Intel® Cloud Builders Guide: Cloud Design and Deployment on Intel® Platforms

Scale-out Storage with EMC* Atmos*

Intel® Xeon® Processor 5500 Series
Intel® Xeon® Processor 5600 Series

where information lives

AUDIENCE AND PURPOSE

For companies who are looking to build their own cloud computing infrastructure, including both Enterprise IT organizations and Cloud Service Providers or Cloud Hosting Providers, the knowledge and experience gained from previous work will help facilitate the decision to use a cloud for delivery of IT services. This reference architecture gathers into one place the essentials of a scale-out storage cloud architecture based on EMC* Atmos* cloud-optimized storage. This reference architecture, based on Intel® Xeon® servers, creates a multi-site, capacity-optimized cloud storage deployment. This paper contains details on: the cloud topology; hardware and software deployed; installation and configuration steps; and tests for real-world use cases that should significantly reduce the learning curve as you build and operate your first cloud infrastructure.

The creation and operation of cloud storage requires significant integration and customization based on existing IT infrastructure and business requirements. As a result, we do not expect that configurations described in this paper can be used "as-is." For example, adaptation to an existing network and identification management requirements are out of scope for this paper. Therefore, we anticipate that the user of this paper will make significant adjustments to the design we present in order to meet specific requirements.

This paper also assumes that the reader has basic knowledge of cloud storage infrastructure components and services.

Table Of Contents

Executive Summary

This paper describes the architecture and implementation details of a small scale-out storage (SOS) cloud solution built jointly by Intel and EMC to demonstrate a private cloud storage deployment across two geographically separate data centers. We built the scale-out storage cloud solution with the EMC Atmos cloud-optimized storage solution with Dell* servers based on Intel® Xeon® technology. In addition to the standard functionality expected from a cloud storage solution such as local and remote storage access interfaces, the dynamic provisioning of storage and storage users, and low power / high capacity storage scaling to petabytes (PBs) and beyond, the solution offers a complete set of tools to automate management of the cloud storage. These tools include a high availability metadata service that supports user-defined policies and self-management of storage placement, reliability, compression, and deduplication. This paper illustrates the cloud storage functionality built and tested over a two-month period with several real-world use cases, such as server backup and automated file archive.

Introduction

The emergence of cloud computing and the explosion of digital content has driven the development of capacity-based, scale-out storage cloud architecture. Figure 1 shows one estimate of the projected growth of digital content that all devices will generate over the next 10 years. The expectation is that digital content growth will approximately double every year. One element that enables this rapid growth is a reduction in the investment per gigabyte to store the data. As much as 15% of the information in the Digital Universe in 2020 could be part of a cloud service.[1] As a result, cloud storage investment needs to decrease at the same rate to make the storage efficiently support the massive projected growth rate.

To achieve the best efficiency, internet portals that optimize for both cloud computing and digital content have created cloud storage architectures based on industry standard x86 servers with directly attached disks. This implementation of storage is commonly referred to as scale-out storage. Portal applications utilize scale-out storage for a number of purposes such as for search indexes and to preview content, social networking information, videos, photos, thumbnails, and standard office documents (for example, pdf, Word*, PowerPoint*). Optimization of storage to achieve the best efficiency is especially significant for portal applications like YouTube*, Flickr*, and Facebook* that offer free storage of consumer video and photo content. Enterprises strive to achieve a similar efficiency as they deploy scale-out storage as repositories for data and documents stored in applications like Microsoft SharePoint* and EMC* Documentum*.[2]

This paper defines a scale-out storage cloud reference architecture based on EMC Atmos. The reference architecture highlights EMC Atmos[3] deployed as a single private cloud storage across the Intel® Folsom and Portland cloud labs. In addition, a number of real-world use cases are validated to highlight the features of the EMC Atmos implementation.

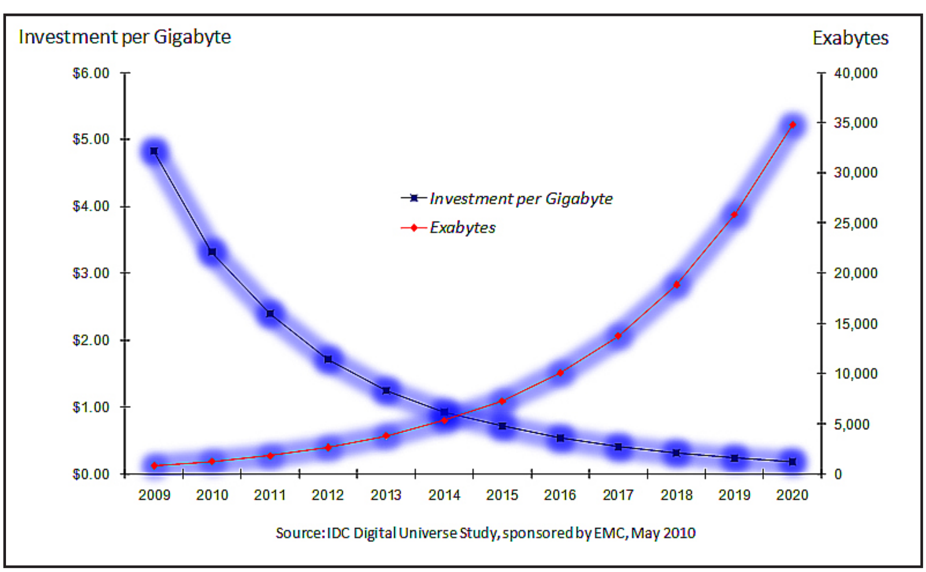

Figure 1: Digital Content Creation Estimates

Scale-out Storage Usage Models

Backup/Archive

As consumer and enterprise data storage needs have risen, a low-cost and high-throughput means to back up and retrieve computer files has become critical. Many cloud storage Software as a Service (SaaS) solutions exist to support consumer PC, smart phone, and tablet backup. These SaaS solutions also support PC and server backup for small and medium business. Enterprises typically deploy backups as a privately managed service which directly connects to private storage based SAN or NAS, or directly to tape. To achieve better efficiency, enterprises now move to replace these solutions with scale-out storage cloud architecture.

IT typically builds backup systems around a client software program that runs on a schedule, typically once a day. For higher reliability, you can also implement backups as continuous data protection. The backup program collects, compresses, encrypts, and transfers the data to the cloud storage—a private cloud, a SaaS cloud storage, or first to a private cloud and then to the SaaS cloud.

Large Object Store

The biggest growth of data in the last ten years has been semi-structured large objects. Photo, video, thumbnails, and documents (for example, pdf, Word, PowerPoint) are the most prevalent examples. The challenge is especially significant for portal applications like YouTube, Flickr, and Facebook that store consumer video and photo content. Enterprises have a similar requirement to store documents with applications like Microsoft SharePoint and EMC Documentum.

Large-Scale Data Warehousing and Analytics Store

Enterprises can analyze data collected as part of their business operations through online databases (business to business and business to cloud), log files, sensors, and general documents. We typically refer to this collection of data as a data warehouse. Scale-out storage is an optimal architecture to both cost effectively store the data, and at the same time make it available for data analytics. The Apache* Hadoop[4] application framework is commonly used on a scale-out storage cloud to perform the data transformation necessary to create the analytics database.

Usage Model Overview in this Paper

This paper only focuses on configuration and testing of backup and archive usage models. Four applications were selected based on discussions with Intel IT on use of a scale-out storage cloud for backup and archiving. The four applications tested are described in the EMC Atmos scale-out storage tests section.

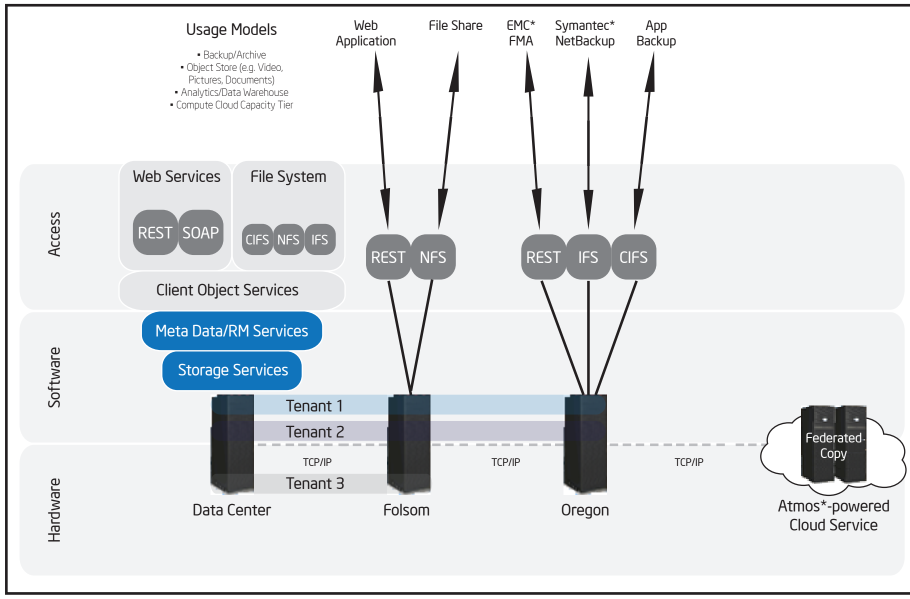

Figure 2: EMC Atmos Scale-out Storage Cloud

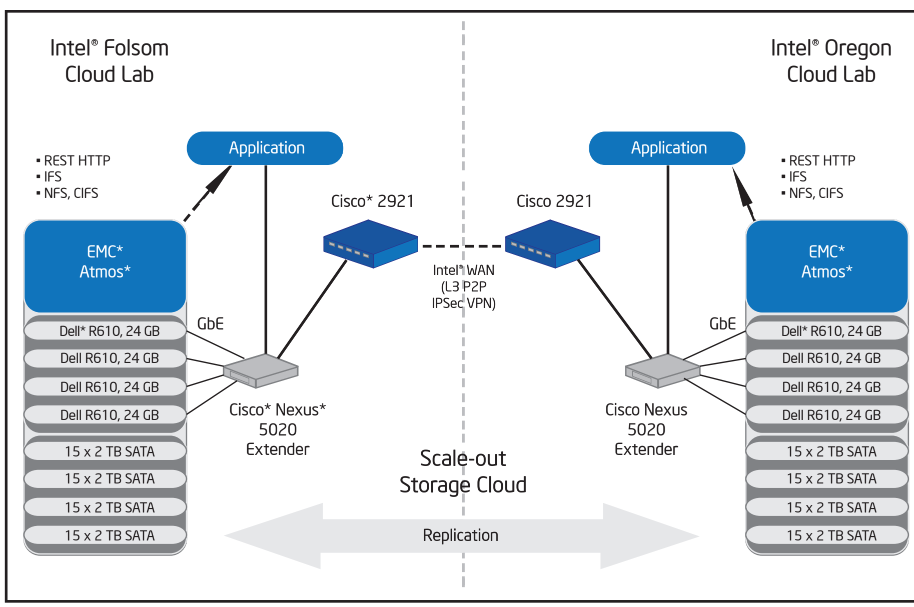

Figure 3: EMC Atmos Deployment in Intel Cloud Labs

EMC Atmos Scale-out Storage Cloud Architecture

EMC Atmos Overview

EMC designed EMC Atmos with features that support massive scalability (See Figure 2). EMC Atmos provides a platform that is designed to support multiple "tenants," and integrates an extremely powerful policy engine and a variety of software features to deliver a multi-use storage platform that can meet the demands of efficient, wide scale storage sought in cloud-based architectures. Atmos features include GeoProtect,[5] which is an intelligent object level protection scheme that utilizes object replication and Erasure coding. With both protection schemes available, the user can integrate the correct EMC Atmos GeoProtect scheme through the policy engine, which enables a business-level policy approach to run on any standard x86 server platform. EMC has integrated Atmos into racks using Dell* R610,[6] Intel® Xeon® servers and EMC's low-cost, high-density disk enclosures and SATA hard disk drives. Atmos addresses management complexity issues presented to a massively scalable, geographic footprint, with its globally distributed architecture and unified namespace. Atmos is typically deployed as a multi-petabyte, multi-site, scale-out storage cloud platform for information storage and distribution across multiple physical storage nodes. As a scale-out storage cloud Atmos alleviates limitations of traditional network attached storage (NAS) or storage area network (SAN) storage platforms by managing content as objects in a virtually unlimited namespace that can span geographies. Atmos software allows this distributed system to be managed as a single entity. In addition, management is simplified and automated with integrated policy-based data services that handle data protection, placement, drive spin-down, compression and de-duplication. With this set of robust functionality, EMC packages the efficiency achieved by large Internet companies into a supported and validated product.

Figure 4: EMC WS2-120

Atmos achieves scaling efficiency because it separates storage metadata from storage objects/data. A scale-out storage architecture has three major components: storage client, storage data node, and metadata store. This architecture enables optimal deployment of Atmos on standard server and storage hardware components (e.g. x86 processors, gigabit Ethernet (GbE) and 10 GbE network interface cards (NICs), hard disk drives, solid state drives, host bus adapters, etc.) that use the standard Linux* operating system. Use of standard server and software components are key tenets to meeting the stated cloud storage requirements.

Test Bed Architecture

For the purpose of this reference architecture, EMC and Intel deployed Atmos as two four-node racks: one in the Intel Folsom, California cloud lab and one in the Intel Hillsboro, Oregon cloud lab as shown in Figure 3. We configured the Atmos systems to be a single scale-out storage cloud connected by the Intel® WAN. We configured the Atmos system with a policy that automatically replicated data stored at either location to the other Atmos rack over the Intel WAN.

EMC Atmos Rack Configuration

We based the racks on the EMC Atmos WS2-120 configuration (see Figure 4). For each deployed rack:

- Each rack has four servers and four hard disk enclosures (half the normal eight servers and hard disk enclosures in a WS2-120 configuration).
- Each of the four servers has one GbE external NIC which is connected to a Cisco Nexus* 5020 Switch. Server NICs can be doubled or upgraded to 10 GbE if needed.
- Each of the four servers has an internal private GbE* NIC connected to a top-of-rack switch that is used for PXE boot during the initial Atmos configuration.
- Each server has a x4 SAS 6 GB/s connection to one EMC disk enclosure.
- Each EMC disk enclosure contains 15 EMC* SATA II 5.4K, 2 TB hard disk drives for a total of 30 TBs of raw storage per hard disk enclosure.

Each server is a Dell* PowerEdge* R610 Rack Server with the following configuration:

- Dual socket Intel® 5200 Chipset platform
- Two Intel® Xeon® Processor E5504
- 24 GB 1066 DDR3 Memory
- X86-64 Red Hat* Linux (Kernel: 2.6.29.6-4.2.smp.gcc4.1.x86_64)
- EMC* Atmos* Appliance 1.3.2.52930

EMC Atmos Installation and Validation

EMC Atmos Software Installation

We first installed the EMC Atmos software on the master server in the rack with the EMC Atmos Appliance CD through the keyboard, visual-display unit, mouse (KVM) interface. We then powered on each of the other three servers one at a time. Each server PXE booted from the master server over the private network and then copied the master server Atmos installation to their local boot disk.

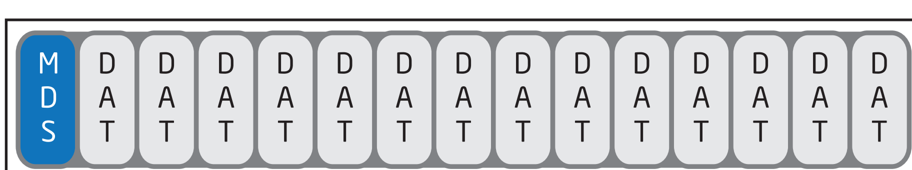

Figure 5: Metadata to Data Storage Config

EMC Atmos Hardware and Software Validation

We executed a test plan to validate hardware and software. The test plan includes validation of:

- metadata database
- services as running and healthy
- installation logs
- Atmos cloud through read and write sample objects
- policies as we viewed the location of object creation

EMC Atmos Configuration

EMC Atmos configuration occurs primarily through a Microsoft* Internet Explorer* version 7 browser. We configured both Atmos racks (i.e. the Atmos scale-out storage cloud) through a single connection and session.

Resource Management Group Configuration

The EMC Atmos resource management group (RMG) defines the racks in a single data center location. We defined a resource management group for the EMC Atmos racks in the Intel Folsom and Hillsboro labs.

1. Launched Microsoft IE with the URL: https://192.168.155.201/mgmt_login. The MasterServerIP address in our cloud lab was 192.168.155.201. We installed the EMC Atmos cloud on a dedicated VLAN.
2. Logged in as "SecurityAdmin." Selected "Initiate a new system" and changed the default "SecurityAdmin" password.
3. Created a resource management group (RMG) for the Intel Folsom cloud lab:

   Atmos RMG Name: FolsomCloudLab
   Atmos Location Name: Folsom
   Host Name Prefix: folsom1
4. Configured the installation segments:

   Segment name: FolsomCloudLab1-IS-1
   IP range:
   192.168.155.201-192.168.155.204
5. Created a resource management group (RMG) for the Intel Oregon cloud lab:

   Atmos RMG Name: OregonCloudLab
   Atmos Location Name: Oregon
   Host Name Prefix: oregon1
6. Clicked "Next." Application configuration setup for metadata versus data. Set the metadata to data ratio to 1:14 (see Figure 5).
7. Configured the installation segments:

   Segment name:
   OreganCloudLab1-IS-1
   IP range:
   192.168.32.200-192.168.155.203

The total storage amount used for the metadata is 8 TB. The rest of the storage (112 TB) is used to store data. If the data is stored through the use of a policy with one local object and one remote object (2 objects in total similar to a mirrored copy in a traditional storage system), then the total available storage is 56 TB.

Tenant Configuration

EMC Atmos provides the tenant and subtenant features to enable policy and administrative partitioning of the cloud. We defined the two Atmos racks with a single tenant, "IntelCloudLab." We allocated the four data nodes (Dell R610 server and EMC hard disk enclosure with hard disk drives (HDDs)) in each location to the tenants. We configured the tenants through the use of policies. We executed the following commands to create the tenants:

1. Launched Microsoft IE with the URL https://192.168.155.201/mgmt_login.
2. Clicked "Create Tenant" in the navigation pane.
3. Configured the tenant:
   Authentication source: local
   Tenant name: IntelCloudLab
4. Clicked "Submit" button to create the tenant.
5. Defined the tenant admin as "SysAdmin."
6. Clicked "Add Access Nodes" to add access nodes. Added the following:

   folsom1-001 through folsom1-003, file system none
   oregon1-001 through oregon1-003, file system none

EMC Atmos Replication Configuration

We performed the following steps to configure the EMC Atmos replication.

1. Clicked "MDS Remote Replica" on the system dashboard to configure metadata store service (MDS) to enable remote replication.
2. Selected "FolsomCloudLab-IS-1" as "installation segment 1" and "OregonCloudLab-IS-1" as "installation segment 2" to enable replication of data across the EMC Atmos nodes in the Folsom and Oregon cloud labs.
3. Added EMC Atmos "Netbackup" policy. Set Replica1 and Replica2 to the following values:

   Type: Sync
   Location: SameAs, $client
   Server Attributes: Optimal, Compression

EMC Atmos Linux Installable File System Installation and Configuration

The EMC Atmos installable file system (IFS), which uses the Linux file system in user space (FUSE[7]), enables high performance and reliable direct file system access on a Linux server to the EMC Atmos cloud storage. We used the following sequence of commands to set up EMC Atmos IFS on two cloud lab Intel Xeon servers running Red Hat Linux 5.5; one server in the Intel Folsom cloud lab and one in the Intel Oregon cloud lab:

1. Installed[8] FUSE (Red Hat) package manager (RPM) retrieved from sourceforge: fuse-2.7.4.tar.gz.[9]
2. Installed “Atmos-1.3.2.52930.x86_64.rpm” by executing “rpm –hiv Atmos-1.3.2.52930.x86_64.rpm” on the Red Hat servers.
3. Configured the EMC Atmos file system by executing “service mauifs configure” on the Red Hat servers.
4. Entered “Folsom” and “Oregon” as respective locations and “IntelCloudLab” tenant id and subtenant id.
5. Entered the four EMC Atmos node IP addresses for Folsom and Oregon respectively.
6. Disabled NFS exporting. Only used the EMC Atmos IFS on the servers.
7. Started EMC Atmos IFS by executing “service mauifs start.”
8. Created four directories structures to enable Atmos to distribute files uniformly across the four Atmos nodes.

   cd/mnt/mauifs
   mkdir NetBackup1
   mkdir NetBackup2
   mkdir NetBackup3
   mkdir NetBackup4

9. Tested the EMC Atmos IFS by executing the following:

   cd NetBackup1
   touch testFile

EMC ATMOS Scale-out Storage Tests

Linux RSync Test

Linux* Rsync[10] is an open source incremental file transfer utility that synchronizes two directory trees across different file systems that can be located on different computers and locations. We used Rsync as a simple test of the EMC Atmos scale-out storage cloud. The test was to backup and restore a Linux user’s home directory’s contents with the following sequence of commands:

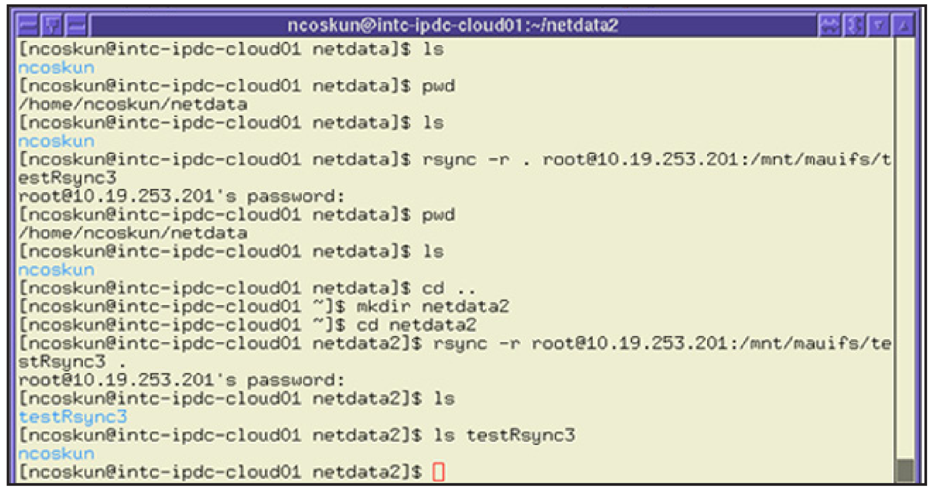

Figure 6: Rsync Results

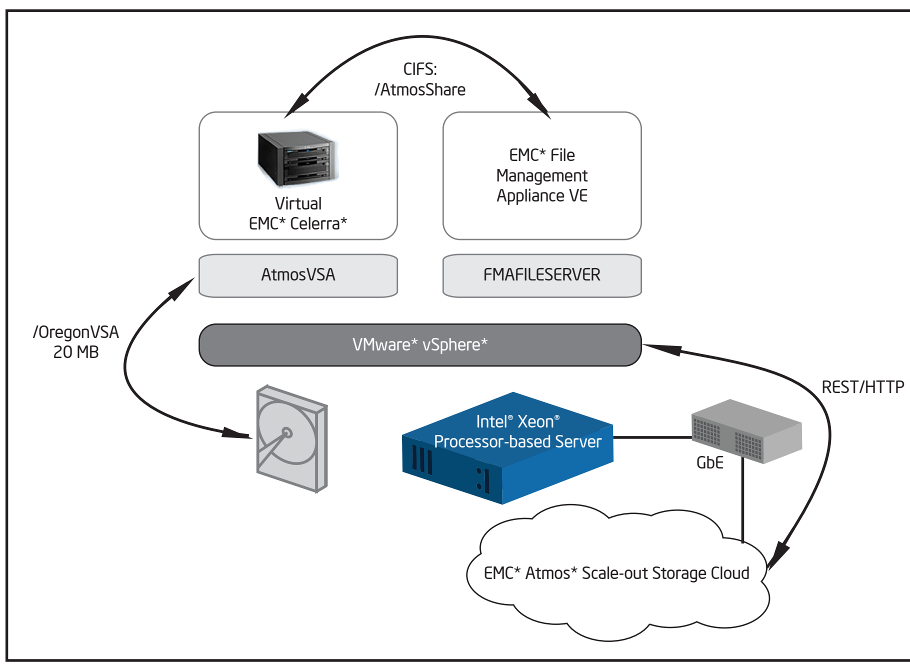

Figure 7: EMC FMA Test Setup

1. Executed “mkdir/mnt/mauifs/NetBackup1/ncoskun” on the Linux server where the Atmos FUSE installation was done. This created a backup directory in the Atmos cloud storage.
2. Executed “Rsync –rncoskun@pdx123.intel.com:/nfs/pdx/home/ncoskun/NetData /mnt/mauifs/NetBackup1/ncoskun” to back up the home directory. The command replicated the user directory to the Atmos cloud storage.
3. Executed “Rsync –r/mnt/mauifs/NetBackup1/ncoskun ncoskun@pdx123.intel.com:/nfs/pdx/home/ncoskun/NetData” to validate that the backup could be retrieved (see Figure 6).

EMC File Management Appliance (FMA)

We used the EMC File Management Appliance (FMA)[11] to archive inactive data from a NAS source to the Atmos scale-out storage cloud. A small (8KB) stub file is left behind on the NAS source, which the user sees as the actual data file, but which actually points to the archived data. When an end user or application attempts to access an archived file in its original NAS location, EMC FMA transparently recalls and presents the requested file. FMA data archival is fully automated through user-defined policies and scheduling.

File tiering or archiving lowers the total cost of ownership of NAS storage because it moves inactive or infrequently accessed data to the Atmos scale-out storage cloud, which thus reclaims capacity on more costly primary storage. We tested the file tiering usage model with the EMC File Management Appliance/VE (FMA/VE), a VMware virtual appliance. FMA/VE was set up to migrate data from a virtual EMC* Celerra* as shown in Figure 7 with a CIFS export. FMA/VE connects directly to the EMC Atmos cloud storage through the REST/HTTP interface.

Configuration of Virtual Celerra

We performed the following steps to configure the EMC virtual Celerra:

1. Imported the virtual Celerra .OVA file into VMware vSphere*. Installed the virtual Celerra with pre-configured hardware settings.
2. Logged into the virtual Celerra console and launched the configuration wizard.
3. Configured the following IP/Subnet/Gateway details:

   Hostname: AtmosVSA
   IP address: 192.168.32.212
   DNS: 192.168.32.211
   Domain name: Oregon.atmos.com

4. a. Launched the virtual Celerra graphical user interface (GUI).

   b. Expanded **Filesystems**.

   c. Selected **Datamover**.

   d. Selected **Server_2**.

5. Clicked **New**.

   Filesystem name: OregonVSA
   Radio button: Storage pool
   Storage space: 5 GB
   Datamover: Server_2

6. a. Clicked **Network**.

   b. Selected **New Network**.

   Ethernet adapter: eth0
   IP address

7. Clicked **CIFS** folder.
8. Clicked **CIFS Servers** tab.
9. a. Clicked **New**.

   b. Selected **Server_2**.

   c. Selected **Windows 2008**.

   Hostname: FMAFILESERVER
   Domain name: Oregon.atmos.com

10. a. Clicked the **Configuration** tab.

    b. Selected **Unicode enabled**.

    c. Started the CIFS service.

    d. Provided a domain admin username.

    e. Provided a password.

11. Selected the CIFS server checkbox for the AtmosVSA IP address..

12. Selected the **CIFS shares** tab.

13. Clicked **New share**. See Figure 8.

    Datamover: Server_2
    Name: Atmosshare
    File system: OregonVSA
    Path: \OregonVSA
    CIFS server: FMAFILESERVER

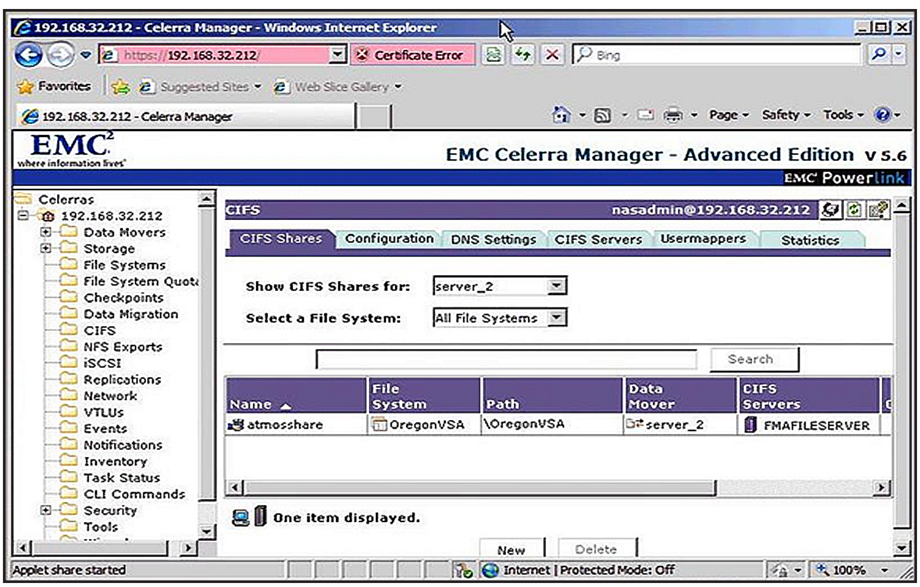

Figure 8: New Celerra Share

14. Logged in to the Celerra appliance using secure shell (SSH).

15. Started the dhsm api service with the following commands:

    #Server_http server_2 –service dhsm –start
    #Server_http server_2 –append dhsm –hosts 192.168.32.214
    #Fs_dhsm –c OregonVSA –d 0 –recall_policy no

Configuration of Atmos

Configuration of Nodes

1. Logged into the Atmos management console as SysAdmin: https://oregon1-001/mgmt_login.

2. Clicked **Edit** on IntelCloudLab. Clicked “Add.”

3. Selected “oregon1-001.”

4. Clicked the **Webservice** radio button.

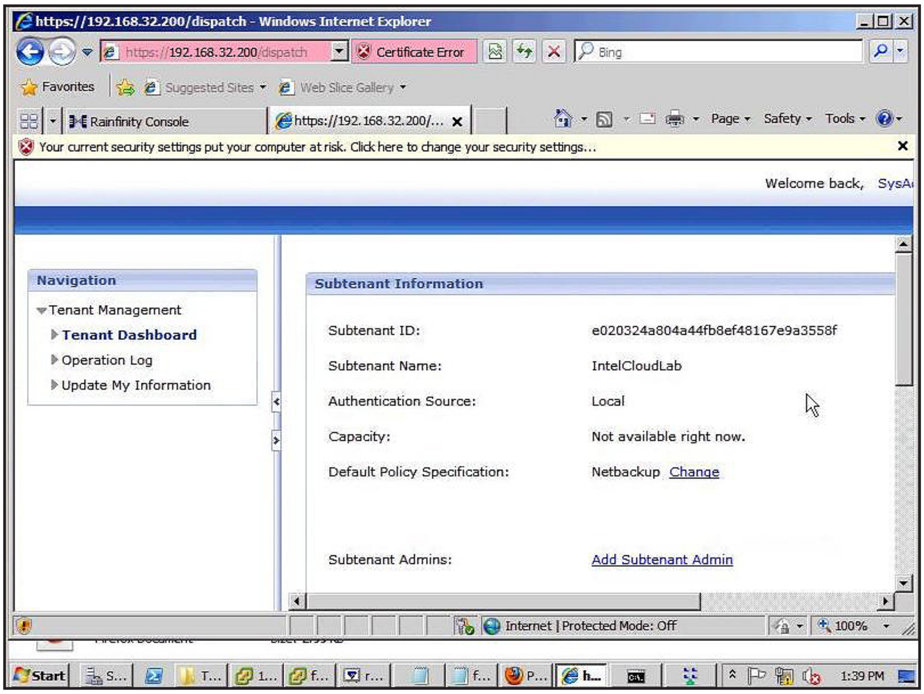

Figure 9: Configuration of Atmos Credentials

Configuration of Credentials

1. Logged into the Atmos tenant page: https://oregon1-001/:

   Tenant name: IntelCloudLab
   Password: SysAdmin

2. Clicked **Edit** on IntelCloudLab user.
3. Clicked **Add** under UID.

   UID: FMA

4. Note the shared secret string and subtenant ID generated by the system. See Figure 9.

Configuration of FMA

1. Imported the FMA virtual edition .OVA file into VMware vSphere. Installed the FMA VE with the pre-configured hardware settings.
2. Logged into the FMA VE console and launched the configuration wizard:

   Date: current date
   Time: current time
   Hostname: FMAFILESERVER
   IP address: 192.168.32.213
   DNS settings: 192.168.32.211
   Domain name: Oregon.atmos.com
   Time configuration: Domain controller NTP server.

3. Typed "!" return to exit the setup wizard.

4. Typed the following commands to start the Atmos service and exited out of command line:

   #atmoscallback start
   #exit

5. Logged into FMA GUI: http://<fmaApplainceIP>/login.
6. Clicked the **Configuration** tab.
7. Clicked **New fileservers** under Server configuration.
8. Selected **New server** and **Celerra** from the drop down list. See Figure 10.

   IP address: 192.168.32.213
   netbios name: FMAFILESERVER
   DART version: 5.6
   Control station: Celerra OS IP
   Domain name: Oregon.atmos.com
   Admin password: nasadmin
   Source tick box: Celerra
   Atmos call back agent: fm.oregon.atmos.com

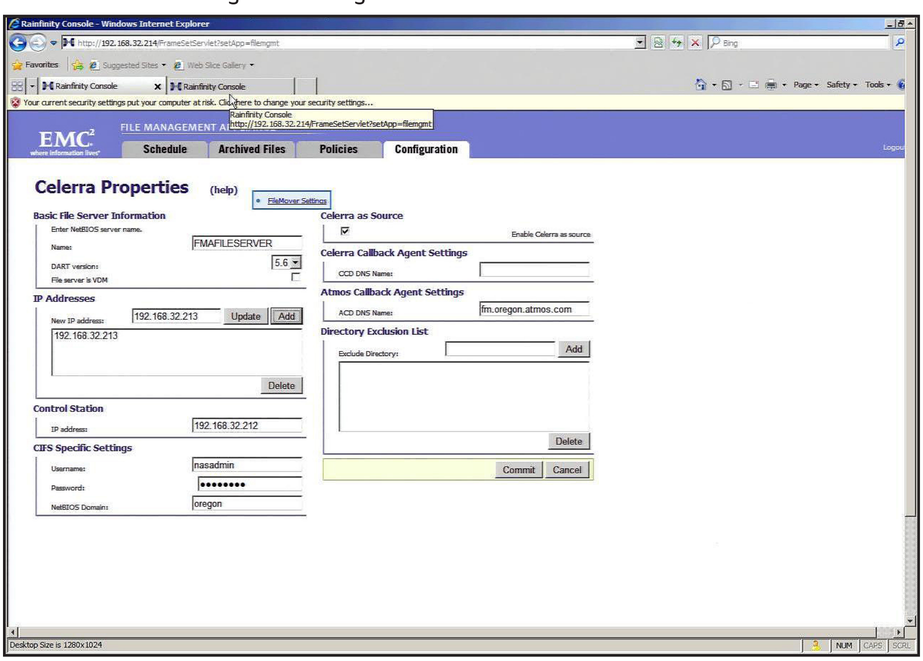

Figure 10: Configuration of Virtual Celerra for FMA

9. Clicked **Commit** and exited the configuration window.
10. Clicked **Filemover settings**.

    Username: dhsm
    Password: nasadmin

11. Clicked **Commit** and **Exit**.

12. Clicked **New fileserver** and selected **Atmos** under the drop down menu. See Figure 11.

    Name: oregon1-001
    DNA Name: oregon1-001.oregon.atmos.com
    Port: 80 HTTP
    Username: <atmos_string\uid> string copied from Atmos subtenant suffixed by UID
    Password: shared secret string created copied from Atmos GUI

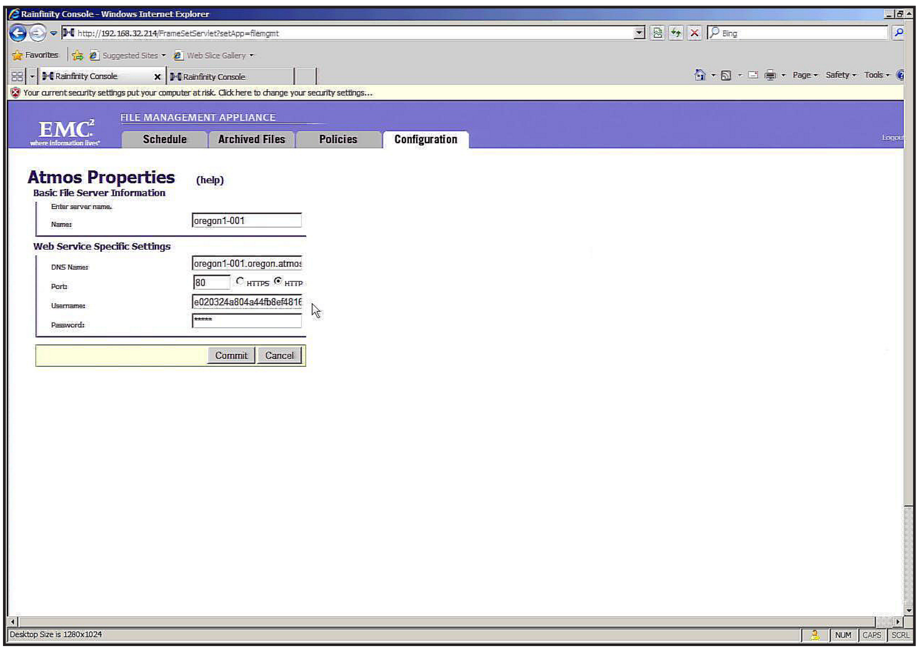

Figure 11: Configuration of Atmos File Server

13. Clicked **Verify** to check that the end-to-end connection on the fileserver is working and validated **Connection successful** message.

14. Clicked the **Policy** tab and click **Create new policy**.

    Name: test
    Policy type: archive
    Retention period: 0
    Delayed period: 0

15. Clicked the **Add rule to** button. See Figure 12.

    Created a rule to move files greater than 8KB to the archive:
    Archive destination: Atmos
    Server: Oregon1-001

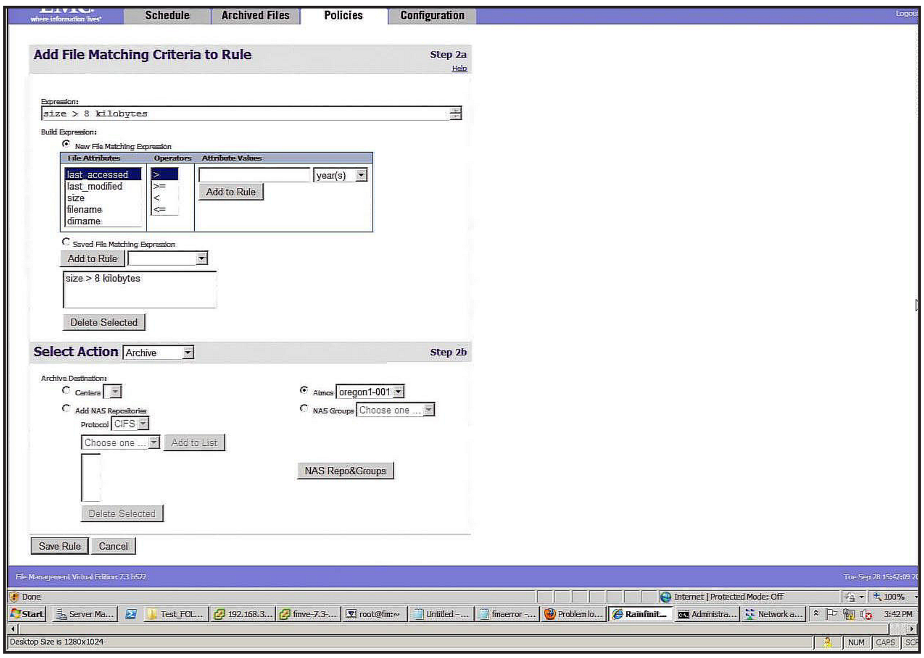

Figure 12: FMA File Matching Criteria to Rule

16. Clicked **Save** and exited rules.
17. Clicked **Save policy** and **Schedule** to exited policy tab.

Running Archive Job

1. Clicked the **Schedule** tab.
2. Clicked the orange icon next to policy "test."
3. Selected **Run now.** The file migration job starts.
4. Observed the green line showing job was in progress.
5. Observed that the line turns to black. This indicated that the job had completed.

6. Clicked **View summary** and viewed the details. See Figure 13.

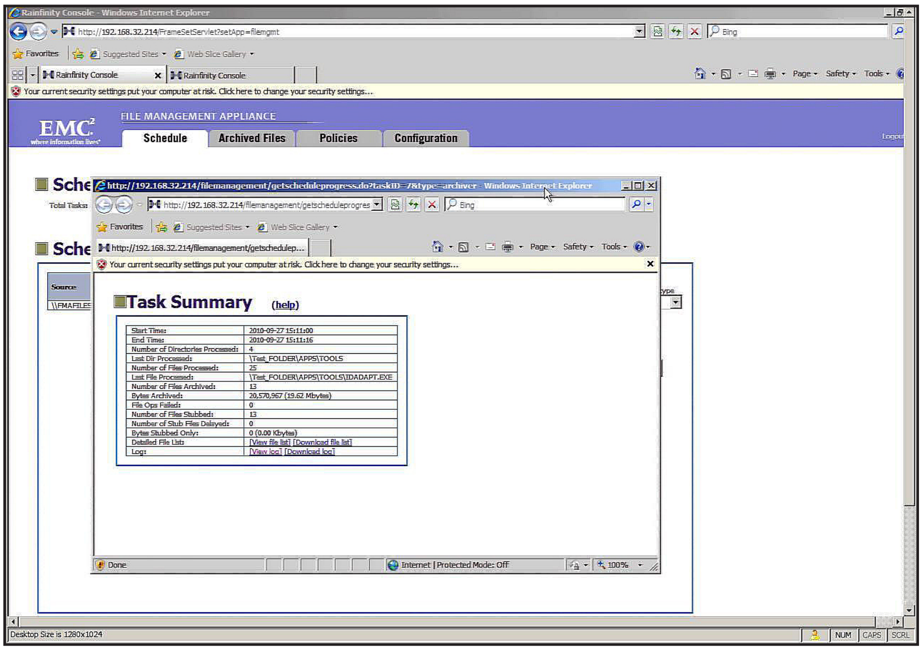

Figure 13: FMA Task Summary

7. Checked that files were archived by opening a folder on virtual Celerra. Archived files are sized at 8 KB FMA stub regardless of original size. See Figure 14.

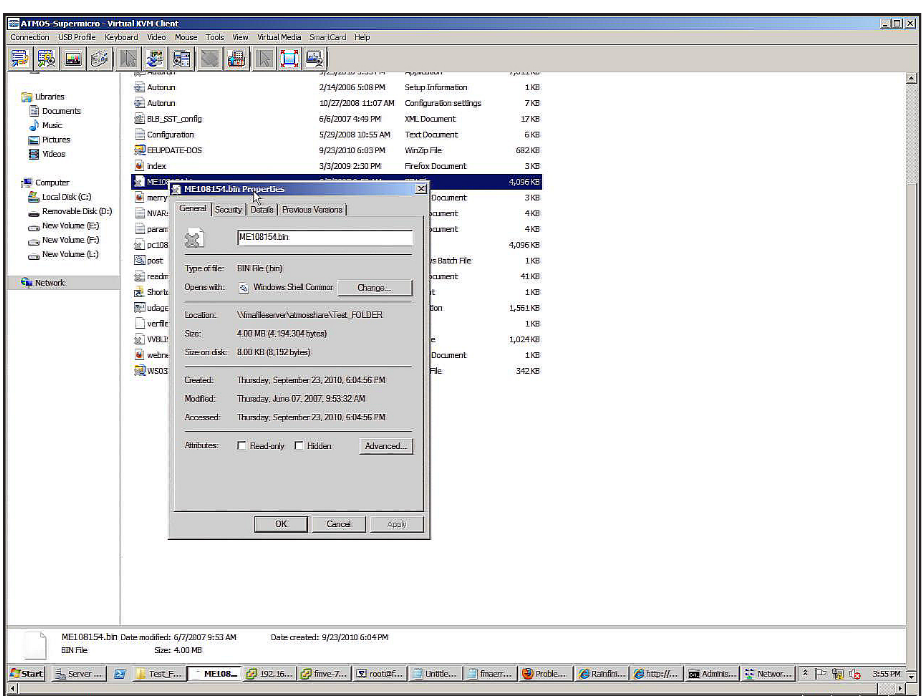

Figure 14: Archived Files on Virtual Celerra

Symantec Netbackup Test

Symantec NetBackup* is a data protection solution for a variety of platforms including Microsoft Windows*, UNIX* and Linux. You can use NetBackup to set up periodic or calendar-based schedules to perform automatic, unattended backups for clients across a network.

Symantec NetBackup Overview

NetBackup[12] includes both the server and client software. Server software resides on the computer that manages the storage devices. Client software resides on computer(s) that contain data to back up. Figure 15 shows NetBackup as we deployed and tested it with EMC Atmos in the Intel labs.

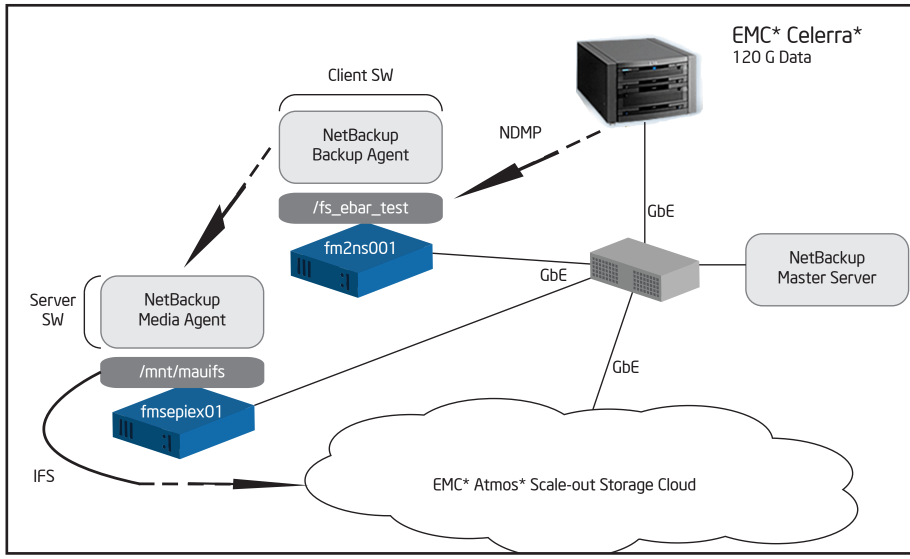

Figure 15: Symantec Netback Deployment in Intel Cloud Labs

NetBackup accommodates multiple servers that work together under the administrative control of one NetBackup master server in the following ways:

- The master server manages backups, archives, and restores. The master server is responsible for media and device selection for NetBackup. The master server contains information about tested backups and configuration.
- The NetBackup media servers provide the connection to the EMC Atmos scale-out storage cloud through the EMC IFS interface (EMC Atmos Linux FUSE). During a backup or archive, the client sends backup data across the network to a NetBackup media server.

During a restore, users can browse and select the files and directories to recover. NetBackup finds the selected files and directories and restores them to the disk on the client.

The NetBackup administration console provides a graphical user interface through which the administrator can manage NetBackup. NetBackup uses the Administration Console to configure, manage, and monitor the storage devices, storage servers, disk pools, storage volumes, catalogs, policies, host properties, backup, and archives, as well as to restore jobs, daemons, processes, and reports.

NetBackup was integrated with EMC Atmos through the Installable File System (IFS)[13] interface. RedHat 5.4 was installed on an Intel Xeon server. The server was configured with two NICs. One NIC was used to connect to the Intel network. The other NIC was connected to the Atmos servers.

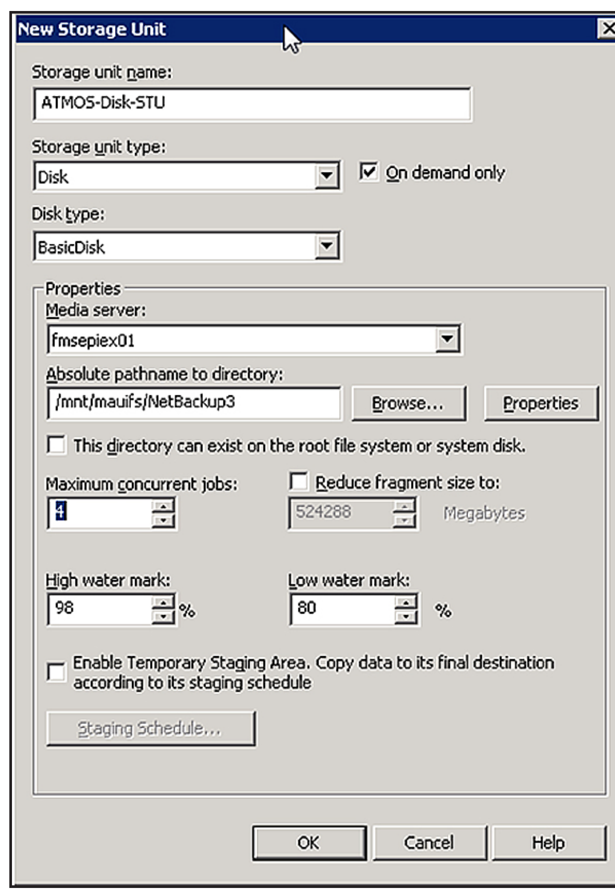

Figure 16

NetBackup Installation Notes

NetBackup media server software manages the storage devices within the NetBackup environment. We installed Linux NetBackup media server on the Red Hat Linux server that contained the Atmos IFS (/mnt/mauifs) installation.

NetBackup Test

We used the following steps to test NetBackup to the Atmos in the lab:

1. Created a NetBackup storage unit for the Atmos cloud storage. See Figure 16.
2. Created a NetBackup policy for the storage unit:

   Policy name: Atmos_Backup_test
   Policy type: NDMP
   Policy storage: Atmos-STU
   Client server hardware and operating system: NDMP, NDMP
   Backup selection: /fs_ebar_test

3. Right clicked "Atmos_fm2ns001_ndmp" and selected **Manual Backup.**
4. Selected **Activity Monitor** to view details. See Figure 17.

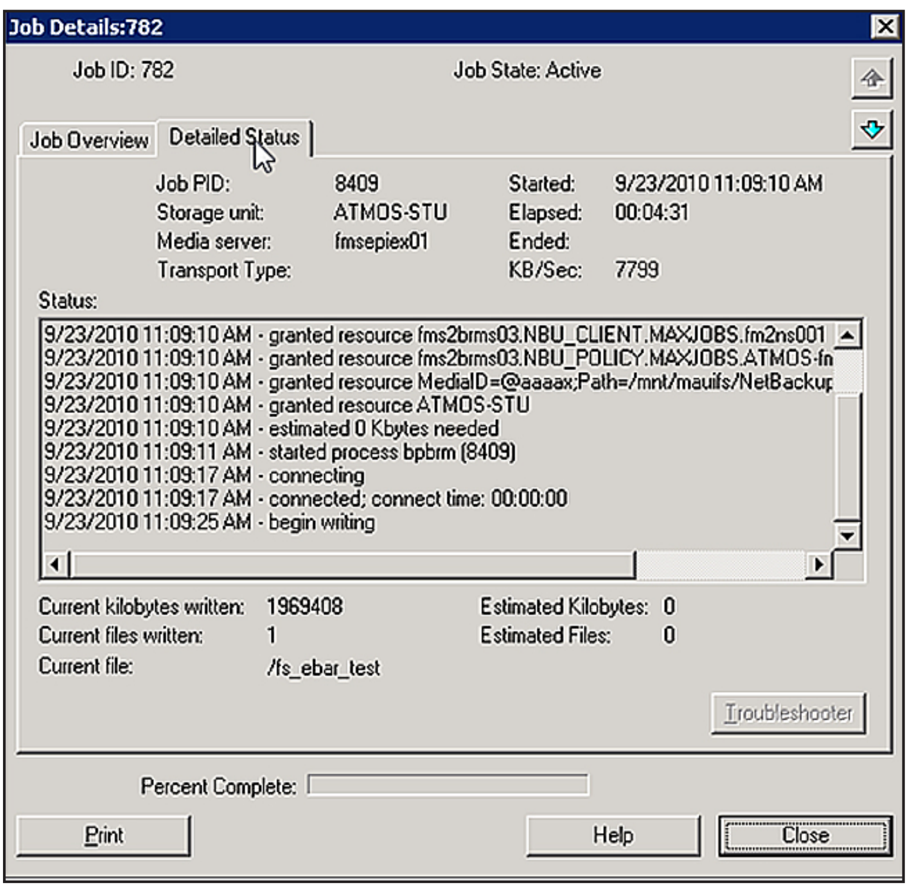

Figure 17

Microsoft SQL Server* Backup Test

The EMC Atmos scale-out storage cloud is ideal to store SQL Server database backup (DB) snapshots. We did a simple test to back up a DVD store to the SQL Server DB with the use of the Atmos CIFS client interface.

EMC Atmos CIFS Subtenant Configuration

Configured Atmos through the use of Internet Explorer (see Figure 18).

1. Logged in to system management console https://<Atmosnodeip>/mgmt_login/ and used SysAdmin as user account.

2. Navigated to **Tenant Management** >**Tenant List**> **Enable CIFS** on the desired nodes through the selection of the appropriate radio button. Also select **Multi Subtenant Access**. Save and exit the menu.
3. Created a new subtenant.
4. Logged into the Tenant Management console https://<Atmosnodeip>/ ; used "IntelCloudLab" as the tenant ID.
5. Clicked **Add** under the subtenant header to create new a subtenant under "IntelCloudLab." This new subtenant hosted CIFS share to store our backup.
6. Provided the name of the subtenant as "MSSQLCIFS" and click **Create**. This step creates the MSSQLCIFS subtenant which appeared under the subtenant list.
7. Clicked **Edit** next to the subtenant to continue with configuration. This step configured the subtenant by default as CIFS.
8. Clicked **Change** on the default policy specification and select "Netbackup" as the preferred policy. This policy ensured the replication of data written to the node on the local site and also to the remote site in Folsom.
9. Clicked the hyperlink named "CIFS" next to the node name. Clicked **Add** to add a new share.
10. Configured the share name as "SqlBackup." We did not make changes to the other parameters.
11. Clicked **Submit** to create a new share.

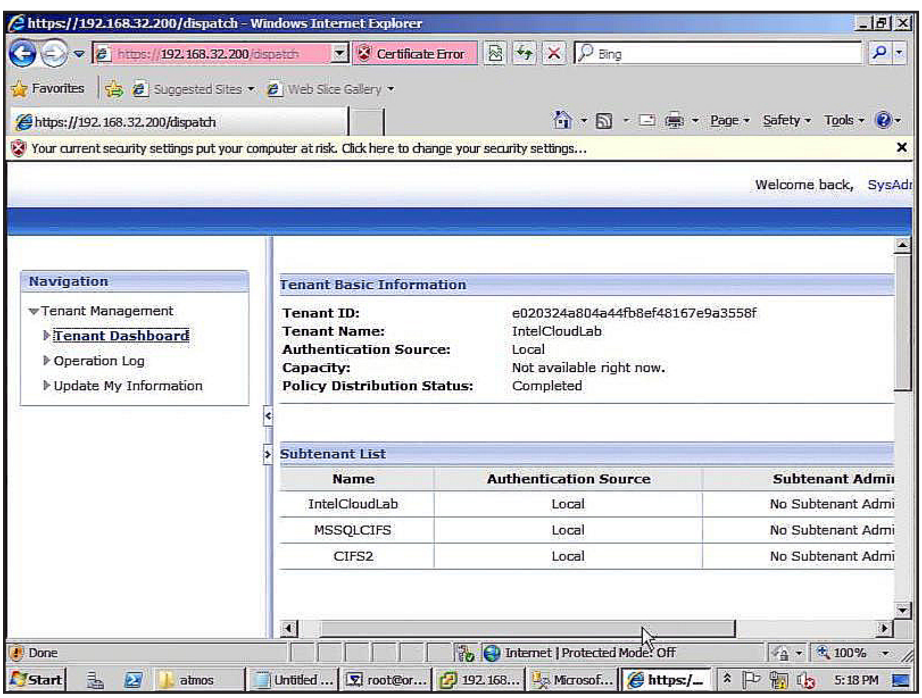

Figure 18: Atmos MS SQL Server Subtenant

MS SQL Server Configuration

We implemented the MS SQL server on the following hardware configuration:

- Super Micro X8DTU Chassis, Intel 5200 Chipset
- Intel® Xeon® X5667 @ 3.07 GHz quad core two sockets
- 12 GB DDR3 memory
- Intel® X-25 160 GB SSD boot drive
- LSI MegaRAID* SAS
- 12 Hitachi* 2 TB drives

- Microsoft Windows Server* 2008 R2 – Enterprise Edition 64-bit
- Microsoft SQL 2005 with SP3*
- Dell DVD Store (DDS) database workload[14]

We set up the MS SQL server with the following steps.

1. Installed Windows 2008 R2 Enterprise edition with the default settings.
2. Installed Microsoft SQL 2005 64-bit edition with Service Pack 3.
3. Unziped the DDS to c:\ .
4. Created the three partitions required by the DDS: “E:”, “F:”, and “L:.”
5. Opened Createdb.sql in a query analyzer and executed it to create the database and log files.
6. Configured Windows 2008 path to the new EMC Atmos share: “\\oregon1-001\SqlBackup\.”

We performed the following steps to back up the DDS database:

1. Opened SQL Management Studio, navigated to **Server Objects** > **Backup Devices** and right clicked **New Backup Device**.
2. Provided device name as AtmosBackup.
3. Provided destination file path as \\Oregon1-001\SqlBackup\Ds1.bak.
4. Opened **Server Objects** > **Backup Devices** and right clicked **Back Up a Database**.
5. Selected “DS2” as the database and backup type as “Full.”
6. Clicked **Ok** to start the backup.

Verify MS SQL Server Backup

We stored the validated SQL Server backup files in Atmos cloud storage with the following sequence of commands:

1. Logged in to the system management console https://<Atmosnodeip>/mgmt_login with SysAdmin as the user account.
2. Copied the TenantID string for IntelCloudLab and the Subtenant ID string for MSSQLCIFS.
3. Opened a SSH session to Atmos Node.
4. Executed “mauiobjbrowswer -t <tenant_id> -s <subtenant_id> -p <file_path>.

5. Validated that the backup file was present on the Oregon Atmos rack (Figure 19) as well as on the Folsom Atmos rack (Figure 20).

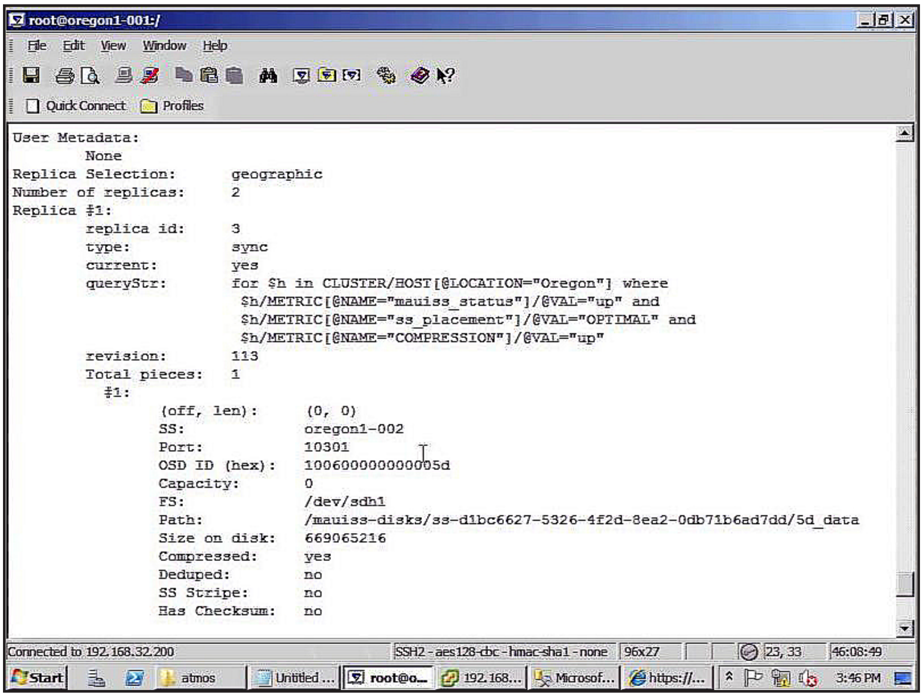

Figure 19: MS SQL Server Backup Replica on Oregon Atmos Rack

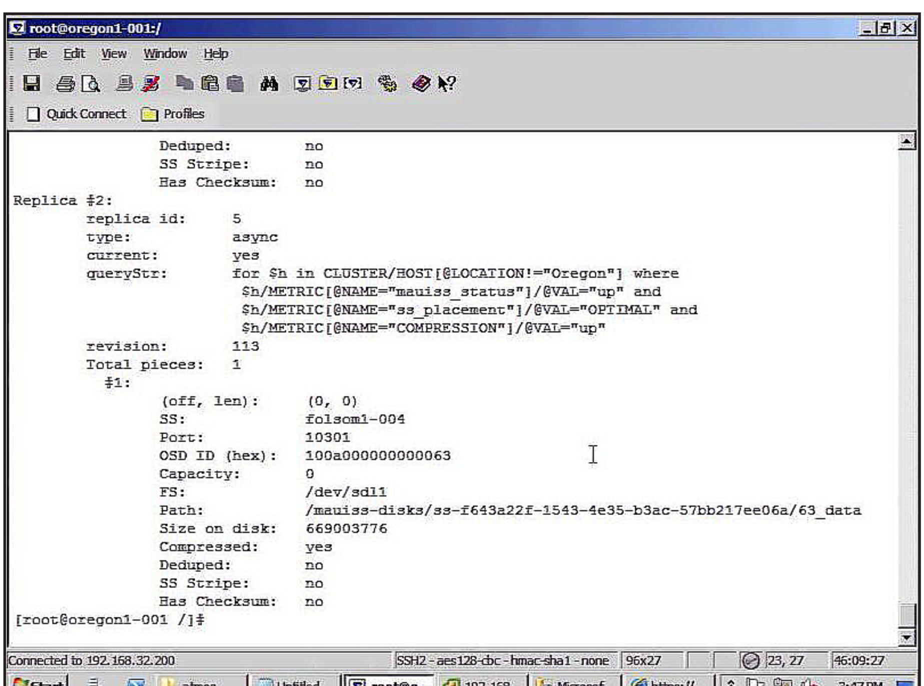

Figure 20: MS SQL Server Backup File on Folsom Atmos Rack

Things to Consider

Networking Architecture

The EMC ATMOS scale-out storage cloud solution documented in this paper utilized Gigabit Ethernet (GbE) connectivity from each node through a GbE switch as the interface between the application and the EMC nodes. You can achieve higher throughput to the cloud storage if you equip the Dell R610 servers with an Intel 10 GbE network interface card and engineer 10 GbE switch access for the applications. As an example, each ATMOS storage node supports 15 2 TB hard disk drives. The SATA drives are capable of achieving close to 100MB/s sustained transfer rate. Assuming this, the sustained sequential transfer from 15 drives would exceed the throughput provide by a 10GbE interface.

Performance of the Metadata Access

The Atmos metadata service performance is currently scaled by adding more SATA hard disk drives. The metadata service performance could likely be more efficiently scaled by using Intel® X25-E or X25-M SATA solid state drives[15] to store the metadata. Solid state drives provide well over 10 times the random read and write throughput of a SATA hard drive.

Application Processing on the Data Nodes

One of the usage models we mentioned in this paper but did not test was the use of scale-out storage for data warehousing/analytics. Portals commonly use the Hadoop application framework with scale-out storage to transform the data. The Hadoop framework runs the transform application (typically called the “map app”) on the same node on which the data would reside. In this case, you would need to size the processor performance of the node appropriately to support both the data access and the application processing. To gain the additional performance needed by the application, you might want to use a higher performance Intel Xeon processor.

Caching of Data at the Compute Server Access

Highest application performance is typically achieved when the storage resides on the same server as the application. This allows the application to take full advantage of the high throughput SATA and PCI Express interfaces provide by the Intel 5400 chipset.[16] One architecture for which this could be supported is equipping the application server with an Intel X25-E or X25-M SATA solid state drive. This drive would be used as a cache for data being read from the Atmos storage nodes. Once the drive is populated, subsequent accesses to the data would then be access from the SSD. The Linux open source community has created two solutions for enabling this usage model: Linux Flashcache[17] and NFS CacheFS.[18]

Power Optimization

Power consumption for large amounts of storage (i.e. Petabytes (10^{15}) or even Zettabytes (10^{21})) can become a significant part of the total cost of ownership for a storage cloud. The scale-out storage cloud test could be augmented by profiling the power consumption of the racks when idle or being tested by the usage models. The power utilization could be optimized through technology like the Intel Intelligent Power Node Manager[19] and Intel lower power processors.

Next Steps

The next step is to profile the performance of the usage models in addition to the basic functionality. In addition, we continue to plan tests for the items mentioned in Things to Consider.

Conclusions

This paper defined and tested a scale-out storage cloud reference architecture based on EMC Atmos. The scale-out storage architecture demonstrated how to maximize a cloud deployment strategy by enabling a multi-use, shared infrastructure.

The Atmos scale-out storage architecture demonstrated how to maximize a cloud deployment strategy by enabling a multi-use, shared infrastructure. Important customer considerations when designing cloud reference architectures are eliminating single use, silo or project oriented infrastructure that is purpose built. One benefit of Atmos enabled shared infrastructure deployments is allowing customers to improve application and resource delivery cycles driving business innovation. Atmos demonstrated support for multi-use, shared infrastructure, which was utilized through multi-tenancy and multiple access methods.

Atmos provides flexible access mechanisms. The configuration and tests validated the REST/HTTP, IFS, and CIFS interfaces supported by Atmos. Atmos, by supporting multiple access methods, provides an agile scale-out storage strategy that enables end users to accommodate a wider variety of use cases.

Adaptive storage capabilities via granular policy control were also demonstrated. An important aspect to scale out storage is effective, efficient utilization. One method to maximize the ROI of scale out storage, as described above, is to deploy multiple use cases across a shared infrastructure. An equally important aspect towards providing effective, efficient utilization is to provide an intelligent mechanism and an array of feature sets that allows each use case to uniquely handle objects. Atmos presents a powerful policy engine combined with several object storage capabilities, such as replication, compression, and erasure coding, to provide efficient storage of objects across many use cases or multiple tenants.

In conclusion, this reference architecture demonstrates EMC Atmos as an agile implementation available today that meets enterprise end-user requirements for efficiency and elasticity in the scale-out storage cloud usage model. It also demonstrates the versatility of Atmos with regards to multiple use cases that are relevant to a cloud solution. This paper further demonstrated the value proposition of scale-storage with Atmos through the availability of multiple access methods, various EMC Atmos GeoProtect schemes, support for multi-tenancy, and multiple data services (such as compression and de-duplication) integrated through a powerful and intelligent meta-data driven policy engine.

Glossary

B2B: Business-to-business (B2B) describes commerce transactions between businesses, such as between a manufacturer and a wholesaler, or between a wholesaler and a retailer. From: http://en.wikipedia.org/wiki/Business-to-business.

B2C: Business-to-consumer (B2C, sometimes also called Business-to-Customer) describes activities of businesses that serve end consumers with products and/or services. An example of a B2C transaction is a person who buys a pair of shoes from a retailer. The transactions that led to the shoes being available for purchase, that is the purchase of the leather, laces, rubber, etc. as well as the sale of the shoe from the shoemaker to the retailer would be considered (B2B) transactions. From: http://en.wikipedia.org/wiki/Business-to-consumer.

CIFS: Common Internet File System: Also known as Server Message Block (SMB) is a network protocol used to provide shared access to files, printers, serial ports, and miscellaneous communications between nodes on a network – typically a set of Microsoft Windows servers and PC clients. See: http://en.wikipedia.org/wiki/Server_Message_Block.

Compression: Data compression is the process of encoding information with fewer bits than the unencoded representation would use, through use of specific encoding schemes. From: http://en.wikipedia.org/wiki/Data_compression.

Deduplication: Data deduplication (or Dedup) is a specialized data compression technique to eliminate coarse-grained redundant data, typically to improve storage utilization. In the deduplication process, duplicate data is deleted, which leaves only one copy of the data to be stored, along with references to the unique copy of data. Deduplication reduces the required storage capacity since only the unique data is stored. From: http://en.wikipedia.org/wiki/Data_deduplication.

Disk enclosure: A disk enclosure is a chassis or shelf designed to hold and power disk drives while providing a mechanism to allow them to communicate to one or more servers. JBODS (just a bunch of disks) is another term for disk enclosures. From: http://en.wikipedia.org/wiki/Disk_enclosure.

Erasure coding: Erasure coding is a forward error correction (FEC) algorithm which uses data stripping to transform k data elements (across scale-out storage data nodes) into a longer message stripe coded with n data elements. This enables the original data to be recovered from a subset of the n data elements on failure. See: http://en.wikipedia.org/wiki/Erasure_code

Host bus adapter (HBA): Connects a host system (a server) to other network and storage devices including hard disk drives and solid state storage. See: http://en.wikipedia.org/wiki/Host_adapter.

Installable File System (IFS): IFS enables users to create their own file systems without editing kernel code through the use of FUSE. See http://en.wikipedia.org/wiki/Filesystem_in_Userspace.

KVM: A KVM switch (with KVM being an abbreviation for keyboard, video or visual display unit, mouse) is a hardware device that allows a user to control multiple servers from a single keyboard, video monitor and mouse.

Metadata: Metadata is loosely defined as data about data. Metadata is a concept that applies mainly to electronically archived or presented data and is used to describe the a) definition, b) structure, and c) administration of data files with all contents in context to ease the use of the captured and archived data for further use. From: http://en.wikipedia.org/wiki/Metadata.

Multi-tenant: The ability to serve multiple tenants or customers with a single scale-out storage platform. For public cloud storage, a tenant is typically one of the customers of the cloud storage (e.g. a small business using the cloud for backup). For private cloud storage, a tenant is typically one application using the storage (e.g. the EMC File Manage Appliance archiving data). With a multi-tenant architecture, the scale-out storage is designed to securely partition its data and configuration enabling each tenant to customize its instance as if it was the only user of the scale-out storage cloud.

NAS: Network Attached Storage is a storage server or appliance that uses file-based protocols such as NFS (network file server) or CIFS to enable clients (typically servers and PCs) to access files over a TCP/IP network. See: http://en.wikipedia.org/wiki/Network-attached_storage.

NDMP: Network Data Management Protocol is a protocol invented by the NetApp and Legato that transports data between NAS devices and backup devices. This removes the need to transport the data through the backup server itself, thus enhancing speed and removing load from the backup server.

NIC: A network interface card is hardware that enables a server to interface to an Ethernet or TCP/IP local area network (LAN). A NIC is not necessarily a card in the server; it could be integrated as LOM (LAN on server motherboard).

Portal: Web portals offer services such as a web search engine, e-mail, news, stock prices, information, databases and entertainment. Portals provide a way for enterprises to provide a consistent look and feel with access control and procedures for multiple applications and databases. See: http://en.wikipedia.org/wiki/Web_portal.

PXE boot: The Preboot eXecution Environment (PXE, and also known as Pre-Execution Environment) is a process to boot a server by remotely accessing the boot image through the use of the NIC and a LAN. See: http://en.wikipedia.org/wiki/Preboot_Execution_Environment.

Replication: Data replication is the process of sharing data so as to improve reliability between redundant storage devices. The replication is transparent to an application or end-user. In a failure scenario, failover of replicas is hidden as much as possible.

REST/HTTP: Representational State Transfer: An architecture that communicates between clients and servers over a TCP/IP network (e.g. Internet). Clients initiate requests to servers; servers process requests and return appropriate responses. At any particular time, a client can either be in transition between application states or "at rest." A client in a rest state is able to interact with its user, but creates no load and consumes no per-client storage on the set of servers or on the network. The client begins to send requests when it is ready to make the transition to a new state. The Hypertext Transfer Protocol (HTTP) is commonly used as the transport layer basis for REST communication. See http://en.wikipedia.org/wiki/Representational_State_Transfer.

Samba: Samba is the standard Windows interoperability suite of programs for Linux and Unix. See: http://www.samba.org/

SAN: Storage Area Network: A storage server or appliance that uses block-based protocols typically based on SCSI to access files over a fibre channel or TCP/IP network. See: http://en.wikipedia.org/wiki/Storage_area_network.

SATA: Serial Advanced Technology Attachment: A storage interface to connect host bus adapters to hard disk drives and solid state drives. Desktop and laptop computers use SATA hard disk drives which typically have the largest capacity at the lowest cost (dollars per Gigabyte). http://en.wikipedia.org/wiki/Serial_ATA.

Scale-out Storage (SOS): SOS is a usage model for storage that enables an enterprise to grow capacity incrementally as it adds more storage nodes (typically as a new server on an IP network). The goal of scale-out storage is to grow capacity with near linear (versus lump sum) investment.

SSH: The Secure Shell client is a program for logging into a remote machine and for executing commands on the remote machine. It is a secure version of the rlogin, rsh, and telenet commands. It provides secure encrypted communications between two untrusted hosts over an insecure network. http://en.wikipedia.org/wiki/Secure_Shell

Spin-down: Spin-down refers to turning off a hard disk drive after a specific period of time to conserve energy..

Additional Information

1. Atmos Conceptual Overview 1.3.2A.pdf
2. Atmos Installation and Upgrade Guide 1.3.2A.pdf
3. Atmos Admin Guide 1.3.2A.pdf
4. Symantec NetBackup Administrator's Guide, Volume 1, Unix and Linux, Release 7.0.1
5. Symantec NetBackup Install Guide for Windows, Release 7.0.1
6. Symantec NetBackup Install Guide for Unix and Linux, Release 7.0.1

Endnotes

1. Source: IDC Digital Universe Study, Sponsored by EMC, May 2010.

2. http://www.emc.com/domains/documentum/

3. http://www.emc.com/products/family/atmos.htm

4. http://hadoop.apache.org/

5. http://www.emc.com/collateral/demos/microsites/mediaplayer-video/emc-atmos-geoprotect.htm

6. http://content.dell.com/us/en/enterprise/d/virtualization/PowerEdge-R610.aspx

7. http://sourceforge.net/projects/fuse/files/fuse-2.X/2.7.4/fuse-2.7.4.tar.gz/download

8. For details refer to Atmos Installation and Upgrade Guide 1.3.2A.pdf

9. See number 6.

10. http://www.samba.org/rsync/

11. http://www.emc.com/products/detail/software/file-management-appliance.htm

12. http://www.symantec.com/business/netbackup

13. http://fuse.sourceforge.net/

14. http://www.delltechcenter.com/page/DVD+Store

15. http://www.intel.com/go/ssd

16. http://www.intel.com/products/server/chipsets/5400/5400-overview.htm

17. http://github.com/facebook/flashcache/

18. http://en.wikipedia.org/wiki/CacheFS

19. http://www.intel.com/technology/intelligentpower

To learn more about deployment of cloud solutions,
visit www.intel.com/cloudbuilders

Disclaimers

Printed in USA 1010/NS/PRW/PDF Please Recycle 324438-001 US