ElasticTree: Saving Energy in Data Center Networks

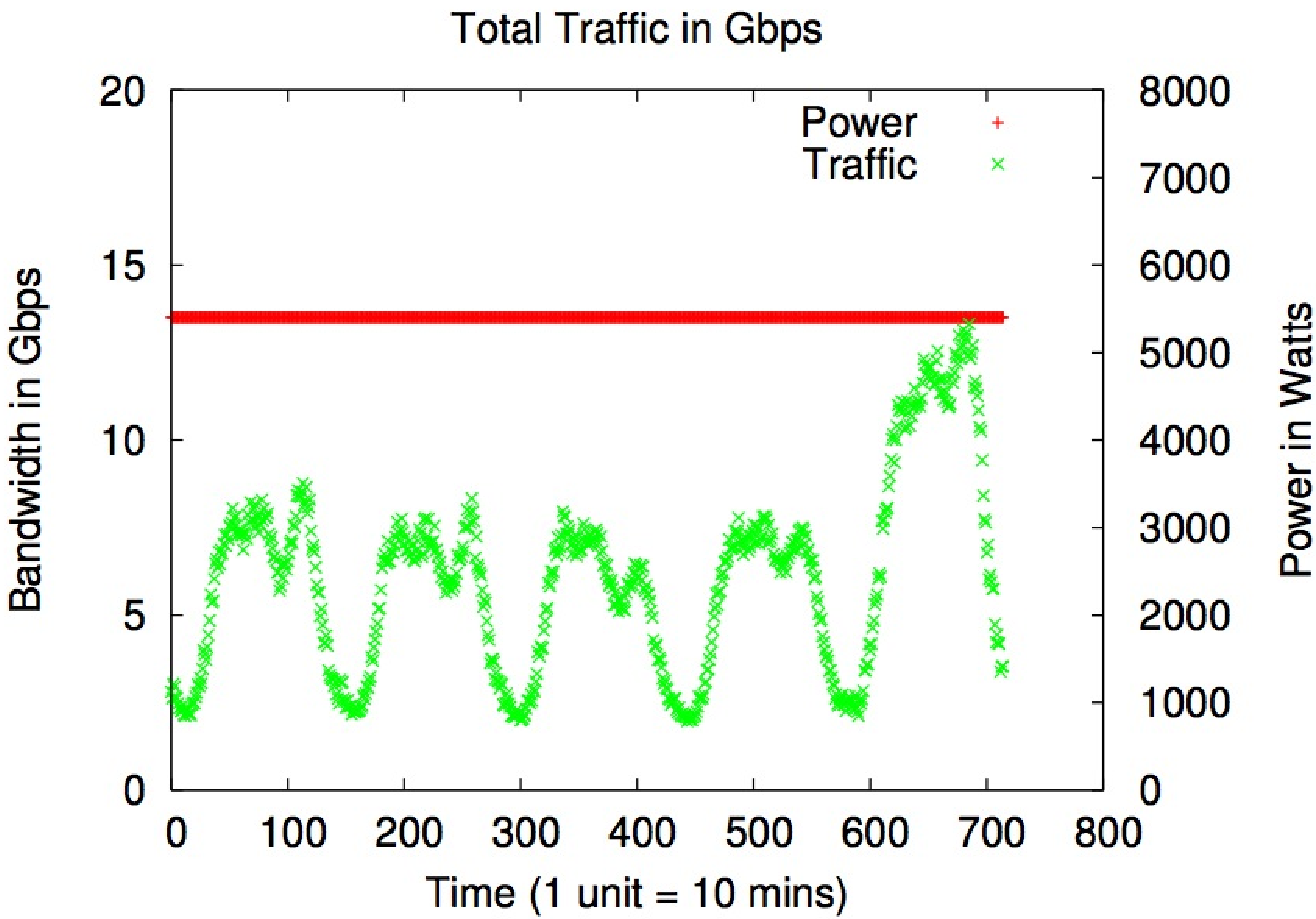
Brandon Heller, Srini Seetharaman,
Priya Mahadevan, Yiannis Yiakoumis,
Puneet Sharma, Sujata Banerjee, Nick McKeown

Presentation by Michał Dereziński

Problem

- Data centers consume huge amounts of energy.
- Networking elements are responsible for 10-20% of power usage in a data center.
- They are typically provisioned for peak workload.
- During lower traffic, most of the network components are idle, still using power.

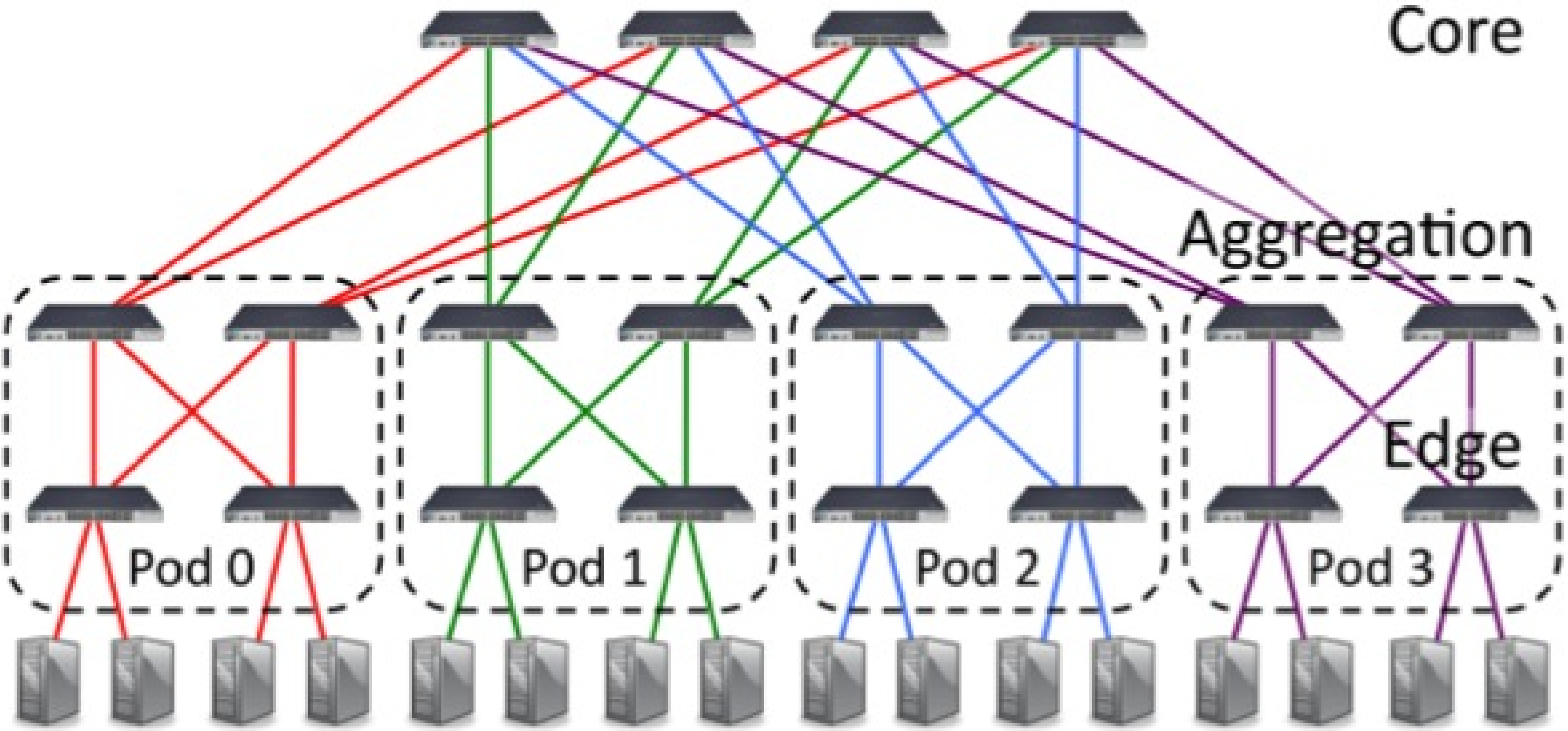
Problem


Total Traffic in Gbps
Power
Traffic
Bandwidth in Gbps
Power in Watts
Time (1 unit = 10 mins)
0
5
10
15
20
0
1000
2000
3000
4000
5000
6000
7000
8000
0
100
200
300
400
500
600
700
800

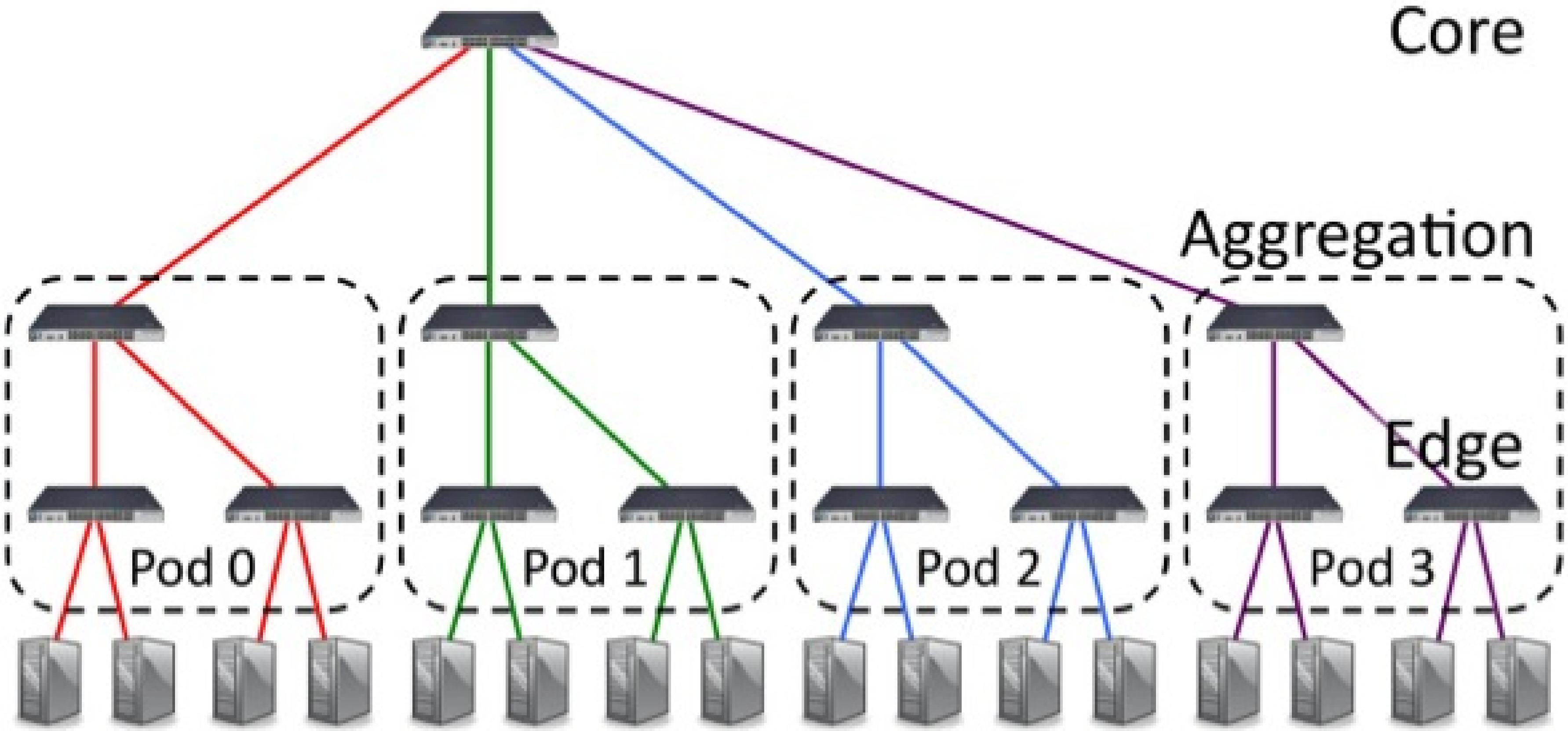
Solution: ElasticTree

- Network-wide energy optimizer.
- Turns off as many unneeded links and switches as possible.
- Monitors data center traffic conditions and dynamically adjusts the network.
- Keeps good performance and fault tolerance while significantly decreasing energy usage.

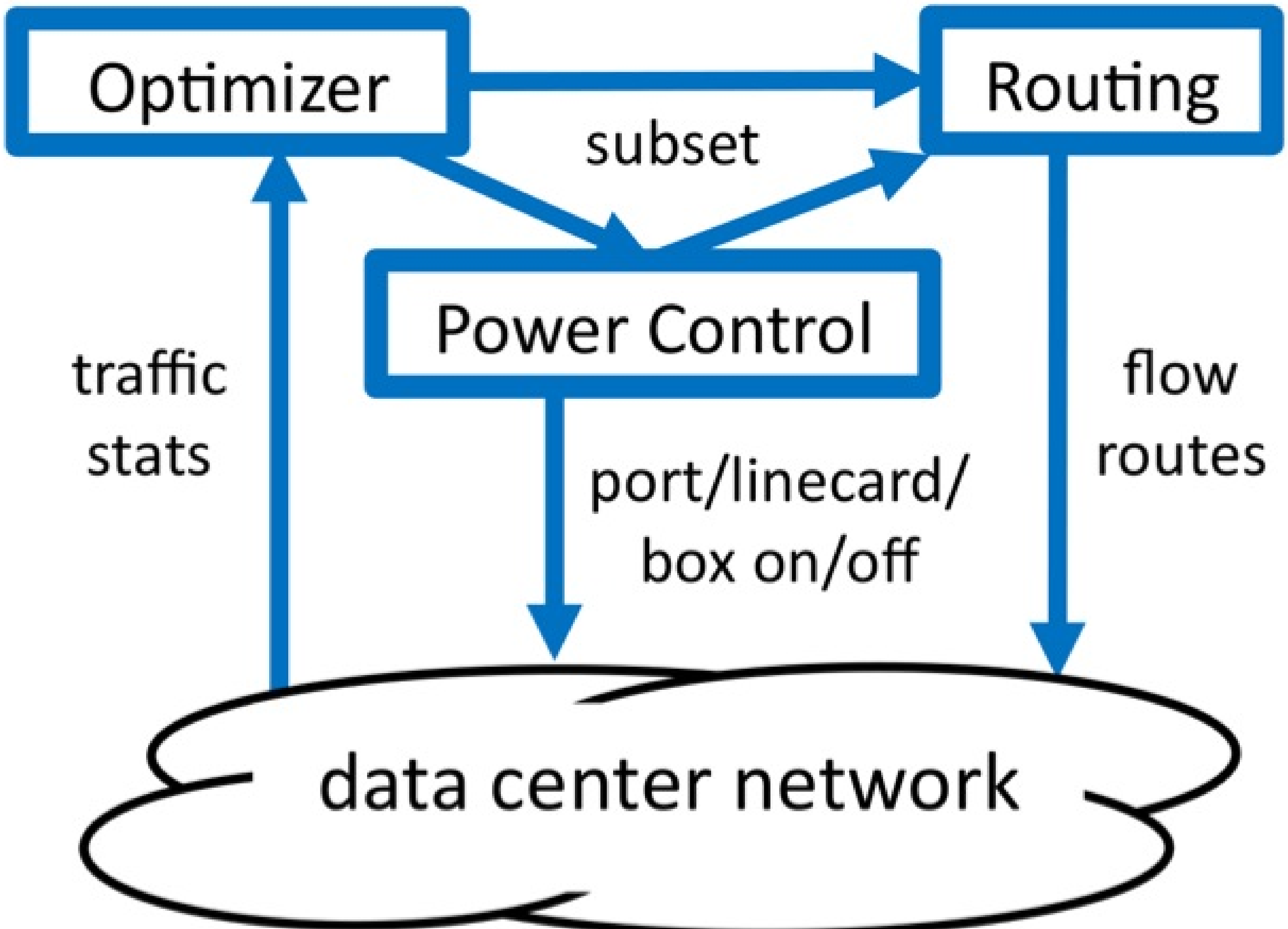
Example
fat tree topology, k=4



Example
fat tree topology, k=4



System Diagram



Optimizers

- Formal Model
- Greedy Bin-Packing
- Topology-aware Heuristic

Type	Quality	Scalability	Input	Topo
Formal	Optimal	Low	Traffic Matrix	Any
Greedy	Good	Medium	Traffic Matrix	Any
Topo-aware	OK	High	Port Counters	Fat Tree

Prototype Test Bed

- **Smaller configuration:**
  - complete k=4 fat tree topology,
  - 20 four-port virtual switches,
  - supporting 16 hosts at 1Gbps apiece.

- **Larger configuration:**
  - complete k=6 fat tree topology,
  - 45 six-port virtual switches,
  - supporting 54 hosts at 1 Gbps apiece.

- NetFPGA traffic generators
- Latency monitor

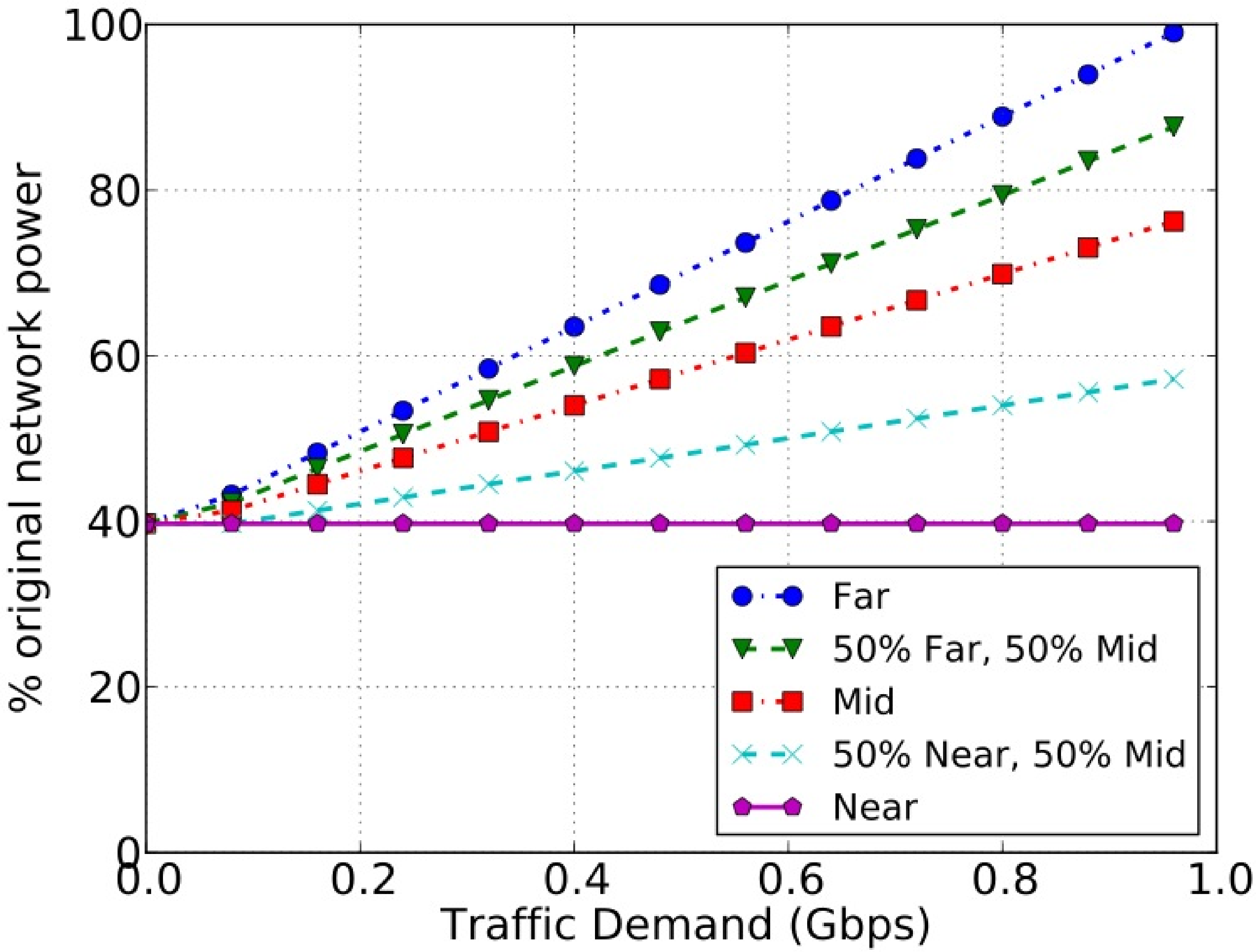
Power Savings Analysis

- Small tests performed on prototypes.
- Larger networks tested through simulations.

- Considered power usage:
  - number of switches powered on,
  - number of ports enabled on them.

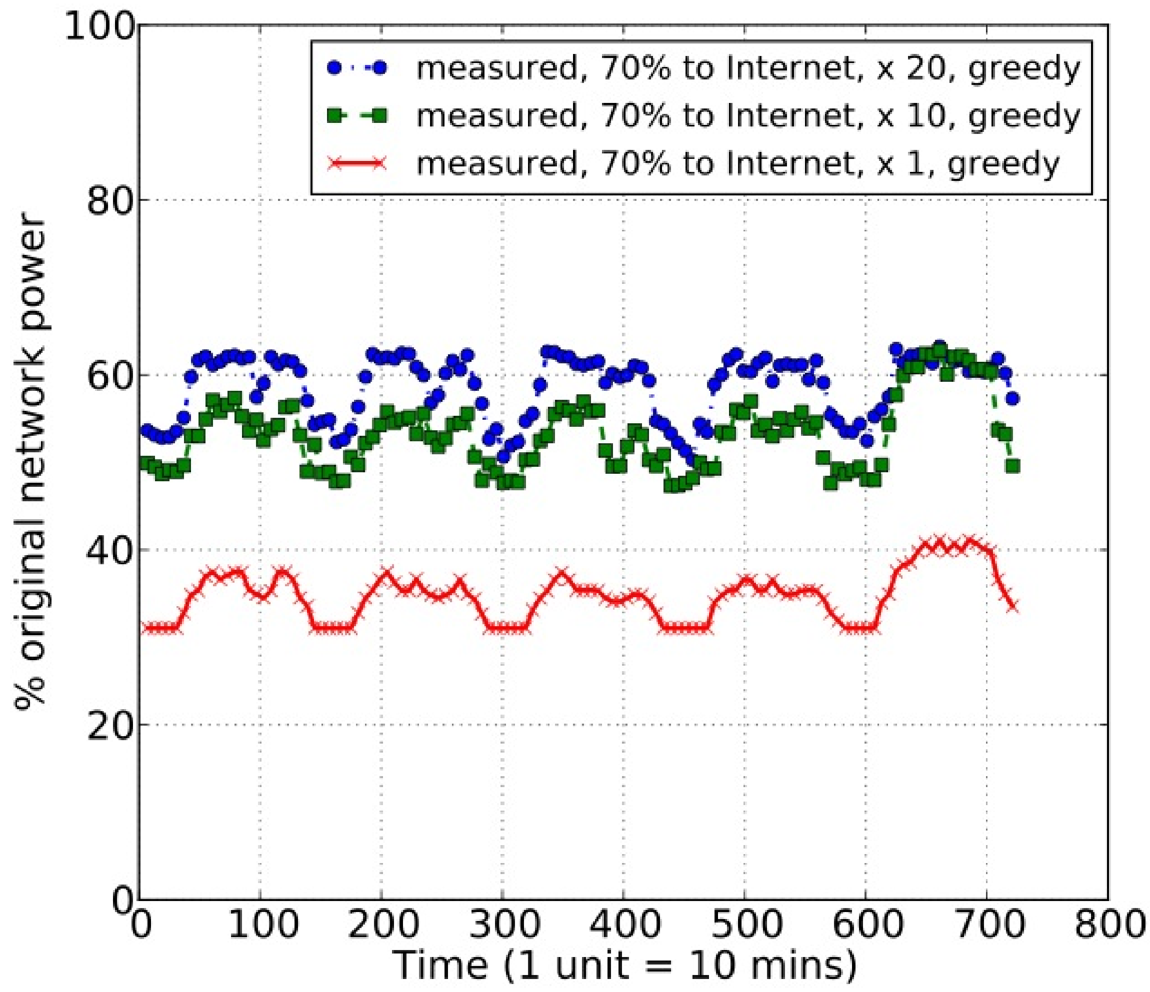
- Ignored power usage:
  - running servers hosting ElasticTree modules,
  - cooling components:
    - additional energy for cooling servers,
    - decreased energy for cooling switches.

Traffic Patterns



Simulations on real traffic data

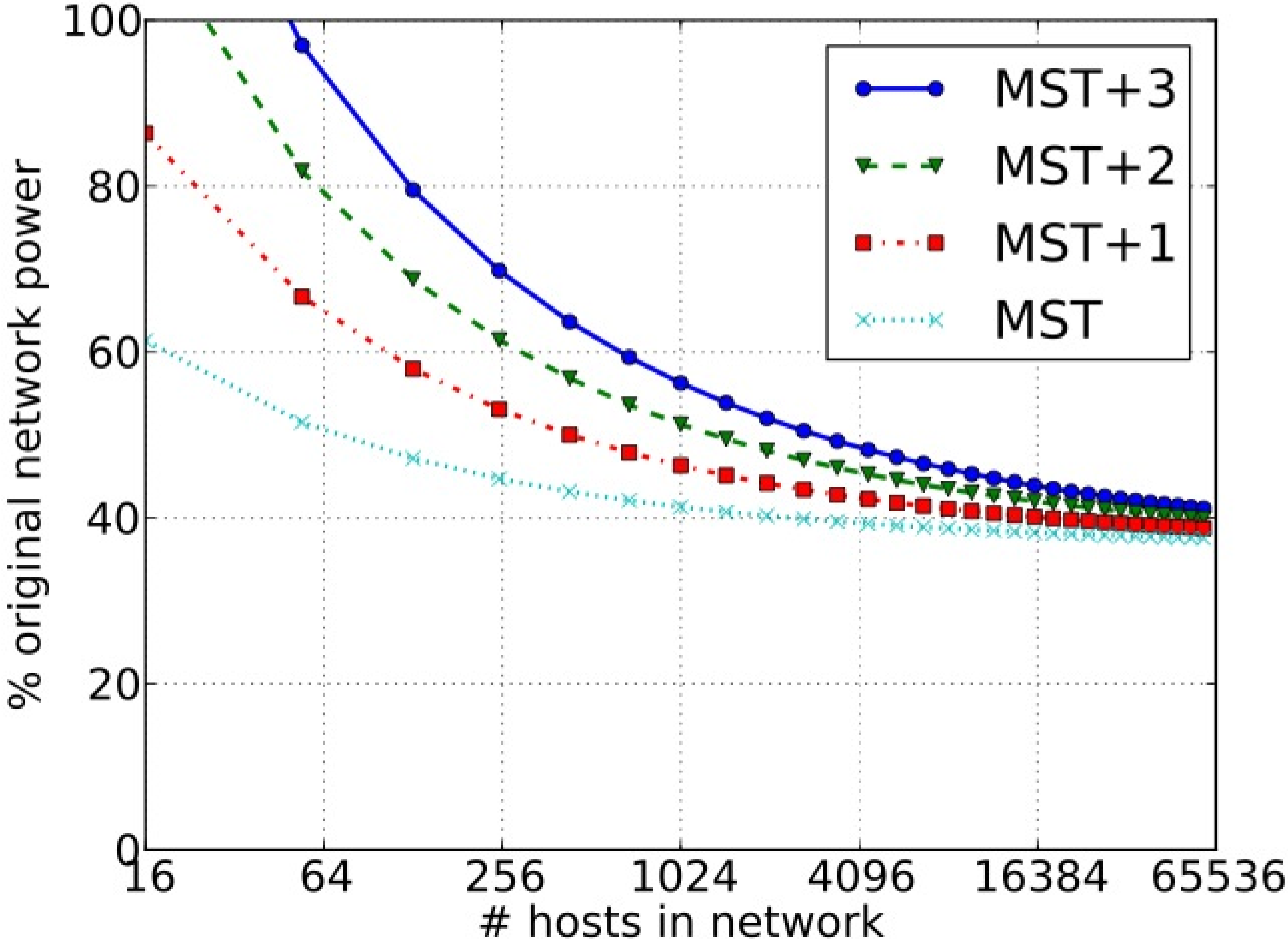
- E-commerce website
- 292 servers
- Fat tree, k=12
- Tested for different levels of overall traffic



Robustness Analysis

- Network topology must be prepared for:
  - traffic surges,
  - network failures.
- Adding a minimum spanning tree to the power optimized topology enables one failure with no loss of connectivity.
- Additional energy cost decreases with the size of the topology.

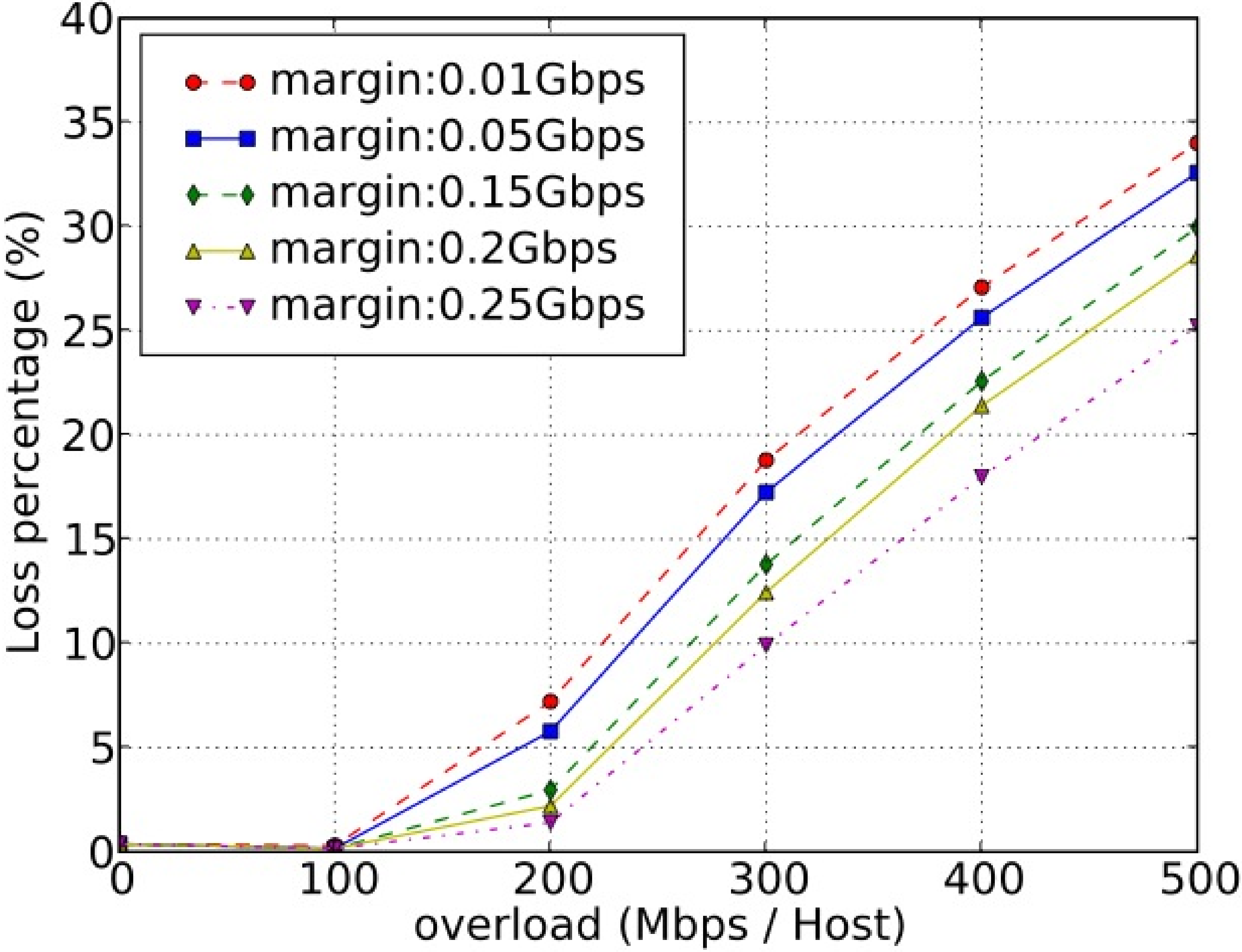
Robustness Analysis



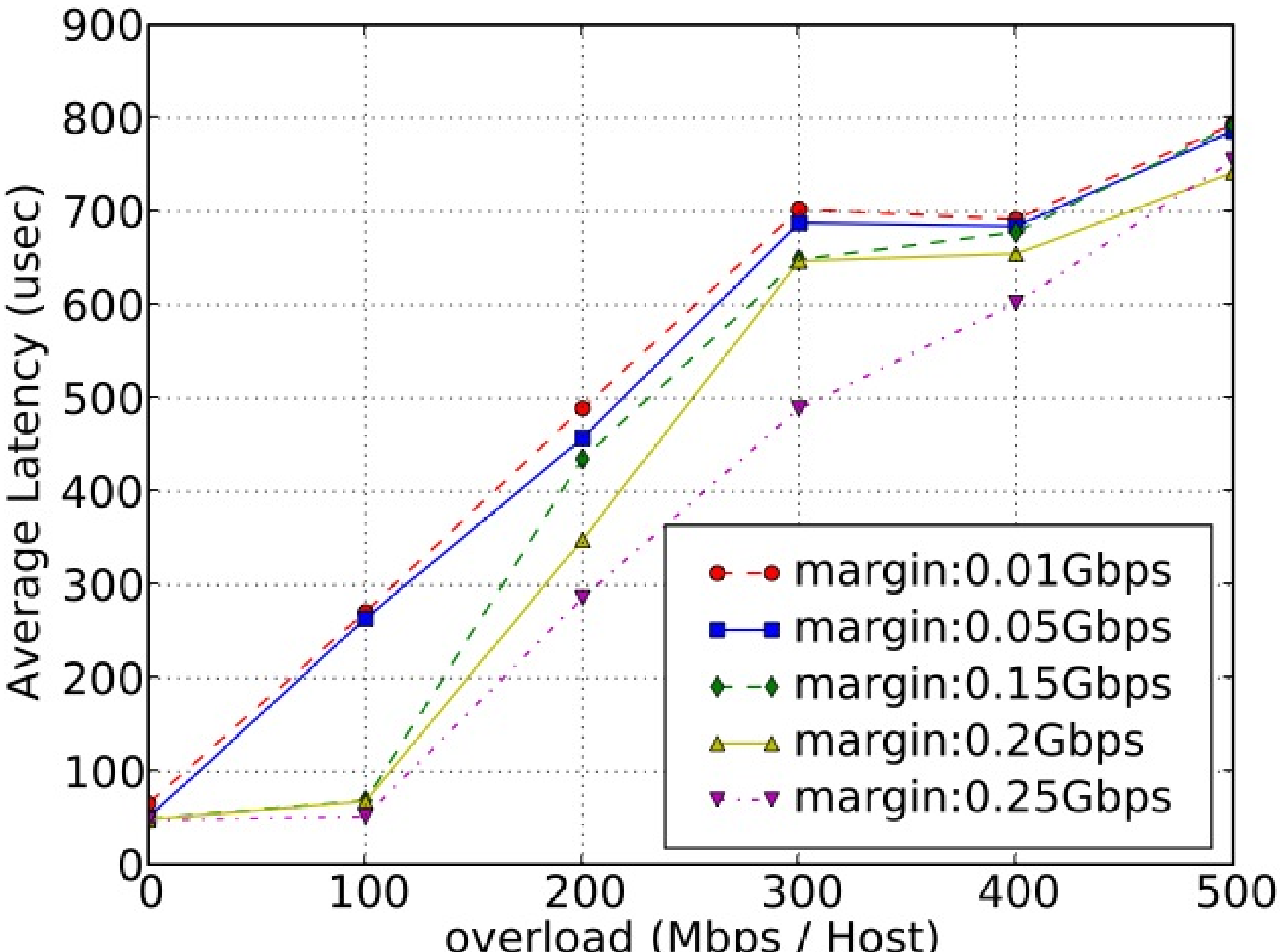
Performance

- Any energy saving policy should have negligible performance penalty.
- ElasticTree needs to deal with processing overheads, traffic bursts and sustained load increases.
- Safety margins are added to the traffic data to improve network latency and decrease the number of dropped packets.

Dropped packets

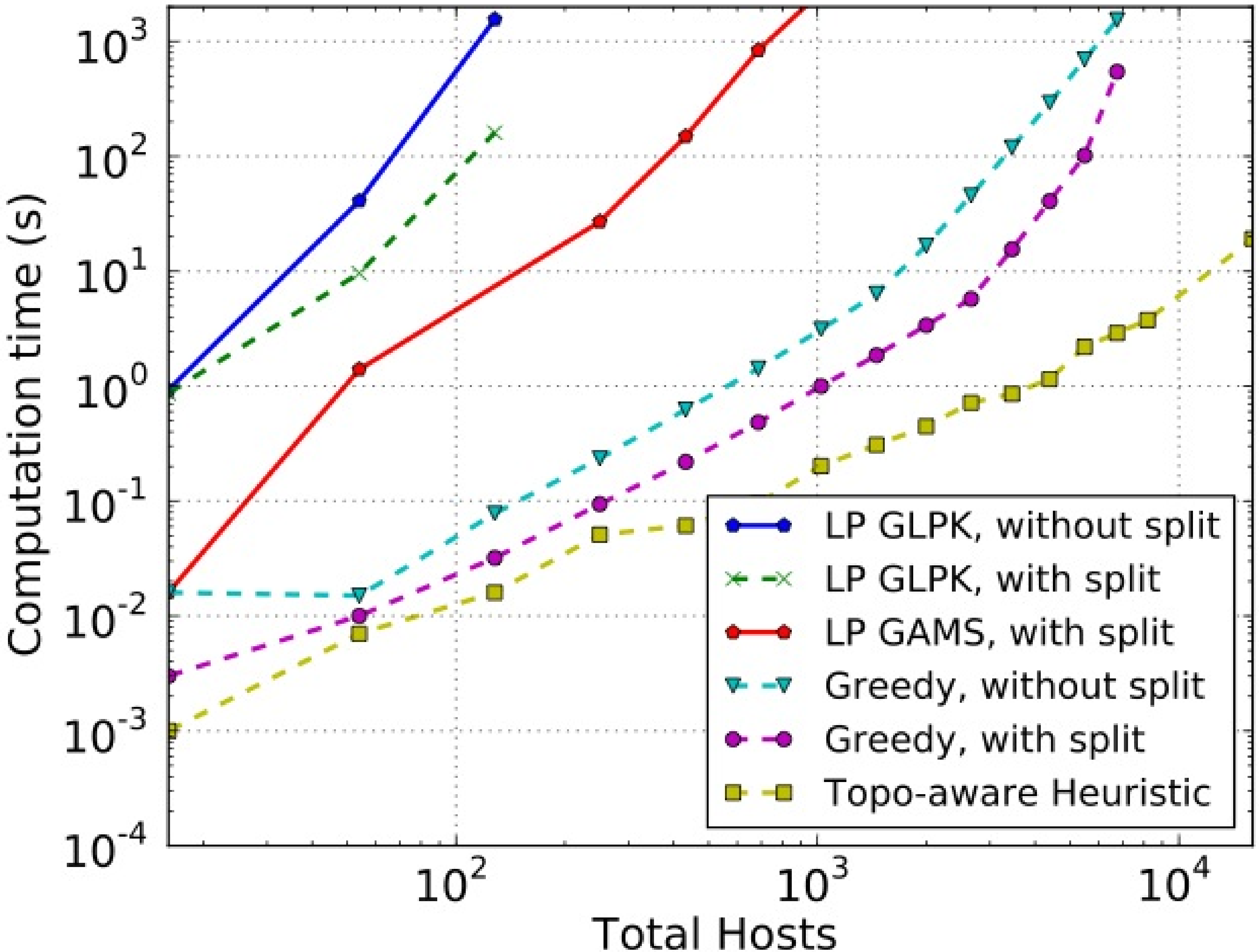


Latency



Average Latency (usec)
900
800
700
600
500
400
300
200
100
0
0
100
200
300
400
500
overload (Mbps / Host)
margin:0.01Gbps
margin:0.05Gbps
margin:0.15Gbps
margin:0.2Gbps
margin:0.25Gbps

Processing Overhead



Practical Considerations

- **Response time.** When increasing network's capacity, turning the switches on takes up most of the time (ranges from 30 seconds to 3 minutes)
- **Traffic prediction** should significantly improve the response time. Initial tests are promising.
- **Fault tolerance.** In case of optimizer failure, power management should be turned off automatically.

Conclusions

- ElasticTree introduces energy proportionality in today's non-energy proportional networks.
- This highly flexible system allows for balancing between performance, robustness and energy.
- Initial results suggest very significant power benefits for networks with varying utilization.

Thank You

Questions?