19.09.06

Deliverable DS2.0.3: Report on Use of GÉANT2 Advanced Network Services - Lambdas and Switched Optical

Deliverable DS2.0.3

Contractual Date:	30/08/06
Actual Date:	19/09/06
Contract Number:	511082
Instrument type:	Integrated Infrastructure Initiative (I3)
Activity:	SA2
Work Item:	N/A
Nature of Deliverable:	R (Report)
Dissemination Level	PU (Public)
Lead Partner	DANTE
Document Code	GN2-06-223v5

Authors: R. Sabatino (DANTE), J. Chevers (DANTE), M. Karapandzic (DANTE, Editor)

Abstract

This document provides a description of the progress of implementation of point-to-point (p2p) dedicated high-speed circuits, namely the lambdas and switched optical services, within the footprint of the GÉANT2 network.

Table of Contents

Table of Figures

0 Executive Summary

With the GÉANT2 network being completed at the time of writing, the use of its p2p service portfolio continues to increase. The service itself is still in the beginning stages of development; the process of ordering and installation of p2p links is being developed.

The current users of p2p services include the Large Hadron Collider (LHC) in CERN, DEISA and others. At the time of writing, eleven requests have been received to implement p2p services on the GÉANT2 network; all these requests are being processed.

In parallel, discussions have been held as to the management of such links in an operational environment, namely the establishment of a specialised NOC and monitoring functions. In the former case, it was decided that the project will establish an End to End Coordination Unit (E2ECU), to act as a single point of contact for all parties involved in the implementation of p2p links; the E2ECU is expected to be operational by January 2007.

Activities JRA4 and JRA1 have conducted work to provide the monitoring of p2p 10Gbps links, with only one metric to be monitored in the initial stages – up/down status. This multi-domain, multi-technology challenge is being addressed by JRA1 and its perfSONAR monitoring framework, which aims to work on establishing a uniform way of collecting data across several domains. JRA4 provides a tool that will query the perfSONAR infrastructure, retrieve status information and provide a view of the link.

It is expected that both current and new users, with 'bandwidth-hungry' applications, on both sides of the Atlantic, will require a significant number of dedicated optical circuits to conduct their research in the future. DANTE continues to conduct regular surveys of NRENs' p2p requirements and thus provide forecasts of future demand.

1 Introduction

The GÉANT network offered a best efforts IP service, including multi-cast, to the connected NRENs. There were some refinements on this, particularly in the area of IP Quality of Service and MPLS tunnels. Initially, there was no formal service definition provided in such an environment. Obligations to operate and maintain the service at stringent levels of reliability and quality were specified and enforced by DANTE, on behalf of the consortium, through its contractual arrangements with commercial providers.

In contrast to GÉANT, GÉANT2 offers a greater range of services where, in addition to an IP service, various point-to-point services, which guarantee elements of network capacity between individual NRENs, are available. The routine provision of such point-to-point services is a new operational and technical experience, not least because more of the underlying infrastructure is directly managed by DANTE and the NRENs than has been the case ever before. As such, service parameters introduced here are what now seem to be reasonable goals and will be subject to revision in the light of experience.

The geographic availability of the service portfolio is precisely defined in respect of the dark fibre footprint of GÉANT2. All NRENs on this footprint are capable of accessing both IP and point-to-point service via the GÉANT+ subscription. In addition, there is the possibility of acquiring additional point-to-point access capacity according to a pre-defined cost structure.

The availability of the service portfolio off the fibre cloud is determined by the access subscription of the individual NREN and the usage of that access. There is no obligation to subscribe to point-to-point services. For wavelength based point-to-point services, this will require an additional subscription. For sub-wavelength services, principally GE services, a number of potential technical options are available, subject to the existence of sufficient access capacity.

This document describes the p2p services on offer and provides a short summary of their implementation so far. Two issues relevant to the services are discussed, namely the End to End Coordination Unit (E2ECU; a 'NOC' for p2p circuits) and the monitoring of p2p links, both crucial for the success of the service.

2 Definition of Advanced (p2p) Services

GÉANT2 offers the following point-to-point (p2p) services:

- Ethernet services (Ethernet Private Line – EPL), of which there are two varieties:
    - Gigabit Ethernet services (GE services)
    - 10 Gigabit Ethernet services (10 GE services)
- SDH/SONET services

To see further details of the possible scenarios for the provision of p2p services, please refer to Appendix A

2.1 Point-to-point Ethernet (Ethernet Private Line) Services

2.1.1 Point-to-point Gigabit Ethernet Service

Gigabit Ethernet services are supported via the SDH/SONET switch (Alcatel 1678 MCC). There are three ways for an NREN to access this service:

- Connecting to a GE port on the switch. The line card will use the Generic Framing Procedure (GFP) to map the Ethernet frames onto SDH.
- Connecting to a 10GE port on the switch. This requires the use of VLANs on the 10GE link to distinguish between the different GE services that are carried over a single physical link. From that perspective the network is not VLAN transparent anymore since the network interprets the VLAN tags. The line card in the switch will use the Generic Framing Procedure (GFP) to map the Ethernet frames onto SDH.

- Connecting to an SDH port on the switch. This requires the NREN to use GFP encapsulation.

In the first instance, it is envisaged that only full-rate gigabit Ethernet services will be offered although, in principle, sub-1G services can be provided.

The switches have the capability to provision the link capacity in the network between the two Ethernet ports with a granularity of VC-4 (~150Mbps). Seven VC-4s (VC-4-7v) are required to fully provision the capacity for full-rate gigabit Ethernet.

The GÉANT2 network will only offer unprotected services. To increase the service availability it will be possible to request for two diversely routed connections. Failure detection and switchover decisions are then the responsibility of the endpoints. This approach will incur the double amount of cost since the double amount of capacity is provisioned.

2.1.2 Point-to-point 10 Gigabit Ethernet Service

2.1.2.1 *10 Gigabit Ethernet LAN PHY Service via Switch*

The NREN can access the service via a 10GE port on the SDH switch. Again, the service provisioning relies on GFP and virtual concatenation. Alternatively, the service can be accessed via an STM-64 port provided that GFP encapsulation is used.

The provisioned capacity in the network can be between 1 and 64 VC-4s. Indeed, even when 64 VC-4s are provisioned there is still a rate mismatch with the rate of 10GE LAN PHY.

The same protection and restoration options apply as for the Gigabit Ethernet (GE) service.

2.1.2.2 *10 Gigabit Ethernet Service over G.709 wavelength*

Another option to implement 10GE services is to use the transmission equipment only. This option will be available from the start of the GÉANT2 network. This service is classified separately from the '10GE via the switch' because some characteristics are different such as:

- No MTU limit for the Ethernet frames
- No rate mismatch between the access port and the capacity provisioned in the network.
- No PAUSE frames are supported. (It is also not required since there is no rate mismatch between the access port and the network.) In actual fact, the service is transparent to PAUSE frames.
- VLAN transparent, i.e. VLAN tags are not interpreted within the GÉANT2 network.

The NREN connects directly to a 10GE transponder of the DWDM system. In the intermediate POPs the transponders of the DWDM terminals have to be connected back to back by manually installing fibre patch cords. This obviously increases the provisioning time of the service. It is suitable for (semi-)permanent connections.

The transmission layer does not offer any protection or restoration mechanisms. Hence, the same protection and restoration options apply as for the gigabit Ethernet service.

2.2 SDH Services

2.2.1 SDH Service via switch (Alcatel 1678 MCC)

To access SDH services the NREN has to connect to an SDH (STM-16/64) port on the switch. The service is provisioned using the Network Management System (NMS) and as a result the provisioning time can be very small.

Sub rate services can be offered ranging between STM-1 and STM-64. In case the NREN sets up multiple SDH services (trails) via its access interface there is a need to configure the SDH cross connects accordingly in the GÉANT2 network. In other words, the network is terminating the SDH multiplex section.

In case the NREN does not have a very accurate timing source in its network it needs to derive the timing from the network. This is for example the case when routers are connected at the ends of the service.

The same protection and restoration options apply as for the gigabit Ethernet service.

2.2.2 SDH service via the transmission system

SDH services can also be supported via the DWDM transmission system. In that case only a 10G port is available for the NREN to connect to. It is essentially a conventional unprotected transparent 10G wavelength service. It is fundamentally different from the ‘SDH service via switch’ in the sense that the service is fully transparent, i.e. the multiplex section is not terminated, and the timing is derived from the NREN.

An overview of the properties of an SDH wavelength service:

- No protection or restoration.
- Full transparency, the multiplex section is not terminated in the network. As a result, NRENs can setup several SDH trails over the 10G wavelength without notifying the GÉANT2 network operator.
- The NREN equipment has to provide the timing (one end providing, the other end receiving).

- The provisioning time will be longer than for switched services due to the manual fibre patching that is required.

The same protection and restoration options apply as for the GE service.

3 Progress of Implementation

A detailed procedure has been developed to handle the provisioning of point-to-point services over GÉANT2. Since any such request must involve a minimum of five separate parties (two user groups, two NRENs, GÉANT2) it is clear that a chain of responsibility must be established. The process was carefully constructed to ensure financial, technical and resource commitments from all parties are transparent and that the status of a request can be easily tracked through its lifecycle.

The document which initiates the point-to-point process at DANTE is the *Service Request Form* (available at http://intranet.geant2.net/server/show/nav.922) which is completed by one of the connecting NRENs with the full knowledge of the counterpart connecting NREN. This form provides sufficient detail for operational, network planning and financial functions of DANTE to make decisions on the technical feasibility, timescale and cost of the requested circuit. This information is then related to the NRENs concerned in the *Service Confirmation Form.* Upon signature by both connecting networks, this document constitutes the formal order for the point-to-point circuit.

To date, 11 separate point-to point service requests have been received, many for full 10G wavelength services to the CERN particle physics site near Geneva. A summary of the current status of point-to-point requests is given in Table 3.1 below.

Date rec'd	Site A	Site B	Service Type	NREN A	NREN B	Service implemented
11/05/2006	SARA	CERN	10G SDH	SURFnet	CERN	
23/05/2006	SURFnet	CERN	10G SDH	SURFnet	CERN	Due 10/06
23/05/2006	SURFnet	CERN	10G SDH	SURFnet	CERN	Due 10/06
06/06/2006	Fyzikalni ustav AV CR Na Slovance 2, Praha	Karlsruhe	1 GE	CESNET	DFN	01/09/06
06/06/2006	CESNET	Netherlight	10G SDH	CESNET	SURFnet	Due 10/06
09/06/2006	Karlsruhe	CERN	10GE	DFN	CERN	
N/A	INFN Bologna	CERN	10GE	GARR	CERN	30/10/06
29/06/06	RAL	CERN	10GE	UKERNA	CERN	01/09/06
28/08/06	Barcelona SC	Juelich	10GE	RedIRIS	DFN	Due 11/06
29/08/06	SARA	Juelich	10G M/C	SURFnet	DFN	Due 11/06
04/09/06	IN2P3	Fermilab	2 x 1GE	RENATER	I2/ESNET	Due 10/06

Table 3.1: Summary of GÉANT2 p2p requests

Where possible, these services have been implemented using equipment in place throughout the GÉANT2 network. In other cases equipment has been promptly ordered upon receipt of the signed service confirmation forms. In these cases equipment delivery times dictate the implementation date of the circuit.

4 Network Operations for Advanced Services

The general structure of a p2p link from an endpoint (E1) to another endpoint (E2) is shown in Figure 4.1, where the domains indicate the operational responsibility for the optical equipment. Each domain is delimited by a well defined number of demarcation points.

Different colours for the fibre within an operational domain indicate possibly concatenated fibres which may be provided by different fibre providers within that domain.

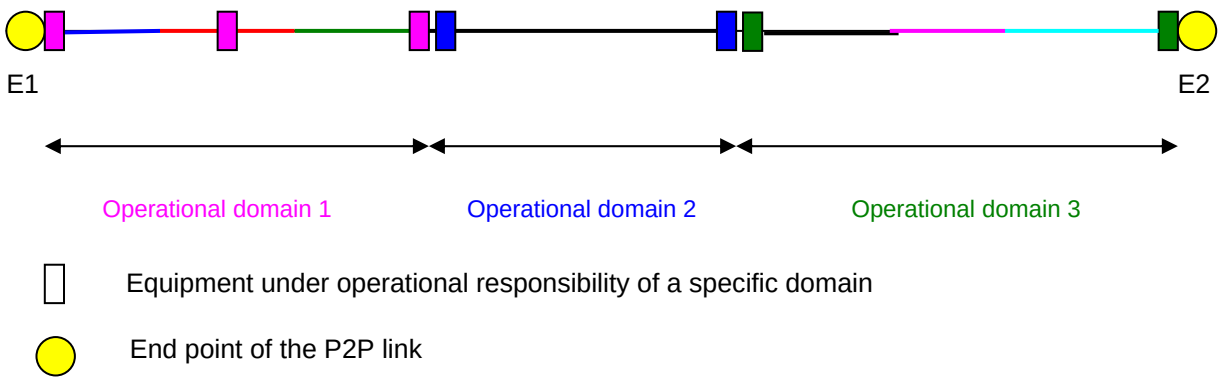

Figure 4.1: Structure of P2P links

A domain is an area of operational responsibility with a homogenous operational concept. The organisation of the management of each domain may differ substantailly fom domain to domain.

A multi-domain environment means that a p2p link is implemented within concatenated domains. Because of the multiplicity of operational responsibilities, and differences of operational models, there is a need to have a coordination function on top of all domains which is in able to give a view on the operational status of the entire

p2p link. This is called the End 2 End Co-ordination Unit (E2ECU). In the cases of supporting projetcs, where a number of p2p links are provsioned to for an Optical Private Network (OPN), the E2ECU is able to give special views of all those links, which belong to a special project (or VPN) like the LCG network or the DEISA network.

4.1 End to End Coordination Unit (E2ECU)

The E2ECU will act as single point of contact for all parties involved in the provisioning of an p2p link to:

- signal faults;
- request information on the status of the p2p link;
- signal scheduled maintenance that will affect the p2p link;
- provide updated information of fault resolution.

Network faults are not fixed by the E2ECU, but always and only by the network where the fault resides. There is the expectation that the network where the fault resides will inform the E2ECU who in turn will inform all other parties involved, on the status of the resolution of faults.

The E2ECU will be able to detect, through the p2p monitoring system being developed jointly by JRA4 and JRA1 (see next section), where a fault resides and will query the networks on the faults it detects.

The parties involved are the NOC personnel of all the networking domains along the end to end path, including networking staff at the end sites.

Appendix A contains a use case of the E2ECU for the LHC application.

The implementation of the E2ECU is currently underway, with discussions between DANTE and the GÉANT2 NOC. The GÉANT2 NOC contract will be extended to perform the functions specified. It is expected that the E2ECU will be fully operational in January 2007.

4.2 Monitoring of End to End (Point-to-Point) Links

In JRA4-WI3 work has been conducted to allow the monitoring of p2p 10Gbps links. In the first instance, only one p2p metric will be monitored, the status, i.e. link UP or DOWN. The p2p link is UP if and only if all intermediary links on the path are UP. If one intermediary link is DOWN, the whole p2p link is DOWN.

The challenge for p2p monitoring of the status of the p2p links is represented by the fact that p2p links are multi-domain and multi-technology, and each domain and each technology offers different ways of extracting

status information of a 10Gbps link. Each network domain will have it's own network performance data repository and access to that repository will differ substantially between networks.

JRA1 has worked on solving this type of problem, and defines a uniform way of storing and presenting network performance data, in a multi-domain environment (perfSONAR). The requirement for the networks involved is to utilise the results of JRA1 to store network performance data. In addition, each network involved in the establishment of a p2p link, is expected to extract status (UP/DOWN) information of their portion of the p2p link and store the information into the perfSONAR framework using the formats defined by JRA1.

JRA4 has developed a tool that will query the distributed perfSONAR infrastructure, extracting the status information provided by the participating networks and provide a p2p view of the link. If there is a fault on the p2p link, it will be able to detect where the fault is on the basis that the network at fault will provide a DOWN status to the perfSONAR infrastructure.

The p2p monitoring tool will be accessible to all parties involved in the provision of a link, so that at any time they can verify the status of the e2elink. This is most useful to the end sites and the E2ECU.

5 Conclusions

The GÉANT2 p2p services have been successfully deployed. There are five different scenarios as to how such services can be implemented in practice. At the time of writing, a total 11 requests have been received from a number NRENs (and corresponding end sites) for the installation of such services across the GÉANT2 footprint.

It is expected that the demand for these services will rise over the duration of the GÉANT2 project.

6 Acronyms

E2ECU	End to End Coordination Unit
EPL	Ethernet Private Line (Alcatel)
DWDM	Dense Wavelength Division Multiplexing
GFP(-F)	(Framed) Generic Framing Protocol
LHC	Large Hadron Collider (CERN)
MCC	Metro Core Connect (Alcatel)
MTU	Maximum Transmission Unit
MPLS	Multi Protocol Label Switching
NMS	Network Management System (Alcatel)
P2p	Point to Point
PHY	Physical Layer Protocol
SDH	Synchronous Digital Hierarchy
VCG	Virtual Concatenation Group
(V)LAN	(Virtual) Local Area Network

Appendix A Potential Technical Service Scenarios

This Appendix provides technical details for several possible scenarios for the provision of p2p services on the GÉANT2 footprint.

A.1 Scenario 1

In this scenario, end-to-end, point-to-point GE services can be offered to the NRENs. The NRENs access GÉANT2 via N x GE interfaces and N x STM-64 interfaces. The scenario is depicted below (in Figure 6.1). NREN-1, accessing GÉANT2 via N x GE interfaces, does not need to use VLAN tagging to identify each instance of GE service. Thus, in principle, each end-to-end service instance should be VLAN transparent meaning that the end users can make use of VLAN tags (i.e. run VLAN trunking over the wide area link) should they wish. If either of the NRENs are themselves using VLAN trunking somewhere along the service path then, unless they can also support VLAN tag stacking in one form or another (e.g. Q-in-Q), end-to-end VLAN-transparency will not be possible. NREN-2, accessing GÉANT2 via N x STM-64, will need to map Ethernet into SDH VCGs using GFP-F encapsulation in order to distinguish between GE services provisioned.

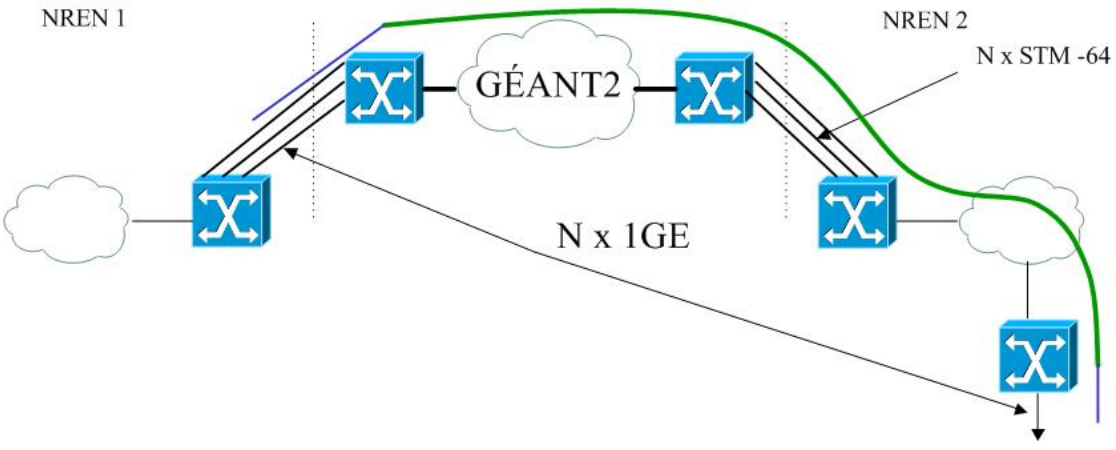

Figure 6.1: Scenario 1 (GE service, accesses via GE & SDH)

In this scenario, end-to-end, point-to-point GE services can be offered to NRENs. The NRENs access GÉANT2 via NxGE interfaces and Nx10GE LAN-PHY interfaces. The scenario is depicted below (in Figure 2.2). NREN-1, accessing GÉANT2 via NxGE interfaces, does not need to use VLAN tagging to identify each GE service. NREN-2, accessing via Nx10GE, will need to use VLAN tagging to identify each point-to-point GE service being carried over the same 10GE access link.

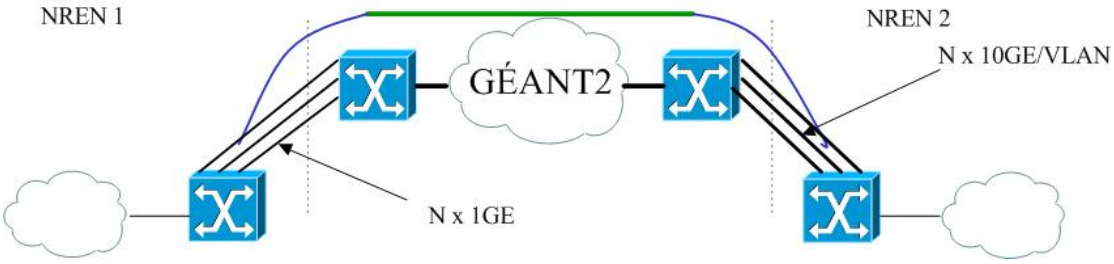

Figure 6.2: Scenario 2 (GE service, accesses via GE & 10GE)

A.2 Scenario 3

In this scenario (depicted below in Figure 2.3), end-to-end, point-to-point GE services are deployed between two NRENs, one having Nx10GE LAN-PHY access to GÉANT2 and one having NxSTM-64 access to GÉANT2. NREN-1 will need to use VLAN tagging to distinguish between the GE services supported whilst NREN2 will need to encapsulate Ethernet frames into SDH VCGs using GFP-F.

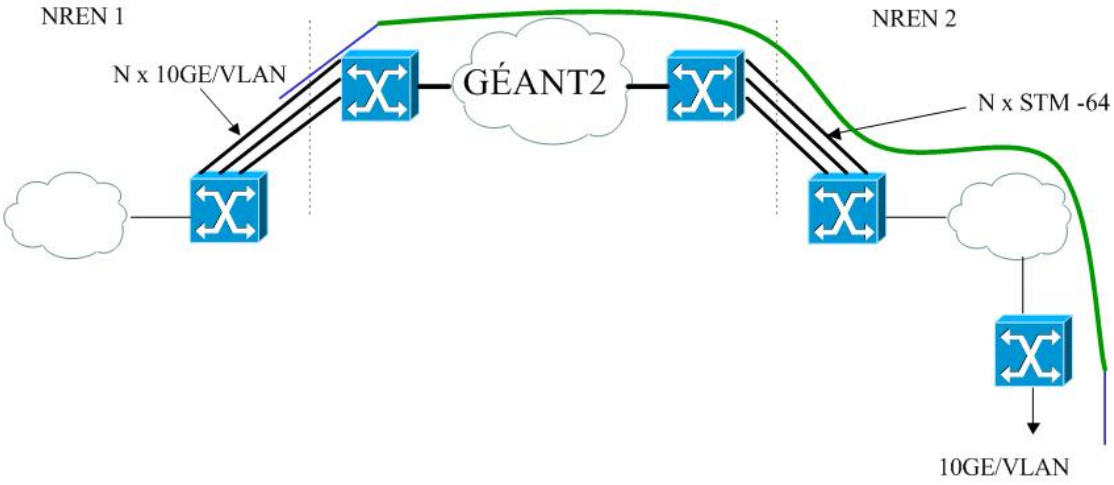

Figure 6.3: Scenario 3 (GE service, accesses via 10GE & SDH)

A.3 Scenario 4

In this scenario (depicted in Figure 2.4), an end-to-end, point-to-point 10GE LAN-PHY service is provisioned between two NRENs, both having Nx10GE LAN-PHY access to GÉANT2. NRENs do not need to use VLAN-IDs to distinguish between 10GE services as these are identified by the physical access port to GÉANT2. Within GÉANT2 either VLANs or VCGs are used to distinguish between end-to-end 10GE services.

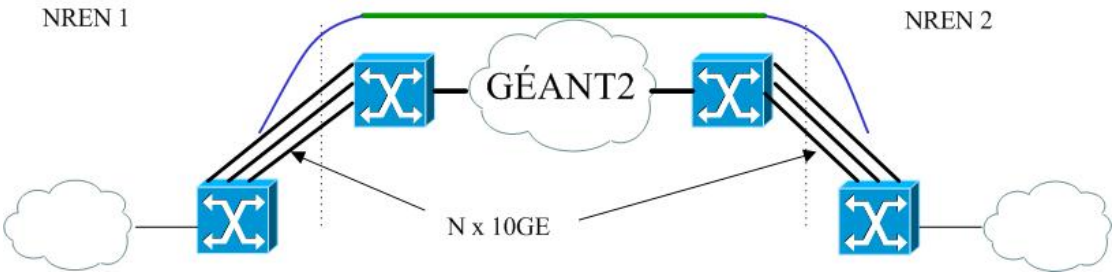

Figure 6.4: Scenario 4 (10GE service, both accesses via 10GE)

A.4 Scenario 5

In this scenario (Figure 6.5), end-to-end, point-to-point 10GE LAN-PHY services are offered between NRENs. NREN-1 has Nx10GE LAN-PHY access to GÉANT2 whilst NREN-2 has NxSTM-64 access to GÉANT2.

NREN-1 does not need to use VLAN-IDs to distinguish between 10GE services as these are identified by the physical access port to GÉANT2. NREN-2 will need to encapsulate Ethernet frames into SDH VCGs using GFP-F.

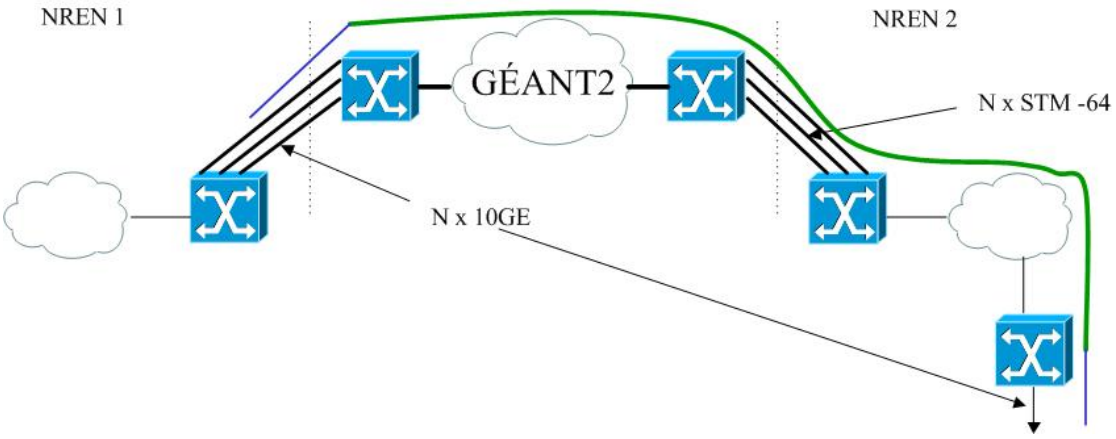

Figure 6.5: Scenario 5 (10GE service, accesses via 10GE & SDH)

Appendix B Use Cases for E2ECU

A number of use cases are detailed here which describe the functions that need to be performed by any entity which co-ordinates the operation of end-to-end links or OPNs based on the optical platform of GÉANT2/NRENs. The use cases are described with the LHC-OPN case as a practical example.. Each use case shows the typical interactions between the parties involved in the provisioning chain.

B.1 Parties involved

The following parties are involved in the typical co-ordination exercise performed by E2ECU, when monitoring and troubleshooting multi-domain point-to-point circuits:

- GÉANT2 NOC
- NREN NOC
- End-site NOC staff (in the case of LHC-OPN, networking staff at T1s and T0);
- Project NOC. Each project or application that avails itself of an OPN provided by GÉANT2 and the NREN may avail itself of its own NOC that manages the OPN. In the case of EGEE and LHC, this is the ENOC.

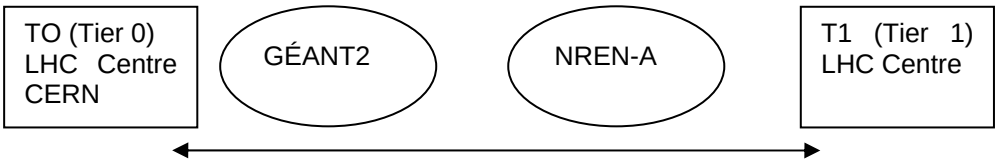

Figure 6.6: Parties involved in a typical E2ECU use case

Assuming the parties involved are as above, five partially overlapping workflow scenarios can be envisaged as typical for most interventions by the E2ECU in a real-time operational environment. They are:

- T1 networking staff notices a fault in connectivity to T0

- The E2ECU takes TT from T1 staff/gives feedback to T1 staff

- NREN A notifies the E2ECU of a fault
- NREN A notifies a scheduled maintenance
- The E2ECU, via the weathermap, notices a fault in NREN-A that has not yet been signalled

Each of the workflow scenarios is explained in more details below:

B.2 Workflow 1: T1 networking staff notices a fault in connectivity to T0

It is envisaged that the following steps would have to occur for this instance to be resolved successfully:

1. the networking staff at T1 will inform the E2ECU via a TT
2. T1 (belongs to NREN A) may contact the NREN A following the national procedures
3. There are the following options:
    a. It is a known fault (or scheduled maintenance) in which case E2ECU will re-issue the trouble ticket, or point to the repository of open tickets.
    b. If it is a new fault, workflow 2 below applies

B.3 Workflow 2: The E2ECU takes TT from T1 staff/gives feedback to T1 staff

Workflow 2 follows on from Workflow 1.

1. E2ECU analyses where the fault is. Two cases apply:
    a. The E2ECU, via the weathermap, is able to automatically locate the fault. The E2ECU will signal the fault to the NREN involved via a trouble ticket and await further information from that NREN. E2ECU will keep other relevant NOCs informed via TTs
    b. It is a new fault and the weathermap does not locate the fault. In this case, the E2ECU will query each NOC involved in the link and ask them to check the status of their portion of the link. Once the fault is located, the NOC where the fault is located will send periodic TTs to E2ECU who will dispatch them to other relevant NOCs.

B.4 Workflow 3: NREN A notifies the E2ECU of a fault

NREN A notices a fault in its own network that affects the p2p link. The following steps are taken:

1. NREN A will notify its user base following its own national procedures.
2. NREN A will notify the E2ECU via a trouble ticket
3. The weathermap will be updated showing a fault in NREN A
4. The E2ECU will notify all other parties affected by the fault via a trouble ticket
5. NREN A will send periodic updates on the resolution of the fault to the E2ECU and these will be passed on by the E2ECU to all other parties affected until the fault is resolved. At that time the weathermap is updated

B.5 Workflow 4: NREN A notifies a scheduled maintenance

NREN A needs to carry out a scheduled maintenance task. In order not to disrupt the service, the following steps should be taken:

1. NREN-A notifies the E2ECU of a scheduled maintenance with an agreed notice period (for example 1 week).
2. The E2ECU re-issues the trouble ticket to all other relevant parties with an agreed notice period
3. The E2ECU stores the information in a DB that can be consulted by all parties via a web portal
4. The weathermap will be able to show that a scheduled maintenance is due on a link within a period of 1 week

B.6 Workflow 5: The E2ECU, via the weathermap, notices a fault in NREN-A that has not yet been signalled

In case the E2ECU is able to identify a fault in the NREN which has not yet been brought to its attention by the NREN itself, the following steps will be taken:

1. The E2ECU contacts the NREN-A to clarify the situation
2. If the fault is confirmed, E2ECU informs GGUS via a TT and other relevant NOCs

3. NREN-A will issue periodic TTs to E2ECU who will dispatch them to GGUS and other relevant NOCs.