CERN network overview

Ryszard Jurga

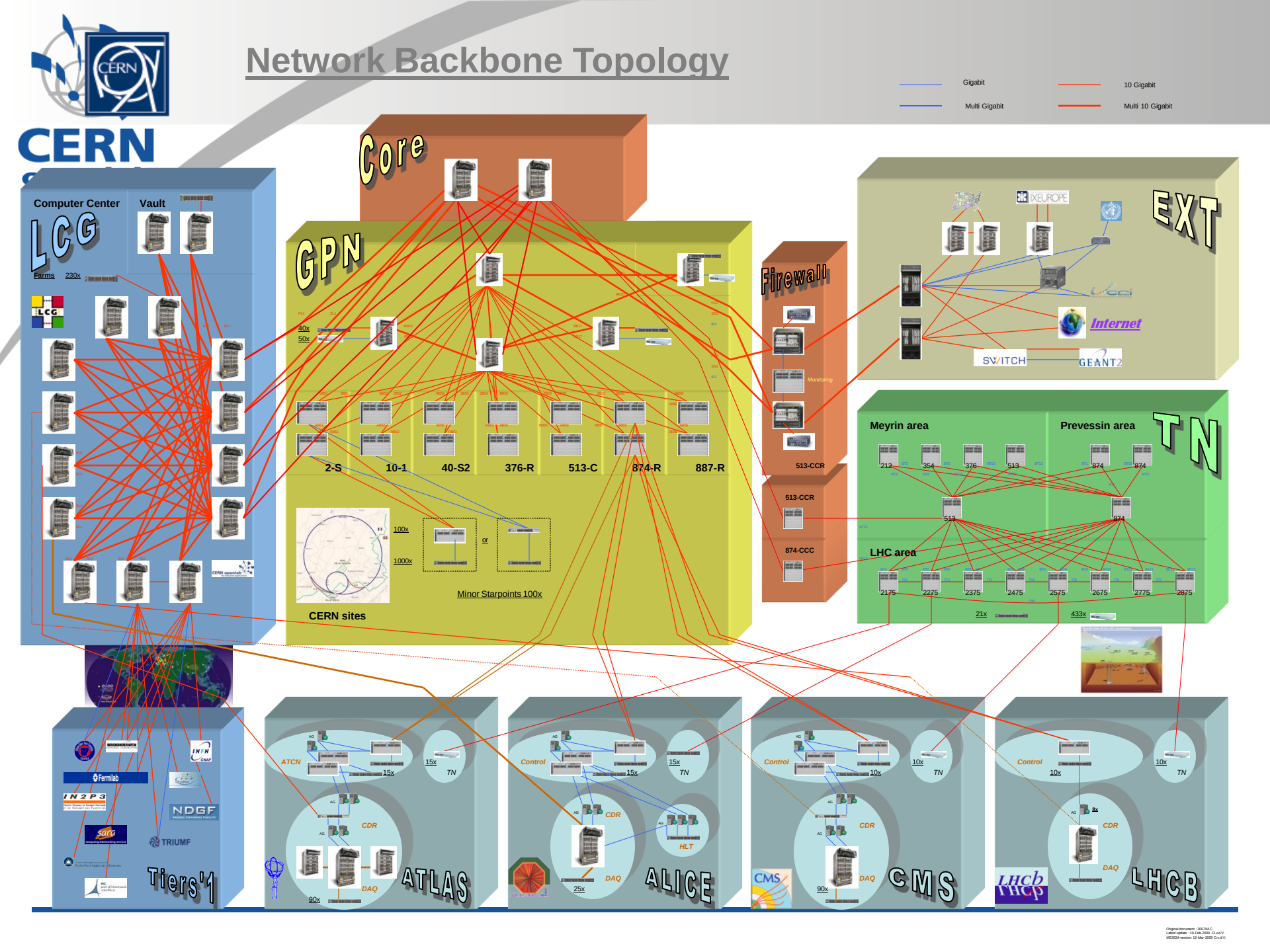
CERN openlab Summer Student Programme 2010

Outline

- CERN networks
  - Campus network
  - LHC network
- CERN openlab NCC

CERN campus network

Network Backbone Topology



Original document : 2007/MC
Latest update : 19-Feb-2009 O.v.d.V.
MS3634-version 13-Mar-2009 O.v.d.V.

The CERN Networks Facts and Figures

- Three distinct multi-ten-gigabit backbones
- 150+ very high performance routers
- 3'700+ subnets
- 2200+ switches (increasing)
- 50'000 active user devices (exploding)
- 80'000 sockets – 5'000 km of UTP cable
- 400+ starpoints (from 20 to 1'000 outlets)
- 5'000 km of fibers (CERN owned)
- 150 Gbps of WAN connectivity
- 4.8 TB LCG core
- Desktops, HPC, VoIP, Process control, etc

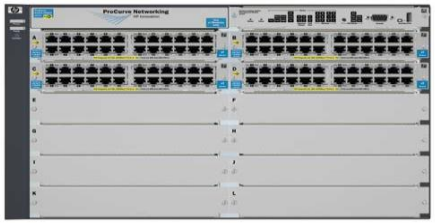
- Extremely Dynamic
  - 1'500+ requests for Moves-Adds-Changes per month
  - ISP like (2x more visitors than staff)

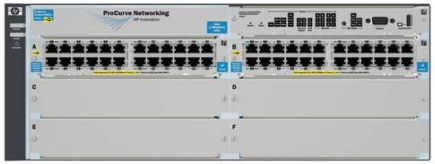
- Multi vendor site using only standards

CERN openlab

Force10 for the CORE

HP for distribution





Network Management (1)

- CERN is “particular”
  - Very large infrastructure all centrally managed
  - Large number of network devices (of different manufacturers)
  - Very limited staff, rationalization
  - 10+ years of development of our own NMS
    - ~5 software developers
    - Extremely high level of automation
    - 500’000+ lines of code/150+ DB tables
  - Only one commercial package
    - Network Supervision: SPECTRUM (CA)

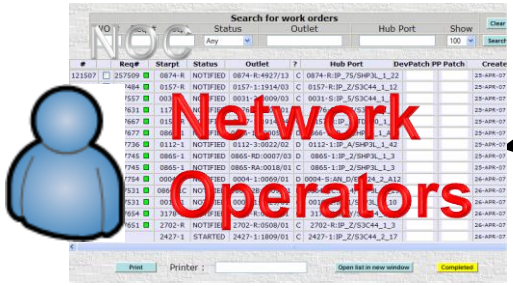
Network Management (2)

Users

SOAP

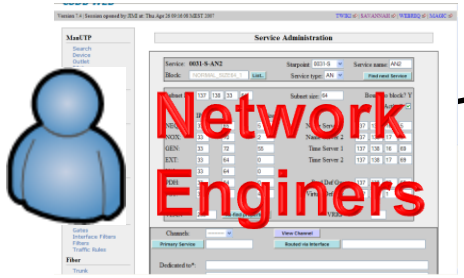
Devices
States
Topology
Addressing
Cables
Fibers
Security (FW)
SLAs
User Requests
Tickets

2500 network devices



Network Operators

Configuration Manager



Network Engineers

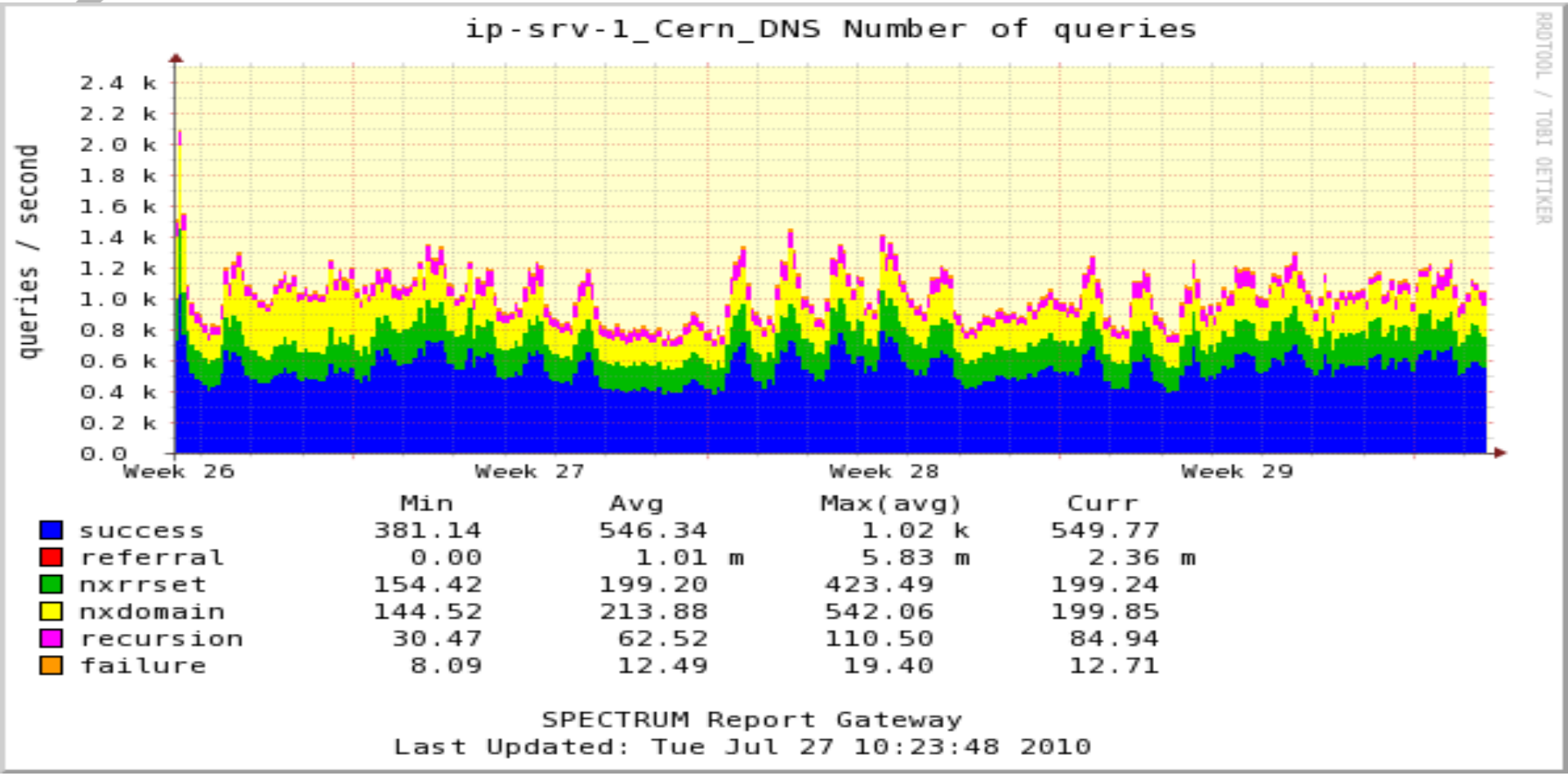
HR
Active Dir.
Others

Firewall

"If it does not fit, it will not happen"

DNS Stats

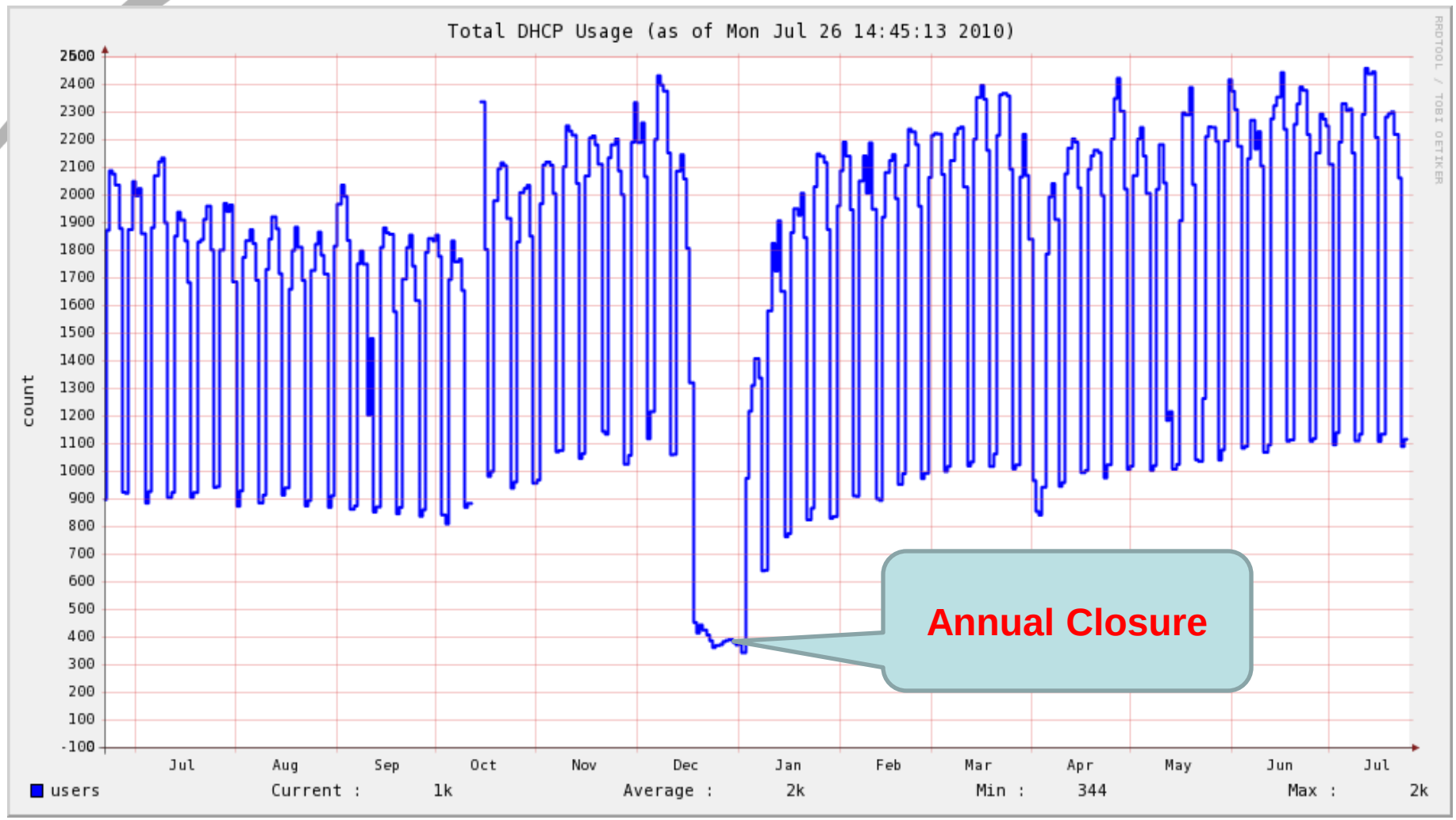
~2500 DNS queries/sec (all CERN DNS servers)



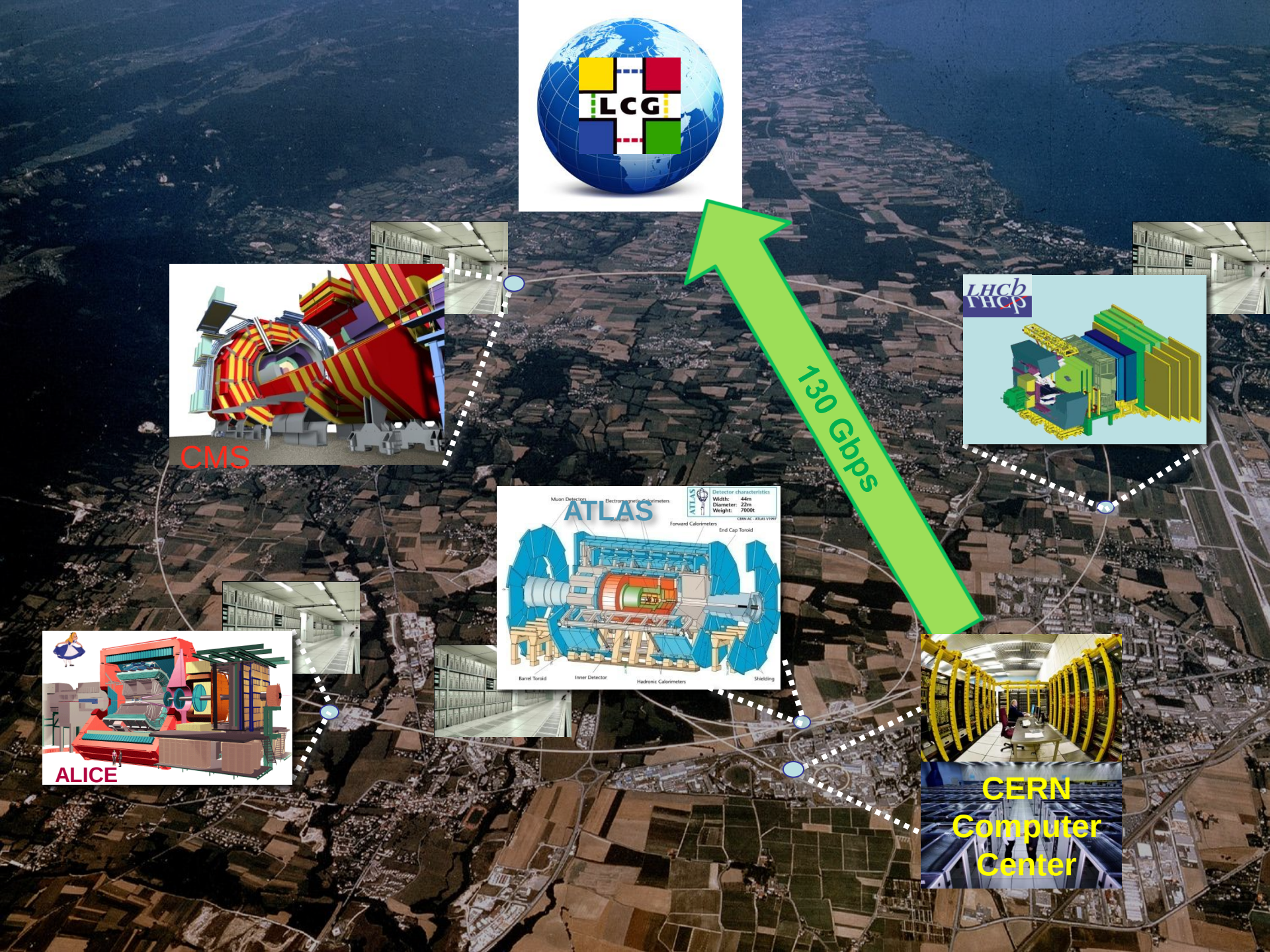
	Min	Avg	Max(avg)	Curr
success	381.14	546.34	1.02 k	549.77
referral	0.00	1.01 m	5.83 m	2.36 m
nxrrset	154.42	199.20	423.49	199.24
nxdomain	144.52	213.88	542.06	199.85
recursion	30.47	62.52	110.50	84.94
failure	8.09	12.49	19.40	12.71

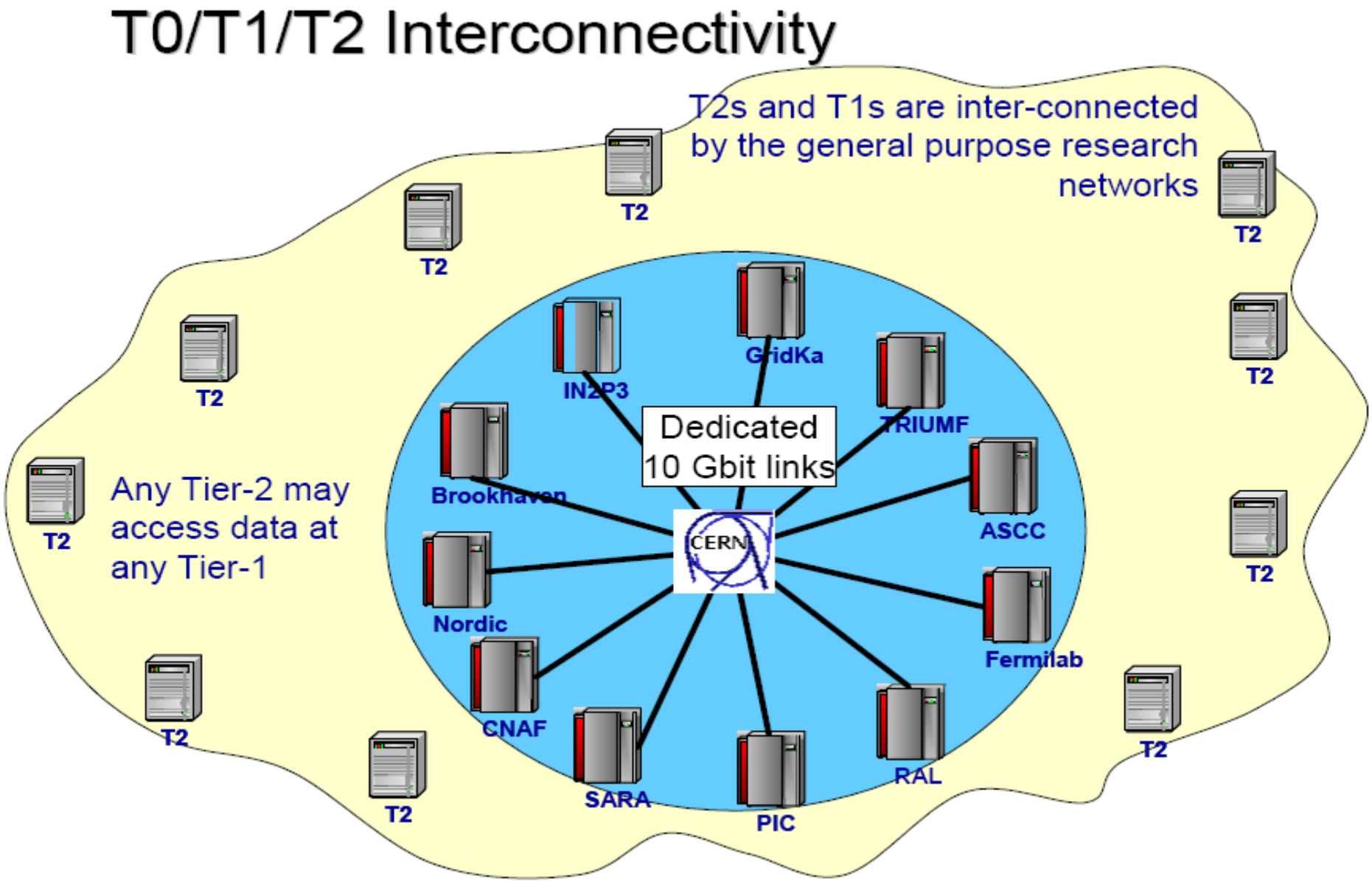
SPECTRUM Report Gateway
Last Updated: Tue Jul 27 10:23:48 2010

DHCP Stats


Total DHCP Usage (as of Mon Jul 26 14:45:13 2010)
count
2500
2400
2300
2200
2100
2000
1900
1800
1700
1600
1500
1400
1300
1200
1100
1000
900
800
700
600
500
400
300
200
100
-100
Jul
Aug
Sep
Oct
Nov
Dec
Jan
Feb
Mar
Apr
May
Jun
Jul
Annual Closure
users
Current : 1k
Average : 2k
Min : 344
Max : 2k
RRDTOOL / TOBI OETIKER


LHC Network

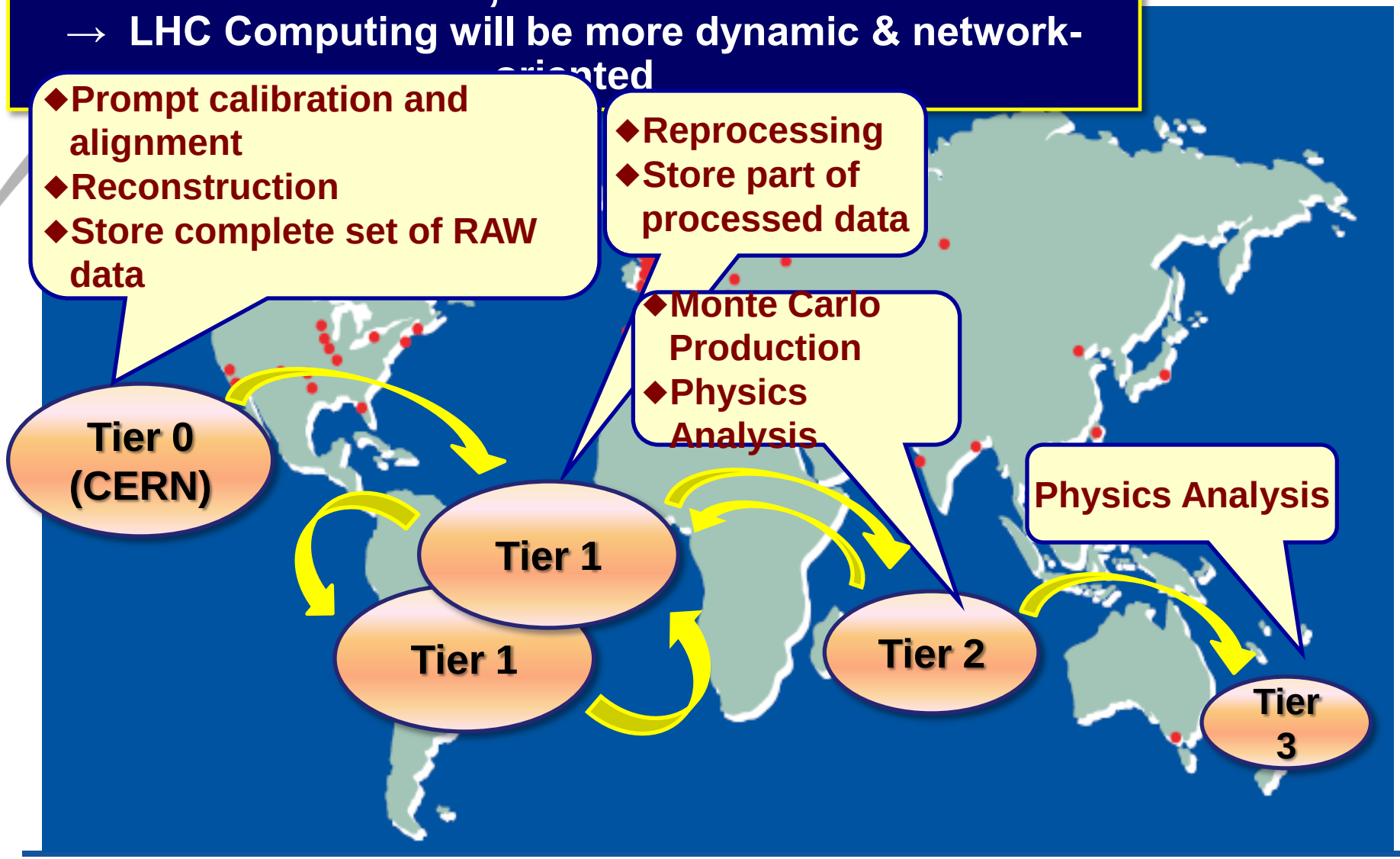

LCG
CMS
LHCb
ATLAS
Detector characteristics
Width: 44m
Diameter: 22m
Weight: 7000t
Muon Detectors
Electromagnetic Calorimeters
Forward Calorimeters
End Cap Toroid
Barrel Toroid
Inner Detector
Hadronic Calorimeters
Shielding
ALICE
130 Gbps
CERN
Computer
Center


T0/T1/T2 Interconnectivity
T2s and T1s are inter-connected by the general purpose research networks
Any Tier-2 may access data at any Tier-1
Dedicated 10 Gbit links
CERN
GridKa
IN2P3
TRIUMF
Brookhaven
ASCC
Nordic
Fermilab
CNAF
SARA
PIC
RAL
T2

The Roles of Tier Centers

11 Tier1s, over 100 Tier2s

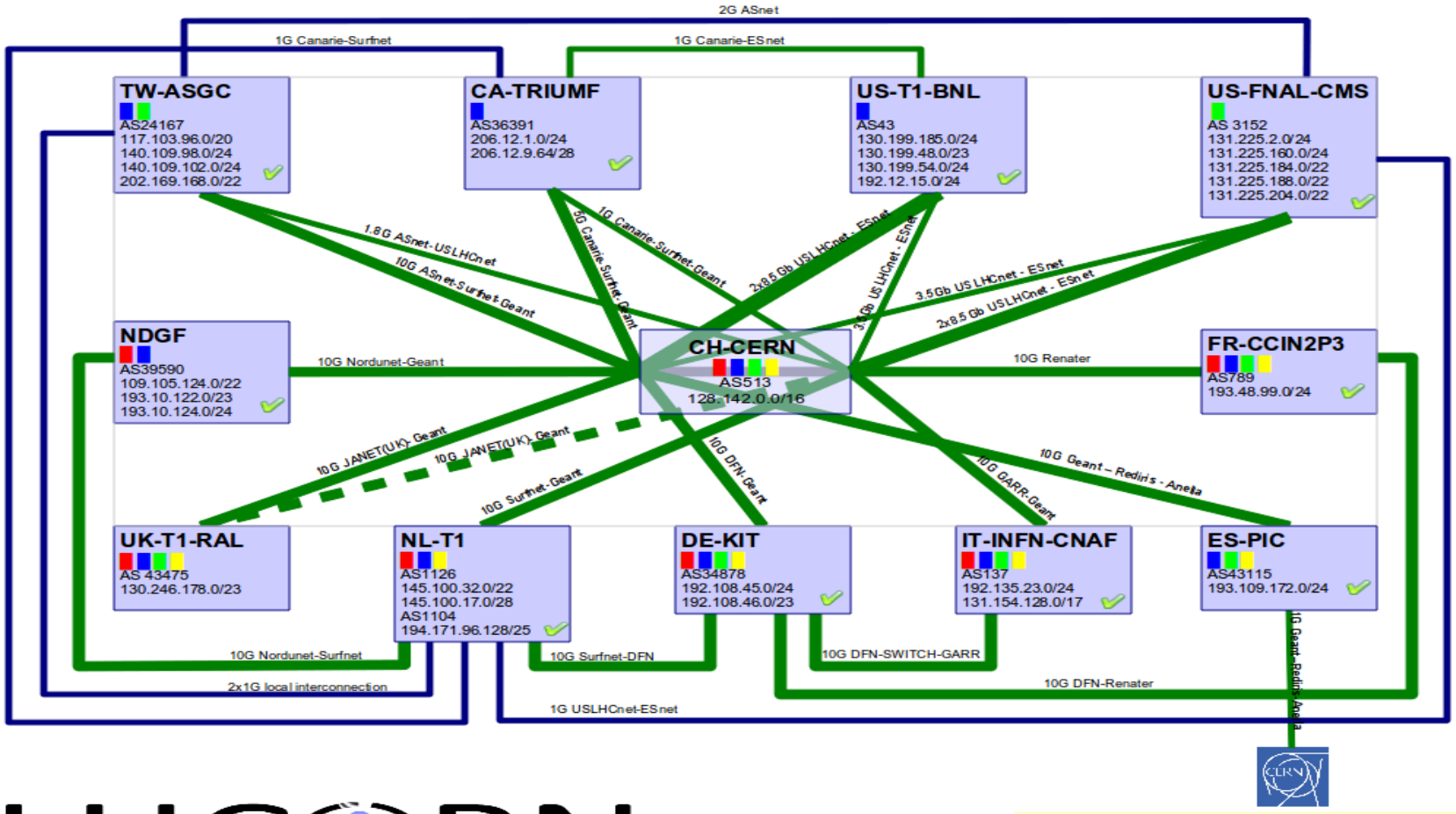
→ LHC Computing will be more dynamic & network-oriented



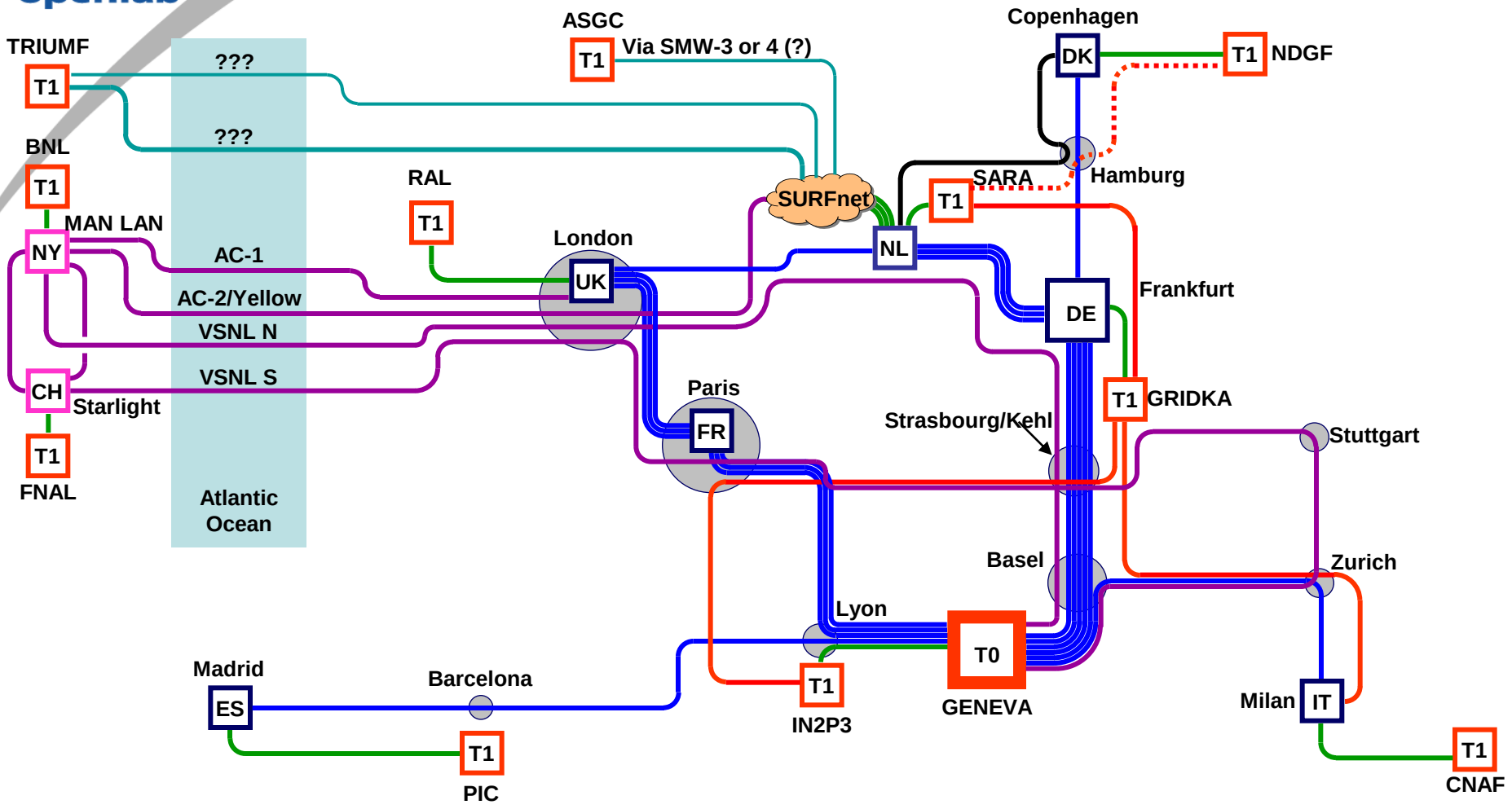
LHC **O**pen **P**rivate **N**etwork

- It consists of any T0-T1 or T1-T1 link which is dedicated to the transport of WLCG traffic
- It's based on 10Gbps light path technologies
- Infrastructure is provided by GEANT, US LHCNet, National Research and Education Networks (NREN) and Commercial links
- 70'000 km of optical paths

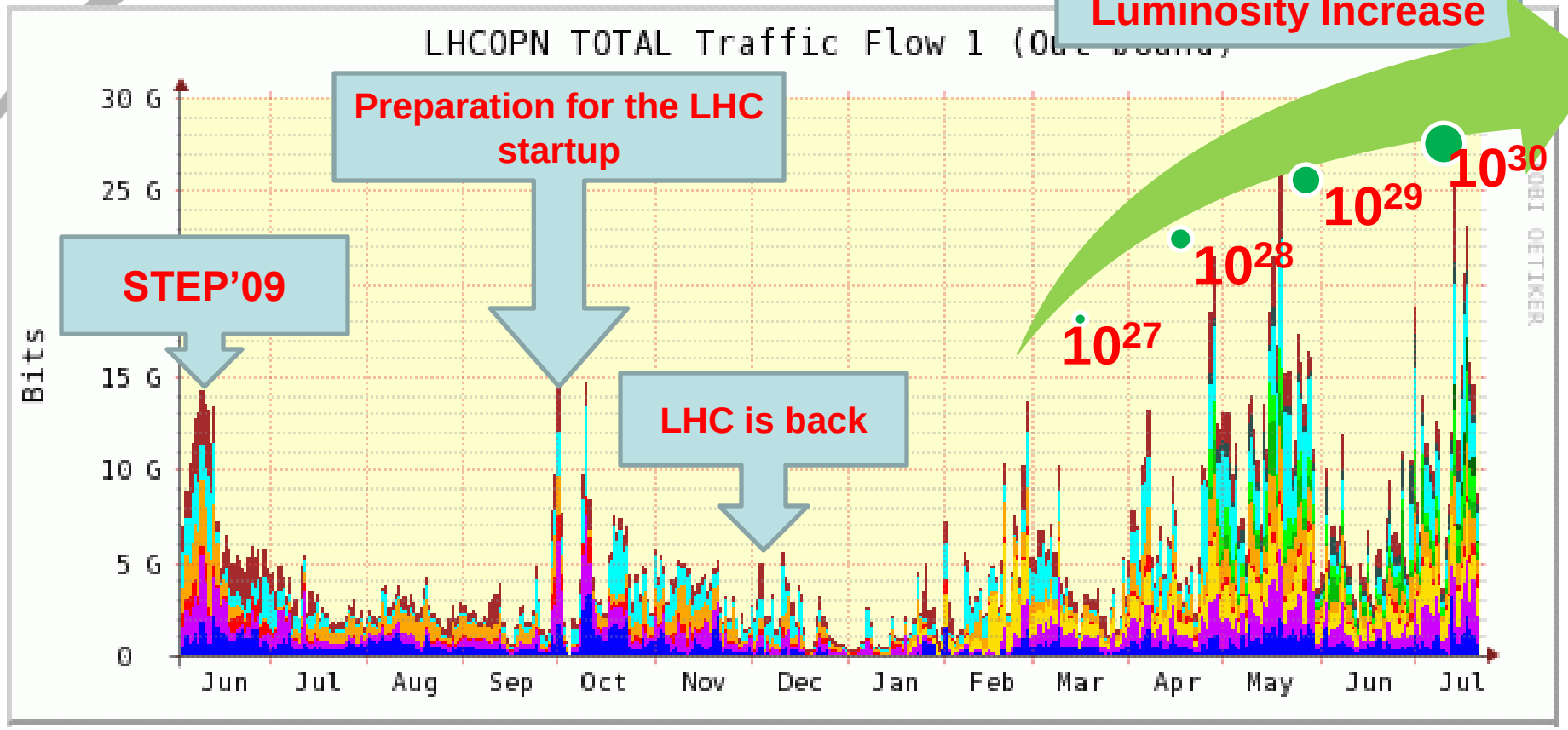
LHCOPN map


2G ASnet
1G Canarie-Surfnet
1G Canarie-ESnet
TW-ASGC
AS24167
117.103.96.0/20
140.109.98.0/24
140.109.102.0/24
202.169.168.0/22
CA-TRIUMF
AS36391
206.12.1.0/24
206.12.9.64/28
US-T1-BNL
AS43
130.199.185.0/24
130.199.48.0/23
130.199.54.0/24
192.12.15.0/24
US-FNAL-CMS
AS 3152
131.225.2.0/24
131.225.160.0/24
131.225.184.0/22
131.225.188.0/22
131.225.204.0/22
1.8G ASnet-USLHCnet
10G ASnet-Surfnet-Geant
1G Canarie-Surfnet-Geant
4G Canarie-Surfnet-Geant
2x8.5 Gb USLHCnet - ESnet
3.5Gb USLHCnet - ESnet
3.5Gb US LHCnet - ESnet
2x8.5 Gb USLHCnet - ESnet
NDGF
AS39590
109.105.124.0/22
193.10.122.0/23
193.10.124.0/24
10G Nordunet-Geant
CH-CERN
AS513
128.142.0.0/16
10G Renater
FR-CCIN2P3
AS789
193.48.99.0/24
10G JANET(UK)-Geant
10G JANET(UK)-Geant
10G Surfnet-Geant
10G DFN-Geant
10G GARR-Geant
10G Geant – Rediris - Anella
UK-T1-RAL
AS 43475
130.246.178.0/23
NL-T1
AS1126
145.100.32.0/22
145.100.17.0/28
AS1104
194.171.96.128/25
DE-KIT
AS34878
192.108.45.0/24
192.108.46.0/23
IT-INFN-CNAF
AS137
192.135.23.0/24
131.154.128.0/17
ES-PIC
AS43115
193.109.172.0/24
10G Nordunet-Surfnet
10G Surfnet-DFN
10G DFN-SWITCH-GARR
10G DFN-Renater
2x1G local interconnection
1G USLHCnet-ESnet
10G Geant – Rediris - Anella
T0-T1 and T1-T1 traffic
T1-T1 traffic only
Not deployed yet
(thick) >=10Gbps
(thin) <10Gbps
= Alice
= Atlas
= CMS
= LHCb
= internet backup available
p2p prefix: 192.16.166.0/24
edoardo.martelli@cern.ch 20100630

T0-T1 Lambda Routing



LHCOPN traffic



Network Competence Center Projects

CINBAD (2007-)

&

WIND (2010-)

CINBAD codename deciphered

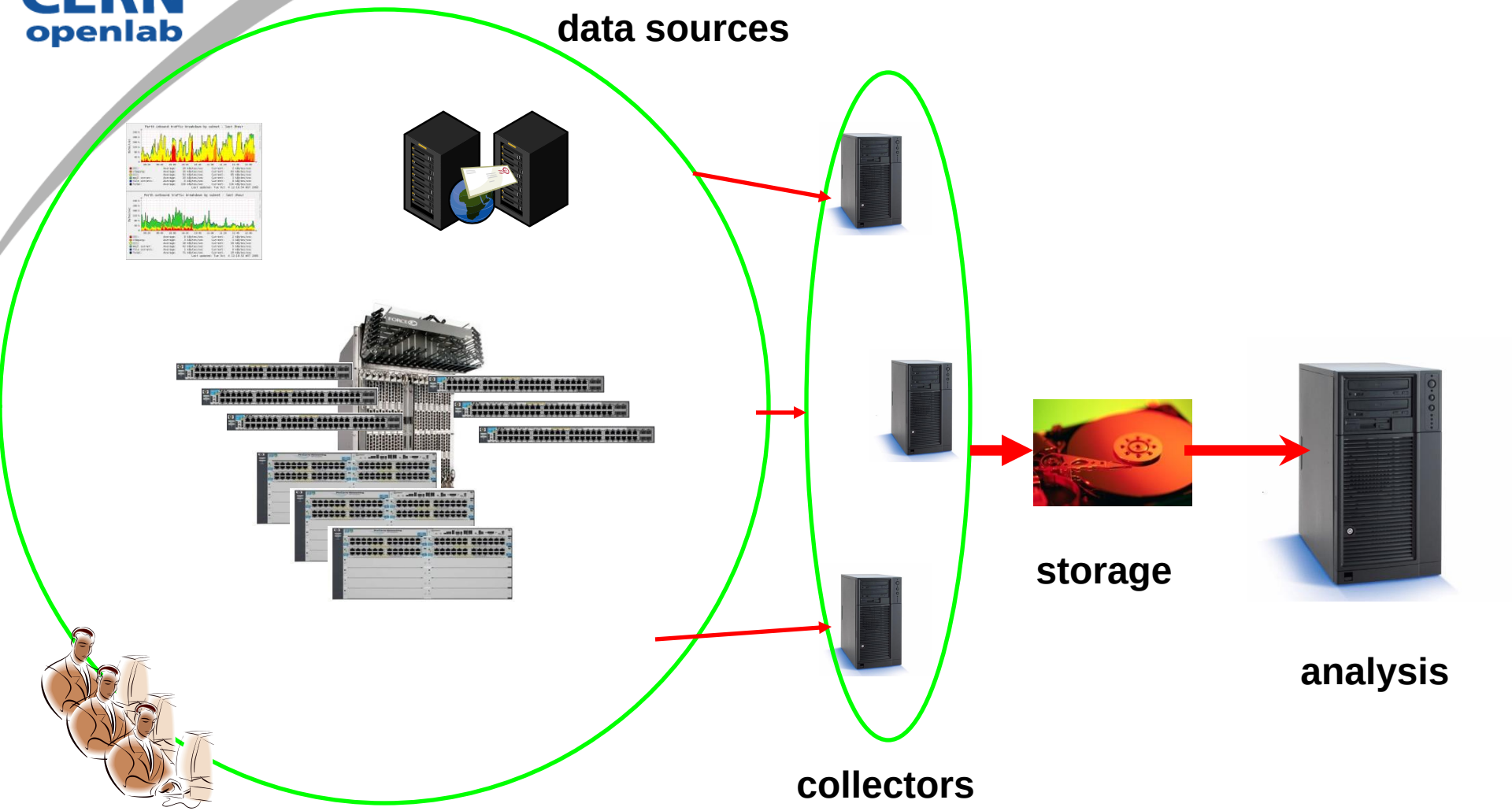
CERN **I**nvestigation of **N**etwork **B**ehaviour and **A**nomaly **D**etection

Project Goal

"To ***understand the behaviour of large computer networks*** *(10'000+ nodes) in High Performance Computing or large Campus installations to be able to:*

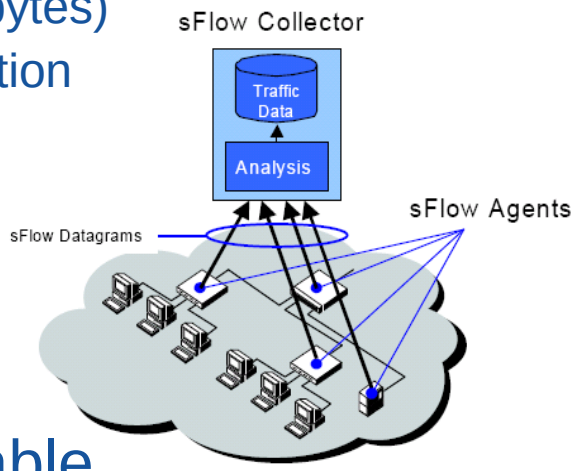
- *Detect traffic anomalies in the system*
- *Be able to perform trend analysis*
- *Automatically take counter measures*
- *Provide post-mortem analysis facilities* "

CINBAD project principle



sFlow – main data source

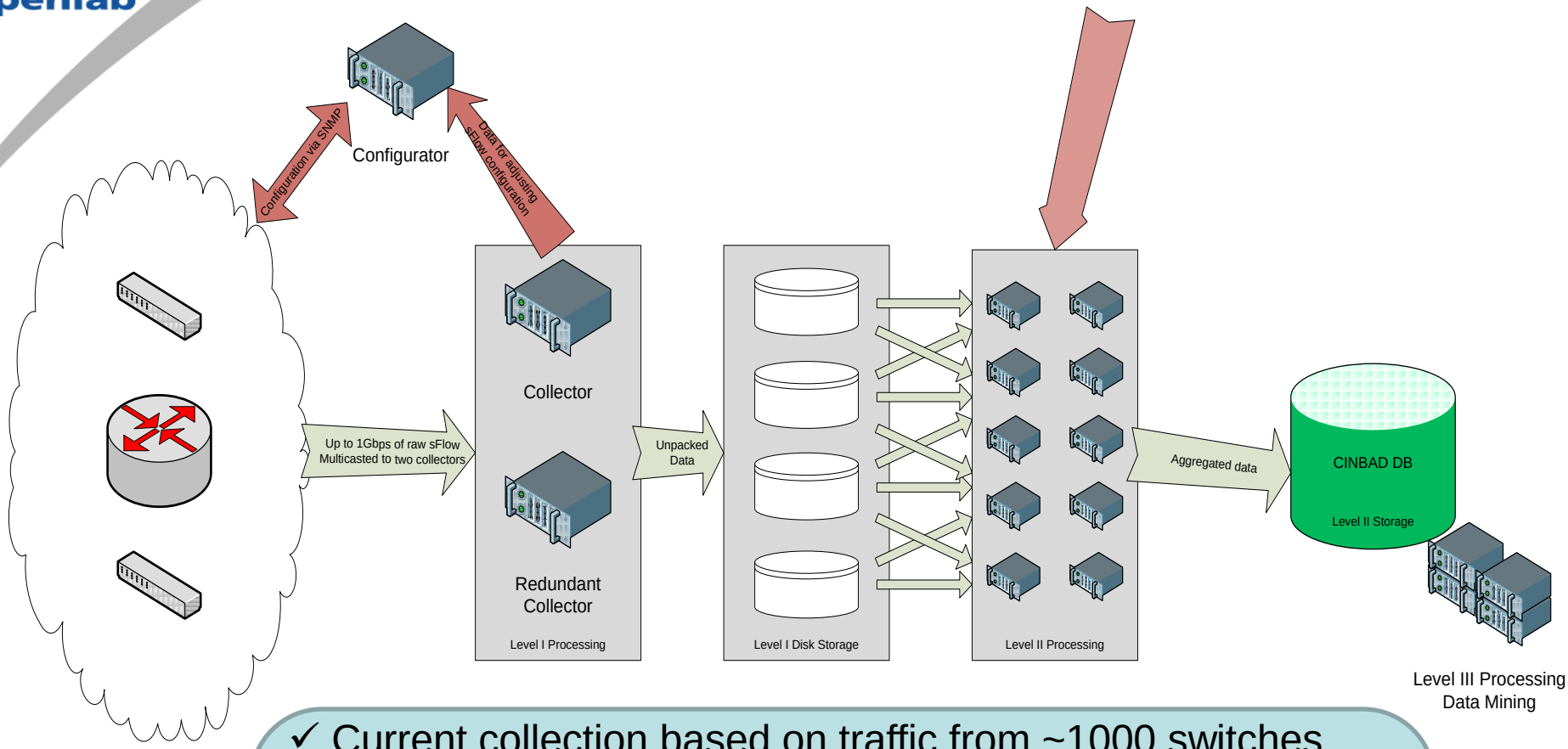
- Based on packet sampling (RFC 3176)
  - on average 1-out-of-N packet is sampled by an agent and sent to a collector
    - packet header and payload included (max 128 bytes)
      - switching/routing/transport protocol information
      - application protocol data (e.g. http, dns)
  - SNMP counters included
  - low CPU/memory requirements – scalable



- For more details, see our technical report

http://openlab-mu-internal.web.cern.ch/openlab-mu-internal/Documents/2_Technical_Documents/Technical_Reports/2007/RJ-MM_SamplingReport.pdf

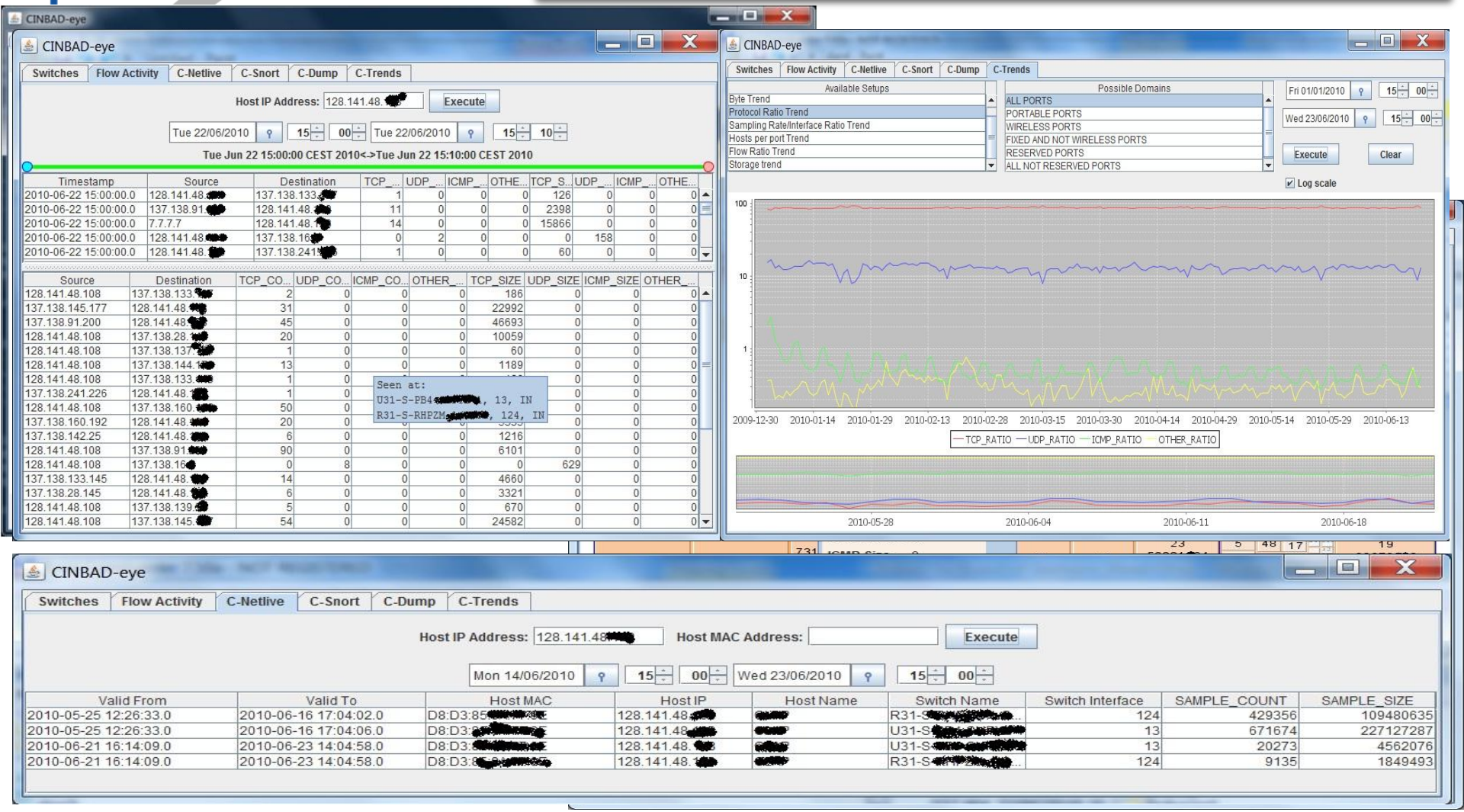
CINBAD sFlow data collection



- ✓ Current collection based on traffic from ~1000 switches
  - ✓ ~6000 sampled packets per second
  - ✓~ 3500 snmp counter sets per second
  - ✓~100GB per day

CINBAD-eye

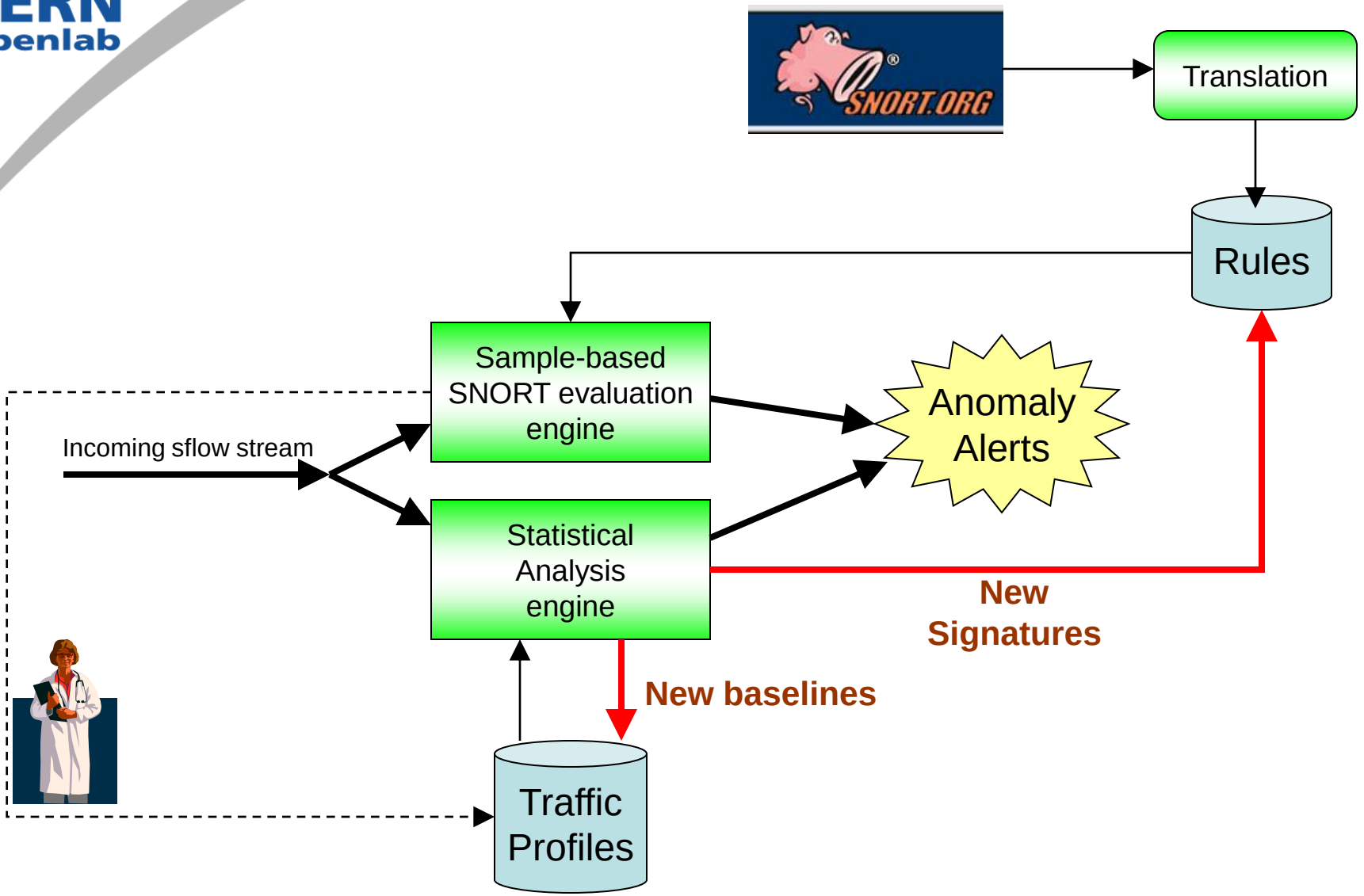
Host Activity and Connectivity, Trends



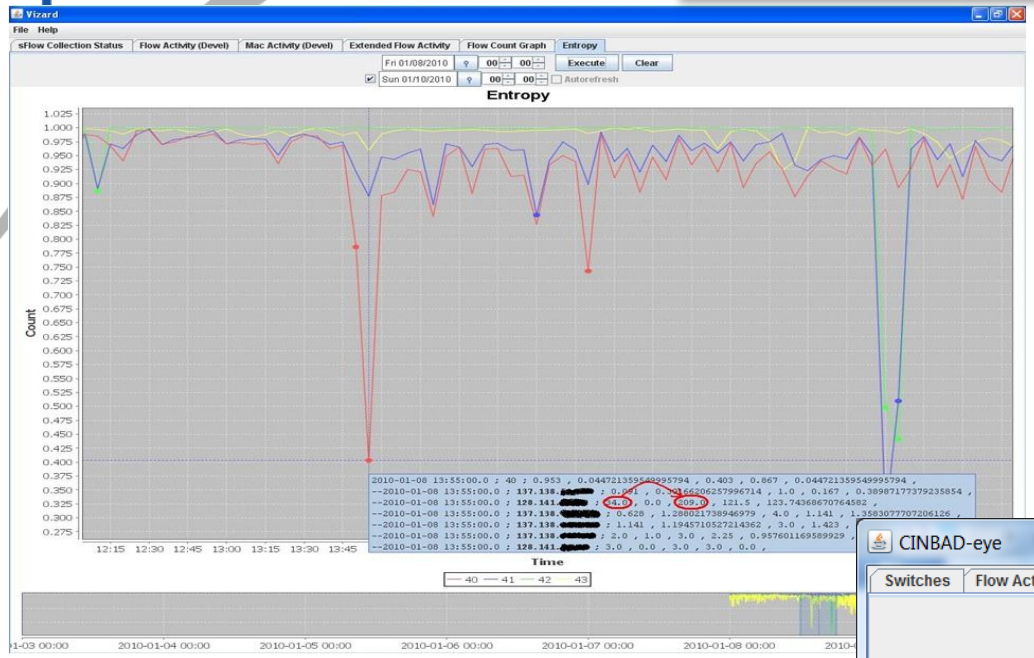
Tools for Anomaly Detection

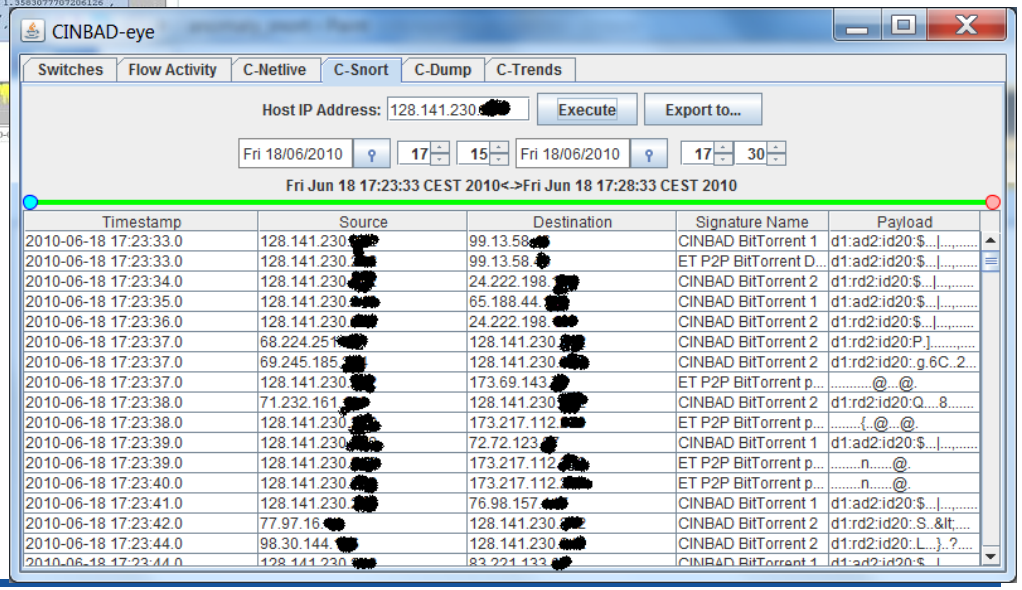
- Statistical analysis methods
  - detect a change from “normal network behavior”
    - selection of suitable metrics is needed
  - can detect new, unknown anomalies
  - poor anomaly type identification
- Signature based
  - we ported SNORT to work with sampled data
  - performs well against known problems
  - tends to have low false positive rate
  - does not work against unknown anomalies

Synergy from both detection techniques



Statistical and Signature-based Anomaly Detection





WIND

WIND codename deciphered

Wireless **I**nfrastructure **N**etwork **D**eployment

Project Goals

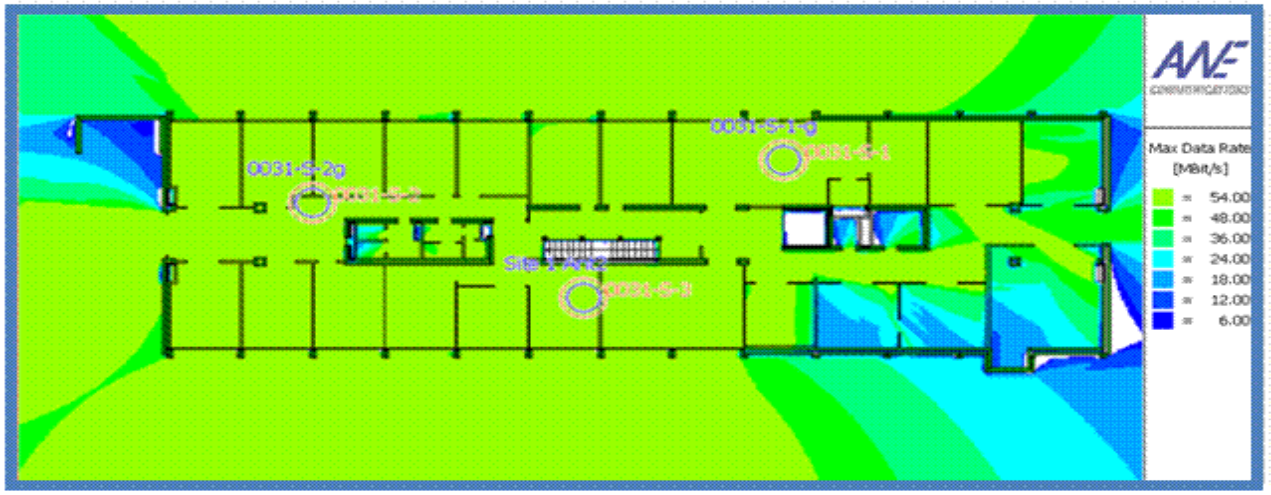
- **Analyze** the problems of large scale wireless deployments and understand the constraints
- **Simulate** behaviour of WLAN
- **Develop** new optimisation algorithms
- **Verify** them in the real world
- **Improve** and **refine** the algorithms
- **Deliver** : algorithms, guidelines, solutions

WLAN deployments are problematic

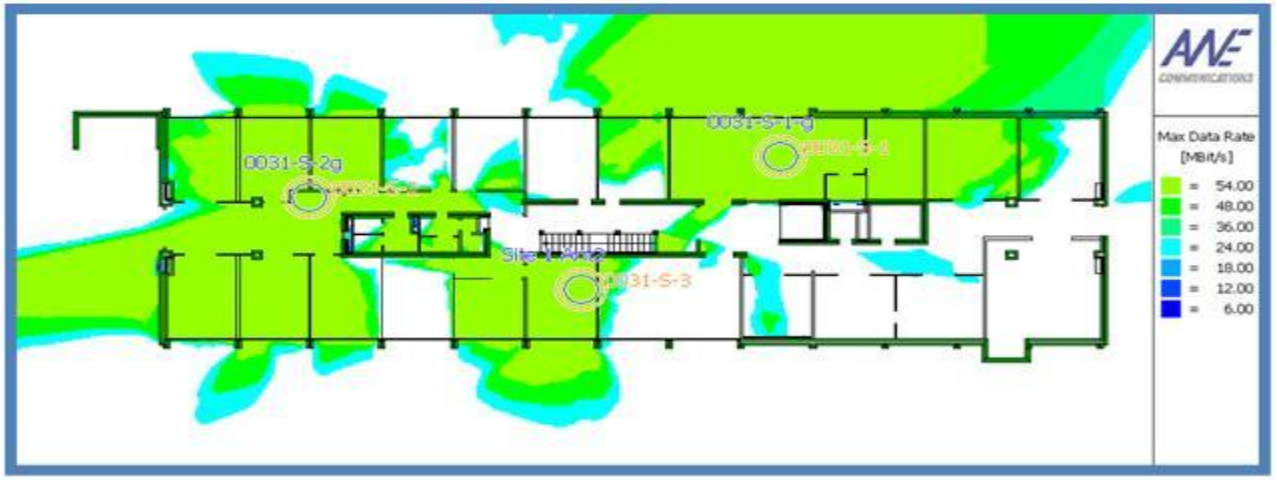
- Radio propagation is very difficult to predict
- Interference an ever present danger
- WLANs difficult to properly deploy
- Monitoring was not an issue when the first standards were developed
- When administrators are struggling just to operate the WLAN, performance optimisation is often forgotten

Problem Example – RF interference

Max data rate in 0031-S: The APs work on 3 independent channels



Max data rate in 0031-S: The APs work on the same channel



Acknowledgments

IT/CS

Thank you!