CCAMP Working Group
Internet Draft
Intended status: Standards Track

V.Beeram
I.Bryskin
W.Doonan
ADVA Optical Networking
J.Drake
G.Grammel
Juniper Networks
Manuel Paul
Ruediger Kunze
Deutsche Telekom

Expires: April 21, 2012
October 21, 2011

# GMPLS-UNI BCP
# draft-beeram-ccamp-gmpls-uni-bcp-00.txt

## Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html

This Internet-Draft will expire on April 21, 2012.

## Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (http://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

## Abstract

This document pools together the best current practices that are being used to apply the GMPLS Overlay model at the User-Network Interface (UNI) reference point (as defined in [G.8080])

## Table of Contents

1. Introduction.....2
2. Multi-Layered Approach.....3
3. Traffic Engineering.....6
   3.1. Augmenting the Client-Layer Topology.....9
      3.1.1. Virtual TE Links.....11
   3.2. Macro SRLGs.....13
   3.3. Switching Constraints.....14
4. Connection Setup.....16
5. Security Considerations.....18
6. Normative References.....18
7. Acknowledgments.....18

## 1. Introduction

Generalized Multiprotocol Label Switching (GMPLS) provides tools to create end-to-end services in various transport technologies. These tools can be used to support service management in different types of deployment models. RFC 4208 discusses how GMPLS can be applied to the overlay model. There are a good number of implementations that have built on the basic concepts discussed in RFC 4208 and have successfully demonstrated interoperability. This document is an attempt to pool together the best current practices that are being

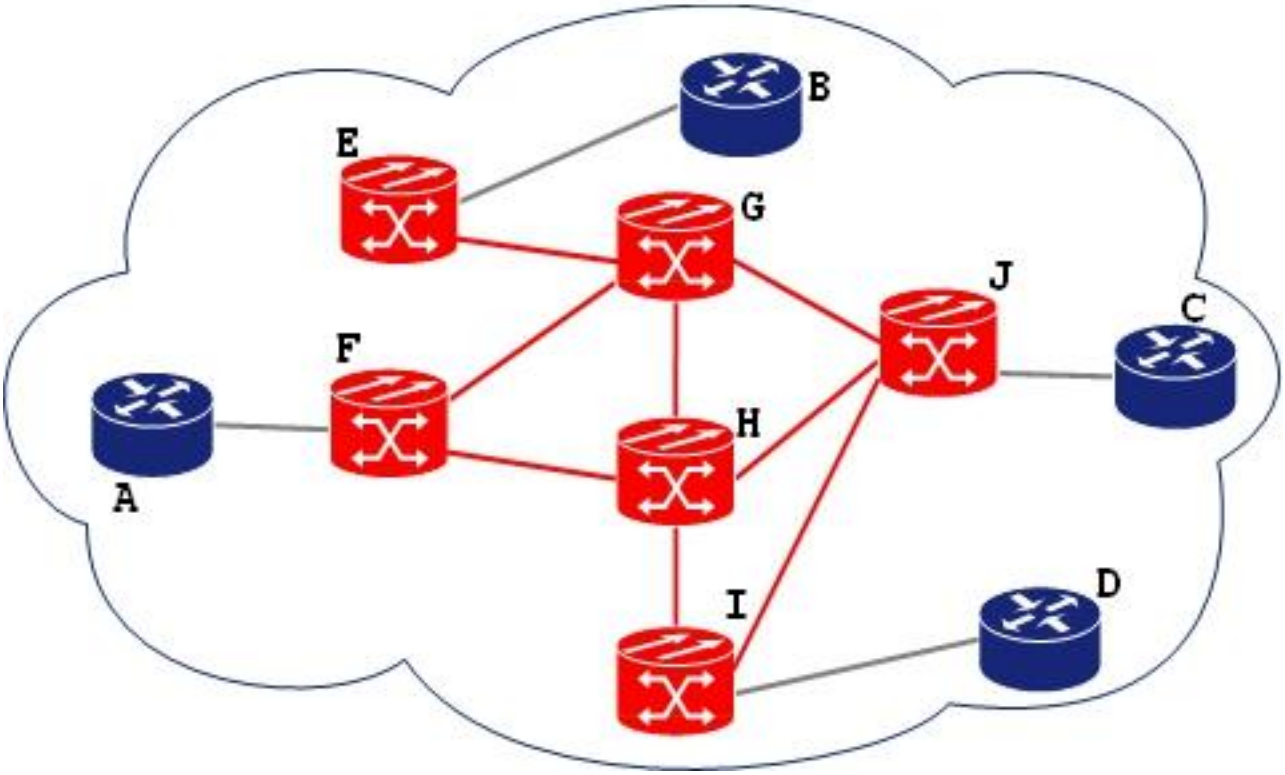
used to apply the GMPLS Overlay model at the User-Network Interface (UNI) reference point (as defined in [G.8080]).

RFC 4208 recommends the use of hierarchical service activation when GMPLS is used for the core network - this document takes this concept further, discusses the notion of "augmenting the client-layer TE topology" and explains how this augmentation enables client-layer networking in an overlay model. The concepts discussed in this document are based primarily on experiences drawn from interoperating GMPLS-enabled IP routers with Optical Transport elements.

## 2. Multi-Layered Approach

When an end-to-end service crosses a boundary between two regions of dissimilar transport technology, it is necessary to execute distinct forms of service activation within each region.

Fig 1: Sample Hybrid Topology




For example, in the hybrid network illustrated in Fig 1, provisioning a transport service between two GMPLS-enabled IP routers on either side of the optical WDM transport topology requires operations in two distinct layer networks; the client-layer network interconnecting the routers themselves, and the server-layer network interconnecting the optical transport elements in between the routers.

Activation of the end-to-end service begins with a path determination process, followed by the initiation of a signaling process from the ingress along the determined path, per the set of figures shown in Fig2.

**Fig 2: Hierarchical Service Activation**

**Fig 2a: Client-Layer Service setup is initiated**

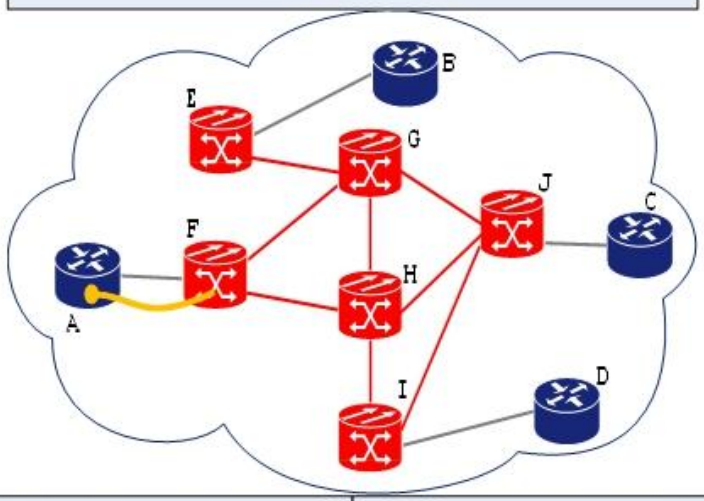



**Fig 2b: Server-Layer WDM Service that caters to the client-layer service is established within the core**

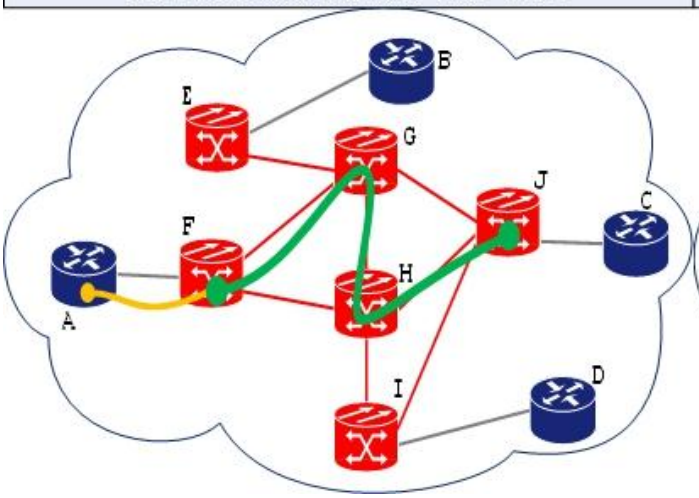
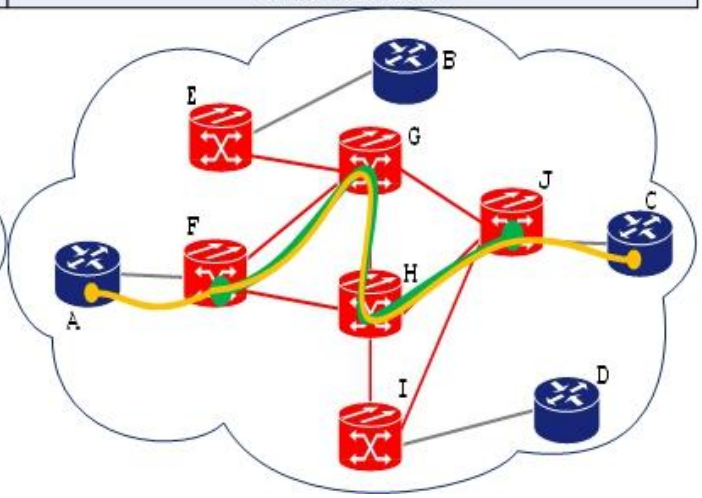



**Fig 2c: Client-Layer Service setup is resumed and the end-to-end connection is established**



## 3. Traffic Engineering

The previous section outlines the basic method for activating end-to-end services across a multi-layer network. As a necessary part of that process an initial path selection process was performed, whereby an appropriate path between the desired endpoints was determined through some means. Further, per expectations set through current practices with regard to service provisioning in homogeneous networks, operators expect that the underlying control plane system will provide automated mechanisms for computing the desired path or paths between network endpoints.

In particular, operators do not expect under normal circumstances to be required to explicitly specify the end-to-end path; rather, operators expect to be able to specify just the endpoints of the path and rely on an automated computational process to identify and qualify all the elements and links on the path between them. Hence when operating a hybrid network such as that described in Fig 1, it is necessary to extend existing traffic engineering and path computation mechanisms to operate in a similar manner.

Path computation and qualification operations occur at the path computation element (PCE) selected by ingress element of an end-to-end service. In order to be able to compute and qualify paths, the PCE must be provided with information regarding the traffic engineering capabilities of the layer network to which it is associated with, in particular what the topology of the layer network is and what layer-specific transport capabilities exist at the various nodes and links in that topology.

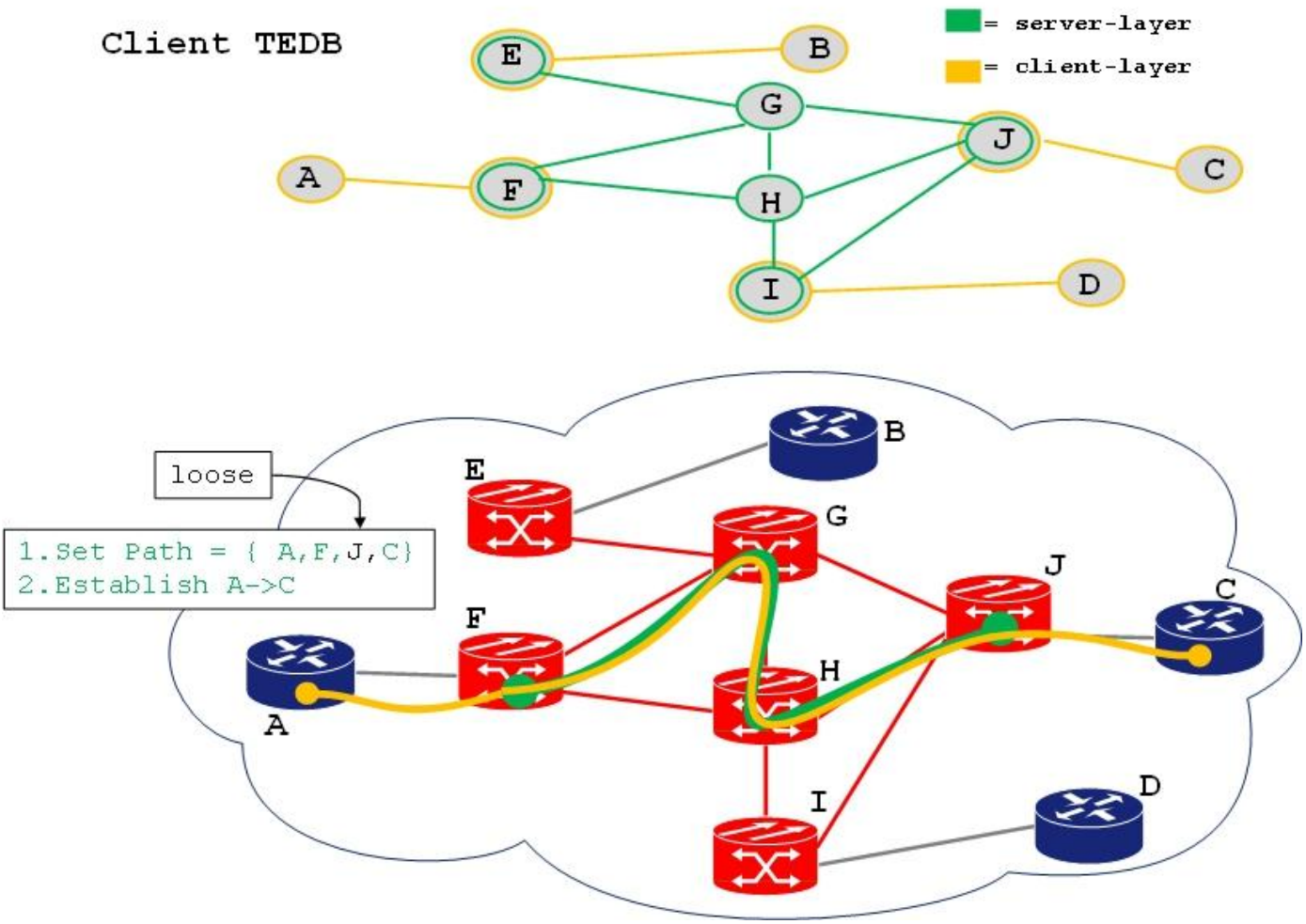
It is important to note that topology information is layer-specific; e.g. path computation and qualification operations occur within a given layer, and hence information about topology and resource availability are required for the specific layer to which the connection belongs to. The topology and resource availability information required by elements in the client-layer is quite distinct from that required by the elements in the server-layer network. Hence, the server-layer traffic engineering links are of no importance for the client-layer network, and it is actually desirable to block their advertisements into the client TE domain by the server-layer border nodes.

In the sample hybrid network (Fig 1) there are multiple optical transport elements supporting the connection between the GMPLS-

enabled IP routers, and hence the physical topology between them includes several nodes and links. However, the optical elements between the IP routers are not able to switch traffic within the client-layer network of routers (e.g. IP/MPLS), as the optical elements are lambda switches, not IP/MPLS switches. Hence while the intervening optical elements may physically exist along the path, they are not a part of the topology available to the IP/MPLS routers for the purposes of traffic engineering in the client-layer network.

An example of what the client-layer Traffic Engineering topology would look like for the sample hybrid network is shown in the top half of Fig 3.

**Fig 3:Traffic Engineering - ERO with "loose hop"**




In this example, the TE topology associated with the client-layer network is indicated by the links and nodes colored yellow, whereas the TE topology associated with the server-layer network is indicated by the links and nodes colored green. The border nodes at the core are visible in both the topologies. The yellow topology is capable of switching traffic within the client-layer, whereas the green topology is capable of switching traffic within the server-layer.

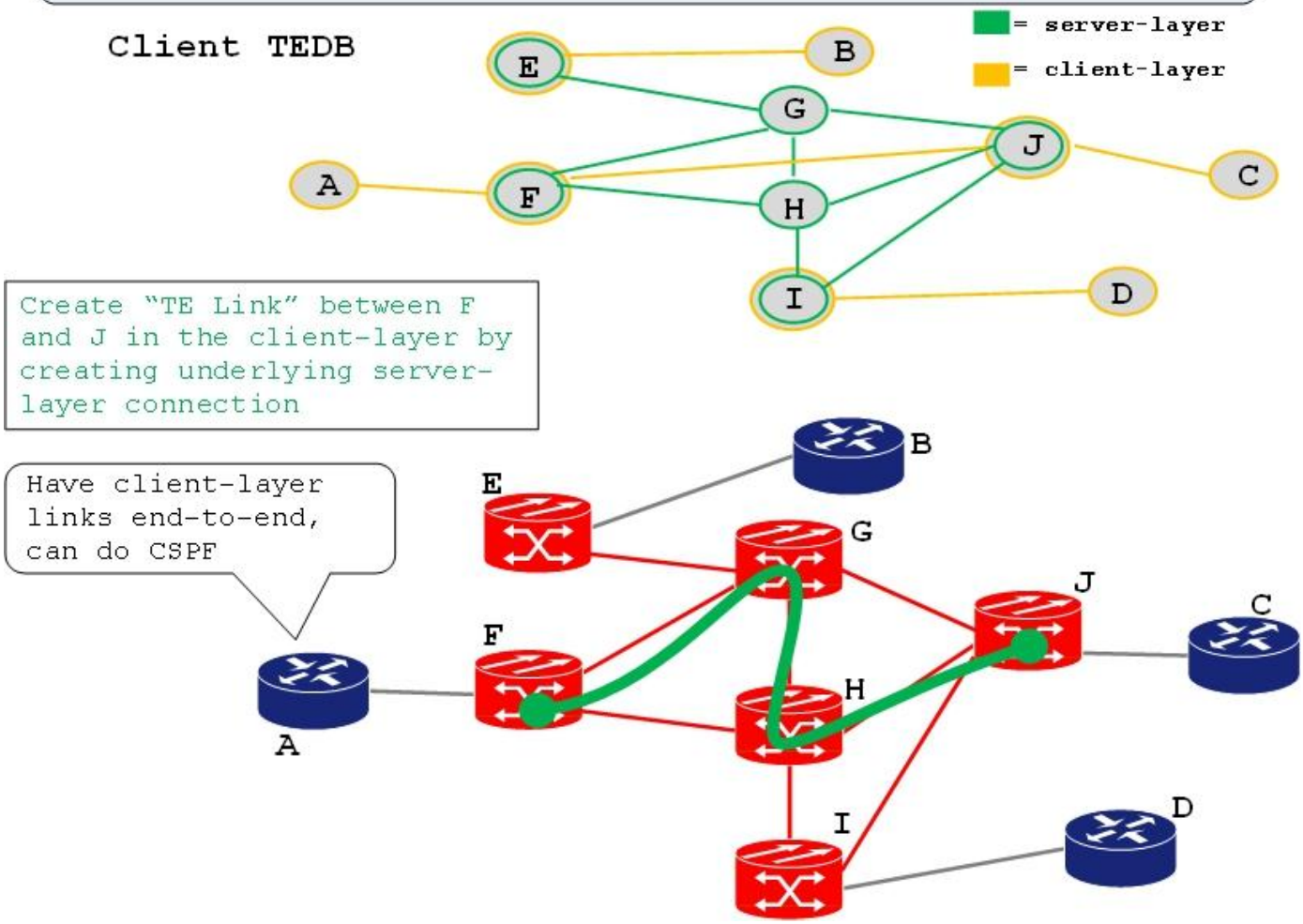
In this example, if the “B” router attempts to determine a path to the “D” router it will be unable to do so, as the yellow topology to which the B and D routers is connected does not include a fully-yellow path between them. The only way to setup an end-to-end path in this case is to use an ERO with a “loose hop” across the server-layer domain as illustrated in Fig 3. This would cause the server-layer to create the necessary segment of the client-layer topology on the fly. However, this approach has a few drawbacks - [a] the necessity for the operator to specify the ERO with the “loose” hop; [b] potential sub-optimal usage of server-layer network resources; and [c] unpredictability with regard to the fate-sharing of the new segment (that is created on the fly) with other links of the client-layer topology.

In order to be able to compute a full path between the two client-layer endpoints, the yellow topology must be sufficiently augmented to indicate where there are paths through the green topology which can provide transport to services in the yellow topology. In other words, in order for a client to compute path(s) across the server-layer domain to other clients, the segment of the client-layer topology over the server-layer domain should be pre-planned and made available (in terms of TE links and nodes that exist in the client-layer) to all the clients. This is discussed in detail in the next section.

## 3.1. Augmenting the Client-Layer Topology

In the example hybrid network, consider a scenario where each GMPLS-enabled IP router is connected to the optical WDM transport network via a transponder. Further consider the situation where the transponder at node F can be connected to the transponder in node J via the optical pathway F-G-H-J. A WDM connection can be provisioned in the server-layer along this path, and then advertised as a TE link in the client-layer. With the availability of this link, the path computation function at node A is able to compute an end-to-end path from A to C.

**Fig 4: Traffic Engineering - End to End Path Computation**




In this case, in order for the TE link to be made available in the client-layer topology, the network resources corresponding to the underlying server-layer connection must be fully provisioned beforehand.

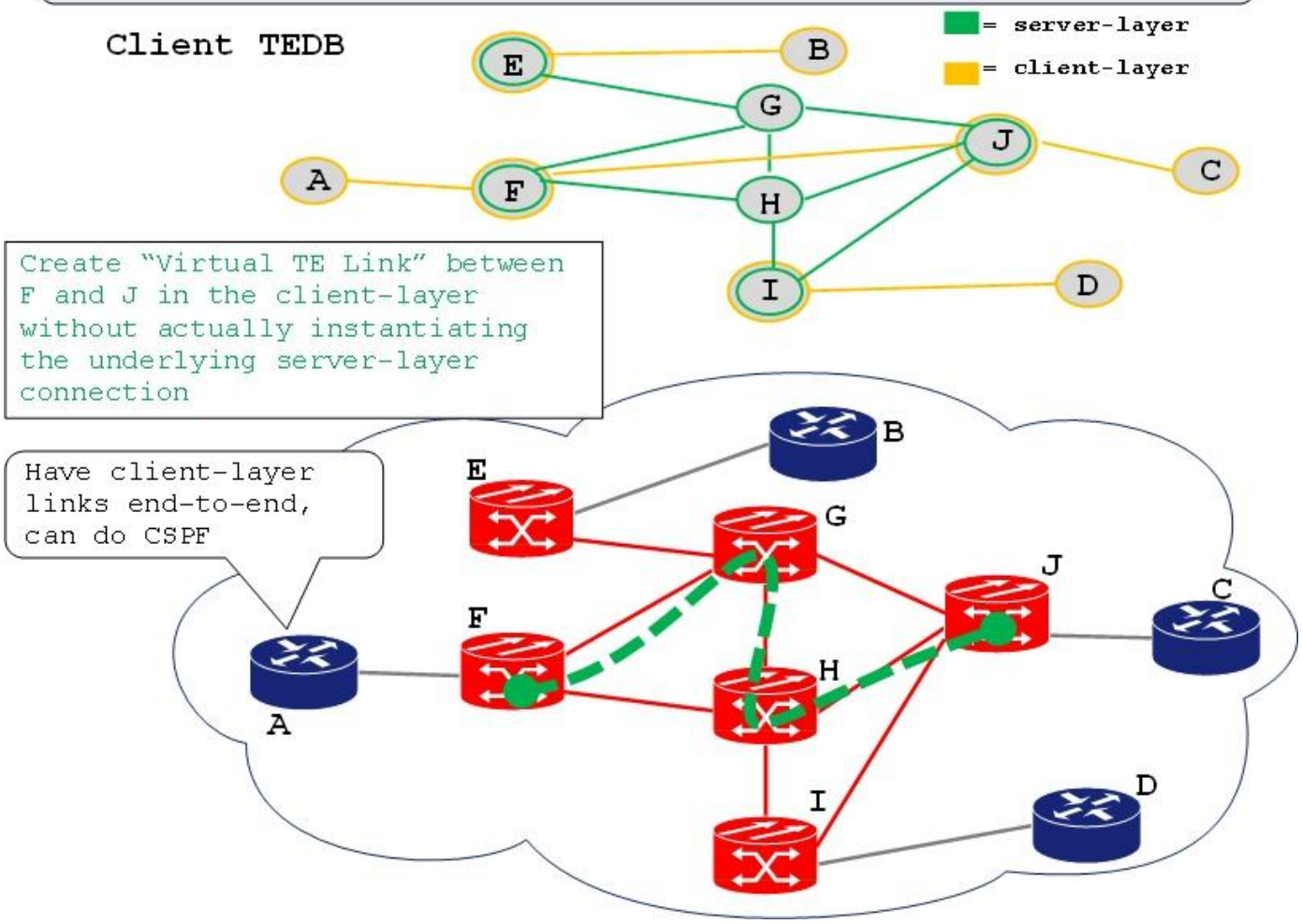
As another example, consider a network configuration where the transponders at nodes E, F, J and I are connected to each other via directionless ROADM components. It is physically possible to

connect any transponder to any other transponder in the network. As there are transport capabilities available in the server-layer network between every element containing an adaptation function to the client-layer network, the operator in this case would not wish to reserve any network resources in the server-layer until the setup of the client-layer connection is initiated. The next section proposes a method to cater to this particular operational requirement.

### 3.1.1. Virtual TE Links

A “Virtual TE Link” is defined as a TE link that is advertised into the client-layer, with available but unreserved resources in the server-layer necessary to bring up the connection that supports that TE link. In other words, “Virtual TE Links” represent specific transport capabilities available in the server-layer network which can support services in the client-layer network. The two fundamental properties of a Virtual TE Link are: [a] It is advertised just like a real TE link and thus contributes to the buildup of the client-layer topology (and thus client-layer elements see no difference between virtual and real links); and [b] It does not require allocation of resources at the server-layer until used, thus allowing the sharing of server-layer resources with other virtual TE links.

**Fig 5: Traffic Engineering - End to End Path Computation (w/ "Virtual TE Links")**




In the example shown in Fig 5, the availability of a lambda channel along the path F-G-H-J results in there being a virtual traffic engineering link between F and J within the client-layer topology (yellow line). With the availability of this link, the path computation function at node A is able to compute an end-to-end path from A to C.

When a virtual TE link gets selected and signaled in the ERO of the client-layer connection, it ceases to be “Virtual” and transforms into a regular TE-link. When this transformation takes place, the clients will notice the change in the advertised available bandwidth of this TE-link. Also, all other Virtual TE links that share resources with the TE-link in question start advertising “zero” available bandwidth. Likewise, the TE network image reverts back to the original form as soon as the last client-layer connection, going through the TE link in question, is released.

### 3.2. Macro SRLGs

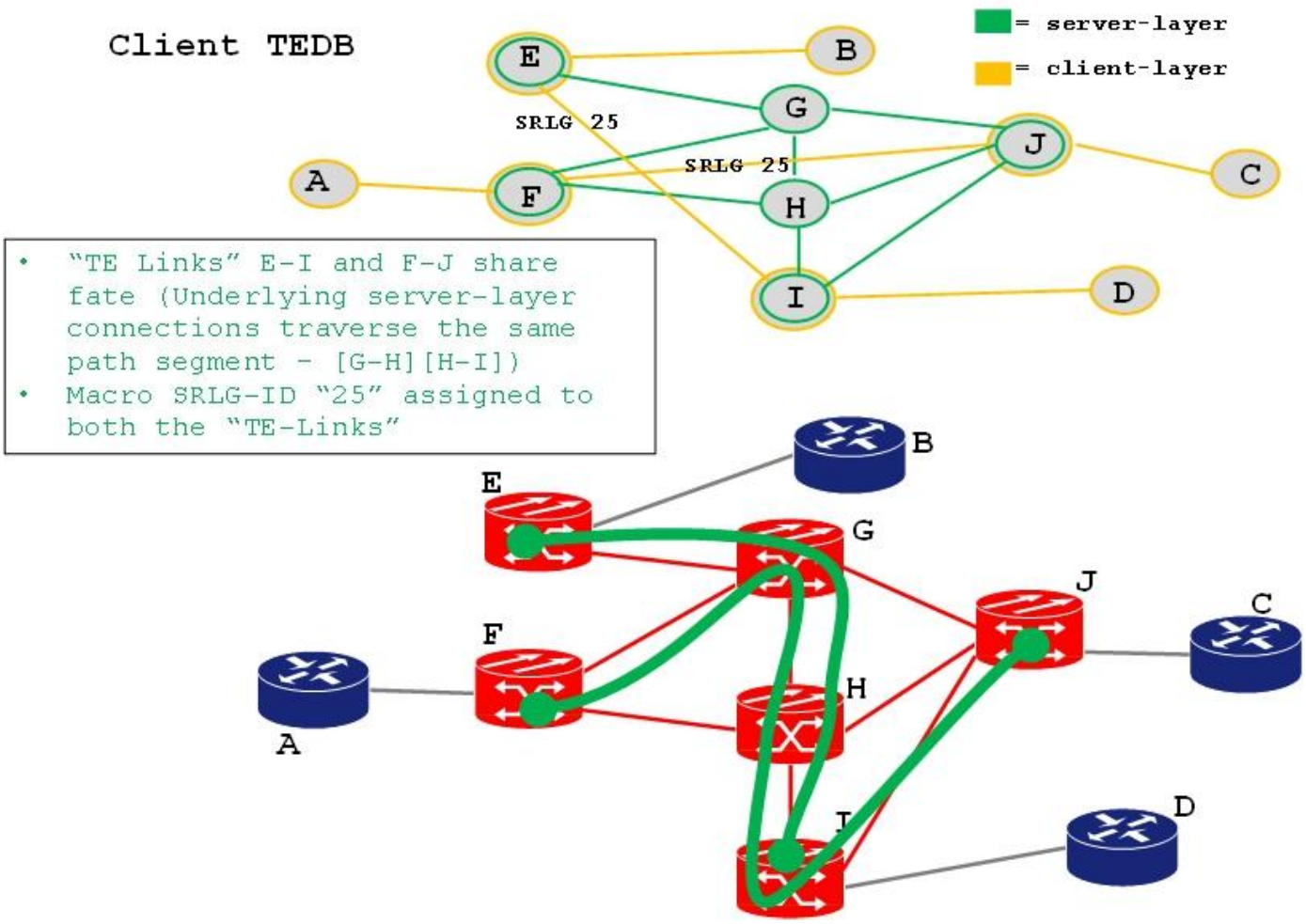
The TE links that are added to the client-layer topology cannot be assumed to be totally independent. It is quite possible for a given TE link to share the same fate with one or more other TE link(s). This is because the underlying server-layer connections (real or potential) can traverse the same server-layer link and/or node, and failure of any such shared link/node would make all such connections inoperable (along with the client-layer links they serve). If diverse end-to-end client-layer connections are to be computed, the fate-sharing information of the TE links needs to be accounted for. The standard way of addressing this problem is to use SRLGs as a part of TE link advertisements.

A traditional SRLG represents a shared physical network resource upon which normal function of a link depends. Such SRLGs can also be referred to as physical SRLGs. Zero, one or more physical SRLGs could be identified and advertised for every TE link in a given layer network. However, there is a scalability issue with physical SRLGs in multi-layer environments. For example, if a WDM layer connection serves an IP layer link, every WDM link and node traversed by the connection must be considered as a separate SRLG. The number of SRLGs to be advertised to client (e.g. IP) layer per link would be directly proportional to the number of hops traversed by the underlying server-layer connection.

The notion of Macro SRLGs addresses this scaling problem. Macro SRLGs have the same protocol format as that of their physical counterpart and can be assigned automatically for each TE link that is advertised into the client-layer as a result of the existence of an underlying server-layer connection (instantiated or otherwise). A Macro SRLG represents a set of shared path segments that are traversed by two or more of the underlying server-layer connections. Each shared path segment can be viewed as a sequence of shared

resources where each individual resource has a physical SRLG associated with it (example depicted in Fig 6). The actual procedure for deriving these Macro SRLGs is beyond the scope of this document.

**Fig 6: Macro SRLGs**

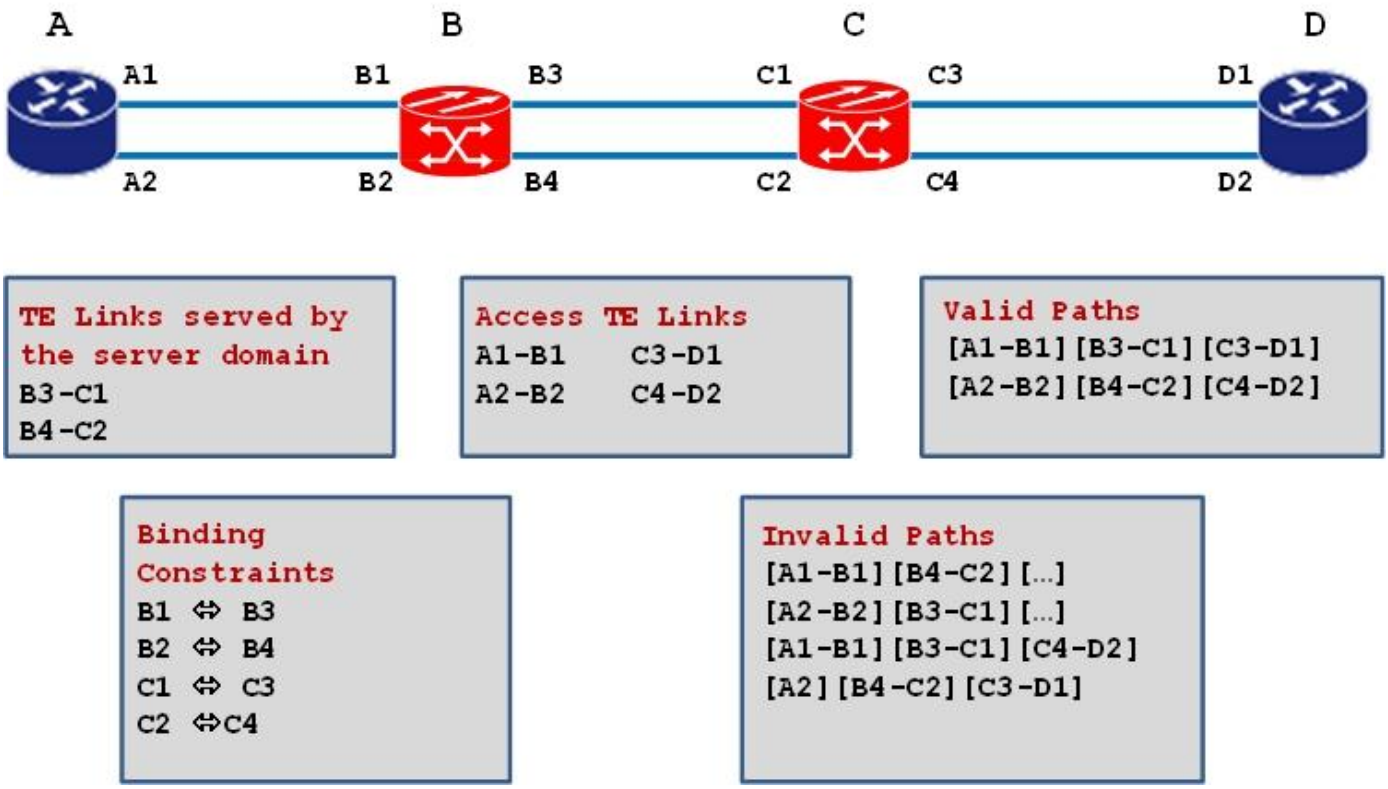



## 3.3. Switching Constraints

Certain types of optical network configurations necessitate the specification of connectivity constraints in the TE advertisements.

If the switching constraints associated with the binding between the TE link served by the server domain and its associated access TE link do not get advertised, there is a risk of an invalid path being picked (Fig 7). This document recommends the use of the extensions specified in [GEN_CNSTR] to address this.

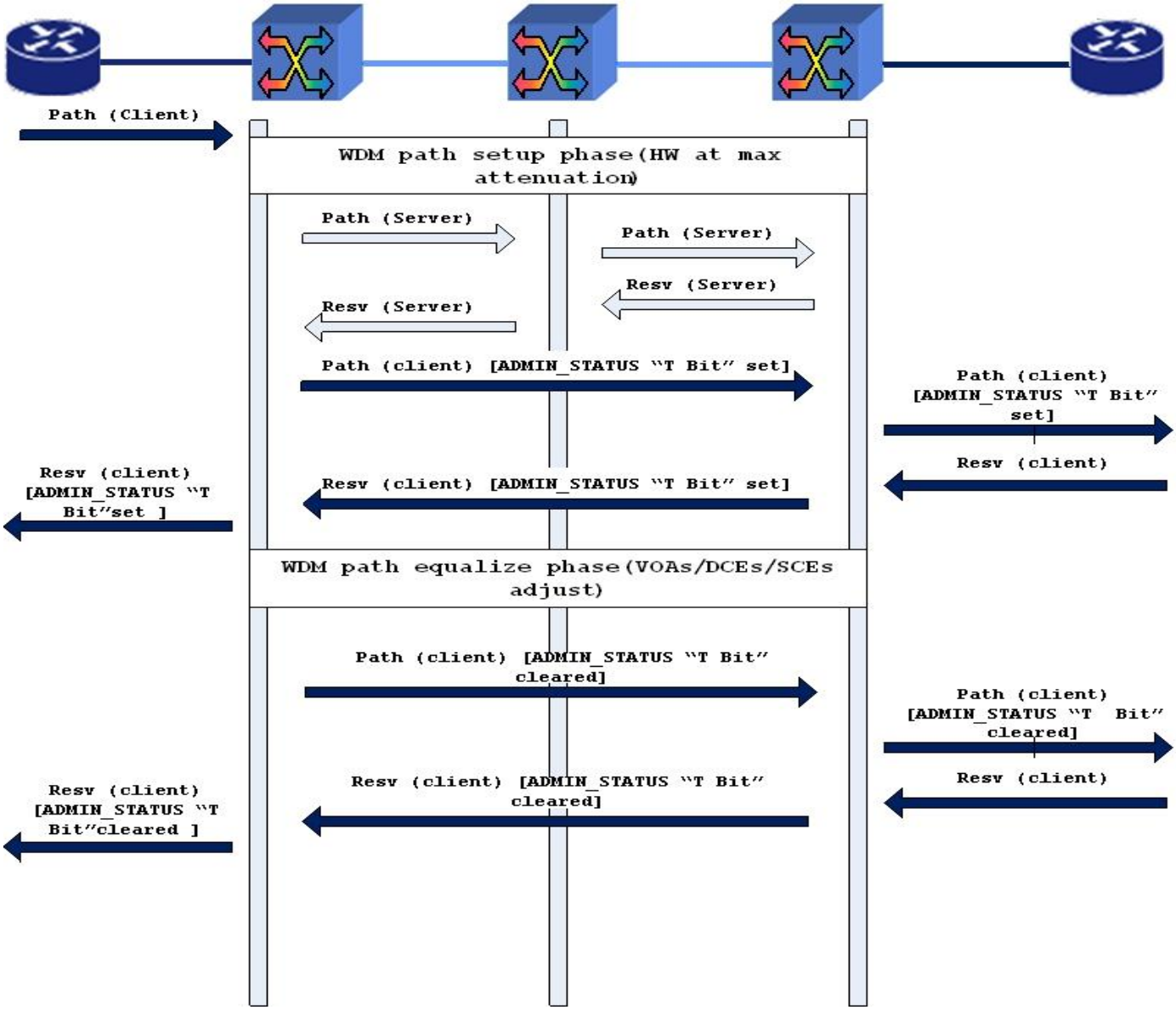
**Fig 7: Switching Constraints**



## 4. Connection Setup

Experience with control plane operations in multi-layer networks indicates there are benefits to coordinating certain signaling operations, in the following manner. Consider the scenario where the core is a WDM network comprising of ROADMs. The set-up time for a service at the optical layer can be fairly long, as it can involve time-consuming power-equalization procedures, amongst other layer-specific operations. This means that at very least, the setup timers for the client-layer service would need to be somehow coordinated with that of the server-layer service. To avoid this operationally awkward issue, a phased connection setup process as depicted in Fig 8 is proposed.

Fig 8: Phased Connection Setup




As long as the server-layer connection is not completely "UP" (for example: Fully Power Equalized), the nodes at the edge of the core would signal the client-layer PATH/RESV messages with the T

(Testing) bit set in the ADMIN_STATUS. The T bit would be cleared in these messages only after the underlying server-layer connection is deemed fully operable.

## 5. Security Considerations

TBD

## 6. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC4208] G. Swallow, J.Drake, H. Ishimatsu, and Y. Rekhter, “GMPLS UNI: RSVP-TE Support for the Overlay Model”, RFC 4208, October 2005.

## 7. Acknowledgments

TBD

## Authors’ Addresses

Vishnu Pavan Beeram
ADVA Optical Networking

Email: vbeeram@advaoptical.com

Igor Bryskin
ADVA Optical Networking

Email: ibryskin@advaoptical.com

Wes Doonan
ADVA Optical Networking

Email: wdoonan@advaoptical.com

John Drake
Juniper Networks

Email: jdrake@juniper.net

Gert Grammel
Juniper Networks

Email: ggrammel@juniper.net

Manuel Paul
Deutsche Telekom

Email: Manuel.Paul@telekom.de

Ruediger Kunze
Deutsche Telekom

Email: Ruediger.Kunze@telekom.de