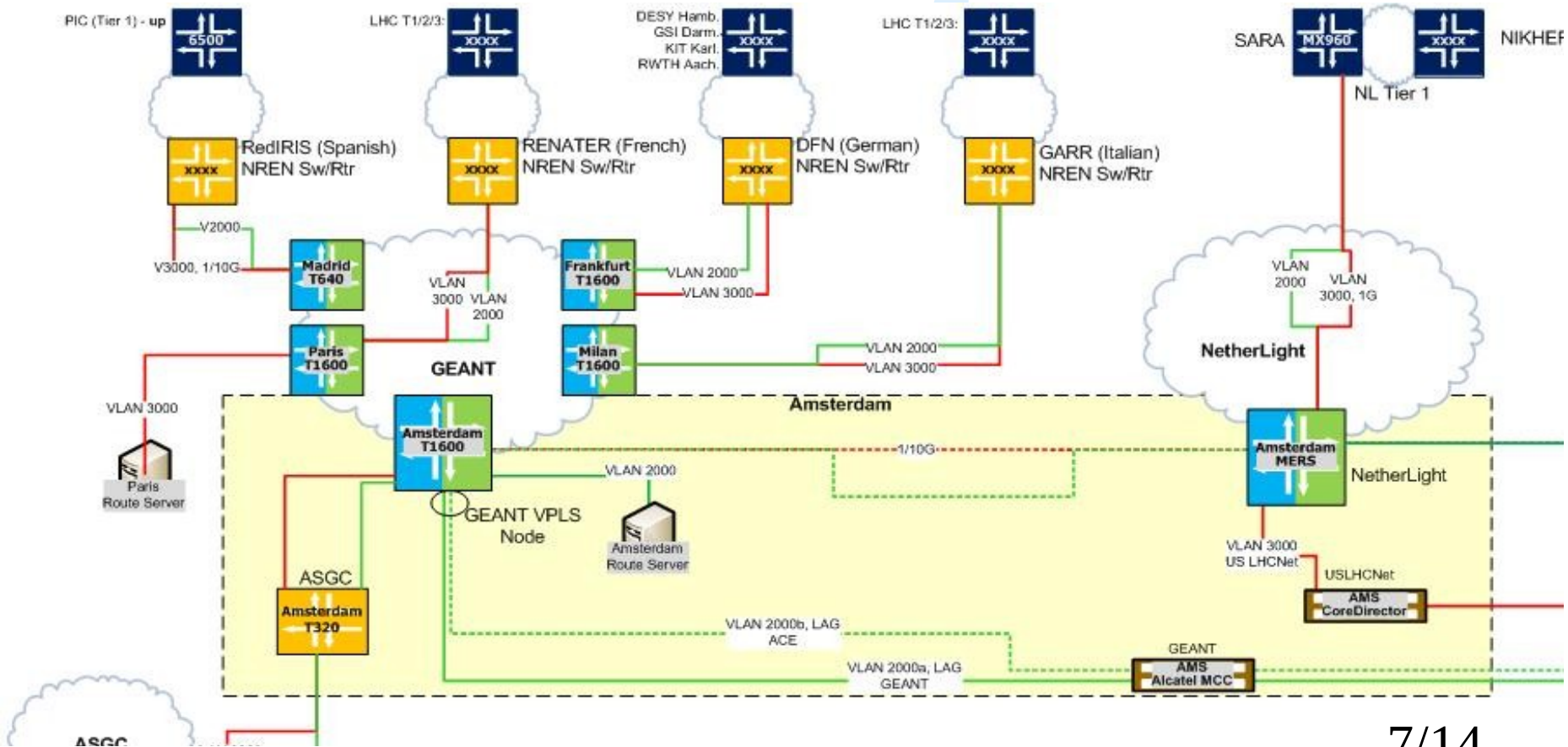
World shared 2 VLAN (2000/3000) prototype. All-All BGP peering through route servers.

Most Tier1 connected and some Tier2/3.

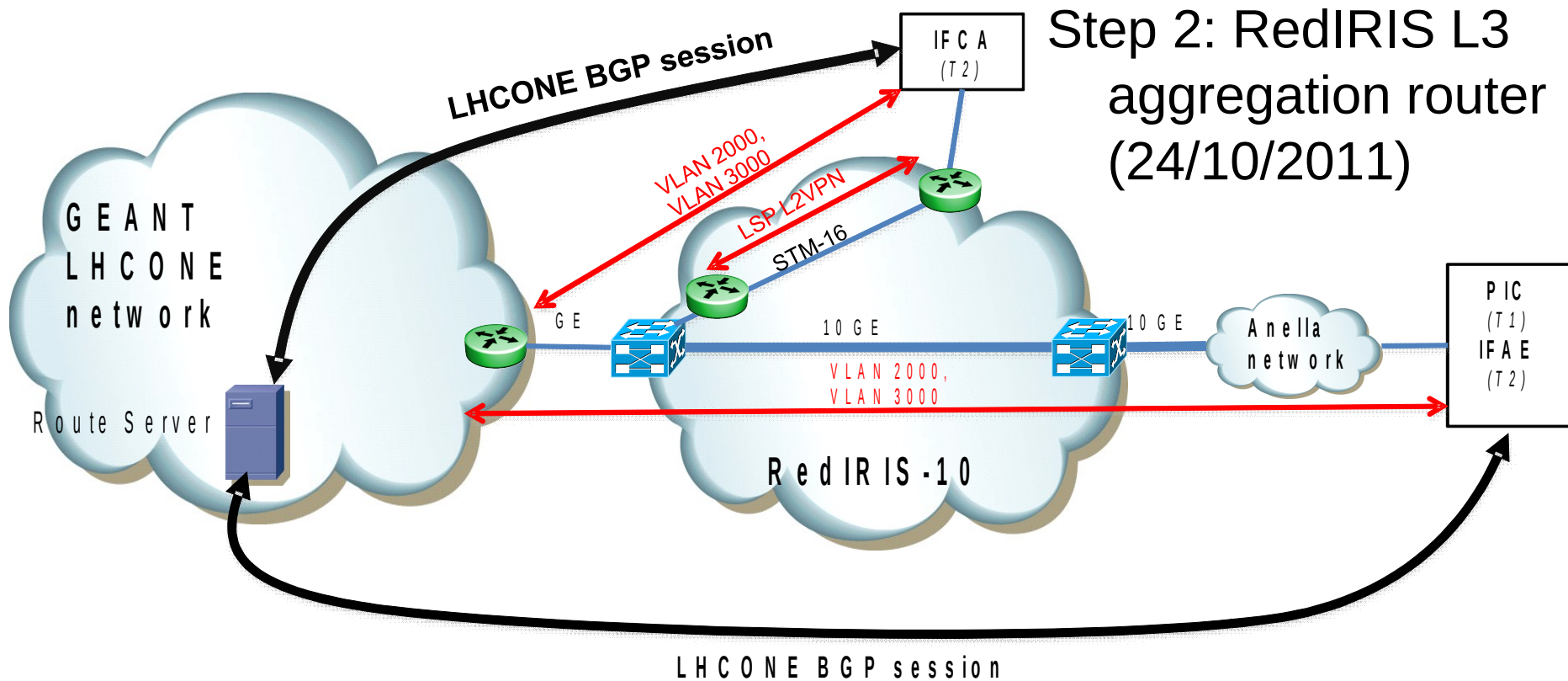


World shared 2 VLAN (2000/3000) prototype. All-All BGP peering through route servers.

Most Tier1 connected and some Tier2/3.



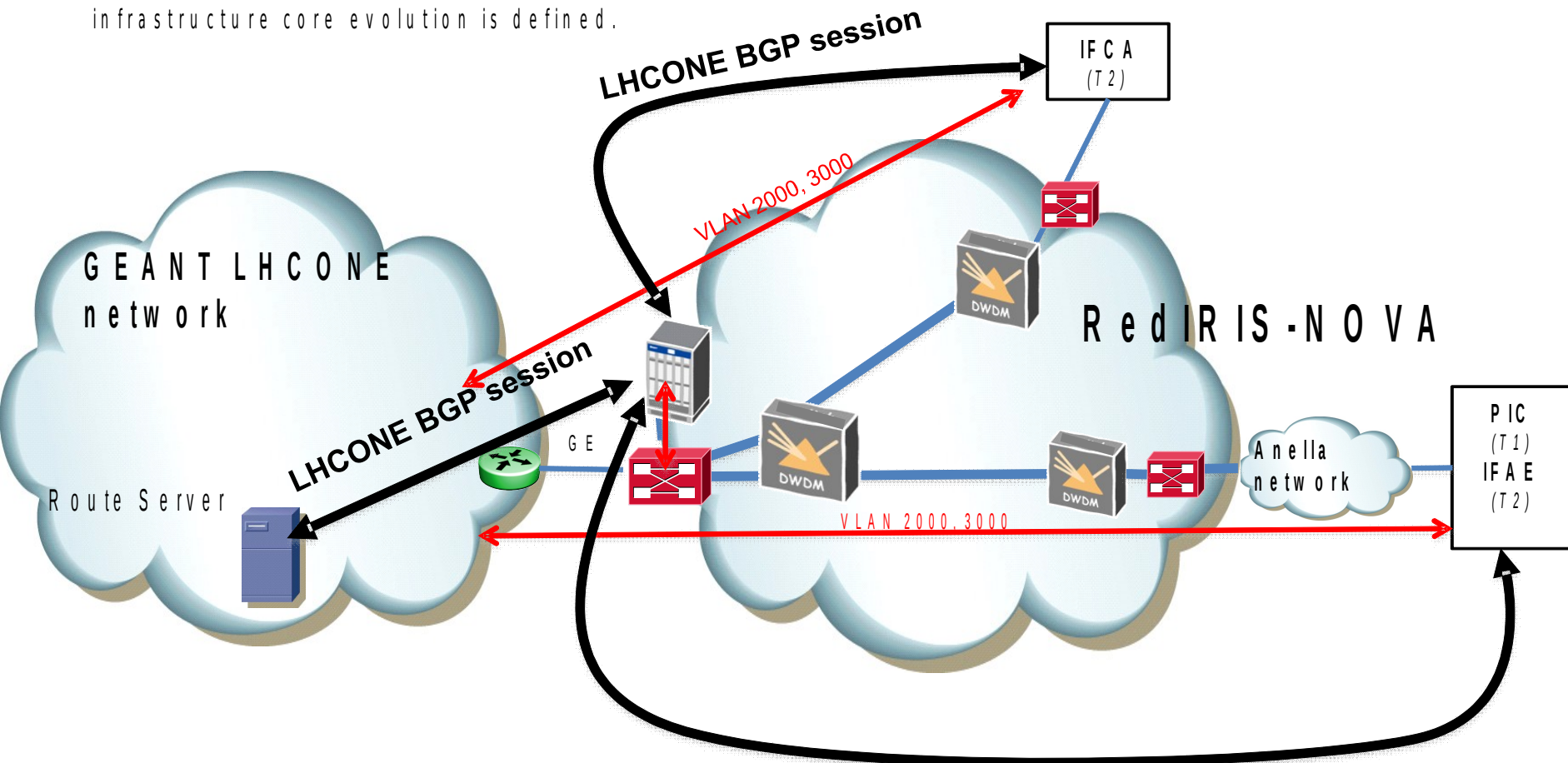
- ES T1 site : PIC at Barcelona
- There are 7 LHC T2 sites at Spain.
 - Two of them in LH CONE prototype network : IFAE (share location with PIC) and IFCA (pending internal IFCA network config)
- L2 Ethernet transport in RedIRIS backbone (in operation)



Step 2: RedIRIS L3 aggregation router (24/10/2011)

LHCONE BGP session

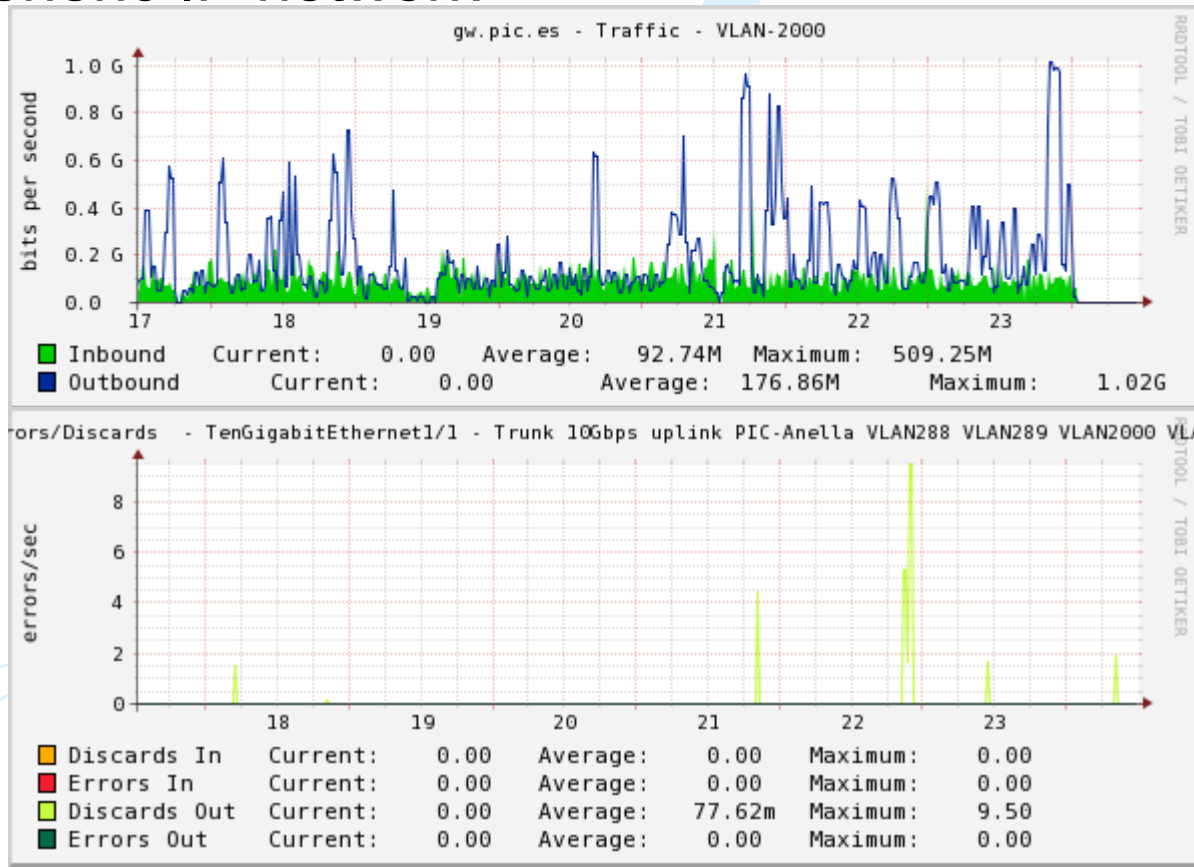
- Move physical connections to **RedIRIS-NOVA dark fiber network**, with dedicated 10G wavelengths (December 2011)
- Physical connection update between GEANT and RedIRIS-NOVA to be decided when LH CONE network infrastructure core evolution is defined.



LHCONE BGP session

PIC T1/IFAE T2 LHCONE production traffic reaching the international temporary 1GE limit.

When the prototype is down traffic automatically flows through RedIRIS Generic IP network



Move the Spanish implementation of the prototype to RedIRIS-NOVA dark fiber infrastructure.

In the last joint LHCONE&LHCOPN meeting it was stated that necessity to move traffic off general purpose IP networks is not as urgent as it was.

WLCG does not see an immediate problem from “doing physics” point of view. LHCONE should focus on laying solid foundations for long-term future (2015+).

In order to have a production infrastructure in place on 2015 we should keep working on the prototype.

Since some instabilities have been detected in the shared VLAN LHCOPN pilot infrastructure it's probable there will be some short-term architectural adjustments on the prototype.

LHCONE prototype's future will be the main topic for the face-to-face architecture meeting on 1st/2nd December
<https://indico.cern.ch/conferenceDisplay.py?confId=159549>

Building the Spanish LHCONE on top of a versatile L1 infrastructure (RedIRIS-NOVA) allows us to adapt to new requirements and most of the LHCONE future proposals

- Multipoint service
 - IEEE 802,1aq (Shortest Path Bridging)/TRILL (Transparent Interconnect of Lots of Links)
 - Hybrid Layer2&3 infrastructure
- Point-to-point service
 - Mixing static & dynamic VLANs

RedIRIS-NOVA provides high throughput links, the connected sites implement whatever intelligence needs to be implemented for the agreed LHCONE infrastructure.

Further discussion and LHCONE roadmap will probably emerge on LHCOPN&LHCONE joint meeting 30-31/1/2012.

More documents at <http://lhcone.web.cern.ch/>

Bos-Fisk paper: <http://lhcone.web.cern.ch/node/19>

For further information contact:

- gerard.bernabeu@pic.es (or network@pic.es)
- alberto.escolano@rediris.es