Green Data Centers
Multi Constraint based load balancing to optimize power consumption in Data Centers

Bharath Venkatesh
Truptesh Malangi Nagesh

Objective

- Developing an algorithm for load balancing between servers within a data center based on factors like:

  - CPU Usage

  - Memory

  - Temperature

  - Average Requests in Queue etc..

- Reduce Overall power consumption of a data center through efficient load balancing

- Basic Assumption: A Server repeatedly handling higher loads consumes more power [True?]

Measurements

- CPU Usage
- Memory Consumption
- Number of Requests
- Temperature
- Type of Request – SDN??

Where and When ??

- Where:

    – Load Balancer inside a datacenter

    – Inter data center?

- When:

    – At the time before the packet is scheduled

    – Pre-processing of the Queue?

How??

- An algorithm that assigns dynamic priority to each of the measurement

- Plan to use CAST algorithm to design the time slots at which the measurements are to be taken

- TODO: Investigate CAST and other methodologies

- Modify the scheduler to schedule based on the obtained priority

Management Loop

Goal:
To minimize overall
Power consumption

System under Test:
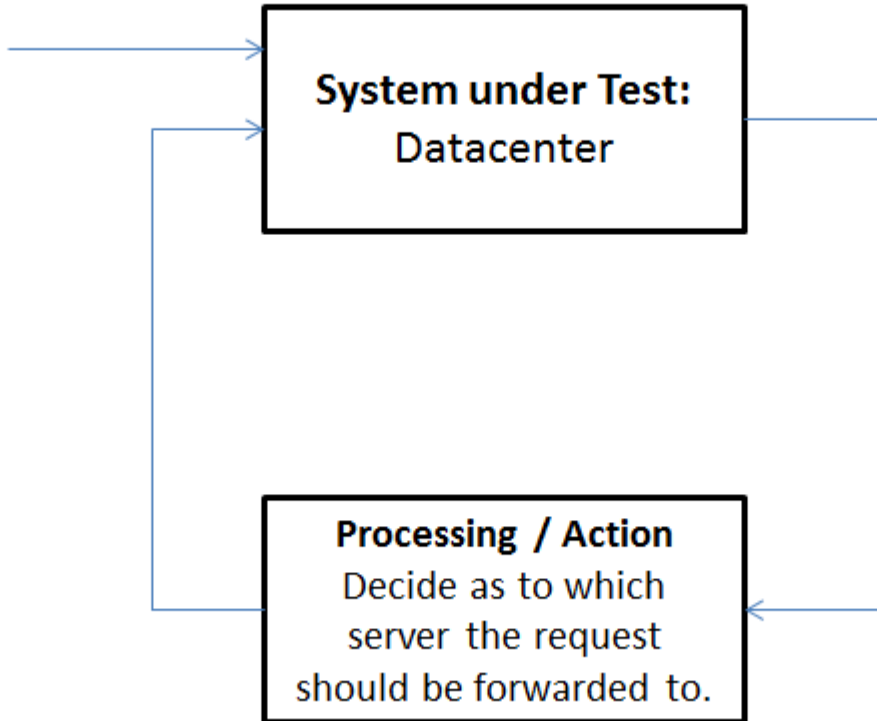Datacenter

Measurements:
Temperature
CPU
Memory
No. of requests in queue

Processing / Action
Decide as to which
server the request
should be forwarded to.

SDN as means of achieving

- ## With SDN:
  - Obtain Measurements using syslog and other api's
- ## In contrast: Without SDN
  - Traditional methods using CLIs to manually obtain these measurements periodically
  - Or run scripts that fetch these values

Inputs Required

- Novelty??
- SDN ???
- Feasibility??
- Usefulness?