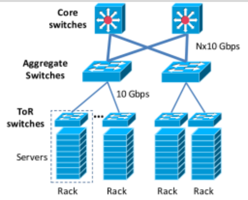
Research 2 (series): Performance Enhancement in Overloaded Networks

Motivation

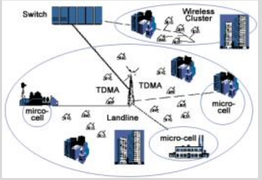
- Extensive analysis over underloaded system

- However, overloaded situation becomes more frequent in IoT but under unsystematic study & results

Server Farm



Datacenter



Mobile System

Communication infrastructure in Smart Grid; Cloud; HPC; Edge computing, etc.

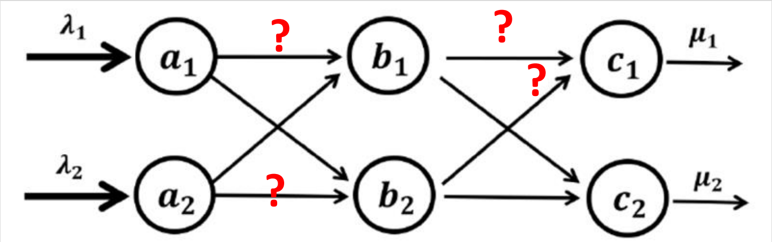
Contributions

1) Model queue dynamics by flow, which generalizes different network settings.
   - overload/underload, shared/split buffer, etc.
2) Propose network policies that optimizes metrics: latency, fairness, throughput, under network overload.

Two papers are published, and one is submitted.

(1) Latency

Set service rates to minimize queueing latency when overload:

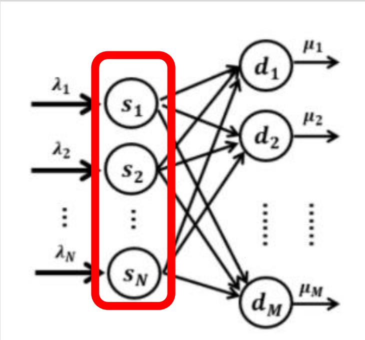


Main Results:

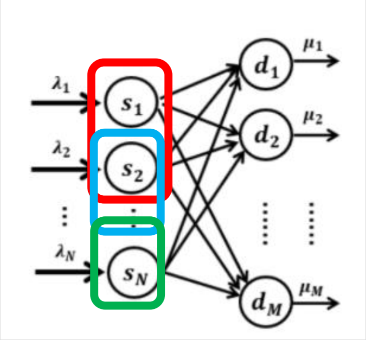
- Setting max rates on all links is generally NOT optimal.
- Properly setting smaller rates reduces latency & saves energy.
- Our algorithm brings **10% reduction** in avg. delay & **50% reduction** in max delay

(2) Fairness

Balancing input loads when egress buffer is bounded:



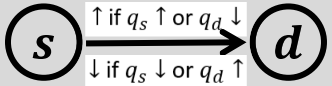
Centralized



Distributed

(3) Stability

A queue-based policy design criterion to stabilize the networks that generalizes a set of policies:

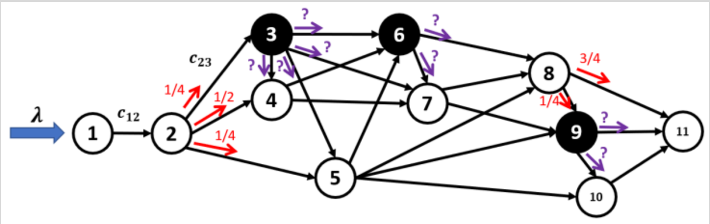


$$\frac{\partial g_{ij}(q_i, q_j)}{\partial q_i} \geq 0, \quad \frac{\partial g_{ij}(q_i, q_j)}{\partial q_j} \leq 0$$

(4) Routing Attack on Causing Overload

Propose algorithms to identify **optimal routing attack** to cause **network overload**:

- Minimize no-loss throughput & Maximize loss
- Critical nodes to protect from overload



Connection to Industrial Research:

- Network control to guarantee high performance under network overload
- Benchmark to forecast network vulnerability to overload under attack