

31-01-19

Deliverable D8.10

Consolidated Connection Services v3.0

Deliverable D8.10

Contractual Date: 31-08-2018
Actual Date: 31-01-2019
Grant Agreement No.: 731122
Work Package/Activity: 8/JRA2
Task Item: Task 1
Nature of Deliverable: O (Other – demo)
Dissemination Level: PU (Public)
Lead Partner: CESNET
Document ID: GN4-2-18-615DC2
Authors: J. Sobieski (NORDUNET), M. Hazlinsky (CESNET)

© GÉANT Association on behalf of the GN4-2 project.

The research leading to these results has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 731122 (GN4-2).

Abstract

This document presents a brief overview of the deployment of Consolidated Connection Services v3.0, its key features – multi-domain circuits delivery and performance guaranteed circuits – and three further results: central logging for OpenNSA, the Management Station data-presentation app and integration with GÉANT operations tools. CCSv3 was successfully demonstrated as a working service over a production network at the Supercomputing 2018 conference in November 2018

Table of Contents

Executive Summary	1
1 Introduction	2
2 CCSv3 Production Deployment	3
3 Multi-Domain Circuit Provisioning	5
4 Circuits with Guaranteed Performance	6
5 Other Results	9
6 Recommendations	11
7 Conclusions	14
References	15
Glossary	16

Table of Figures

Figure 4.1: The QoS demonstration – GTS testbed	8
Figure 5.2: ONSA Management Station web GUI – live log feed	9

Table of Tables

Table 2.1: List of nodes and ports covered by CCS (as at 10 December 2018)	4
--	---

Executive Summary

As its name indicates, Consolidated Connection Services (CCS), developed by GN4-2 Joint Research Activity 2 Network Services Development, Task 1, consolidates a number of GÉANT services – GÉANT Plus, Bandwidth on Demand and the dynamic circuit provisioning utilised by the GÉANT Testbeds Service. These all generate RFC 4448 Ethernet-over-MPLS virtual circuits, differing only in their user interface and standards compliance. CCS is compliant to the Open Grid Forum Network Service Interface (NSI) standard, and extends circuit provisioning across multiple NSI-compliant domains. In addition, the service has the capability to deliver circuits with guaranteed performance.

The purpose of deliverable D8.10 is to report on the development effort that has led JRA2 Task 1 to achieve its objective of providing CCS v3.0 as a working service over a production network. Towards the fulfilment of this objective, the key features of the service were demonstrated at the Supercomputing 2018 conference in Dallas, Texas, November 12–15, 2018.

This document presents a brief overview of the production deployment of CCSv3, its key features and three further results identified as requirements during development: central logging for OpenNSA, the Management Station data-presentation app and integration with GÉANT operations tools, which are now running as production beta.

1 Introduction

This document presents a brief overview of the Consolidated Connection Services v3.0 (CCSv3) deployment in the GÉANT core network. The main features of CCSv3 were deployed in the first week of November 2018 and successfully demonstrated at the Supercomputing 2018 conference in Dallas, Texas, November 12–15, 2018.

The CCSv3 service is intended to be the successor of GÉANT Plus and Bandwidth on Demand (BoD) with, in addition, new features such as performance guarantees. CCS consolidates a number of near-duplicate services: BoD, GÉANT Plus and the dynamic circuit provisioning utilised by the GÉANT Testbeds Service (GTS). These three services all generate RFC 4448 Ethernet-over-MPLS virtual circuits, differing only in their user interface and standards compliance. Further, CCS has deployed actual performance-guaranteed capacity allocation – quality of service (QoS) – so that these circuits can now request and are guaranteed a user-specified capacity. The consolidated service is also compliant to the Open Grid Forum Network Service Interface (NSI) standard [\[NSI\]](#), and extends the circuit provisioning across multiple NSI-compliant domains. CCS provides a single service development and management solution that replaces all three prior separate services.

The two most important new features introduced with CCSv3 are:

1. Multi-domain circuits delivery, which required control plane peering with other NSI-enabled domains and improved pathfinding in the OpenNSA aggregator.
2. Support for QoS-enabled Ethernet circuits. This feature required an improvement in the reservation procedure in order to manage capacity available on a selected path. It also required improved pathfinding to explore/allocate alternative paths in case of no bandwidth on the primary path. Other work had to be done on defining a configuration procedure for the Juniper MX platform and implement it in the backend.

Other results were achieved which were not originally listed as goals, but were identified as requirements during the development process and are now part of the current CCSv3 running as production beta. The most important of these are:

- Central logging for OpenNSA in order to improve operational usability.
- The Management Station, a presentation app which shows real-time operational data for each circuit.
- Integration with GÉANT operations tools such as OpsDB and Cacti.

All these features and results are described in the following sections, together with an overview of CCSv3 production deployment.

2 CCSv3 Production Deployment

To date during GN4-2, Consolidated Connection Services (CCS) has been provisioned entirely as “best effort” circuits (no quality of service (QoS)) over the production GÉANT core network. The service is based on OpenNSA, an open source implementation of the NSI protocol. In order to enable this deployment, the OpenNSA hardware interface (the southbound interface (SBI)) module for the Junos Space API has been implemented. CCS provisions approx. 350 circuits each month, primarily for the GÉANT Testbeds Service (GTS). The best effort circuits are provisioned over Juniper MX960 core routers.

In autumn 2017, CCS was adapted to dynamically provision optical transport network (OTN) circuits in 1 Gbps granularity. This capability was demonstrated at Supercomputing 2017 (12–17 November 2017 in Denver, Colorado) on Ciena equipment. The OTN capability was initially considered for the GÉANT circuit performance guarantees, but the proposal was dropped when the software adaptation to the Infinera kit proved intractable due to software compatibility issues and the short expected remaining lifetime of the Infinera switched OTN infrastructure.

A solution to providing performance guarantees (quality of service (QoS)) was proposed that used Juniper MX204 routers with a hierarchical QoS (HQoS) feature – the necessary hardware to support appropriate policing, shaping and precedence queuing [[QoS](#), [HQoS](#)]. This solution was validated in the lab during summer 2018, and deployment began in autumn 2018. The performance-guarantee capability was demonstrated on circuits between London and Hamburg in the pilot deployment of GTS in November 2018. This also coincided with GÉANT’s presence at the SC18 conference [[SC18](#)].

The current version of CCS, v3.0, has been deployed to enhance the GÉANT Testbeds Service. CCS v1.0 and v2.0 are still being integrated into operations tools. The final goal is to turn CCS into the primary circuit provisioning system for GÉANT. At the end of GN4-2, CCS serves four GÉANT core points of presence (PoPs) that constitute the GÉANT Testbeds Service. These locations have been augmented with QoS-capable switching elements, on an aggressive schedule. London, Amsterdam, Hamburg and Milan were completed in 2018, with the remaining sites of Madrid, Paris, Prague and Bratislava (making eight) to complete promptly after the holiday blackout for changes is lifted, in early 2019.

The CCS topology is being extended by other external network-to-network interface (ENNI) and user network interface (UNI) ports into other GÉANT PoPs to promote NSI peering with NRENs and open exchange points (see also Section 3).

CCSv3 now covers the following locations and devices:

Location	Client ports	Core ports (internal SDPs)	External ports (Peering domains)
Amsterdam	21	13	1
Bratislava	17	9	0
Hamburg	25	13	0
London	24	13	0
Madrid	17	9	0
Milan	17	9	0
Paris	25	9	0
Prague	25	9	1
Total locations: 8	Total client ports: 171	Total core ports: 84 (56 are for MPLS- based SDPs)	Total external ports: 2 (5 more are expected soon)

Table 2.1: List of nodes and ports covered by CCS (as at 10 December 2018)

3 Multi-Domain Circuit Provisioning

The main advantage of basing the system on the NSI protocol is that the final solution can automatically deliver multi-domain circuits when peering with other NSI-enabled domains is established. Currently, CCSv3 peers with two foreign domains, CESNET and Netherlight. Peering with additional GÉANT NRENs, as well as with other networks via Open Exchange Points, is now largely an operational configuration task with CCSv3 that will be commence in early 2019. Other known, interested early adopters are DFN, NORDUnet, the ANA consortium, ESnet, and RENAM. It is also recommended that the GÉANT Open exchange points also deploy the CCS v3 MD and QoS capabilities as quickly as possible.

In order to be able to provision multi-domain circuits using OpenNSA, the pathfinding logic had to be significantly improved: a completely new solution was required for cases when the remote end of the circuit is in a foreign domain and the aggregator does not have all necessary information about the topology of that remote domain, so cannot calculate the exact path to the remote port. In such cases the solution now queries neighbour domains; if they can deliver a circuit from the NNI between the GÉANT network to the remote end of the requested circuit, the aggregator will let that neighbour provide it and then deliver the remaining local part of the circuit.

Multi-domain circuits were successfully publicly demonstrated at the Supercomputing 2018 conference (SC18) in Dallas, Texas, November 12–15, 2018 [[SC18](#)].

4 Circuits with Guaranteed Performance

The next important goal on the roadmap for CCSv3 was the capability to deliver circuits with guaranteed performance.

Guaranteed performance (QoS) enables applications to have a predictable level of end-to-end service. This is of particular interest to network researchers for whom performance evaluation is a critical aspect of their work. Without a known baseline performance of testbed resources – especially transport circuits—it is impossible to determine if a link is being utilised efficiently or effectively. Traditional over-engineering strategies of best effort IP infrastructure – particularly where user flows are expected to consume a large percentage of backbone link capacity – are expensive, and do not address the global end to end needed to assure predictable, and repeatable, performance. The emergence of MPLS-based traffic engineering, commoditised HQoS hardware features (used commonly in large-scale commercial services), increasingly integrated networked applications (with compute, storage, transport (circuits), instruments/sensors), and emerging service capabilities, such as network slicing require QoS and even reference NSI [\[5G PPP\]](#).

The primary (and initial) use case for CCS QoS-enabled circuits is GTS. As the QoS is fully deployed and available for GTS by CY2019-Q1, we expect to see QoS provisioning to be used substantially. As the multi-domain CCS/NSI peerings increase to create a larger end-to-end service region, we expect QoS will become a go-to feature to simplify end to end provisioning and to offer users predictable service levels.

As the NSI service definition for an Ethernet circuit (EVPN) includes parameters for stating how much bandwidth is requested, OpenNSA was already able to receive and process such requests. However, there were still significant blockers in the code of OpenNSA which prevented the introduction of this feature as a part of production CCS. The following issues had to be resolved in order to deliver guaranteed performance capability:

- Implement infrastructure with support for per-circuit quality of service (QoS).
- Define and validate the configuration procedure for the chosen technology.
- Integrate the scripted provisioning procedure into the OpenNSA backend for the chosen technology.
- Improve the reservation process of OpenNSA in order to be able to check whether the requested bandwidth is available along the selected path.

The current GÉANT MPLS core network is not equipped with the necessary hardware features to fully support hierarchical quality of service (HQoS). After investigation and evaluation of different possibilities, it was decided to use Juniper MX204 routers to create dedicated QoS-enabled infrastructure. As the GÉANT Testbeds Service (GTS) had the initial requirement for deterministic

performance and circuit isolation, it was decided to deploy the QoS capability initially covering all eight GTS locations.

The topology configuration consisting of both MX960s and MX204s – some supporting QoS and some not – was integrated into the OpenNSA backend.

The key requirements to be met by the QoS provisioning solution were:

- Ensure QoS circuits always got their guaranteed capacity.
- If the QoS capacity was not being utilised by the QoS application, allow best effort traffic to burst into that available capacity up to the full link rate.
- Ensure that multiple QoS circuits were provided with a weighted fair queuing ability to pace/shape their traffic fairly among themselves.
- Police excess traffic at ingress.
- Allow a link to be configured so as to ensure there is always some minimum best effort capacity
- Ensure the pathfinder can seamlessly distinguish QoS-capable network elements from legacy network elements that do not support QoS provisioning

The complete solution, together with the improved reservation process, has been tested in the JRA2 development laboratory and then deployed into production as the new MX204 routers have been installed.

Creating testbeds in GTS with QoS-enabled links was successfully publicly demonstrated at SC18.

The validation scenario established five testbeds – each with a source VM in London, a destination VM in Hamburg, and a single circuit connecting these VMs. Three of these testbeds (A, B, and C) had best-effort connecting circuits. In Testbed D, the circuit was provisioned with a guaranteed 3 Gbps, and in Testbed E the connecting circuit was provisioned with 4 Gbps guaranteed capacity.

These testbeds were forced (in CCS) to transit a single 10Gbps link between London and Hamburg in order to test the queuing performance under load. The ensuing tests (see Figure 4.1 below) started an iPerf TCP stream in each testbed in sequence. As Testbed A began, it quickly climbed to 10 Gbps. When Testbed B was added, the streams fell back to just under 5 Gbps each, and Testbed C settled in at approximately 3Gbps for all three best-effort circuits in parallel. As the QoS Testbed E was started, it quickly claimed 4 Gbps, and the best-effort streams fell back to circa 2Gbps each. As Testbed D was added it quickly claimed its 3Gbps, without noticeable degradation of Testbed E's 4Gbps. The best-effort circuits all dropped back to approximately 1Gbps. The behaviour of the five testbeds was consistent, regardless of the sequence in which the flows commenced.

Additional testing should be continued to more accurately understand the nature and cause of the stream variance and why this occurs (which could be due iPerf nuances, or to potentially other behaviour of the TCP reactions to interfering flows). UDP tests should be run with proper pacing applied (similar behaviour would be expected on both TCP and UDP tests).

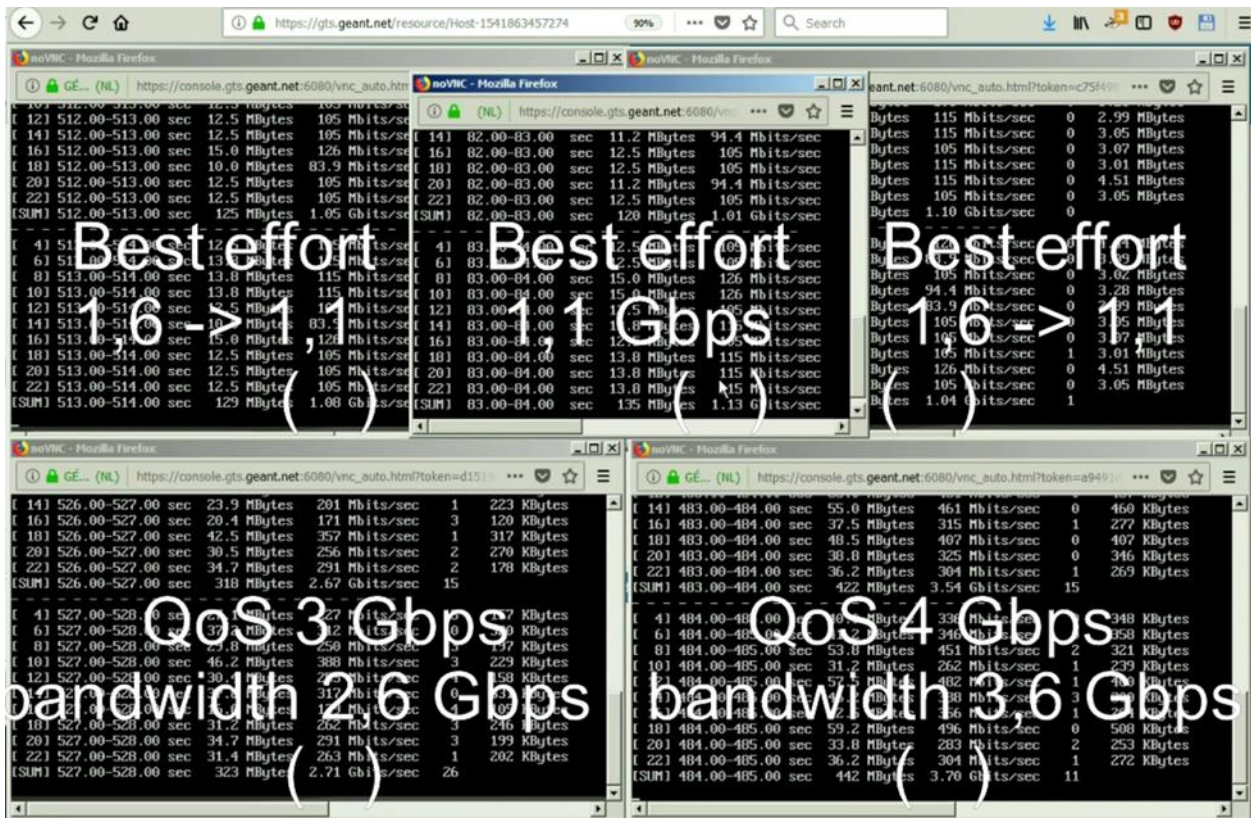


Figure 4.1: The QoS demonstration – GTS testbed

The demonstration of two circuits with guaranteed performance and three, best-effort circuits showed how the best-effort circuits experienced bandwidth fluctuations, while the QoS circuits had stable bandwidth when provisioned over the same 10GE link.

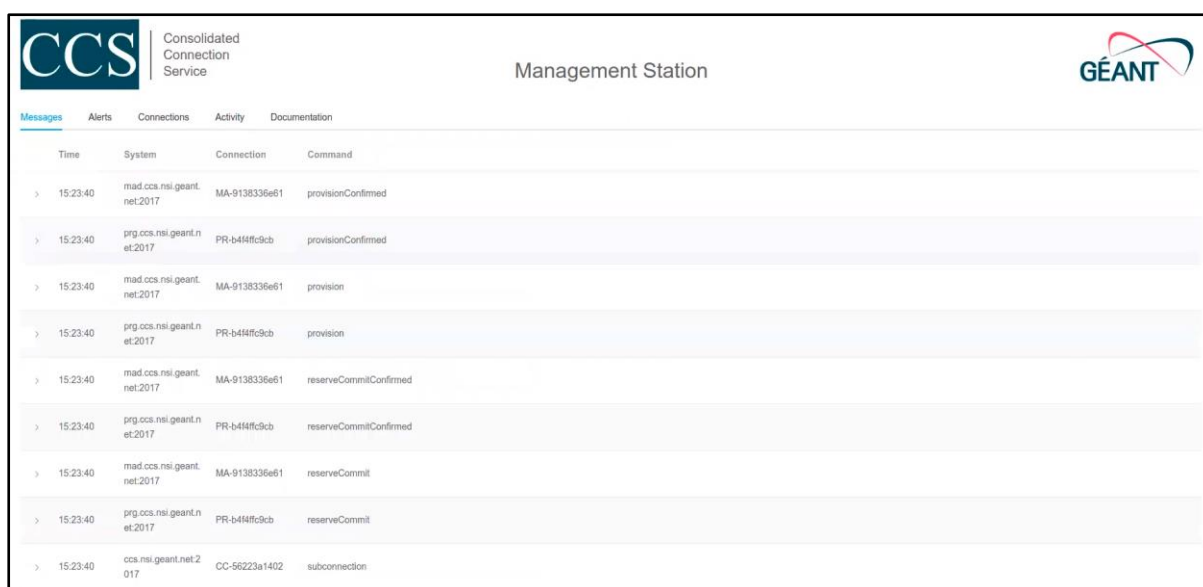
5 Other Results

As stated in the introduction, other results were achieved which were not originally on the roadmap but were identified as requirements during CCS development. The most important of these are:

- Central logging for OpenNSA in order to improve operational usability.
- The Management Station, a presentation app which shows real-time operational data for each circuit.
- Integration with GÉANT operations tools such as OpsDB and Cacti.

The central logging application allows the collection of all state data about all running OpenNSA agents as well as state data about individual service instances (circuits) at any moment of their lifecycle. This allows the solution to react to possible error events or send status information to any other system where it might be useful (or needed) such as Network Operations Centre (NOC) monitoring systems or the service portal.

The ONSA Management Station has been introduced as a lightweight application that can present state data about all OpenNSA agents and services in real time from the central logging system. It also contains a simple form for circuit request creation and a documentation section for OpenNSA and the ONSA Management Station.



Time	System	Connection	Command
> 15:23:40	mad.ccs.nsl.geant.net.2017	MA-9138336e61	provisionConfirmed
> 15:23:40	prg.ccs.nsl.geant.net.2017	PR-b44ff9cb	provisionConfirmed
> 15:23:40	mad.ccs.nsl.geant.net.2017	MA-9138336e61	provision
> 15:23:40	prg.ccs.nsl.geant.net.2017	PR-b44ff9cb	provision
> 15:23:40	mad.ccs.nsl.geant.net.2017	MA-9138336e61	reserveCommitConfirmed
> 15:23:40	prg.ccs.nsl.geant.net.2017	PR-b44ff9cb	reserveCommitConfirmed
> 15:23:40	mad.ccs.nsl.geant.net.2017	MA-9138336e61	reserveCommit
> 15:23:40	prg.ccs.nsl.geant.net.2017	PR-b44ff9cb	reserveCommit
> 15:23:40	ccs.nsl.geant.net.2017	CC-56223a1402	subconnection

Figure 5.2: ONSA Management Station web GUI – live log feed

The central logging app also implements a simple messaging system which is able to send status updates to a registered third-party system. This has been implemented in response to a GÉANT Operations team feature request for a notification to be sent to the operations database (OpsDB) as well as to the monitoring system (Cacti) whenever a new service is activated or deactivated.

The ONSA Management Station and central logging are now considered beta versions and are being tested with real production data.

6 Recommendations

CCS is a collection of software packages, deployed and configured to deliver a service. Automated multi-domain provisioning can play an enormous role in the NREN community. The GN4-2 effort concentrated on the GÉANT core deployment but the efforts and resulting service software reach beyond this. All of the NRENs, campuses and labs can deploy CCS.

The CCS developers were asked to address GÉANT-specific requests, such as integration with the existing ticketing system, however, capabilities delivered and the resulting functional capabilities apply across all deploying service domains. Looking ahead, there are opportunities to improve these services, as listed below.

- **Operational expansion of the NSI/CCS service region.** As a pilot service, CCS was constrained to find a balance between a large enough service footprint to show viability, and a small footprint to minimise capex and opex costs as the service was refined. This covered eight GTS locations. Now, as CCS migrates into a fully supported production GÉANT service, the GEANT core network must expand the reach of CCS, and promote the CCS adoption and deployment in the partner NRENs and campus networks. Service region expansion does not require new features, but rather is an operational process of adding new ports into the set of Service Termination Points (STPs) served by CCS. These ports are the NREN Access ports connecting GÉANT and its NRENs. Other ports include Open Exchange Points, such as London Open, NetherLight, MANLAN, etc. Additional internal GÉANT ports serving other aspects such as perfSONAR and VMware servers.

CCS already has approximately 160 ports (mostly GTS ports), comprising almost 2 Tbps of capacity in its STP topology that define its current edge-to-edge reach. Adding NRENs will allow CCS to serve those circuit customers that require services other than GTS (e.g. transit circuits from NREN to NREN or NREN to CSP, etc.)

- **Operational expansion of the CCS/QoS service region.** The QoS feature has only been deployed in locations serving GTS platforms. This QoS capability must be extended to/from the NREN access ports. GTS testbeds, or any other circuits that require performance guarantees for multi-domain require edge-to-edge, QoS-capable switching. Multi-domain testbeds extending beyond the GÉANT edge require at least one NREN access port to be CCS/QoS enabled (i.e. the NREN port must be moved to directly attach to an MX204 QoS capable router at the edge.) In the eight existing GTS locations, the first few NRENs could easily re-home one of their access ports to the MX204 in that location. In the other (non-GTS) locations no QoS-capable platforms have been deployed. Looking ahead, we should be prepared to offer QoS – enabled CCS Access ports to all NRENs if/when they ask to be part of the CCS service region.

The MX204s are inexpensive and have proven QoS performance, but are limited in their port-count scalability (they need more ports). Alternatively, HQoS capable blades could be

deployed in the larger MX960 routers, or evolve to a future switching architecture that supports hierarchical QoS.

- **Simplified installation.** In order to promote adoption and deployment of CCS in the NRENs, initial deployment should be painless: not require special hardware or take more than a few hours to achieve a functioning service in a new network. The cost of entry to NRENs should also be far lower than it is at present. Support for software switching elements on commodity server platforms, such as Open vSwitch (OVS) or similar, should be developed as one of the data plane hardware options to allow NRENs or other service providers to quickly and inexpensively deploy the circuit provisioning service. In the past, such software implementations contained bugs and were lacking important features such as shaping and policing or MPLS support. Mature software switches are available now that can provide 10-20 Gbps throughput on even old commodity servers. The ability to provision circuits across a network service domain will be crucial to deploying other higher layer advanced services – even if an intermediate transit network does not deploy those high-level services, multi-domain services will rely upon multi-domain automated provisioning of performance guaranteed and best effort circuit services. Simplified entry is critical.
- **OVS as a first-class data plane switching element.** OVS in compute nodes is a nontrivial feature. It integrates CCS and the OpenStack VM services, allowing circuits to be provisioned directly VM to VM inside the computeNode without transiting the local MX router. This “bridged” circuit is a more elegant way to integrate VM STPs, but introduces a number of issues, most notably around dynamic creation/deletion and advertising of VM ports and their physical mapping/QoS attributes and modification properties.
- **Policy engine.** Enhanced policy control allows the service provider to authorise different users or projects according to internal or community-based agreements. For instance, students/researchers limited to certain aggregate capacity on circuits, or certain peer networks (e.g. Internet2 or NetherLight) may be allowed a much larger capacity and longer hold times in return for reciprocal arrangements. Policy engines should be interoperable with other services like VM services, storage services or GTS to align policy across interoperating services. This process requires- higher-level management discussions.
- **Define a simple initial policy regimen** among global NSI peering domains that allows circuit services to evolve.
- **Enhanced NSI “Aggregator” multi-domain path selection agent.** With NSI, multi-domain provisioning is tractable, but requires thorough and detailed understanding of recursive path computation (tree and chain path selection, etc.) Distributed path computation can be computationally and time intensive, so it is important to identify heuristics for pruning and evaluating those heuristics with real testing.
- **SDN BoD, or a similar implementation of label/VLAN switching (with QoS) on SDN hardware** (such as the Corsa OpenFlow switches, or emerging P4 whiteboxes) should be supported. This solution would reduce the incompatibility among various products in the NREN networks. As the intent of CCS is to consolidate a number of GÉANT Connection Services, the next stage of CCS evolution should take the abstracted Connection Service model and apply it to services, such as Wave/Lambda circuits. The NSI provisioning model supports this.

A more intriguing challenge would be to build upon the multi-domain foundation of NSI/CCS and the high performance virtual switching capabilities of GTS to deliver multi-point L2 services across multiple domains.

In general, as networking evolves advanced services like CCS, GTS, PVM, etc that are tuned to the R&E requirements will require in-house (or at least community based) software support. It is not simply about fixing bugs, but about building the foundation for more advanced and sophisticated services that will require CCS.

7 Conclusions

The purpose of deliverable D8.10 is to report on the development effort that has produced the Consolidated Connection Services v3.0 (CCSv3) software suite.

The current CCS v3 is the culmination of over four years of iterative improvements and field use by GTS and other NRENs around the world. This software has been deployed as a working service over the GÉANT core production network. CCS has been used heavily by the GTS over several years, and now the GTS v6 production service will be relying upon CCS v3 for all of the circuit provisioning requirements. With the authorisation enhancements implemented in CCS, it is now generally available for other client applications to leverage. As has been described in this document, CCS v3 delivers several important features:

- NSI standard API conformance.
- Multi-domain automated circuit provisioning.
- QoS-enabled Performance guarantees for Ethernet circuits.

Towards the fulfilment of its objective, all of these features were successfully demonstrated at the Supercomputing 2018 conference in Dallas, Texas, November 12–15, 2018.

Additional results were achieved which were not originally on the CCS roadmap but were identified as requirements during development. The most important of these, central logging for OpenNSA, the Management Station and integration with GÉANT operations tools, are now part of the current CCSv3, running as production beta.

References

- [5G PPP]** <https://5g-ppp.eu/wp-content/uploads/2018/01/5G-PPP-5G-Architecture-White-Paper-Jan-2018-v2.0.pdf>
- [HQoS]** Hierarchical Class of Service Overview
https://www.juniper.net/documentation/en_US/junos/topics/concept/hierarchical-cos-overview.html
- [NSI]** Network Service Interface 2.0
<https://www.ogf.org/documents/GFD.212.pdf>
- [SC18]** https://www.geant.org/News_and_Events/Pages/GEANT-at-SC18-with-the-worlds-High-Performance-Computing-community.aspx

Glossary

API	Application Programming Interface
BoD	Bandwidth on Demand
CCS	Consolidated Connection Services
ENNI	External Network-to-Network Interface
EVPN	Ethernet VPN
GE	Gigabit Ethernet
GTS	GÉANT Testbeds Service
GUI	Graphical User Interface
HQoS	Hierarchical Quality of Service
IETF	Internet Engineering Task Force
JRA2	GN4-2 Joint Research Activity 2 Network Services Development
MPLS	Multi-Protocol Label Switching
NNI	Network-to-Network Interface
NOC	Network Operations Centre
NREN	National Research and Education Network
NSI	Network Service Interface
OpsDB	Operations Database
OTN	Optical Transport Network
OVS	Open vSwitch
PoP	Point of Presence
QoS	Quality of Service
RFC	Request for Comments
SBI	Southbound Interface
SC18	Supercomputing 2018
SDP	Service Distribution Path
Task 1	JRA2 Task 1 Consolidated Connection Services
UNI	User Network Interface
VPN	Virtual Private Network