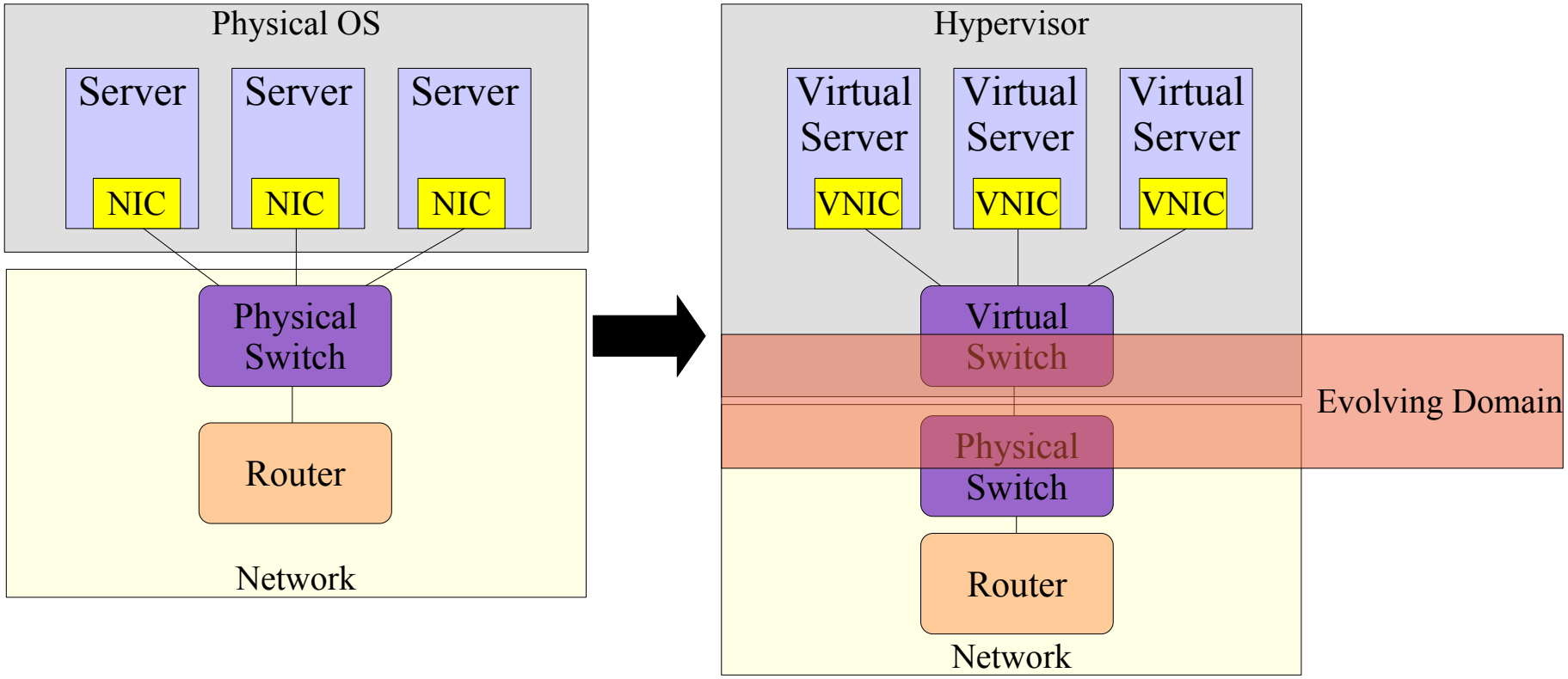
Virtual Networks: Host Perspective

IETF-77 Anaheim, CA
Virtual Network Research Group
March 23rd, 2010

Sunay Tripathi
Sunay.Tripathi@Oracle.Com

Evolving Virtualization Landscape



- New Challenges
  - Defining the Virtual Network and its scope
  - Identifying the Virtual Machines on the network and policy enforcement
  - Added complexity with new layers

Components of a Virtual Network

- Assigning MAC address to virtual NICs (VNIC) and VSwitches
  - A randomly created MAC address is preferred to aid virtual machine migration
  - L2 networks are becoming bigger in data centers (1000s of hosts) and the hosts and becoming more powerful capable of hosting 100s of VM so MAC addresses can collide fairly often
  - For a Vswitch to be managed and be a identifyiable entity, it needs to have a MAC address too
- Identifying Virtual Machine on the network
  - Need mechanisms to find current physical location
- Policies associated with VNICs and VM migration
  - MAC address, B/W limits, ACLs, host resources (CPUs, MIBs, stats, etc) need to be transferred to destination hypervisor during Virtual Machine migration
  - A centralized Port Profile Manager is not preferable since it creates another point of failure

Scope of a Virtual Network

- Naming and Identifying a Virtual Network
- Do they span just a layer 2 network or they span multiple IP networks that can be geographically separated
  - Perhaps we can classify them into two three types including a simple network that just spans a L2
- Migrating, snapshotting a Virtual Network

Security in a Virtualized Network

- L2 network open to new attacks
  - With SR-IOV virtualized NICs, a VM has ability to send bridge PDU, OSPF packets, etc and attack the L2 networks in new ways
    - Some OSes along with NICs can protect themselves but others can't
    - Some switch can deal with per VM security while others can't
    - Who does the protection can be a business decision so both modes need to be supported
- Performance and Security
  - Doing security checks twice doesn't improve performance
  - Clear protocols needed to negotiate who is doing the enforcement so we don't end up doing it twice (EVB group has some drafts)
  - At the same time, need to guarantee that it has been done atleast once
- Challenges
  - The environment gets very dynamic in terms of Number of Virtual Machines and migration
  - The policy enforcement, negotiation needs to scale in this environment

Isolation and B/W sharing

- For some people, Virtual Network means VLANs
  - VLANs do provide functional separation of the broadcast domains but have no resources attached to them
  - The hypervisors have QoS mechanisms that can be set on per VNIC basis
  - Some switches can do QoS per VM basis
- Challenges
  - Too many VMs and VNICs make up a Virtual Network
  - Configuring them individually is too challenging and error prone
  - Need a way to tie a group of VMs to a VLAN or extended VLAN and a mechanism for hypervisor and switch to negotiate B/W sharing mode

Diagnostics and Observability

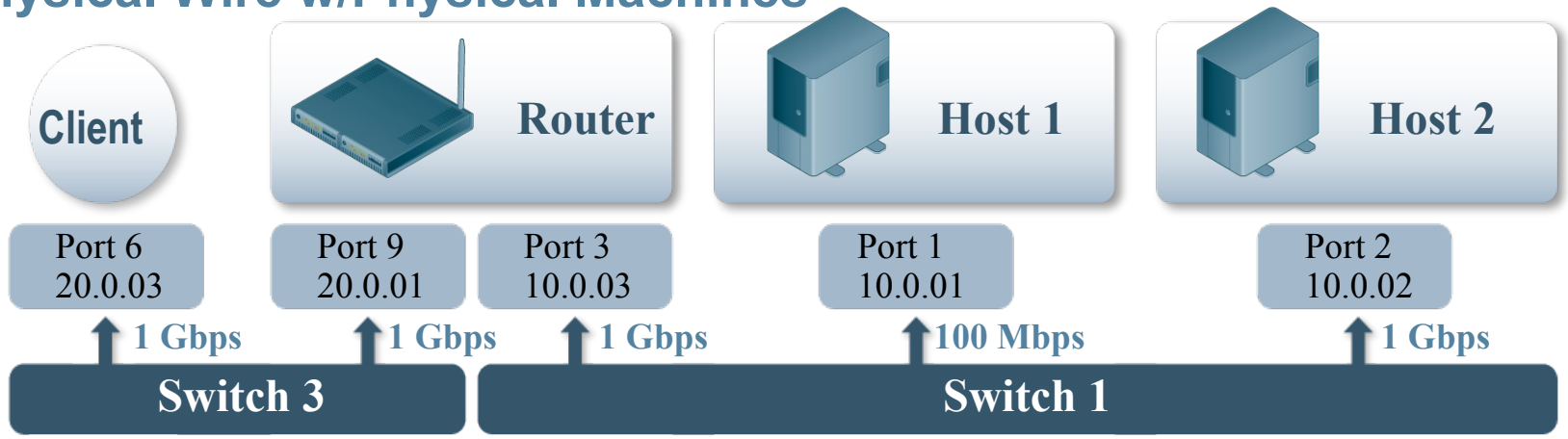
- Statistics in a Virtualized Network
  - Need per VNIC statistics – some OSes (like OpenSolaris) do it while others don't
  - Need per Virtual Network aggregated statistics

OpenSolaris Network Virtualization

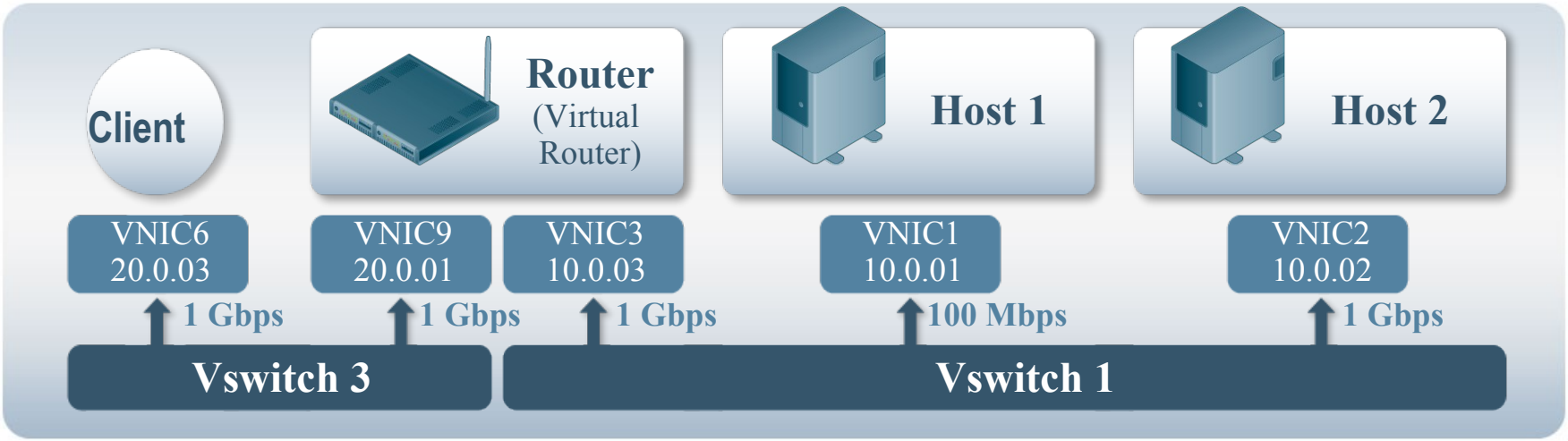
- Implemented via project Crossbow
  - Supports VNICs and Vswitches with NIC H/W assist
  - Per VNIC MIBs
  - Supports configurable link speeds (QoS) between VMs
  - Works with link aggregation and IPMP
  - VNICs can have VLAN tags assigned to them
  - VNICs have dedicated NIC, CPU, kernel threads and queues and are fully isolated from each other within the system

Crossbow: Virtual Network in a Box

Physical Wire w/Physical Machines



Virtual Wire w/Virtual Machines



More details

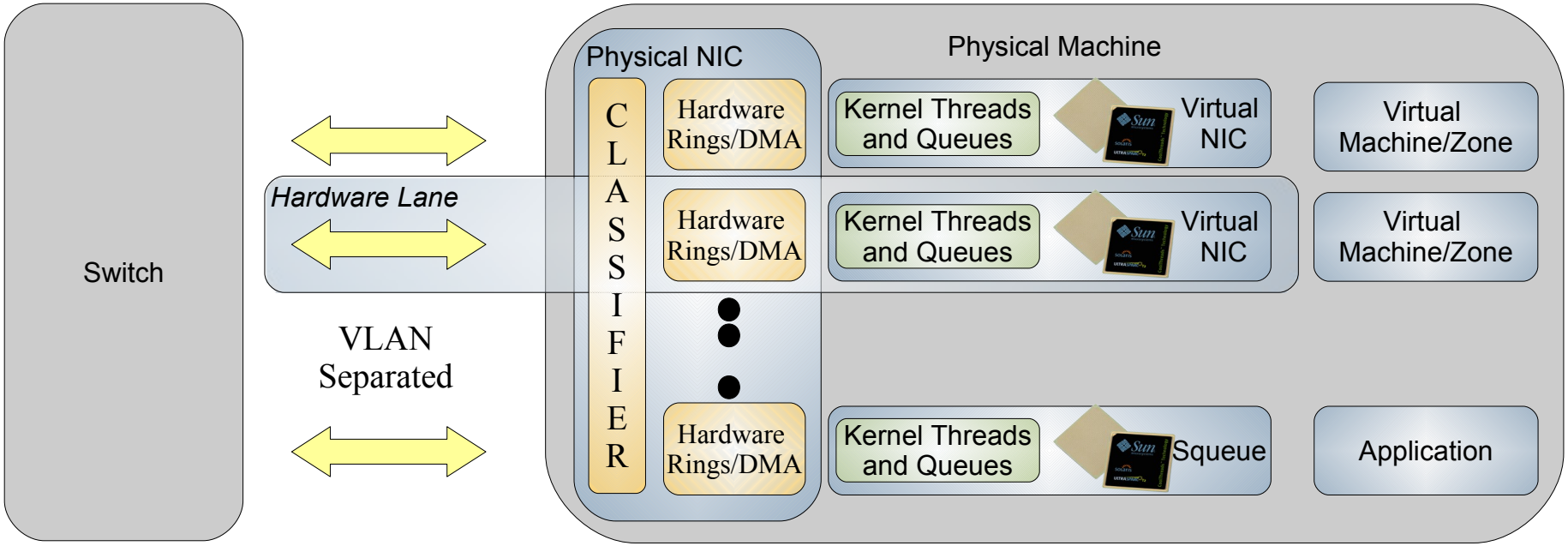
- Related Links
  - CrossBow: http://opensolaris.org/os/project/crossbow
  - VNM: http://opensolaris.org/os/project/vnm
  - Networking: http://opensolaris.org/os/community/networking
- Research Papers
  - Sigcomm VISA 2009 - “Crossbow: From H/W Virtualized NICs to Virtualized Networks”
  - Sigcomm WREN 2009 - “Crossbow: A vertically integrated QoS stack”
  - Usenix LISA 2009 - “Crossbow Virtual Wire: Network in a Box”
- *All the papers can be accessed via http://blogs.sun.com/sunay*

BACKUP

Crossbow '*Hardware Lanes*'

Ground Up Design for multi-core and multi-10GigE

> Linear Scalability by using '*Hardware Lanes*' with dedicated resources
> Network Virtualization and QoS designed in the stack
> More Efficiency due to '*Dynamic Polling and Packet Chaining*'

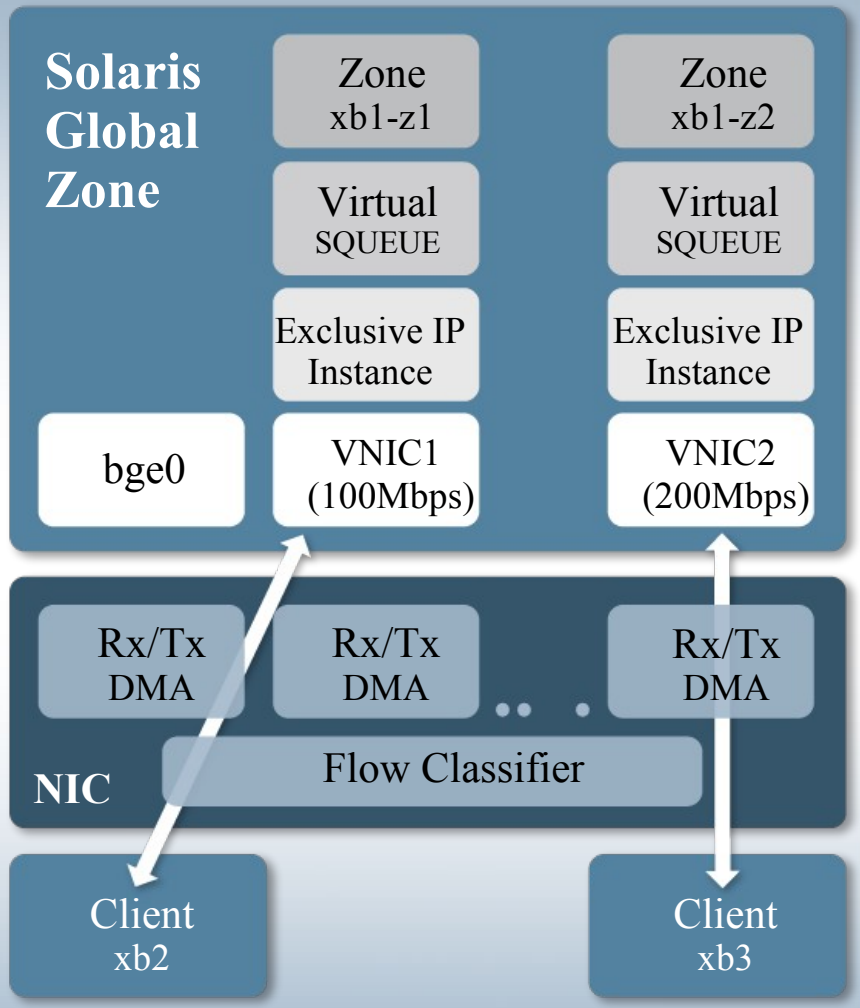


Virtual Network Containers

Virtualization
- Flows
- Virtual NICs & Virtual Switches
- Virtual Wire

Resource Control
- Bandwidth Partitioning
- NIC H/W Partitioning
- CPUs/pri assignment

Observability
- Real time usage for each Link/flow
- Finer grained stats per Link/flow
- History at no cost



Virtual NIC (VNIC) & Virtual Switches

Virtual NICs

- Functionally physical NICs:
  - IP address assigned statically or via DHCP and snooped individually
  - Appear in MIB as separate '*if*' with configured link speed shown as '*ifspeed*'
  - VNICs can be created over Link Aggregation on can be assigned to IPMP groups for load balancing and failover support
- VNICs Can have multiple hardware lanes assigned to them
- Can be created over physical NIC (without needing a Vswitch) to provide external connectivity with switching done in NIC H/W
- VNICs have configurable link speed, CPU and priority assignment
- Standards based End to End Network Virtualization
  - VLAN tags and Priority Flow Control (PFC) assigned to VNIC extend Hardware Lanes to Switch
- No configuration changes needed on switch to support virtualization

Virtual Switches

- Can be created to provide private connectivity between Virtual Machines

Virtual Machines