

# SmartSwitch: Blurring the Line Between Network Infrastructure and Cloud Applications

Wei Zhang and **Timothy Wood**, *George Washington University*  
K.K. Ramakrishnan, *Rutgers University*  
Jinho Hwang, *IBM Research*

# Networks are Changing

## Scale and organization:

- Cloud data centers, mobile users
- Large-scale, highly dynamic

## Hardware and control plane

- Software Defined Networking (SDN)
- Network Function Virtualization (NFV)

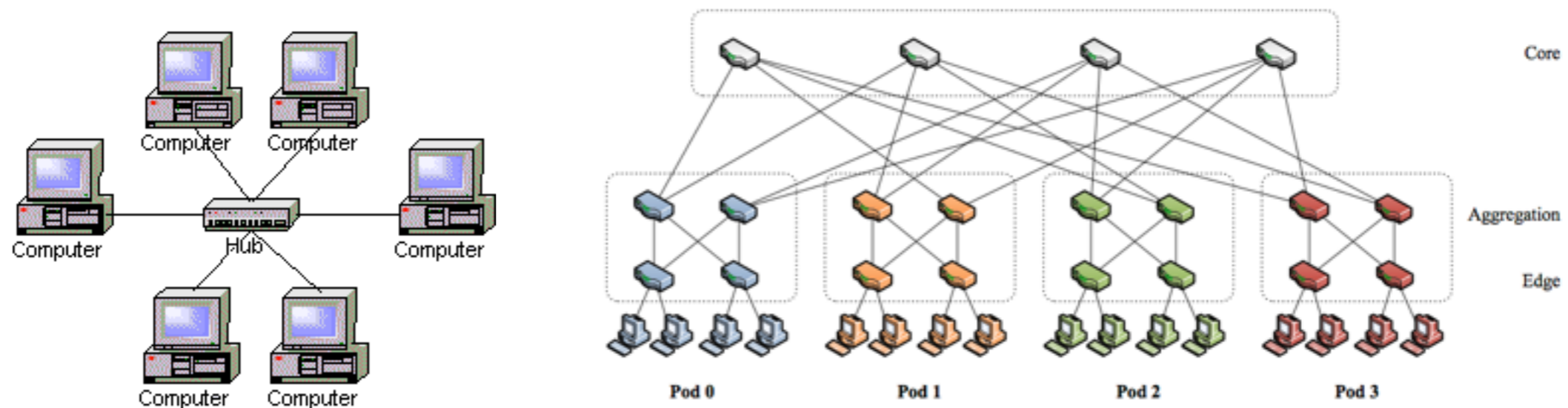


Figure 1: Sample fat tree topology.

# High Performance VM Networking

## Hardware Routers and Switches

- Expensive, single purpose
- Controllable with SDNs, but not flexible



# High Performance VM Networking

## Hardware Routers and Switches

- Expensive, single purpose
- Controllable with SDNs, but not flexible



## PacketShader [Han, SIGCOMM '10]

- Use commodity servers and GPUs
- 39 Gbps processing rates



## Netmap [Rizzo, ATC '12] and DPDK

- Libraries to provide zero-copy network processing on commodity 10gbps NICs



## ClickOS [Martins, NSDI '14] and NetVM [Hwang, NSDI '14]

- VM based network services
- Flexible deployment and composition



# Problems and Opportunities

Dynamic workloads, applications, and users

- Migrating between servers and data centers
- Networks need the flexibility to adapt

Network is more than just transport

- Security, QoS, accounting, caching, transcoding
- Network must be powerful

# Problems and Opportunities

Dynamic workloads, applications, and users

- Migrating between servers and data centers
- Networks need the flexibility to adapt

Network is more than just transport

- Security, QoS, accounting, caching, transcoding
- Network must be powerful

Opportunity to rethink!

Virtualization provides flexible deployment, efficient resource usage, and greater power

# Outline

Background & Motivation

High Performance Networking in VMs

## **Smart Switch**

- Application-aware Networking
- Network-integrated Storage
- Computation in the Network

Prototype Evaluation

Conclusions

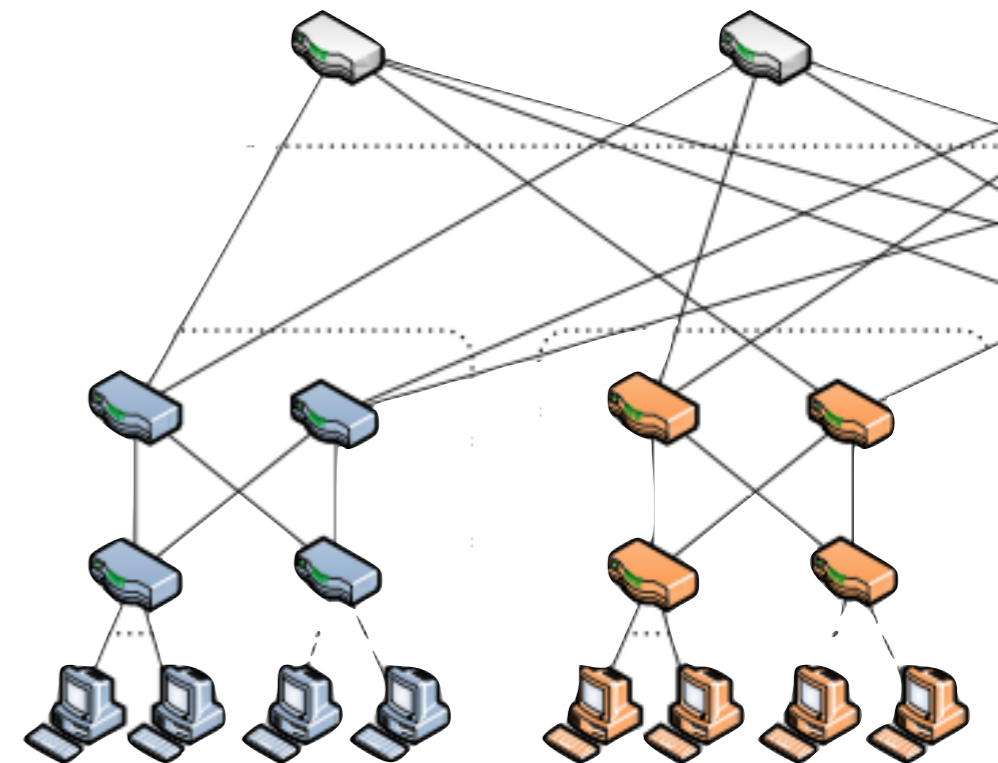
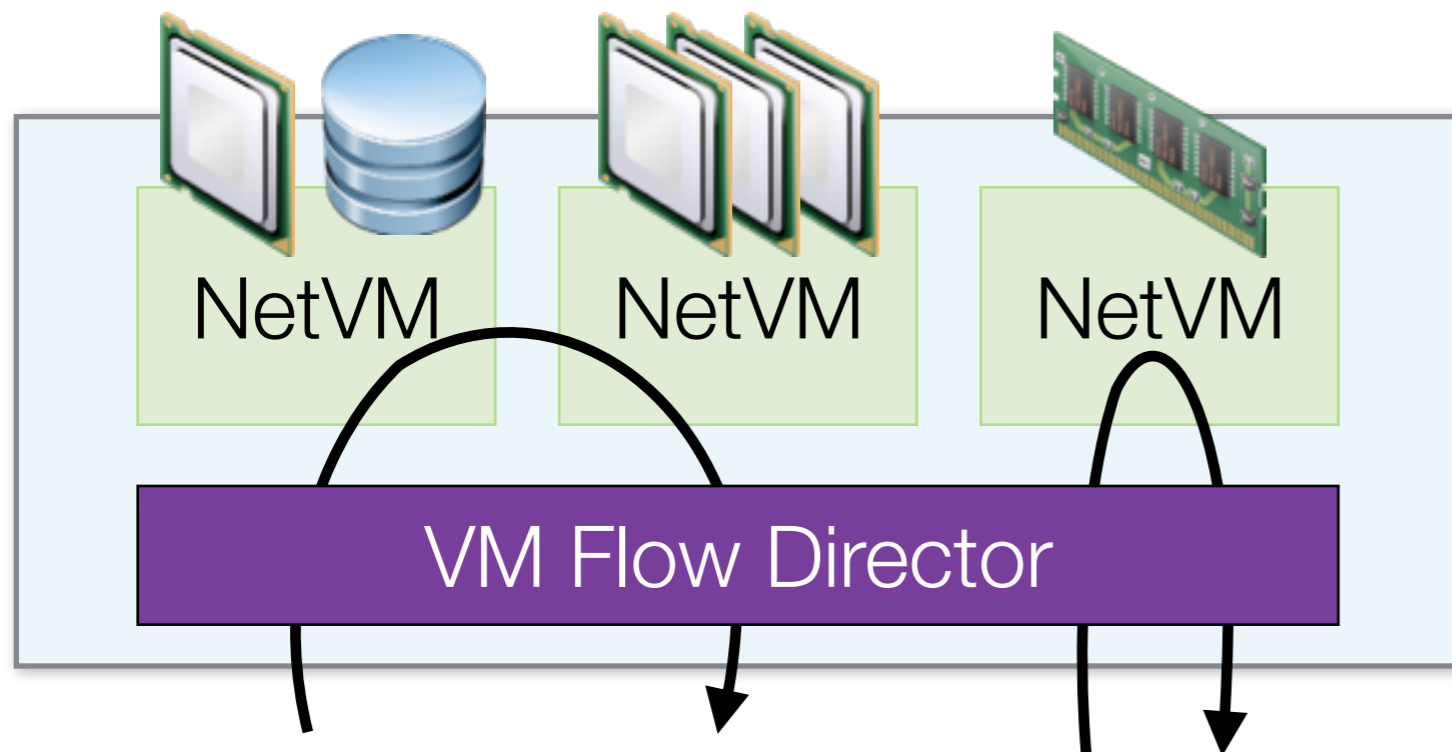
# Smart Switch Platform

## Virtualized High Performance Networking

- Each VM runs a network service or application component
- Dynamically assigned hardware resources

## VM Flow Director

- Assigns incoming packets to VMs based on flexible rules
- Complex services can be composed of chains of VMs





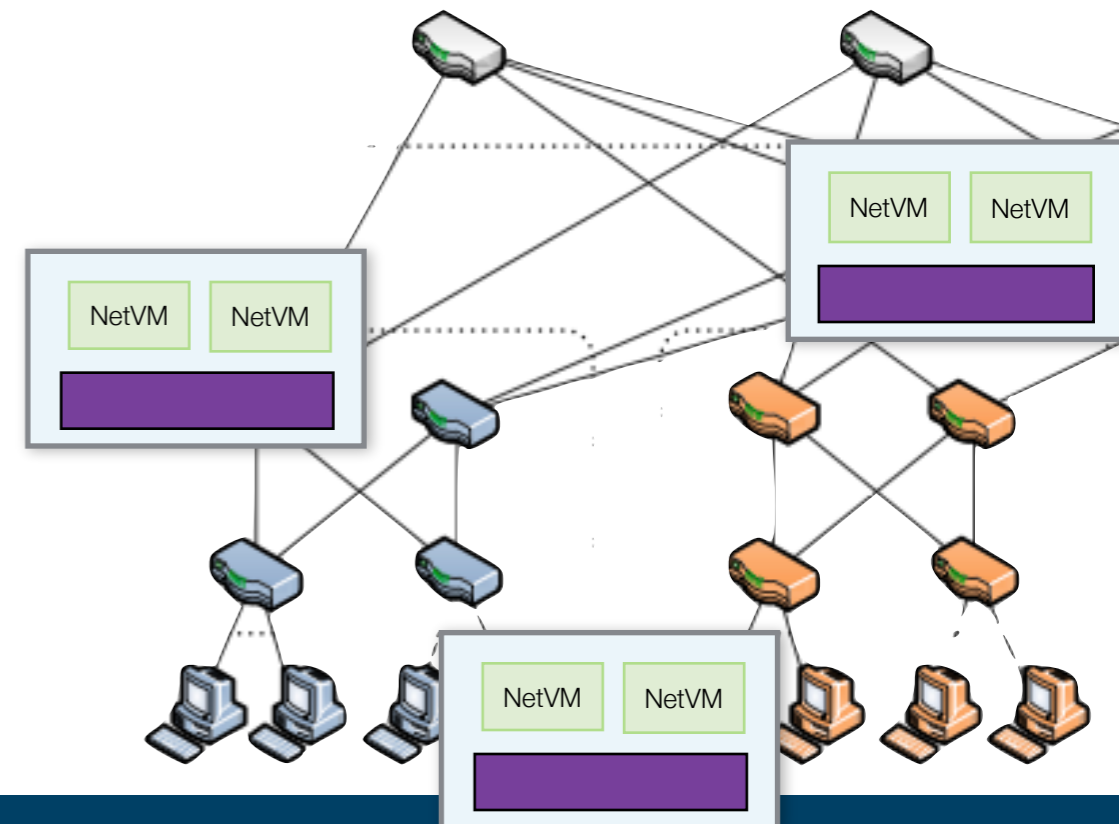
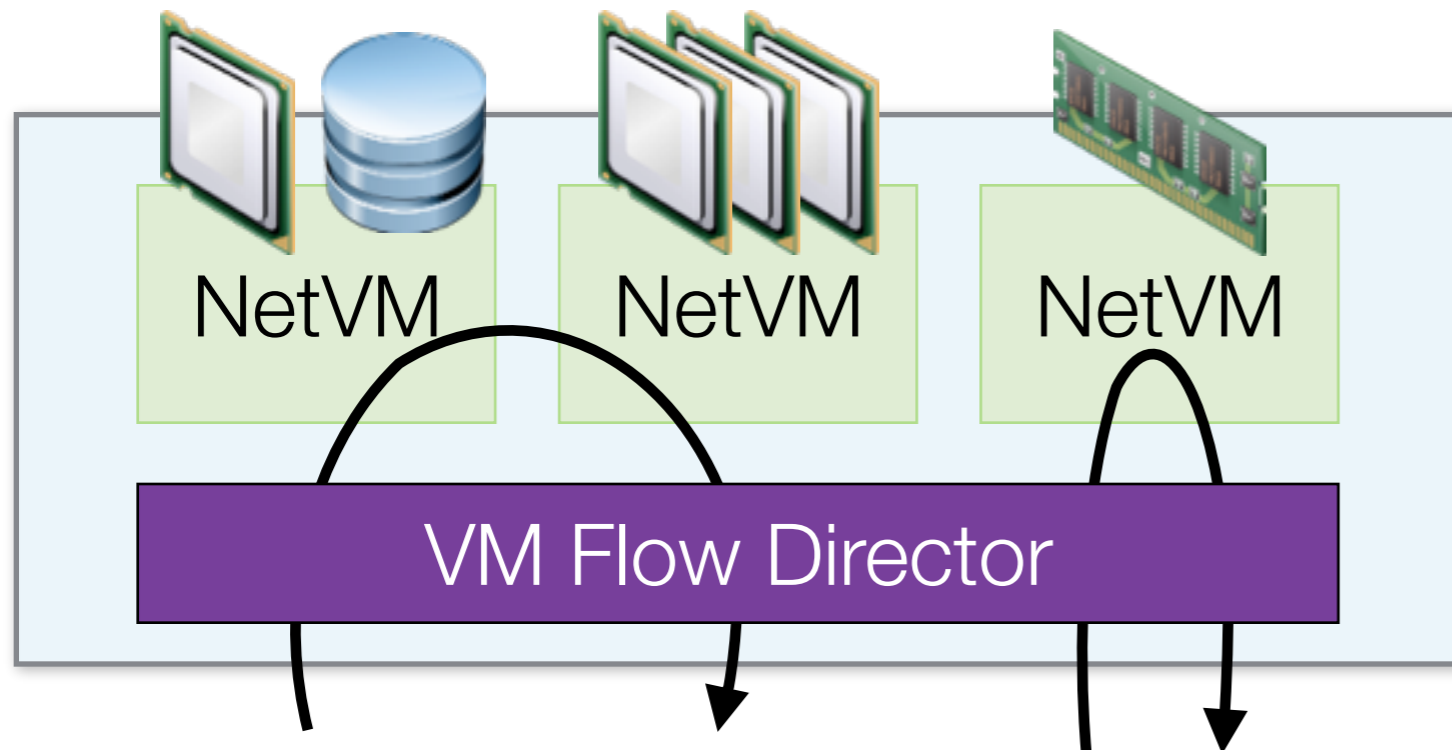
# Smart Switch Platform

## Virtualized High Performance Networking

- Each VM runs a network service or application component
- Dynamically assigned hardware resources

## VM Flow Director

- Assigns incoming packets to VMs based on flexible rules
- Complex services can be composed of chains of VMs



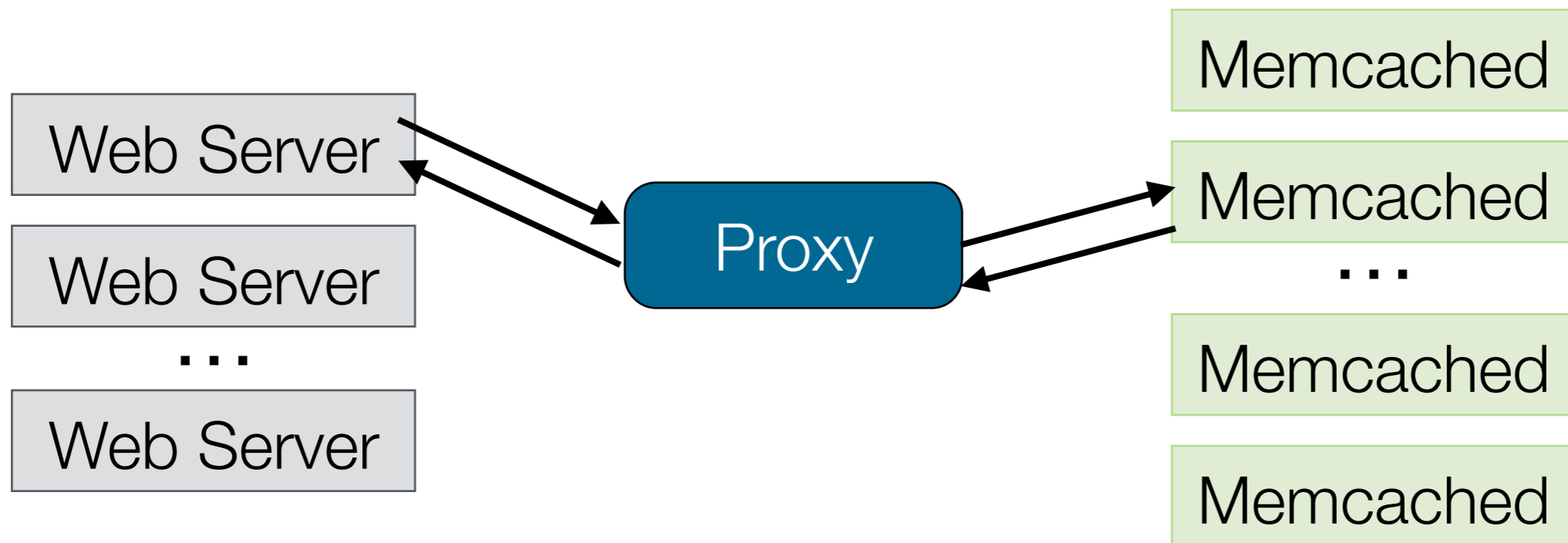
# Large Scale Memcached

Shard data across memcached nodes

- Consistent hashing [Facebook, NSDI 13]

Use proxies to direct requests

- Moxi, Twemproxy, etc



Proxy can get overloaded, adds latency to each request

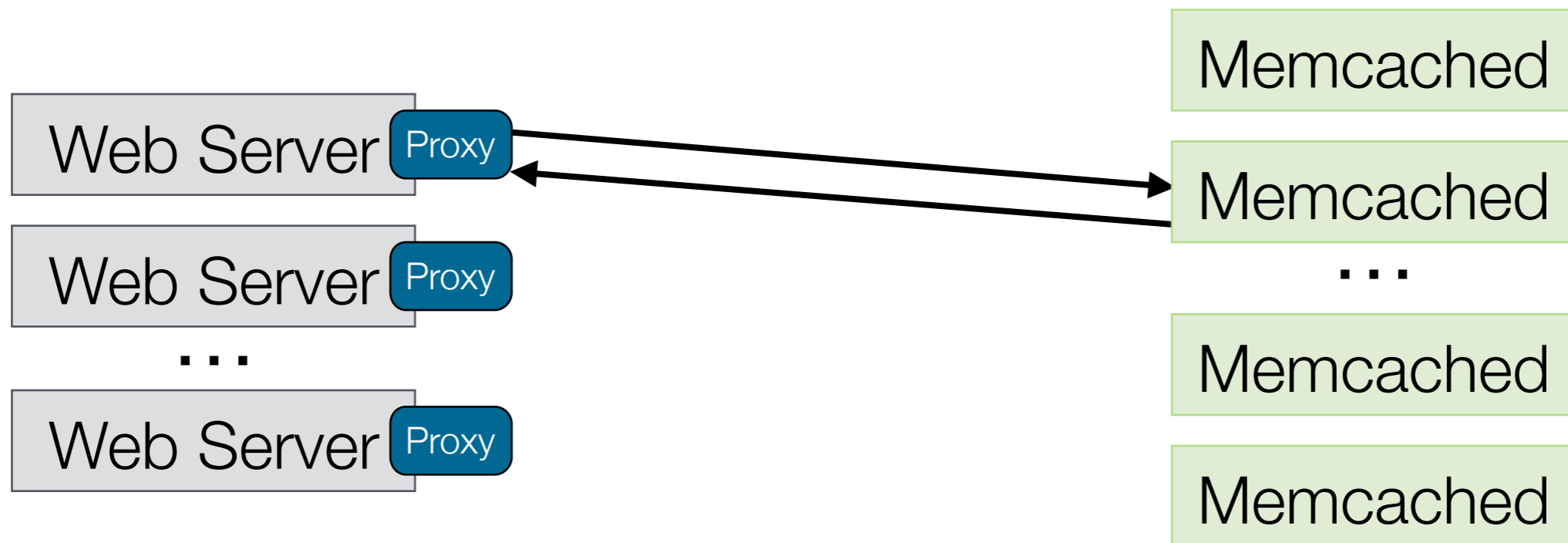
# Large Scale Memcached

Shard data across memcached nodes

- Consistent hashing [Facebook, NSDI 14]

Use proxies to direct requests

- Moxi, Twemproxy, etc

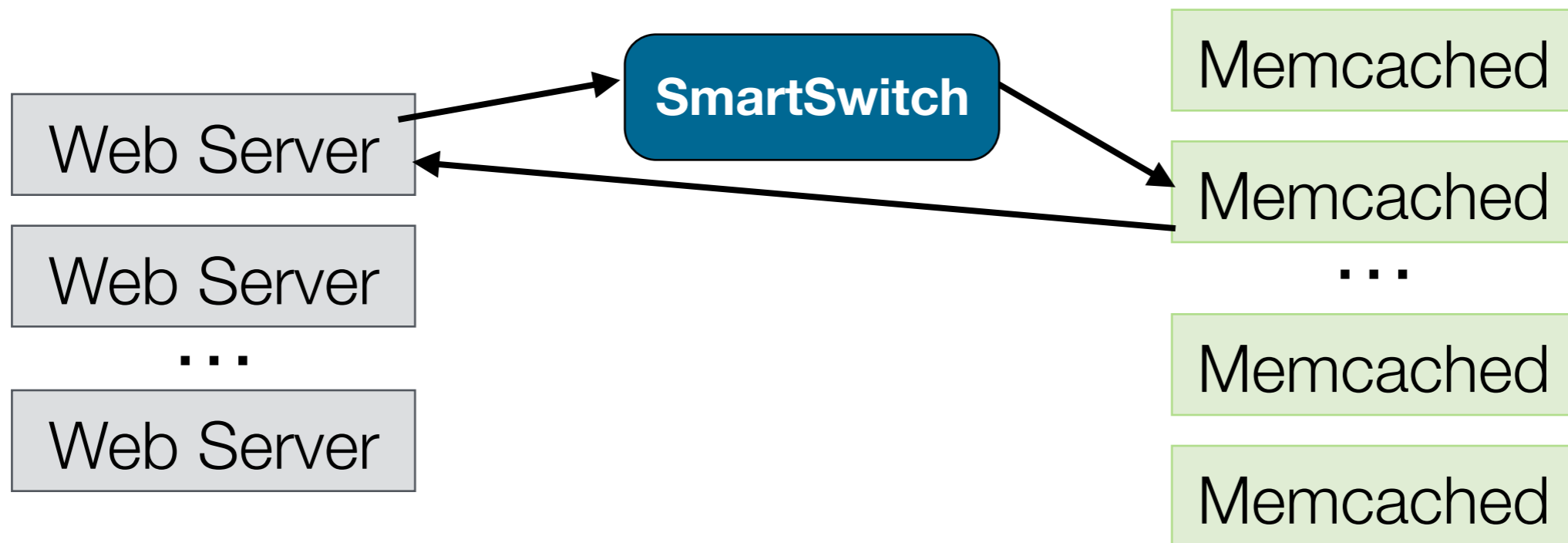


Difficult to update mapping, no centralized control

# Large Scale Memcached

There's probably a switch between the web servers and the memcached nodes...

Memcached-aware SmartSwitch parses packet body and routes packet to the appropriate server.



Return path needs no processing, reducing overhead

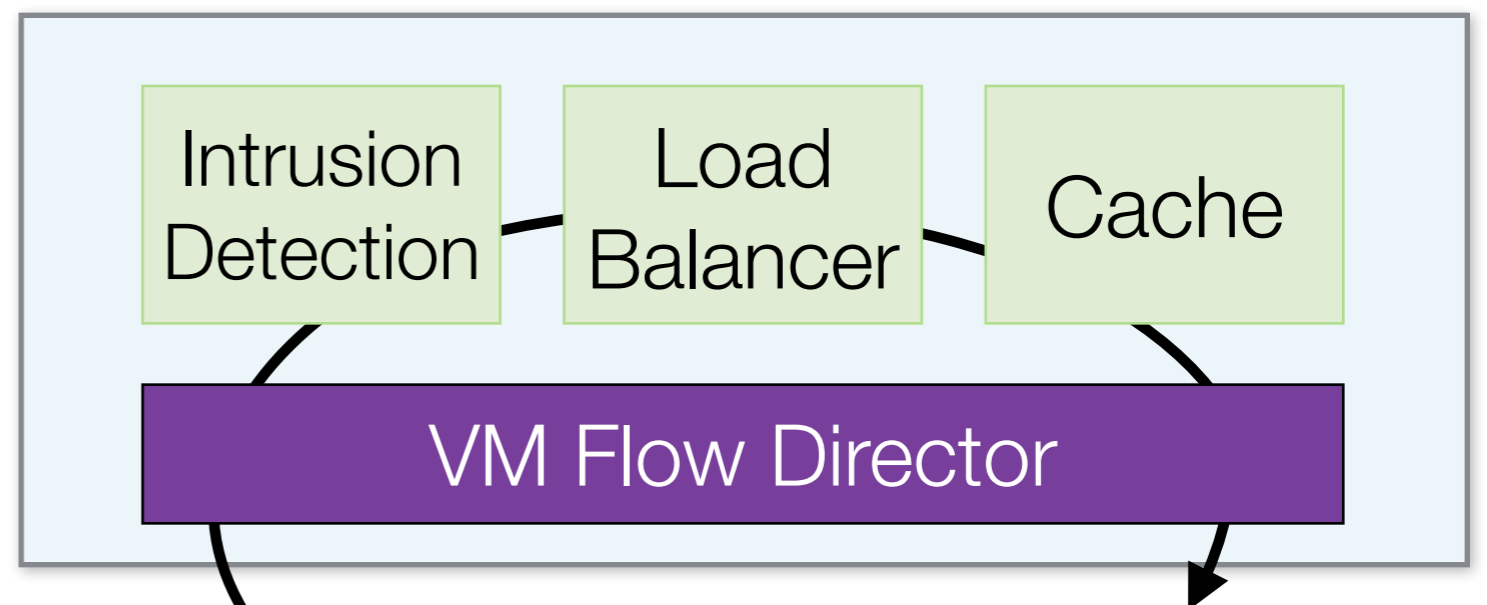
# Application-aware Networking

Network components like switches and routers can be customized to application needs

Run inside standard VMs

- Familiar programming environment
- Each component is isolated
- Easily deployed when and where needed

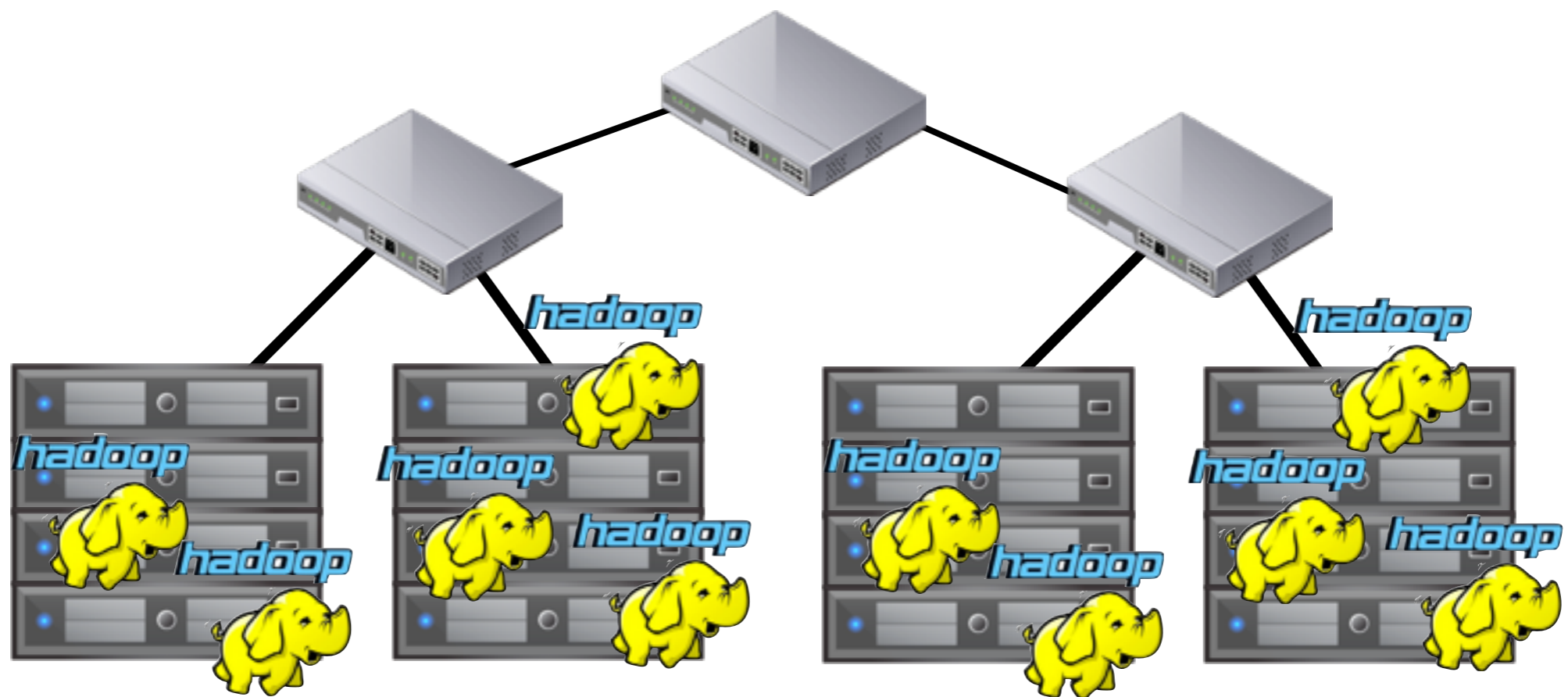
Complex services can be created by chaining multiple VMs



# Storage in the Network

Commodity servers can easily support both fast network cards and large storage capacity

- PCI-express based SSDs can operate at close to 10gbps rates

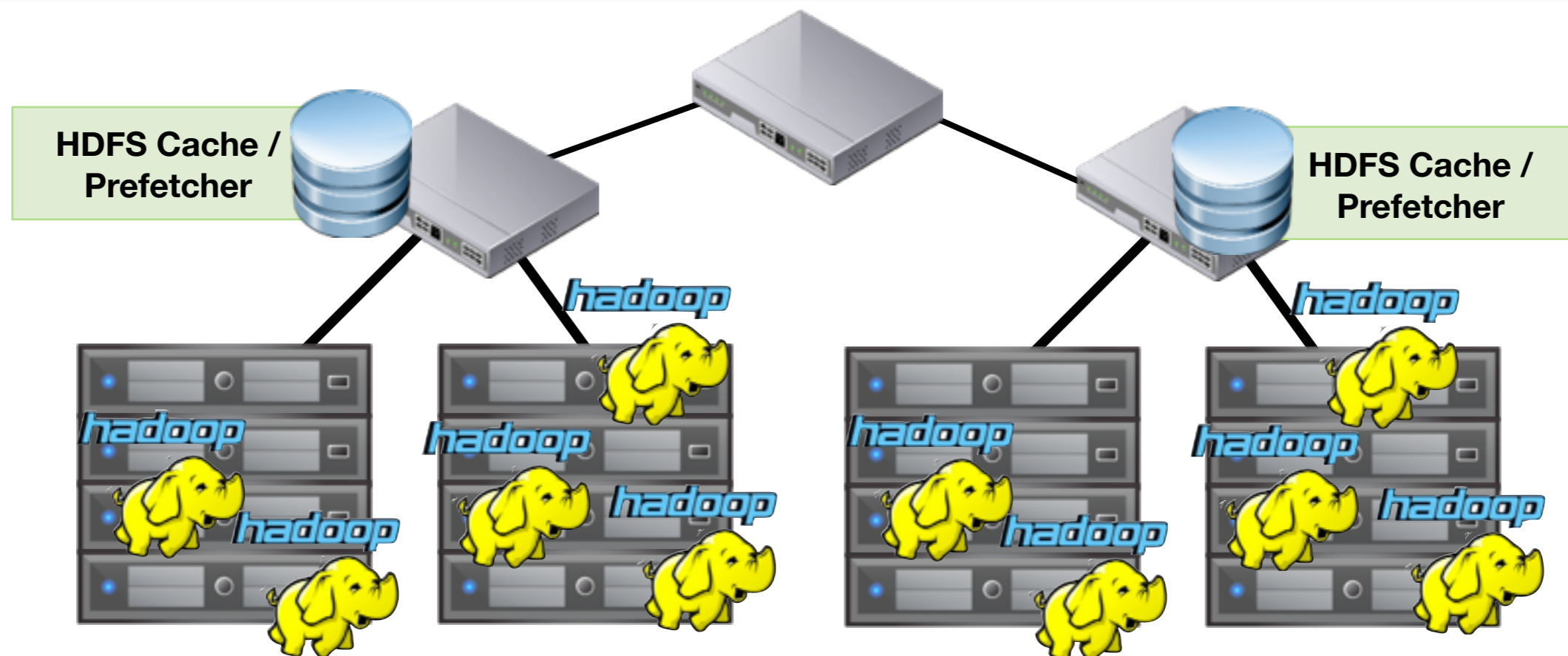


# Storage in the Network

Commodity servers can easily support both fast network cards and large storage capacity

- PCI-express based SSDs can operate at close to 10gbps rates

Storage doesn't just need to be at network leaf nodes

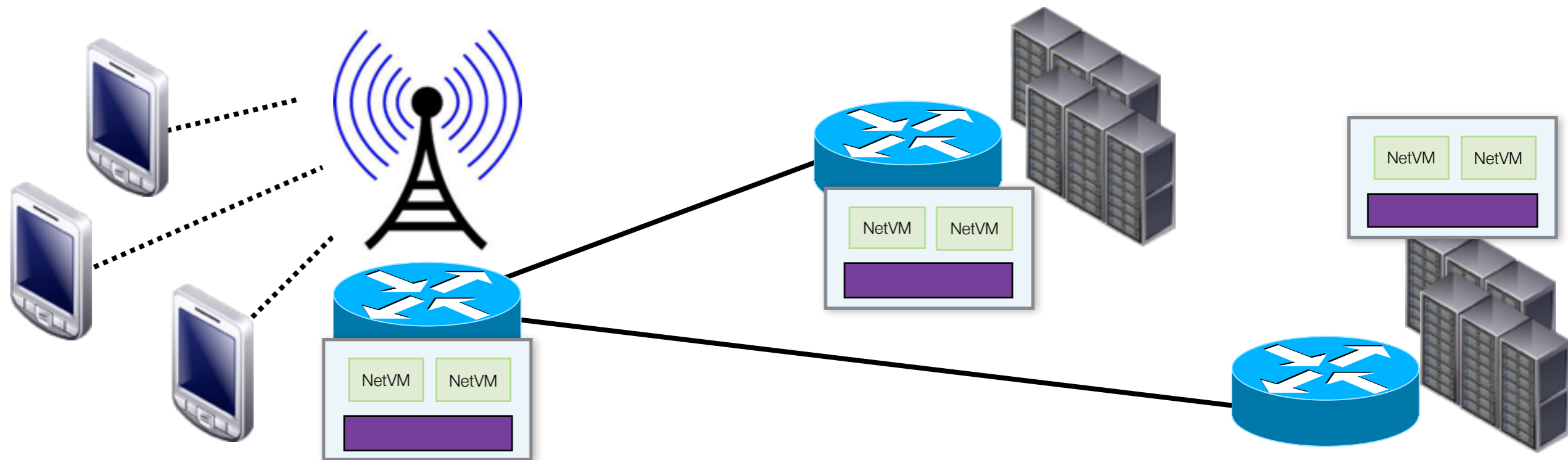


# Network-based Computation

Applications can be spread across computing and networking infrastructure

- Video transcoding at edge routers
- Caching at cell tower

Dynamically instantiated, moved, replicated as needed





# Outline

Background & Motivation

High Performance Networking in VMs

Smart Switch

- Application-aware Networking
- Network-integrated Storage
- Computation in the Network

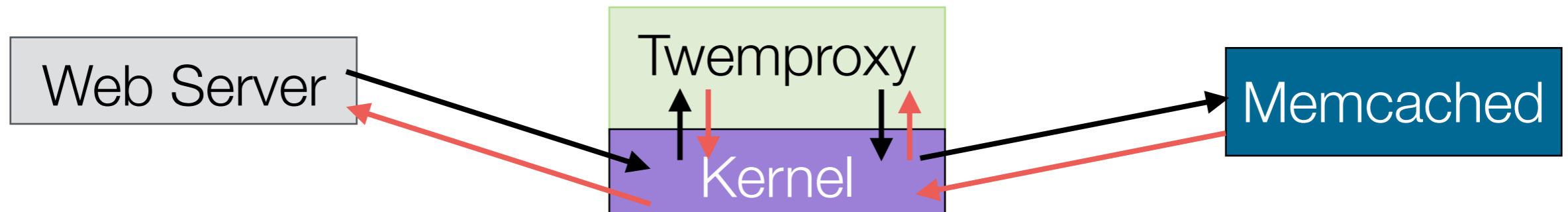
**Prototype Evaluation**

Conclusions

# MemSwitch Prototype

## Twemproxy

- Establishes connection with web server and memcached nodes

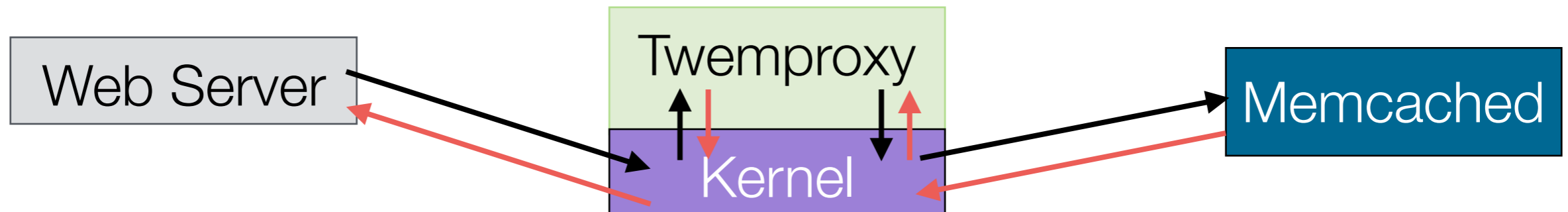


- 4 packet copies + 4 DMAs

# MemSwitch Prototype

## Twemproxy

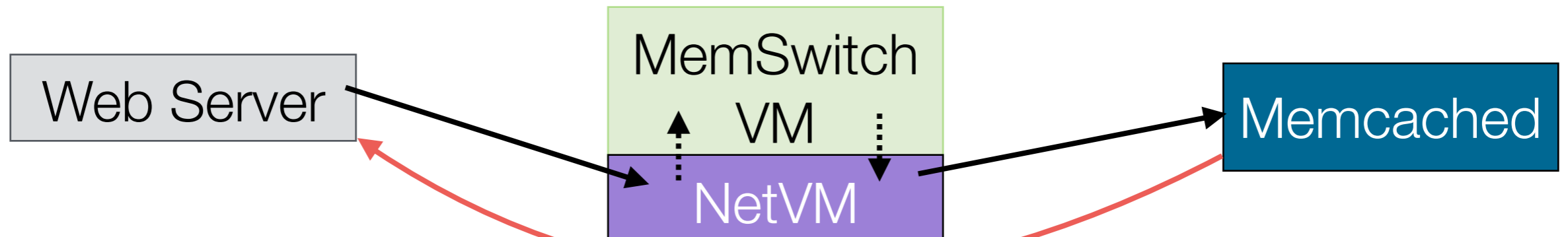
- Establishes connection with web server and memcached nodes



- 4 packet copies + 4 DMAs

## MemSwitch

- Examine packet to find key and rewrite destination address

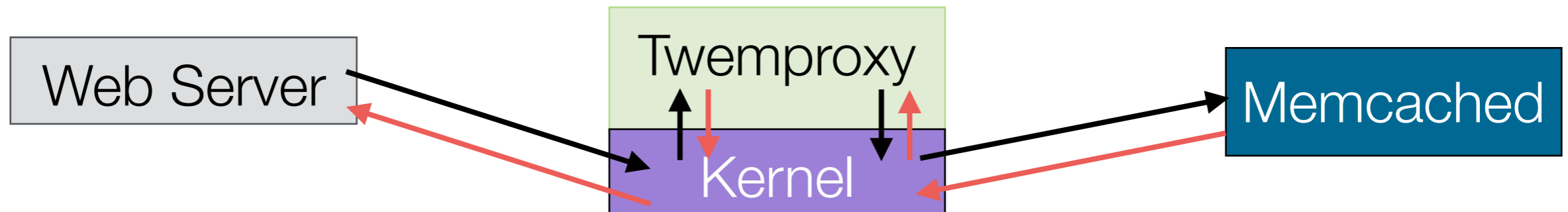


- Zero packet copies + 2 DMAs
- Reply directly to web server

# MemSwitch Prototype

## Twemproxy

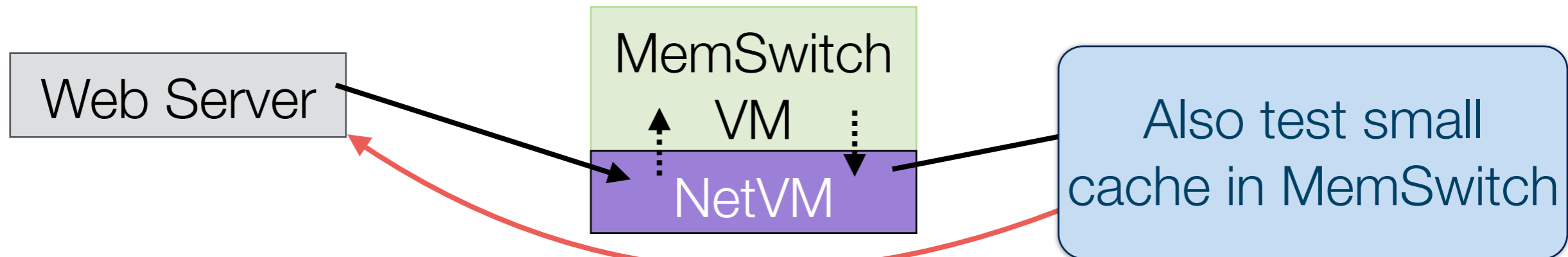
- Establishes connection with web server and memcached nodes



- 4 packet copies + 4 DMAs

## MemSwitch

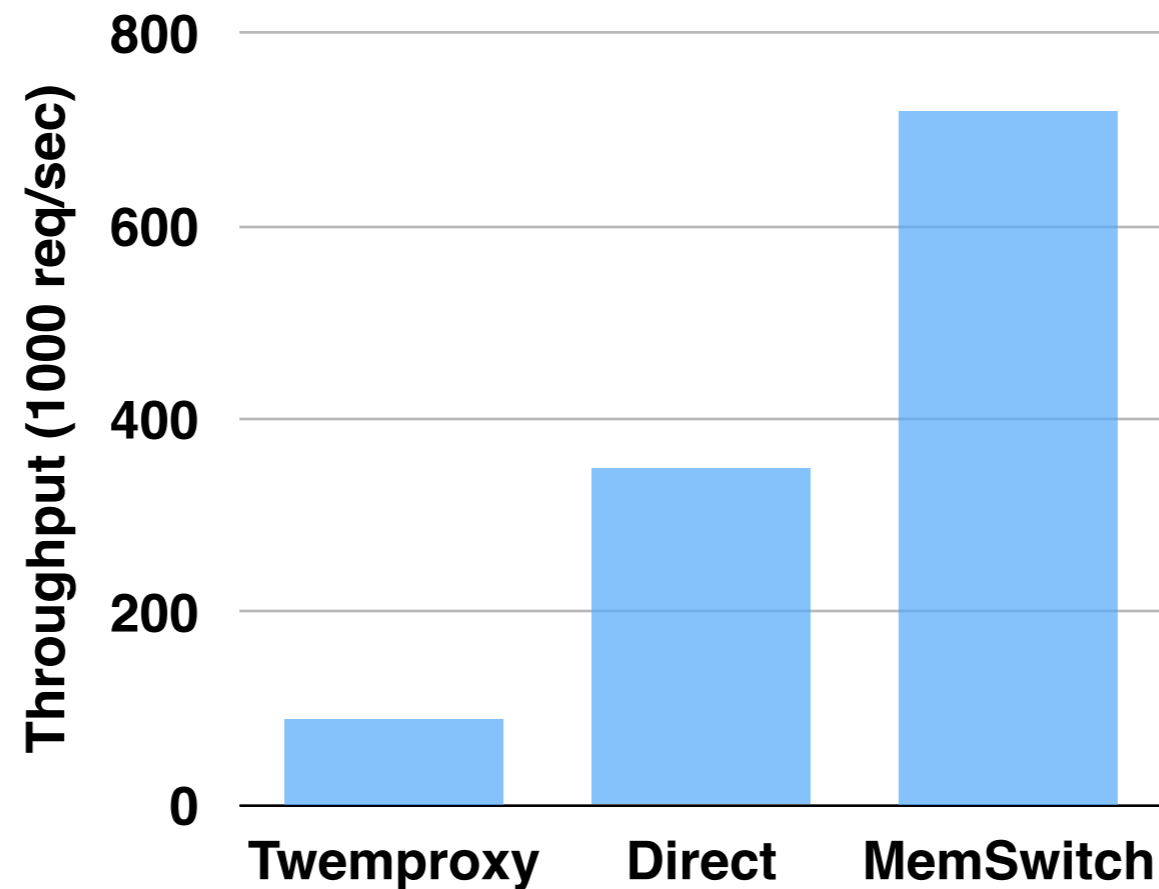
- Examine packet to find key and rewrite destination address



- Zero packet copies + 2 DMAs
- Reply directly to web server

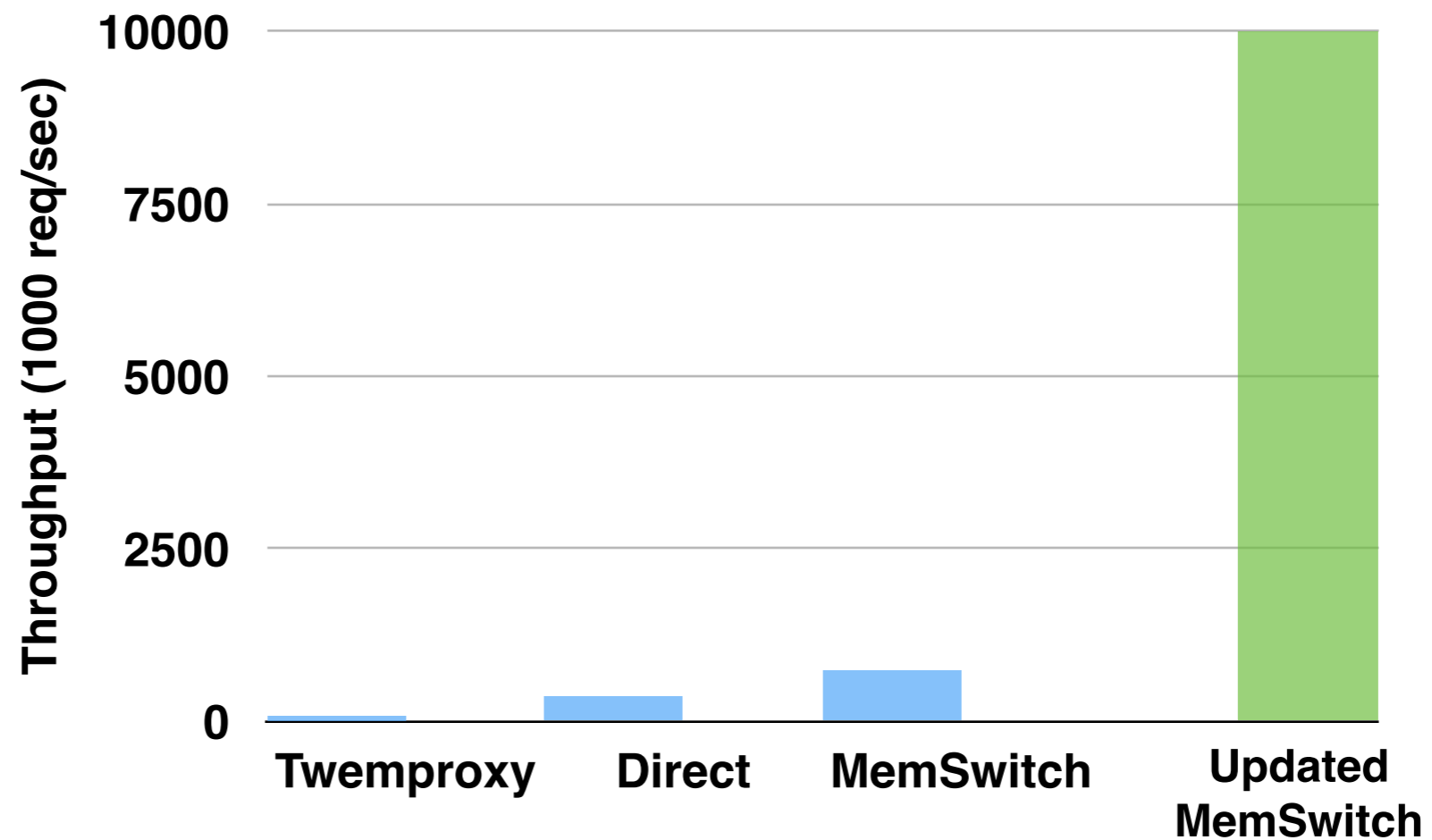
# Performance

Proxy request redirection throughput



# Performance

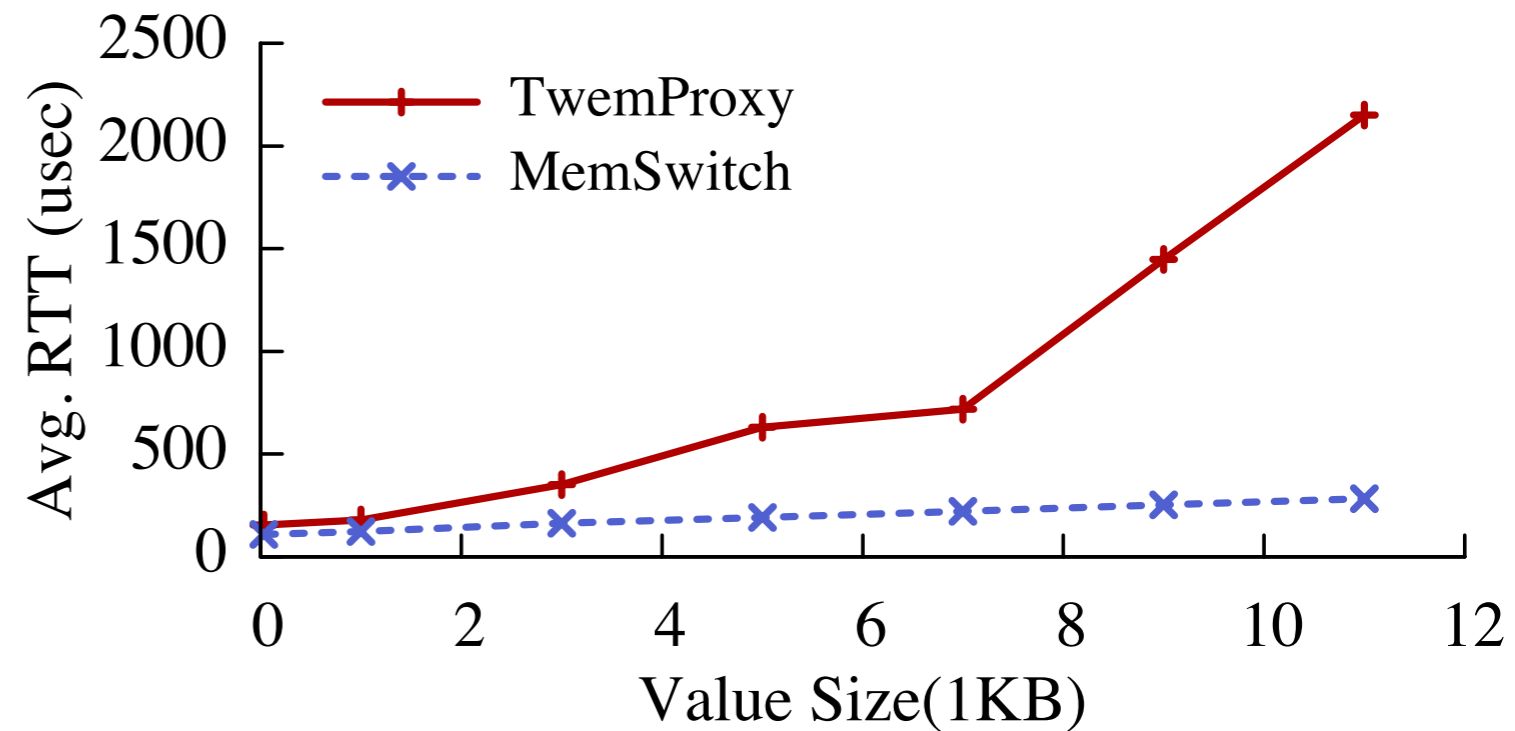
Proxy request redirection throughput



10 Million req/sec using one CPU core and one 10gbps NIC port

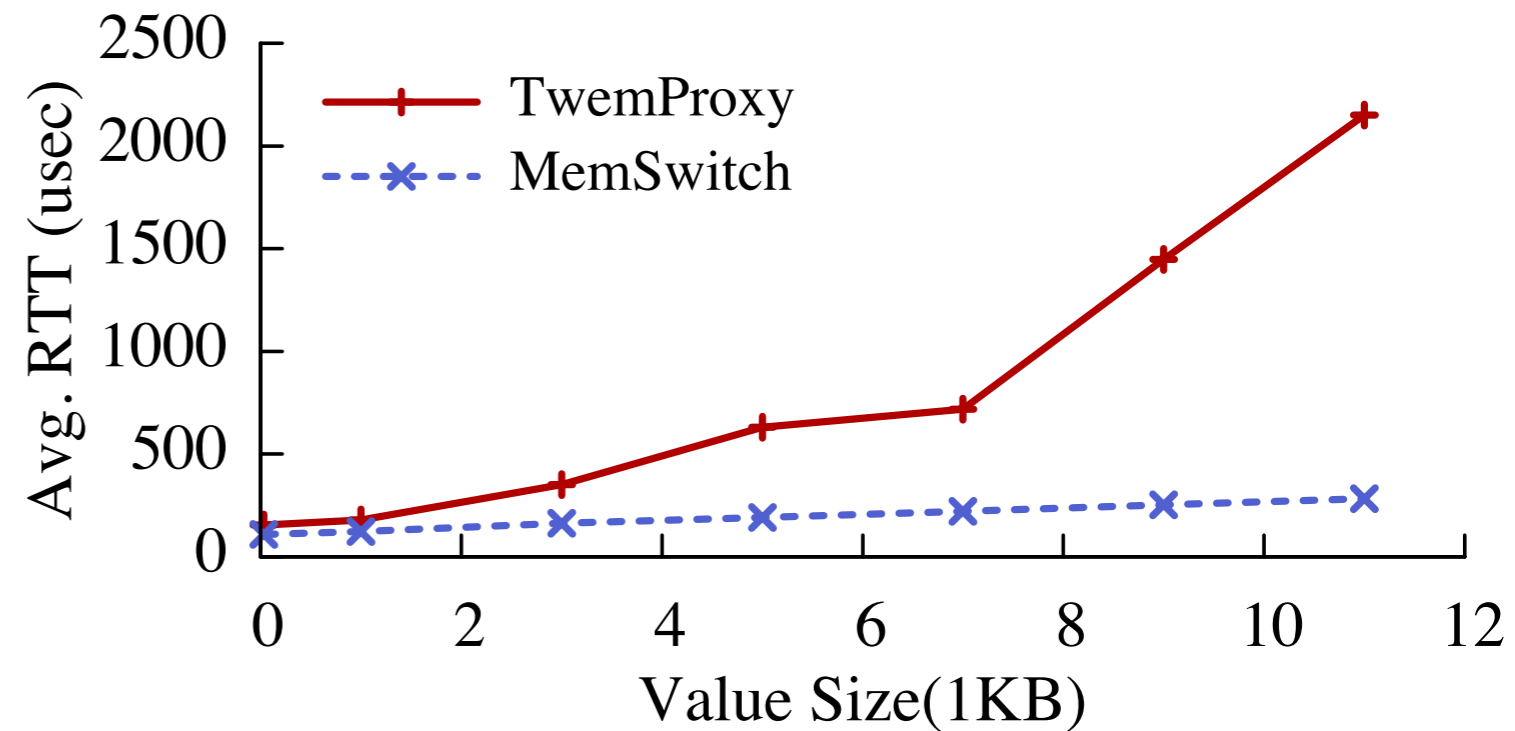
# Overheads

	Latency (usec)
Direct	175
Twemproxy	315
MemSwitch	208
w/cache	19



# Overheads

	Latency (usec)
Direct	175
Twemproxy	315
MemSwitch	208
w/cache	19

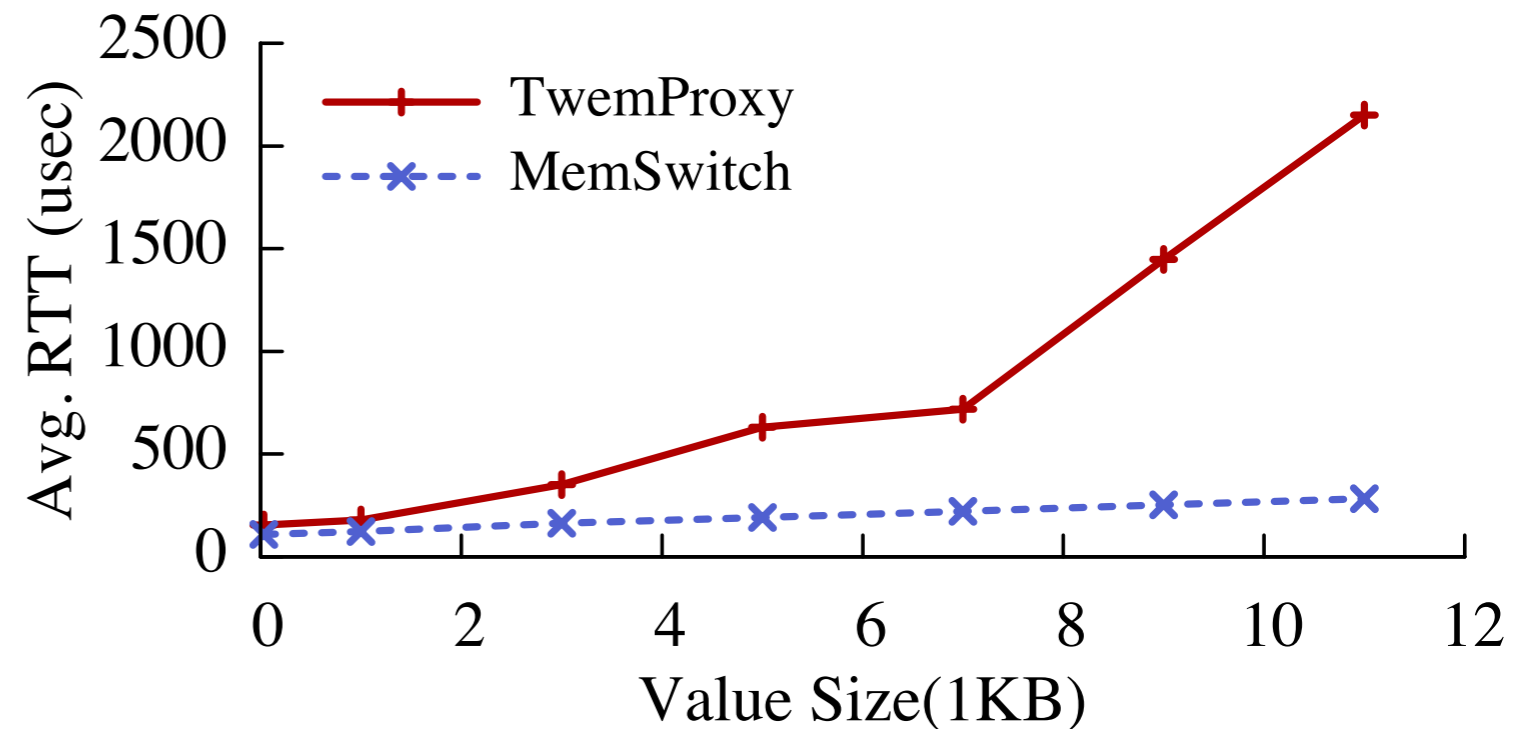


Small latency overhead



# Overheads

	Latency (usec)
Direct	175
Twemproxy	315
MemSwitch	208
w/cache	19



Small latency overhead

Not affected by data size since replies are sent directly to web server.

# Conclusions & Future Work

Networks are becoming more dynamic

- How to add flexibility to network processing elements?
- Where should functionality be deployed?

## SmartSwitch

- Efficient network processing in VMs
- Customize network services for applications
- Add storage and computation to network services
- Easy and flexible to deploy and manage

## Remaining Challenges

- How to manage a data center or WAN deployment of SmartSwitches?
- What is the right programming model to expose network functions to applications and vice versa?