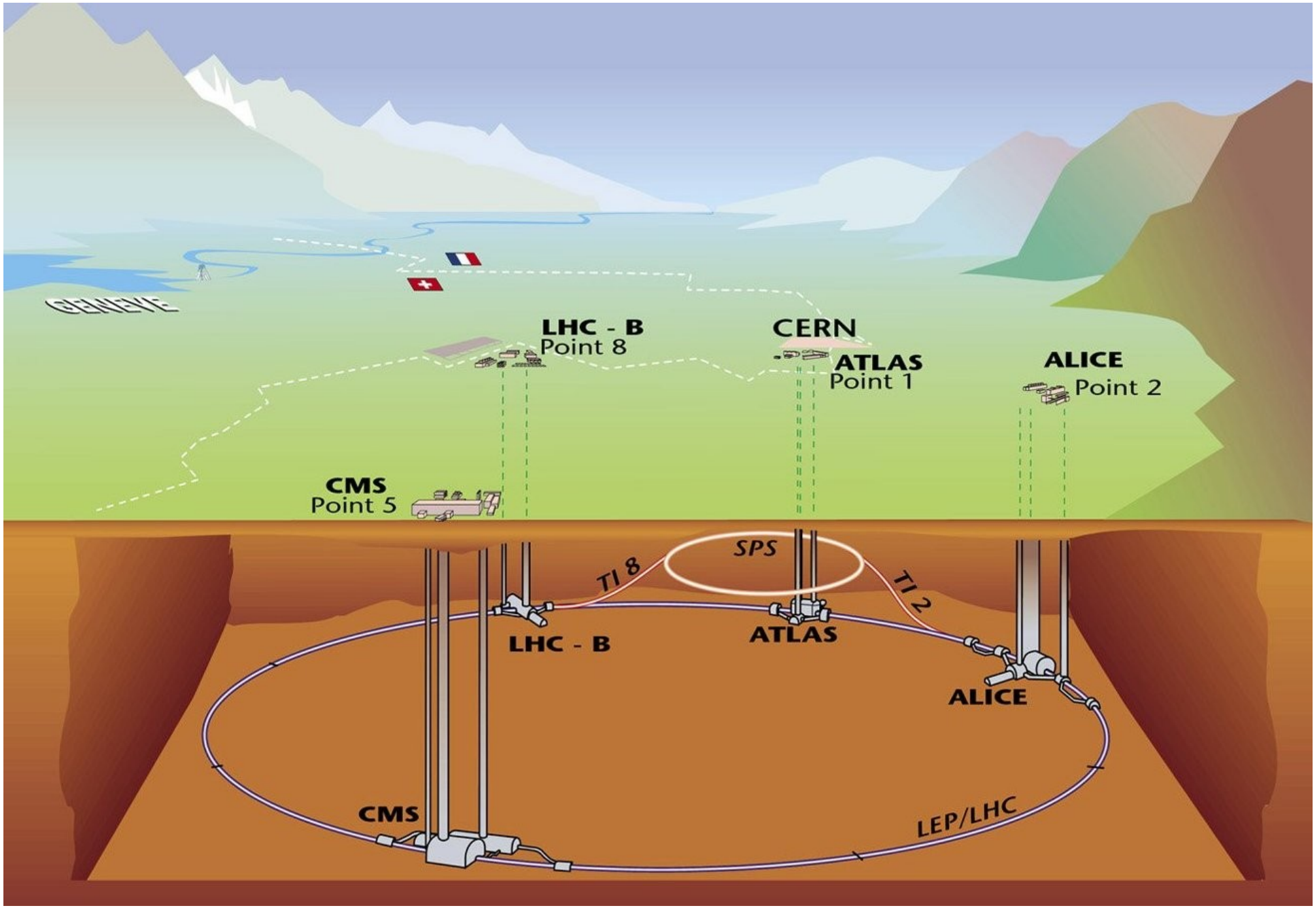
LHC network architecture

Geant2 technical workshop

Cambridge 9 January 2006

edoardo.martelli@cern.ch

The LHC


GENEVE
LHC - B
Point 8
CERN
ATLAS
Point 1
ALICE
Point 2
CMS
Point 5
SPS
TI 8
TI 2
LHC - B
ATLAS
ALICE
CMS
LEP/LHC

LHC and LCG

The **Large Hadron Collider (LHC)**, is the largest scientific instrument on the planet. It will produce roughly 15 Petabytes (15 million Gigabytes) of data annually, which thousands of scientists around the world will access and analyse.

The **LHC Computing Grid (LCG)** is the computer infrastructure that will provide data storage and analysis capabilities to the entire high energy physics community that will use the LHC.

LCG four-tiered model

The data from the LHC experiments will be distributed around the globe, according to a four-tiered model. A primary backup will be recorded on tape at CERN, the **Tier-0** centre of LCG. After initial processing, this data will be distributed to a series of **Tier-1** centres, large computer centres with sufficient storage capacity for a large fraction of the data.

The Tier-1 centres will make data available to **Tier-2** centres, each consisting of one or several collaborating computing facilities, which can store sufficient data and provide adequate computing power for specific analysis tasks. Individual scientists will access these facilities through **Tier-3** computing resources, which can consist of local clusters in a University Department or even individual PCs, and which may be allocated to LCG on a regular basis.

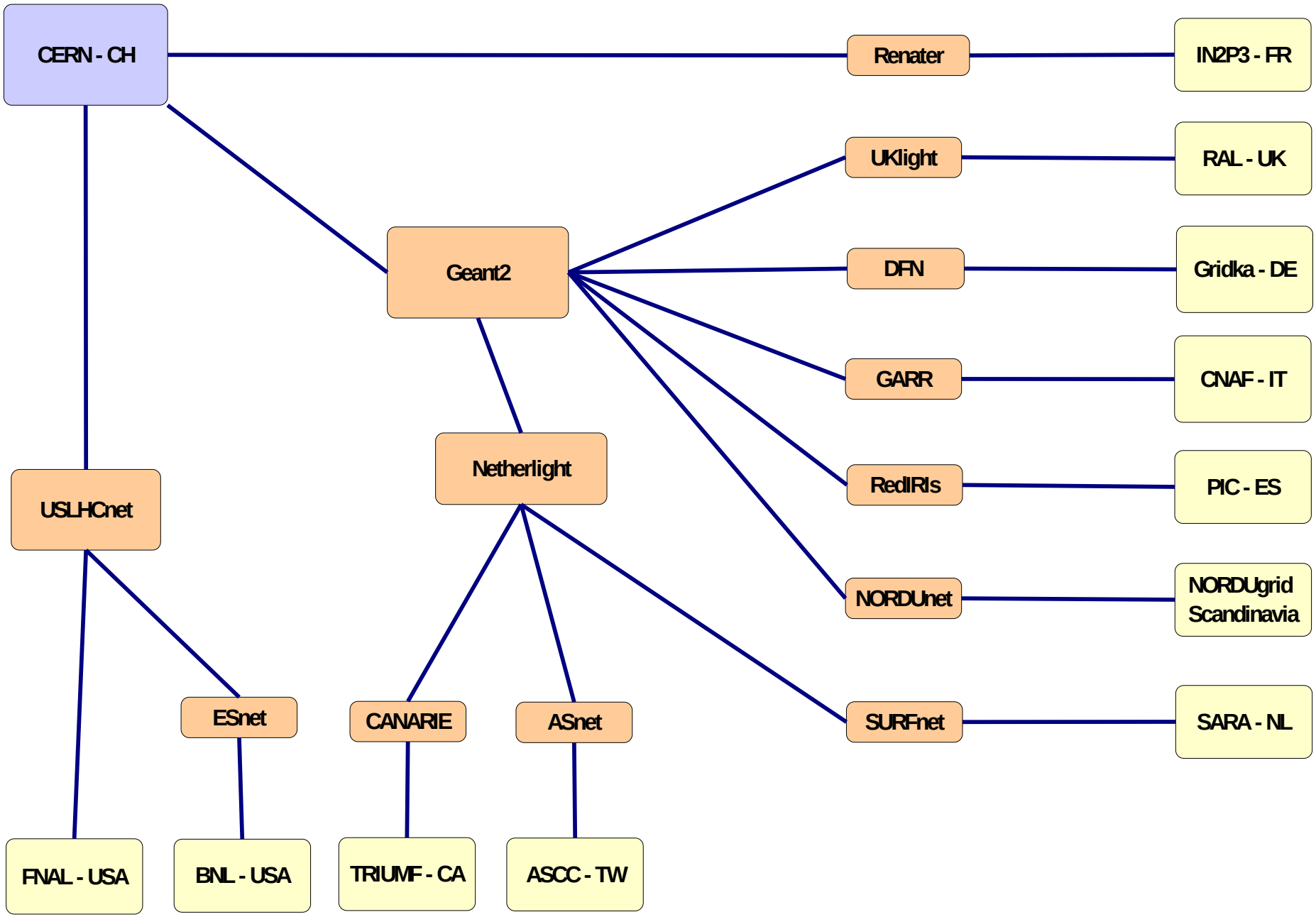
LHCOPN – LHC Optical Private Network

The **LHCOPN** is the network that will connect the Tier0 to all the Tier1s.

To cope with the amount of data produced by the Tier0, every Tier1 will be connected through a dedicated 10 Gbps link, plus a backup path of the same bandwidth.

The use of the LHCOPN will be restricted to the members, namely the Tier0 and the eleven Tier1s

LHCOPN


CERN - CH
Renater
IN2P3 - FR
UKlight
RAL - UK
Geant2
DFN
Gridka - DE
GARR
CNAF - IT
Netherlight
RedIRIs
PIC - ES
USLHCnet
NORDUnet
NORDUgrid Scandinavia
ESnet
CANARIE
ASnet
SURFnet
SARA - NL
FNAL - USA
BNL - USA
TRIUMF - CA
ASCC - TW
transit network
Tier0
Tier1

Connectivity

Layer 1 – 2 architecture

Every Tier1 will be connected to the Tier0 with a direct 10Gbps link (*lightpath).*

Those lightpaths will be of different kinds:

- single or concatenated layer 1 links (STM64, LANPHY, WANPHY)
- layer 2 VLANs

Tier1s should also provide a dedicated backup link to Tier0, but during the startup phase (Service challenges) backup will be provided via routed paths (GN2, Esnet, Abilene, NRENs..).

Tier1 to Tier1 links will be established between pairs that need to exchange large amount of traffic. However transit via the Tier0 can be provided.

Layer3 architecture

IPv4 will be used as layer 3 network protocol. IPv6 can be used in the future (not supported everywhere yet).

Public IP addresses must be used at every Tier (especially needed for backup and Tier2 reachability). Prefixes allowed into the LHCOPN are called **LHC prefixes** and the list of them is kept updated in the RIPE whois database (route-set RS-LHCOPN).

BGP will be used to manage the routing. Peerings will be established between each Tier1 and the Tier0. Peerings can also be established between Tier1s connected by a direct link.

BGP guidelines

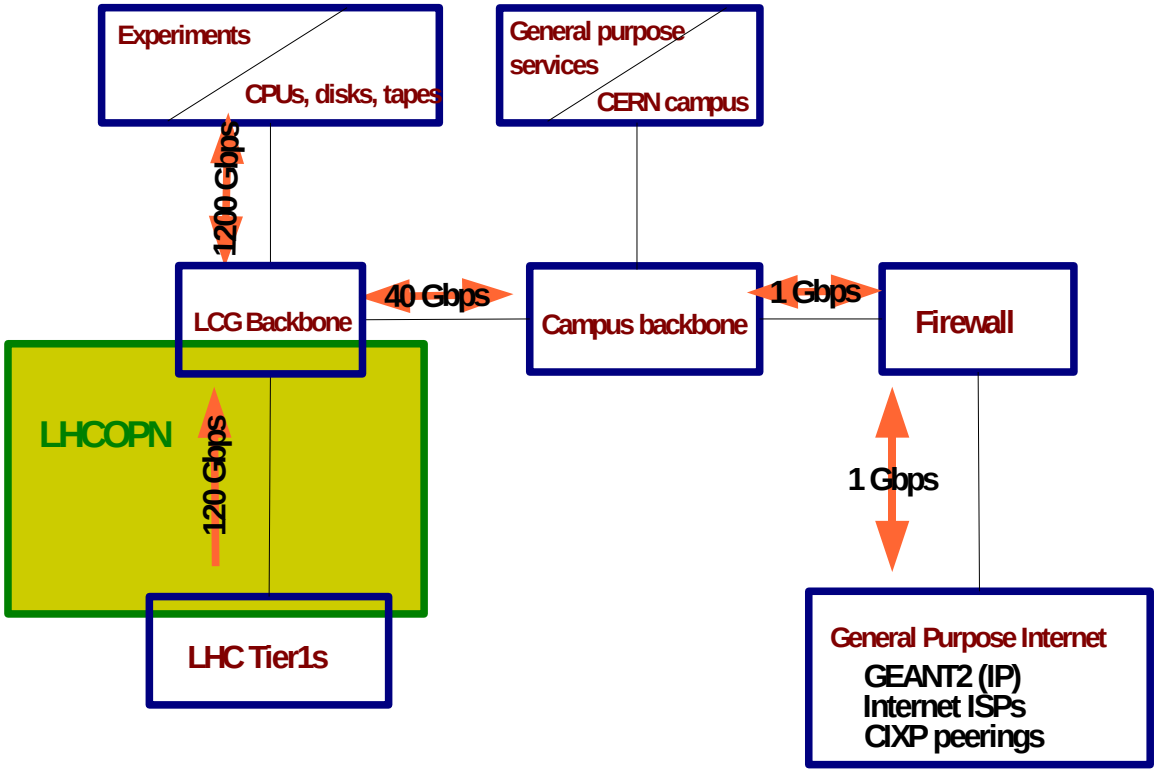
- Every Tier must use the AS number to which its LHC prefixes belong to.
- Every T1 must announce its own LHC prefixes to T0.
- T0 accepts all the LHC prefixes from every Tier1s (for transit over T1-T1 links).
- T0 announces its own LHC prefixes to every peering T1.
- T0 re-announces all the LHC prefixes received to all the T1s (for transit via T0).
- T1 accepts T0's prefixes, plus any other necessary prefix (to achieve transit).
- T0 and T1s should announce their LHC prefixes to their upstream continental research networks (NRENs, GÉANT2, Abilene, ESnet..) in order to provide backup and allow connectivity towards the T2s. Special care must be taken by each Tier to not leak out BGP prefixes that belong to other Tiers.
- Usage of static routes is not advisable.
- No default route must be used in T1-T0 routing.

Security

Security concern

In most cases the LHCOPN will bypass the security system (firewall, IDS...) already in place at every Tier; that's because the current technologies cannot deal with the requested bandwidth.

CERN example:



Security

The use of the LHCOPN will be limited to the LCG traffic: only packets with source and destination addresses comprised in the LHC prefixes can travel along the network.

If possible, traffic will be further filtered at TCP level allowing only packets with source and destination ports that belong to any recognized LCG applications.

Security - means

Every site is responsible to implement a level of security that is compatible with the local policies.
Every site must implement an adequate level of protection in order to increase the overall security.

This will be basically implemented on the border routers by mean of:
- Access Control Lists for filtering packets on IP and (possibly) TCP values.
- BGP filters for announcing and accepting only LHC prefixes.
- Avoid of default routes pointing to the LHCOPN

Operations

Operations

The LHCOPN needs to be monitored and operated continuously in order to provide an effective and reliable service.

All the members agreed that the normal way of resolving problems between neighbouring networks won't be enough, but a central co-ordination entity will be needed.

Two models of central co-ordination were discussed:

Helpdesk model: monitors the network and contacts NRENs for fault resolution

NOC model: in addition to help-desk, it manages the LHCOPN

Operations – Help desk

The help desk has RO access to T0 and T1 elements, either directly or via a proxy

It monitors the LHCOPN for faults

It is a single point of contact for problem notification, configuration change requests (carried out by T0 or T1s),

It contacts owners of networking elements (T0, T1s, NRENs, GEANT2) for fault resolution and change management.

Issues TTs

Operations – NOC

The NOC is an extension of help desk: in addition it carries out configuration of routers and switches forming the LHC-OPN.

Needs separation of LHC-OPN equipment from internal equipment at T0 and T1s

Operations – The EGEE work

The EGEE project has foreseen and produced the specifications for a user support centre; it has been called **GGUS (**Global Grid User Support).

The **ENOC** is Network Co-ordination Service. It can

- look after network issues for EGEE and LCG

- receive network TTS from NRENs, analyse them and provide relevant information to the GGUS who will then interact with the users.

- monitor the e2e status of the lightpaths and trigger the appropriate corrective actions

This work was done in co-operation with GN2 partners.

Operations – ENOC

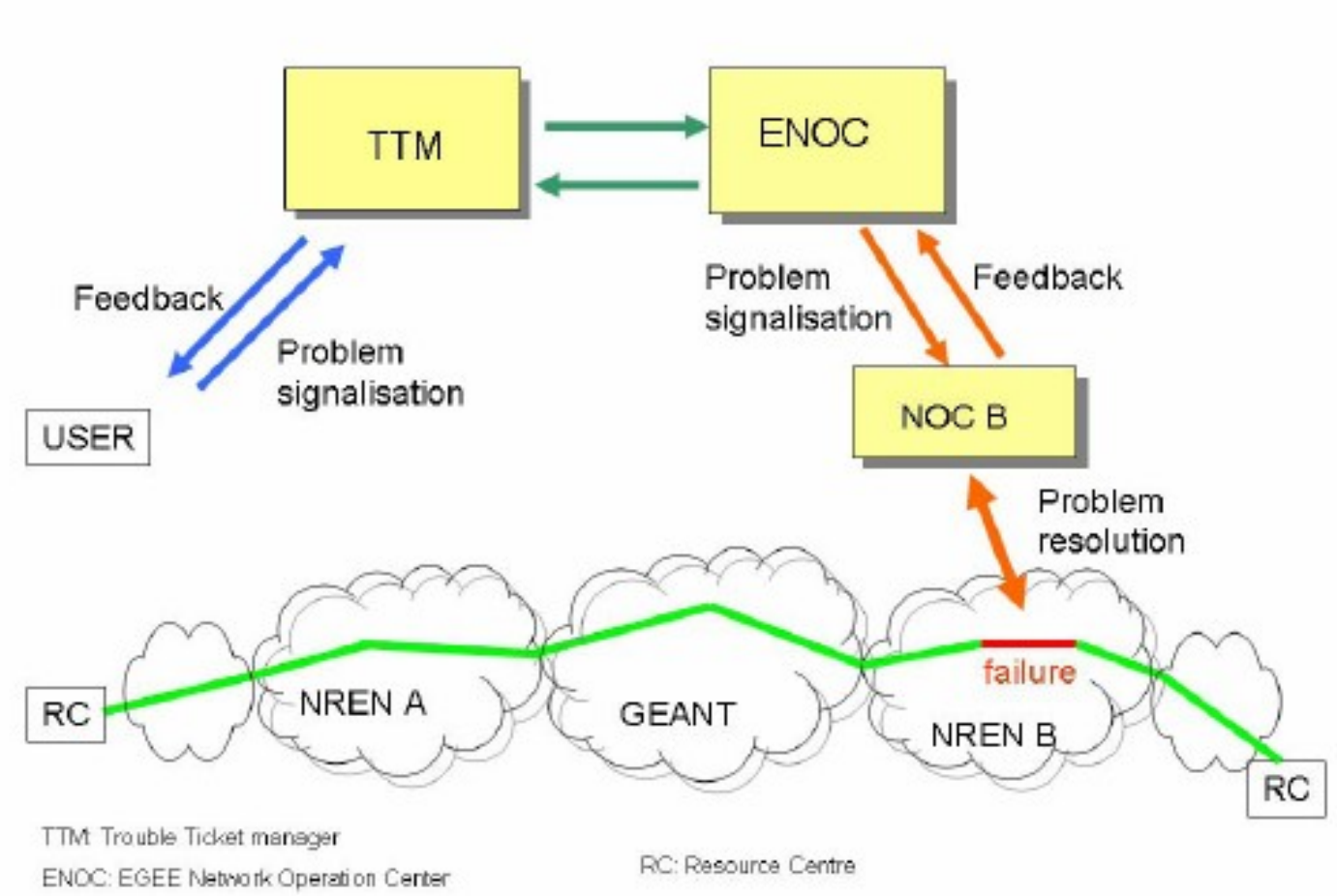


Figure 2. Summary of interactions between involved entities

Work in progress

Under construction