Software Defined Optical Networking
0. Syllabus

Prof. Dr. Andreas J. Kassler
kassler@ieee.org

1. Software Defined Networking
2. SDN and MPLS
3. SDN and TE
4. SDN and Optical Networking

Thanks to many contributions from colleagues including J. Rexford, N. McKeown, G. Parulkar

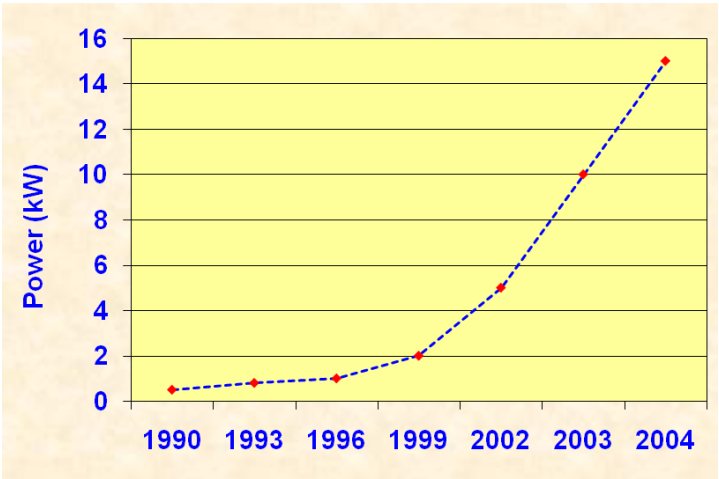
SDN and Optical Networking
Motivation

- Today: Two different networks: MPLS/IP and Optical Transport
  - Provisioned, managed and controlled independently, Significant CAPEX and OPEX
  - Reconfiguration complex (order of weeks)
  - Duplication of functionalities in multiple layers, Complex
- No interaction
  - Static IP links, Static circuits or Wavelengths in transport network
- IP backbone
  - Router connections hardwired by Wavelengths
  - Both MPLS and Optical Transport Networks are overprovisioned
    - Multiple protection modes at MPLS requires up to 4 times BW in optical
    - Peak rate provisioning is required at optical transport … why?... How to scale?

Given these considerations, **WAN links are typically provisioned to 30-40% average utilization**. This allows the network service provider to mask virtually all link or router failures from clients. Such overprovisioning delivers admirable reliability at the very real costs of **2-3x bandwidth over-provisioning** and high-end routing gear.
S. Jain, et. al., "B4: Experience with a Globally-Deployed Software Defined WAN," SIGCOMM 2013

SDN and Optical Networking
Motivation

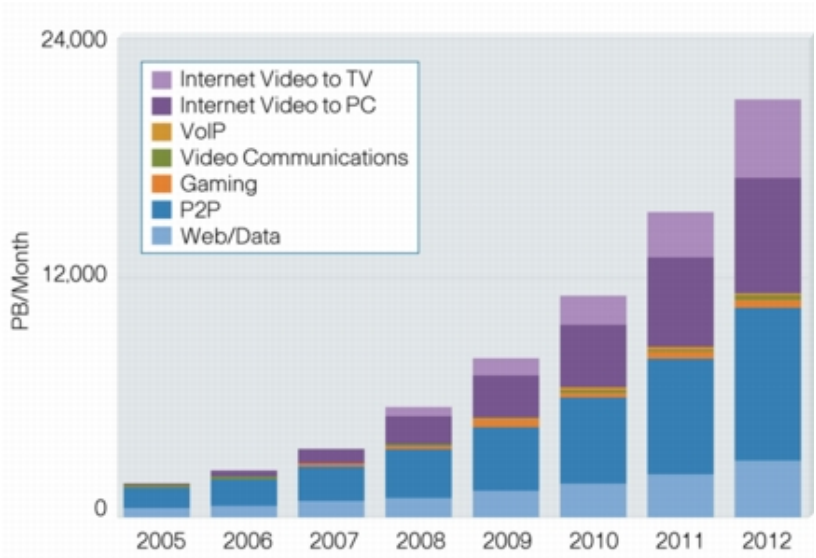
- Bigger routers?
  - Expensive
  - Power hungry
- Functionality
  - Complex and unreliable



Between 2,000 and 4,000 Cisco routers went down for about 7 hours in the NTT East network after a switchover to backup routes triggered the routers to rewrite routing tables, according to a bulletin from CIBC World Markets. The outage disconnected millions of broadband Internet users across most of eastern Japan.

The routing table rewrite overflowed the routing tables and caused the routers' forwarding process to fail, the

- Traffic mix
  - Dominated by video
  - Is packet based processing good?
- Overprovisioning masks the problem
  - Packet switching cannot provide guarantees in terms of bandwidth, delay, jitter and loss



SDN and Optical Networking
Optical Switching

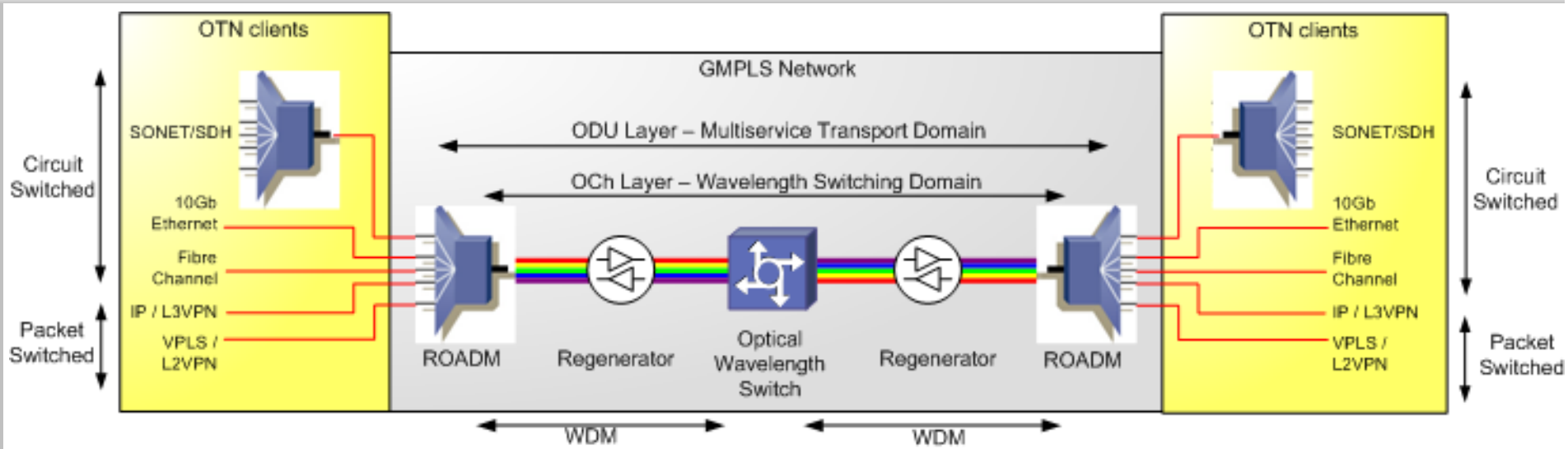
- Optical Switches Needs
  - High capacity, Low power, Reduced cost, High availability
- We want Dynamic Circuit Switching
  - Recover faster from failures
  - Guaranteed bandwidth & Bandwidth-on-demand
  - Good for video flows
  - Guaranteed low latency & jitter-free paths
  - Help meet SLAs – lower need for over-provisioned IP links
- Interactions with higher layers desired to
  - Support dynamic services
  - Automatic provisioning (now mostly manually controlled)

SDN and Optical Networking
GMPLS

- GMPLS
  - original goal: Unified Control Plane (UCP) across packet & circuit (2000)
  - Today – the idea is dead
    - Packet vendors and ISPs are not interested
    - Transport n/w SPs view it as a signaling tool available to the mgmt system for provisioning private lines (not related to the Internet)
  - After 10 yrs of development, next-to-zero significant deployment as UCP
- GMPLS Issues
  - control plane complexity escalates when unifying across packets and circuits
    - makes basic assumption that the packet network remains same
      - IP/MPLS network – many years of legacy L2/3 baggage
    - transport network remain same - multiple layers and multiple vendor domains
  - use of fragile distributed routing and signaling protocols, COMPLEXITY
  - does not take into account the conservative nature of network operation
    - can IP networks really handle dynamic links?
    - can transport network service providers give up control to an automated control plane?

SDN and Optical Networking

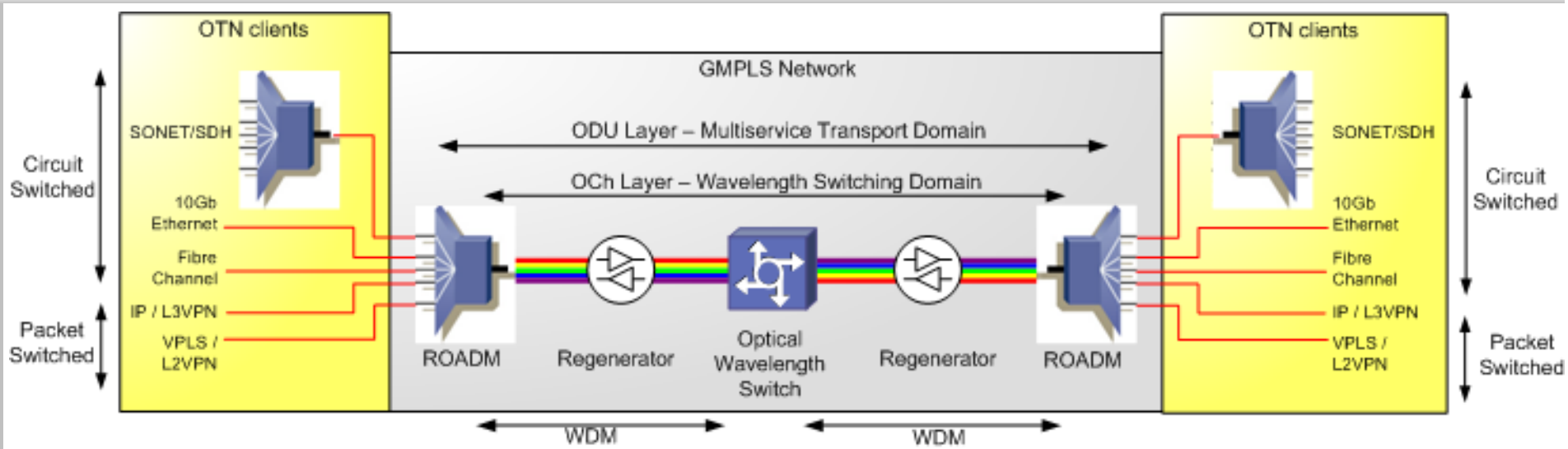
Converged Packet/Optical Circuit Networks and optical circuit switching



- Optical Transport Network (OTN) ITU-T G.709
  - adds SONET/SDH-like features to WDM equipment
  - creates a transparent, hierarchical network designed for use on both WDM/WSON devices and TDM devices
  - OTN control plane is based on GMPLS
    - calculates an optimal path for each client connection
    - Provides protection
    - Automatically recalculation of path if resources are added

SDN and Optical Networking

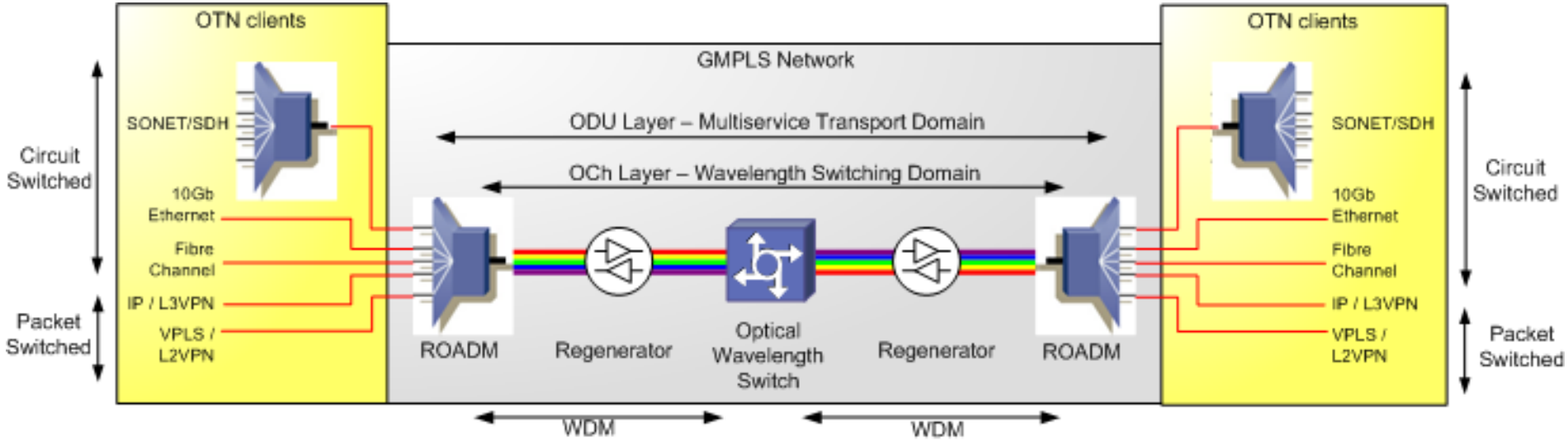
Converged Packet/Optical Circuit Networks and optical circuit switching



- Optical Transport Network (OTN) ITU-T G.709 Frame Format: 6 layers
  - OPU: Optical Channel Payload Unit → analogous to the 'Path' layer in SONET/SDH
  - ODU: Optical Data Unit. This level adds optical path-level monitoring, alarm indication signals and automatic protection switching
  - OTU: Optical Transport Unit. This represents a physical optical port (such as OTU2, 10Gbps), and adds performance monitoring (for the optical layer) and the FEC (Forward Error Correction)

SDN and Optical Networking

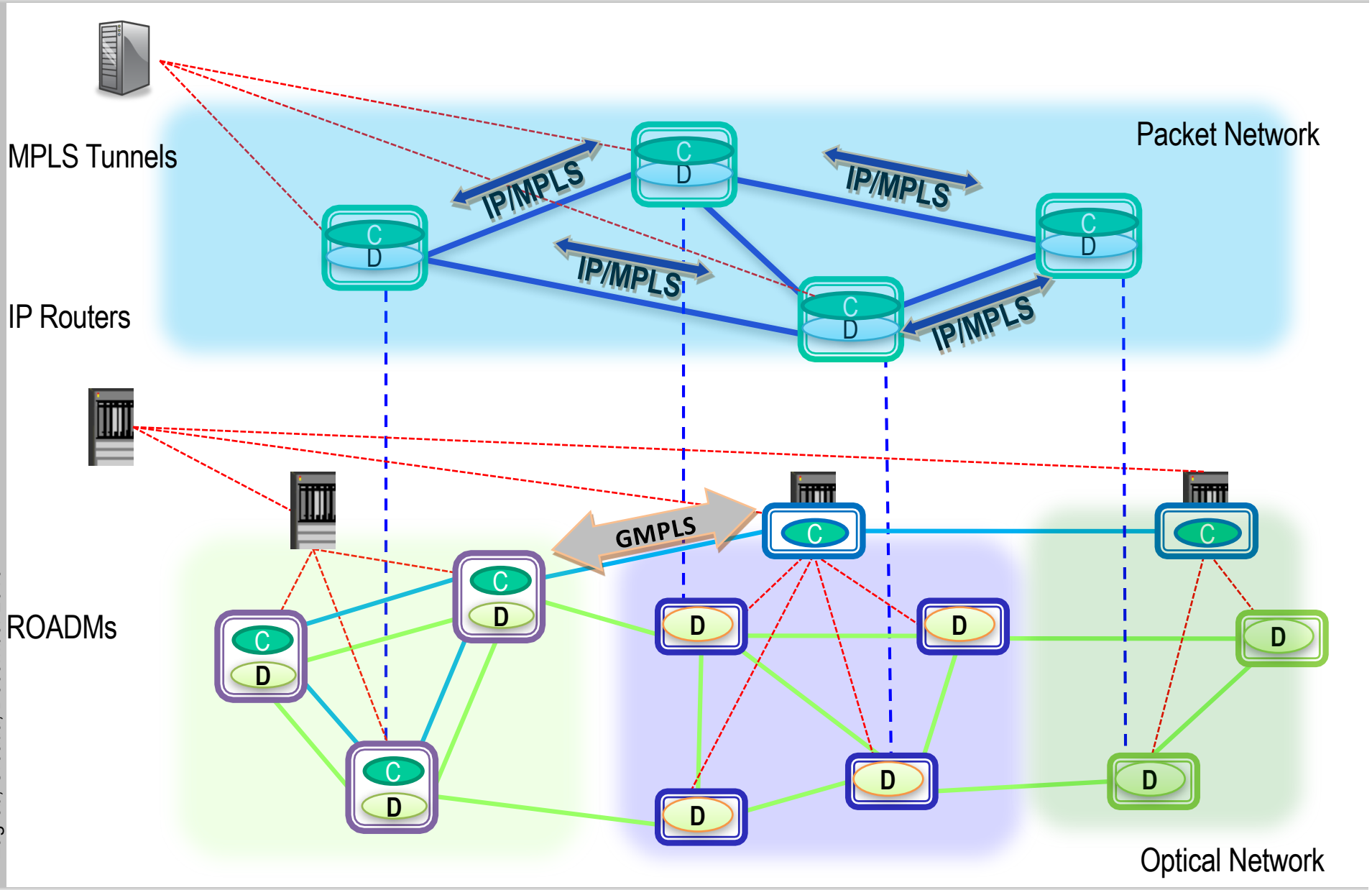
Converged Packet/Optical Circuit Networks and optical circuit switching



- Optical Transport Network (OTN) ITU-T G.709 Frame Format: 6 layers
  - OCh: Optical Channel. This represents an end-to-end optical path.
  - OMS: Optical Multiplex Section. This deals with fixed wavelength DWDM (Dense Wavelength Division Multiplexing) between OADMs (Optical Add Drop Multiplexer).
  - OTS: Optical Transport Section. This deals with fixed wavelength DWDM between relays.

  - Looks complicated?

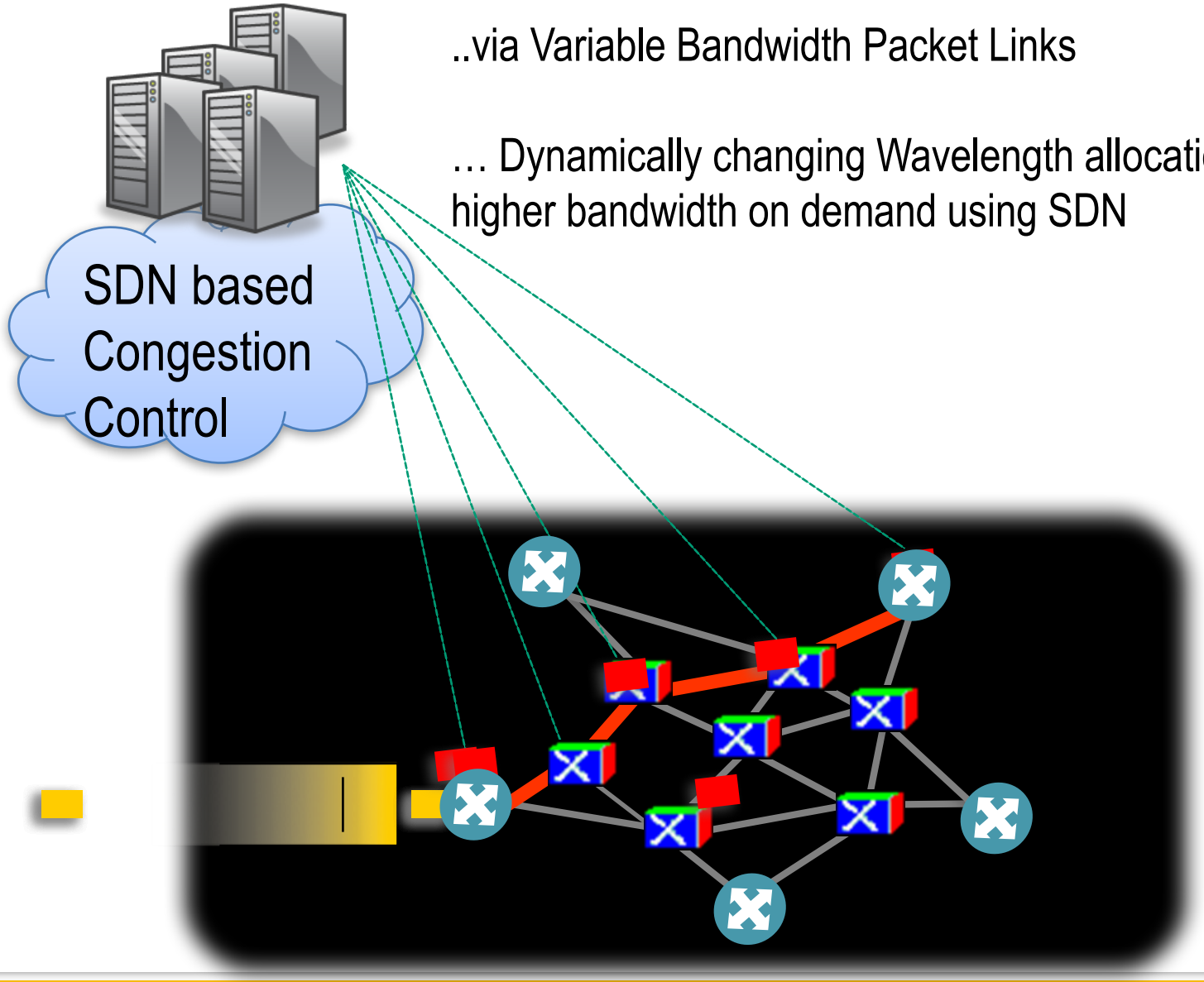
SDN and Optical Networking
Current Situation



SDN and Optical Networking
Example Use Case

..via Variable Bandwidth Packet Links

… Dynamically changing Wavelength allocation to provide higher bandwidth on demand using SDN



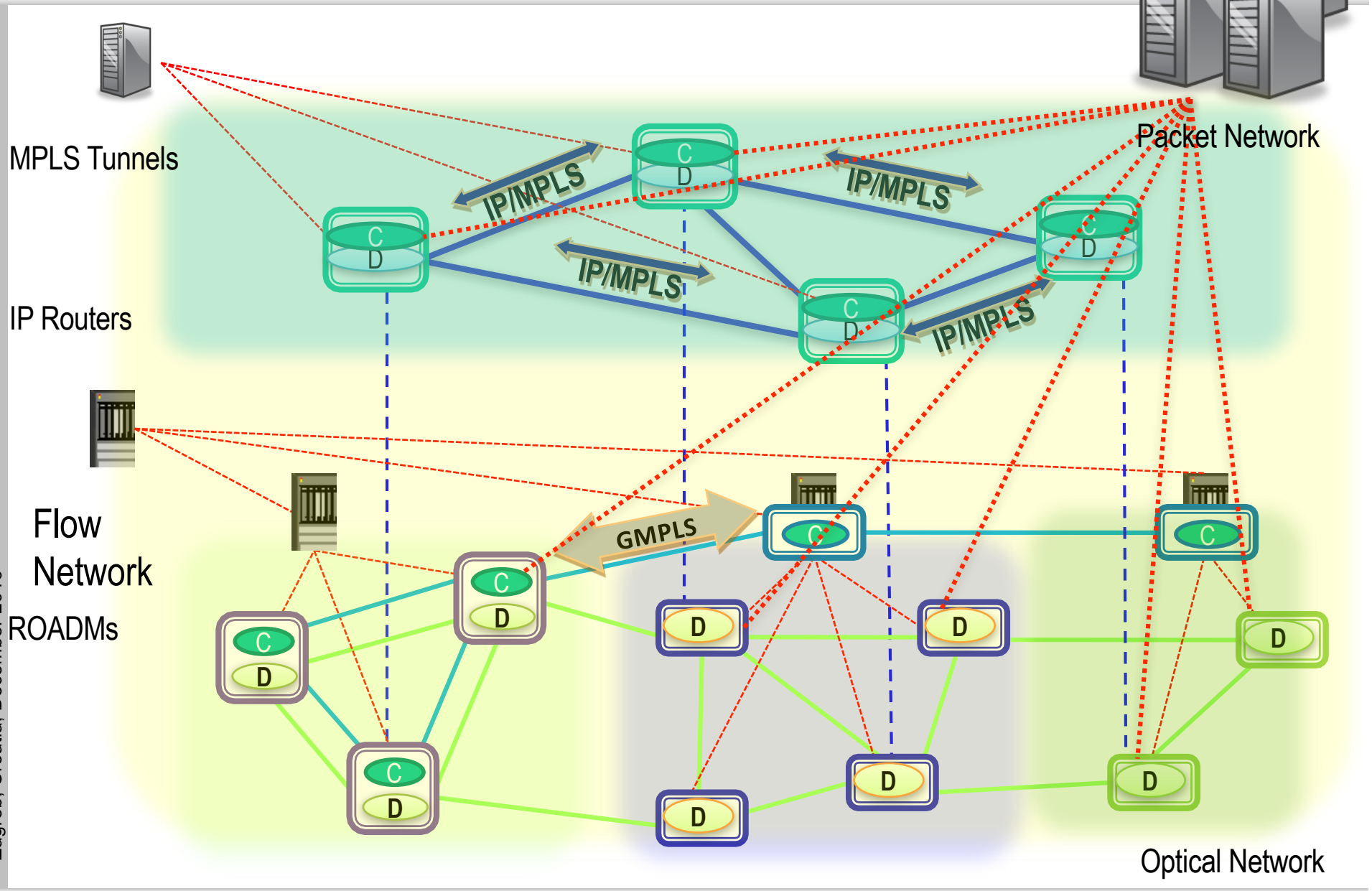
SDN and Optical Networking

Converged Packet/Optical Circuit Networks and optical circuit switching

- New Architecture
  - Want to manage both domains commonly
  - Benefits from dynamic interaction between packet and circuit switching: Example??
- Requires
  - Common control mechanisms – logical centralized, common usage model
- Difficulties
  - Different architecture makes integrated operation hard, ROADMs use e.g. TL1 protocol (or vendor specific) for configuration
  - Need to extend OpenFlow to support Layer 0 (photonic) and Layer 1 (SONET/SDH,OTN) networks
  - Need to enable to program switch ports and fabrics that operate on fibers, wavelengths, and timeslots as well as packet headers while still being simple match/action tables
  - Need to support protection, performance monitoring, administration, and maintenance (OAM) capabilities

SDN and Optical Networking

Converged Packet and Circuit Networks



Packet and Circuit Flows Commonly Controlled & Managed

Simple, network of Flow Switches

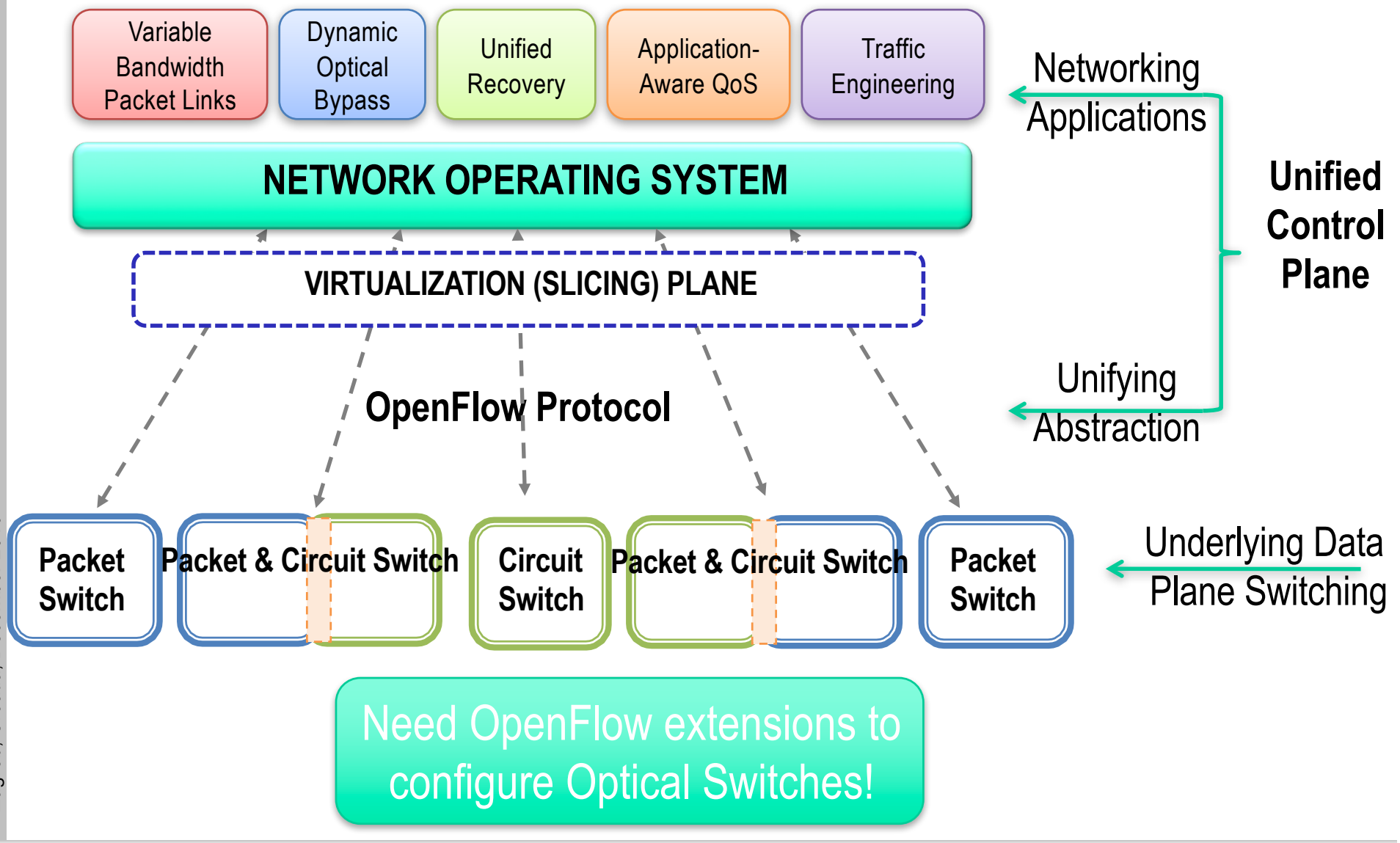
Flow Network

- Reduced overprovisioning of resources with converged Control Plane
- Fast Reconfiguration of Optical Transport Network Capacity according to packet traffic requests
- Seamless Traffic Rerouting in case of failures
- Agile Service Provisioning

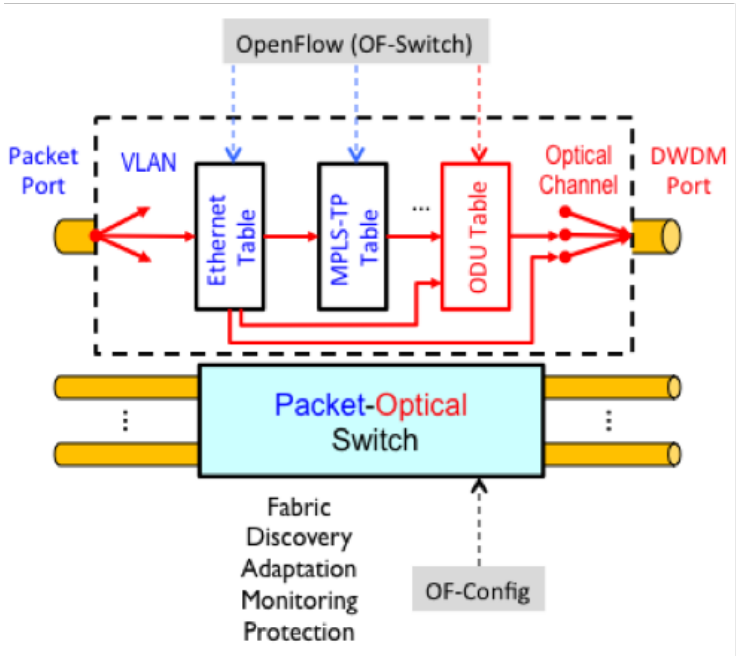
Simple, Unified, Automated Control Plane

… that switch at different granularities: packet, time-slot, Wavelength& fiber

SDN and Optical Networking
Integrated Architecture



Extended OpenFlow Switch Model



- Extensions required
  - L0/L1-capable port descriptions
    - Including neighbor discovery / ID
  - Match table capable of matching/changing L0/L1 "fields",
  - Match & Group tables supporting NE-based protection and OAM
  - Dynamic multilayer control of switching and adaptations, e.g.
    - Ethernet > ODUk > Optical Channel
    - Ethernet > ODUj > ODUk > Optical Channel
    - Ethernet > MPLS LSP > ODUk > Optical Channel

ODU-k = Optical Data Unit – k (k stands for bitrate, e.g. k=4 --> 100 Gbps)

Optical Extensions for OF 1.4

- Until OF 1.3, only support for Ethernet
  - No support for Circuit Switching
  - OF1.4: Optical Port properties
    - configure and monitor the transmit and receive frequency and power of a laser
    - either Ethernet optical port or optical ports on circuit switches.
    - ofp_port_desc_prop_optical for description

```
struct ofp_port_desc_prop_optical {
   uint16_t type;    /* OFPPDPT_3OPTICAL. */
   uint16_t length;  /* Length in bytes of this property. */
   uint8_t pad[4];  /* Align to 64 bits. */
   uint32_t supported;     /* Features supported by the port. */
   uint32_t tx_min_freq_lmda;   /* Minimum TX Frequency/Wavelength */
   uint32_t tx_max_freq_lmda;   /* Maximum TX Frequency/Wavelength */
   uint32_t tx_grid_freq_lmda;  /* TX Grid Spacing Frequency/Wavelength */
   uint32_t rx_min_freq_lmda;   /* Minimum RX Frequency/Wavelength */
   uint32_t rx_max_freq_lmda;   /* Maximum RX Frequency/Wavelength */
   uint32_t rx_grid_freq_lmda;  /* RX Grid Spacing Frequency/Wavelength */
   uint16_t tx_pwr_min;         /* Minimum TX power */
   uint16_t tx_pwr_max;         /* Maximum TX power */
};
```

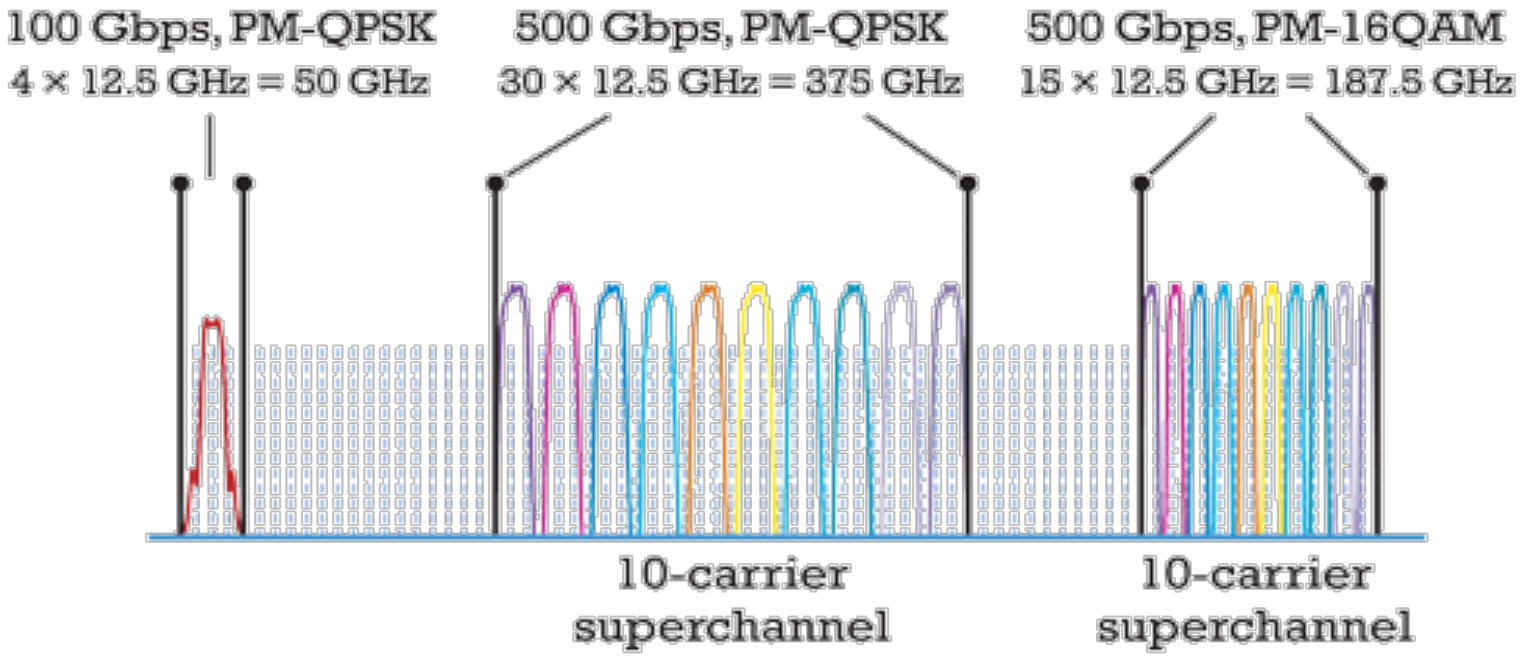
Optical Extensions for OF 1.4

- Until OF 1.3, only Ethernet
  - No support for Circuit Switching
  - OF1.4: Optical Port properties
    - configure and monitor the transmit and receive frequency and power of a laser
    - either Ethernet optical port or optical ports on circuit switches.
    - ofp_port_desc_prop_optical for description

```
/* Features of optical ports available in switch. */
enum ofp_optical_port_features {
   OFPOPF_RX_TUNE = 1 << 0,  /* Receiver is tunable */
   OFPOPF_TX_TUNE = 1 << 1,  /* Transmit is tunable */
   OFPOPF_TX_PWR = 1 << 2,  /* Power is configurable */
   OFPOPF_USE_FREQ  = 1 << 3,  /* Use Frequency, not wavelength */
};
```

Optical Extensions for OF 1.4

- Until OF 1.3, only Ethernet
  - No support for Circuit Switching
  - OF1.4: Optical Port properties
    - configure and monitor the transmit and receive frequency and power of a laser
    - either Ethernet optical port or optical ports on circuit switches.
    - ofp_port_mod_prop_optical for config



Optical Extensions for OF 1.4

- Until OF 1.3, only Ethernet
  - No support for Circuit Switching
  - OF1.4: Optical Port properties
    - configure and monitor the transmit and receive frequency and power of a laser
    - either Ethernet optical port or optical ports on circuit switches.
    - ofp_port_stats_prop_optical for monitor

```
/* Optical port stats property. */
struct ofp_port_stats_prop_optical {
   uint16_t type;    /* OFPPSPT_OPTICAL. */
   uint16_t length;  /* Length in bytes of this property. */
   uint8_t pad[4];  /* Align to 64 bits. */
   uint32_t flags; /* Features enabled by the port. */
   uint32_t tx_freq_lmda; /* Current TX Frequency/Wavelength*/
   uint32_t tx_offset; /* TX Offset */
   uint32_t tx_grid_span; /* TX Grid Spacing */
   uint32_t rx_freq_lmda; /* Current RX Frequency/Wavelength*/
   uint32_t rx_offset; /* RX Offset */
   uint32_t rx_grid_span; /* RX Grid Spacing */
   uint16_t tx_pwr; /* Current TX power */
   uint16_t rx_pwr; /* Current RX power */
   uint16_t bias_current; /* TX Bias Current */
   uint16_t temperature; /* TX Laser Temperature */
};
```

SDN for Converged Packet and Circuit Networks

- Cross-connect table in Circuit Switch
  - Can be exploited for flow abstraction
  - Integrated operation between the tables

Allow to specify internal and virtual ports used in circuit switches.
These internal port numbers can be used to specify "mapper" ports that map Ethernet packets to TDM time-slots
virtual ports can be used to define Virtual Concatenation Group (VCG) numbers used for VCAT technology in TDM switches

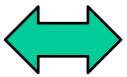
Packet Flows

Switch Port	MAC src	MAC dst	Eth type	VLAN ID	IP Src	IP Dst	IP Prot	TCP sport	TCP dport	Action

Exploit the cross-connect table in circuit switches

Circuit Flows

In Port	In Lambda	VCG	Starting Time-Slot	Signal Type



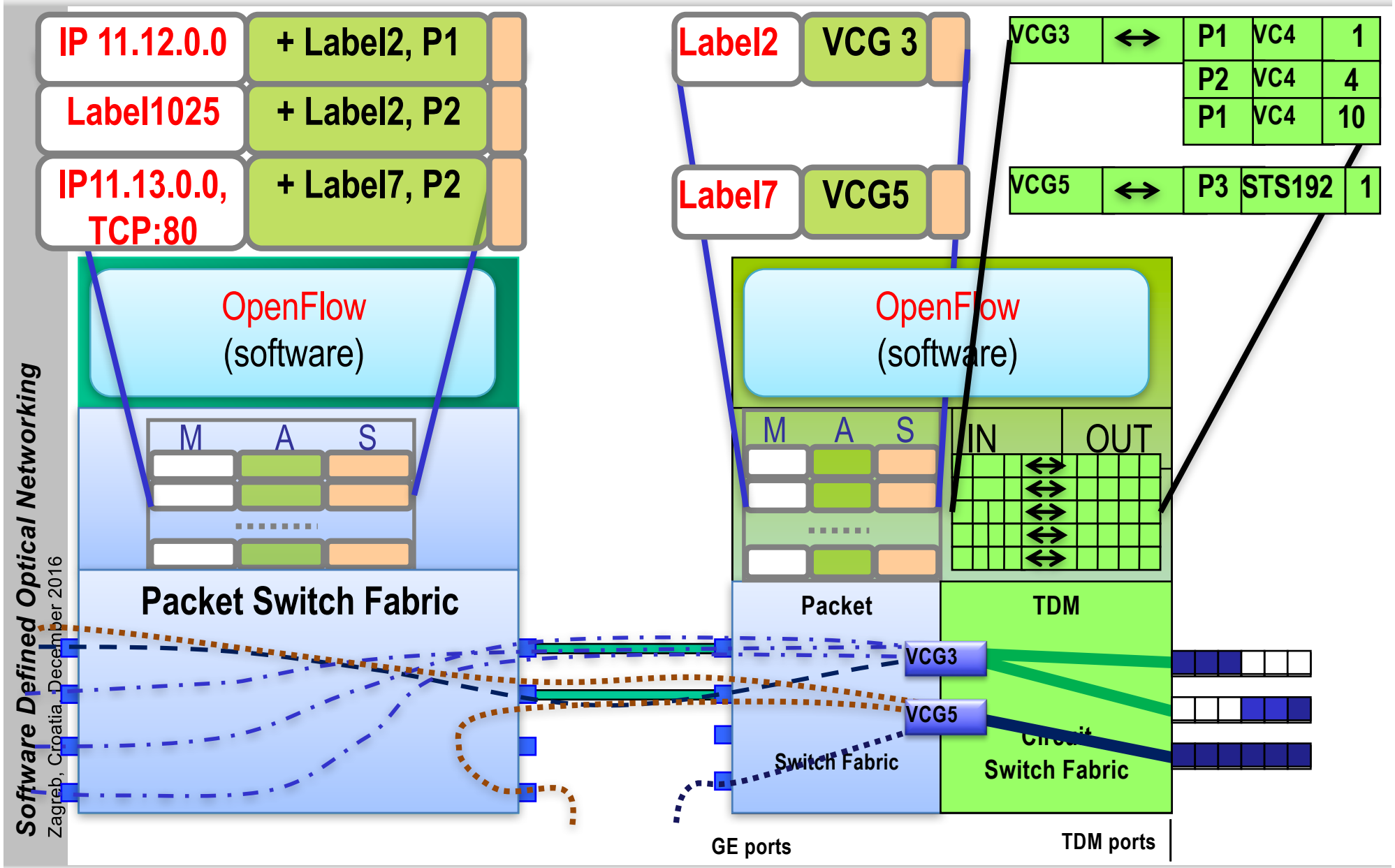
Out Port	Out Lambda	VCG	Starting Time-Slot	Signal Type

The Flow Abstraction presents a unifying abstraction

… consider both as flows in a flow-switched network

SDN and Optical Networking

SDN for Converged Packet and Circuit Networks



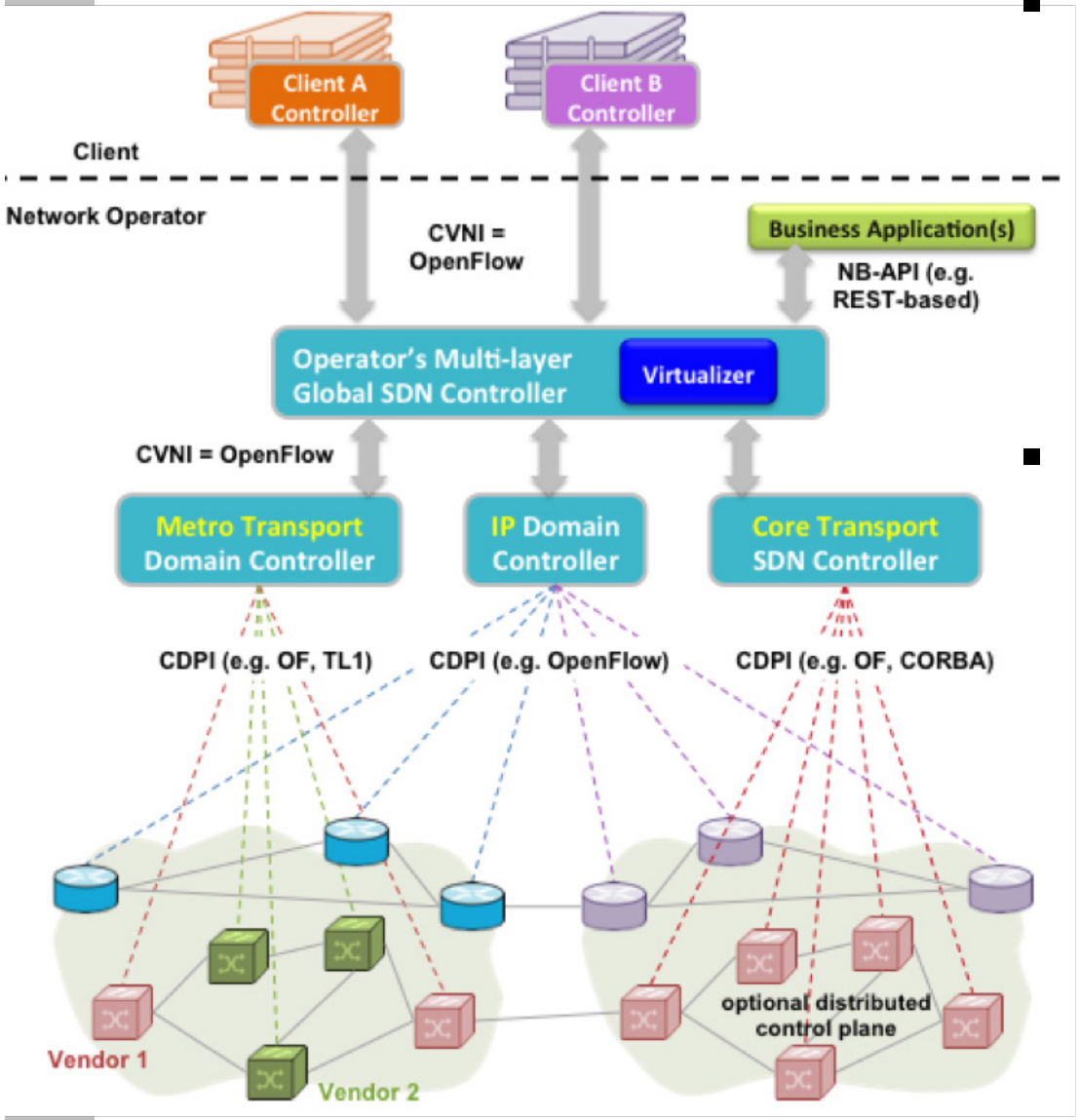
SDN and Optical Networking
Practical Issues

- Transport Service Providers
  - Do not want to give up manual control of their networks
    - To automated control plane
    - No matter how intelligent
  - How to convince them?

- Converged Operation of packet and circuit networks is good
  - For companies that own both types of network e.g. AT&T
  - What about e.g. Google (owns only packet infrastructure)
    - Do not buy circuit switches (too costly)
  - How to convince them?

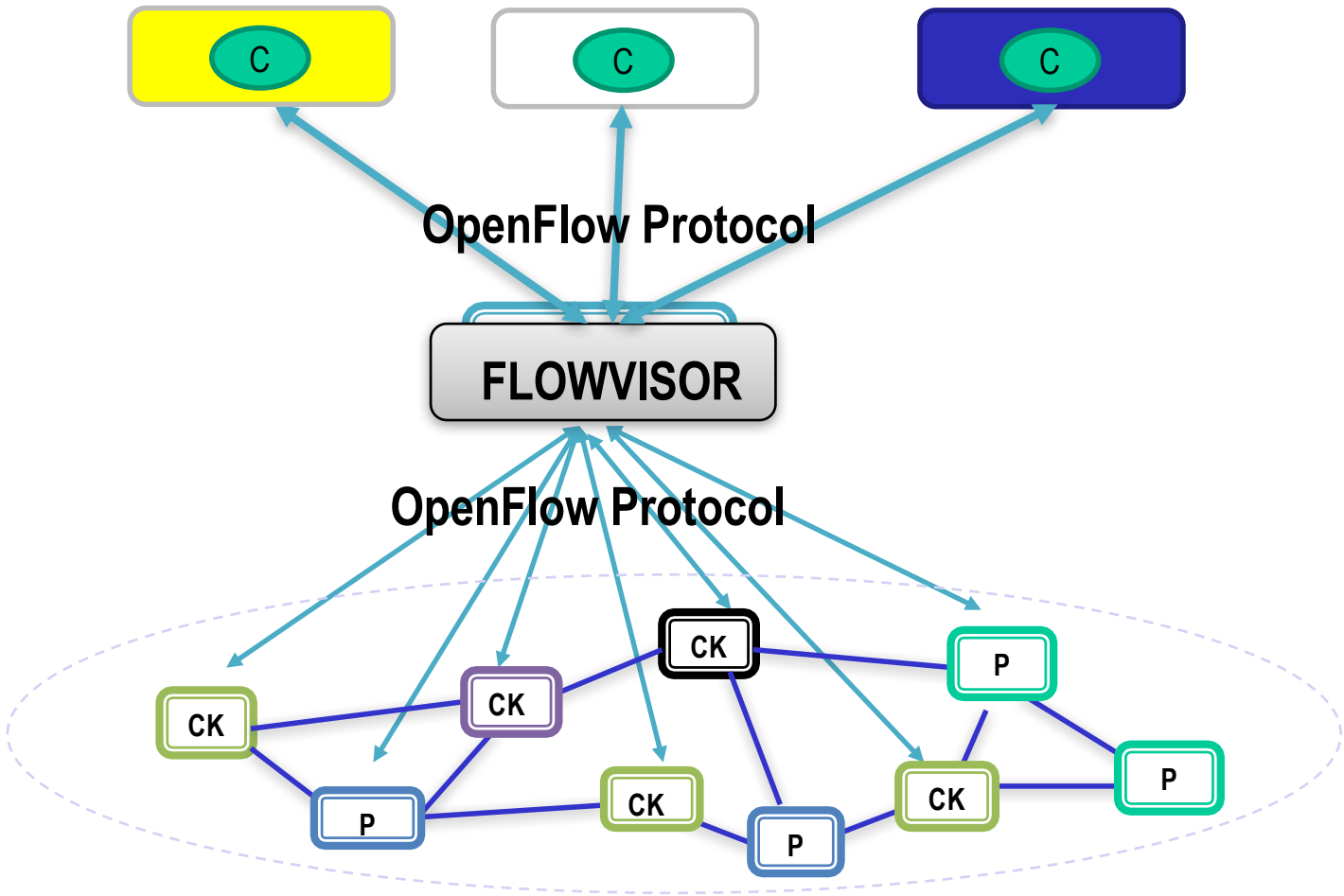
- Virtualization has benefits

SDN and Optical Networking
Virtualization



- Network Virtualization allows to
  - Partition an operator transport network into multiple virtual ones (VTNs)
  - Each VTN contains certain virtual network elements (V-NEs) and interconnecting virtual links (VLs) that are under the control of the respective clients
- Two interfaces
  - CDPI- control data plane interface
    - direct interaction between the SDN controller and the transport network elements. Can be anything, OF or GMPLS
  - CVNI - control virtual network interface
    - Client Controller uses OF to interact with provider controller offering a virtualized view of each client slize

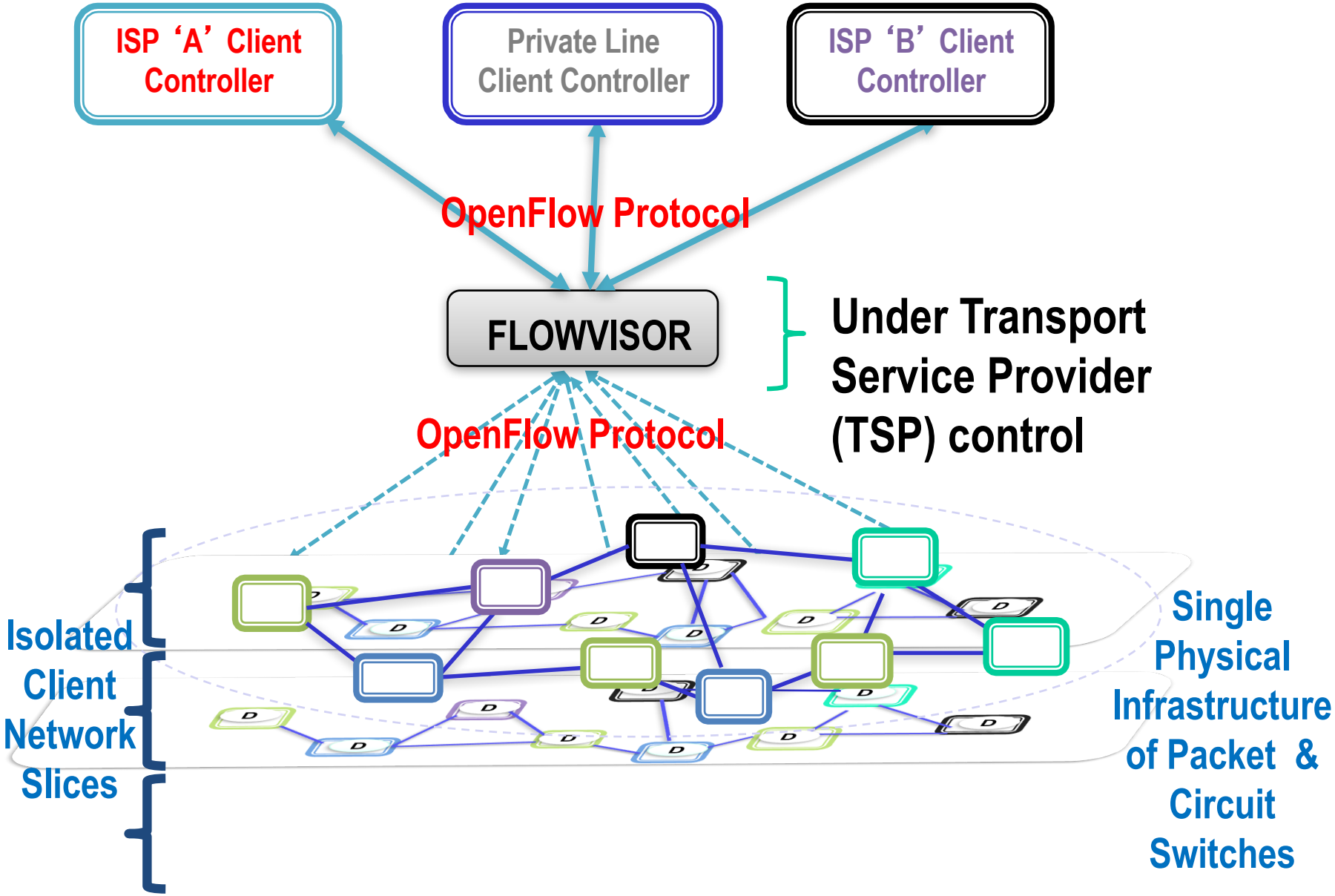
SDN and Optical Networking
Unified Virtualization



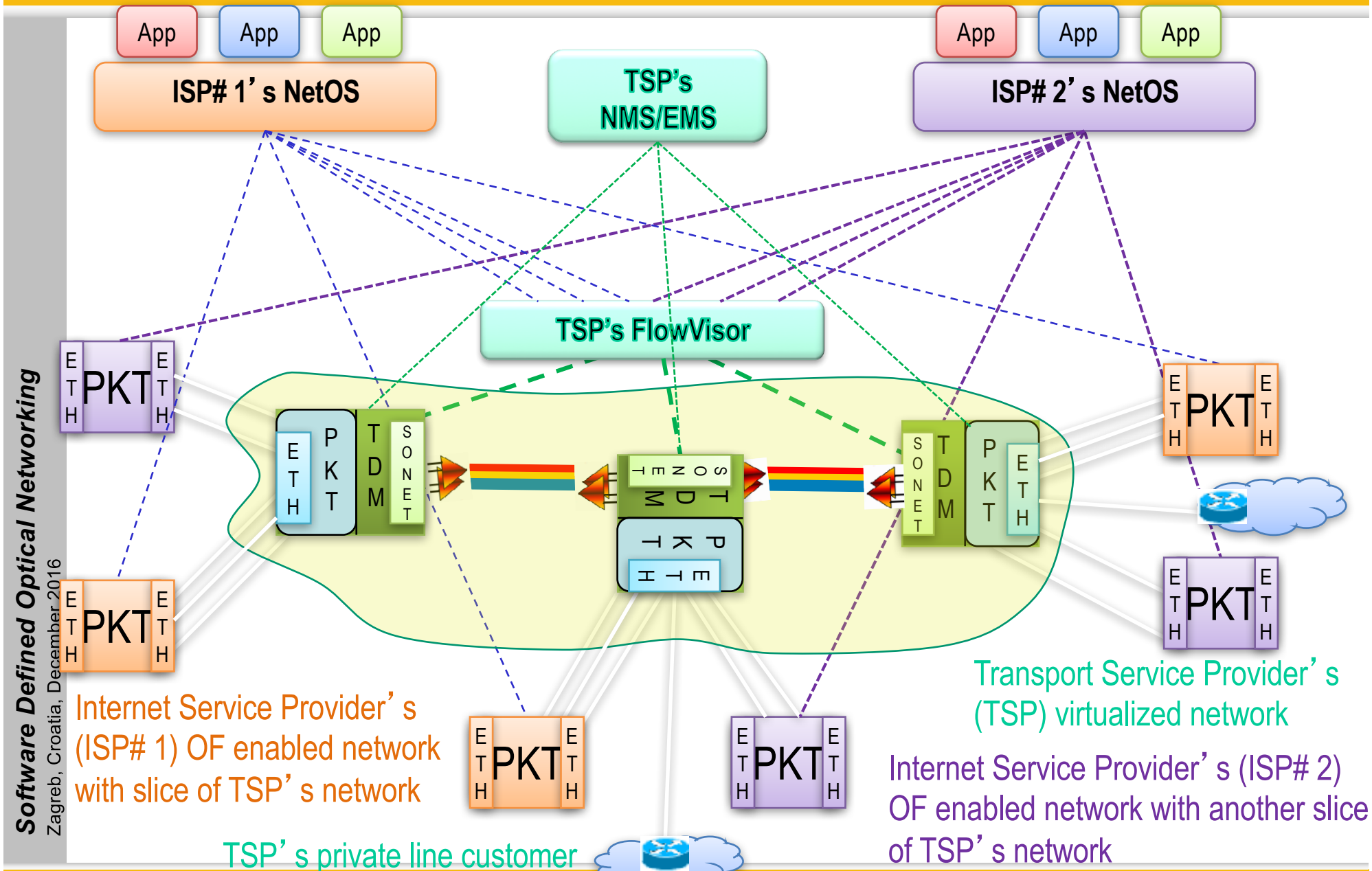
CK = Optical Circuit Switch
P = Packet Switch

SDN and Optical Networking

Different ISP's Scenario



SDN and Optical Networking
Example



SDN and Optical Networking
Benefits

1. TSP can virtualize its network with the FlowVisor while maintaining operator control via NMS/EMS.
   - FlowVisor will manage slices of the TSP's network for ISP customers, where { slice = bandwidth + control of part of TSP's switches }
   - NMS/EMS can be used to manually provision circuits for Private Line customers

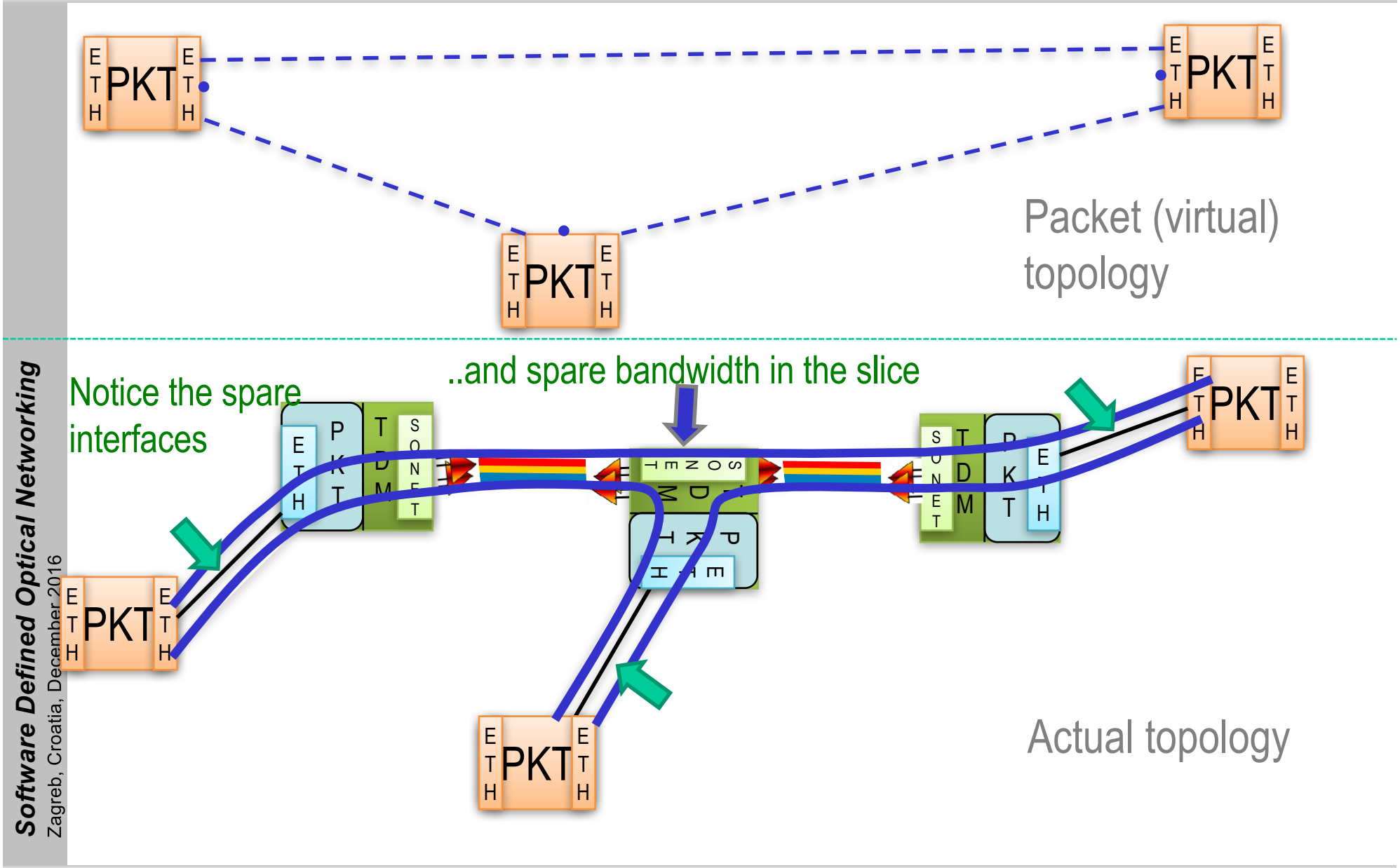
2. Every customer (ISP# 1, ISP# 2, Pline) is isolated from other customer's slices.
   - ISP#1 is free to do whatever it wishes within its slice
     - eg. use an automated control plane (like OpenFlow)
     - bring up and tear-down links as dynamically as it wants
   - ISP#2 is free to do the same within its slice
   - Neither can control anything outside its slice, nor interfere with other slices
   - TSP can still use NMS/EMS for the rest of its network

SDN and Optical Networking
Business Models

- ISP# 1 pays for a slice = { bandwidth + TSP switching resources }
  - Part of the bandwidth is for static links between its edge packet switches (like ISPs do today)
  - and some of it is for redirecting bandwidth between the edge switches (unlike current practice)
  - The sum of both static bandwidth and redirected bandwidth is paid for up-front.
  - The TSP switching resources in the slice are needed by the ISP to enable the redirect capability.
- ISP# 2 pays for a slice = { bandwidth + TSP switching resources }
  - Only the bandwidth for static links between its edge packet switches is paid for up-front.
  - Extra bandwidth is paid for on a pay-per-use basis
  - TSP switching resources are required to provision/tear-down extra bandwidth
  - Extra bandwidth is not guaranteed

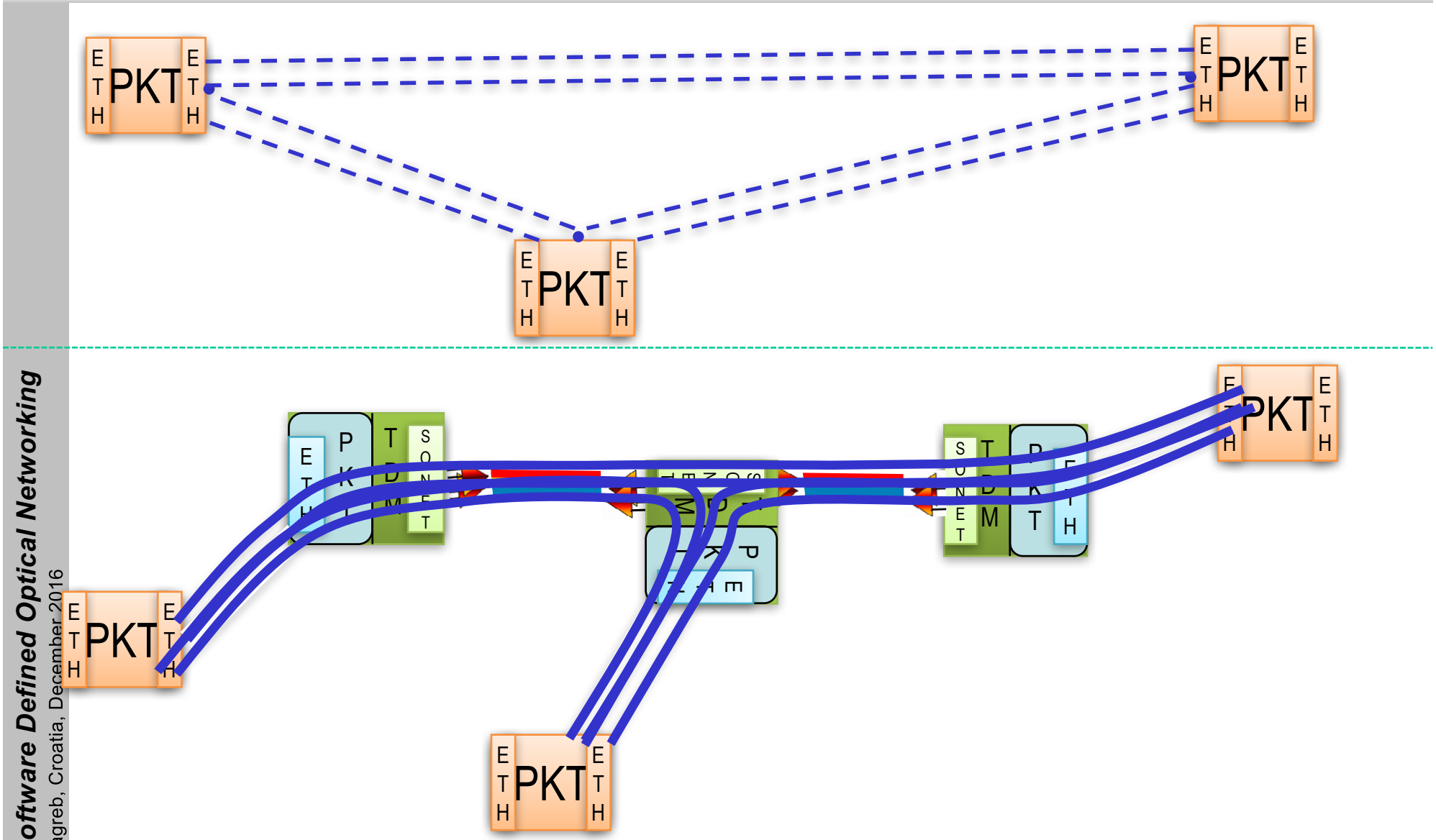
SDN and Optical Networking

ISP-1



SDN and Optical Networking

ISP-1



ISP# 1 redirects bw between the spare interfaces to dynamically create new links!!

SDN and Optical Networking
Business Models

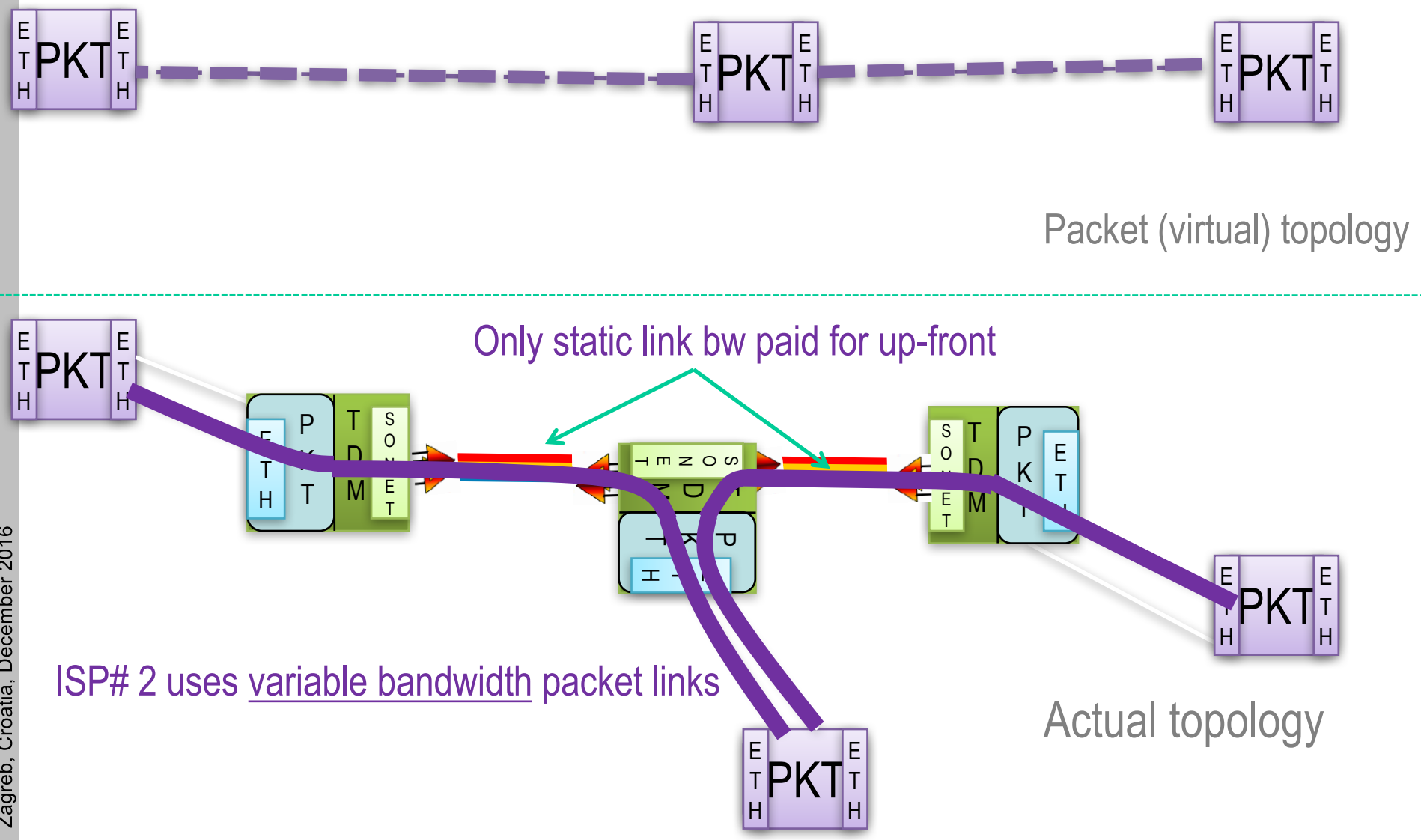
- ISP# 1 Rationale
  - Spare interfaces on Edge Switches
  - Cost less than bandwidth in core
  - Wants to share expensive core bandwidth between cheaper edge ports because that's more cost effective
  - Provides flexibility to create dynamic packet links where and when needed
  - Compare
    - 3 static links + 1 dynamic link = 3 ports/edge switch + static and dynamic bandwidth in core
    - 6 static links = 4 ports/edge switch + static core bandwidth
    - This gap increases with number of edge switches

SDN and Optical Networking
Business Models

- ISP# 1 pays for a slice = { bandwidth + TSP switching resources }
  - Part of the bandwidth is for static links between its edge packet switches (like ISPs do today)
  - and some of it is for redirecting bandwidth between the edge switches (unlike current practice)
  - The sum of both static bandwidth and redirected bandwidth is paid for up-front.
  - The TSP switching resources in the slice are needed by the ISP to enable the redirect capability.
- ISP# 2 pays for a slice = { bandwidth + TSP switching resources }
  - Only the bandwidth for static links between its edge packet switches is paid for up-front.
  - Extra bandwidth is paid for on a pay-per-use basis
  - TSP switching resources are required to provision/tear-down extra bandwidth
  - Extra bandwidth is not guaranteed

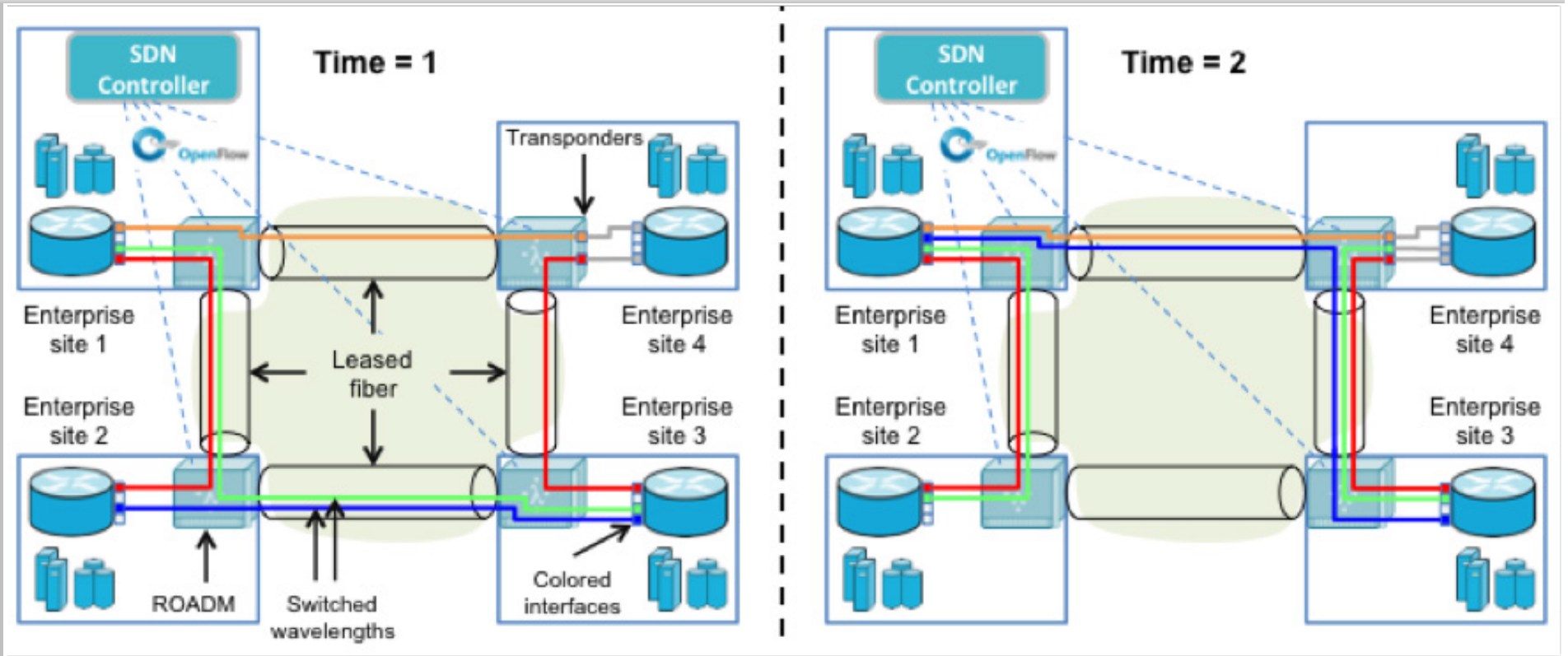
SDN and Optical Networking

ISP-2 Network



- ISP# 2 Rationale
  - uses variable bandwidth packet links
  - Has more bandwidth at the edge (say 10G) and pay for less bandwidth in the core up-front (say 1G)
  - Cost efficiency
    - ISP's today would pay for the 10G in the core up-front and then run their links at 10% utilization.
    - Instead they could pay for say 2.5G or 5G in the core, and ramp up when they need to or scale back when they don't – pay per use

SDN and Optical Networking
Use Case: Private Optical Networks

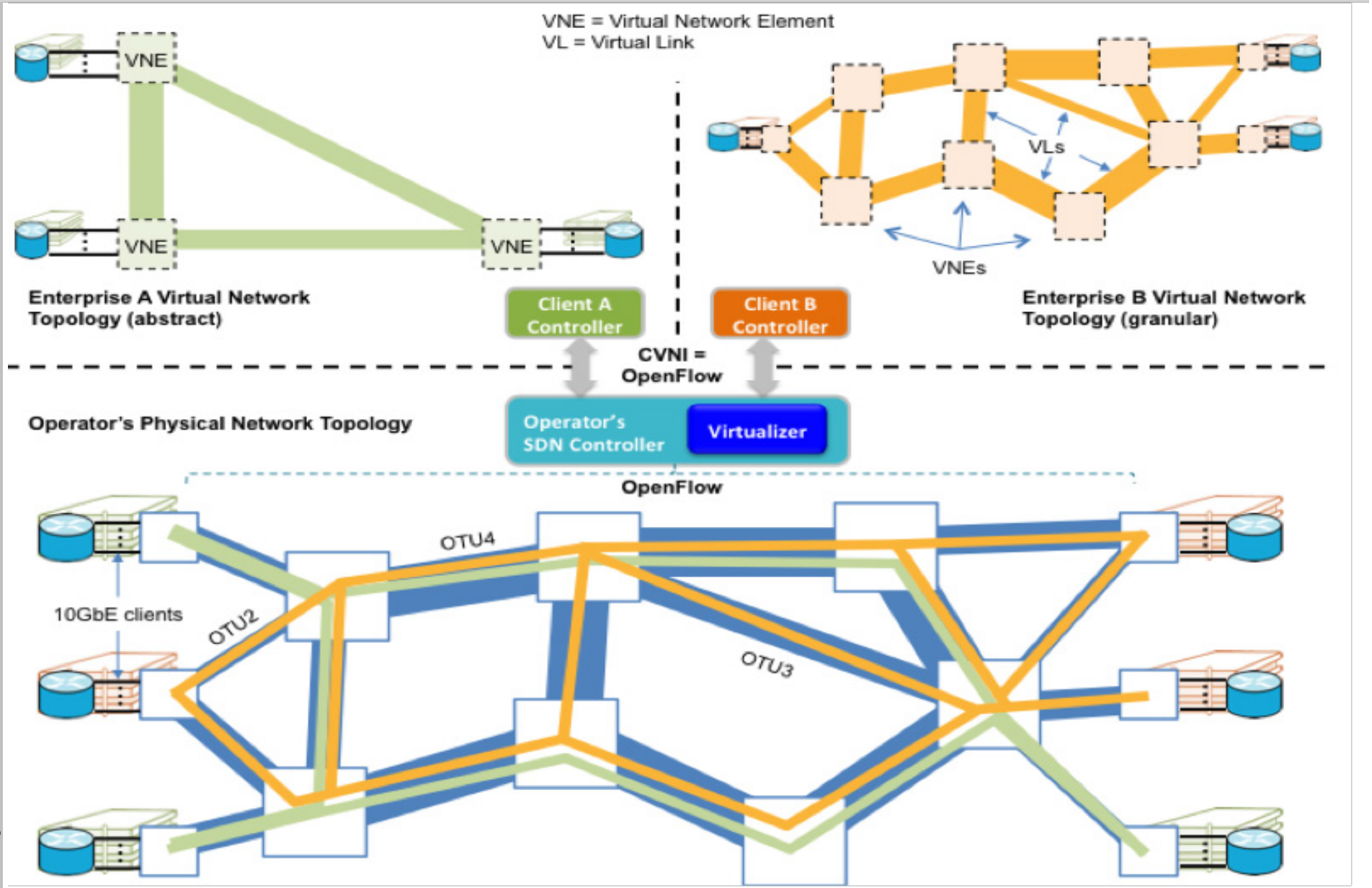


- Implement Bandwidth on Demand (requires expensive optical equipment)
  - Redirect wavelengths through ROADMs
  - Retune transceivers on optical transponders/muxponders or colored interfaces on OpenFlow-enabled switches/routers to a new wavelength
  - Activate new connections and deactivate old connections

SDN and Optical Networking
Use Case: Optical VPNs

- Extending VPN concept to optical layer
  - IP/MPLS VPNs should provide deterministic performance guarantees → Fiber can do that
  - Customer should be able to allocate capacity as if it were their own private network.
- Challenges
  - providing independent control and management across the end-to-end service
  - Provisioning O-VPN services across various layers and vendor Network Elements
- SDN based Network Virtualization allows
  - network operators to expose the underlying multi-layer, multi-vendor network as a virtual topology
  - To control the visibility and performance information (e.g., latency) of this topology
  - to create virtual slices of the network that deliver dedicated capacity
  - Provide the client with self-managed control of their end-to-end virtual networks
    - E.g. develop their own algorithms to manage connection paths

SDN and Optical Networking
Use Case: Optical VPNs

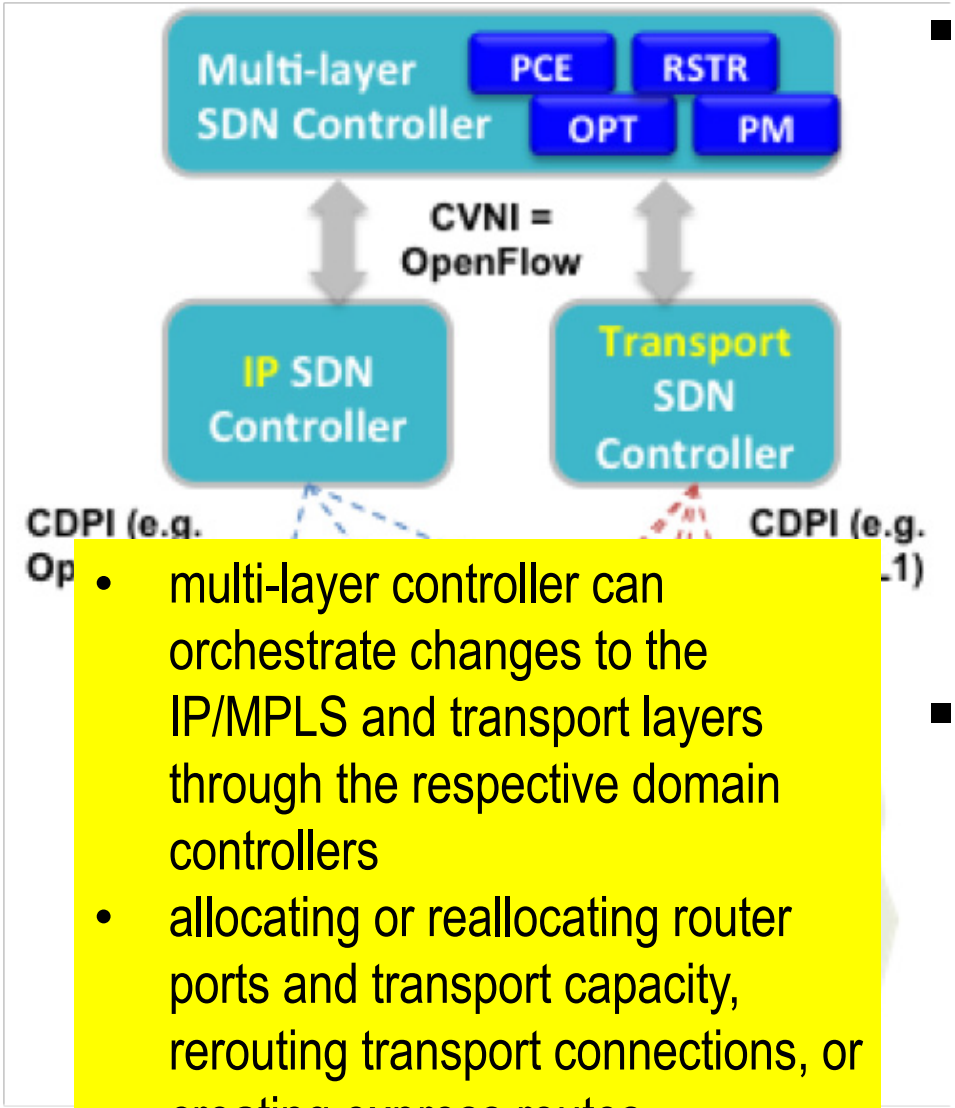


SDN and Optical Networking
Use Case: Multi Vendor IP/MPLS Plus Transport Optimization

- Optical Transport Networks must become more dynamic
  - Higher traffic demands require to follow better capacity demand
  - Example: coordinating multi-layer IP/MPLS and transport provisioning and restoration to address shifting aggregate IP flows
- State of the Art
  - IP/MPLS topologies assume a static transport layer
  - Physical router topologies and transport connections are sized based on 6- to 12-month traffic forecasts plus some extra margin
  - IP/MPLS traffic is 1+1 protected
  - Transport network provides two diverse optical paths to the routed network.
  - Efficiency < 40% Why???
- Problems
  - MPLS and Optical Transport have disjoint provisioning processes and network management systems
  - Path computation and traffic engineering are employed at each layer separately.
  - Control plane information for each layer is also held separately

SDN and Optical Networking

Use Case: Multi Vendor IP/MPLS Plus Transport Optimization

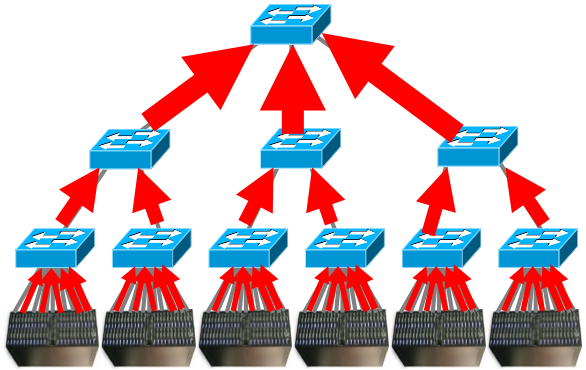


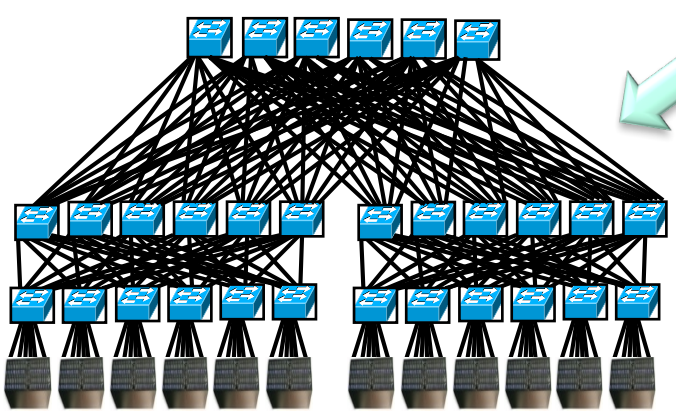
- multi-layer controller can orchestrate changes to the IP/MPLS and transport layers through the respective domain controllers
- allocating or reallocating router ports and transport capacity, rerouting transport connections, or creating express routes.

- SDN based dynamic transport layer
  - IP/MPLS and optical transport layers to be controlled synergistically
  - IP/MPLS pseudowire or
  - MPLS-TP tunnel to GFP-F to ODU to DWDM physical port
  - Can be performed on router with colored interfaces or a packet-optical transport node
- Each domain controller
  - provides detailed information about topology, latency, provisioned services, and performance to the multi-layer controller for path computation and restoration management

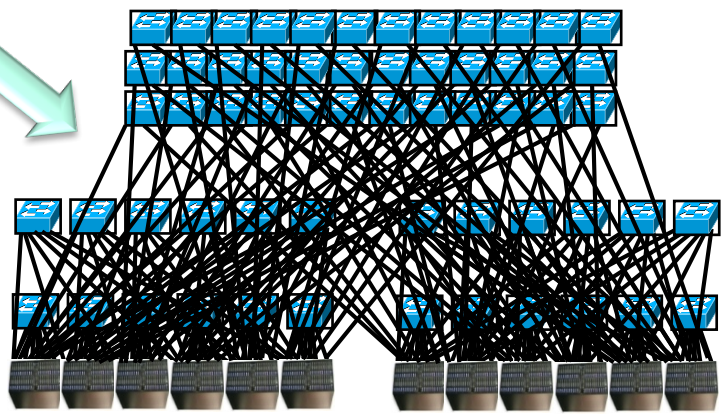
Optical Datacenter Networks

Current solutions for increasing data center network bandwidth





FatTree



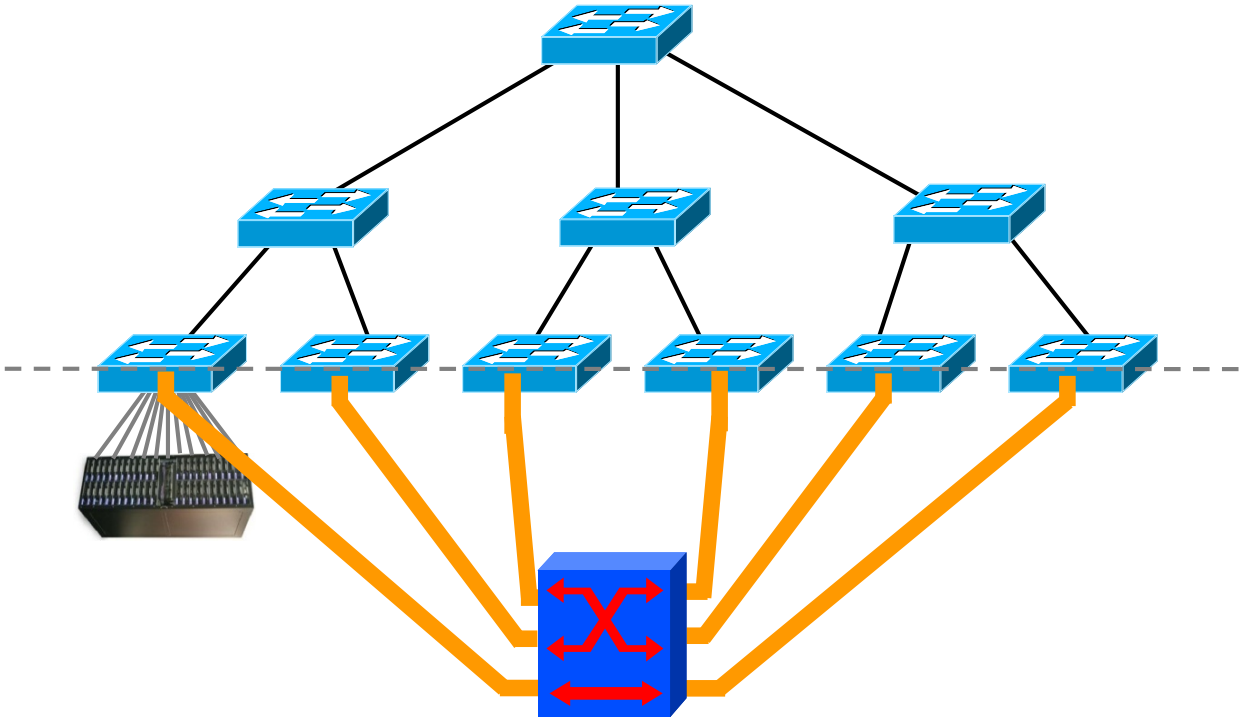
BCube

1. Hard to construct

2. Hard to expand

Optical Datacenter Networks

An alternative: hybrid packet/circuit switched data center network



- Idea:
  - Augment Datacenter Network with Optical Circuit Switch
  - Software Based Design

Optical Datacenter Networks

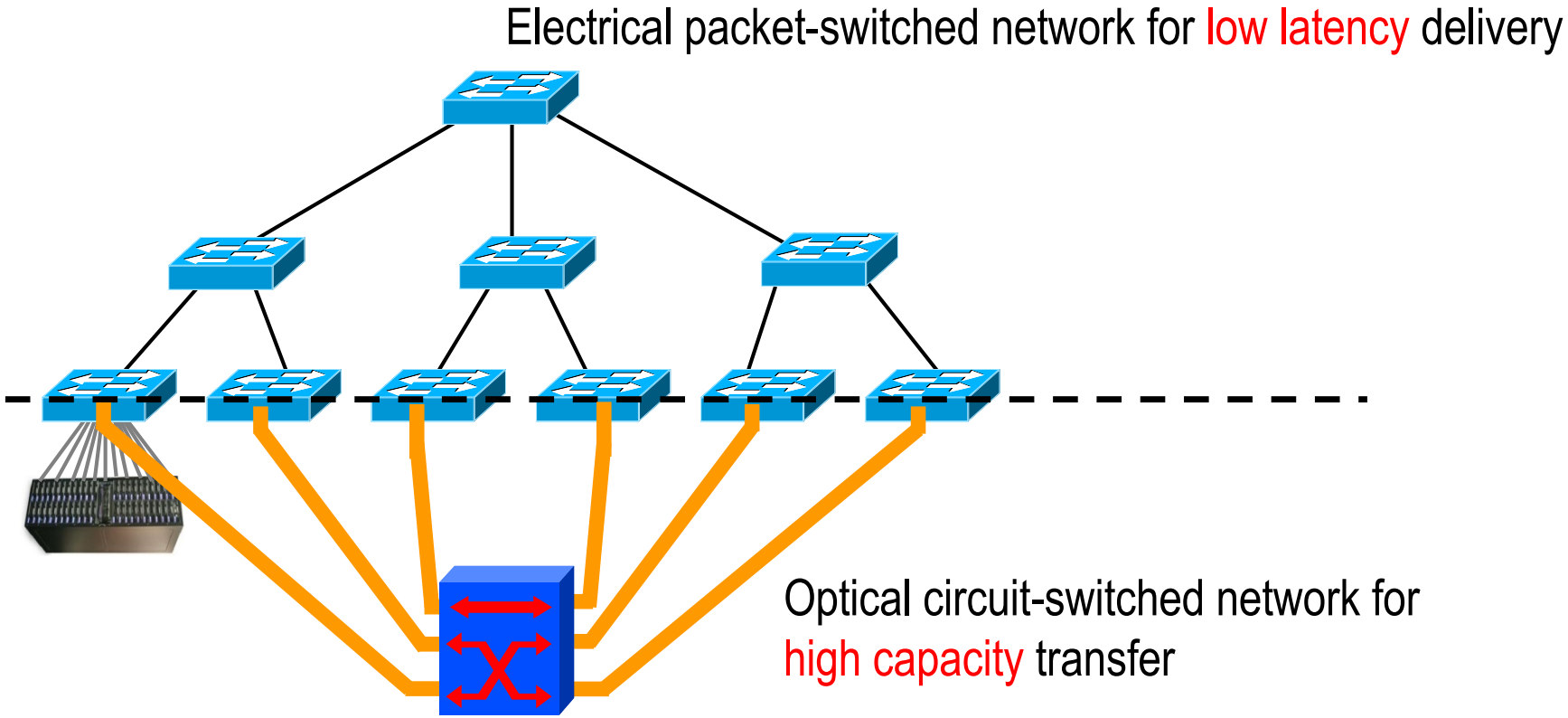
Optical circuit switching is promising despite slow switching time

- [IMC09][HotNets09]: *"Only a few ToRs are hot and most their traffic goes to a few other ToRs. …"*
- [WREN09]: *"…we find that traffic at the five edge switches exhibit an ON/OFF pattern… "*

Full bisection bandwidth at packet granularity may not be necessary

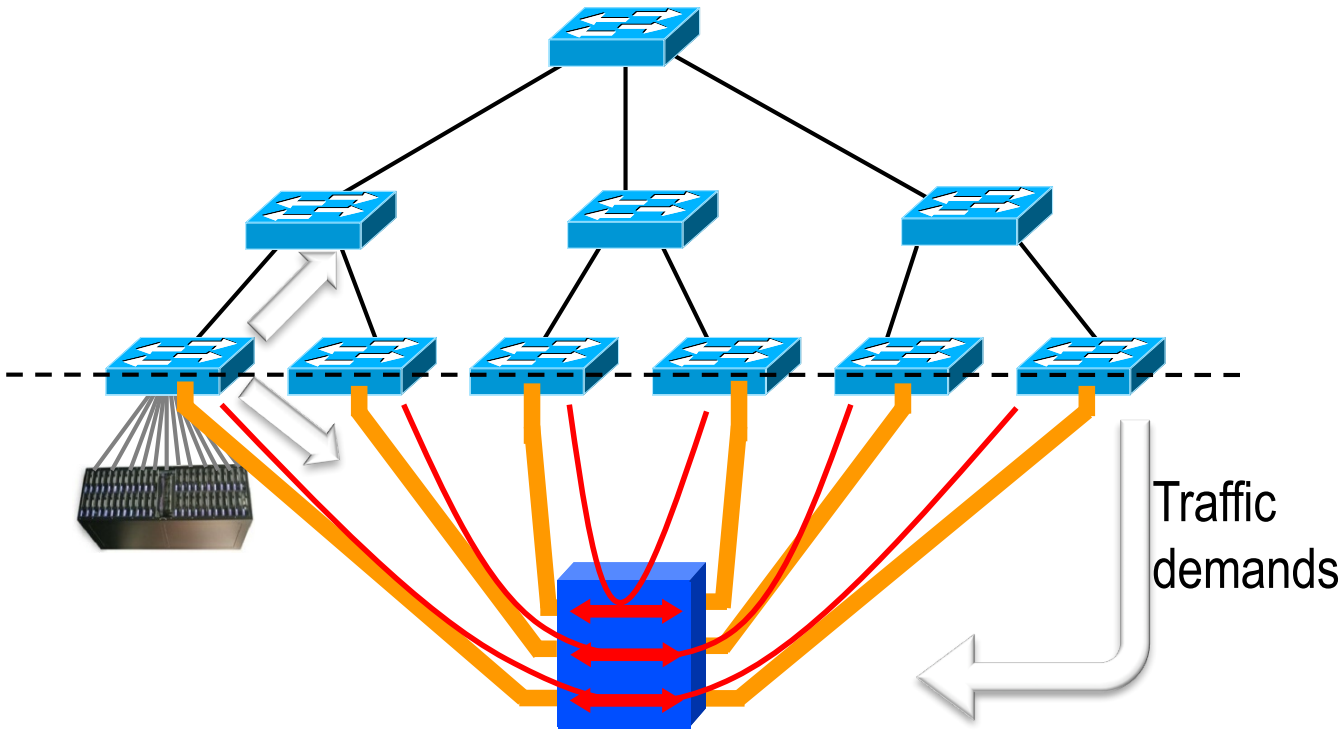
Optical Datacenter Networks

Hybrid packet/circuit switched network architecture



- Optical paths are provisioned rack-to-rack
  - A simple and cost-effective choice
  - Aggregate traffic on per-rack basis to better utilize optical circuits

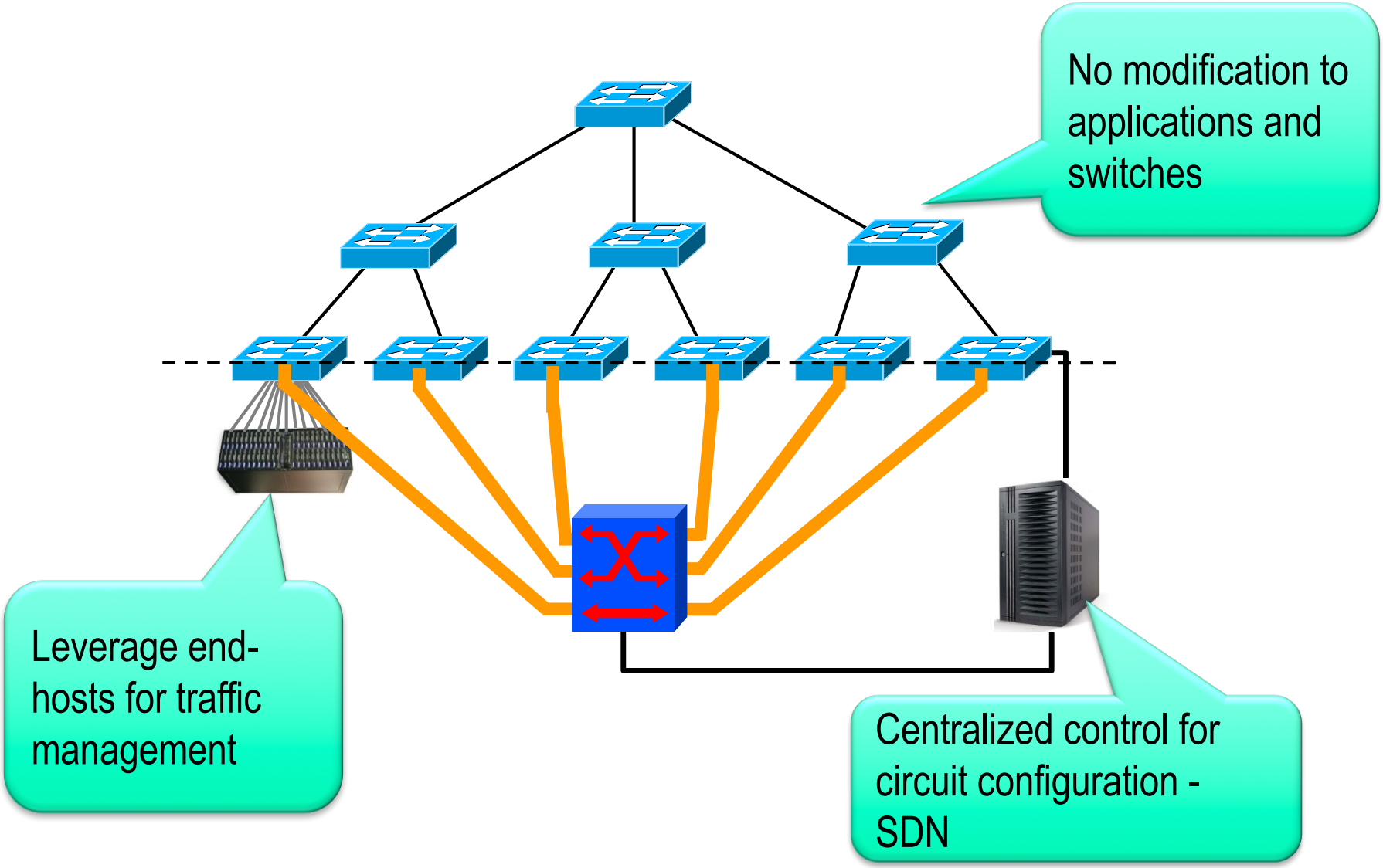
Optical Datacenter Networks
Design requirements



- Control plane:
  - Traffic demand estimation
  - Optical circuit configuration

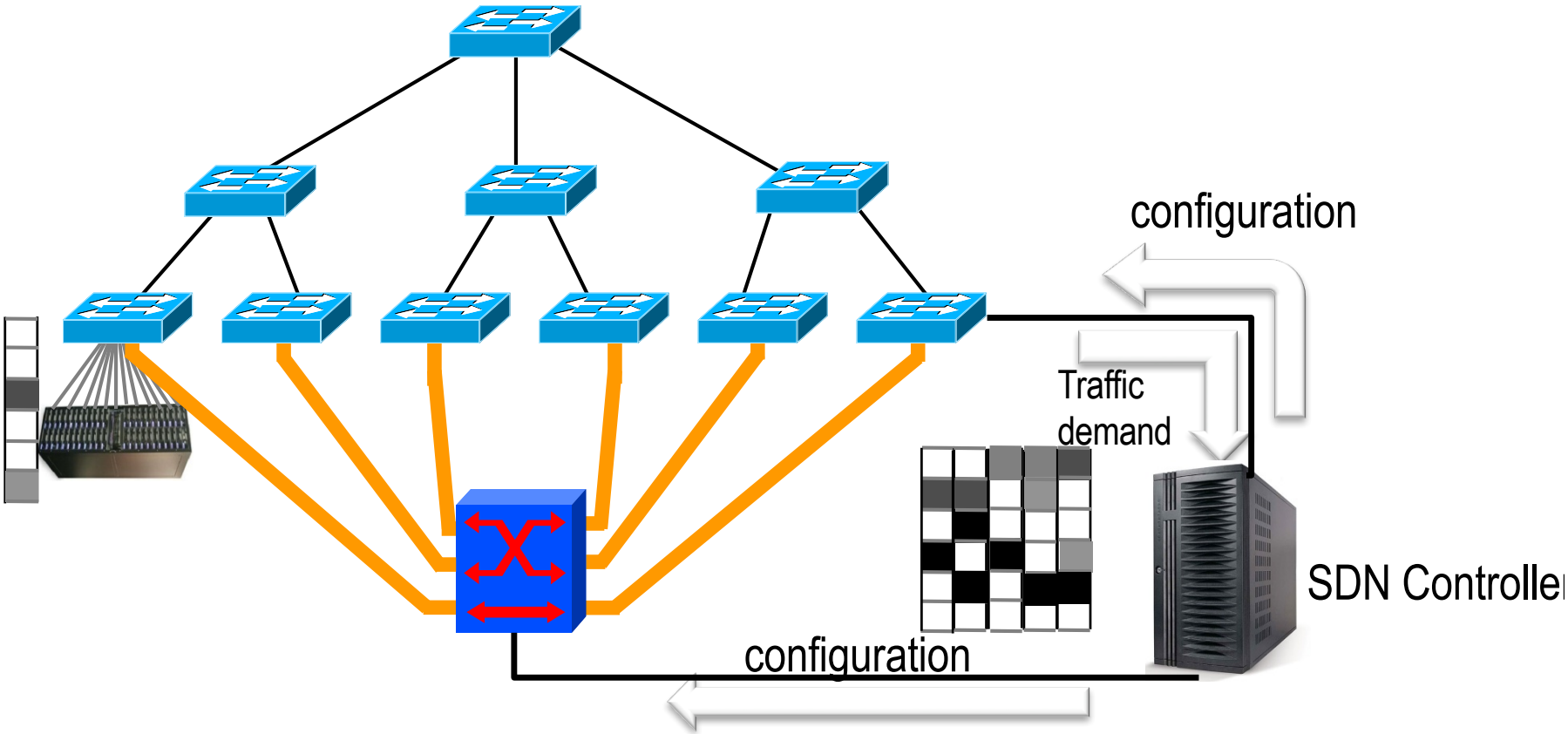
- Data plane:
  - Dynamic traffic de-multiplexing
  - Optimizing circuit utilization (optional)

Optical Datacenter Networks
c-Through (a specific design)



Optical Datacenter Networks

c-Through - optical circuit configuration



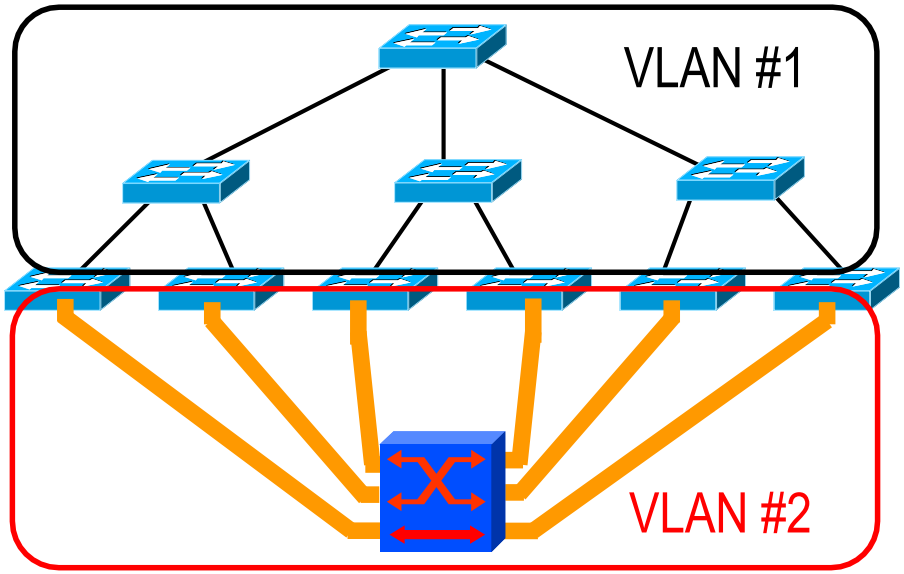
Use Edmonds' algorithm to compute optimal configuration

Many ways to reduce the control traffic overhead

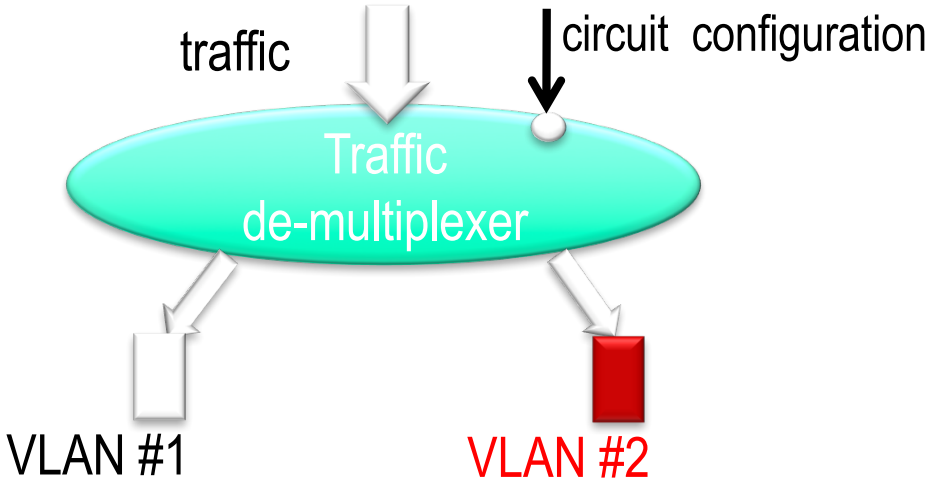
Optical Datacenter Networks

c-Through - traffic de-multiplexing

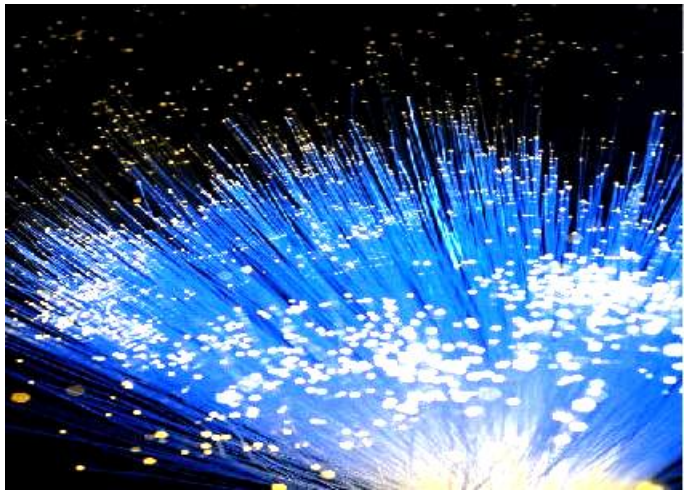
- VLAN-based network isolation:
  - No need to modify switches
  - Avoid the instability caused by circuit reconfiguration



- Traffic control on hosts:
  - Controller informs hosts about the circuit configuration
  - End-hosts tag packets accordingly

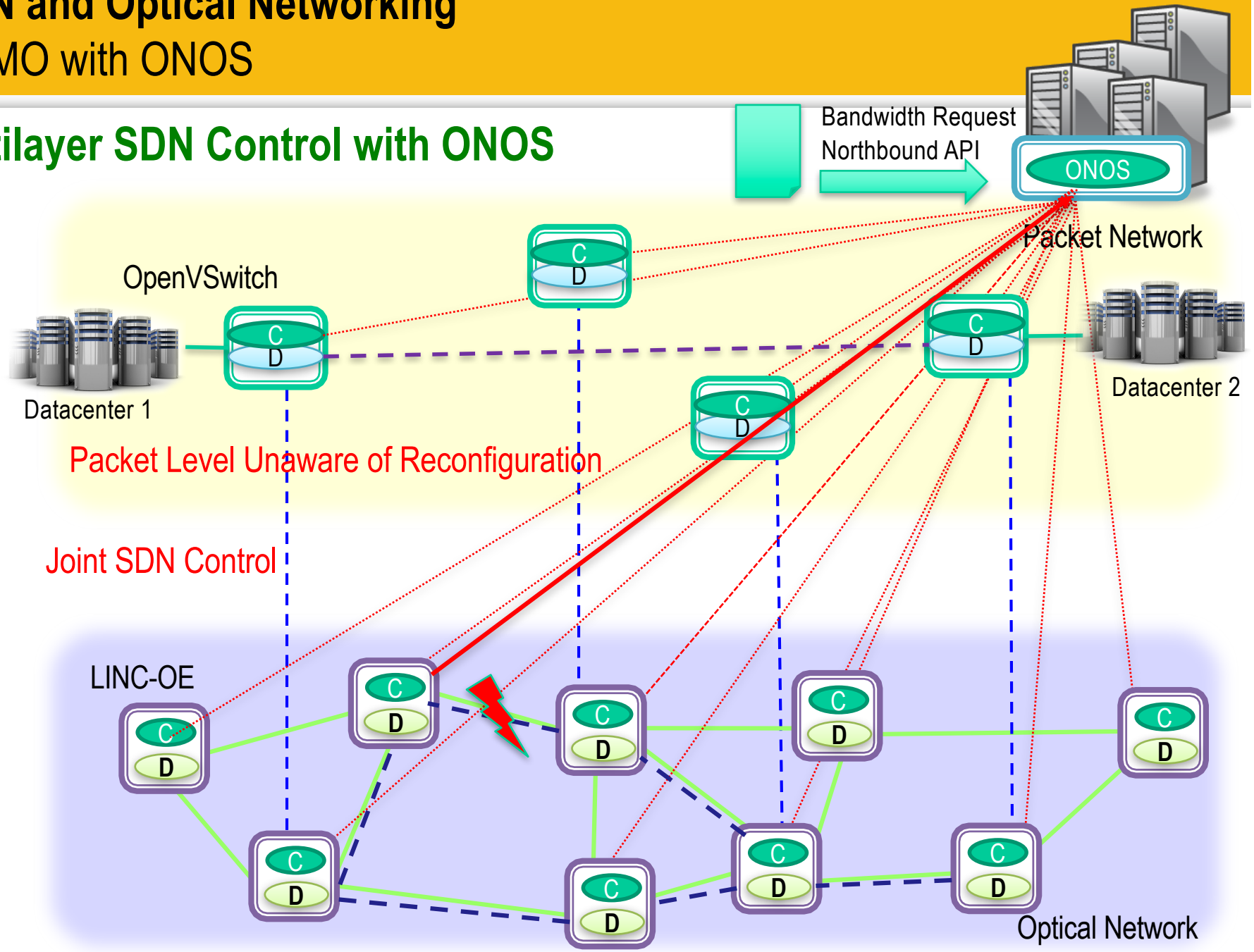


Optical Datacenter Networks
Summary



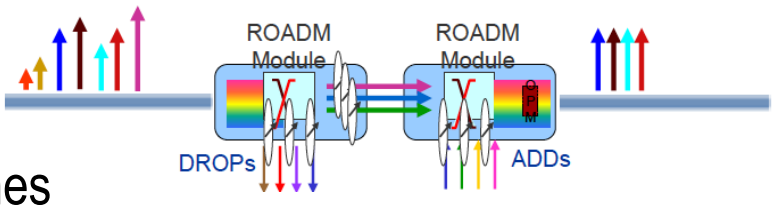
- Hybrid packet/circuit switched data center network
  - c-Through demonstrates its feasibility
  - Good performance even for applications with all to all traffic
- Future directions to explore:
  - The scaling property of hybrid data center networks
  - Making applications circuit aware
  - Power efficient data centers with optical circuits

Multilayer SDN Control with ONOS



LINC-OE Forwarding Model

- LINC-OE
  - Emulates optical layer topology from predefined table
  - Includes characteristics of optical cross connect and Packet to Optical Link Interface (ADD/DROP) configurable from OF1.3+
  - Supports failure scenarios
  - Mininet can reconfigure ports, links and switches



- Forwarding Model
  - Match/action abstraction for ROADMs: add/drop/forward
  - Match is about wavelength provision

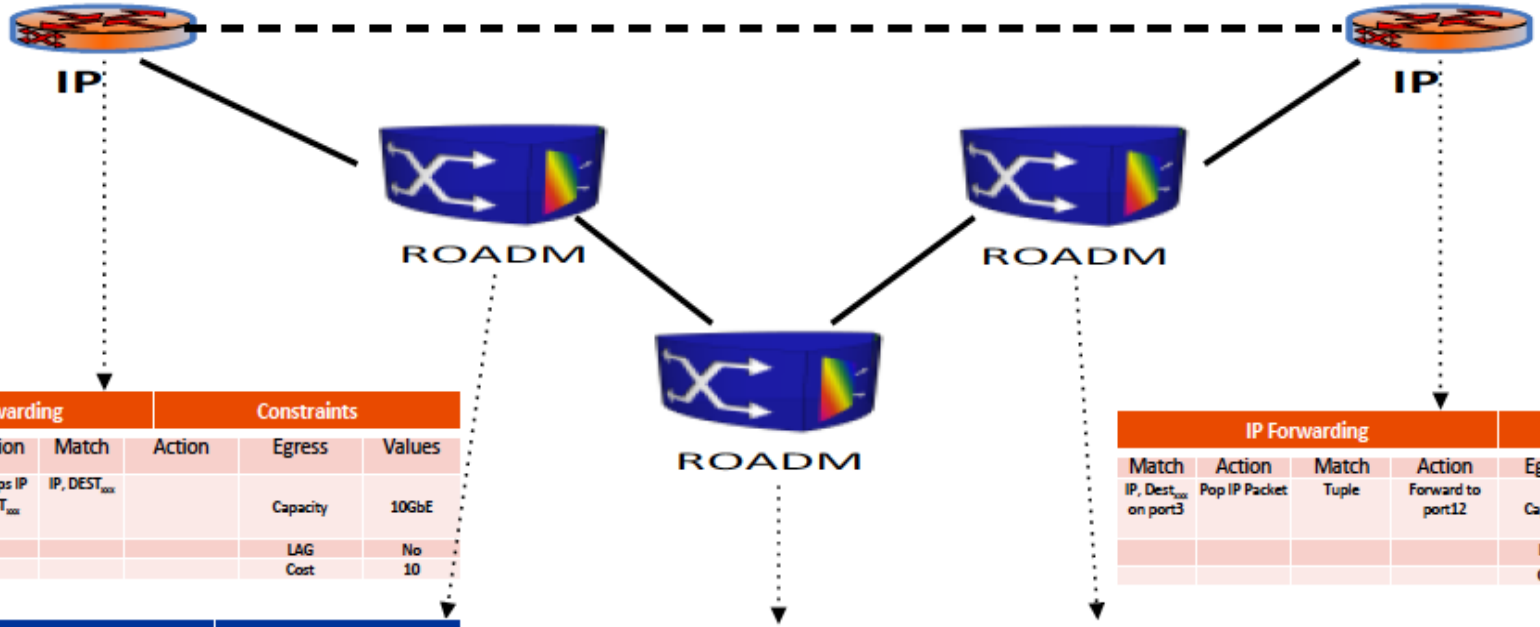
Add	
Match	Packet port and traffic type
Action	Transponder uses lambda and output to optical port

Forwarding	
Match	Optical port and lambda
Action	Output to optical port (easy to extend when considering regenerators)

Drop	
Match	Optical port and lambda
Action	Transponder uses lambda and output to packet port

SDN and Optical Networking

Packet and ROADM Layer Forwarding Model



IP Forwarding				Constraints	
Match	Action	Match	Action	Egress	Values
Tuple on port3	Encaps IP $DEST_{xxx}$	IP, $DEST_{xxx}$		Capacity	10GbE
				LAG	No
				Cost	10

Lambda Forwarding				Constraints	
Match	Action	Match	Action	Egress	Values
TPND port 5	PushOCH Lambda 66	Lambda 66	Forward to WDM port1	Capacity	100Gb
				MUXPDR	Yes
				Cost	250
				Loss	2.5
				NF	4.75
				PMD_{ext}	0.1
				$Disp_{ext}$	40

Lambda Forwarding		Constraints	
Match	Action	Egress	Values
l66 on port2	Forward WDM port6	Capacity	100Gb
		REGEN	No
		Cost	250
		Loss	2
		NF	4.25
		PMD_{ext}	0.05
		$Disp_{ext}$	25

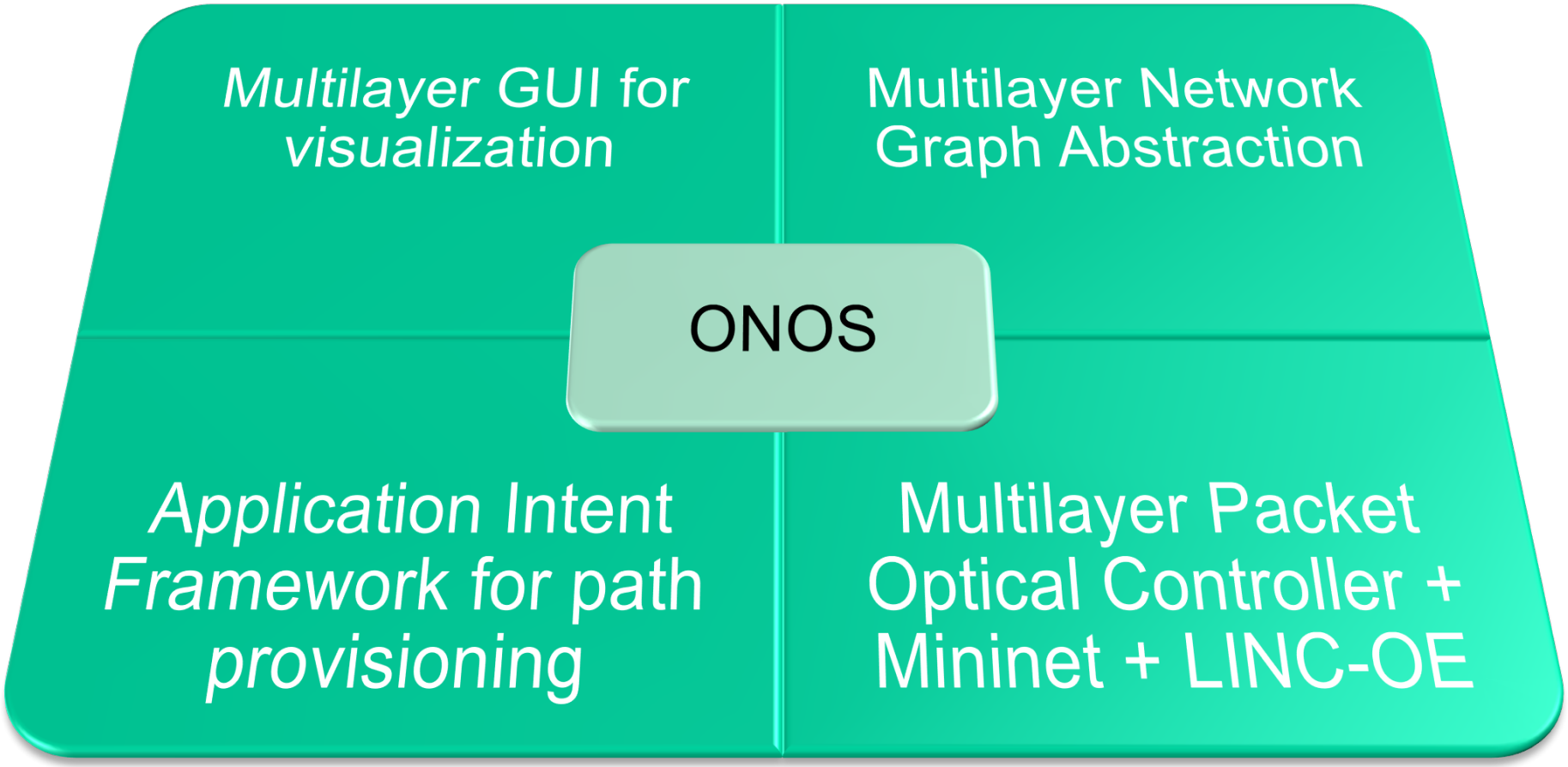
IP Forwarding				Constraints	
Match	Action	Match	Action	Egress	Values
IP, $Dest_{xxx}$ on port3	Pop IP Packet	Tuple	Forward to port12	Capacity	10GbE
				LAG	No
				Cost	10

Lambda Forwarding				Constraints	
Match	Action	Match	Action	Egress	Values
l66 on port3	Forward TPND port3	l66 on TPND port3	POP OCH and egress	Capacity	40Gb
				LAG	NA
				Cost	100

SDN and Optical Networking
Implementation Control Plane - ONOS

- Southbound for configuration of ROADMs
  - Followed ONF Optical Transport Working Group
  - OF1.3+ experimenter messages (some have been defined in OF1.4)
- Converged Topology
  - Control of packet and optical layers
  - Can add additional headers (e.g. OTN)
- Discovery
  - LLDP for L3 discovery (automatic)
  - Static configuration of L0 topology
- Path calculation over multi-layer graph
  - Constraints and resource management
  - Wavelength continuity, bandwidth, latency
- Restoration
  - Optical link failure causes disappearing packet links
  - First try packet layer restoration, if unsuccessful perform optical layer restoration

SDN and Optical Networking
DEMO with ONOS



thank you
danke
謝謝
ngiyabonga
teşekkür ederim
gracias
tapadh leat
спасибо
Баярлалаа
рахмат
bedankt
obrigado
dziękuję
hvala
mauruuru
sukriya
kop khun krap
go raibh maith agat
mochchakkeram
arigatō
takk
dakujem
merci
grazie
terima kasih
감사합니다
ありがとう
ευχαριστώ
merci
merci
merci