UNIVERSITEIT GENT

FACULTY OF ENGINEERING AND ARCHITECTURE

Resilient backbone networks for multi-site data centers: Exploiting anycast (re)routing for multi-period traffic

Ting Wang[1], Brigitte Jaumard[1] and Chris Develder[2]

1: CSE, Concordia University, Montreal, Canada

2: INTEC – IBCN, Ghent University – iMinds, Ghent, Belgium

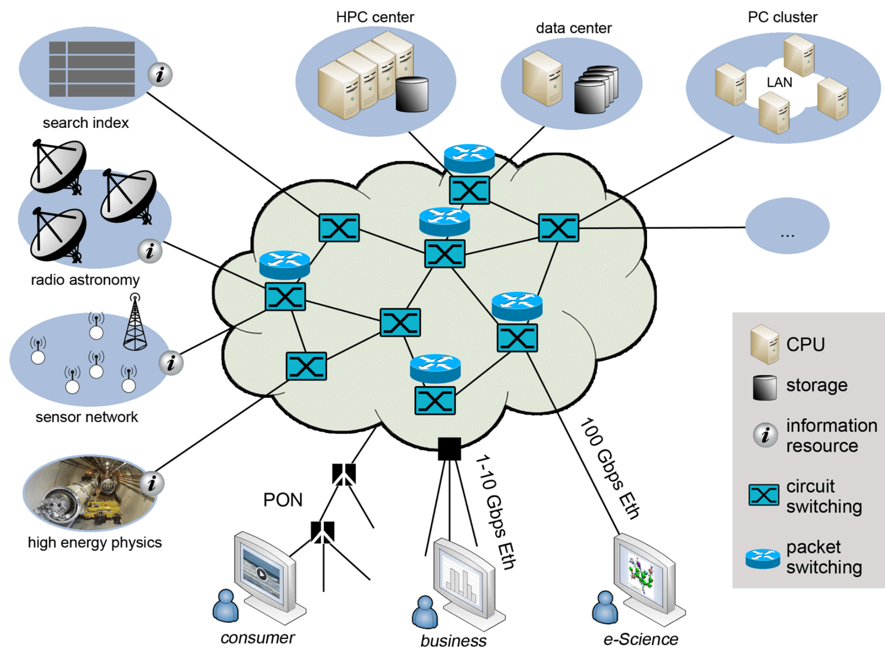
iMinds FUTURE INTERNET DEPT.

INTEC

Optical clouds

Optical networks crucial for increasingly demanding cloud services, e.g.,

- Computing:
  - High energy physics
  - Amazon EC2, Microsoft Azure
- Online storage:
  - Dropbox, Google Drive, etc.
- Collaboration tools:
  - MSOffice 365, Google Docs
- Video streaming:
  - Netflix, YouTube

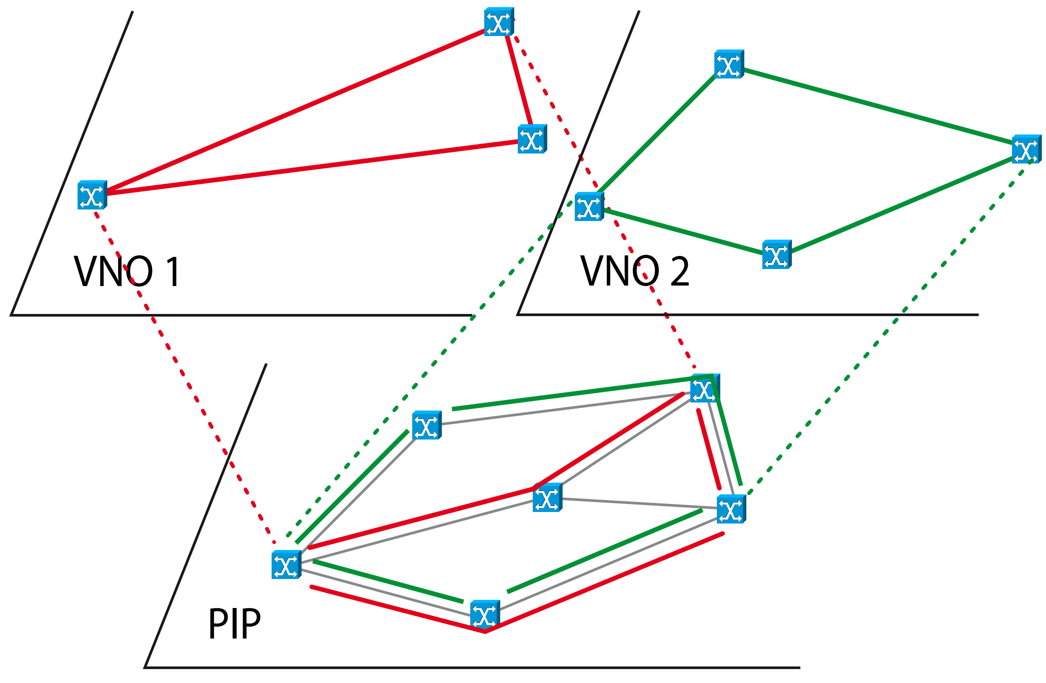


C. Develder, et al., "*Optical networks for grid and cloud computing applications*", Proc. IEEE, Vol. 100, No. 5, May 2012, pp. 1149-1167.

Network virtualization

Physical network is logically partitioned in isolated virtual networks

- **Virtual Network Operators (VNO)** operate logically separate networks
- **Physical Infrastructure Providers (PIP)** have full control over infrastructure (fibers, OXCs)



J.A. García-Espín, et al., *"Logical Infrastructure Composition Layer: the GEYSERS holistic approach for infrastructure virtualisation"*, in Proc. TERENA Networking Conference (TNC 2012), Reykjavík, Iceland, 21-24 May 2012.

Overview

1. Introduction
2. Problem statement
3. Model & solution approach
4. Case study
5. Conclusions

Overview

Resiliently provisioning virtual cloud networks

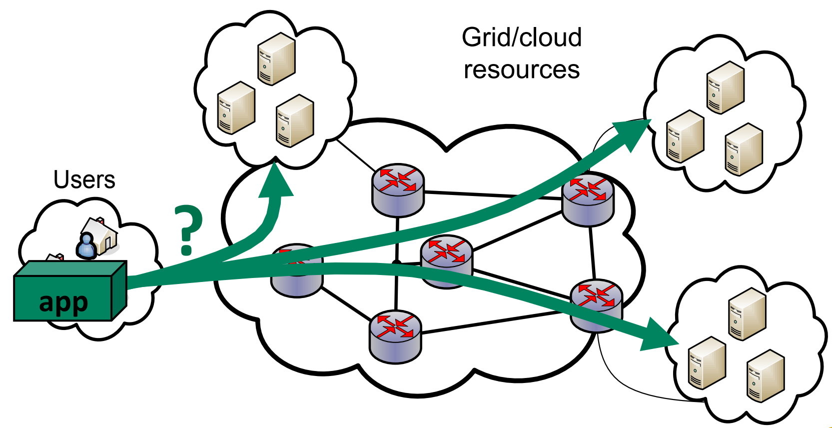
How to choose **the virtual to physical mapping**, such that

- Services remain available in case of network failures
- Bandwidth for providing services is minimal

?

Note:

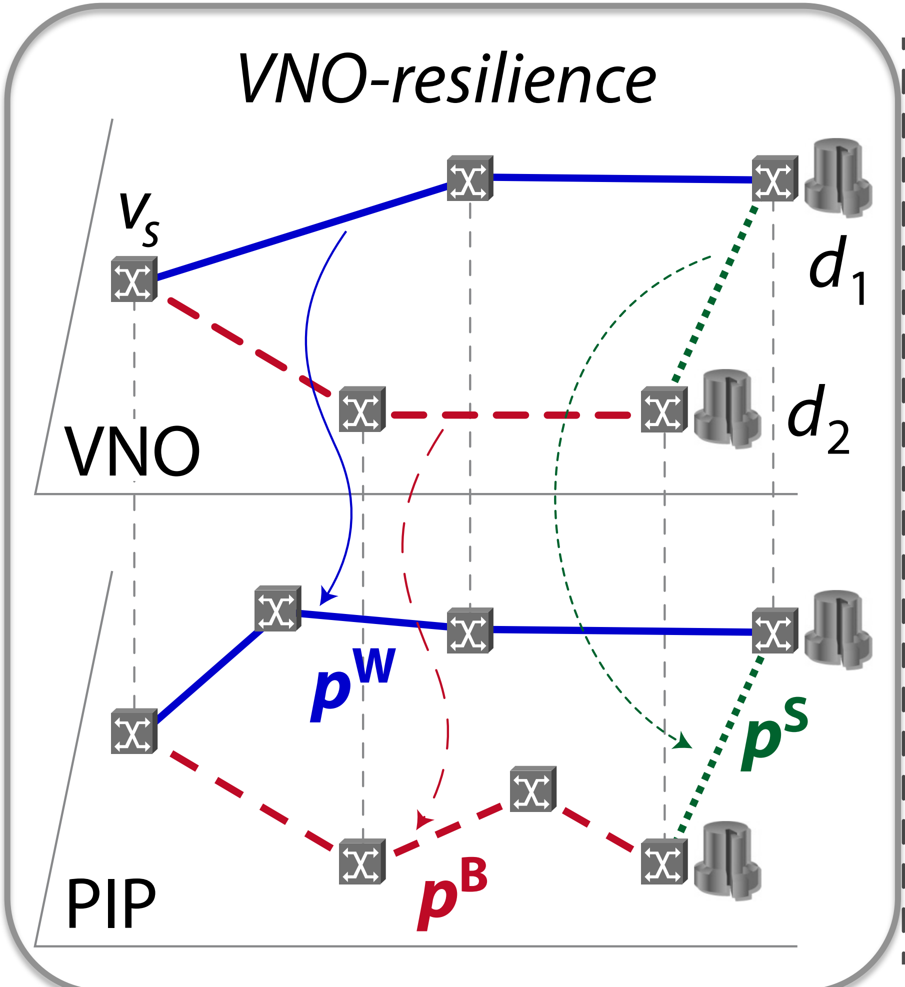
- **Anycast**: requests coming from users can be served by any server
- Cloud services offered by VNO
- Cloud services run on top of PIP



B. Jaumard, A. Shaikh and C. Develder, *"Selecting the best locations for data centers in resilient optical grid/cloud dimensioning (Invited Paper)"*, in Proc. 14th Int. Conf. Transparent Optical Netw. (ICTON 2012), Coventry, UK, 2-5 Jul. 2012.

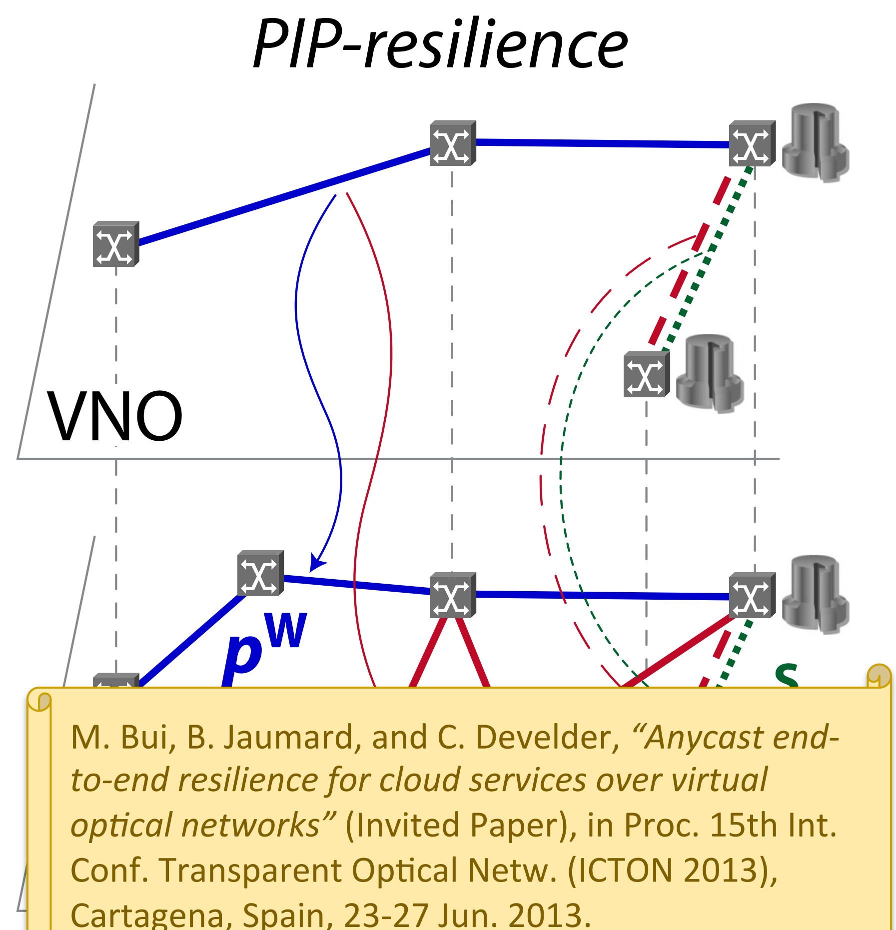
Two proposed protection schemes:

VNO-resilience



This paper

PIP-resilience



M. Bui, B. Jaumard, and C. Develder, *"Anycast end-to-end resilience for cloud services over virtual optical networks"* (Invited Paper), in Proc. 15th Int. Conf. Transparent Optical Netw. (ICTON 2013), Cartagena, Spain, 23-27 Jun. 2013.

Related work: Static traffic scenarios

- **Traditional dimensioning (no virtualization, no resilience)**:
  - Develder *et al*. 2009: Anycast, flexibility in choosing data center
- **Resilient** dimensioning problem:
  - Shaikh *et al.* 2011, Develder *et al.* 2013: scalable method, no synchronization between working and backup DCs
- Routing cloud requests and **mapping a VNet** to physical infrastructure separately:
  - Lee *et al.* 2009, Yu *et al.* 2010: Survivable VNet embedding, but *assume VNet is given*
  - Jiang *et al.* 2012, Alicherry *et al.* 2012: Optimal server selection and routing of anycast services in the physical layer for intra- and inter-DC networks but *no resilient network design in the virtual layer*
- **VNet planning** problem:
  - Barla *et al.* 2012, Barla *et al.* 2013: using mixed integer linear programming, but *no synchronization between working and backup DCs*
  - Bui *et al.* 2013 (ICTON): first model that incorporates **synchronisation path**, but *still static traffic!*
  - Bui *et al.* 2014 (ICTON): first model for **multi-period** scenario, but just considers 1 transition from a period T to T+1
  - Develder *et al.* 2015 (ICTON): first true **multi-period** model, cyclic **sequence of periods**

Problem statement

- Study **time-varying traffic**:
  - Traffic pattern changes from one period (*t*) to the next (*t*+1)
  - Optimize routes jointly for a sequence of periods

- Key research question:
  Benefit (in network resource usage) of **changing routes** for multi-period traffic, i.e., that continues from *t* to *t*+1?
  - Does it help to only change backup paths?
  - … or do we need to change working as well?

- Further analysis:
  - Impact of traffic: (i) varying fraction of traffic spanning multiple periods, (ii) varying number of regions with different traffic timing
  - Scalability: parallel solution scheme for column generation model

Problem statement

Given:
- Cloud network topology: $G = (V, L)$, with V = nodes, L = links
- Locations of the (candidate) data centers, $V_D \subseteq V$
- Topology nodes are partitioned in **time zones**
- Time is divided in **multiple periods** (time slots) w/ different traffic

Find:
- Choice of primary and backup **DC locations** for each service,
- Primary, backup and synchronization **paths** in period t+1,
- in **each of the time slots** it lasts

Such that: Total network **bandwidth cost is minimized**

Where routing is:

- *Scenario I:* unchanged,
- *Scenario II:* only changed for backup/synchronization paths
- *Scenario III:* freely changed (i.e., also allowed for working path)

Overview

1. Introduction
2. Problem statement
3. Model & solution approach
4. Case study
5. Conclusions

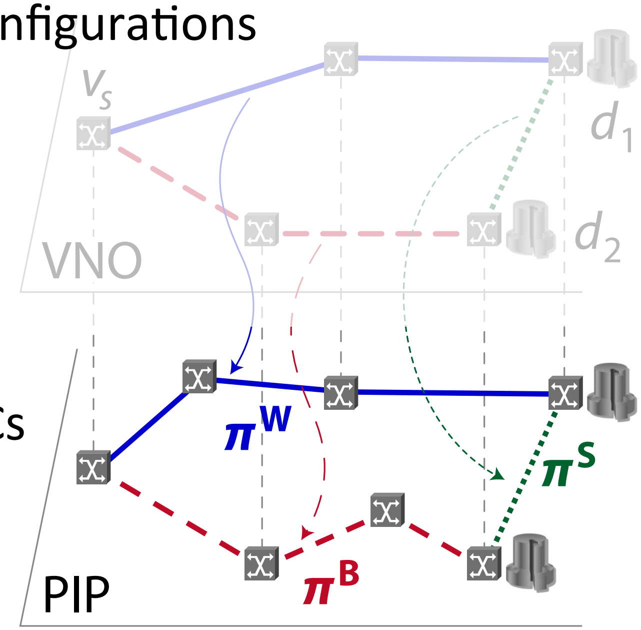
Overview

1. Introduction
2. Problem statement
3. Model & solution approach
4. Case study
5. Conclusions

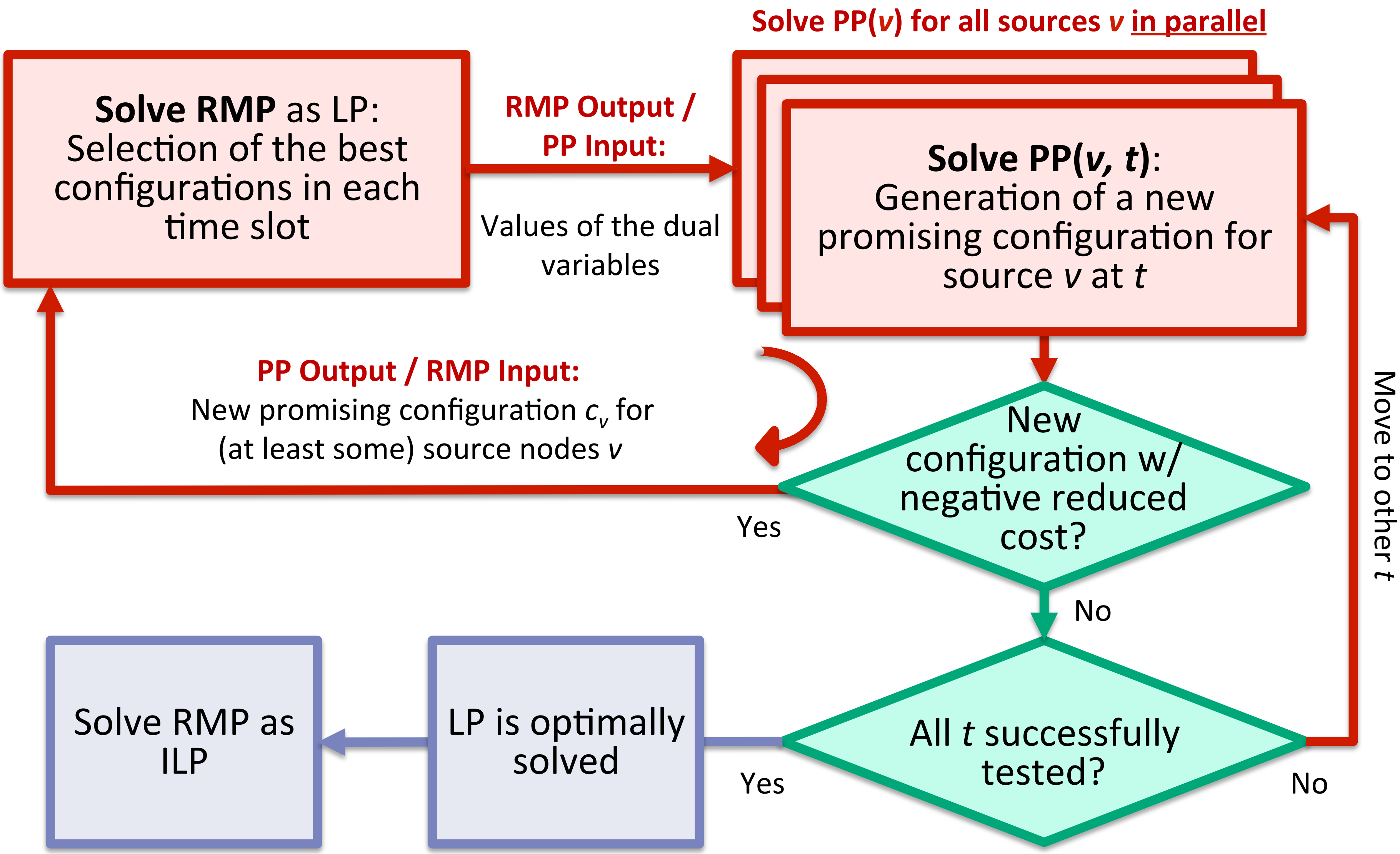
Solution: Column generation model

- Column generation idea:
  - Many different “configurations”
  - Start from a restricted subset of such “configurations”
  - Iteratively find additional configurations that reduce the cost:
    (1) **Restricted Master Problem (RMP)** to use best existing configurations
    (2) **Pricing Problem (PP)** to construct new configurations

- A configuration =
  - **Working** path from source to primary DC
  - **Backup** path from source to secondary DC
  - **Sync** path between the primary & backup DCs



Column generation solution algorithm



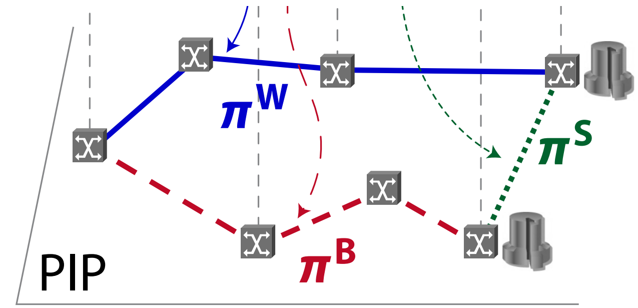
C. Develder, et al., "Resilient networks for multi-site data centers [...]", OSA Phontonic Netw. & Devices 2016

Restricted Master Problem (RMP)

$$\min \sum_{\ell \in L} \underbrace{\left(\beta_\ell^{\mathrm{W}} + \beta_\ell^{\mathrm{B}} + \beta_\ell^{\mathrm{S}}\right) \cdot \|\ell\|}_{\mathrm{BW}_\ell}$$

$$+ \mathrm{PENAL}^{\mathrm{DISRUPT\_BS}} \sum_{v \in V} \sum_{t \in T^\star} x_v^{\mathrm{BS},t}$$

$$+ \mathrm{PENAL}^{\mathrm{DISRUPT\_W}} \sum_{v \in V} \sum_{t \in T^\star} x_v^{\mathrm{W},t}$$



Case (ii):
minimize # disruptions of B/S path of multi-period traffic

Case (iii):
minimize # disruptions of W path of multi-period traffic

Constraints:

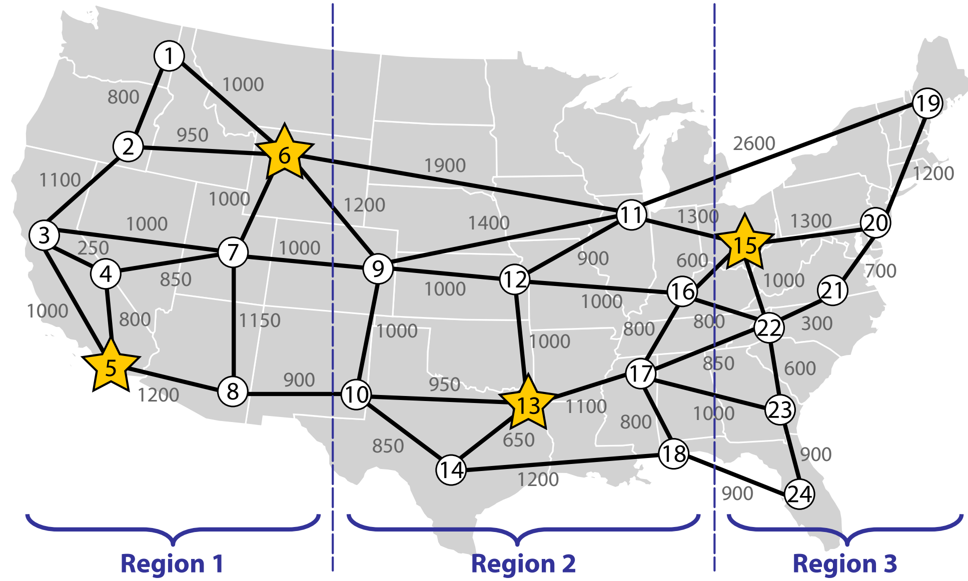
- Assure all requests are granted
- Count configuration changes $x^{\mathrm{BS},t}$, $x^{\mathrm{W},t}$
- Compute W, B, S bandwidths

Overview

Overview

1. Introduction
2. Problem statement
3. Model & solution approach
4. Case study
5. Conclusions

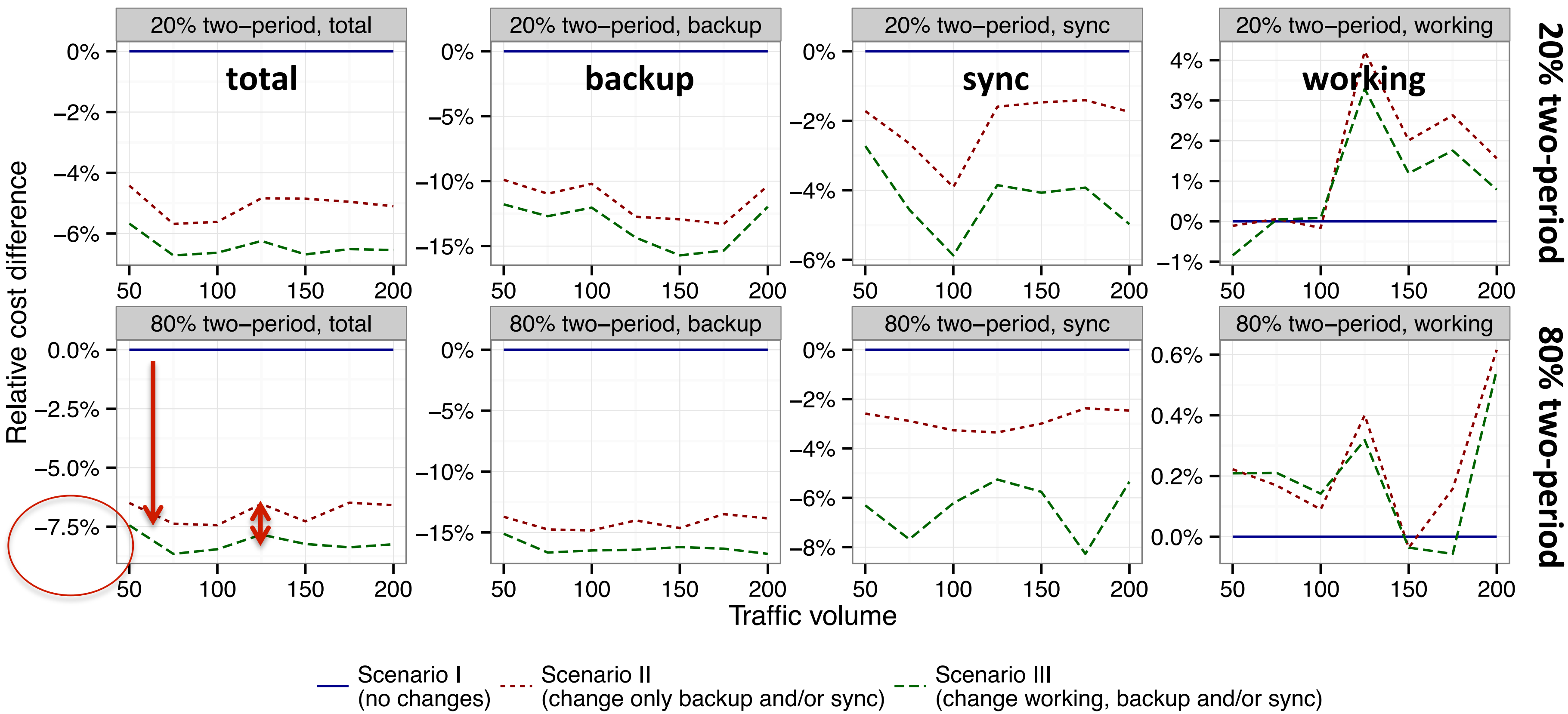
Case study



- Topology:
  - 24 nodes, 43 links
  - Data centers in ★: CA, WY, TX, OH
- Traffic: 3-region case
  - **Total traffic:** 33.3% region 1, 37.5% region 2, 29.2% region 3
  - **Three periods:** A: 14%, B: 38%, C: 48%
    - Region 1: A, B, C
    - Region 2: B, C, A
    - Region 3: C, A, B
  - **Duration:**
    - *Pattern #1:* 20% two-period, 80% single period traffic
    - *Pattern #2:* 80% two-period, 20% single period traffic

Results: Net total bandwidth savings?

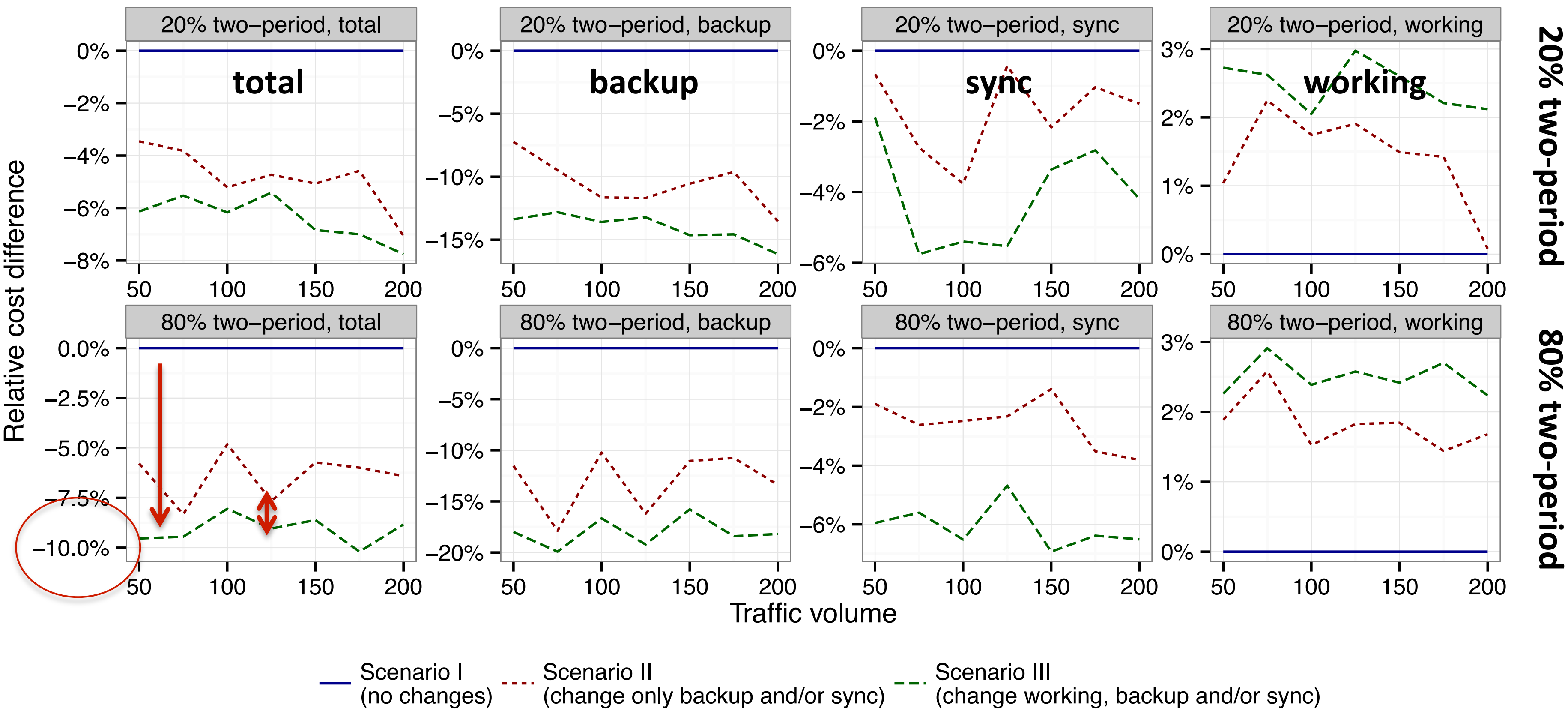
1. Relative total cost savings up to nearly 8% (pattern #2, i.e., more multi-period traffic)
2. Capacity savings are realized mainly by **sharing of backup** (backup savings >15%)
3. Saving by only changing backup/synchronization (Scenario II) almost as good as when also changing working (Scenario III)



C. Develder, et al., "Resilient networks for multi-site data centers [...]", OSA Phontonic Netw. & Devices 2016

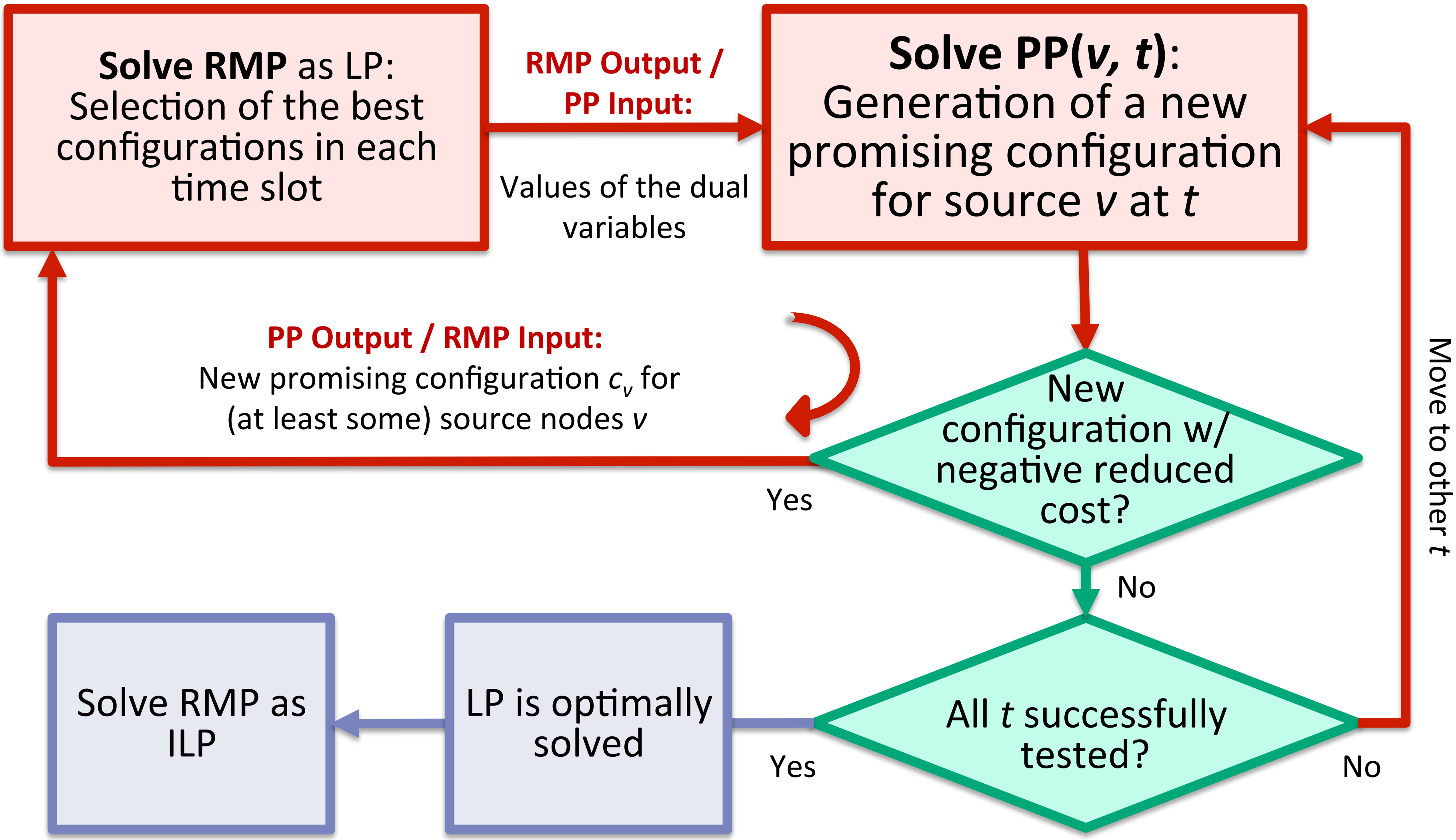
Results: Net total bandwidth savings for 4 regions

1. Relative total cost savings up to nearly **10%** (pattern #2, i.e., more multi-period traffic)
2. Capacity savings are realized mainly by **sharing of backup** (backup savings >15%)
3. Saving by only changing backup/synchronization (Scenario II) almost as good as when also changing working (Scenario III)



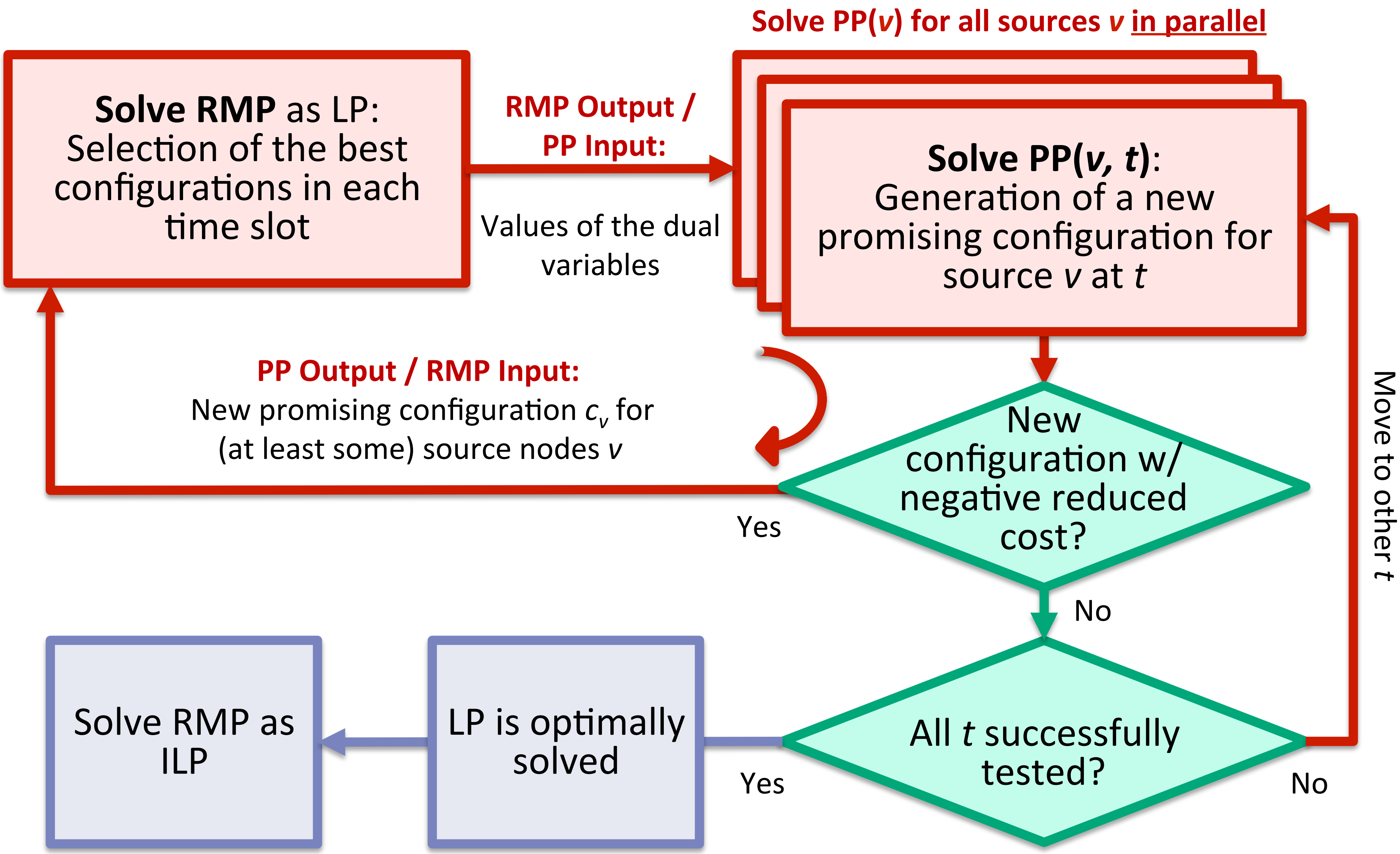
C. Develder, et al., "Resilient networks for multi-site data centers [...]", OSA Phontonic Netw. & Devices 2016

Solution scheme: serial vs parallel



C. Develder, et al., "Resilient networks for multi-site data centers [...]", OSA Phontonic Netw. & Devices 2016

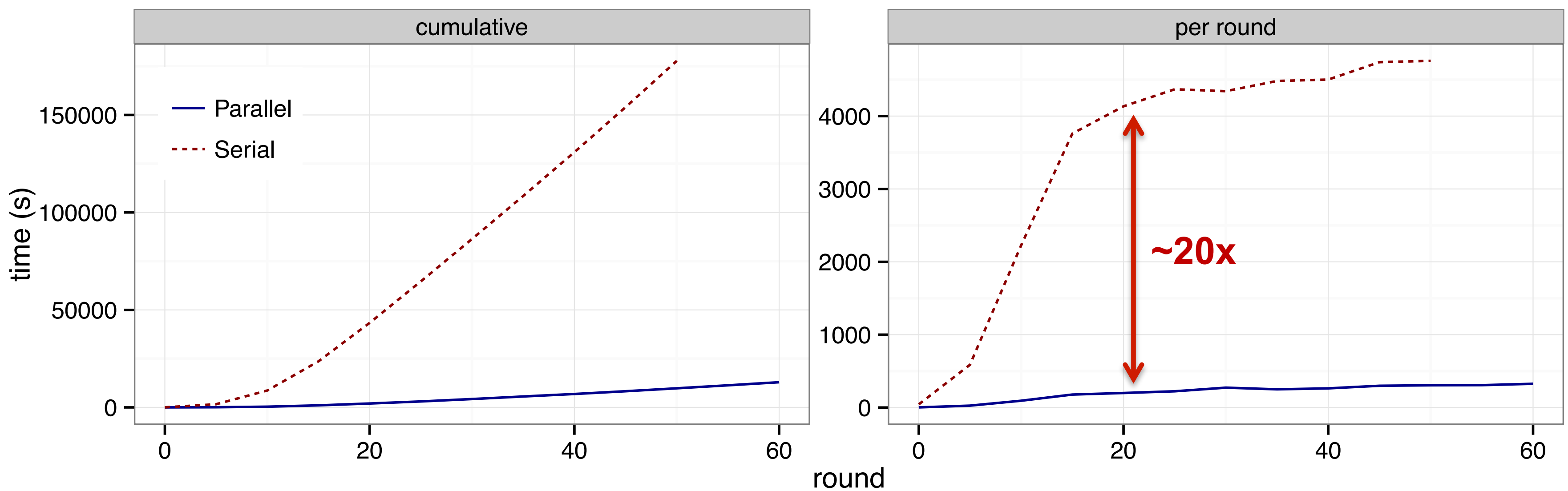
Solution scheme: serial vs parallel



C. Develder, et al., "Resilient networks for multi-site data centers [...]", OSA Phontonic Netw. & Devices 2016

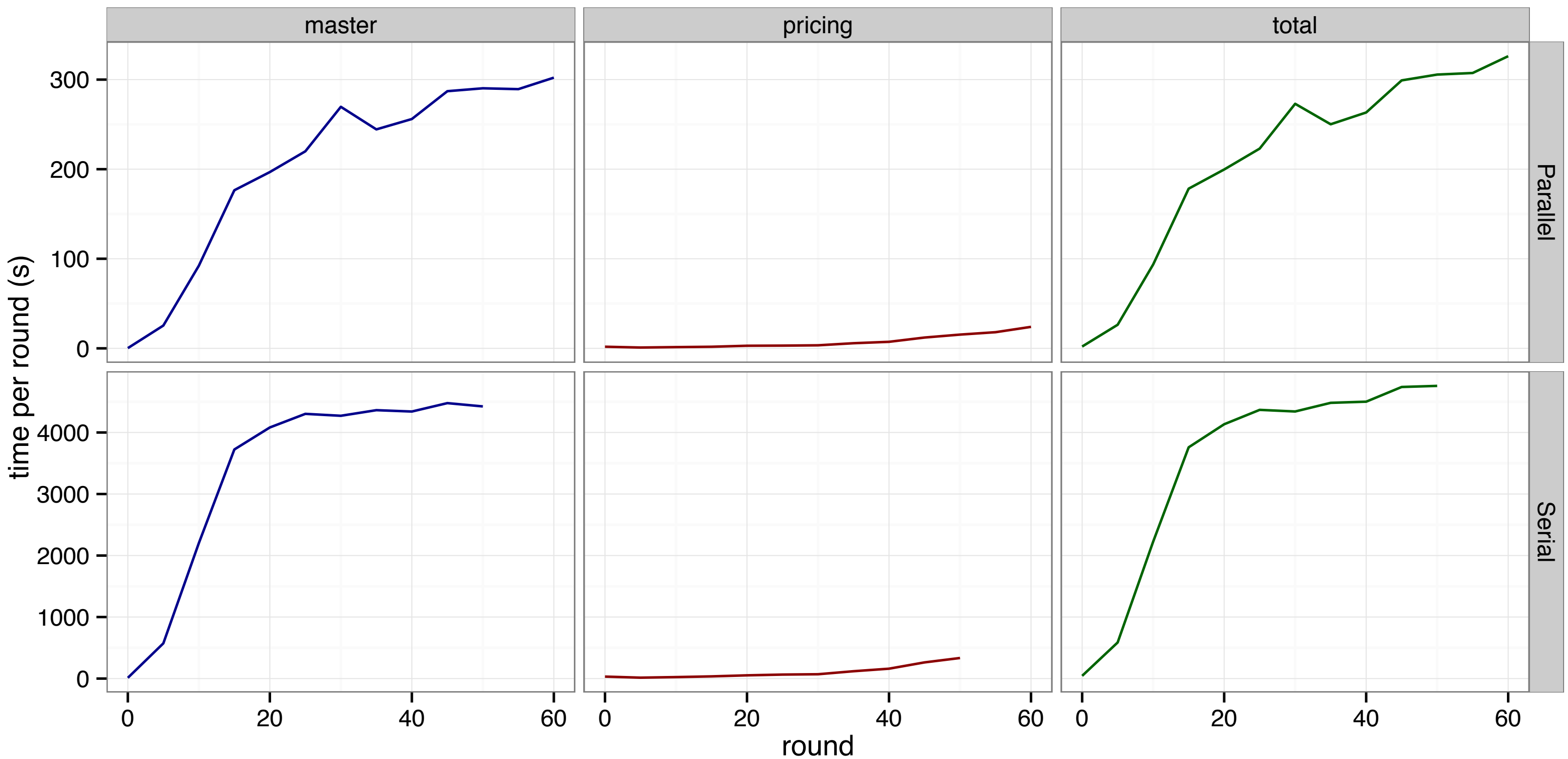
Scalability: Time savings by parallel PP solving

Only re-solve RMP after adding multiple configurations (i.e., for multiple source nodes)!



Scalability: Time savings by parallel PP solving

Only re-solve RMP after adding multiple configurations (i.e., for multiple source nodes)!



Overview

Overview

1. Introduction
2. Problem statement
3. Model & solution approach
4. Case study
5. Conclusions

Conclusions

- Scalable column-generation method (w/ parallel solving of multiple PPs) for resilient VNet planning of time-varying traffic, **over all periods together**
- Our (relatively limited) case study shows that:
  - Changing routing from one period to the next saves several % of the total bandwidth cost (mostly backup cost savings)
  - … but we need only to change about 50% of them
  - … and only changing backup/synchronization seems to suffice
  - Savings seem to increase for (i) more multi-period traffic, (ii) more regions
- Future work: Optimize DC locations
  (e.g., 'scattered' vs 'paired', see ICTON 2013)

Thank you … any questions?

Chris Develder

chris.develder@intec.ugent.be

Ghent University – iMinds