Consistent SDN Flow Migration aided by Optical Circuit Switching

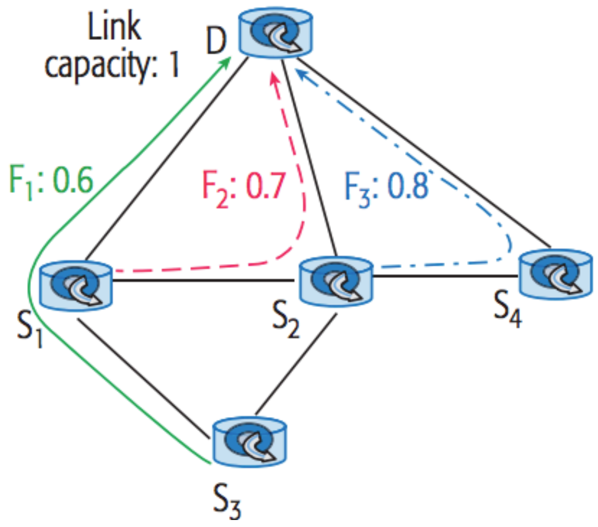
Rafael Lourenço

December 2nd, 2016

UCDAVIS
UNIVERSITY OF CALIFORNIA

What is Flow Migration?

Installation/update of new rules on [multiple] [asynchronous] SDN switches to realize a new network policy [replacing an older one]



Initial

[1] On the resource trade-off of flow update in software-defined networks
Yujie Liu, Yong Li, Yue Wang, Ying Zhang, Jian Yuan (Tsinghua University and Ericsson Research)
IEEE Communications Magazine 2016

Reasons for Flow Migration in Data Center SDNs

- VM Migration – flows previously being routed to physical host machine A need to be now routed to new physical host machine B [2]
- Load Balancer reconfiguration [2]
- Element firmware upgrade – switch must be drained, updated, and reinserted in the network [2]
- Element failure repair – even with low level protection, it might result in an undesirable unbalanced use of network resources [2]
- Element onboarding – new flows originate/end in new element [2]
- Change in security policy – e.g., reroute flows through firewall [3]

[2] "zUpdate: Updating Data Center Networks with Zero Loss"
Hongqiang Harry Liu, Xin Wu, Ming Zhang, Lihua Yuan, Roger Wattenhofer, David A. Maltz (Yale, Duke, and Microsoft) SIGCOMM 2013
[3] "Survey of Consistent Network Updates."
Klaus-Tycho Foerster, Stefan Schmid, and Stefano Vissicchio. (ETH, Aalborg Univ, Univ of Louvain) *arXiv preprint* 2016.

Flow Migration Desirable Traits

- *Fast* – updates occur frequently, commonly involving transition of a large number of flows, thus must be fast to quickly adapt to new situations and reduce possibility of congestion during migration [1]
- *Low overhead* – utilization of resources cannot exceed resource limitations during migration (link and switch table capacity) [1]
- *Consistent* – at a minimum, no packet should undergo a mixture of old and new switching policies [1]

[1] On the resource trade-off of flow update in software-defined networks
Yujie Liu, Yong Li, Yue Wang, Ying Zhang, Jian Yuan (Tsinghua University and Ericsson Research)
IEEE Communications Magazine 2016

Flow Migration Consistency Space

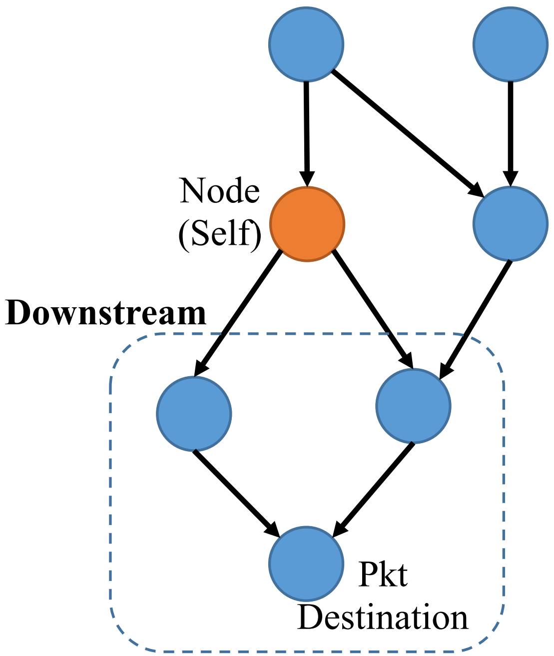
- **Eventual consistency** No consistency is provided during updates. If the new set of rules computed by the controller are consistent, the network will be eventually consistent
- **Blackhole freedom** No packet should be blackholed during updates. Blackholes occur if a packet arrives at a switch when there is no matching rule to handle it
- **Packet coherence** The set of rules seen by a packet should not be a mix of old and new rules
- **Congestion freedom** The amount of traffic arriving at a link should not exceed its capacity

[4] "Consistent Updates in Software Defined Networks: On Dependencies, Loop Freedom, and Blackholes." Klaus-Tycho Förster, Ratul Mahajan, and Roger Wattenhofer. (ETH Zurich) *Proc. 15th IFIP Networking* (2016).

Relationship between Types of Consistency

	None	Self	Downstream subset	Downstream all	Global
Eventual consistency	Always guaranteed				
Blackhole freedom	Impossible	Add before remove			
Loop freedom (Section V)	Impossible (Lemma 5)		Rule dep. forest		
Packet coherence	Impossible (Lemma 6)			Per-flow ver. numbers	Global ver. numbers
Congestion freedom	Impossible (Lemma 7)				Staged partial moves

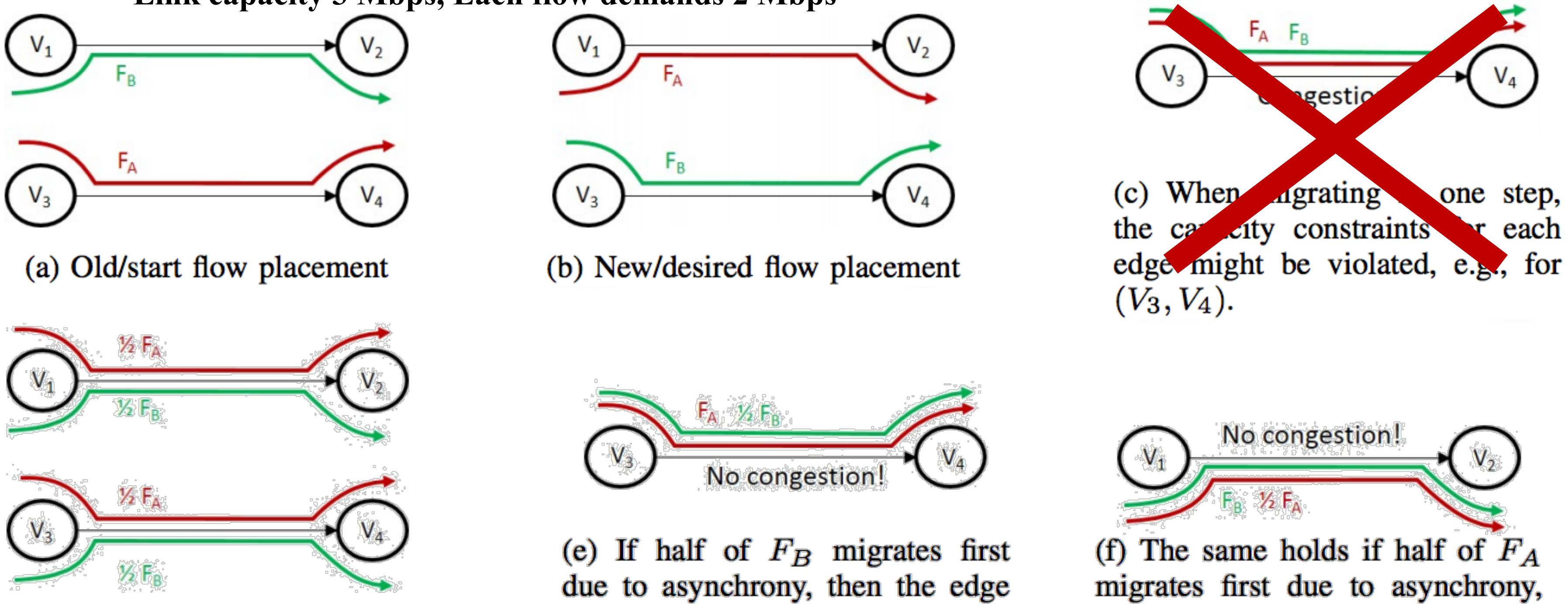
- Packet coherence ⇒ loop freedom
- Congestion freedom ⇒ loop freedom
- Packet coherence ⊥ congestion freedom
- Blackhole freedom ⊥ loop freedom



[4] "Consistent Updates in Software Defined Networks: On Dependencies, Loop Freedom, and Blackholes."
Klaus-Tycho Förster, Ratul Mahajan, and Roger Wattenhofer. (ETH Zurich)
Proc. 15th IFIP Networking (2016).

Congestion Free Migration

Link capacity 3 Mbps, Each flow demands 2 Mbps



(a) Old/start flow placement

(b) New/desired flow placement

(c) When migrating in one step, the capacity constraints for each edge might be violated, e.g., for (V_3, V_4).

(d) However, with this intermediate placement, one can migrate congestion free in two steps.

(e) If half of F_B migrates first due to asynchrony, then the edge capacity of 3 of (V_3, V_4) is not violated.

(f) The same holds if half of F_A migrates first due to asynchrony, the edge capacity of 3 is not violated.

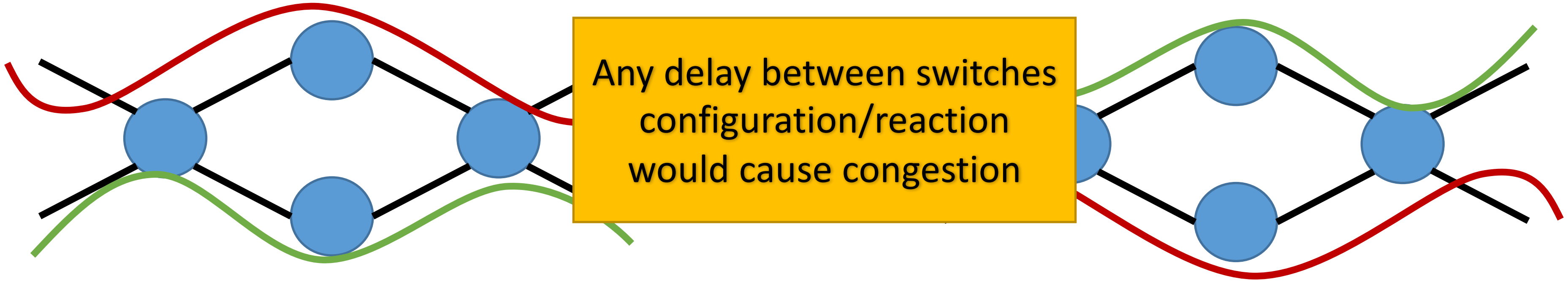
[5] " On Consistent Migration of Flows in SDNs"
Sebastian Brandt, Klaus-Tycho Foerster, Roger Wattenhofer. (ETH Zurich) *IEEE INFOCOM April 2016*

UCDAVIS

Unfeasible Congestion Free Migration

- In [5], it is proposed a poly-time algorithm to decide if there exists a congestion free migration from flow placement A to B (*NP-hard if all flows are integers or unsplittable*)
  - **One such case occurs** when trying to migrate between flow placements A and B, such that all source-destination pairs remain the same, there is no change in any flow demands (basically, merely re-routing flow paths in the network), and all links are completely full in both A and A'.

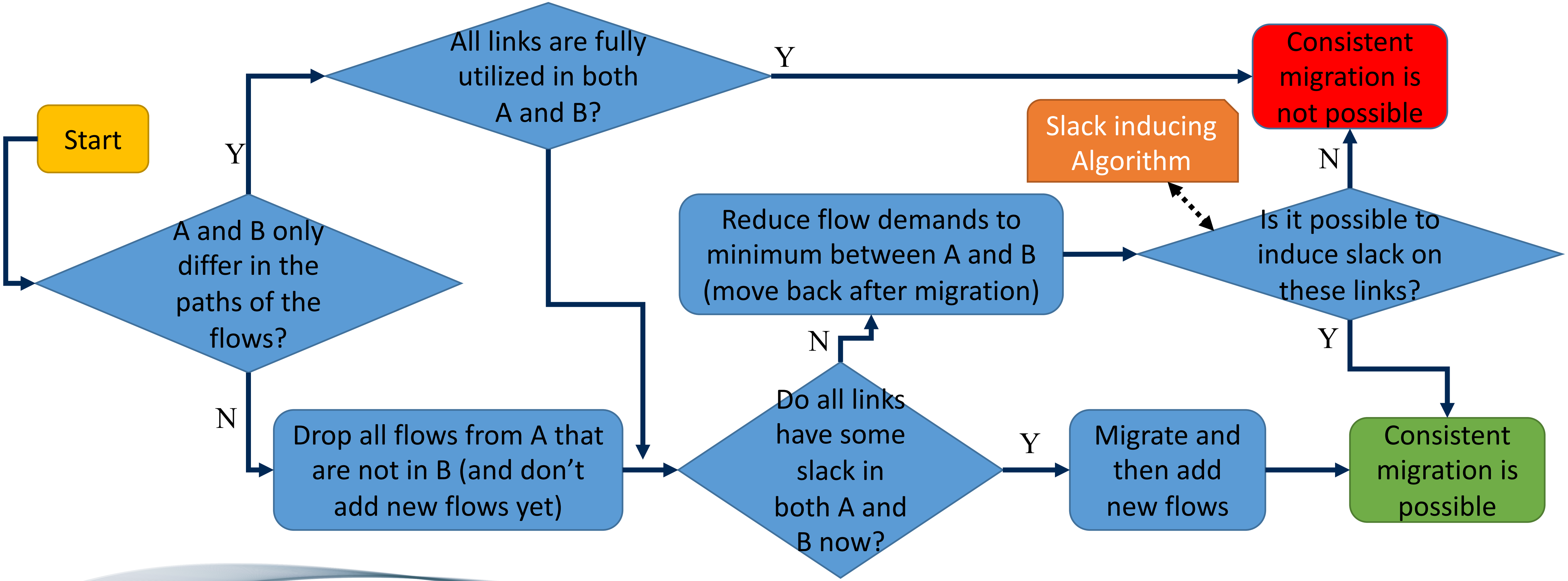
Link capacity 10 Gbps, Each flow demands 10 Gbps



Flow placement A **Flow placement B**

[5] " On Consistent Migration of Flows in SDNs"
Sebastian Brandt, Klaus-Tycho Foerster, Roger Wattenhofer. (ETH Zurich) *IEEE INFOCOM April 2016*

Congestion Free Migration Feasibility Verification



UCDAVIS

[5] " On Consistent Migration of Flows in SDNs"
Sebastian Brandt, Klaus-Tycho Foerster, Roger Wattenhofer. (ETH Zurich) *IEEE INFOCOM April 2016*

Using OCS to aid in Flows Migration

- Research on sub-microsecond switching times in OCS:
  - 50x50 MEMS OCS with switching time of 0.85us [8]
  - 62x62 MEMS OCS with switching time of 0.9us [9]
  - 2x2 CMOS Mach-Zehnder OCS with switching time of 4ns [10]
- Switching speed depends on OCS architecture aside from components (e.g., MEMS vs CMOS)
  - Certain architectures have varying switching times according to input/output selection

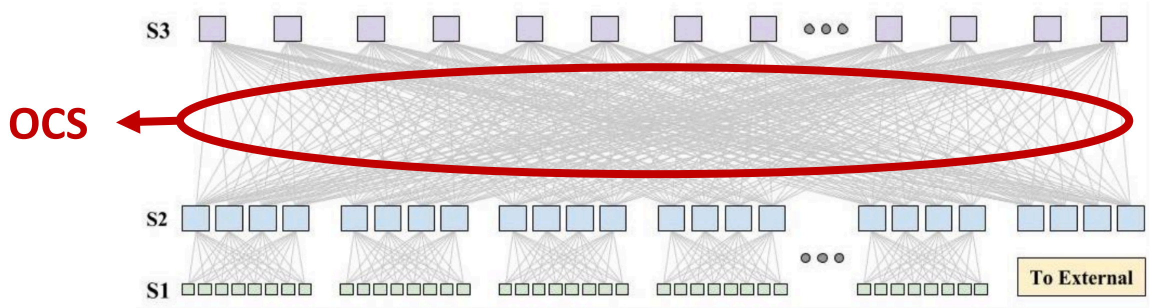
[8] "50x50 digital silicon photonic switches with MEMS-actuated adiabatic couplers." Seok, Tae Joon, et al. Optical Fiber Communication Conference. Optical Society of America, 2015. APA

[9] Seok, Tae Joon, et al. "64× 64 Low-loss and broadband digital silicon photonic MEMS switches." *Optical Communication (ECOC), 2015 European Conference on*. IEEE, 2015.

[10] "CMOS Photonic Nanosecond-Scale Switch Fabrics." Dupuis, Nicolas. *Optical Fiber Communication Conference*. Optical Society of America, 2016.

Using OCS to aid in Flow Migration

- Optical Circuit Switches are used for network management of the inter-pod (inter-block) topology [6, 7] in Data Center Networks

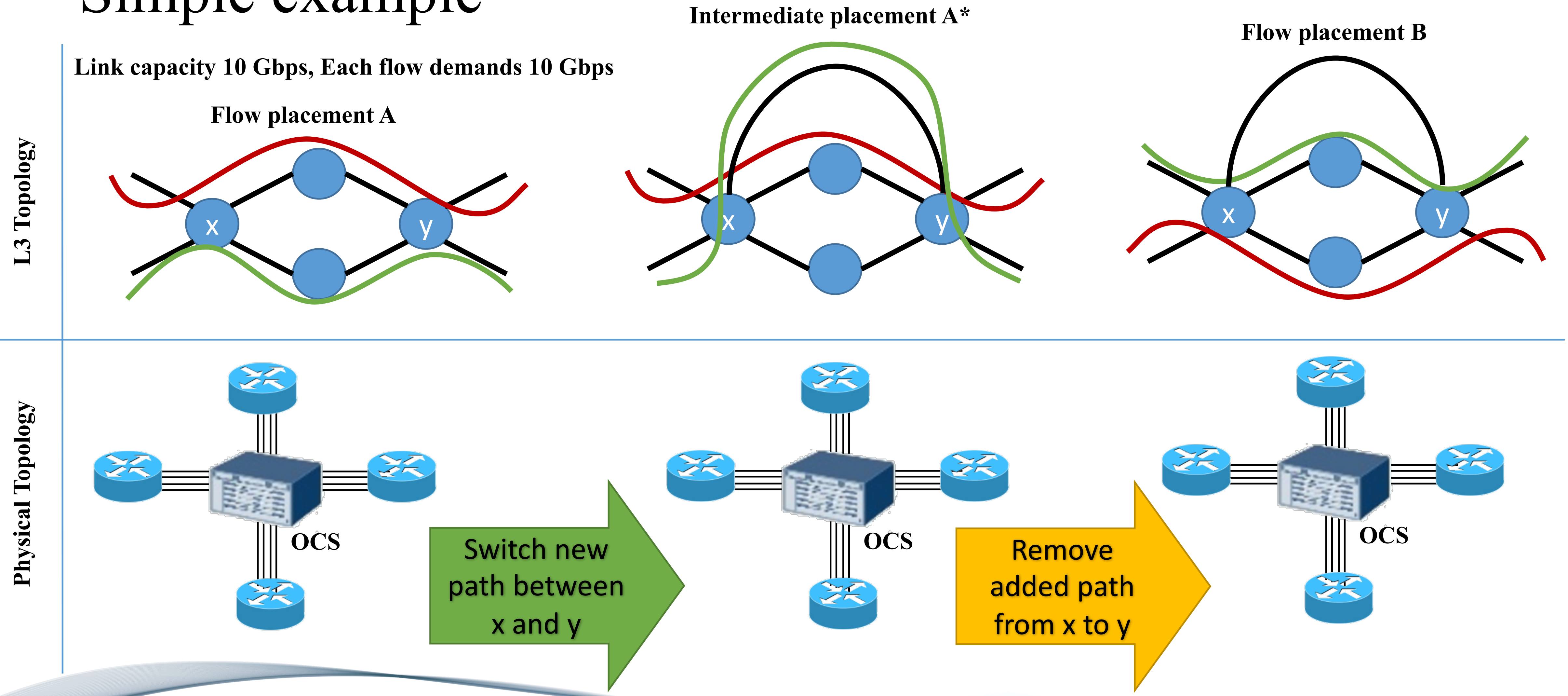


[6] ”*Integrating microsecond circuit switching into the data center*.” George Porter, Amin Vahdat, et al. Vol. 43. No. 4. ACM, 2013.
[7] "OSA: an optical switching architecture for data center networks with unprecedented flexibility." Chen, Kai, et al. *IEEE/ACM Transactions on Networking* 22.2 (2014): 498-511.

UCDAVIS

Simple example

Link capacity 10 Gbps, Each flow demands 10 Gbps



UCDAVIS

Problem Statement

Given:

- Physical topology T
- Current flow placement A
- Final desired flow placement B that cannot be consistently migrated to

Output:

- List of changes to be made to the physical topology such that they allow the final desired flow placement B to be reached from the current A, by going through an [set of] intermediate flow placement A* mapped on intermediate [set of] topology T*

Goal:

- Find the fastest possible migration from A to B, by minimizing the number of topological (OCS) modifications (and fastest OCS switching times for non-uniform OCS switch)

Constraints:

- Only some physical links can be OCS-re-configured given that no active flows may be harmed
- Only topological changes that might allow migration from A to an intermediate flow placement A*, and then from an intermediate A* to B are desirable

UCDAVIS

Next Steps

- Decide DC topology (fat tree, spineles, etc)
- Find useful example
- Decide how to approach problem

UCDAVIS