

# Closer to the Cloud

## - A Case for Emulating Cloud Dynamics by Controlling the Environment

Ashiwan Sivakumar

Shankaranarayanan P N

Sanjay Rao

School of Electrical and Computer Engineering

Purdue University

# Latency-Sensitive Applications on the Cloud

- Complex
- Multi-tiered
- SOA
- Stringent SLA

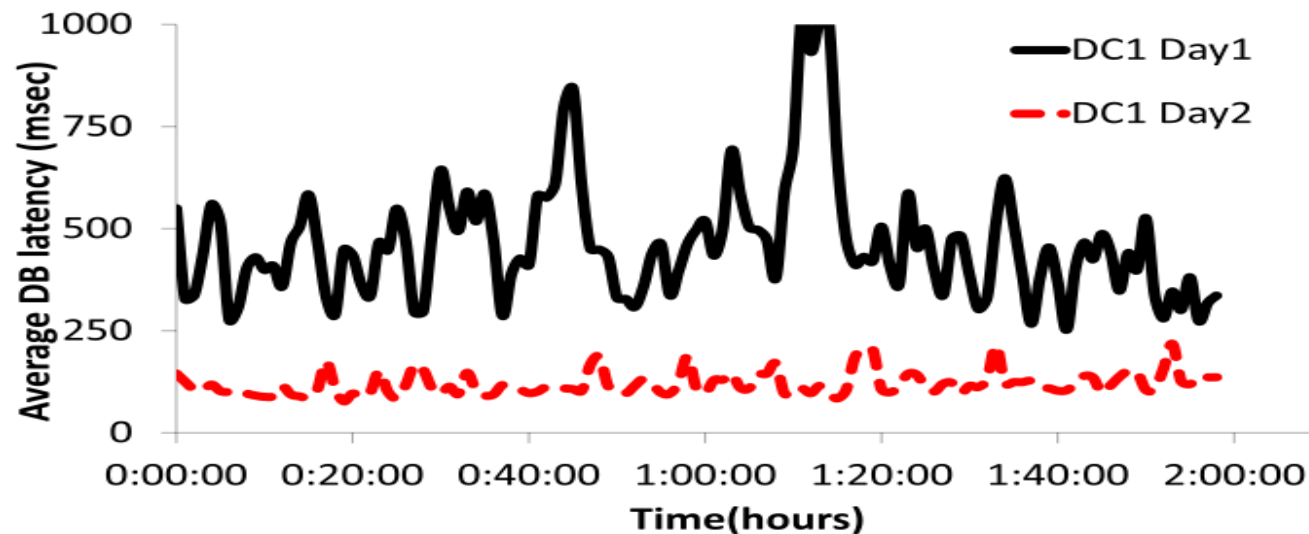


- \$\$\$ savings
- Geo-redundancy
- Elasticity
- Scalability



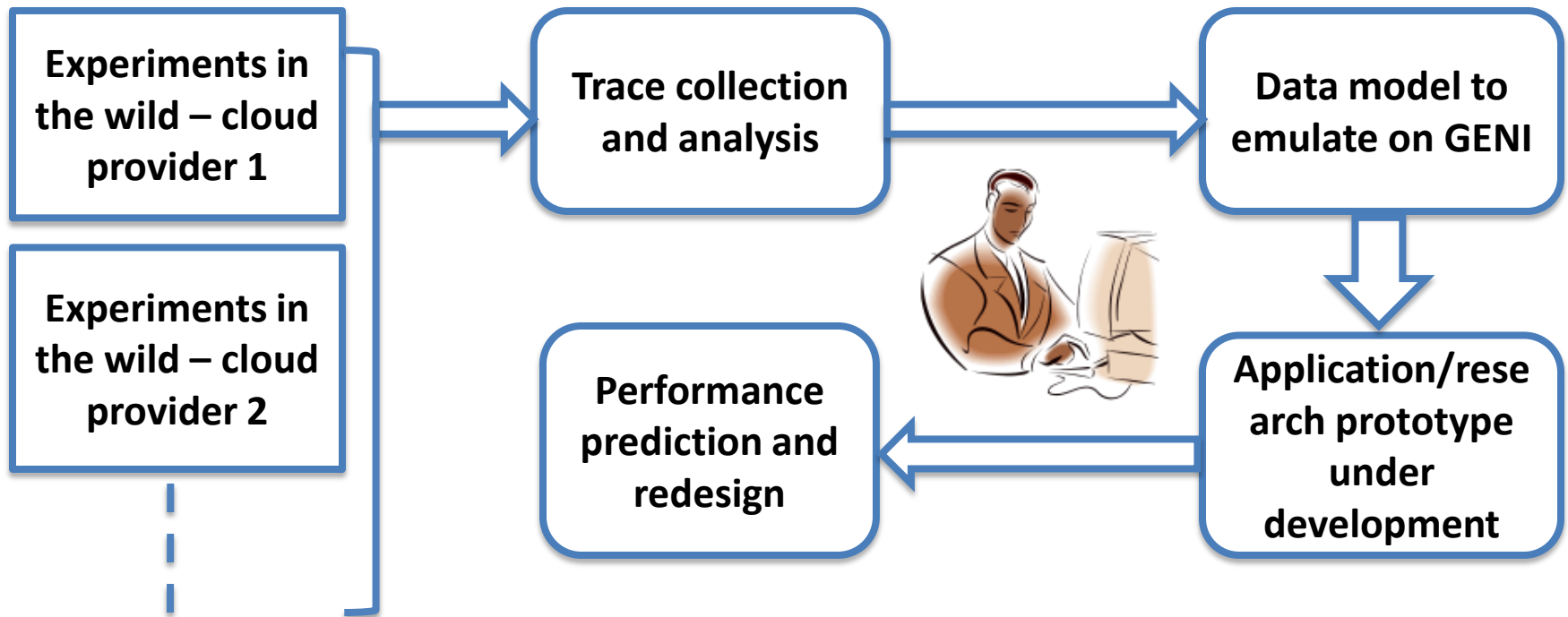
Test beds

# Performance Fluctuations in Microsoft Azure



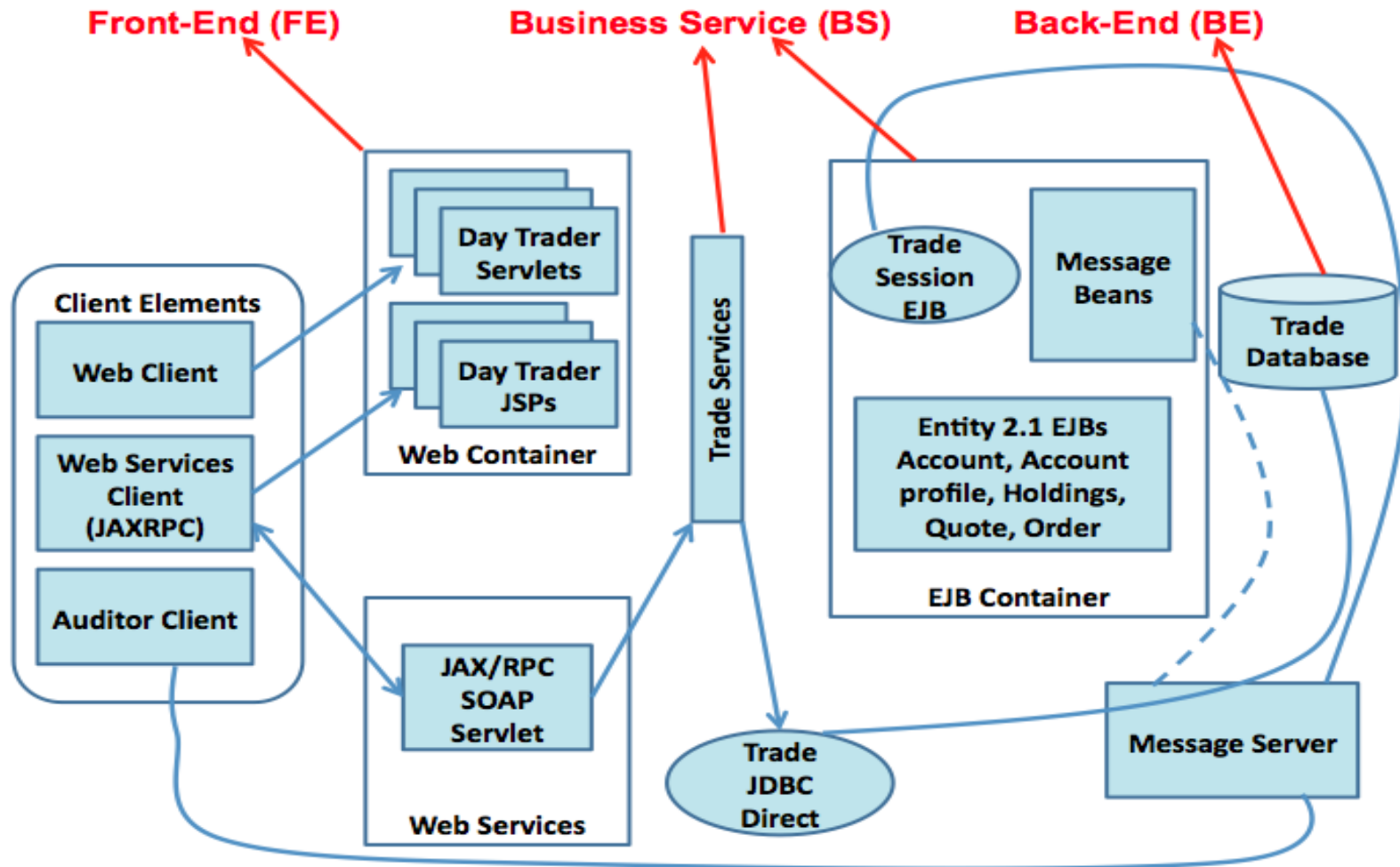
- Short-term and long-term variation in DB latency
- DB performs 100% worse on Day 1 than on Day 2
- Variation in RTT of TCP streams observed in Amazon EC2
- Our study confirms - hotcloud10 (CloudCmp), SoCC10 (YCSB)

# Control over the Environment : A Technique

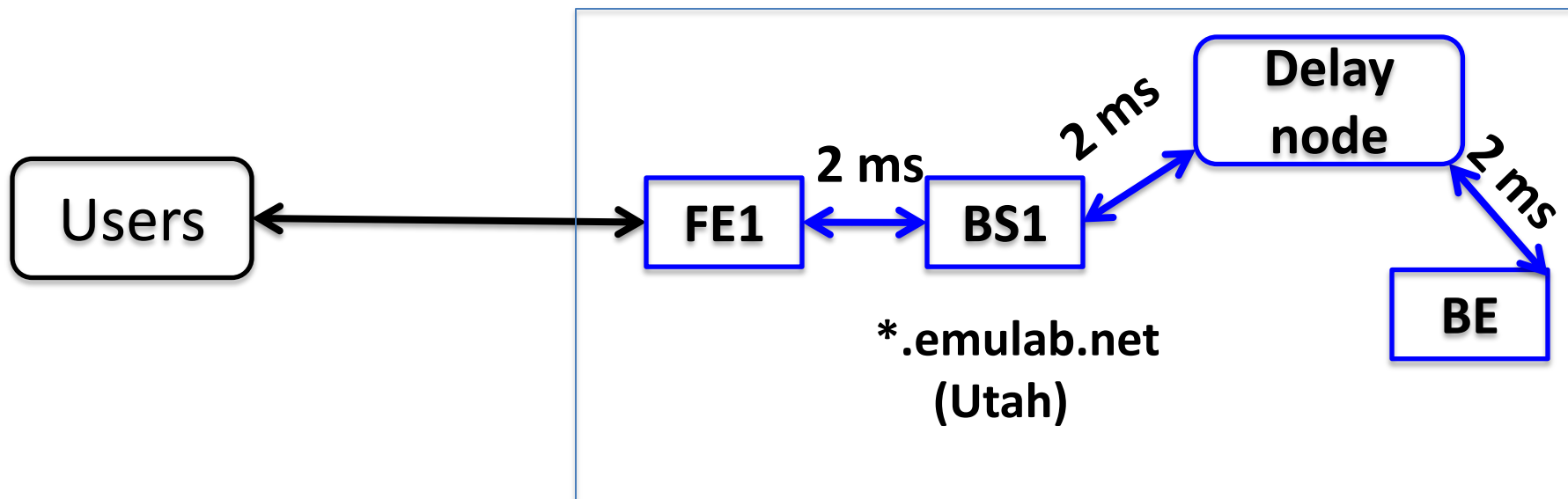


- Commercial clouds ?
  - No control over the environment and shared infrastructure
- Cloud test beds like Open Cirrus, Eucalyptus?
  - Can emulate cloud dynamics and federated, but no control over the environment

# Sample Enterprise Application Day Trader



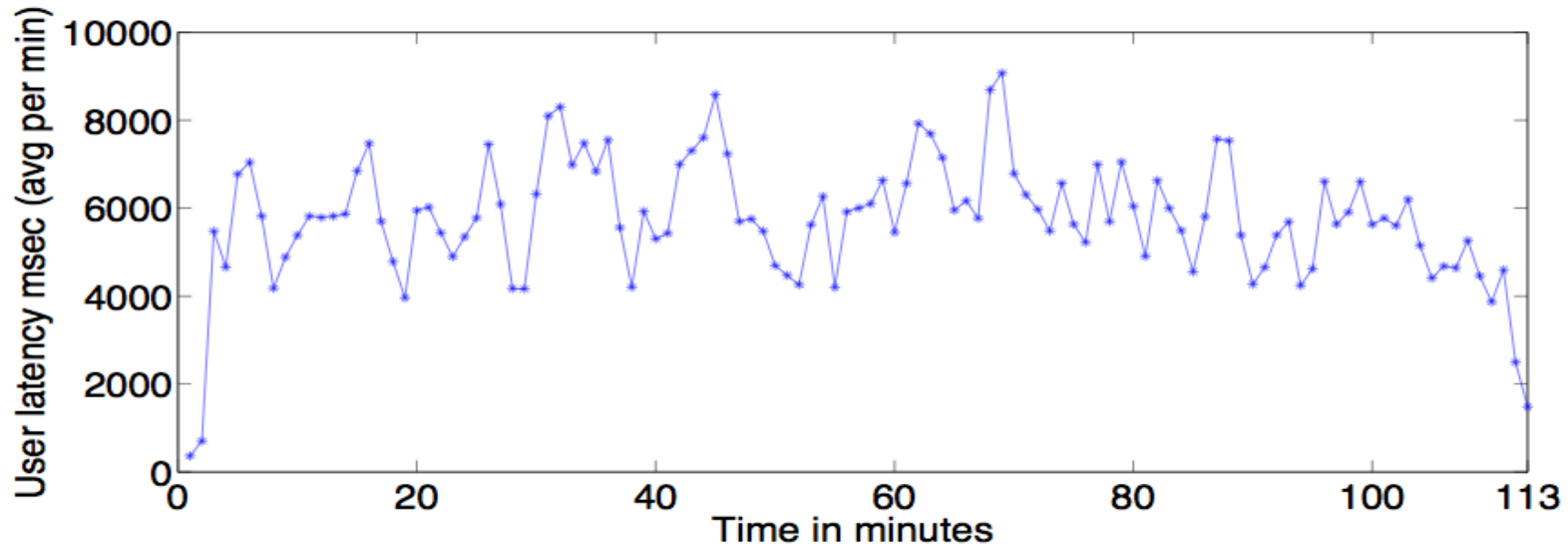
# Single Data Center Deployment on GENI



## ➤ Test bed

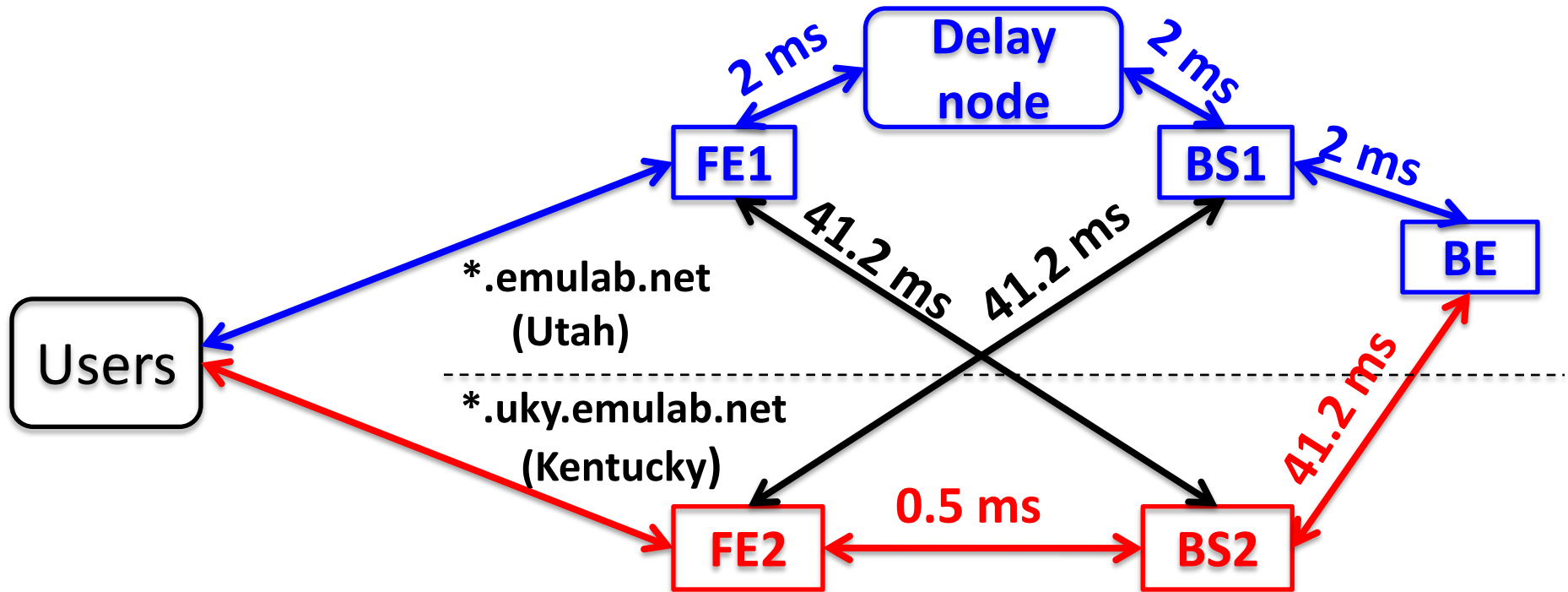
- User workload generator (grinder) running in Purdue
- Workload from DaCapo benchmark for Enterprise apps
- Application instances running on nodes from Utah
- Delay node between BS1 and BE

# Predicting Performance of Applications by Replaying Trace



- DB latency trace from Azure (slide 3) replayed between BS and BE
- Total application response time for each user request plotted
- More or less follows network delay between BS and BE
- Also measured error percentage – always within 1 %

# Multi Data Center Deployment on GENI

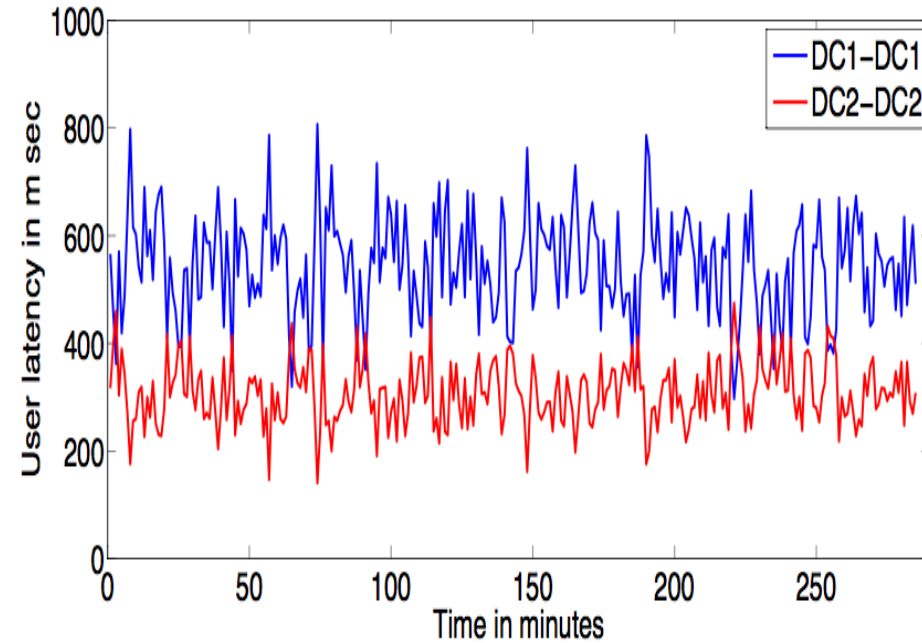
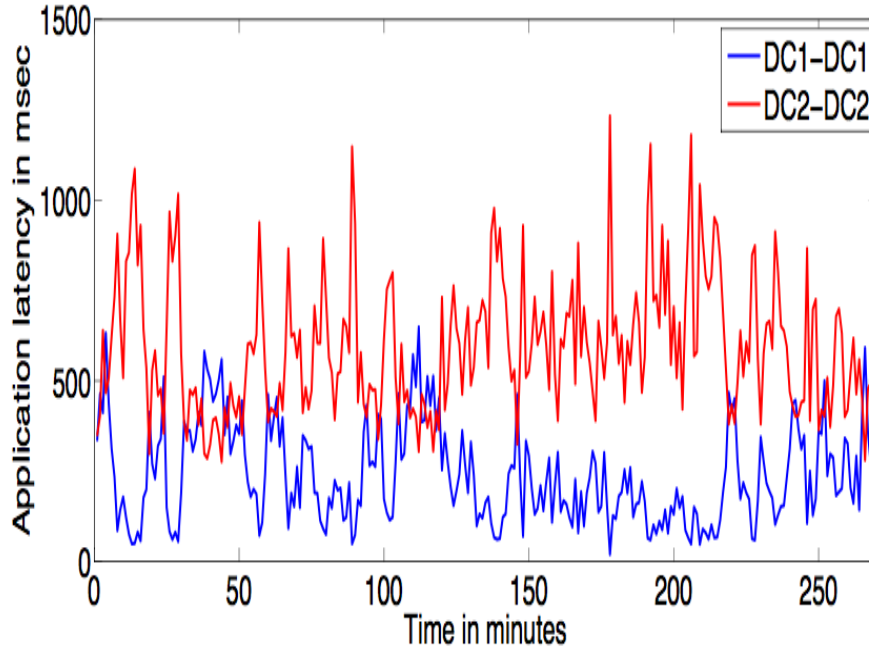


## ➤ Test bed

- User workload generator in Purdue
- Two instances of the application components
- One deployment in Utah and another in Kentucky
- Delay node between FE1 and BS1 in Utah



# What –if Analyses on Application



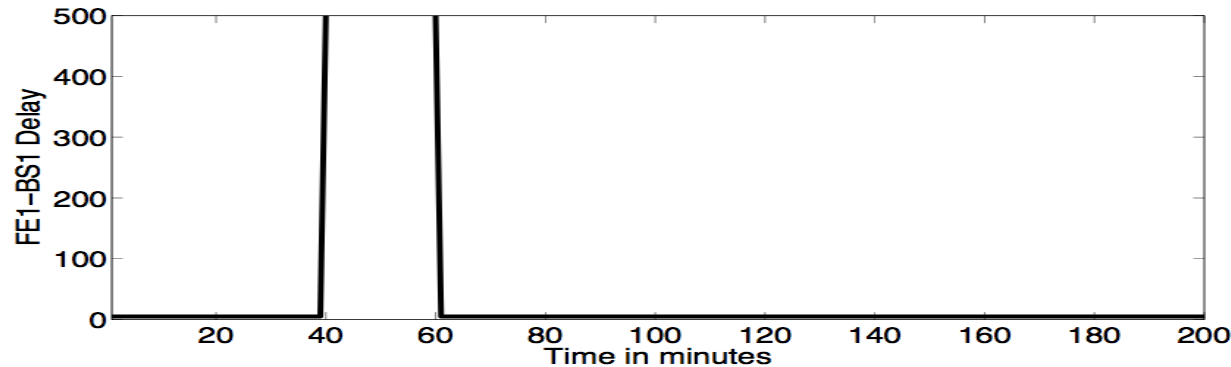
- **No delay** : DC2 slower than DC1
- **Delay 250 ms** : DC1 slower than DC2
- DC1 – DC1 : FE1 – BS1 in Utah & DC2 – DC2 : FE2 – BS2 in Kentucky

# An Adaptive System - Dealer

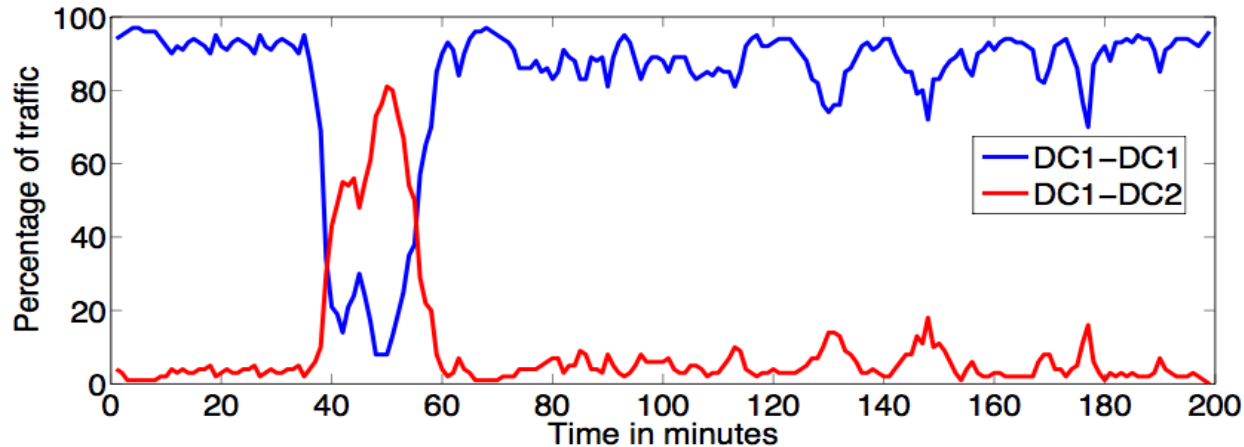
- Cloud environment is highly dynamic
  - Problem isolated to one component in a DC
- Current redirection mechanisms
  - Abandon entire deployment
- **Dealer** : Adapts to cloud dynamics
  - Components deployed on multiple DCs
  - Fine-grained component level redirection
  - Optimally suggests path based on latencies

# Evaluation of the Adaptive System

## ➤ Control Input : Step



## ➤ Output : Dealer path change



## ➤ DC1 – DC1 : FE1 – BS1 & DC1 – DC2 : FE1 – BS2

# Conclusions

- To open up new avenues in cloud research
  - Control over test bed environment – critical
- Can be used as a technique by developers and researchers for
  - emulating real cloud dynamics to
    - predict application performance
    - compare cloud providers
    - what-if analyses on application
    - develop solutions for performance problems

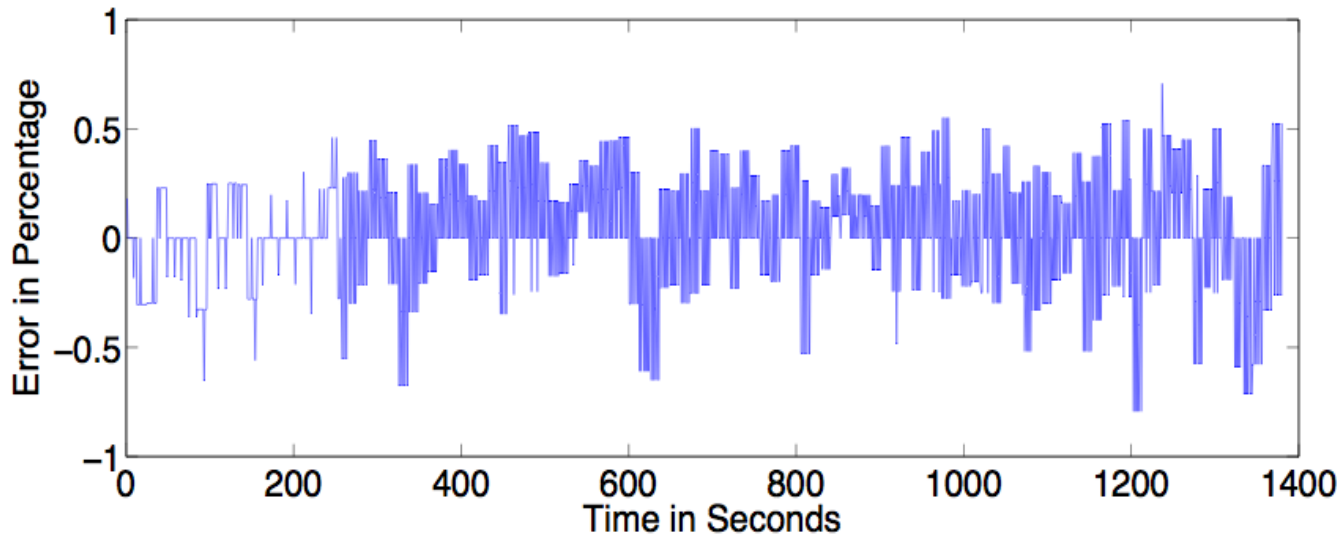
Q & A

Back up slides

# Future Directions

- Dynamically change the delay
  - Multi data-center deployment like Emulab
- Emulating cloud data stores
  - Blobs, Queues, Big-Table provided as service
- Installing and running Cassandra on GENI

# Accuracy and Repeatability Requirements



- Error percentage :
  - $(\text{observed latency} - \text{expected latency}) / \text{expected latency}$  in %
  - observed – ping , expected – delay applied
  - always within +/- 1% error bound
- Require such high degree of accuracy and repeatability



# SQL Azure Performance Issue Snapshot (6 Days)

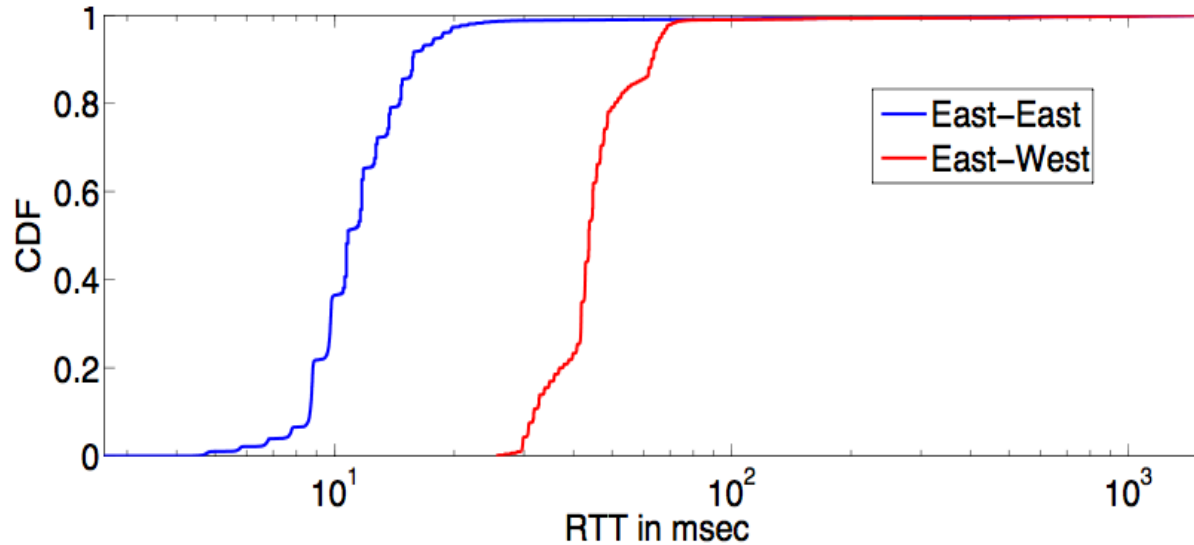
The screenshot shows the Windows Azure Platform Status page in a Mozilla Firefox browser. The page displays the status of various services. A table lists Windows Azure Storage services across different regions, all showing 'Service is running normally.' Below this, the 'Status History' section is selected, showing a performance issue for SQL Azure Database in North Central US starting on Sep 21. A detailed incident report is visible, stating: '5:28 PM UTC A small number of customer databases have been impacted by a performance issue triggered by load on the system. We are currently working on a solution to improve the service and reduce the impact to our customers. If you are impacted by this, please contact us so that we may assist you.' The page also includes a legend for service status icons and a footer with links to features, scenarios, and support.

Service [Sub-Region]	Sep 26	Sep 25	Sep 24	Sep 23	Sep 22	Sep 21	Sep 20
SQL Azure Database [East Asia]	✓					✓	✓
SQL Azure Database [North Central US]	⚠					⚠	✓
SQL Azure Database [North Europe]	✓					✓	✓
SQL Azure Database [South Central US]	✓					✓	✓
SQL Azure Database [Southeast Asia]	✓					✓	✓
SQL Azure Database [West Europe]	✓					✓	✓

Page Last Updated: 26 Sep 2011 5:09 PM UTC

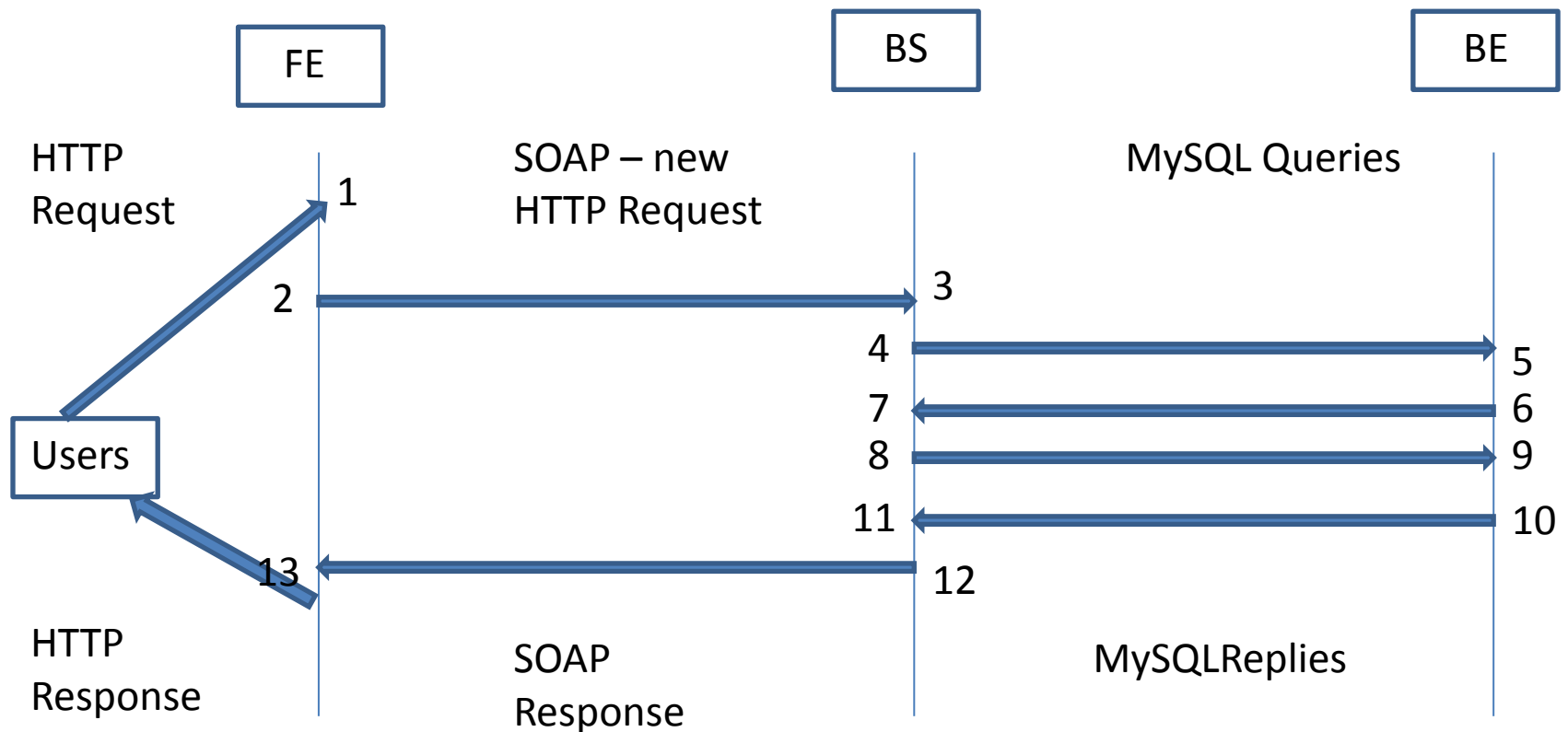
Legend: ✓ Normal service availability, ⚠ Performance degradation, ✖ Service interruption, ⓘ Additional information

# Performance Fluctuations in Amazon EC2



- Simple Experiment
  - RTT measured for a TCP stream
  - Both inter and intra DC show variation

# Day Trader Communication Pattern



- Response time = component latency + communication latency
- Multiple calls between components for a given user request