Network Data Management Protocol (NDMP) White Paper

Summary

What is the primary goal of enterprise storage management? To back up and restore information in an intelligent, secure, timely, cost-effective manner over all enterprise-wide operating systems.

Companies need high-performance backup and the ability to back up data to local media devices, without network traffic. While the data itself may be distributed throughout the enterprise, its cataloging and control must be centralized. The emergence of network-attached storage and dedicated file servers makes storage management more challenging.

The current practice for all storage management vendors is to adapt their architecture to the operating systems they support. The result is an implementation with layers of OS dependencies that require the user to change workflow and make concessions to departmental systems in the enterprise.

The Network Data Management Protocol (NDMP) is a new approach. Recognizing that these issues must be addressed, Network Appliance and Intelliguard have pioneered a network-based protocol that can be used for communications by all centralized backup applications and all agents on file servers. NDMP will create the first opportunity to provide truly enterprise-wide heterogeneous storage management solutions. The paradigm now shifts, permitting platforms to be driven at a departmental level and backup at the enterprise level.

NDMP will allow administrators to back up critical data using any combination of compliant network-attached servers, backup devices, and management applications. The protocol addresses the user's need for centralized control of enterprise-wide network data management while minimizing network traffic.

NDMP User Benefits

NDMP meets the strategic need to centrally manage and control distributed data, while minimizing network traffic. NDMP, as an embedded protocol, separates the data path and the control path, so network data can be backed up locally yet managed from a central location. NDMP has been pioneered by Intelliguard and Network Appliance and endorsed by industry leaders as a protocol that empowers users to simplify the management and protection of business-critical data.

The Problem

Data storage requirements are doubling annually. Storage management costs are estimated at $7 per MB per year (Strategic Research 1996). The need for comprehensive, efficient data management is obvious. Centralized, enterprise-wide control of mission-critical data backup is imperative in increasingly heterogeneous environments. Data is spread across multiple hardware

and operating system platforms. The emergence of network-attached storage, or dedicated file server "appliances," means that the network-storage environment is growing more complex.

An enterprise backup solution must be scalable, meet performance requirements, and provide excellent vendor support. In addition, it must support backup of all platforms on which the enterprise's mission-critical data is stored.

The User Challenge

The network storage environment is heterogeneous. The operating systems are constantly being upgraded. Dedicated file server "appliances" store mission-critical data. There is a wide variety of backup media devices and technologies, such as 8mm and DLT. The backup software solution must be fully compliant with all platforms on which the data to be backed up is stored and with the media devices being used.

The Information Systems department has to ensure that upgrade paths are well planned in order to protect mission-critical data by ensuring timely backup. In complex environments, version control and software distribution are the stuff of which nightmares are made.

Then there is the challenge of backing up network-attached storage. Such systems, by their nature, are dedicated appliances and do not support software applications such as backup. Performing NFS mounts of their file systems and backing up over the network has been the traditional method of backup. This is an inefficient, time- and resource-consuming activity.

The Backup Software Vendor Challenge

Extensive development and testing resources are required by a backup software vendor to ensure complete interoperability and compatibility between various platforms and versions. The vendor works with multiple file system vendors to understand the specific OS internals and seeks to maximize data transfer performance.

The effort required to port the backup software and to maintain ongoing platform support is substantial. Tape library or jukebox support typically requires device drivers.

These necessary activities extend product development timelines. Resources are diverted from feature enhancements and support to compatibility assurance.

The Server Vendor Challenge

Server vendors have to meet the demands of their customers by supporting a broad range of backup media devices and major backup software applications. Systems with large storage capacity require very high-performance local backup capabilities. Server vendors work with multiple external vendors and often invest in proprietary APIs to ensure interoperability with backup hardware and software. These extensive development and support efforts eat into development and support cycles that could otherwise be used to improve performance and develop valuable new features.

The Network-Attached Storage Vendor Challenge

In the case of network-attached storage, the general-purpose server vendor challenge is compounded. Network-attached storage "appliances" are designed to optimize a single function, namely file service. As such, they do not have the general-purpose operating system required for porting a backup software solution. Indeed, the backup software is likely to have a great deal more lines of code than the dedicated OS itself. Backing up and restoring data over the network is slow and ties up the network.

The Solution

Backup is an issue that must be addressed by the server vendor, backup software vendor, and backup device vendor communities collectively. The common objective is to provide centrally managed, enterprise-wide data protection for the user in a heterogeneous environment. A common backup architecture must be defined, and the problem of backup must be partitioned between vendor types. This is achieved by defining and promoting an open standard network backup protocol. NDMP is this protocol.

Vendor compliance with NDMP provides users with plug-and-play ability in a heterogeneous environment. Best-of-breed solutions, ideally suited for the specific demands of a given network environment, can be chosen from server, backup device, and backup software vendors, and interoperability is assured.

By partitioning the problem between vendors, each vendor implements solutions in compliance with one well-defined standard network protocol. Interoperability efforts and widespread vendor support are eliminated. The vendor is free to focus on core competencies, thus improving the user offering and decreasing time-to-market.

The Basic Elements of Backup

The enterprise backup is a highly complex procedure. The data to be backed up must be defined, and complex interactions with the backup media device and extensive cataloguing and control must be managed. Good enterprise wide backup solutions are designed to assure data protection and efficient restoration of mission-critical data in the event of data loss. This means a great number of control and management features are implemented at the front end, at backup.

In a generic sense, the complex task of backup can be broken into the following tasks:

- Discovery
- Configuration
- Scheduling
- Media management
- Tape device control
- Autoloader device control
- Data client software
- User interface

The Common Backup Architecture

We have spoken about the three vendor types associated with backup, namely the file server, backup device, and backup software vendors. Now let's take a look at how file system data and control data flows among the three.

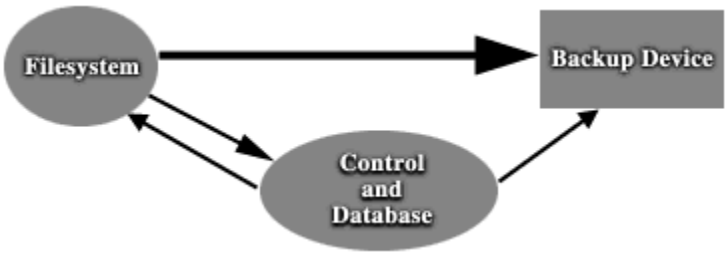

Figure 1: Common Backup Architecture

Enterprise data is stored in file systems. During a backup, this data is copied to the backup device (e.g., tape in a tape library or jukebox). The backup software controls what is being backed up and manages the database or catalog of the data being backed up.

While this is the common architecture of backup today, individual backup software vendors implement their own "protocols" to manage this data flow. While the architecture remains constant, the system calls differ by software package and, sometimes, by server platform.

NDMP and the Common Architecture

NDMP is an open network protocol that defines common functional interfaces used for these data flows. With NDMP, vendors use common interfaces for common architecture data flows. File system data flows from the file system to the backup device using a common interface, regardless of the platform or device. Control or file meta data is passed to and from the backup software using common interfaces, regardless of the software package.

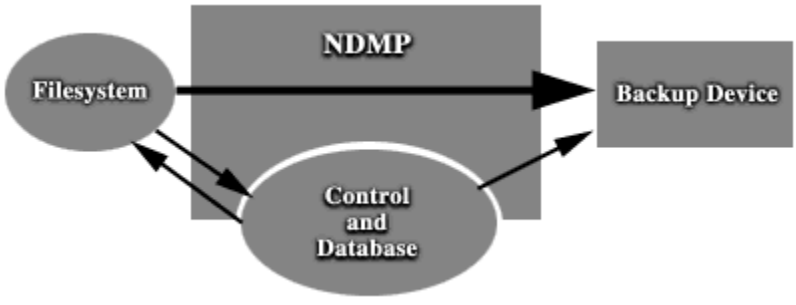

Figure 2: Common Architecture of Backup with NDMP

The Network Data Management Protocol (NDMP): NDMP is a network protocol that specifies the communication between the server and the backup software. Communication is defined using

a series of defined interfaces. These are XDR encoded messages that are exchanged over a bidirectional TCP/IP connection.

The architecture is a client/server model, and the backup software is considered a client to the NDMP server. For every connection between the client on the backup software host and the NDMP host there is a virtual state machine on the NDMP host that is controlled using NDMP. This virtual state machine is referred to as the NDMP server.

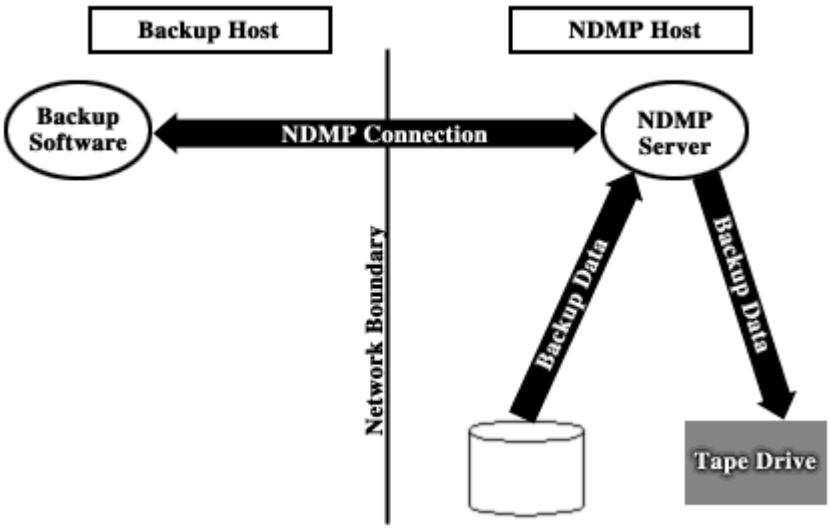

Figure 3: Simple NDMP Configuration

Server Implementation of NDMP

The server vendor implements NDMP by including connect, configure, data, tape, and SCSI interfaces in the operating system code.

Connect Interface:
: This allows authentication of the client and negotiates the version of the protocol being used.

Configure Interface:
: This interface allows the backup software to discover the configuration of the NDMP server.

Data Interface:
: This interface deals with the format of the backup data. Backup software initiates backups and restores using this interface.

Tape Interface:
: The tape interface provides complete control of a tape drive.

SCSI Interface:
: The SCSI interface allows low-level control of SCSI devices such as jukeboxes.

Backup Software Vendor Implementation of NDMP

Backup software adopting NDMP implements notify, file history, and logging interfaces.

Notify Interface:

This notifies the backup software that the NDMP server requires attention.

File History Interface:

This interface allows the NDMP server to make entries in the file history of the backup software. The file history building is important for later file retrieval.

Logging Interface:

This interface provides messages that are used by the operator to monitor backup progress and diagnose problems.

Backup Device Vendor Implementation of NDMP

Jukebox vendors need only ensure that they remain compliant with SCSI standards. Device drivers are not required in an NDMP environment.

The Standardization of NDMP

The full specification for NDMP was submitted to the Internet Engineering Task Force (IETF) in October 1996. Through the Internet Draft and Request For Comment (RFC) processes, the specification will continue to evolve and gain widespread industry support.

The full specification and additional NDMP information is published at www.ndmp.org/info.

NDMP was co-developed by Legato and Network Appliance. The two companies authored the specification and implemented the protocol as a proof of concept.

With a proven concept and working specification in place for NDMP, Intelliguard and Network Appliance approached a number of other companies asking for support. Initially, 17 vendor companies endorsed the initiative to make NDMP an open standard backup protocol. Now that the protocol has been submitted to the IETF, and more companies are hearing about the initiative, support for NDMP is growing. NDMP-compliant products will be available to users as vendors implement the protocol in their future product releases.

Intelliguard and Network Appliance already ship NDMP-compliant backup software and file servers.

Conclusion

NDMP provides a logical partitioning of the backup activity, through a series of well-defined interfaces that addresses the flow of file system and control data in the backup and restore process. System and backup software vendors add a limited amount of code in their software, as defined by NDMP.

NDMP-compliant servers and devices ship "backup ready." Once attached to the network, the NDMP-compliant backup software can provide backup protection to the server as part of an enterprise-wide solution. NDMP compliance provides true plug-and-play interoperability to users. Users can choose the best enterprise-wide backup software solutions and hardware to meet the demands of their particular environment.

NDMP-compliant server vendors can concentrate resources on improving file system internals and tape transfer mechanisms, assured that these enhancements can be utilized by all NDMP-compliant software solutions. As such, enhancements can be utilized by all of the server vendor's customers.

Backup software vendors implementing NDMP are free to redirect resources to feature enhancements that are available to all of their customers, regardless of platform.

NDMP is openly available and will continue to evolve through the IETF and the NDMP Task Force. As NDMP evolves, it is anticipated that increasing functionality will be added. NDMP continues to gain vendor support, and NDMP-compliant products will be available very soon. Continued vendor support and the release of compliant products will make NDMP a standard whose time has arrived.

While vendor development efforts are simplified by NDMP, the real winners are our mutual customers—the I.S. user and management community.