Motivation & Approaches for Edge Virtual Bridging or NIC Aggregation

Chuck Hudson (HP)   **c.hudson@hp.com**

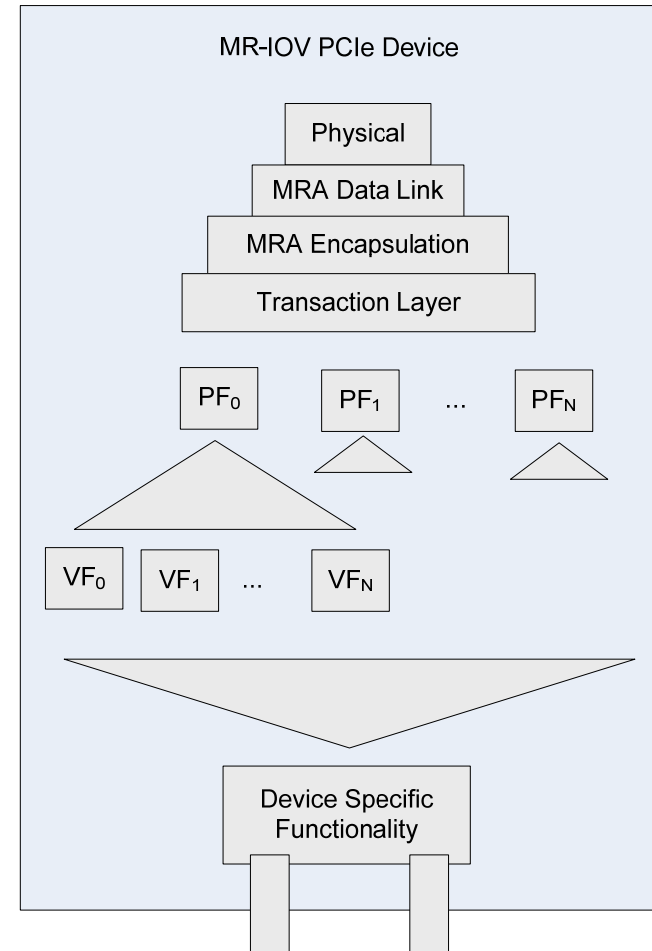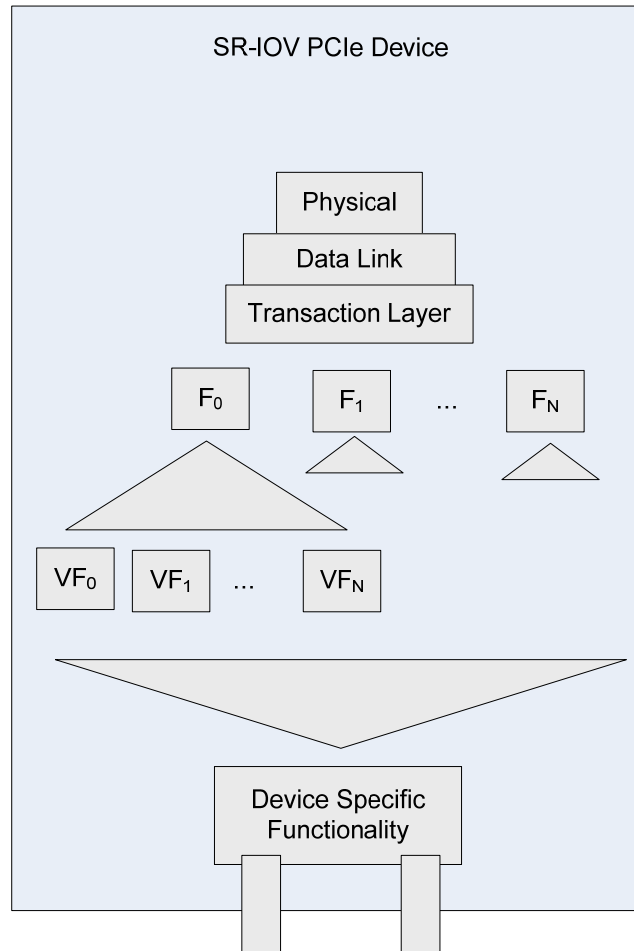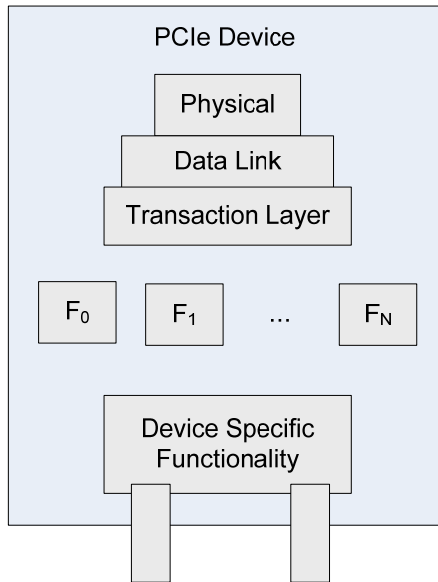**Paul Congdon (HP)**   **ptcongdon@ucdavis.edu**

Joe Pelissier (Cisco)   **jopeliss@cisco.com**

Possible Motivations

1. Need simple aggregation capability to be integrated into new NICs

    a. PCI Virtual Functions (VFs) bypass vSwitches

    b. Shared PCI IO virtualization (IOV) Ethernet devices

2. Allow bridge control of vNIC network connectivity

3. Create stand-alone edge devices that surrender bridging features to a higher level bridge device

PCI Device Evolution

PCIe Device

Physical

Data Link

Transaction Layer

F_0 F_1 ... F_N

Device Specific Functionality

SR-IOV PCIe Device

Physical

Data Link

Transaction Layer

F_0 F_1 ... F_N

VF_0 VF_1 ... VF_N

Device Specific Functionality

MR-IOV PCIe Device

Physical

MRA Data Link

MRA Encapsulation

Transaction Layer

PF_0 PF_1 ... PF_N

VF_0 VF_1 ... VF_N

Device Specific Functionality

Motivation 1-a
PCI Virtual Functions (VFs) Bypass vSwitches

- Each VM has some number of vNICs

- PCI VFs provide direct vNIC access to hardware queues and improve IO performance

- Direct access bypasses the software vswitches

- Something needs to replace the software vswitches to control Ethernet frame flow and packet replication.

- Full embedded bridges are problematic

  – Expensive (gates & management processor complexity)

  – Puts NICs on 802.1 feature tread mill

- Need a solution that allows for simple NIC hardware

Motivation 1-b
Shared PCI IO Virtualization Devices

- PCI has technologies that allow device sharing

  - A single physical device can contain the NICs (really vNICs) for multiple physical servers.

- Something must be done to converge these (v)NICs onto a single link to the network infrastructure.

- Full embedded bridges are problematic

  - Expensive (gates & mgmt processor cycless)

  - Puts NICs on 802.1 tread mill

- Need a solution that allows for simple NIC hardware

Motivation 2
## Allow bridge control of vNIC network connectivity

- Virtual Machine environments must provide network access to the individual guest operating systems

- VM environments are typically owned by "server admins"

- Allowing consistent control of network access for single-OS physical servers and individual virtual machines simplifies data center management.

- Need a solution that allows edge bridges to control virtual machine network connections in the station.

Motivation 3
Provide increased fan-in / fan-out for core switches

- Extend core/distribution switch to edge
  - Control and functionality remain in core/distribution devices
  - Provides full capability scaling in a more cost effective manner
  - Reduced number of managed nodes / Cost of ownership
  - Simplified and more flexible VM migration

Summary of Possible Technical Approaches
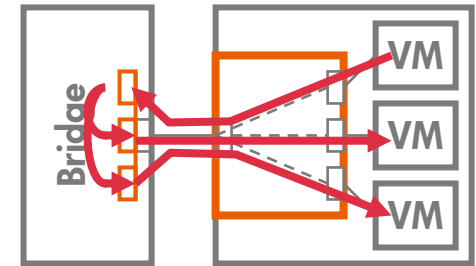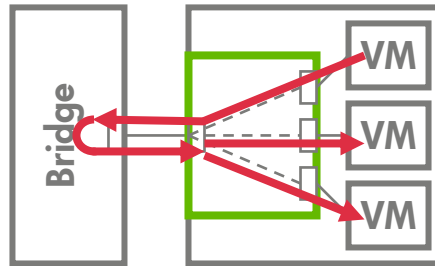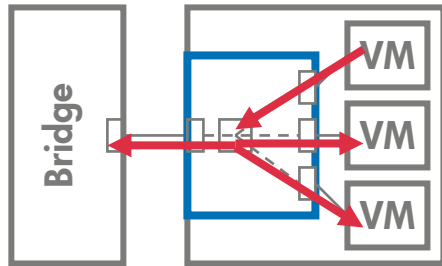
Virtual Ethernet Bridge (VEB)
uses MAC+VID to steer frames

Tag-less VEPA
uses MAC+VID to steer frames

Tagged
uses <u>new</u> tag to steer frames

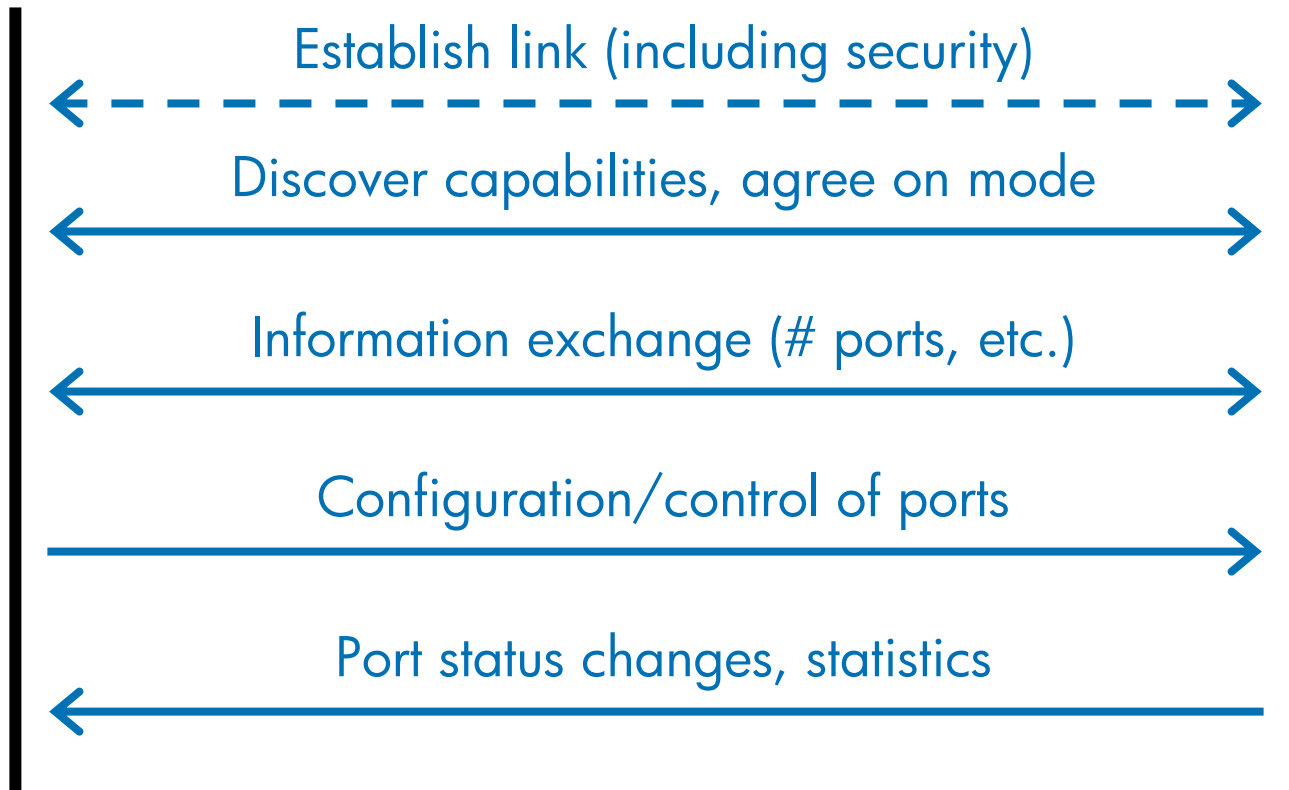| Virtual Ethernet Bridge (VEB)	Tag-less VEPA	Tagged
• Emulates 802.1 Bridge	• Extends 802.1 Bridge	• Extends 802.1 Bridge
• Limited controls	• Advanced controls	• Advanced controls
• Managed by station	• Managed by bridge	• Managed by bridge
• Works with all existing bridges	• Works with many existing bridges	• Works with few or no existing bridges
• No changes to existing frame format.	• No changes to existing frame format.	• Changes to existing frame format.
• Open-ended changes to NIC	• Limits NIC changes	• Limits NIC changes

Elements of a solution

- Frame format and flow
  - Tagged vs. tag-less approach
  - If tagged, the tagging approach
- Requirements for controlling bridge
- Requirements for virtual bridging device
- Discovery/control protocol

Discovery/Control

Controlling Bridge

Virtual Bridging Device

Establish link (including security)

Discover capabilities, agree on mode

Information exchange (# ports, etc.)

Configuration/control of ports

Port status changes, statistics

new-dcb-hudson-niv-motivation-0109.ppt

Call to Action

- Obtain consensus on the need for a virtual bridging effort

- Invite interested parties to participate in the generation of a PAR and 5 criteria proposals