# CHARON-VAX application note

## AN-036 Testing various Shared Disk VAX-cluster

Author: Ralf van Diesen Updated: 16 September 2005

### CLUSTERING WITH DIRECT ACCESS USING SHARED STORAGE

VAX/VMS clusters can be formed with different interconnects, SCSI, Memory Channel, CI, DSSI and LAN. All these interconnects have their own limits.

| SCSI | Memory Channel | CI | DSSI | LAN |
|---|---|---|---|---|
| 3 systems | 4 systems | 16 systems | 4 systems | 96 systems |
| Max. 20 MB/s | Max 100 MB/s | Max. 17.5 MB/s | Max. 4 MB/s | Max. 12.5 MB/s<br>-FDDI<br>Max. 1.25 MB/s<br>-Ethernet |
| 25 meters | 6 meters | 90 meters | 25 meters | 40 Km - FDDI<br>2 Km - Ethernet |
| Alpha Only | Alpha Only | VAX & Alpha | VAX & Alpha | VAX & Alpha |

Combine several of these interconnect types to create a new way of VAX/VMS clustering with CHARON-VAX and modern hardware. The result is:

- Cluster with a potentially large number of nodes.
- Maximum speed (100 MB/s)
- Long distances (one Ethernet segment; several Km)
- Using different storage solutions (SCSI, ATA, SATA, SAN, iSCSI etc.)

### HOW TO BUILD A CLUSTER USING CHARON-VAX

Use the MSCP disk controller of CHARON-VAX/XM (Plus) and CHARON-VAX/XL (Plus) for Windows to create a VAX/VMS shared disk cluster. Note that an MSCP controller theoretically supports up to 256 nodes in a cluster.

Using the CHARON-VAX MSCP controller brings up a fast and flexible solution for shared storage; iSCSI is a suitable technology to share disks over Ethernet. Store data on an iSCSI target and use iSCSI initiators running on the VAX-emulator equipment to access the data. The iSCSI target can be a hardware storage solution (such as the PeerStorage$^{\text{TM}}$ Array 100$^{E}$ from Equallogic), or a storage host running iSCSI target software (such as Wintarget). The Windows host platforms with the standard Microsoft iSCSI initiator installed will see virtual physical disk(s) in the device list that can be directly assigned to the MSCP controller in CHARON-VAX (like \\.\PhysicalDriveX).

See application note 22 for further details on Building VMS disk cluster systems with CHARON-VAX.

- **Sharing physical disks with iSCSI target software.**

  Using iSCSI target software (Wintarget) as iSCSI target a multi node VAX/VMS cluster can be created. This way the VAX-emulator host platforms will access a virtual disk that can be assigned to the MSCP controller of the VAX-emulator. This type of cluster has been tested with 5 nodes and 5 disks mounted cluster wide. (http://www.stringbeansoftware.com)

- **Sharing logical volumes on iSCSI target hardware.**

  Using iSCSI target hardware (PeerStorage$^{TM}$ Array 100$^{E}$) as iSCSI target a multi node VAX/VMS cluster can be created. Just like the iSCSI target software solution the Windows host platforms of the CHARON-VAX access a virtual disk that can be assigned to the MSCP controller of CHARON-VAX. In addition to the software target (Wintarget), the PeerStorage$^{TM}$ Array 100$^{E}$ supports RAID 10, RAID 50, mirroring between multiple targets and load balancing across 3 internal Gigabit Ethernet ports. Clustering with this storage unit has been tested with 5 nodes and 5 disks mounted cluster wide. (http://www.equallogic.com)

- **Sharing container files on Windows with MSCP controller.**

  Creating a multi node VAX/VMS cluster with an MSCP controller and CHARON-VAX container files as shared disks works the same as container files on the SCSI controller but now all disks can be mounted cluster wide. This has been tested with 5 nodes and 5 disks mounted cluster wide.

## PERFORMANCE

- All tests are performed on a AMD quad Opteron 850 2.4 Ghz with 2 GB Memory.
- The server used for Wintarget and container file sharing is a P4 2.5 Ghz with 512 MB Memory.
- Size of used iSCSI volumes is 10GB each.

The two main aspects of disk I/O performance are the throughput when copying large files (that mainly depends of the I/O infrastructure performance) and the speed at which individual file records can be read/written. At typical (emulated) VAX speeds and modern storage hardware, record oriented I/O is mostly limited by the VAX/VMS application environment, as the test results below show:

***Copying a large file from dua0: to dua1:***

On the Equallogic storage unit the average transfer speed is: 16 MB/sec.

On Wintarget the average transfer speed is: 3 MB/sec.

Using container file sharing the average transfer speed is: 1.5 MB/sec.

***Copying a large file from dua0: to dua0:***

On the Equallogic storage unit the average transfer speed is: 14 MB/sec.

On Wintarget the average transfer speed is: 3 MB/sec.

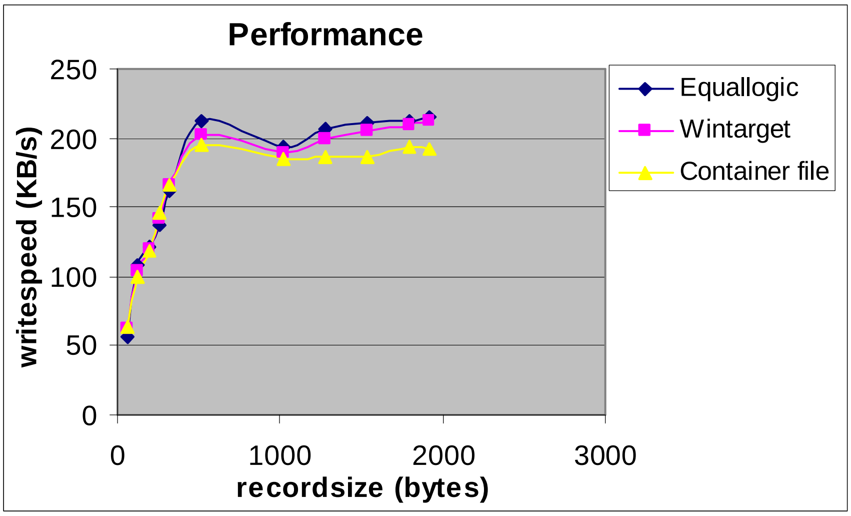
Using container file sharing the average transfer speed is: 1.2 MB/sec.

Note that in the case of Wintarget and shared container file measurements the I/O speed depends a lot on the drive parameters. With fast SCSI hardware “disk to disk copy” will probably be somewhat faster than with the disks tested, but file copy performance of 'software based' sharing remains largely below that of dedicated storage hardware.

### *I/O performance versus record size*

Only a minimal difference is noticeable for operations on a single file, reflecting the VAX I/O limitations. Work on repeating the same tests simultaneously from a large number of nodes will reflect much better the effective storage I/O bandwidth of each solution, but has not yet been completed.

Using the “disk_bench_var.com” performance tool the write speed into a single file relative to the record size is displayed in the following graph:



## SOME ALTERNATIVE METHODS THAT MAY WORK FOR CHARON-VAX:

### SCSI storage unit and CHARON-VAX host system SCSI controller:

Creating a two node VAX/VMS-cluster with shared SCSI.

To connect two CHARON-VAX host systems to shared SCSI disks, a dual port SCSI storage unit with simultaneous access from both systems is necessary, such as the HP StorageWorks MSA 500 with two (physically separated) SCSI ports. Both host systems are connected to this storage unit via their respective SCSI controllers to access the shared SCSI drives.

*Note that not all dual port SCSI storage units (especially the ones specifically designed for Windows systems) provide fully separated SCSI ports; in this case a host system SCSI bus reset can impact the other system and result in disk timeouts causing a disk to go offline in VMS.*

The MSCP controllers in the two CHARON-VAX emulators are mapped to the shared SCSI disks. See AN-022 for the correct setup. Since VAX SCSI controllers do not support the Tagged Command Queuing (TCQ) protocol, it is not possible to include a physical VAX system in this (or any other) cluster setup.

*When using the MSA 500 storage unit, disable “Selective Storage Presentation” using HP ACU (Array Configuration Utility). This characteristic is stored in the MSA controller’s non-volatile memory. The larger MSA 1000 can provide the same configuration using fibre channel access ports.*

### Using a shared VMS container file on a Windows remote share:

This is a marginal solution that can be used for test and demonstration purposes but not for a production system. The VAX SCSI controllers (and hence the emulated VAX models using a SCSI controller) do not support Tagged Command Queuing (TCQ). This restriction limits the number of shared VMS disk (images) to one, effectively the disk from which both systems boot.

To realize this the container file must be located in a directory that each CHARON instance can access. Sharing a folder on a server and mapping this folder on each Windows system will do. Locating the shared container file on one of the host systems instead of on a separate server will result in performance loss due to the CHARON-VAX system load and likely cause a disconnect.

As long as only one disk image is shared, more nodes can be added. This type of cluster was tested with 4 nodes and running disk tests writing the same file from 4 nodes for days without any data corruption occurring.

## FAIL SAFE CONSIDERATIONS

### iSCSI-target hardware (Equallogic):

- Disconnecting the network cable and reconnecting it again will restore the data connection. VMS will see the device offline until it's reconnected again.
- The Equallogic storage unit supports RAID 10, RAID 50 and multiple storage units can be mirrored (even over the Internet).
- Load balancing over 3 Gigabit Ethernet ports.

### iSCSI-target software (Wintarget):

- Disconnecting the network cable and reconnecting it again will restore the data connection. VMS will see the device as offline until it is reconnected again.
- No facility to configure multiple iSCSI targets for failsafe switchover. Limited to applications where failsafe data access is not required.

### Container file on a remote Windows share with MSCP controller:

- Disconnecting the network cable and reconnecting it again will result in a disk time-out in VMS. The data connection will not be restored and VMS needs to be restarted before it's available again.

### SCSI storage unit:  (MSA 500/1000):

- The HP MSA500/1000 have both been tested successfully with CHARON-VAX.

### Container file on a remote Windows share with emulated SCSI controller:

- Disconnecting the network cable and reconnecting it again will result in a disk time-out in VMS. The data connection will not be restored and VMS needs to be restarted before it's available again.

## PRO'S AND CON'S

Note: in all cases, the VMS SCS protocol uses a NIC configured in CHARON-VAX. Since the SCS protocol is limited to one Ethernet segment, the VMS cluster is limited to one segment as well. Where remote shares or iSCSI is used, the VMS data transfer passes via a separate Ethernet controller configured in the Windows host system. This implies specific measures to safeguard the CHARON-VAX host systems against unauthorized access. Read Application note 29 for Recommendations Regarding Security of CHARON-VAX Host Platforms.

### iSCSI-target hardware (Equallogic) and CHARON-VAX MSCP Controller

- All disks can be mounted cluster wide
- Not limited in the number of cluster nodes (VMS limit 256)
- Failsafe on network and storage, high performance RAID support independent of iSCSI initiators.
- Very high throughput, easy to expand by adding more units.

- Cost effective for high-end, multi node, automated backup configurations.

**iSCSI-target software (Wintarget) CHARON-VAX MSCP Controller:**

- All disks can be mounted cluster wide
- Not limited in nodes (VMS limit 256)
- Failsafe on network, but implies the use of a single point of failure Windows system as disk server.
- Inexpensive
- Windows server required as target (data storage).
- No ability to increase disk I/O access bandwidth by adding target systems.

**Windows remote share container file with CHARON-VAX MSCP Controller:**

- All disks can be mounted cluster wide
- Not limited in nodes (VMS limit 256)
- Not failsafe on network, not suitable for production purposes.
- Inexpensive
- Extra PC required as file server

**SCSI storage unit using a CHARON-VAX MSCP controller via a host system SCSI port:**

- All disks can be mounted cluster wide.
- Limited to two VAX/VMS nodes at 'SCSI distance' (about 5 meters).
- Failsafe on SCSI connection, but requires dual ported SCSI storage unit.
- Cost effective for midrange dual node production systems.
- Does not require to connect the Windows host systems to Ethernet, reduces security concerns.

**Windows remote share container file with CHARON-VAX SCSI Controller:**

- Limited to ONE Cluster wide disk
- Not limited in nodes (VMS limit 256)
- Not failsafe on network
- Inexpensive
- Completely unsuitable for production use.

[30-18-036]