Cutting Electricity Cost For Service Provider Networks

Muhammad Saqib Ilyas

FDC

Zartash Afzal Uzmi

Tariq Mahmood Jadoon

Ihsan Ayyub Qazi

Muhamad Fareed Zaffar

Aamir Qayyum

Agenda

- Background and motivation
- Opportunity and key idea
- Case studies:
  – Data centers (e.g., Facebook and Google)
  – Cellular networks (e.g., Sprint and Verizon)
- Conclusions and future work

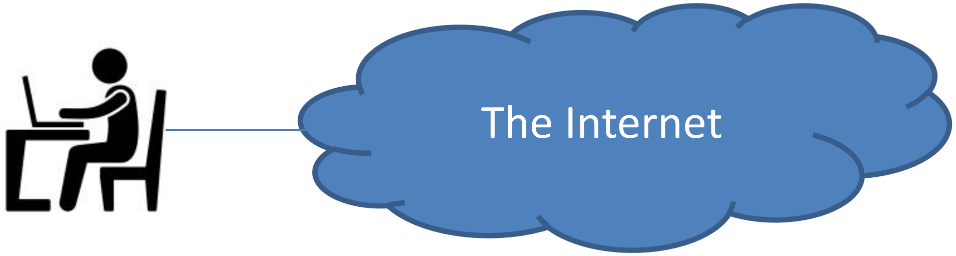
Background

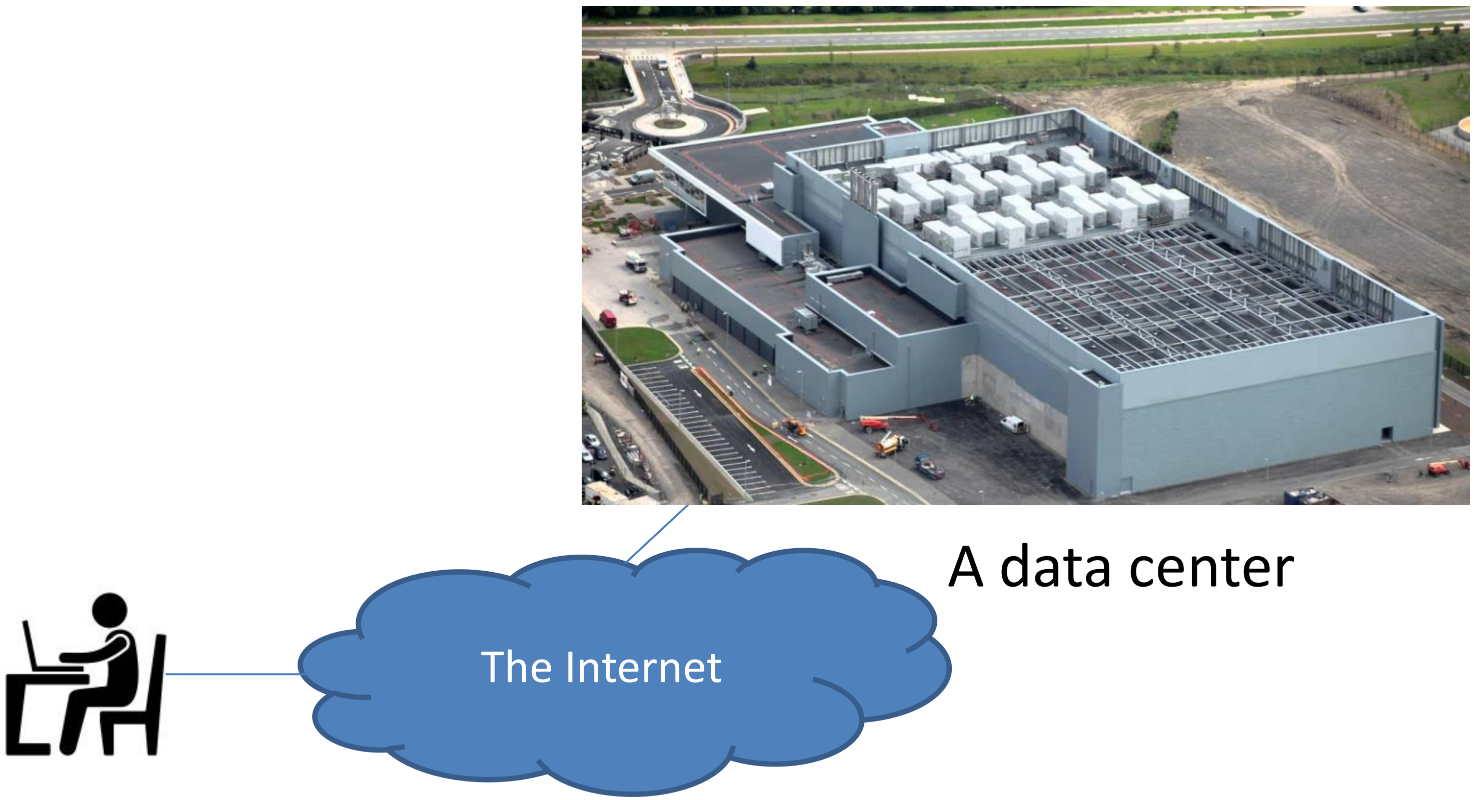
Source: freeimages.com

Background

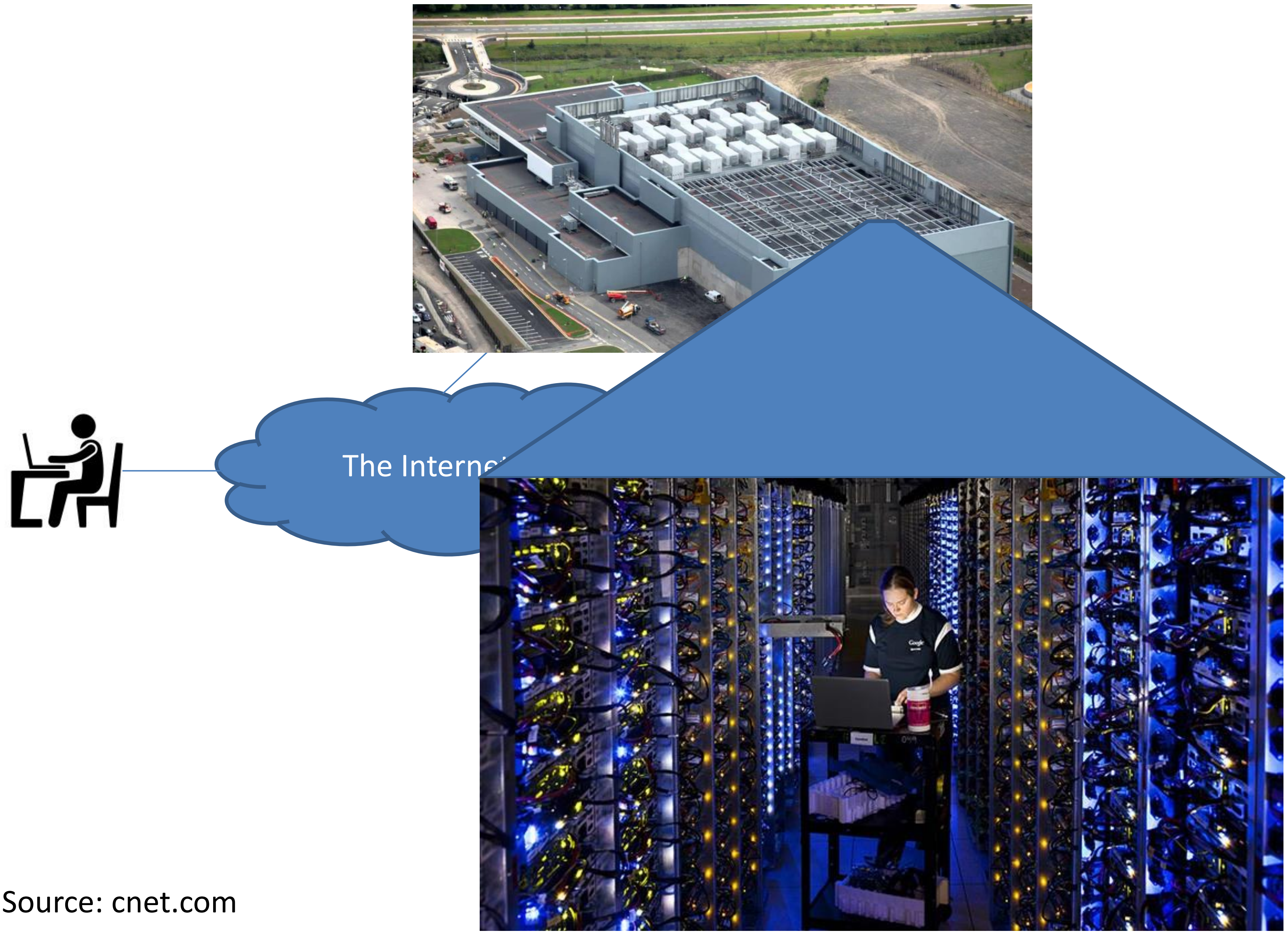
Source: freeimages.com

Background



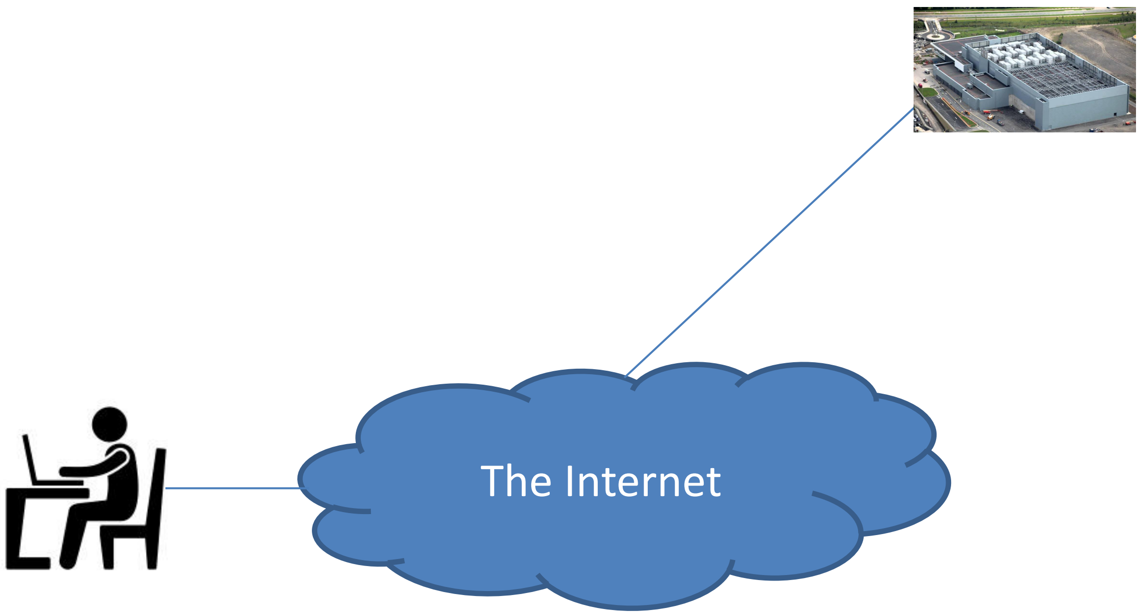


Source: youtube.com

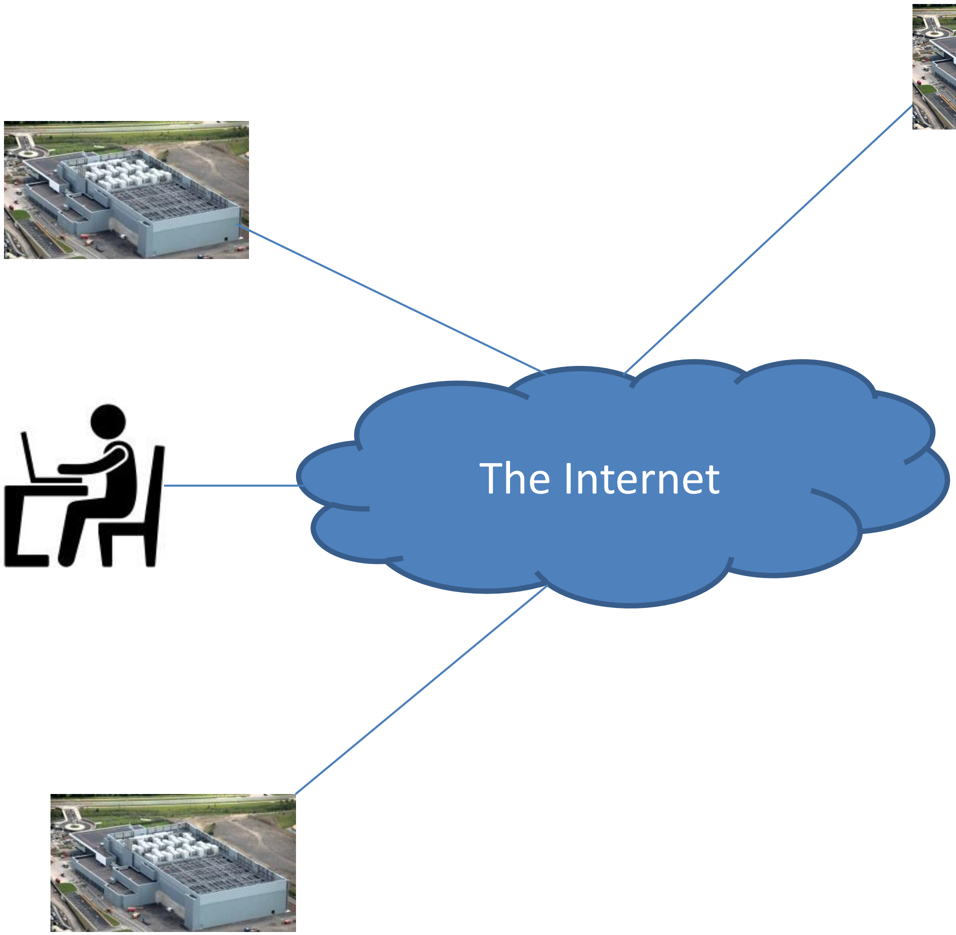


Source: cnet.com

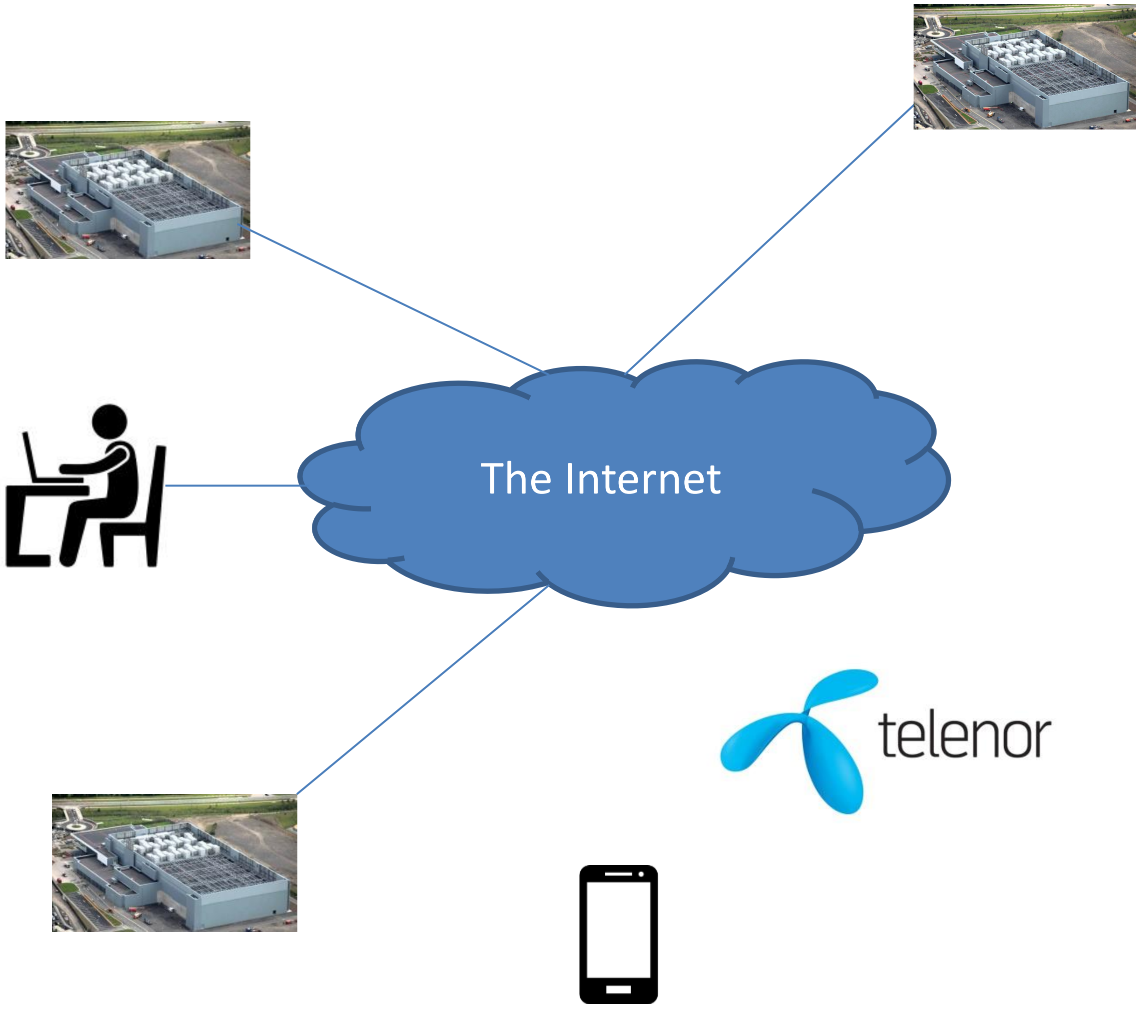
Background



Background

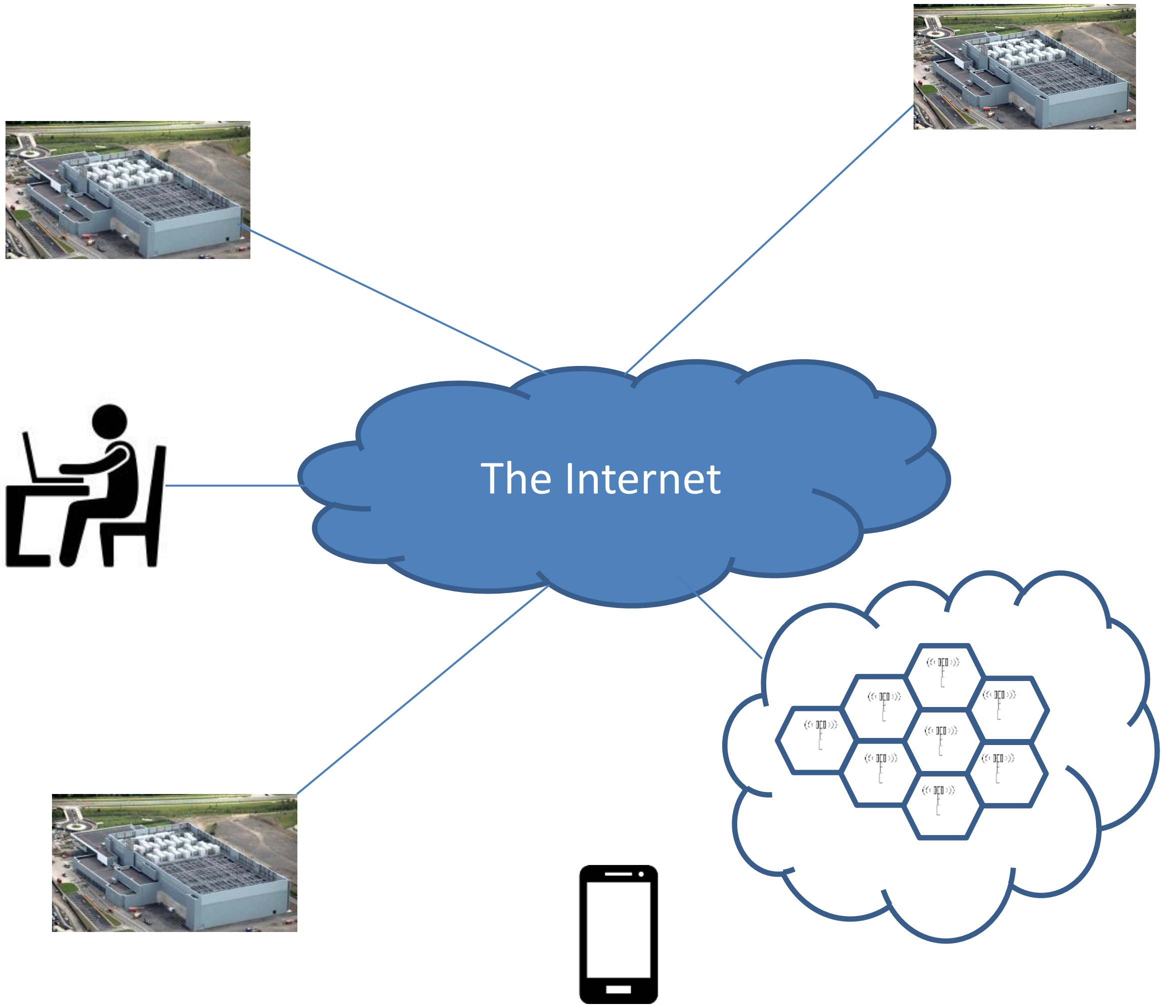


Background



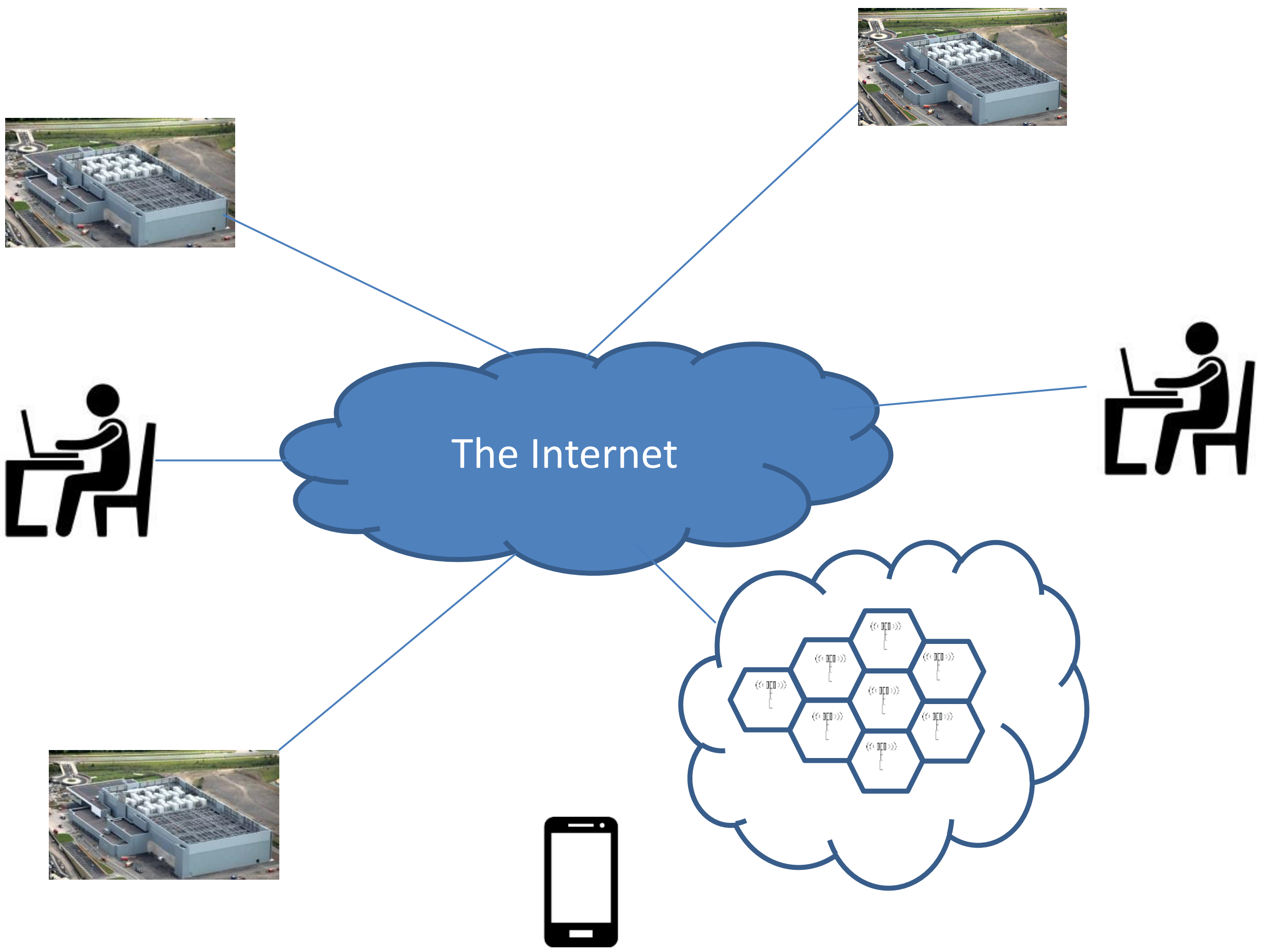
Source: flaticon.com

Background

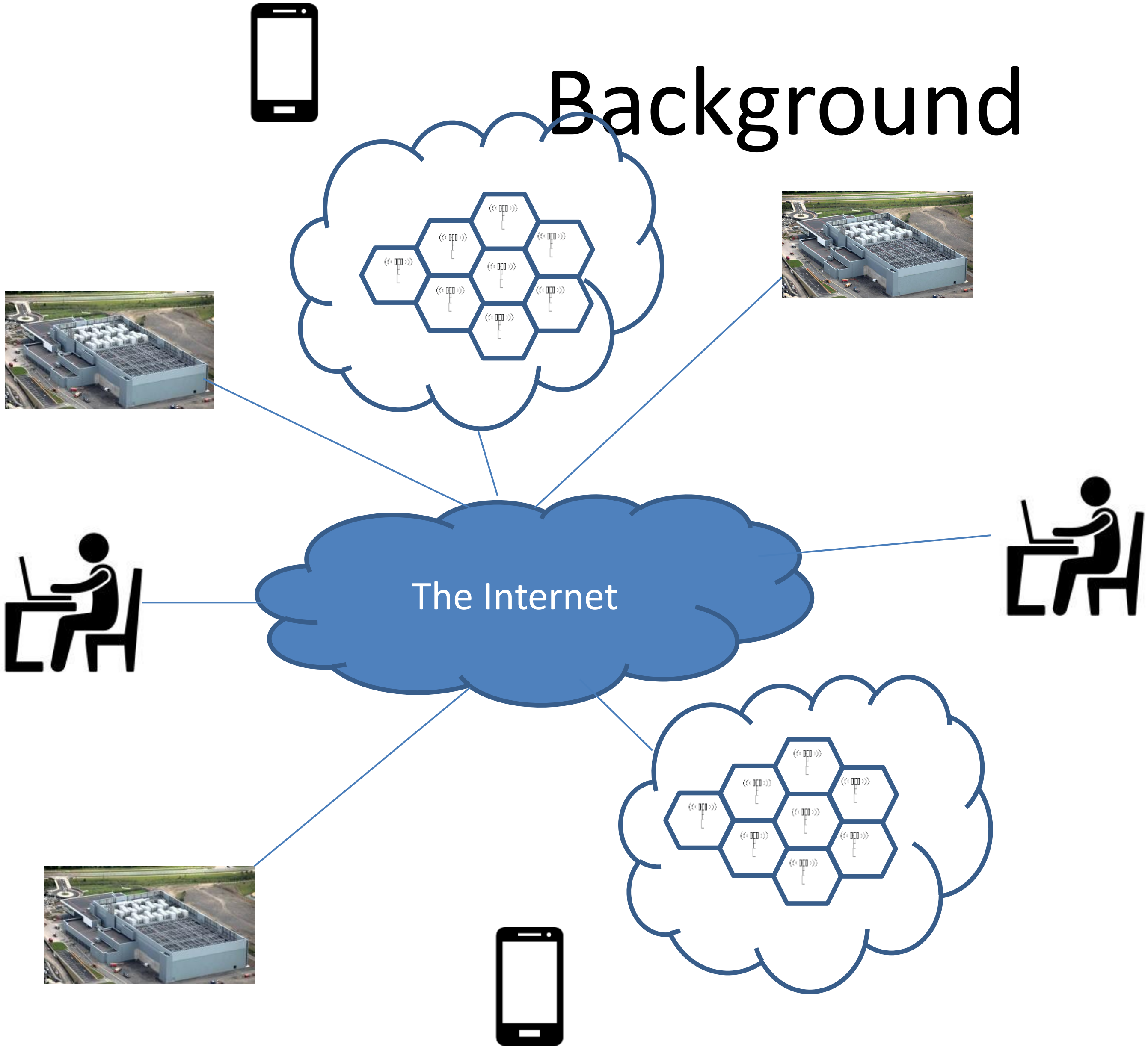


Source: vcetsiras.blogspot.com

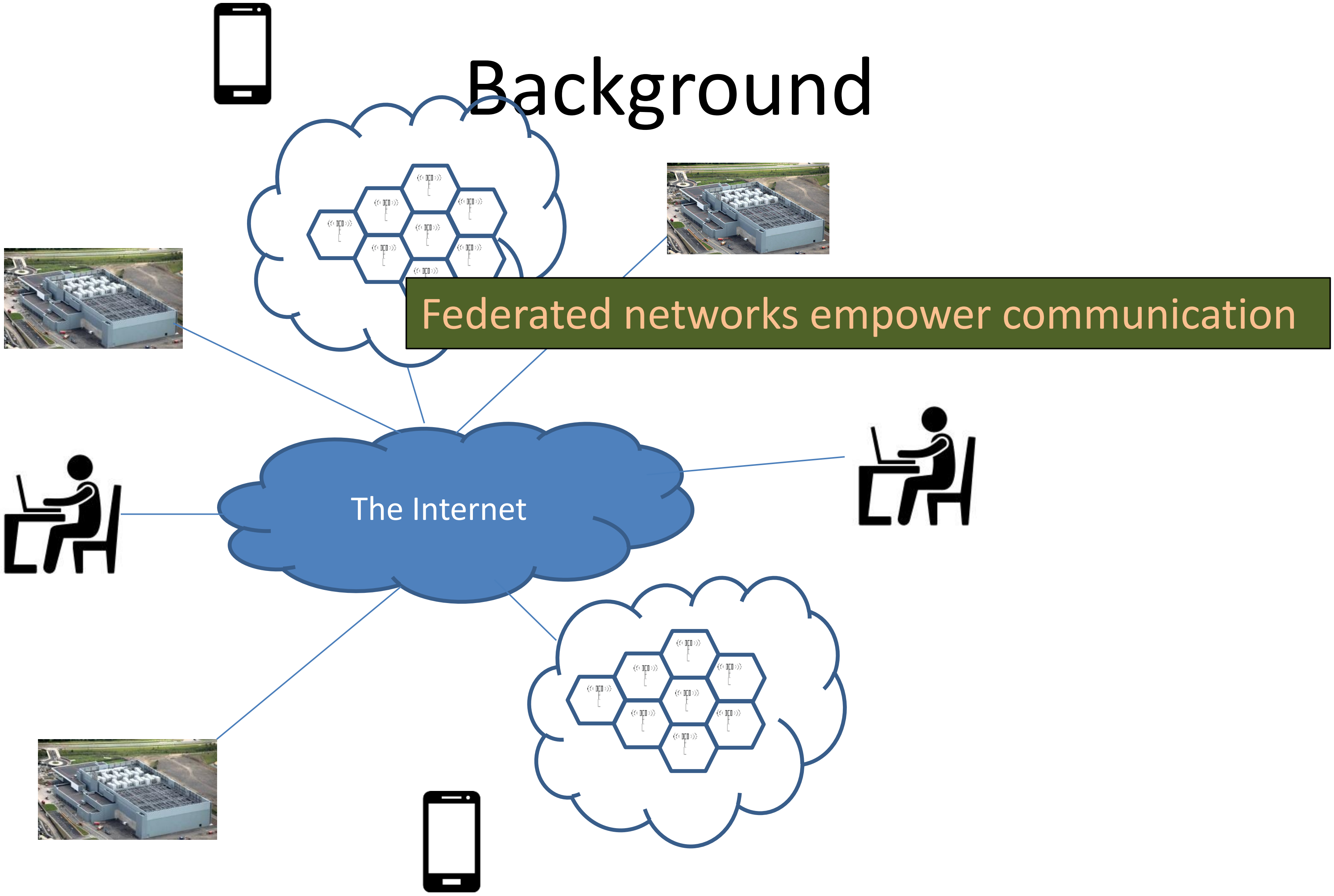
Background



Background



Background



Network Scale

Image source: http://bit.ly/1awWnLn

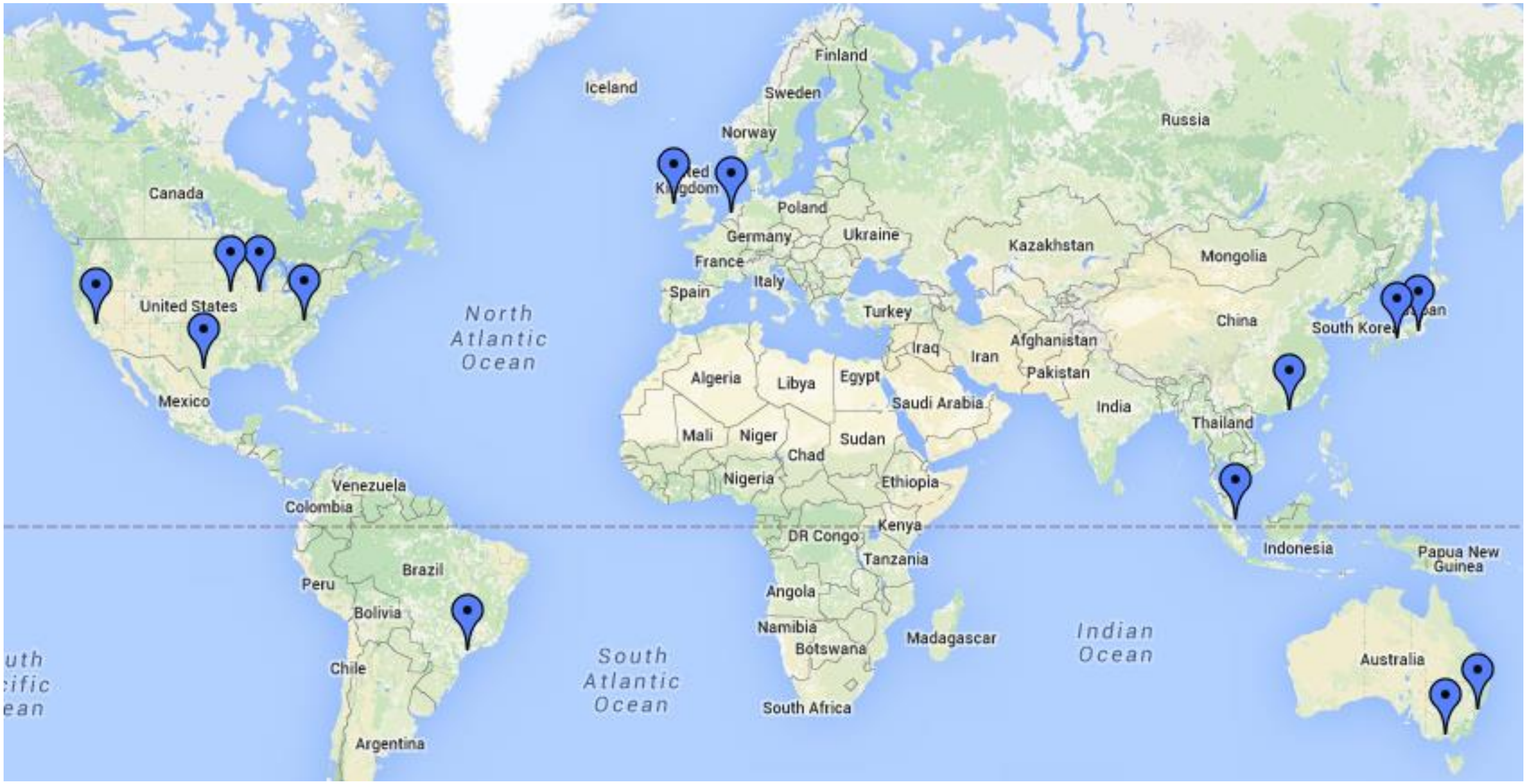
Network Scale

1 Data Center ~ 50,000 - 80,000 servers

Google’s data center locations
http://bit.ly/1WbIvbe

Network Scale

1 Data Center ~ 50,000 - 80,000 servers

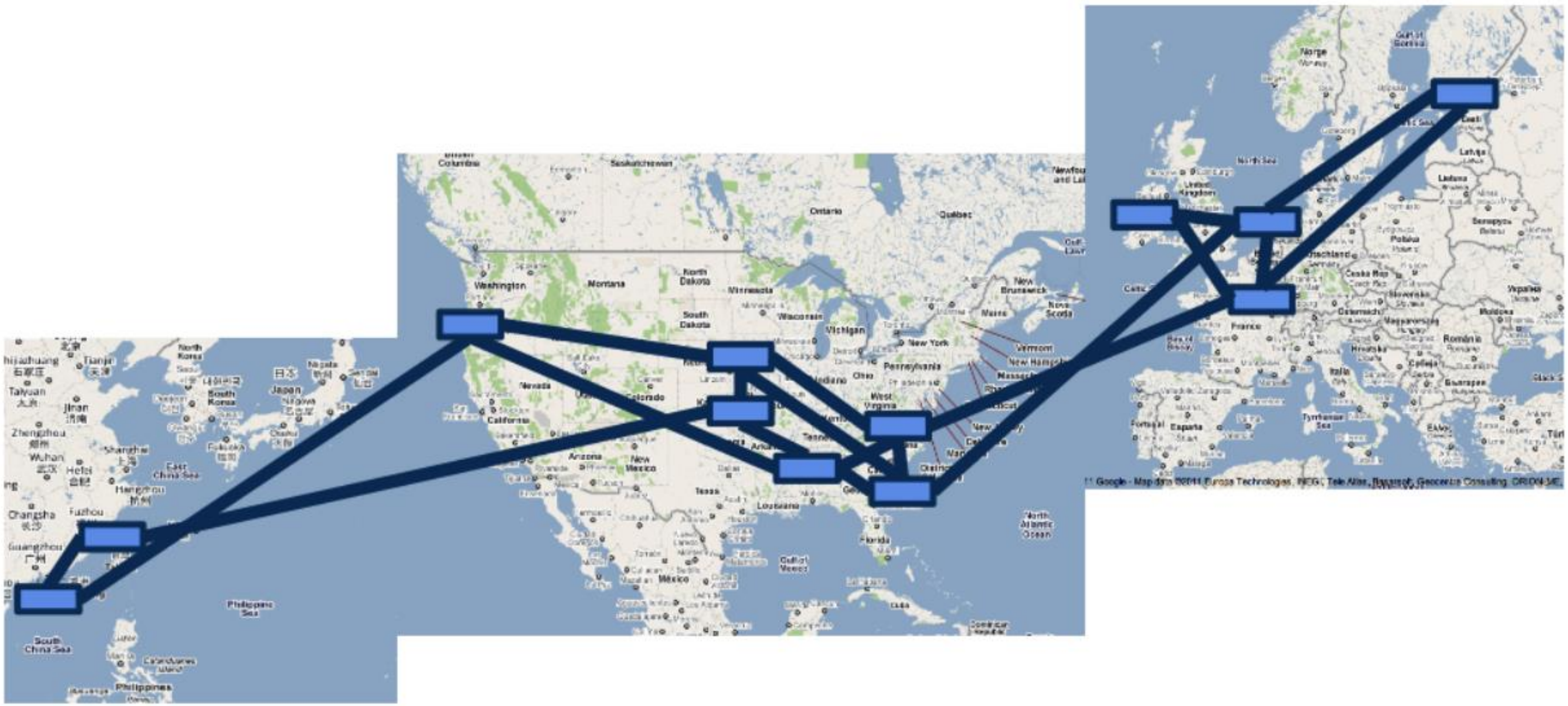


Microsoft Azure's data center locations
http://bit.ly/1mqvi26

Network Scale

Parameter	Microsoft	Amazon	Google
Servers	Millions	2.8 – 5.6 M	900000
Data Centers	> 26	~ 87	14

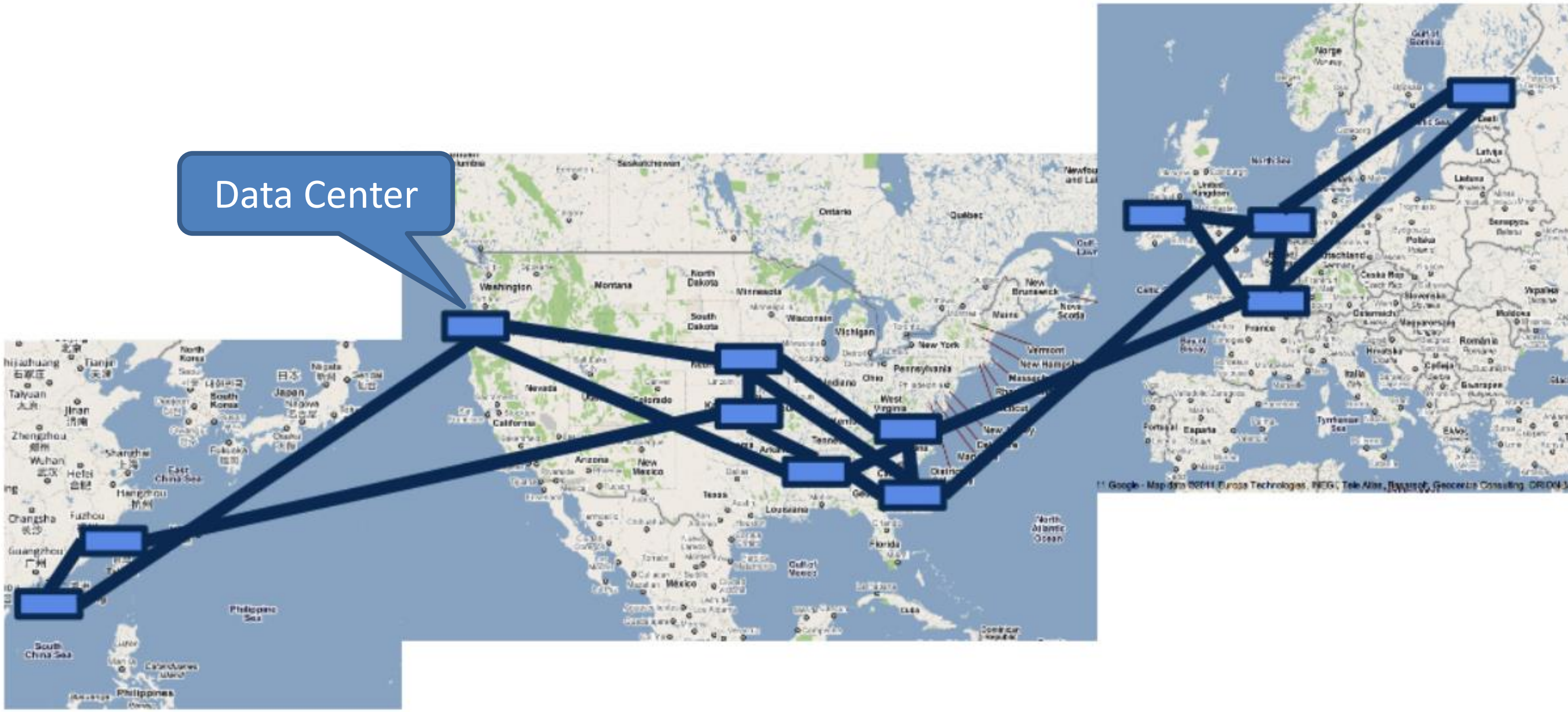
Network Scale



Google's B4 SDN

Image Source: Jain et. al, "B4: Experience with a globally-deployed software defined WAN"

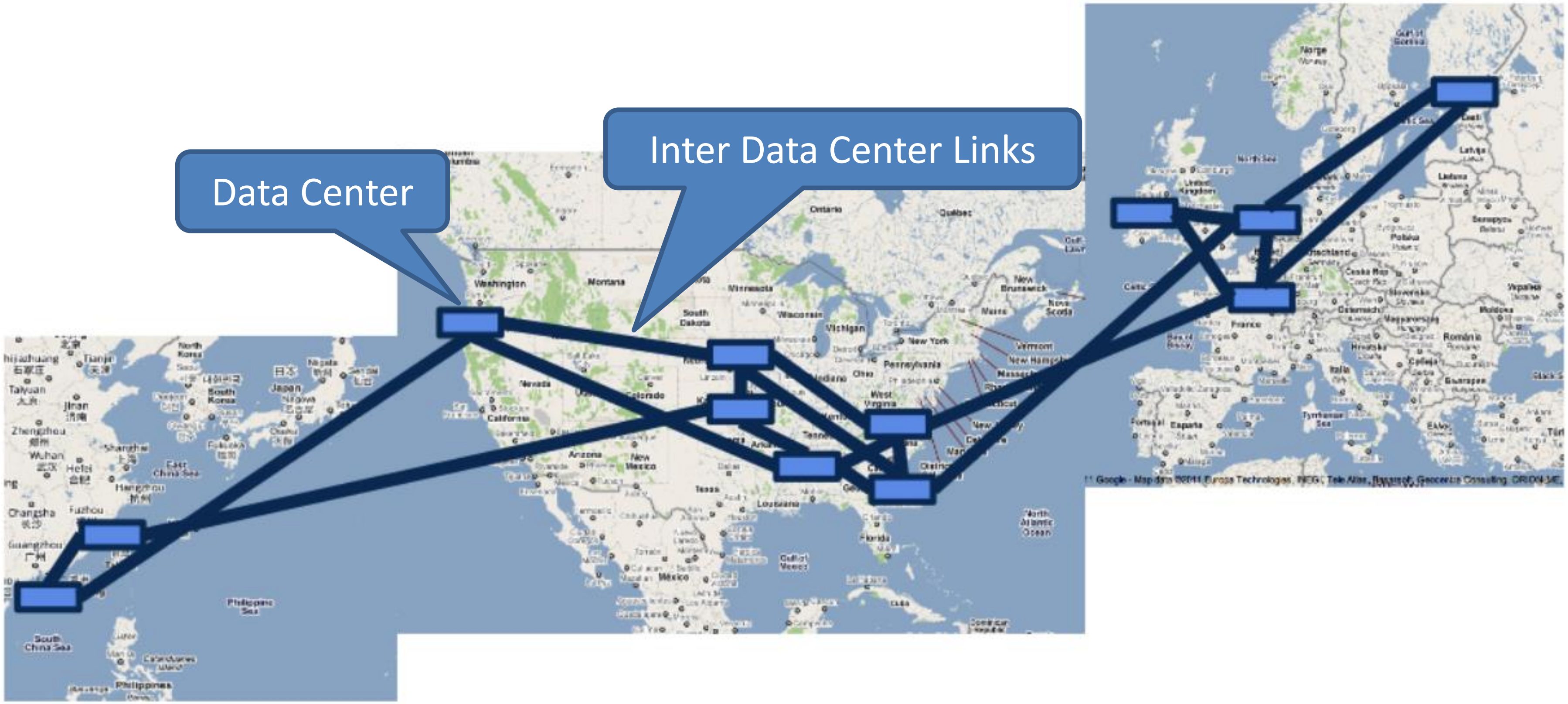
Network Scale



Google’s B4 SDN

Image Source: Jain et. al, “B4: Experience with a globally-deployed software defined WAN”

Network Scale



Google's B4 SDN

Image Source: Jain et. al, "B4: Experience with a globally-deployed software defined WAN"

Network Scale

Telenor: 8000 cell sites

Network Scale

Massive infrastructure

Network Scale

Massive infrastructure ➡ Massive power draw

Network Scale

Network Scale

Motivation

Motivation

Annual DC Opex

$951 M

Source: datacenterknowledge.com

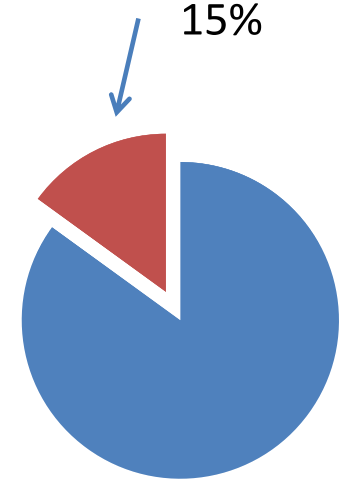
Motivation

Annual DC Opex

$951 M

Electricity Cost

$143 M



Source: Greenberg et. al, “Cost of a cloud: Research Problems in Data Center Networks”, CCR 2009

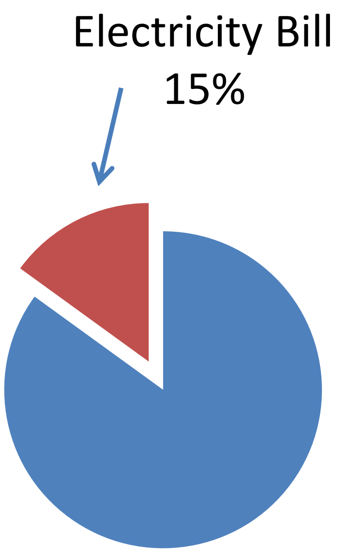
Motivation

Annual DC Opex

$951 M

Electricity Cost

$143 M



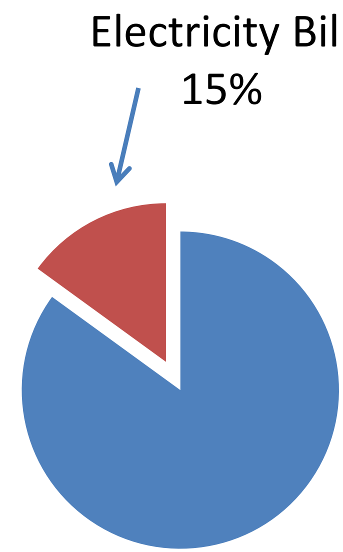
Motivation

Annual DC Opex

$951 M

Electricity Cost

$143 M



Electricity Cost 2012

$81 M

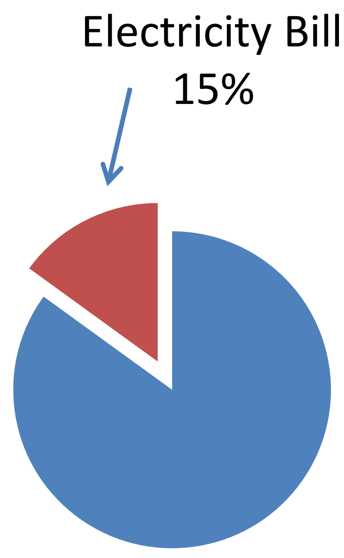
Motivation

Annual DC Opex

$951 M

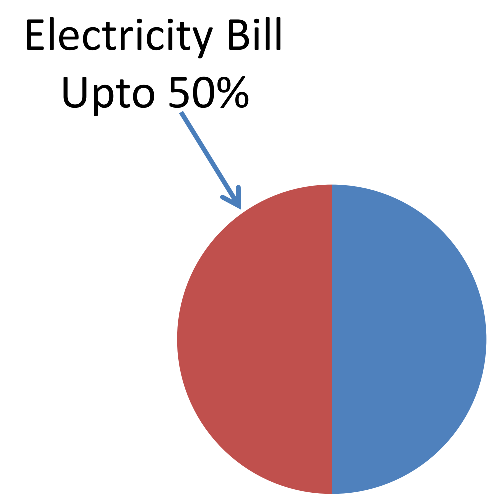
Electricity Cost

$143 M



Electricity Cost 2012

$81 M



Source: GREENNETS

Motivation

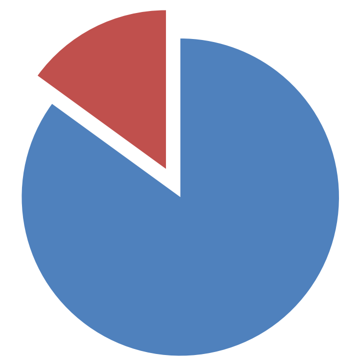
Annual DC Opex

$951 M

Electricity Cost

$143 M

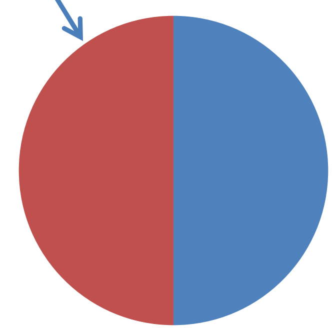
Electricity Bill
15%



Electricity Cost 2012

$81 M

Electricity Bill
Upto 50%

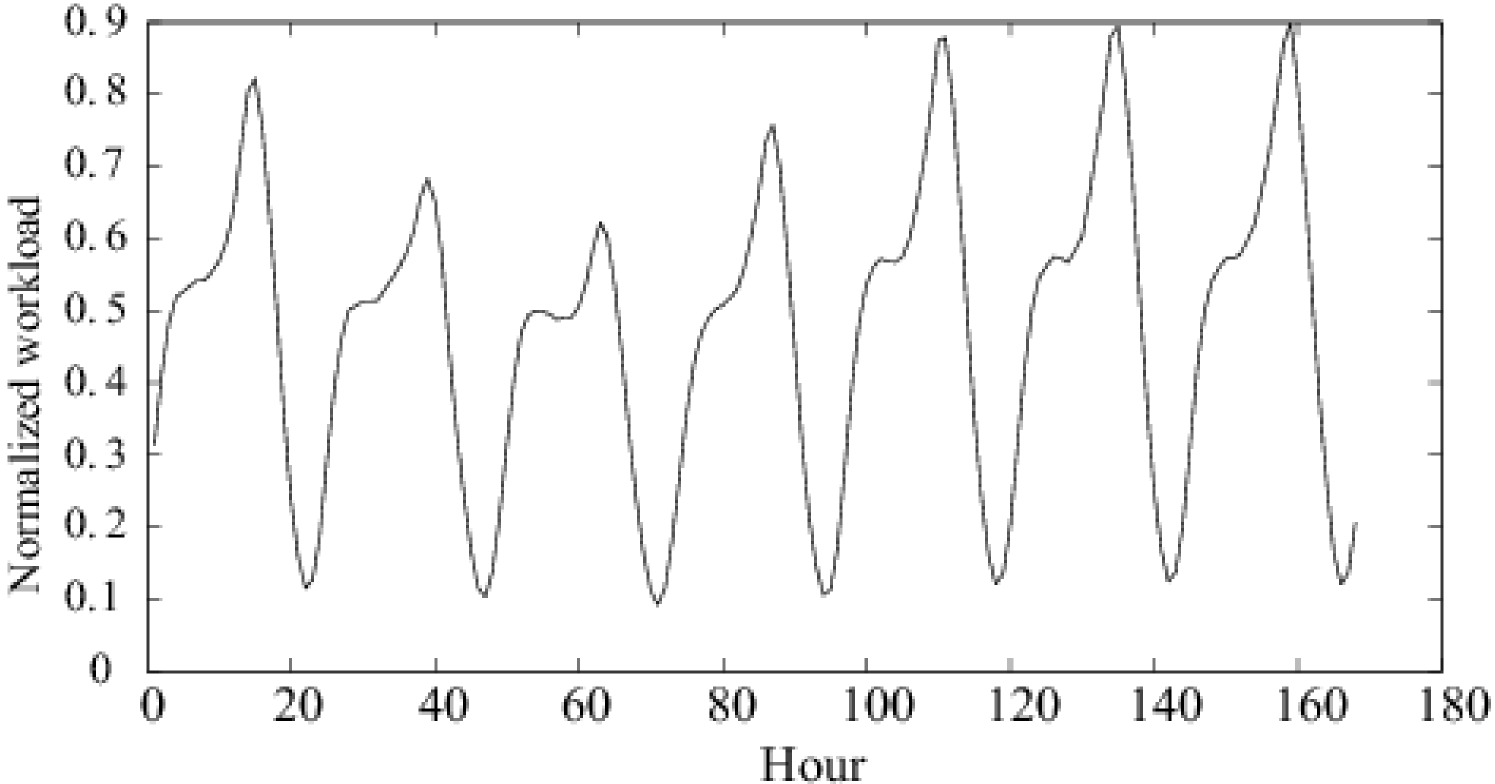


Significant electricity costs

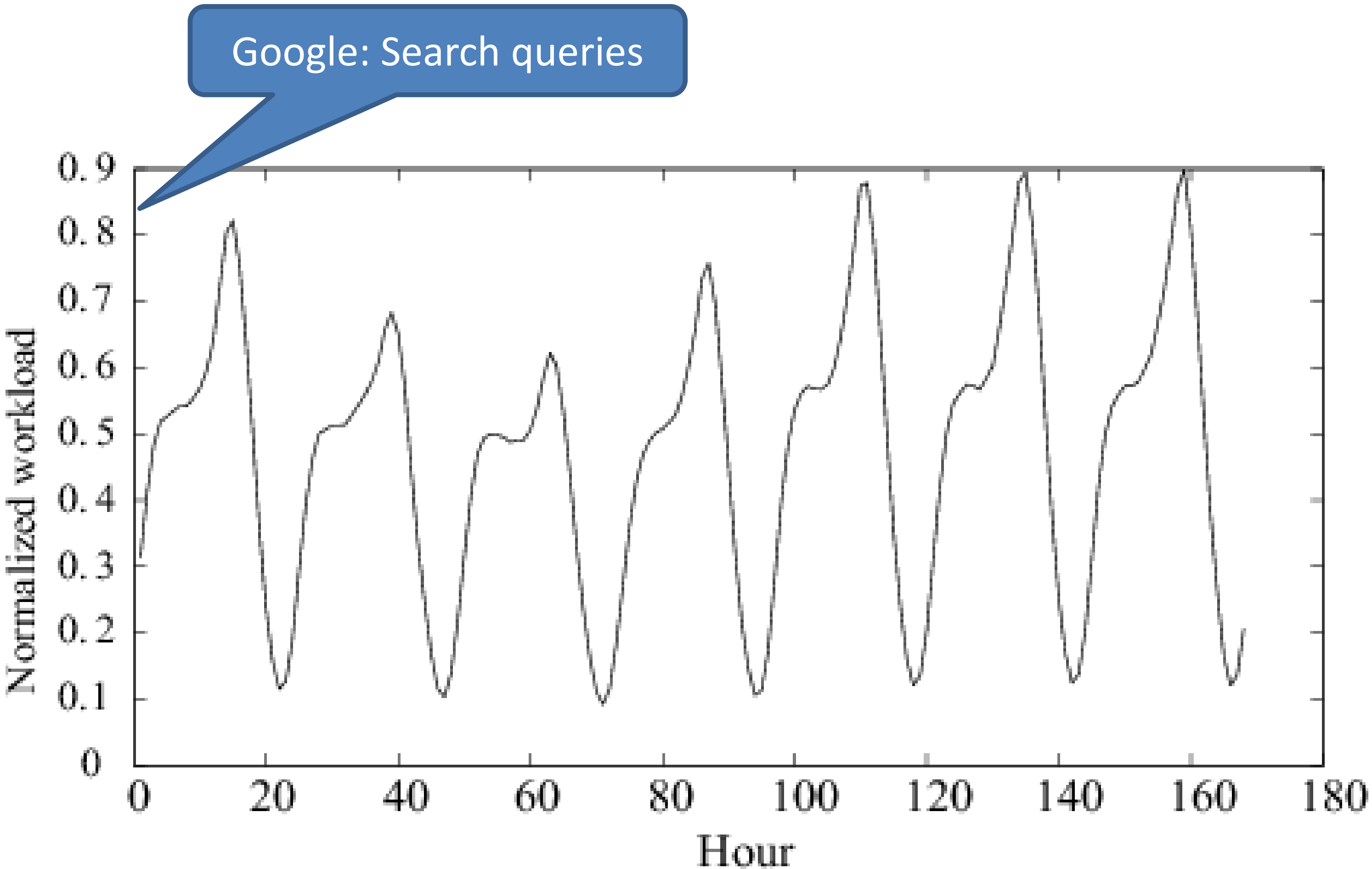
Agenda

- Background and motivation
- **Opportunity and key idea**
- Case studies:
  - Data centers (e.g., Facebook and Google)
  - Cellular networks (e.g., Sprint and Verizon)
- Conclusions and future work

Opportunity

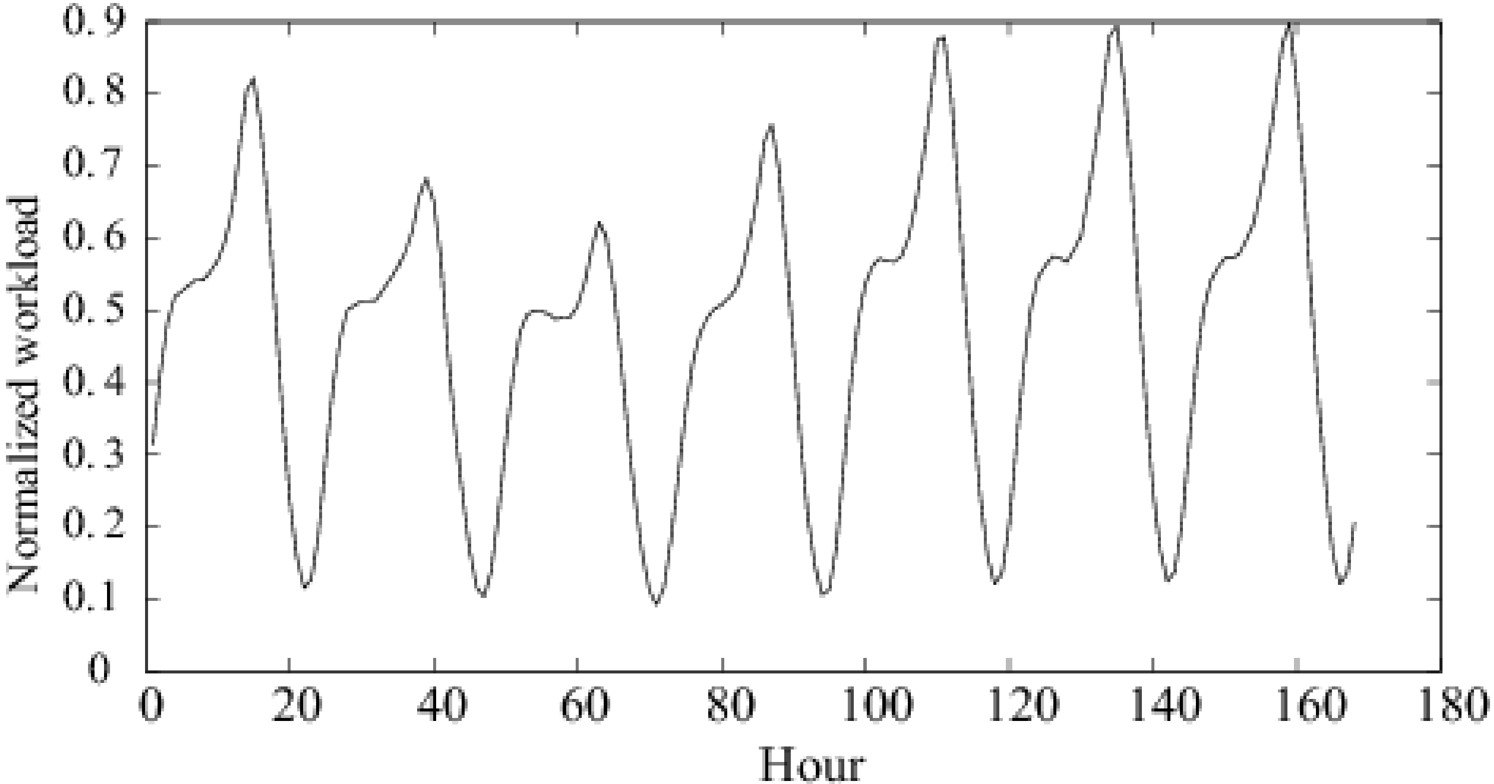


Opportunity



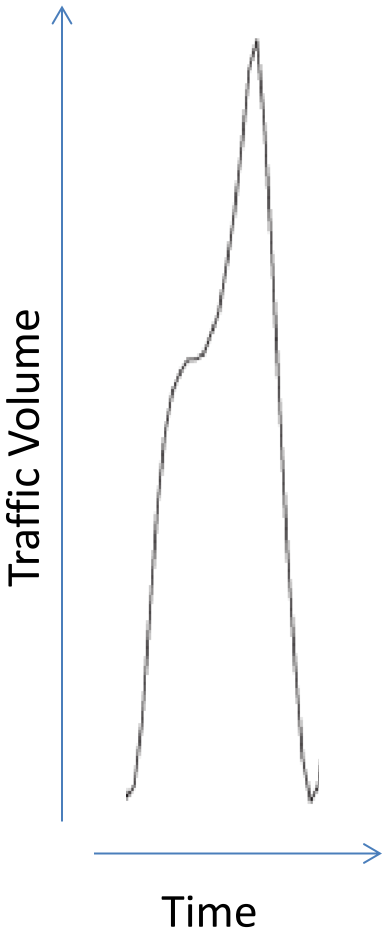
Opportunity

Diurnal cycles



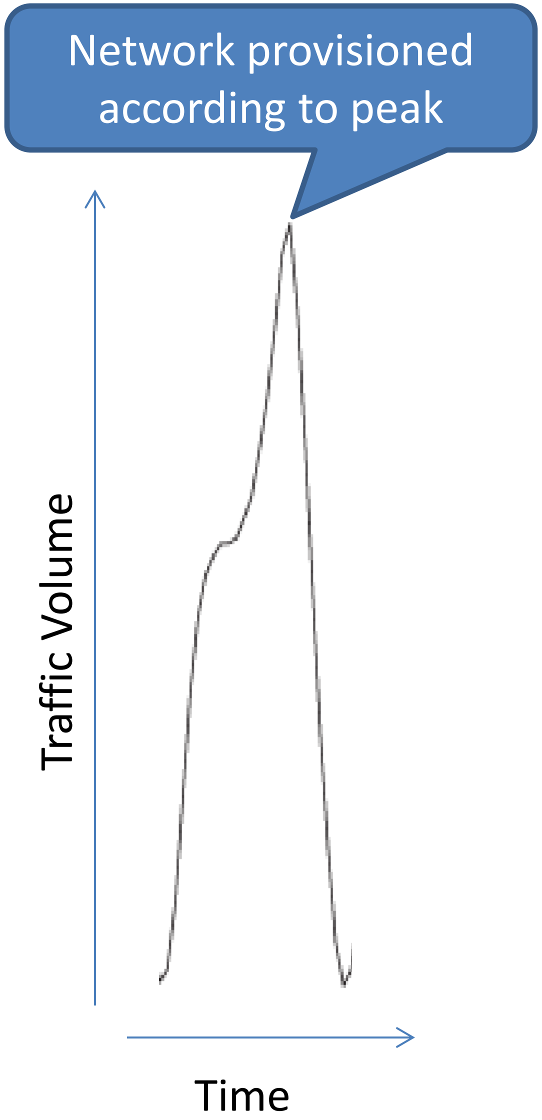
Barroso et. al, “The Case for Energy Proportional Computing”, IEEE Computer, 2007
Peng et. al, “Traffic-Driven Power Savings in Operational 3G Cellular Networks”, MOBICOM 2011

Opportunity



Peak and trough are quite pronounced

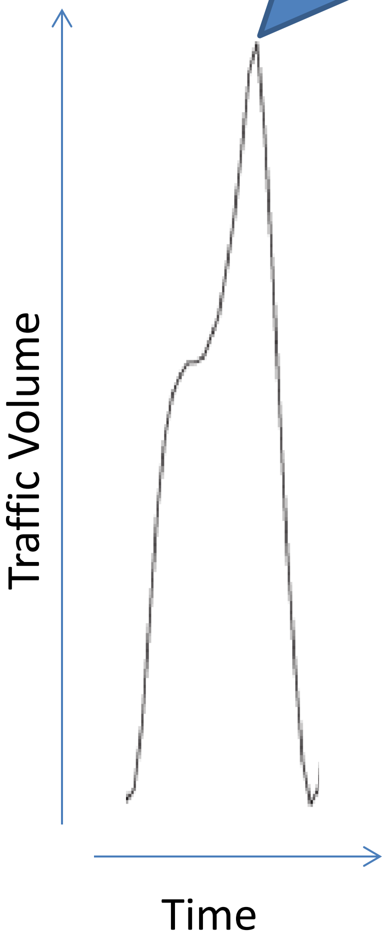
Opportunity



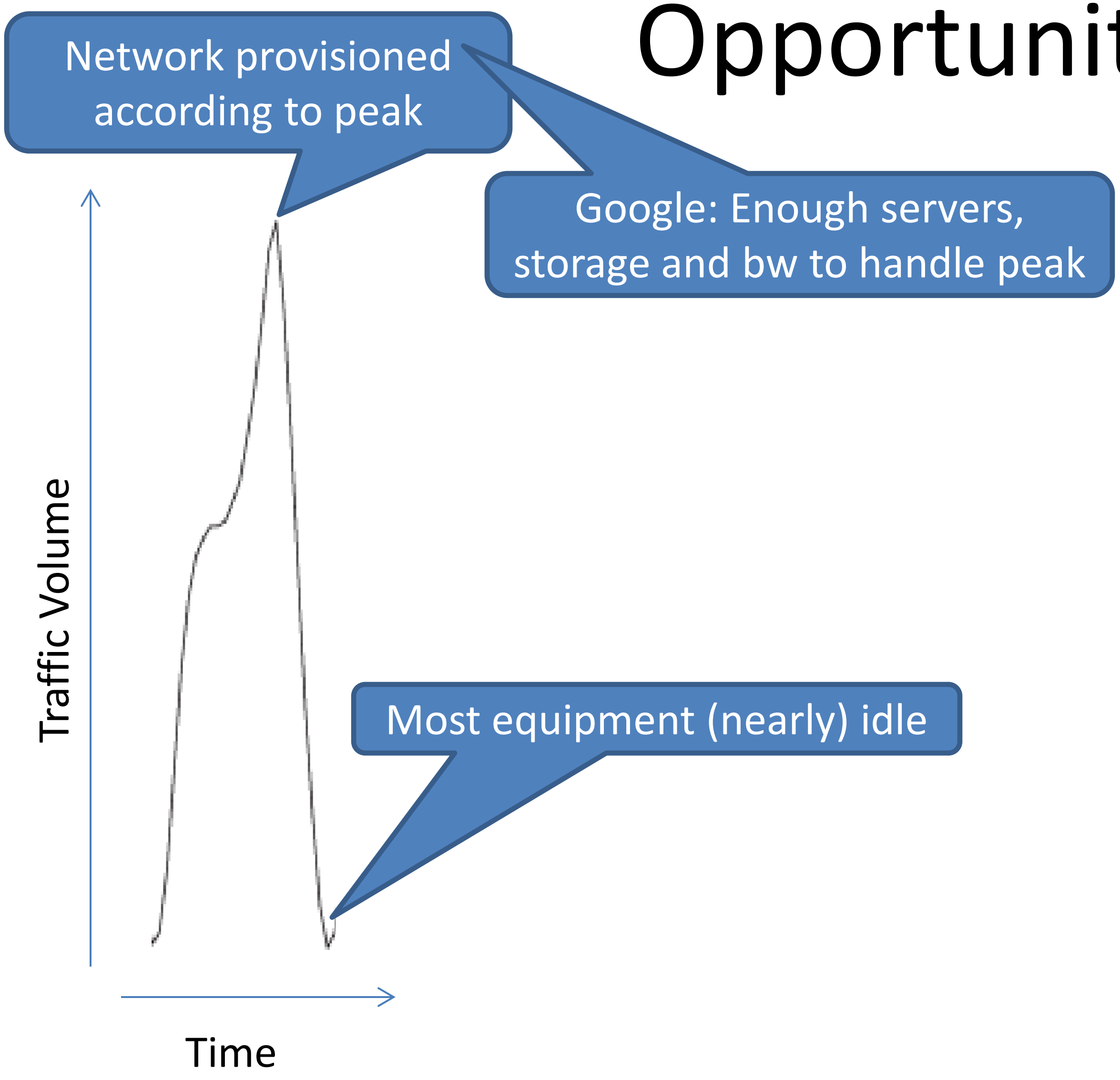
Opportunity

Network provisioned according to peak

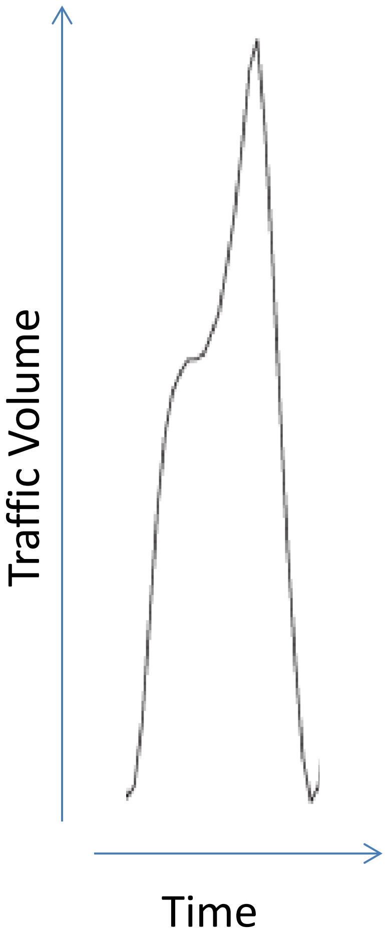
Google: Enough servers, storage and bw to handle peak



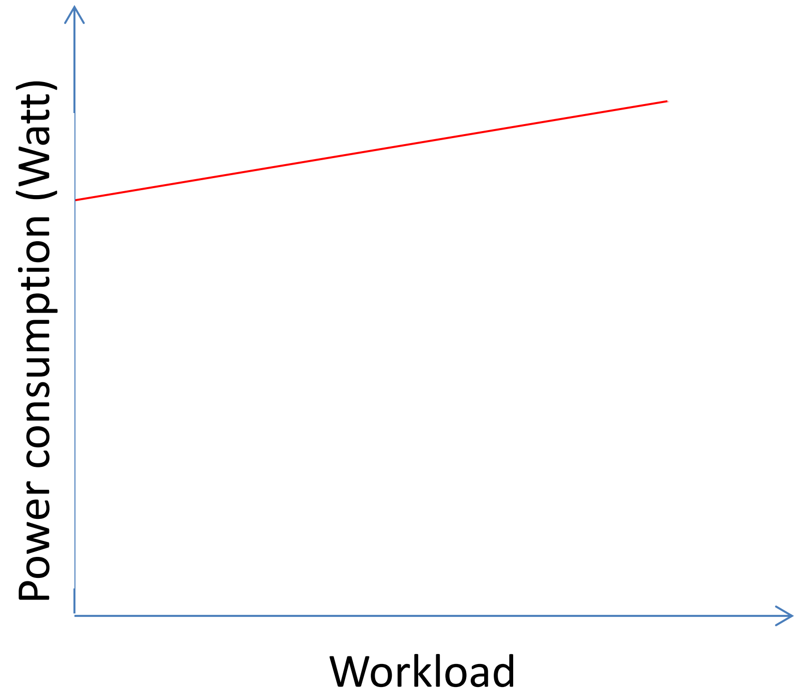
Opportunity



Opportunity

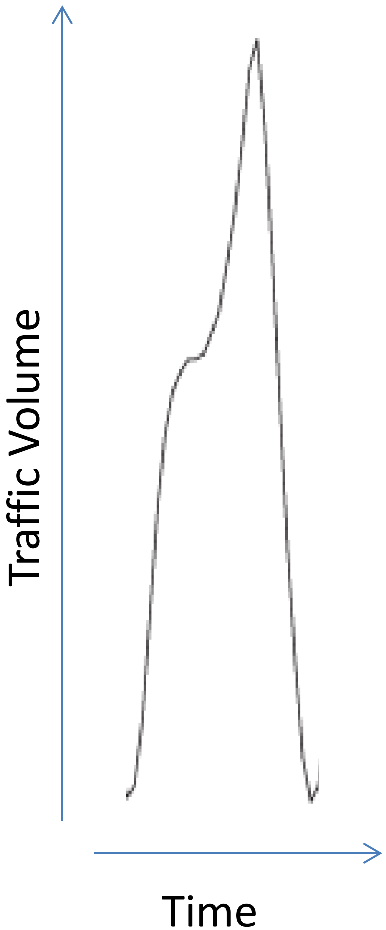


DC Idle power
75%-85% of peak power

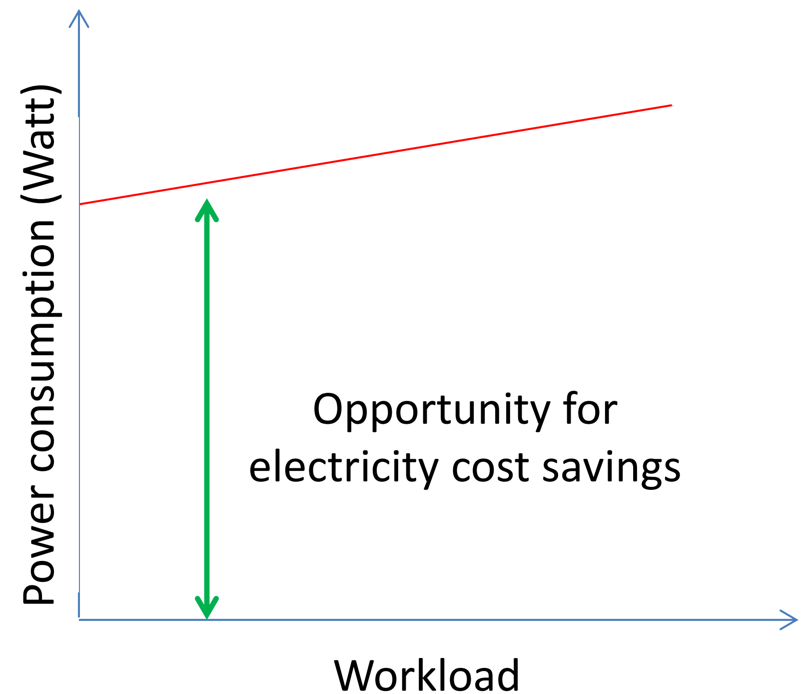


Source: Emerson, “Energy Logic: Reducing Data Center Power Consumption…”

Opportunity



DC Idle power
75%-85% of peak power

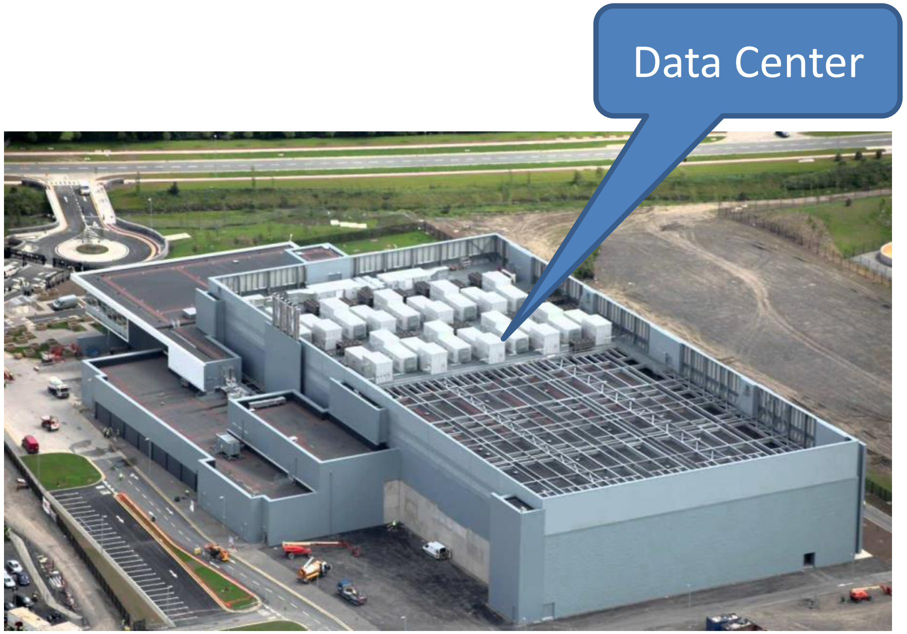


Source: Emerson, "Energy Logic: Reducing Data Center Power Consumption…"

Key Idea

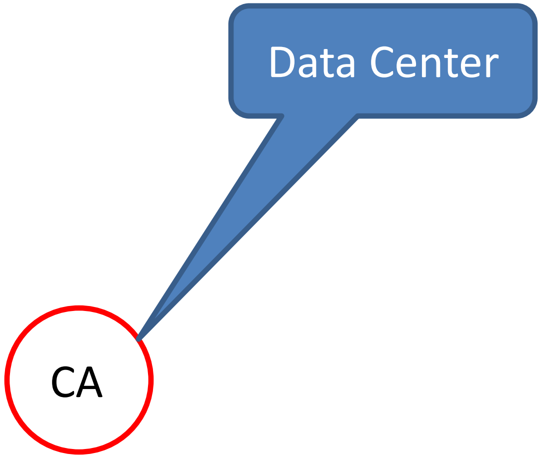
Deactivate idle equipment

Key Idea



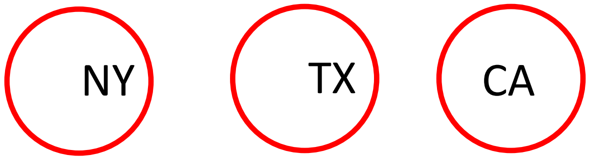
Key Idea

CA: California



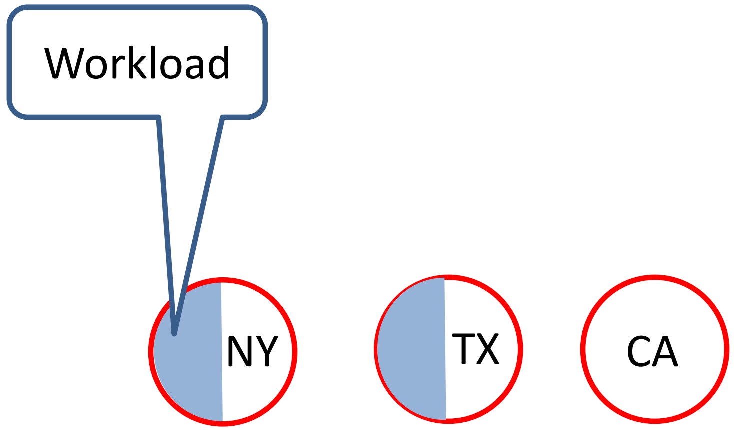
Key Idea

CA: California
NY: New York
TX: Texas



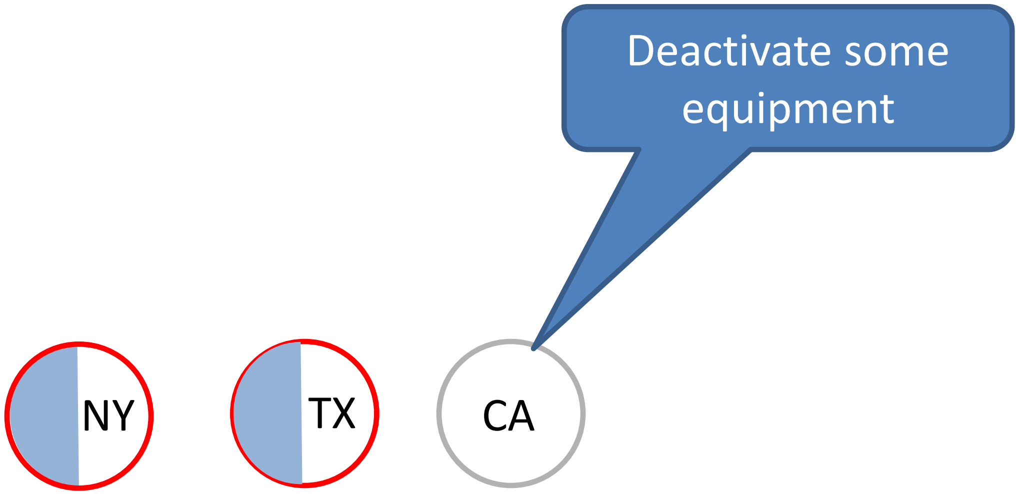
Key Idea

CA: California
NY: New York
TX: Texas



Key Idea

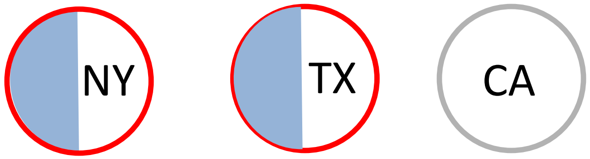
CA: California
NY: New York
TX: Texas



Key Idea

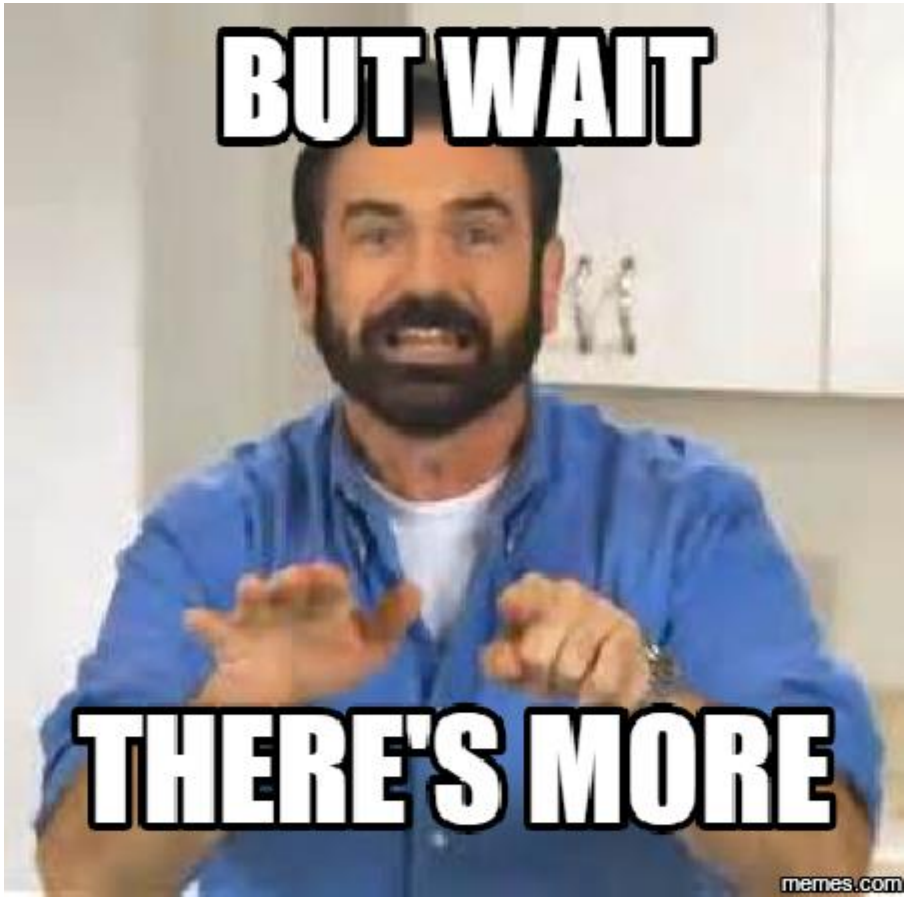
CA: California
NY: New York
TX: Texas

Resource pruning cuts electricity cost

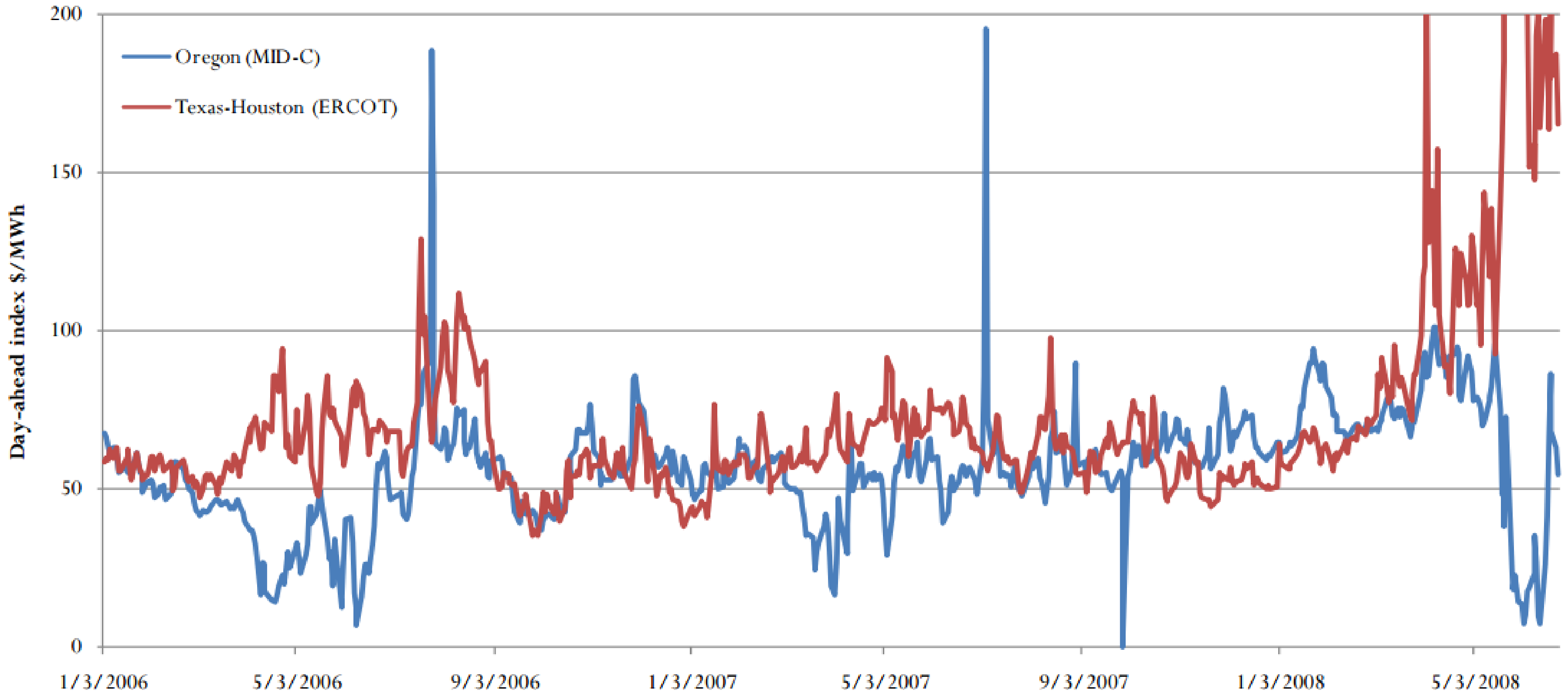


Key Idea

CA: California
NY: New York
TX: Texas

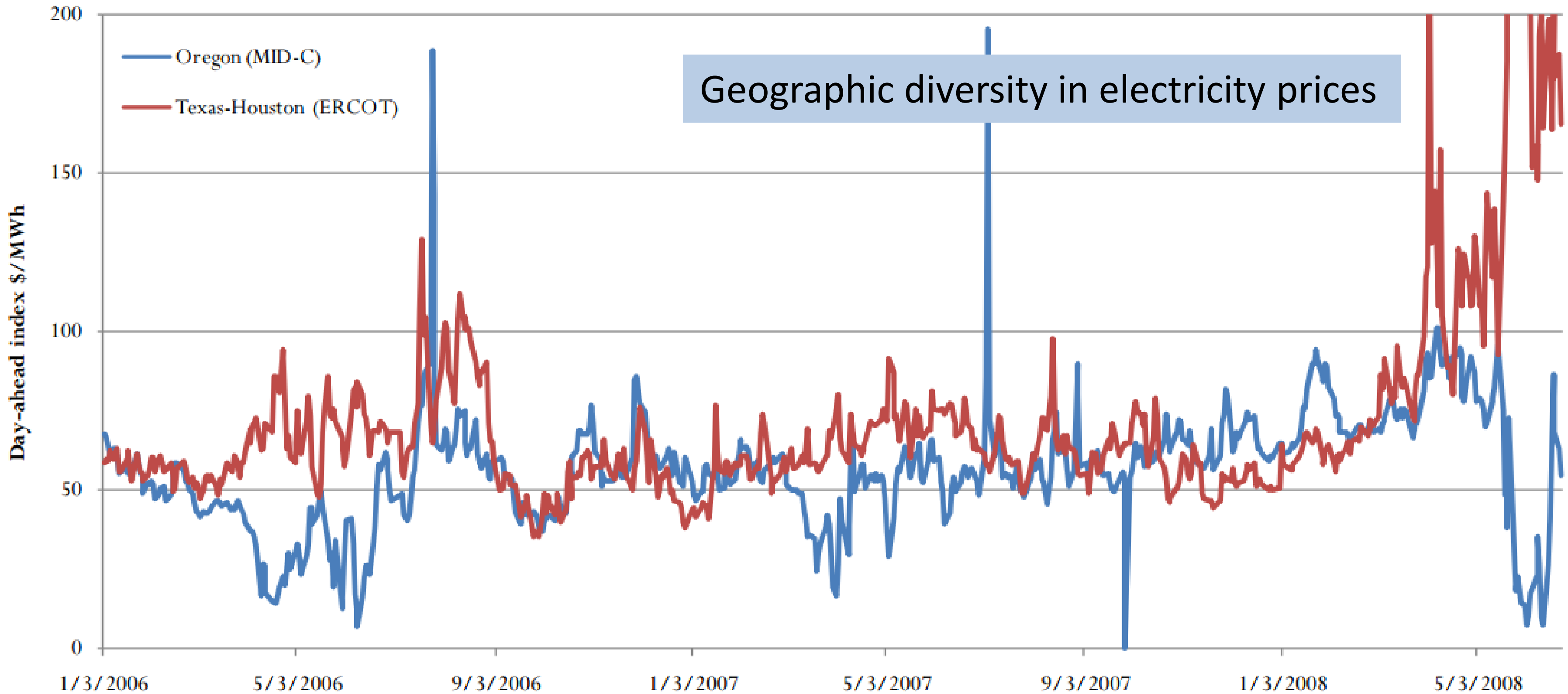


Key Idea

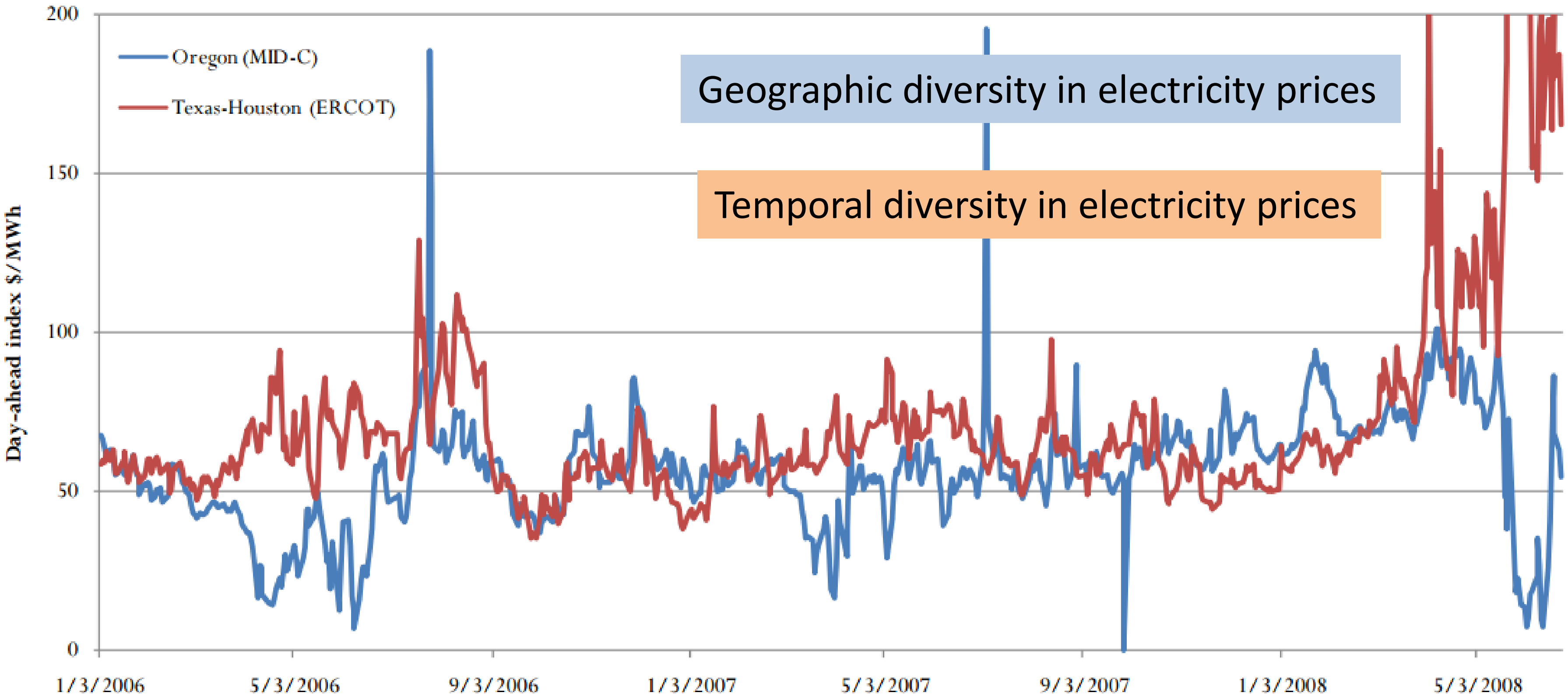


A. Qureshi et. al, “Plugging into energy market diversity”, HOTNETS 2008

Key Idea


200
150
100
50
0
Day-ahead index $/MWh
Oregon (MID-C)
Texas-Houston (ERCOT)
Geographic diversity in electricity prices
1/3/2006
5/3/2006
9/3/2006
1/3/2007
5/3/2007
9/3/2007
1/3/2008
5/3/2008

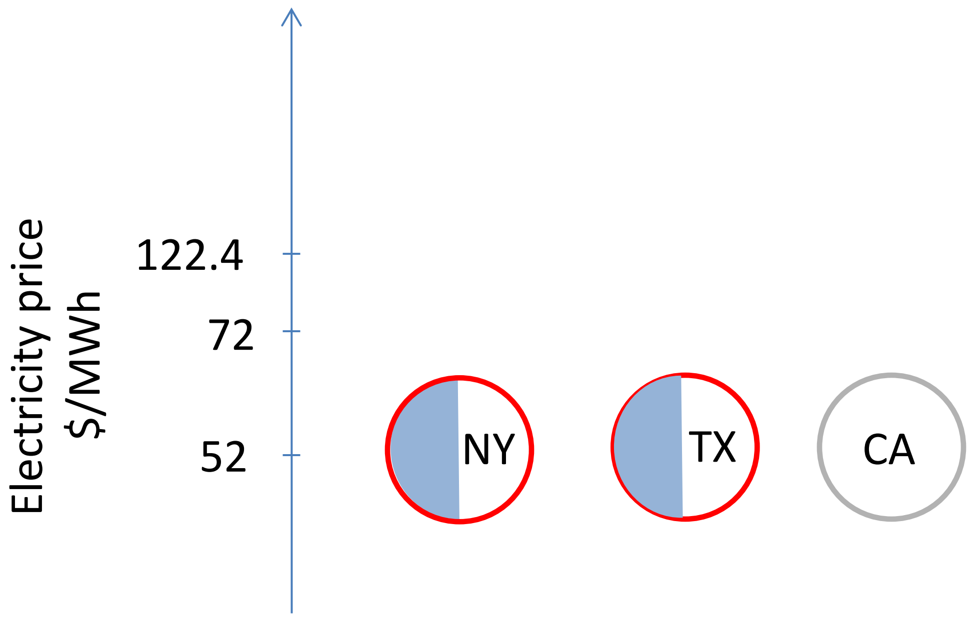
Key Idea



Oregon (MID-C)
Texas-Houston (ERCOT)
Geographic diversity in electricity prices
Temporal diversity in electricity prices
Day-ahead index $/MWh
200
150
100
50
0
1/3/2006
5/3/2006
9/3/2006
1/3/2007
5/3/2007
9/3/2007
1/3/2008
5/3/2008

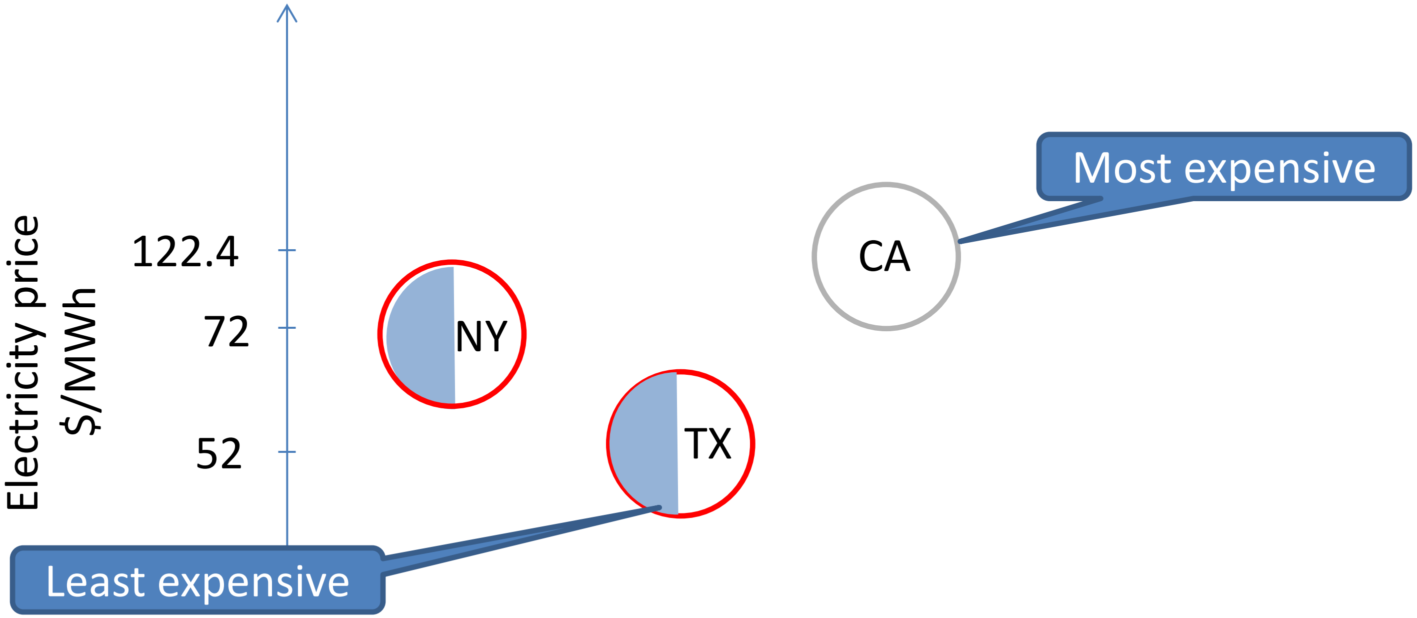
Key Idea

CA: California
NY: New York
TX: Texas



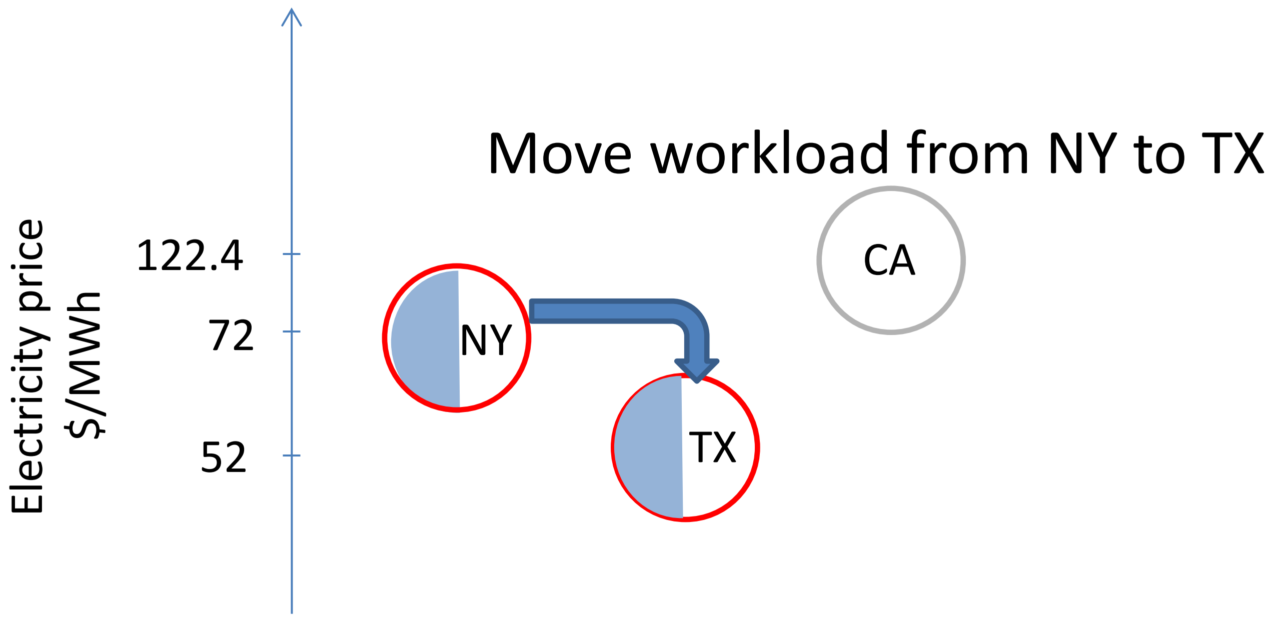
Key Idea

CA: California
NY: New York
TX: Texas



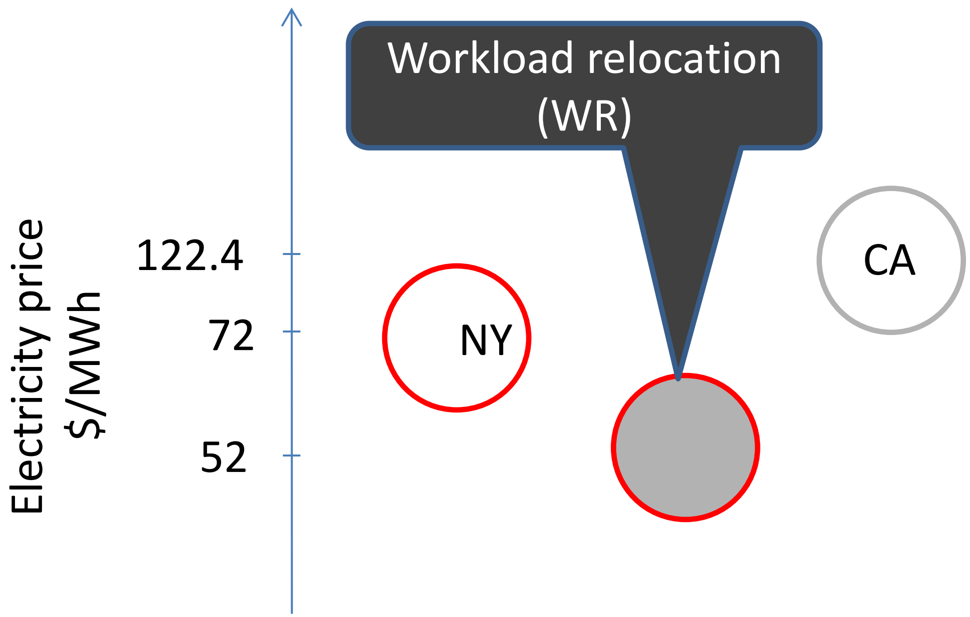
Key Idea

CA: California
NY: New York
TX: Texas



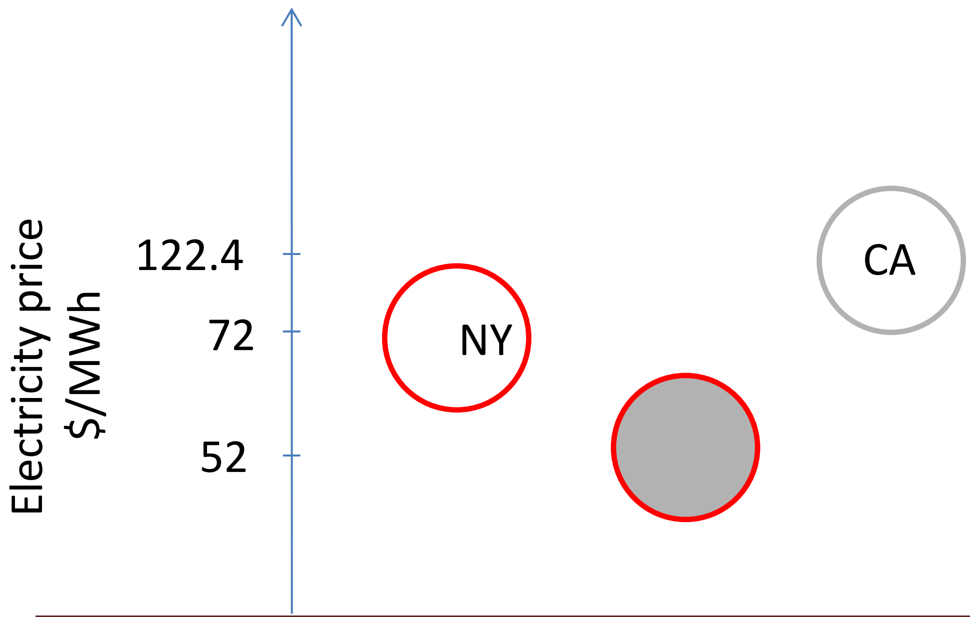
Key Idea

CA: California
NY: New York
TX: Texas



Key Idea

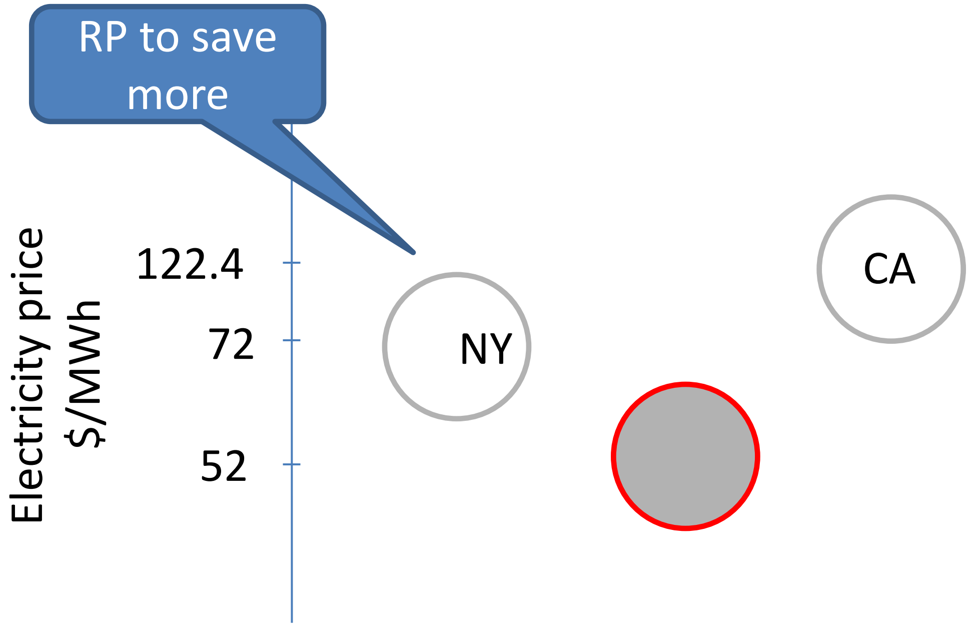
CA: California
NY: New York
TX: Texas



Workload relocation cuts electricity cost ***further***

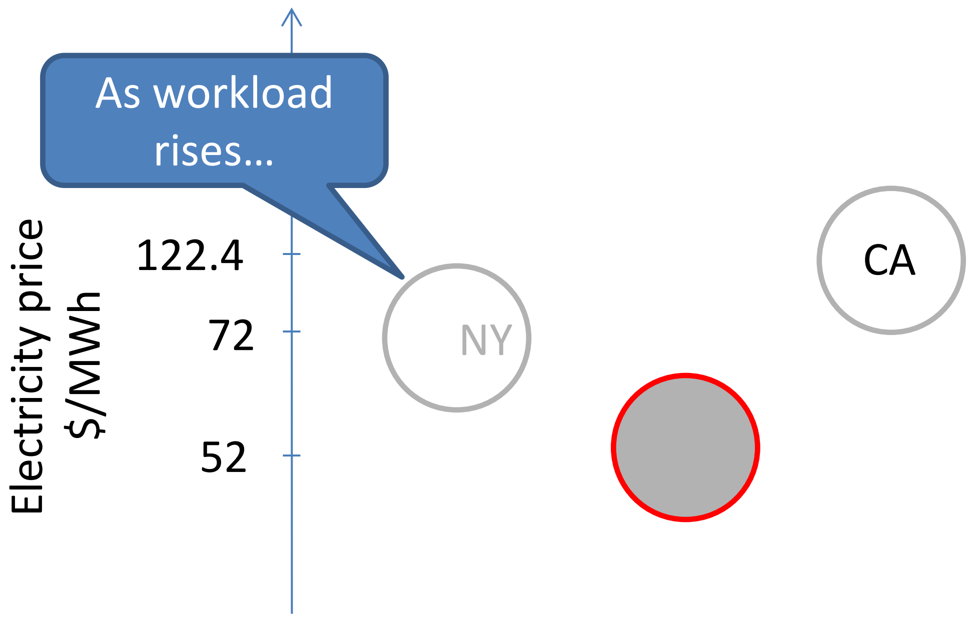
Key Idea

CA: California
NY: New York
TX: Texas



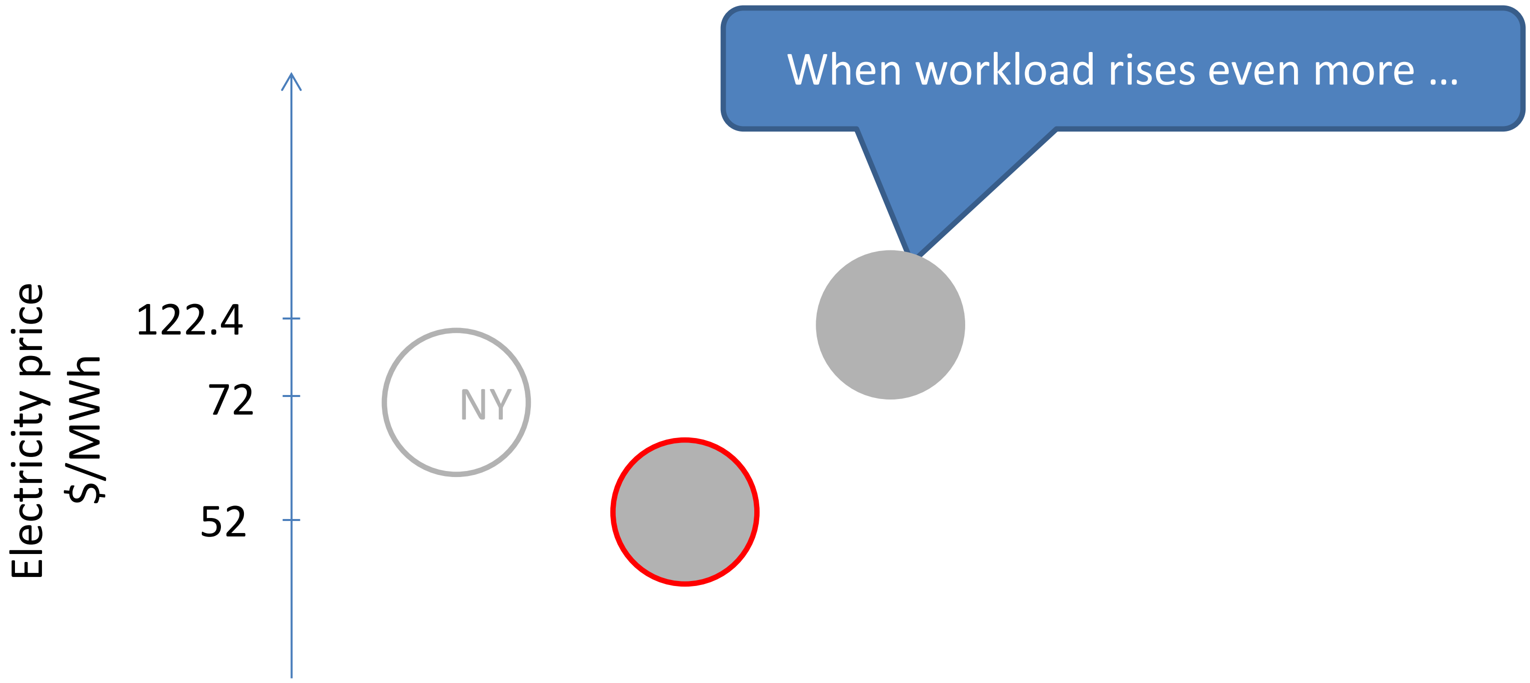
Key Idea

CA: California
NY: New York
TX: Texas



Key Idea

CA: California
NY: New York
TX: Texas



Key Idea

RP and WR can cut electricity costs

Key Idea

RP and WR can cut electricity costs

Ain’t no such thing as a free lunch

Transition Costs

- Workload relocation overheads
  - E.g., Cost of data transfers
    - Expensive inter-data center links

Transition Costs

- Workload relocation overheads
  - E.g., Cost of data transfers
    - Expensive inter-data center links
- Equipment activation and de-activation overheads
  - E.g., Energy spent while resuming and sleeping

Transition Costs

- Workload relocation overheads
  - E.g., Cost of data transfers
    - Expensive inter-data center links
- Equipment activation and de-activation overheads
  - E.g., Energy spent while resuming and sleeping
- Must consider transition costs while optimizing

Transition Costs

- Workload relocation overheads
  - E.g., Cost of data transfers
    - Expensive inter-data center links
- Equipment activation and de-activation overheads
  - E.g., Energy spent while resuming and sleeping
- Must consider transition costs while optimizing
- Relocate Energy Demand to Better Locations (RED-BL)

This Thesis

Towards systematic minimization of network electricity cost

using Workload Relocation (WR) and Resource Pruning (RP)

while considering transition costs

Contributions

- Optimal state trajectory formulation [1]
- Considered transition costs [1]
- Evaluation of cost savings under various scenarios using real traces [1]
- Evaluated impact of prediction accuracy on savings [1]

[1] "RED-BL: Energy Solution for Loading Data Centers", Infocom 2013 mini-conference

Contributions

- Evaluation of RED-BL with:
  - Partial data center shutdown [2]
  - Sleep modes [2]
- Sliding window re-optimization [2]
- Granular deactivation of data center equipment [2]
- Showed that RED-BL (for data centers) is the NP-Complete unit commit problem [2]

[2] "RED-BL: Evaluating Dynamic Right Sizing for Data Center Networks", Elsevier Computer Networks, 2014

Contributions

- Application of state trajectory optimization to cellular networks [2, 3]
- Showed that RED-BL for cellular networks is NP-Hard [4]
- Evaluated RED-BL using traces from a live network [3, 4]

[3] "Electricity Cost Efficient Workload Mapping", IEEE INFOCOM 2013 Computer Communications Workshop

[4] "Low-Carb: Reducing Energy Consumption in Operational Cellular Networks" IEEE GLOBECOM 2013

Agenda

- Background and motivation
- Opportunity and key idea
- Case studies:
  - **Data centers (e.g., Facebook and Google)**
  - Cellular networks (e.g., Sprint and Verizon)
- Conclusions and future work

Case Study – I : Background

Source: http://bit.ly/1mrIi7o

Case Study – I : Background

- Data center operator
  - Geographically distributed data centers

Case Study – I : Background

- Data center operator
  - Geographically distributed data centers
- Data center equipment

IT Load	Non-IT Load
Servers	Lighting
Storage	Cooling
Network	Power distribution

Case Study – I : Background

- Data center operator
  - Geographically distributed data centers
- Data center equipment

IT Load	Non-IT Load
Servers	Lighting
Storage	Cooling
Network	Power distribution

- Power drawn is affine function of workload

Case Study – I : Background

- Data center operator
  - Geographically distributed data centers
- Data center equipment

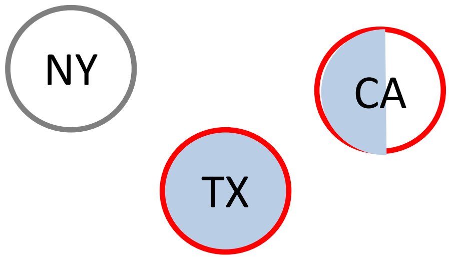
IT Load	Non-IT Load
Servers	Lighting
Storage	Cooling
Network	Power distribution

- Power drawn is affine function of workload

Let’s recap how we can use WR and RP

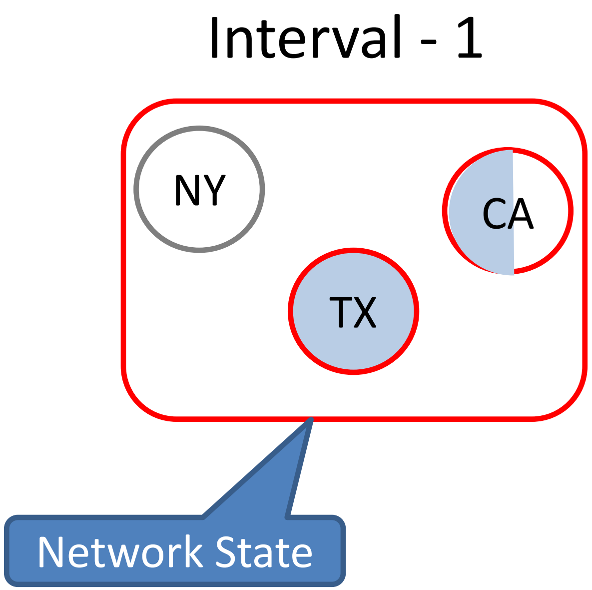
Problem Model

Interval - 1

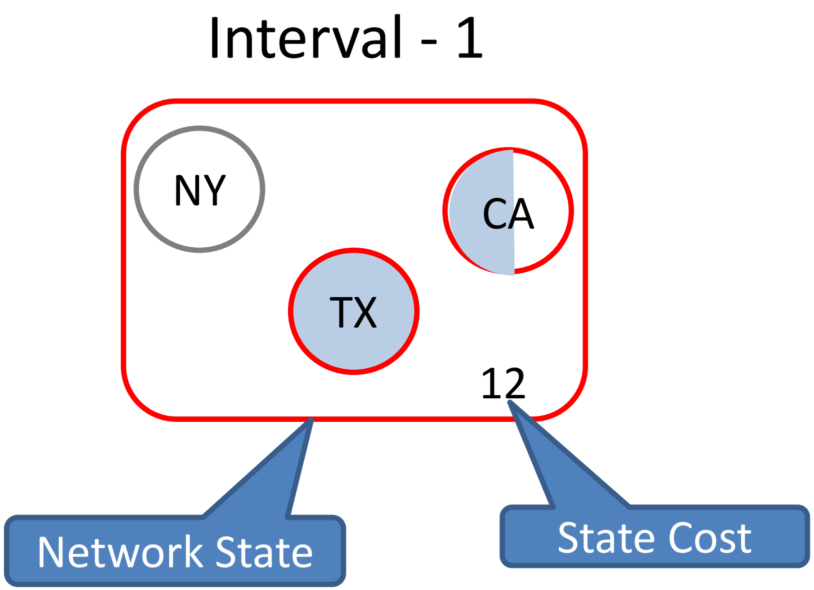


Electricity price driven workload assignment

Problem Model



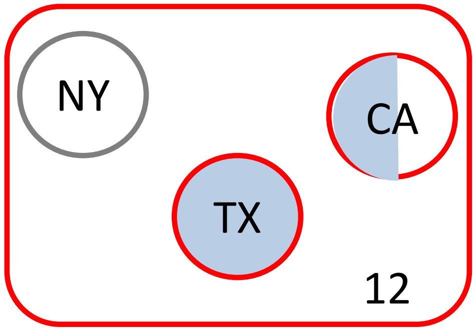
Problem Model



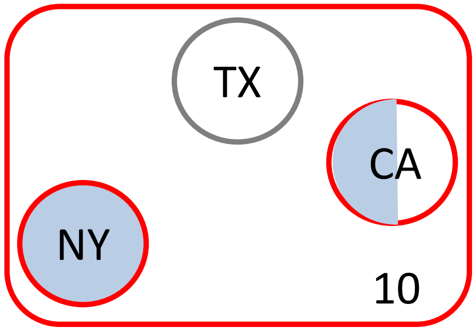
Sum of all data centers' electricity cost

Problem Model

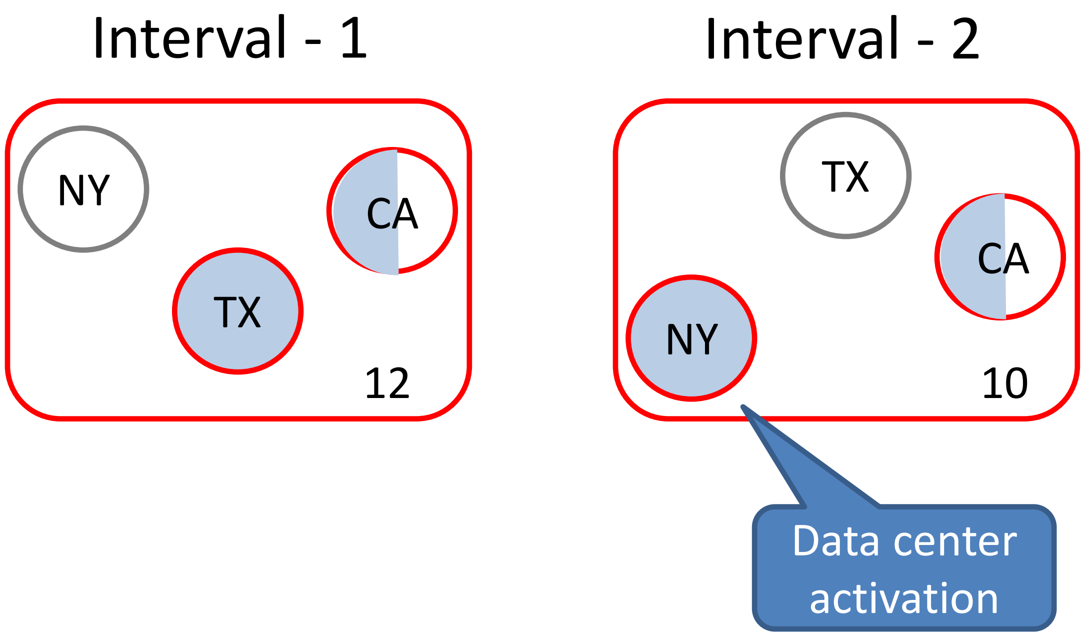
Interval - 1



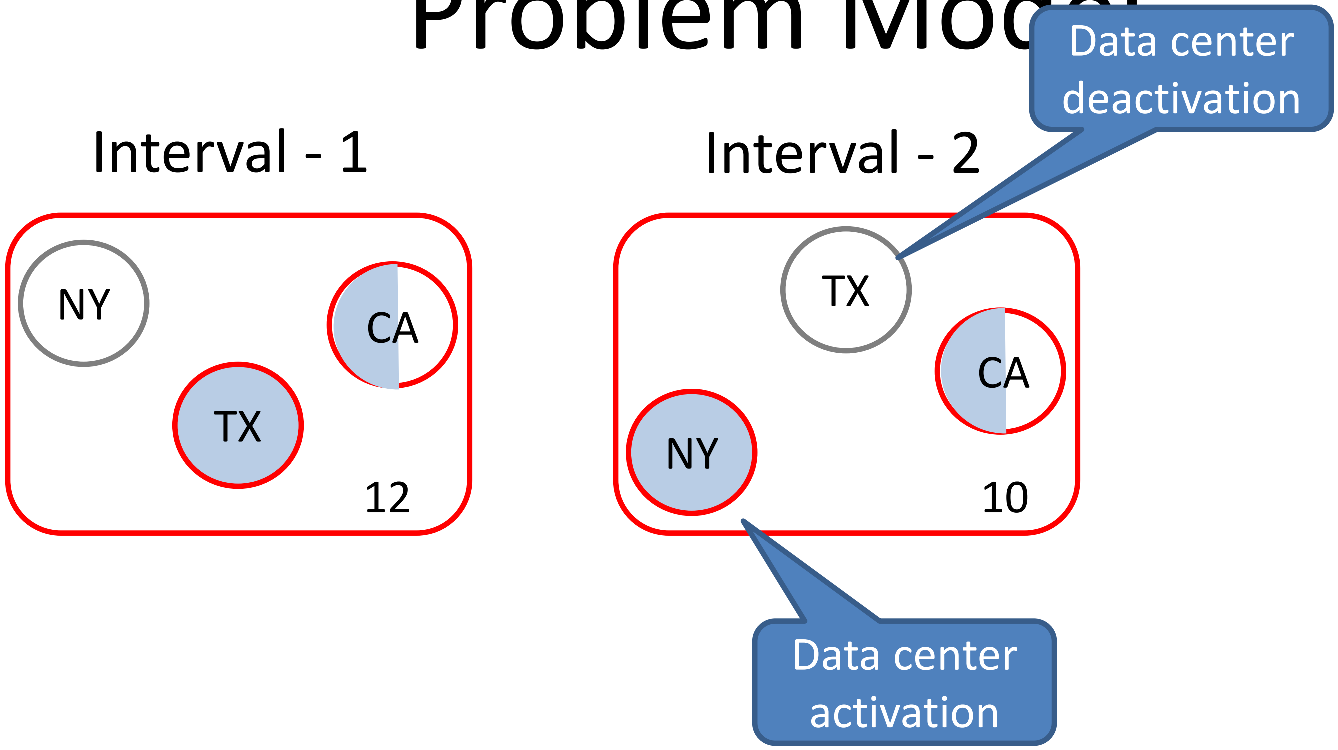
Interval - 2



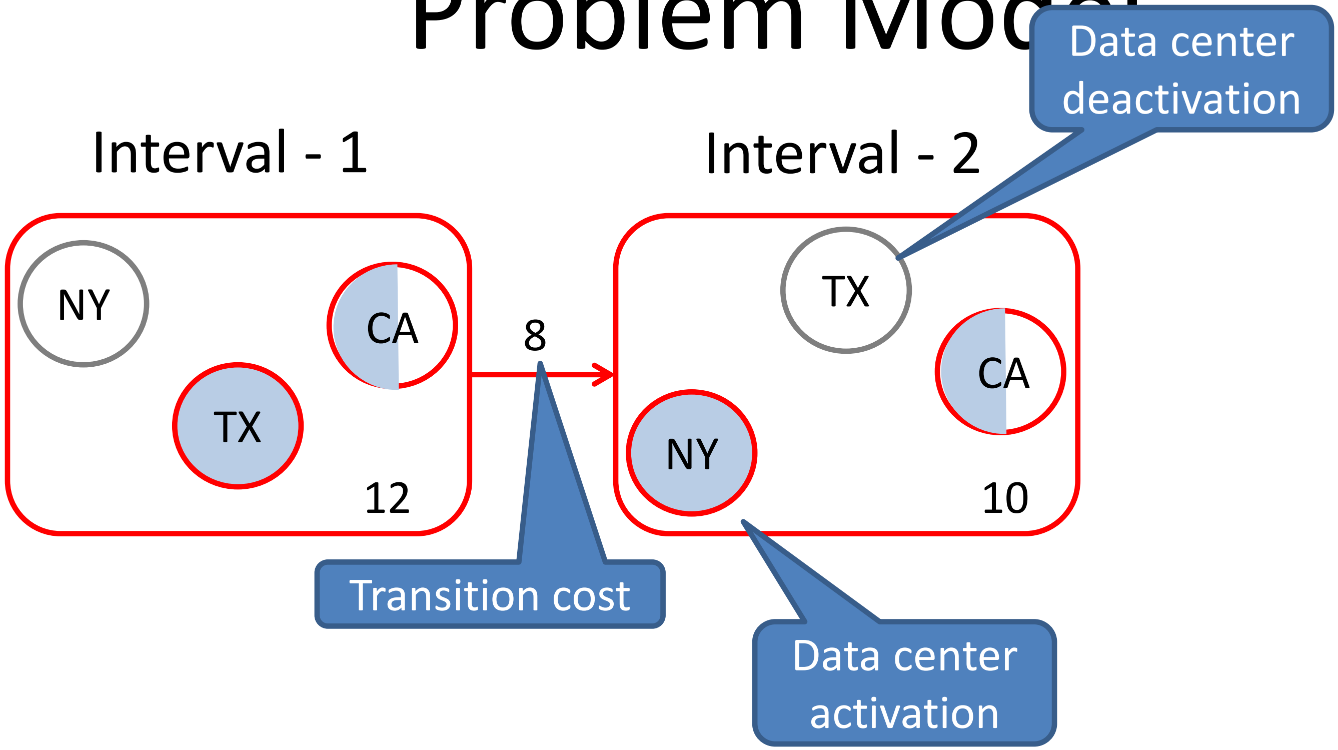
Problem Model



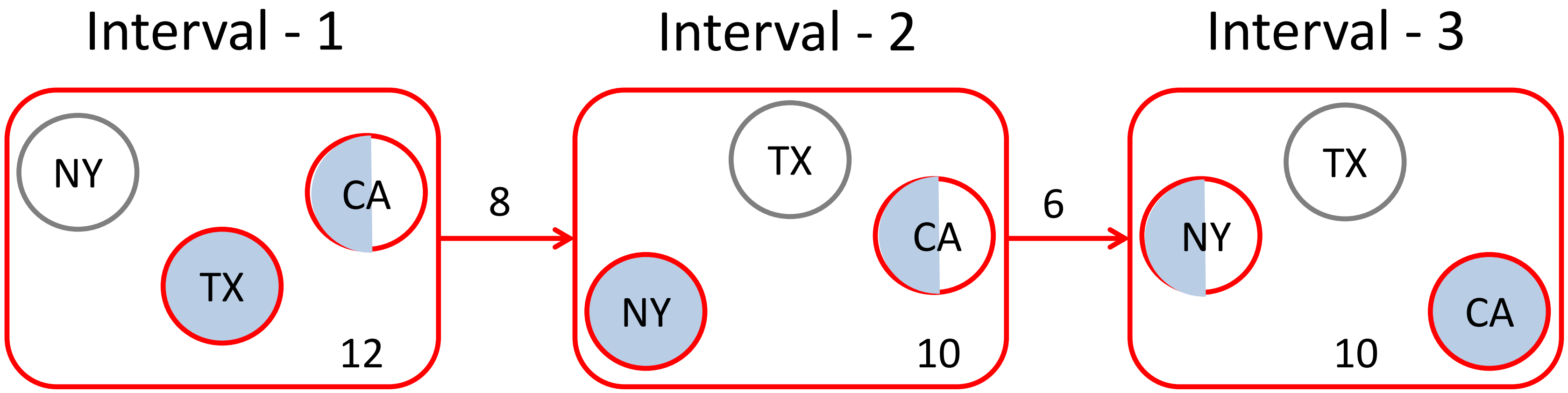
Problem Model



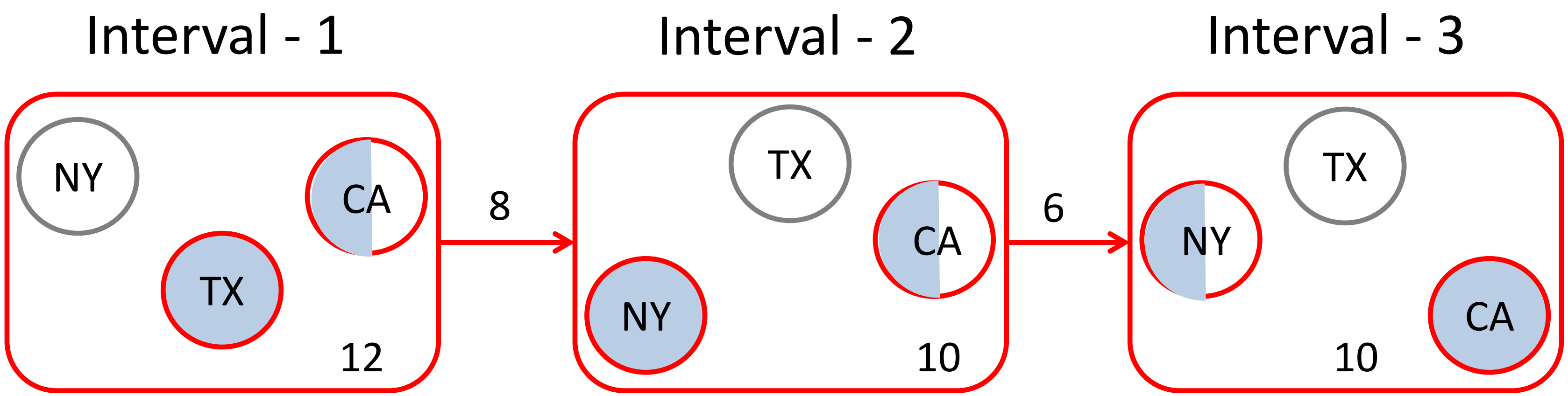
Problem Model



Problem Model

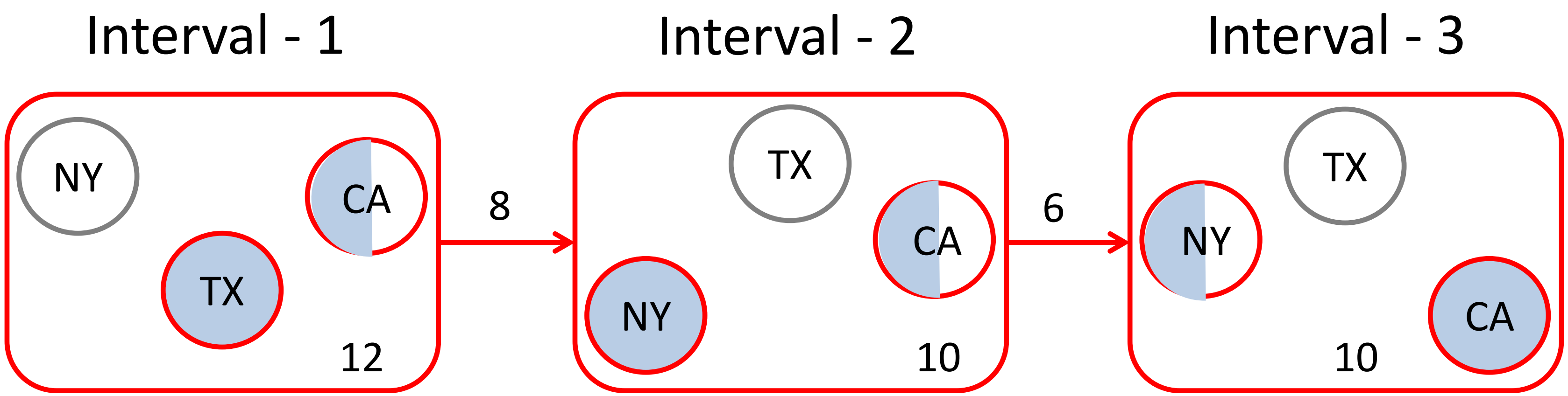


Problem Model



Locally optimal

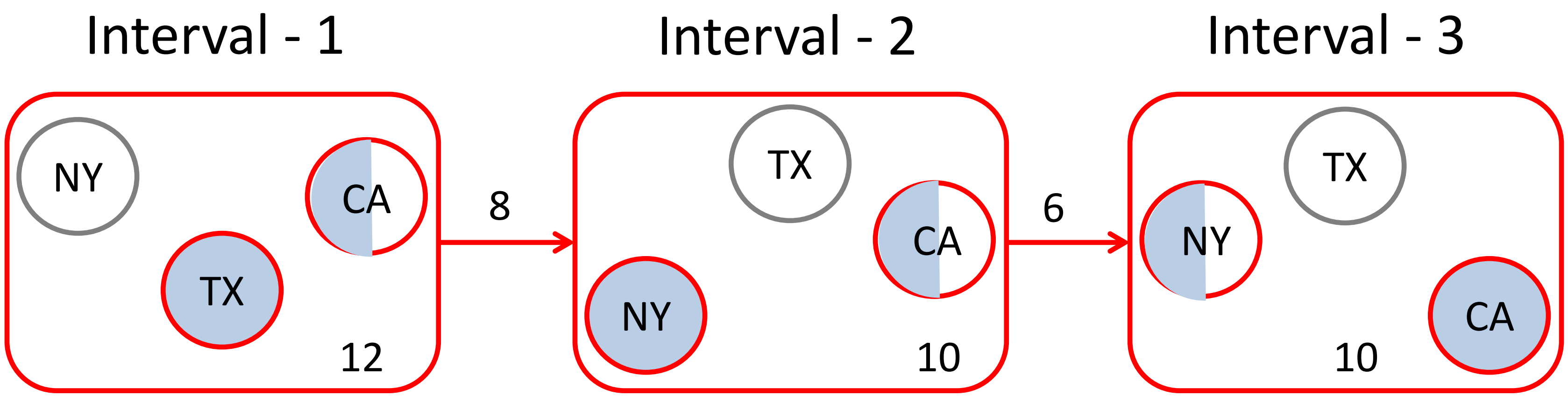
Problem Model



Locally optimal

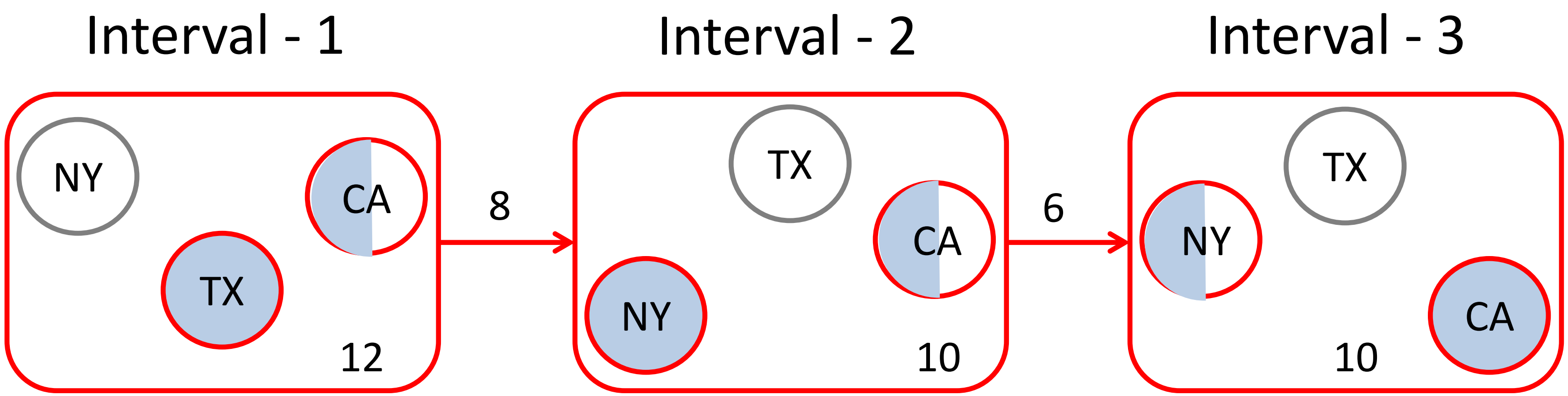
Might not be globally optimal

Problem Model

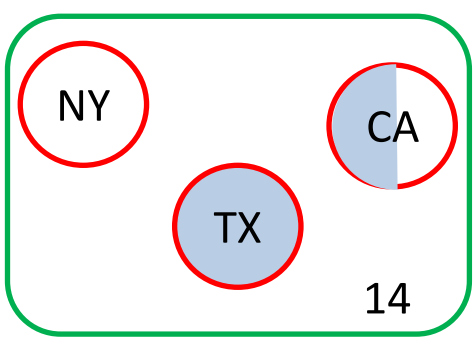


An alternative workload mapping

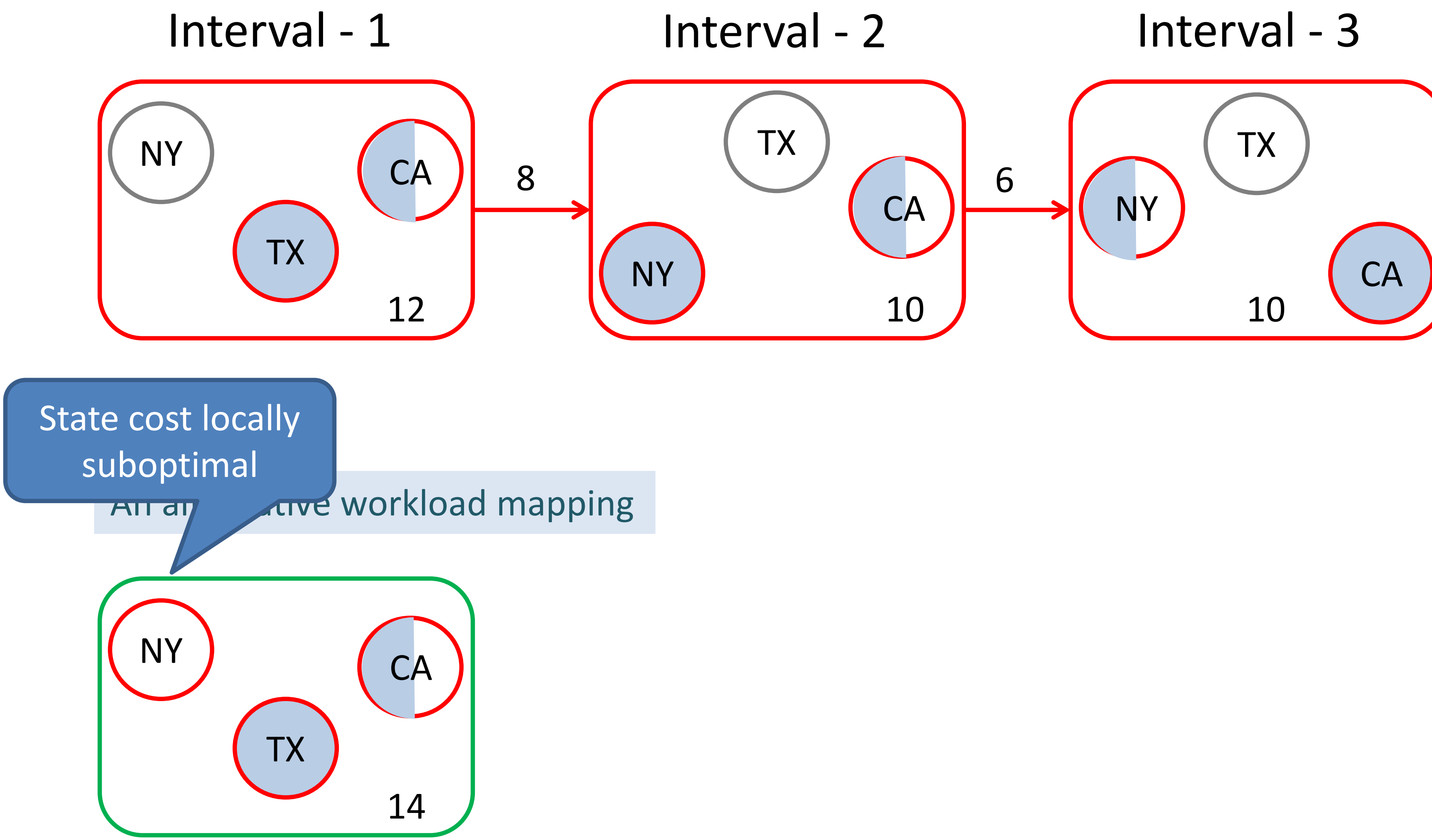
Problem Model



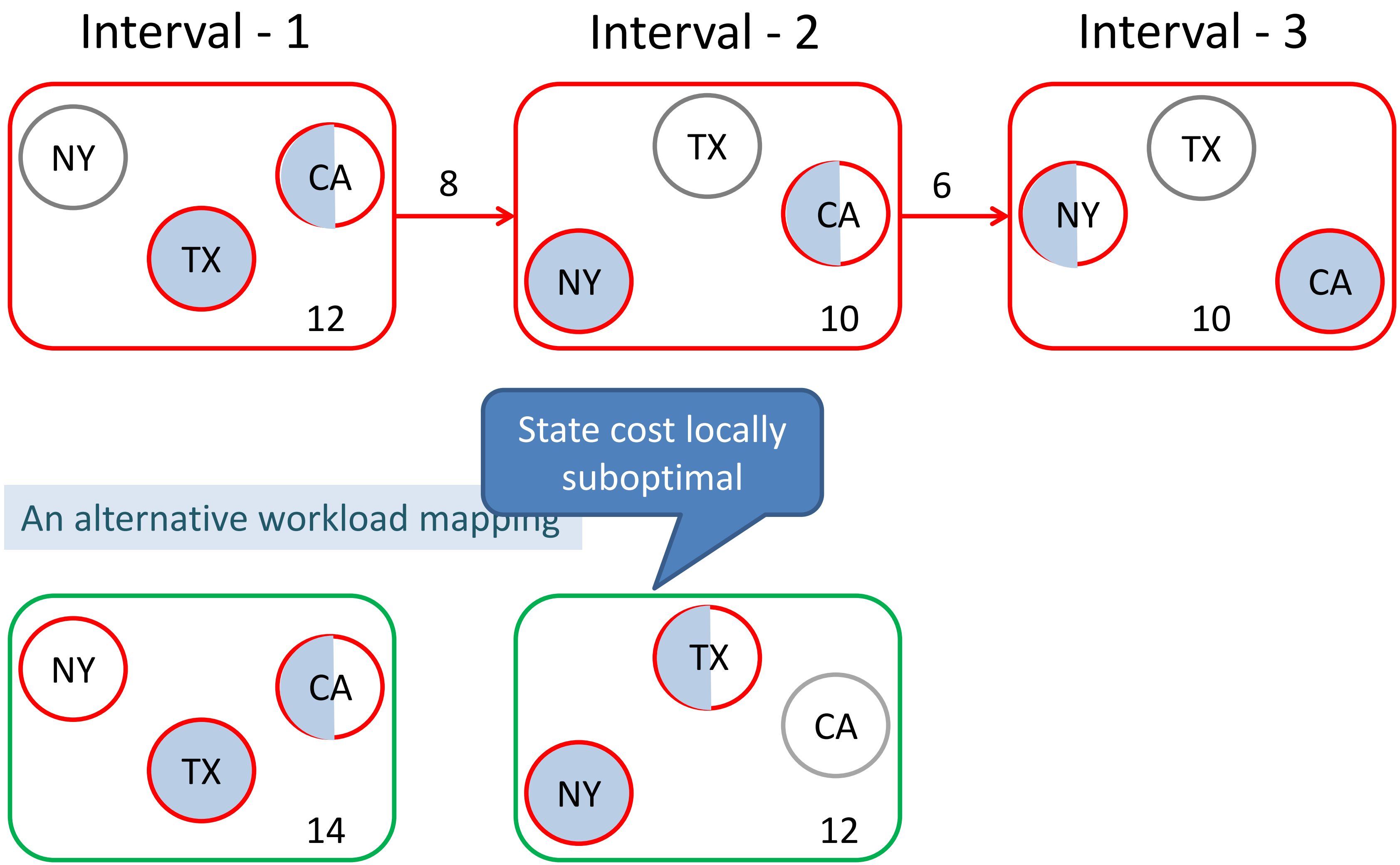
An alternative workload mapping



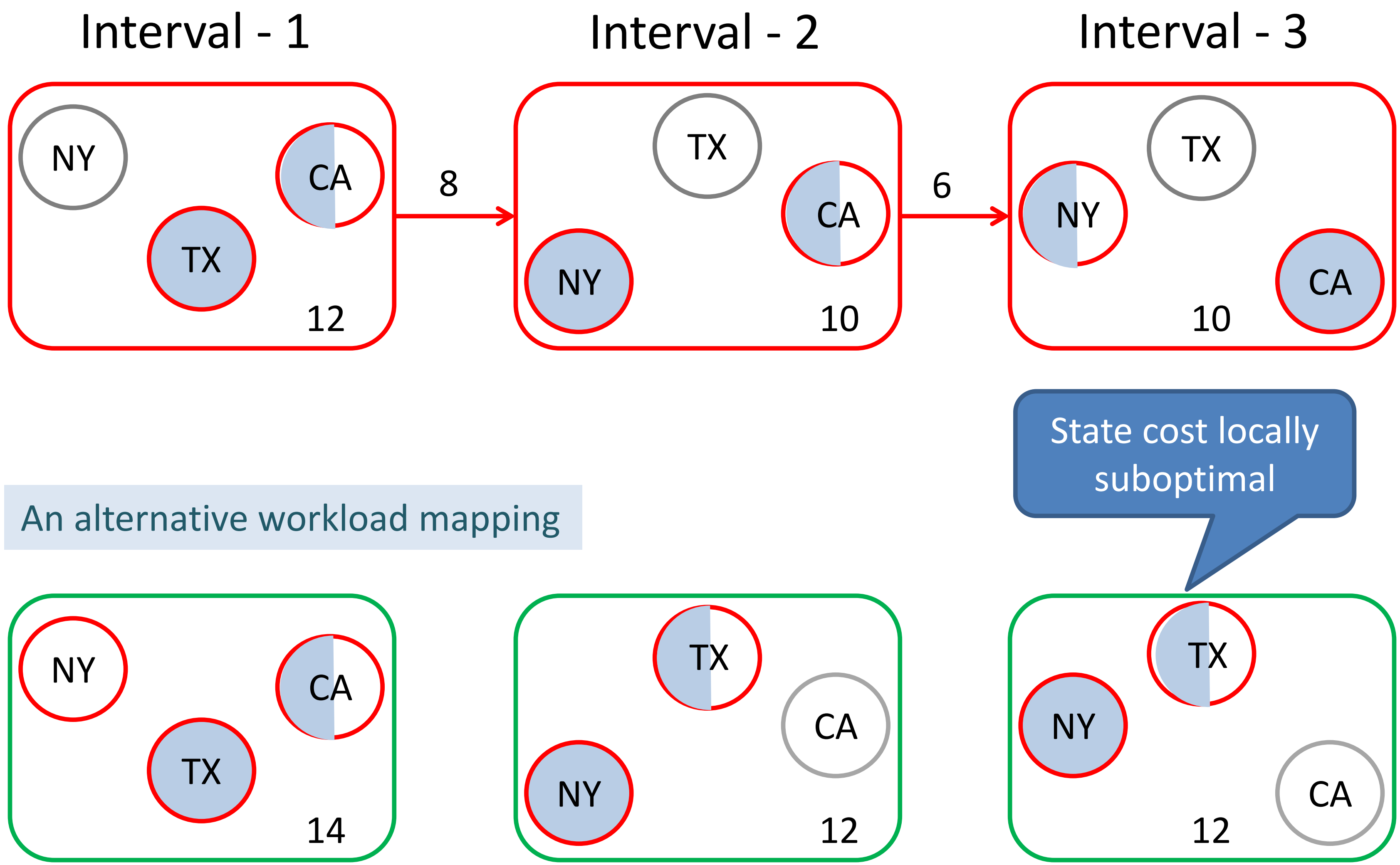
Problem Model



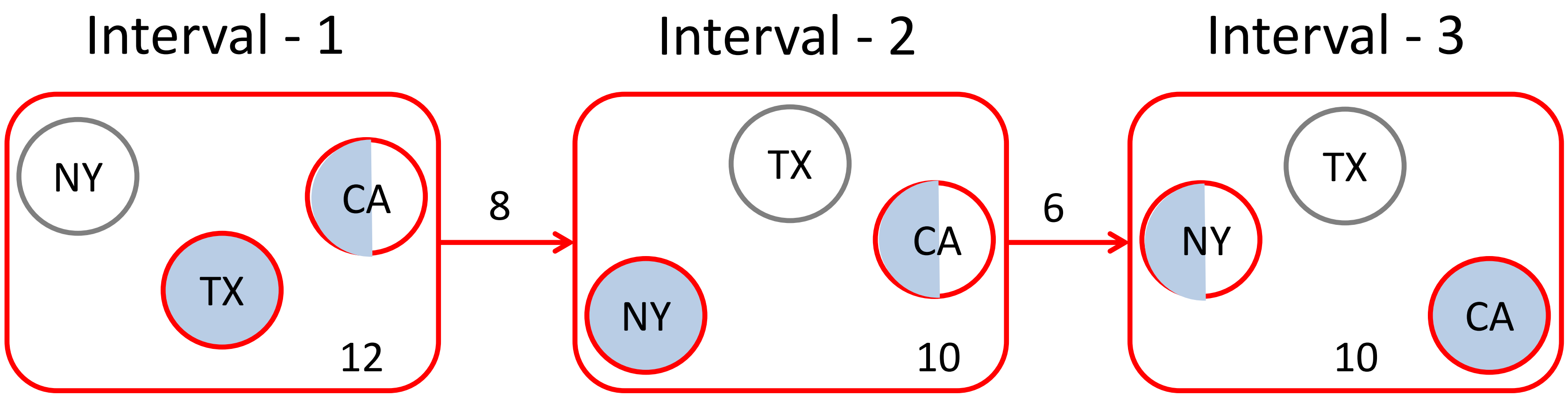
Problem Model



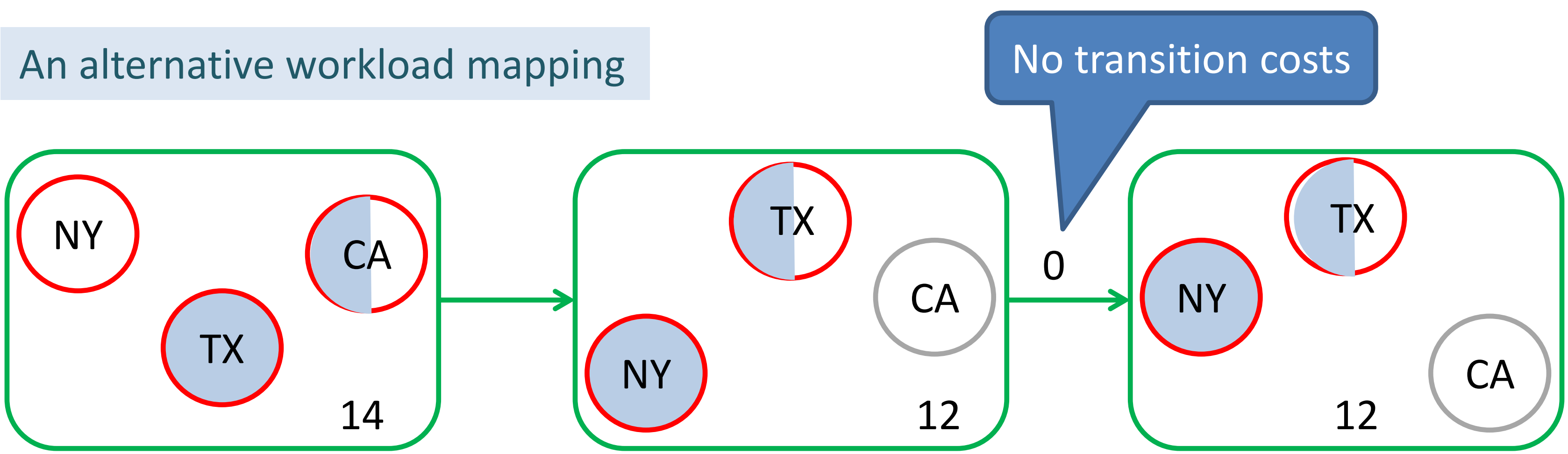
Problem Model



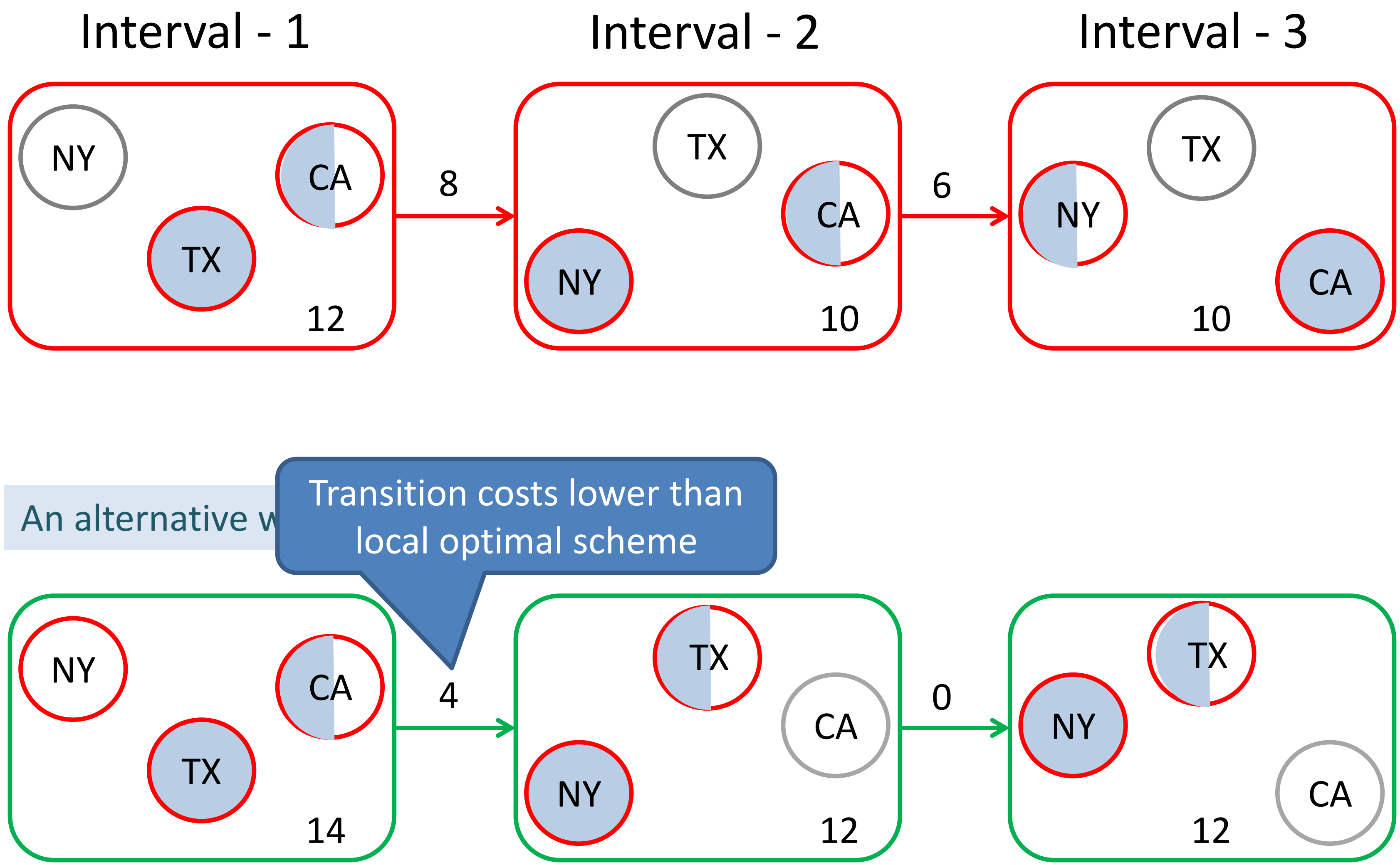
Problem Model



An alternative workload mapping

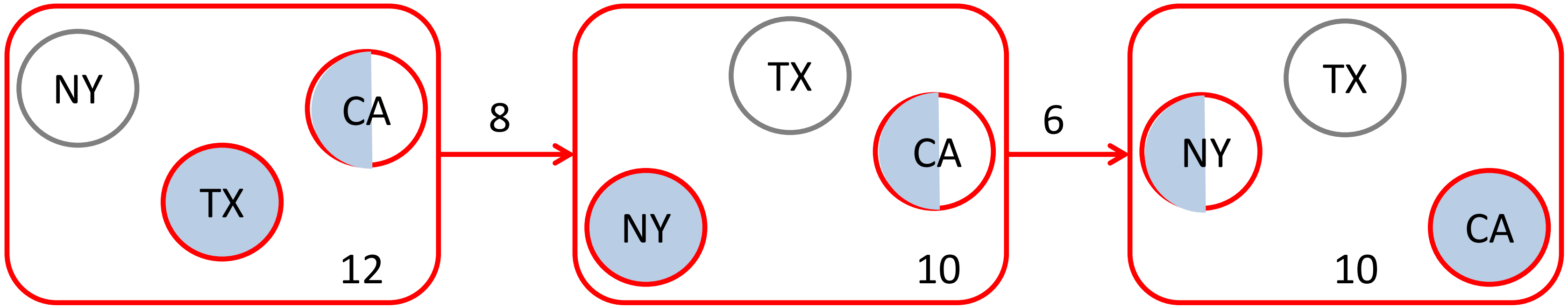


Problem Model



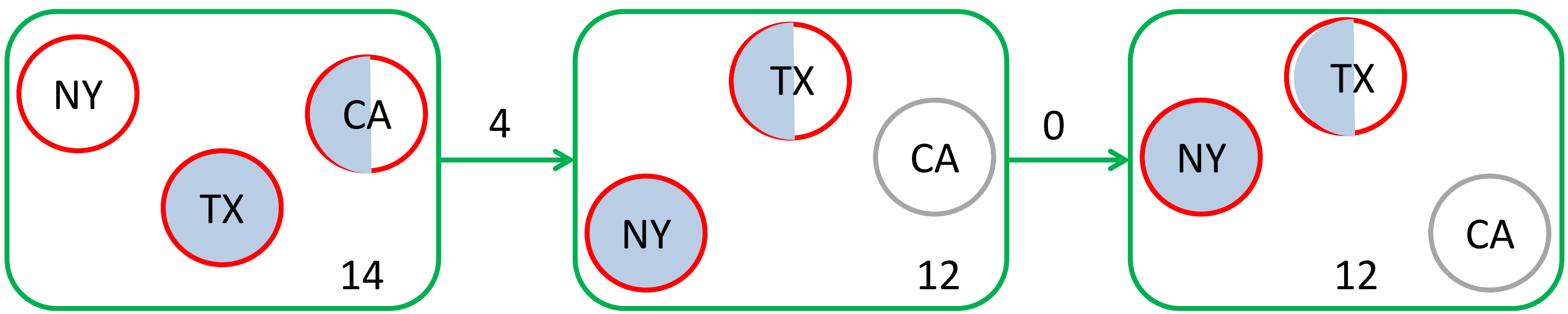
Problem Model

Interval - 1 Interval - 2 Interval - 3

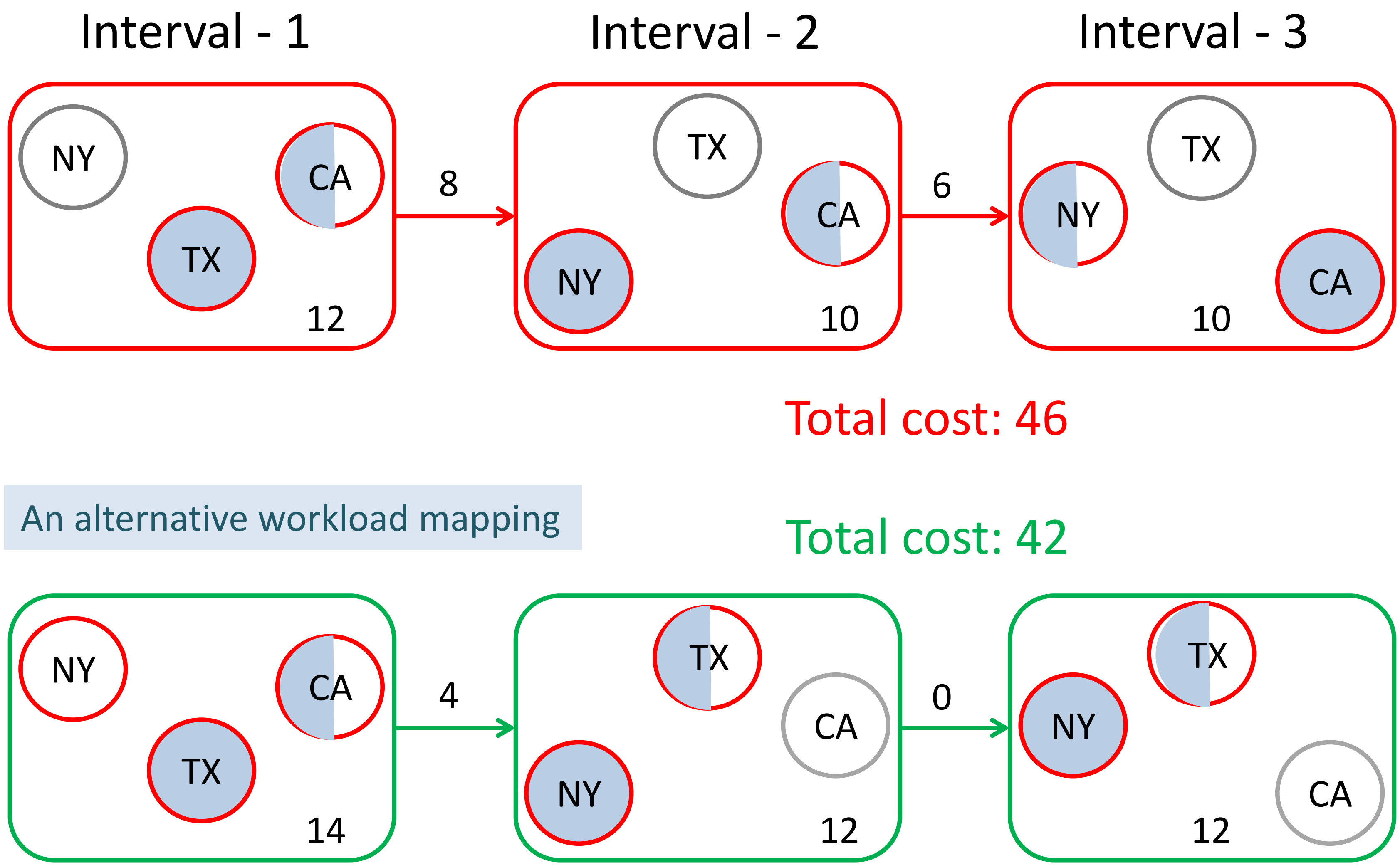


An alternative workload mapping

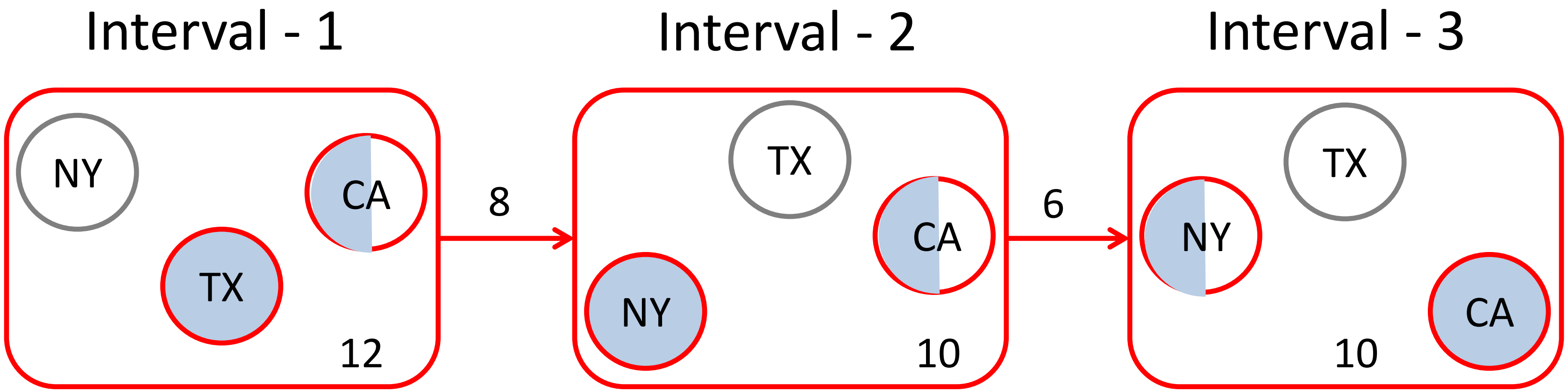
Total cost: 42



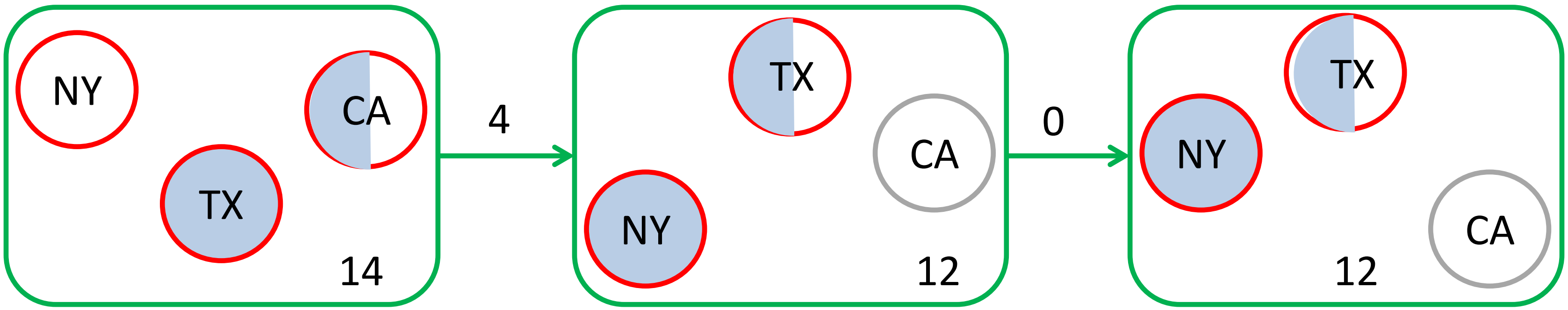
Problem Model



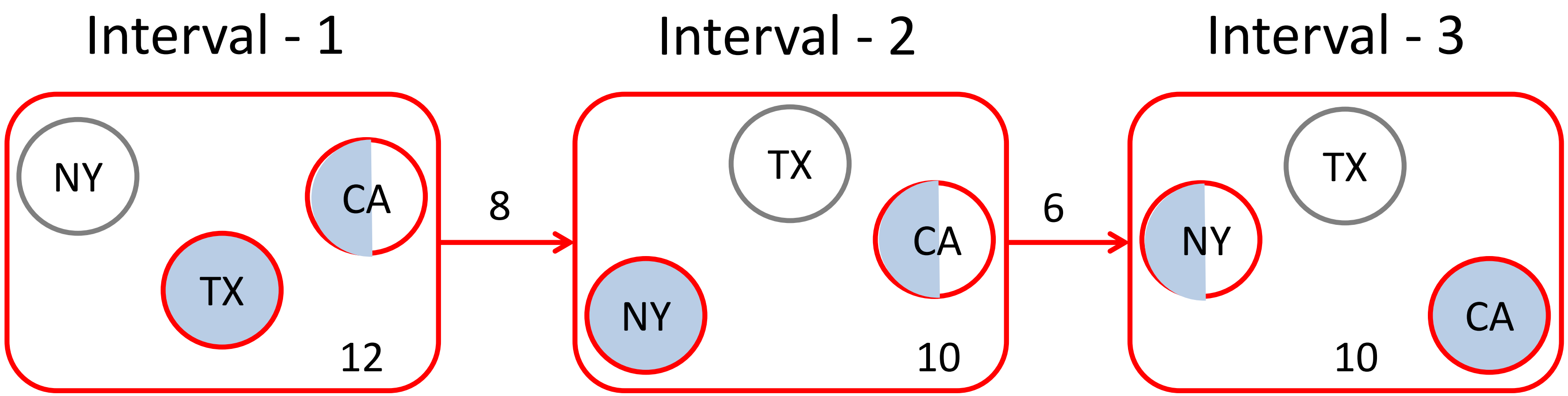
Problem Model



Optimal State Trajectory Problem

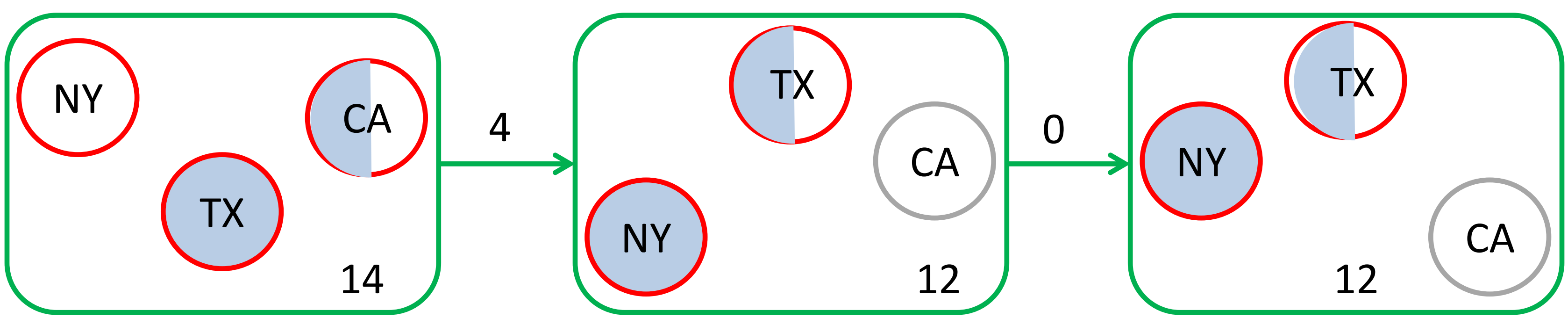


Problem Model



Optimal State Trajectory Problem

Relocate Energy Demand to ***Better*** Locations (RED-BL)



Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j \left(p_i^j \lambda \left(f + (1-f) \frac{x_i^j}{c_i} \right) + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}} \right)$$

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n}\sum_{i=1}^{m} c_i e_i^j \Big(\underbrace{p_i^j \lambda\big(f + (1-f)\frac{x_i^j}{c_i}\big)}_{\text{State energy}} + \underbrace{b_i^j\sigma + s_i^j\delta}_{\text{Transition energy}}\Big)$$

Optimization Formulation

Unit price of electricity

$$\text{minimize} \sum_{j=1}^{n}\sum_{i=1}^{m} c_i e_i^j \underbrace{\left(p_i^j \lambda \left(f + (1-f)\frac{x_i^j}{c_i}\right)\right.}_{\text{State energy}} + \underbrace{\left.b_i^j \sigma + s_i^j \delta\right)}_{\text{Transition energy}}$$

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n}\sum_{i=1}^{m} c_i e_i^j \Big(\underbrace{p_i^j \lambda \Big(f + (1-f)\frac{\overset{\text{Workload}}{x_i^j}}{c_i}\Big)}_{\text{State energy}} + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}}\Big)$$

Optimization Formulation

Workload

Data center capacity

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j \underbrace{(p_i^j \lambda (f + (1-f)\frac{x_i^j}{c_i})}_{\text{State energy}} + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}})$$

Optimization Formulation

Fraction of data center that is active

Workload

Data center capacity

$$\text{minimize} \sum_{j=1}^{n}\sum_{i=1}^{m} c_i e_i^j \Big(\underbrace{p_i^j \lambda \big(f + (1-f)\frac{x_i^j}{c_i}\big)}_{\text{State energy}} + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}}\Big)$$

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j \big(\underbrace{p_i^j \lambda \left(f + (1-f)\frac{x_i^j}{c_i}\right)}_{\text{State energy}} + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}} \big)$$

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j \underbrace{(p_i^j \lambda (f + (1-f)\frac{x_i^j}{c_i})}_{\text{State energy}} + \underbrace{b_i^j \sigma + s_i^j \delta}_{\text{Transition energy}})$$

Sum over all data centers

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j \underbrace{\left(p_i^j \lambda \left(f + (1-f)\frac{x_i^j}{c_i}\right)\right.}_{\text{State energy}} + \underbrace{\left. b_i^j \sigma + s_i^j \delta \right)}_{\text{Transition energy}}$$

Sum over all intervals

Sum over all data centers

Subject to several constraints (please see the thesis)

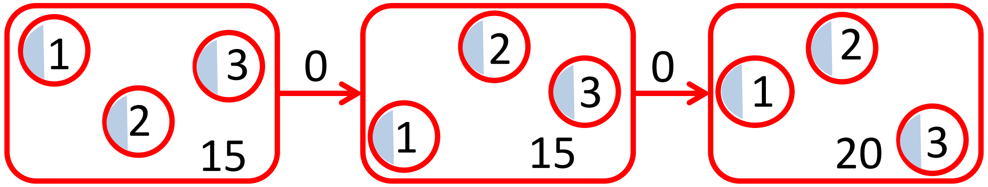
Experimental Setup

- Workload from 3 popular Facebook apps
- Electricity prices from 33 US locations
- Simulated a week-long deployment plan
- Compared RED-BL against various schemes

Comparison Benchmarks

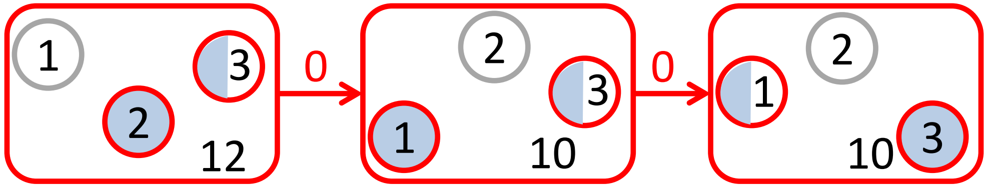
Comparison Benchmarks

UNIFORM: Equally distribute workload



Comparison Benchmarks

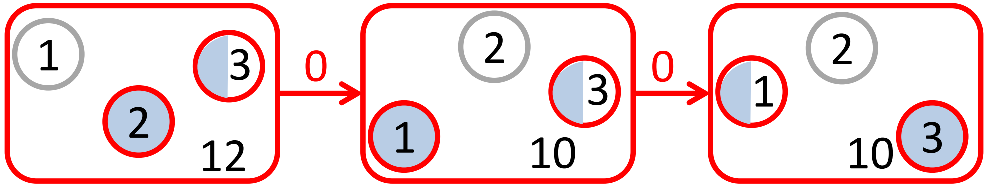
LO: Local Optimal Ignoring Transition Costs



Comparison Benchmarks

LO: Local Optimal Ignoring Transition Costs

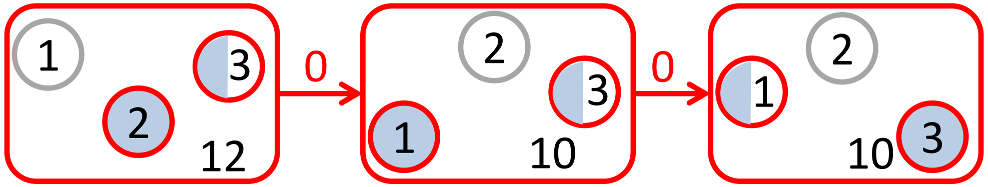
Theoretical lower bound



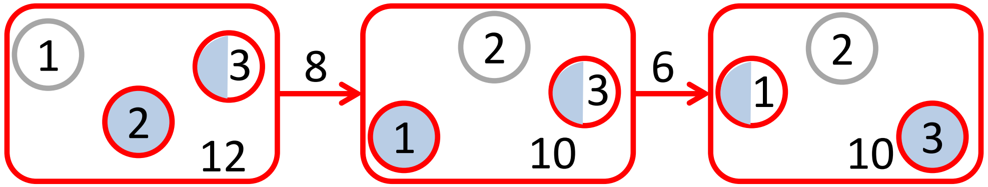
Comparison Benchmarks

LO: Local Optimal Ignoring Transition Costs

Theoretical lower bound



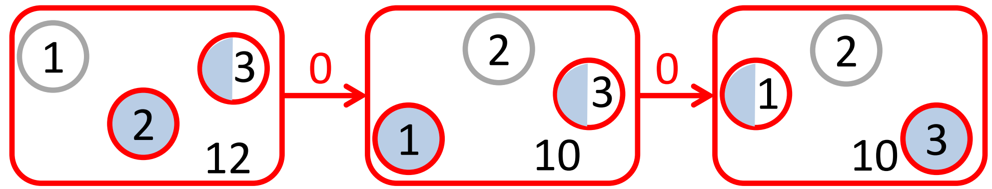
LD: Local Optimal with Deactivation



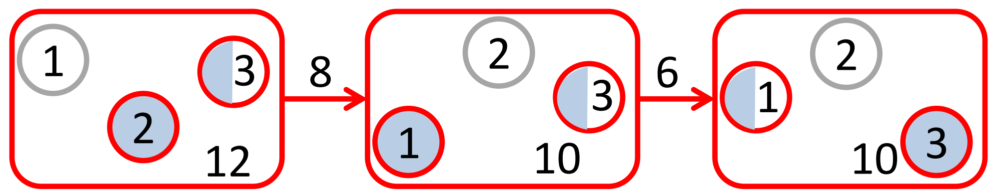
Comparison Benchmarks

LO: Local Optimal Ignoring Transition Costs

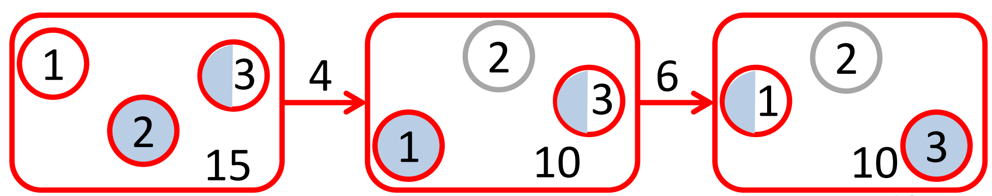
Theoretical lower bound



LD: Local Optimal with Deactivation



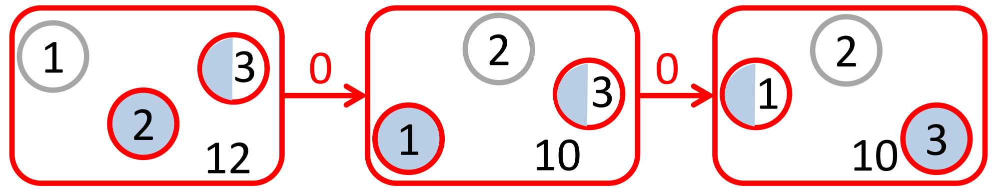
LS: Local Optimal with Selection



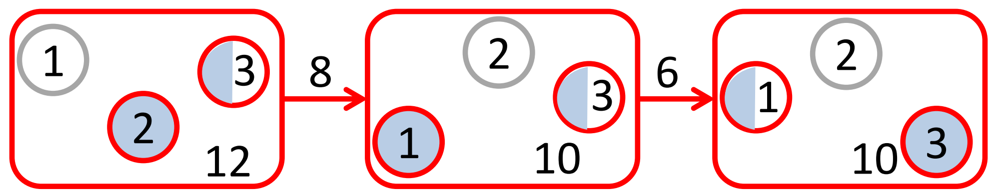
Comparison Benchmarks

LO: Local Optimal Ignoring Transition Costs

Theoretical lower bound

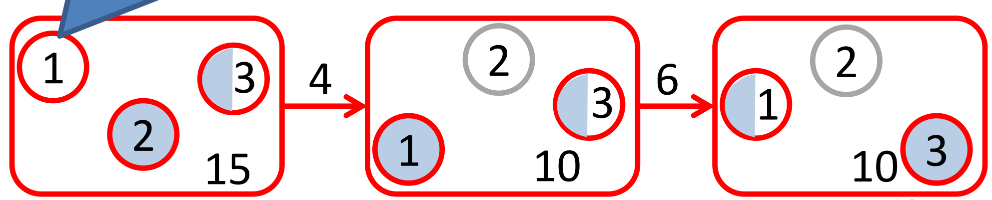


LD: Local Optimal with Deactivation



Sometimes idling is better

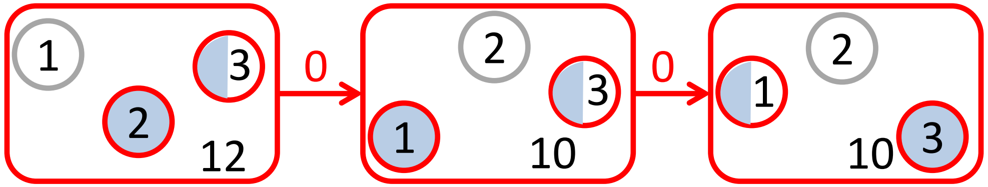
LS: Local Optimal with Selection



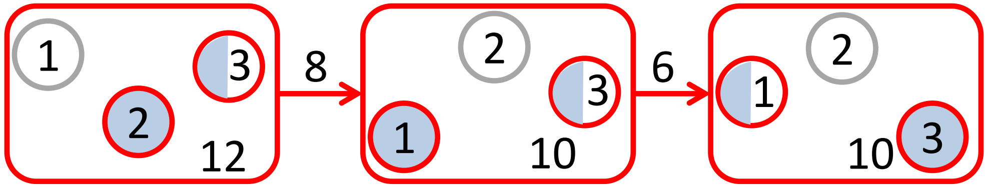
Comparison Benchmarks

LO: Local Optimal Ignoring Transition Costs

Theoretical lower bound



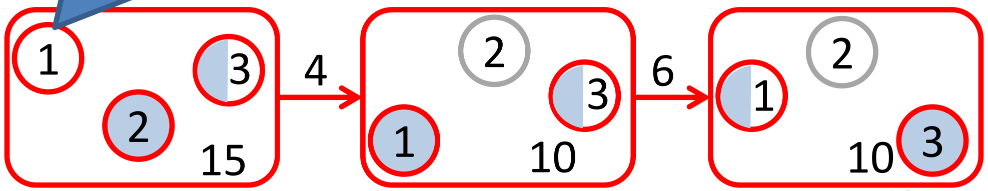
LD: Local Optimal with Deactivation



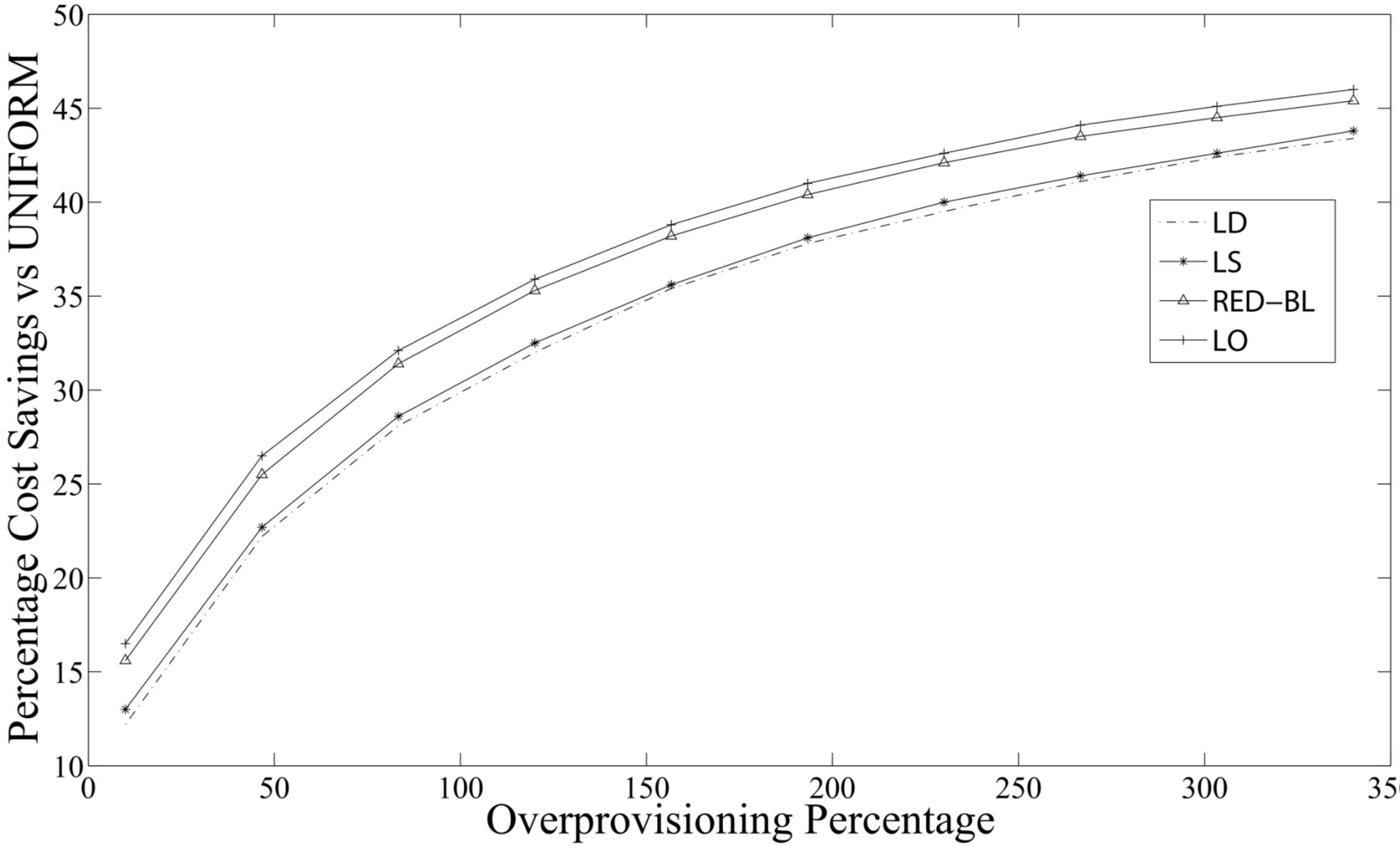
Sometimes idling is better

LS: Local Optimal with Selection

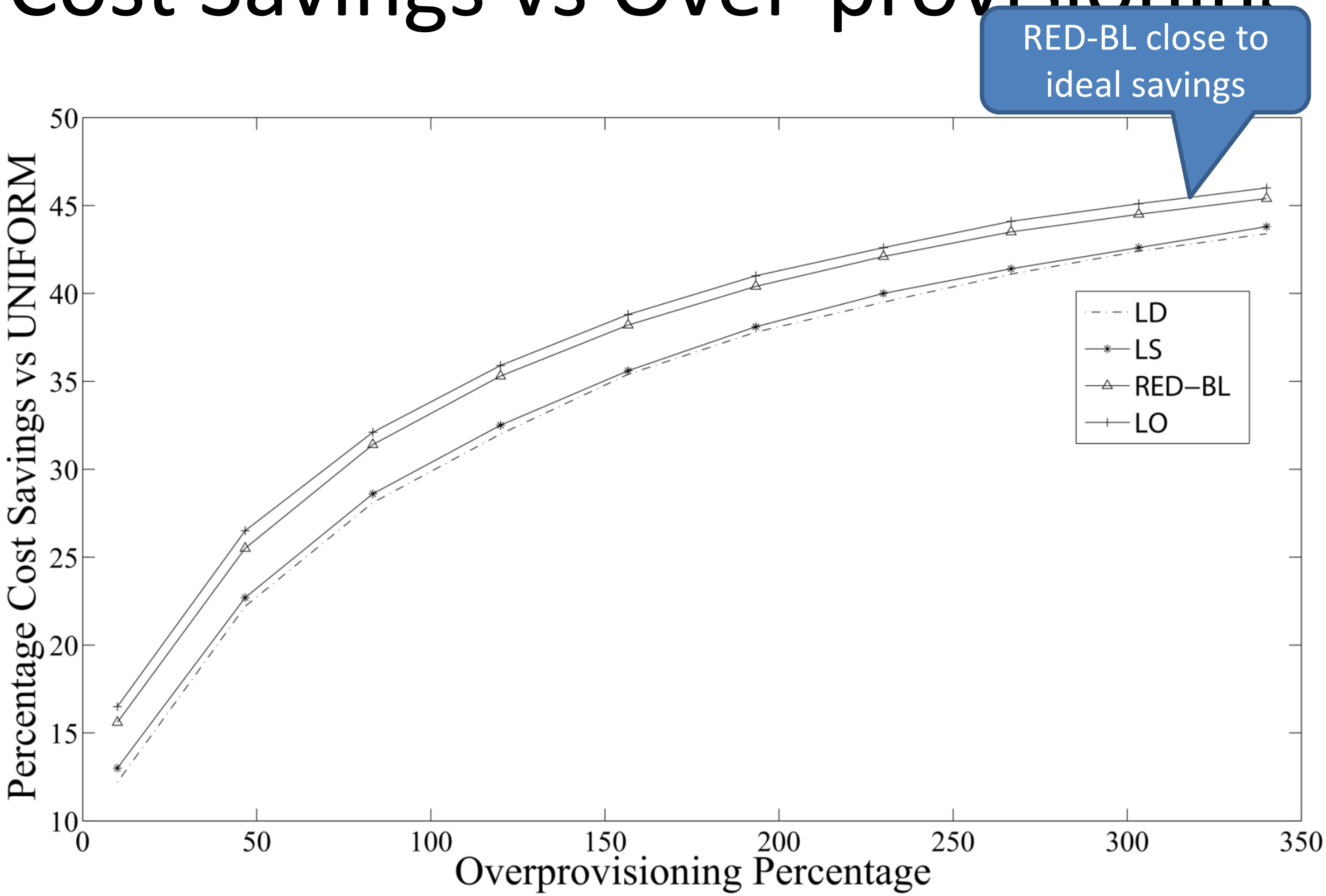
Best practical variant of local optimal



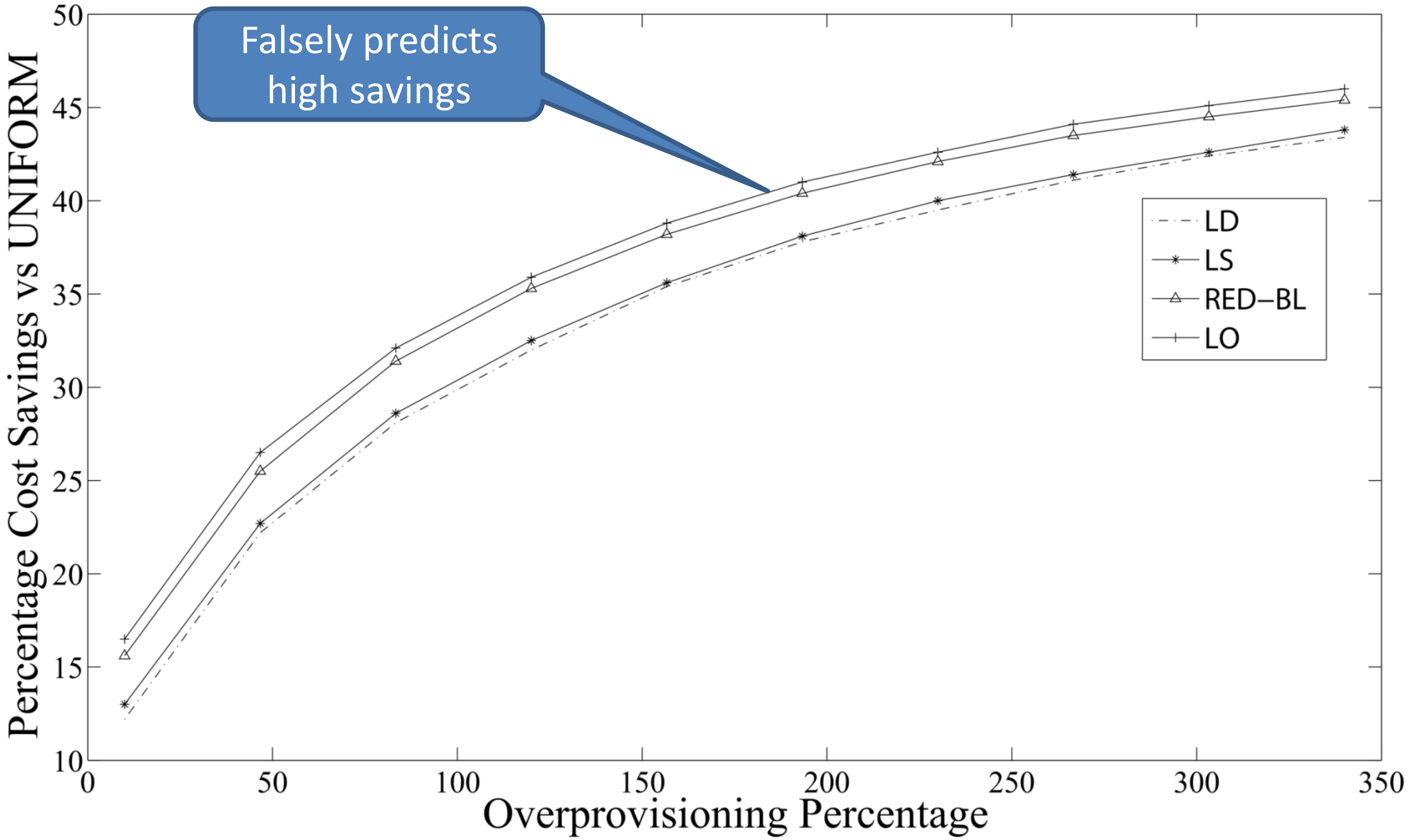
Cost Savings vs Over-provisioning



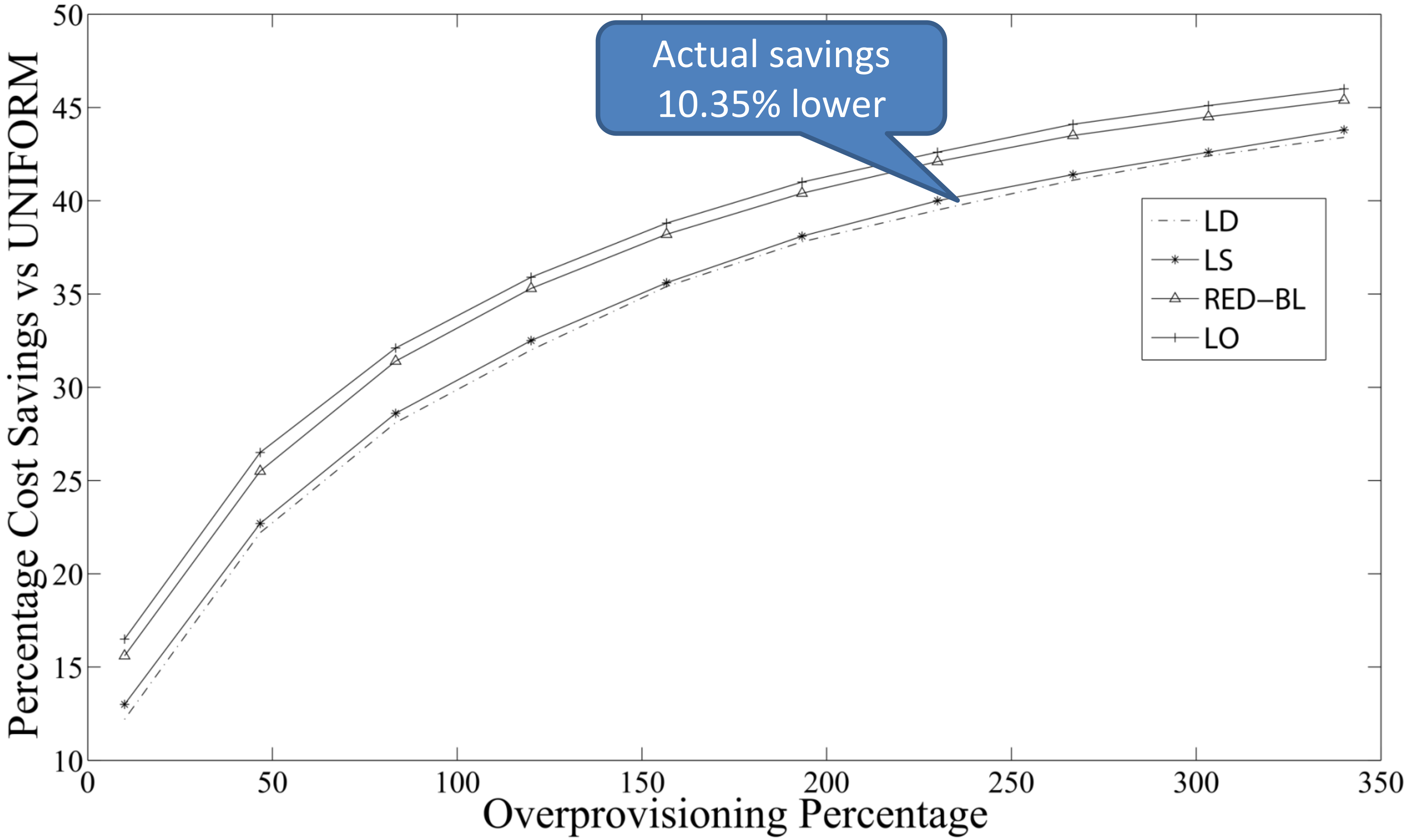
Cost Savings vs Over-provisioning



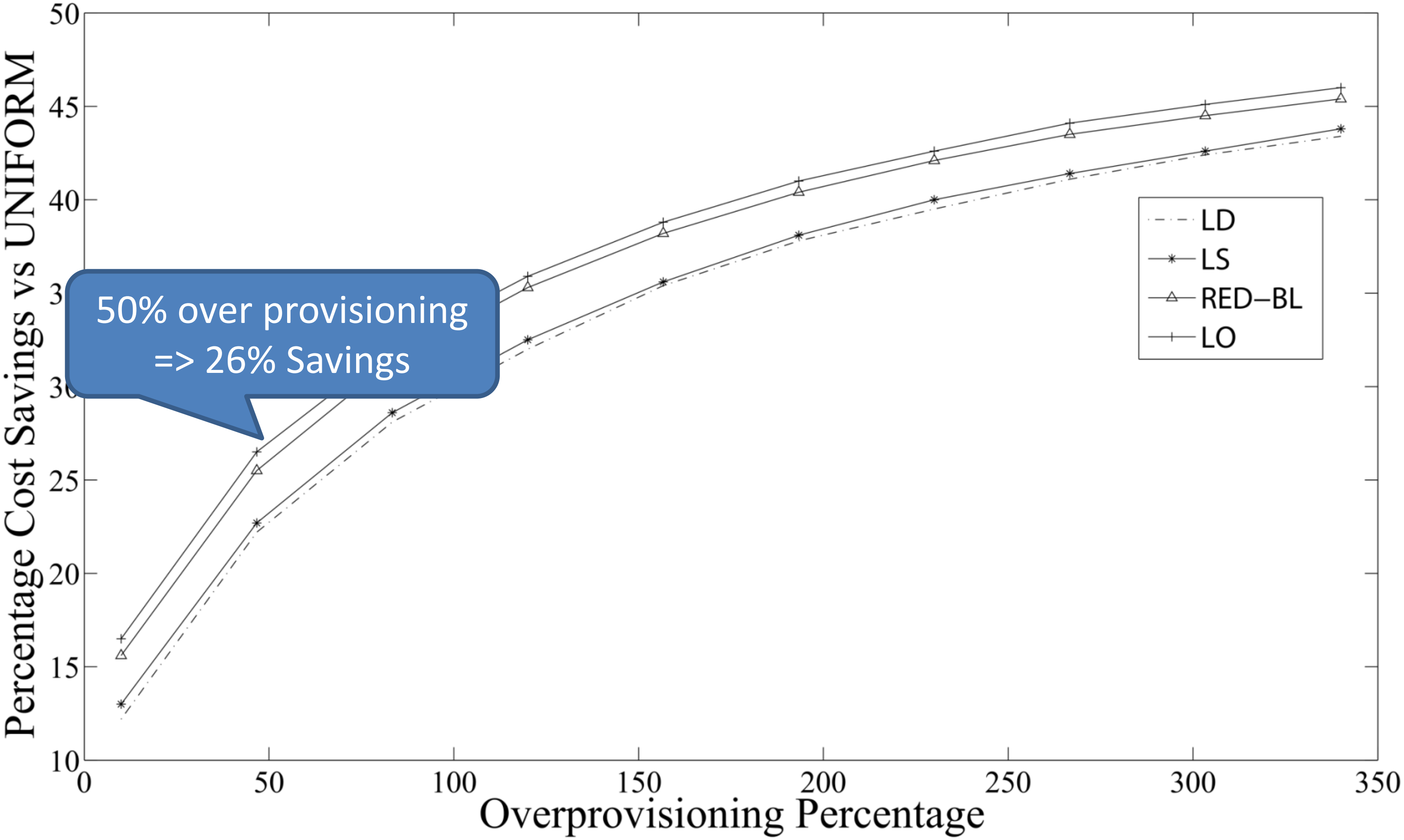
Cost Savings vs Over-provisioning



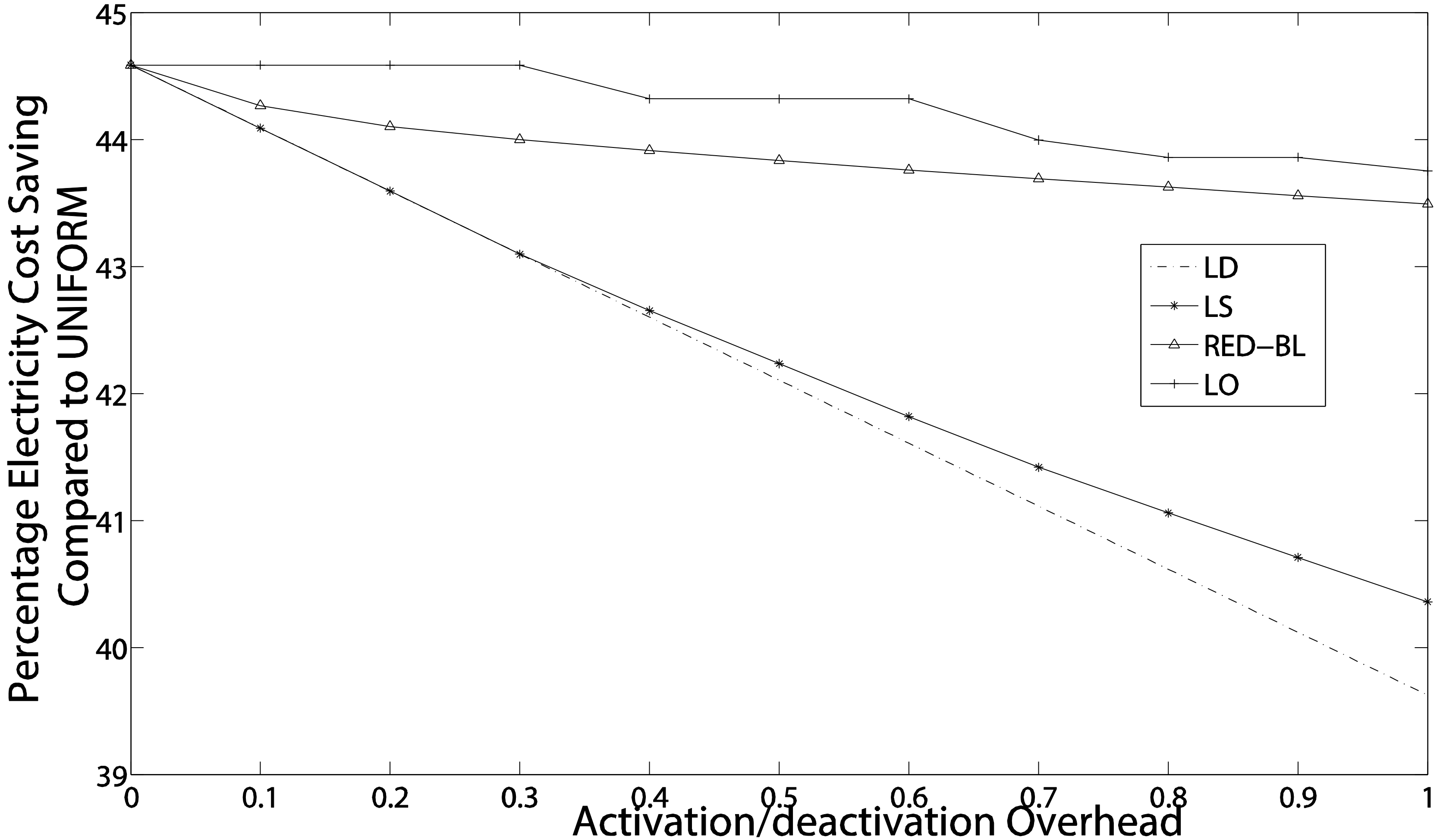
Cost Savings vs Over-provisioning



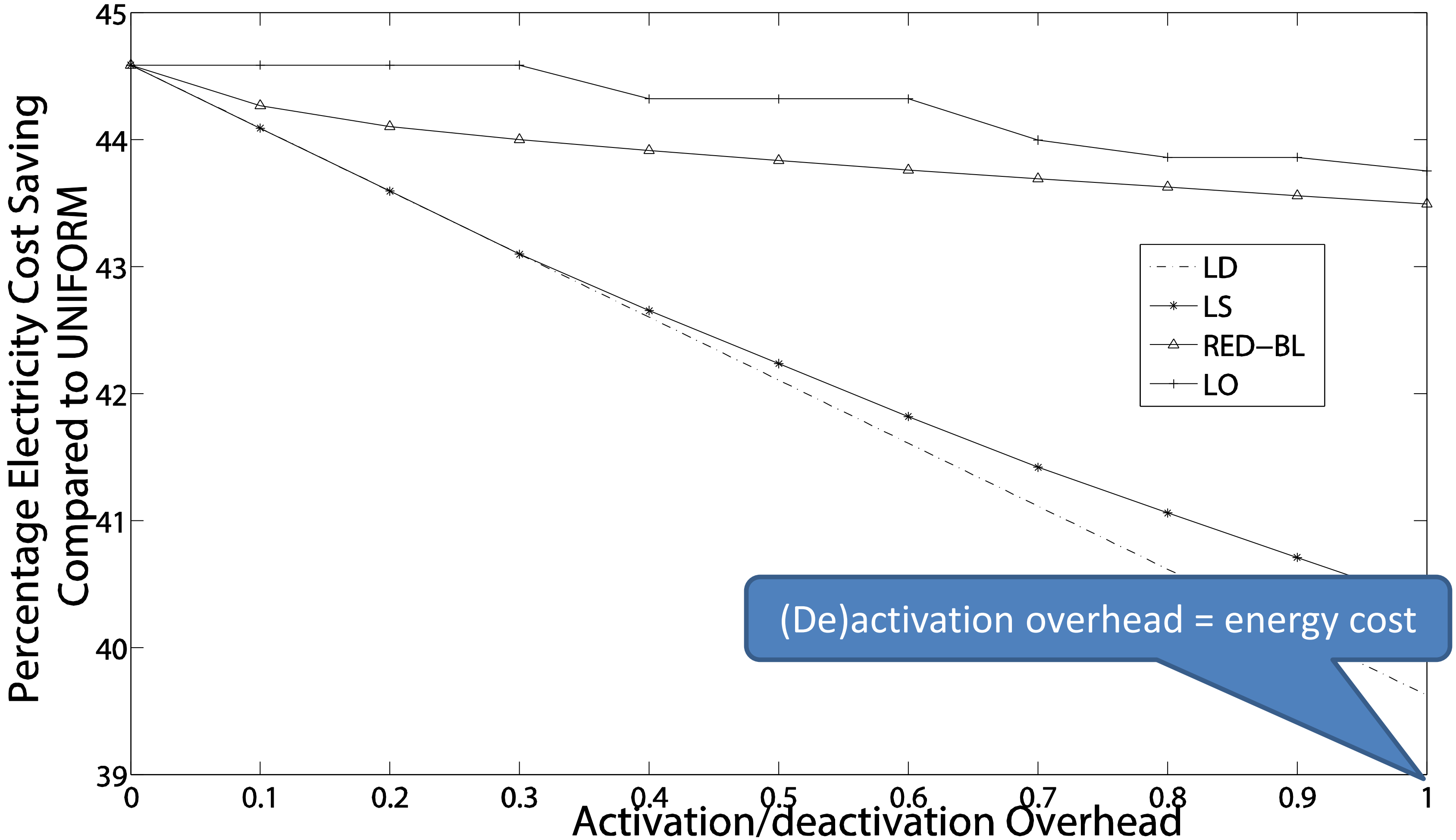
Cost Savings vs Over-provisioning



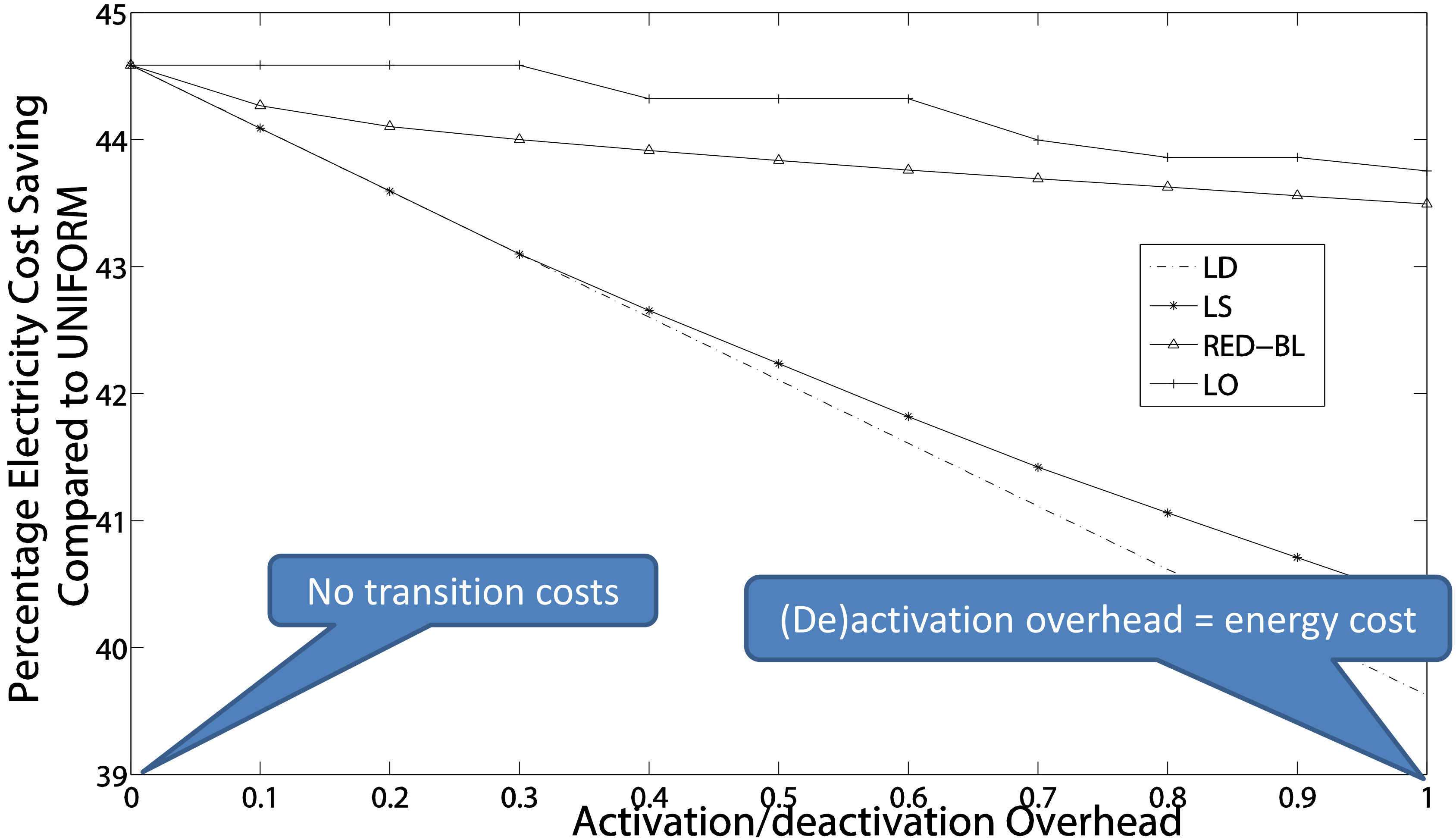
Electricity Cost vs Transition Cost



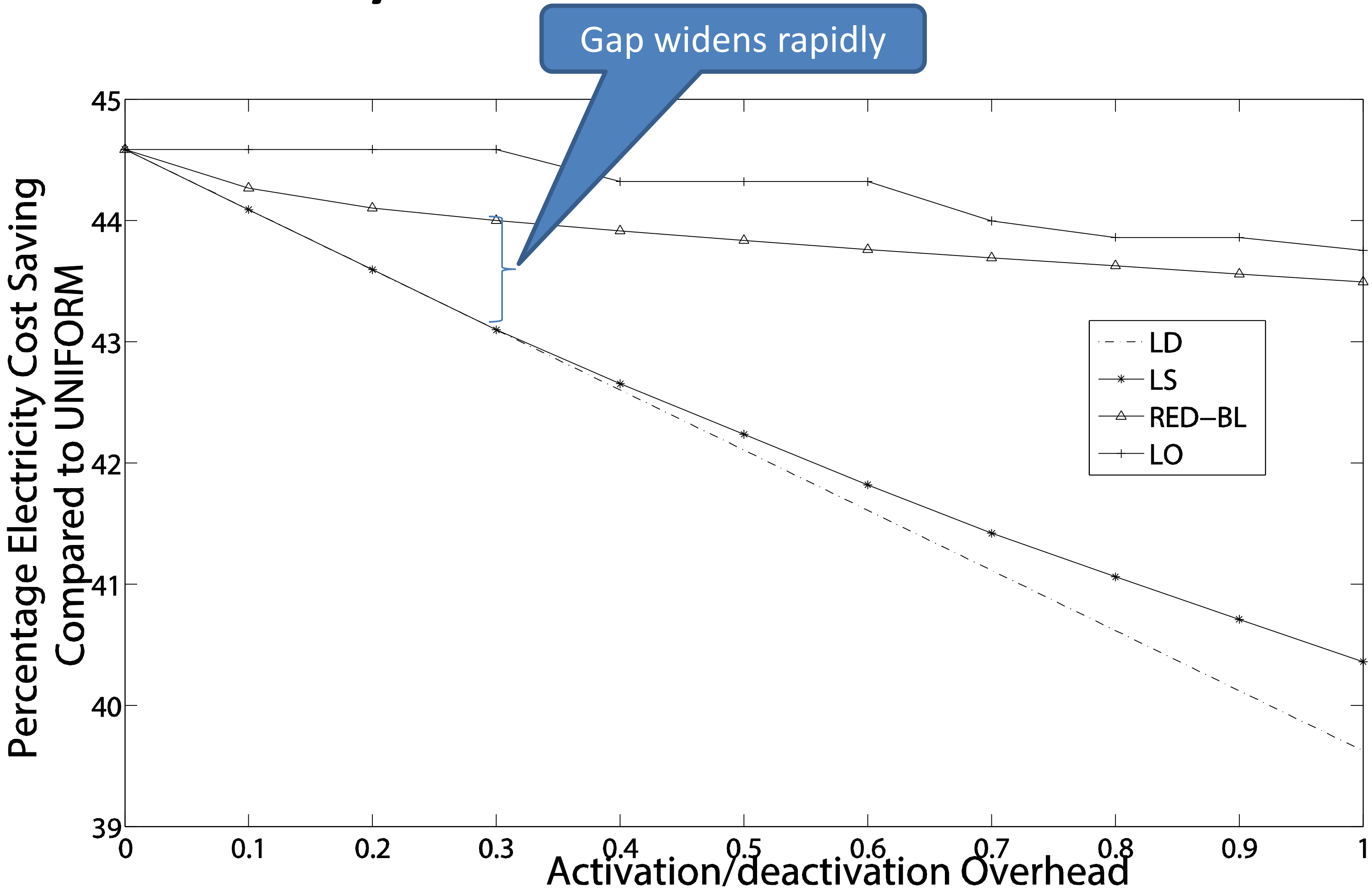
Electricity Cost vs Transition Cost



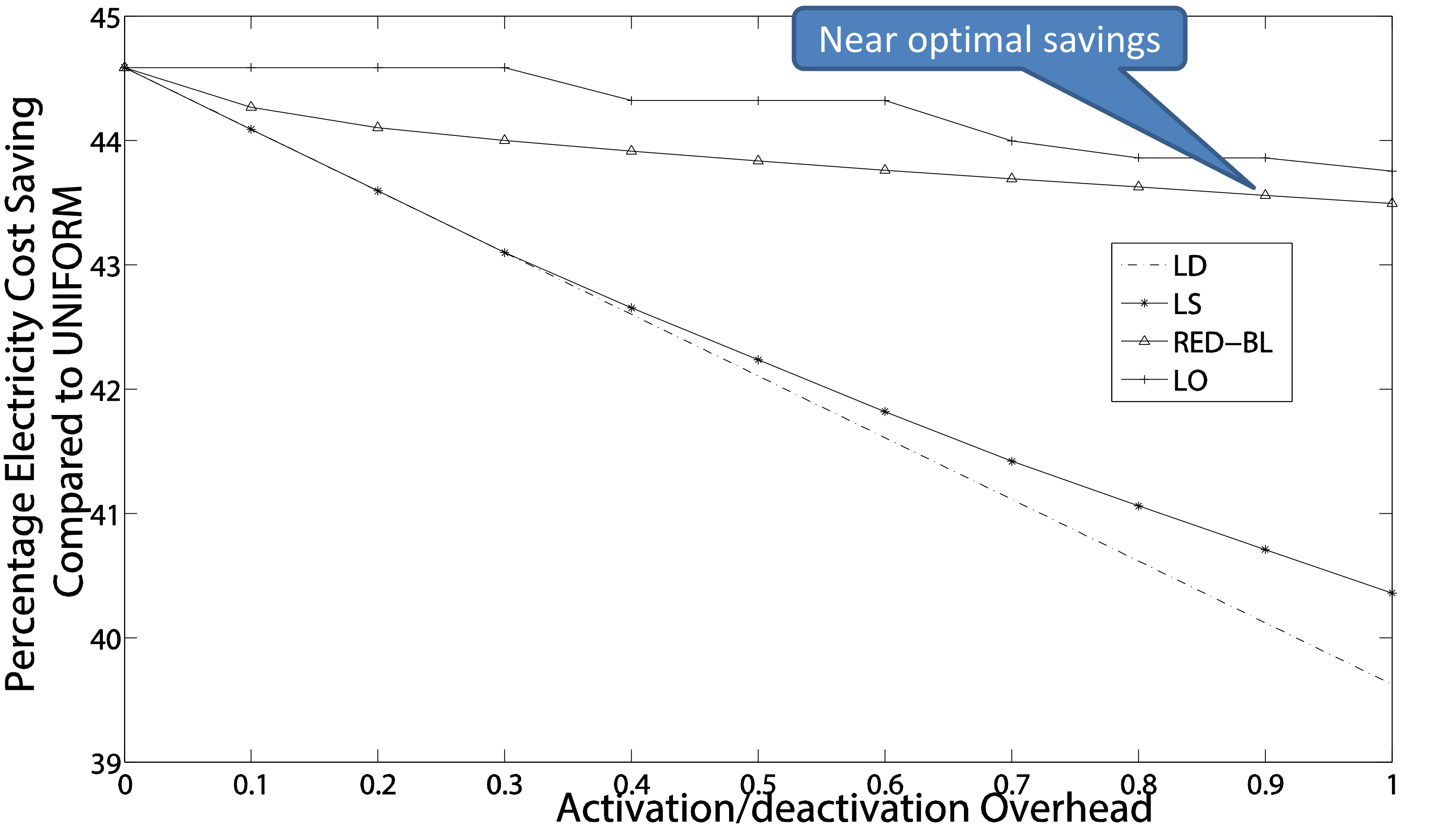
Electricity Cost vs Transition Cost



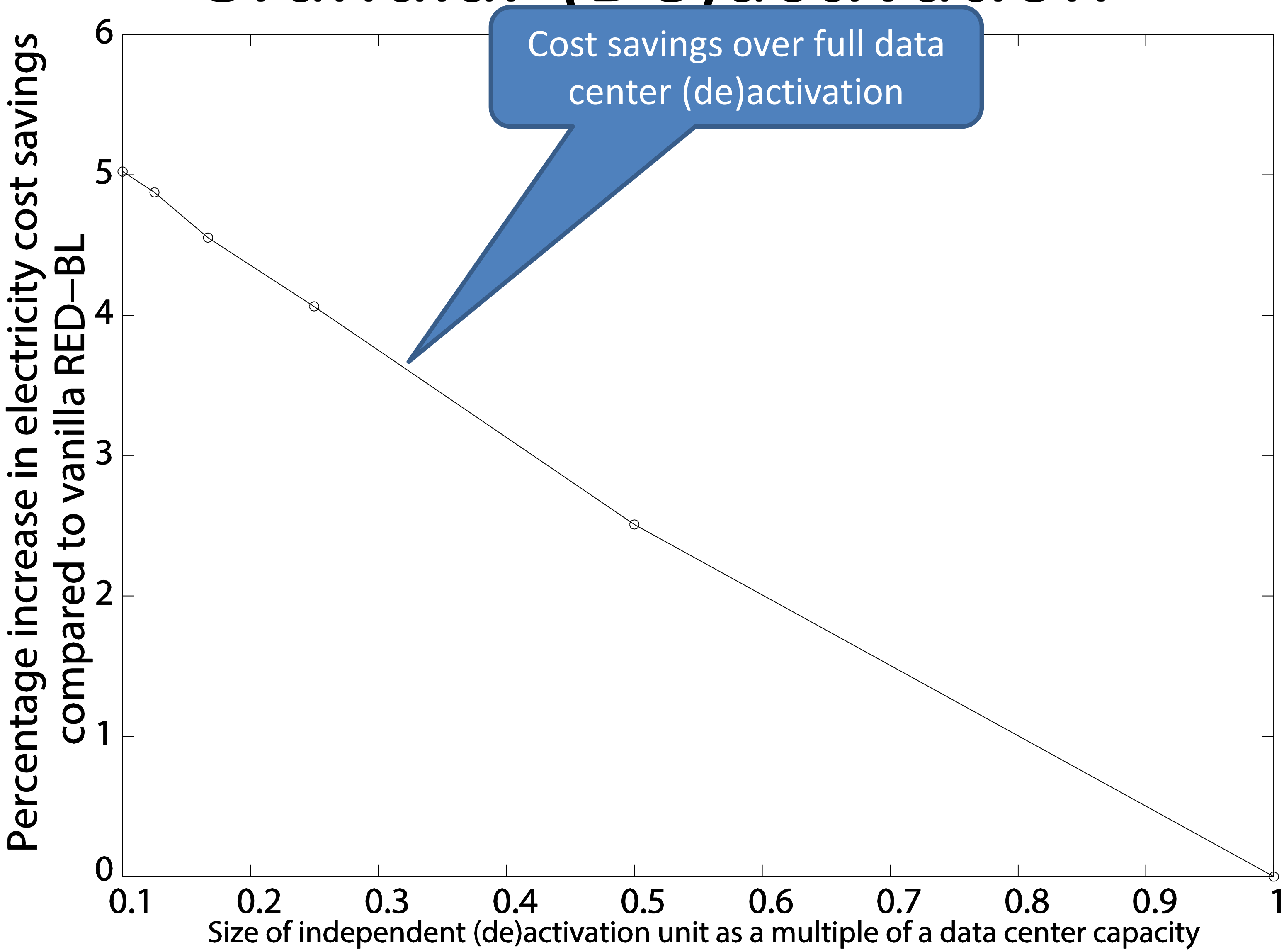
Electricity Cost vs Transition Cost



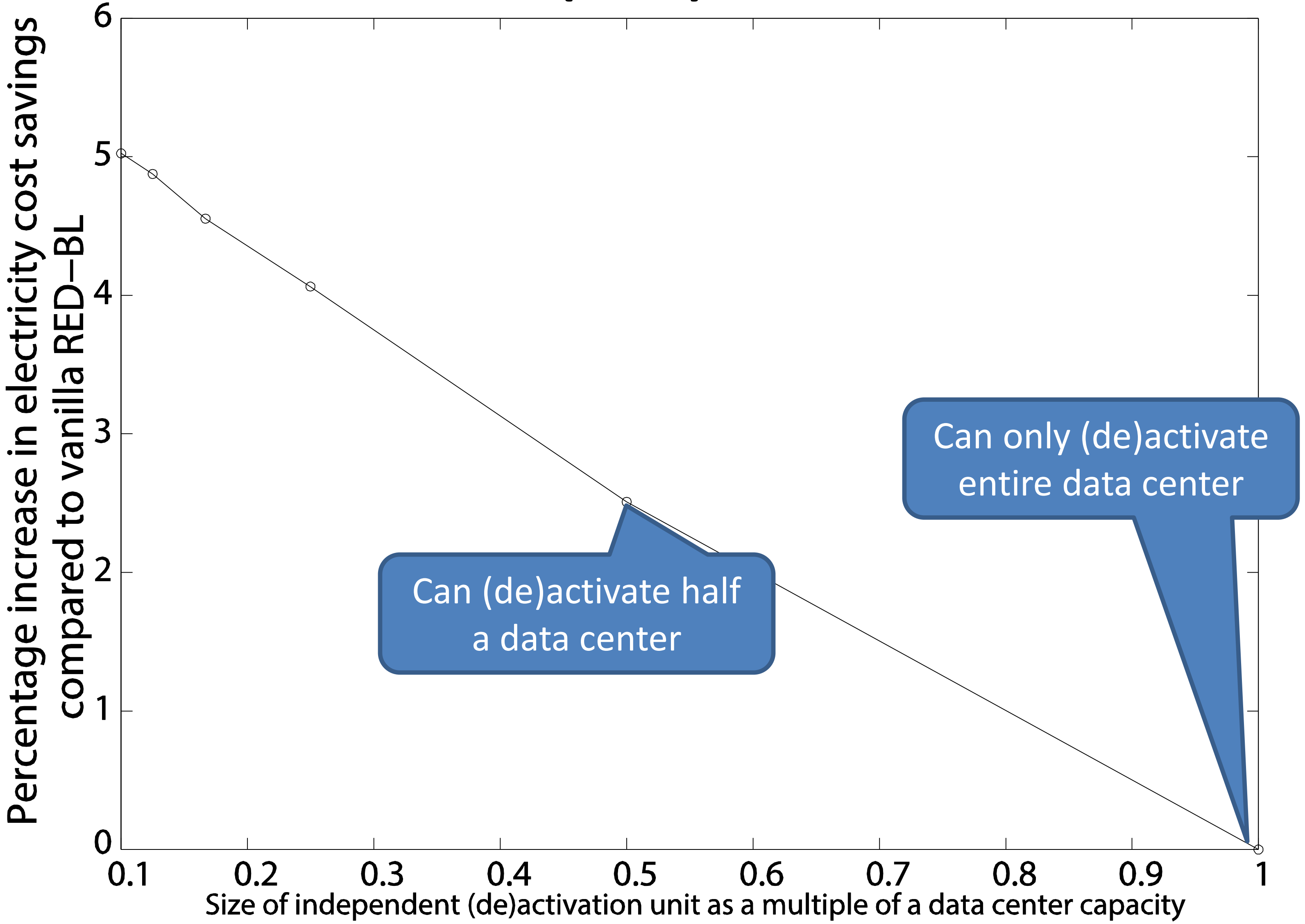
Electricity Cost vs Transition Cost



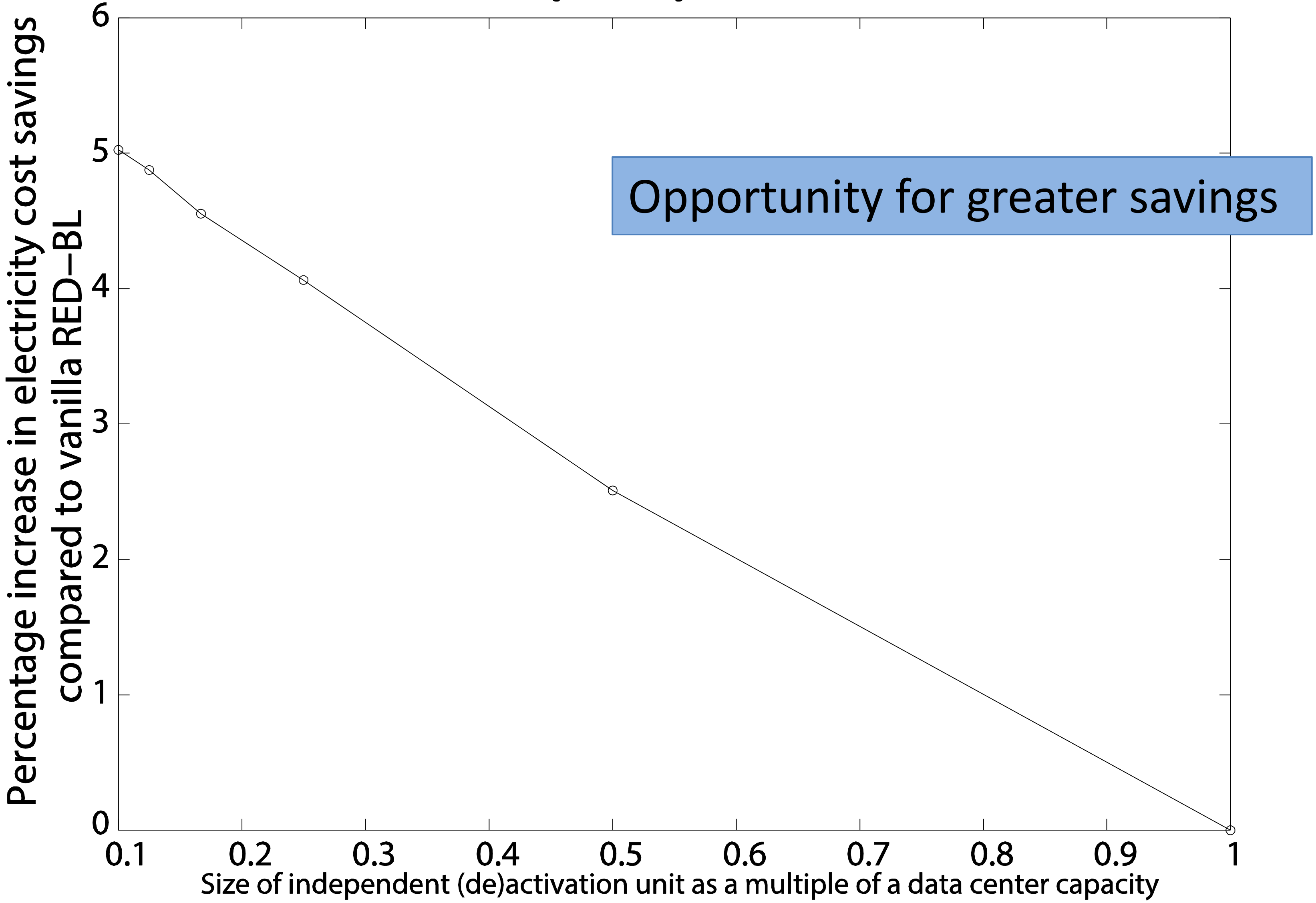
Granular (De)activation



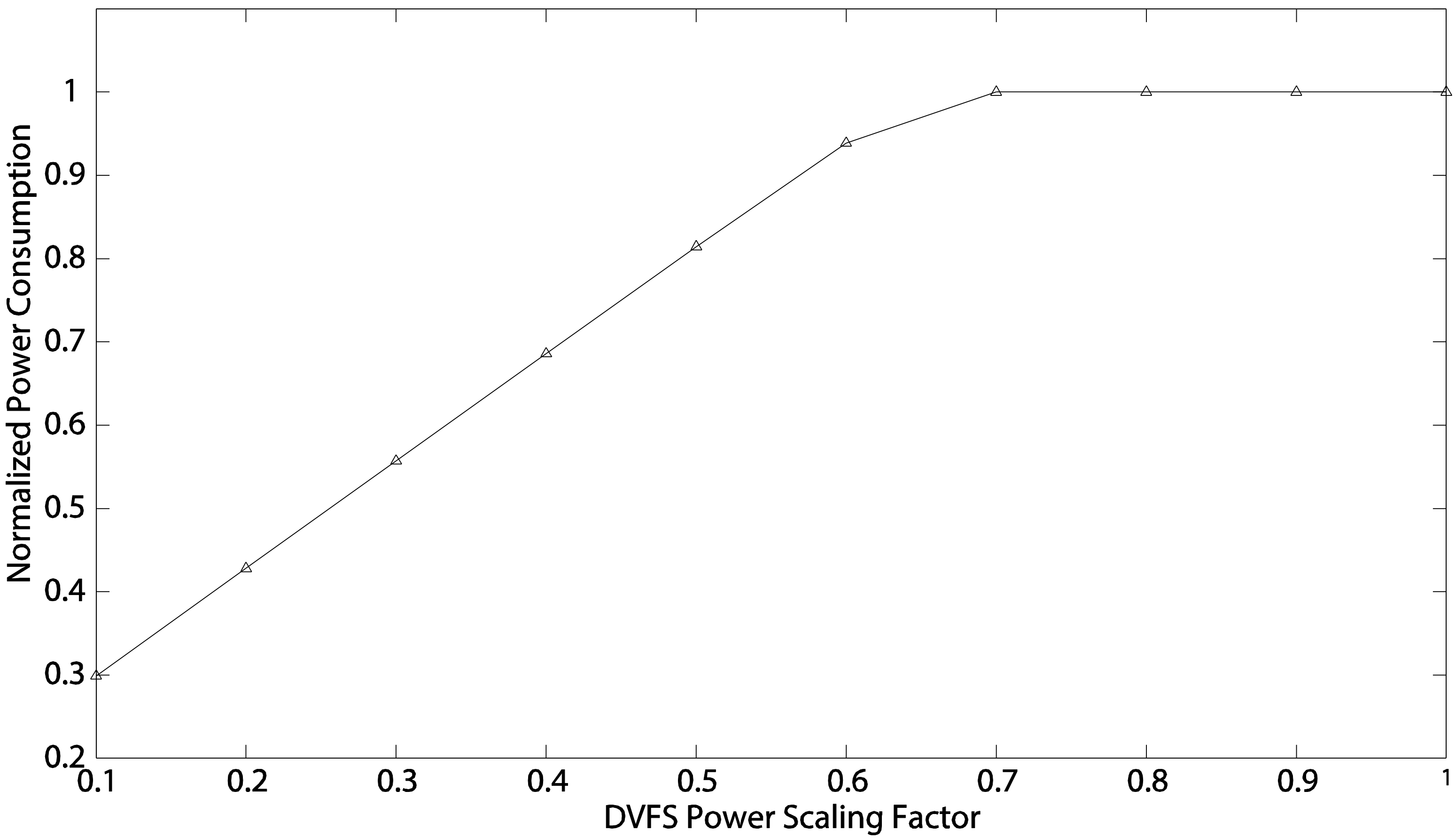
Granular (De)activation



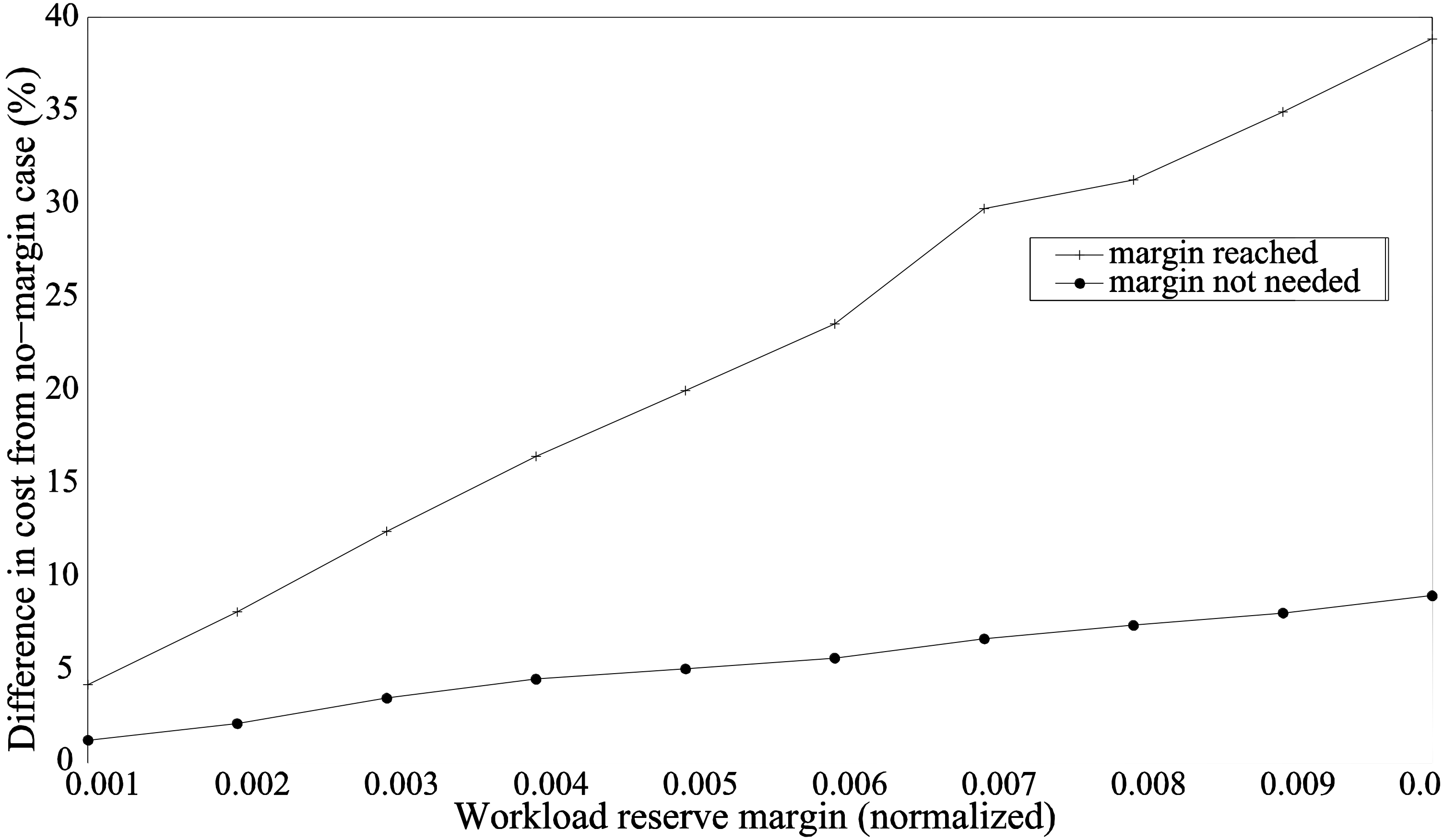
Granular (De)activation



DVFS Instead of Deactivation



Reserve Margin



Summary – Case Study I

- Significant cost savings are possible using RED-BL

Summary – Case Study I

- Significant cost savings are possible using RED-BL
- Finer granularity of resource (de)activation increases savings

Summary – Case Study I

- Significant cost savings are possible using RED-BL
- Finer granularity of resource (de)activation increases savings
- Low-power mode instead of shutdown also helps considerably

Summary – Case Study I

- Significant cost savings are possible using RED-BL
- Finer granularity of resource (de)activation increases savings
- Low-power mode instead of shutdown also helps considerably
- Cost savings decrease sharply with reserve margin

Summary – Case Study I

- Significant cost savings are possible using RED-BL
- Finer granularity of resource (de)activation increases savings
- Low-power mode instead of shutdown also helps considerably
- Cost savings decrease sharply with reserve margin

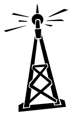
Can we apply this optimization “machinery” to other networks?

Agenda

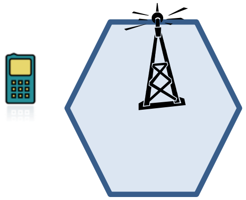
- Background and motivation
- Opportunity and key idea
- Case studies:
  - Data centers (e.g., Facebook and Google)
  - **Cellular networks (e.g., Sprint and Verizon)**
- Conclusions and future work

Case Study II
Cellular Networks

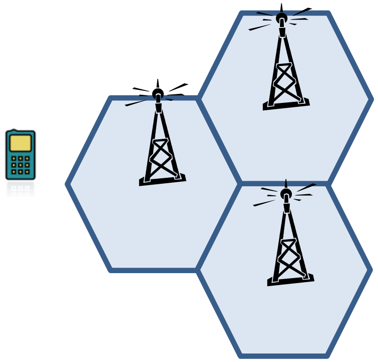
Case Study II
Cellular Networks



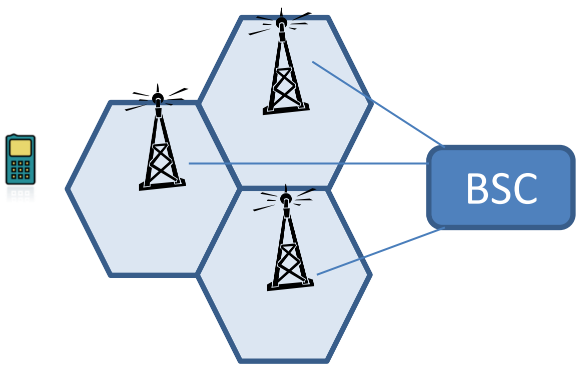
Case Study II
Cellular Networks



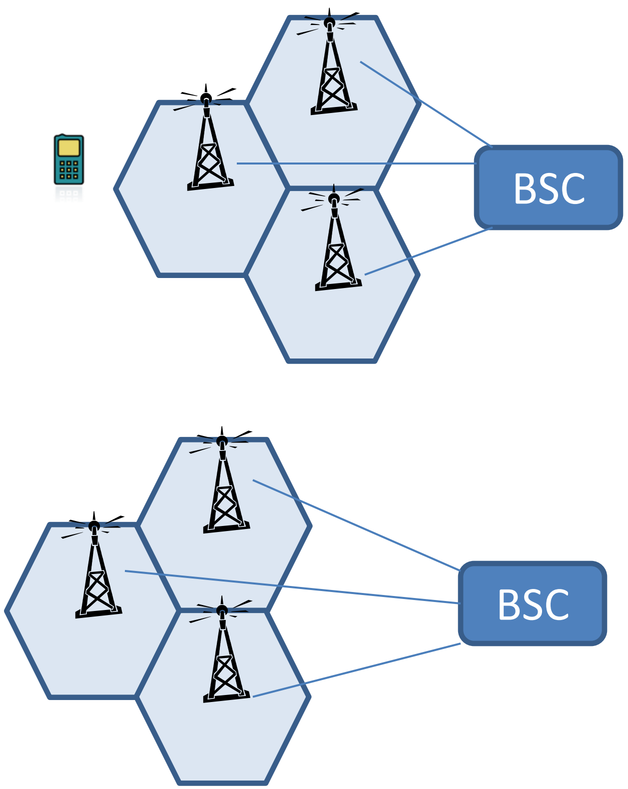
Case Study II
Cellular Networks



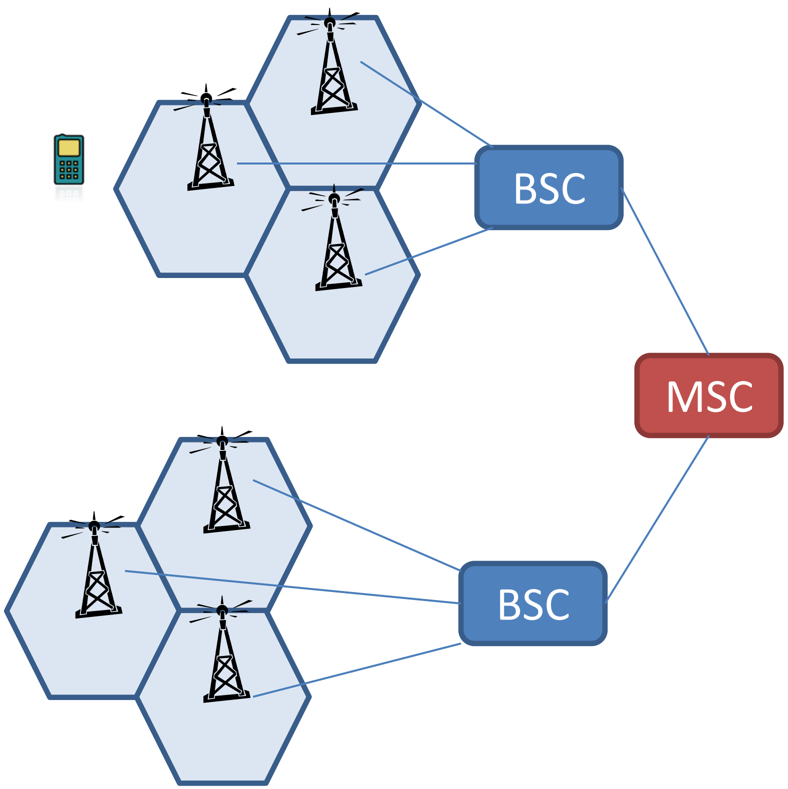
Case Study II
Cellular Networks



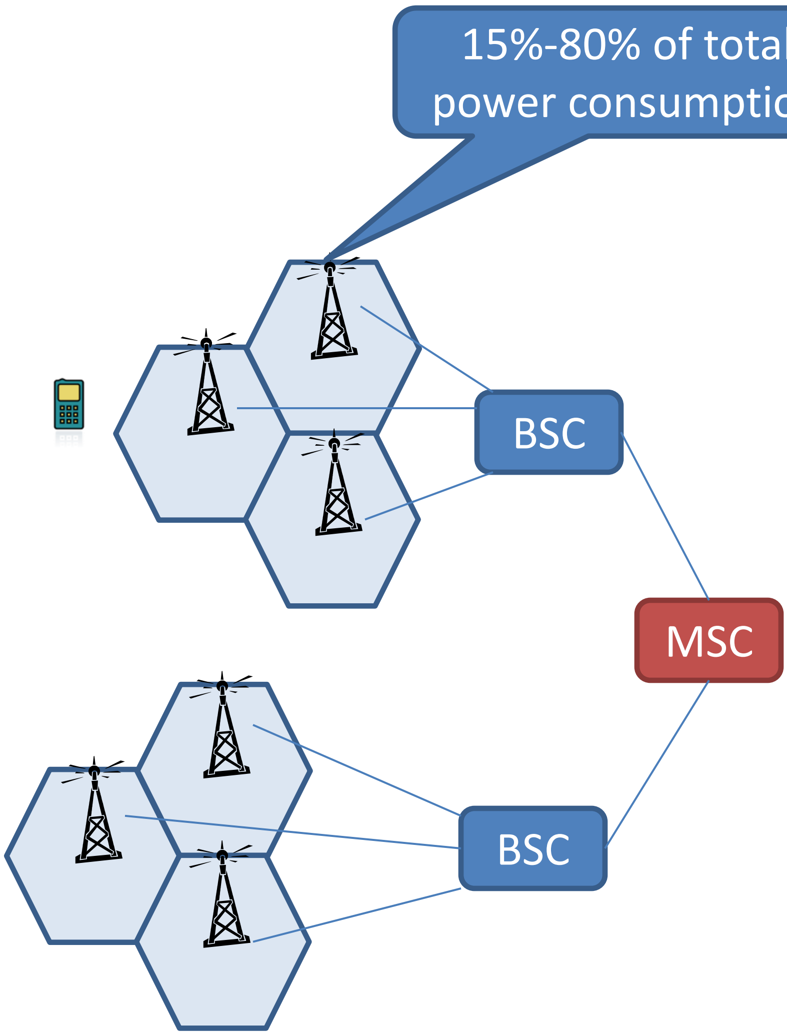
Case Study II
Cellular Networks



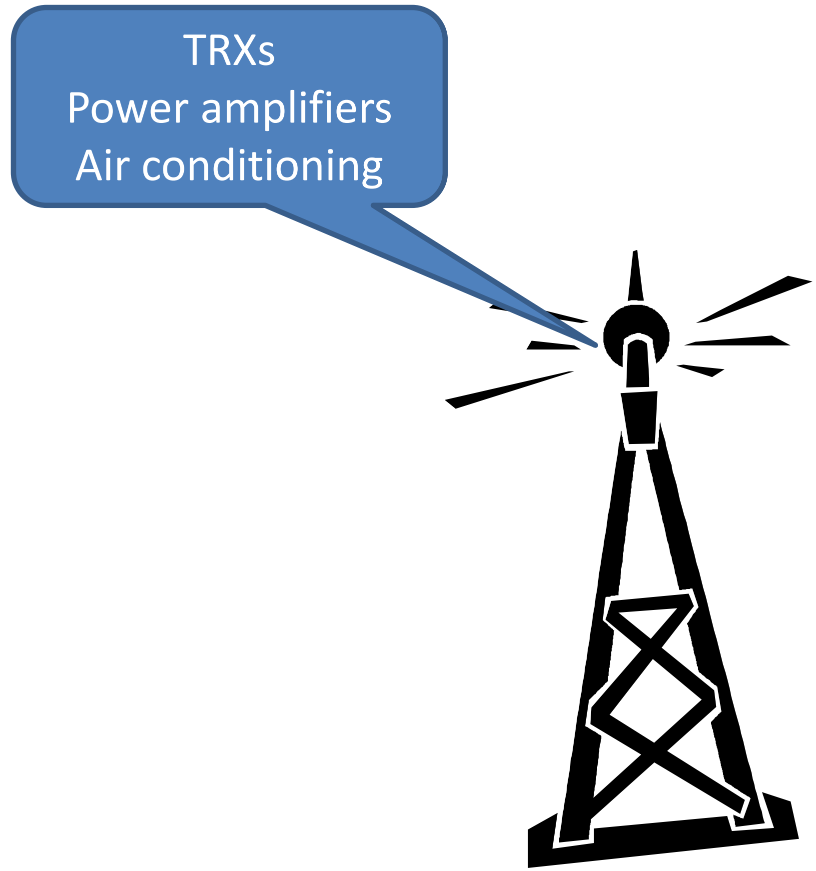
Case Study II
Cellular Networks



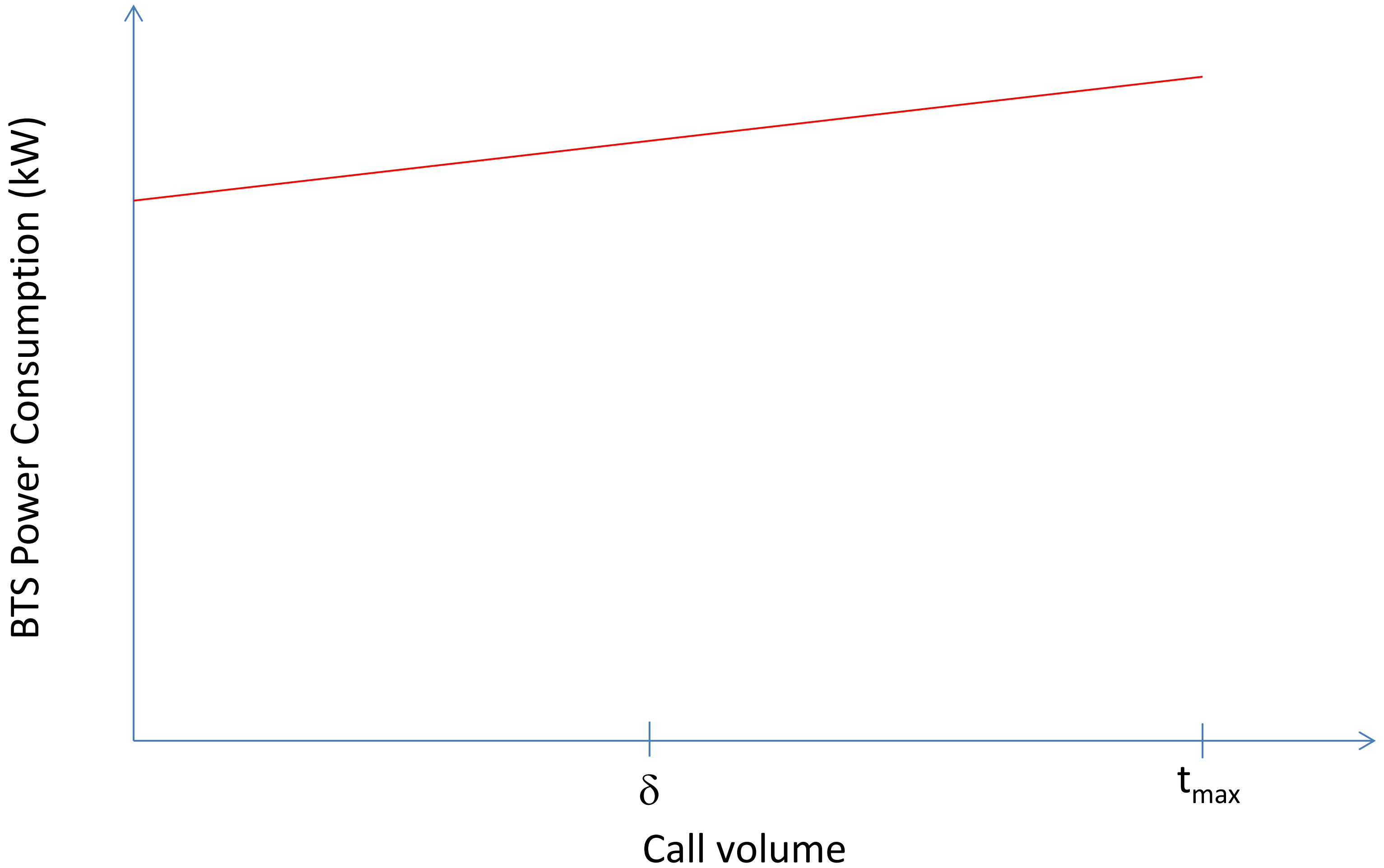
Case Study II
Cellular Networks



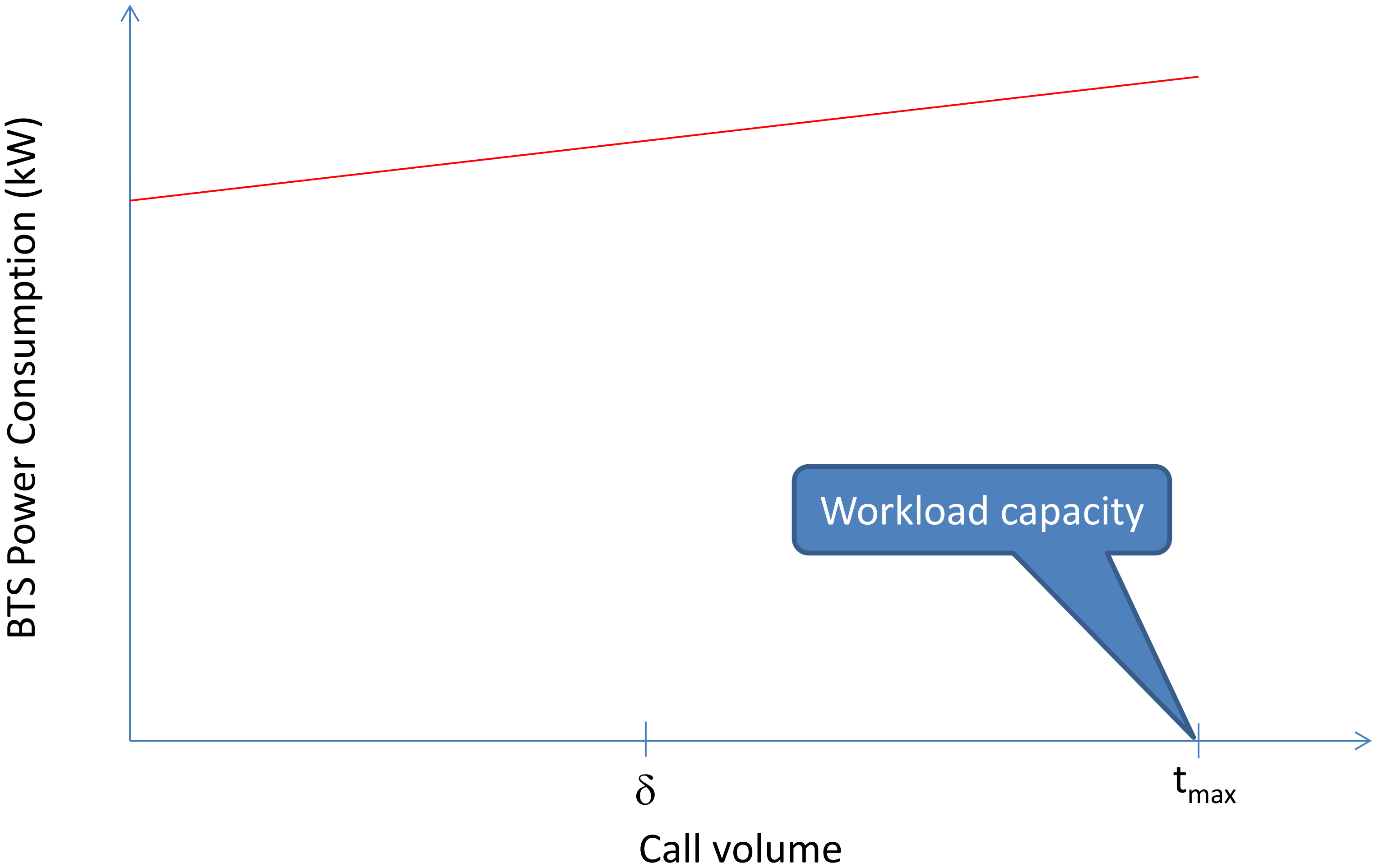
Case Study II
Cellular Networks



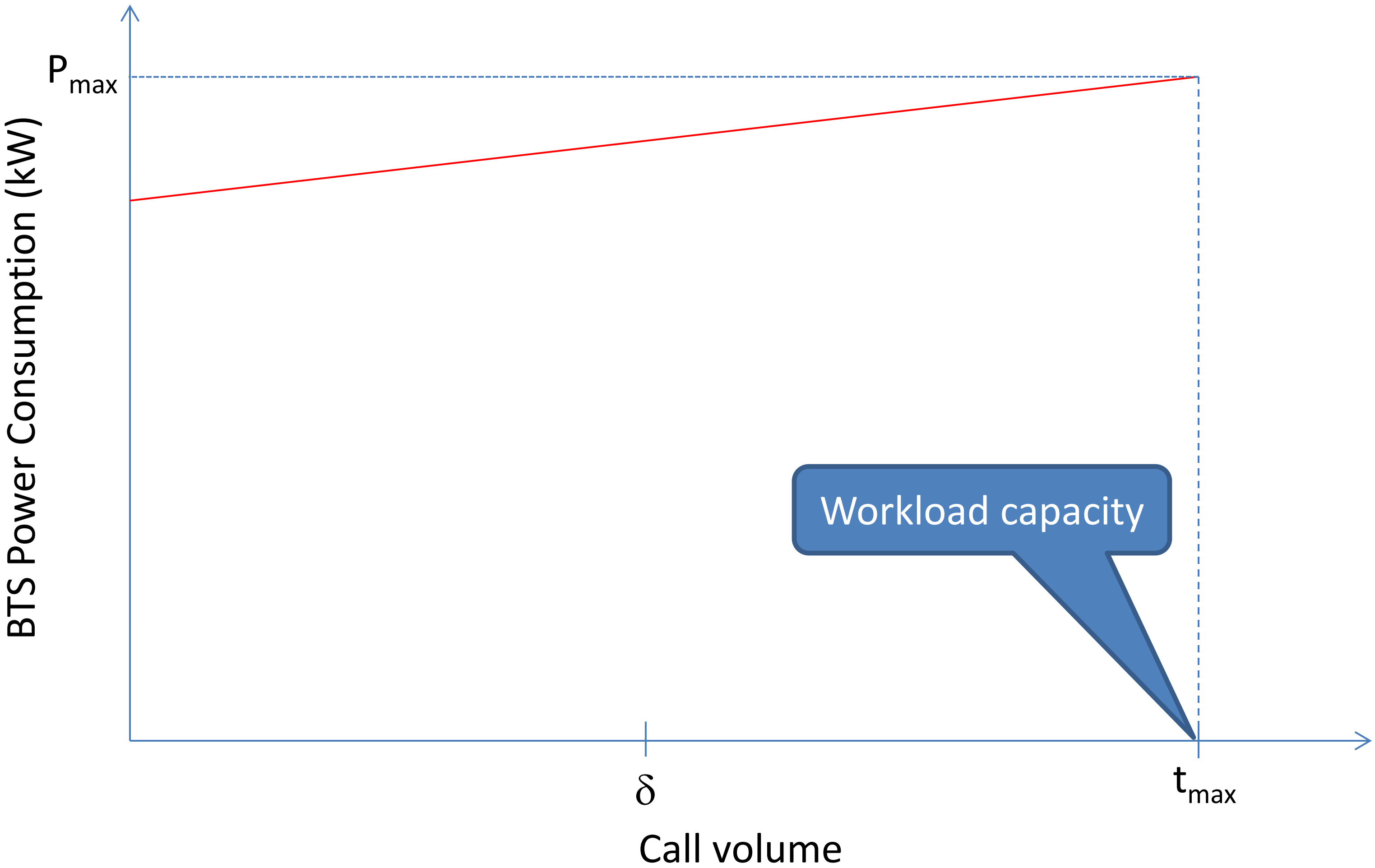
Resource Pruning



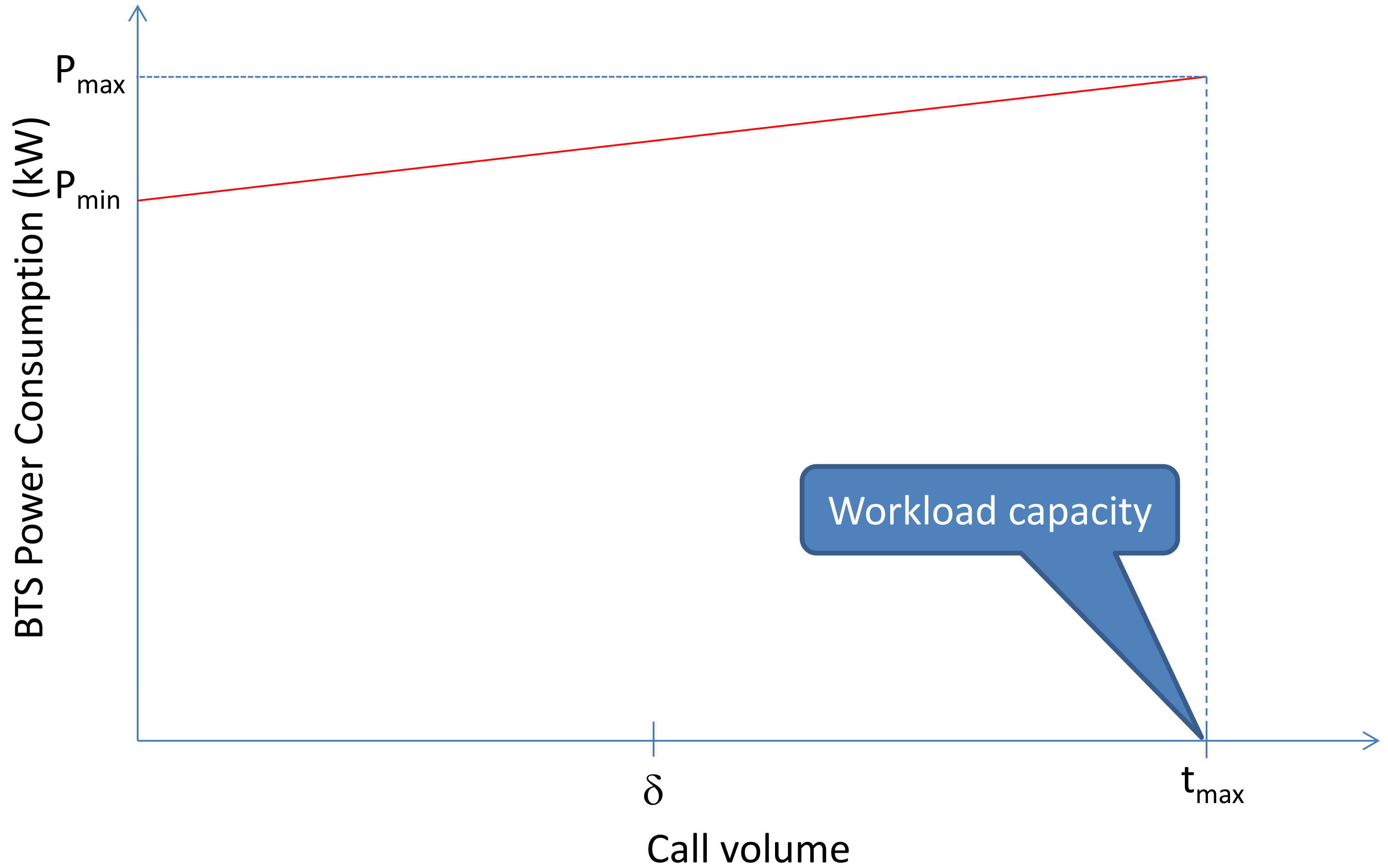
Resource Pruning



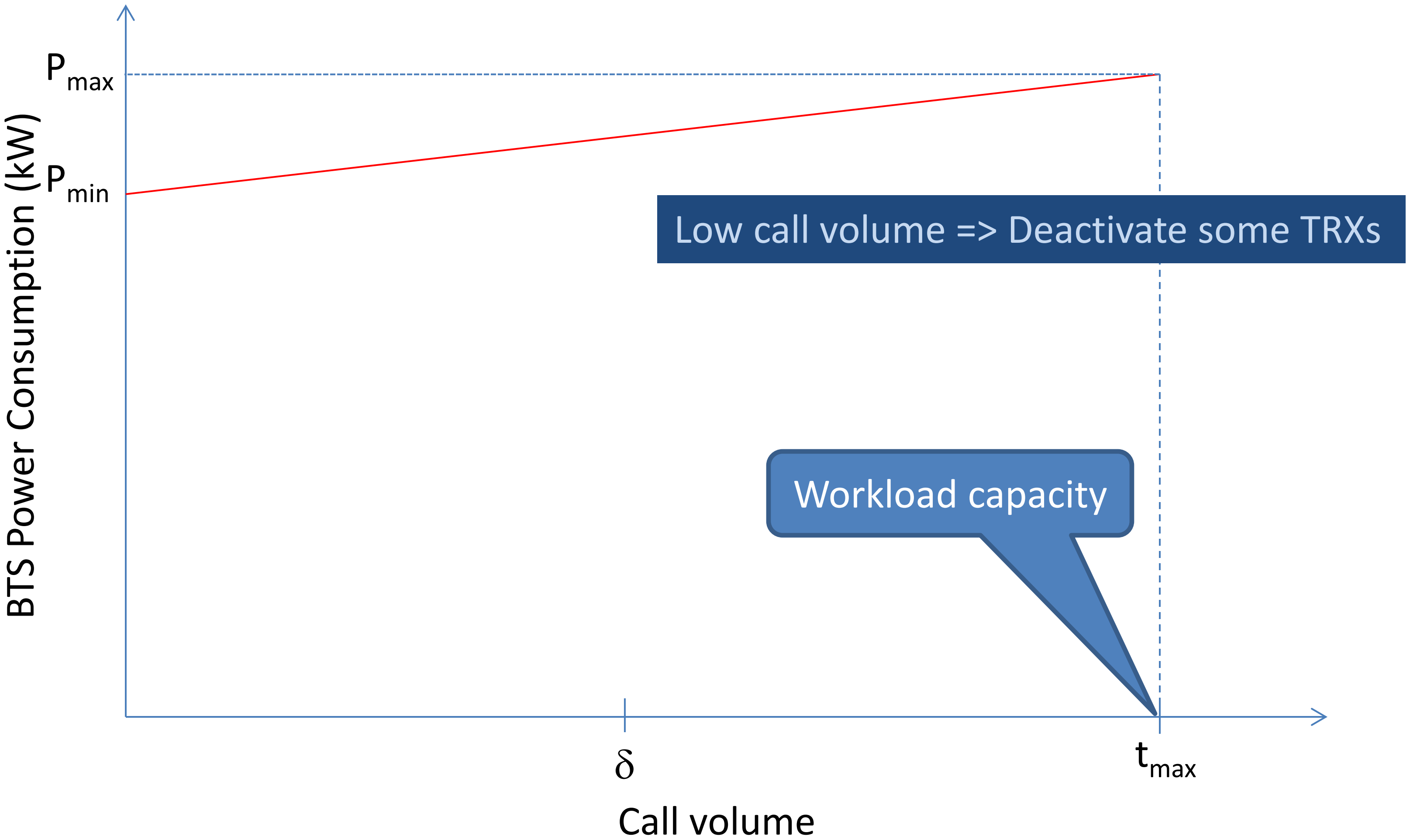
Resource Pruning



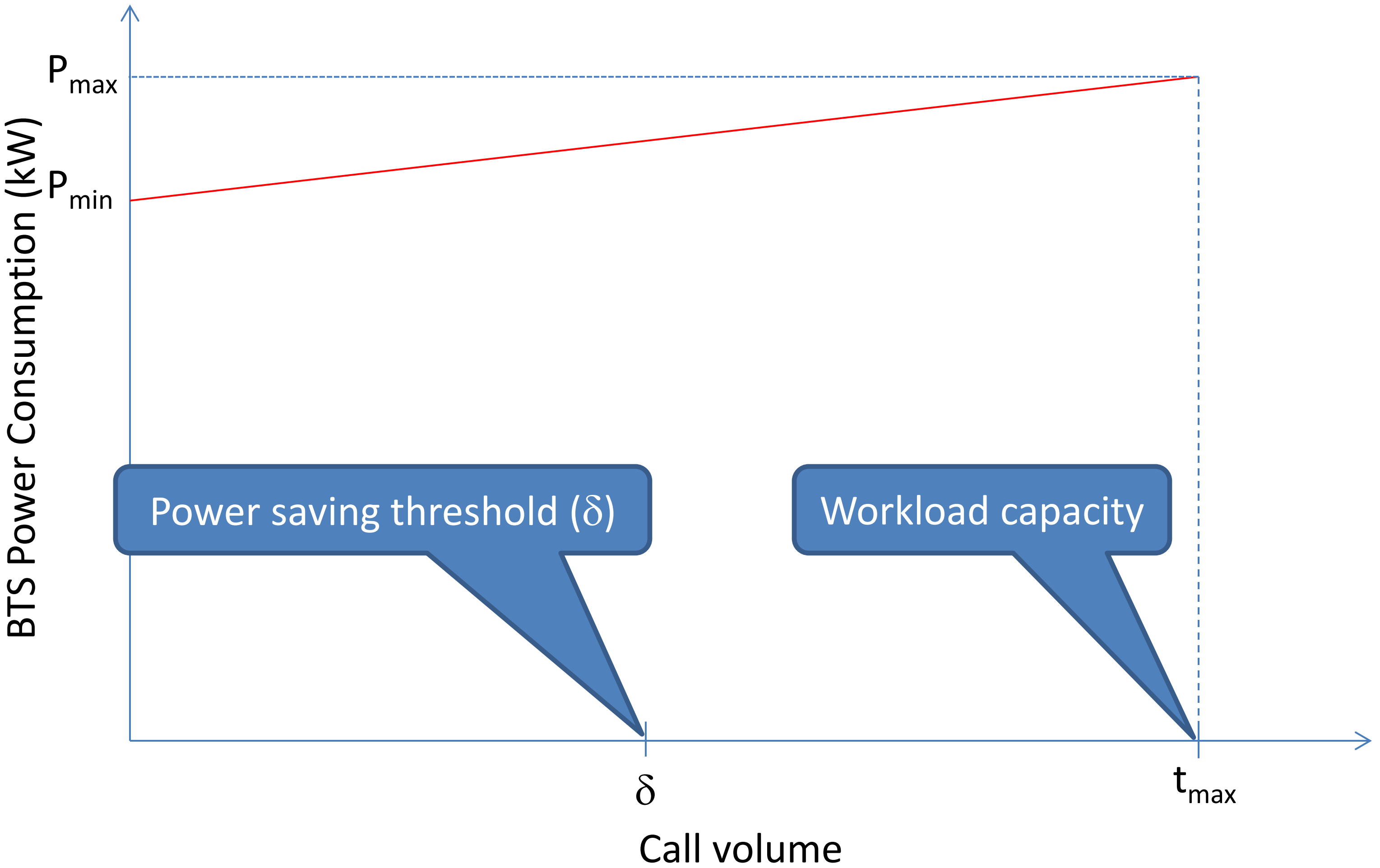
Resource Pruning



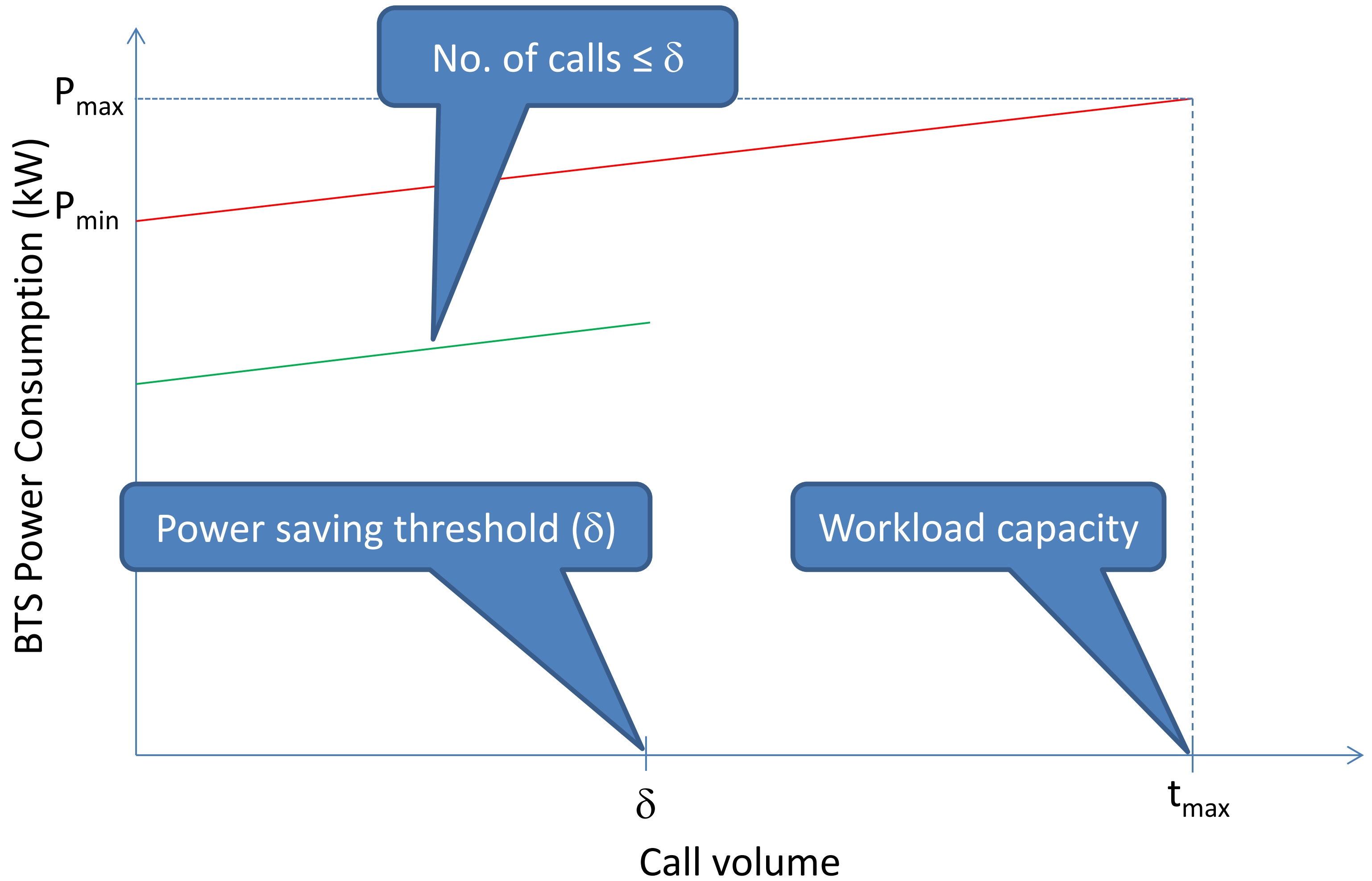
Resource Pruning



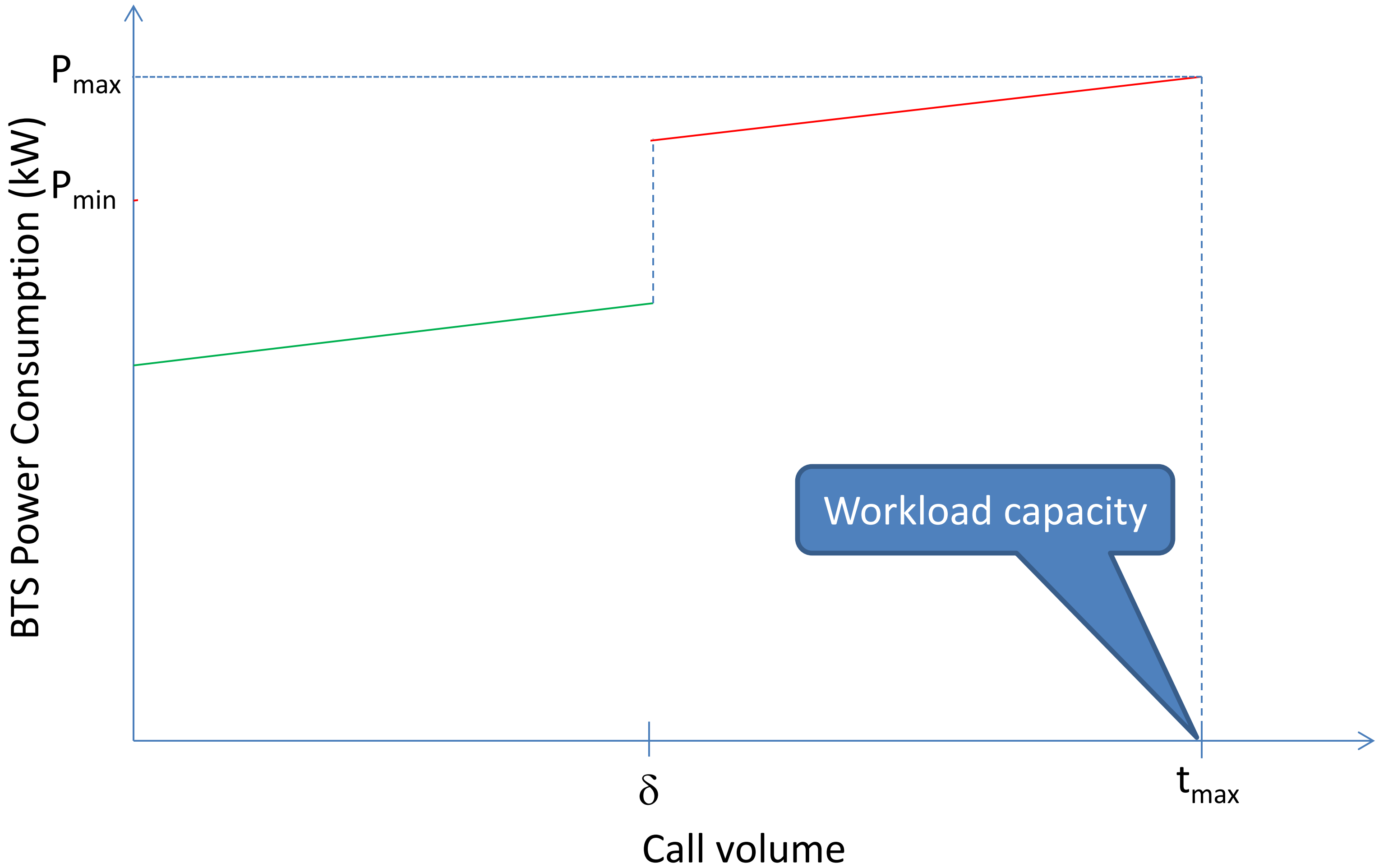
Resource Pruning



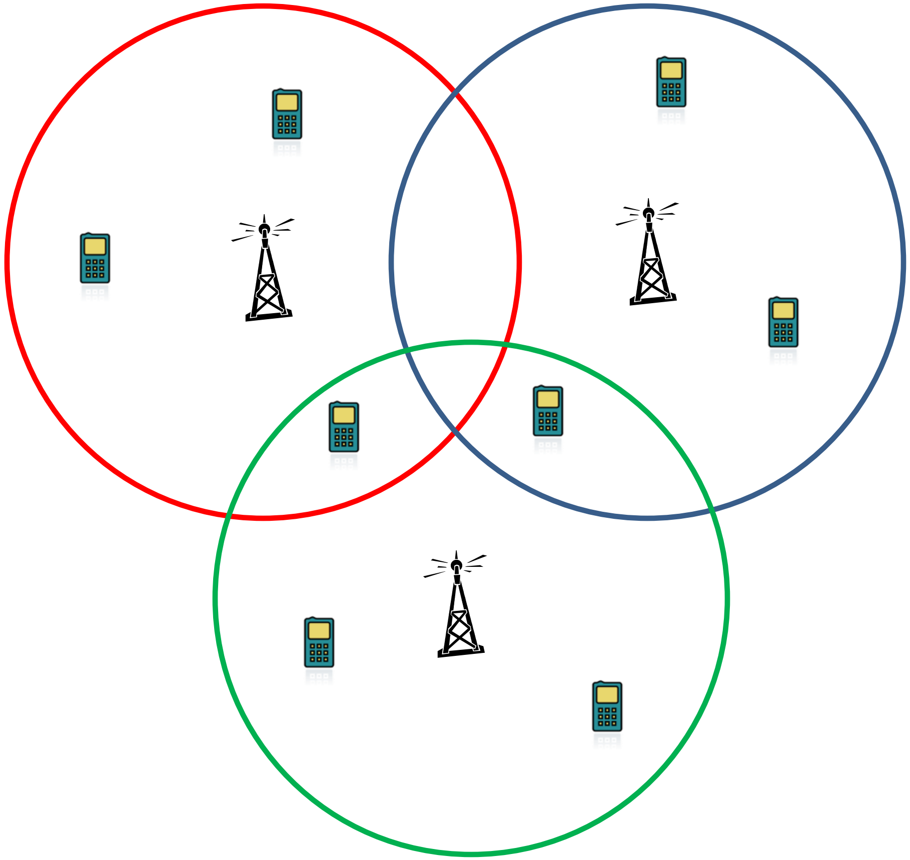
Resource Pruning



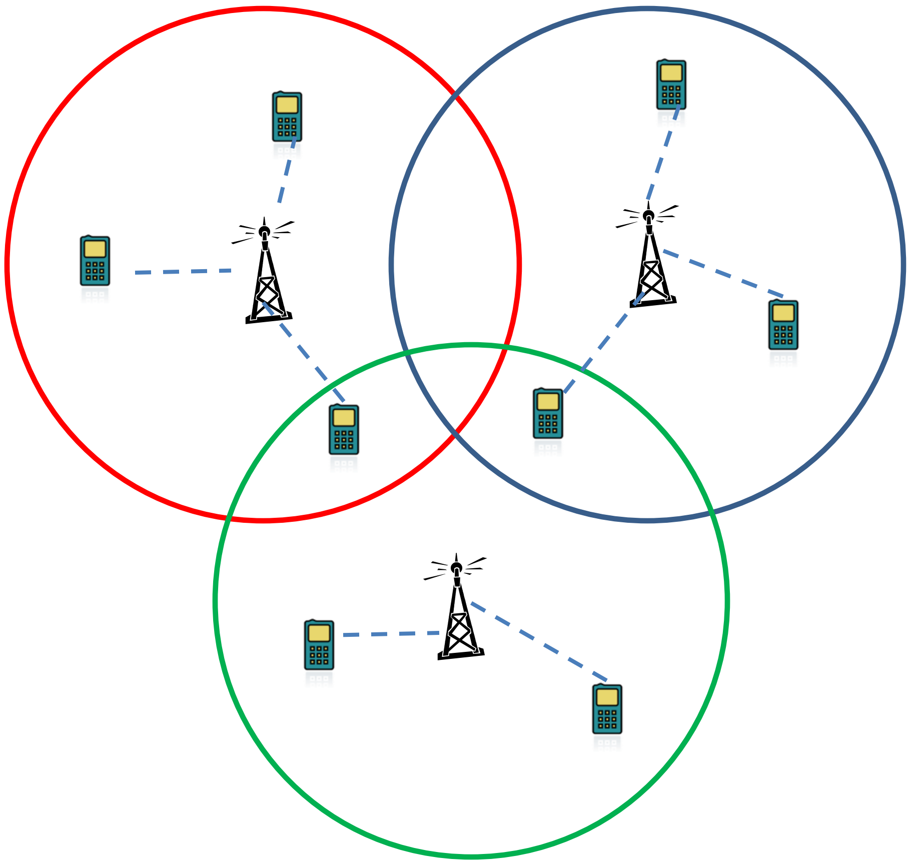
Resource Pruning



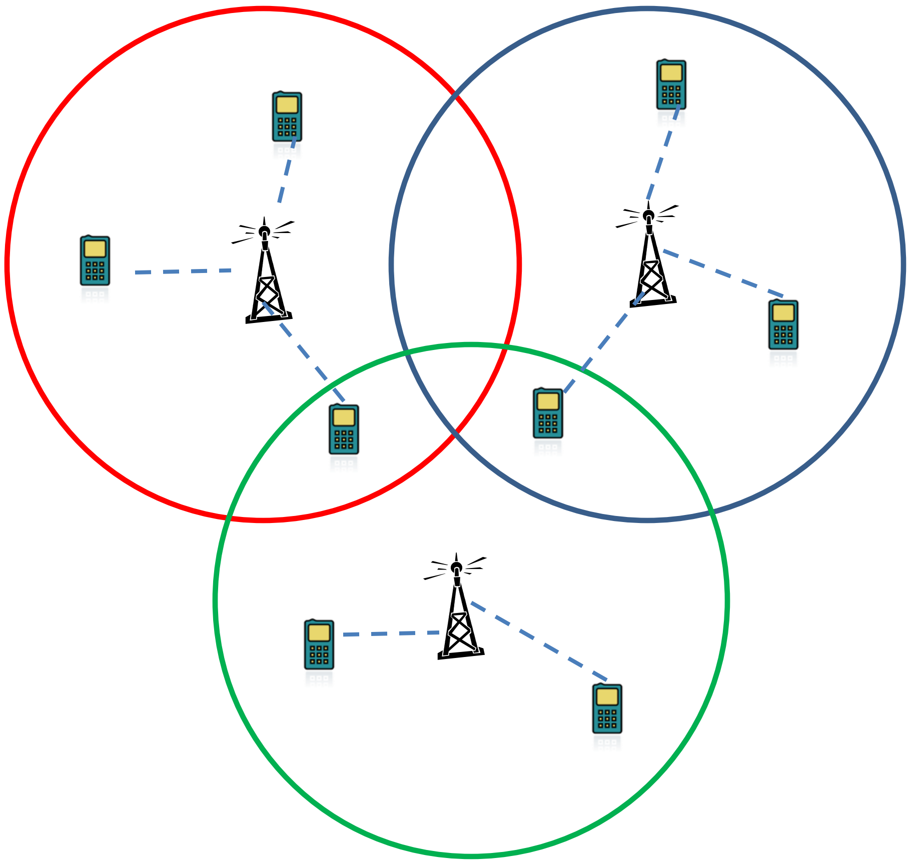
Does workload relocation help?



Does workload relocation help?

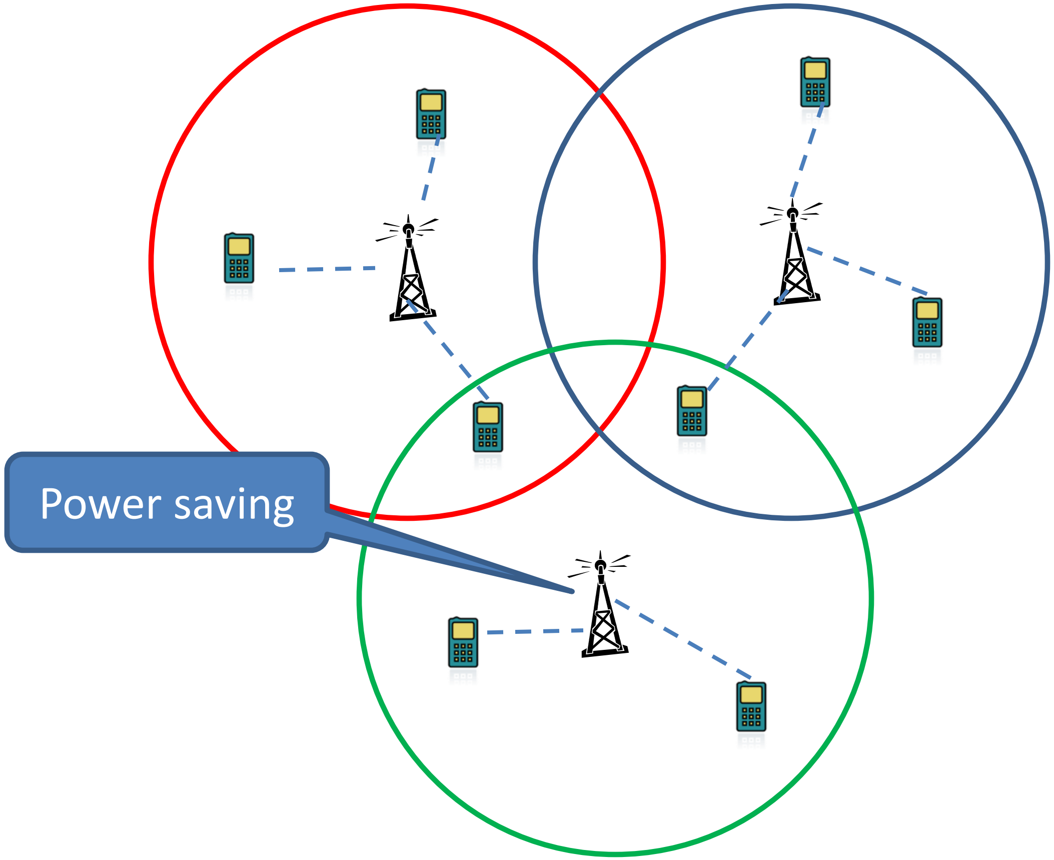


Does workload relocation help?



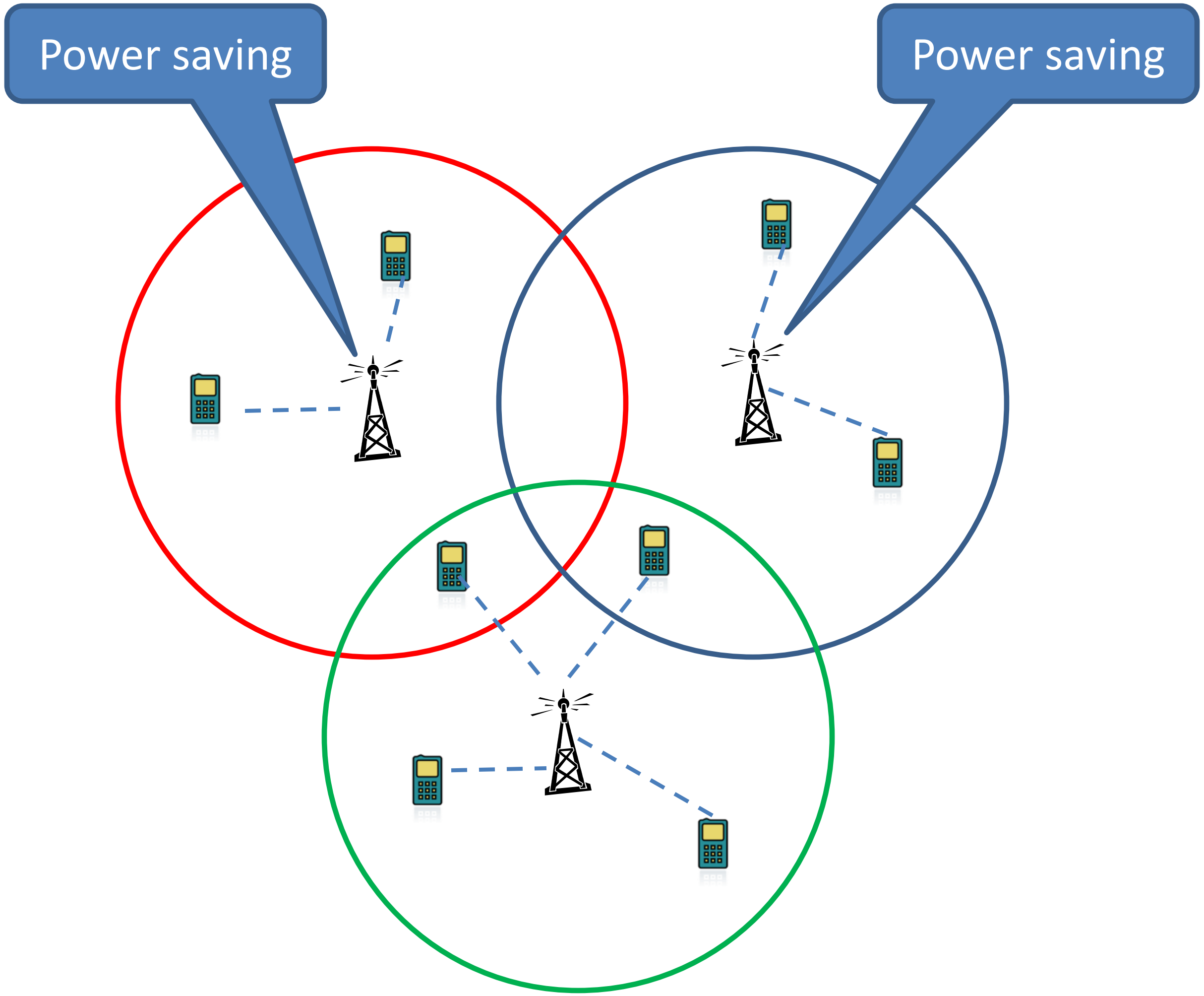
Assume that power saving is enabled if upto two calls are being served

Does workload relocation help?



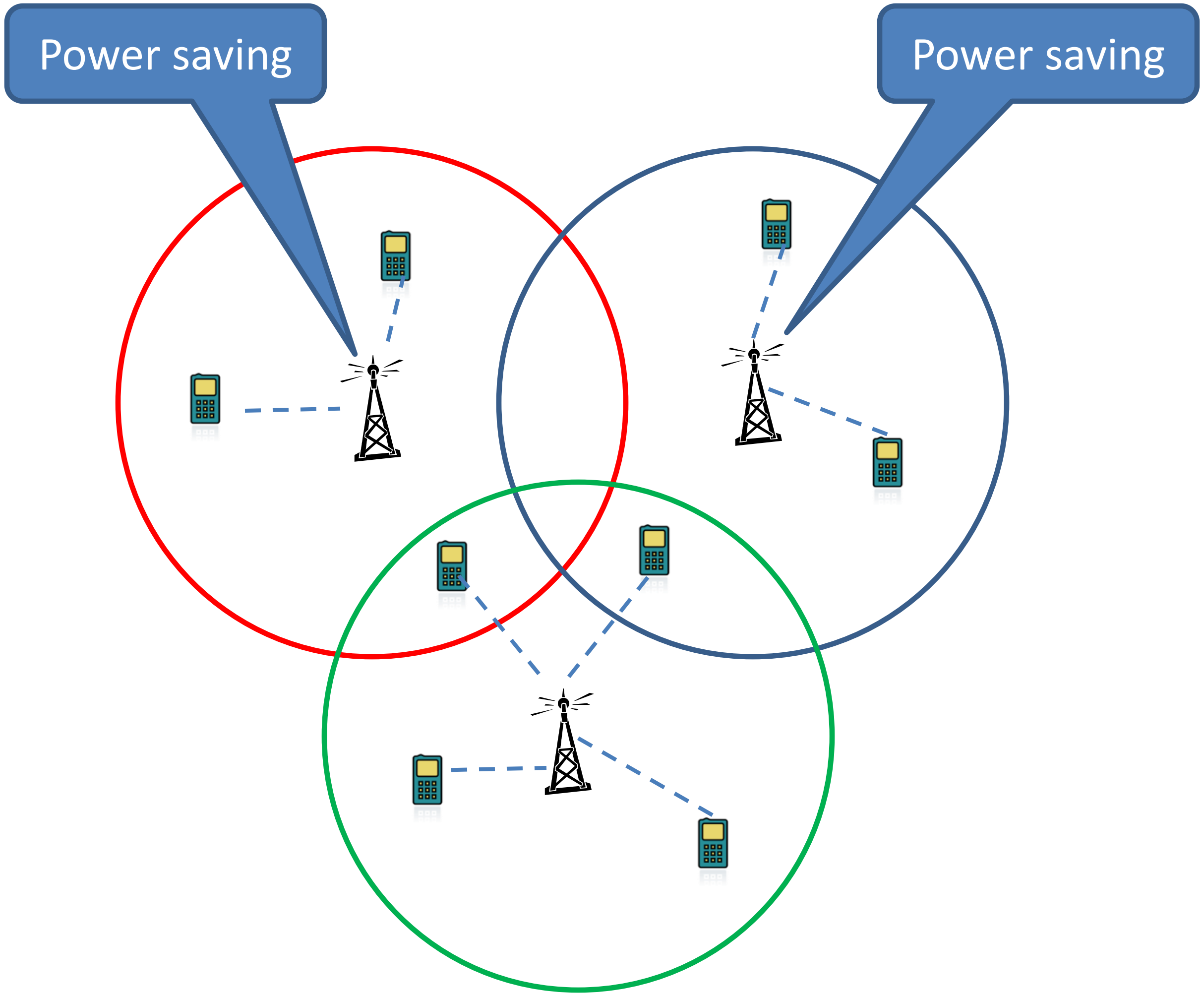
Assume that power saving is enabled if upto two calls are being served

Does workload relocation help?



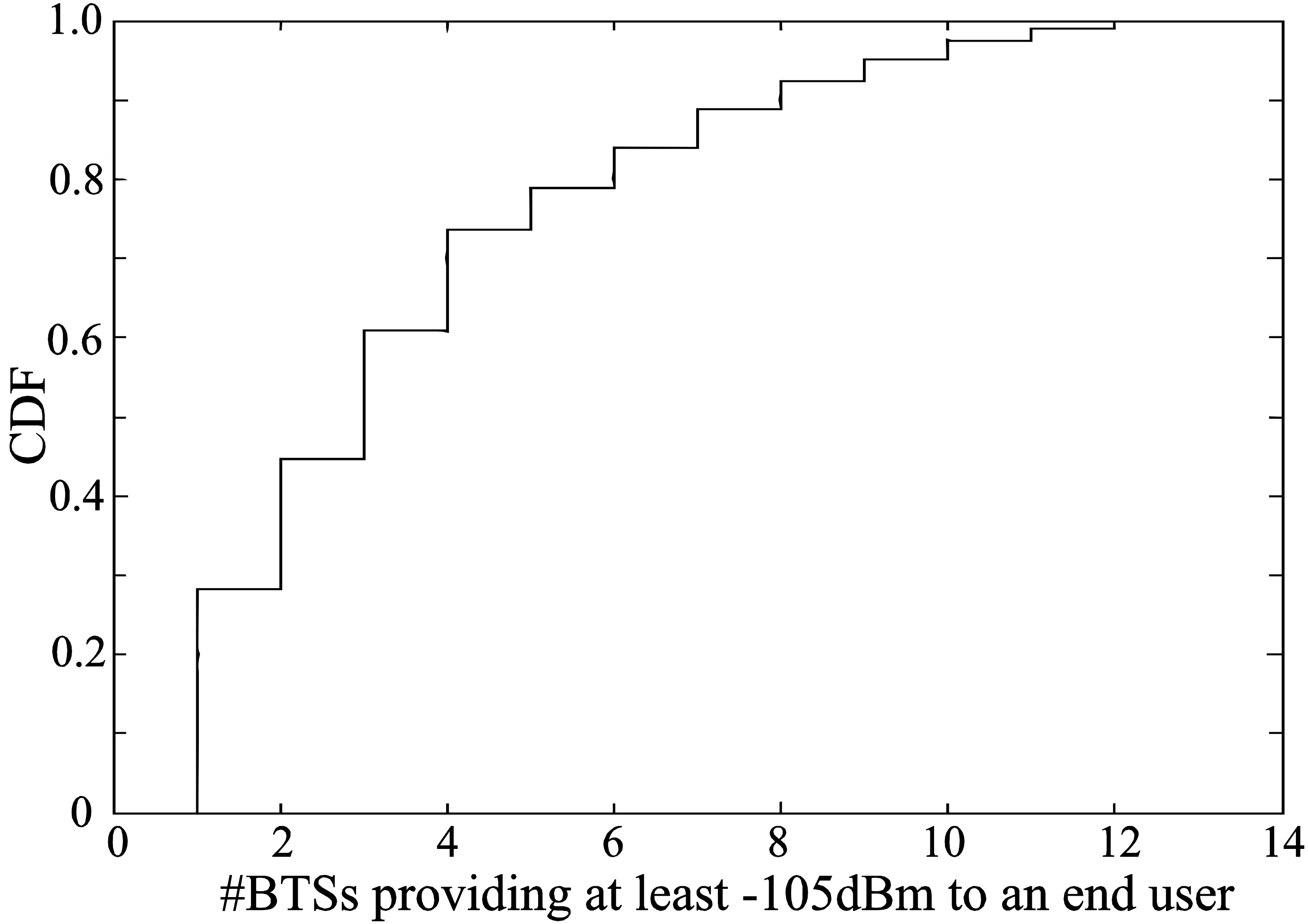
Assume that power saving is enabled if upto two calls are being served

Does workload relocation help?



Handing off some calls may enable greater power savings

Is Workload Relocation Possible?



Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource		

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource		Servers

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation		

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation		Client redirect

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning		

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning		Server shutdown / idle / hibernate

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning	BTS Power Saving	Server shutdown / idle / hibernate

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning	BTS Power Saving	Server shutdown / idle / hibernate
Transition costs		

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning	BTS Power Saving	Server shutdown / idle / hibernate
Transition costs		(De)activation overheads

Drawing Parallels With Case Study I

Parameter	Cellular network	Data centers
Network resource	TRX	Servers
Workload relocation	Call hand off	Client redirect
Resource pruning	BTS Power Saving	Server shutdown / idle / hibernate
Transition costs	Negligible	(De)activation overheads

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda (f + (1-f) \frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)$$

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n}\sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda(\text{[illegible]} 1 - f)\frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)$$

Optimization Formulation

$$\cancel{\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda (\ldots 1 - f) \frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)}$$

$$minimize \sum_{j=1}^{m} p_i^j$$

For every interval, minimize # TRXs

Optimization Formulation

$$\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda (\quad 1 - f) \frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)$$

$$minimize \sum_{j=1}^{m} p_i^j$$

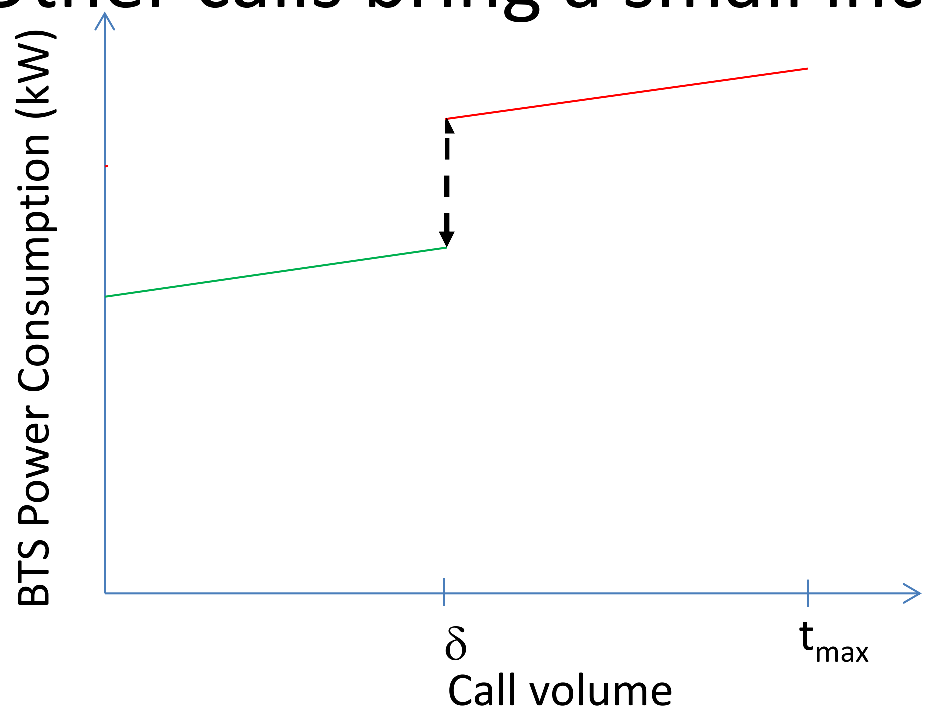
Optimization Formulation

$$\cancel{\text{minimize} \sum_{j=1}^{n} \sum_{i=1}^{m} c_i e_i^j (p_i^j \lambda (\cdots 1 - f) \frac{x_i^j}{c_i}) + b_i^j \sigma + s_i^j \delta)}$$

$$minimize \sum_{j=1}^{m} p_i^j$$

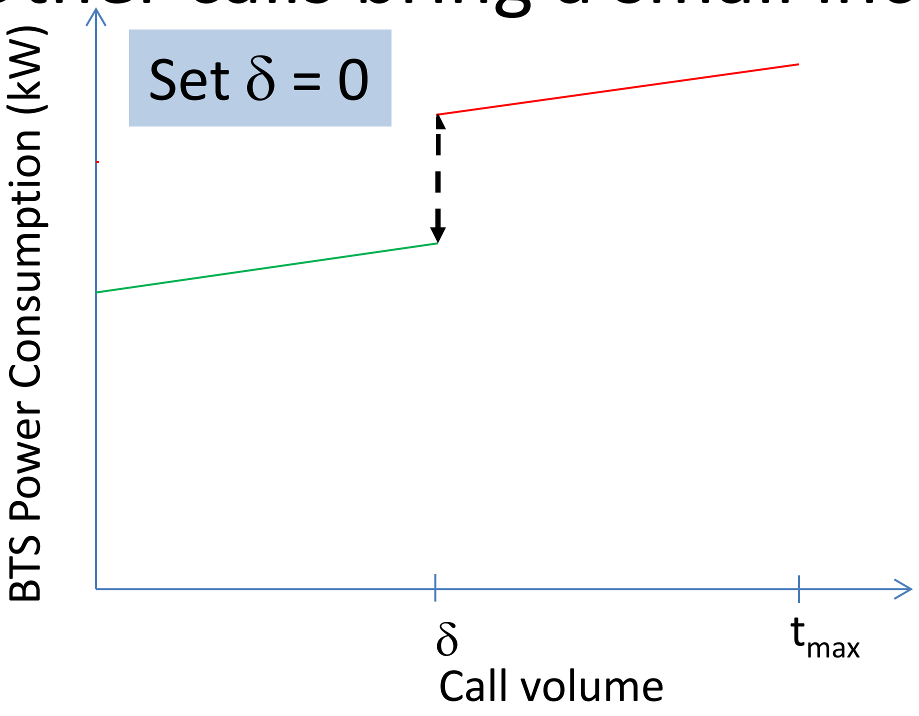
Complexity 1/2

- (δ + 1)th call incurs a sudden jump in power
- Other calls bring a small increase in power (ε)



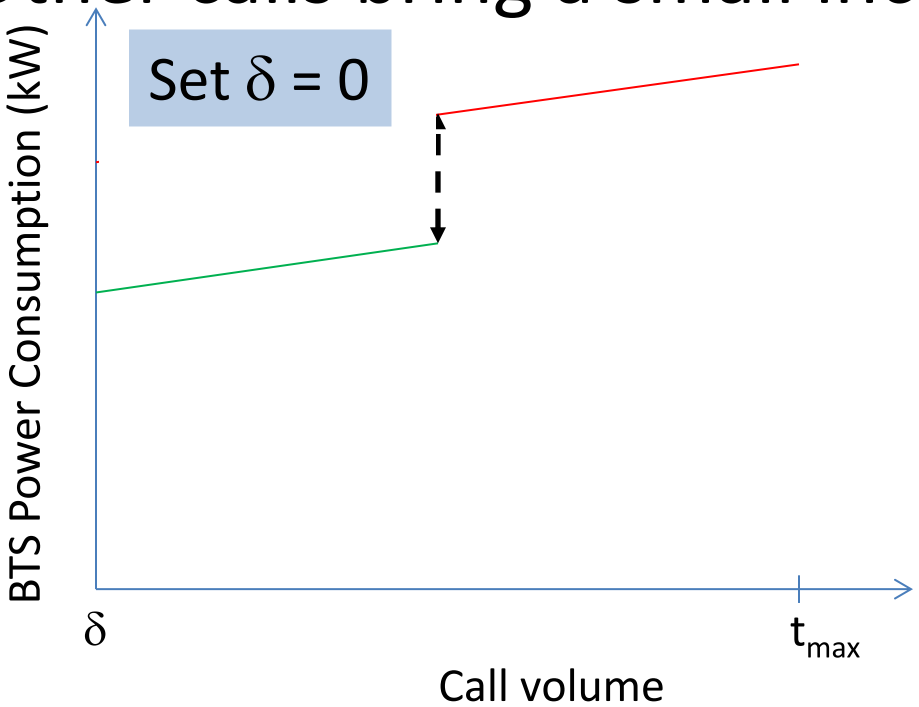
Complexity 1/2

- $(\delta + 1)$th call incurs a sudden jump in power
- Other calls bring a small increase in power (ε)



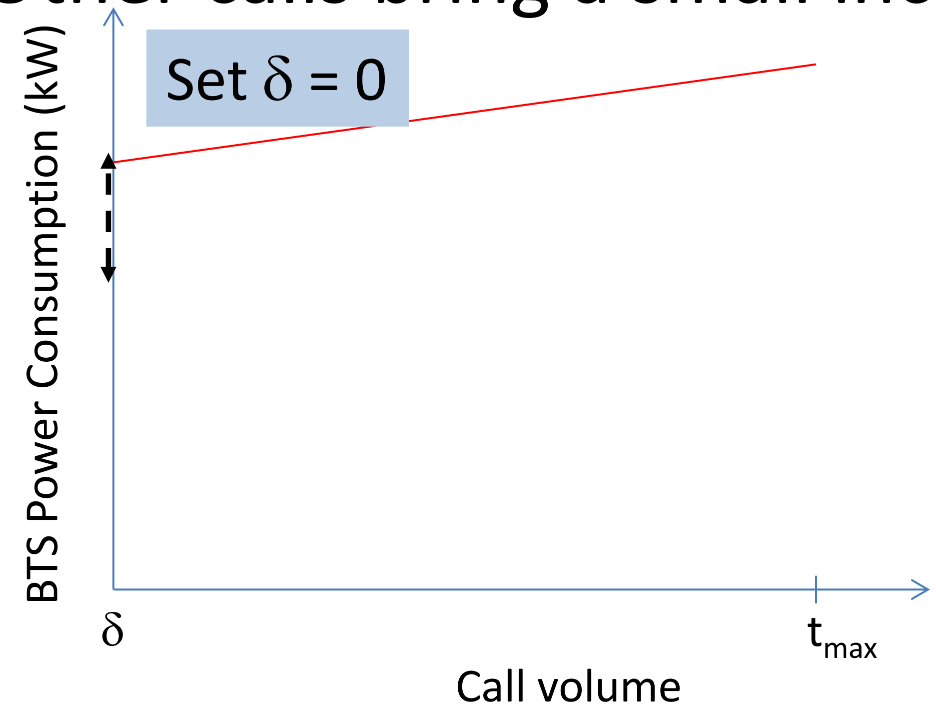
Complexity 1/2

- (δ + 1)th call incurs a sudden jump in power
- Other calls bring a small increase in power (ε)

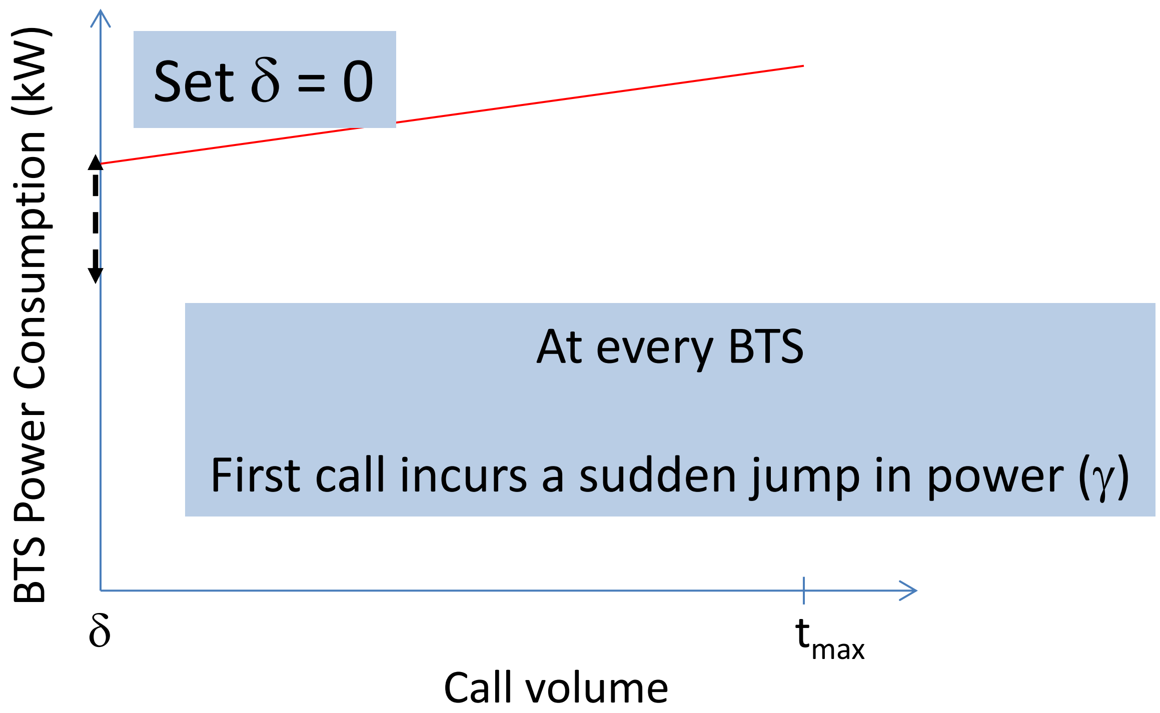


Complexity 1/2

- (δ + 1)th call incurs a sudden jump in power
- Other calls bring a small increase in power (ε)



Complexity 1/2

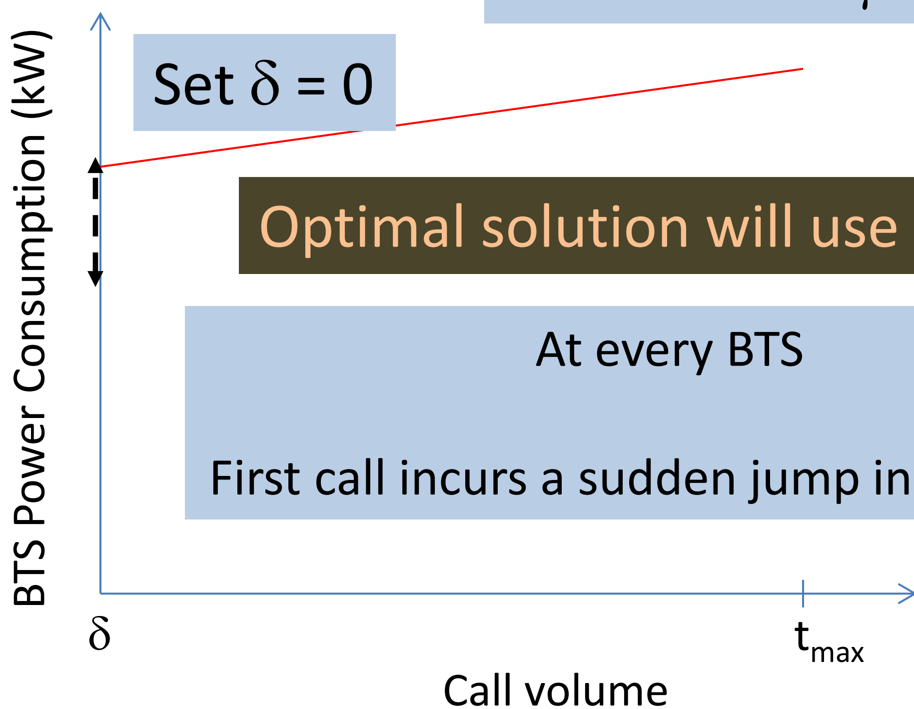


Complexity 1/2

BTS Cost:

0 calls: 0

n calls: $\gamma + n\varepsilon$



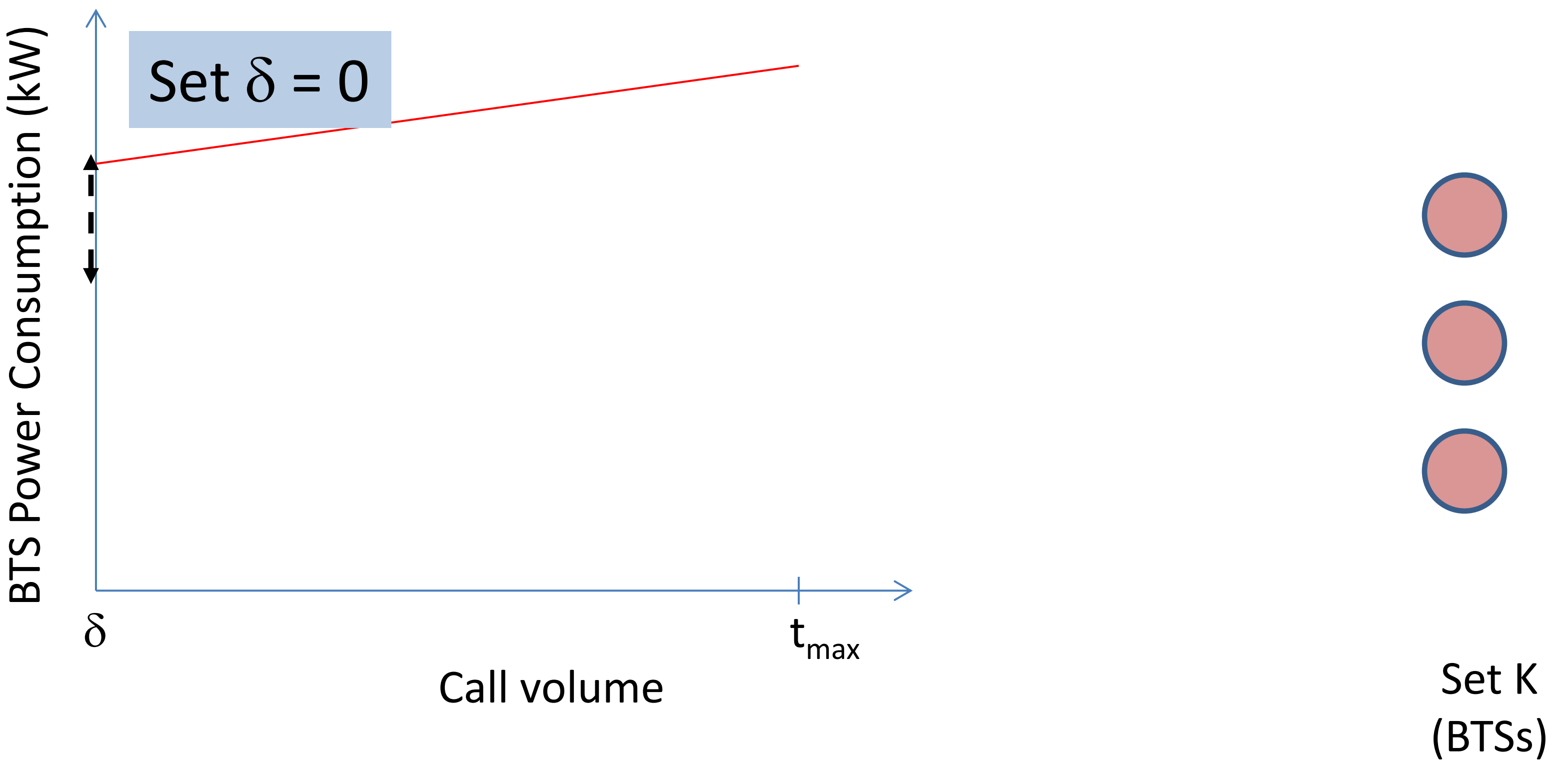
Set $\delta = 0$

Optimal solution will use fewest possible BTSs

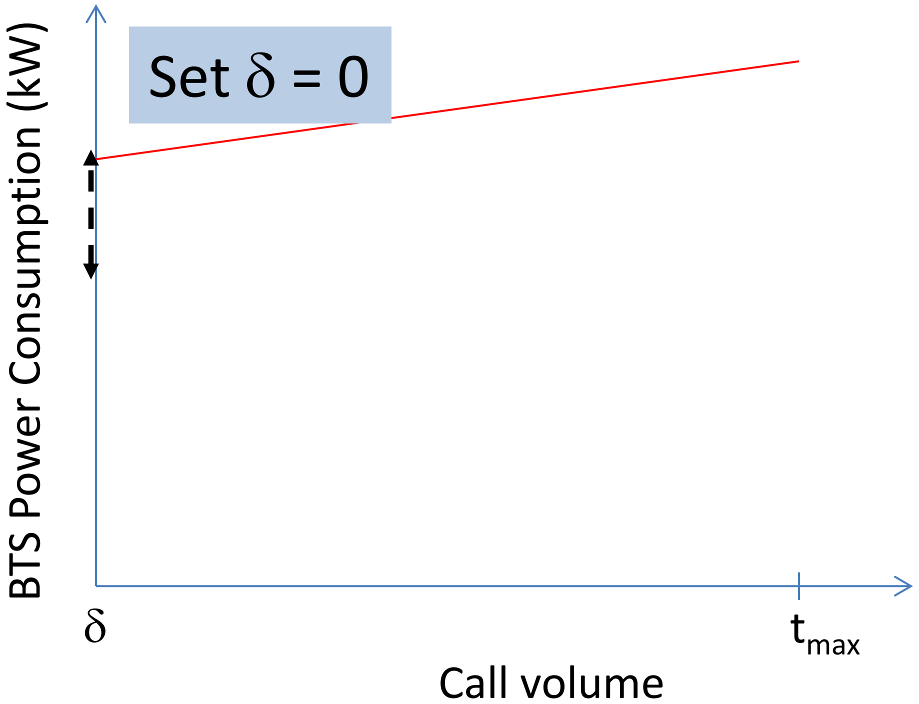
At every BTS

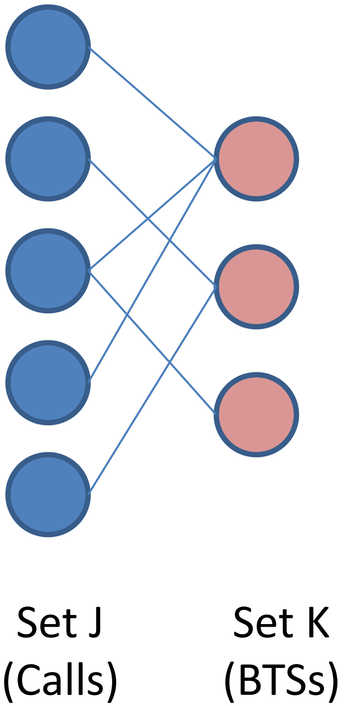
First call incurs a sudden jump in power (γ)

Complexity 1/2

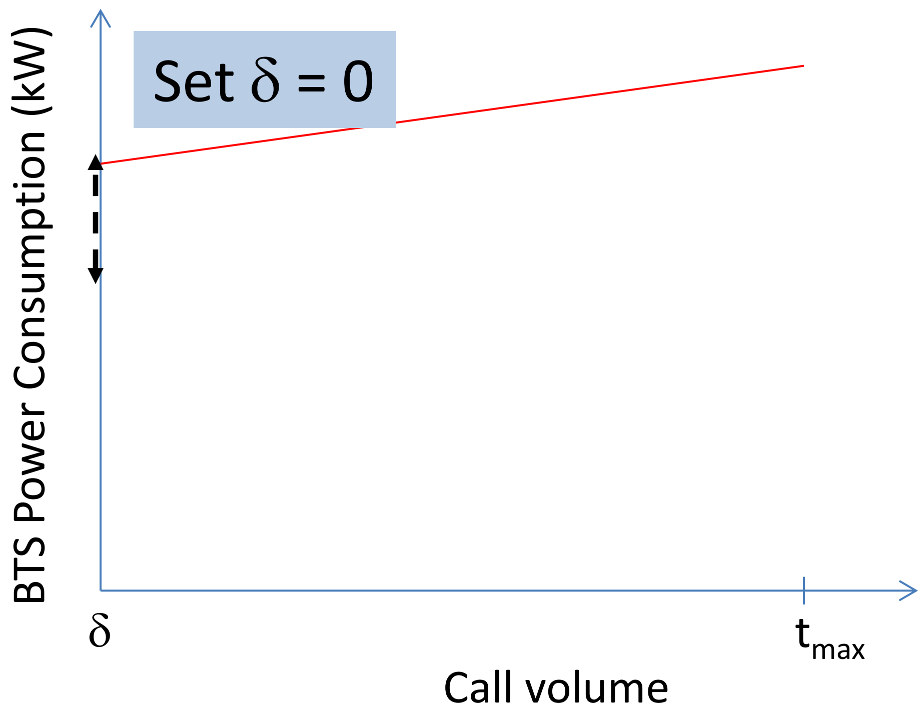


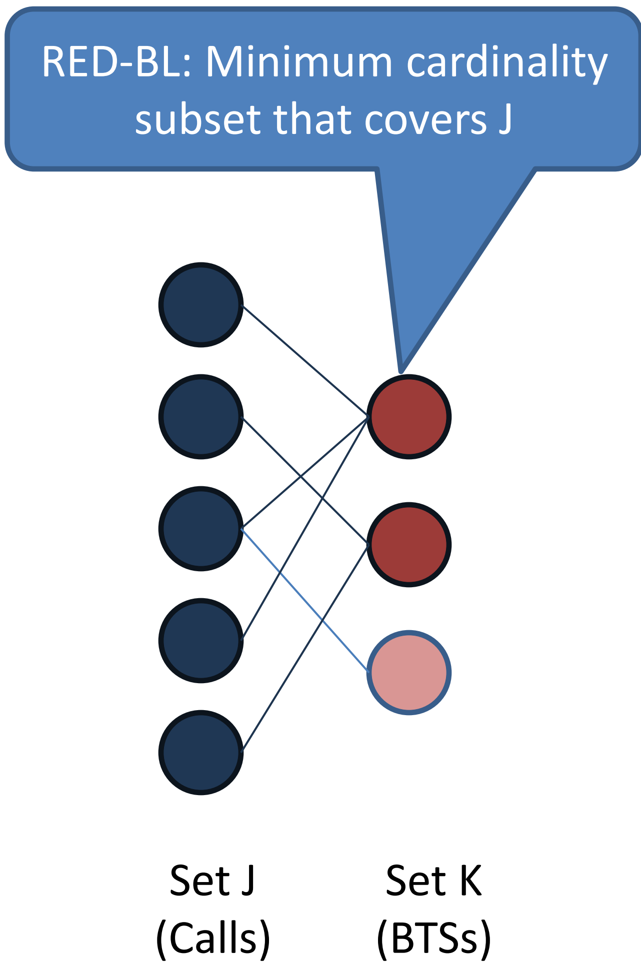
Complexity 1/2



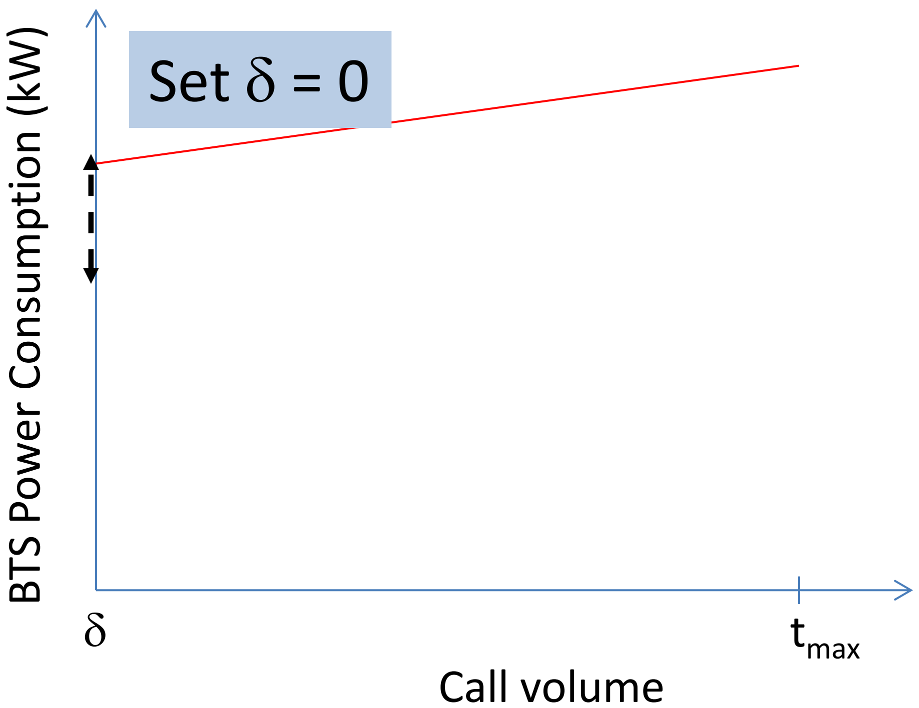


Complexity 1/2

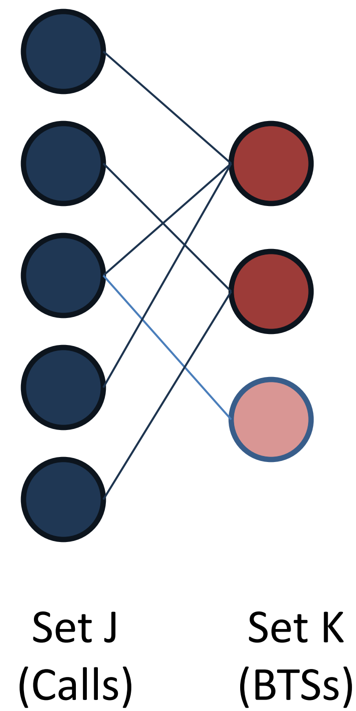




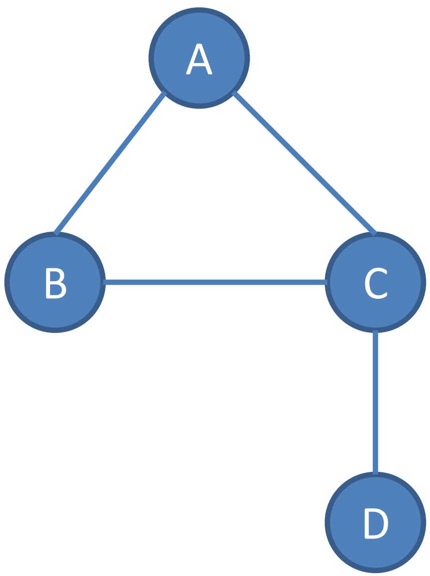
Complexity 1/2



Call this graph problem:
Modified MDS

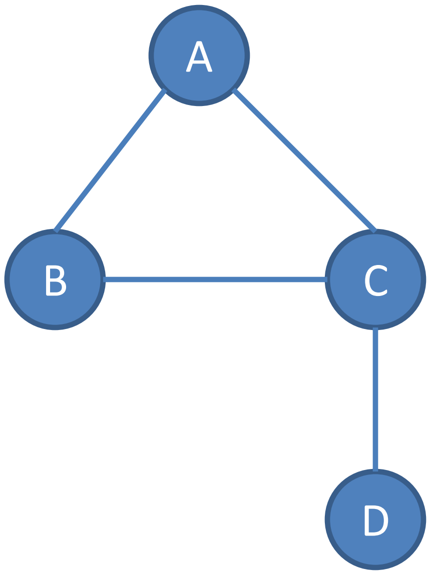


Complexity 2/2



Minimum Dominating Set (MDS)

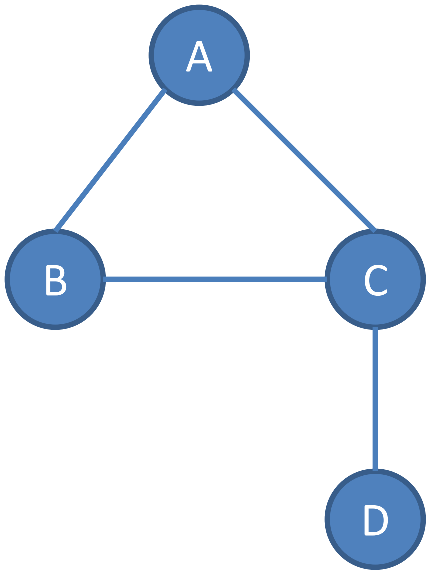
Complexity 2/2



Minimum Dominating Set (MDS)

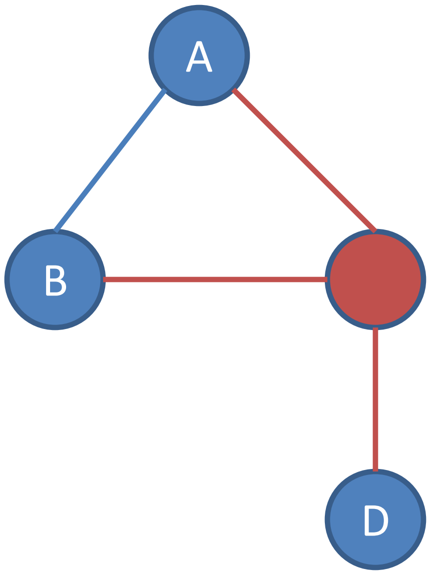
Smallest subset V’ of vertices V such that every other vertex is adjacent to at least one vertex in V’

Complexity 2/2



Minimum Dominating Set (MDS)

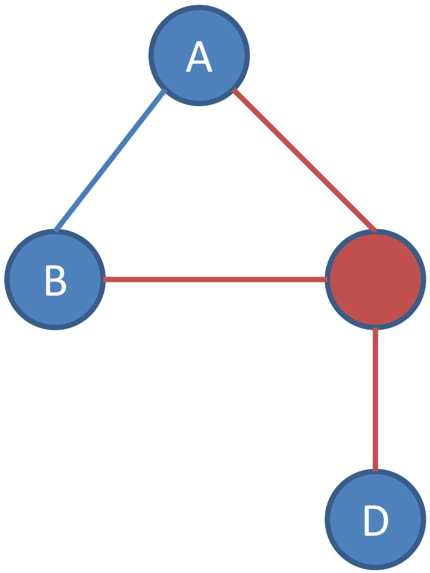
Complexity 2/2



Minimum Dominating Set (MDS)

Complexity 2/2

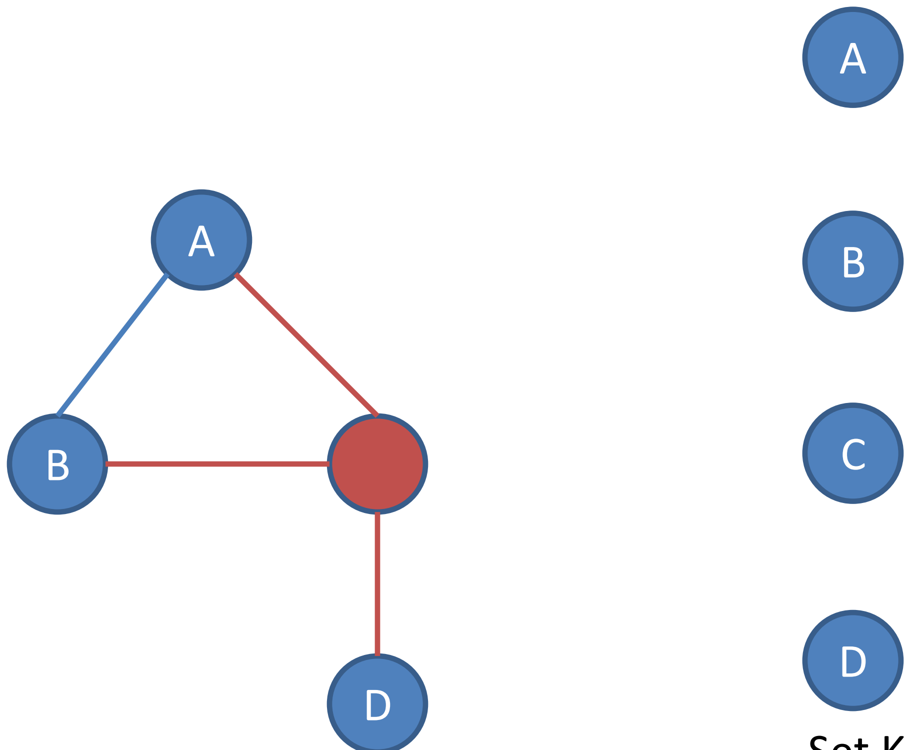
1. Make two copies of vertices



Minimum Dominating Set (MDS)

Complexity 2/2

1. Make two copies of vertices

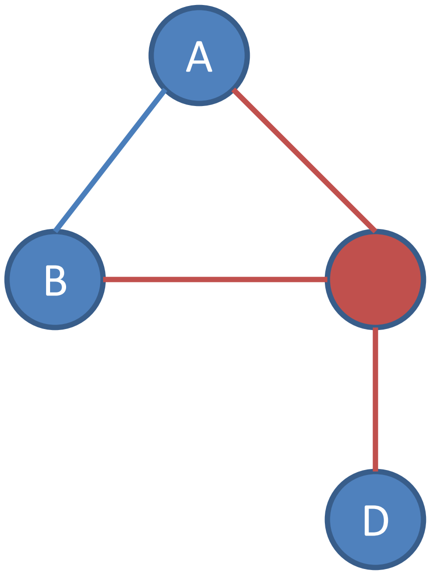


Minimum Dominating Set (MDS)

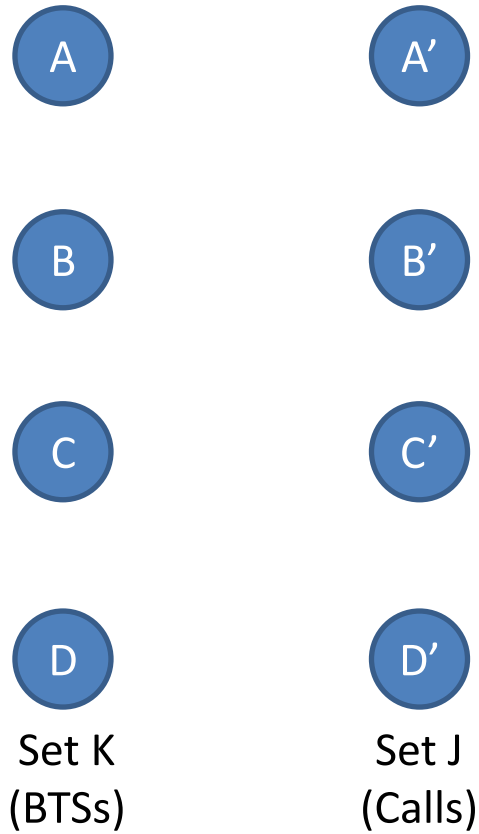
Set K
(BTSs)

Complexity 2/2

1. Make two copies of vertices

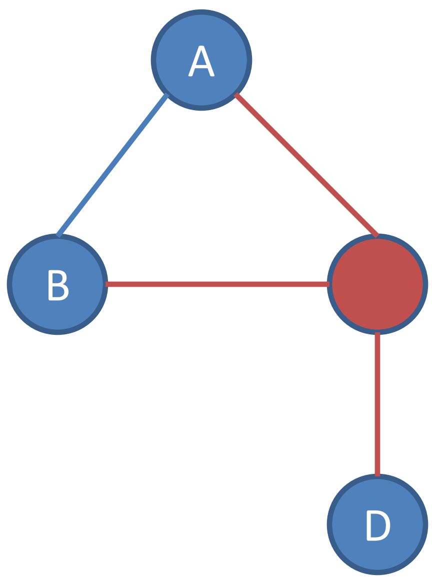


Minimum Dominating Set (MDS)

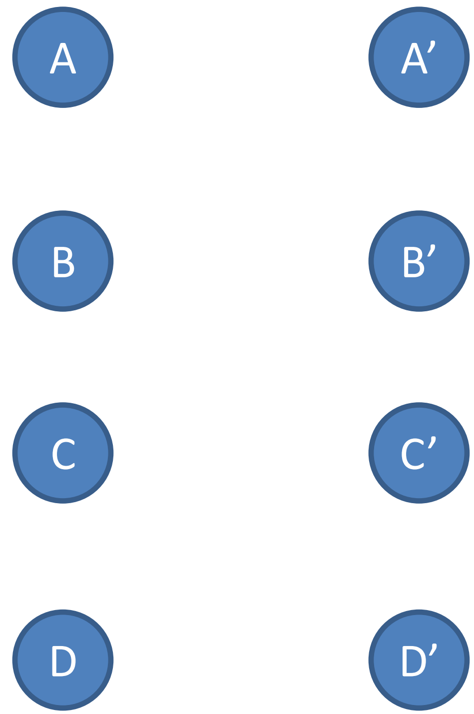


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy



Minimum Dominating Set (MDS)

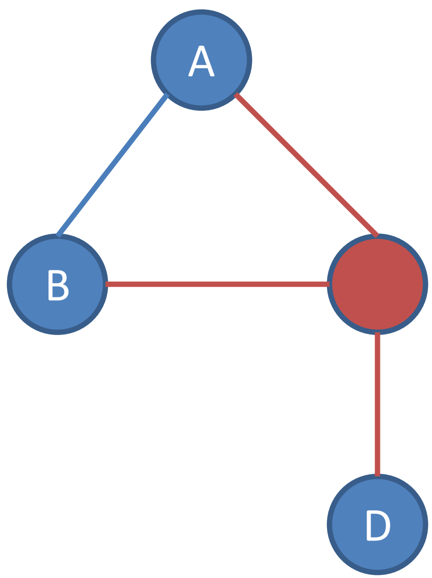


Set K (BTSs)

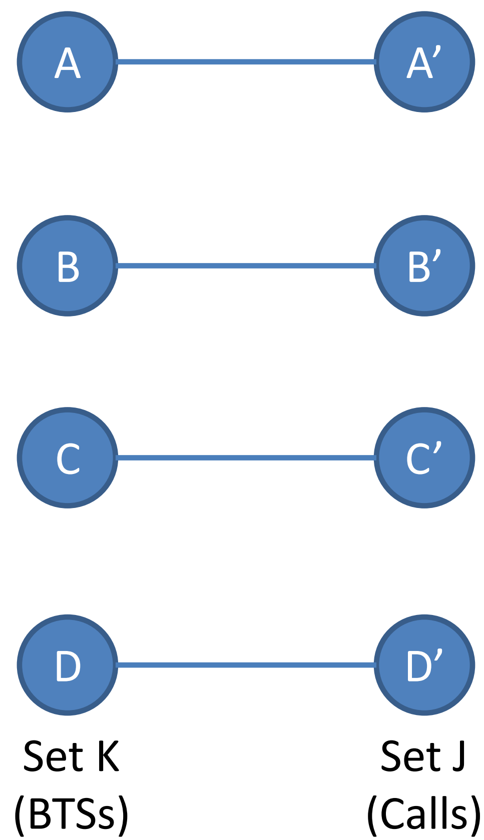
Set J (Calls)

Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy

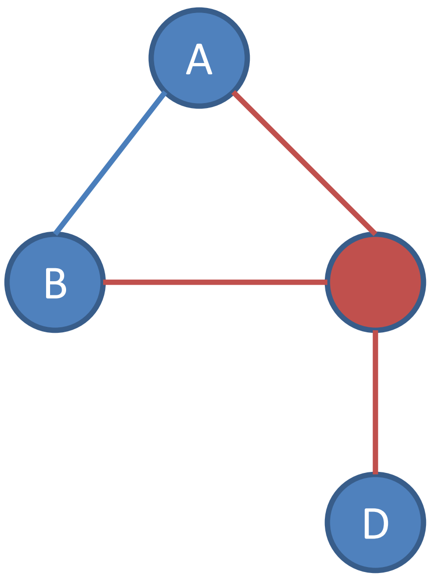


Minimum Dominating Set (MDS)

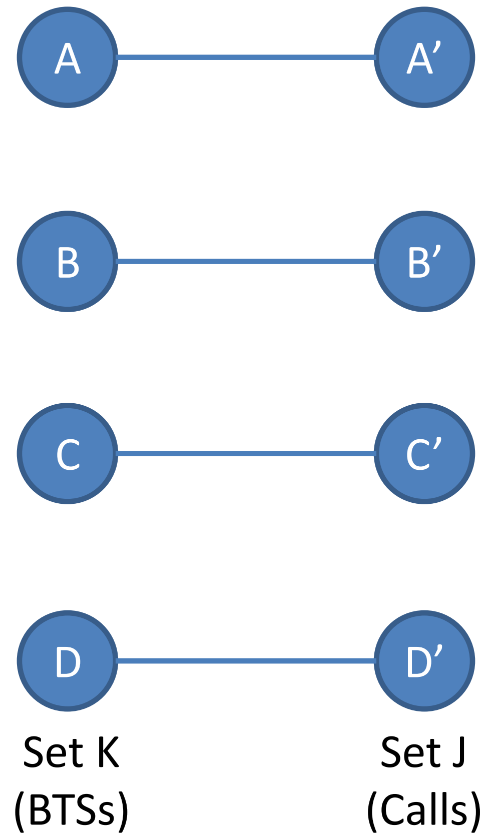


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph

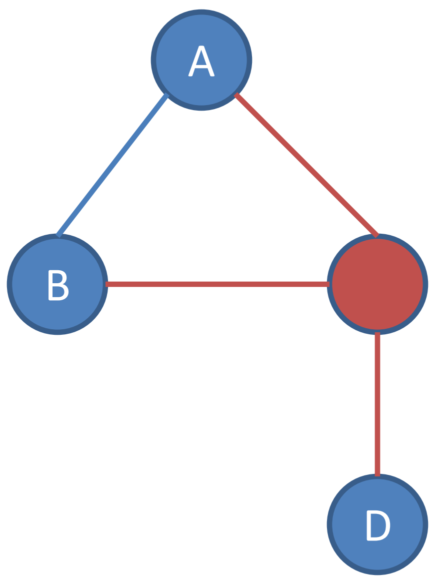


Minimum Dominating Set (MDS)

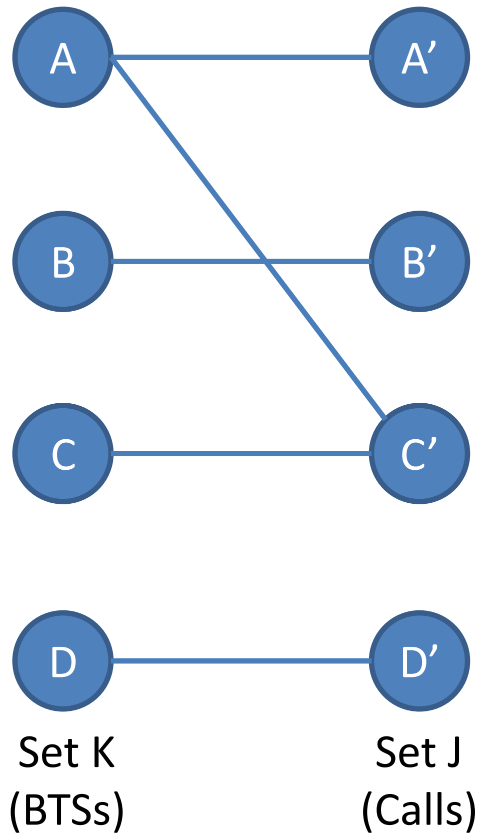


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph

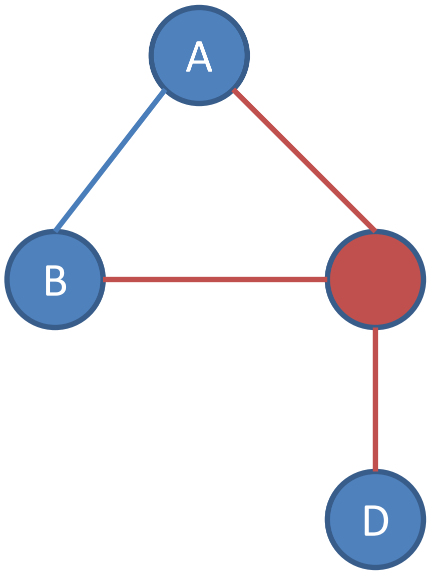


Minimum Dominating Set (MDS)

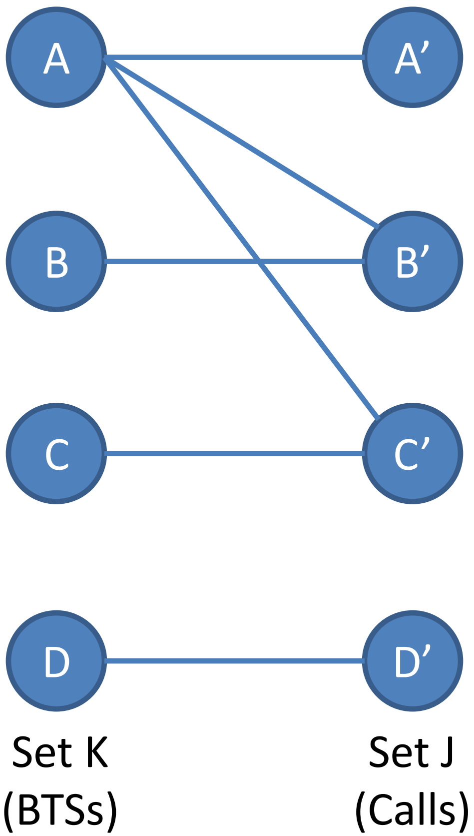


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph

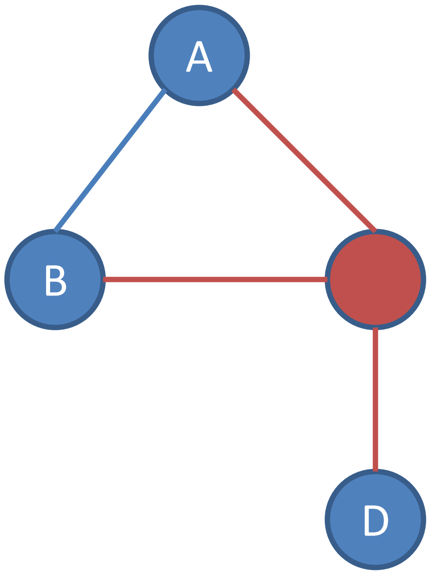


Minimum Dominating Set (MDS)

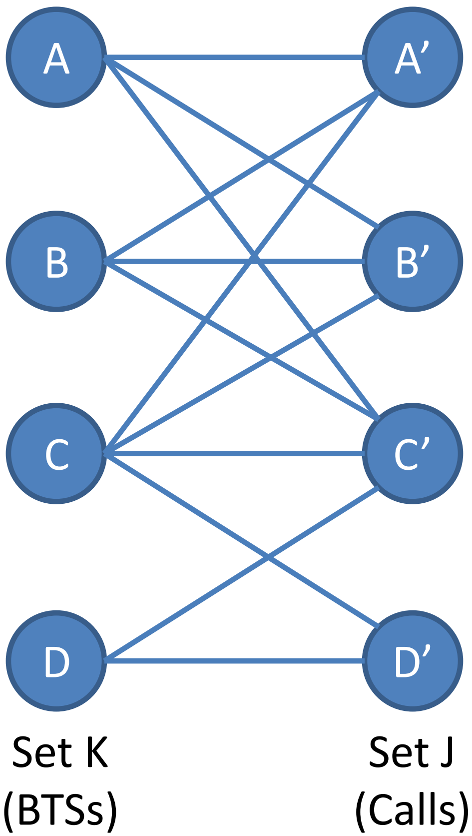


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph

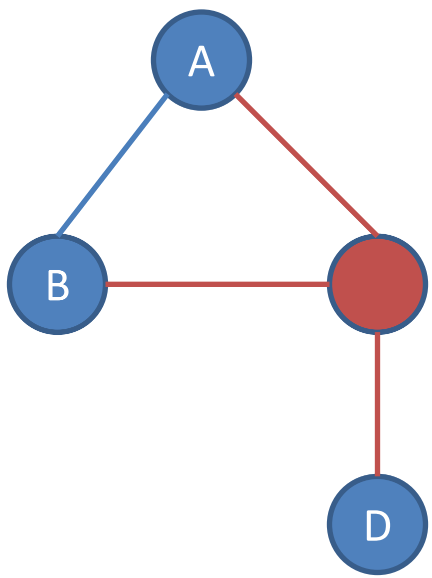


Minimum Dominating Set (MDS)

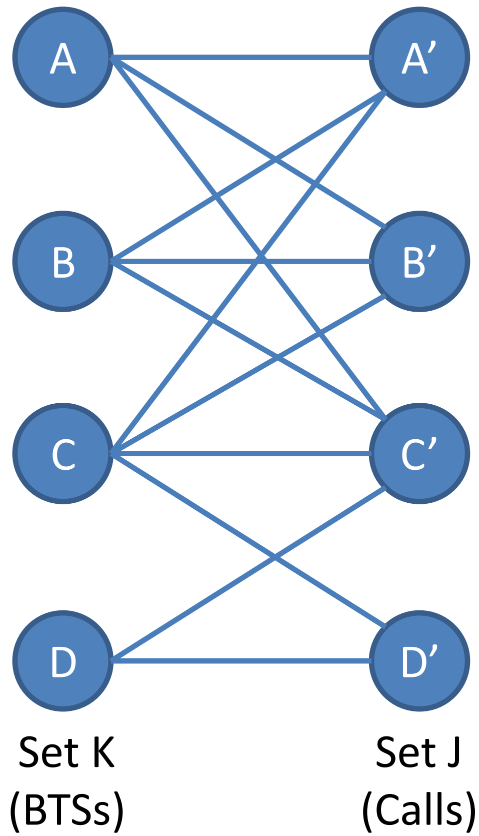


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph
4. Solve Modified MDS on new graph

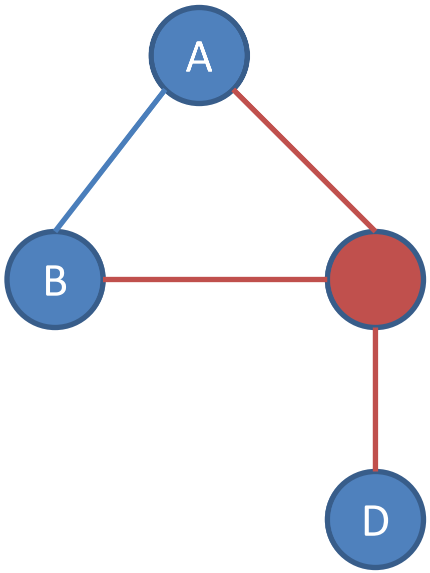


Minimum Dominating Set (MDS)

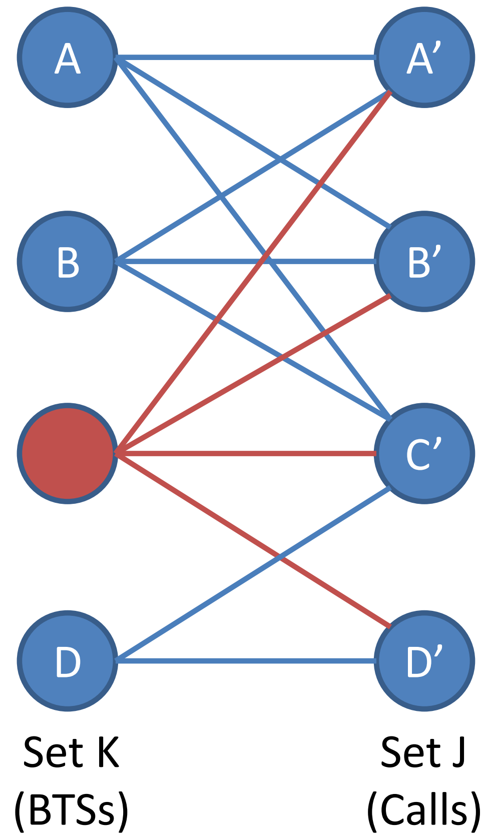


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph
4. Solve Modified MDS on new graph

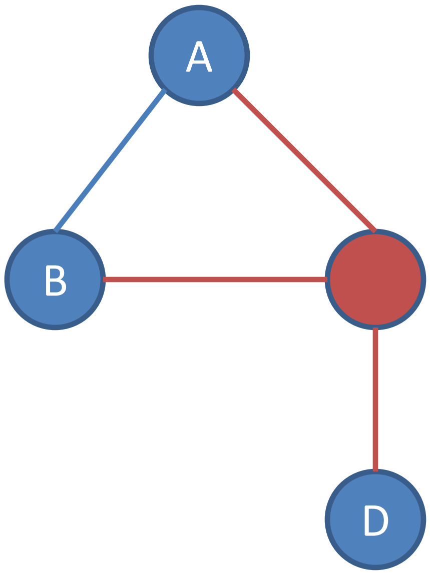


Minimum Dominating Set (MDS)

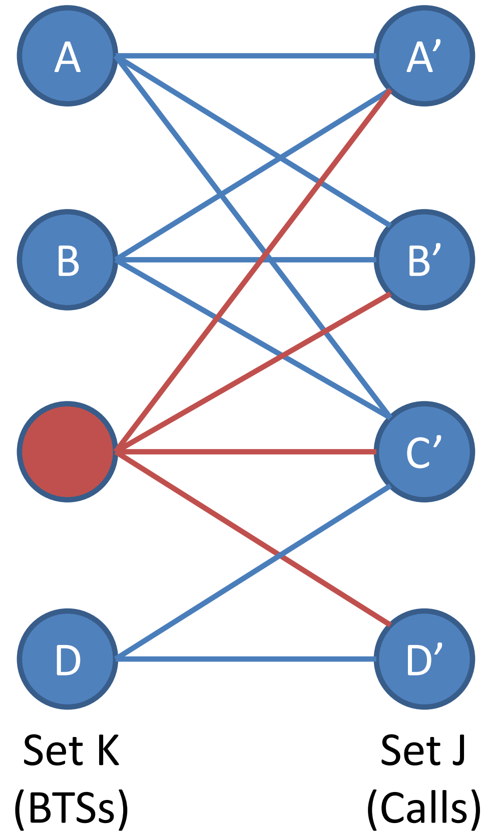


Complexity 2/2

1. Make two copies of vertices
2. Place edges between each vertex and its copy
3. Replicate other edges in original graph
4. Solve Modified MDS on new graph

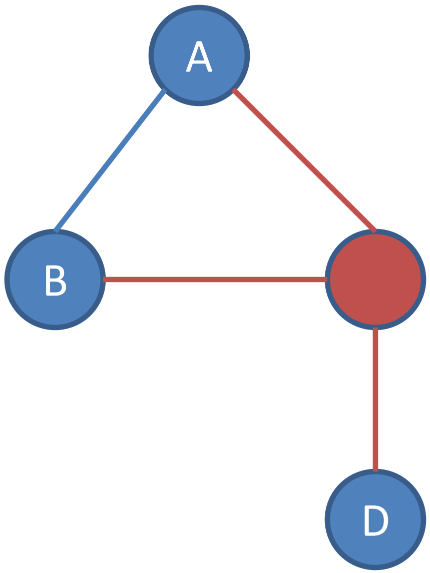


Minimum Dominating Set (MDS)

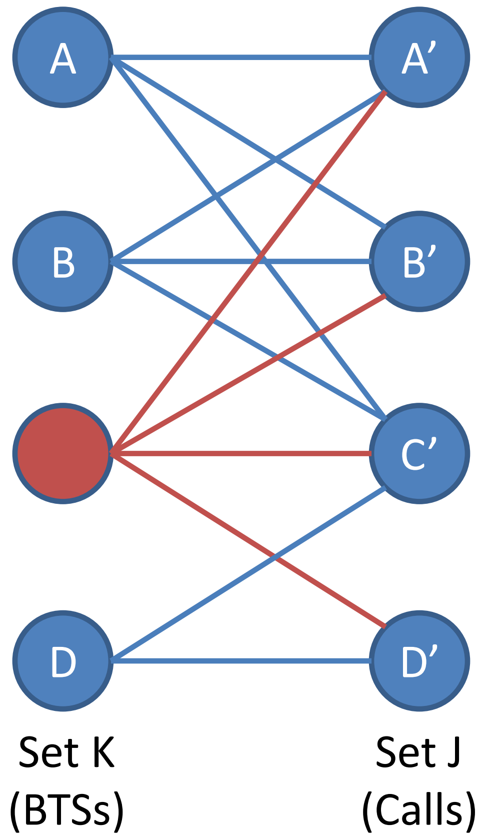


MDS Solved!

Complexity 2/2



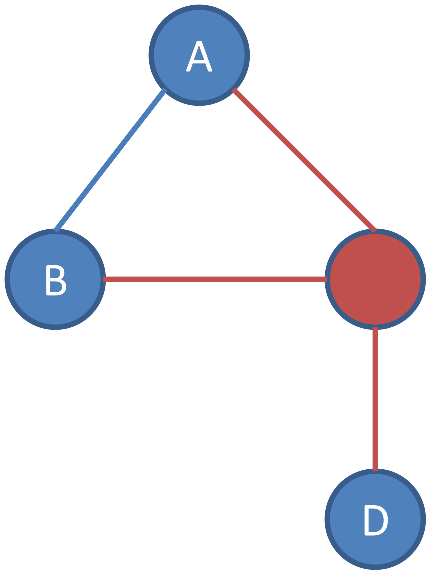
Minimum Dominating Set (MDS)



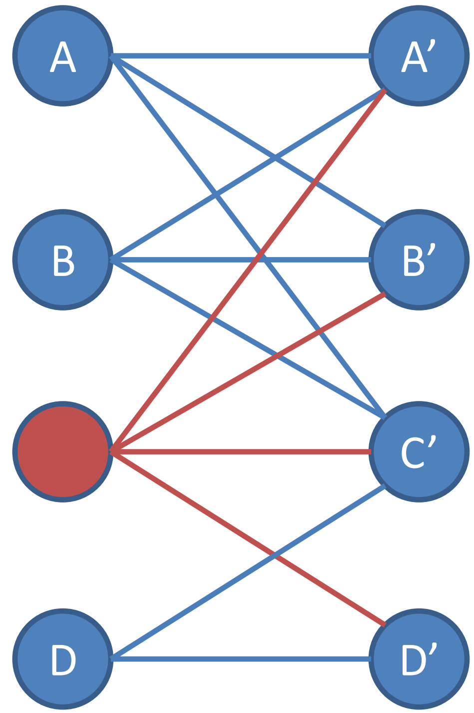
MDS Solved!

Complexity 2/2

RED-BL for cellular



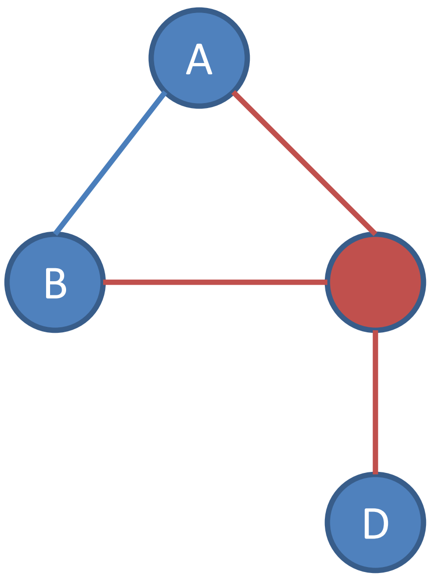
Minimum Dominating Set (MDS)



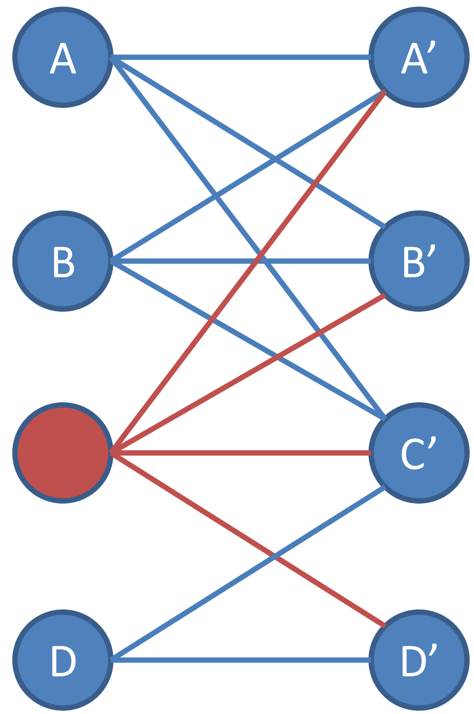
MDS Solved!

Complexity 2/2

RED-BL for cellular solves Modified MDS



Minimum Dominating Set (MDS)



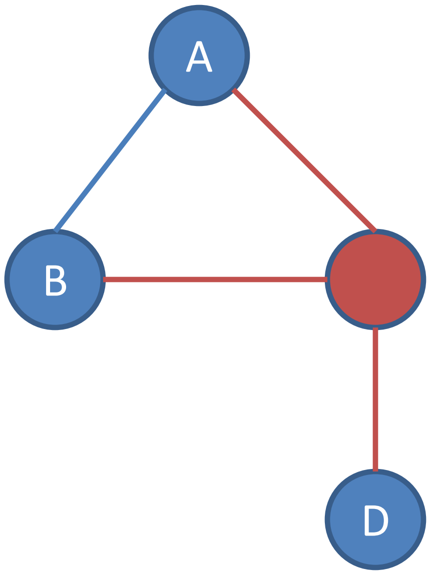
Set K
(BTSs)

Set J
(Calls)

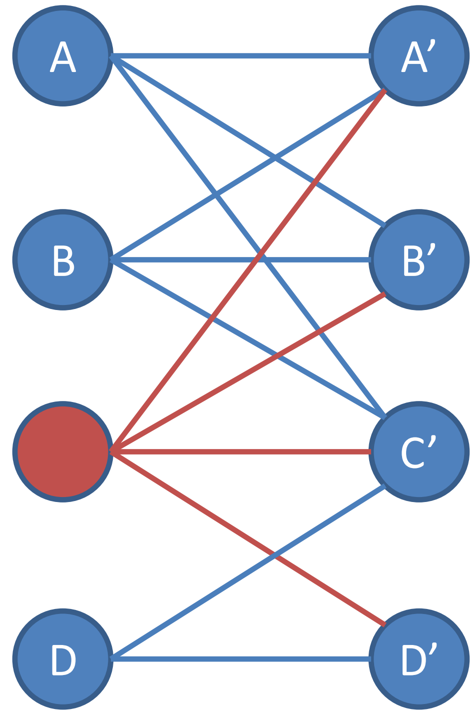
MDS Solved!

Complexity 2/2

RED-BL for cellular solves Modified MDS solves MDS



Minimum Dominating Set (MDS)



Set K
(BTSs)

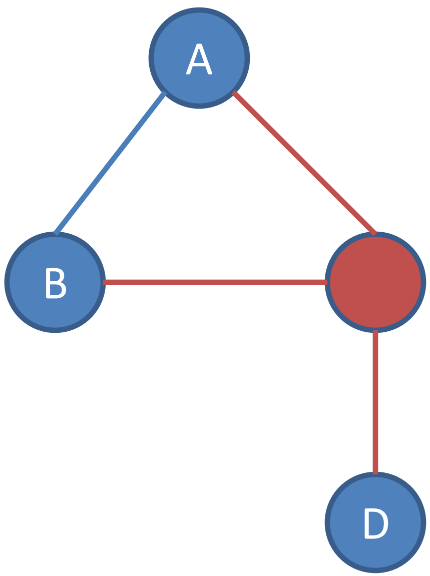
Set J
(Calls)

MDS Solved!

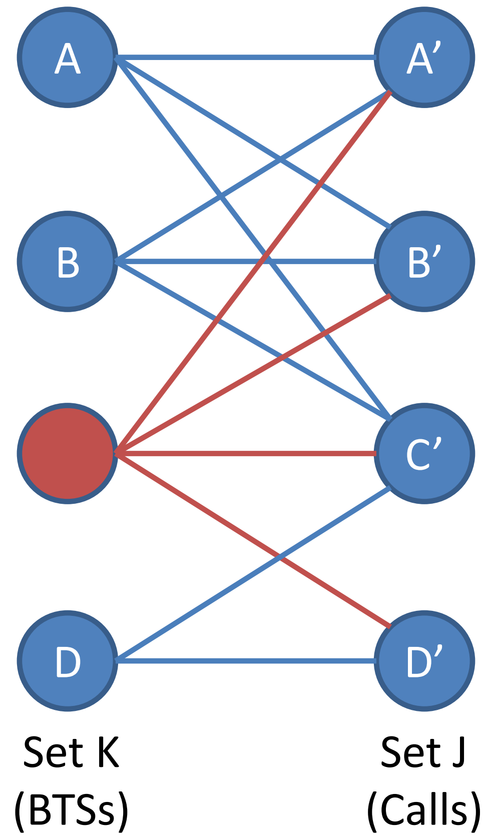
Complexity 2/2

RED-BL for cellular solves Modified MDS solves MDS

NP-Hard



Minimum Dominating Set (MDS)



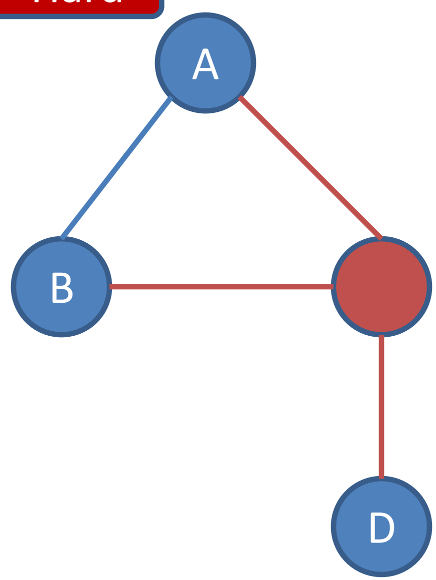
MDS Solved!

Complexity 2/2

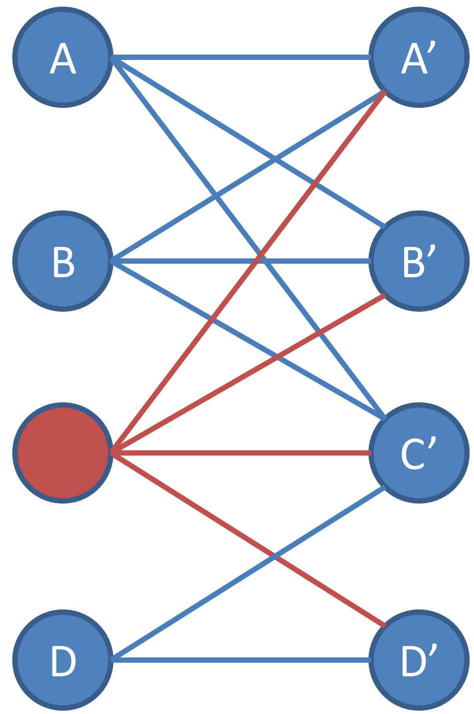
RED-BL for cellular solves Modified MDS solves MDS

NP-Hard

NP-Hard



Minimum Dominating Set (MDS)



Set K
(BTSs)

Set J
(Calls)

MDS Solved!

Experimental Setup

Experimental Setup

- Call volume traces for 2 days at 26 urban BTSs

Experimental Setup

- Call volume traces for 2 days at 26 urban BTSs
- Trace driven simulation:
  – Periodically obtain optimal call placement
  – Place BTSs with low-traffic in power-saving mode

BTS Power Consumption Models

Parameter	Value		
	Model 1	Model 2	Model 3
Idle Power (W)	1425	2401.8	2341.5
Peak Power (W)	1500	3887.5	2973.9
Power Saving per TRX (W)	20	50	100

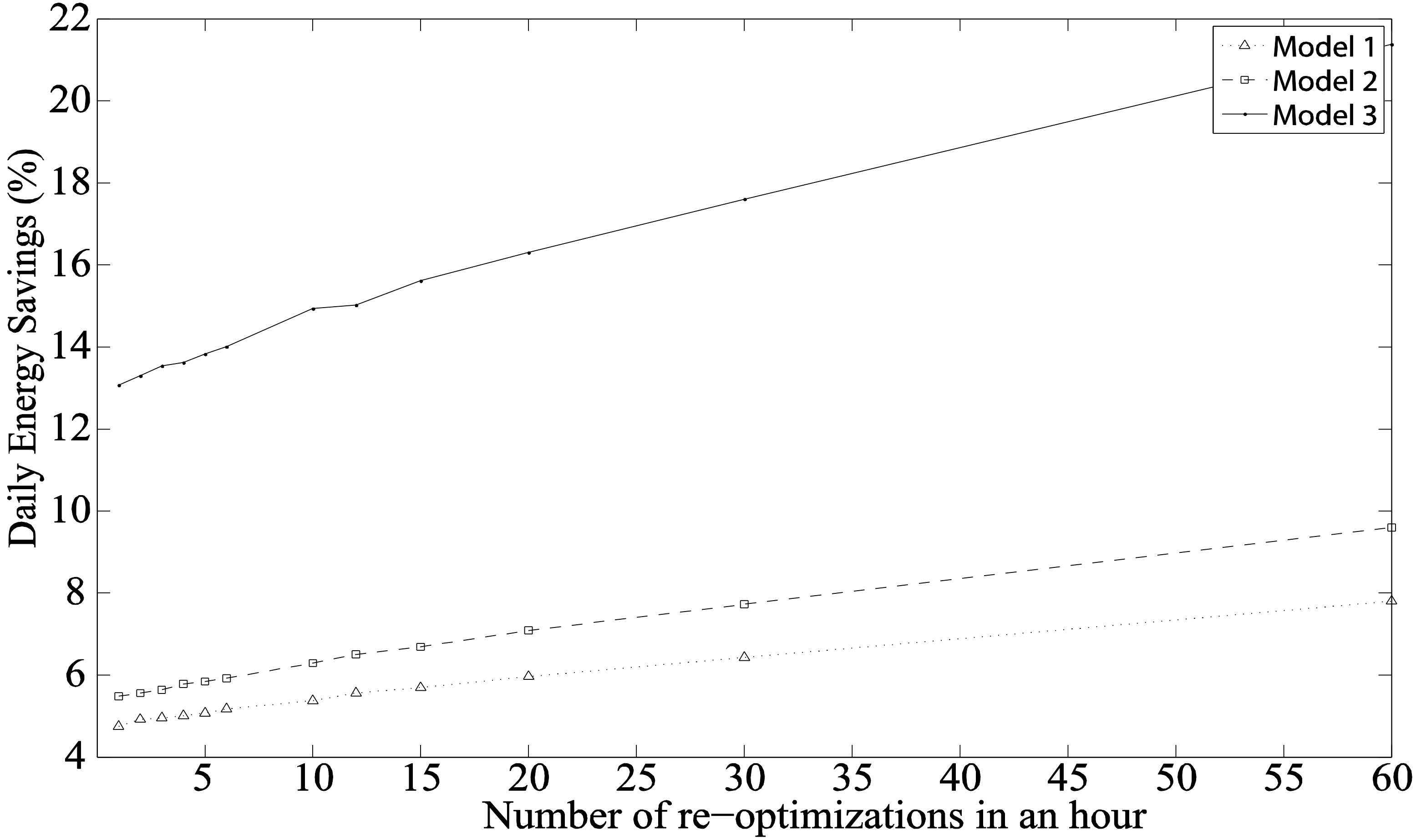
Results: Power-Saving Feature Only

Energy savings	Model 1	Model 2	Model 3
Percentage	4.73%	5.43%	12.89%
Daily energy savings (kWh)	43.28	109.68	217.12
Country-wide daily savings -31000 sites (MWh)	51.6	130.77	258.87

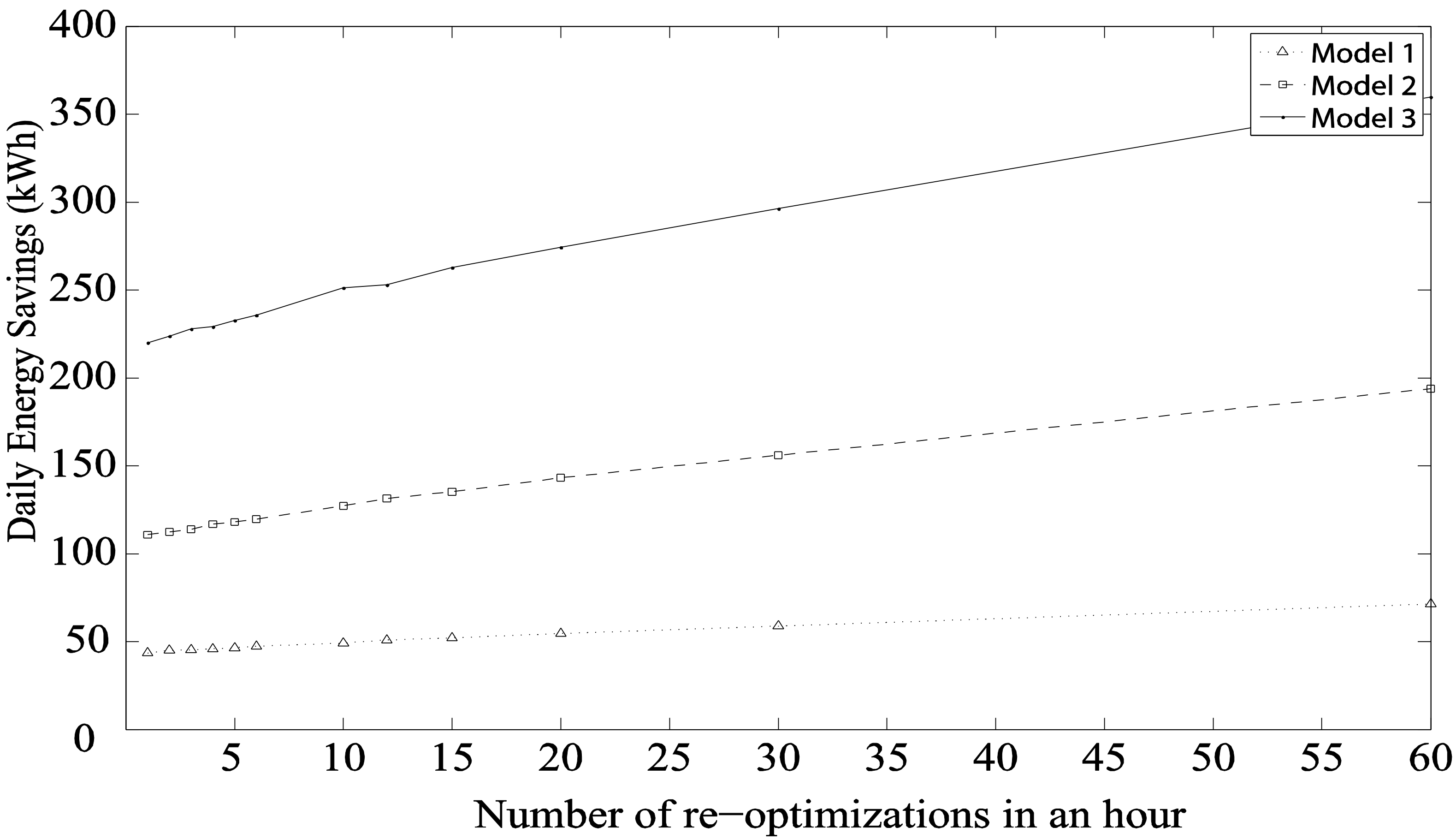
Results: Power-Saving Feature Only

Energy savings	Model 1	Model 2	Model 3
Percentage	4.73%	5.43%	12.89%
Daily energy savings (kWh)	43.28	109.68	217.12
Country-wide daily savings -31000 sites (MWh)	51.6	130.77	258.87

Results: Power-Saving + Handoff Absolute Energy Savings (%)



Results: Power-Saving + Handoff Absolute Energy Savings (kWh)



Effect of Granular Deactivation

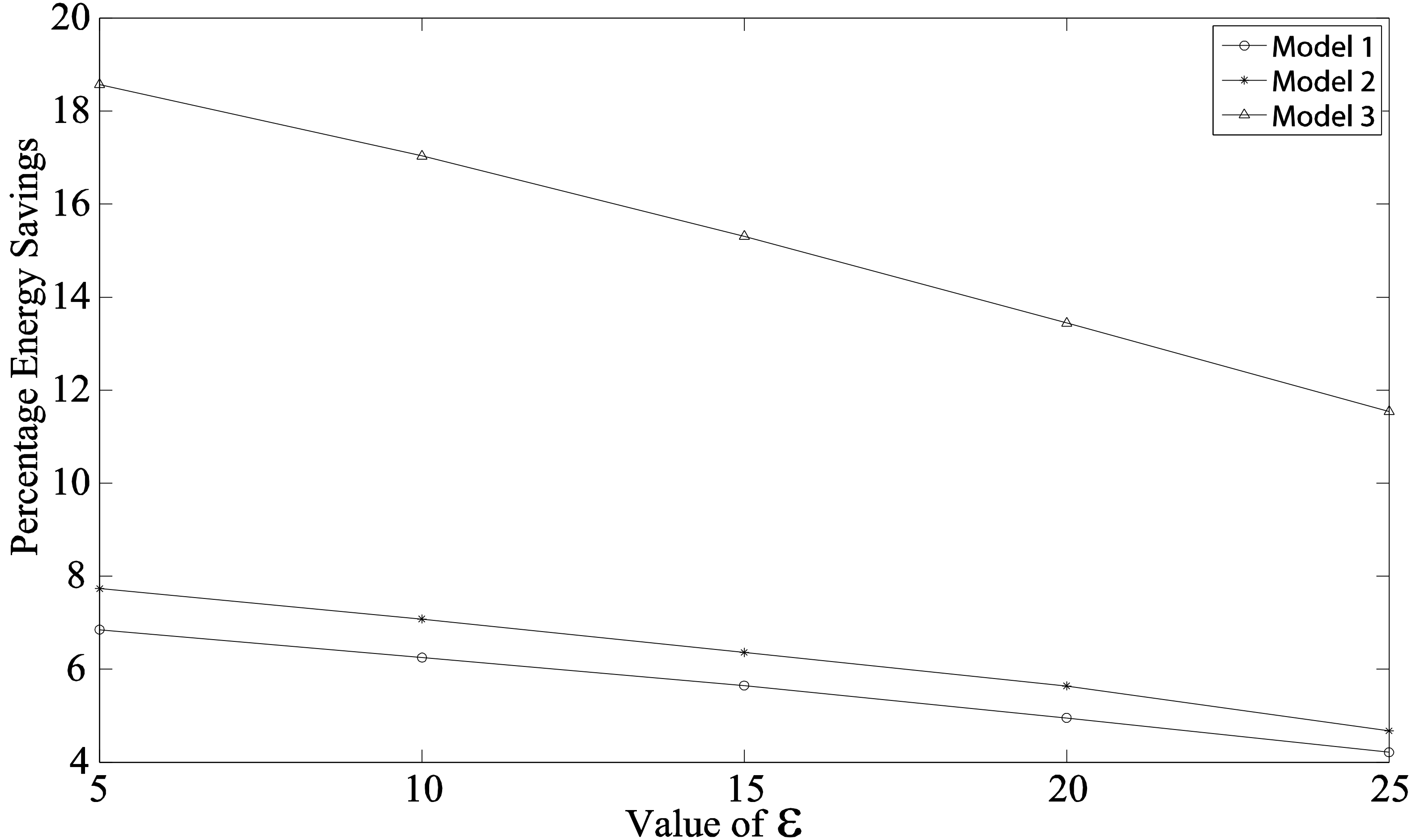
Granularity	Model 1	Model 2	Model 3
2-state	5.38%	6.29%	14.94%
3-state	6.81%	7.73%	18.62%
6-state	8.70%	9.65%	23.37%

Effect of Granular Deactivation

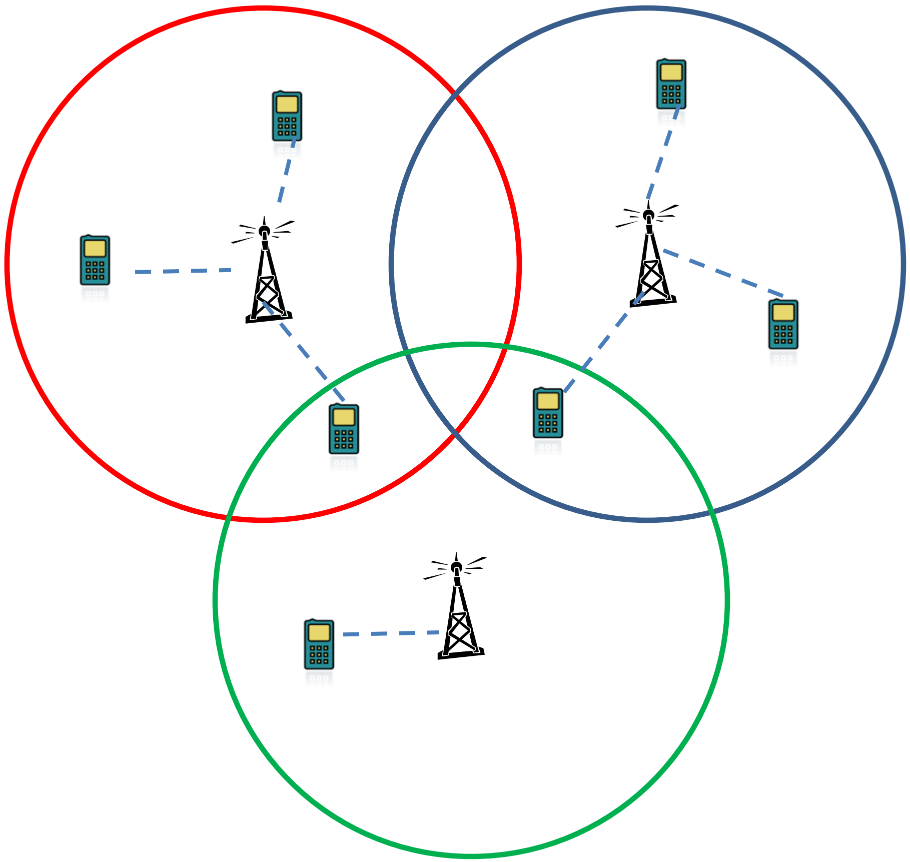
Granularity	Model 1	Model 2	Model 3
2-state	5.38%	6.29%	14.94%
3-state	6.81%	7.73%	18.62%
6-state	8.70%	9.65%	23.37%

Savings increase with finer granularity

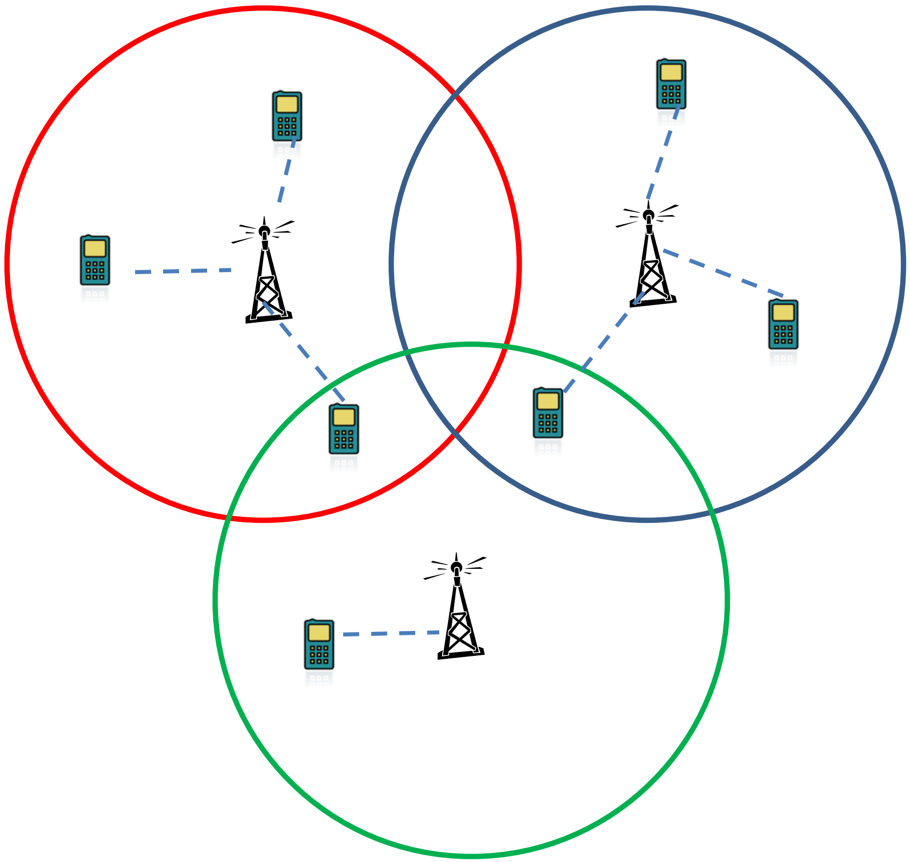
Effect of Late Deactivation



A Randomized Algorithm

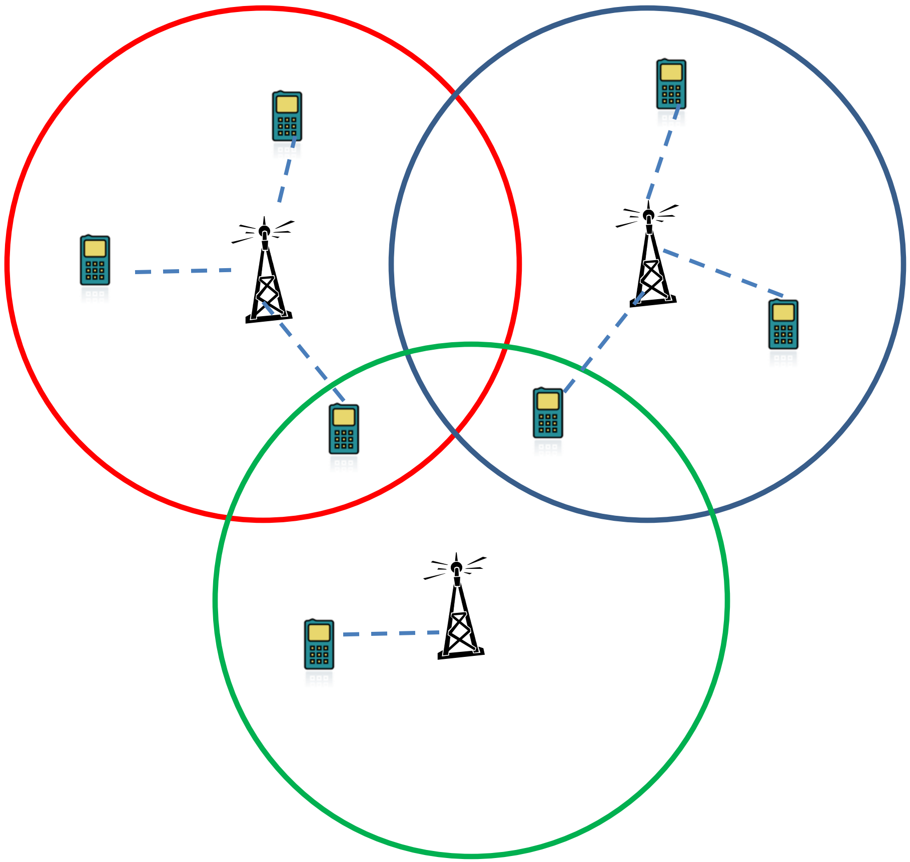


A Randomized Algorithm



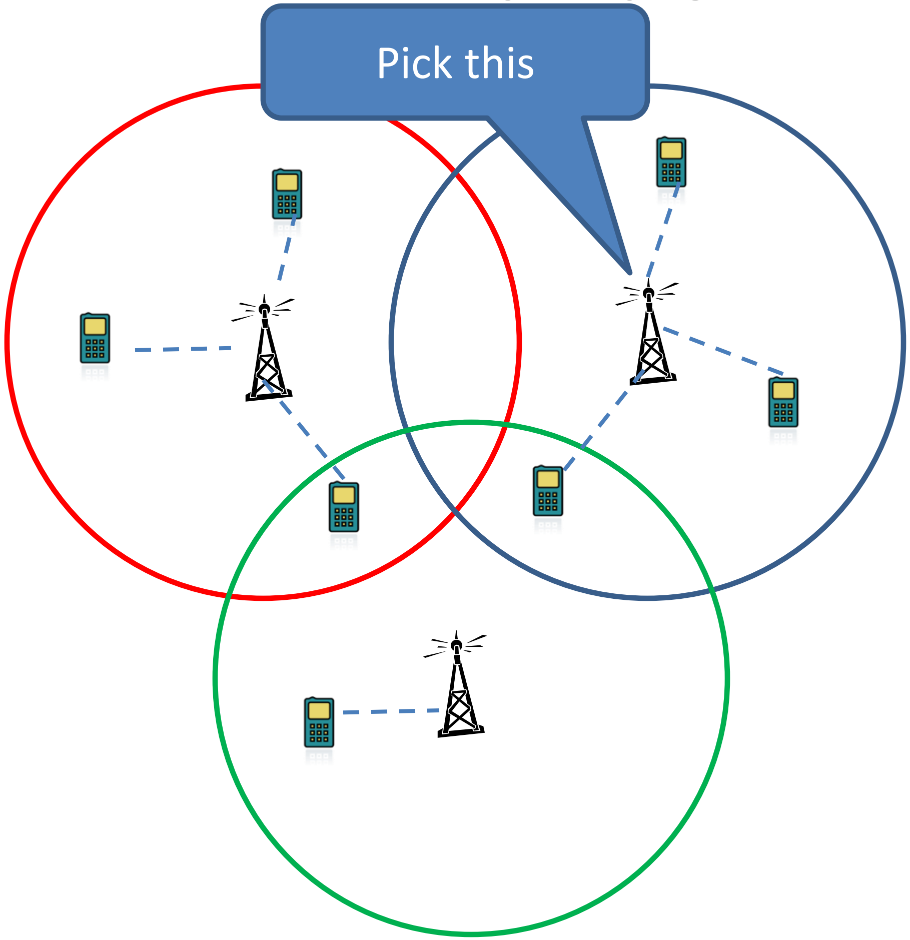
- While there are BTSs in high-power mode

A Randomized Algorithm



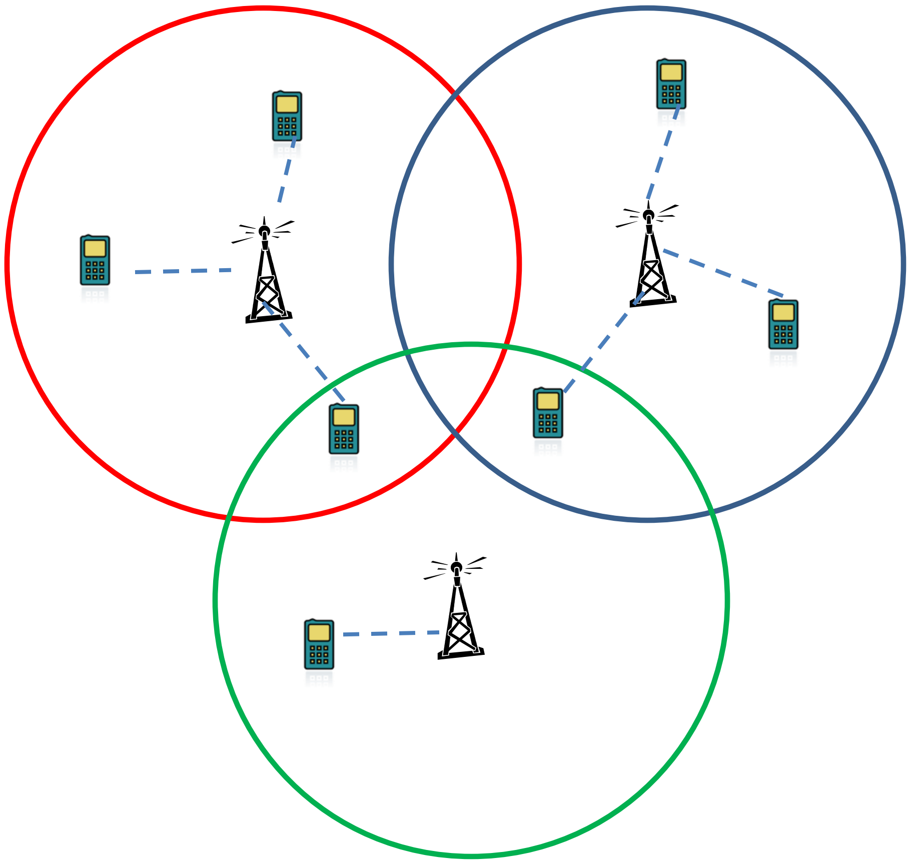
- While there are BTSs in high-power mode
  - Pick a random BTS

A Randomized Algorithm



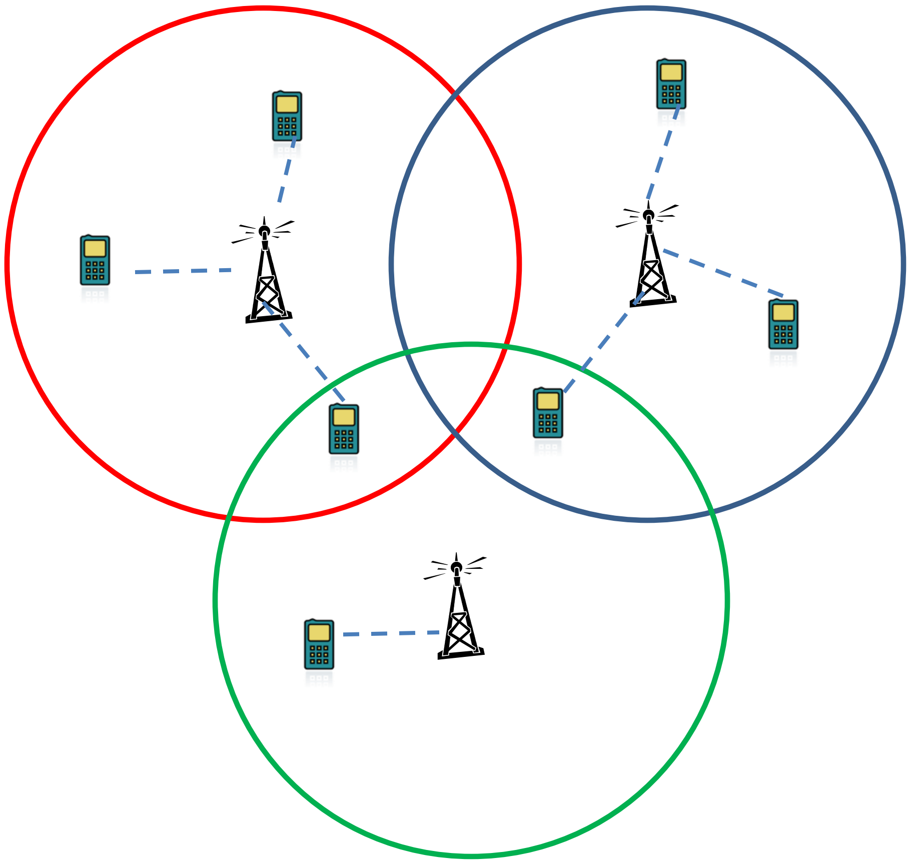
- While there are BTSs in high-power mode
  - Pick a random BTS

A Randomized Algorithm



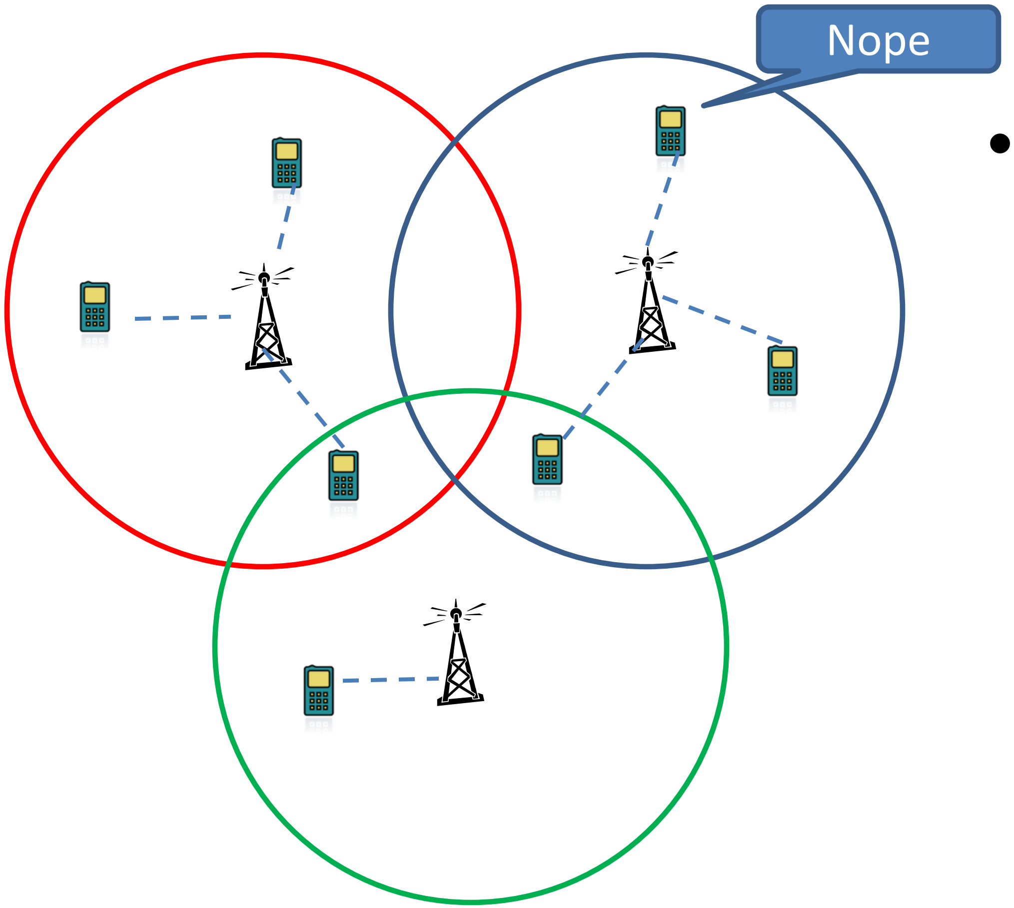
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS

A Randomized Algorithm



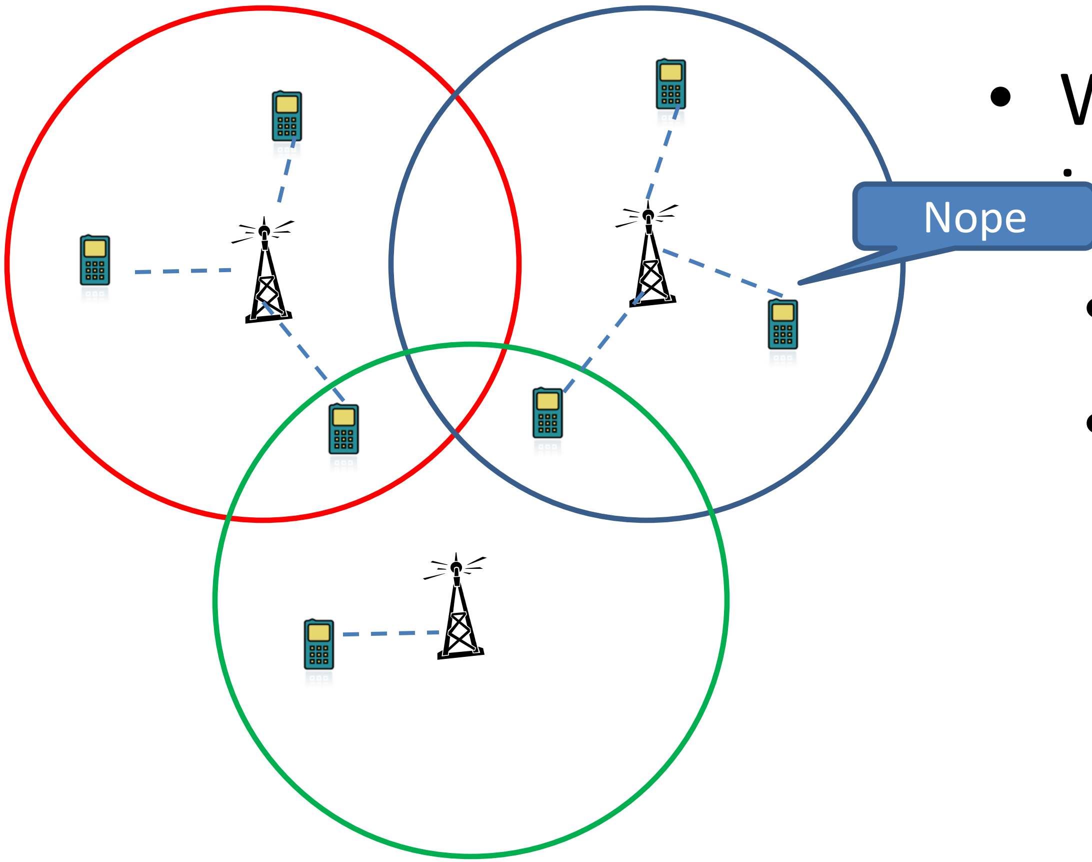
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

A Randomized Algorithm



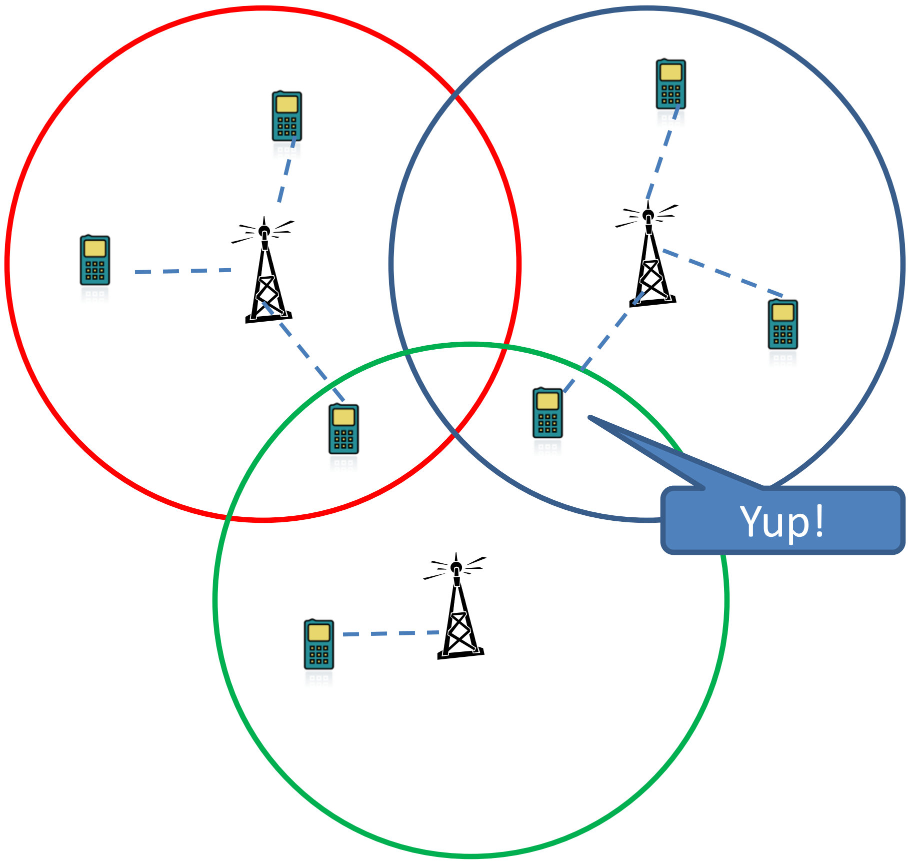
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

A Randomized Algorithm



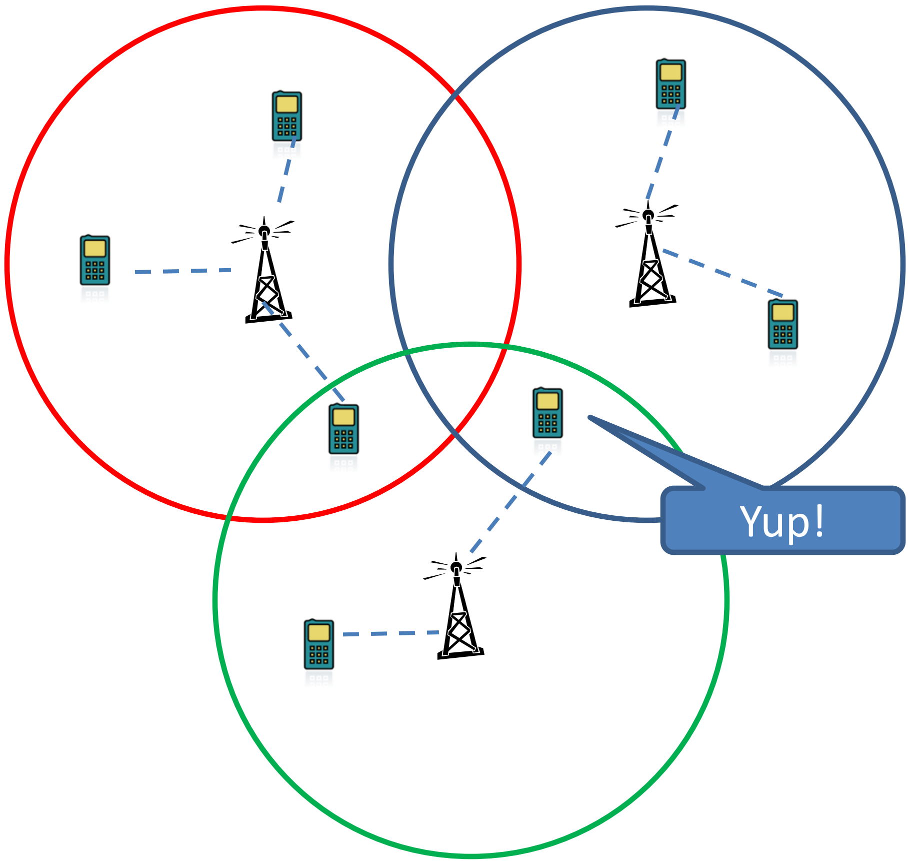
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

A Randomized Algorithm



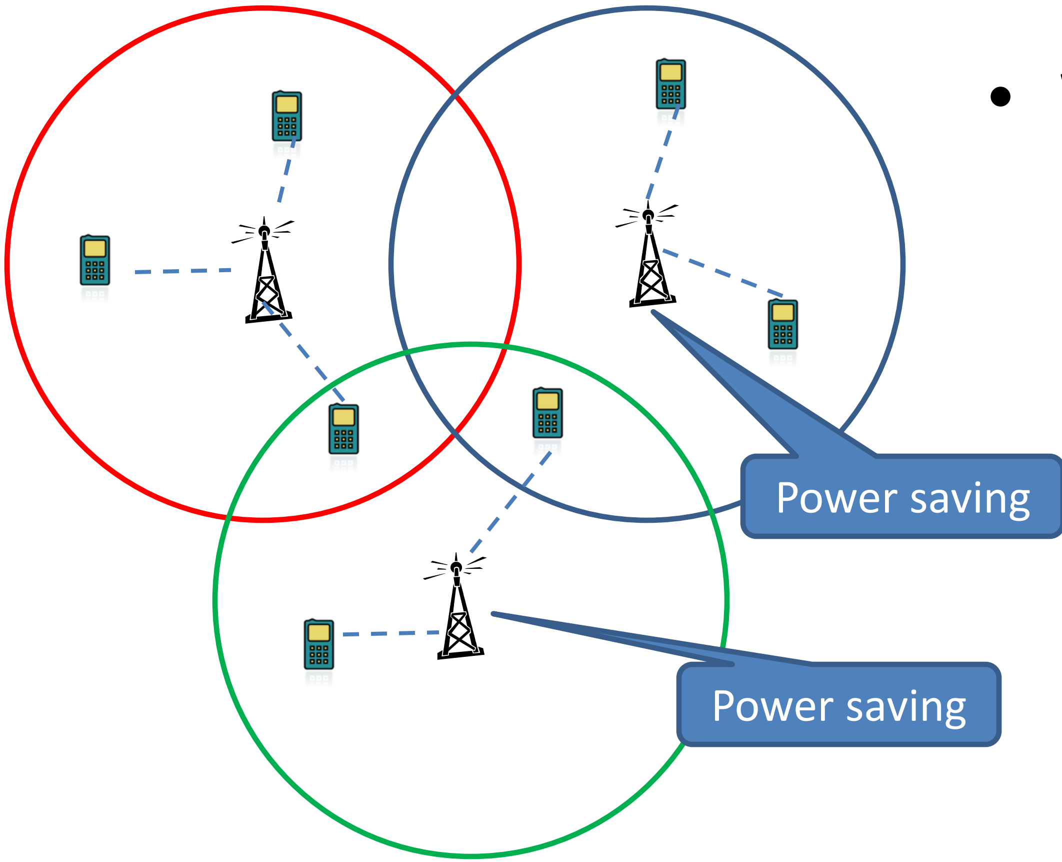
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

A Randomized Algorithm



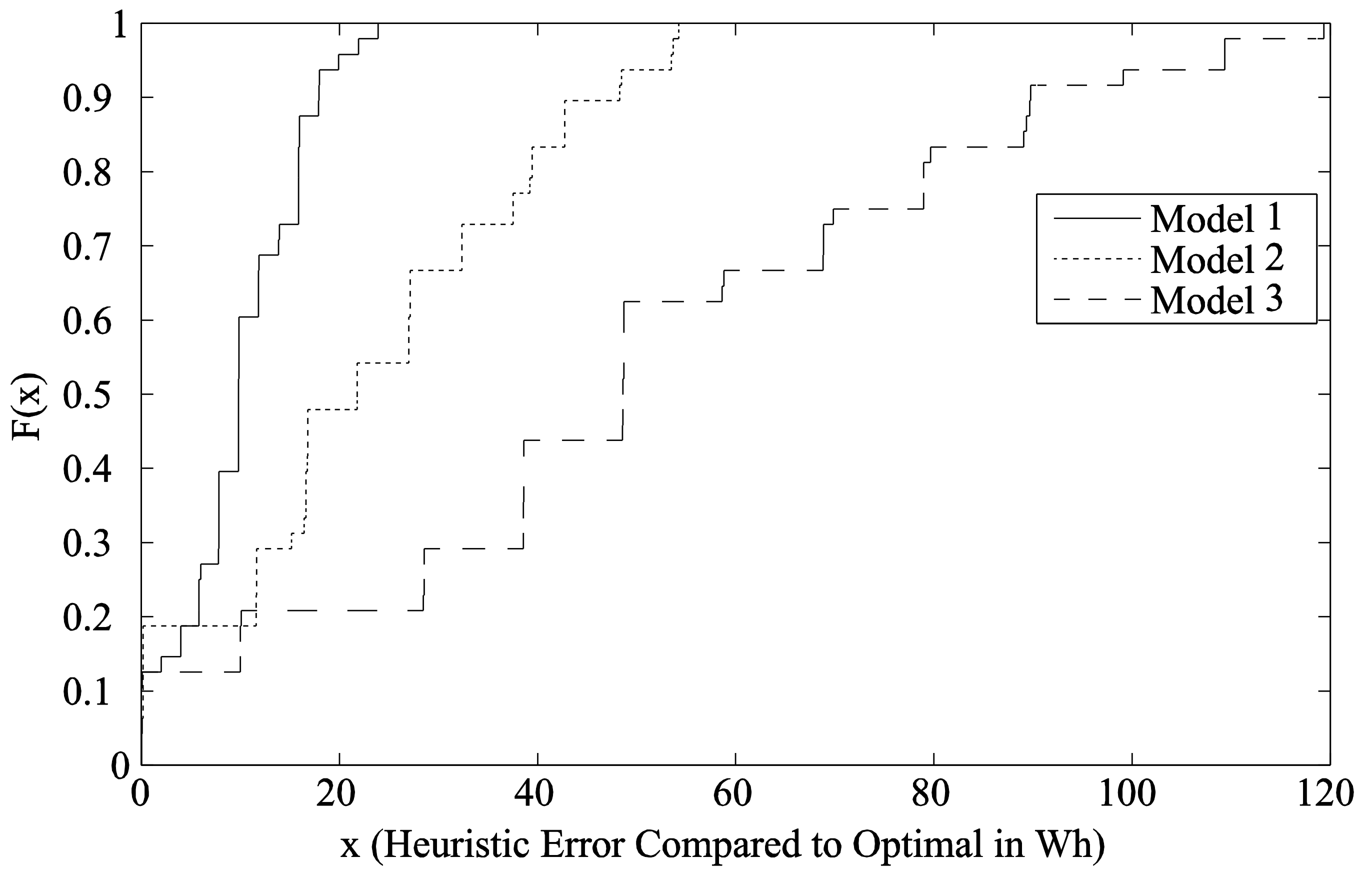
- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

A Randomized Algorithm



- While there are BTSs in high-power mode
  - Pick a random BTS
  - For each call being handled by this BTS
    - Hand-over to a candidate BTS in low-power mode

Performance of Heuristic Algorithm



Case Study II - Summary

- Traffic has limited geo-flexibility compared to data centers
- No geo-diversity in electricity prices
- Activation of power savings feature in hardware helps
- RED-BL achieves greater savings even for relatively conservative settings

Agenda

- Background and motivation
- Opportunity and key idea
- Case studies:
  - Data centers (e.g., Facebook and Google)
  - Cellular networks (e.g., Sprint and Verizon)
- **Conclusions and future work**

Conclusions

- Opportunities for electricity cost savings using:
  - Workload relocation
  - Resource pruning
- WR and RP are applicable to:
  - Data centers
  - Cellular networks
  - Others (need to explore)

Conclusions

- Modeled electricity cost minimization as an optimal state trajectory problem
- Showed the problem to be NP-Hard in the two case studies
- Studied the sensitivity of the problem to various parameters

Conclusions

- Data centers and cellular networks:
  - Sets of geo-diverse resources
- Contrasts:
  - Availability of geo-diversity in electricity prices
  - Geo-flexibility in traffic
  - Magnitude of transition costs

Future Work

- Factor in other forms of transition costs:
  - Cost of change in latency
  - Cost of replication
  - Cost of increase in call blocking probability
- Implementation on software BTS
- Incorporation into an OA&M framework
- Adaptation to recent generations of cellular networks
- Consider expensive diesel-generated power in cellular BTSs

Questions and Answers

If you can read this

Murphy’s law was violated

List of Papers

- Published:
  - A simulation study of GELS for Ethernet over WAN, GLOBECOM 2007
  - RED-BL: Energy solution for loading data centers, INFOCOM Mini-Conference, 2012
  - Electricity cost efficient workload mapping, INFOCOM Computer Communications Workshop, 2013
  - Low-Carb: Reducing energy consumption in operational cellular networks, GLOBECOM 2013
  - RED-BL: Evaluating dynamic right sizing for data centers, Computer Networks, vol. 72, 2014
- Submitted:
  - Low-Carb: A practical scheme for improving energy efficiency in cellular networks

Questions 1

- Why not have one data center at the cheapest locations?
  - There is no single cheapest location
  - Diversity for:
    - Disaster
    - Latency

Question 2

- Why can’t DVFS be used?
  - It can certainly be used
  - It does not achieve fine grained energy proportionality
    - Granularity of VF scaling is coarse
    - Other components are also energy proportional