SCSI Features, Futures & Issues

Glenn C. Everhart
Everhart@star.enet.dec.com

Digital Equipment Corp., VMS Engineering

New Features Being Added

* Wider IDs (16 IDs per bus if the bus is wide)

* Cluster support with new drivers (PKS, PKQ)

* Fast 20 support

* Device Naming changes (to support simpler configurations
in SCSI Clusters)

* Mode sense code in VAX like Alpha code.

* Synchronous, Modesense, and TCQ improvements

* Better SCSI arbitration compensation

* SCSI-3 and Fibre Channel work going on, further out

What This Means To You

* More 3rd party devices will "just work"
* Initial ~ 11 errors on some disks go away
* Some new options exist to avoid errors
* SCSI cluster configurations easier to set up
* Opticals may no longer need special drivers
* Less device timeout errors under extreme load
* Fibre and SCSI-3 will mean larger areas, more devices

What Problems Have There Been
(3rd Party Devices)

* VMS supplies class drivers, one for all devices of a kind

* SCSI must negotiate several things before a transfer can be set up:

  > What devices are at this ID?
  > What synchronous/asynch counts may be used?
  > Is the device a wide-bus device?
  > How should errors be handled?
  > Can the device handle tagged commands?
  > What geometry or density parameters exist?

* In addition it must assume that for most purposes a "device" is an ID and a LUN

* Each of these gives rise to problems with some devices
(Digital qual deals with these for Digital devices; 3P support is the issue.)

Issue: What devices are at this ID?

Problems:

* Some devices hang, respond to all LUNs alike, get confused when commands to multiple LUNs are sent to the devices

* Much more common on tapes due to PC-oriented manufacturers and cost factors. CDs also often fail this way.

Approaches:

* Code knows some devices that cannot handle LUN commands. Prevents autoconfigure from ever trying multiple LUNs for those.

* SYSMAN IO exclude command. Allows devices to be excluded from autoconfigure, configure them manually (e.g. in systartup_vms)

Issue: What synchronous/asynch counts may be used?

Problems:

* Some 3P devices report synch support, but it doesn't work
* Unusual responses to synch negotiation
* Devices that run synch for user data, asynch for command data

Approaches:

* Again, software knows some "rogue" devices.
* VMS drivers will accept much larger range of responses, will only use synchronous mode after a completely successful negotiation.
* System will try more combinations of command modes to find one the device can accept.

>>> Wide negotiation is very similar to synchronous

Issue: How should errors be handled?

Problems:

* VMS drivers have rejected some safe error options in the past.

* Bad block recovery has lost data on old SCSI-1 disks. (VMS rejects these.)

Approaches:

* Most safe error recovery options now accepted. (One unsafe one is "RC" which allows the device to report phony data if set.) VMS will not allow this. Where data integrity and logging are **not** compromised, VMS will allow the device as a local one. (Mostly affects opticals, some intelligent controllers.)

* Devices on a shared cluster bus must still support more restrictive recovery, and must support device error correction (since VMS does not coordinate error recovery of bad blocks).

* Mode pages can be missing completely if their information can be found otherwise or can be defaulted. Again, this mostly benefits opticals.

* Transient error recovery is much more robust in new VMS drivers.

Issue: How should unfair SCSI arbitration be dealt with?

Problems:

* SCSI arbitration is inherently unfair...higher IDs ***always*** win

* Commands must be timed out in practice

* Busy devices can "starve" lower IDs

Approaches:

* TQE driven code checks for "at risk" I/O, stops higher ID’s I/O if timeout is imminent.

* On a shared bus, periodic INQUIRY to all devices will ensure that all initiators "see" the at-risk I/O. (This is current thinking.)

* If this is an issue now, put the initiators at low SCSI IDs.

Issue: Can the device handle tagged commands?

Problems:

* Tagged commands are needed for clusters because they give an ordering guarantee.

* Queue tag values supported may not be all of 0-255 range.

* Mis-implemented by some 3rd party manufacturers at times. Lack of fairness counts, misunderstood tag meanings, and tag ranges all have been reported problem areas. (Digital device qual tests these things.)

* First was used in VMS 6.2; some devices apparently "broke"

Approaches:

* "Rogue" devices are known to software.

* Recognize DQUE bit, so that tagged queueing can be turned off at the device in most cases.

Issue: What geometry or density parameters exist?

Problems:

* Some devices report no geometry information, or information which cannot be used for a legal VMS geometry.

* Some tape types use vendor unique density or compaction controls.

Approaches:

* Device geometry is now fabricated if the device doesn't give it. Device size is still needed.

* While device block size must be 512 bytes, non-512-byte block disks may at least permit $qio access.

* Tape density issues still under investigation. Compaction has been made more likely to work due to firmware updates. (Issue only with some 3rd party manufacturers; compaction is standard on Digital tapes.)

What Can You Do when your device won't work?

* Check your cables and terminators!!!

* SCSI_INFO in the SYS$ETC directory will report the first mode pages

* Error log information

* SCSI documents at *ftp.symbios.com* site if you want references. (These are draft standards only, but ordering info for final ones is there also.)

* SCSI_MODE tool will allow mode page data to be written provided you know the SCSI offsets. Do ***not*** use this unless you know what you're doing!!!

* Look for block size (should be 200 hex), device type in mode headers if you have the standard handy.

* Remember to check DIP switches on the device, set appropriately for a multiuser system.

Conclusion

* VMS is becoming much more 3rd Party Device - friendly, so long as data integrity OK.

* (Hopefully you begin to understand some of what Digital's I/O device qualification does to filter out weird behavior and bugs.)

* Drivers to support faster I/O (speed and width!) and more devices

* Cluster configuration SCSI device naming will be much easier

* SCSI-3 and Fibre Channel are on the way. This should support

  > Very large SCSI I/O farms
  > Much larger physical geometry
  > SCSI bus as the cluster bus (and possibly for other protocols)
  > Device naming changes to support much larger device farms
  > VMS will still be the ***only*** system where clustering will allow fully concurrent device access from multiple hosts.