Multi-Region Networks (MRN)
Evolution toward a GMPLS Integrated Control Plane

IP + OPTICAL Network Conference 2005

21-22 February 2005, Tokyo, Japan

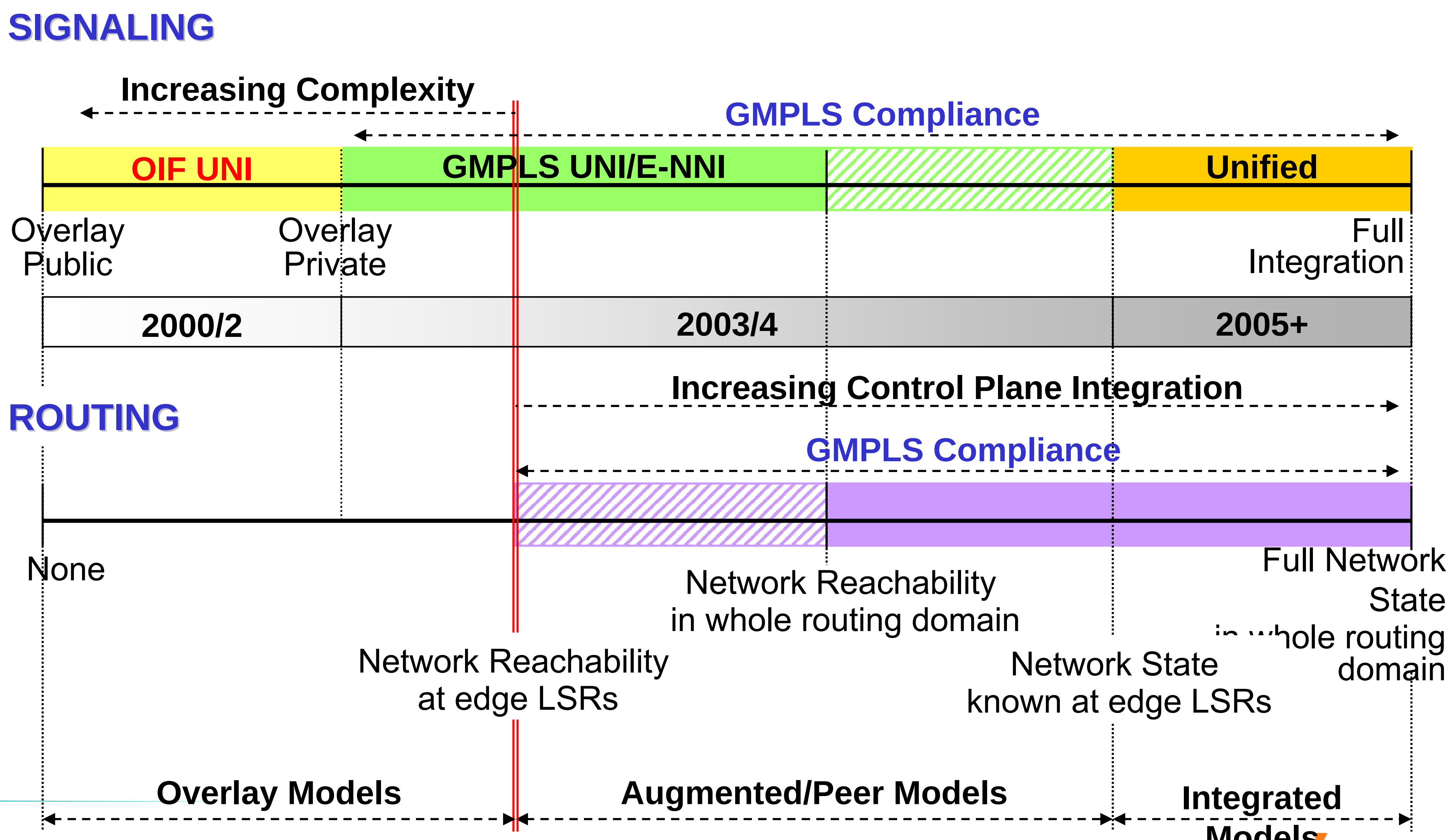
Dimitri PAPADIMITRIOU

ALCATEL

Introduction

- Generalized Multi-Protocol Label Switching (GMPLS) with unified signaling and integrating routing approach for resources provisioning and recovery ⇒ control-plane optimization for multiple switching layer networks
- Targeted network architectures referred to as vertical integration: nodes hosting multiple switching layers and controlled by a single instance of the control plane
- When such nodes are port of the same routing area the resulting network is referred to as a Multi (LSP-)Region Network (MRN)
- This approach differs from horizontal integration related to interworking between partitions (routing areas, autonomous systems) of the network

GMPLS Evolution and Control Plane Interconnection



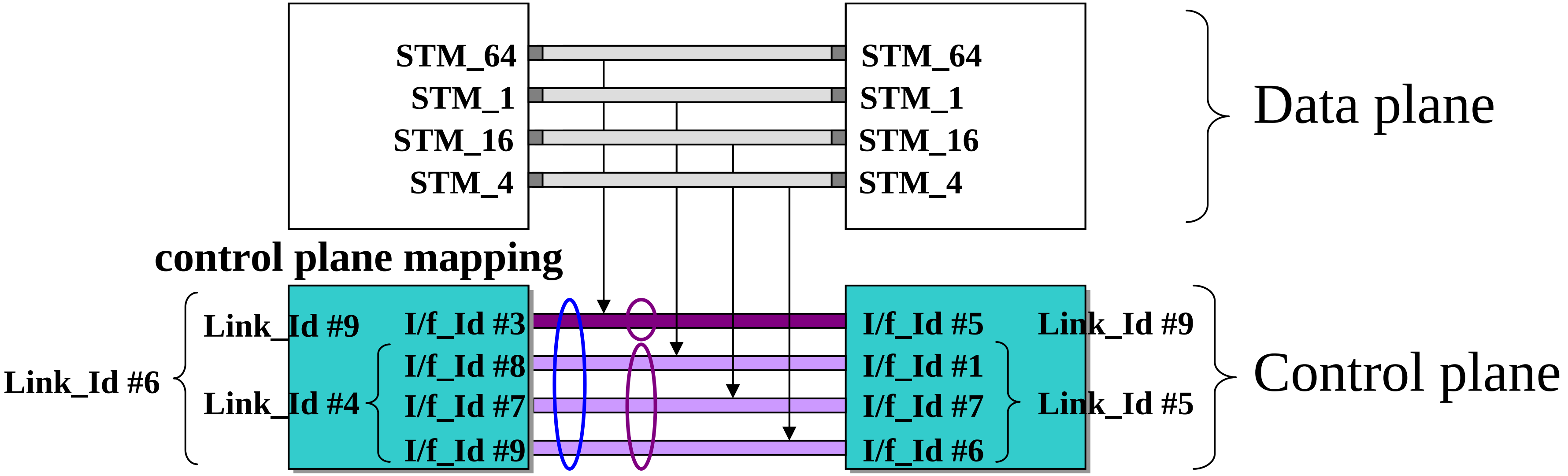
LSP Region vs Data Plane Layer

> (data plane) layer ≠ (control plane) LSP region
- Data plane layer: defined in ITU-T G.805 (G.803, G.872), IEEE 802.3, etc.
- LSP Region: defined in [LSP-Hierarchy]

> LSP region is a technology domain (identified by the Switching Capability) for which data plane resources (= data links) are represented into the control plane as an aggregate of TE information associated to a set of links (= TE links)
- Example: VC-11 to VC4-64c capable TE links are part of the same TDM Region

> When the control plane is able to manage the resources of multiple of such technology domains ⇒ unified TE model
- Example: control plane instance managing the packet (PSC) and SDH/Sonet circuit (TDM) LSP regions

Data Link and TE Link Concepts



> **Step 1**: associate to each data plane link a data link identifier [local interface_id; remote interface_id] in the control plane

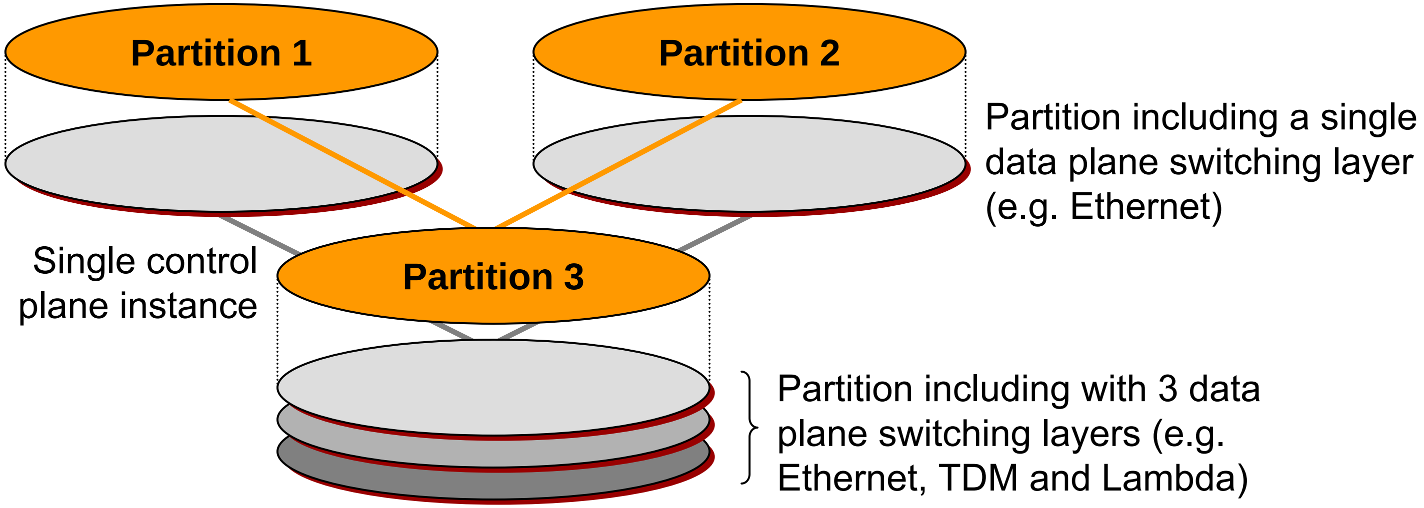
> **Step 2**: TE link that populates the TEDB are identified as [local link_id; remote link_id] and can include a single data link (= single component) or a set of data links (= multiple components)

> **Step 3**: TE link can comprise a set of TE links (recursive), each of these TE links as defined in Step 2

Vertical vs Horizontal Integration

> Vertical Integration: collaborative mechanisms within a single control plane instance driving multiple switching capabilities hosted by a single system

> Horizontal Integration: each entity constituting the network environment includes at least one common (data plane) switching layer and the control plane topology extends over several partitions (routing systems) being either areas or autonomous systems



Overlay Model Limitations (1)

> Separate routing instances between technology domains implies
  - No reachability information ($\Leftrightarrow$ address allocation/resolution)
  - No a priori routing adjacencies (b/w source and destination client)

  $\Rightarrow$ Need for an out-of-band mechanism to bootstrap the system

> Per layer Traffic Engineering
  - Path computation performed per layer using separate TE Database
  - Only manual triggered connection creation/modification

    $\Rightarrow$ default operational model is provisioned
  - State of the client connection is unknown within the network (no distinction b/w a protecting versus protected connection $\Rightarrow$ no restoration)

> Strict separation of the signalling domains, in turn requires split of (end-to-end) sessions for a single end-to-end connection

$\Rightarrow$ This leads to limitations in using overlay models (in turn this limits the gain from using GMPLS-based control plane)

Overlay Model Limitations (2)

- Client routing adjacencies are "unknown" ⇒ precludes dynamic connection triggering (or requires out-of-band messaging b/w edges), hence, only suitable for pre-provisioned (server layer) approach
  - ⇒ Bandwidth on demand applicability restricted to sub-channels (logical or physical)
- Server layer unaware of client-initiated connection status
  - ⇒ Usually part of client layer control plane topology (thus client control plane performance are strongly correlated to the data plane performance)
  - ⇒ Impossible to provide soft-reservation of network capacity (thus precludes client control-plane driven shared recovery mechanism)
- Network address allocation (network to client) does not allow for dynamic client learning of reachable end-points
  - ⇒ Requires address resolution mechanisms (logical to physical) for switched connections or (physical to logical) for soft-permanent connections (in turn, this precludes best exit point selection for multi-homed clients)
- Other issues: multi-homing, session split, recovery timing efficiency

ALCATEL

Augmented Model Limitations

> The augmented model is "intermediate" between the overlay and the unified model; it allows for
  - Exchange of a limited amount of routing information (mainly reachability) between the client and the server layer's network control plane
  - (but does not mandate) using different addressing spaces and full or partial opacity of the server layer addressing space
    - When applied without a common addressing between the IP/MPLS and optical network, solves the "address resolution" problem only (through exchange of reachable endpoint addresses) but does not solve the "unknown adjacency" problem because adjacencies between client and network edge nodes are not used for carrying client link state information (i.e., client layer routing information is still transported across the server layer's data plane)
    - When applied with a common addressing a tighter integration of the respective routing instances is provided (in that case, restricting exchange of information to reachability boils down to convey a restricted amount of topological information)

> Augmented model is appropriate for horizontal integration within large single carrier corporations that contain multiple Autonomous Systems (AS)

ALCATEL

Unified Model (1)

- Client and server domain act as peers (a.k.a. peer model)
  - Common control plane addressing space and separation between control and transport plane addressing space is still possible (unnumbered TE links)
  - Single signaling and routing protocol instance (routing topology, reachability, topology and resource/status information)
  - “Connections” viewed as Forwarding Adjacency LSPs (FA-LSP) and advertised as FA links (a.k.a virtual TE link since no associated routing adjacency)
- Unified services model (based on MPLS-TE principles generalization)
  - Network elements and services they deliver are treated alike (e.g. optical node ≡ LSR with optical interfaces)
  - System appears as one single network offering connection-oriented services from packet to optical Label Switched Paths (LSPs)
- Link-state routing protocol (OSPF/IS-IS) with traffic engineering extensions provides the required information allowing for
  - Distributed on-line path computation (non co-located on-line computation not precluded)
  - End to end LSP establishment/modification between edge IP/MPLS routers using GMPLS signaling protocol (RSVP-TE)

Unified Model (2)

> Generic approach provides:

- Automation: typically operations without human intervention, sub-classification (Real Time, non-Real Time)
- Distribution (control/transport separation): per node control (entity) and per function (signalling mandatory, routing, etc.) also distribution ratios can be defined
- Network Resource Optimization: using optical Traffic Engineering (routing) attributes and mechanisms
  - First to deliver optimised multi-region traffic engineering and resource management and delivering multi-region recovery capabilities (and diversely routed connection without impacting signaling protocol)
  - As opposed to client/server models, here packet-oriented equipment can now decide how bandwidth is allocated and connections routed within optical networks (constraint-based routing)

Signaling Models

Signaling Model	Public UNI (OIF)	GMPLS UNI	GMPLS E-NNI	Unified Signaling
Service Invocation	Direct and Indirect	Direct	Direct	Direct
Symmetry/ scope	Asymmetrical/ Local (Client/ Server)	Asymmetrical/ End-to-end (Peer)	Symmetrical/ End-to-end (Peer)	Symmetrical/ End-to-end (Peer)
Routing protocol	None	Optional	Mandatory	Link state preferred
Routing information	None	Optionally end-point reachability information (filtering and summarization)	At least end-point reachability information (filtering and summarization)	End-point and internal reachability information and TE attributes
Address space	Must be distinct	May be common	May be common	Common
Discovery	Optional and only local	Optional and may be network-wide	Optional and may be network-wide	Through routing and network-wide
Security	No trust	Limited trust	Limited trust	High trust
	Signaling must be domain specific (a separate signaling protocol instance must be running in the network)	End-to-end signaling that may be domain specific (a separate signaling protocol instance may be running in the network)	End-to-end signaling that may be domain specific (a separate signaling protocol instance may be running in the network)	Inherently multi-layer so is also referred to as end-to-end integrated signaling

Note: Unified signaling supports GMPLS UNI/E-NNI capabilities, and GMPLS UNI/E-NNI supports Public UNI capabilities

ALCATEL

Beyond the GMPLS Unified Model

- Within the peer model, no restriction implying exact matching of LSP regions and routing systems resulting from network partitioning (horizontal integration)
- However, interactions between control planes are mainly seen as interactions between systems hosting a single switching layer with limited coupling from routing information exchange:
  - With different flavors of edge/core nodes relationship (GMPLS UNI)
  - Between areas/levels and/or autonomous systems (multi-area, multi-AS)
- This does not cover current networks evolution towards delivery of multi-service over a multi-switching network (e.g., SDH + Ethernet)
- Therefore, GMPLS models should cover the vertical integration case dealing with multiple switching capable nodes resulting in multi-region networks (MRN)
- Examples: PSC+LSC (e.g. IP/MPLS + Lambda), L2SC+TDM (e.g. ETH + SDH), PSC + L2SC + LSC, etc.

MRN – Control Plane Model

> MRN Control Plane: Unified Signaling + Integrated Routing

	GMPLS	MRN Specifics
Architecture	Single Switching Capable nodes (PSC, L2, TDM, LSC)	+ Multiple Switching Capable nodes
Signaling	RSVP-TE (single and multi-area, multi-AS)	+ Multi-region RSVP-TE signaling (e.g., extension to explicit routing mechanisms)
Routing	OSPF(-TE), IS-IS(-TE): single-/multi-area BGP (Multi-AS) case	+ Multi-region TE attributes (e.g., adaptation) no extensions to BGP
Survivability	End-to-end / segment Pre-planned / Dynamic LSP re-routing	+ Multi-region recovery (not discussed)
Traffic Engineering	Single region TE attributes default TE metric inheritance	+ Multi-region TE attributes TE metric inheritance

ALCATEL

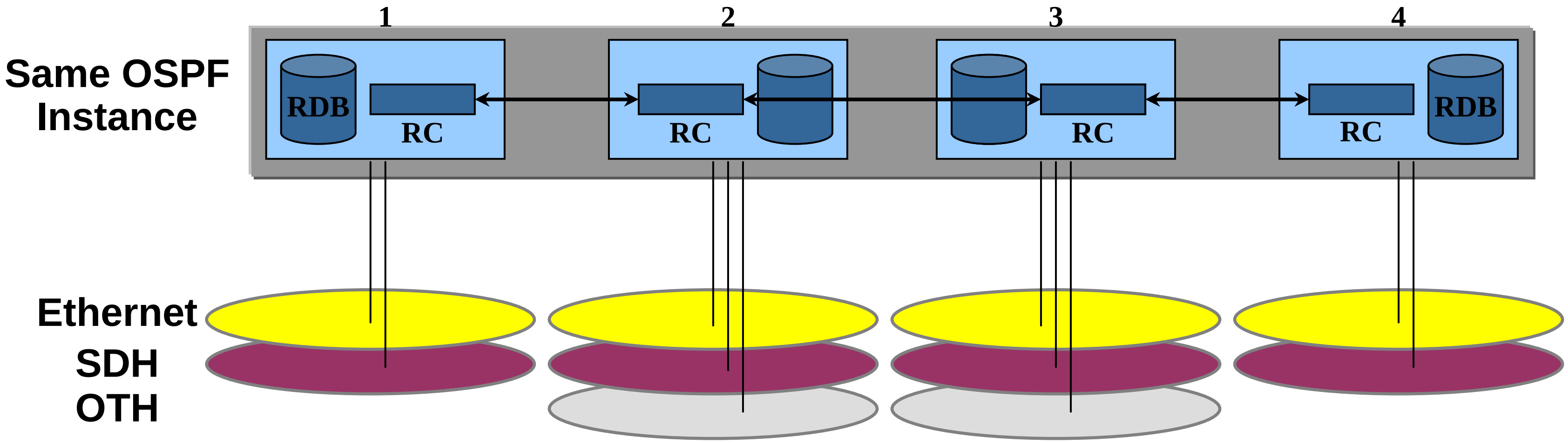
Integrated Routing Approach - Overview

- Considering one administrative authority (e.g., a single carrier) for a network including multiple switching capability systems, GMPLS should provide an effective solution avoiding multiplication of control plane instances per node
- Corresponding to the data plane integration, GMPLS is the unique candidate which is able to target this function
- The MRN approach is based on following concepts:
  - Reduction of the complexity of the control plane interactions: single control plane instance
  - Reduction of addressing complexity: single addressing space
  - Usage of LSP hierarchy mechanism: Forwarding Adjacencies
  - Consistent Traffic Engineering (TE) within the scope of a single policy to improve resource optimization (using a coordinated routing approach)

ALCATEL

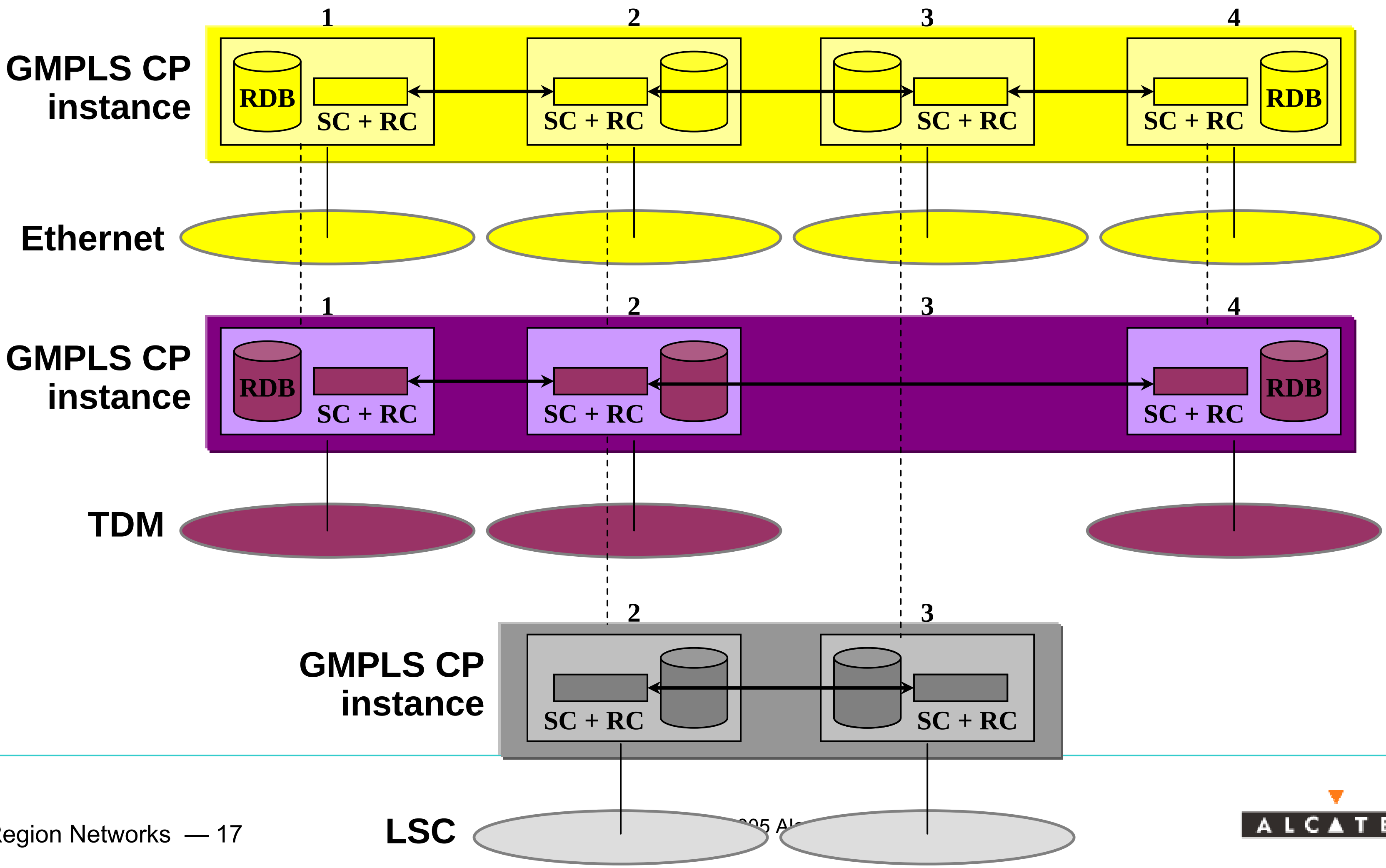
Integrated Routing Definition

- Scope: 1:N (#control plane instances:#switching layers)
- Single GMPLS controller instance with single Routing Controller (RC) and Routing Database (RDB) that processes and maintains all routing information (TE link topology/attributes) for all TE links terminating the corresponding Switching Capability (SC)

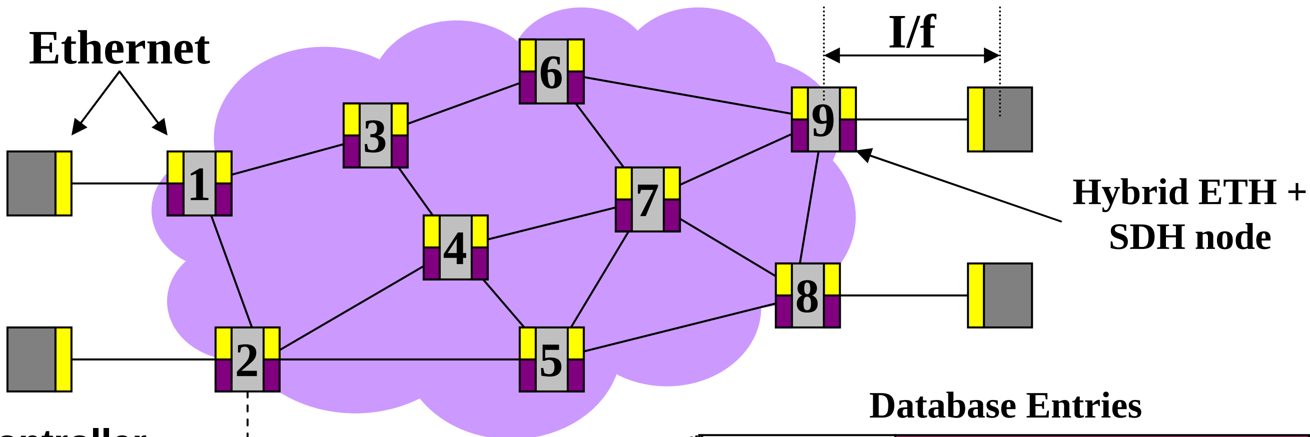


Comparison with Overlay

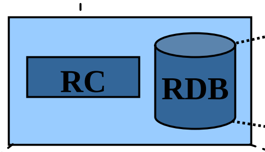
- Scope: N:N (#control plane instances:#switching layers)
- Multiple control plane instance one per network switching layer:



Multi-Layer TE Routing - Example



<u>Single GMPLS (routing) controller</u> maintains L2SC and TDM-SC TE link topology and reachability within single OSPF instance

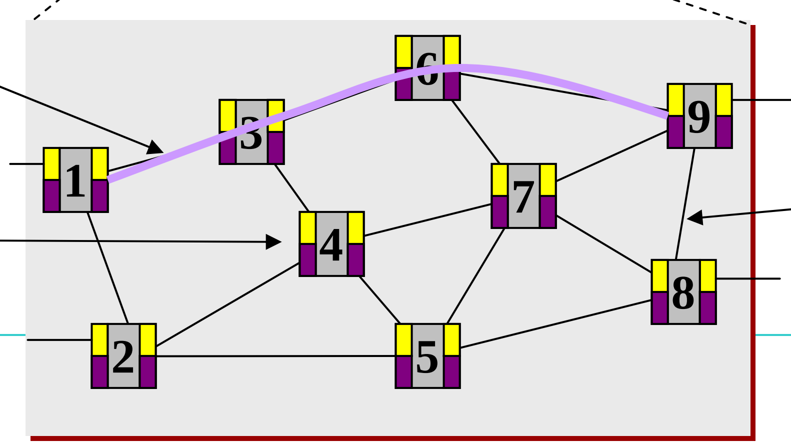


Database Entries

TDM	TE Link 1
L2SC	TE Link 1
…	
L2SC	TE Link 7

TDM LSP (Forwarding Adjacency) with Ethernet L2SC Capability

Interface w/ two switching capabilities
Ethernet (L2SC)
SDH (TDM-SC)



TE Links with Ethernet - L2SC and TDM-SC

Forwarding Adjacency (FA) - Concept

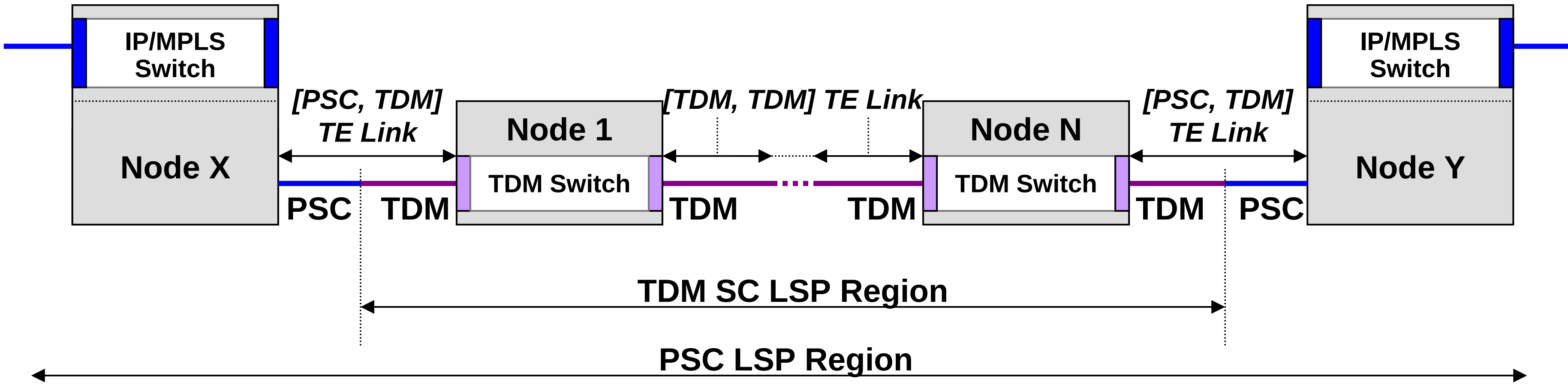
> FA Label Switched Path (LSP):

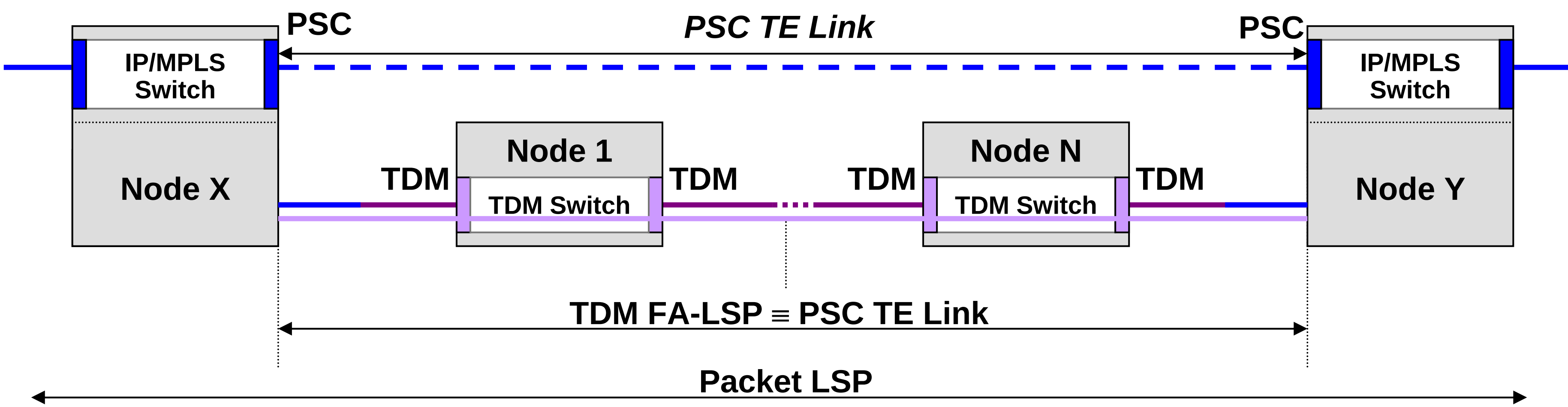
- Provisioned or triggered end-to-end LSP representing a (virtual) TE link at the data plane client layer $\Rightarrow$ data plane adjacency
- Do never imply any routing adjacency (but may imply a signaling adjacency)
- FA-LSPs can be un-/numbered and bundled
- Inheritance mechanism for TE attributes

> Control Plane scalability enhancement:

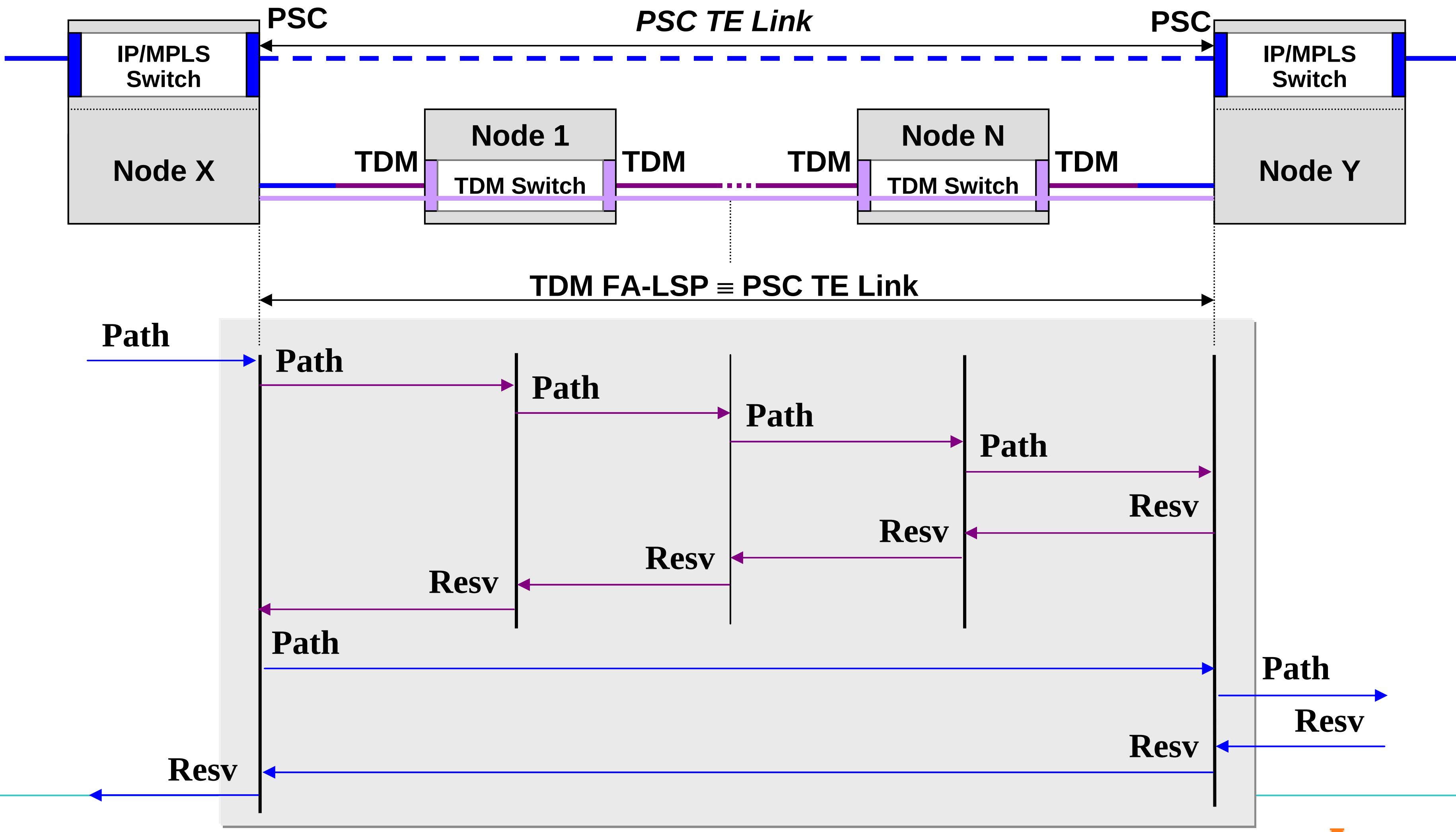
- Signaling: no Path/Resv state maintained at intermediate nodes for LSPs nested within the FA-LSP (nesting LSP)
- Routing: #data plane adjacencies independent from #control plane adjacencies (fixed by CP topology)

Forwarding Adjacency (FA) – Example (1)





Forwarding Adjacency (FA) Signaling – Example (2)



ALCATEL

What is Missing from GMPLS ?

> A profile of existing GMPLS building blocks

- Routing over Forwarding Adjacencies (FA) with TE metric inheritance policies to avoid undesirable effects (e.g. full mesh)
  - TE attributes related to packet and circuit constraints may be orthogonal (e.g. TE metric)
  - consideration for LSP region specifics and multi LSP-region advertisement of resources (one FA TE link built from two FA-LSPs)

> Extensions for cross-region considerations

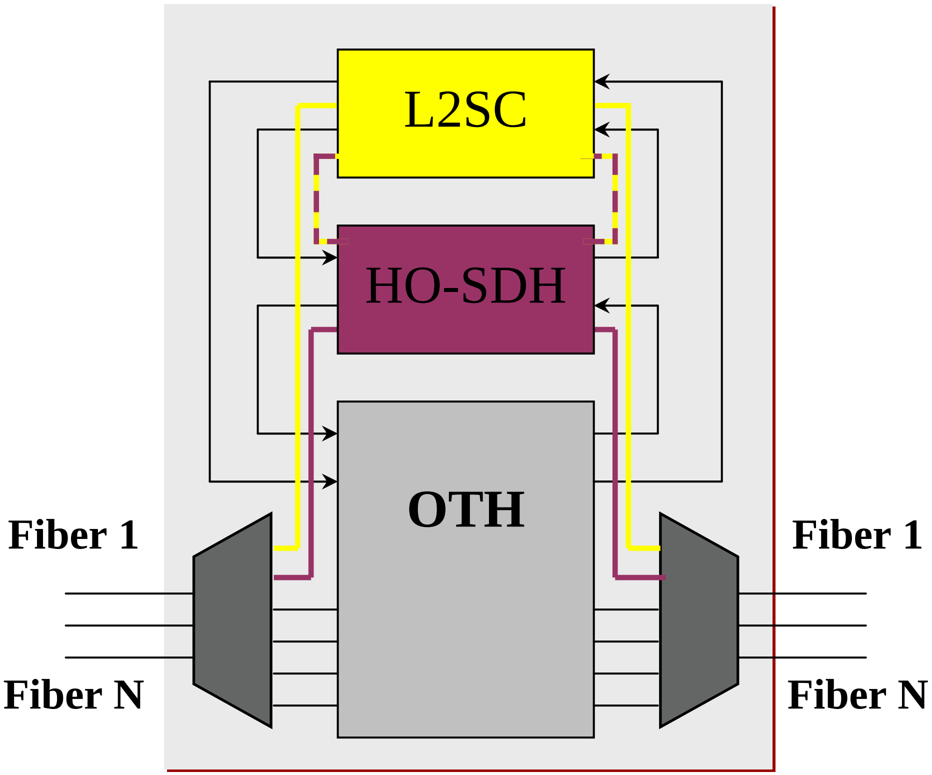
- By definition Multiple SC nodes are not able to process all the transit traffic at each switching capability
- This results in potential contention for use (terminate a LSP for example) of a given SC (interface capability is not sufficient)

⇒ Introduction of interface adaptation capability descriptor (IACD)

ALCATEL

Interface Adaptation Capability Descriptor (IACD)

- The Interface Switching Capability descriptor (ISCD) alone does not allow remote LSR to deduce intermediate termination capabilities of Multiple SC systems
- The Interface Adaptation Capability descriptor (IACD) complete sub-TLVs of the TE link TLV of (opaque TE LSA defined in [RFC3630] and extended in [GMPLS-OSPF])
- Termination issues (ambiguous/blocking case)



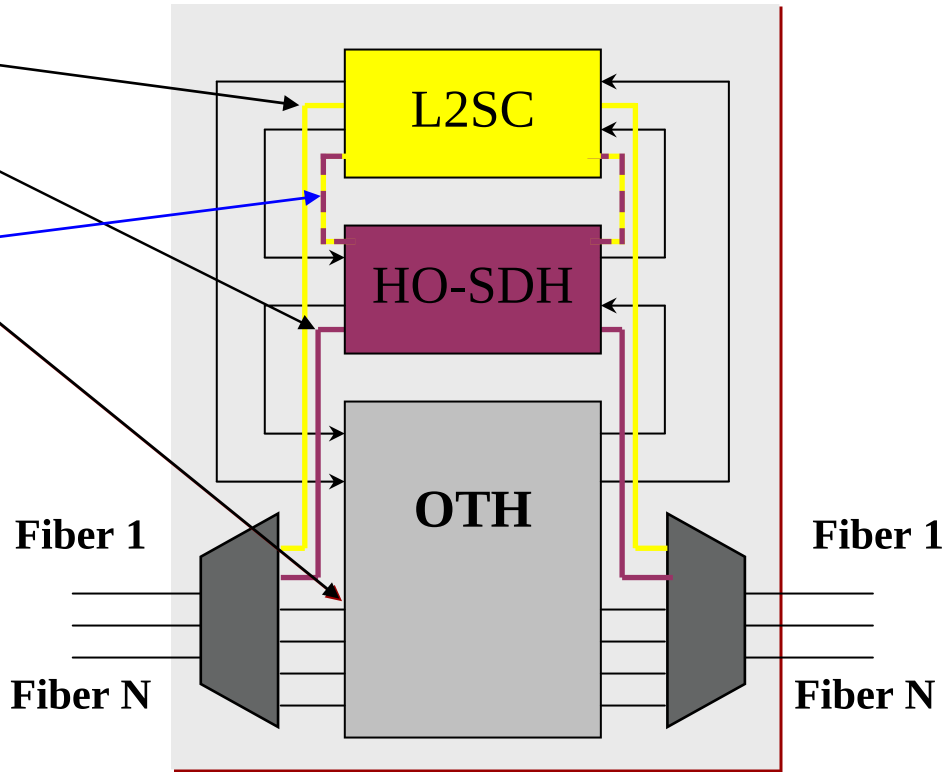
ALCATEL

Interface Adaptation Capability Descriptor (IACD)

> As part of TE Link_1 advertisement
- ISCD sub_TLV 1 for L2SC
- ISCD sub_TLV 2 for HO-SDH
- ISCD sub_TLV 3 for OTH
- IACD sub_TLV 1 for L2SC to HO-SDH

> if no bundling of [1…N]

> As part of TE Link_N advertisement
- ISCD sub_TLV 1 for L2SC
- ISCD sub_TLV 2 for HO-SDH
- ISCD sub_TLV 3 for OTH
- IACD sub_TLV 1 for L2SC to HO-SDH



What is Missing from GMPLS ?

> From [LSP-HIER]: if a node at the edge of a region receives a Path message, it determines whether located at the edge of an LSP region with respect to the ERO carried in the Path message:
  - It must then determine the other region edge wrt ERO, using the TE database and extract from the ERO the subsequence of hops from itself to the other end of the region
  - Then, the node compares the subsequence of hops with all existing FA-LSPs it originated:
    - if a match is found and that FA-LSP has enough unreserved bandwidth for the LSP being signaled, and the PID of the FA-LSP is compatible with the PID of the LSP being signaled, the LSR uses that FA-LSP. The Path message for the original LSP is sent to the egress of the FA-LSP

      Note: before sending the Path message, the ERO is adjusted by removing the subsequence of the ERO that lies in the FA-LSP, and replaced with just the end point of the FA-LSP
    - if no existing FA-LSP is found, the LSR sets up a new FA-LSP. That is, it initiates a new LSP setup just for the FA-LSP

ALCATEL

Multi-Region Signaling

> Applying this procedure, in MRN environments can lead to setup one-hop FA-LSPs between each node

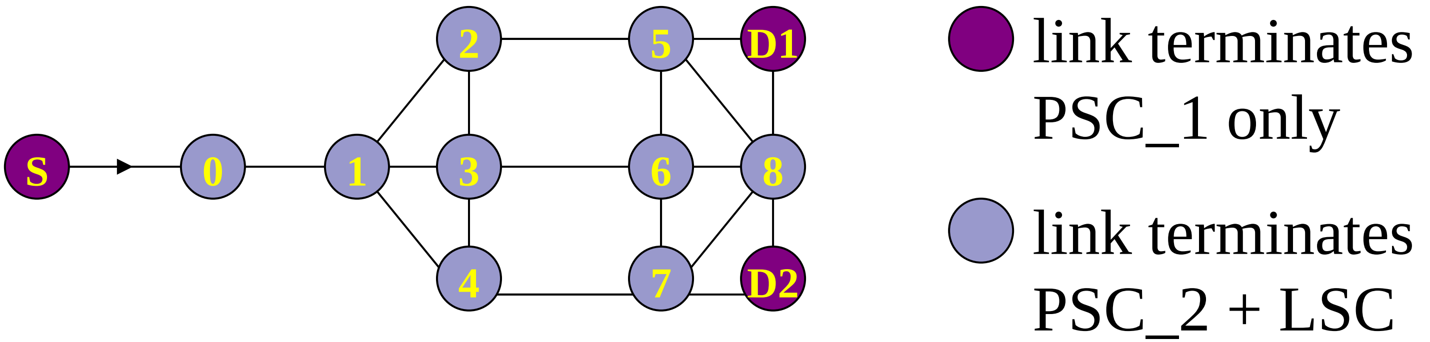
> Therefore, considering that path computation is capable to take into account TE information (in particular, the SC terminating on links along to the path), it is relevant to provide enough information to indicate 1) SC to be included/excluded and 2) for which TE link

> Limiting modifications to existing RSVP-TE procedures are required for this purpose e.g. indication of these SC values in a new sub-object of the eXclude Route Object (XRO)

- Such information can be specified by explicitly indicating which SCs have to included or excluded before initiating the procedure described (see previous slide)

> This solves the ambiguous selection of SC to be used along a given segment of the path (and thus provides the possibility to optimize resource usage on a multi-layer basis)

Multi-Region Signaling - Example



> Outgoing explicit route from node S: [0,1,3,6,8,D1] to setup PSC_2 LSP from S to D1

> At node 0: route selection is PSC_2 LSP [1,3,6,8,D1]

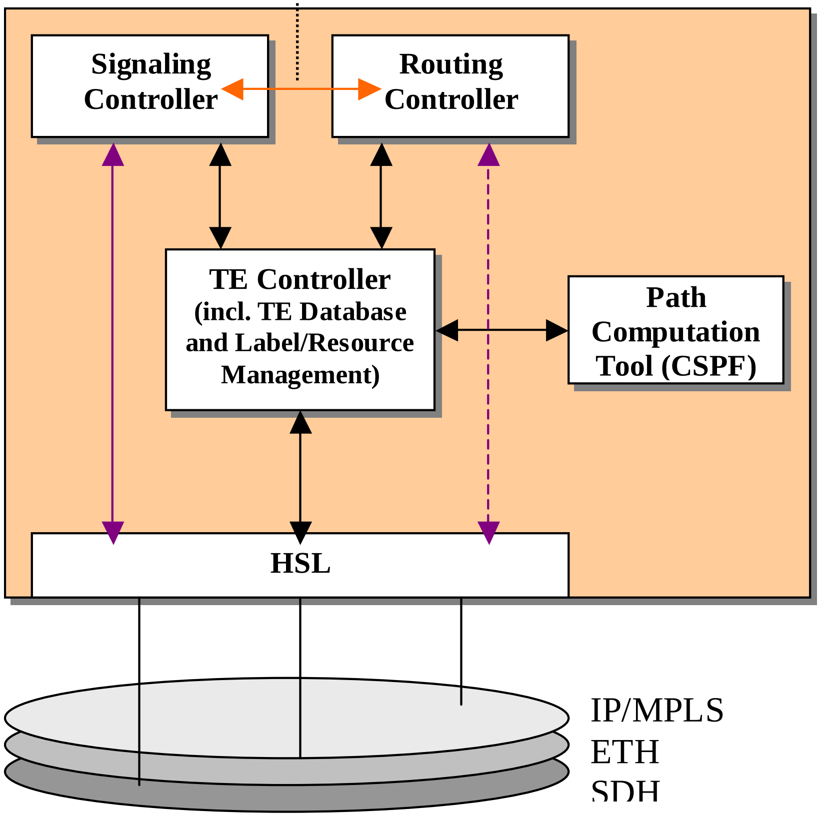
> At node 1: solution scope for route selection is either LSC LSP [3] [3,6], [3,6,8] or [3,6,8,D1] before continuing PSC_2 LSP signaling

MRN Controller - Functional Architecture (1)

> Control plane organized around the <u>TE link controller</u> interacting with the signaling controller and the routing controller + direct access to the IP control channels (IPCC)
  - any controlled entity ≡TE link i.e. LSP, FA-LSP (FA link), un/bundle TE link
  - prevents from maintaining 1:1 relationship between TE links and routing adjacency

> <u>Routing controller</u>
  - Process TE, topology and reachability information from multiple SC (TEDB without any specialization)
  - Can still make use of advanced features (incl. two step increments, fast convergence, hitless restart, redundancy) w/o impacting genericity

<u>cut through</u>: allow routing of the signaling messages and fast re-convergence of signaling adjacencies in case IP connectivity failure



MRN Controller – Functional Architecture (2)

> This architecture, has no impact on
  - control plane modularity
  - modularity of the capabilities delivered by one of its components

> Example:
  - Assume a legacy IP/MPLS - SDH node and introduction of Ethernet functionality; however, the routing and signaling instances do not yet support L2SC interfaces
  - To avoid any disruption during software upgrade, OSPF(-TE) hitless restart and RSVP(-TE) hitless restart procedures are applied

    $\Rightarrow$ upgrade operations can be performed without impacting any active TE link
  - TE controller can transiently process all data link failures that may occur during upgrade/downgrade operation as long as the Hardware Shielding Layer (HSL) interacts with an SC unaware API

Standardization: IETF CCAMP Working Group

Requirements

draft-shiomoto-ccamp-gmpls-mrn-reqs

Result of the previous version split in 2 parts

Analysis

draft-xyz-ccamp-gmpls-mrn-analysis

Informational document analysing current GMPLS protocol suite capabilities wrt MRN requirements

Functional Specification

draft-vigoureux-ccamp-gmpls-mrn-functional

Document detailing functional and operational aspects (planned for next meeting)

GMPLS Protocol Extensions

draft-papadimitriou-ccamp-gmpls-mrn-extensions

Result of the previous version split in 2 parts

ALCATEL

Conclusion

> As GMPLS is experiencing its first deployment, the unified signaling with integrated routing may also apply for “real life” models based on devices supporting multiple switching capabilities

> Starting with a mono-carrier (single administrative authority) framework, few protocol extensions (Interface Adaptation Capability Descriptor) are required to support multi-region networks

> MRN is an enabler for optimized networking (TE, protection and rerouting, reduced complexity,…) and packet/transport integration

> Further developments may also cover more sophisticated features such as (re-)grooming or Hybrid Photonic Networks

Thanks for your attention

Dimitri PAPADIMITRIOU
ALCATEL