Experimental Demonstration of Multi-vendor and Multi-domain Elastic Optical Network with data and control interoperability over a Pan-European Test-bed


Abstract— The operation of multi-domain and multi-vendor EONs can be achieved by interoperable Sliceable Bandwidth Variable Transponders (SBVTs), a GMPLS / BGP-LS-based control plane and a planning tool. The control plane is extended to include the control of SBVTs and EXCs, enabling the end-to-end provisioning and recovery of network services. A multi-partner testbed is built to demonstrate and validate the proposed end-to-end architecture. Interoperability among S-BVTs is experimentally tested between different implementations. In this case, transponders are configured using the proposed control plane. The achieved performance with hard-decision and soft-decision FECs using only the information distributed by the control plane is measured against the performance of single-vendor implementation, where proprietary information is used, demonstrating error-free transmission up to 300 km.

Index Terms—Interoperability, SBVT, EON, FEC, standardization, DSP, data-plane

I. INTRODUCTION

Elastic Optical Networks (EONs) are envisioned as the base of next generation transport networks. Recent studies [1] show that EONs will be able to cope with the imminent capacity crunch thanks to their flexibility and scalability resulting in an efficient bandwidth allocation. The pillars allowing this evolution are the flexible bandwidth-variable transponders (BVTs), capable of transmitting / receiving signals with configurable physical parameters (i.e., bitrate, modulation format, FEC, etc.) and the spectrum-selective switches (SSSs), capable of switching frequency slices in multiples of 12.5 GHz. By using these two elements, EONs provide the potential to enable fully configurable multi-bitrate lightpaths, thus increasing service-oriented flexibility and overall network capacity.

Key factors for the successful deployment of EONs are the presence of solutions for the subdivision of the network in different domains (administrative, geographical, vendor, etc.) and for the interworking among the different implementations. These are seen as the main requirements emerging from the network operators that have to cope with the growing of traffic and the emerging of new services on one hand and with market competition on the other. It is therefore clear that adequate control plane operations are required to achieve lightpath provisioning, along with additional EON-specific procedures, such as elastic operation and re-optimization (e.g., defragmentation) in such a complex environment.

In this paper, we present a multi-domain, multi-vendor EON integrating both control and data plane. Protocol extensions are evaluated in a distributed multi-partner control plane test-bed, fully enabling the interoperability of several implementations of Sliceable Bandwidth Variable Transponders (S-BVTs), which are experimentally validated.

The S-BVTs of each implementation are configured using only the information distributed by the control plane. The performance, in terms of BER, is compared against the single-vendor implementation case, where the information used by the S-BVTs is not restricted to the one disseminated by the control plane and proprietary FEC and DSP algorithms can be employed.
II. CONTROL ARCHITECTURE AND PROCEDURES

The control plane architecture and interfaces designed for the multi-domain EON extend the ones previously presented [3] in order to include the control of SBVTs, as illustrated in Fig. 1. The control plane enables the end-to-end provisioning and recovery of network services (either a flexi-grid network media channel or a constant bit rate service between transceivers with different bit rates and dynamic ODUFlex grooming).

In this section, we detail the functional elements of the architecture, we sketch the control plane procedures and we summarize the protocol extensions that have been proposed.

From a bottom-up approach, each domain deploys an extended GMPLS control plane including, notably, the OSPF-TE protocol for topology dissemination and the RSVP-TE protocol for the signaling of the Label Switched Paths (LSPs). On top of the GMPLS control plane, each domain deploys an active stateful Path Computation Element (AS-PCE), for the purposes of both optimal path computation and service provisioning within its domain. Thus, multi-domain path computation and provisioning is carried out by means of a Hierarchical Path Computation Element (H-PCE), with the parent PCE (pPCE), coordinating the procedures between children PCEs (cPCE): the interface between pPCE and domain cPCEs (based on PCEP protocol) is thus used by the pPCE for path computation and instantiation. This involves the use of the PCEP protocol request (PCReq) and response (PCRep) messages, as shown in Fig. 2. Once the ERO is computed, the pPCE forwards this message to the cPCEs so each cPCE requests the head end node of its domain to initiate the GMPLS signalling process, using the RSVP-TE Path and Resv messages. During the signalling processes, the nodes end up configuring the underlying optical hardware: flexi-grid ROADMuns are configured with the frequency slot parameters and S-BVTs are configured with the allocated transmission parameters (bit rate, modulation format, FEC, etc.)

Fig. 1: Multi-domain Control Plane Architecture

Fig. 2: Control plane workflow for the single instantiation of a service, path computation is performed by the H-PCE
endpoints and allocated frequency slot parameters are compatible across the involved domains.

The second workflow is characterized by the use of the planning tool to perform the computation of a batch of services, described in Fig. 3. The establishment of a batch of services is again driven by the NMS which, this time, requests the computation and provisioning to the planning tool, via a dedicated REST interface. The planning tool performs the joint computation of the paths for each service and, once computed, proceeds with their instantiation using the interface provided by the ABNO controller.

Control plane extensions affected all the protocols of the GMPLS suite together with the ones adopted as northbound interfaces between the cPCEs and the pPCE (i.e. PCEP and BGP-LS). The main purpose of these extensions is the handling of the media layer in which switching is based on a frequency slot described as central frequency and a slot width. A new 64-bits label format includes the grid type (assigned a value of 3 to indicate ITU-T Flex), the channel spacing (a value of 5 indicates a 6.25 GHz granularity), the n parameter that expresses the slot central frequency according to the formula Frequency (THz) = 193.1 + n * 0.00625 and m is the slot width as in Slot Width (GHz) = 12.5 * m. The new format is used in all the objects carrying a label (GENERALIZED_LABEL, SUGGESTED_LABEL, LABEL_SET, ERO, etc.)

Bandwidth and traffic specification (when referring to optical spectrum) conveys the m parameter (notably during the path computation and signaling). It can be observed that ‘m’ is a parameter both of the GMPLS Flexigrid label and of the Flexigrid TSpec and Flowspec. The overlap comes from the fact that in a Flexigrid system the label value, that defines what is switched, indicates the slot width, therefore affecting also the bandwidth supported by an LSP. A new sub-object of the EXPLICIT_ROUTE object was defined to describe MF-OTPs or SBVTs capability of generating multiple optical flows. It is formed by a list of TLVs describing the sub-carrier attributes and appears only at the beginning and the end of the ERO to convey specific information about the configuration of the MF-OTPs at the path endpoint [7]. Both OSPF-TE and BGP-LS were extended similarly, to describe the capabilities of the S-BVTs and for the dissemination of the availability of nominal central frequencies using a bitmap encoding.

III. TEST-BED DESCRIPTION

In order to demonstrate the feasibility and performance of a multi-domain multi-vendor EON network, a pan-European test-bed with both control and data capabilities has been developed by the IDEALIST project [1]. This test-bed, illustrated in Fig. 4, is built by the interconnection of different components, both hardware and software, physically distributed within labs. The test-bed encompasses three Flexi-grid domains with different capabilities, one hierarchical PCE, an ABNO Controller and the PLATON planning tool. The domains are interconnected resulting in the inter-domain topology shown in Fig. 5. Two kinds of domain interconnections are envisioned. In the first one, S-BVTs are attached to the border nodes of each domain and both transponders are back-to-back connected. In the second one, the border nodes are transparently interconnected by a fiber, allowing end-to-end media channels to be set up. Note that,
practical reasons, real data plane connectivity only happens inside the domains, while the interconnection is logical. In addition, it is worth mentioning that the data plane domains, as shown below, are multi-partner, and equipment developed by multiple parties was integrated in the same laboratories.

A. Flexi-Grid Domain

The flexi-grid domain defined as (102) in Fig. 5 includes two nodes (Fig. 6). Each node includes:

- a flexible and configurable digital cross connect (OTN switching fabric) for client mapping and centralized control of S-BVT. To this end, the hardware controller has been modified to adapt the commands from the restful server to the registers of the FPGA-based S-BVT;
- real-time S-BVT modules carrying multiple OTU2 tributaries into a single flexible OTU container (which follows the “beyond 100G” OTN standards recommendation thanks to programmable FPGAs and a multi-flow optical front end that can adapt its symbol rate such as 107 Gb/s or 53.5 Gb/s per carrier [8] and its number of carriers according to the reach, physical impairments or capacity demand. The FPGA platform is a Xilinx Virtex 7 580HT with high speed output serializer/deserializer GTZ interfaces up to 28 Gbit/s;
- an all-optical matrix that provides flexibility through architecture-on-demand (AoD) in terms of synthesis of fiber switching cross-connections (~1dB/cross-connection, 192x192 Polatis switch series 6000), optical bandwidth switching with bandwidth-variable WSSs (4x16 Finisar Waveshaper)[9];
- an offline coherent receiver (Fig. 6), able to recover the transmitted signals using an all-optical coherent front-end (polarization beam splitters, 90° optical hybrid and balanced receivers) and an offline DSP (OMA Keysight N4391A).

The GMPLS control plane is thus able to configure the underlying hardware (via the connection control interface) by using dedicated REST interfaces, as mandated by the signaling process. This covers the cross-connection configuration, the WSS filters and the S-BVT. A deployed middleware translates the high-level REST interfaces to the actual low-level hardware interface. In particular, it can configure the number of physical OTU2 tributaries/lanes interfaced to the elastic BVT, and the required symbol-rate. So if there are five OTU2 lanes to be carried, the BVT is configured at 13GBd PM-QPSK. The middleware contain RESTful Servers e.g. running as a northbound interface on top of a SDN OpenDayLight (ODL) controller with their southbound interface based on extended OpenFlow Protocol (OFP). The transport equipment (e.g., Fiber Switch and BV-WSS) has an OFP agent capable to translate the received OFP commands to specific equipment APIs (e.g., TL1).

B. Multi-vendor Flexi-grid domain

The second domain (101) (Fig. 4 bottom right) includes integrated data and control plane.

The data plane setup consists of four nodes flexgrid network (based on configurable spectrum selective switches- SSS), a CNIT/Ericsson DSP unit at the TX and two different ones (CNIT/Ericsson and Coriant) at the RX as part of an optical coherent test-bed. The TX is able to provide a super-channel with different configurable number of carriers and capacity (i.e. 1 carrier for 100G, 3 carriers for 400G and 7 carriers for 1T). At 1 Tb/s, we adopted PM-16QAM Nyquist-shaped signals shaped by a roll-off = 0.05 and symbol rate = 23 GBd. A 64 GSa/s Digital-to-Analog Converter (DAC) was used with 3dB bandwidth of 13 GHz and ENOB 5.5. In order to mitigate DAC bandwidth limitations, digital pre-emphasis was applied to ensure proper rectangular shape. Two single polarization IQ-modulators were used to modulate the even and odd sub-carriers of the super-channel. Subcarrier spacing was set to 25 GHz. Next, the single polarization signals went through the polarization multiplex emulation stage where they were first split into two orthogonal polarizations, delayed with respect to each other for de-correlation and finally combined by a PBC resulting in a polarization multiplexed signal. The even and odd sub-carriers were multiplexed via 3dB coupler resulting in a multiple sub-carrier super-channel. The test-bed was equipped with an optical re-circulating loop consisting of...
recommendations [13]. The AC attribute, exported by BGP-LS and PCEP/RSVP-TE protocols, in line with ITU Application Code (AC) and Transponder Class (TC) attributes configuration.

The same policy is enforced in the RSVP-TE signaling at the network scenario of application (i.e., in terms of wavelength frequencies and width, modulation, FEC). Path computation [10], multi-action re-optimization, adaptation and equalization and suppress accumulated out of band ASE noise. In addition, a polarization scrambler (PS) was also employed to evenly distribute polarization dependent loop effects. After the loop circulation, the desired sub-carrier was selected via a tunable band-pass filter, then demodulated by employing coherent phase- and polarization-diversity detection, and finally setting the local oscillator (LO) at the nominal central frequency of the selected sub-carrier. The received optical signal is mixed with the LO through a polarization-diversity 90° hybrid optical coupler, whose outputs are sent to four couples of balanced photo-diodes. The four photo-detected signals are sampled and digitized through a 20 GHz 50 GSa/s real-time oscilloscope. The sampled signals were saved for off-line receiver DSP. The line rate of 1.28 Tb/s allows us to assume a 23% FEC-OH plus 5% framing OH. Such value would guarantee proprietary soft-decision FEC with 23% OH and a pre-FEC BER of $3.7 \times 10^{-2}$.

The control plane is performed via Linux-based C++ GMPLS controllers located at each physical node of the network. Controllers run PCEP establishing session with the domain active stateful cPCE and RSVP-TE to perform signaling. Controllers are connected with the data plane test-bed (i.e., SSS, TX and RX) by means of USB, serial and GPIB interfaces. Based on RSVP-TE messages, controllers are able to automatically configure SSS (i.e., filter shape as the reserved frequency slot), transponders parameters (i.e., symbol rate, number of carriers, sub-carrier central frequencies) and DSP parameters (i.e., modulation format, FEC).

Domain active stateful child PCE, developed in C++ in a Linux box, performs intra-domain topology export advertisement by means of BGP-LS, impairment-aware path computation [10], multi-action re-optimization, adaptation and instantiation by means of PCEP [11][12].

Multi-vendor interoperability between CNIT/Ericsson and Coriant DSPs was achieved by the exchange of novel Application Code (AC) and Transponder Class (TC) attributes in BGP-LS and PCEP/RSVP-TE protocols, in line with ITU recommendations [13]. The AC attribute, exported by BGP-LS within the TE link attribute extensions, defines the whole network scenario of application (i.e., in terms of wavelength range, type of fiber, dispersion compensation, presence of amplifiers, system rate, etc.). The TC attribute identifies the transponder features and compatibility by implicitly defining a large set of available and allowed physical tx/rx parameters values and ranges. In particular, the TC has been encoded as a novel subTLV specification inside the MF-OTP extensions [7] of PCEP and RSVP-TE specifying the super-channel description (i.e., in terms of sub-carriers nominal central frequencies and width, modulation, FEC). Path computation allows transponder-transponder end point assignment only if they both belong to the same TC or to compatible TCs. The same policy is enforced in the RSVP-TE signaling at the endpoint nodes, during the actual transponder selection and configuration.

C. Emulated Flexi-grid domain

The third domain, defined as (105), is composed of six nodes running a GMPLS control plane on dedicated Linux boxes. No data plane is employed: therefore, media layer devices (i.e. SSS) are emulated within the same boxes running the GMPLS controllers. An active stateful child PCE, enhanced with BGP-LS speaker features, collects local topology information and coordinates with the pPCE for path calculations and LSPs provisioning inside the domain. Persistent PCEP sessions between the cPCE and all the border nodes are configured to accomplish this.

D. Hierarchical Stateful PCE and ABNO Controller

The parent PCE (pPCE) is built using the open source Netphony Hierarchical Stateful PCE suite [14], developed in Java, and is deployed in a virtual machine that is connected to the control plane network. The pPCE uses the Netphony BGP-LS speaker and TED [15][16] to obtain the domain interconnection and the TE details of both Border Nodes and Inter-domain Links. Thus, when the pPCE is used for end-to-end computations, runs an algorithm to select the domains and performs a distributed RSA when end-to-end media channel is requested. When the inter-domain links are a transponder interconnection, the pPCE does not need to perform the RSA. A redis database is used to store persistently the LSP and Traffic Engineering databases. The ABNO controller, built with the Netphony Open Source ABNO suite [17], runs in the same virtual machine as the pPCE. A workflow is implemented that defines the interactions with the pPCE and can be triggered either from the Planning tool or from the Network Management System. A web server with the Graphical User Interface is also deployed to facilitate the operation of the network.

E. Planning tool

Upon the reception of a bundle of connections to be served, the PLATON planning tool solves an optimization problem so as to minimize the resources used for the whole bundle. To that end, the planning tool explores a fixed number of solutions by randomly sorting the bundle set. For each individual connection request, the planning tool solves the Routing Spectrum and Modulation Format Assignment (RSMA) problem (an extension of the RSA problem, where the modulation format is as well assigned)[18]. The RSMA problem can be stated as follows:

Given:

- a graph $G(N, E)$ representing the network topology. The set $N$ includes the set of optical nodes. The set of optical links $E$ connecting nodes in $N$, where each link $e$ is of a given length.
- the availability of every frequency slice (slice width denoted as $\Delta_s$) in the optical spectrum of every link in $E$;
- the set $T(n)$ of SBVT installed in every node $n$ in $N$. The characteristics of the SBVTs are known, including available capacity and sub-carrier modules;
- a set $F$ of modulation formats supported by the SBVTs,
where \( m(f) \) represents the spectrum efficiency of \( f \). \( F \) is ordered by \( m(f) \);

- a connection request to be served defined by the tuple \(<\text{src}, \text{dst}, \text{bw}>\), where \( \text{src} \) and \( \text{dst} \) are source and destination nodes in \( N \), and \( \text{bw} \) is the requested bitrate.

**Output:** the SBVTs and the RSMA for the request.

**Objective:** minimize the used resources.

Table I presents the pseudocode of the proposed algorithm to solve the RSMA problem. A set of shortest paths are computed between end nodes selecting one of the set SBVTs with enough available capacity and sub-carrier modules (line 1 in Table I); each path includes its physical route \( k \) (sequence of hops), and the width of the largest continuous slot in that route, \( n \). Each path is afterwards checked to verify the width of largest slot available (line 4). Next, the best modulation format is selected from set \( F \) provided that the reach works for the length of the route (lines 6-9). The set of paths satisfying the previous constraints, if any, is sorted first by the length of its route and the best path is selected (line 12). A slot of the proper width is selected (line 13) and the computed lightpath is eventually returned (line 14).

### Table I RMSA Algorithm

<table>
<thead>
<tr>
<th>IN: G(N, E), &lt;src, dst, bw&gt;</th>
<th>OUT: &lt;k, c, f&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ( Q = {k, n_k} \leftarrow \text{KSP}(G, &lt;\text{src}, \text{T}(\text{src}, \text{bw}), \text{dst}, \text{T}(\text{dst}, \text{bw}))&gt; )</td>
<td></td>
</tr>
<tr>
<td>2. ( Q' \leftarrow \emptyset )</td>
<td></td>
</tr>
<tr>
<td>3. for each ( q ) in ( Q ) do</td>
<td></td>
</tr>
<tr>
<td>4. if ( q.n \leq \text{minSlotWidth} ) then continue</td>
<td></td>
</tr>
<tr>
<td>5. ( q.f \leftarrow 0 )</td>
<td></td>
</tr>
<tr>
<td>6. for each ( f ) in ( F ) do</td>
<td></td>
</tr>
<tr>
<td>7. if ( \text{len}(q.k) \leq \text{len}(f) ) AND width(bw, m(f)) \leq q.n then</td>
<td></td>
</tr>
<tr>
<td>8. ( q.f \leftarrow f )</td>
<td></td>
</tr>
<tr>
<td>9. break</td>
<td></td>
</tr>
<tr>
<td>10. if ( q.f \neq 0 ) then ( Q' \leftarrow Q' \cup {q} )</td>
<td></td>
</tr>
<tr>
<td>11. if ( Q' = \emptyset ) then return ( \emptyset )</td>
<td></td>
</tr>
<tr>
<td>12. sort(Q', {q.k, ASC}); ( q \leftarrow \text{first}(Q') )</td>
<td></td>
</tr>
<tr>
<td>13. ( q.c \leftarrow \text{selectSlot}(k, \text{width}(bw, m(f))) )</td>
<td></td>
</tr>
<tr>
<td>14. return ( q )</td>
<td></td>
</tr>
</tbody>
</table>

IV. EXPERIMENTAL EVALUATION

To demonstrate the architecture, an end-to-end connection of 107Gb/s, with one segment in domain 102, and a second segment with a cross-vendor connection between S-BVTs in domain 101, is setup.

**A. Control Plane Evaluation**

A Wireshark capture of the PCEP messages flow of the computation and instantiation is shown in Fig. 7 with a total set-up time of 6.12 seconds. In the trace, it can be seen that the instantiation time is dominated by the hardware configuration, 2.29 seconds in domain 101 and 5.96 seconds in domain 102.

**B. ALU-BRISTOL Optical Performance**

The Flexi-grid domain (102) has been characterized in terms of optical performances for both SBVTs and off-line processing from Alcatel-Lucent (ALU) and University of Bristol (UNIVBRIS). Fig. 8 shows the resulting BER vs. OSNR curves. For ALU→UNIVBRIS, ALU’s transmitter is used as well as the off-line processing of UNIVBRIS at the receiver. Conversely, for UNIVBRIS→ALU, the SBVT transmitter of UNIVBRIS is used with the DSP design of ALU. As shown in Fig. 8, at the HD-FEC limit, a penalty as low as 1dB is observed in the BER curves, while the maximum penalty is observed below 2dB for the highest OSNR. This performance confirms an achievable interoperability in between different elastic SBVTs and different DSPs. Fig. 8 also includes DP-QPSK constellations for the back-to-back and after transmission through the two AoD nodes and 175km of standard single-mode fiber (SSMF). Even after optical transmission of a signal at total bitrate of 107Gb/s and configuration of Flexi-grid WSS with 37.5GHz filtering window, the constellation remains clear for both polarization states, showing the interoperability between the ALU SBVT transmitter and UNIVBRIS SBVT receiver. The switching time has also been measured in domain (102). At the restful server level, the time between the http request and the http reply, including complete configuration of all elements below the server, has been measured at 0.9s for the SBVT + OTN matrix, while it is ~4s for the AoD nodes.

**C. Multi-vendor Flexi-grid Interoperability Performance**

Fig.9 right reports the transmission performance of the single-vendor (CNITTX → CNITRX - SV) and cross-vendor (CNITTX → CORAX - CV), where COR stands for Coriant. Both solutions are deployed in the (101) domain at the maximum capacity (i.e., 1 Tb/s per super-channel). The aforementioned pre-FEC BER threshold of \( 3.37 \times 10^{-2} \) will be considered only for the SV scenario (i.e., a proprietary and more powerful SD-FEC can be adopted). In both cases we used blind-DSP algorithms [19], because an interoperability scenario would not allow the usage of algorithms that require knowledge of the link [20] or of training sequences [21].

In case we opt for the cross-vendor (CV) solution, we can rely only on already standardized FEC. For example, the ITU standard G975.1, appendix I.7 hard-decision super-FECstd (HD-FECstd) with 20% could use a pre-FEC BER threshold = \( 1 \times 10^{-4} \). This value could increase, in case new SD-FEC
become standard, up to a reasonable pre-FEC BER threshold = 2×10^{-2}. Based on the pre-FEC thresholds, we can draw the following conclusions concerning the experiment carried out by employing the test-bed depicted in Fig. 4. In case we assume the usage of already standardized HD-FECstd a transmission up to ~750 km in case of SV-transmission is guaranteed. This halves once we operate in cross-vendor-mode. Though, the system performance has been significantly reduced, such a scenario would still guarantee the error-free transmission, with CV-transmission, over the majority of European links. Moreover, if we could adopt a standardized transmission, with CV-transmission, over the majority of proprietary FEC, the transmission distances would be ~1800 km [22], and once again about half in case of CV transmission. These values are summarized in Table II.

### Table II Performance Results

<table>
<thead>
<tr>
<th>Case</th>
<th>HD-FECstd</th>
<th>SD-FECstd</th>
<th>SD-FECstd</th>
</tr>
</thead>
<tbody>
<tr>
<td>SV</td>
<td>750</td>
<td>1100</td>
<td>1800</td>
</tr>
<tr>
<td>CV</td>
<td>300</td>
<td>600</td>
<td>900</td>
</tr>
</tbody>
</table>

In addition to the results of the conducted experiment, we display in Fig. 9 also the ones obtained within a second experiment that has been reported in [19][20]. In this previous analysis, the vendor configuration was mirrored by having (CORTX → CORRX - SV) and cross-vendor (CORTX → CNTUX - CV). In [22][23] the channel configuration was slightly different (32 Gbd and 38 GHz channel spacing). The performance of this experiment is reported by the curves with filled markers in Fig. 9 and in this comparison, it is clearly visible that the two single vendor scenarios and two cross-vendor-ones, between them, achieve similar performance. From a DSP perspective, the degraded performance of CV could be associated to the lack of knowledge of system, such as for example TX / RX I/Q skew compensation. Such comparison confirms that the CV-transmission, with standard HD-FEC, can reach ~300km, and that therefore if longer distances are needed, a solution concerning standardization on SD-FEC must be agreed within the ITU panels.

### V. CONCLUSIONS

The paper has demonstrated a fully end-to-end interoperable EON network at control and data plane levels. The interoperability of several implementations of S-BVTs, with Hard-decision and Soft-decision FECs, is evaluated. The control architecture is able to configure the SBVTs so the multi-vendor transmission reach is failure-free up to 300 km with current FEC standards.

### REFERENCES