

AccessionIndex: TCD-SCSS-T.20121208.102

Accession Date:

Accession By:

Object name: HP ExDS High-Availability Storage System

Vintage: c. 2009

Synopsis: Highly redundant RAID system with 232TB usable storage.

**Description:**

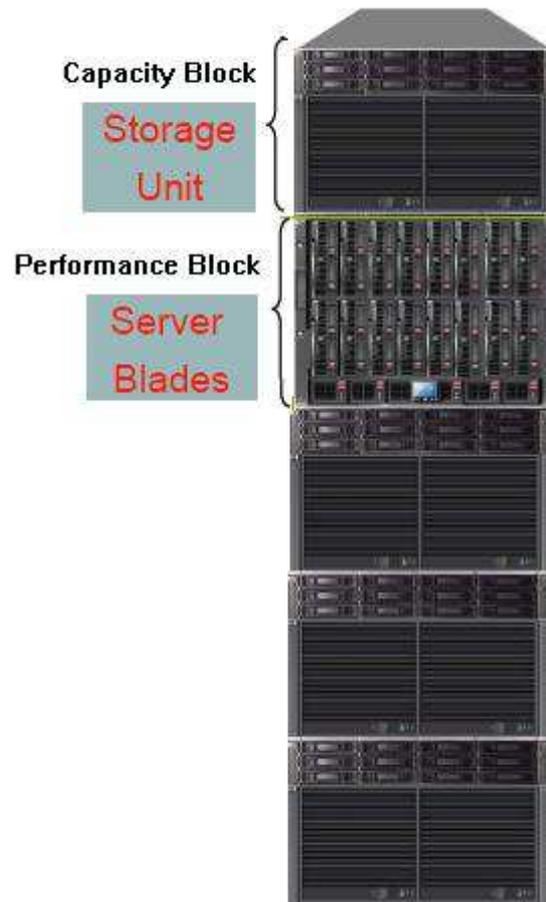
The Hewlett Packard (HP) ExDS-9100 is a highly-density storage system that presents a highly-available filesystem across 328TB of raw or 232TB of usable storage. This was installed as a back-end of the csTCDie grid site datastore for frequently-accessed mutable data, see elsewhere in this catalog. Problems with its 10Gbps virtual network switch made it unreliable.

The ExDS is designed around three major entities, *Capacity Blocks*, essentially high-density storage units, and *Performance Blocks*, essentially HP server blades, and a high-performance high-availability I/O infrastructure of redundant SAS switches and RAID array controllers.



Figure 1: HP ExDS system and constituent blocks, three-quarter views (courtesy HP)

The ExDS in this collection has four *Capacity Blocks* and one *Performance Block*.



*Figure 2: HP ExDS system and constituent blocks*

The Performance Block consists of HP C-Class blades running HP's PolyServe cluster file system (which is optimised for high-availability storage management), and potentially can also host user applications. These can hold up to 16 blades, with fairly flexible interconnect bays for Ethernet, SAS, etc, and a built-in administrative facility.



*Figure 3: HP ExDS blade chassis, front and rear views  
Generic images courtesy Tim Reddin, HP Galway*

There are 4 x ProLiant BL460c blades, each with two quad-Xeon-5300 Sequence processors with 8 x DIMM slots, two hot-plug disks, two 1Gbps Ethernet interfaces

and two I/O expansion slots. The blade enclosure includes an I/O infrastructure of dual-redundant SAS switches with an aggregate bandwidth of 3.2GB/s. Each server has a P700M mezzanine SAS host bus adapter with two serial-attached-SCSI (SAS) ports. Each SAS port is connected via the backplane to a SAS switch (TBM). There are two SAS switches so that each server is connected by a dual-redundant SAS fabric. Each disk array has two redundant controllers, and each of the controllers is connected to each SAS switch.

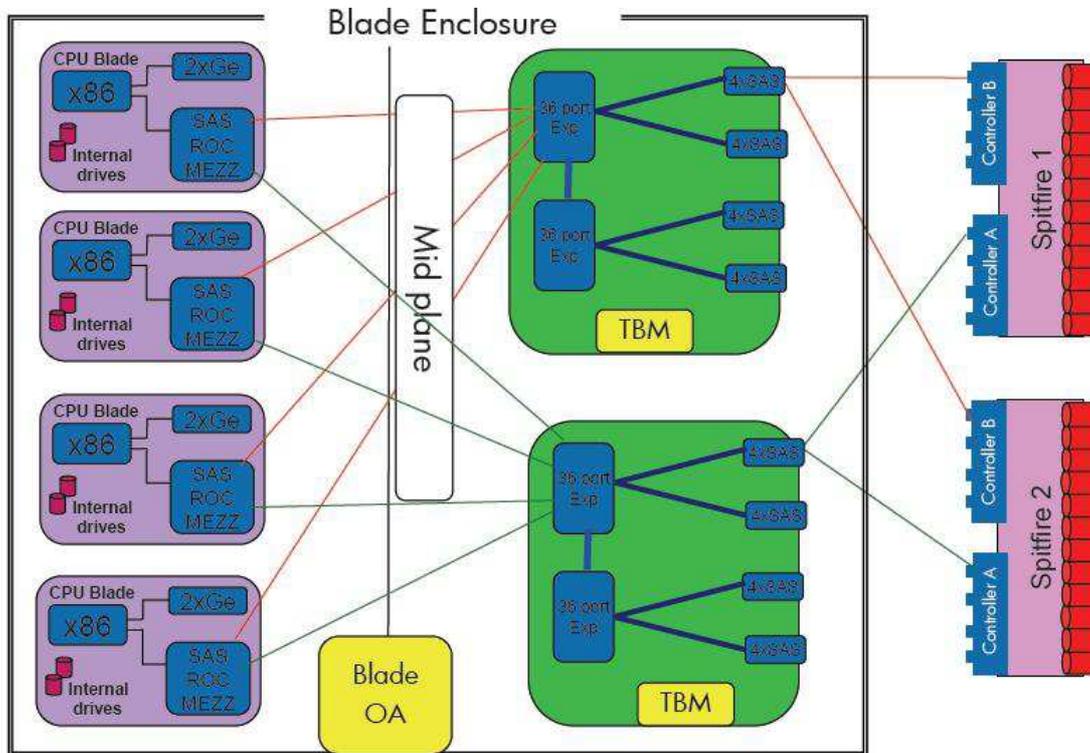


Figure 4: HP ExDS I/O fabric  
Images courtesy Tim Reddin, HP Galway

The Capacity Blocks consist of a 2U HP *Spitfire* SAS array controller (at the top of the block) with a battery-backed cache, plus a 5U HP *Freighter* chassis holding a dense JBOD (Just a Bunch Of Disks) array of 70 SAS *nearline* disk drives. 12 more SAS *nearline* disks are mounted within the controller chassis.



Figure 5: HP ExDS storage chassis, front and rear views  
Generic images courtesy Tim Reddin, HP Galway

The disks are all dual-ported for high-availability (i.e. each disk has two SAS interfaces). The design has dual arrays, multiple I/O paths, and fault containment.

The underlying storage for a given filesystem is located on a single storage array. By default HP pre-configures each storage array with two filesystems, each of 32TB (29TB for users). All the disk partitioning was pre-created by HP with RAID6 plus spares, also the volumes and the global filesystems, as well as monitoring of the health of arrays, LUNs, disks, and their environment. All the elements are managed centrally: the servers, the 1Gbps and 10Gbps Ethernet switches, SAS switches, and the power and cooling.

The blade chassis also includes the main network switching, with four HP 10Gbps virtual-connect-Ethernet (*VC-Enet*) modules in two redundant pairs. These can present virtual host (vhost) IP addresses that are bound to a network interface on one of the hosts in the ExDS. They each have a 10Gbps CX4 copper uplink (i.e. there are two redundant pairs of uplinks). Unfortunately the 10Gbps ports of these switches exhibit errant behaviour that appears to be related to flow control. However, they also include 8 x 1Gbps ports (downlinks), which work fine, and can be bound together to act as a single multi-link high-throughput channel. This approach is being explored.

The ExDS rack also includes a slide-out keyboard/monitor/mouse (KVM) unit.

<b>Accession Index</b>	<b>Object and Identification</b>
TCD-SCSS-T.20121208.102.01	AN540A HP ExDS9100 Base Rack.
TCD-SCSS-T.20121208.102.02	AN543A HP ExDS9100 Capacity Block.
TCD-SCSS-T.20121208.102.03	AN543A HP ExDS9100 Capacity Block.
TCD-SCSS-T.20121208.102.04	AN543A HP ExDS9100 Capacity Block.
TCD-SCSS-T.20121208.102.05	AN543A HP ExDS9100 Capacity Block.
TCD-SCSS-T.20121208.102.06	AN541A HP ExDS9100 Performance Chassis.
TCD-SCSS-T.20121208.102.07	AN542A HP ExDS9100 Performance Block. (ProLiant BL460c blade with 16GB DRAM)
TCD-SCSS-T.20121208.102.08	AN542A HP ExDS9100 Performance Block. (ProLiant BL460c blade with 16GB DRAM)
TCD-SCSS-T.20121208.102.09	AN542A HP ExDS9100 Performance Block. (ProLiant BL460c blade with 16GB DRAM)
TCD-SCSS-T.20121208.102.10	AN542A HP ExDS9100 Performance Block. (ProLiant BL460c blade with 16GB DRAM)
TCD-SCSS-T.20121208.102.11	HP VC-Enet 10Gbps Virtual Connect Ethernet Module.
TCD-SCSS-T.20121208.102.12	HP VC-Enet 10Gbps Virtual Connect Ethernet Module.
TCD-SCSS-T.20121208.102.13	HP VC-Enet 10Gbps Virtual Connect Ethernet Module.
TCD-SCSS-T.20121208.102.14	HP VC-Enet 10Gbps Virtual Connect Ethernet Module.
TCD-SCSS-T.20121208.102.15	HA838A1 HP Factory Express HW Sys Custom SVC.
TCD-SCSS-T.20121208.102.16	HA839A1 HP Factory Express Svr Sys Custom SVC.
TCD-SCSS-T.20121208.102.17	AG053A HP TFT7600 Rackmount Kbd 17" UK Monitor .
TCD-SCSS-T.20121208.102.18	252663-833 HP 32A HV Core Only Corder PDU.
TCD-SCSS-T.20121208.102.19	HA864A1 HP Custom Placement within Rack Service.
TCD-SCSS-T.20121208.102.20	HA865A1 HP Rack Signal Cable Interconnect Svc.

## References:

1. Wikipedia: <http://en.wikipedia.org/>
2. Brian Coghlan, John Walsh, Stephen Childs, Geoff Quigley, David O'Callaghan, Gabriele Pierantoni, John Ryan, Neil Simon, Keith Rochford, *The Back-end of a 2-Layer Model for a Federated National Datastore for Academic Research VOs that Integrates EGEE Data Management*, Journal of Grid Computing, 8, (2), 2010, p341 – 364.
3. Tristan Suerink, NIKHEF, The Netherlands, personal email to Dr. Brian Coghlan, 14-Sep-2012 12:21. regarding HP 10Gbps VC-Enet switch problems, "I've tested the VC-Enet switch very extendedly. If you take a Brocade FCX-848 or an Juniper EX4200 switch and use that switch to land the 10G and then LACP multiple 1G's to the VC-Enet will solve the flow control problems of the VC-Enet".