Towards a Network Aware VM Migration: Evaluating the Cost of VM Migration In SDN-Based Cloud Computing Network

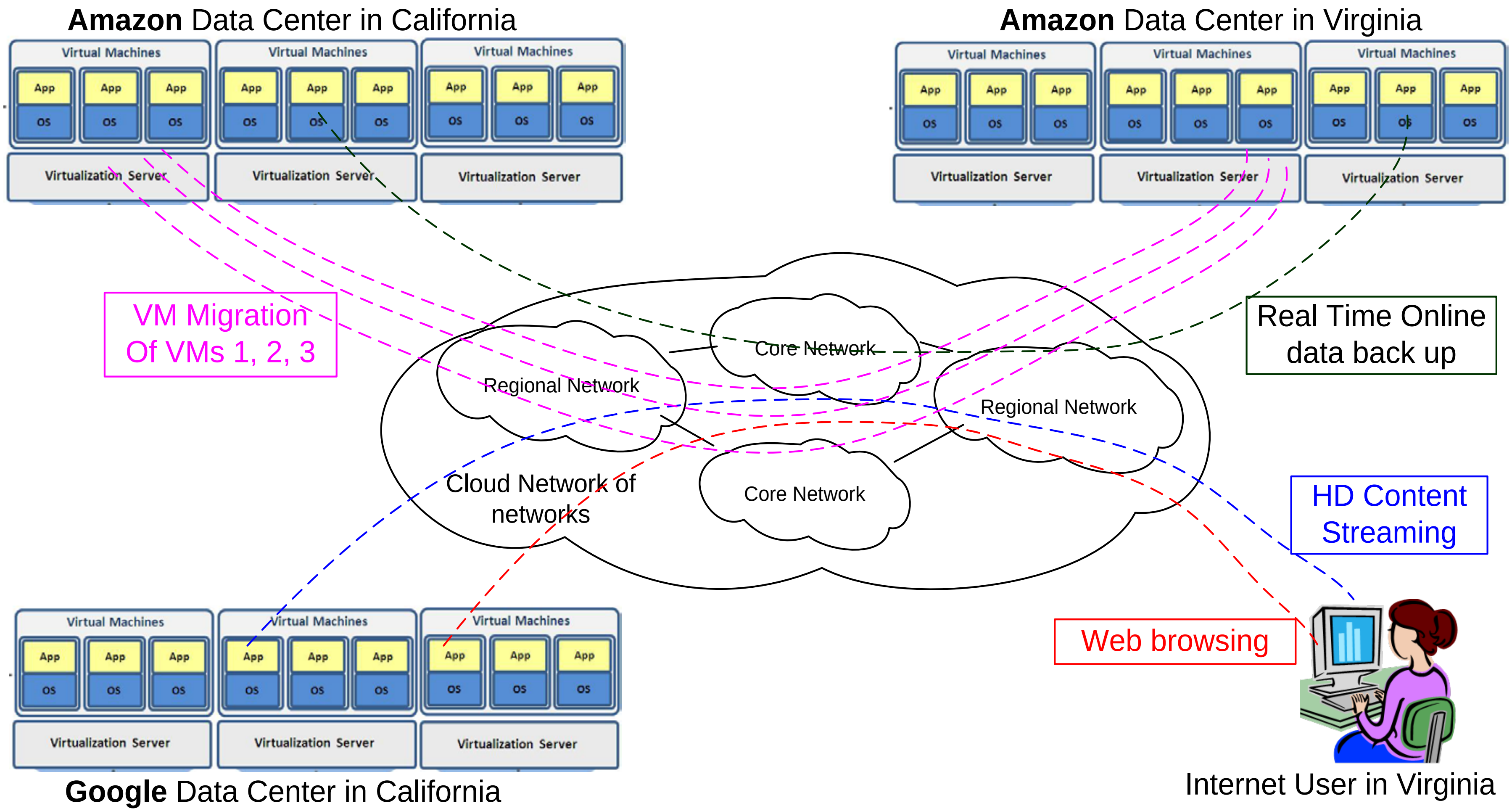
Sachin Shetty
Associate Professor
Department of Electrical and Computer Engineering
Tennessee State University
Nashville, TN

June 16, 2016

Motivation

- Migrating VMs is a good defense mechanism but comes at a network performance cost
  - Consume nearly the entire bandwidth
  - Affects performance of competing applications
- Cloud applications deadlines must be met [Kohavi et. al. SIGKDD, 2007]
  - 100 ms additional latency on Amazon EC2 = 1% drop in sales
  - 500 ms increase in Google search time = 20% drop in revenues
- VM Migration process is unreliable and only 87% of migrations successfully finish [Tomas et. al., TSP, 2015]

The Cloud Network – Resource Sharing



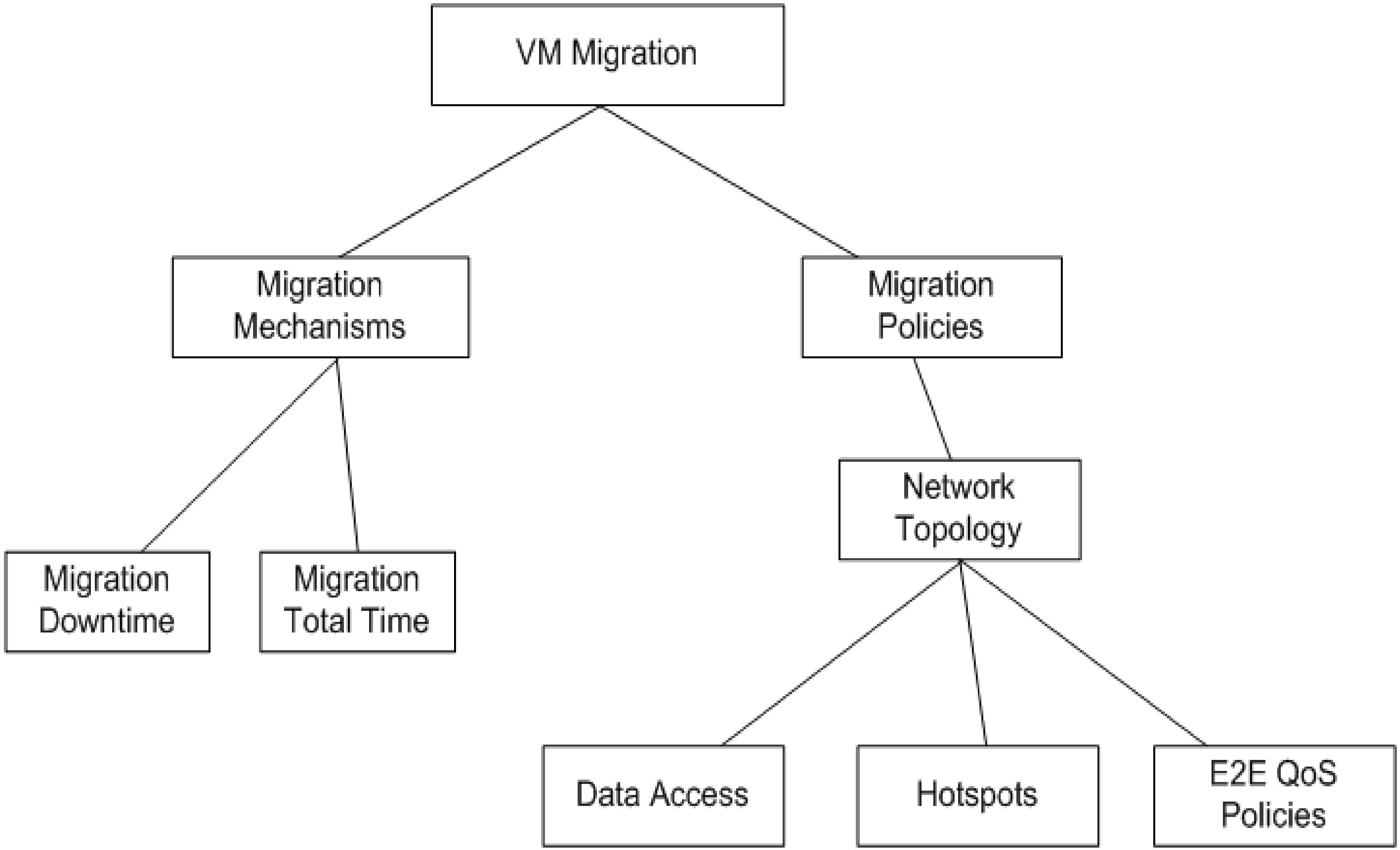
Research Challenge

- Need for a framework to quantify the cost of VM Migration
- Nneed for a mechanism that intelligently allocates network resources during VM Migration
- Ensure a fair share of resources during VM Migrations
- Minimize degradation of network performance

The Problem

- Live VM Migrations are expensive
  - Nearly consumes entire bandwidth.
  - Impacts performance of competing flows
- Vendors
  - Recommend separate network for VM mobility – costly
- Know the cost of VM Migration
- Fair share of resources during VM Migration
  - Migration policies, Network topology, Data access, QoS policies
  - Cost estimation model $N = M \frac{1-(R/L)^{n+1}}{1-(R/L)}$ – simulated

Related Efforts



Related Work - Remedy

- Builds on pre-copy migration technique [V. Mann, Networking 2012]
  - Used by VMware vMotion and KVM Live Migration
  - Given the Memory size M of a VM in MB,
  - Page dirty rate R of a VM in MB/s,
  - Bandwidth of the link used for migration L in MB/s,
  - Switchover goal time T and
  - Minimum required progress amount X in MBs,

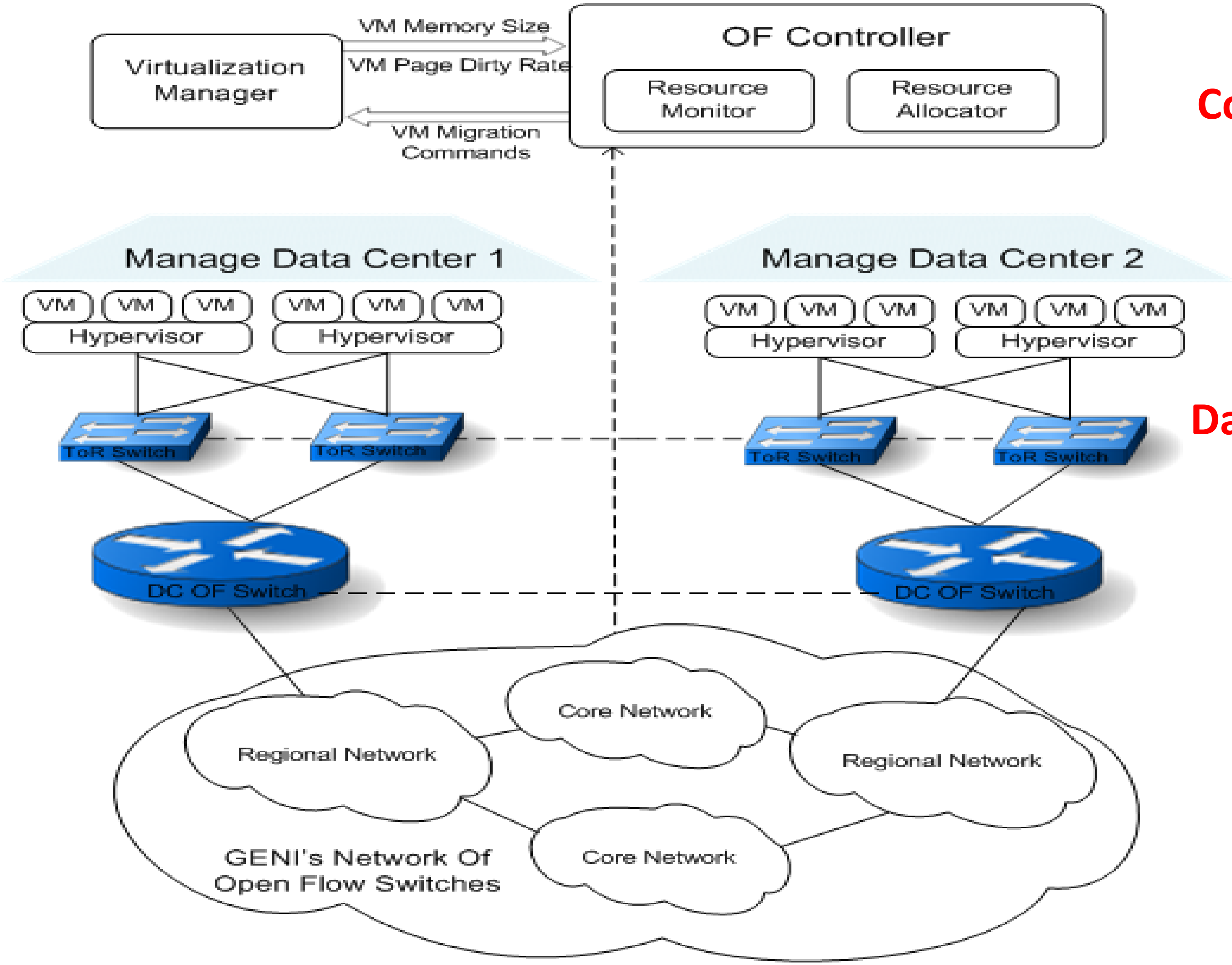
Related Work - Remedy

- Remedy cost Model is described as:
  - The number of pre-copy cycles: $n = \min\left(\left\lceil \log_{R/L} \frac{T.L}{M} \right\rceil, \left\lceil \log_{R/L} \frac{X.R}{M(L-R)} \right\rceil\right)$
  - Total traffic generated by the migration: $N = M \frac{1-(R/L)^{n+1}}{1-(R/L)}$
  - Total migration time: $Tm = N / L$
  - Time spent on a stop-copy transfer $W(L) = \frac{M}{L}\left(\frac{R}{L}\right)^{n}$
    - Downtime experienced

Goals and Objectives

- **Goal:** Compute cost of VM Migration and ensure successful VM Migration without degrading performance of network applications in a SDN based cloud computing network
- **Objectives:**
  - Empirically evaluate Remedy cost estimation model on GENI testbed
  - Openflow Controller to enforce QoS and fair share of resources during VM migration
  - Ensure VM Migration completes successfully
  - Improve the performance of competing flows

Methodology



Controller

Data Center

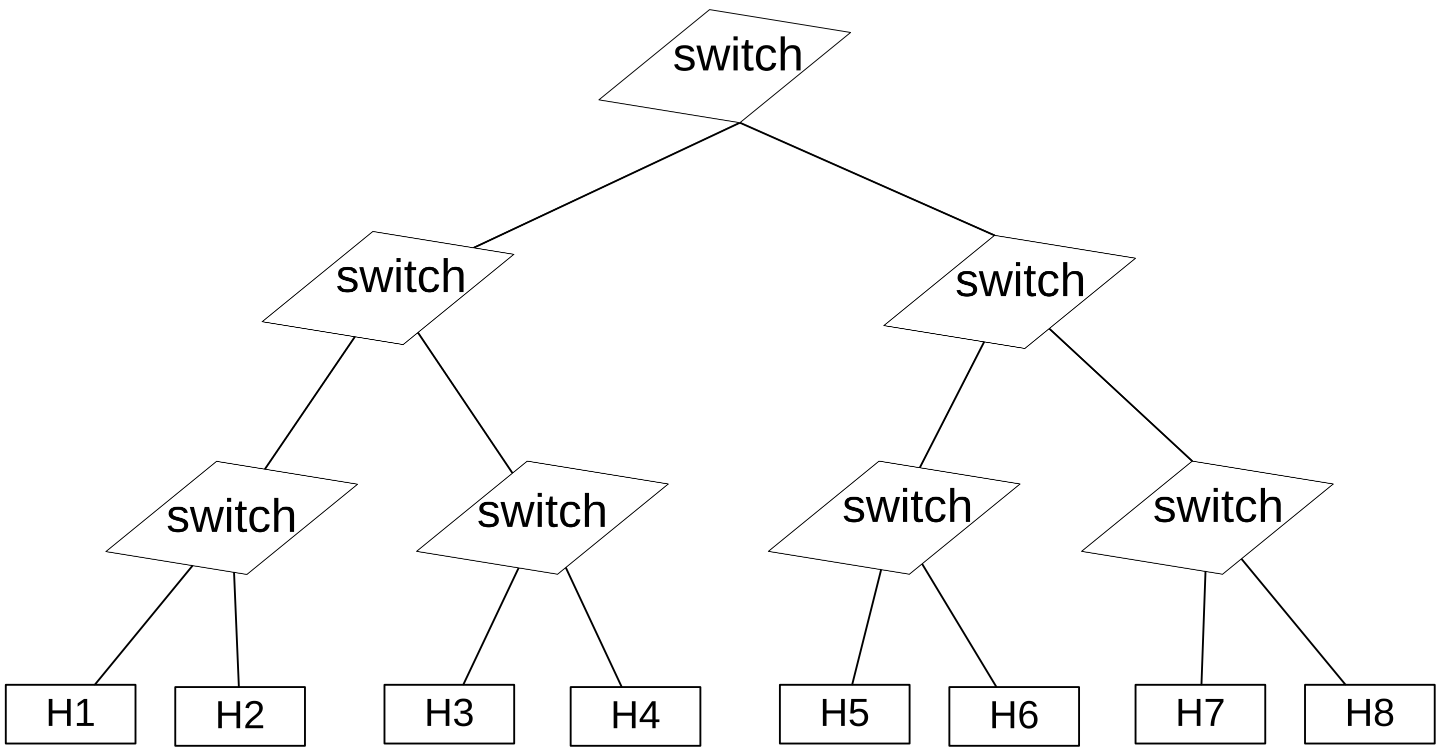
Core

Methodology

- Design and implement an OpenFlow controller
    - Monitors network utilization
    - Calculates the cost of VM Migration on all network paths
        - Cost of VM migration, Network topology and Network traffic
        - Optimal bandwidth for VM Migration flows
        - Enforce end-to-end QoS policies to reserve optimal Bandwidth
- Emulate data center networks by utilizing computing and networking resources on GENI infrastructures

Experiment Setup

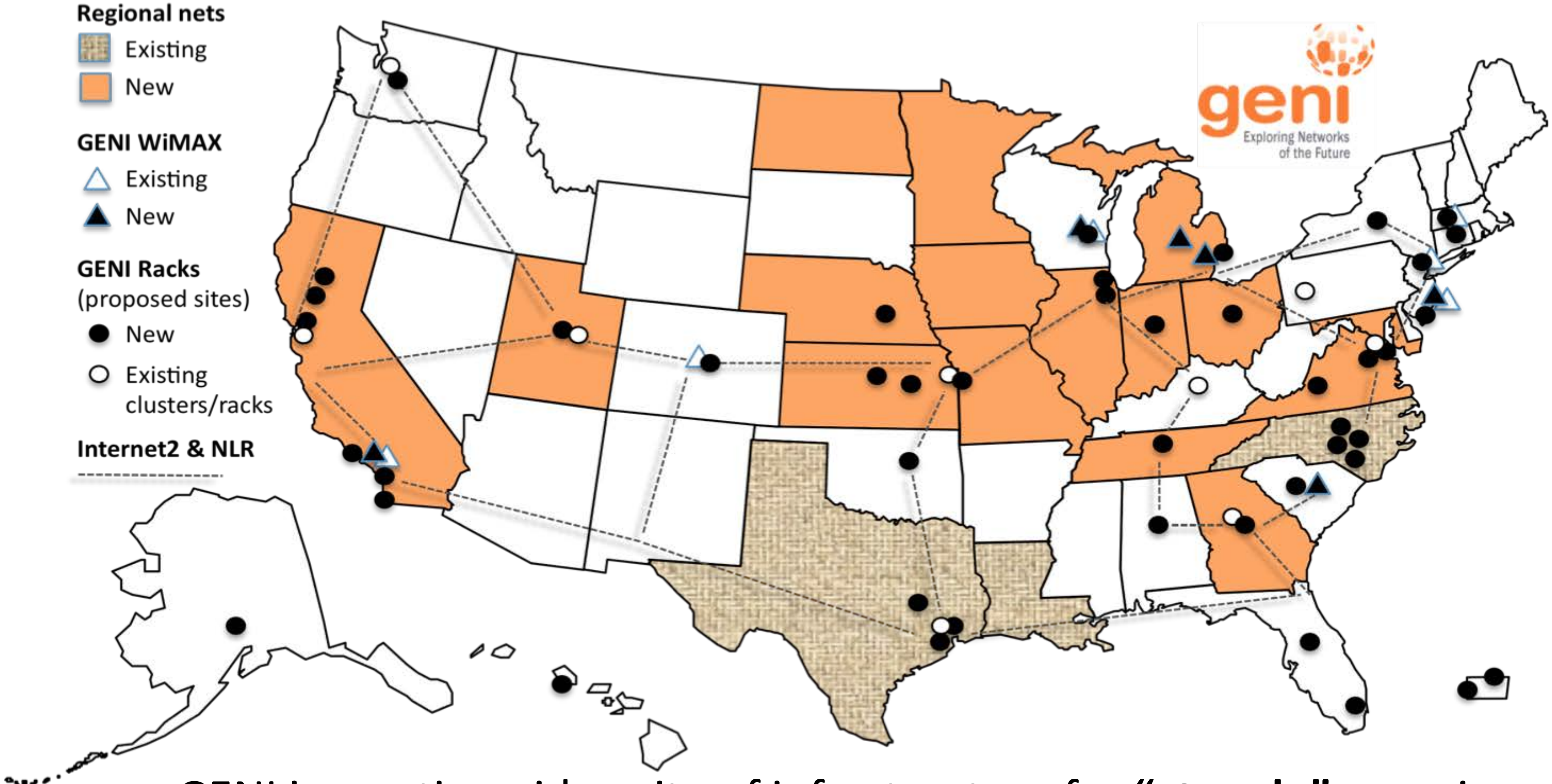
- Reserving resources: Data center, Core, OF controller
  - 8 kvm hosts and 7 ovs switches
  - NFS server for shared storage



Experiment Setup

- Virtualization Manager
  - Libvirt toolkit
  - Ubuntu 12.04 Desktop VM, 2048 MB Memory on H1
  - Openvswitch bridged networking
  - Manage, Clone, Live-Migrate VMs across the data center
- VM Workload and Network Flows
  - RUBiS Workload
    - Auction site prototype modeled after eBay.com
    - 50 concurrent clients: average page dirty rate of 500 pages/second
  - Ping flows
  - iperf flows

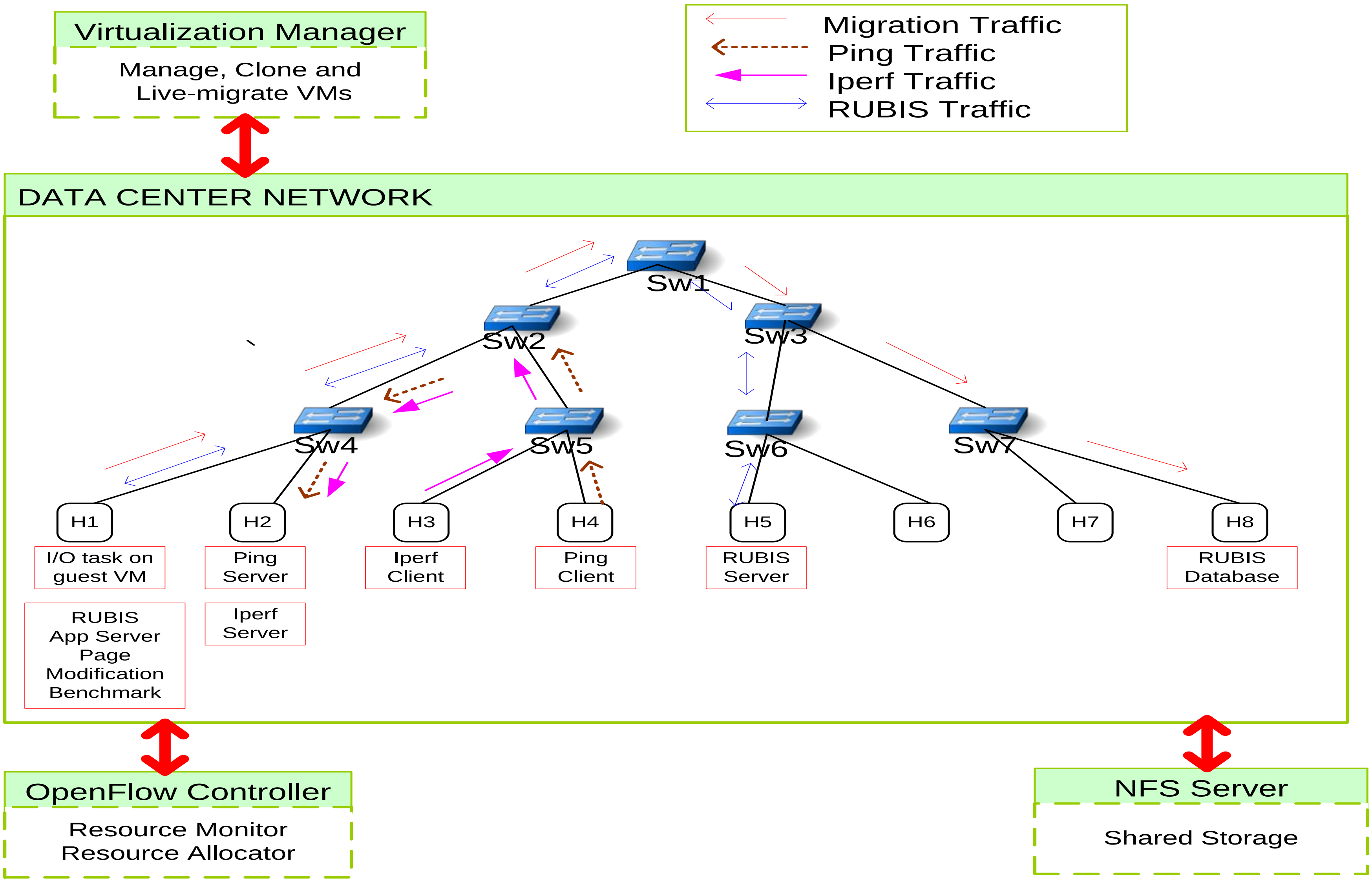
Global Environment for Network Innovations (GENI)



- GENI is a nationwide suite of infrastructure for **"at scale"** experiments in networking, distributed systems, security, and novel applications.
- GENI provides compute resources that can be connected in experimenter specified Layer 2 topologies.

Source: GENI Exploring Networks of the Future, www.geni.net

Evaluation in GENI Testbed



Experiment: Resource Manager

- Bandwidth Monitor
  - Discover all links in the network
  - Poll for port and flow statistics from all switches
  - Discover heaviest flows and links
  - Discover available bandwidth in the network
- Memory Monitor
  - Virtualization manager: VM memory size, page dirty rate

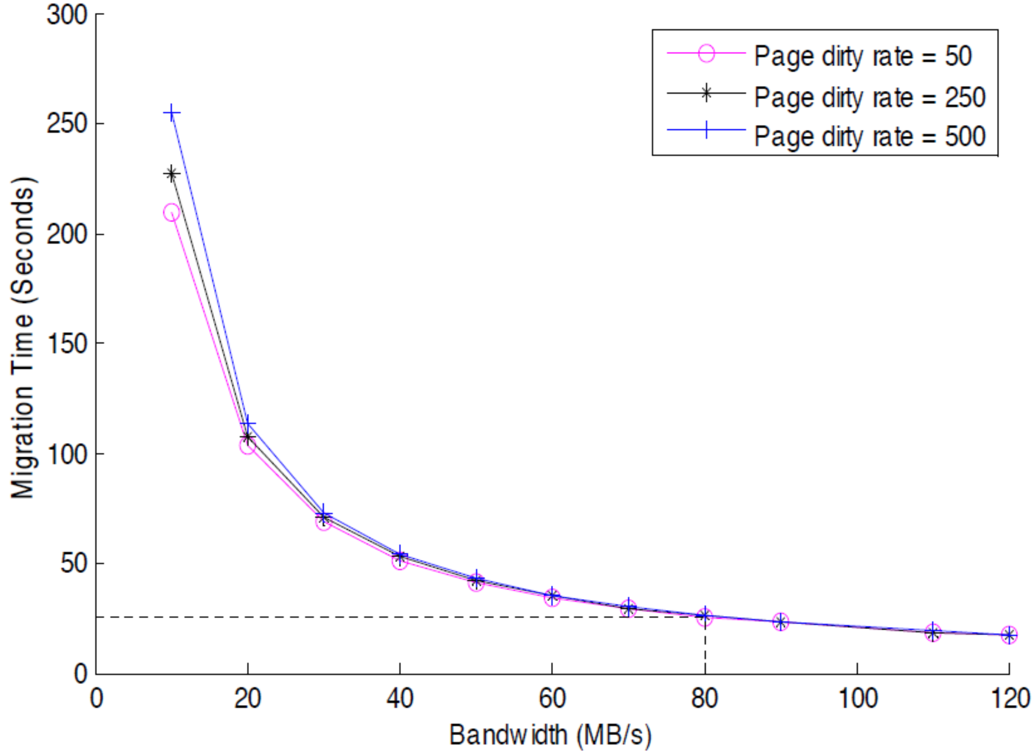
Experiment: QoS Enforcer

- OVS capabilities: create 2 queues on each port of OpenFlow switch
  - Egress traffic shaping

- Cost of Migration Estimation – Migrate H1 to H8
  - User specified VM migration deadline (25 seconds)
  - VM's page dirty rate (500MB/s)
  - Memory size (2048 MB),
  - User specified minimum progress amount (100MB)
  - Switchover goal time (100ms),
  - Minimum bandwidth used by the migration (80MB/s).

Experiment: QoS Enforcer

- Bandwidth of dedicated queue = estimated minimal value.
- OF rule to Enqueue migration traffic to dedicated queue.
- Predicted Migration Times



Results

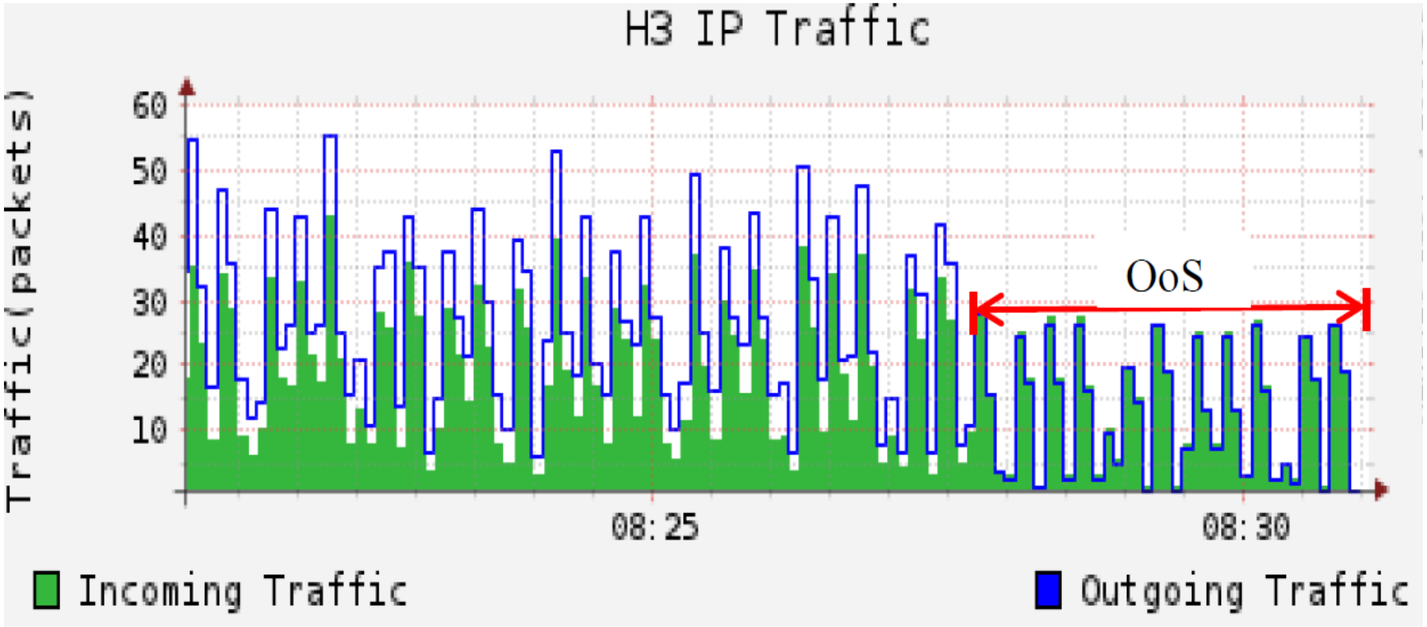
- Actual Migration Times and Average Ping Latencies

Reserved Bandwidth (MB/s	Migration Elapsed Time (s)	Average Ping Latency (ms)
No reservation	19.465	19.059
80	19.576	10.066
90	19.436	5.865
100	19.116	3.31
120	18.896	2.4844

- Increasing bandwidth reservation has minimal effect on migration time, but heavily improves ping flow performance (47.18%)

Results

- Impact of QoS enforcer on iperf flow



- Incoming traffic throttled in the absence of QoS policy
- Enforcing QoS policies in terms of bandwidth reservation relieves the network of possible overloads during migration

Results – Evaluating Remedy

- Which parameters are more critical?
  - Page dirty rate R,
  - Available bandwidth L,
  - User specified parameters progress amount X,
- Conditions under which the model works/fails
  - $L \geq 1$Gbps
  - $R \leq 3000$ pages/seconds
- Higher R, Lower L : Model fails to perform
  - Irrespective of VM memory size

Varying Page dirty rate (R), (1 GB VM)

Migration deadline = 330 s, L=0.1Gbps

R (Pages/s)	Remedy Bandwidth (Mbps)	Our Bandwidth (Mbps)
50	30	33
250	38	41
500	45	51
1000	60	70
2000	90	103

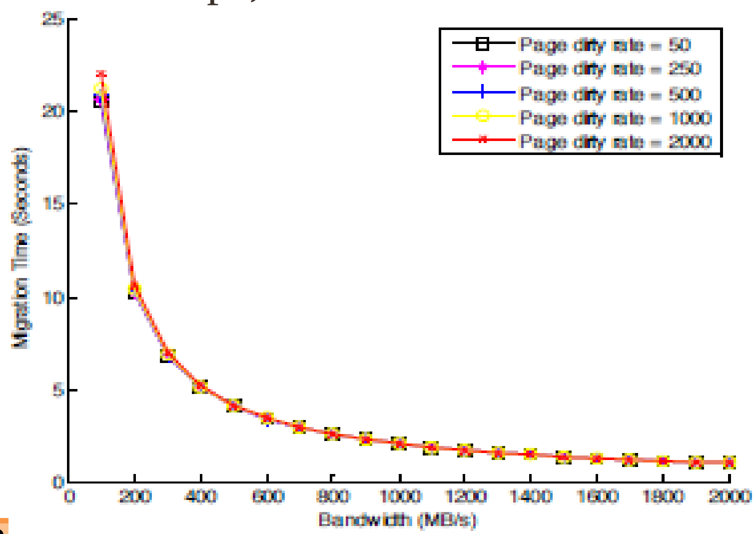
Migration deadline = 80, L=1Gbps

R (Pages/s)	Predicted Bandwidth (Mbps)
50	112
250	120
500	128
1000	144
2000	160

Migration deadline = 8 s, L=10Gbps

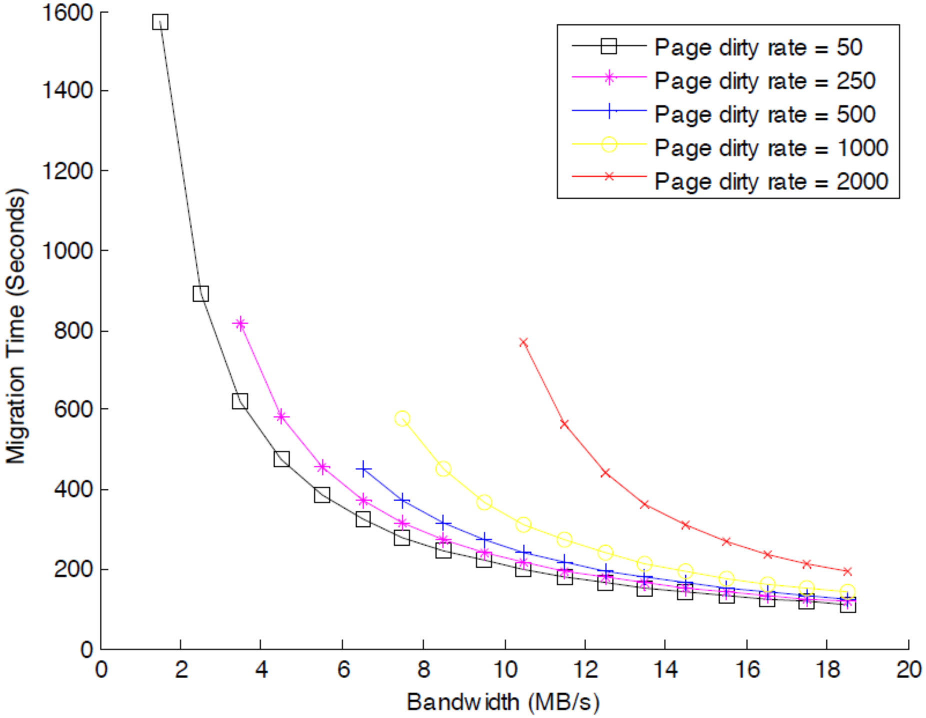
R (Pages/s)	Predicted Bandwidth (Mbps)
50	1144 same
250	1144 same
500	1144 same
1000	1144 same
2000	1144 same

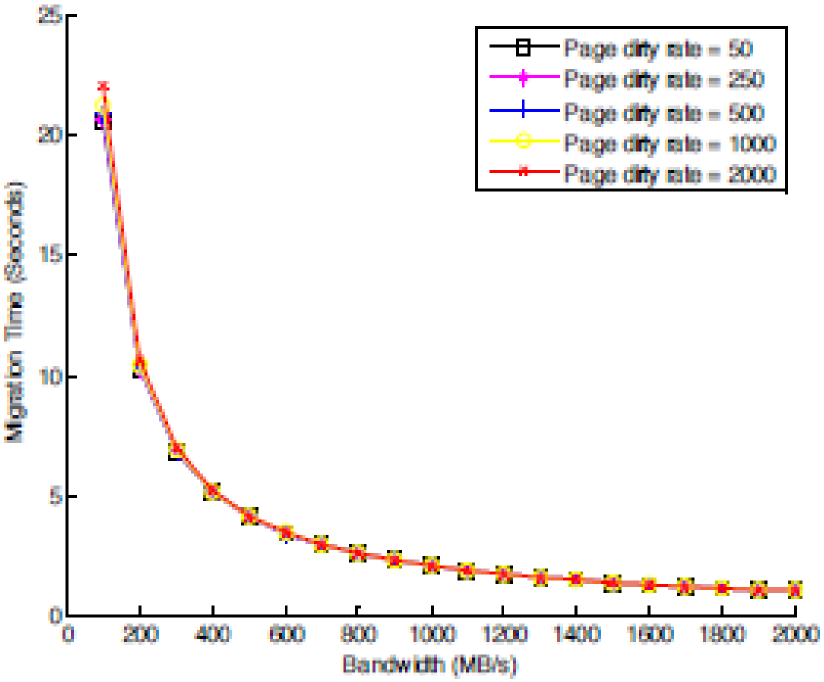
L=15Gbps, M=2GB



Available bandwidth, L

- Dictates range of user migration deadline



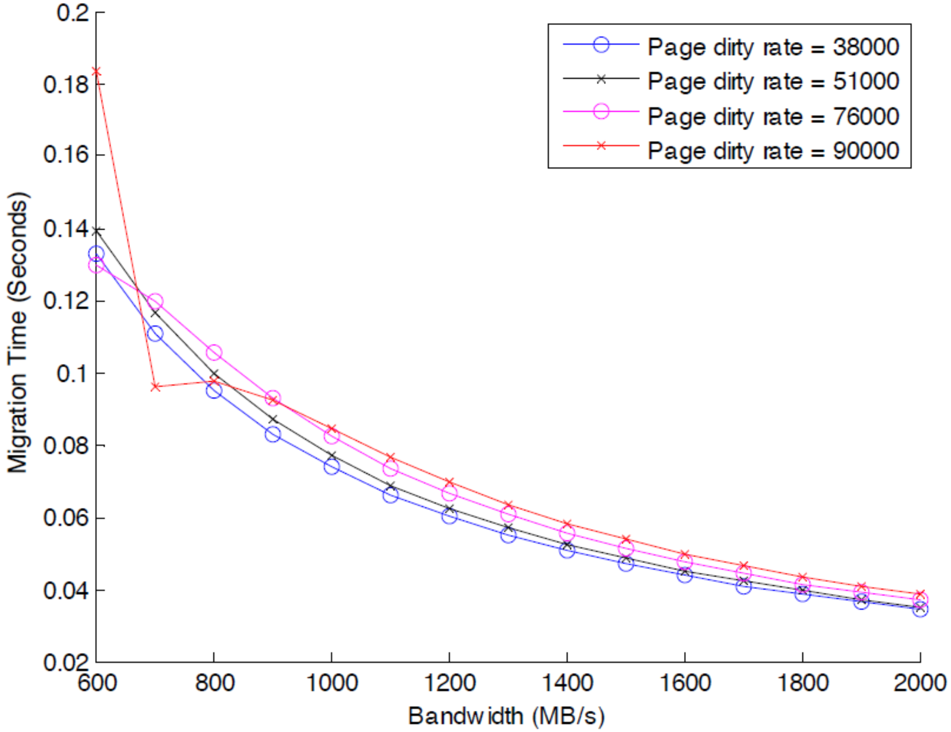


- For 20 MB/s link, and R=2000 pages/s
  - You can only reserve 11-18 MB/s
  - Allowable deadline 200-800 s
- For 2000 MB/s link, and all values of R
  - You can reserve 0-2000 MB/s
  - Allowable deadline 2-23 s

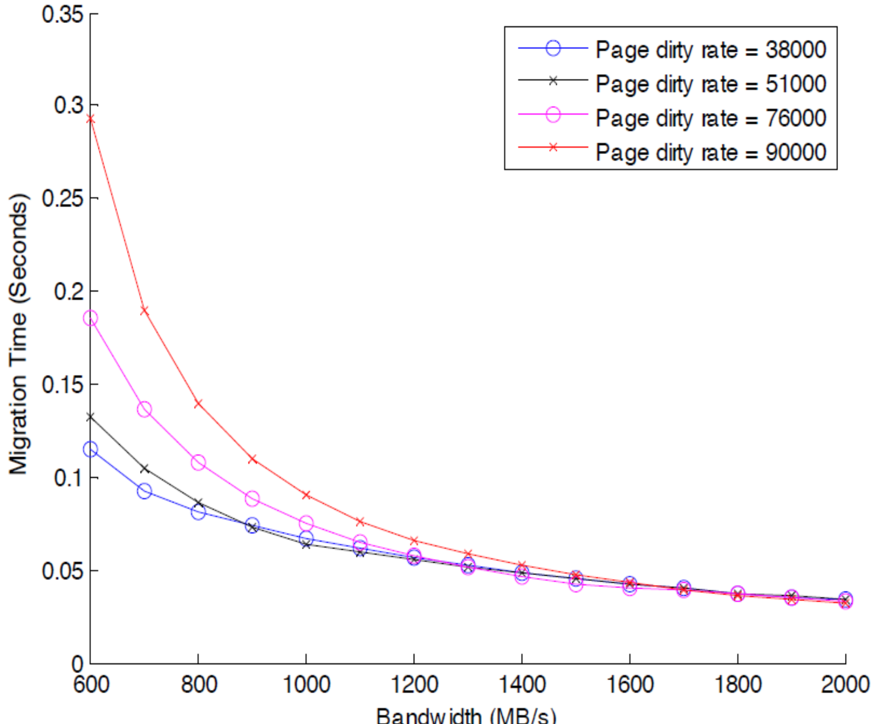
User specified progress amount, X

- Corrects discrepancies in the model caused by:
  - Insufficient bandwidth, L
  - Higher page dirty rates, R

X=50MB, M=64MB, T=0.1s



X=250MB, M=64MB, T=0.1s



User specified progress amount, X

- Range of X under which the model works/fails (R=10,000 pages/s)

Memory size (MB/s)	Range of X To select MB/s	Range of X NOT To select MB/s	Range of X To select MB/s
64	0-30	30-500	> 500
512	0-100	100-1250	> 1250
1024	0-250	250-2000	> 2000
2048	0-500	500-5000	> 5000

Ongoing and Future work

- Significance of cost estimation model parameters
  - Modify parameters to suit types of applications, optimize model
- Performance of model on a wide area network
  - Variable link bandwidths – QoS, OF
  - Migration across Multiple Data Centers
    - GENI - USA
    - iMinds - Europe
- Deploy the model to compute cost of VM Migration based Moving Target Defense technique in the cloud computing network

REFERENCES

1. R. Kohavi et al., “Practical guide to controlled experiments on the web,” *ACM SIGKDD*, 2007
2. Mann, Vijay, et al. "Remedy: Network-aware steady state VM management for data centers." *NETWORKING 2012*. Springer Berlin Heidelberg, 2012
3. Kukral, Tomas, et al. "VM migration measurement and failure detection.“ Telecommunications and Signal Processing (TSP), *2015 38th International Conference on IEEE*, 2015.
4. Sachin Shetty, Xuebiao Yuchi and Min Song, "Moving Target Defense for Distributed Systems," Springer, ISBN 978-3-319-31031-2, April 2016

Thank you!