Exploiting Temporal Domain for a Transparent Fine-Grained Optical Network

Zhizhen Zhong

Tsinghua University & UC Davis

zhongzz14@mails.tsinghua.edu.cn , zzzhong@ucdavis.edu

02 Feb. 2018

Networks Lab Group Meeting

清華大學 Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

Outline

- The need for a transparent fine-grained optical network
- Past solutions on transparent fine-grained optical switching
- Our plan: Optical Time Slice Switching (OTSS)
- OTSS use cases
- Conclusion and discussions

Outline

- **The need for a transparent fine-grained optical network**
- Past solutions transparent fine-grained optical switching
- Our plan: Optical Time Slice Switching (OTSS)
- OTSS use cases
- Conclusion and discussions

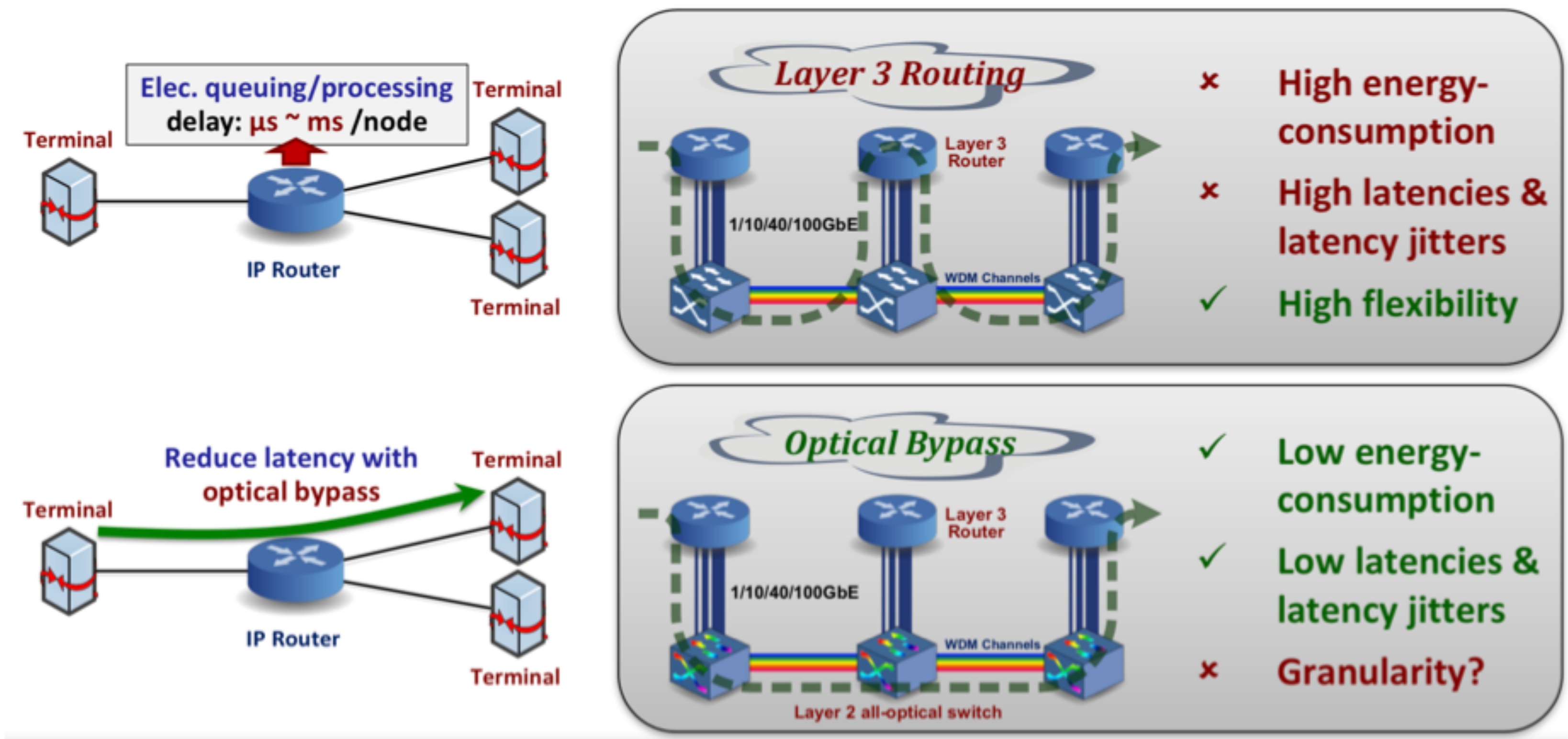
清華大學 Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

3

Transparent optical network: latency, energy and security

- **Electronic queuing and processing** delay increase **end-to-end latency**
- Introducing **all-optical switching** technology may reduce latency

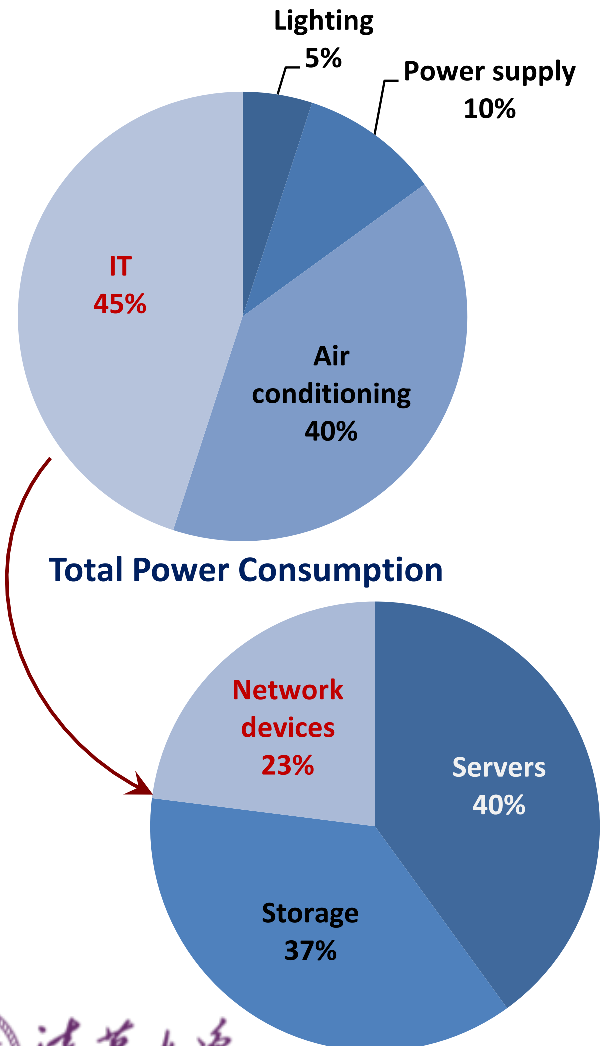


Tsinghua University

UC DAVIS UNIVERSITY OF CALIFORNIA

Advantages on energy efficiency of optical switching

DC power consumption



2010*: *235 billion kWh

2020*: *1 trillion kWh = 10 Three Gorges Power Stations

Electronic switching 13.5W/10Gbps port [1]

Optical switching  0.24W/10Gbps port [1]

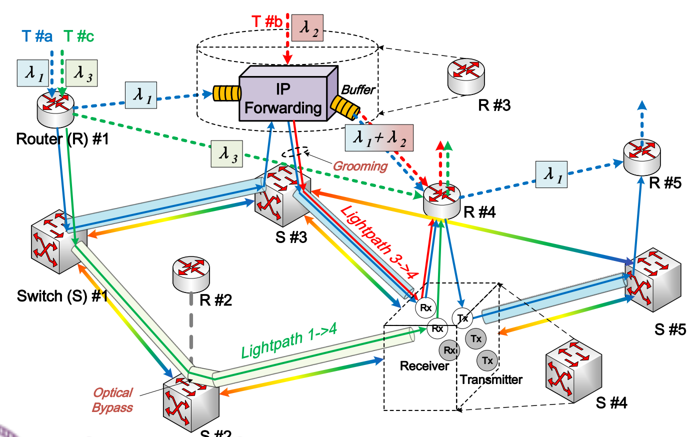
Tsinghua University

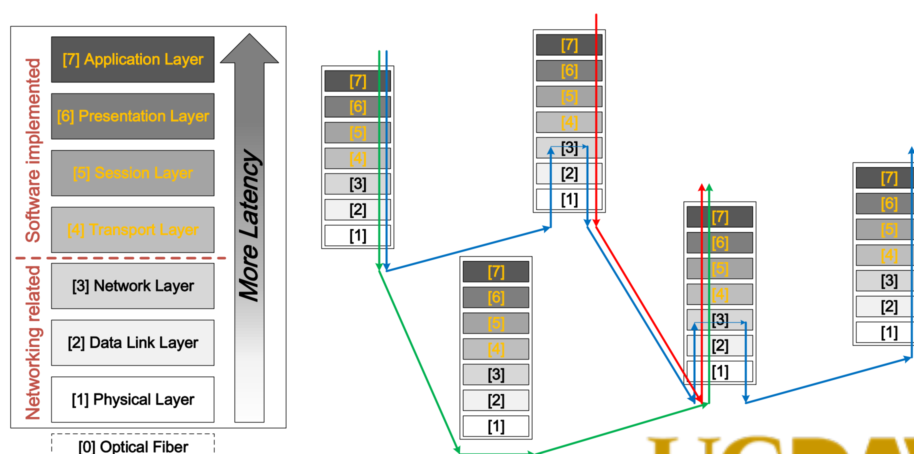
UCDAVIS UNIVERSITY OF CALIFORNIA

[1] Farrington N, et al. SIGCOMM Computer Communication Review. 2011 Oct 22; 41(4):339-50.

Needs for a transparent fine-grained optical network

- Optical networks: enormous transmission bandwidth.
  - High-order modulation (PAM4): increase per-channel capacity.
  - space-division-multiplexing: increase spatial channels.
- Mismatch between **application demands** and **optical channel capacity**.
  - Traffic grooming is the first proposal.
  - Drawback of grooming: energy, latency, security, etc.





Rethinking Optical Networks

- Resource granularity, flexibility and usability.
- Match between request and resource. Avoid waste.
- Flexi-grid technologies: provides more channels with fine granularities.

清華大學 Tsinghua University

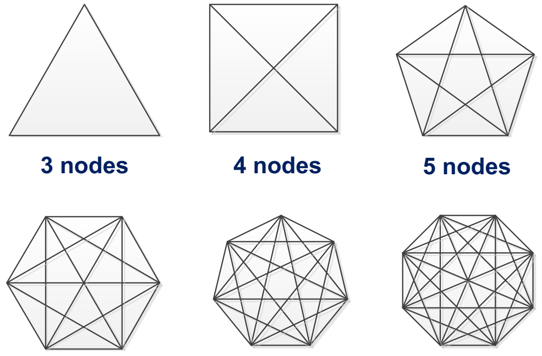
UCDAVIS UNIVERSITY OF CALIFORNIA

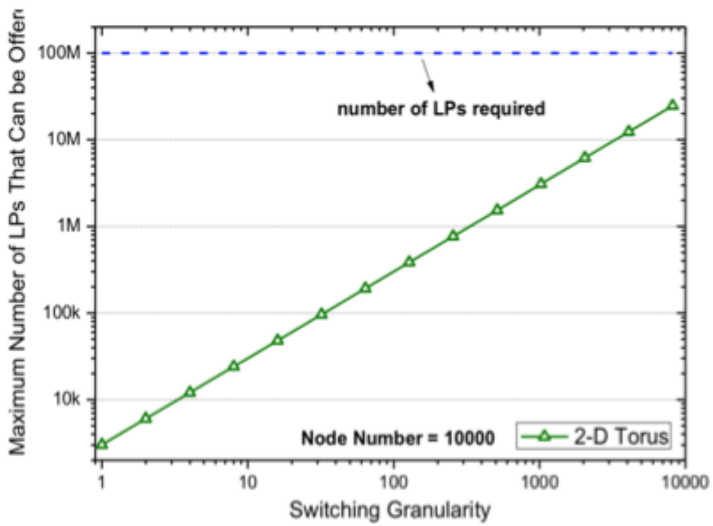
Outline

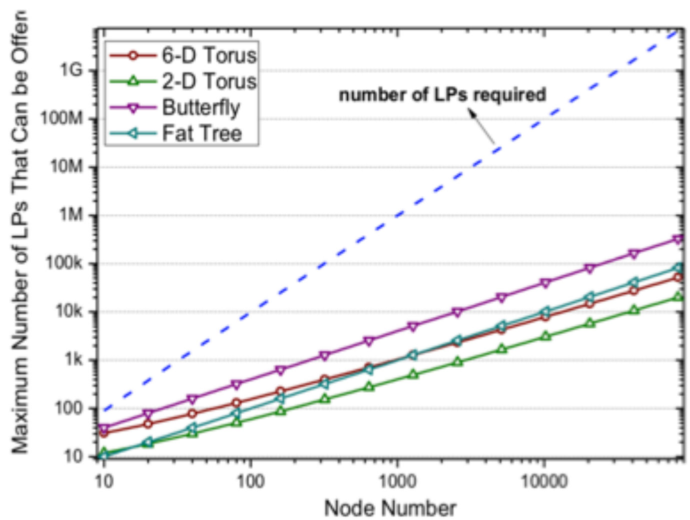
- The need for a transparent fine-grained optical network
- **Past solutions transparent fine-grained optical switching**
- Our plan: Optical Time Slice Switching (OTSS)
- OTSS use cases
- Conclusion and discussions

Transparent connections vs. network scale

- Number of LPs **required vs.** that **can be offered.**
  - All-to-all communication.
  - Set up **Dedicated end-to-end lightpath** for each node pair.
  - **Topologies**: 2-D/6-D torus, Butterfly, Fat tree
    Nodal degree: d= 4 (for 6-D torus: d=12).







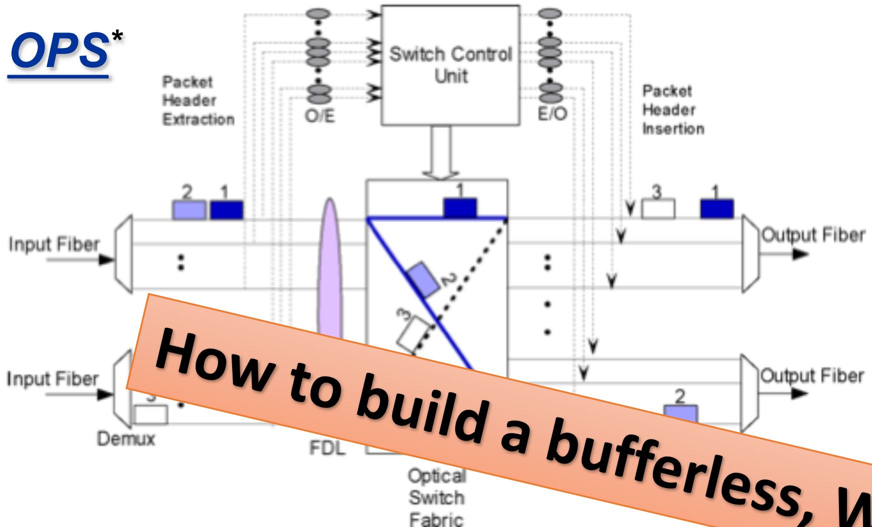
Nan Hua, **Zhizhen Zhong**, Xiaoping Zheng, "Enabling Low Latency at Large-Scale DataCenter and High-Performance Computing Interconnect Networks Using Fine-GrainedAll-Optical Switching Technology," ***ONDM***, May 2017.

Tsinghua University

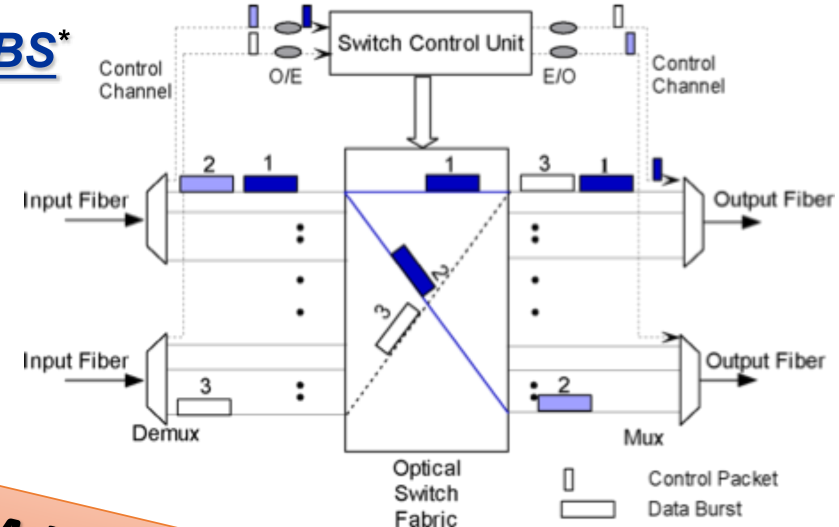
UC DAVIS
UNIVERSITY OF CALIFORNIA

OPS & OBS

OPS*



OBS*



How to build a bufferless, WDM-like TDM optical network?

- Temporal-domain switching techniques, ...nd **OBS**, can provide **much finer granularity** than freq... domain switching techniques.
- **OPS** requires **optical buffer** (not mature) to avoid collisions.
- For **OBS**, **burst loss rate** could be **very high** at a heavy load without buffer.

Outline

- The need for a transparent fine-grained optical network
- Past solutions transparent fine-grained optical switching
- **Our plan: Optical Time Slice Switching (OTSS)**
- OTSS use cases
- Conclusion and discussions

11

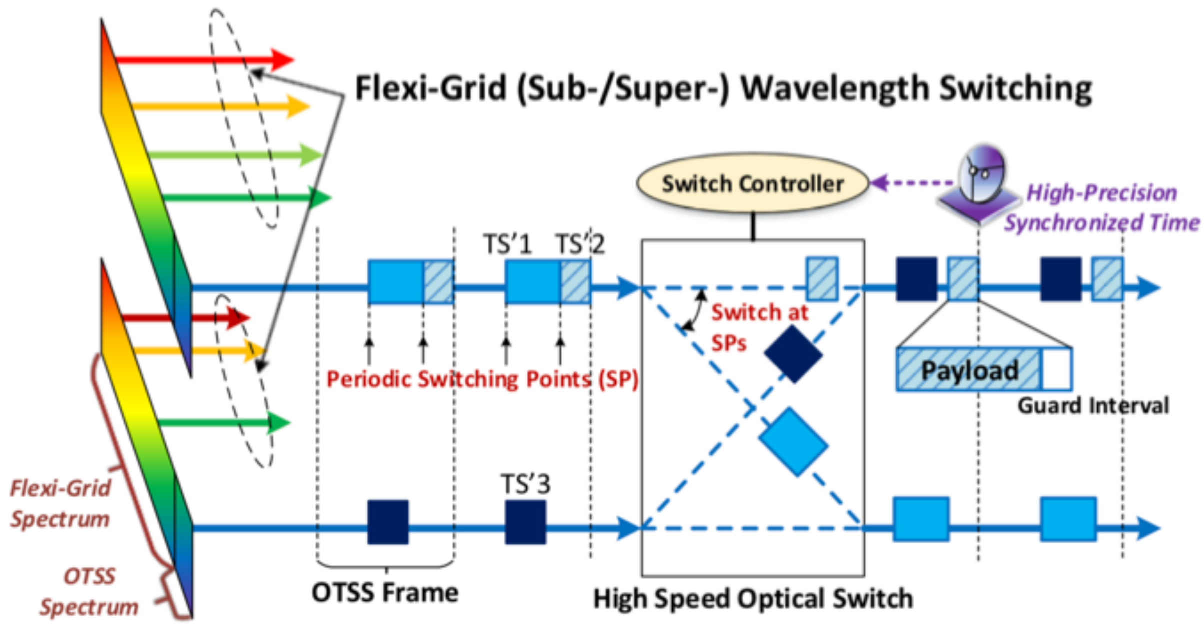
Build a transparent, bufferless, fine-grained, WDM-like network

- Why WDM can avoid collision?
  - Wavelength channels are separated by a global coordinate.
  - (frequency! All the same in different nodes)
- Time synchronization: a global coordinate in temporal domain
  - Definite time, all nodes are synchronized for a global coordinate.
- **Temporally-statistical multiplexing for asynchronous transmission based on synchronized global time.**
  - **We call it: Optical Time Slice Switching (OTSS).**

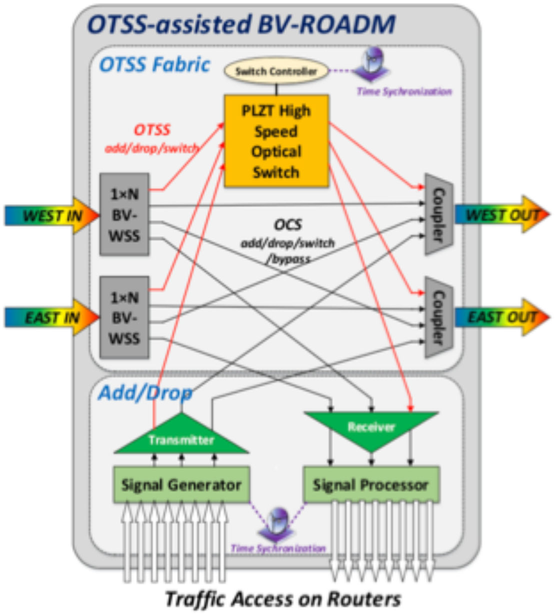
12

OTSS principle

- Designing a WDM-like TDM switching paradigm
  - WDM: all nodes have same frequency coordinate.
  - OTSS: all node should have same time coordinate!



OTSS switching principle



OTSS node architecture

Zhizhen Zhong, Nan Hua, Zhu Liu, Wenjing Li, Yanhe Li, and Xiaoping Zheng, "Evolving Optical Networks for Latency-Sensitive Smart-Grid Communications via Optical Time Slice Switching (OTSS) Technologies," ***IEEE CLEO-PR / OECC / PGC***, *Aug. 2017*. **(IEEE Photonics Society 1st place Poster Award)**

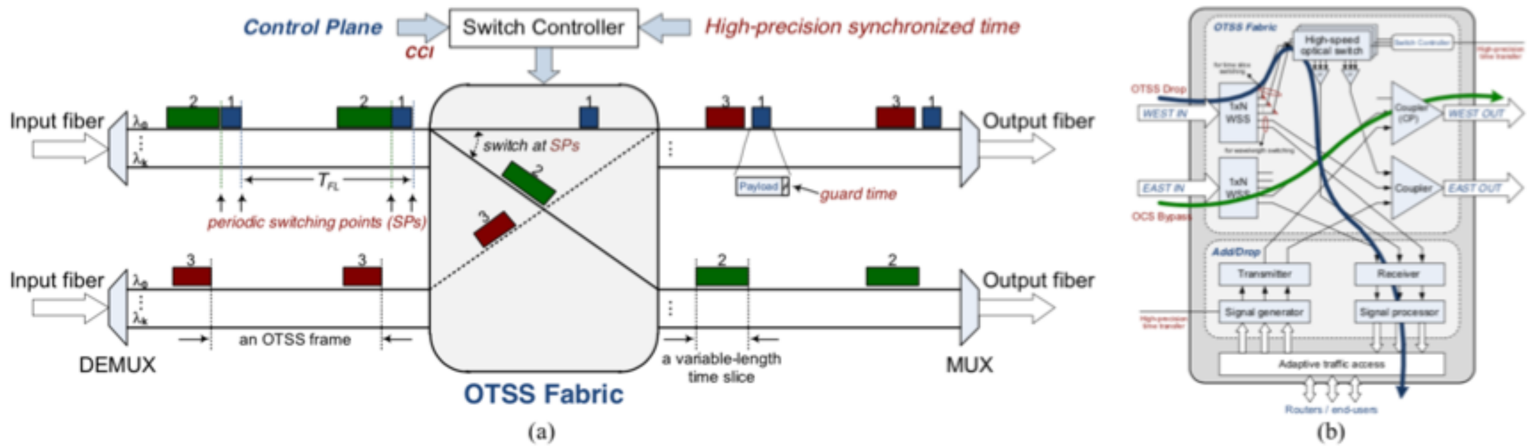
Tsinghua University

UC DAVIS UNIVERSITY OF CALIFORNIA

13

OTSS principle: node switching

- The optical transmission channels are organized into repetitive OTSS frames.
- Each OTSS frame contains one or several variable-length time slice(s).
- When a time slice arrives, the switch controller sends control signals to the OTSS fabric at the precise time to direct the time slice to the expected output port.



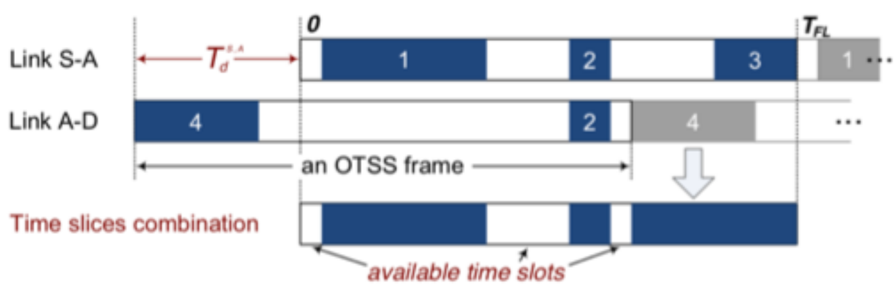
Hua, Nan, and Xiaoping Zheng. "Optical time slice switching (OTSS): an all-optical sub-wavelength solution based on time synchronization." In *Asia Communications and Photonics Conference*, pp. AW3H-3. Optical Society of America, 2013.

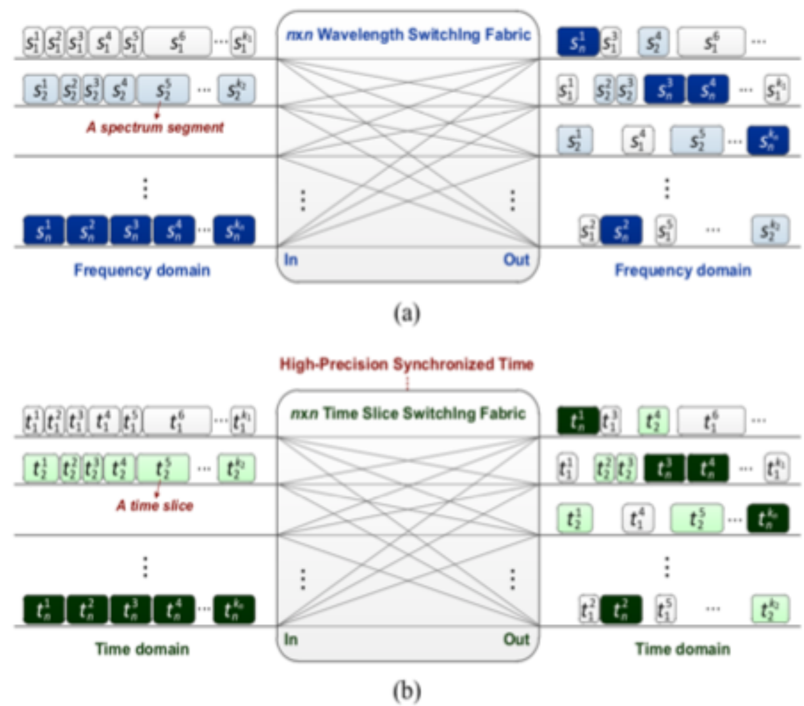
Tsinghua University 清華大學

UCDAVIS UNIVERSITY OF CALIFORNIA

OTSS data plane: Routing and Time slice allocation (RTA) problem

- **Time-slice contiguity constraint**: a request's time slice should be contiguous on temporal domain.
- **Time-slice continuity constraint**: time slice should be continuous along different links.
- **Time slice shifting constraint**: signal propagation delay may result in time slice shifting in OTSS frame.





Flexi-grid wavelength switching and optical time slice switching

OTSS data plane: Routing and Time slice allocation (RTA) problem

➢ Mathematical formulations and Algorithms design

Parameters:

- $G(N, E)$: network topology in a unidirectional graph, where N and E denotes the set of nodes and MMF fiber links, respectively.
- R: set of traffic requests.
- s_r, d_r, b_r: source, destination, and bandwidth of traffic request r, $r \in R$.
- X: accumulated crosstalk threshold for direct-detection receivers.
- M: set of supporting modes by a MMF.
- $d(i,j)$: length of fiber link (i,j) of fiber link, $(i,j) \in E$.
- C: maximum link capacity.
- $Y(m_1, m_2)$: modal crosstalk of mode m_1, m_2, in dB.
- Max: a maximum number.
- η_1, η_2 : parameters for optimization sequence, $\eta_1 \gg \eta_2$.
- T: OTSS frame length.
- S: length of the minimum time slice.

Variables:

- $\lambda_{(i,j)}^{r,m,t}$: binary, which equals one if request r uses mode m and time slot t on fiber link (i,j).
- $\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t}$: binary, which equals one if request r_1 on mode m_1, r_2 on mode m_2 have crosstalk on time slot t of fiber link (i,j).
- $\theta_{(i,j)}^{r_1,r_2,m_1,m_2}$: binary, which equals one if request r_1 on mode m_1, r_2 on mode m_2 have crosstalk on fiber link (i,j).
- ρ_r : binary, which equals one if request r is accepted.

Objective: Maximize network throughput first, then minimize network resource usage.

$$Maximize:\ \eta_1 \cdot \sum_{r\in R} \rho_r \cdot b_r - \eta_2 \sum_{r\in R}\sum_{m\in M}\sum_{t\in T}\sum_{(i,j)\in E} \lambda_{(i,j)}^{r,m,t} \quad (1)$$

Constraints:

$$\sum_{j\in N}\sum_{m\in M}\sum_{t\in T}\lambda_{(i,j)}^{r,m,t} - \sum_{j\in N}\sum_{m\in M}\sum_{t\in T}\lambda_{(j,i)}^{r,m,t} = \begin{cases} \rho_r \cdot b_r, i = s_r \\ -\rho_r \cdot b_r, i = d_r \\ 0, i \neq s_r, d_r \end{cases}, \quad \forall r \in R \quad (2)$$

$$\sum_{m\in M}\sum_{j\in N}\lambda_{(s_r,j)}^{r,m,t} = \sum_{m\in M}\sum_{i\in N}\lambda_{(i,d_r)}^{r,m,t}, \forall r \in R, t \in T \quad (3)$$

$$\sum_{m\in M}\sum_{j\in N}\lambda_{(z,j)}^{r,m,t} = \sum_{m\in M}\sum_{i\in N}\lambda_{(i,z)}^{r,m,t}, \forall r \in R, t \in T, z \in N \setminus \{s_r, d_r\} \quad (4)$$

$$\sum_{m\in M}\sum_{j\in N} m \cdot \lambda_{(s_r,j)}^{r,m,t} = \sum_{m\in M}\sum_{i\in N} m \cdot \lambda_{(i,d_r)}^{r,m,t}, \forall r \in R, t \in T \quad (5)$$

$$\sum_{m\in M}\sum_{j\in N} m \cdot \lambda_{(z,j)}^{r,m,t} = \sum_{m\in M}\sum_{i\in N} m \cdot \lambda_{(i,z)}^{r,m,t}, \forall r \in R, t \in T, z \in N \setminus \{s_r, d_r\} \quad (6)$$

$$\sum_{r\in R}\lambda_{(i,j)}^{r,m,t} \leq 1, \quad \forall m \in M, t \in T, (i,j) \in E \quad (7)$$

$$\sum_{t\in T}\left(\left\langle \lambda_{(i,j)}^{r,m,t} \neq \lambda_{(i,j)}^{r,m,(t+1)\%N_m} \right\rangle \times 1\right) \leq 2, \quad \forall r \in R, m \in M, (i,j) \in E \quad (8)$$

$$\sum_{t\in T}\left(\left\langle \sum_{m\in M}\lambda_{(i,j)}^{r,m,t} \neq \sum_{m\in M}\lambda_{(i,j)}^{r,m,t+1} \right\rangle \times 1\right) \leq 2, \quad \forall r \in R, (i,j) \in E \quad (9)$$

$$\sum_{m'\in M}\sum_{t'\in T}\lambda_{(i,j)}^{r,m',t'} - (\lambda_{(i,j)}^{r,m,t} - 1) \cdot Max \cdot C \cdot S/T \geq b_r, \forall r \in R, m \in M, t \in T, (i,j) \in E \quad (10)$$

$$\sum_{r_2\in R}\sum_{(i,j)\in E}\sum_{m_1\in M}\sum_{m_2\in M}\theta_{(i,j)}^{r_1,r_2,m_1,m_2} \cdot d(i,j) \cdot Y(m_2, m_1) \leq X, \quad \forall r_1 \in R \quad (11)$$

$$\sum_{t\in T}\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t}/Max \leq \theta_{(i,j)}^{r_1,r_2,m_1,m_2} \leq \sum_{t\in T}\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t}, \forall r_1, r_2 \in R, m_1, m_2 \in M, (i,j) \in E \quad (12)$$

$$\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t} \geq \lambda_{(i,j)}^{r_1,m_1,t} + \lambda_{(i,j)}^{r_2,m_2,t} - 1, \forall r_1, r_2 \in R, m_1, m_2 \in M, t \in T, (i,j) \in E \quad (13)$$

$$\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t} \leq \lambda_{(i,j)}^{r_1,m_1,t}, \forall r_1, r_2 \in R, m_1, m_2 \in M, t \in T, (i,j) \in E \quad (14)$$

$$\beta_{(i,j)}^{r_1,r_2,m_1,m_2,t} \leq \lambda_{(i,j)}^{r_2,m_2,t}, \forall r_1, r_2 \in R, m_1, m_2 \in M, t \in T, (i,j) \in E \quad (15)$$

Tab. 2. Description of STSC(ψ_i, ψ_j, T_d^i)

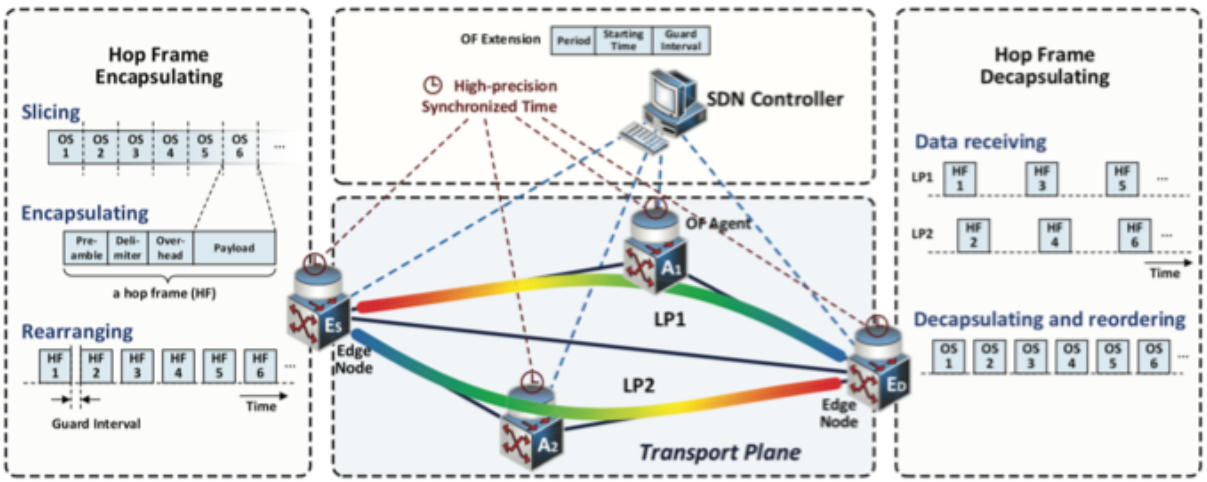
Input: ψ_i, ψ_j, *the* propagation *delay of* e_i (T_d^i).
Output: *the combination results* ψ_c.
1) Shift the time slices of link e_j :
  for $k = 1$ *to* K_j ***do***
    $t_{j_k}^{start} \leftarrow t_{j_k}^{start} - T_d^i$; $t_{j_k}^{end} \leftarrow t_{j_k}^{end} - T_d^i$;
2) Combine ψ_i *and* ψ_j *, and output the combination result:*
  output $\psi_c \leftarrow \psi_i \cup \psi_j$;
  return;

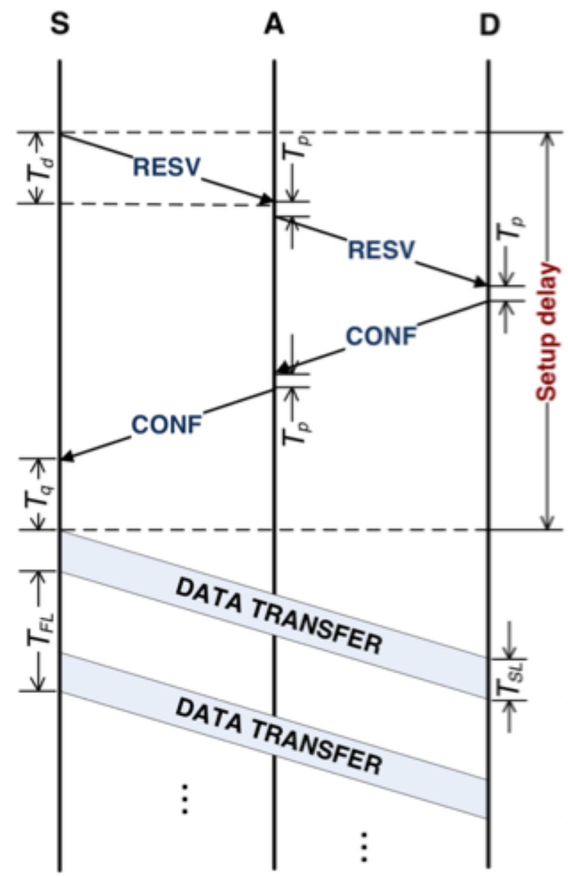
Tab. 3. Description of FR-TSA(G, ψ, p, ϕ_B)

Input: $G(N,E)$, $\psi(E)$, *a fixed path* $\vec{p} = (e_1, e_2, ..., e_H)$, *connection request* ϕ_B *with required bandwidth of* B.
Output: the start time of the allocated time slot s (t_s^{start}).
1) Combine the time slices of the links on $\vec{p}$:
  $T_d^c \leftarrow 0$; $\psi_c \leftarrow NULL$;
  for $i = 1$ *to* H ***do***
    $\psi_c \leftarrow$ STSC(ψ_c, ψ_i, T_d^c) ; $T_d^c \leftarrow T_d^c + T_d^i$;
2) Allocate a time slot and output its start time:
  for $k = 1$ *to* $K_c - 1$ ***do***
    if $t_{c_{k+1}}^{start} - t_{c_k}^{end} \geq B$ ***then***
      output $t_s^{start} \leftarrow t_{c_k}^{end}$; ***return***;
  output *NULL*; ***return***;

16

OTSS control plane: unified control and signaling





Signaling procedures of RSVP

- A software-defined unified control architecture with time synchronization.
- RSVP signaling for distributing different temporally switching command to different nodes.

Li, Yao, Nan Hua, Yiqiao Song, Shangyuan Li, and Xiaoping Zheng. "Fast lightpath hopping enabled by time synchronization for optical network security." *IEEE Communications Letters* 20, no. 1 (2016): 101-104.

Tsinghua University

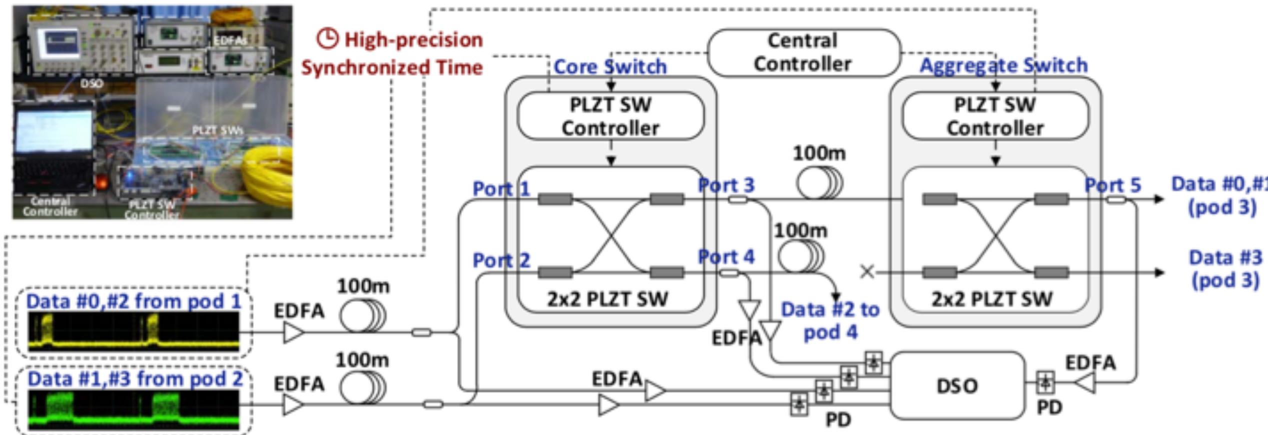
UCDAVIS UNIVERSITY OF CALIFORNIA

Enabling technology: time synchronization

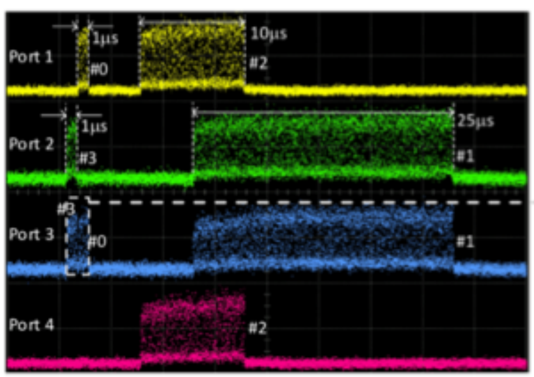
- **Time synchronization: mature**
  - precision should be finer than the smallest time slice in OTSS.
  - GPS receiver: 50ns.
  - IEEE 1588 or 1588v2: 100ns
- **Fast optical switch: mature**
  - PLZT high speed switch: 10ns switching speed.
  - MO high speed switch: 100us switching speed.

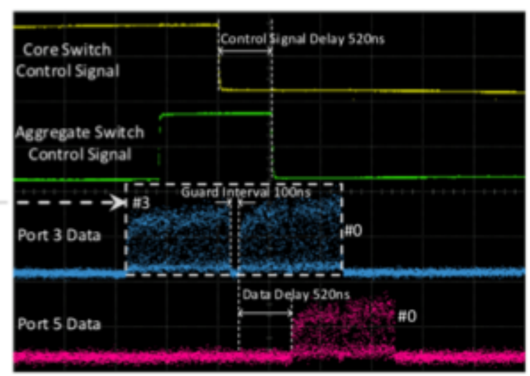
18

OTSS experimental validation



Experimental demonstration





Time slice switching (separation and aggregation)

Transmission and control signal delay

Tsinghua University

Li, Yao, Nan Hua, and Xiaoping Zheng. "Fine-grained all-optical switching based on optical time slice switching for hybrid packet-OCS intra-data center networks." In *Optical Fiber Communication Conference*, pp. W3J-5. 2016.

UC DAVIS UNIVERSITY OF CALIFORNIA

19

Outline

- The need for a transparent fine-grained optical network
- Past solutions transparent fine-grained optical switching
- Our plan: Optical Time Slice Switching (OTSS)
- **OTSS use cases**
- Conclusion and discussions

20

Overview of switching paradigms

	OCS	*OBS*	*OPS*	*OTSS*
Switching type	Circuit (wavelength)	Burst	Packet	Circuit (time slice)
Granularity	Coarse	Moderate	**Fine**	Moderate
Optical buffer	**Not Required**	**Not Required***	Required	**Not Required**
Bandwidth utilization	Low	**High**	**High**	**High**
Transfer guarantee	**Guaranteed**	Not guaranteed	Not guaranteed	**Guaranteed**
Control overhead	**Low**	Moderate	High	**Low**
Processing requirements	**Low**	Moderate	High	**Low**
Requirements of switching speed	**Low (*ms*)**	Moderate (μs~*ms*)	High (*ns*)	Moderate (μs~*ms*)
Time synchronization	**Not Required**	**Not Required**	**Not Required**	Required

*Buffer can improve OBS performance

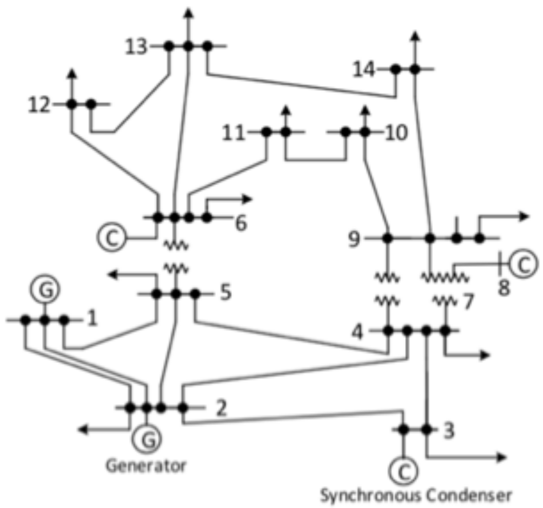
- **optical packet/burst switching**: delay-sensitive flows with relatively small sizes (query, coordination and control state messages).
- **Optical circuit/λ switching**: delay-insensitive bandwidth-hungry data transfer (file backup and virtual machine migration).
- **Optical time slice switching**: combine the advantages of capacity and energy consumption of optical fiber/λ switching and the flexibility of electrical packet switching.

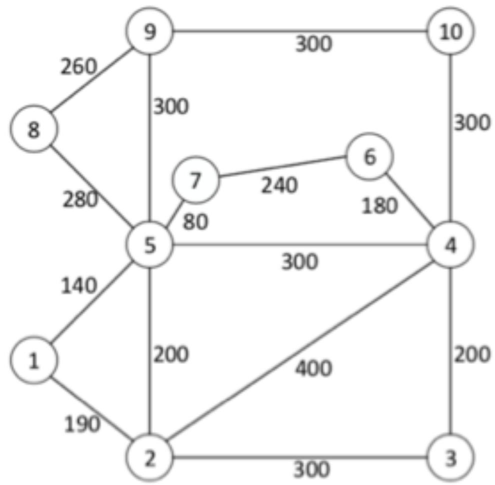
清華大學 Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

21

1. Fine-grained Communications for Smart Grid





- **Small bandwidth**
- **Stringent latency**

Application	Paradigm	Bandwidth	Latency
Teleprotection	P2P	∼ 500 Kb/s	8-10 ms
Load Shedding for Underfrequency	P2P & HS	∼ 500 Kb/s	10 ms
SCADA	P2P & HS	∼ 800 Kb/s	100-200 ms
Smart Metering	HS	∼ 500 Kb/s	250-1000 ms
File Transfer	Random	200-1000 Mb/s	≥ 1000 ms

Zhizhen Zhong, Nan Hua, Zhu Liu, Wenjing Li, Yanhe Li, and Xiaoping Zheng, "Evolving Optical Networks for Latency-Sensitive Smart-Grid Communications via Optical Time Slice Switching (OTSS) Technologies," ***IEEE CLEO-PR / OECC / PGC****, Aug. 2017*. **(IEEE Photonics Society 1st place Poster Award)**

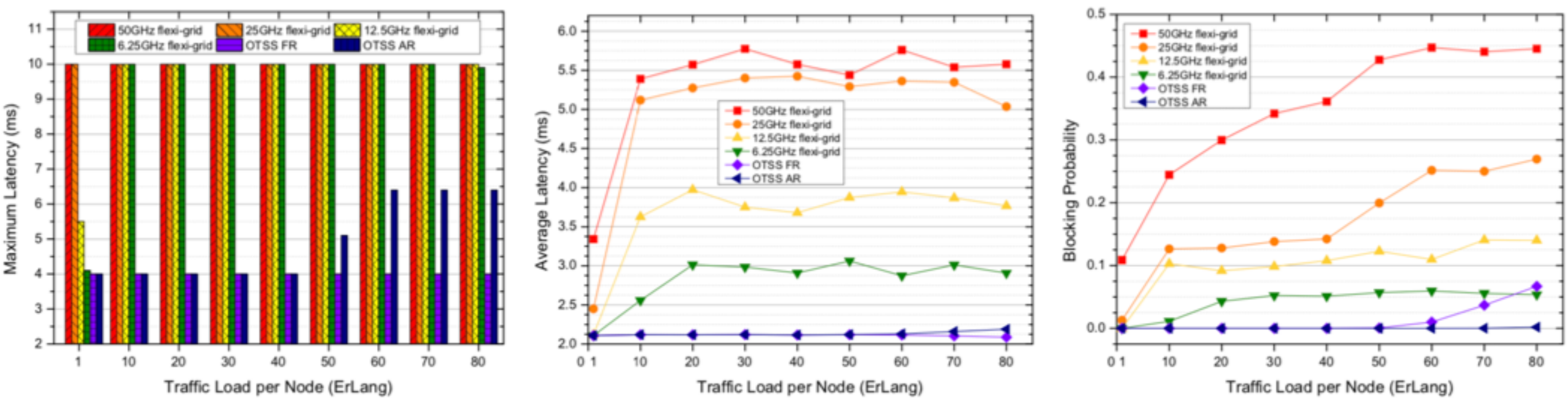
Tsinghua University

UC DAVIS UNIVERSITY OF CALIFORNIA

22

1. Fine-grained Communications for Smart Grid

- We study the 10 ms latency bound for security-related smart-grid traffic.
- Those requests larger than 10 ms, can not be served.
- Compare OTSS and flexi-grid, using the same amount of spectrum.



(a) Maximum latency vs. traffic load. (b) Average latency vs. traffic load. (c) Blocking probability vs. traffic load.

Zhizhen Zhong, Nan Hua, Zhu Liu, Wenjing Li, Yanhe Li, and Xiaoping Zheng, "Evolving Optical Networks for Latency-Sensitive Smart-Grid Communications via Optical Time Slice Switching (OTSS) Technologies," ***IEEE CLEO-PR / OECC / PGC***, *Aug. 2017*. **(IEEE Photonics Society 1st place Poster Award)**

Tsinghua University

UC DAVIS UNIVERSITY OF CALIFORNIA

23

2. Transparent Fine-grained MDM Network

- **OTSS + Mode Division Multiplexing**
- **OTSS provide fine-grained transparent channels breaking mode-coupling wavelength channels constraint.**
- **Can be adopted in datacenters.**

Zhizhen Zhong, Nan Hua, Yufang Yu, Zhongying Wu, Juhao Li, Haozhe Yan, Shangyuan Li, Ruijie Luo, Jialong Li, Yanhe Li, and Xiaoping Zheng, "Throughput Scaling for MMF-Enabled Optical Datacenter Networks Based on Time-Slicing-Based Crosstalk Mitigation," to be presented, ***OFC***, Mar. 2018.

24

Background

Datacenter is the basic infrastructure for future information-based society.

Cloud Computing
5G mobile networks
Video streaming
Interconnected car
Internet of things

1. More datacenters being built
2. Datacenter itself evolves to be larger

清華大學
Tsinghua University

UC DAVIS
UNIVERSITY OF CALIFORNIA

Motivation

Roadmap to Ubiquitous Datacenters:
Single-Mode or Multi-Mode for intra-DC interconnection networks?

Intra-DC optical networks:
- Short reach (100m~1km fiber length)
- Large connectivity (millions of network nodes)
- Large channel capacity (10Gb/s, 100Gb/s, 400Gb/s)

Cisco 10G SFP+ Transceiver*	
Single-Mode SFP+	$7
Multi-Mode SFP+	$6

Cisco 40G SFP+ Transceiver*	
Single-Mode QSFP+	$340
Multi-Mode QSFP+	$55

Cisco 100G SFP+ Transceiver*	
Single-Mode QSFP28	$2800
Multi-Mode QSFP28	$400

*Price from Fiberstore: www.fs.com

SMF and MMF are both mature technologies. Cost matters!

For future widely-located large-scale cloud/fog datacenters, MMF is a better choice.

26

Motivation

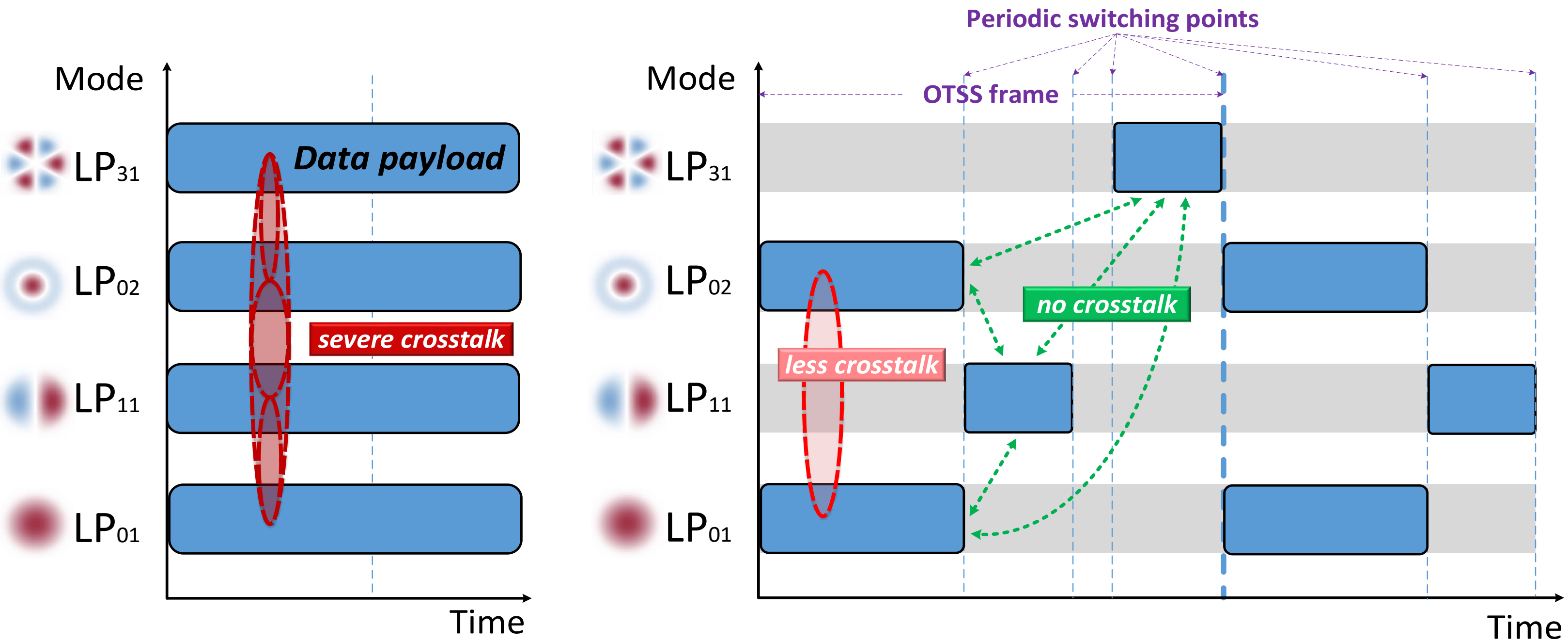
MMF suffers severe modal crosstalk in direct detection transmission systems.

- Modal crosstalk accumulate along propagation path, and can induce OSNR degradation at the receiver end [4,5].
- Such crosstalk must be prevented, as it cannot be fully undone by electrical signal processing after direction detection[6].

Is there a way to apply MMF into DC?

[4] B. Franz, "Mode Group Division Multiplexing in Graded-Index Multimode Fibers," Bell Labs Technical J., 2013.
[5] F. Yaman, et al., "Impact of Modal Crosstalk and Multi-Path Interference on Few-Mode Fiber Transmission," OFC, 2012.
[6] K.-P. Ho, et al., "Linear propagation effects in mode-division multiplexing systems," J. Lightw. Technol., 2014.

Time-Slicing-Based MDM



- **Basic idea**: stagger utilized modes in temporal domain via synchronized time slices.
- **Design Principle**: avoiding using modes with high crosstalk in the same time slice, while changing utilized modes in different time slices by switching at selected switching points, assisted by precise synchronized time.

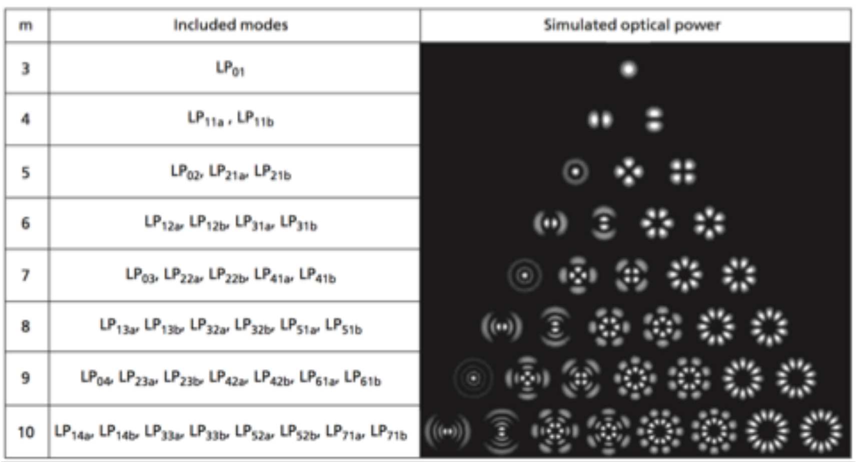
Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

28

Theoretical Analysis

Modal crosstalk, caused by random mode coupling, depending on fiber fabrication, imperfections, bending or twisting [7].



m	Included modes	Simulated optical power
3	LP_{01}	
4	LP_{11a}, LP_{11b}	
5	LP_{02}, LP_{21a}, LP_{21b}	
6	LP_{12a}, LP_{12b}, LP_{31a}, LP_{31b}	
7	LP_{03}, LP_{22a}, LP_{22b}, LP_{41a}, LP_{41b}	
8	LP_{13a}, LP_{13b}, LP_{32a}, LP_{32b}, LP_{51a}, LP_{51b}	
9	LP_{04}, LP_{23a}, LP_{23b}, LP_{42a}, LP_{42b}, LP_{61a}, LP_{61b}	
10	LP_{14a}, LP_{14b}, LP_{33a}, LP_{33b}, LP_{52a}, LP_{52b}, LP_{71a}, LP_{71b}	

Accumulated crosstalk [8,9]:

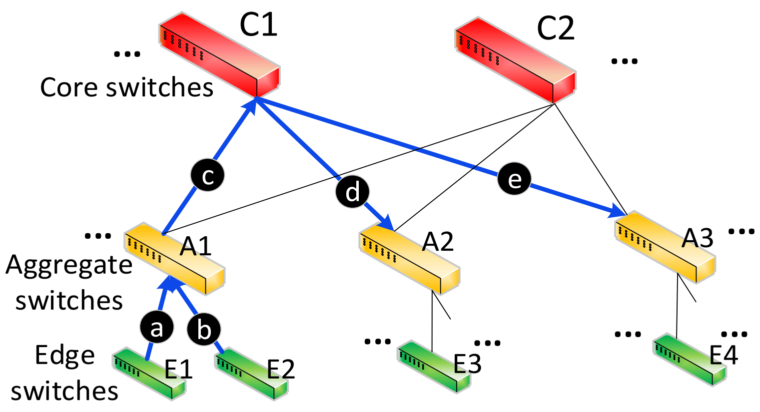
$$XT = \tanh hz$$

where h is the mode coupling parameter (m^{-1}) and z is the fiber length (m).

[7] K.-P. Ho, et al., “Linear propagation effects in mode-division multiplexing systems,” J. Lightw. Technol., 2014.
[8] P. J. Winzer, “Penalties from In-Band Crosstalk for Advanced Optical Modulation Formats,” ECOC 2011.
[9] F. M. Ferreira, et al., “Semi-Analytical Modelling of Linear Mode Coupling in Few-Mode Fibers,” J. Lightw. Technol., 2017.

29

Optimization Setup



Tab. 1, crosstalk (dB/100m) of selected modes.

XT	m_1	m_2	m_3	m_4
m_1	-	-26.0	-21.2	-43.0
m_2	-17.7	-	-15.8	-19.7
m_3	-19.5	-14.3	-	-15.6
m_4	-21.5	-16.7	-17.5	-

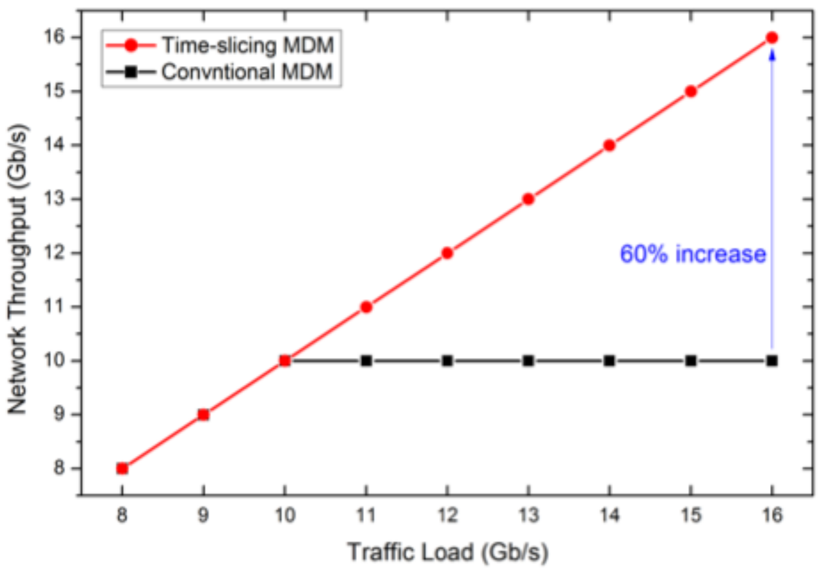
- IBM CPLEX solver
- Accumulated XT threshold:-13dB
- Fiber length: 100m
- Traffic: uniformly generated between edge switch pairs, 1Gb/s
- Modal channel: 10Gb/s
- OTSS frame: 20ms
- Min time slice: 5ms
- Only on 1550nm wavelength

Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

30

Optimization Results



- Time-Slicing MDM can achieve at least 60% higher throughput than conventional MDM scheme.
- This throughput increase come from flexible allocation of time slices.
- This increase will be more significant when adopting more wavelength and modes.

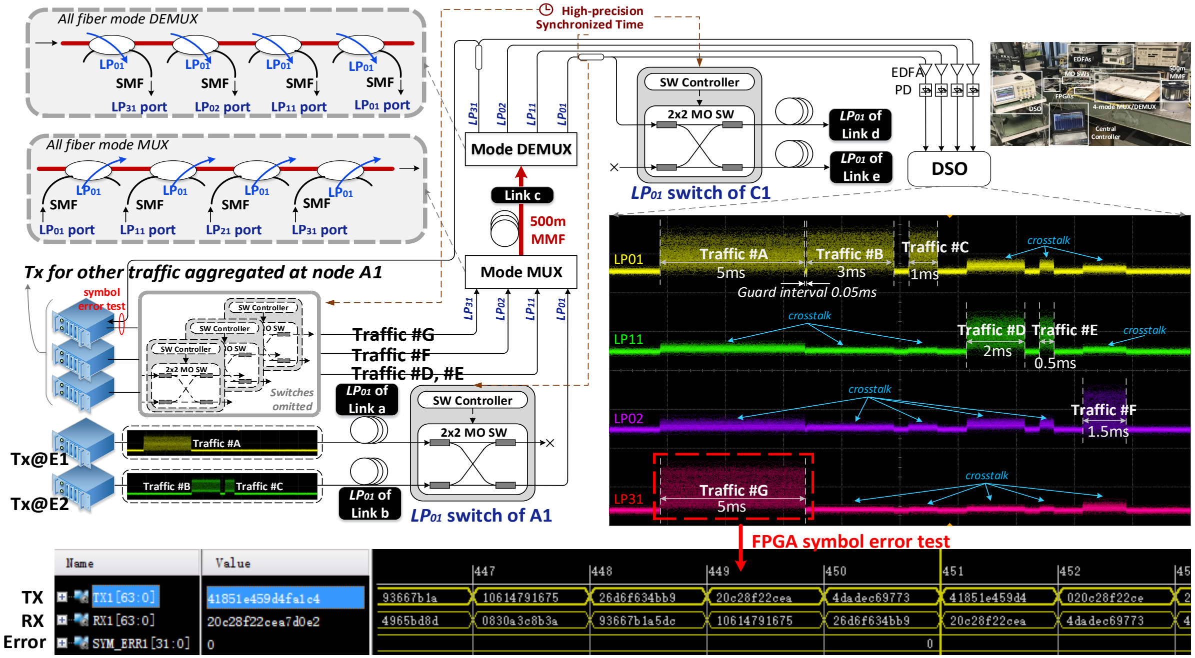
Zhizhen Zhong, Nan Hua, Yufang Yu, Zhongying Wu, Juhao Li, Haozhe Yan, Shangyuan Li, Ruijie Luo, Jialong Li, Yanhe Li, and Xiaoping Zheng, "Throughput Scaling for MMF-Enabled Optical Datacenter Networks Based on Time-Slicing-Based Crosstalk Mitigation," to be presented, ***OFC***, Mar. 2018.

清華大學 Tsinghua University

UCDAVIS UNIVERSITY OF CALIFORNIA

31

Experimental Demonstration



- 4-mode MDM transmission systems of OM3 MMF with core/cladding diameter of 50/125µm.
- Xilinx VC709 FPGA generate PRBS codes at 10Gb/s at 1550nm wavelength.
- OTSS frame length 20ms, min tine slice 500µs, guard interval 50µs.
- Magneto-electronic (MO) 2 × 2 optical switch.

Zhizhen Zhong, Nan Hua, Yufang Yu, Zhongying Wu, Juhao Li, Haozhe Yan, Shangyuan Li, Ruijie Luo, Jialong Li, Yanhe Li, and Xiaoping Zheng, "Throughput Scaling for MMF-Enabled Optical Datacenter Networks Based on Time-Slicing-Based Crosstalk Mitigation," to be presented, ***OFC***, Mar. 2018.

Overview of my recent publications on OTSS

- **Zhizhen Zhong**, Nan Hua, Zhu Liu, Wenjing Li, Yanhe Li, and Xiaoping Zheng, "Evolving Optical Networks for Latency-Sensitive Smart-Grid Communications via Optical Time Slice Switching (OTSS) Technologies," ***IEEE CLEO-PR / OECC / PGC***, *Aug. 2017*. **(IEEE Photonics Society 1st place Poster Award)**
- **Zhizhen Zhong**, Nan Hua, Yufang Yu, Zhongying Wu, Juhao Li, Haozhe Yan, Shangyuan Li, Ruijie Luo, Jialong Li, Yanhe Li, and Xiaoping Zheng, "Throughput Scaling for MMF-Enabled Optical Datacenter Networks Based on Time-Slicing-Based Crosstalk Mitigation," to be presented, ***OFC***, Mar. 2018.
- Nan Hua, **Zhizhen Zhong**, Xiaoping Zheng, "Enabling Low Latency at Large-Scale DataCenter and High-Performance Computing Interconnect Networks Using Fine-GrainedAll-Optical Switching Technology," ***ONDM***, May 2017.
- Yufang Yu, Nan Hua, **Zhizhen Zhong**, Jialong Li, Ruijie Luo, Zelin Zheng, and Xiaoping Zheng, "Fast-Reconfigurable Optical Interconnect Architecture Based on Time-Synchronized Node Coordination for High Performance Computing," ***ACP***, Nov. 2017.

33

Outline

- The need for a transparent fine-grained optical network
- Past solutions transparent fine-grained optical switching
- Our plan: Optical Time Slice Switching (OTSS)
- OTSS use cases
- **Conclusion and discussions**

Tsinghua University

UNIVERSITY OF CALIFORNIA

34

Conclusion

- Time-division-multiplexing (TDM) is the main way to a transparent fine-grained optical network.
- Time synchronization over network is the key to bufferless switching and networking.
- OTSS is proposed as transparent fine-grained bufferless networking paradigm for next-generation optical networks.
- New problems arises as how to allocate resource for OTSS networks, and novel solutions have been proposed
- Several use cases where OTSS show significant advantages over conventional networks.

35

Thank you for attention!

Zhizhen Zhong

Tsinghua University & UC Davis

zhongzz14@mails.tsinghua.edu.cn , zzzhong@ucdavis.edu

02 Feb. 2018

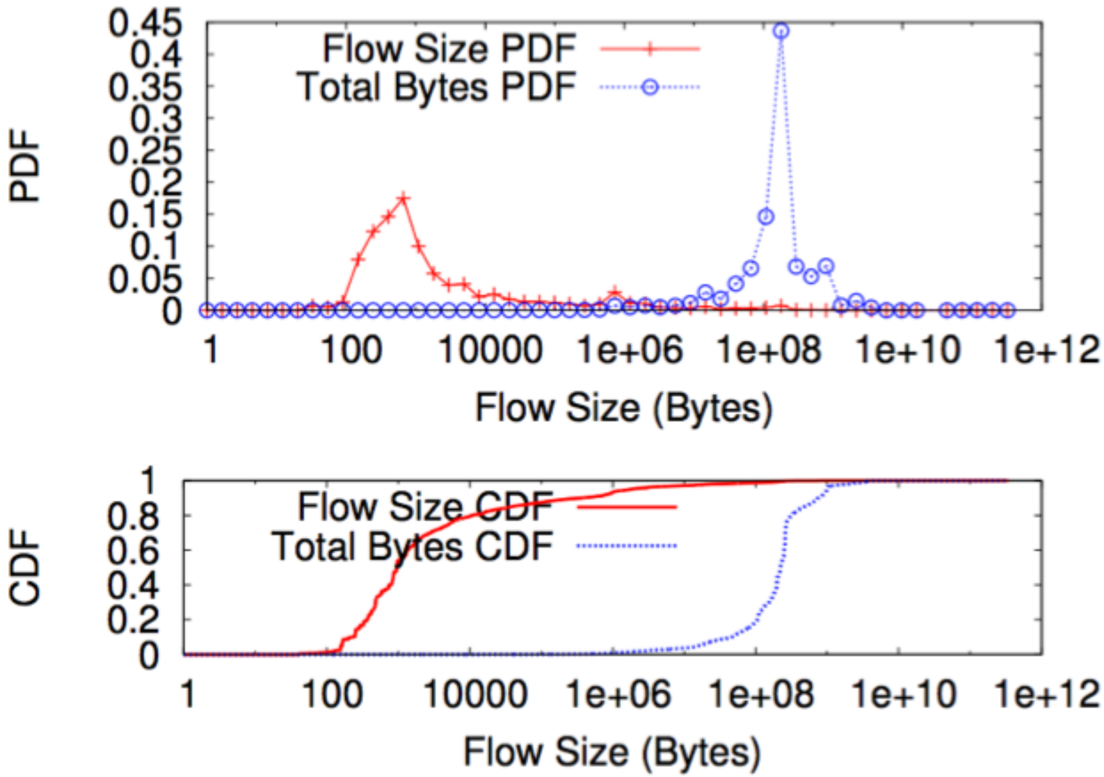
Networks Lab Group Meeting

清華大學 Tsinghua University

UC DAVIS UNIVERSITY OF CALIFORNIA

Appendix: DC traffic model

Microsoft Datacenter traffic*:



Mice flows are numerous: **99%** of flows are smaller than 100MB. However, more than 90% of bytes are in flows between 100MB and 1GB.

* A. Greenburg, "VL2: A Scalable and Flexible Data Center Network," ACM SIGCOMM 2009.