OPENFABRICS ALLIANCE

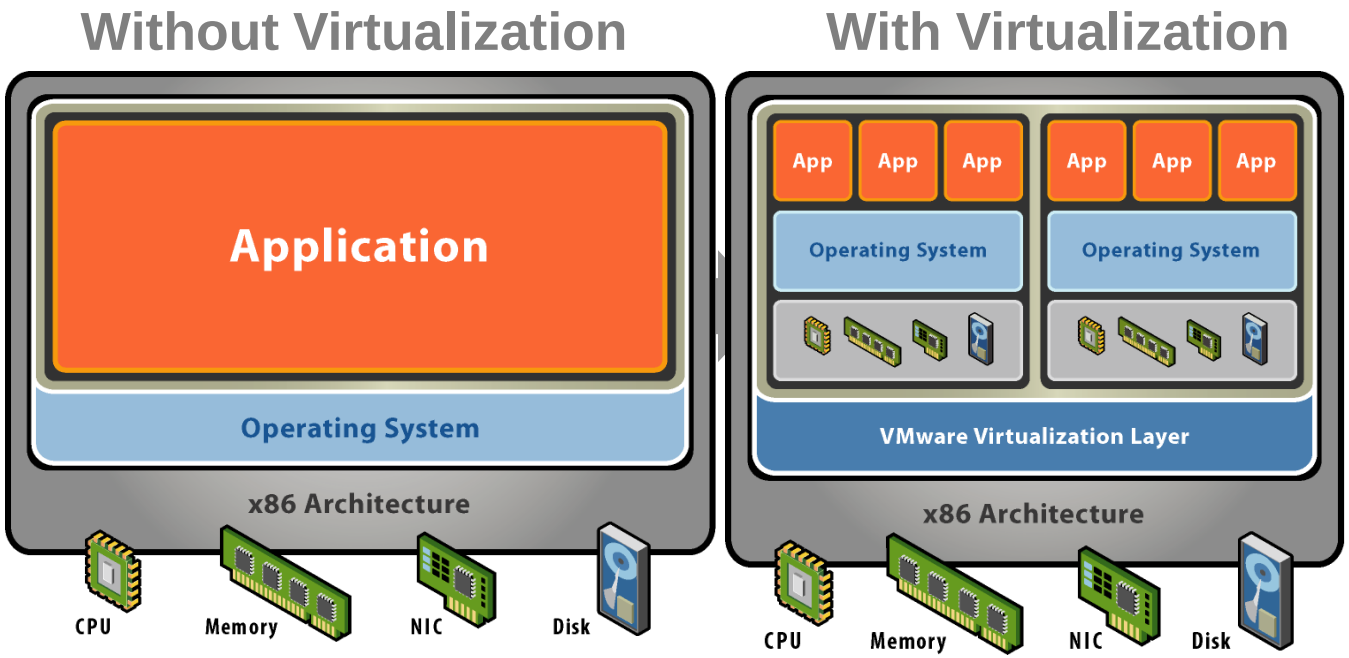
The Software Defined Datacenter

Josh Simons

Office of the CTO

VMware, Inc.

Server Virtualization



- Hardware virtualization presents a complete x86 platform to the virtual machine
- Allows multiple applications to run in isolation within virtual machines on the same physical machine
- Virtualization provides direct access to the hardware resources to give you much greater performance than software emulation

Virtual Platform Evolution

OPENFABRICS ALLIANCE

2001-2005

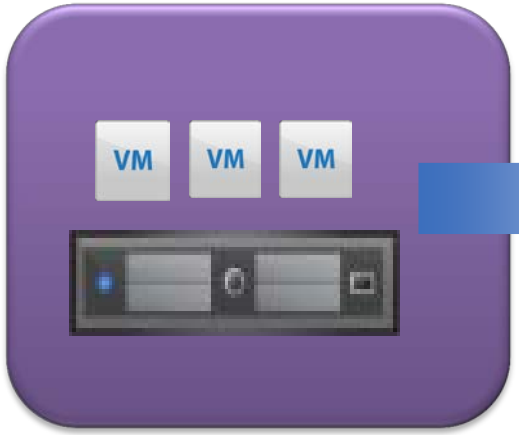
Server Virtualization

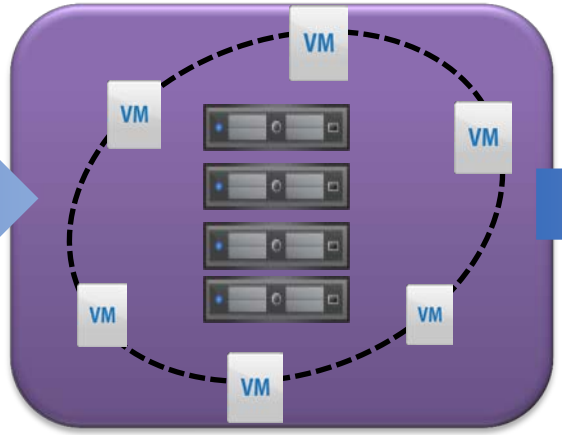
2006-2010

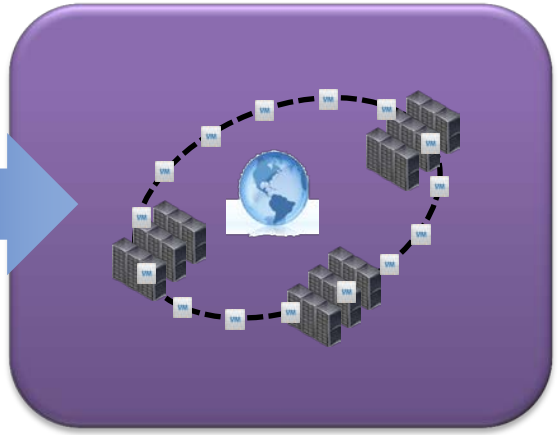
Cluster / Distributed Virtualization

2011+

Hybrid Cloud







VMs, Physical Storage, Network and Security

Pools of compute, Physical Storage, Network, and Security

Virtual DataCenters, Virtual Storage, Network, & Security

OPENFABRICS ALLIANCE

Past

$10,000
10 weeks

Present

$1,800
5 days, 2 minutes

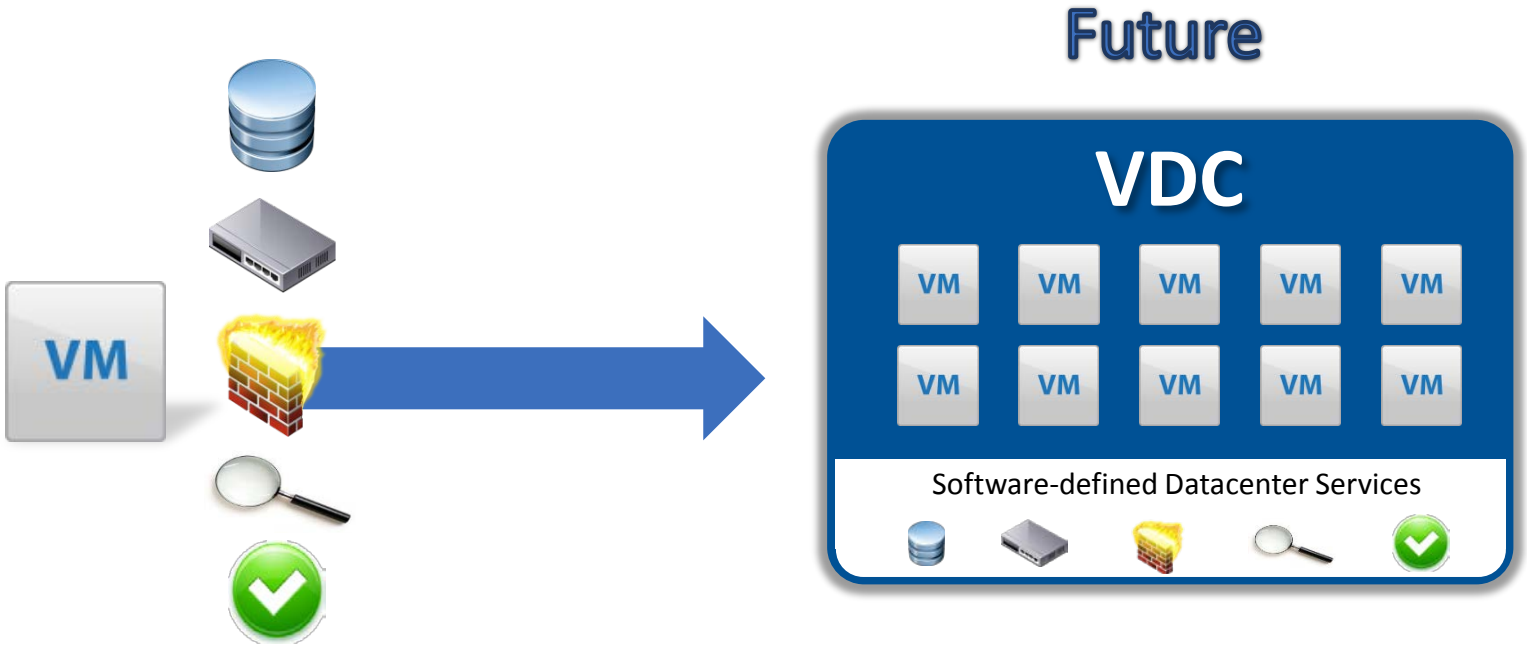
Enterprise storage

VLAN networks

Firewall, load-balancer

IDS, security, monitoring

Availability



5 days, 2 minutes

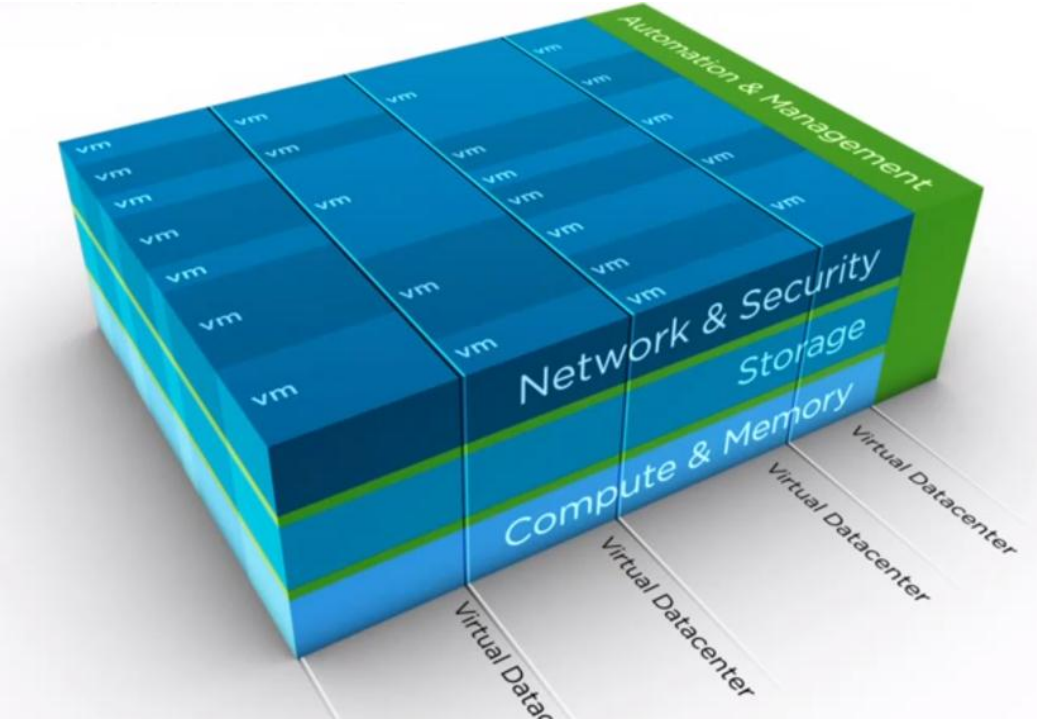
3 minutes

Software-defined Datacenter

Cloud

A way of offering computing services that prioritizes

- Self-service
- Elasticity
- Pay-by-use
- Agility



Software-defined Data Center

The *architecture for cloud* where:

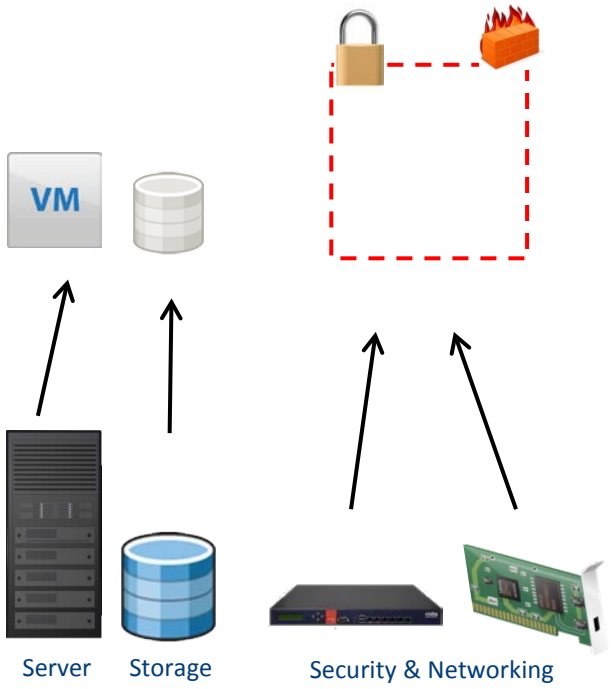
- All infrastructure is virtualized
- Delivered as a service
- Control of this datacenter is entirely automated by software

SDDC Mantra

OPENFABRICS ALLIANCE

Abstract

Decouple from physical



Pool

Scale infinitely, manage as one

Automate

Provision/manage per app's needs

Software Defined Networking

OPENFABRICS ALLIANCE

Telephony

650.555.1212

Identifier = Location

Wireless Telephony

650.555.1212

Networking

192.168.10.1

VM VM VM

Identifier = Location

EMULEX We network storage

ARISTA

VXLAN

BROCADE

CISCO

intel

BROADCOM

192.168.10.1

VM

Low Latency Requirements

Virtual Infrastructure RDMA

- Distributed services within the platform, e.g.
  - vMotion (live migration)
  - Inter-VM state mirroring for fault tolerance
  - Virtually shared, DAS-based storage fabric
- All would benefit from:
  - Decreased latency
  - Increased bandwidth
  - CPU offload

Guest OS RDMA

- RDMA access from within a virtual machine
- Scale-out middleware and applications increasingly important in the Enterprise
  - memcached, redis, Cassandra, mongoDB, …
  - GemFire Data Fabric, Oracle RAC, IBM pureScale, …
- Big Data an important emerging workload
  - Hadoop, Hive, Pig, etc.
- And, increasingly, HPC

SDN and RDMA

- SDN
  - Decouple logical network from physical hardware
    - Encapsulate Ethernet in IP → more layers
  - Flexibility and agility are primary goals
- RDMA
  - Directly access physical hardware
    - Map hardware directly into userspace → fewer layers
  - Performance is primary goal
- Is there any hope combining the two?
  - Converged datacenter supporting both SDN management and decoupling along with RDMA

Glimmers of Hope

- While 50+% of datacenter workloads are now virtualized, many customers have a mix of virtual and physical hardware
  - while percentage will continue to increase, there are some workloads that will likely remain un-virtualized
- The need for low-latency, high-bandwidth interconnect in the enterprise is a clear trend
  - e.g., scale-out DBMS, Big Data, etc.

►SDDC (and SDN) must accommodate these realities in the future datacenter

More (Philosophical) Glimmers

- SDN splits the control and data planes and uses a centralized controller to program switch fabric
  - Central fabric management familiar to RDMA community
- SDN exploring use of more physical telemetry to offer better application performance (while still maintaining network abstraction)
  - Metrics, topology sensing
- SDDC wants to treat datacenters as pools of interchangeable resources – all traffic becomes East/West – driving datacenters towards high bisection and low, uniform latency topologies

RoCE Thought Experiment

- Can RoCE be used as the basis of an SDN environment that also supports RDMA?
- RoCE satisfies both of SDN's basic requirements
  - IP connectivity to all hosts – needed by management controller
  - IP for Ethernet encapsulation
- Hypervisor-Hypervisor traffic
  - Support both IP and RDMA traffic
  - Use of RDMA to accelerate virtual-platform services, e.g. live migration

RoCE Thought Experiment (2)

- VM-VM traffic
  - IP connectivity available
  - RDMA via passthrough, or
  - Paravirtualized RDMA (for endpoint mobility)
- Simultaneous sharing of interconnect for:

  HV-HV / VM-VM / IP / RDMA
- Visibility of RDMA traffic to SDN framework?
- RoCE approach seems to support what is necessary for SDN management and data paths while also allowing hosts and VMs to access RDMA, as required
- IPoIB? Perhaps a similar answer…

Summary

- Software Defined Datacenter is a critical component for delivering the full value of cloud computing
- SDN is the means by which networking will be decoupled from underlying hardware
- RDMA is clearly important for the future Enterprise datacenter
- We would value discussions with experts to craft a solution that supports RDMA within a larger SDDC/SDN context

Thank You