Tag-less Virtual Ethernet Port Aggregator (VEPA) Proposal

January 2009
Chuck Hudson (HP)  c.hudson@hp.com
Paul Congdon (HP)  ptcongdon@ucdavis.edu
Motivation

• Enable robust bridge features to individual virtual machines
  − Network controls / ACLs
  − Network monitoring & security
  − Private VLANs

• Coordinated management of the network edge
  − Physical servers
  − Virtual servers

• Simplify data center management

• Rapid industry adoption
# Summary of Possible Technical Approaches

## Virtual Ethernet Bridge (VEB)
- Emulates 802.1 Bridge
- Limited controls
- Managed by station
- Works with all existing bridges
- No changes to existing frame format.
- Open-ended changes to NIC

## Tag-less VEPA
- Extends 802.1 Bridge
- Advanced controls
- Managed by bridge
- Works with many existing bridges
- No changes to existing frame format.

## Tagged
- Extends 802.1 Bridge
- Advanced controls
- Managed by bridge
- Works with few or no existing bridges
- Changes to existing frame format.
- Limits NIC changes

---

**Virtual Ethernet Bridge (VEB)** uses MAC+VID to steer frames.

**Tag-less VEPA** uses MAC+VID to steer frames.

**Tagged** uses new tag to steer frames.
Tag-less 101
Virtual Ethernet Port Aggregator (tagless)

1. Provides Multiple VEPA Device Ports (vdp) as vNICS to Virtual Machines
2. Each VDP is configured as individual NIC (i.e. MAC addr, Multicast addr, VLAN tags, or passthru). VEPA aggregates configurations.
3. May support all traditional NIC features (e.g. TCP Checksum, RSS, Large Segment Send)
4. Does NOT perform Local Bridging. Not a Virtual Ethernet Bridge (VEB)
5. Sends all outbound traffic to the wire
6. Replicates received mcast/bcast traffic
7. VLAN aware
8. May provide QoS and BW management
9. Invoked by special Bridge mode negotiation

Note: This proposal does NOT require new tags, but could work with them.
VEPA Forwarding

1. A→C
2. A→B
3. A→Bcast
4. C→Bcast

Bridge Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach
VEPA Forwarding

1. A->C
2. A->B
3. A->Bcast
4. C->Bcast

* = Bridge Port Configured for VEPA attach

<table>
<thead>
<tr>
<th>Address</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
</tbody>
</table>
VEPA Forwarding

1. A->C
2. A->B
3. A->Bcast
4. C->Bcast

**Bridge Address Table**

<table>
<thead>
<tr>
<th>Address</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach
VEPA Forwarding

1. A->C
2. A->B
3. A->Bcast
4. C->Bcast

Bridge Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>Port</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach
VEPA Elements
Key VEPA Terms

- **VEPA-Enabled Bridge Port**
- **End Node**
- **Base VEPA Device Ports**
- **VEPA Mgr**
- **VEPA Uplink Port**
- **Station**
- **Bridge**
Additional VEPA Terms

VEPA ‘Portlets’ (optional)

VEPA Uplink Port

VEPA-Enabled Bridge Port

VEPA Egress ‘Portlets’ for Broadcast, Multicast, & flooded frames

VEPA Mgr

Bridge

VEPA

Base VEPA Device Ports

End Node

End Node

End Node

End Node

VEPA Expander Port

VEPA Expander

Station

VEPA Mgr

VEPA Expander Uplink Port

Expander VEPA Device Ports
Basic Tag-less VEPA Construction

• Each VEPA has
  – A single, active VEPA Uplink Port
  – 1 to n VEPA Device Ports
  – 0, 1, or more VEPA Expander Ports
  – Station VEPA Manager & VEPA Address Table

• Connected to VEPA-enabled Bridge Port
  – VEPA ‘Portlets’ (optional)
  – Egress ‘Portlets’ (optional)
  – Bridge VEPA Manager

• A station may have multiple VEPAs
VEPA Device Ports

- Each VEPA Device Port
  - May be implemented as a PCI virtual function
  - Has one or more statically-identified MAC addresses
  - Movement of MAC addresses coordinated through VEPA Managers
- VEPA Device Ports are ‘NIC Configuration Aware’
  - Of MAC addresses
  - Of MAC listening entries (multi-cast and unicast)
- Configured via Station VEPA Manager
  - 1 or more specific MAC addresses (by station)
  - VLAN tagging behavior*
  - Priority tagging behavior*
- Forwards incoming frames to VEPA uplink
  - May set VLAN/Priority based on settings
- Receives frames from VEPA uplink
  - May remove VLAN/Priority tag based on settings

* Can be set by station or bridge
VEPA Uplink Port

- Single VEPA Uplink per VEPA
  - May be LAG
  - Has a MAC address (for capability exchange)
  - May implement ETS queues

- Settings
  - VEPA MAC address
  - Acceptable frame types
    - Only VLAN tagged
    - Untagged, Pri tagged
    - All frames
  - PVID
  - Egress VLAN IDs (aggregate of the VDP VIDs)
VEPA Expander

- Usually software (operating mode of vswitch)
- Extends beyond limits of HW VEPA
  - # of VEPA Device Ports
  - # of VEPA Address Table Entries
- Consists of
  - One VEPA Expander Uplink Port
  - One to m Expansion VEPA Device Ports
  - Expander VEPA Address Table
- Forwards frames from VDPs to VEPA Uplink
- Sends (replicating as necessary) from Expander Uplink Port to expansion VDPs
- Linked to Station VEPA Manager
  - Configuration of VEPA Device Ports
  - Contribute to VEPA Capability Exchange
VEPA-enabled Bridge Port…

- The port is enabled for ‘turn-around’ forwarding of
  - Multicast
  - Broadcast
  - Flooded Unicast
- Unicast destinations per forwarding table
- The bridge may implement controls and features via
  - VEPA Device Port configuration (VLAN ID, Private VLANs, Priority Settings, MAC filtering)
  - Portlets (ACLs, Statistics)
  - Address table entries (IGMP snooping)
VEPA ‘Portlets’

- **Optional** (can be simulated by rules engines)
- Useful in simplifying ACLs & statistics collection
- VEPA Portlets
  - Associated with VDP MAC address(s)
  - Identifies incoming frames by SRC MAC
  - Identifies outgoing unicast frames by DST MAC
- Egress Portlets
  - Extra controls & statistics on broadcast, multicast, and flooded frames
VEPA Address Table Management
Address Table Management

- Managed by Station VEPA Manager
  - Information coordinated with bridge via VEPA Capability Exchange
- Static settings (no learning)
- Driven by NICs
  - VM NIC driver register for unicast/multicast listens
    - Fully-supports Locally-Assigned MAC Addresses (LAA)
  - Station VEPA manager receives request
  - Station VEPA manager creates/updates table entries
- Multicast entries may be driven by Bridge (IGMP snooping)
  - Bridge intercepts join/leave messages
  - Creates/updates/deletes address table multicast entry
VEPA Address Table Example

**Example: Base VEPA Address Table**

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0001 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach

**Diagram:**

- **Station**: VM A, VM B, VM C, VM D
- **VEPA**: VM E, VM F
- **Bridge**: 1*, 2
- **Legend**:
  - Green: VLAN 1 Tag Mask = UUUUT
  - Blue: VLAN 2 Tag Mask = UUUUT

---

**Note:**

- VLAN 1 Tag Mask = UUUUT
- VLAN 2 Tag Mask = UUUUT
VEPA Address Table Example

A -> Bcast

1. Dst Lookup = 10101
2. Src Lookup = 10000
3. Delivery Mask = 00101

(dst & ~src)

Note: Bridge should echo IGMP packets too

* = Bridge Port Configured for VEPA attach

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0001 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

VLAN 1 Tag Mask = UUUUT
VLAN 2 Tag Mask = UUUUT
VEPA Address Table: Multicast Entries

A -> Mcast1

Station

VM A VM B VM C VM D

VM E

VM F

Exander

1. Dst Lookup = 10101
2. Src Lookup = 10000
3. Delivery Mask = 00101
(dst & ~src)

1 2

Bridge

* = Bridge Port Configured for VEPA attach

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0001 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

VLAN 1 Tag Mask = UUUUT

VLAN 2 Tag Mask = UUUUT
**VEPA Address Table: Unknown addresses**

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach

Unknown Multicast entries allow for multicast handling when there are excessive entries, promiscuous multicast listens, and steering of multicast entries to expander port.

Unknown unicast entries needed to steer packets to expander port(s). Also allows for support of promiscuous listen or monitoring ports.

**VLAN 1 Tag Mask = UUUUT**

**VLAN 2 Tag Mask = UUUUT**

---

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach
Multiple VLANs on VDP

* = Bridge Port Configured for VEPA attach

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask  (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>C</td>
<td>2</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0001 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

VLAN 1 Tag Mask = UUTUT
VLAN 2 Tag Mask = UUTUT
VDPs in Dual Listening Mode

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0001 0</td>
</tr>
<tr>
<td>H</td>
<td>1</td>
<td>1010 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0101 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

Used by MS Cluster Server that sends frames with a unicast address that is never used as a source.

Caused by VMs A & C registering H as a listening MAC address (if allowed by Station VEPA manager).

* = Bridge Port Configured for VEPA attach.

VLAN 1 Tag Mask = UUUUT
VLAN 2 Tag Mask = UUUUT
VDP in Monitor Mode

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1010 0</td>
</tr>
<tr>
<td>B</td>
<td>2</td>
<td>0110 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>2</td>
<td>0011 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Bcast</td>
<td>2</td>
<td>0111 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>2</td>
<td>0110 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>2</td>
<td>0111 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0010 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>2</td>
<td>0010 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0010 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>2</td>
<td>0010 1</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach

Set as monitor

Shows how a VDP could be configured to monitor any or all incoming frames.

VLAN 1 Tag Mask = UUTUT
VLAN 2 Tag Mask = UUTUT
VEPA Default Configuration (no VLAN tags)

Example: Base VEPA Address Table

<table>
<thead>
<tr>
<th>Address</th>
<th>VLAN</th>
<th>Mask (ABCD *)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>1000 0</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>0100 0</td>
</tr>
<tr>
<td>C</td>
<td>1</td>
<td>0010 0</td>
</tr>
<tr>
<td>D</td>
<td>1</td>
<td>0001 0</td>
</tr>
<tr>
<td>Bcast</td>
<td>1</td>
<td>1010 1</td>
</tr>
<tr>
<td>Mcast1</td>
<td>1</td>
<td>0100 1</td>
</tr>
<tr>
<td>Mcast2</td>
<td>1</td>
<td>0101 1</td>
</tr>
<tr>
<td>Unk Mcast</td>
<td>1</td>
<td>0000 1</td>
</tr>
<tr>
<td>Unk Ucast</td>
<td>1</td>
<td>0000 1</td>
</tr>
</tbody>
</table>

* = Bridge Port Configured for VEPA attach

Uplink configured as untagged
No VLAN or priority tagging

VLAN 1 Tag Mask = UUUUU
Configuration
VEPA Capability Exchange

• Between Station VEPA Manager and Bridge VEPA Manager
• Exchange VEPA capabilities, configuration
• Re-occurs as needed to keep bridge station up to date
  – Add, move, change of End Nodes
• Initial sequence
  – Establish link
  – Authenticate the link
    • Based on the VEPA Uplink’s MAC address
    • Should allow for: MAC Auth, 802.1x, MACSEC
  – Link Aggregation Control Protocol (LACP) as appropriate
  – VEPA Capability Exchange
VEPA Capability Exchange

- Station → Bridge
  - VEPA Capabilities
    - Mode: Request/require: Tag-less, VEPA Tagged
    - # of base device ports
    - # of VEPA table entries
    - Level(s) of control
  - VEPA General Settings
    - Bridge vs. Station Control of VLAN ID
    - Bridge vs. Station Control of priority
  - Device Ports (Port Listing)
    - Port Number
    - Port Type (Base, Expander)
    - MAC addresses (as assigned by Station)
    - Settings
      - Acceptable Frame Types
      - PVID
      - VLAN IDs**
      - Ingress VID Filtering
      - Priority Settings
  - Address Table Entries (Typically Multicast)
    - Address, VLAN ID, Receiver Ports/Mask
  - Updates

- Station ← Bridge
  - VEPA Capabilities
    - Mode: Request/require: Tag-less, Tagged
    - Total # of supported device ports
    - Total # of supported address entries
  - VEPA General Settings
    - Echo: Control of VLAN ID
    - Echo: Control of priority
  - Device Port Setting (Port Listing)
    - Port Number
    - Echo/control settings
      - Acceptable Frame Types
      - PVID
      - VLAN IDs**
      - Ingress VID Filtering
      - Priority Settings
  - Address Table Entries (Typically Multicast)
    - Address, VLAN ID, Receive Ports/Mask
    - ...
  - Updates

** Could be done with VLAN port membership vectors
## VEPA Device Port Settings

### Settings

**MAC addresses**

- Acceptable frame types
  - Only VLAN tagged
  - Untagged, Pri tagged
  - All frames

**PVID**

- (Egress) VLAN IDs

**Ingress VID Filtering**

- Priority Setting
  - Default value
  - Set to default

**Ingress MAC Filtering**

<table>
<thead>
<tr>
<th>No VLAN Tag</th>
<th>Force Priority</th>
<th>VM has 3 VIDs</th>
<th>VM has 1 VID</th>
<th>private VLAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>one+</td>
<td>one+</td>
<td>one+</td>
<td>one+</td>
<td>one+</td>
</tr>
</tbody>
</table>
Summary
## Approach Comparison

<table>
<thead>
<tr>
<th>Area</th>
<th>VEB</th>
<th>Tagless VEPA</th>
<th>Tagged</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Key Elements</strong></td>
<td>Station VEB SW VEB</td>
<td>Station VEPA SW VEPA Expander Optional Portlets</td>
<td>Station VEPA + tag processing SW VEPA Expander Requires Virtual Switch Ports</td>
</tr>
<tr>
<td><strong>Station-side Learning</strong></td>
<td>Static, NIC-driven address table</td>
<td>Static, NIC-driven address table (used on Ingress) Special treatment of promiscuous ports</td>
<td>No MAC address learning in VEPA (Uses static tag address table)</td>
</tr>
<tr>
<td><strong>Station-side Forwarding (in)</strong></td>
<td>Standard via use of static address table</td>
<td>Based on static address table</td>
<td>Based on static tag-to-port table</td>
</tr>
<tr>
<td><strong>Bridge-side Learning</strong></td>
<td>Standard</td>
<td>Standard</td>
<td>Standard + (must be aware of virtual ports)</td>
</tr>
<tr>
<td><strong>Bridge Ingress Forwarding</strong></td>
<td>Standard</td>
<td>Requires ‘turn-around’ mode</td>
<td>Requires ‘turn-around’ mode (tied to virtual bridge ports)</td>
</tr>
<tr>
<td><strong>Frame Replication</strong></td>
<td>Station-side replication</td>
<td>Station-side replication</td>
<td>Bridge-side replication (or station-side with extensions)</td>
</tr>
<tr>
<td><strong>QoS</strong></td>
<td>Set per VF (?) Single set of ETS queues</td>
<td>Set per VF Single set of ETS queues</td>
<td>Set per Virtual Switch Port Single set of ETS queues</td>
</tr>
<tr>
<td><strong>Statistics</strong></td>
<td>Limited (station-side collection)</td>
<td>Limited (station-side collection) ++</td>
<td>Limited (station-side collection) ++</td>
</tr>
<tr>
<td><strong>ACLs</strong></td>
<td>Limited</td>
<td>ACLs per ‘portlet’</td>
<td>ACLs per virtual switch port</td>
</tr>
<tr>
<td><strong># of VMs</strong></td>
<td>Nearly unlimited (via vswitch)</td>
<td>Nearly unlimited (via expander)</td>
<td>Determined by number of virtual bridge ports</td>
</tr>
<tr>
<td><strong>Bridge traffic monitoring</strong></td>
<td>Limited</td>
<td>Full</td>
<td>Full</td>
</tr>
<tr>
<td><strong>Private VLAN Support</strong></td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Summary of Possible Technical Approaches

**Virtual Ethernet Bridge (VEB)**
- Emulates 802.1 Bridge
- Limited controls
- Managed by station
- Works with all existing bridges
- No changes to existing frame format.
- Open-ended changes to NIC

uses MAC+VID to steer frames

**Tag-less VEPA**
- Extends 802.1 Bridge
- Advanced controls
- Managed by bridge
- Works with many existing bridges
- No changes to existing frame format.
- Limits NIC changes

uses MAC+VID to steer frames

**Tagged**
- Extends 802.1 Bridge
- Advanced controls
- Managed by bridge
- Works with few or no existing bridges
- Changes to existing frame format.
- Limits NIC changes

uses new tag to steer frames