

ESTABLISHED IN 1993 IN SANTIAGO DE COMPOSTELA (SPAIN)



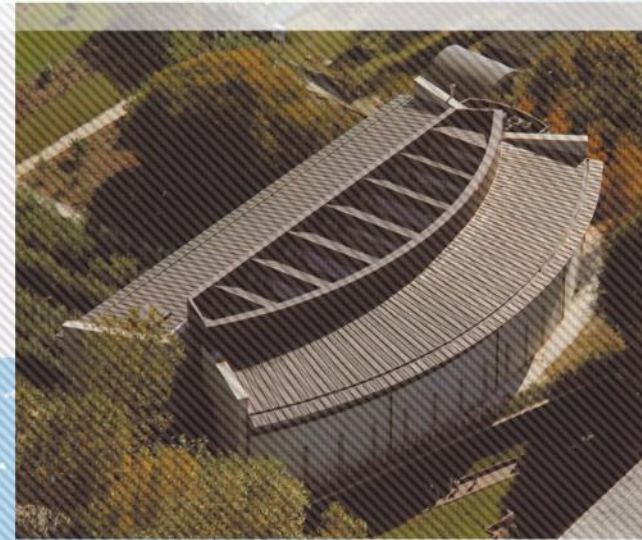
# CESGA



Infraestructuras  
Científicas y Técnicas  
Singulares



SANTIAGO DE COMPOSTELA



# MISSION

“Contribute to the advance in **Science and Technique** by means of research and the application of high performance computing and communications [...] for the **benefit of the Society**”

70%



XUNTA  
DE GALICIA

30%



CSIC

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS



# RECETGA Users

## Universities



UNIVERSIDADE DA CORUÑA  
Universidade de Vigo

## CSIC




MBG  
MISIÓN BIOLÓGICA DE GALICIA

IIM  
Instituto de Investigaciones  
Maritimas de Vigo

CSIC  
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

IIAG

## Xunta



E.S.O.  
Ensinanza  
Secundaria (\*)

## Public Bodies



CEIDA  
CENTRO DE ESTUDIOS  
UNIVERSITARIOS DE GALICIA EN  
AVILA DE GALICIA

FEUGA  
FUNCIÓN EMPRESA - UNIVERSIDAD GALICIA

VIGO  
CORUÑA  
SANTIAGO

## IEO



CENTRO  
OCEANOGRÁFICO  
DE VIGO

CENTRO  
OCEANOGRÁFICO  
DE CORUÑA

## Health



Complexo Hospitalario  
Universitario de  
Santiago de Compostela

## RTD Centres



CTC  
centro tecnolóxico da came

CIS  
CENTRO DE  
INNOVACIÓN  
E SERVIZOS  
DE GALICIA

ANFACO

Intecmar  
INSTITUTO TECNOLÓXICO  
PARA O DESENVOLVIMENTO  
E O INVERSIÓN DE GALICIA

tecnopole

Cetmar

## Other

ICTS  
Projects  
Agreements

(\*) RECETGA users  
(no infrastructure  
users)

# CESGA Services



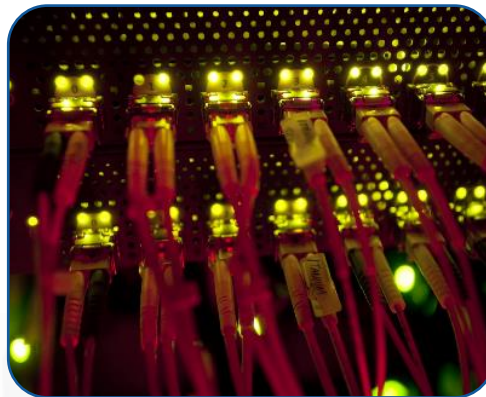
HIGH PERFORMANCE  
COMPUTING



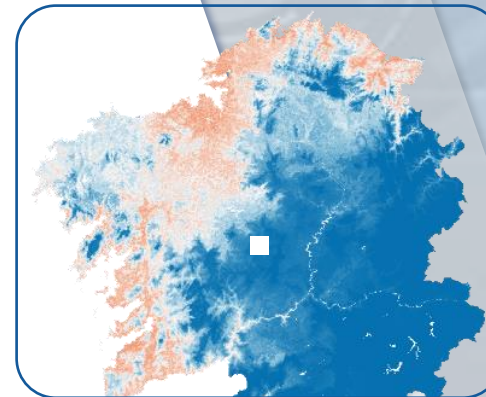
STORAGE



APPLICATIONS



COMMUNICATIONS



GIS



ELEARNING

It became nº 100 of the world in the prestigious TOP500 list

galicia



**FINISTERRAE I**



# CEEGA Infrastructures



**FINISTERRAE I**  
6000 Cores | 40TB RAM | 20GB/s I/O



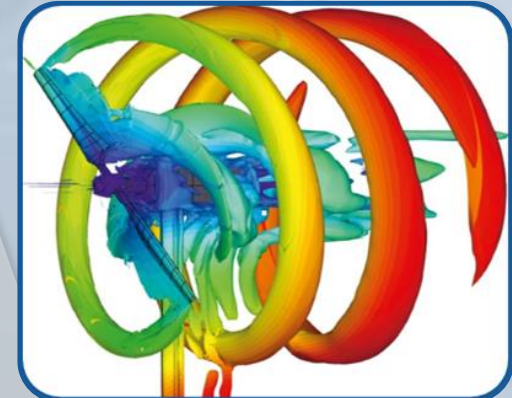
**STORAGE**  
1PB NFS, 750TB Lustre, 820TB BDATA, 100 TB Gluste



**CLOUD**  
1500 cores, 72 nodes

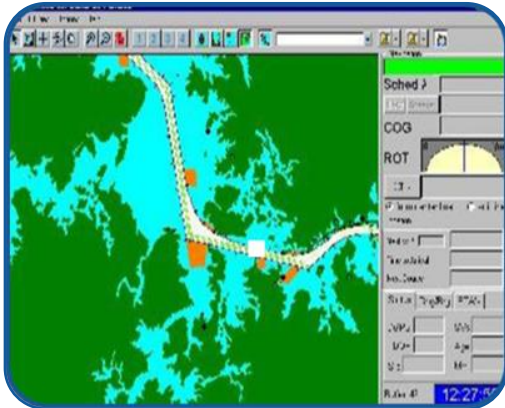


**BIGDATA**  
1 PB (1000 TB) | 34GB/s IO



**VISUALIZATION**  
80 cores

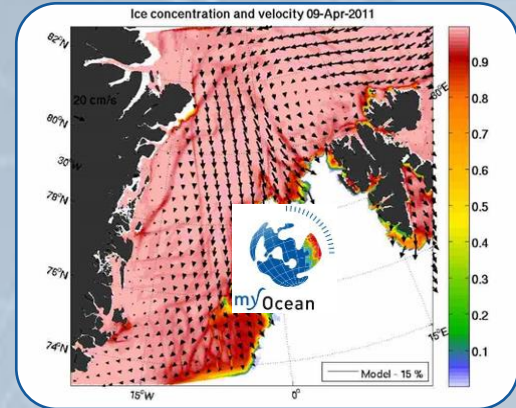
# CESGA Use Cases



Large scale electromagnetism simulation for improved radar systems

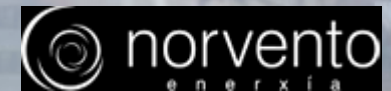
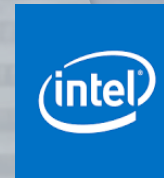
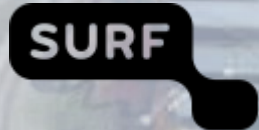
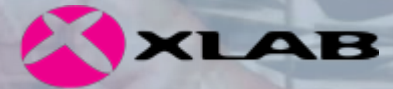


Numerical Simulation for Engineering (CTAG, Galician Automotive Technology Centre)



European IBI Area: Ocean Forecasting Service

# Latest collaborations





# Communications Department

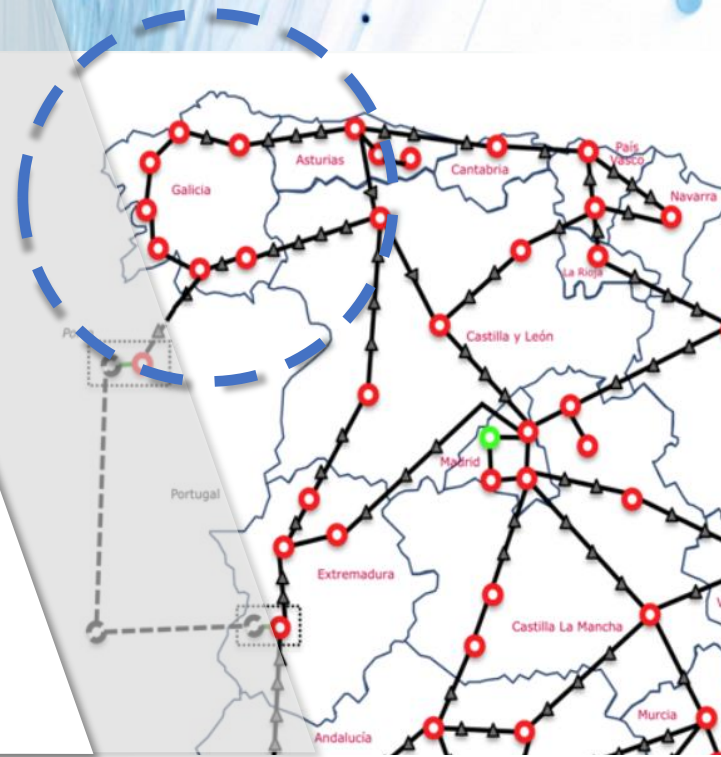
- Manages the **Datacenter** network
- Manages the Galician Regional Research Network (**RECETGA**)





**RECETGA**

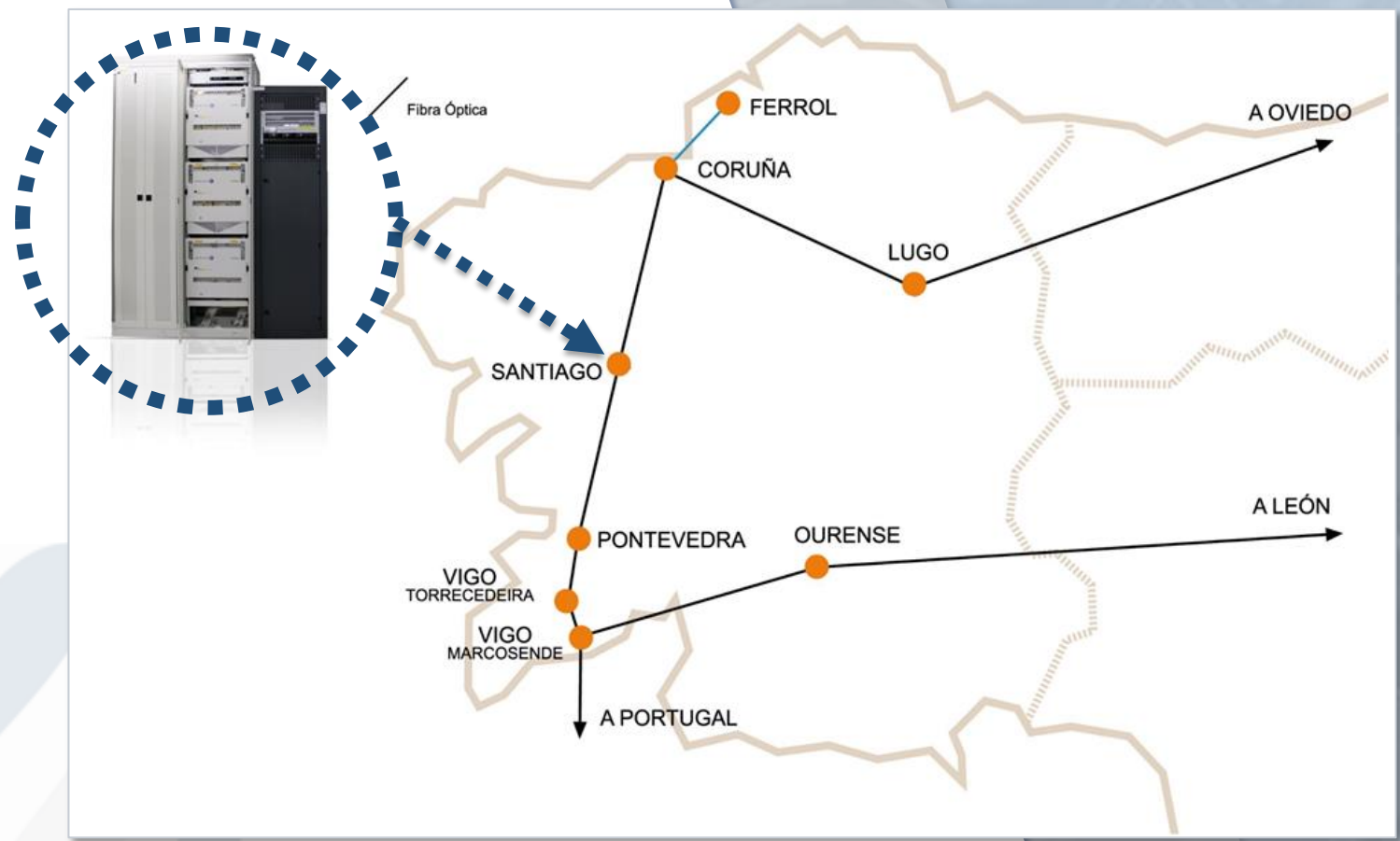
**Galicia Science and  
Technology Network**



# RedIRIS-RECETGA Dark fibre deployment



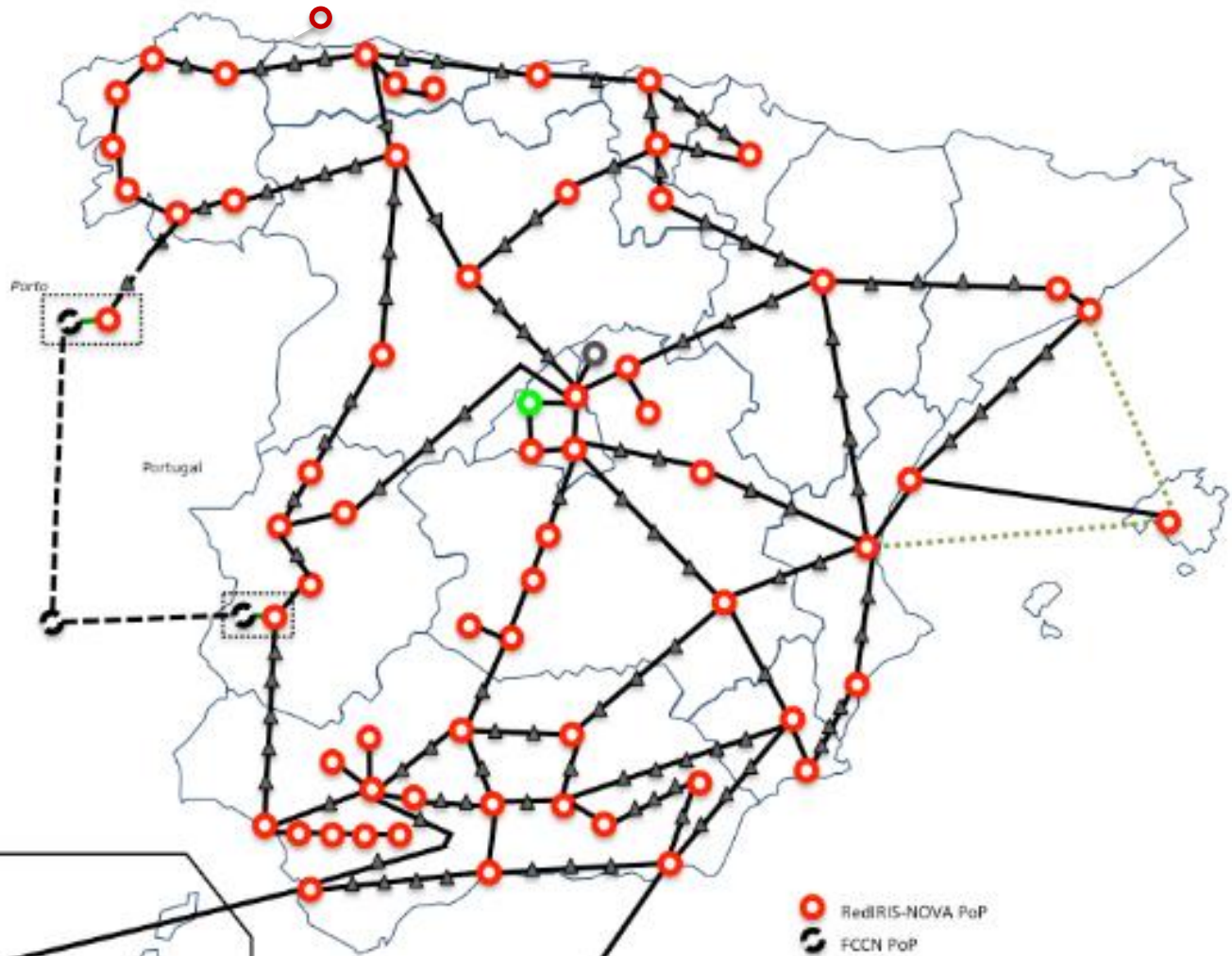
## RECETGA. 2008



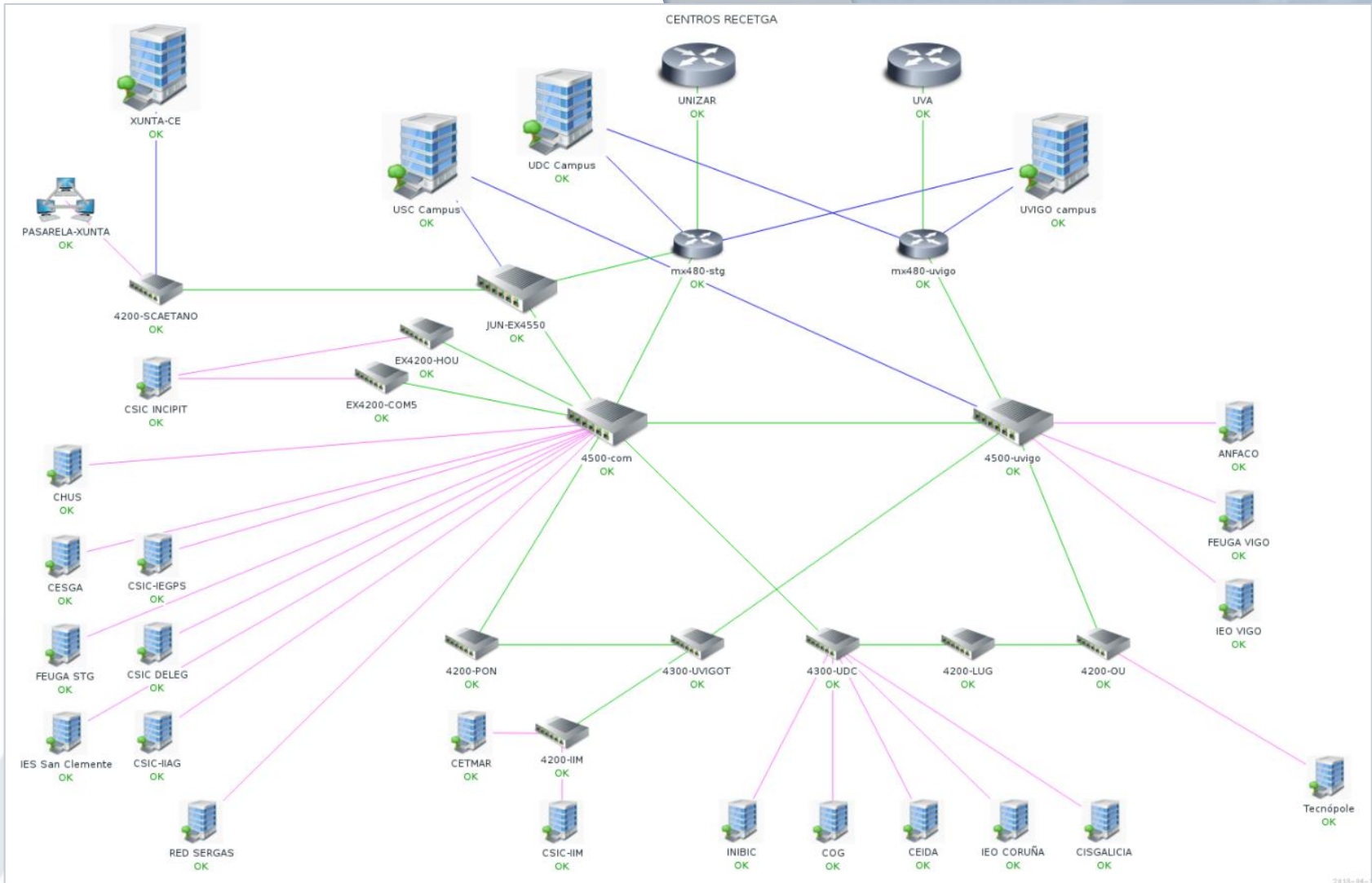
# RedIRISNova. Dark fibre footprint



Versión Noviembre 2015

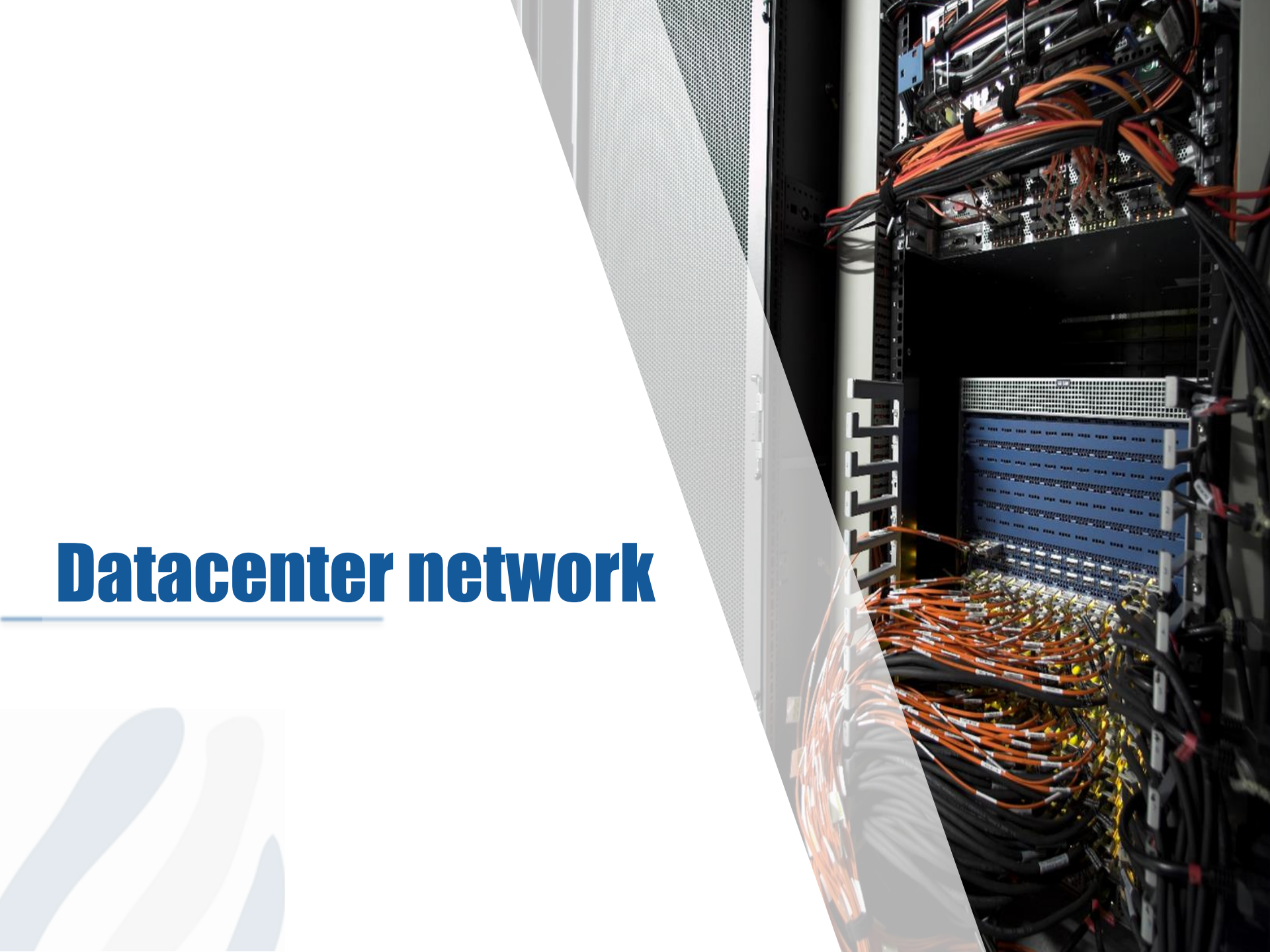


# RECETGA's NMS map

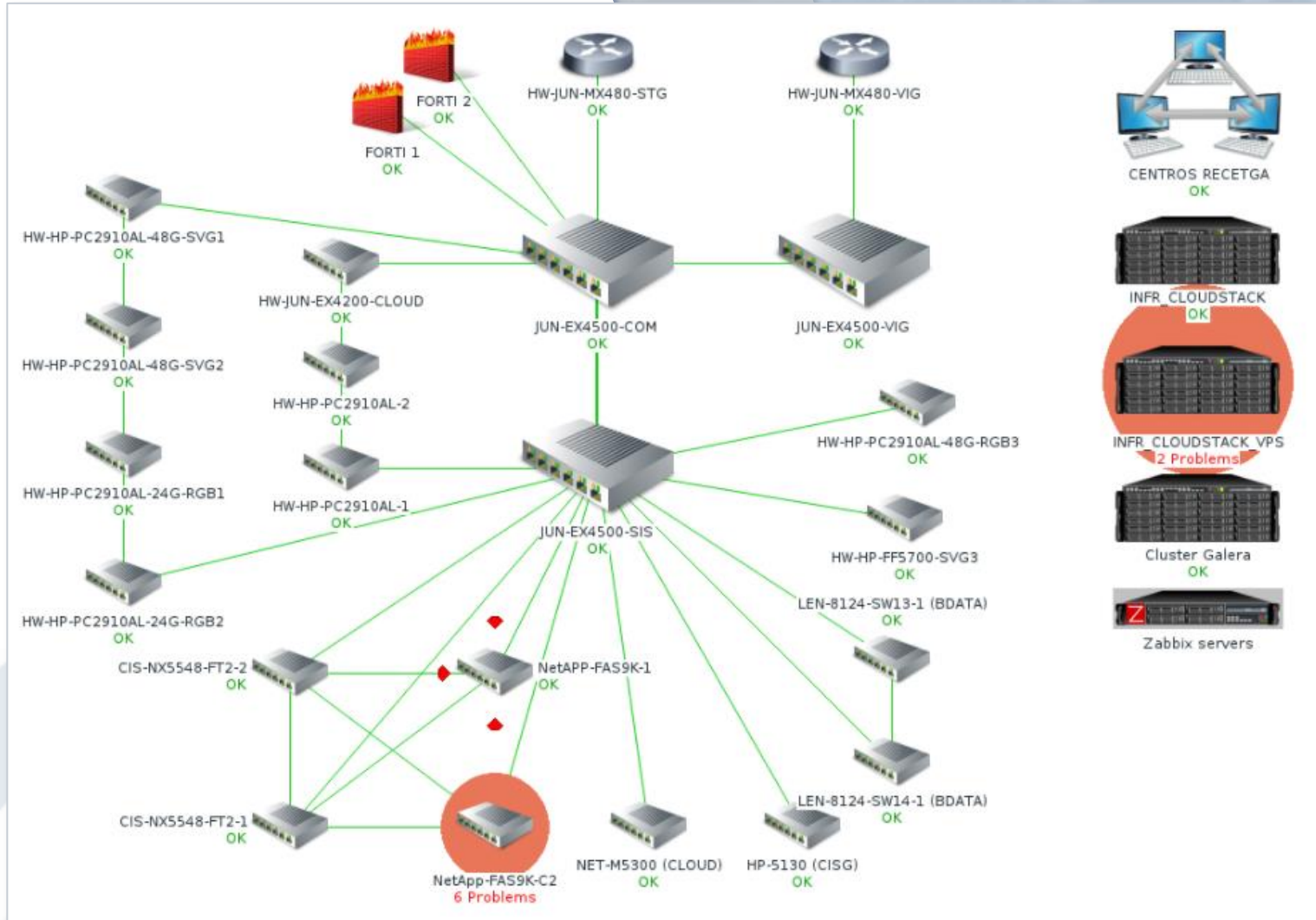


# Datacenter network

---



# RECETGA's Datacenter map



# SERVICES portfolio

## RECETGA/RedIRIS

- **Connectivity**
  - Lambdas
  - Routing IPv4/IPv6
  - VPNs
  - Multicast
- **Monitoring**
- **PERT**
- **CERT**

## RECETGA/RedIRIS

- **Mobility**
  - Eduroam
- **Email**
  - Quality: Lavadora
  - Lists, email suites
- **Digital id/Federation**
- **Data transfer**
  - Filesender
  - Aspera
  - GridFTP

## CESGA

- **Specialized**
  - Computing
  - Storage
  - Cloud
  - BigData
  - Scientific Applications
- **Generic**
  - Hosting/Housing
  - Videoconference
  - LMS
  - SSL VPN
- **Support**
- **Projects**

(\*) Servicios por afiliación a la comunidad académica y de investigación RedIRIS



# RECETGA NOC tools



## Network Monitoring

- Zabbix
- Pandora FMS
- Own code

## Other

- Request tracker
- Rancid
- Data transfer tools

# CESGA monitoring & analysis tools



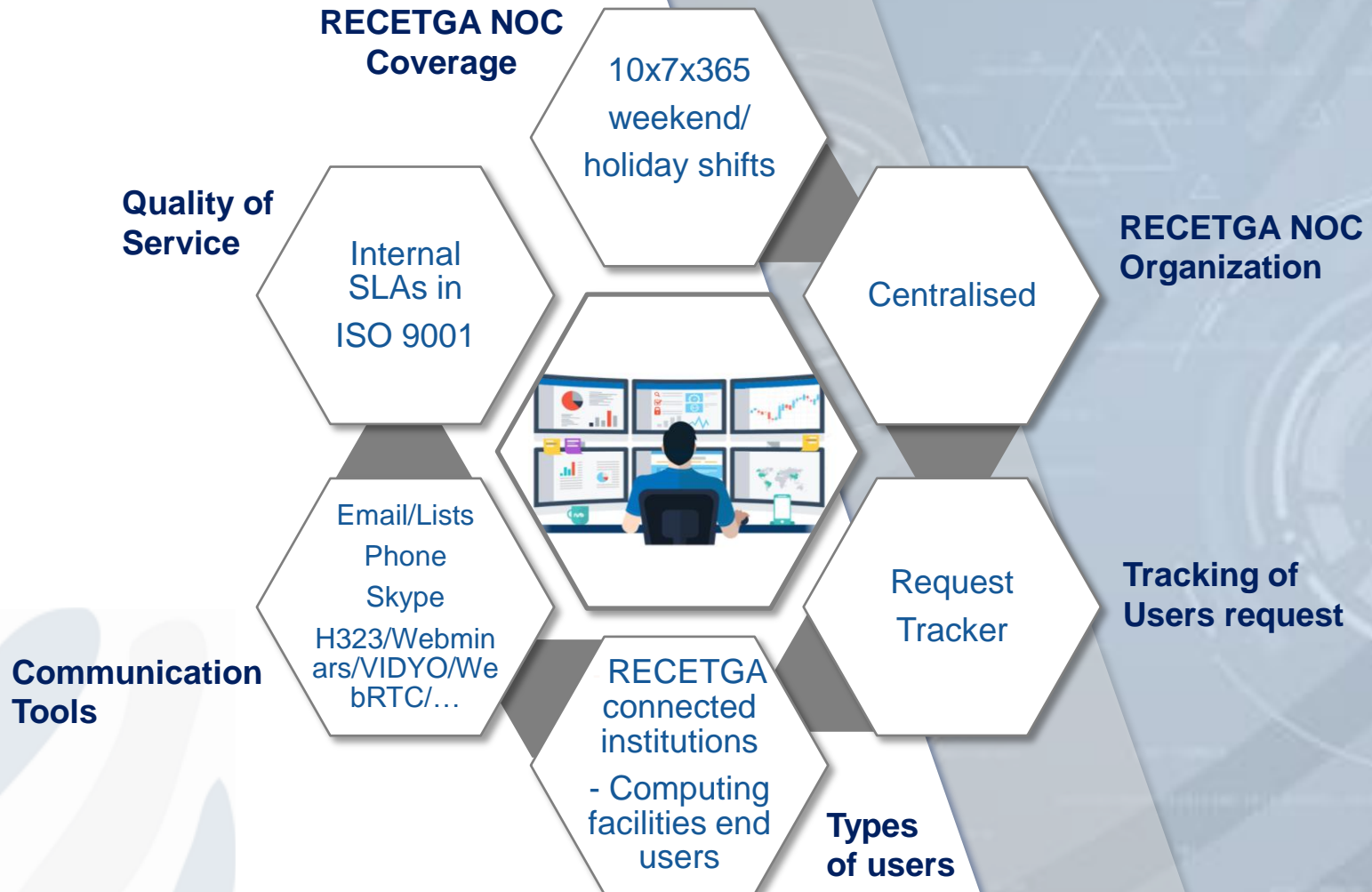
## Systems monitoring

- Icinga2
- Collectd
- Bosun

## Metrics & Logs analysis

- Bigdata platform
  - Opentsdb
  - Hbase
  - Spark
  - Elastic Search+Kibana
  - Grafana

# NOC internals



**Do the users move  
their data fast?**

---



# Max theoretical elephant BW (p=0.046%)

It seems they DONT(\*)

(\*)but in tuned scenarios

max. theoretical BW  
ELPHNT (p=0.0046%)



Mathis. et al Formula

# Transference issues

## Cases

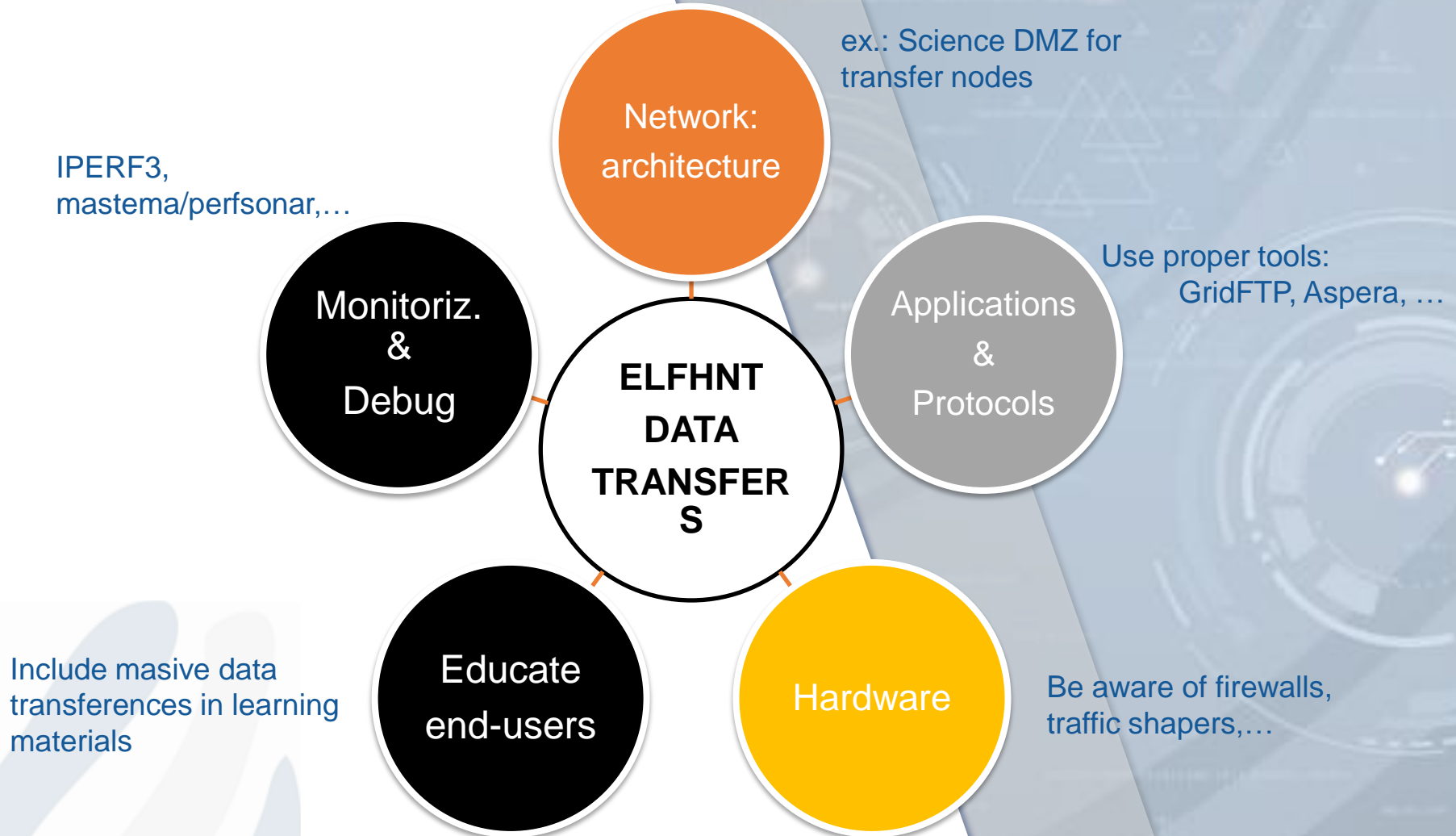
- Puertos del Estado (State Ports)
  - Meteo France – CESGA
  - Used tool: ftp
- Genomics
  - CESGA – UUEE
  - Used tool: sftp
- RES users
  - CESGA – UM, or BSC
  - Used tool: scp
- Educational Video Recordings
  - CESGA-UV, UC,..
  - Used tool: rsync

## Difficulties

- We know the problem and how to solve it but:
  - Users do transfer to/from not tuned endpoints (home, wifi, office,...)
  - User do not know proper tools
  - Institutions sometimes are not concerned about massive data transferences as.. “it works”

(\*) Servicios por afiliación a la comunidad académica y de investigación RedIRIS

# Improve elephant data transferences



# Anomaly detection at scale

---

Ref: Work conducted by  
Javier López Cacheiro



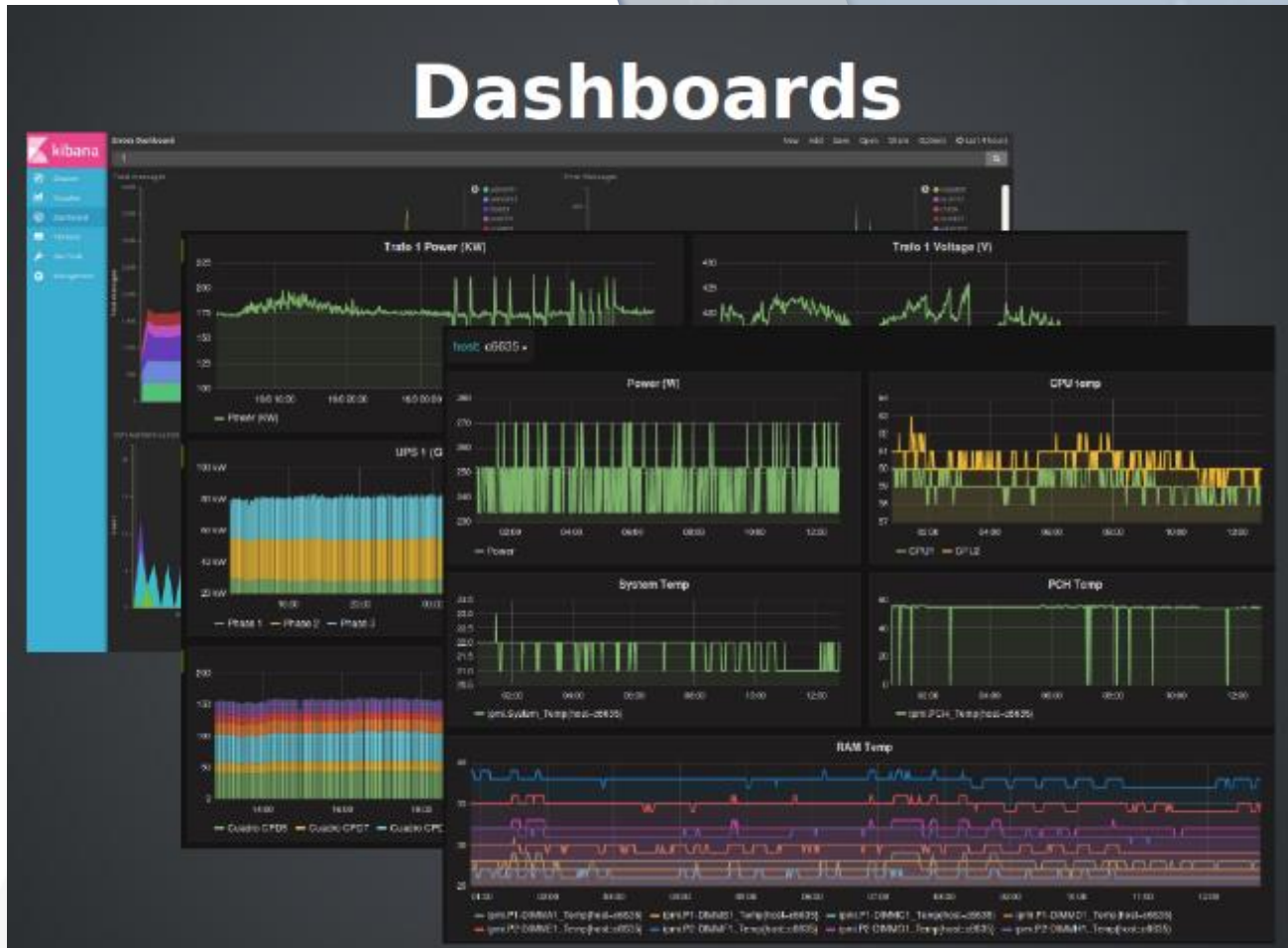


# Measure → Collect & Store

- System logs
- Process accounting
- Server metrics
- IPMI sensors
- SMART disk info
- Chillers & AHUs
- UPSs
- Power Meters
- Network traffic
- Application module usage
- SLURM Accounting
- Temperature/Humidity sensors

**33487** metrics | **10 million** time series

# Understand → Analyze & Visualize



# Control → Alert & Monitoring

Current Incidents    Overdue    Muted    ▼    ↻

---

### Service Problems

CRITICAL 12:28	c7235: Log Alerts Singularity: sexec (U=12529, P=22457)> Retval = 255	! 🔗
CRITICAL 12:28	c7236: Log Alerts Singularity: sexec (U=12529, P=28244)> Retval = 255	! 🔗
CRITICAL 1d 22h	c0511: ssh Server answer:	!
CRITICAL Sep 27	puertos2.cesga.es: Radares 3, K0, 20170731, ftp_off, PHY_RADARH, ID1	!
CRITICAL Aug 3	gestion-sis.cesga.es: Partition /mnt/EMC/Store_uscfm DISK CRITICAL - free space: /mnt/EMC/Store_uscfm 115354 MB (0.71% inode=100%):	! ●

### Recently Recovered Services

OK 1m 6s	c0621: Load OK - load average: 32.00, 31.97, 31.70
OK 20m 28s	c1102: ping4 PING OK - Packet loss = 0%, RTA = 0.44
OK 46m 3s	c7101: IB Status (0x00000000000000000000000000000000 00000000000000000000000000000000 00000000000000000000000000000000 00000000000000000000000000000000 00000000000000000000000000000000 0000)
	c7102: IB Status (0x00000000000000000000000000000000

# Improve → Anomaly detection

## Anomaly Detection

- Types of anomalies in time series
  - Outliers
  - Change points
  - Anomalous time series
- Generic Anomaly Detection Systems

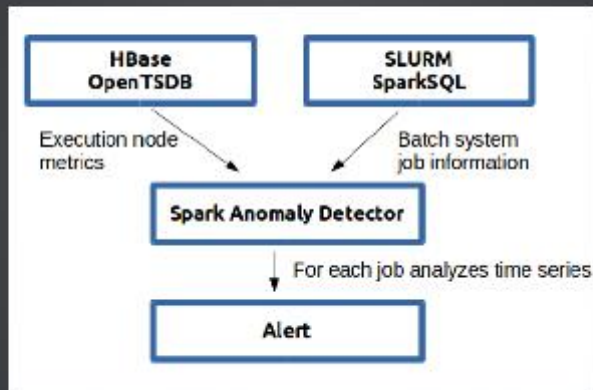
# Use case: Server anomalous performance

**Problem:** Parallel Jobs are cancelled because some of the nodes have poor performance. Computation is lost.

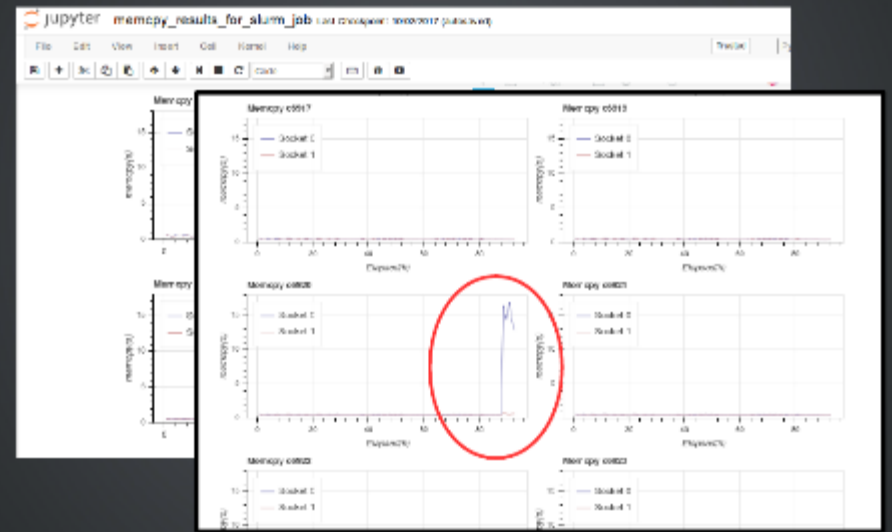
**Detection:** Analyze&Visualize server metrics to spot the anomalous node

**Objective:** Automatically detect low performance nodes

## Anomalous Performance Detection



## Analyze & Visualize

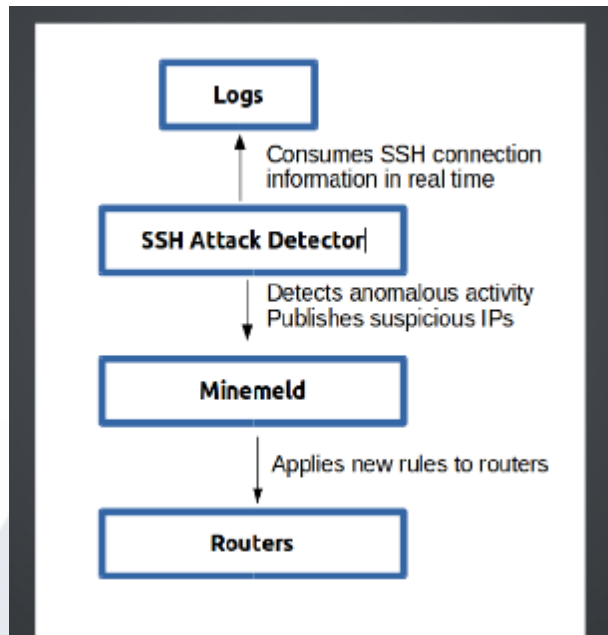



# Use case: SSH attack detection

**Problem:** Daily our public servers are scanned and attacked

**Detection:** Correlate real-time SSH connection information to detect attacks.

**Objective:** Automatically update router configuration to stop the attacks.



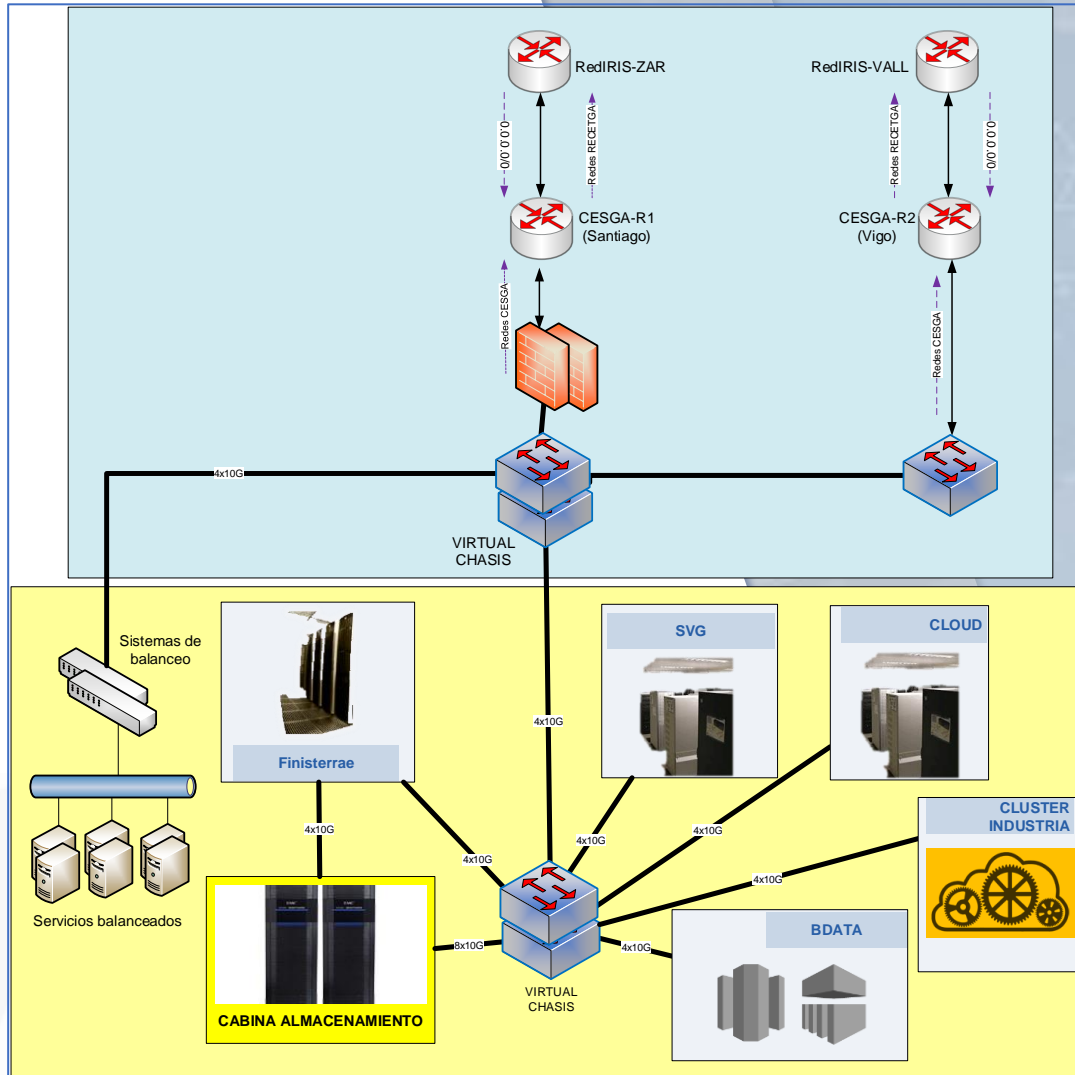


#WeAreHPC

**THANK YOU**  
natalia [a] cesga [.] es

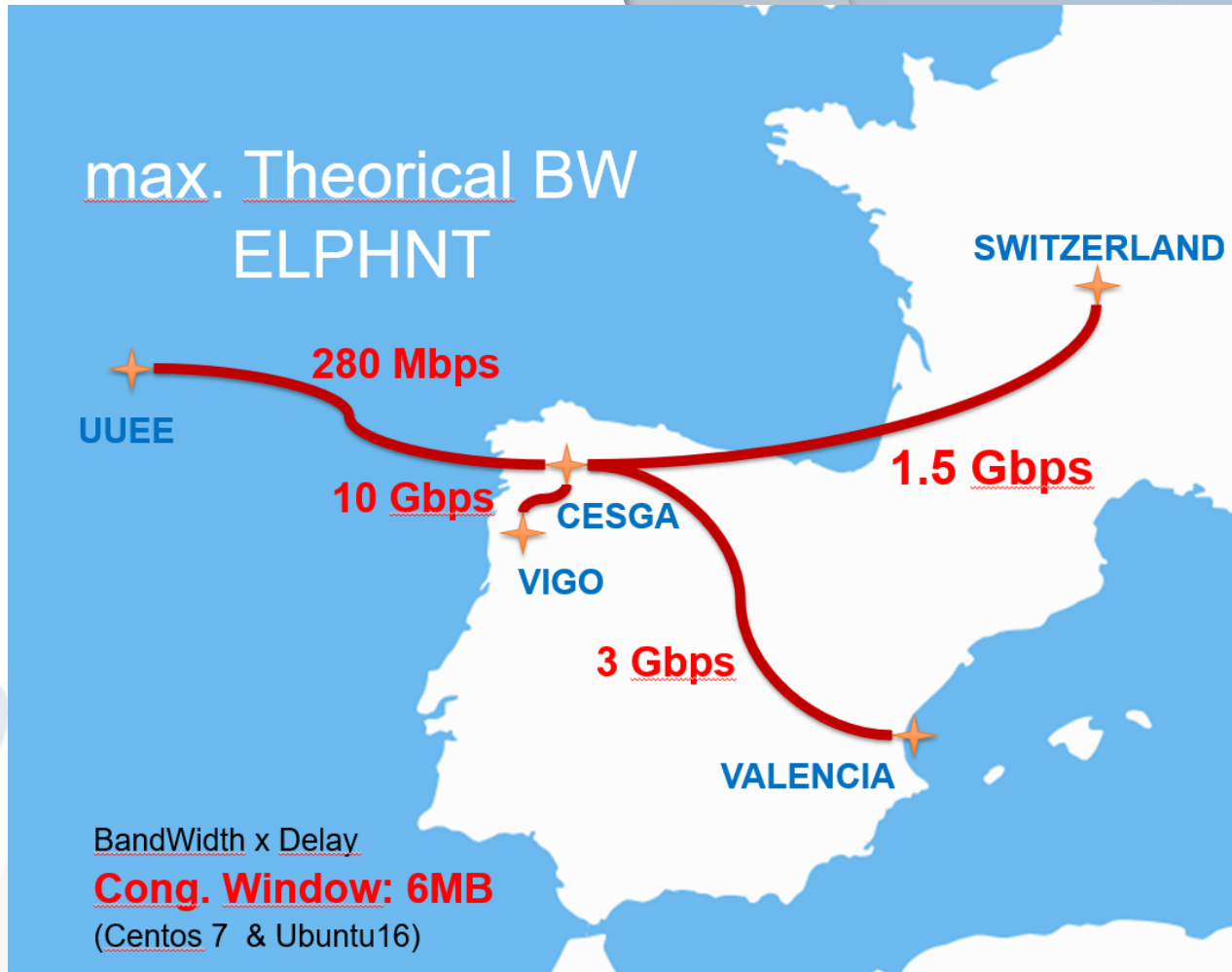


# CESGA's internal network





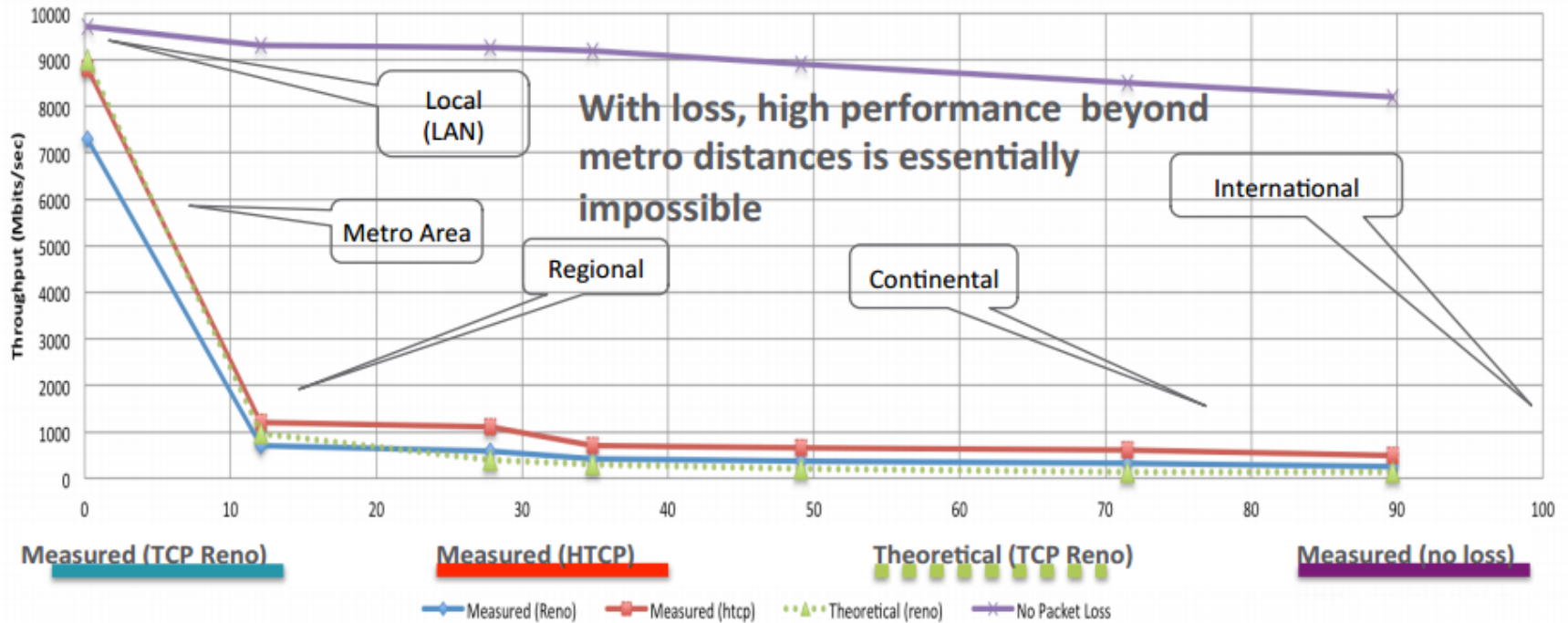
# Max theoretical elephant bandwidth



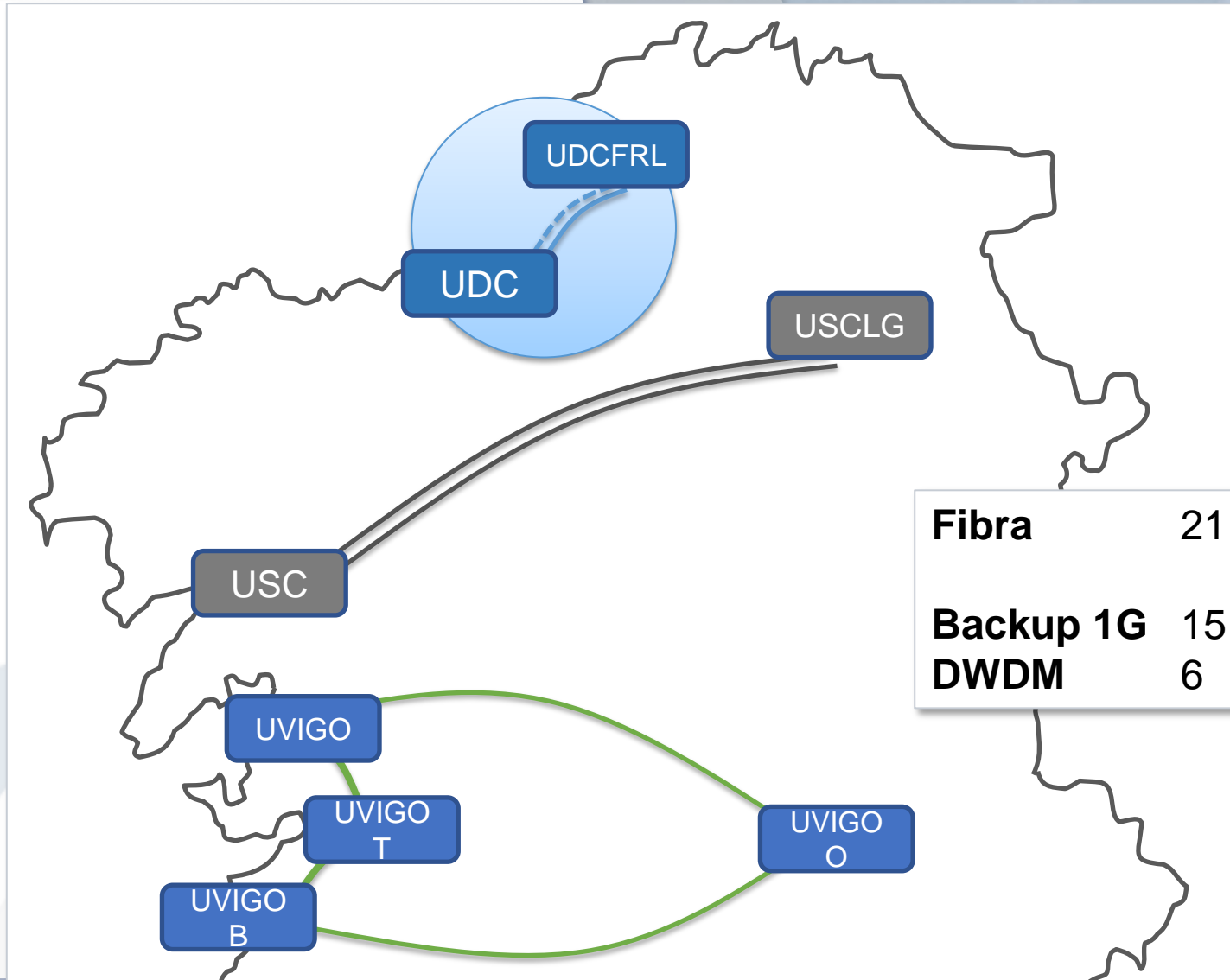
# What happens when there is packet loss?

## ESNET case

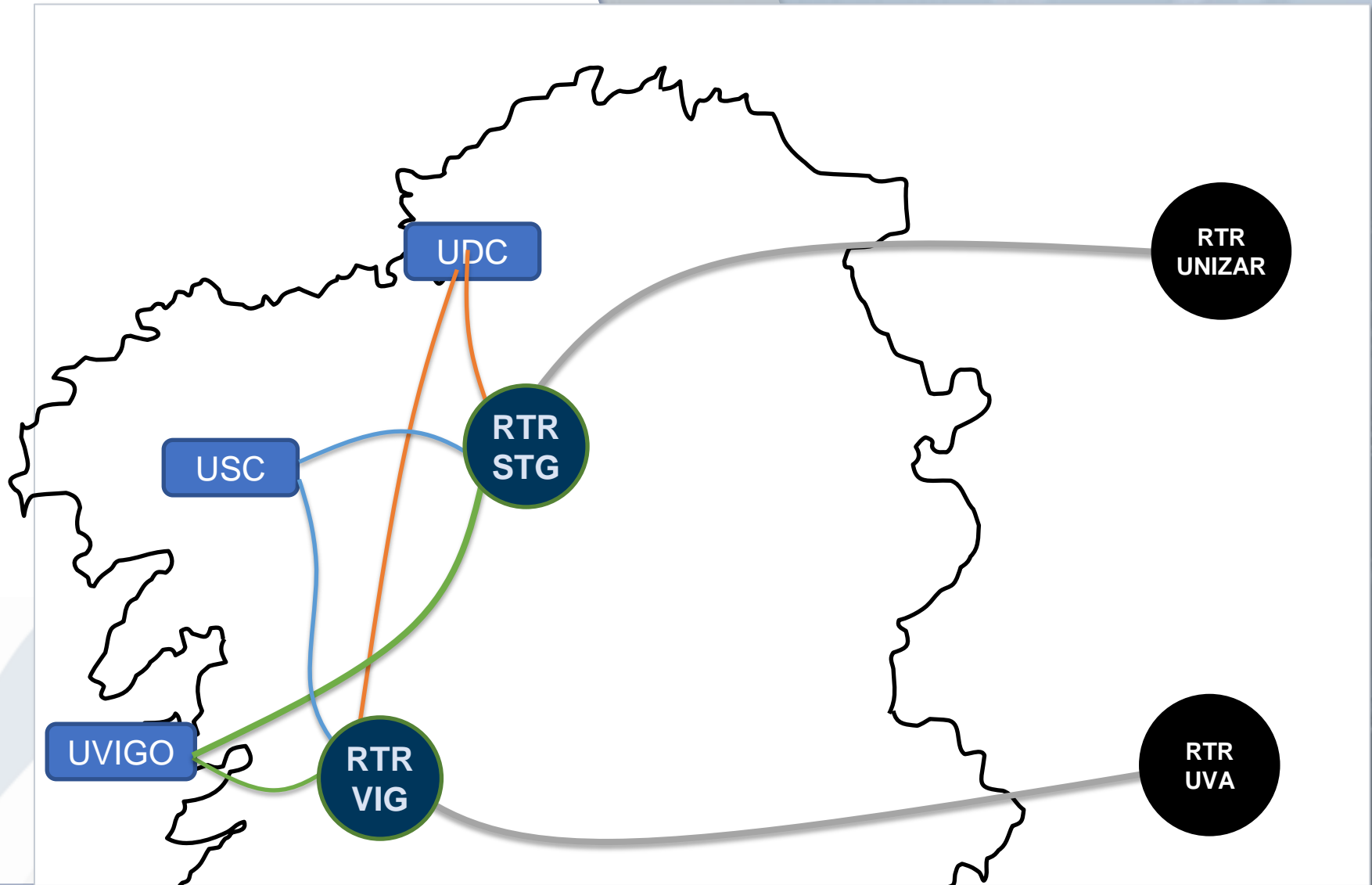
Throughput vs. Increasing Latency with .0046% Packet Loss



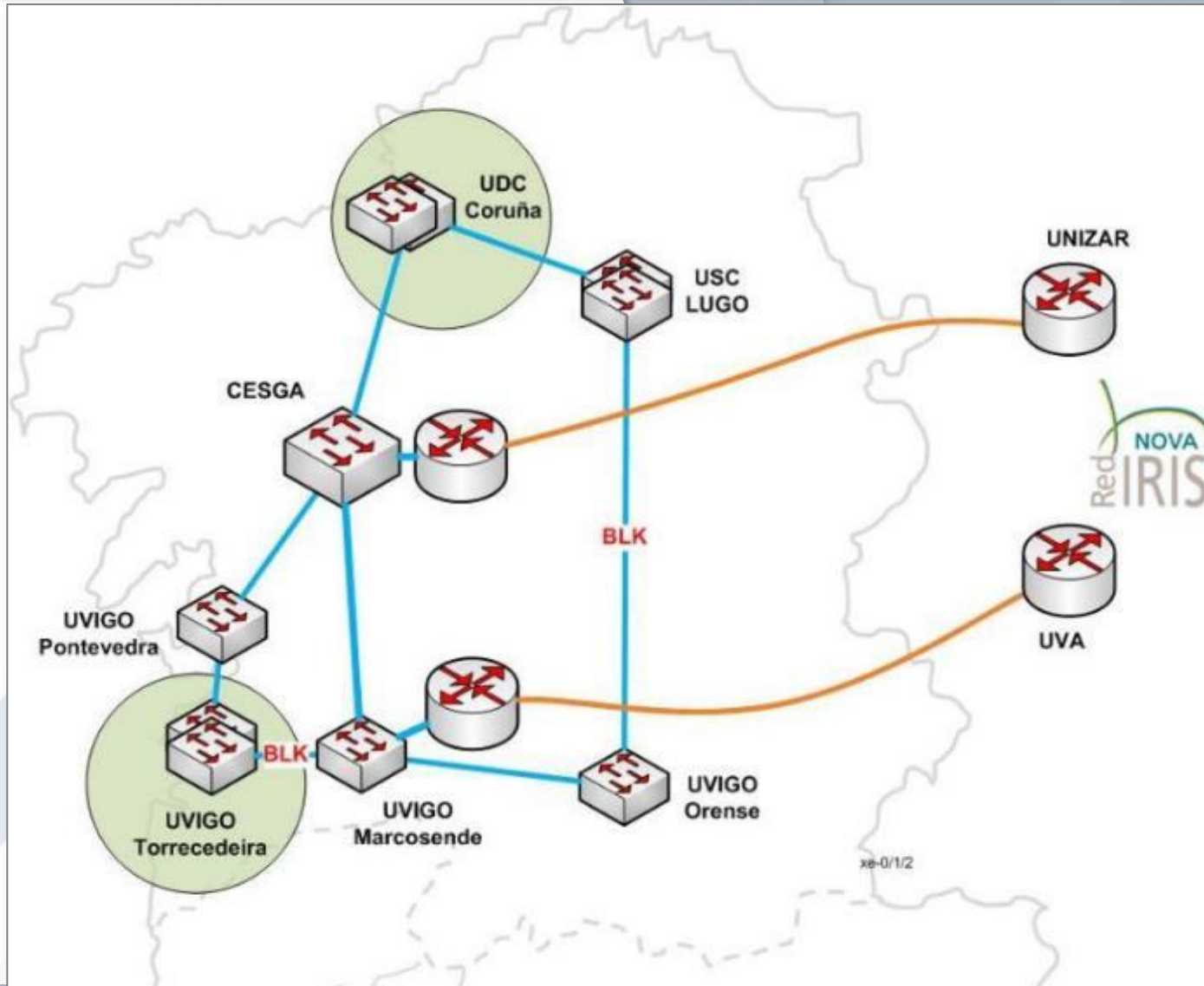
# RECETGA. Universities optical fibre circuits



# RECETGA IP routing service



# RECETGA layer 2 network



# CESGA's equipment is multivendor



## Diferents

- Configurations
- Features
- Parts

## Adapt

- Backup scripts
- Config scripts