

Towards Virtual Networks for Virtual Machine Grid Computing

Ananth I. Sundararaj

Peter A. Dinda




Prescience Lab

Department of Computer Science

Northwestern University

<http://virtuoso.cs.northwestern.edu>

Outline

- Virtual machine grid computing
- Virtuoso system
- Networking challenges in Virtuoso
- Enter VNET
- VNET  Adaptive virtual network
- Related Work
- Conclusions
- Current Status

Aim Deliver arbitrary amounts of computational power to perform distributed and parallel computations

Traditional Paradigm

New Paradigm

Grid Computing
Resource multiplexing using OS level mechanism

Grid Computing using virtual machines

Problem1:
Complexity from resource user's perspective

Solution


Virtual Machines
What are they?

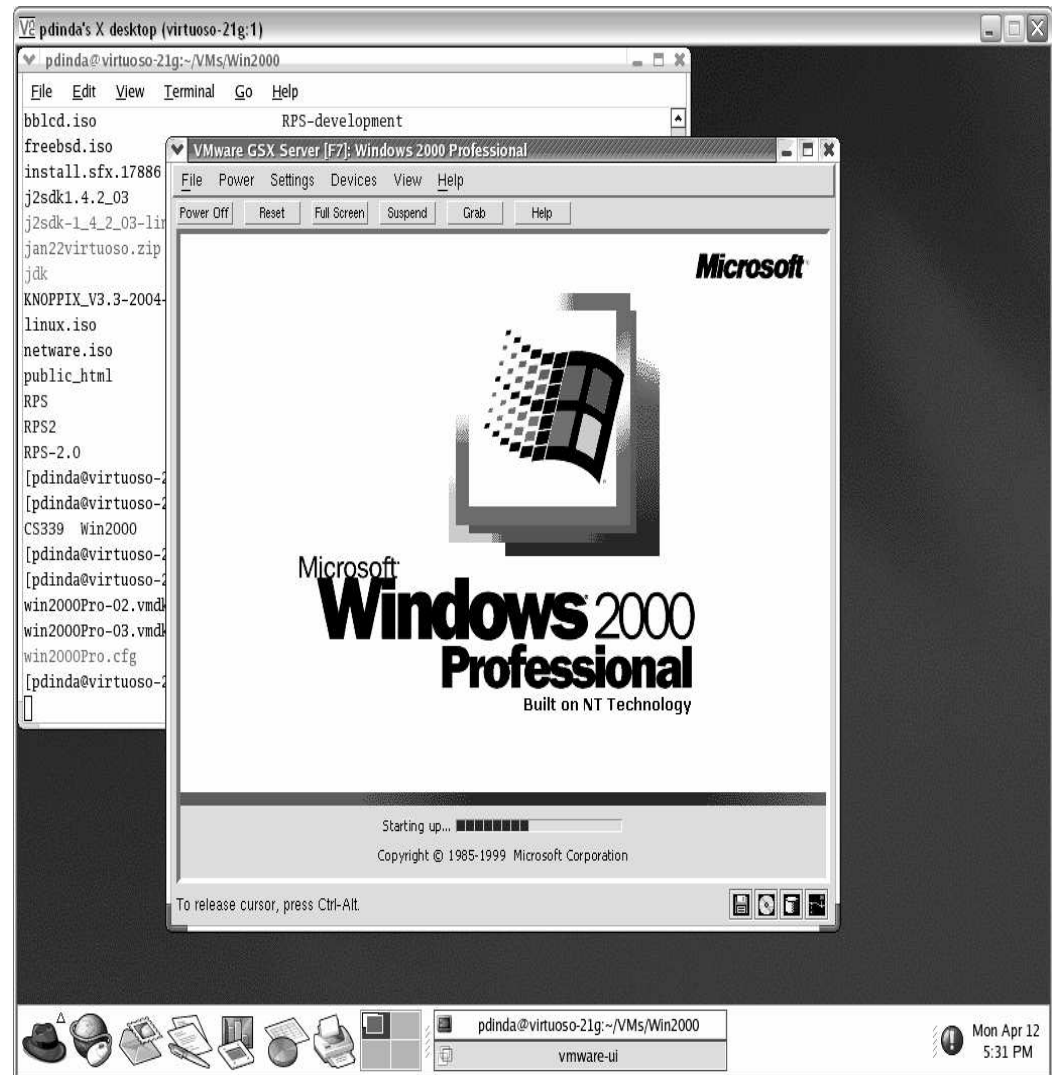
Problem2:
Complexity from resource owner's perspective

How to leverage them?

Virtual Machines

Virtual machine monitors (VMMs)

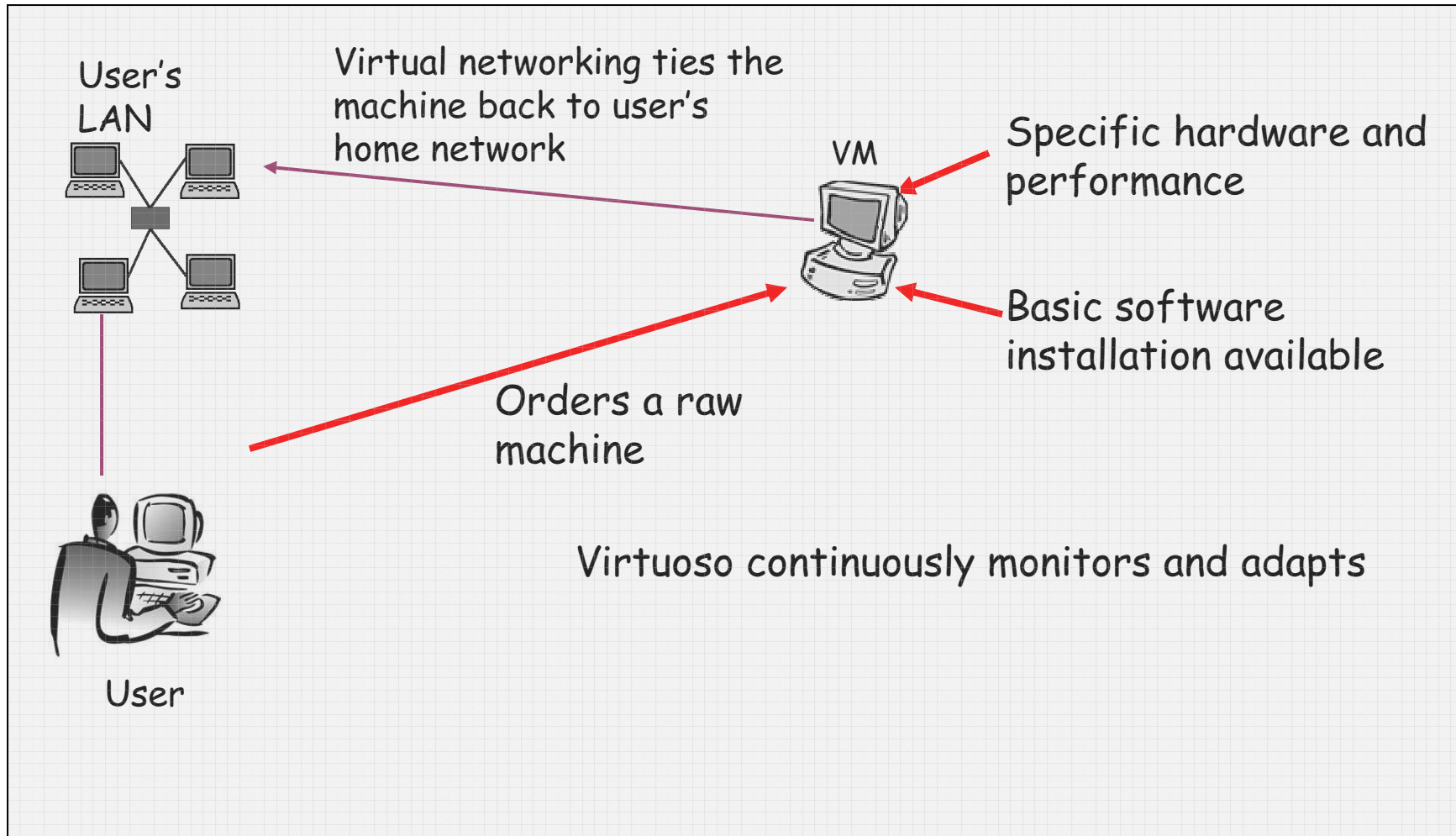
- Raw machine is the abstraction
- VM represented by a single image
- VMware GSX Server 



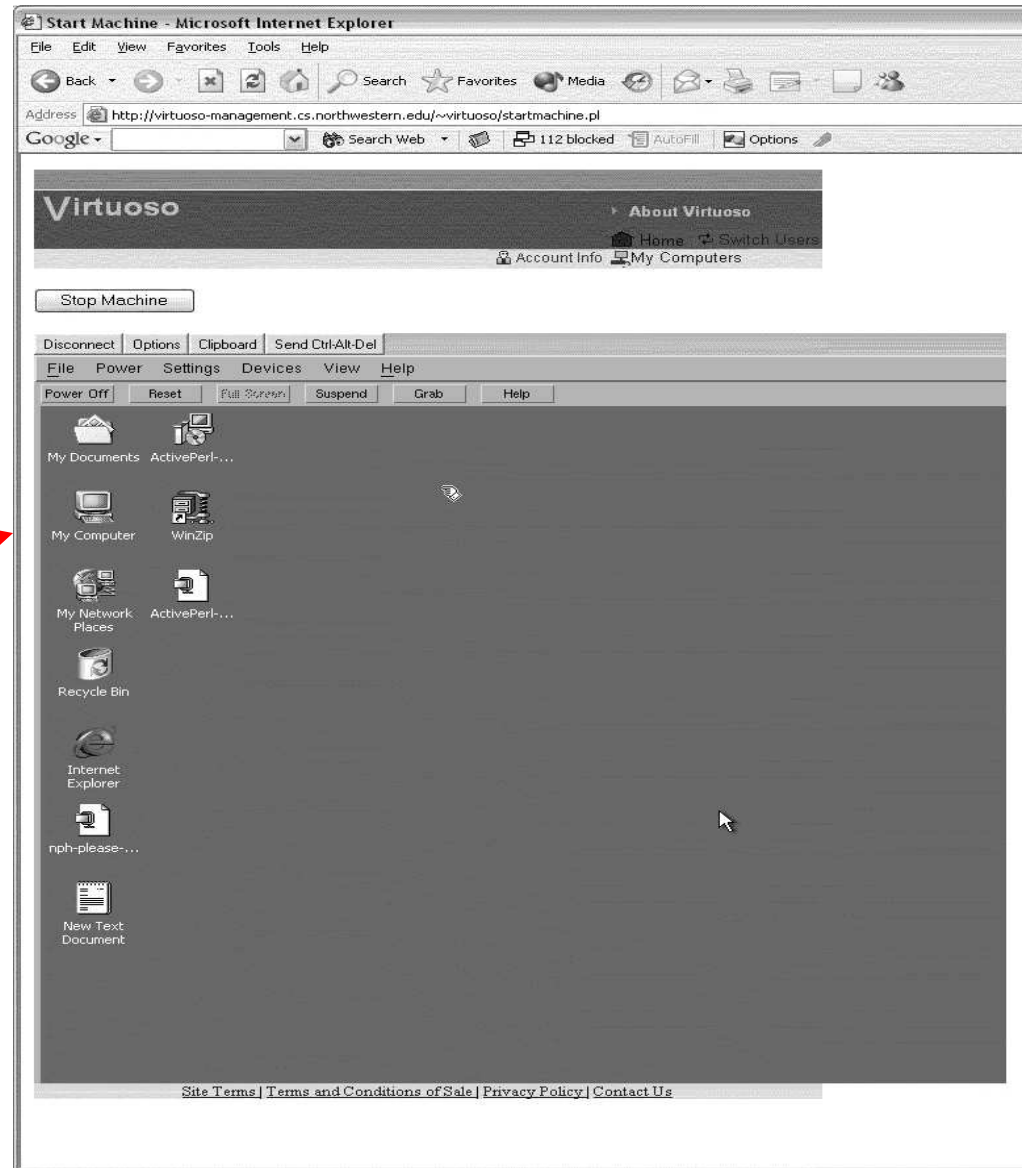
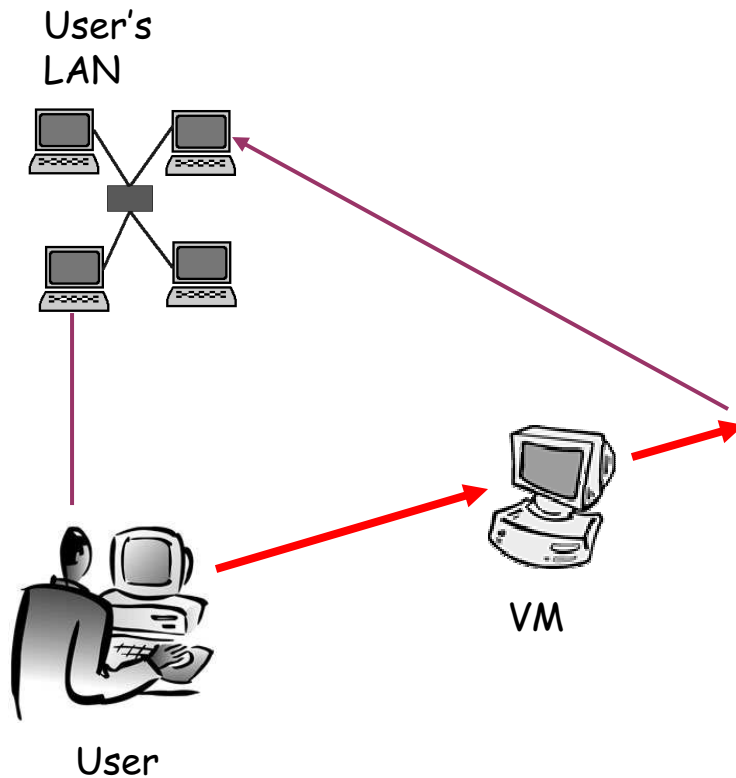
Virtual machine grid computing

- Approach: Lower level of abstraction
 - Raw machines, not processes, jobs, RPC calls
 - R. Figueiredo, P. Dinda, J. Fortes, *A Case For Grid Computing on Virtual Machines*, ICDCS 2003
- Mechanism: Virtual machine monitors
- Our Focus: Middleware support to hide complexity
 - Ordering, instantiation, migration of machines
 - **Virtual networking**
 - remote devices
 - Connectivity to remote files, machines
 - Information services
 - Monitoring and prediction
 - Resource control

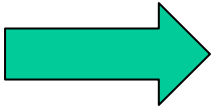
The Simplified Virtuoso Model



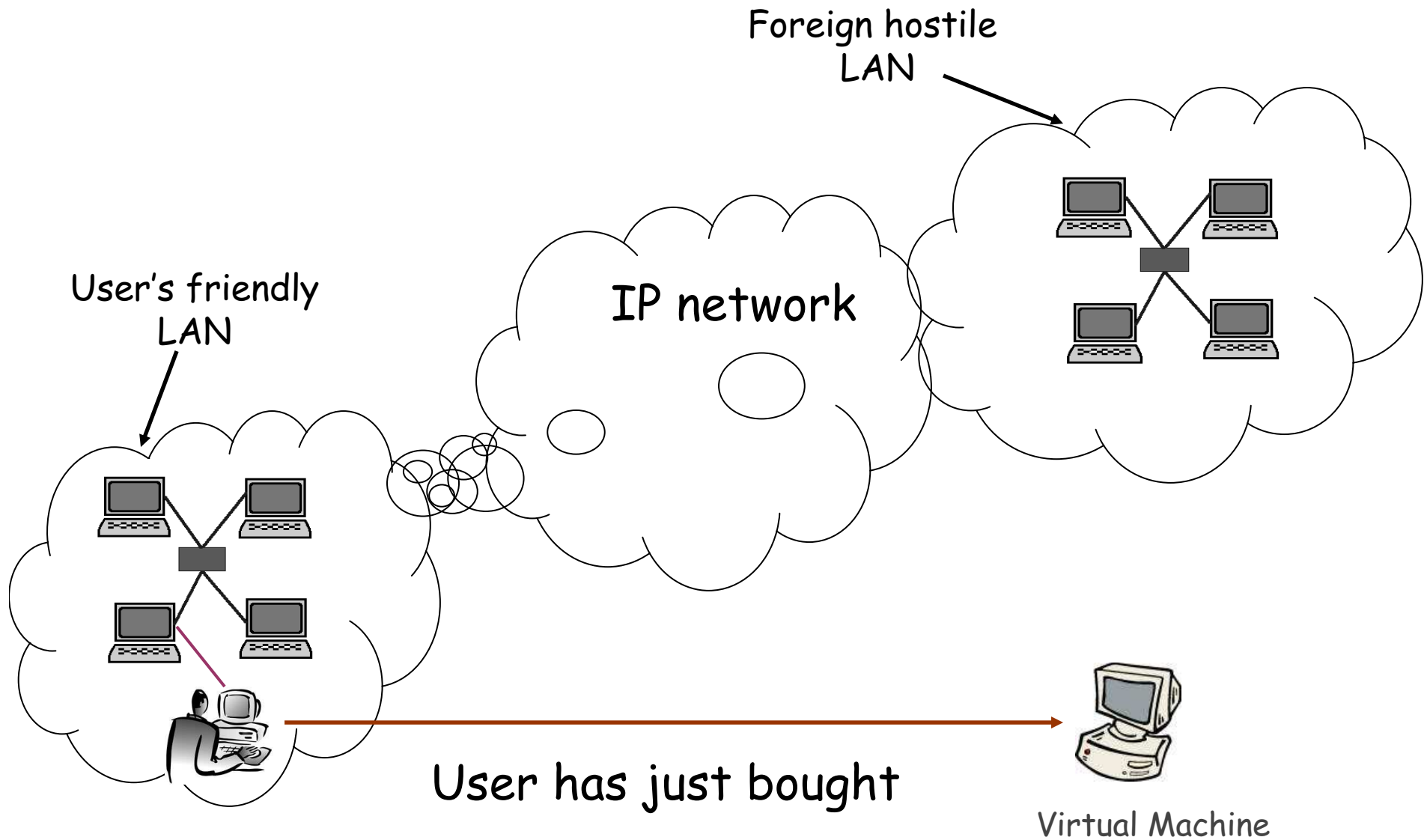
User's View in Virtuoso Model



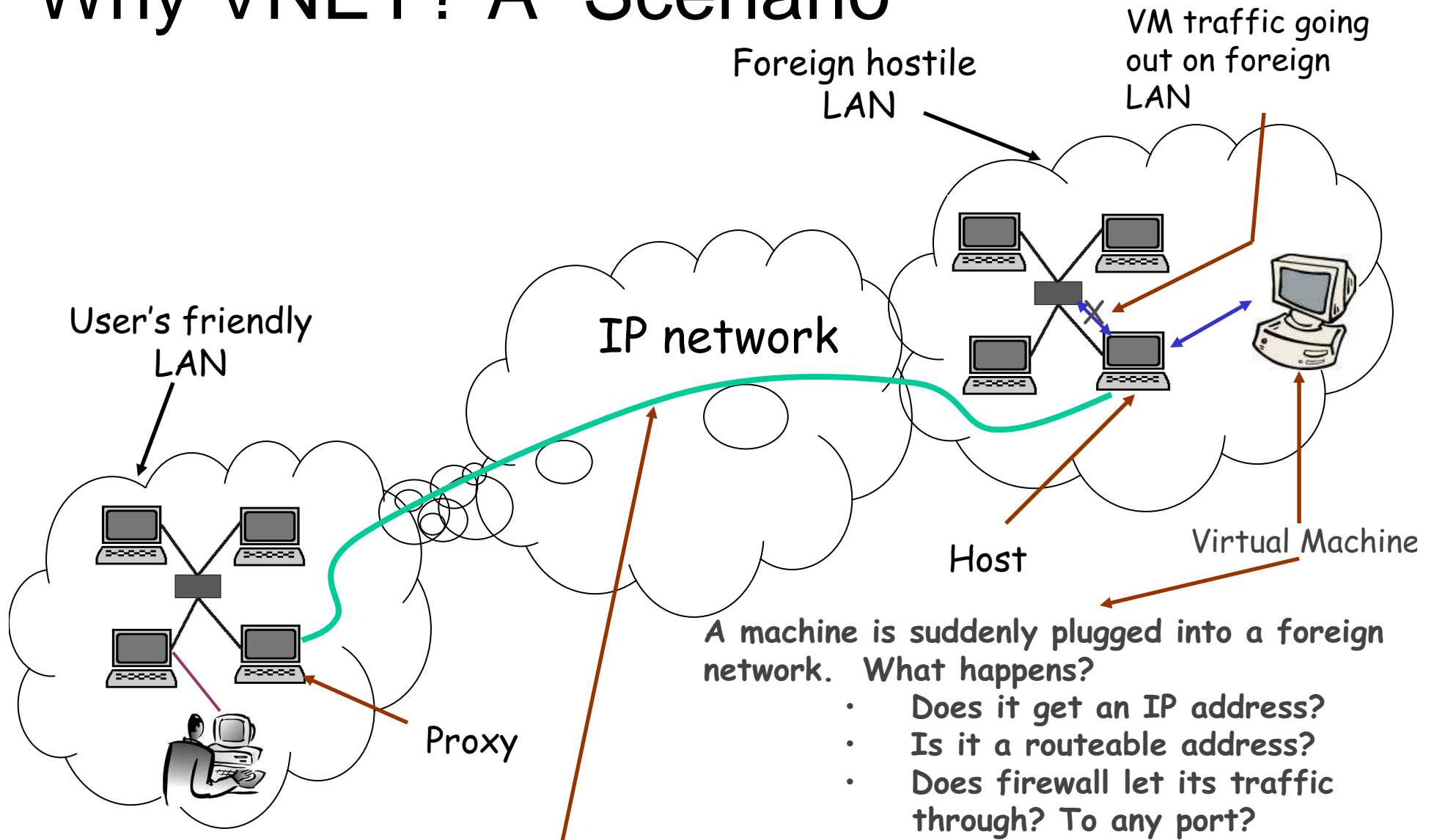
Outline

- Virtual machine grid computing
- Virtuoso system
- Networking challenges in Virtuoso
- Enter VNET
- VNET  Adaptive virtual network
- Related Work
- Conclusions
- Current Status

Why VNET? A Scenario



Why VNET? A Scenario




A machine is suddenly plugged into a foreign network. What happens?

- Does it get an IP address?
- Is it a routeable address?
- Does firewall let its traffic through? To any port?

VNET: A bridge with long wires

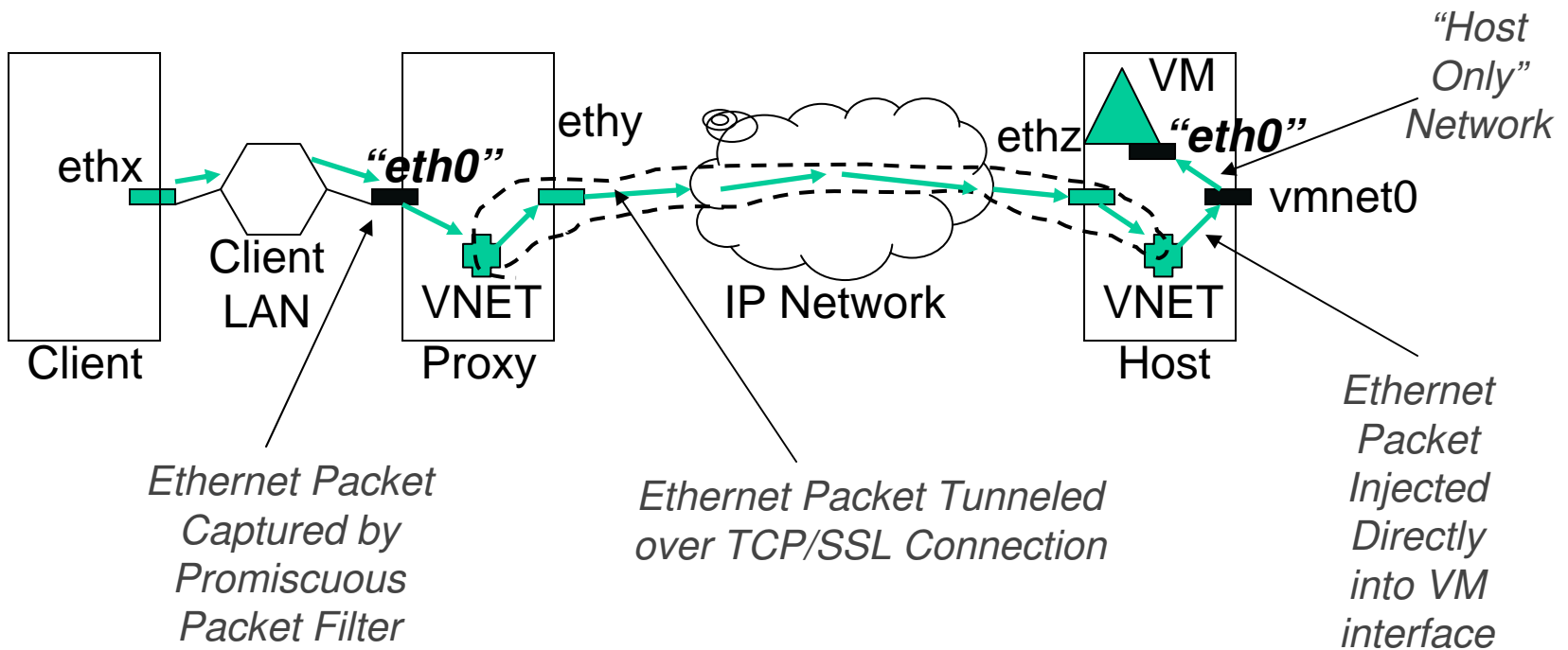
Outline

- Virtual machine grid computing
- Virtuoso system
- Networking challenges in Virtuoso
- Enter VNET
- VNET  Adaptive virtual network
- Related Work
- Conclusions
- Current Status

A Layer 2 Virtual Network for the User's Virtual Machines

- Why Layer 2?
 - Protocol agnostic
 - Mobility
 - Simple to understand
 - Ubiquity of Ethernet on end-systems
- What about scaling?
 - Number of VMs limited (~1024 per user)
 - One VNET per user
 - Hierarchical routing possible because MAC addresses can be assigned hierarchically

VNET operation



Traffic outbound from the user's LAN

Performance Evaluation

Main goal
Convey the network management problem induced by VMs to the home network of the user

However
VNET's performance should be

- In line with physical network
- Comparable to other options
- Sufficient for scenarios

Metrics

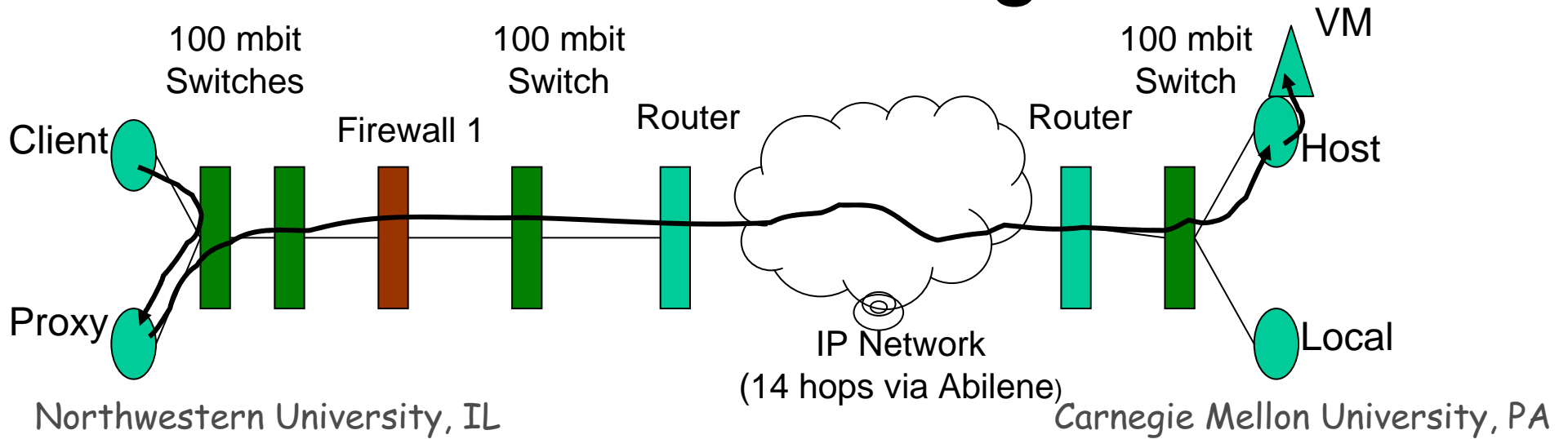
Latency

- | Why? | How? |
|--|--|
| <ul style="list-style-type: none">• small transfer• Interactivity | <ul style="list-style-type: none">• ping• hour long intervals |

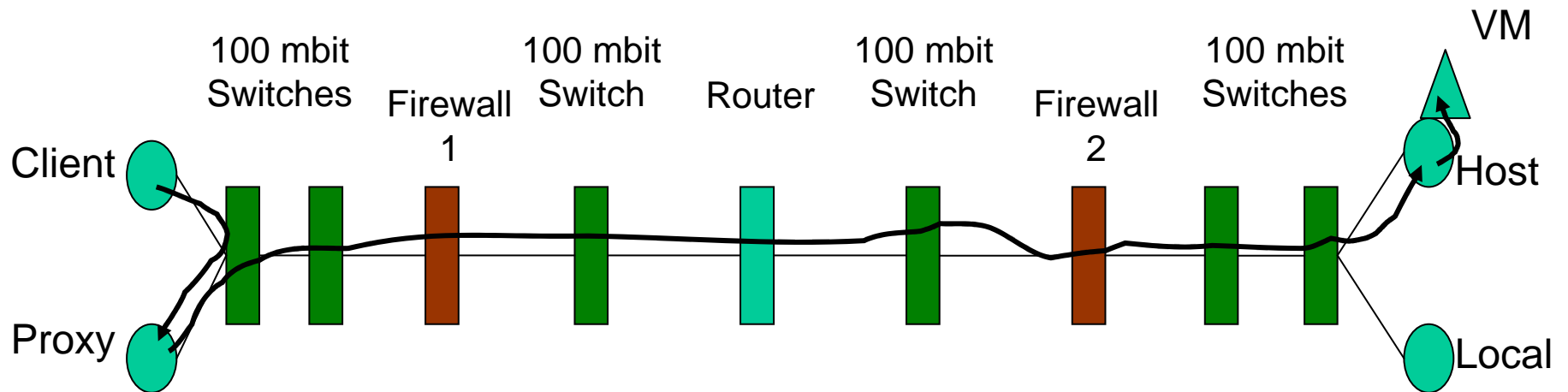
Bandwidth

- | Why? | How? |
|---|---|
| <ul style="list-style-type: none">• Large transfer• low throughput | <ul style="list-style-type: none">• <i>ttcp</i>• socket buffer• 1 GB of data¹⁴ |

VNET test configuration

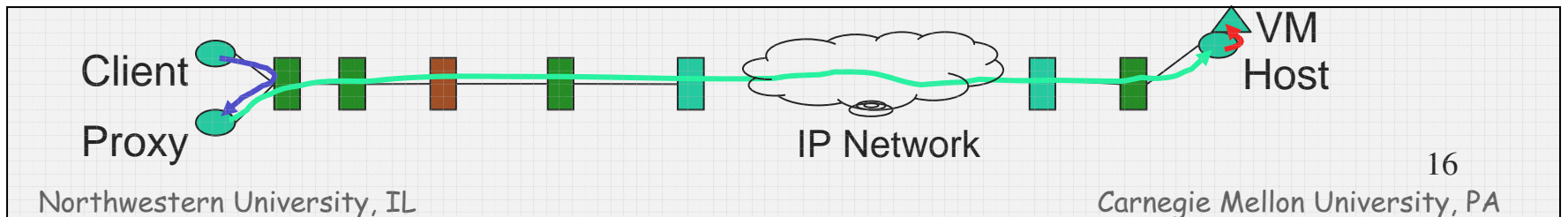
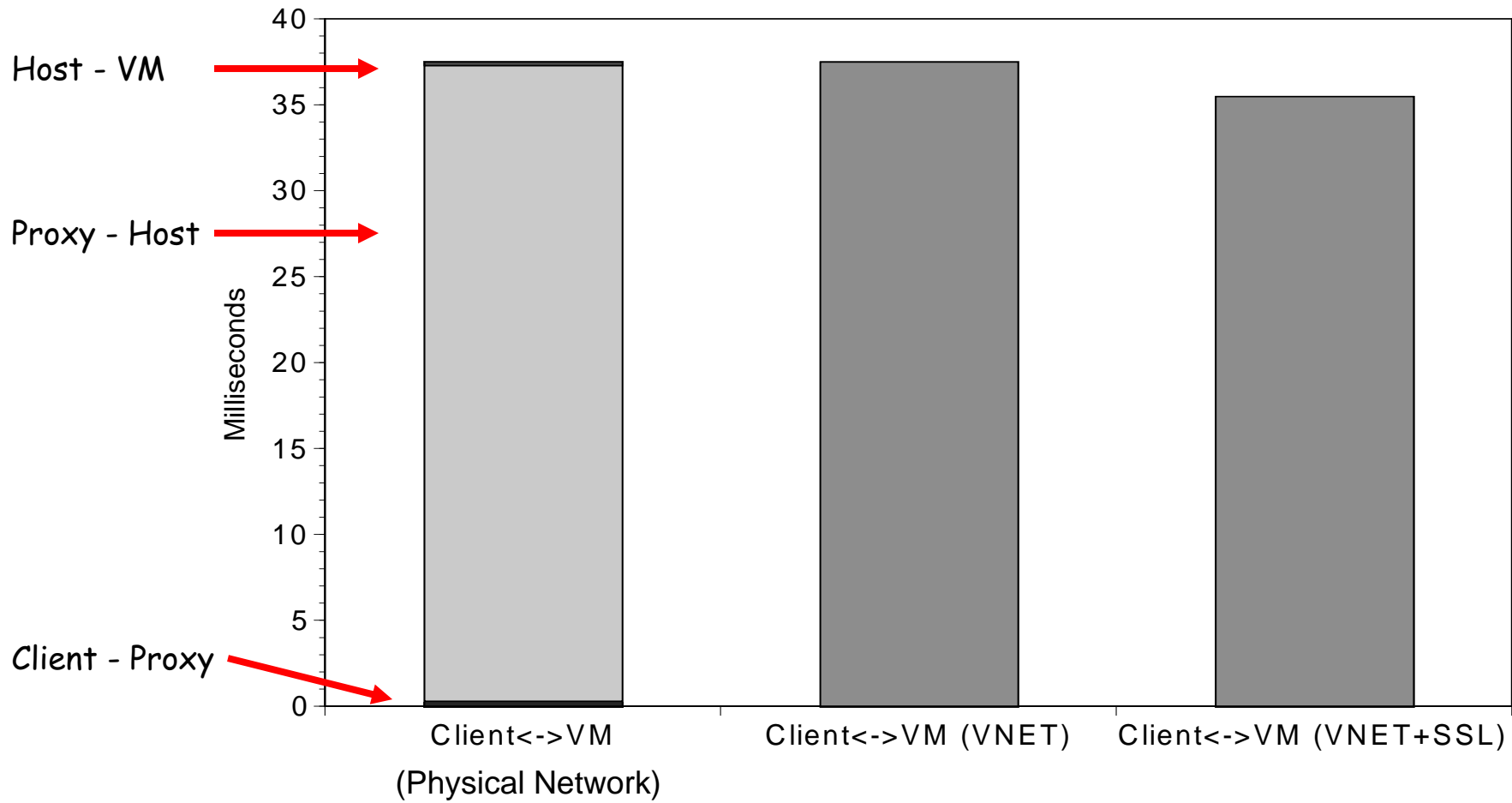


Wide area configuration

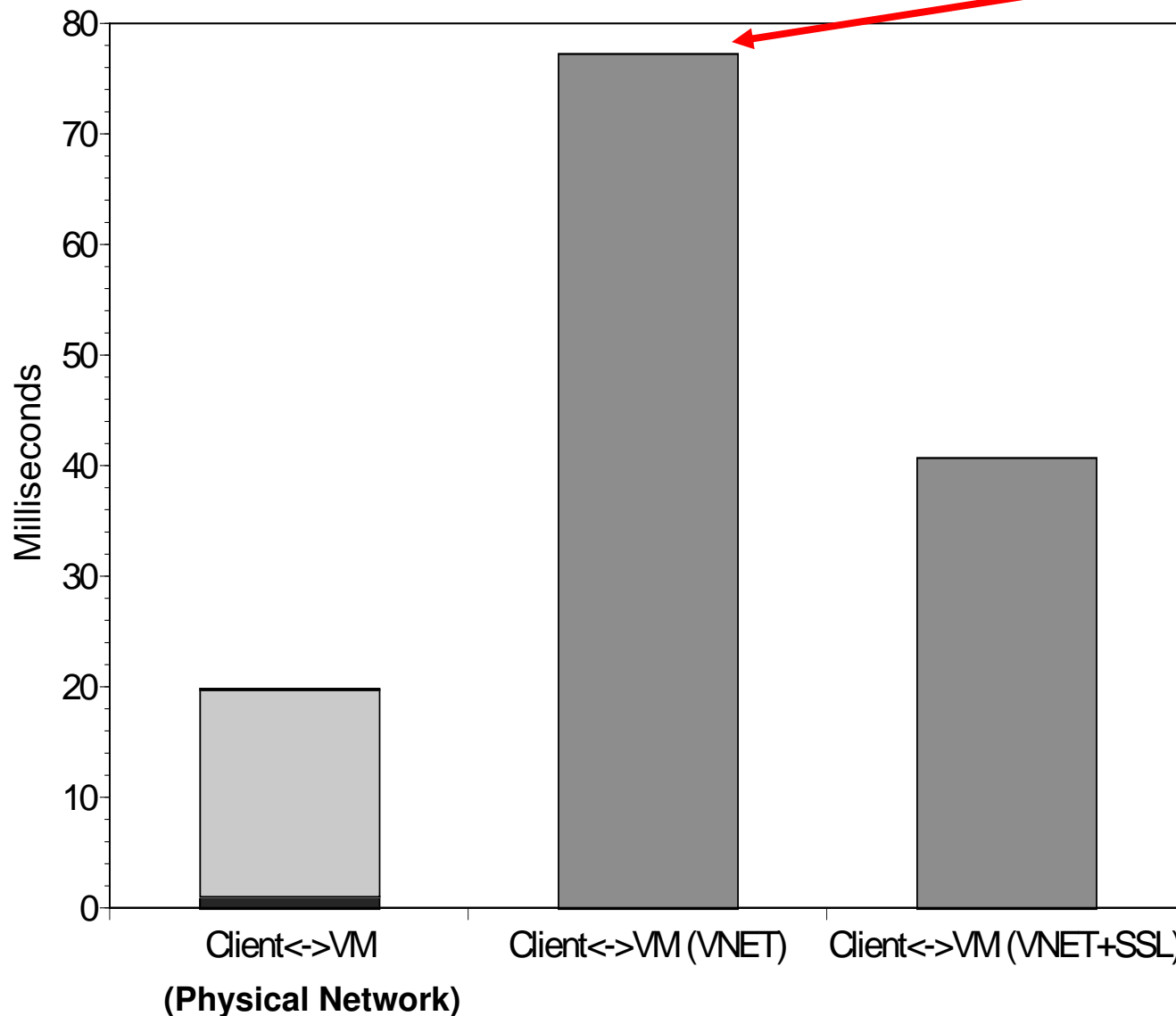


Local area configuration

Average latency over WAN



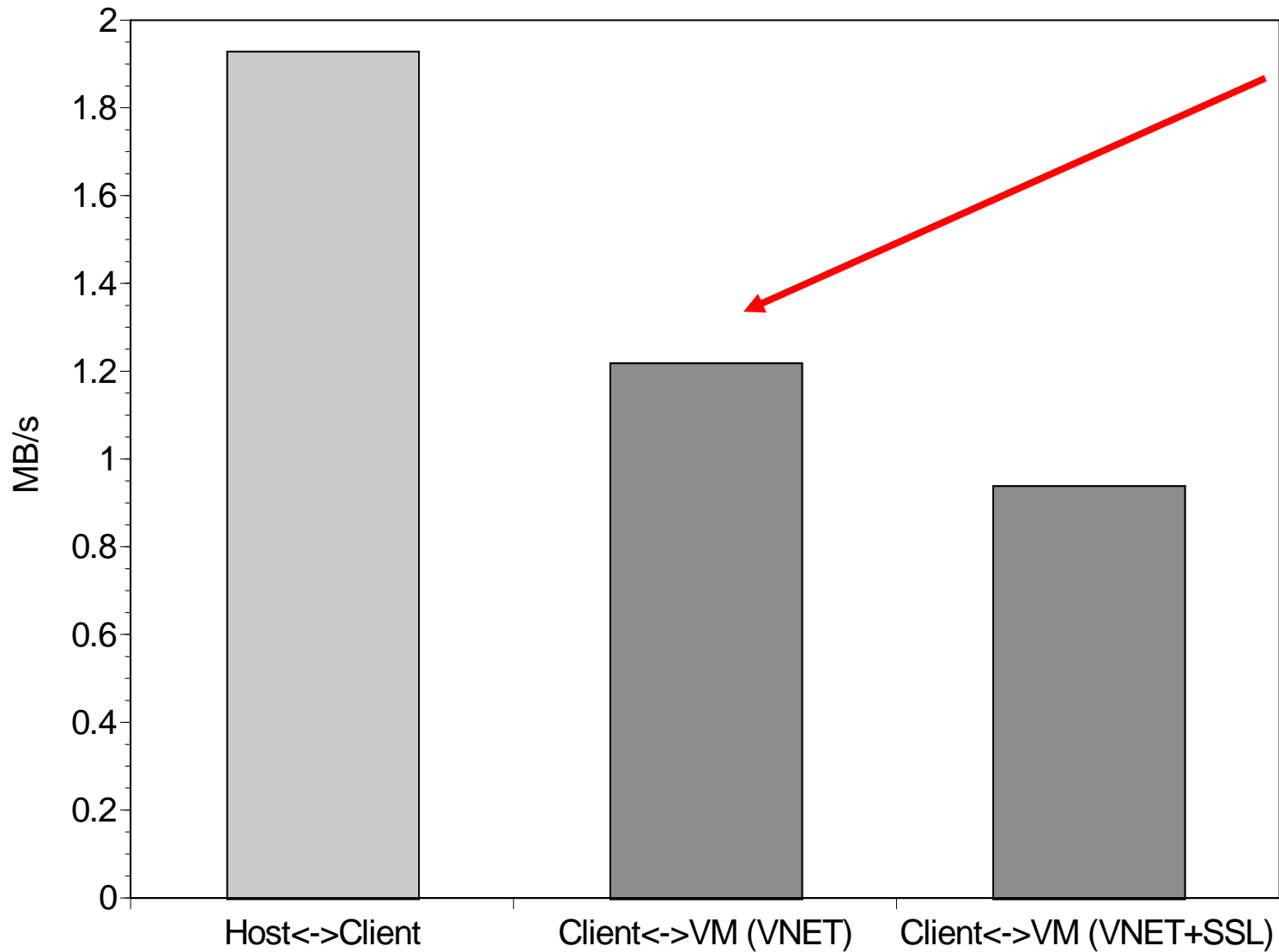
Standard deviation of latency over WAN



What: VNET increases variability in latency

Why: TCP connection between VNET servers trades packet loss for increased delay

Bandwidth over WAN



Expectation:

VNET to achieve throughput comparable to the physical network

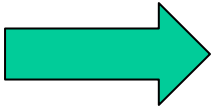
What do we see:

VNET achieves lower than expected throughput

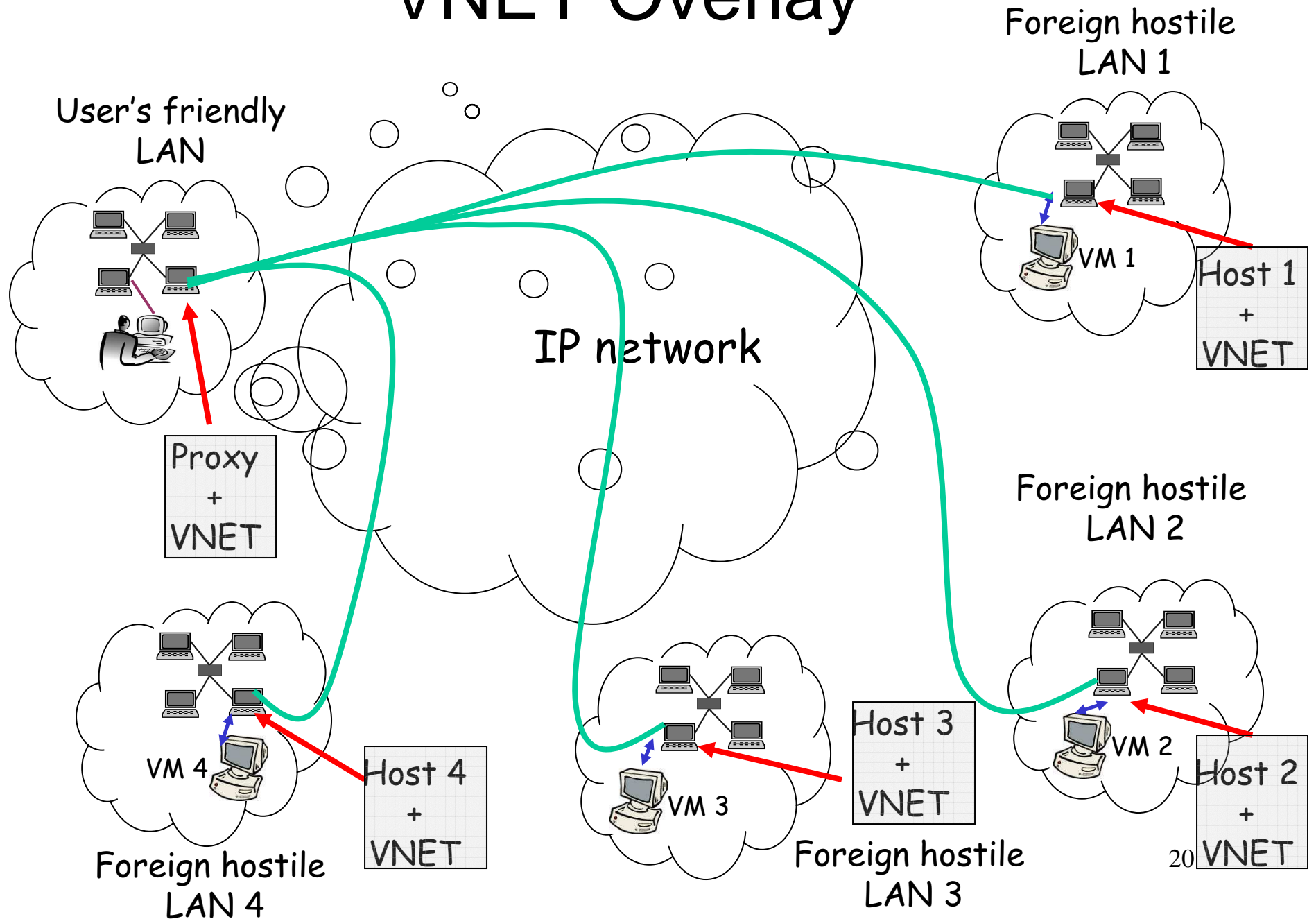
Why:

VNET's is tricking TTCP's TCP connection

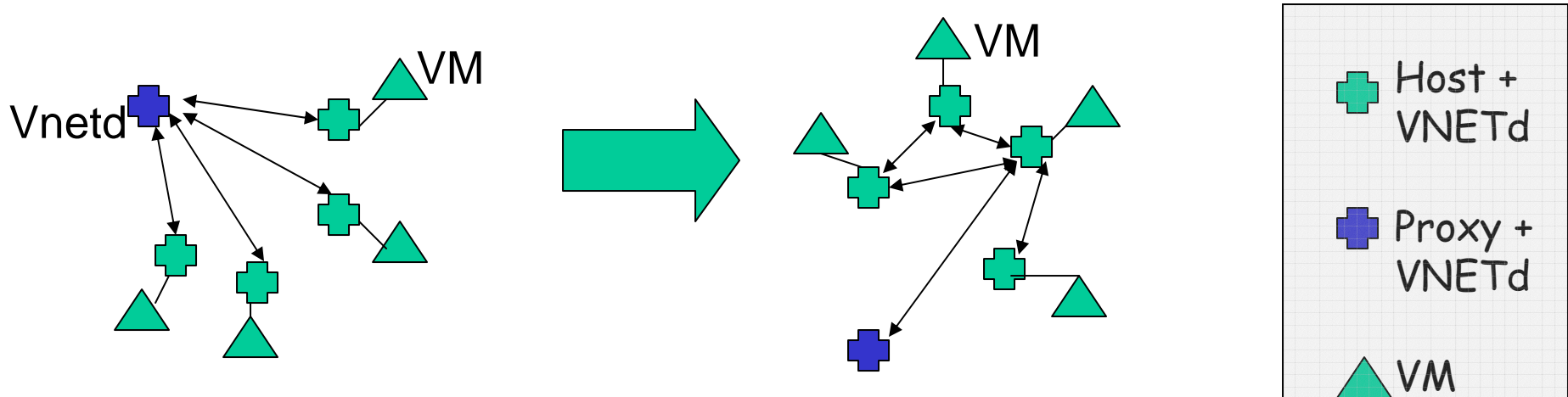
Outline

- Virtual machine grid computing
- Virtuoso system
- Networking challenges in Virtuoso
- Enter VNET
- VNET  Adaptive virtual network
- Related Work
- Conclusions
- Current Status

VNET Overlay



Bootstrapping the Virtual Network



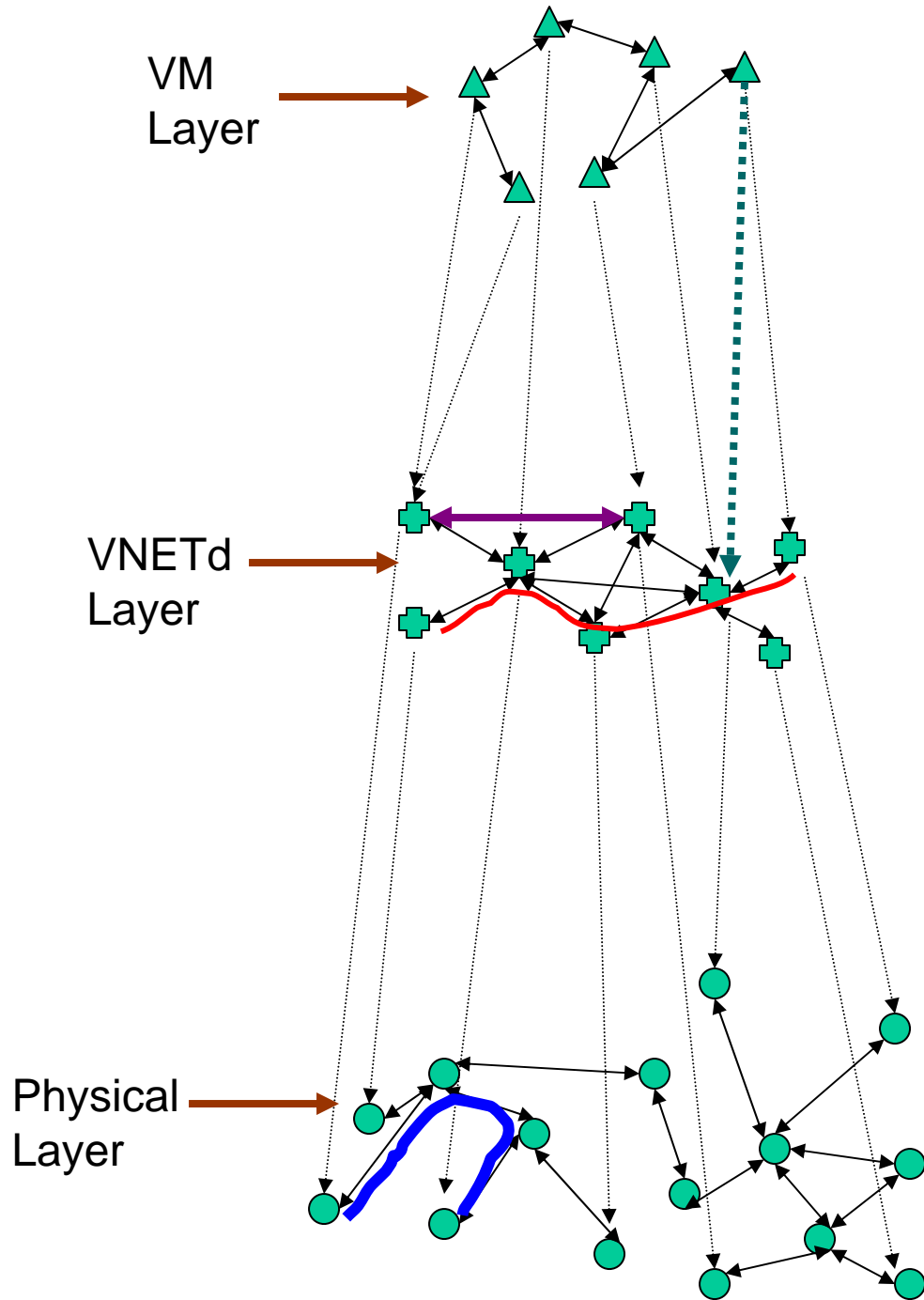
- Star topology always possible
- Topology may change
 - Links can be added or removed on demand
 - Virtual machines can migrate
- Forwarding rules can change
 - Forwarding rules can be added or removed on demand

Application communication topology and traffic load; application processor load


Vnetd layer can collect all this information as a side effect of packet transfers and invisibly act

- VM migrates
- Topology changes
- Routing change
- Reservation

Network bandwidth and latency; sometimes topology



Outline

- Virtual machine grid computing
- Virtuoso system
- Networking challenges in Virtuoso
- Enter VNET
- VNET  Adaptive virtual network
- Related Work
- Conclusions
- Current Status

Related Work

- Collective / Capsule Computing (Stanford)
 - VMM, Migration/caching, Hierarchical image files
- Denali (U. Washington)
 - Highly scalable VMMs (1000s of VMMs per node)
- **SODA and VIOLIN (Purdue)**
 - Virtual Server, fast deployment of services
- VPN
- Virtual LANs, IEEE
- Overlay Networks: RON, Spawning networks, Overcast
- Ensim
- Virtuozzo (SWSoft)
 - Ensim competitor
- Available VMMs: IBM's VM, VMWare, Virtual PC/Server, Plex/86, SIMICS, Hypervisor, VM/386

Conclusions

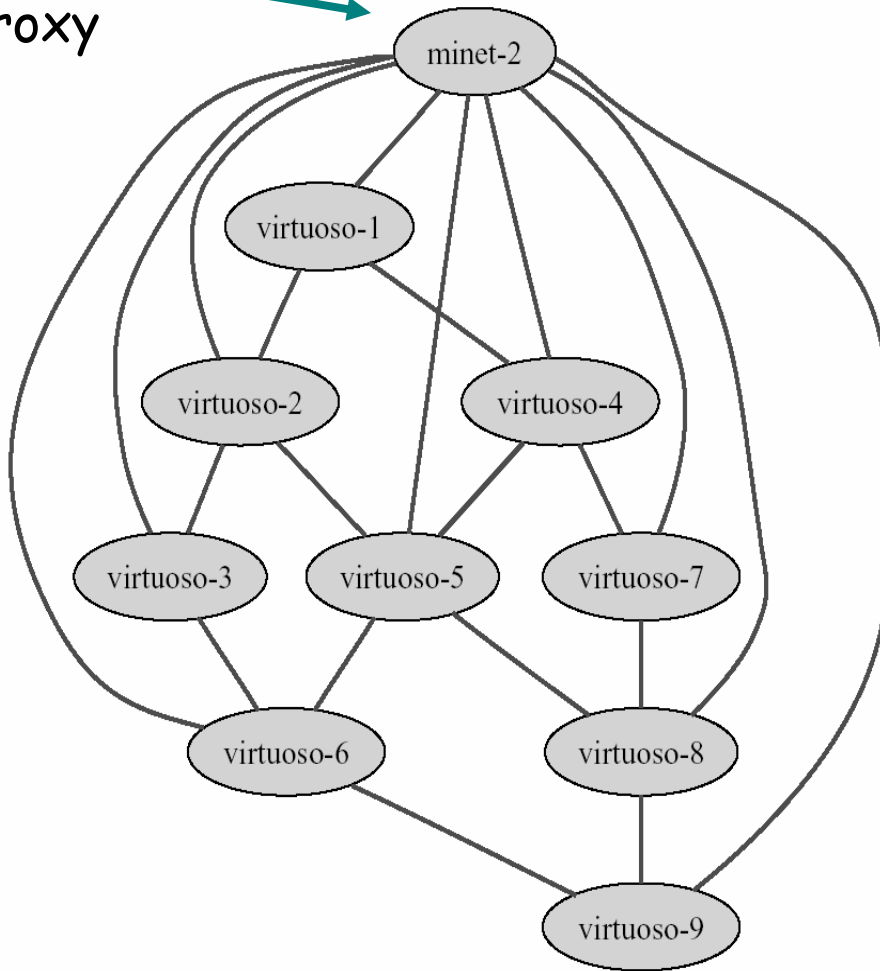
- There exists a strong case for grid computing using virtual machines
- Challenging network management problem induced by VMs in the grid environment
- Described and evaluated a tool, VNET, that solves this problem
- Discussed the opportunities, the combination of VNET and VMs present, to exploit an adaptive overlay network

Current Status

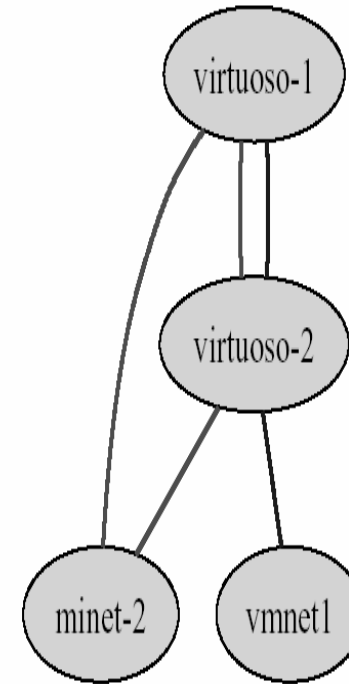
- Application traffic load measurement and topology inference [Ashish Gupta]
- Support for arbitrary topologies and forwarding rules
- Dynamic adaptation to improve performance

Current Status Snapshots

Pseudo proxy



Required Topology of VNET Daemons -- Mesh topology



Blue indicates the route taken for 00:50:56:00:11:00-00:50:56:00:12:00

Red indicates the topology edges



- For More Information
 - Prescience Lab (Northwestern University)
 - <http://plab.cs.northwestern.edu>
 - Virtuoso: Resource Management and Prediction for Distributed Computing using Virtual Machines
 - <http://virtuoso.cs.northwestern.edu>
- VNET is publicly available from
 - <http://virtuoso.cs.northwestern.edu>