

Owan:
**Optimizing Bulk Transfers with
Software-Defined Optical WAN**

Xin Jin¹, **Yiran Li**², Da Wei², Siming Li³, Jie Gao³,
Lei Xu⁴, Guangzhi Li⁵, Wei Xu², Jennifer Rexford¹

¹ Princeton University, ² Tsinghua University,
³ Stony Brook University, ⁴ Sodero Networks, ⁵ AT&T Labs

The Demand of Bulk Transfers over WAN



More demanding

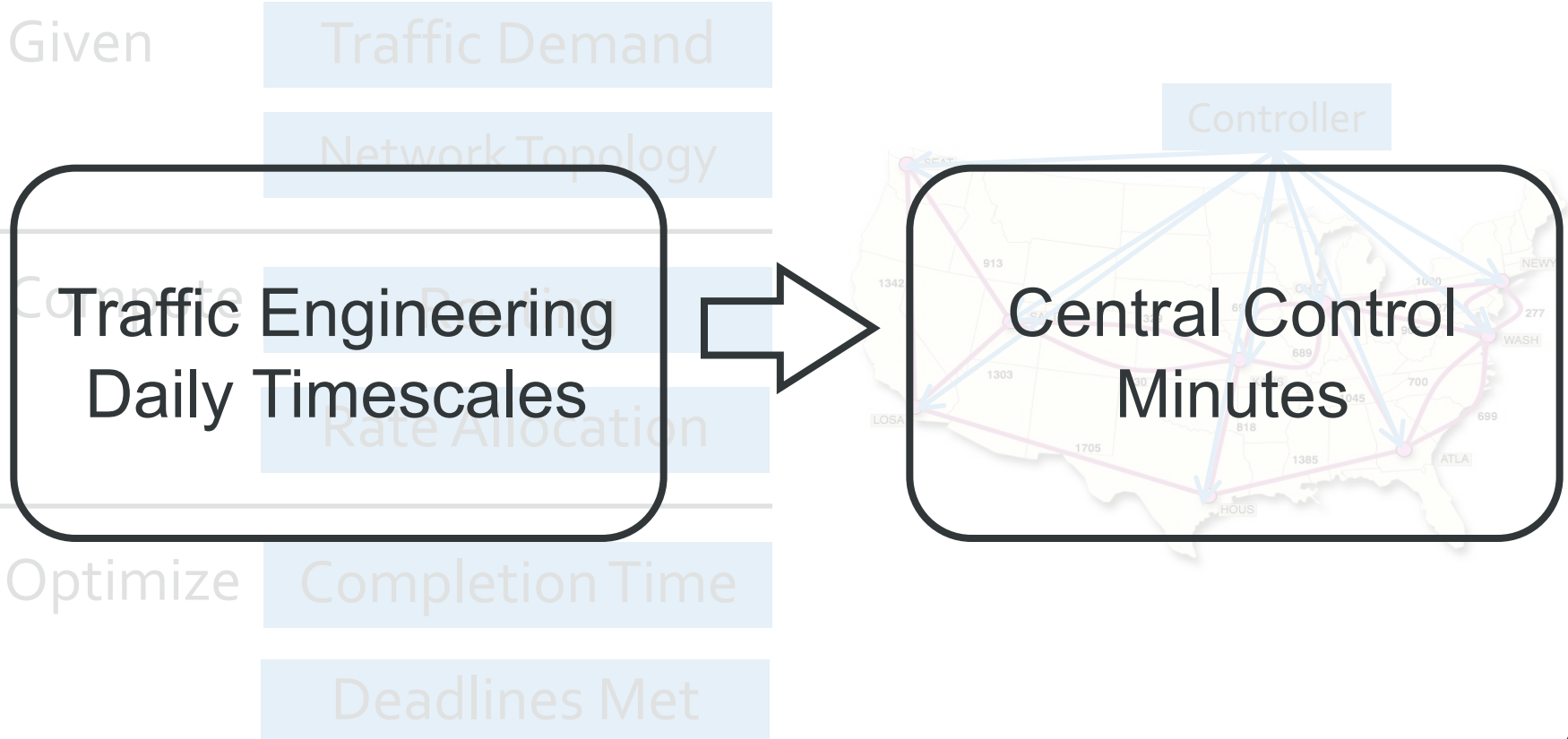
1. Transfer large size
2. Minimize completion time

More willing to

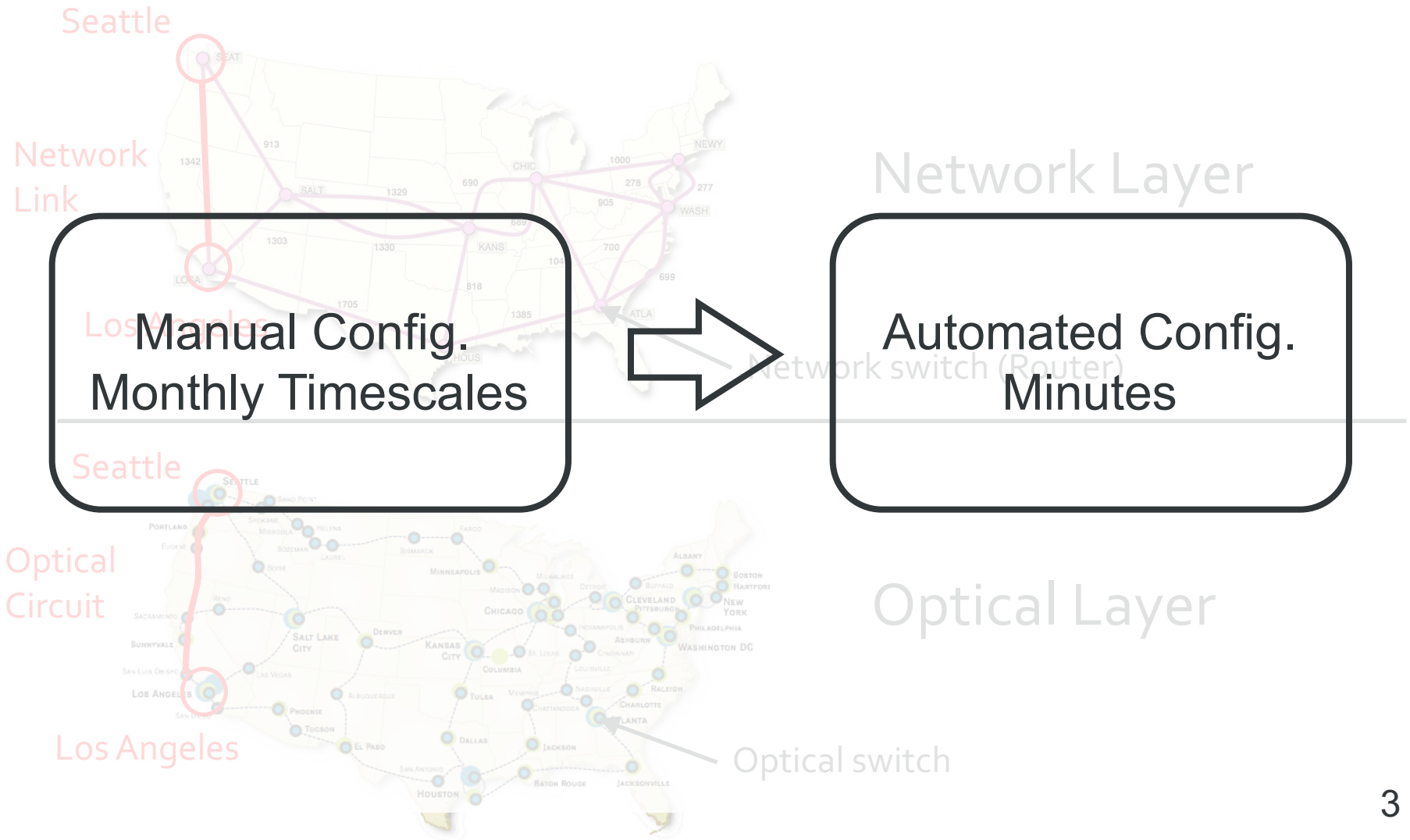
1. Provide demand information
2. Control its transfers

Software-Defined Networking (SDN) in WAN

Global traffic engineering with centralized control, e.g., Google B₄, Microsoft SWAN



Network Layer over Optical Layer

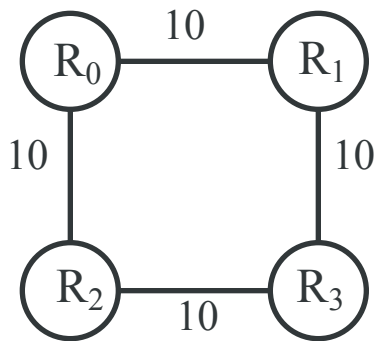


Technology Trends

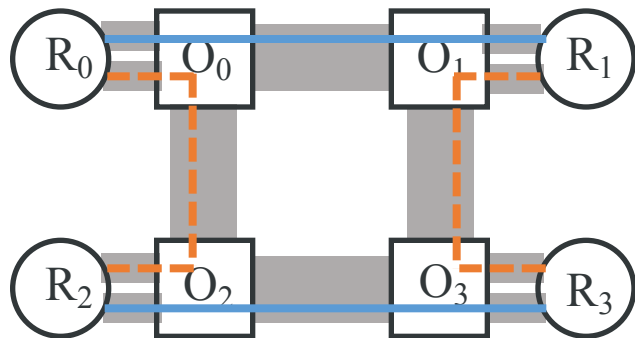
- Bulk-transfer applications with demand information
- Fast centralized control with SDN
- Fast reconfigurable optics

Reconfigure Optical Layer to Change Network-Layer Topology

Configuration A



Network Layer

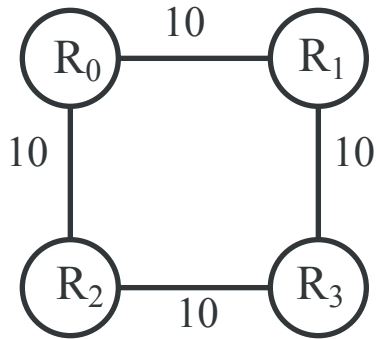


Optical Layer

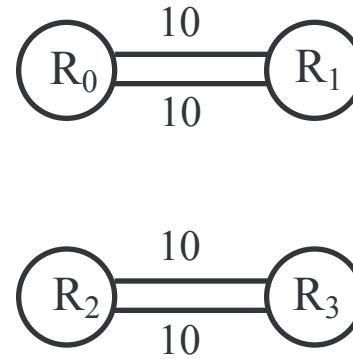
↑ Router ↑ Optical Switch

Reconfigure Optical Layer to Change Network-Layer Topology

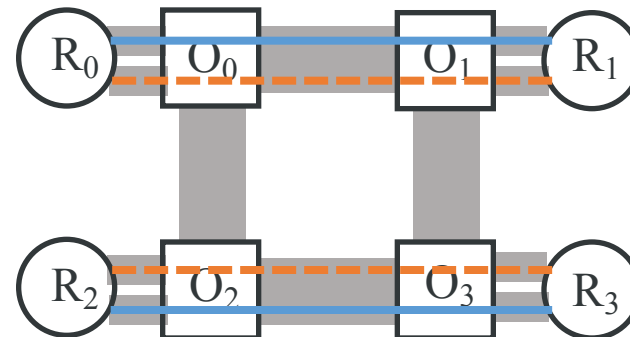
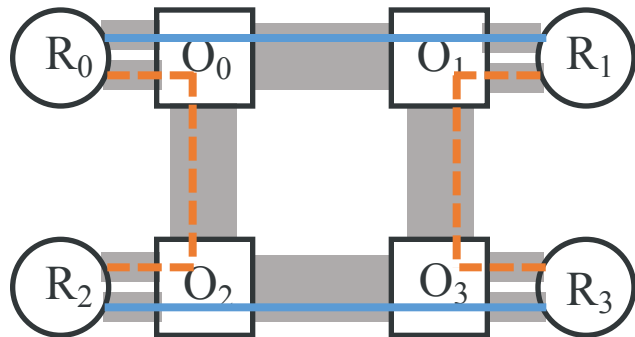
Configuration A



Configuration B



Network Layer

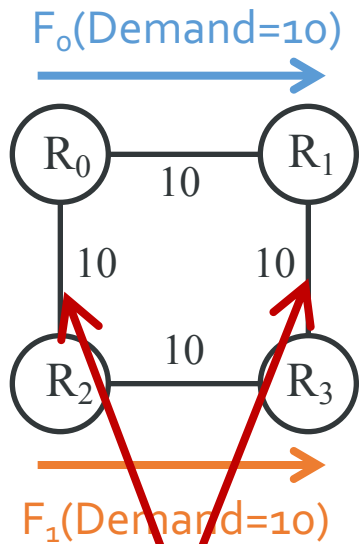


Optical Layer

↑ Router ↑ Optical Switch

Reduce Average Transfer Completion Time

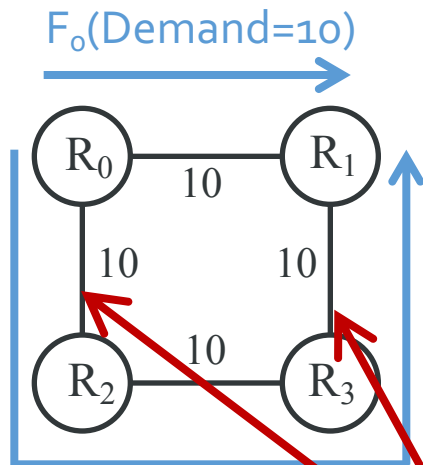
Routing



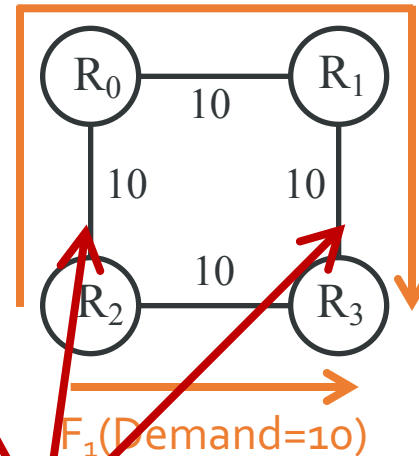
Unused Capacity

Routing + Rate allocation

Step 1

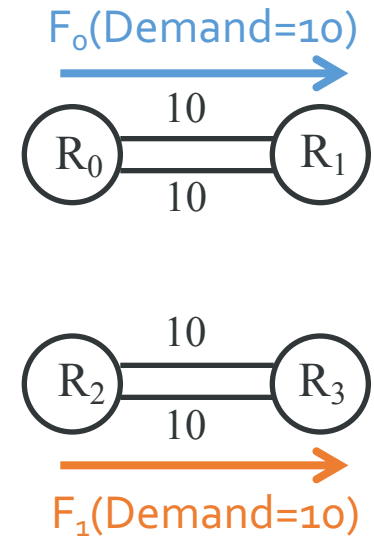


Step 2



Inefficiently Used Capacity

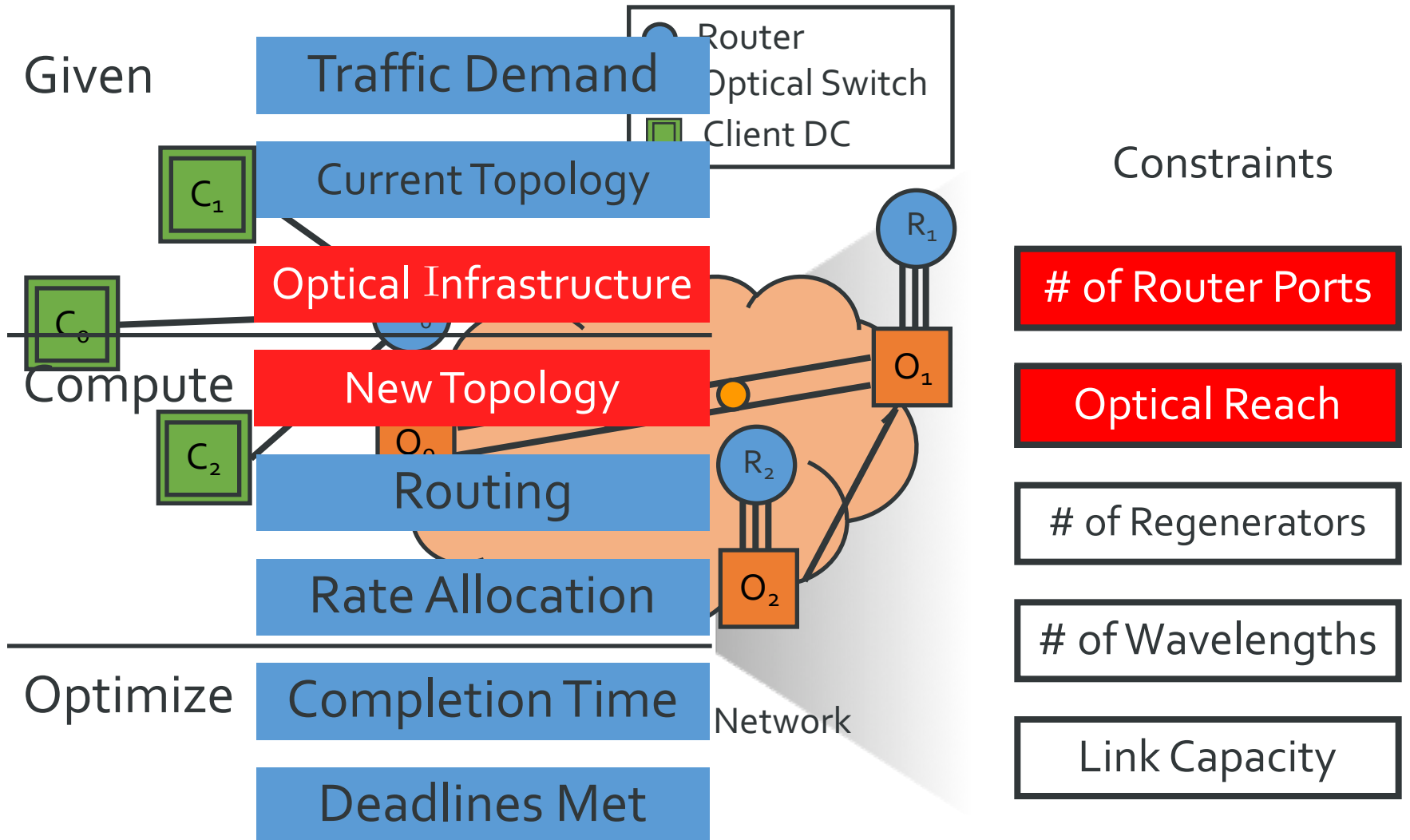
Routing + Rate + Topology



$$\text{Avg.} = \frac{0.5 + 0.5}{2} = 0.5$$

Joint Optimization and Challenges

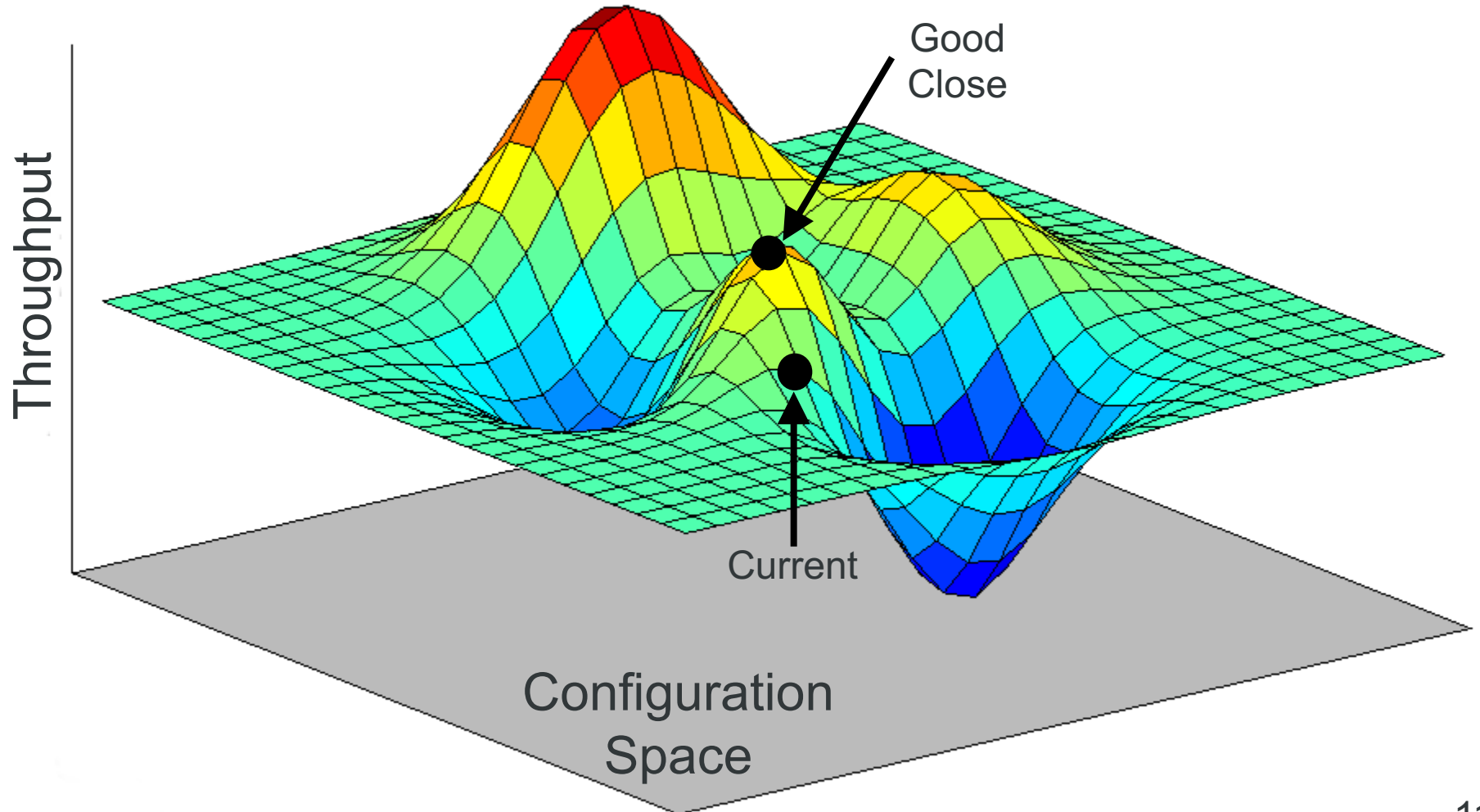
Joint Optimization



Challenges

- **Efficient joint optimization**
 - Routing
 - Rate allocation
 - Topology
- **Transition gracefully**
 - Minimize disruption during update

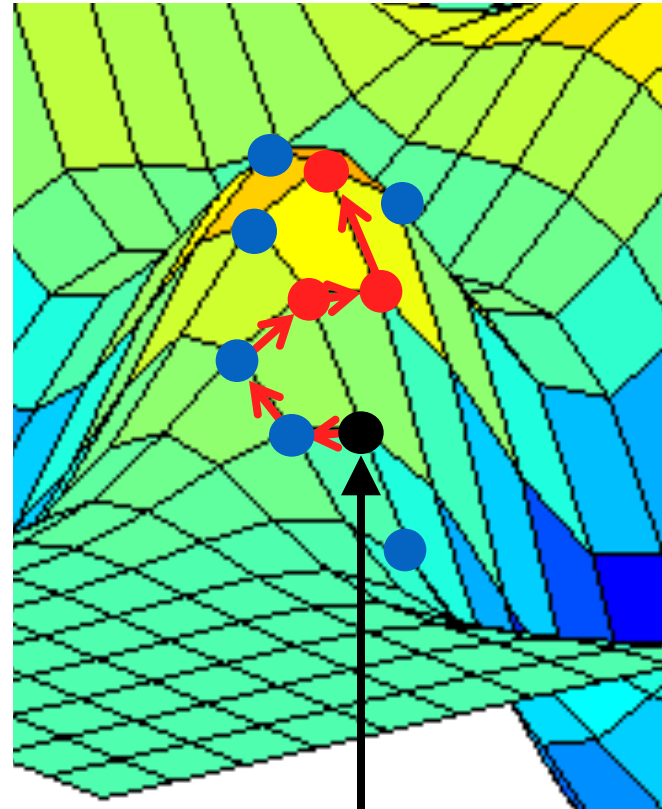
Finding Good Configuration with Small Change



Simulated Annealing Algorithm

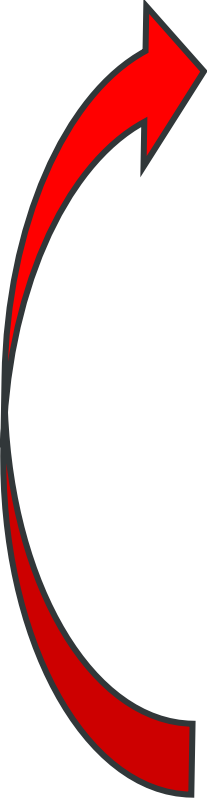
Choose
Random
Neighbor

Evaluate Neighbor



Current

Owan's Solution Overview



Choose
Random
Neighbor

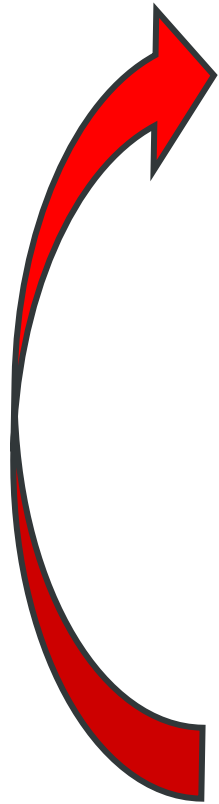
- Joint optimization efficiently
- Avoids disruption

Evaluate Neighbor

Consistent Update

Owan Algorithm

Random Neighbor Topology



Random
Neighbor Topo.

1. Make random local change
2. Select optical circuits

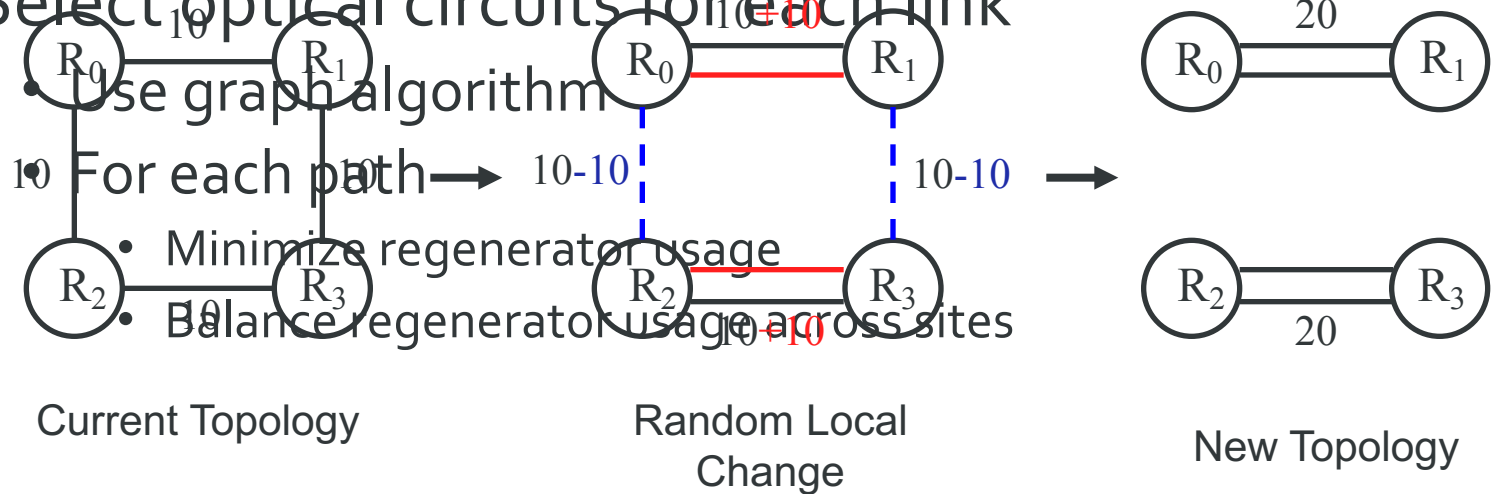
Optimize
Network Layer

Evaluate Neighbor

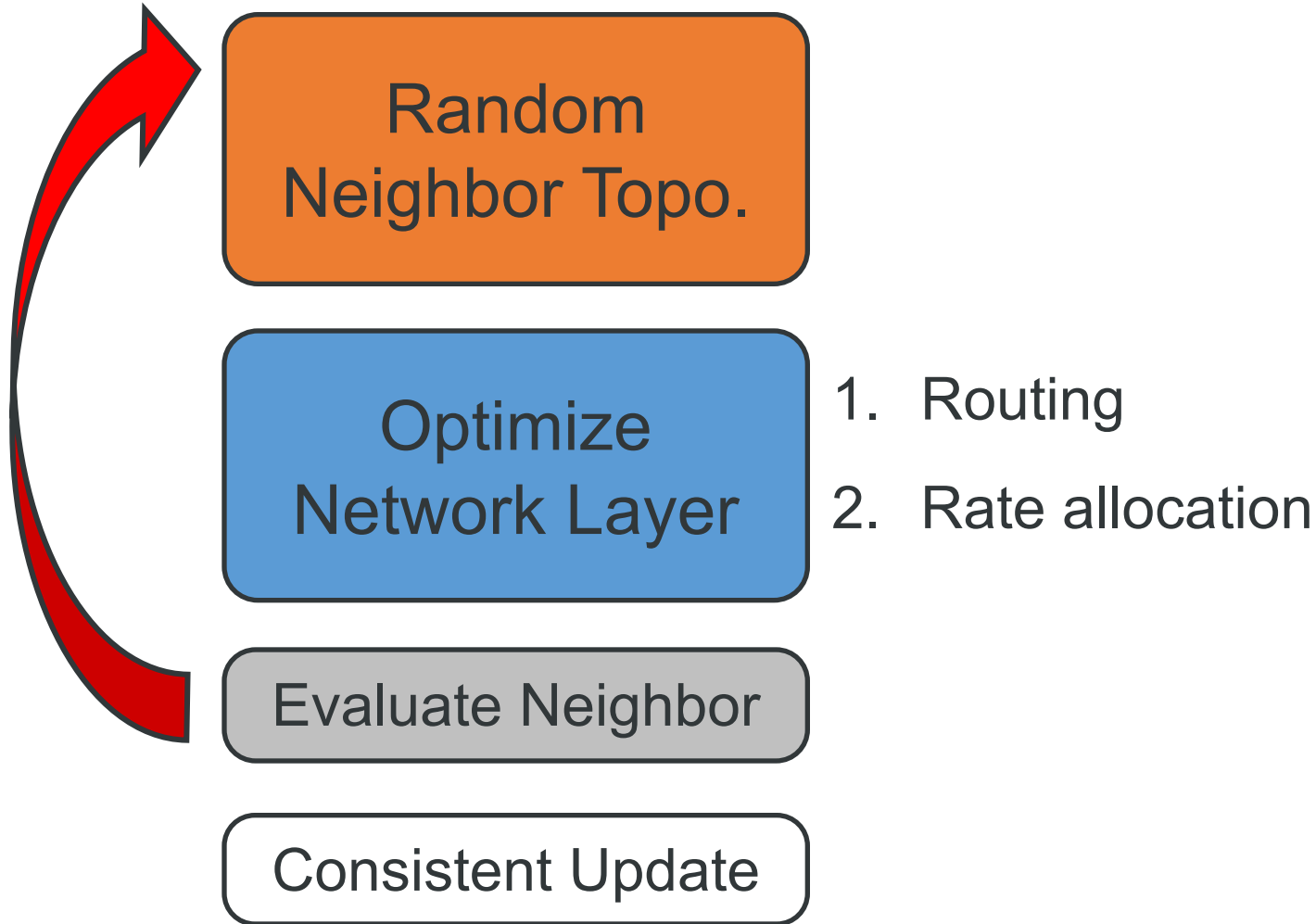
Consistent Update

Random Neighbor Topology

- Make random local change
 - Minimize changes to the network
 - Satisfy the port number constraints
- Select optical circuits for each link

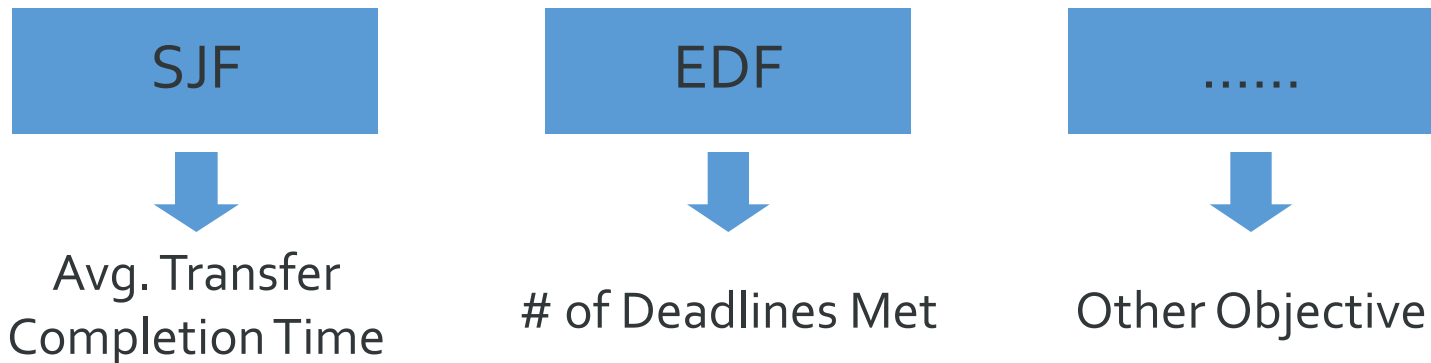


Optimize Network Layer



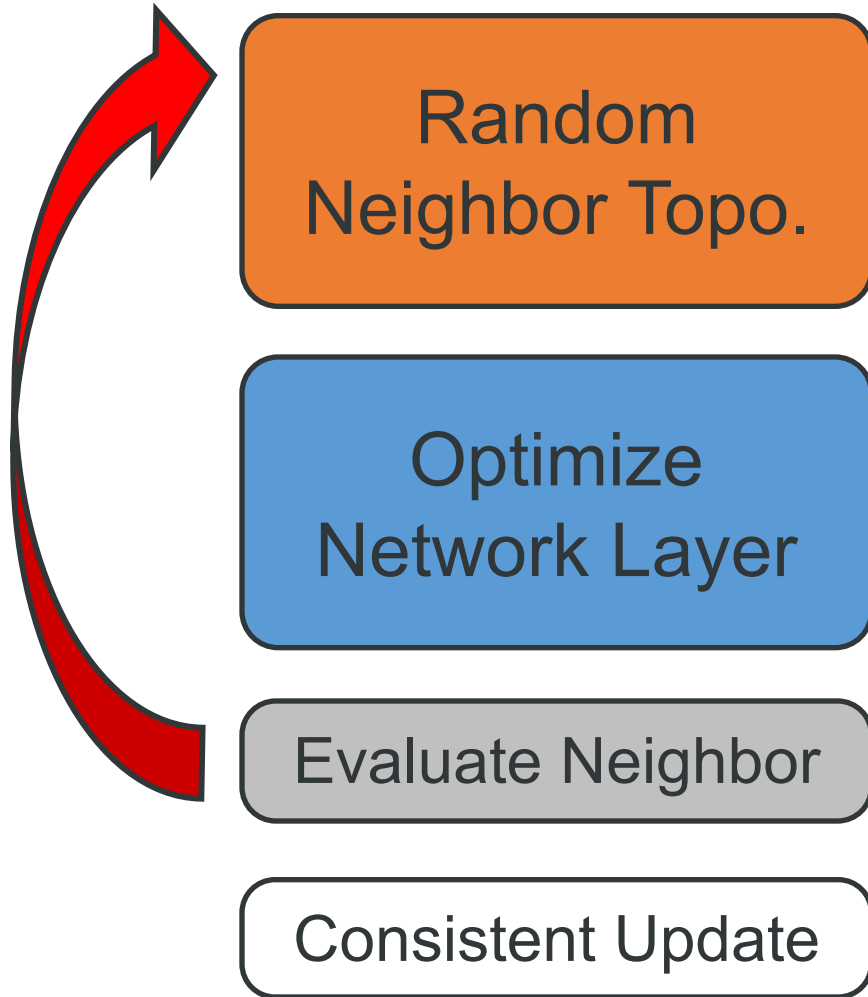
Schedule Transfers on the New Topology

- Order transfers with classic scheduling disciplines



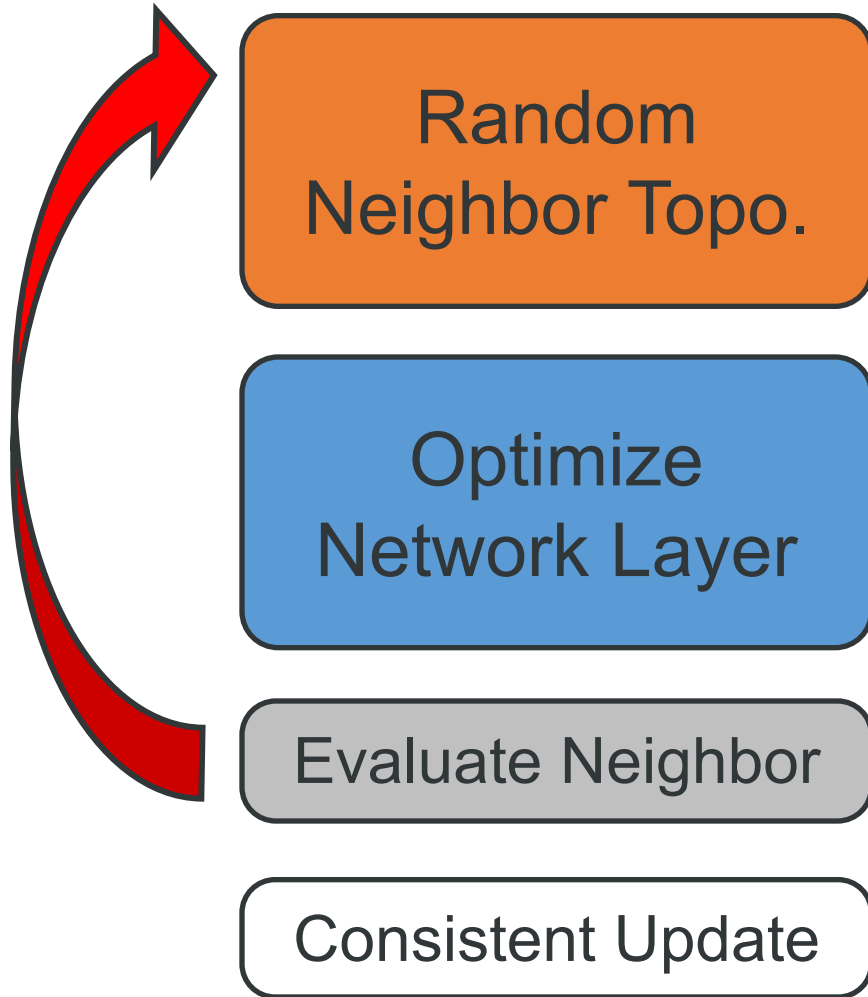
- Prioritize short paths in rate allocation

Evaluate Neighbor Topology



- Throughput: sum of rates

Consistent Update

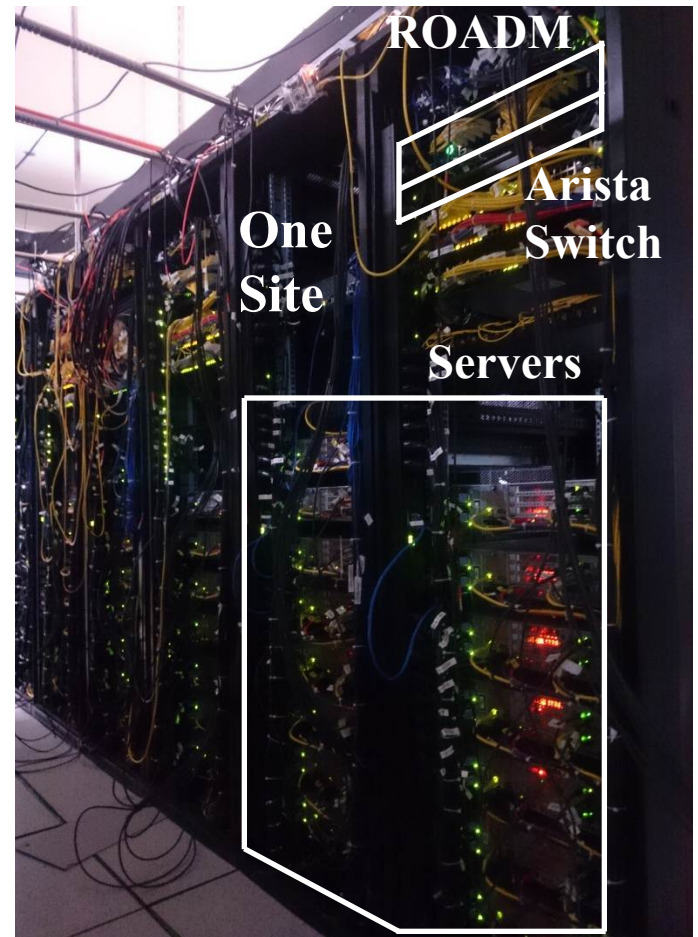


- Dependencies of operations

Implementation and Evaluation

Testbed Implementation

- 9 Sites
- Emulating Internet2 network
- 135 servers
 - Two 6-core Intel E5-2620v2
 - 10GE



Evaluation

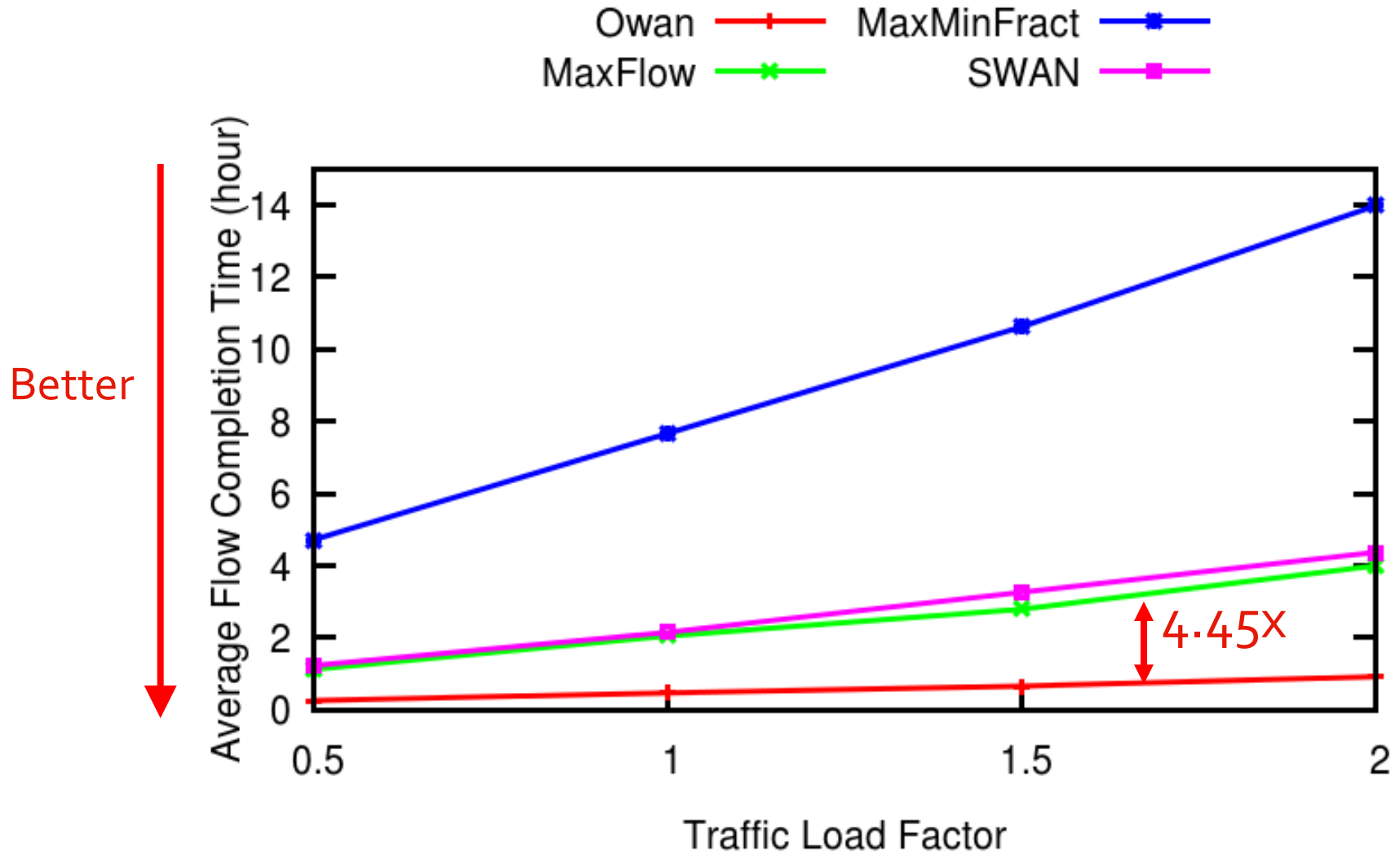
- Workload
 - Generate transfers for 2 hours
 - Draw transfer size from exponential distribution
 - Mean 500GB/5TB for testbed/simulation
- Evaluation
 - Testbed experiments, with 9 sites
 - Large-scale simulations, with about 40 sites
- Results
 - Average transfer completion time: 3.5-4.4x
 - Number of transfers that meet deadlines: 1.1-1.3x

Deadline-Unconstrained Traffic

- Performance metric
 - Transfer completion time
- Other approaches
 - MaxFlow
 - MaxMinFract
 - SWAN[1]

[1] Hong, Chi-Yao, et al., Achieving High Utilization with Software-Driven WAN, SIGCOMM 2013

Better Average Completion Time

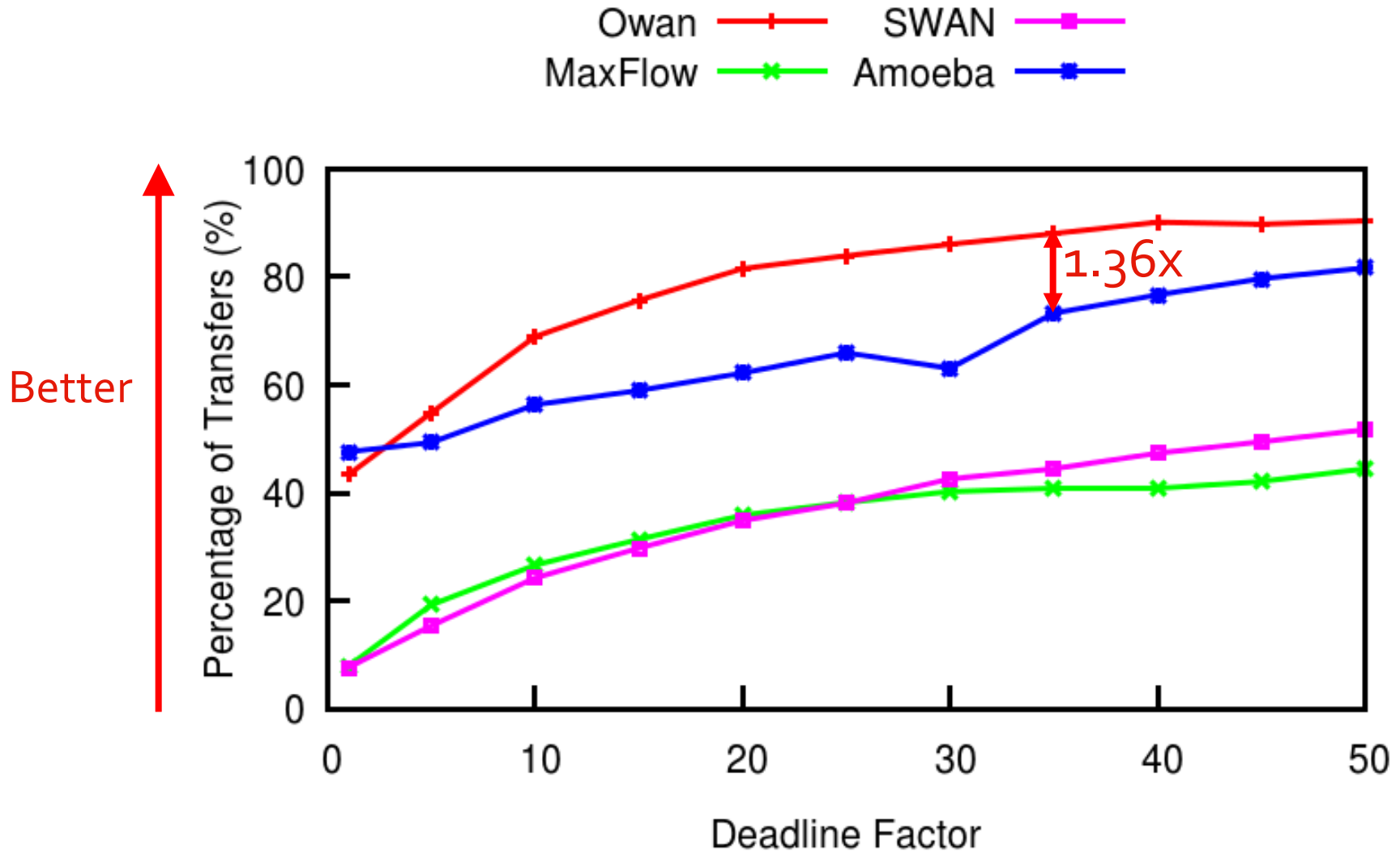


Deadline-Constrained Traffic

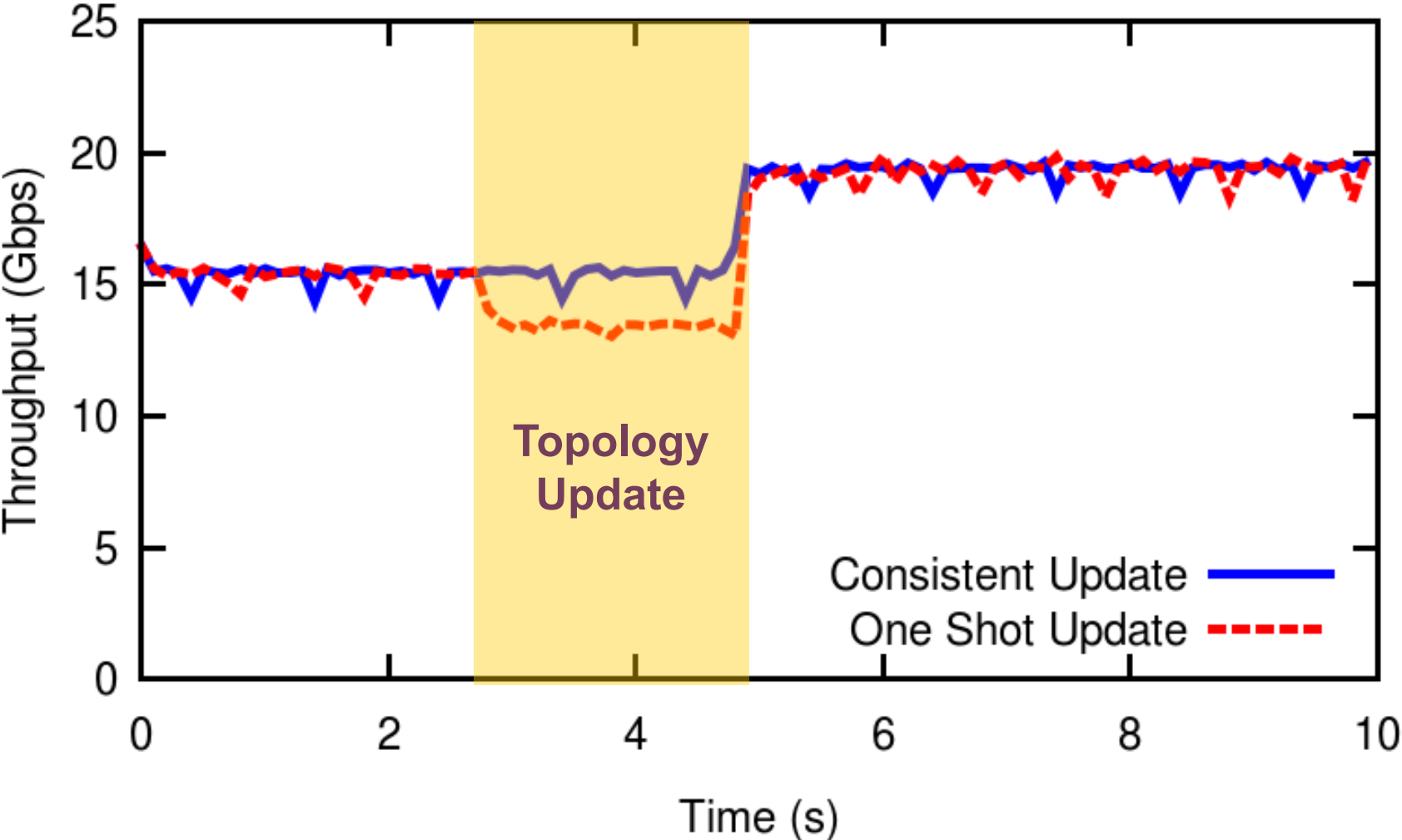
- Performance metric
 - Percentage of transfers that meet deadlines
 - Amount of bytes that finish before deadlines
- Other approaches
 - Deadline-unconstrained approaches
 - Amoeba[1]

[1] Zhang, Hong, et al., Guaranteeing deadlines for inter-datacenter transfers, EuroSys 2015

More Transfers Meet Deadlines



Consistent Update Avoids Disruptions



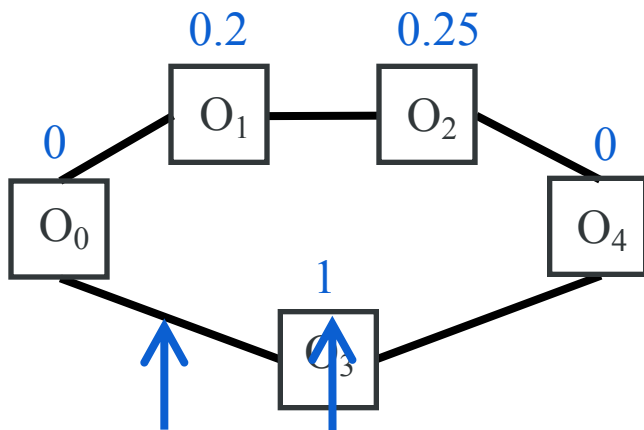
Conclusions

- Optical control improves WAN performance
- Efficient algorithms for joint optimization
- Transition gracefully

Thanks!
Q&A

Build Optical Circuits for Each Link

- Build regenerator graph
- Balance regenerator consumption



Goal: Find path with
min total node weight



Shortest path problem
on directed graph

Distance \leq Optical Reach
Inverse of # Regenerators

Cross-Layer Optimization at Each Time Slot

