Thoughts on Network Interface Virtualization

Joe Pelissier

Andrew Lambeth

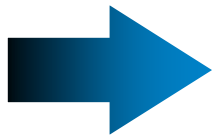
new-dcb-pelissier-NIC-Virtualization-0908

Current state of evolution

Standalone Servers and Individual Adapters

Servers connect to bridges with an OS interface per port

Network policies applied to bridge ports which are equivalent to OS interfaces



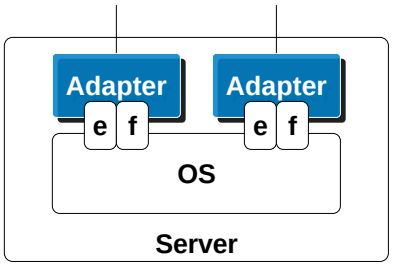
Virtualized Servers and Consolidated IO

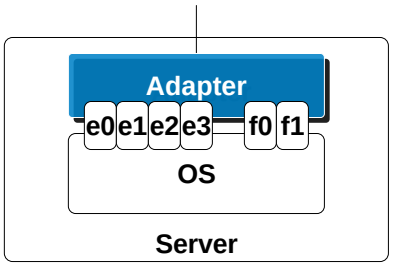
Virtual machines connect to bridges in hypervisors

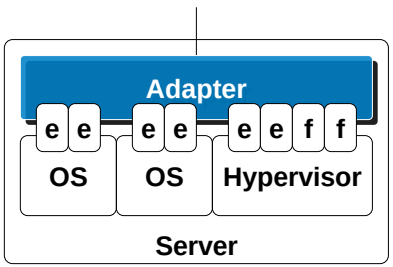
SR-IOV adapters with integrated bridges

Multiple interfaces from multiple OSes per bridge port

Use Cases for Virtualized Adapters







Functionally consolidated IO devices

- **Ethernet & Fibre Channel**

Multiple interfaces for single OS servers

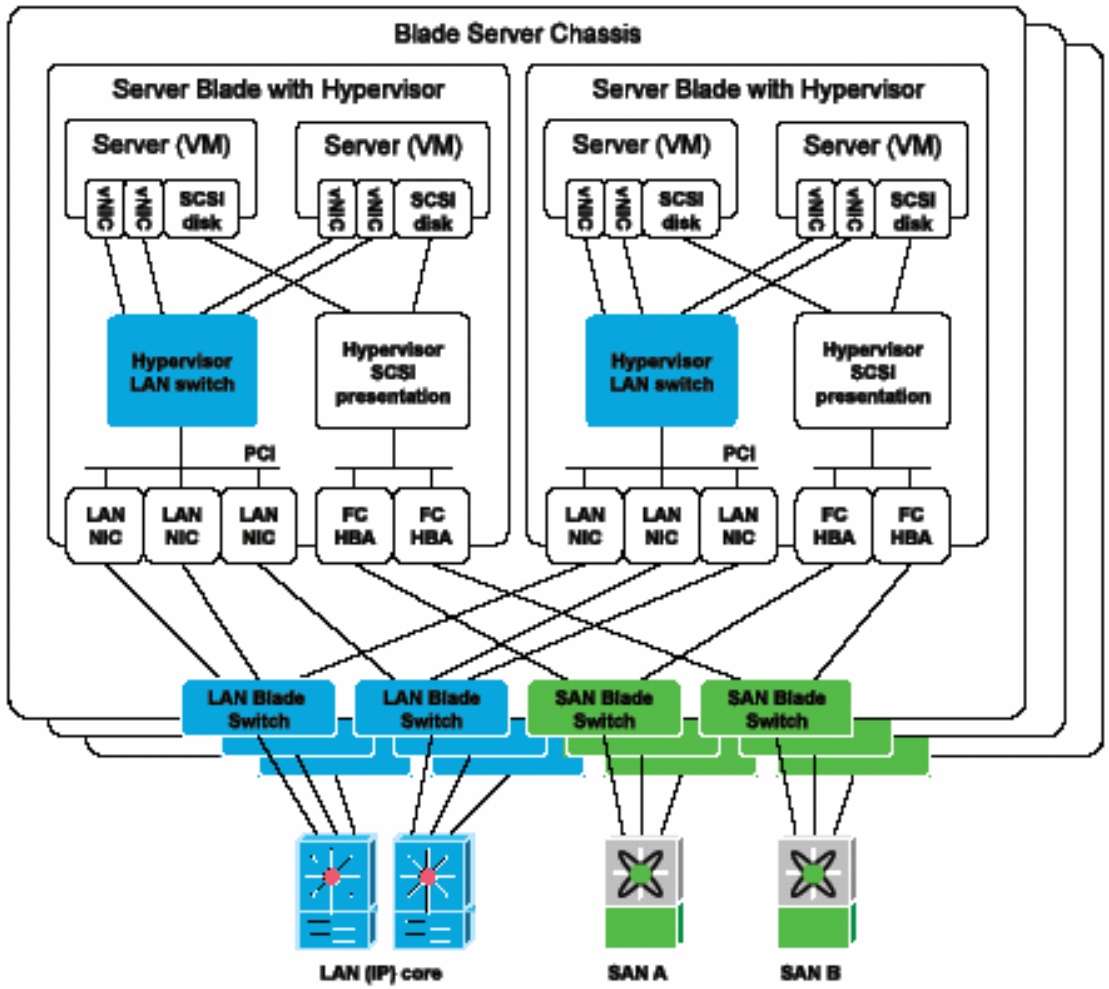
- **Multiple LANs…**
  - e0 – system management VLAN
  - e1 – DMZ VLAN
  - e2 – oracle traffic
  - e3 – all other traffic
- **… and SANs**
  - SAN A and SAN B

Interfaces to virtualized servers

- **Reduces overhead of data movement**

Virtualization Results in Complex Bridge Hierarchy

The Hypervisor LAN Switches (soft bridges) are taking on increasingly complex functionality
-ACLs
-VLANs
-Security
-Congestion Notification
-Priority Flow Control
-Enhanced Transmission Selection
-Etc.



Evolving Issues

- **When VM migration is common, uniform policy enforcement is important:**
  - **Across VMs to be same as across physical servers**
  - **Embedded softbridges complicate policy enforcement and reduce scalability**
- **Enhancing Scalability in a virtualized environment**
  - **Especially when policy enforcement and DCB technologies are involved**

A Possible Solution…

Delegate complex and performance critical data path functions into upstream networking devices (bridges)

- Ensures feature consistency to all traffic
- Fewer bridges – simpler and more consistent management
- Better performance and scalability

NICs provide value add data movement and ULP features

- TCP offload, RDMA, FC/SCSI DDP, IPC queue pairs, etc

Hypervisors provide features based on visibility of host state

Requirements

- **Develop an on the wire protocol for a bridge to indicate the virtual network interface (and potentially VLAN) to which a frame is destined**
  - **Conversely, to indicate from which virtual network interface (and potentially VLAN) a frame was received**
  - **Must operate in conjunction with other DCB technologies**
- **Develop a protocol enabling an attached device (typically a hypervisor) to establish and modify MAC Address (and VLAN) to virtual network interface relationships**
  - **Movement of VMs to be efficient and transparent to the rest of the network**
    - Implies capability to move the physical location of a virtual port at any time
- **Define appropriate associated management objects**
- **Do not change core bridge functionality**
  - **Physical bridge ports are expanded to support multiple virtual ports**
  - **Bridge functionality for the virtual ports remains unchanged**

Next Steps

- **Start building consensus around a PAR and 5C**
- **Build consensus working relationship between Internetworking and DCB**
  - **Currently no changes to basic bridge functionality is anticipated**
  - **Focused primarily more with server / bridge interaction**
  - **DCB seems to be the right group for this activity**
    - With close involvement on Internetworking and all other interested parties
- **A first cut at possible PAR text follows…**

Purpose

- **With the adoption of virtualization in the datacenter, the access layer has moved off of the conventional bridging hardware which has traditionally served this purpose and onto hosts in the form of software switches implemented in the hypervisor. Because of mobility of VMs and their increasing concentration on hosts due to multicore architectures, the access layer must continue to be tightly integrated with the hypervisor while maintaining enterprise class bridging capabilities. One method of increasing the scalability and performance of such environments is to migrate these functions from the hosts back into the conventional bridging hardware while maintaining the tight integration with the hypervisor. This project proposes to develop the functions and protocols necessary to achieve such a migration while maintaining the required integration.**

Scope

- This project would specify protocols, procedures and managed objects that support of virtualized network interfaces on IEEE 802 links. This specifically includes defining on the wire indication of virtualized source and target interfaces as well as protocols to establish and modify the relationship between these interfaces, their MAC addresses, and VLANs. These protocols, procedures, and managed objects will operate independently of and in conjunction with the other DCB technologies currently under development in IEEE. The intent is to expand existing bridge functionality (including functionality currently under development within IEEE) to operate with virtualized network interfaces; however, changes to these basic functionalities is not anticipated.

Some notes on Scope

- **The scope text on the previous slide refers specifically to the IEEE 802.1 activities**
- **There are other important components of this activity outside the scope of 802.1, for example:**

  Interaction between NIC and hypervisor

  Interactions between hypervisor and guest operating systems

Need for Project

- **Current IEEE 802.1 bridge standards do not provide the mechanisms necessary to enable the required integration between a physical bridge and virtualized servers. As a result, bridge functionality has migrated from physical bridges to software implementations within these servers. It continues to be desirable to deploy enterprise class bridging capabilities in a consistent manner whether the endpoints are virtual or physical, rather than separately for each type of endpoint. Consequently, enhancements to current IEEE 802.1 bridge standards are necessary to enable the tight integration necessary to allow the bridging functions to migrate back to the physical bridges.**

Thank You!