

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

Author: Ralf van Diesen

Date: 21-07-2004

### 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 -FDDI Max. 1.25 MB/s -Ethernet
25 meters	6 meters	90 meters	25 meters	40 Km - FDDI 2 Km - Ethernet
Alpha Only	Alpha Only	VAX & Alpha	VAX & Alpha	VAX & Alpha

Combining the best of these interconnects using CHARON-VAX and the most sophisticated hardware to create a new modern way of VAX/VMS clustering results in:

Virtually, no limit of connecting nodes (just the VMS limit of 256)

Maximum speed (100 MB/s)

Long distances (40 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. Data is stored on an iSCSI target and iSCSI initiators running on the VAX-emulator equipment access the data. The iSCSI target can be a hardware storage solution (such as PeerStorage™ Array 100<sup>E</sup> 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™ Array 100<sup>E</sup>) 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™ Array 100<sup>E</sup> supports RAID 10, RAID 50, mirroring between multiple targets and load balancing across 3 internal Gigabit Ethernet ports. Clustering with this box 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 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 is 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.

Comparison between the disk I/O performance of the Equallogic box, Wintarget and container file sharing:

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

On the Equallogic box the speed is: 16 MB/sec.

On Wintarget the speed is: 3 MB/sec.

On Container file sharing the speed is: 1.5 MB/sec.

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

On the Equallogic box the speed is: 14 MB/sec.

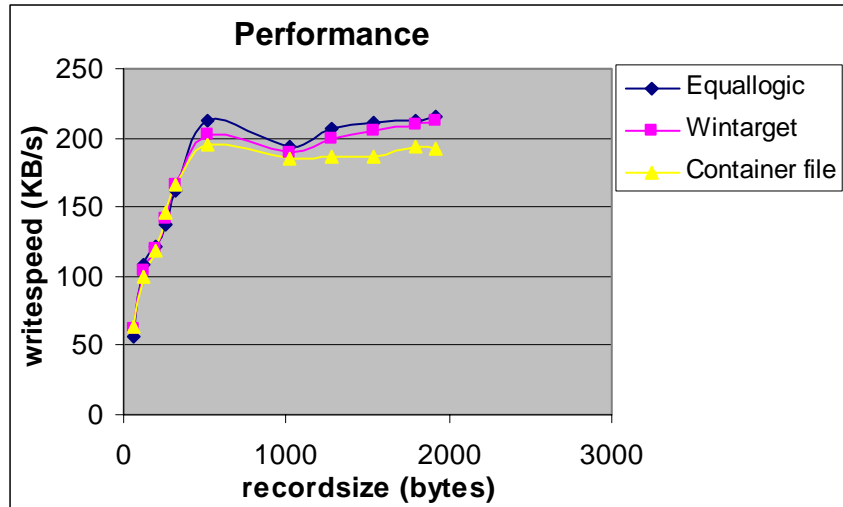
On Wintarget the speed is: 3 MB/sec.

On Container file sharing the speed is: 1.2 MB/sec.

The disk controller used in the Wintarget and shared container file measurements was a fairly old controller so the "disk to disk copy" will probably be faster with up-to-date disks.

## CHARON-VAX application note AN-036

Using the “disk\_bench\_var.com” performance tool the write speed depending on the record size is displayed in the following graph:



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

#### SCSI box and CHARON-VAX SCSI controller:

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

To connect two nodes to shared SCSI disks a dual port SCSI box with simultaneous access from both nodes is necessary, such as the HP StorageWorks Modular SAN Array 500 (MSA 500). This storage box contains several SCSI-drives and has 2 (physically separated) SCSI ports. Both nodes are connected to this box with a SCSI card and they can both access the drives like local disks. Choosing the right storage box is important, because the SCSI busses are not always completely (physically) separated. In this case resetting of the SCSI bus will result in disk timeouts and the disk will go offline in VMS. Most dual ported SCSI boxes are originally designed for Windows clustering which is totally different from VAX/VMS clusters and will not work. If CHARON-VAX is mixed in a VAX cluster with a real physical VAX, SCSI clustering is the only solution.

#### Notes:

1. If you use the MSA 500 SCSI box, you must disable “Selective Storage Presentation” using HP ACU (Array Certification Utility). This characteristic is stored in the MSA controller’s non-volatile memory.
2. The larger modular storage array of HP, the MSA 1000, provides more storage capacity and multiple fibre channel access ports.

#### Sharing Container files with CHARON-VAX SCSI Controller:

Creating a multi node VAX/VMS-cluster with container files as shared disks eliminates the problem of undesirable resetting of the SCSI bus. As the disks are emulated SCSI disks (container files) no physical resetting of the bus will occur. To create such a cluster each node must be able to boot from the same container file. To realize this the container file(s) 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. Sharing the disks on one of the nodes instead of on a separate server will result in performance loss on this node and is not recommended. Using the emulated SCSI controller has a limitation; only the system disk can be mounted cluster wide because there is no support for Tagged Command Queuing (TCQ) in the emulated SCSI controller. This restriction will prevent this solution to replace a real cluster since there are almost no clusters with only the system disk mounted cluster wide. It is suitable though for smaller VAX-

## CHARON-VAX application note AN-036

---

emulator clusters. This type of cluster is tested with 4 nodes and running disk tests writing the same file from 4 nodes for days proved that no data corruption occurred.

### FAIL SAFE CONSIDERATIONS

#### **iSCSI-target software (Wintarget):**

- Disconnecting the network cable and reconnecting it again will restore the data connection. VMS will see the device offline until it's reconnected again.

#### **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 box supports RAID 10, RAID 50 and multiple boxes can be mirrored (even over the internet)
- Load balancing over 3 Gigabit Ethernet ports.

#### **Container file 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 box (MSA 500):**

- Not tested

#### **Fibre Channel Box (MSA 1000):**

- Not tested

#### **Container file with 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

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

- All disks can be mounted cluster wide
- Not limited in nodes (VMS limit 256)
- Failsafe on network and storage
- Fast
- Expensive
- No extra PC required
- Easy to expand
- Long distances possible, SCSI protocol can't be routed and limits the cluster to 1 segment

### **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
- Fast
- Inexpensive
- Extra PC required as target
- Easy to expand
- Long distances possible, SCS protocol can't be routed and limits the cluster to 1 segment

### **Container file with MSCP Controller CHARON-VAX MSCP Controller:**

- All disks can be mounted cluster wide
- Not limited in nodes (VMS limit 256)
- Not failsafe on network
- Inexpensive
- Extra PC required as file server
- Easy to expand
- Long distances possible, SCS protocol can't be routed and limits the cluster to 1 segment

### **SCSI box and CHARON-VAX SCSI controller:**

- All disks can be mounted cluster wide
- Limited to nodes, max 2
- Failsafe on SCSI connection
- Inexpensive
- Only small distances are possible +- 25 meters

### **Container file with CHARON-VAX SCSI Controller:**

- Limited to 1 Cluster wide disk
- Not limited in nodes (VMS limit 256)
- Not failsafe on network
- Inexpensive
- Extra PC required as file server
- Easy to expand
- Long distances possible, SCS protocol can't be routed and limits the cluster to 1 segment

[30-18-036]