Multi-resource Energy-efficient Routing in Cloud Data Centers with Network-as-a-Service

Lin Wang*, **Antonio Fernández Anta°**, Fa Zhang*, Jie Wu+, Zhiyong Liu*

*Institute of Computing Technology, CAS, China

°IMDEA Networks Institute, Spain

+Temple University, USA

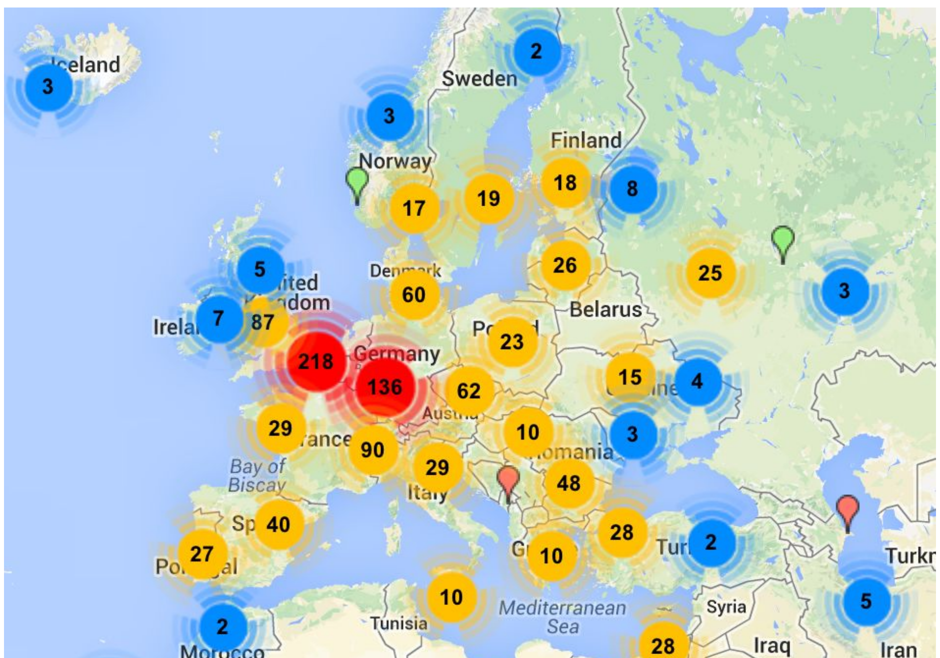
Developing the Science of Networks

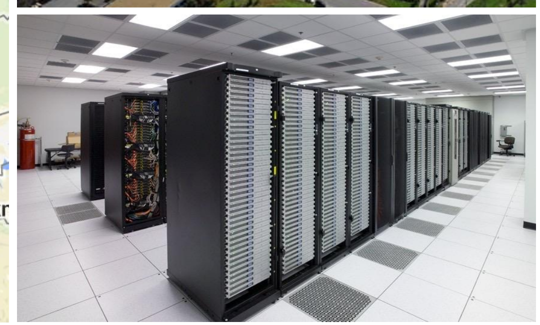
Data center and data center network

- Background
- Motivation
- Problem description
- Algorithms
  - Multi-resource green routing
  - Topology-aware heuristic
- Numerical validations
- Conclusions and future work

Data center and data center network

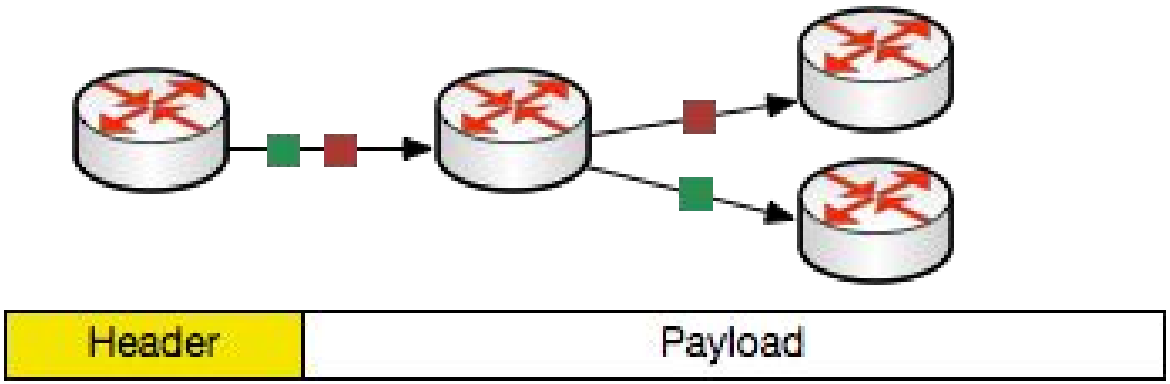
- Data centers have been ubiquitously deployed for providing computation and storage capabilities for cloud computing
- Data center network: the internal network for interconnecting the numerous servers in a data center





Traditional networking model

- Layer 2/3 functions such as forwarding and routing

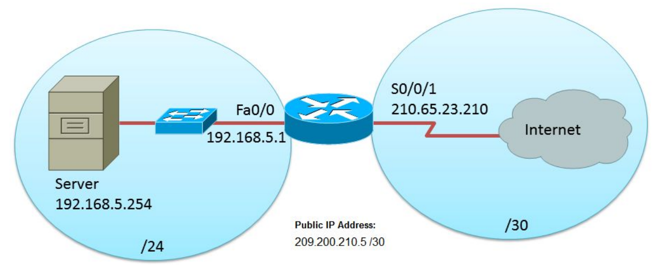


Destination	Egress port
10.0.0.2	1
10.0.0.3	2
...	...

- Link bandwidth is the most important criterion for performance evaluation

Middleboxes are relevant

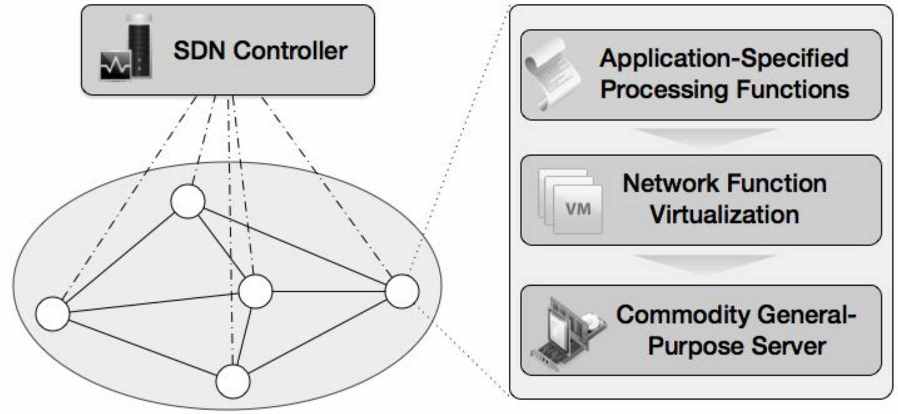
- Middleboxes: providing other network functions
  - Firewall, proxy, deep packet inspection, load balancer, NAT, WAN optimizers etc.
  - Comparable number to switches



- The process of deploying middleboxes is inflexible and prone to misconfiguration
- There are no available protocols and mechanisms to explicitly insert these middleboxes on the path between endpoints

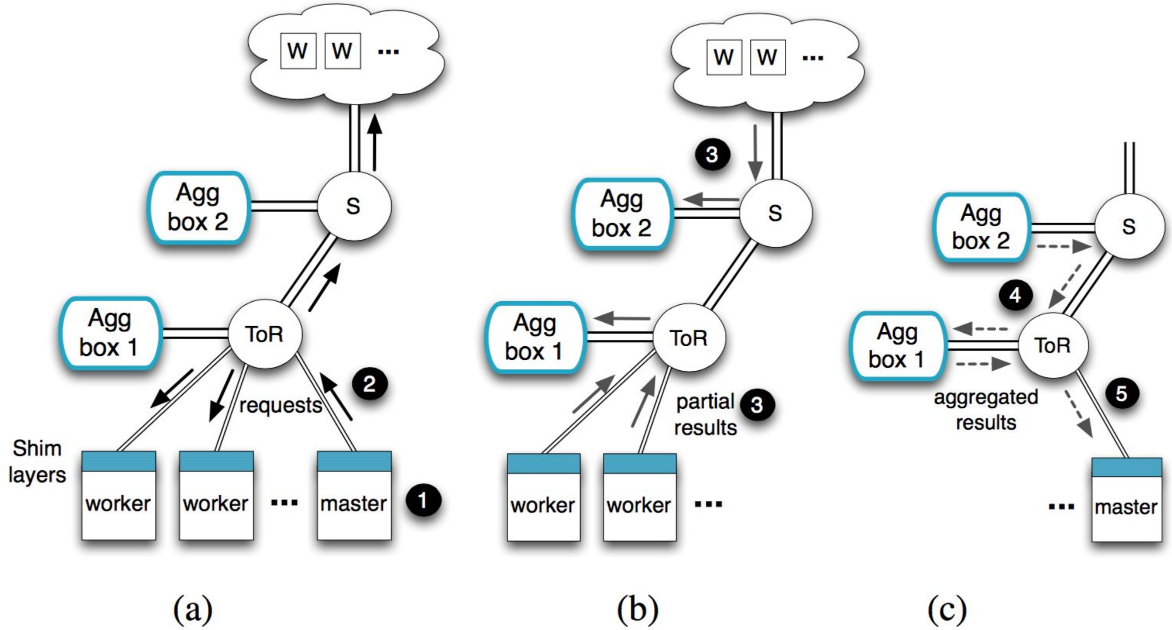
SDN & NFV lead to Network-as-a-Service

- Software-defined networking
  - Separating the network control plane and the data plane
  - Global visibility and logically centralized control
- Network function virtualization
  - Low cost with commodity hardware
  - More flexibility with software control
- Network-as-a-Service



What's new?

- In-network packet processing becomes reality
  - Application-specific on-path aggregation
  - NetAgg [Mai *et al*. CoNEXT 2014]

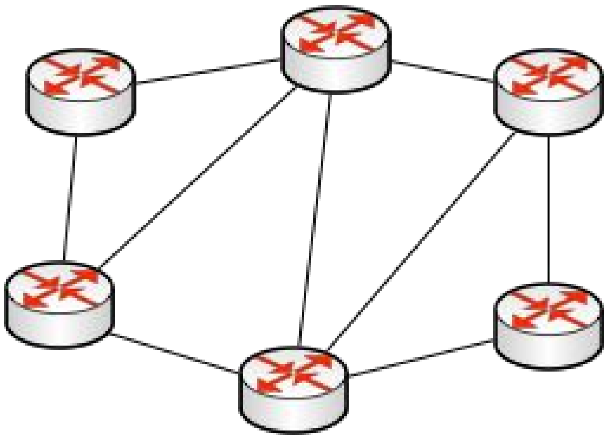


- Design processing pipelines for different processing logics
- Optimization problems will be different under this new networking model

Network optimization

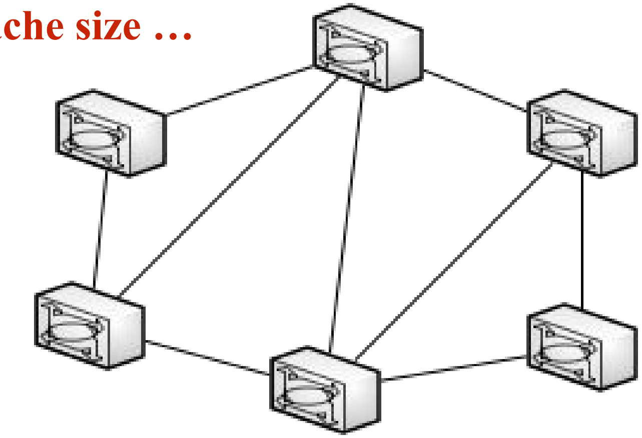
- From single-resource to multi-resource settings
- Old optimization methods are not efficient or even not applicable

link bandwidth



Single resource

link bandwidth
processing capacity
memory
cache size …



Multiple resources

Why energy efficiency matters

- Energy consumption comparison
  - Power consumption of a server is almost three times that of a switch

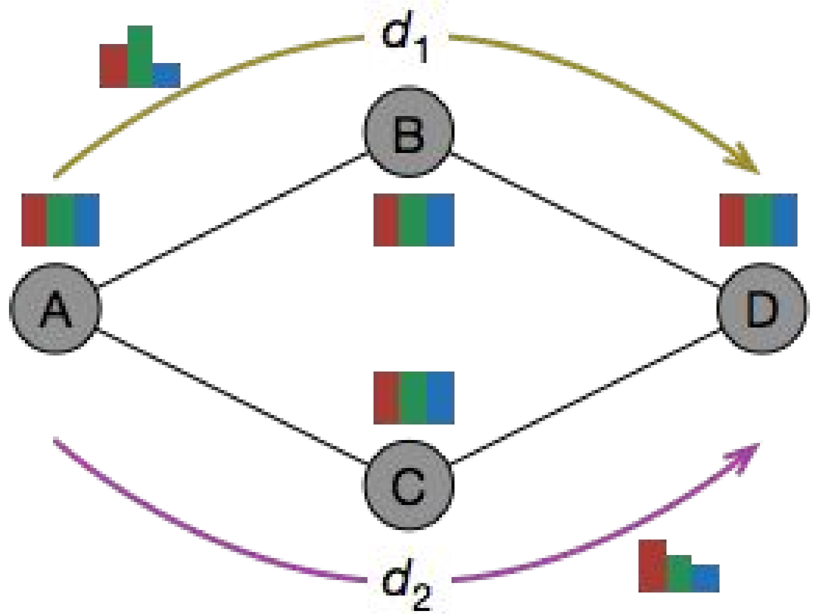
Cisco Nexus 3548: 265W
HP 5900AF-48XG: 260W
Juniper QFX 3600: 345W

Dell PowerEdge R715: 1100W
HP ProLiant DL80: 900W
Lenovo ThinkServer RD550: 750W

- Device level energy-saving mechanism: power-down
- Network global energy-saving strategy: traffic engineering
  - Consolidating network flows to a subset of network devices and turning idle devices into low-power modes

Modeling

- Modeling the network
  - A network $G = (V, E)$
  - K different types of resources, namely CPU, memory…
  - Capacity C_k, normalized to 1
- A set of flow demands $D = \{d_1,…,d_M\}$ where $d_m = (v_m{}^s, v_m{}^t, R_m)$, $\mathbf{R}_m = (r_{m,1}, r_{m,2},…, r_{m,K})$, and $r_{m,k} \in [0, 1]$



Multi-resource energy-efficient routing

- Solution: path P_m for each flow d_m such that $|A_v| \leq 1$ for $v \in V$ where $A_v = \sum_{m:\, v \in Pm} \boldsymbol{R}_m$ is the aggregation of the resource demand vectors of flows that are routed through v.
- Objective: minimize the set of nodes that are used to carry flows

$$(\mathbb{P}_1) \quad \text{minimize} \quad \sum_{v \in \mathcal{V}} y_v$$

subject to

$$\Big|\Big| \sum_{m \in \{1,2,\dots,M\}} \vec{R}_m \cdot x_{m,v} \Big|\Big|_\infty \leq 1 \qquad v \in \mathcal{V}$$

$$x_{m,v} \leq y_v \qquad v \in \mathcal{V}, 1 \leq m \leq M$$

$$x_{m,v}, y_v \in \{0, 1\} \qquad v \in \mathcal{V}, 1 \leq m \leq M$$

$$x_{m,v} : \ \text{flow conservation}$$

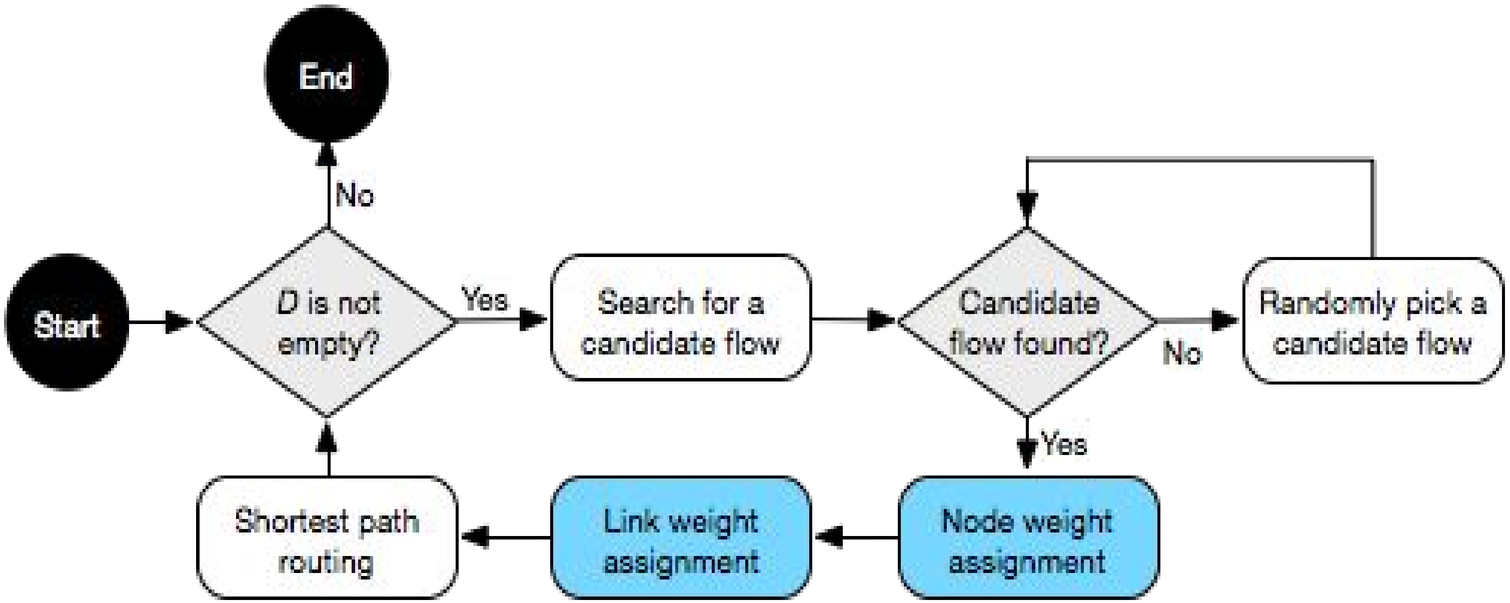
Complexity analysis

- *K* = 1: capacitated network design
  - Link version: polylogarithmic approx. with polylogarithmic congestion [Andrews *et al*. FOCS 2010]
  - Node version: a $O(\log^5 n)$-approx. with $O(\log^{12} n)$ congestion [Krishnaswamy *et al*. STOC 2014]
- *K* > 1: multi-dimensional node capacitated network design
  - **Theorem** Solving the multi-resource energy-efficient routing problem is NP-hard.

    Proof sketch: build a polynomial time reduction from the Vector Bin Packing (VBP) problem which is NP-hard.
  - **Theorem** There is no asymptotic PTAS for the multi-resource energy-efficient routing problem unless P=NP.

Multi-resource green (MRG) algorithm

- Key observations:
  - Flows preferably follow paths that consist of more active nodes (that already carry some traffic)
  - Load balance among all resource dimensions could be the new measuring method for resource efficiency
- A greedy routing scheme (Multi-resource Green, MRG)
- Time complexity: $O(|E|M^2)$



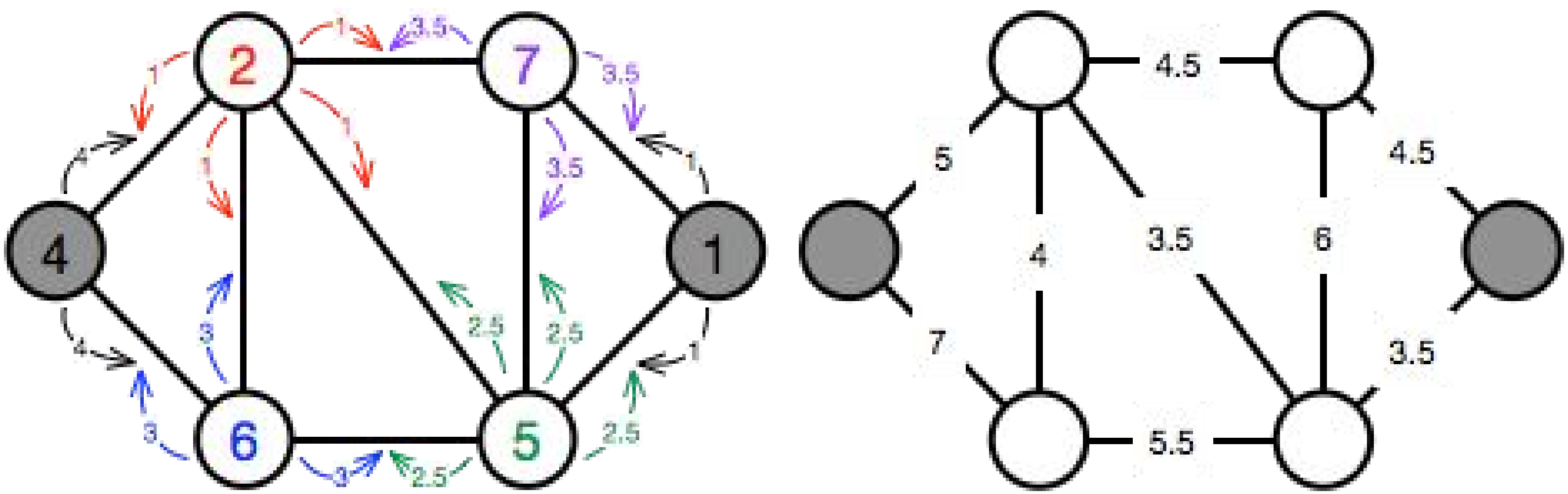
Node weight assignment: inversion counting

- **Definition** Given two vectors $\boldsymbol{X} = (x_1,\dots,x_n)$ and $\boldsymbol{Y} = (y_1,\dots,y_n)$, an **inversion** is defined as the condition $x_i > x_j$ and $y_i < y_j$, for $1 \leq i,j \leq n$.
- **Property** Given two vectors in n dimensions, the total number os inversions is upper bounded by $n(n-1)/2$.

A:	1	2	3	4	
B:	3	1	4	2	# of inversions = 3
C:	4	3	2	1	# of inversions = 6

Link weight assignment

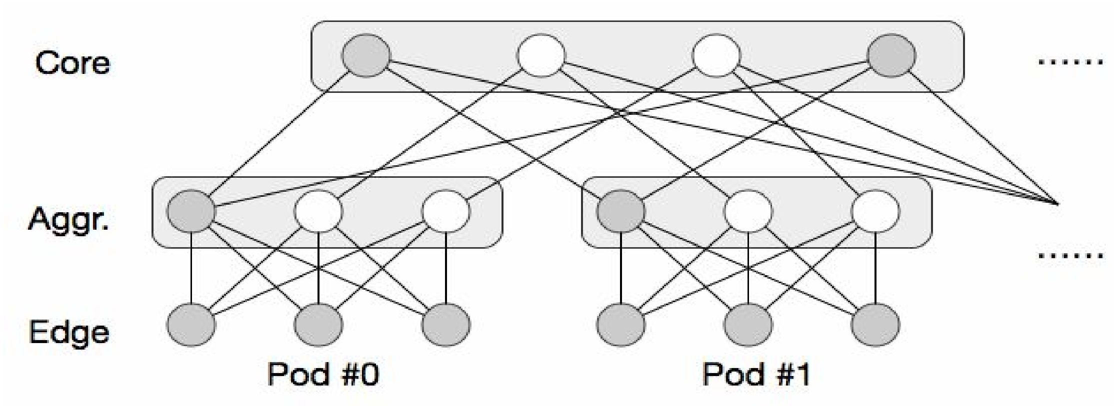
- Node weights to adjacent link weights
  - For *src* and *dst*, node to link directly
  - For intermediate nodes, divide by two



The min-weight routing problem remains the same.

Topology-aware heuristic: Hierarchical green routing

- Taking advantage of the hierarchy of data center network topologies (e.g., fat-tree)

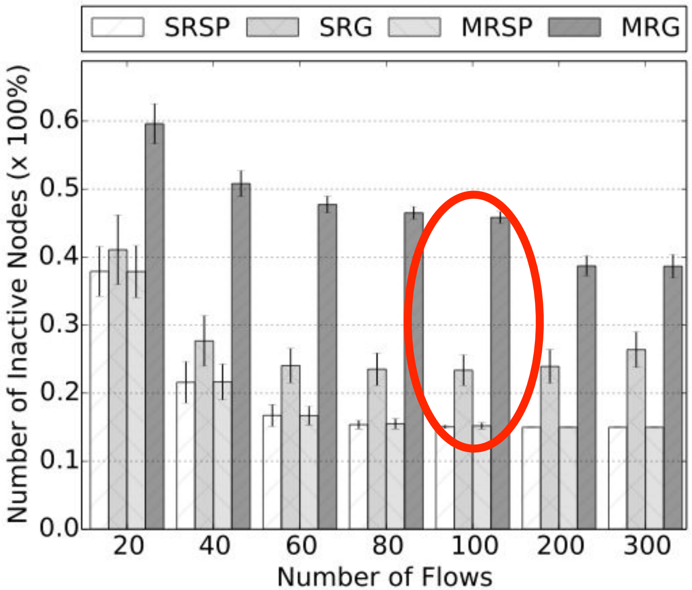


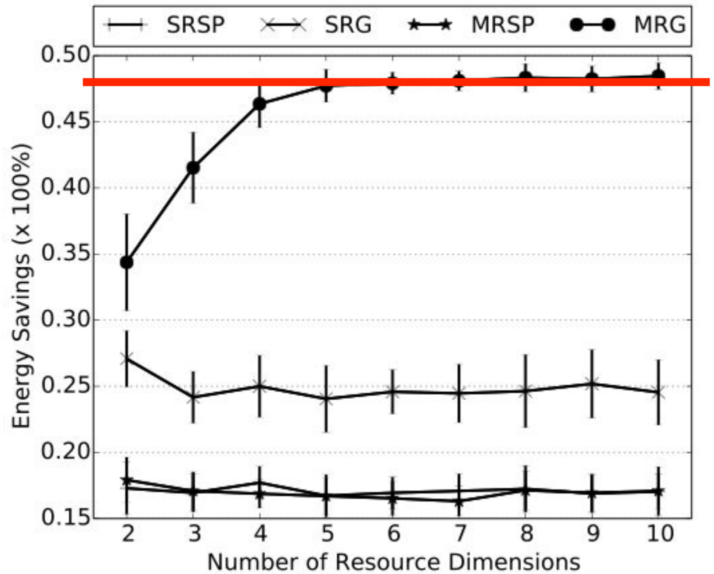
- HGR: solving a series of vector bin packing instances using a norm-based greedy algorithm [Panigrahy *et al*. ESA 2011]
  - Bin-centric
  - In each iteration, choose the item that minimizes the weighted l_2-norm of the bin residual capacity and the demand
- Time complexity $O(M^2)$, speedup $\Omega(|E|)$

Numerical validations

- Python implementation
- Topology: fat-trees in different scales
- Flow demands: randomly generated
  - Endpoints: uniformly at random
  - Resource requirements: normal distribution (positive)
- Comparison
  - Single-Resource Shortest Path (SRSP)
  - Multi-Resource Shortest Path (MRSP)
  - Single-Resource Green (SRG)
  - Multi-Resource Green (MRG)

Performance of MRG

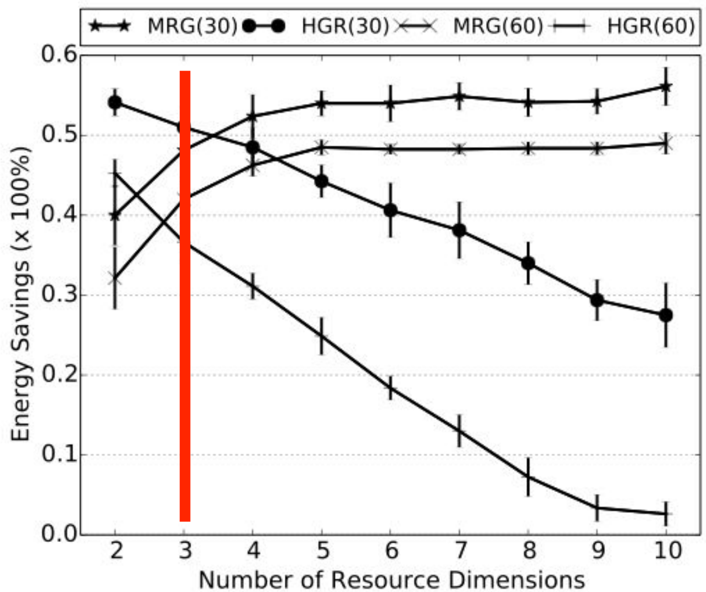


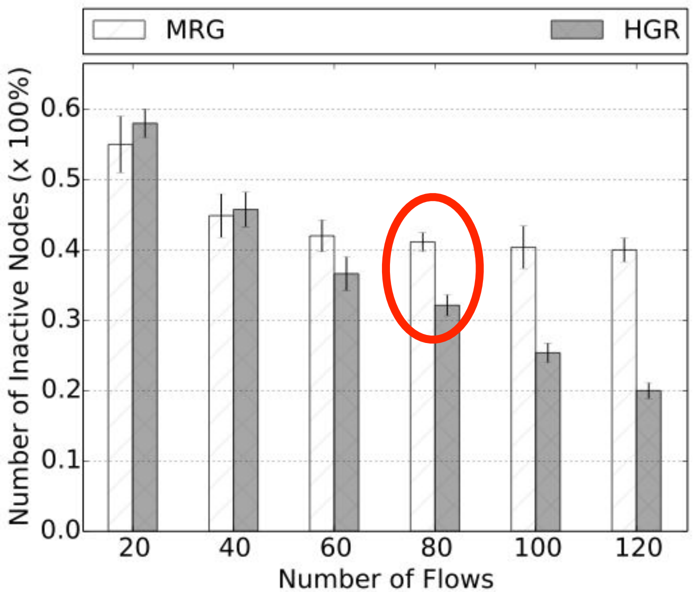


- The MRG algorithm outperforms the others with a factor of over 25% in energy efficiency
- The MRG algorithm converges to a stable energy saving level with respect to the number of resource dimensions

Performance of HGR

typical number of resource dimensions = 3





**Less than 10% energy savings degradation,
but having a speedup of over 180**

Running Time (second)						
# of flows	*20*	*40*	*60*	*80*	*100*	*120*
MRG	5.37	16.63	37.00	58.26	92.93	101.89
HGR						

Conclusions

- The new networking paradigm pushes network optimization models from single-resource to multi-resource
  - Multi-resource traffic engineering requires new techniques
  - The network energy efficiency problem becomes more prominent with the Network-as-a-Service model
- We study the multi-resource energy-efficient routing under the Network-as-a-Service model
  - Problem formulation and complexity analysis
  - A greedy algorithm and a topology-aware heuristic
  - Up to 25% more energy efficiency could be achieved
- Our solution could be extended and applied to many practical networking scenarios

Future lines

- Model extension
  - Online: dynamic flow joining and leaving
  - Heterogeneity: different resource demands on different in-path nodes
  - Both algorithms can be extended to those cases
- Practical application scenarios
  - Named data networking (prefix matching, data caching)
  - Server-centric data center network architectures
    - BCube [Guo *et al*. SIGCOMM 2009]
    - SWCube and SWKautz [Li *et al*. INFOCOM 2014]
  - Network function orchestration

THANK YOU!