



# OpenVMS V8.3-1 H1 and Beyond

**Andy Goldstein**  
**OpenVMS Engineering, HP**




V8.3-1H1

# V8.3-1H1 Overview

- Hardware update release
- HP Integrity only
- Hardware support
- System management features
- Performance improvements, bug fixes, and other loose ends

# System Support

- Blades:
  - Manageability: WBEM & HPSIM Providers
  - Virtual Connect for Ethernet and Fibre Channel
  - Storage (BL40c) Blades support
  - C3000 Enclosure support (shorter enclosure)
  - 4 Port Gigabit Ethernet Mezzanine Card
- Integrity systems with Montvale 

## Montvale Upgrades

BL860c

rx2660

rx3600

rx6600

rx7640

rx8640

Superdome/sx2000

# Storage Adaptors

- HP SmartArray P800
  - Backplane RAID controller
  - 16 internal/external SAS ports
  - 2.5GB/s PCI-e 8x
- MSA-60 RAID array controller
  - 2U rack mount
  - 12 3.5" SAS/SATA drives
- MSA-70 RAID array controller
  - 2U rack mount
  - 25 SFF SAS/SATA drives

# More Storage Blades

- Ultrium 448c LTO-2 tape blade
- Support post-V8.3-1H1 pending qual...
  - Ultrium 920c LTO-3 tape blade
  - PCI-X/PCI-e sidecar blade

# Network / SAN Adaptors

- 2 port GB ethernet – PCI-e
  - AD337A copper
  - AD338A fiber
- 4 port GB ethernet – PCI-e mezzanine
  - 447883-B21
- 10GB ethernet – PCI-X
  - AD385A
- 4GB fibrechannel – PCI-e
  - AD299A 1 port
  - AD335A 2 port

# Network / SAN Adaptors

- Post-V8.3-1H1, pending qual
- 4 port GB ethernet – PCI-e
  - AD339A copper
- 4GB fibrechannel / GB ethernet – PCI-e
  - AD221A copper
  - AD222A fiber
- Virtual connect for 4 port GBe mezzanine (447883-B21)



# No iolock8 Fibre Channel Drivers



- Fibre Channel port drivers will now use port specific locks instead of iolock8 for synchronization
  - Frees up iolock8 for other uses
  - Allows more parallelism for FC drivers
  - FC IO now scales with more processors
  - Implemented for both Qlogic and Emulex families of drivers
  - Implemented for both Alpha and Integrity
    - Alpha release TBD

# iSCSI Technology Demonstration Kit



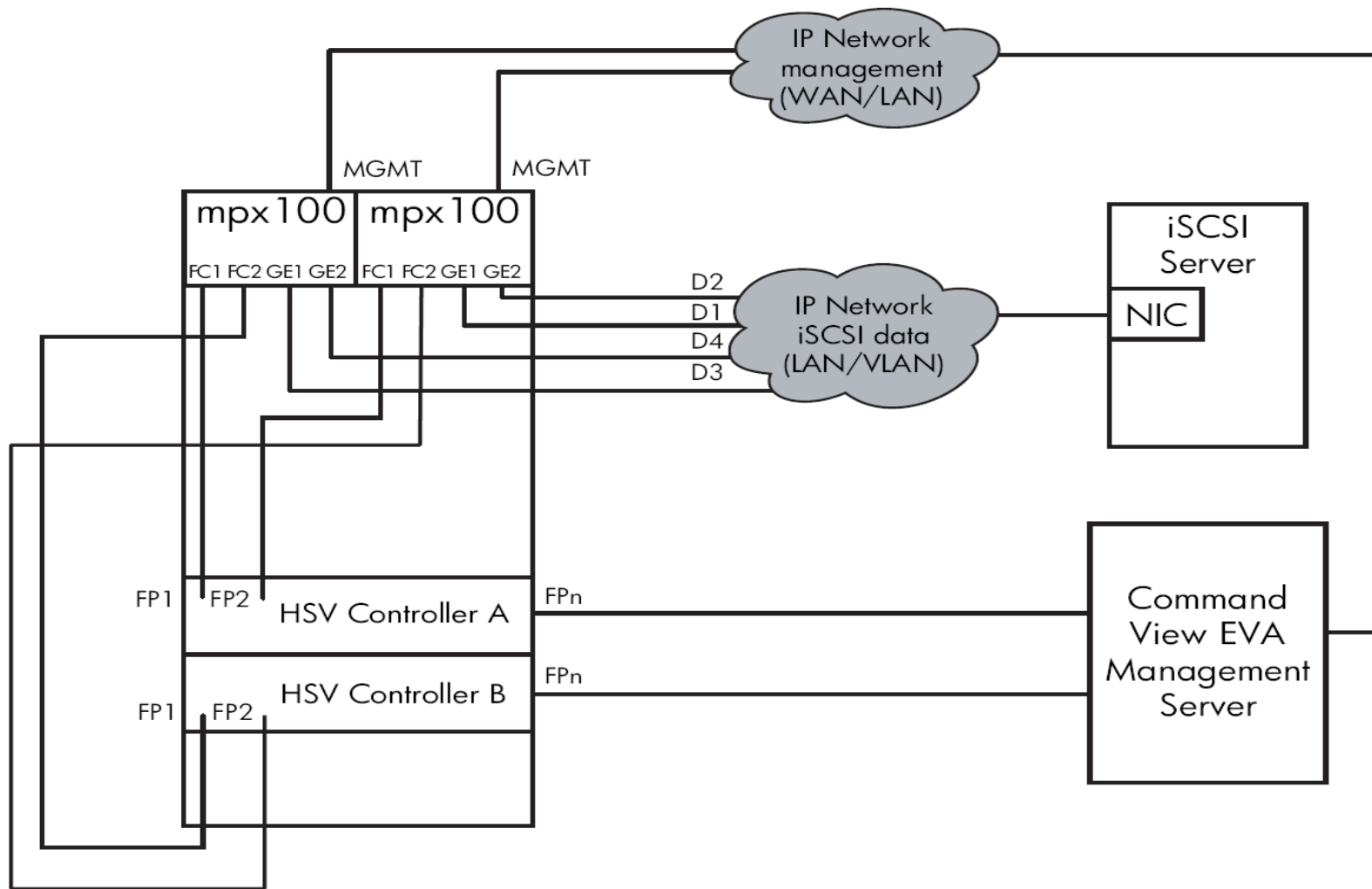
- iSCSI is a means to tunnel storage traffic through TCP/IP based LANs
  - Low cost \$\$\$
    - Can use low cost LAN infrastructure
      - Dedicated infrastructure recommended
    - Can use vanilla NICs in workstations or servers
  - Long distance
    - If you can ping it, you can put your storage there...
  - High Host CPU cost
    - Using software based initiators burns CPU cycles to parse IP packets and to copy data
    - Hardware based initiators are \$\$\$ and not yet mature
  - Modest Performance
    - Uses Gbe LANs today
    - High CPU cost limits high MB/sec applications
    - Long distance and potential routers will limit performance
    - Low cost storage arrays will limit performance

# iSCSI Technology Demonstration Kit

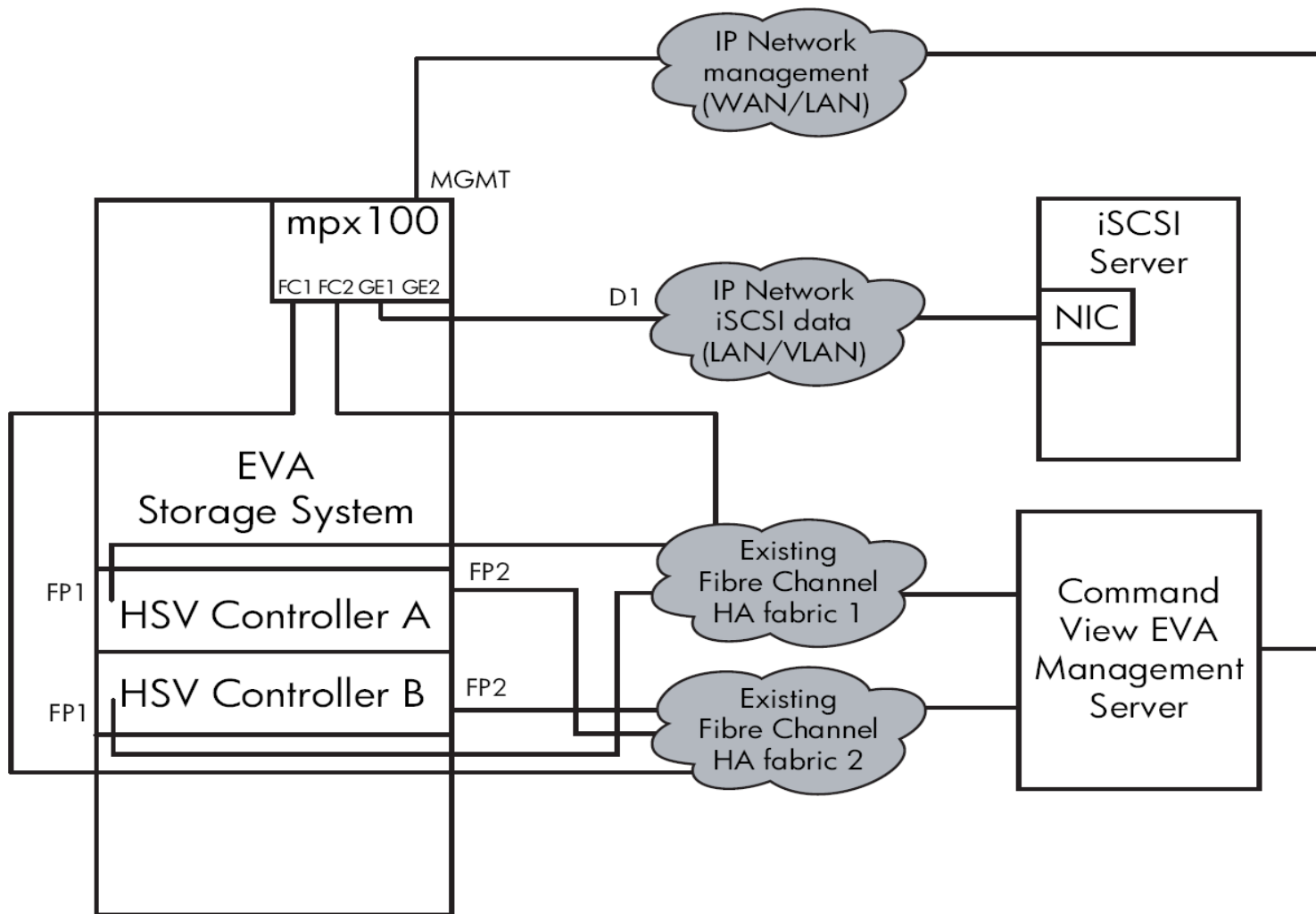


- VMS V8.3-1 includes support for a software based iSCSI initiator
  - Limited initial target support - EVA / MPX100
  - No boot support
  - Requires VMS TCP/IP stack (no 3<sup>rd</sup> party stack support)
  - Both Alpha and Integrity supported

# HP StorageWorks EVA iSCSI Connectivity Option - Direct Connect



# HP StorageWorks EVA iSCSI Connectivity Option - Fabric Connect



# Other I/O Support

- VGA console
  - Either built-in video or optional PCI graphics cards
  - No access to SYSBOOT or XDELTA during boot
- VMedia – remote access to installation media on management server
- USB 2.0 – high speed access to console media
- Edgeport serial line MUX
  - Same product ID, new chip

# WBEM Infrastructure & Providers

- Management of c-class blade systems from SIM management agents
  - Operating system
  - Computer system
  - Process and processor statistics
  - Indicators (event monitors)
  - Firmware version
  - Fan & power supply
  - Management processor
  - CPU instance
  - Memory instance
  - Enclosure

# Provisioning

- Remote installation/upgrade of OpenVMS from HP SIM management station
  - Up to 8 concurrent installations
  - Sourced from Infocenter or VMedia
  - Supports HP Integrity rx3600, rx6600, and BL860c Server Blades
- Remote installation/upgrade of OpenVMS via VMedia
  - Initiated from EFI console
  - Single installation at a time per VMedia image
  - Supports HP Integrity rx2660, rx3600, rx6600, rx7620, rx7640, rx8620, rx8640, and BL860c



# MSA Utility

- Management operations on MSA smart array controller from VMS host
- Command line based
- Add/delete units
- RAID set characteristics
- Reset controller
- Update firmware
- etc.

# Linker Features

- Larger I/O segment size
- Optimized use of external call stubs
- Unwind descriptors co-located with P2 space code
- Misc. bug fixes

# Misc. Software Features

- LDAP external authentication support
- New time zones
- New iconv international text converters
- SDA improvements
- Performance improvements

V8.next

# Hardware Support

- 2TB logical unit support
- 8GB Fibrechannel infrastructure
- iSCSI hardware assist
- Next generation storage controllers
- New blade storage controllers
- New systems as they appear...

# Performance Improvements

- Exception / unwind
- Lazy FP save/restore
- RAD support
- iCache flush
- Process VA deletion
- Dedicated lock manager
- Threads support
- XFC deferred writeback

# New Configurations

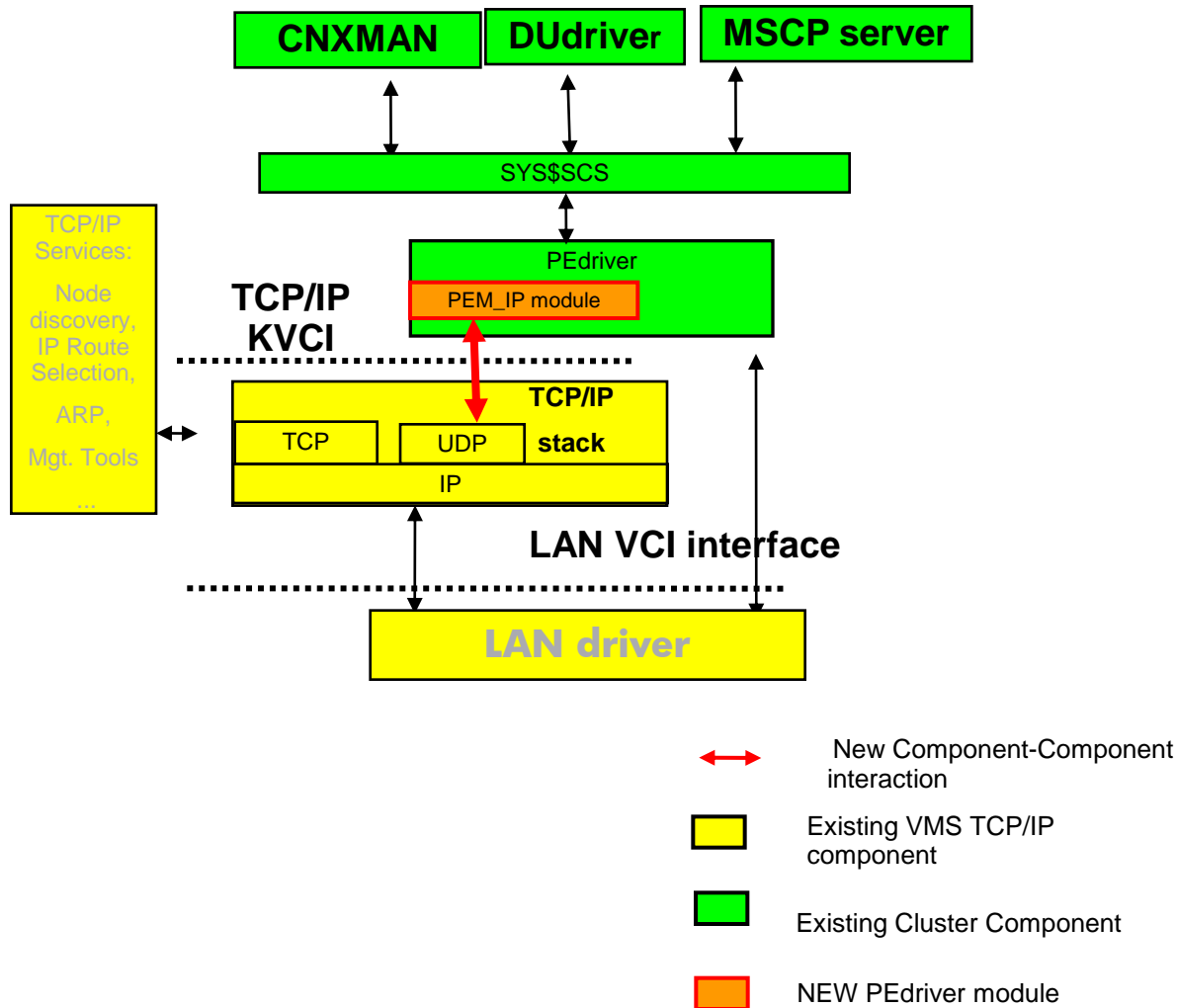
- OpenVMS under HPVM
  - just another cpu...
- VMSClusters over IP...

# Clusters Over TCP/IP

- Why
  - VMS Clusters requires special network management because cluster communications is a “bridged” as opposed to “routable” network protocol
  - Causes extra burden on Network Admins
  - Network switch vendors dropping “bridged” support – new CISCO switches are IP only
  - Eases DT configuration setup – IP internal and external
- What – package cluster communications in IP packets enabling VMS cluster to operate in pure IP only network

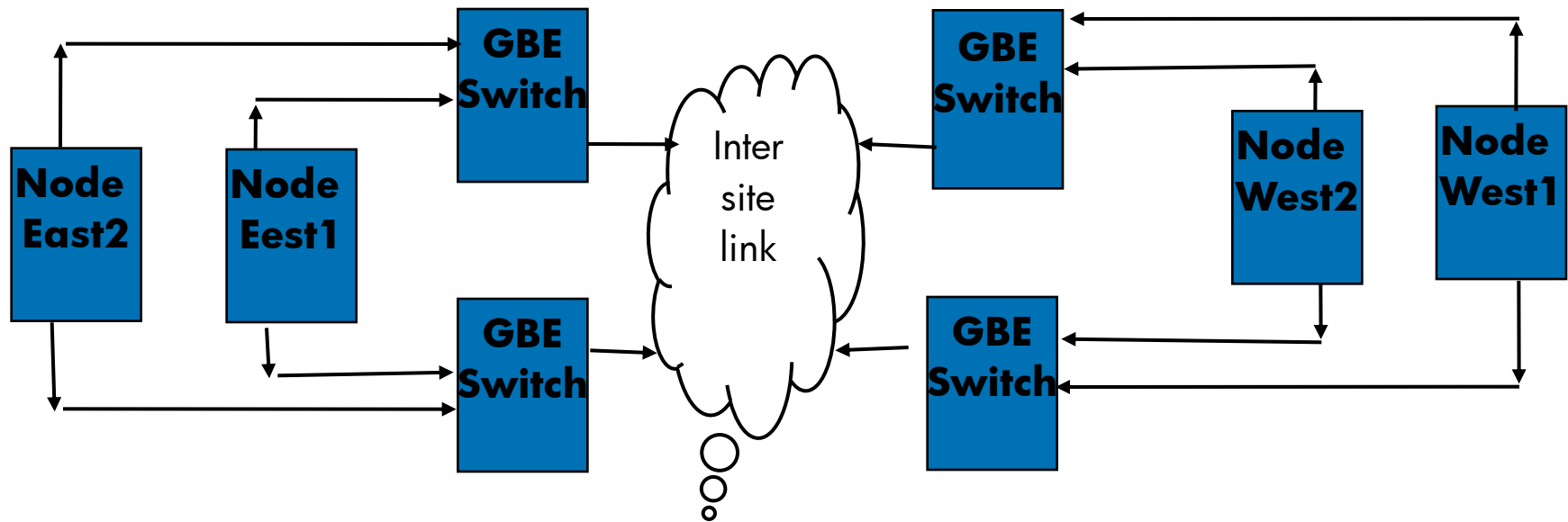


# IPCI solution – PE driver over UDP



# LAN interconnect – Extended LAN

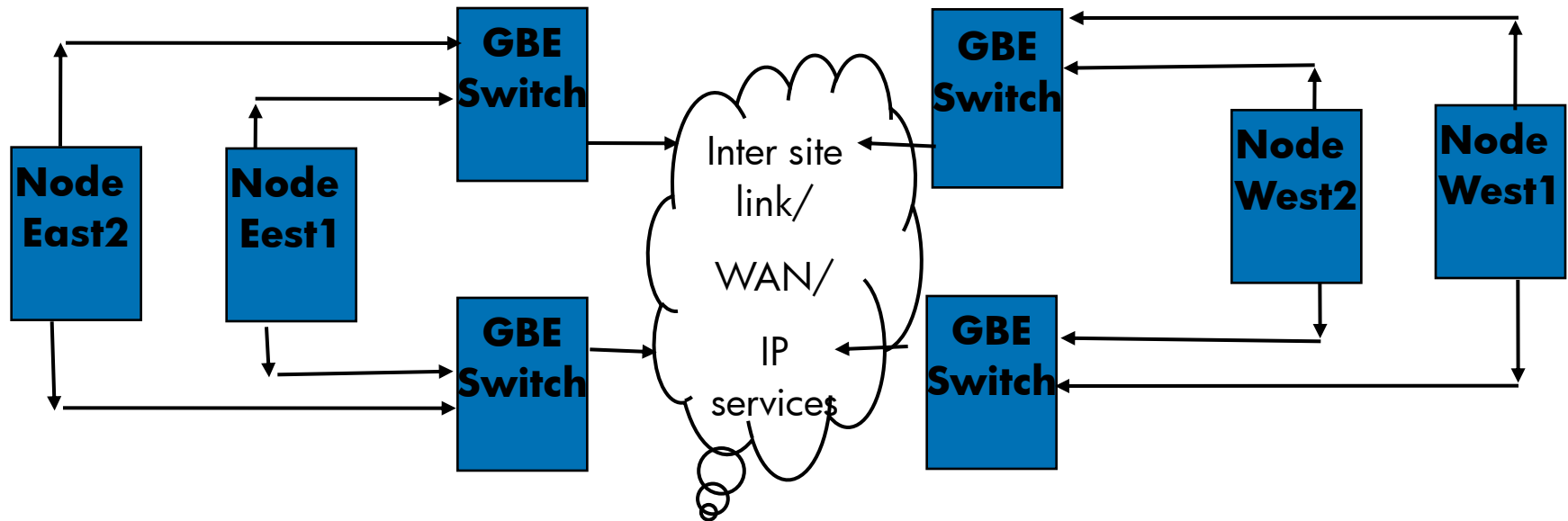
**Cluster using LAN bridging/Extended LANs**



**Node East1,East2 West 1,West2 should be part of SAME LAN for cluster communications.**

# Cluster using IPCI

**Cluster using IPCI for communication**



**Node East1,East2 West 1,West2 can be part of same or different LAN for cluster communications using IPCI.**

**Cluster communication using IP**

# Unix Portability

- RMS – symlinks, etc.
- Shared stream I/O
- Shared memory APIs
- ioctl() enhancements
- Semaphores

# RMS – Posix Pathnames

- Logical name support
- First field in an absolute path is searched:
  - Logical name
  - Directory or file name in the Posix root
  - VMS device name
- Examples
  - /SYS\$LOGIN/login.com
  - /DKA0/test/file.txt
  - /bin/GCC.EXE
  - /SYS\$LIBRARY/SYS\$LIB\_C.TLB

# RMS – Logical Names in Symlinks

```
$ dir sys$system:net*.dat
```

```
Directory SYS$SYSROOT:[SYSEXE]
```

```
NETCIRC.DAT;1          NETCONF.DAT;1          NETLINE.DAT;1          NETLOGING.DAT;1  
NETNODE_LOCAL.DAT;1  NETNODE_REMOTE.DAT;354  NETOBJECT.DAT;1
```

```
Total of 7 files.
```

```
Directory SYS$COMMON:[SYSEXE]
```

```
NETWORKS.DAT;1
```

```
Total of 1 file.
```

```
$ create/symlink="/sys$system" link_sys
```

```
$ dir [.link_sys]net*.dat
```

```
Directory DKA300:[TEST.LINK_SYS]
```

```
NETCIRC.DAT;1          NETCONF.DAT;1          NETLINE.DAT;1          NETLOGING.DAT;1  
NETNODE_LOCAL.DAT;1  NETNODE_REMOTE.DAT;354  NETOBJECT.DAT;1  
NETWORKS.DAT;1
```

```
Total of 8 files.
```

# RMS – Loop Detection

- **With GNV installed...**

```
$ dir sys$sysdevice:[*...]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT.mnt]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT.mnt.SYSDISK]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT.mnt.SYSDISK.PSX$ROOT]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT.mnt.SYSDISK.PSX$ROOT.mnt]
```

```
Directory of SYS$SYSDEVICE:[PSX$ROOT.mnt.SYSDISK.PSX$ROOT.mnt.SYSDISK]
```

oops...

- Included in V8.3-1 and future RMS remedial kits

# RMS – Other Posix Features

- Wildcard Posix file names
  - “^UP^test/a/\*.c”
  - No directory wildcarding
- Follow symlinks in wildcard directory search
  - Optional!
- Support secondary RMS inputs (DNM & RLF) for Posix filenames (maybe)
- Misc bug fixes



# Shared Stream I/O

- Shared-write access to stream files
  - Standard UNIX I/O default
- Applicable across the nodes of a cluster
- Can be used on any stream file type
- Accessible via the C RTL
  - Set per file by an option on file open APIs
  - Set per process by C RTL feature switch

# Shared memory

- Implemented via the C RTL
- POSIX shared memory
  - `shm_open()`, `shm_unlink()`
- System V shared memory
  - `shmat()`, `shmctl()`, `shmdt()`, `shmget()`
- Possibly `ftok()`
- Not applicable across nodes of a cluster

# ioctl() enhancements

- Support more than socket operations
- Terminal I/O: allow read termination when a given number of characters is entered (TCGETA, TCSETA)
- UDP sockets: I\_SETSIG (signal on data in)
- Ethernet: SIOCGIFMTU (check MTU size)
- Disk: GETBLKSIZE (disk size in blocks)

# Semaphores

- Implemented via the C RTL
- POSIX semaphores
  - `sem_open()`, `sem_post()`, `sem_wait()`, etc.
- System V semaphores
  - `semop()`, `semctl()`, `semget()`
- Thread-aware

# TCP/IP V5.7

- IPSEC
- Packet Processing Engine
- Other Improvements

# Performance, Performance, Performance

- 10 GbE is here
- Issues with High-Bandwidth Feeds
  - Protocol Overhead
  - Memory latency (mCPI)
    - Multiple bus-crossings moving data from NIC to App (5-6x on receive!)
  - Interrupt load
  - Instruction Path Length
  - Multiple-passes over data
  - CPU contention
  - Cache misses
  - Easy to saturate a CPU with packet processing
  - Performance improvements yoy (RFC 4297)
    - CPU – 50%, Memory – 35%, Network – 40%

# TCP/IP PPE Concepts

- Modelled on OpenVMS Dedicated Lock Manager
- Without PPE
  - TCP/IP runs on a shared CPU
  - When that CPU is nearing saturation, TCP/IP may become the bottleneck
- PPE Dedicates a CPU to TCP/IP
  - no more sharing
- PPE may be enabled/disabled dynamically
  - Potential for custom procedures to monitor performance and enable PPE as needed

# TCP/IP PPE Benefits

- Avoid context switches
- Avoid cache conflicts with application processes
- **Polling** for work avoids interrupt overhead
- No need to fork



# PPE Configuration Requirements

- Hardware
  - Requires more than one active CPU
  - PPE monitors this and will become dormant if only one CPU is left active or wake-up when more than one CPU becomes available
  - Better suited to systems with many CPUs
- Software
  - Configure the BGO: as the only driver on the nominated CPU
    - i.e. all other fast-path drivers must be moved to other CPUs

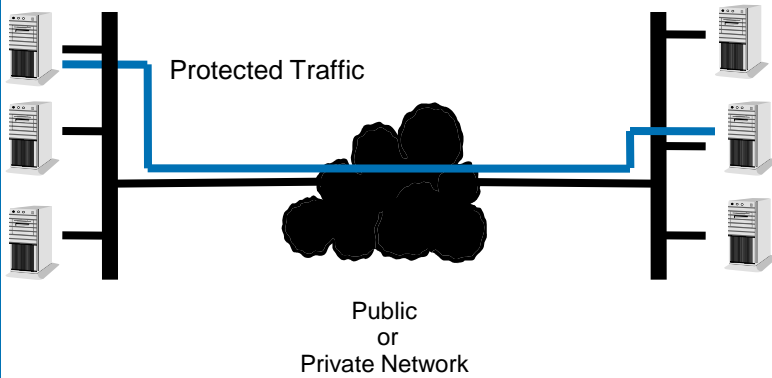
# PPE Futures

- NIC Polling
  - This implementation polls only for TCP/IP request packets
  - Disable NIC interrupts and poll NIC for work
- New API possibilities
  - Zero-copy
  - Extended Sockets API
- 10GbE is here – 100GbE is touted for 2010!
  - Need new ideas if TCP/IP is to keep pace

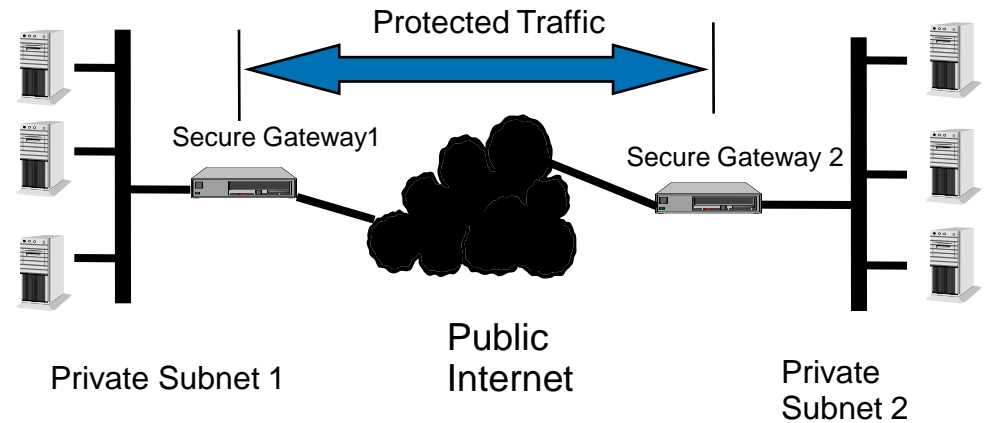
# What is IPsec?

- Provides security at the IP layer
- Strong security that can be applied to all traffic
- Transparent to applications and end users
  - No need to train users on security mechanisms
- Protects all upper layer protocols
- Secures traffic between any two IP systems
  - Can be used end-to-end, router-to-router, or host-to-router
- Extensions to the IP protocol suite
  - Applies to IPv4 and IPv6
- Encryption and Authentication
- Key management and Security Association creation and management

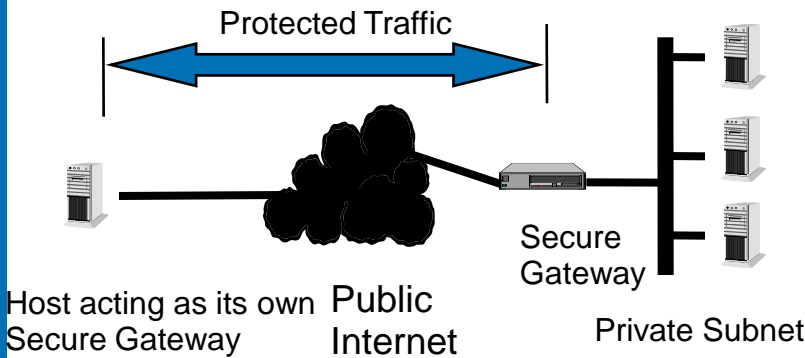
# IPsec Security



IPsec for Host-to-Host



IPsec for Virtual Private Networks



IPsec for Remote Access

# IPsec Support

- Based on the IPsec implementation from SafeNet Inc. <http://www.safenet-inc.com/> called "QuickSec"
- IPsec consists of
  - Interceptor - a platform-specific module that provides the interface between OpenVMS IP kernel and IPsec Engine module
  - Engine – a Loadable IPsec kernel module which provides crypto-processing of packets
  - Policy Manager/IKE - an application which provides processing of security policies formulated by the system manager and exchanges security policies information with remote hosts
  - Management – a set of management utilities (such as key generation, etc.)
  - Configuration tool – a basic IPsec configuration tool which processes security policies formulated by a system manager

# Command line interface (like HP-UX): ipsec\_config



- ADD HOST
  - Specify the IPsec behavior for IP packets sent or received by the local system as an **end host**
  - There is a “default” policy in the ipsec\_config.db
    - Pass packets in clear text
  - Specify source & destination addresses/ports/services or protocol; transforms; precedence; manual keys
  - Can force tunnel mode if the local system is a tunnel endpoint

# Another ipsec\_config command: Add Internet Key Exchange (IKE)



- ADD IKE
  - On HP-UX, represents an IKE Policy record
    - On OpenVMS, information is part of the tunnel object data
  - Used to configure dynamic keys
  - Not needed for configuring manual keys
  - IPsec uses the parameters in an IKE policy when using the IKE protocol to establish ISAKMP/Main Mode SAs with remote systems
    - Needed before IPsec can negotiate IPsec SAs
  - Specify the remote peer, precedence, auth type (PSK or certificates), hash alg, encryption alg, etc.

# One more ipsec\_config command: Add Authentication Records



- ADD AUTH
  - On HP-UX, represents an authentication record
    - HP-UX sends update to IKE daemon
    - On VMS, the information is stored in a pre-shared key object which is part of the tunnel object data
  - Needed only for dynamic keys, and further, only if you configured pre-shared keys as an IKE authentication method
  - Not needed if using certificates
  - Can be used to configure IKE ID information
  - This is where you specify the pre-shared key



# Sample configuration commands

## **ipsec\_config commands for host-to-host using a preshared key**

On Host A (requires equivalent entries on Host B):

```
$ipsec_config add host hostb
```

```
  -source 10.1.48.7
```

```
  -destination 10.1.48.8
```

```
  -action ESP_AES128_HMAC_SHA1
```

```
$ipsec_config add ike hostb
```

```
  -remote 10.1.48.8
```

```
  -authentication PSK
```

```
$ipsec_config add auth hostb
```

```
  -remote 10.1.48.8
```

```
  -preshared hello,world
```

# IPSEC Status

- EAK available now
  - <http://h71000.www7.hp.com/openvms/products/ipsec/index.html>
- EAK includes PPE
- Firewall function included in IPSEC code base
  - No separate firewall product planned

# TCP/IP Misc Features

- Scriptable/automated config for mass deployment
- Improved NFS setup documentation
- IPv6 Logo Testing
- UNIX-style SIGIO asynch I/O API
- TCPIP\$PEERNAME utility
- LPD configurable port
- Complete \$ACM support in FTP, POP, IMAP, and REXEC (and SSH re-fit)

# TCP/IP Misc Features

- FTP compatibility for major browsers
- FTP anonymous-light
- RENAME /FTP (and finish DELETE /FTP)
- NFS: continued enhancements
- SFTP access control improvements
  - NETWORK, BATCH, TERMINAL access, etc.
- Multi-threaded SMTP receiver
- IMAP long line (>255 chars.) handling
- IMAP new message polling

# System Management

- More providers
- Provisioning enhancements
- Infrastructure update

# WBEM Providers

- Indications
  - Reports to HPSIM based on events and thresholds
  - Storage controllers and media  
(e.g., device offline, shadow set faults)
- Networking Providers
  - DNS
  - LAN
  - Ports

# WBEM Providers

- Insight Power Manager
  - Monitor and graph power usage
  - Policy-based power regulation
    - High power – Maximum CPU performance
    - Low power – Maximum power savings
    - Dynamic – Adjust according to CPU load
    - OS control – Let OS control CPU
- Additional Blade hardware support

# Other Management Features

- WBEM Services
  - OpenPegasus V2.7
  - Support for IPV6
- Provisioning
  - Enable licenses via HPSIM
  - Cluster & TCP/IP configuration via HPSIM
- GiCAP
  - Ability to shift capacity units among systems
- OpenView enhancements under evaluation



# Contact info

- Gaitan D'Antoni (OpenVMS Technical Director)
  - [Gaitan.Dantoni@hp.com](mailto:Gaitan.Dantoni@hp.com)
- Leo Demers (OpenVMS Product Management)
  - [Leo.Demers@hp.com](mailto:Leo.Demers@hp.com)
- Andy Goldstein (Speaker)
  - [Andy.Goldstein@hp.com](mailto:Andy.Goldstein@hp.com)



i n v e n t