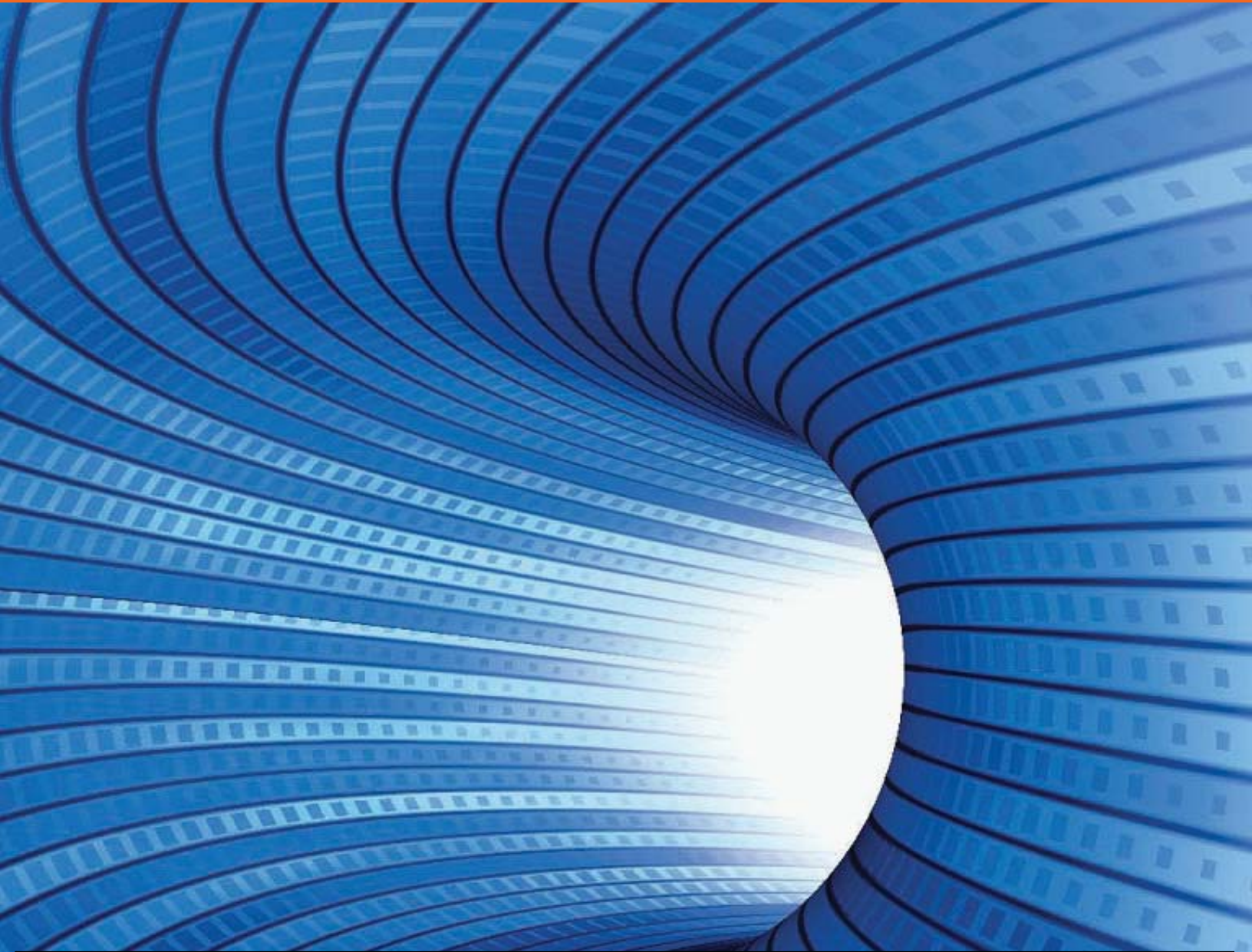


Data De-duplication/ Data Domain



real testing | real data | real results

Table of Contents

- 1. Introduction 1
- 2. Background..... 1
 - 2.1 Offered Capabilities..... 2
- 3. Key Evaluation Criteria 8
- 4. Evaluation Setup..... 9
- 5. Evaluation Results 10
 - 5.1 Evaluation Findings..... 10
 - 5.1.1 Software Installation and Setup 10
 - 5.1.2 Management 12
 - 5.1.3 Testing Results..... 14
 - 5.1.4 Functional Testing..... 15
- 6. Summary..... 18
- 7. Appendix A..... 19

High Performance Computing Modernization Program Air Force Research Lab DoD SuperComputing Resource Center

Data De-Duplication
Data Domain, an EMC Company

1. Introduction

The project will include an in-depth investigation of various data de-duplication (de-dupe) technologies that will identify the following: capabilities, user and center impacts, security issues, and inter-operability issues within a single location.

Data Domain systems are disk-based de-duplication appliances, arrays and gateways that provide data protection and disaster recovery (DR) in the enterprise.

All Data Domain systems run the Data Domain operating system (DD OS) derivative of Red Hat Linux, which provides both a command line interface (CLI) and the Enterprise Manager (a graphical user interface (GUI) for configuration and management.

A Data Domain system is designed to make backup data available with the performance and reliability of disks at a cost competitive with tape-based storage. Data integrity is assured with multiple levels of data checking and repair.

For the purposes of this investigation we will use the Data Domain System as an online storage device similar to what is used for home directories of the HPC users at the AFRL DSRC.

2. Background

The Department of Defense (DoD) High Performance Computing Modernization Program (HPCMP) has formed a Storage Initiative (SI) team to investigate the program's current storage architectures across the centers. Today, HPCMP is responsible for five major centers and two disaster recovery sites. One of the key areas of concern to the Storage Initiative team is the storing, managing and organizing of user data.

A data de-duplication solution could provide a cost savings in storage needs for user data files. The HPCMP Storage Initiative team partnered with the Data Intensive Computing Environment (DICE) Program Management Team to conduct a technical evaluation of the Data Domain product within a HPC environment to gain a better understanding of the functionality and integration requirements.

Starting with the first Data Domain Restorer shipped in 2003, Data Domain has helped customers reduce or eliminate the use of tape infrastructures with a very efficient disk and network-based data protection alternative. Today, Data Domain customers are expanding the use of their systems to include a broad range of near line workloads and use cases beyond data protection. In essence, Data Domain de-duplication storage represents a new generation of data versioning and replication efficiency. Data Domain is focused on bringing new storage and network efficiencies to the enterprise data storage arena on a single, integrated product platform to improve storage economics and simplify data management for its customers.¹

The Data Domain DD690 product was tested by the Data Intensive Computing Environment (DICE) Program Management Team to function in the HPC environment. This data de-duplication solution from Data Domain consists of a hardware storage appliance that front ends one or more Data Domain storage shelves.

¹ Source: www.datadomain.com/company

2.1 Offered Capabilities

Table 2.1 below describes the capabilities of the Data Domain DD690 product. The data is generally available on Data Domain's website (www.datadomain.com), through documentation or questions directed to Data Domain personnel.

Table 2.1 Capabilities Summary

Data Domain DD690

General	
Name & version of Data De-duplication software	Data Domain Operating System 4.6 (DD OS 4.6) (Red Hat Linux based)
General architecture	<p>Embedded OS Intel motherboard, based on Linux kernel.</p> <p>Data Domain uses the SISL™ (Stream-Informed Segment Layout) scaling architecture to minimize the number of disk accesses required in the de-duplication process and to perform an in-line de-duplication process.</p> <p>Data Domain employs its Data Invulnerability Architecture (DIA) for rigorous data protection and to ensure a “copy of last resort” capability.</p>
Can function in a heterogeneous environment?	Yes, the administration team can configure the DD690 to communicate remotely with any device on the network that can mount NFS or CIFS. VTL and NetBackup OpenStorage (OST) are also available.
Connectivity & Security	
File systems Supported	CIFS & NFS are supported by Data Domain.
OS Supported	Compatible OS to any OS that supports CIFS, NFS, VTL or OST.
Relational Databases Supported	Data Domain utilizes a proprietary database within the DD690 for metadata. No third party database needed.
Web Browsers Supported	Internet Explorer 6.0, on Windows XP Pro Microsoft Internet Explorer 7.0, on Windows XP Pro FireFox 2.0, on Windows XP Pro FireFox 2.0, on Linux
Archives Supported	<p>Data Domain has qualified its systems with leading archive applications including:</p> <ul style="list-style-type: none"> • CommVault Data Archiver • IBM Tivoli Storage Manager • Symantec Enterprise Vault • AXS-One Compliance Platform • Mimosa Systems NearPoint • The Kazeon Information Server^a platform

	<ul style="list-style-type: none"> Index Engines solutions F5 Acopia
Email support (syslogger, SNMP, email)	<p>SNMP and SMTP protocol</p> <p>All Data Domain systems have an automatic call-home system reporting capability, called Auto-Support, which provides email notification of complete system status. This non-intrusive alerting and data collection capability enables proactive support and service without administrator intervention, further simplifying ongoing management.</p>
Any future support for IPv6	The DataDomain Operating System DDOS currently supports kernel level pings of IPv6. The IPv6 software is present in DDOS, however it has not been enabled for general usage. In the future a method will be provided to configure IPv6 within DDOS. There are no current plans for complete support of IPv6 including configuration within the GUI or CLI.
Service and Ports Used	NFS v3 over TCP, Telnet, FTP, SSH, HTTP Default ports are used.
Can the system detect intrusion (unauthorized access to file) and react either automatically (log entry) or report to the site system administrator?	<p>Failed attempts to gain access to the appliance are logged.</p> <p>Example: Aug 17 09:06:52 ddrmc1-4 sshd[9186]: error: PAM: Authentication failure for sysadmin from 10.5.3.3</p>
Performance	
Available benchmark data	Official benchmarking is available from Data Domain if a Non Disclosure Agreement is signed. A throughput of 60-65 MB/s write and up to 95 MB/s read were achieved on a very busy platform. It had third party HSM software running and sharing platform resources that otherwise would not be recommended in production.
Throughput #'s for read/write transactions	DD690 can handle 60 simultaneous transactions with a max of 60 write transactions or 50 read transactions.
Determine maximum file size - can handle file sizes of 25 TB.	File size is limited to 288 PB currently. File size is only limited by the disk space available to store the files.
Scalability	
Optimizing scalability	<p>Up to 48 TB raw capacity</p> <ul style="list-style-type: none"> Up to six 8 TB expansion shelves Up to three 16 TB expansion shelves Support for a mix of 8 TB and 16 TB expansion shelves up to 48 TB raw capacity <p>Data Domain also has a product called the DD880 that scales to 96 TB.</p>

Reliability and Availability	
Single points of failure	Motherboard
Remote capabilities	Can be managed remotely with Telnet, SSH, and HTTP. E-mail alerts can be configured automatically.
Is there any mechanism for detecting data corruption?	In the background, the Online Verify operation continuously checks that data on the disk drives is correct and unchanged since the earlier validation process.
File data corruption must be reported, corrected and completely understood if it is the fault of the data de-duplication framework	All data written is read into cache and verified before final commit to disk. Any correctable errors encountered will be corrected but not logged. Only uncorrectable errors will be logged. The Data Domain Data Invulnerability Architecture provides the industry's best defense against data integrity issues. Continuous recovery verification along with extra levels of data protection detect and protect against data integrity issues during the initial backup and throughout the data lifecycle. Unlike any other enterprise array or file system, each appliance ensures recoverability is verified and then continuously re-verified.
Metadata corruption should be reported, corrected and completely understood if it is the fault of the data de-duplication framework?	When writing to disk, the DD OS creates and stores self-describing metadata for all data received. After writing the data to disk, the DD OS then creates metadata from the data on the disk and compares it to the original metadata. The correctable errors are corrected and not logged only uncorrectable errors are logged.
Does the system provide failover capabilities for each component?	The DD690 can withstand a dual drive failure, plus obtains redundant fans, power supplies, controllers, dual path. The ES20 expansion shelves offer the same redundancy. RAID6 with spares, dual fans, power and controllers.
Capacity Planning and Performance Analysis	
What tools are available to determine future capacity needs?	Logs and/or email alerts can provide valuable information regarding current metrics. This will enable the admin to fully understand and determine future capacity needs.
Does the DD690 have the capability to add capacity on demand?	Yes, Data Domain offers capacity on demand.
What is the complexity of adding or removing storage devices?	There are two types of Data Domain appliances: self contained storage and gateway storage. In both cases adding additional storage is straight forward. In the first case, an appliance that does not have the maximum number of expansion shelves can add a shelf. The additional shelf can

	<p>be added without shutting down the appliance with the “disk add enclosure command”. First, the shelf is installed in the same rack and SAS cables attached. Next, the “disk add enclosure” command is used to incorporate the new storage into the DDFS.</p> <p>In the second case, the gateway storage, (which was not included in our testing), additional LUNs are created in the storage array according to that vendors specifications. Next, the LUNs are discovered on the Data Domain with the “disk scan” command. Finally, the LUNs are added to the DDFS with the “disk add” command.</p> <p>In both cases, removing storage devices requires a re-installation of the DDOS.</p>
What is the complexity of migrating data from disk to tape?	This is done by third party ITSM software preferably configuring the Data Domain system as a Virtual Tape Library.
Are there tools to see how the DD690 is performing (transactions / minute, # of users / transaction, etc.)?	<p>There are two methods to gauge the current performance of the Data Domain appliance. First, with the GUI from the home page select “System Stats” and the GUI will launch a single page with 6 separate graphs indicating:</p> <ol style="list-style-type: none"> 1. CPU average % 2. Ethernet adapter KB/second 3. NFS recv % send % 4. Disk KiB/second read and write 5. Replication (if active) KB/s in and out 6. FS operations: NFS operations/second & CIFS operations/second <p>The second method is to use the CLI and the “system show performance” command. 1. See Appendix A</p>
Data De-Duplication Management	
Are ACL’s from HPC system managed separately?	For NFS, the ACL’s are managed through the HPC system. For CIFS the ACL’s are managed by Data Domain Appliance.
Is there a command line interface?	Yes, a command line management interface is available. Additionally, a Data Domain Enterprise Manager GUI and SNMP are available.
Is there a GUI provided for administration?	<p>Yes, but this feature does not have all the command line options.</p> <p>2. See Appendix A: Sample “show statistics” display from Enterprise Manager GUI.</p>
Is there a GUI provided for clients?	Not for NFS clients as tested
What types of reports are available for administrators?	Usage, Trends, Errors, etc. Administrators can setup alerts and add the appropriate users to be notified by email.

What documentation is provided for the installation and setup?	Data Domain provides PDF's for setup and configuration.
What are the staffing requirements?	After setup and configurations have been made there is minimal staffing requirement. Little time is necessary for upgrades and fixes to be applied.
Does the DD690 support Quotas? Hard? Soft?	Hard and soft quotas can be enabled on client systems that support NFS quotas.
Can I manage direct attached storage – delete, move, replicate or consolidate based on a specific action?	No, except for the Gateway product, Data Domain appliances only manage the expansion disk shelves attached to the appliance. Replication can be set up between 2 or more Data Domain appliances.
Diagnostic and Debugging Tools	
Success or error results from all operations, including those on remote systems, must be available to the administrator at the conclusion of the operation either through the functionality of the tool or some other work around (logged in log files, etc.).	The success or failure of each operation is clearly indicated for both the GUI and CLI methods of command execution. Logs are also available to the admin.
How are notifications sent (success or failure)?	For scheduled and unscheduled events, alerts can be reported daily through email and/or automatic email notice of hardware failure. Current and resent alerts are also displayed on the GUI displaying the alerts screen.
Is there any mechanism to detect failure due to access permissions prior to executing an operation?	Yes, the user can only see the commands available for that class of user. The help command or "?" will list all the available commands for that user. There are only 2 classes of users.
Are troubleshooting or diagnostic tools available?	Yes. All system messages are logged. These messages can be sent to a customer's log server with the "log host add" command. The messages are categorized by severity. The customer uses this message output to determine the situation and can consult the "Action" section for a suggested solution to the issue. A complete list of messages are included in each DDOS release's ErrorMessageCatalog.xml. 3. See Appendix A
Are there trend or pattern reports available on storage usage?	Yes. Usage, trends, errors, etc. can be configured by the administrator to send alerts. Simply add the appropriate users to be notified by email.
What documentation is provided for the hardware and software troubleshooting?	Data Domain offers training along with knowledge base for troubleshooting issues.
Service and Support	
Support Services Available	Remote monitoring available 24 x 7 x 365 on-site service. Multiple service levels for enterprise-class support are available.
How many service employees are employed?	1,000+
How often is software and firmware updated and who performs the upgrades?	Semi annually, remotely preformed by client or support team member from Data Domain.

Do software and firmware updates require an outage?	Yes. The DDOS file system is taken offline during the upgrade process. A reboot of the appliance is always required. The file system then comes online automatically.
Training & Professional Services	
Training Services Available	Data Domain offers online learning.
Cost	
What is the typical list price for your Data De-Duplication solution?	35¢/GB Acquisition
What are the ongoing costs over the life of your solution?	12% of acquisitions cost annually.
Vendor Information	
Are there any Department of Defense (DoD) references available?	Yes, Army IMCEN, Army INSCOM, HillAFB
Number of deployed sites?	100
Site Management	
The system must be able to provide troubleshooting tools to site level administrators (i.e., replicated site unavailable).	The Data Domain Enterprise Manager GUI allows users to choose the network for which a Data Domain system could be chosen for interaction and retrieve status.
The system should provide the ability to report trends or patterns back to each of the sites regarding storage usage for the system administrators to manage.	Available. This is done by using the autosupport function to generate reports configured by the administration team. Autosupport function can also be scheduled to run periodically and email the results.
Administrator Interface Tool	
Does the tool provide a method to compress files and directories?	Yes, The DD690 uses a local compression algorithm developed specifically to maximize throughput as data is written to disk. The default algorithm allows shorter backup windows for backup jobs, but uses more space. Local compression options allow you to choose slower performance that uses less space, or you can set the system for no local compression.
Does the tool provide a method to replicate files and directories?	Yes, replication is configurable between two or more Data Domain appliances. The Longest Replication - Hong Kong to DC 1500 ms latency limit.
Does the tool provide a method to encrypt files?	Not available with the DD690. With the Gateway version of the data domain appliance, a method is available to encrypt data at rest. A NetApp DataFort appliance is placed between the Data Domain's FC interface and the backend SAN. This seamlessly inserts a layer of strong encryption, authentication, access controls and compartmentalization between the Data Domain appliance and its backend disk storage.
Does the tool report specific types of files which are typically duplicated (unix file type – directory, link, etc.)?	Yes, The CLI interface will allow you determine the De-duplication rate of a file, directory or complete file system through the 'filesys show compression' command. Specific file type identification that is

	typically duplicated is not available. 4. See Appendix A
Does the tool report the number of files and their storage location in which de-duplication is most effective?	Not available on the Data Domain appliance but a 'What If' tool is available that provides an estimate of how well the data can be de-duplicated within the appliance.
Diagnostic tools specific to the data de-duplication system should be available to the administrator - determine such things as cache full.	Yes, logs and reports are available for review with troubleshooting. 5. List of logs in Appendix A
Hardware Interfaces	
The system should be able to manage direct attached system (HPC native files).	No, the DD690 is configured to use NFS/CIFS to connect with the appropriate pre-configured devices.
Communications Interfaces	
The system must interface with HPC system remotely.	Yes, access is available via HTTPS and SSH across the network as well as serial port connection.
The communication software must provide an understanding for future support on IPv6.	The DataDomain Operating System DDOS currently supports kernel level pings of IPv6. The IPv6 software is present in DDOS, however, it has not been enabled for general usage. In the future, a method will be provided to configure IPv6 within DDOS. There are no current plans for complete support of IPv6 including configuration within the GUI or CLI.
Hardware/Software Requirements	
Does the system function in a heterogeneous hardware environment (i.e., hardware agnostic)? It should not be dependent on what it can interface with.	Not the DD690. However, Data Domain offers a Gateway product line that adds De-duplication technology to some supported third party Fiber Channel storage.
The system must support Linux and Unix operating system(s) – POSIX compliant.	Yes
Operational Requirements	
Data should retain its normal structure in order to maintain interoperability with other systems.	Yes, when data is moved or copied to the DD690 a de-duplication process is executed. If the data is accessed on that file system it will be restored to its original structure.
De-duplication should have minimal impact on additional storage.	Yes, some latency may occur during the defragmentation. Additional latency can occur during high thread usage.
Audit Trail	
Does the system keep an audit trail of administrator transactions?	Yes, logs are available for review.

3. Key Evaluation Criteria

The evaluation assessed:

1. The ability to administer systems, files and directories
2. The ability for the software to interface with HPC hardware (HSM, HPC, Tape libraries, etc.)
3. The security and privacy of other users – permission management
4. The ability to report storage gain / loss from eliminating redundancy

5. The ability to replicate and encrypt data
6. The flexibility of command line interface versus Graphical User Interface (GUI)

4. Evaluation Setup

The Data Domain DD690 has many possibilities for configuration based on storage needs. These needs vary from a Virtual Tape Library to an intermediate staging before migration to tape or near line storage. In this situation, the DD690 will act as the network file server for online storage of home directories. Software features include: Global Compression, Data Invulnerability Architecture including end-to-end verification (ongoing) and integrated dual disk parity RAID-6, snapshots, telnet, FTP, SSH, email alerts, scheduled capacity reclamation, Ethernet failover and aggregation, Data Domain OpenStorage, Replicator and Retention Lock optional software.

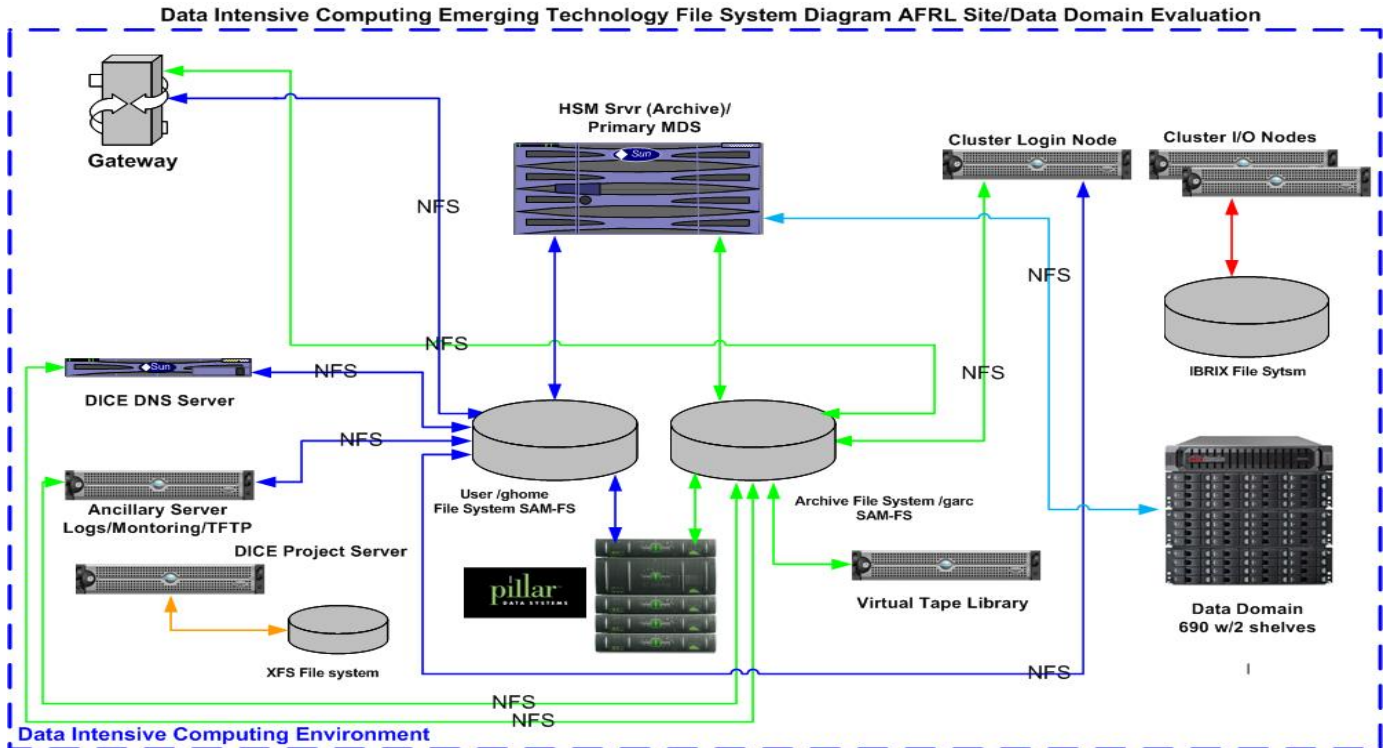
The DD690 was setup within the DICE environment at AFRL and configured to provide the inline de-duplication process and storage. The DD690 was configured as the NFS server to Solaris and Linux NFS clients. The DD690 configuration tested has 2 expansion shelves each with sixteen 1 TB SATA Disk Drives for total raw capacity of 32 TB. Each shelf has a RAID 6 configuration with 2 hot spares, which created a single file system approximately 21 TB in size.

On behalf of AFRL DSRC, the following unclassified data was researched and tested to perform a de-duplication:

- Scientific benchmark data from AFRL DSRC
- Scientific data from Artic Region Super Computing Center (ARSC)
- Scientific weather data from National Oceanic and Atmospheric Administration (NOAA)

The Data Domain DD690 was configured on the public network for NFS and a separate private network for management.

Figure 1 below depicts the solution setup used at the AFRL site.



5. Evaluation Results

5.1 Evaluation Findings

5.1.1 Software Installation and Setup

System Setup

Data Domain along with the DICE team installed the DD690 with two shelves containing 32 GB of raw SATA disk capacity. The hardware platform consisted of the following hardware:

- 2U rack mountable, DD690 controller redundant fans, N+1 power supplies, serial port, 2 copper 10/100/1000 Ethernet ports and optional dual port copper or optical 1Gb Ethernet and dual port copper or single port optical 10 Gb Ethernet
- 2 3U ES20 Disk Shelves with redundant power/cooling units and 16 1 TB SATA Disk Drives in each shelf.

The following information is required for a Data Domain System installation:

- Fully-qualified host name for the system
- The DNS domain name
- For each Ethernet interface
 - Is it using DHCP?
 - If not using DHCP interfaces IP address and Netmask
- The routing gateway IP address
- The DNS server IP address

- An administrator password – initial password based on S/N of device.
- Hostname (or * for all) from which to administer the system
- An administrator's email address (or admin group alias)
- A description of the system's physical location (Optional)
- A mail server (SMTP) hostname
- A time zone (default is US/Pacific)

After the system was racked and powered, configuration of the system was done via a serial console at 9600 baud, 8 data bits, no parity and 1 stop bit. A configuration wizard was used to complete the following tasks:

- License the software
- Configure the Network File System
- NFS Clients allowed; A Solaris and Linux systems were selected as NFS clients for this device.

All systems tested in the DICE environment have to be examined for security vulnerabilities that may exist in the hardware or software. In this case, the scanning of the DD690 from the network revealed down rev versions of OpenSSH, Apache, and other services within the DDOS software. As a result those services had to be turned off in order for network access and testing. Once those services were turned of, all administration had to be conducted onsite through serial port.

NFS was configured for clients to access /backup on the DD system. The following steps provide an example of client configuration.

1. Create a mount point (directory) such as /dd/nfsdd01/backup and create an administrative mount point, such as /dd/nfsdd01/ddvar for each client.

2. test mounts:

```
mount -F nfs -o hard,intr,vers=3,proto=tcp,rsize=32768,wsiz=32768
nfsdd01:/backup /gdatadup
```

```
mount -F nfs -o hard,intr,vers=3,proto=tcp,rsize=32768,wsiz=32768
nfsdd01:/ddvar /ddvar
```

3. Automate mounts: Add the following lines to the file /etc/vfstab. This will mount the directories upon every reboot.

```
system:/backup - /gdatadup nfs - yes hard,intr,vers=3,proto=tcp,
size=32768,wsiz=32768
```

```
system:/ddvar - /ddvar nfs - yes hard,intr,vers=3,proto=tcp,
rsize=32768,wsiz=32768
```

Additional server configuration options:

Give access to additional backup servers

```
# nfs add /backup {*|client-list}
[options]
```

Add users to the email list that reports system problems

```
# alerts add email-list
```

Add users to the system report email list

```
# autosupport add email-list
```

From a remote machine, add an authorized SSH public key

```
# ssh-keygen -d
```

```

# ssh -l sysadmin rstr01 "adminaccess add ssh-keys"\
< ~/.ssh/id_dsa.pub
Enable FTP or TELNET (SSH is enabled by default)
# adminaccess enable {ftp|telnet|ssh}
Add remote hosts to use FTP or Telnet
# adminaccess add {ftp|telnet}
fqdn-host-list
Add a user
# user add name [priv {admin|user}]
Change a user's password
# user change password username

```

User Setup

A Data Domain system has two classes of user accounts:

- The *user* class is for standard users who have access to a limited number of commands. Most of the user commands can only display information.
- The *admin* class is administrative users who have access to all Data Domain system commands. The default administrative account is *sysadmin*. You can change the *sysadmin* password, but cannot delete the account.

Three user accounts were created for testing purposes on the DD690:

- *sysadmin* – used for backup in case dice-admin password is lost (recommended action)
- *dice-admin* – used for administrative purposes
- *dice-user* – used for second tier admin i.e. backup admin

Problems Encountered / Resolution

Security guidelines set forth by AFRL required all network services to be up to date on known security issues. In order for a product to be used in a proof of concept or production a security advisory scan is performed against the product. This security scanning of the DD690 appliance revealed older versions of Apache, OpenSSH, OpenSSL, FTP, and Telnet that have known security issues. As a result of these findings, those services were disabled in order to be placed on the AFRL network. This also resulted in all administration of the DD690 appliance occurring over the serial console on site.

5.1.2 Management

Client

The client (end-user) will see no changes in their daily operations unless the network becomes bogged down causing latency in data delivery. Clients will continue to send data to their home directories and the data will be de-duplicated within the DD690.

Administration

The DD690 allows administration through a web-based GUI called Enterprise Manager and through the CLI. The Data Domain system CLI allows initial system configuration, changes to individual system settings, and display system and operation status. Use the

Enterprise Manager to perform initial system configuration, make some configuration updates after initial configuration, and display system and component status as well as the state of system operations.

Due to the down rev OpenSSL and OpenSSH on the DDOS software appliance, the web interface as well as the SSH network access were disabled, thus limiting administration to only the CLI through the serial port. The command line interface is available through a serial port or keyboard and monitor attached directly to the Data Domain system. As a result of this management access limitation, all administration was conducted only on-site, and no evaluation of the GUI was performed.

The CLI was tested for configuration, user management, space management, system monitoring, log file maintenance, alerts, reporting and some disk management. It contained a very useful help command and everything performed as expected and was very informative.

Server Operation

De-duplication and Compression

The DD690 appliance uses global and local compression to help reduce the storage size of files. Data Domain uses the term “global compression” for the de-duplication. A local compression is then applied on top of the global compression to further reduce the storage space needed to store the file. There are four types of local compression algorithms available for use. These algorithms are developed specifically to maximize throughput as data is written to disk. The default algorithm allows shorter backup windows for backup jobs, but uses more space. Other options allow you to choose slower performance but less space, or you can set the system for no local compression.

- lz - The default algorithm that gives the best throughput. Data Domain recommends the lz option.
- gzfast - A zip-style compression that uses less space for compressed data, but more CPU cycles. Gzfast is the recommended alternative for sites that want more compression at the cost of lower performance.
- gz - A zip-style compression that uses the least amount of space for data storage (10% to 20% less than lz), but also uses the most CPU cycles (up to twice as many as lz).
- none - No data compression.

Changing the algorithm on the active file system will affect only new data being moved into the appliance. The current data remains as is until a clean operation checks the data. In the process of changing local compression algorithm, the file system will need to be disabled and then enabled again. Doing this while the clients have the NFS mount pount active did not seem to cause any problem for the client.

Once data is deleted on the DD690 appliance it becomes inaccessible but remains on the disk until a clean operation is executed. The clean operation reclaims physical storage and can be run at a scheduled point in time either daily or monthly. Although

cleaning uses a noticeable amount of system resources, cleaning is self-throttling and gives up system resources in the presence of user traffic.

Other features

The DD690 appliance uses an append-only write policy that guards against overwriting valid data. Continuously in the background, the *Online Verify* operation checks that data on the disks is correct and unchanged since the earlier validation process.

Also available is a feature called autosupport. The autosupport feature sends a daily report that shows system identification information and consolidates the output from a number of Data Domain system commands. This can be configured to be mailed to the administrators and even to Data Domain support for monitoring. The default autosupport output is a lengthy report and can take considerable time to execute given that you have many millions of files stored.

5.1.3 Testing Results

De-duplication and Compression Results

The de-duplication testing involved three unclassified data sets as mentioned in the setup above. These data sets were copied onto the de-duplication NFS mount point using the 3 separate Unix tar commands in the following consecutive order:

- Scientific benchmark data from AFRL (648 MB in size)
- Scientific weather data from NOAA (20 GB in size)
- Scientific data from Artic Region Super Computing Center (424 GB in size)

After all the data sets were copied to the NFS file system and the resulting file sizes were noted, the files were deleted and a clean operation was performed. Then the local-compression algorithm was changed and the test was run with that new compression algorithm.

Table 5.1.3 De-duplication and Compression Results

	No Compression (De-duplication only)	lz compression	Gzfast compression	gz compression
AFRL data	1.44 to 1	3.04 to 1	4.7 to 1	5.7 to 1
NOAA data	1.10 to 1	1.31 to 1	1.47 to 1	1.48 to 1
ARSC data	1.49 to 1	1.65 to 1	1.91 to 1	1.92 to 1

5.1.4 Functional Testing

Table 5.1.4 below describes the test requirement, evaluation results and evaluation ranking for the critical and high requirements defined by the Storage Initiative team.

The following criteria are used for the evaluation ranking:

Met	The solution offered the minimum functionality.
Surpass	The solution offered more than expected functionality.
Missed	The solution offered less than minimum functionality.

Requirement	Evaluation Results	Evaluation Ranking
Functionality		
Attribute Requirements		
The system must provide a method to recover a file, directory, or metadata that has been identified as corrupt.	The system provides continuous data verification with automatic recovery via the Online Verify function.	Met
Interfaces		
Interface Tool (Only the CLI in this case)		
The tool must be able to perform configuration modifications (add systems, files, directories).	The tool configures which clients have permission to mount the DD690 NFS exported file systems.	Met
The tool must be able to modify specifics regarding files and directories to perform de-duplication (eliminate certain files/directories).	Once the NFS file system is mounted, users can create their own sub-directory, create, edit, or copy files into the directory. Once the files are stored to NFS de-dupe mount point the files will be de-duplicated.	Met
The tool must provide the ability to un-de-duplicate files and directories.	Copying a file from the NFS de-dupe mount point to a different working directory invokes files to be un-de-duplicated.	Met
The tool must be able to operate on single file/directory, lists of files/directories or directory tree.	Using the standard HPC Unix tools to copy single or multiple files including recursive options that work on directories can be acted upon.	Met
The tool must be able to report percentage of storage gained from eliminating redundancy.	Using the CLI command, "fileysys show compression /mount point /file name" outputs the compression ratio, original file size, compression file size and metadata size. This function can work with a directory as well.	Met
Hardware Interfaces		
Communications Interfaces		
The communication software must be configured to run on IPv4.	The system was configured using IPv4.	Met
Hardware / Software Requirements		
The system must be able to have Data De-dupe actions available from the HPC file systems.	De-dupe actions are available when using Unix commands to move or copy files from HPC file systems to de-dupe NFS mount point on HPC client.	Met

The system must interface with a remote or network file system link to HPC systems (like NFS) or a direct attached file system with a client.	NFS was the interface tested within the DICE environment.	Met
The system must support Linux and Unix operating system(s) – POSIX compliant.	The system was tested using Solaris and Linux.	Met
The solution must be commercially available at the time of the installation.	The software is currently available for purchase today.	Met
The metadata system must provide failover and/or recovery capabilities.	As noted in the Data Domain documentation, the system as a whole can provide failover/recovery capability when configured as such.	Met
Operational		
The data portion of all duplicate files flagged must be identical to the original file.	A file originating in the home directory was copied to the NFS De-dupe file system and then copied back to a working directory. A diff command then showed that the original file was identical to the copy in the working directory.	Met
Determine the limitation in the number of files, where metadata is stored, and hash table growth.	<p>Testing successfully created 20 million files into the NFS De-dupe filesystem.</p> <p>Data Domain recommends storing no more than 100 million files on a system. A larger number of files effects performance, but is not a problem otherwise. Some processes, such as file system cleaning, may run much longer with a very large number of files. For example, the enumeration phase of cleaning takes about 5 minutes for one million files and over 8 hours for 100 million files.</p> <p>The system does not have a set number of files as a capacity limit. Available disk space is used as needed to store data and the metadata that describes files and directories. In round numbers, each file or directory has about 1000 bytes of metadata.</p> <p>A Data Domain system with 5 TB of space available could hold about 5 billion <i>empty</i> files. The amount of space used by data in files directly reduces the amount of space available for metadata, and the number of file and directory metadata entries directly reduces the amount of space available</p>	Met

	for data.	
Security and Privacy		
The Data De-dupe administrator (if not system administrator) must be able to perform functions without having system level root privileges.	The Data De-dupe administrator is its own user type defined only on the de-duplication hardware.	Met
The system must be able to detect data corruption and react either automatically or report to the site system administrator.	The system continuously does automatic detection for corruption. Problems are logged and can be reported through e-mail.	Met
The system must pass CSA scans.	CSA scans discovered that the current versions of OpenSSH, OpenSSL, telnet, FTP, and Apache installed on DDOS 4.6 were down rev and contained known vulnerabilities. All of those services were disabled for testing.	Missed
The system must provide the ability that all operations are subject to access permissions, authorizations of target objects, and user privileges on accounts for all systems involved in any operation.	The system follows client NFS ACL privileges.	Met
<p>System functions requiring elevated privilege must be properly documented to allow understanding and limitation of the risks.</p> <p>System configurations must meet the DoD ports and protocols guidance and management processes. Static passwords are not permitted by DoD.</p> <p>One of the following three methods must be met to provide/demonstrate proper security protections:</p> <p>HPCMP prefers Kerberos protected services to secure information transfer and communications along with SecureID or Common Access Card/Public Key Infrastructure (CAC/PKI) for single sign on authentication.</p> <p>or</p>	The DDOS does it have Kerberos or CAC/PKI capabilities. The DDOS also does not have automatic expiration of passwords.	Missed

<p>Independent Certification: Meet the requirements set forth by NAIP CCEFS documentation.</p> <p>http://www.niap-ccevs.org/cc-scheme/pp/pp_os_ca_v1.d.pdf</p> <p>or</p> <p>Systems that do not use Kerberized services and/or SecureID/PKI authentication must be documented and approved to operate in the HPCMP architecture (prior to installation) according to the HPCMP Access Guidelines.</p>		
<p>Assurance that the software is properly using and protecting those privileged actions and credentials is required.</p>	<p>The privileged de-dupe admin commands cannot be accessed or seen by any general user.</p>	<p>Met</p>
<p>All communications between services must be properly authenticated and protected from intrusion (i.e., replicate over WAN).</p>	<p>NFS and CIFS are the only service available between client and de-dupe server. They follow the proper authentication standards.</p>	<p>Met</p>
<p>Audit Trail</p>		
<p>The system must keep log files on Data De-dupe server.</p>	<p>The administrator has access to log files for audit proposes.</p>	<p>Met</p>
<p>General Performance</p>		
<p>The system must be able to handle multiple simultaneous transactions without effecting performance (single stream vs. multiple stream)</p>	<p>The system handled 60 write streams or 50 read streams during testing. Some delay was encountered while doing ls -l in home directory. (A 23-second delay was noted during testing.)</p>	<p>Met</p>
<p>The system must be able to handle file sizes of 2 TB.</p>	<p>Tested successfully using Unix commands</p>	<p>Met</p>

6. Summary

The Data Domain DD690 has proven to be a great tool in dealing with the problem of out of control data growth. In this test case scenario it really comes down to the datasets and how well they can be de-duplicated. In this study, DICE has found that the scientific data sets (defined in [section 4](#)) did not de-dupe as well as desired. Although the DD690 proved to be a powerful tool the datasets continue to have a low de-duplication factor. When tested with Data Domains three different local compression algorithms in addition to the de-duplication, the data showed improvement. However, similar compression

algorithms are available today for many Unix flavors. For a summary on the different compression ratios and test results view [De-duplication and Compression](#) under Testing Results.

Two key requirements were also missed in the testing. The ability to adapt CAC/PKI or Kerberos logins or automatic password expiration and the security advisory scans revealed vulnerable services due to older revisions of open source software. Upgrades to the DDOS could be performed in order to update the out of date software but are currently not on short term roadmap from Data Domain.

Overall testing of the Data Domain DD690 appliance and its functionality proved successful. The throughput performance of the DD690 was very good and could exceed the supported streaming limits.

For storing data, Data Domain recommends no more than 100 million files on a system. A larger number of files effects performance, but is not a problem otherwise. Some processes, such as file system cleaning, may run much longer with a very large number of files. For example, the enumeration phase of cleaning takes about 5 minutes for one million files and over 8 hours for 100 million files.

In testing, the DICE team created over 20 million files on the Data Domain 690, most of which were in the same directory. The DICE team does not recommend creating a single directory with a very large number of files. It will effect system performance for the simple Unix command such as ls and ls -l. The system does not have a set number of files as a capacity limit. Available disk space is used as needed to store data and the metadata that describes files and directories. In round numbers, each file or directory has about 1000 bytes of metadata.

A Data Domain system with 5 TB of space available could hold about 5 billion *empty* files. The amount of space used by data in files directly reduces the amount of space available for metadata, and the number of file and directory metadata entries directly reduces the amount of space available for data.

7. Appendix A

Sample:

1. Are there tools to see how the DD690 is performing (transactions / minute, # of users / transaction, etc.)?

Sample output:

```
-----Throughput (MB/s)-----  Utilization----  Compression  -----Cache Miss-----  --Streams--  State-  ----Utilization----  --Latency--
Date   Time   Read Write Repl Network Repl Pre-comp  proc recv send idle  gcomp lcomp  thra unus ovhd data meta rd/wr/rw/xx 'CDBVMS'  CPU   disk   in ms
-----in/out-----in/out-----
2009/08/26 13:30:14  79.7  5.3  0.00/ 0.00  0.00/ 0.00  77% 10% 10%  1% 407.5  1.0  0%  0%  0%  1%  0%  50/ 1/ 0/ 1  ----- 16%/ 20%[0]  5%[13]  8.7/ 2.5
2009/08/26 13:40:14  90.4  6.1  0.00/ 0.00  0.00/ 0.00  72% 12% 12%  2% 409.7  1.0  0%  0%  0%  0%  0%  50/ 1/ 0/ 1  ----- 18%/ 21%[0]  5%[13]  5.8/ 1.4
2009/08/26 13:50:14  95.9  7.1  0.00/ 0.00  0.00/ 0.00  70% 14% 13%  1% 416.8  1.0  0%  0%  0%  0%  0%  50/ 1/ 0/ 0  ----- 18%/ 22%[0]  5%[04]  4.6/ 1
```

2. Sample “show statistics” display from Enterprise Manager GUI.



3. Are troubleshooting or diagnostic tools available?

Here is a sample message with its corresponding description and suggested action to take to remediate the situation.

ID: MSG-CM-0004 - Severity: CRITICAL - Audience: customer

Message: Container set %1: File deletes disabled.

Description: The restorer cannot delete files from container set %1 because deletions require more than the available metadata space.

Action: Run cleaning to recover space. If necessary, delete snapshots to free space before cleaning.

4. filesys show compression output

```
dice-admin@dd690# filesys show compression /backup/dicerab/ARSCgz.tar
Total files: 1; bytes/storage_used: 2.0
  Original Bytes:      423,877,843,112
  Globally Compressed: 278,137,158,996
  Locally Compressed: 215,499,350,544
  Meta-data:          924,869,328
```

5. log list output

```
dice-admin@dd690# log list
Last modified      Size      File
-----
Sun Aug  2 00:45:01 2009  0 KB     access_log (empty)
Sun Jul 26 00:45:01 2009  0 KB     access_log.1 (empty)
Thu Jul 23 11:39:15 2009  79 KB    access_log.2
Thu Jul  9 17:31:54 2009  0 KB     access_log.3 (empty)
```

Tue Aug 4 12:10:01 2009 5 KB bios.txt
 Sun Aug 2 00:45:01 2009 0 KB boot.log (empty)
 Sun Jul 26 00:45:01 2009 0 KB boot.log.1 (empty)
 Wed Jul 22 21:45:01 2009 0 KB boot.log.2 (empty)
 Thu Jul 9 17:31:22 2009 0 KB boot.log.3 (empty)
 Tue Aug 4 12:06:55 2009 0 KB boot_fsck.err.log (empty)
 Tue Aug 4 12:06:55 2009 0 KB boot_fsck.log
 Tue Aug 4 12:04:25 2009 141 KB bootlog.txt
 Wed Aug 5 15:11:56 2009 445 KB ddfs.info
 Wed Aug 5 17:33:21 2009 4620 KB ddfs.memstat
 Tue Aug 4 07:56:00 2009 4620 KB ddfs.memstat.1
 Thu Jul 23 10:50:44 2009 4620 KB ddfs.memstat.2
 Wed Aug 5 17:34:41 2009 6210 KB ddsh.info
 Thu Jul 9 16:56:09 2009 0 KB destroy.20277.log
 Thu Jul 23 14:36:33 2009 37 KB destroy.31047.log
 Mon Aug 3 00:54:26 2009 13 KB disk-error-log.out
 Tue Aug 4 12:10:18 2009 60 KB dply_log
 Thu Jul 23 14:39:38 2009 38 KB dply_sfdisk_log
 Thu Jul 9 17:32:25 2009 0 KB em.info (empty)
 Thu Jul 9 17:32:23 2009 0 KB em_error.info (empty)
 Thu Jul 9 17:31:22 2009 0 KB enc (empty)
 Sun Aug 2 00:45:01 2009 0 KB error_log (empty)
 Sun Jul 26 00:45:01 2009 0 KB error_log.1 (empty)
 Thu Jul 23 12:46:48 2009 62 KB error_log.2
 Wed Jul 22 21:36:25 2009 0 KB error_log.3
 Tue Aug 4 12:07:24 2009 0 KB glibc.log
 Tue Aug 4 12:08:24 2009 10 KB kern.error
 Tue Aug 4 12:10:18 2009 1368 KB kern.info
 Wed Aug 5 17:30:01 2009 4469 KB memory_usage.log
 Sun Aug 2 00:45:01 2009 8966 KB memory_usage.log.1
 Sun Jul 26 00:45:01 2009 4174 KB memory_usage.log.2
 Wed Jul 22 22:45:01 2009 54 KB memory_usage.log.3
 Wed Aug 5 17:34:41 2009 74 KB messages
 Sun Aug 2 00:00:02 2009 124 KB messages.1
 Sun Jul 26 00:00:02 2009 86 KB messages.2
 Wed Jul 22 21:42:27 2009 11 KB messages.3
 Wed Aug 5 17:33:54 2009 668 KB messages.engineering
 Wed Aug 5 17:33:54 2009 307 KB messages.support
 Wed Aug 5 17:34:40 2009 9498 KB perf.log
 Sun Aug 2 00:45:01 2009 19406 KB perf.log.1
 Sun Jul 26 00:45:01 2009 6997 KB perf.log.2
 Thu Jul 23 14:45:01 2009 19 KB perf.log.3
 Wed Aug 5 17:34:23 2009 629 KB platd_log
 Tue Aug 4 12:07:20 2009 0 KB platd_sem (empty)
 Tue Aug 4 12:07:20 2009 0 KB platd_state (empty)
 Tue Aug 4 12:07:20 2009 0 KB platd_state_change_sem (empty)
 Wed Aug 5 17:31:01 2009 1063 KB qos.info
 Wed Aug 5 09:45:03 2009 1902 KB qos.info.1.gz
 Tue Aug 4 12:07:22 2009 1 KB rc.sysinit.log
 Wed Aug 5 17:33:54 2009 0 KB secure.log
 Sun Jul 26 00:45:01 2009 0 KB secure.log.1 (empty)
 Thu Jul 23 11:34:43 2009 3 KB secure.log.2
 Wed Jul 22 21:40:53 2009 3 KB secure.log.3
 Tue Aug 4 12:07:20 2009 0 KB sem (empty)
 Wed Aug 5 17:34:41 2009 5002 KB sms.info
 Wed Aug 5 00:45:01 2009 10394 KB sms.info.1

```

Mon Aug 3 09:45:01 2009 10520 KB sms.info.2
Sat Aug 1 22:45:01 2009 10508 KB sms.info.3
Fri Jul 31 11:45:01 2009 10530 KB sms.info.4
Thu Jul 30 00:45:02 2009 10510 KB sms.info.5
Tue Jul 28 13:45:01 2009 10516 KB sms.info.6
Mon Jul 27 02:45:01 2009 10396 KB sms.info.7
Wed Aug 5 17:00:01 2009 63 KB space.log
Tue Aug 4 12:08:52 2009 1 KB upgrade_log
Thu Jul 23 11:34:32 2009 0 KB windows/0.0.0.0.log
Thu Jul 23 11:39:15 2009 44 KB windows/clients.log
Thu Jul 23 11:19:37 2009 1363 KB windows/log.nmbd
Thu Jul 23 13:53:50 2009 0 KB windows/log.smbd
Thu Jul 23 13:47:32 2009 0 KB windows/nmbd.log
Thu Jul 23 13:53:50 2009 0 KB windows/smbd.log
-----

```

Lucas Arts requested that this undocumented command be added to the DDOS after the successful proof of concept.

```
dice-admin@dd690# system show droid
```

