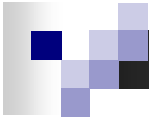# In VINI Veritas
Realistic and Controlled Network Experimentation
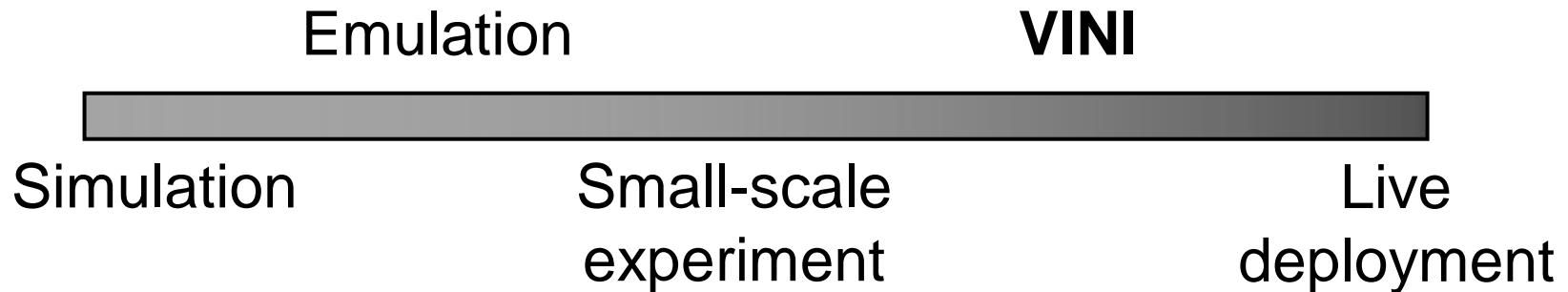
Andy Bavier   Nick Feamster*   Mark Huang

Larry Peterson   Jennifer Rexford

Princeton University   *Georgia Tech
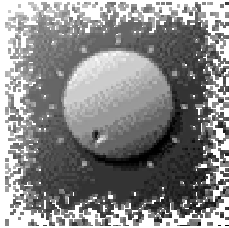
How to Validate an Idea?

Emulation   **VINI**

Simulation   Small-scale   Live
                        experiment   deployment

- Fixed, shared among many experiments
- Runs real routing software
- Exposes realistic network conditions
- Gives control over network events
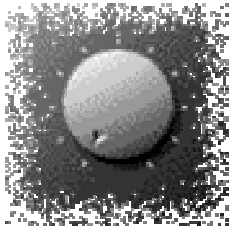- Carries traffic on behalf of real users

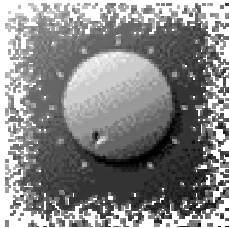# "Controlled Realism"

Arbitrary, emulated

Actual network

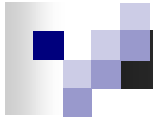**Topology**

Synthetic or traces

Real clients, servers

Traffic

Inject faults, anomalies

Observed in operational network

Network Events

- Start with a controlled experiment

- Relax constraints, study effects

- Result: an operational virtual network that's
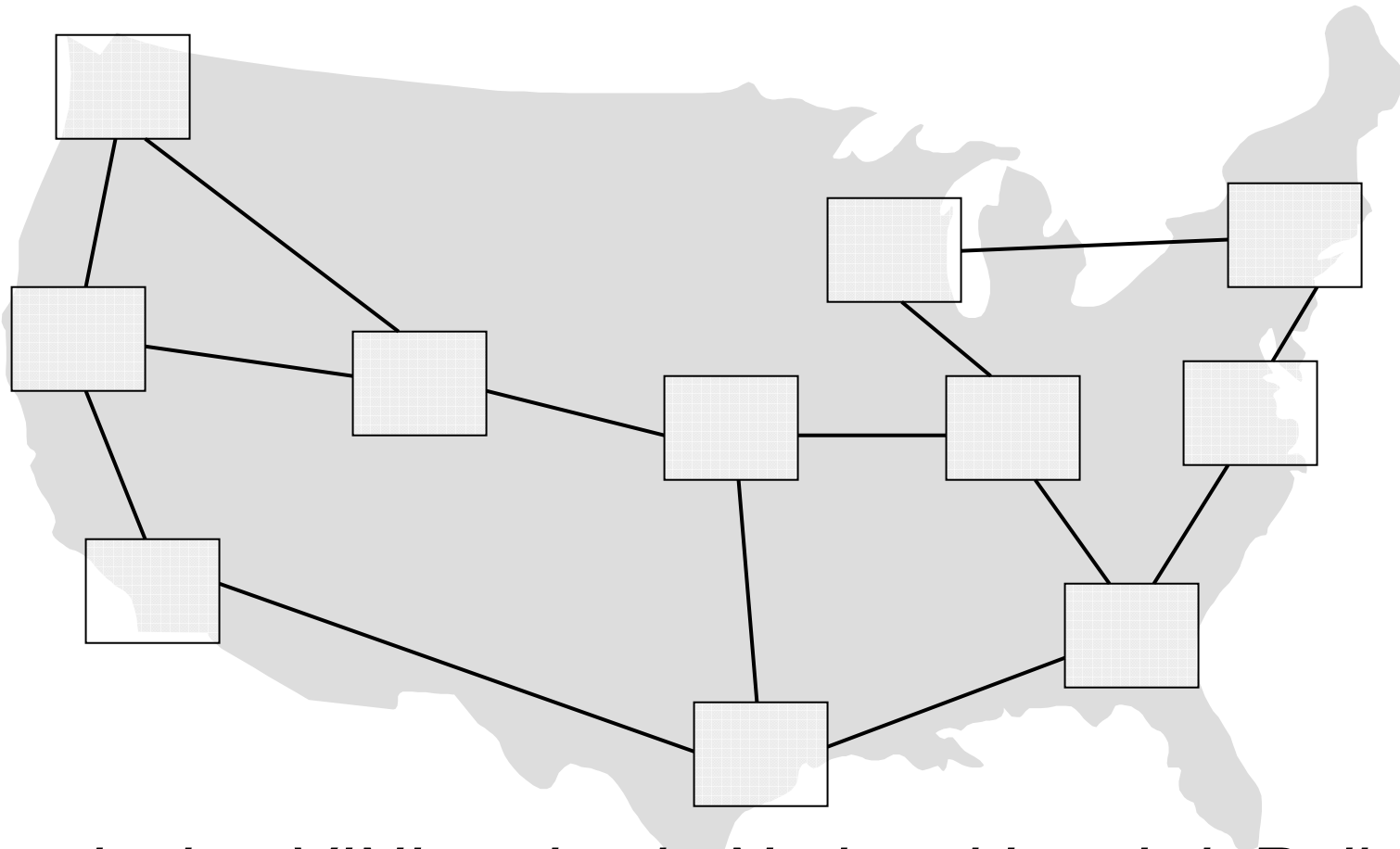 - ☐ Feasible
  - ☐ Valuable
 - ☐ Robust
  - ☐ Scalable, etc.

Overview

- **VINI requirements**
  - ☐ Fixed, shared infrastructure
  - ☐ Flexible network topology
  - ☐ Expose/inject network events
  - ☐ External connectivity and routing adjacencies
- **Strategy for building VINI**
- **PL-VINI: prototype on PlanetLab**
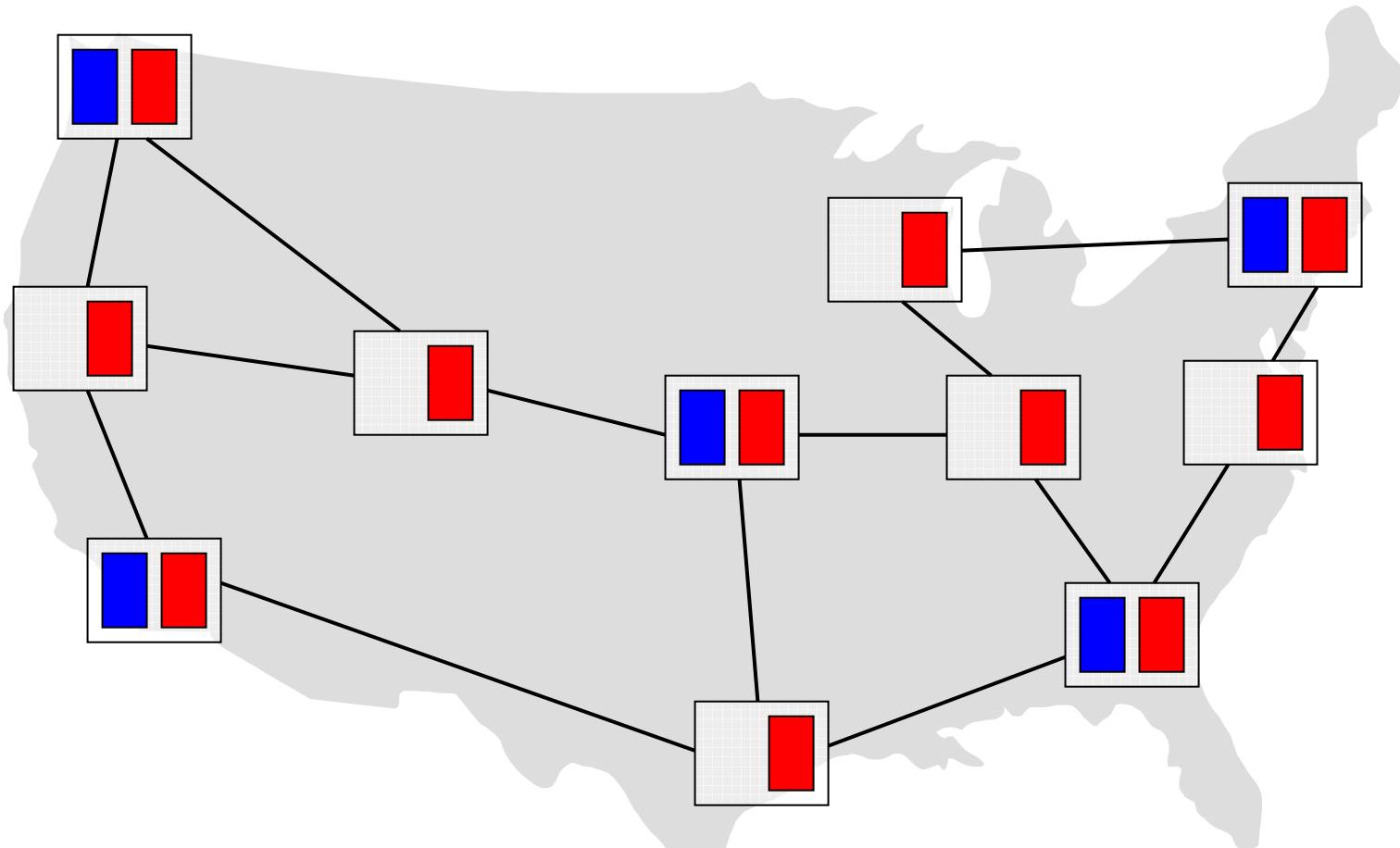- **Experimental results**
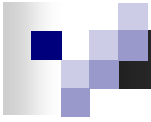- **Timeline**
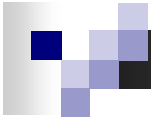
Fixed Infrastructure

Deploying VINI nodes in National LambdaRail,
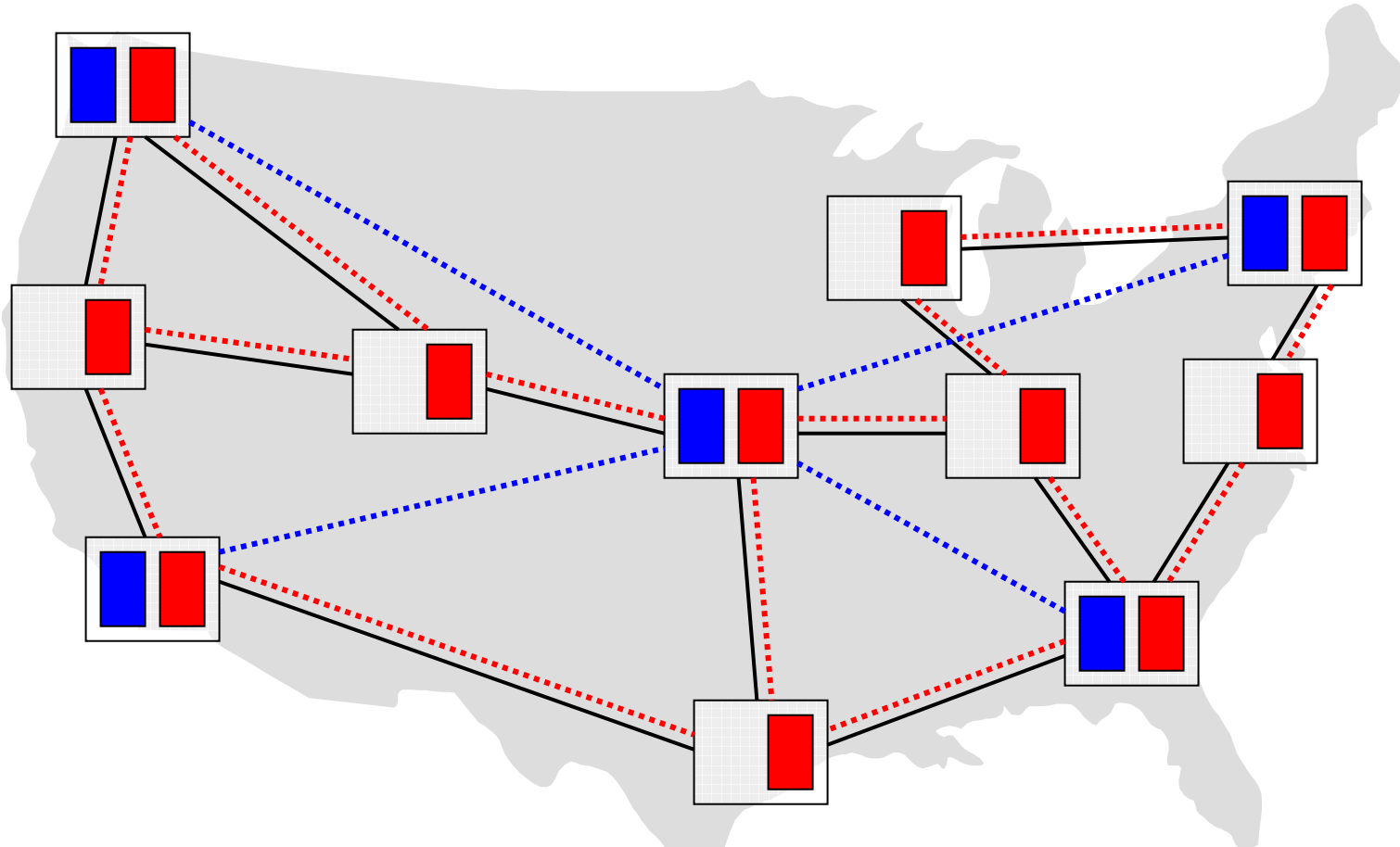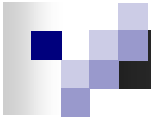*Abilene with Gigabit links*
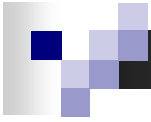
Shared Infrastructure

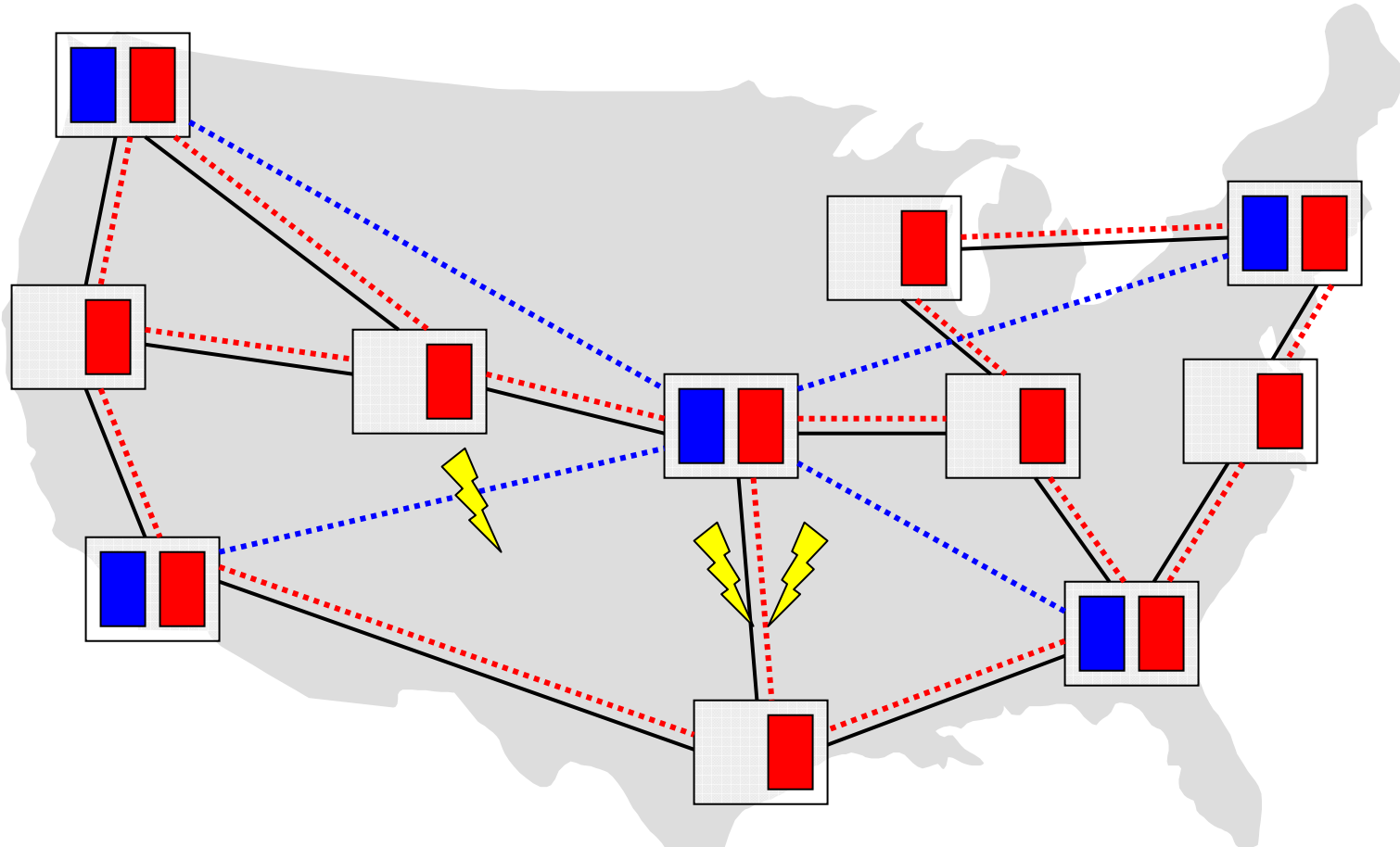*Experiments given illusion of dedicated h/w*

Flexible Topology

VINI supports arbitrary virtual topologies

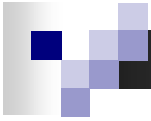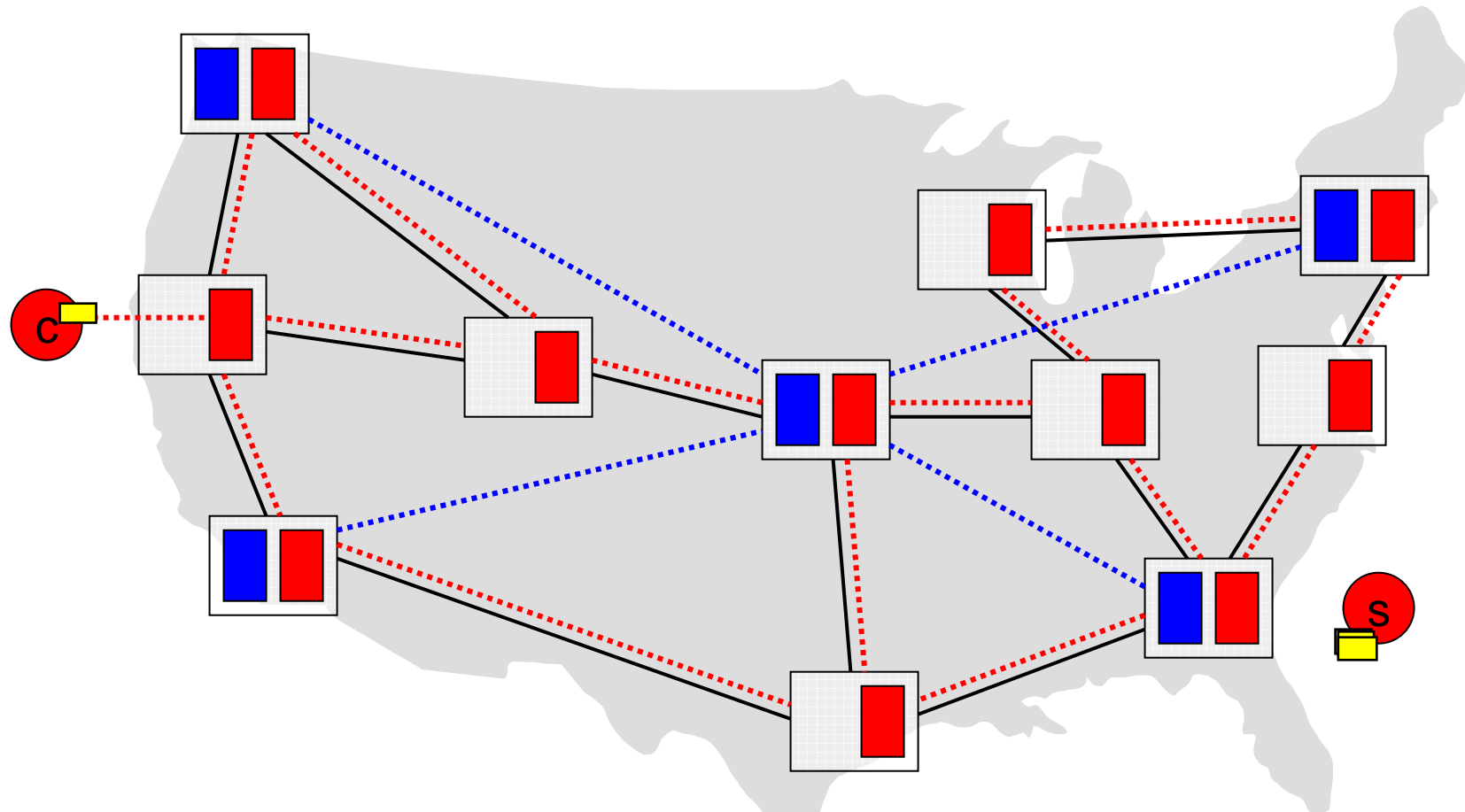# Network Events

VINI exposes, can inject network failures

External Connectivity

Experiments can carry traffic for real end-users

External Routing Adjacencies
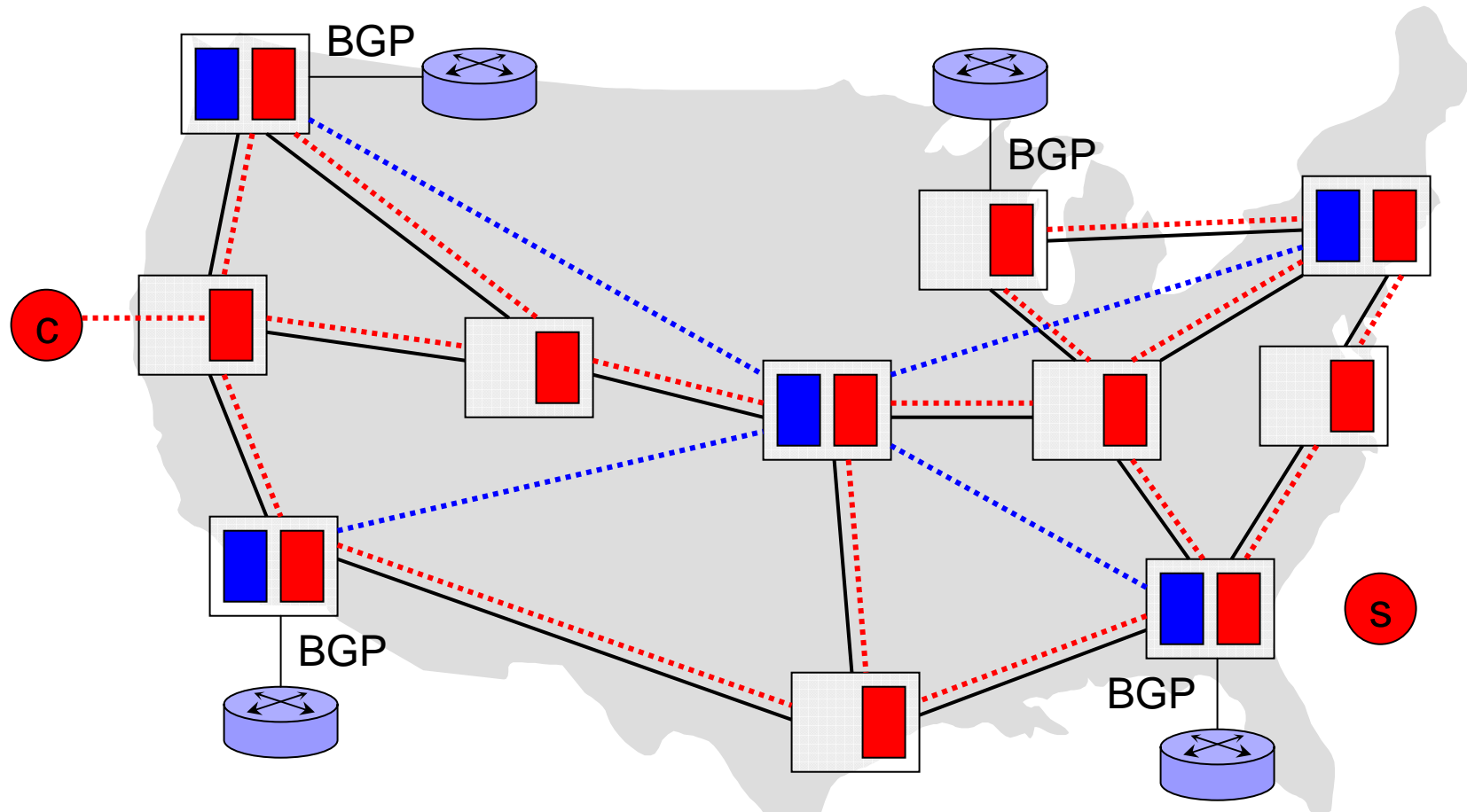
Experiments can participate in Internet routing

PlanetLab ⇨ VINI

- Build VINI from *PlanetLab,* a global testbed for distributed services
  - Begun in 2002
  - 700 nodes at 336 sites in 35 countries
  - 600 projects and 2500 researchers
  - Serves 3-4 TB/day to ~1M clients
- MyPLC: PlanetLab software distribution
  - Anyone can run their own private PlanetLab

PlanetLab Experiments

- Simultaneous experiments in separate VMs
  - Each has "root" in its own VM, can customize
- Reserve CPU, network capacity per experiment

| Node Mgr | Local Admin | VM$_1$ | VM$_2$ | ... | VM$_n$ |

Virtual Machine Monitor (VMM) (Linux++)
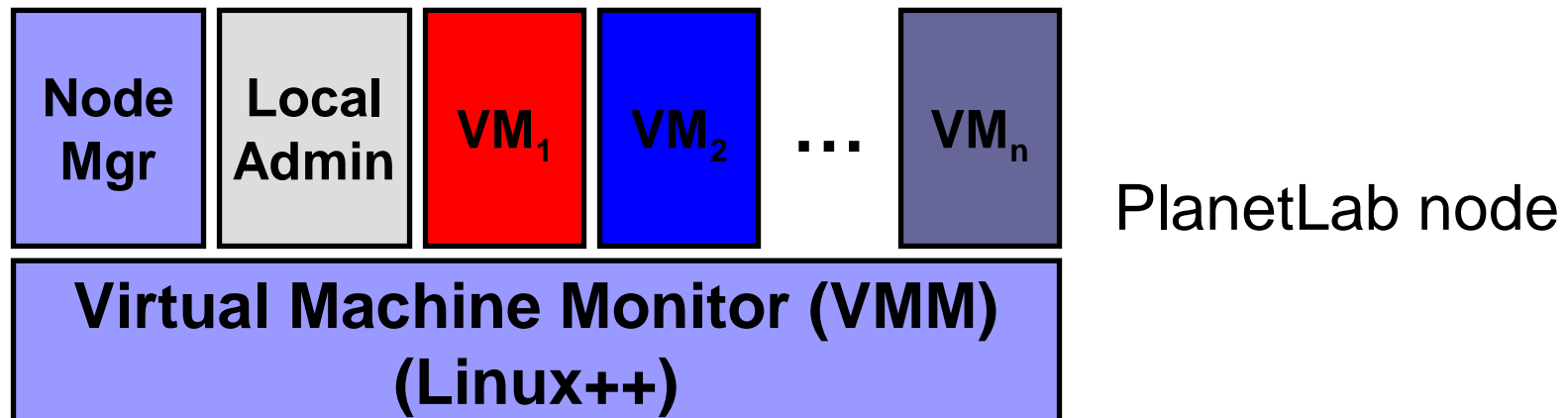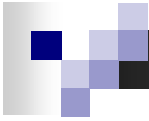
PlanetLab node

PL-VINI: Prototype on PlanetLab

- Feasible? ⇨ prototype on public PlanetLab
- Enable experiment: *Internet In A Slice*
  - ☐ XORP open-source routing protocol suite (NSDI '05)
  - ☐ Click modular router (TOCS '00, SOSP '99)
- Clarify issues that a VINI must address
  - ☐ Unmodified routing software on a virtual topology
  - ☐ Forwarding packets at line speed
  - ☐ Illusion of dedicated hardware
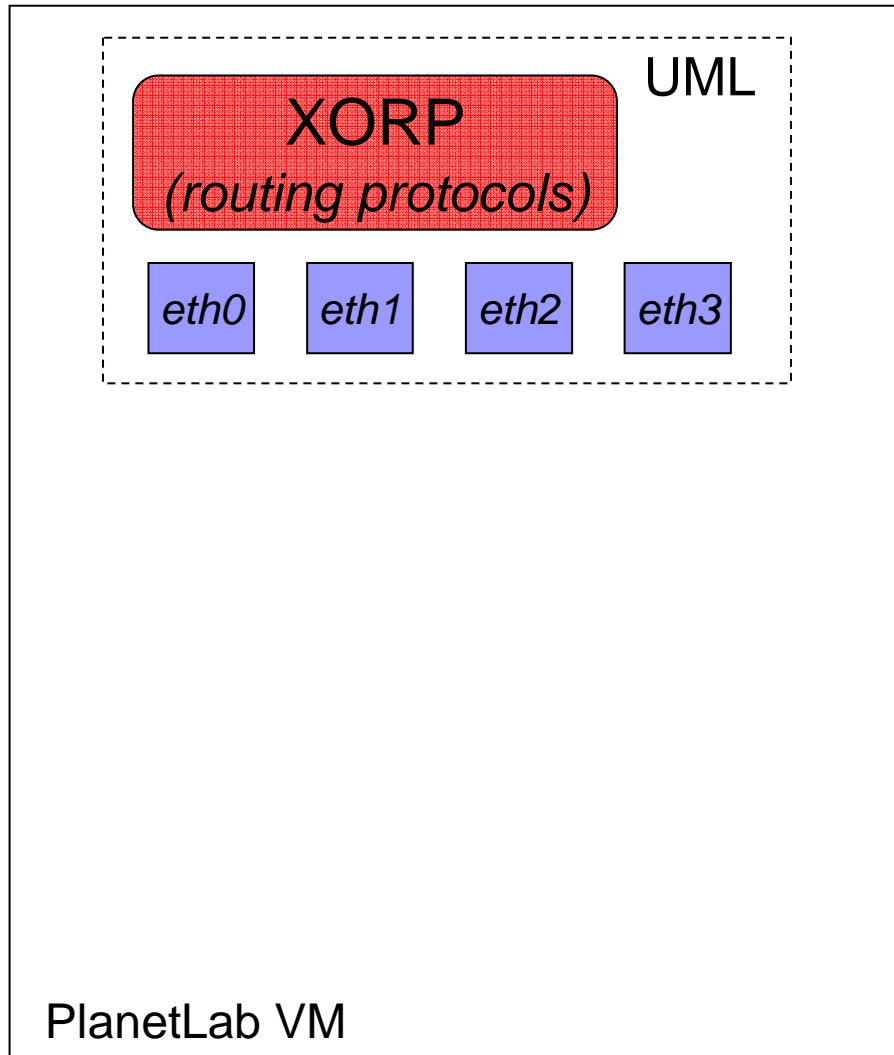 - ☐ Injection of faults and other events

XORP: Control Plane

XORP
(routing protocols)

PlanetLab VM

- Goal: real routing protocols on virtual network topologies

- BGP, OSPF, RIP, PIM-SM, IGMP/MLD

- XORP can run in a PlanetLab VM

User-Mode Linux: Environment

- Interface ≈ network
- PlanetLab limitation:
  - □ Experiments cannot create new interfaces
- Run routing software in UML environment
- Create virtual network interfaces in UML

Click: Data Plane
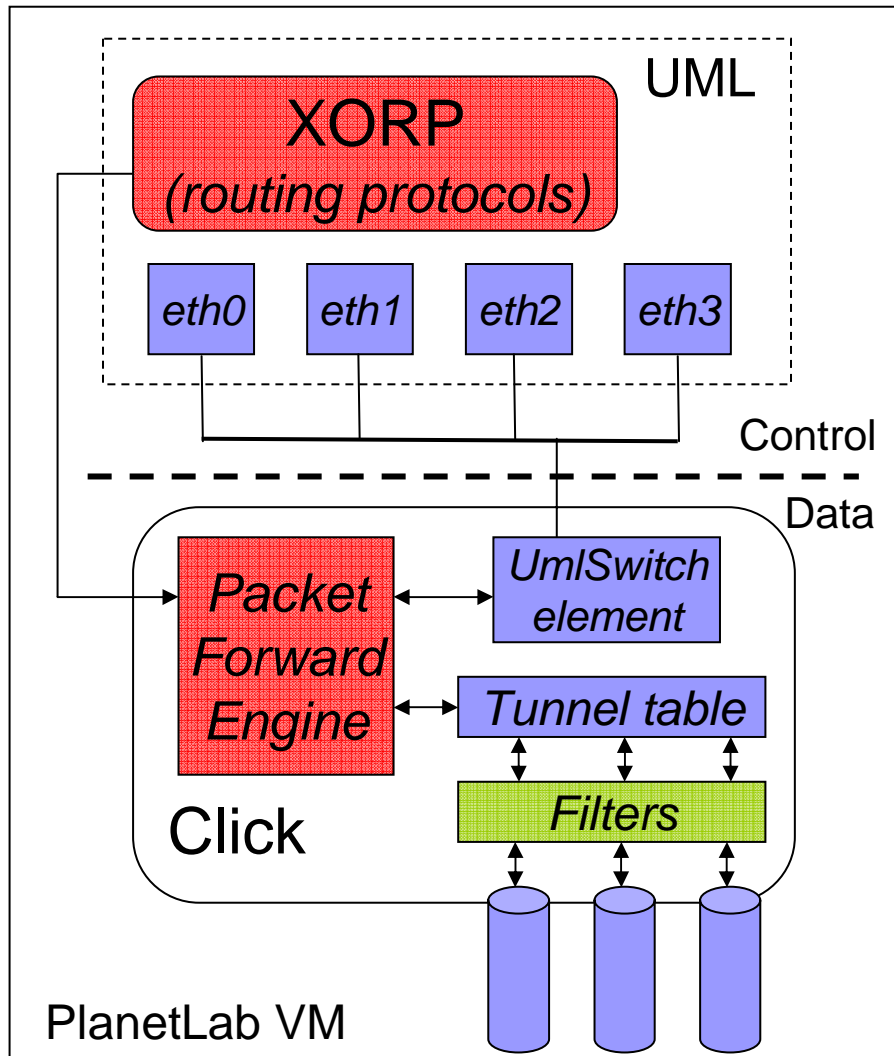
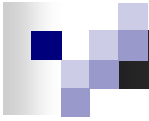- Performance
  - Avoid UML overhead
  - Move to kernel, FPGA
- Interfaces ⇨ tunnels
  - Click UDP tunnels correspond to UML network interfaces
- Filters
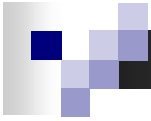 - "Fail a link" by blocking packets at tunnel

Resource Isolation

- Issue: Forwarding packets in user space
  - PlanetLab sees heavy use
  - CPU load affects virtual network performance
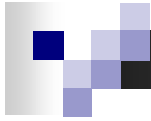
| Property	Depends On	Solution
Throughput	CPU% received	PlanetLab provides CPU reservations
Latency	CPU scheduling delay	PL-VINI: boost priority of packet forward process

Intra-domain Route Changes
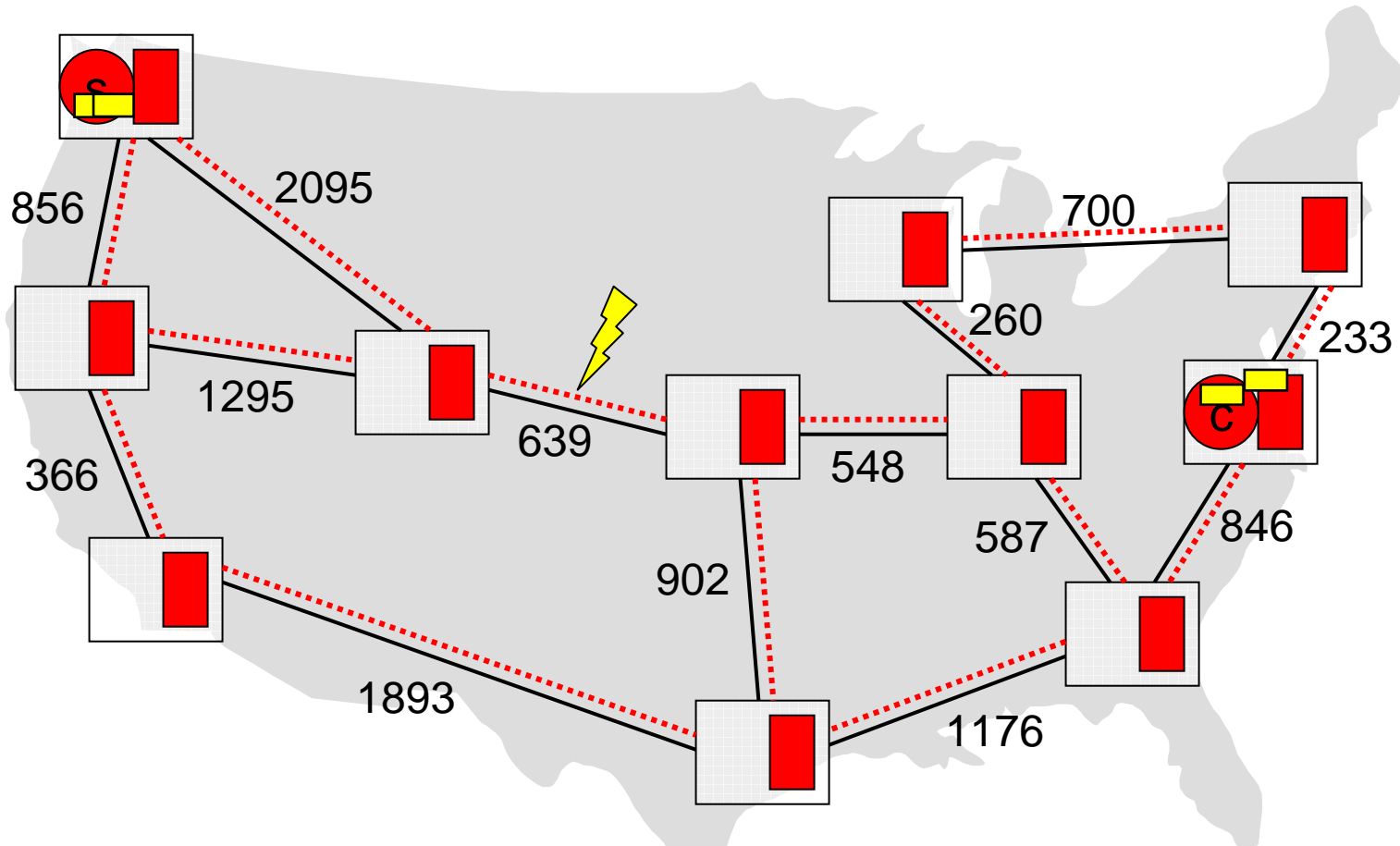

Watch OSPF route convergence on Abilene

Experiment Summary

- Observe performance during failure
  - Ping between client and server
  - Tcpdump to observe effects on packets
- Experiment results
  - Change in RTT after failure
  - Transient loss & reordering during convergence
 - Fine-grain effects on TCP congestion control
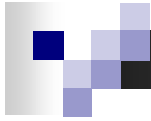
Working on New Experiments

- **Non-IP protocol**
  - Ethernet bridging
  - New ways of scaling Ethernet
- **Border Gateway Protocol**
  - Convergence for internal BGP
  - Evaluation of the Routing Control Platform
- **Integration with wireless**
  - Tunnels to the Orbit testbed at Rutgers
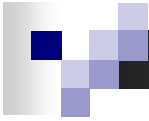 - Wireless mobility over an OSPF backbone
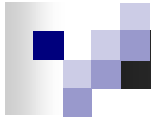
Ongoing Work on VINI

- Sharing with the community
  - User and developer guides
  - Tarball with the VINI code
- Lowering the barrier to running experiments
  - Scripts for configuring "Internet in a slice" experiments
  - Distributed monitoring software (e.g., tcpdump)
- Admission control and embedding
  - Booking of resources for an experiment
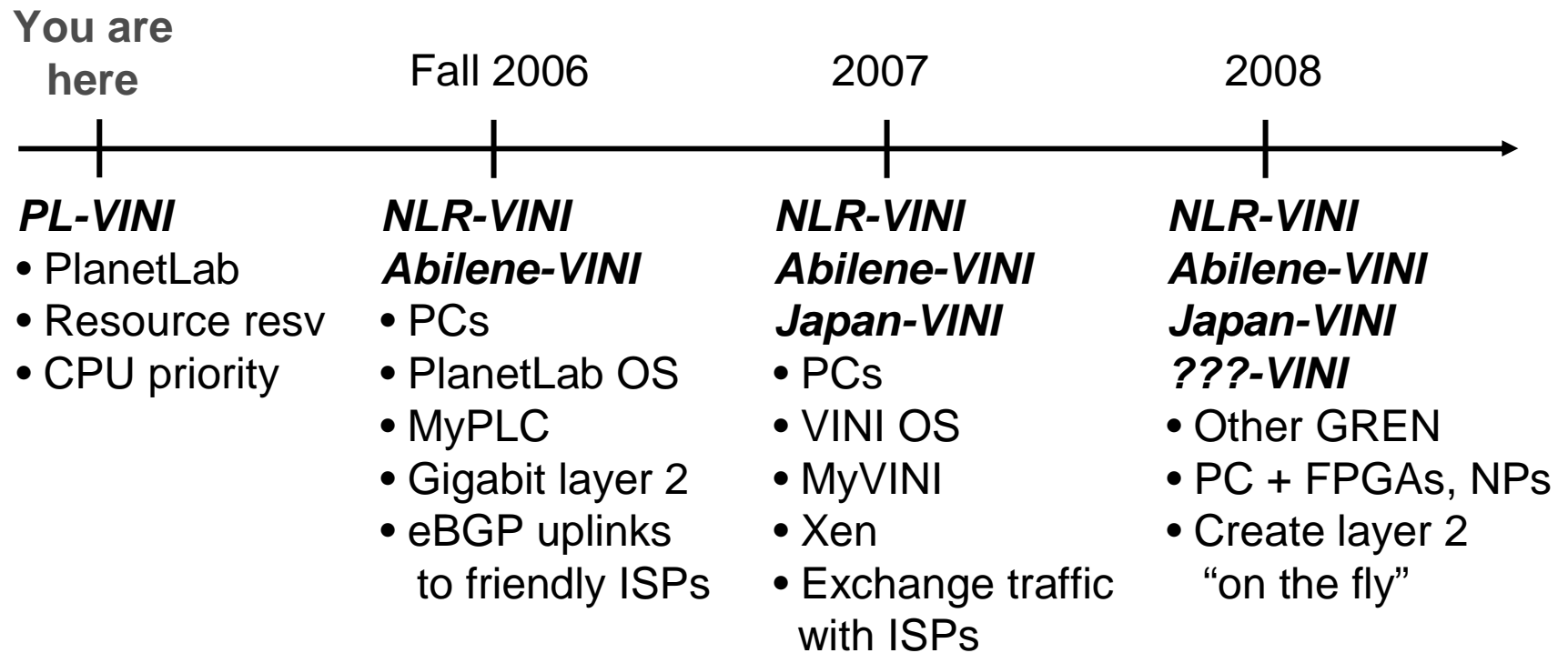 - Simple embedding service to allocate slices
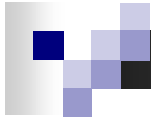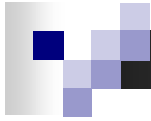
Ongoing Work on VINI

- Exposing real failures
  - Detecting underlay failures in Abilene and NLR
  - Exposing the experiments to the failures
- Cut-through functionality
  - Packets "cutting through" a VINI node
  - E.g., for experiments that want to embed a virtual topology
- BGP multiplexer
  - Prototype of the BGP multiplexer
  - To share a single BGP session with neighbors
- Hardware support for packet forwarding
  - NetFPGA work of Nick McKeown
 - Programmable router of Jon Turner

Timeline

You are here   Fall 2006   2007   2008

PL-VINI

- PlanetLab
- Resource resv
- CPU priority

NLR-VINI
Abilene-VINI

- PCs
- PlanetLab OS
- MyPLC
- Gigabit layer 2
- eBGP uplinks
  to friendly ISPs

NLR-VINI
Abilene-VINI
Japan-VINI

- PCs
- VINI OS
- MyVINI
- Xen
- Exchange traffic
  with ISPs

NLR-VINI
Abilene-VINI
Japan-VINI
???-VINI

- Other GREN
- PC + FPGAs, NPs
- Create layer 2
  "on the fly"

Other features?

Conclusion
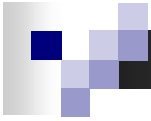
- VINI = evolution of PlanetLab
- Installing VINI nodes in NLR, Abilene
- Download and run *Internet In A Slice*
- MyPLC ⇨ MyVINI as code diverges
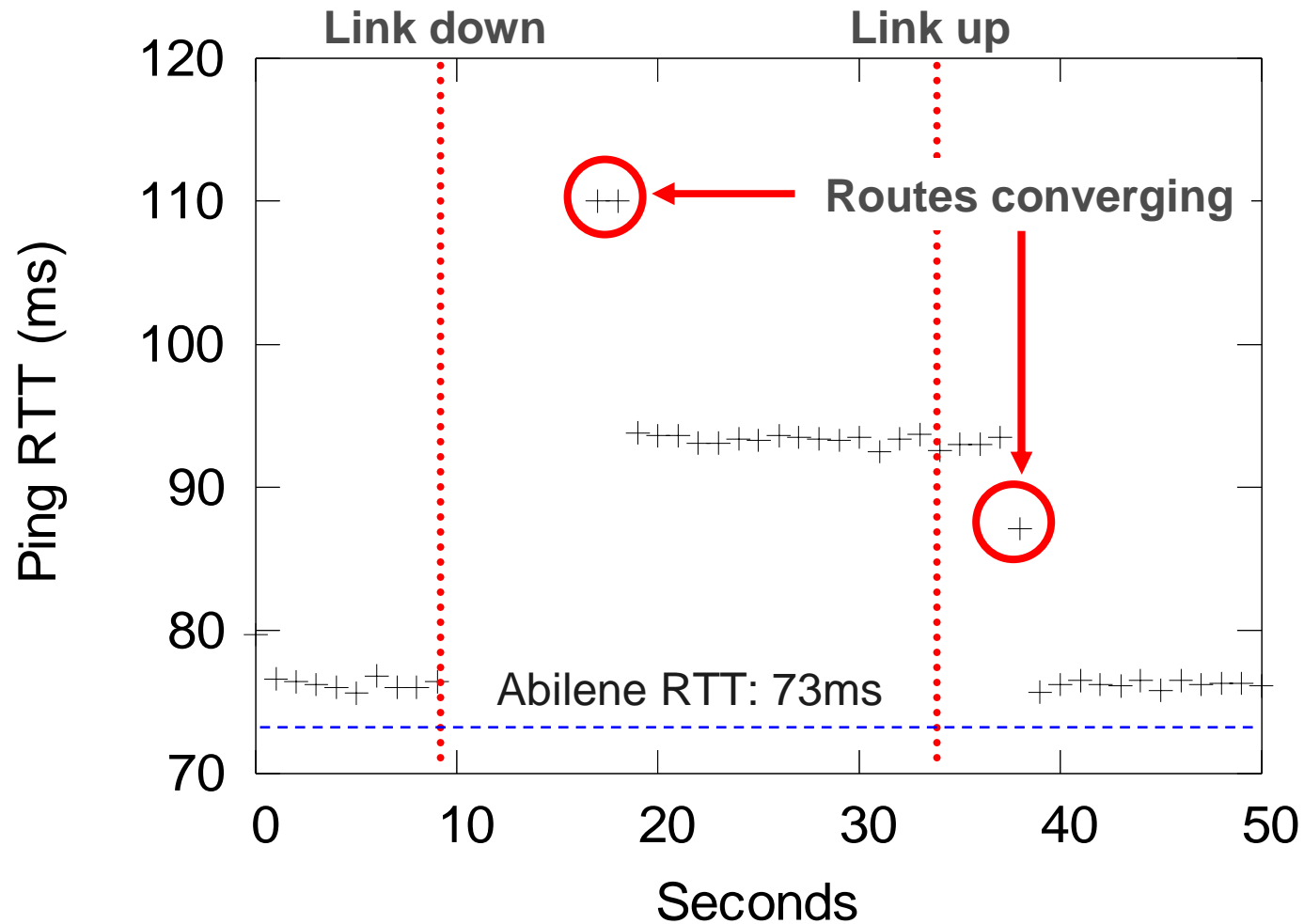 - ☐ Build, run, modify your own VINI
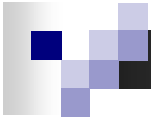 - ☐ We expect there to be many VINIs

http://www.vini-veritas.net
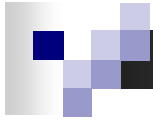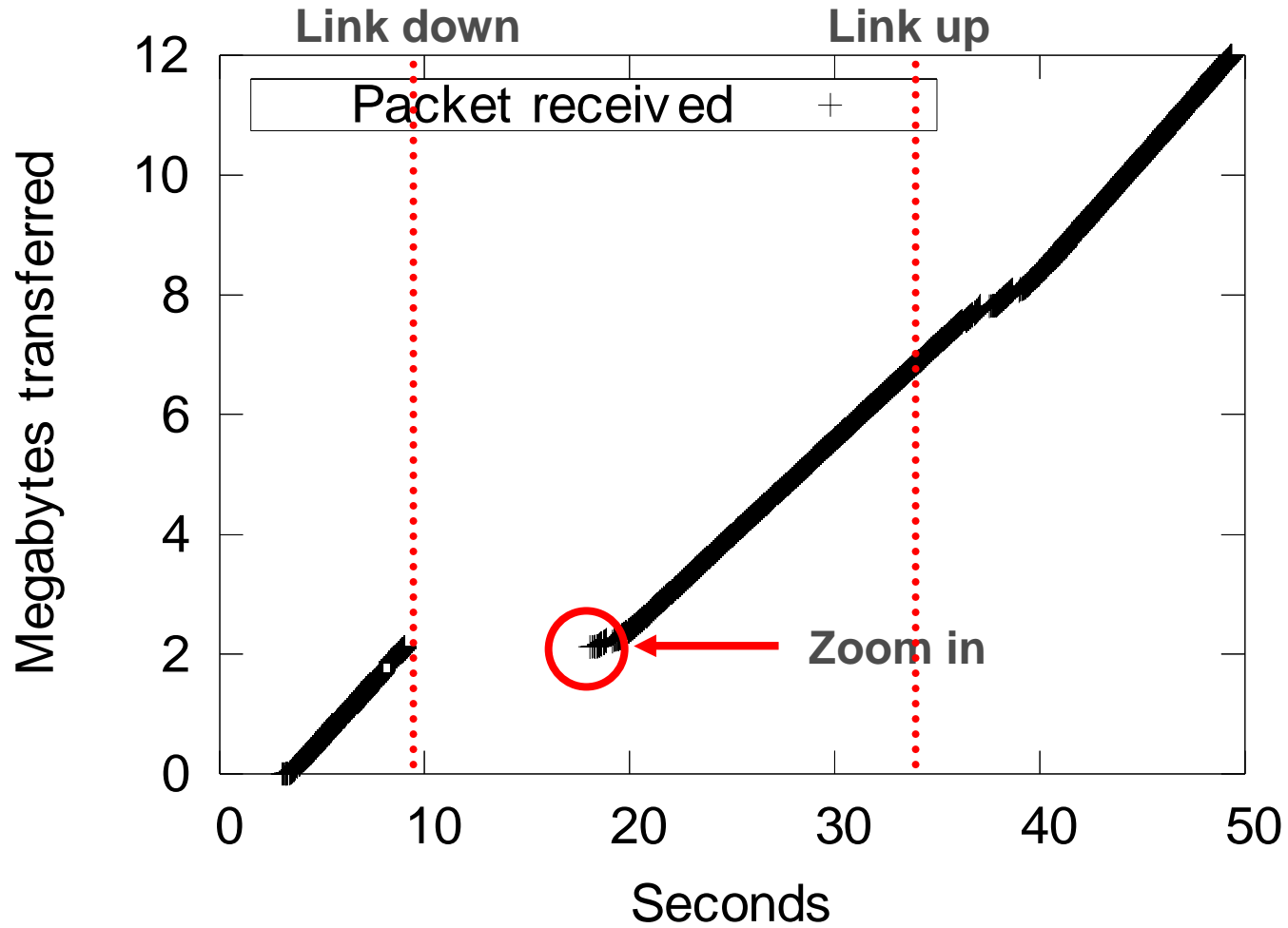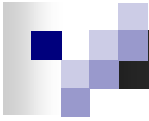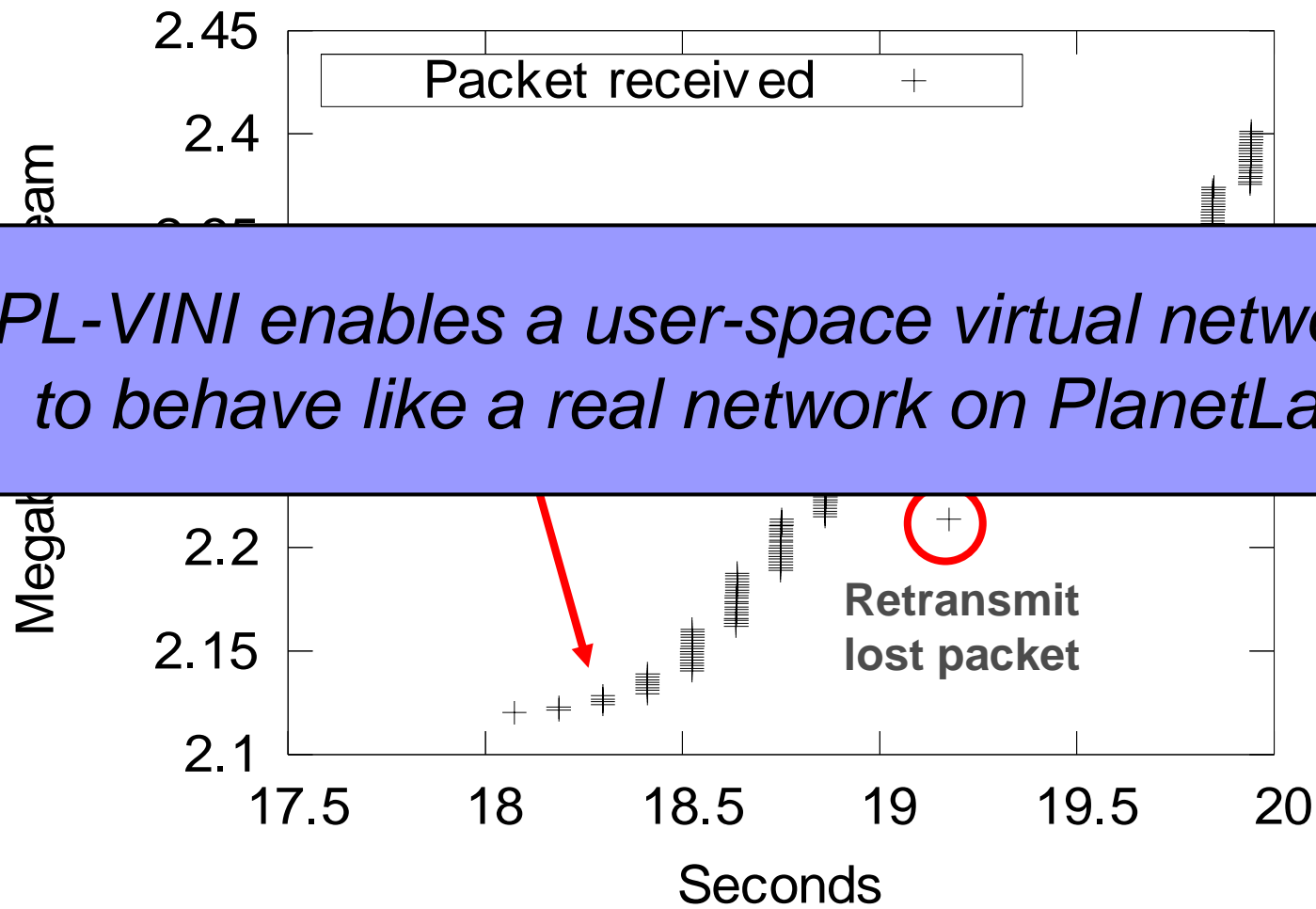
- Backup slides…

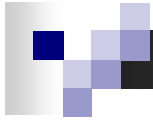# Ping During Link Failure

TCP Throughput

Arriving TCP Packets

PL-VINI enables a user-space virtual network to behave like a real network on PlanetLab

Attracting Real Users

- Could have run experiments on Emulab
- Goal: Operate our own virtual network
  - Carrying traffic for actual users
  - We can tinker with routing protocols
- We expect that:
  - PlanetLab services will subscribe to VINI network architectures to access Gb/s
  - Experiments will advertise routes via BGP