



Education

Disk and Tape Backup Mechanisms

Michael Fishman,
Chief Technology Officer
Backup Platforms, EMC

SNIA Legal Notices

- The material contained in this tutorial is copyrighted by the SNIA
- Member companies and individuals may use this material in presentations and literature under the following conditions:
 - ◆ Any slide or slides used must be reproduced without modification
 - ◆ The SNIA must be acknowledged as source of any material used in the body of any document containing material from these presentations
- This presentation is a project of the SNIA Education Committee

About SNIA and the DMF

About the Storage Networking Industry Association (SNIA)

- ◆ SNIA's primary goal is to ensure that storage networks become complete and trusted solutions across the IT community
- ◆ For additional information about SNIA see www.snia.org
- ◆ SNIA's "Dictionary of Storage Networking Terminology" is online at www.snia.org/dictionary

About the SNIA Data Management Forum (DMF) www.snia-dmf.org

- ◆ The DMF is a sub-group of SNIA acting as the worldwide authority on Data Management, Data Protection and ILM
- ◆ The DMF is a collaborative storage industry resource available to anyone responsible for the accessibility and integrity of their organization's information.

| DMF | | |
|---|---|--|
| Data Protection Initiative (DPI) | Information Lifecycle Management Initiative (ILMI) | Long term Archive and Compliance Storage Initiative (LTACSI) |
| Defining new approaches and best practices for data protection and recovery | Developing, teaching and promoting ILM practices, implementation methods, and benefits | Addressing challenges in developing, securing, and retaining long-term digital archives |

Abstract

Disk and Tape Backup Mechanisms

Extending the enterprise backup paradigm with disk-based technologies allow users to significantly shrink or eliminate the backup time window. This tutorial focuses on various methodologies that can deliver an efficient and cost effective disk-to-disk-to-tape (D2D2T) solution. This includes approaches to storage pooling inside of modern backup applications, using disk and file systems within these pools, as well as how and when to utilize virtual tape libraries (VTL) within these infrastructures.

- **Fundamental concepts for modern, high-performance, scalable systems for Data Protection using networked storage**
- An Overview of Backup Mechanisms for Storage Networks
- Conclusions
- Appendix

Data protection is about data availability

There are a wide variety of tools available to us to achieve that goal, including backup, restoration, replication and recovery, but it is critical to keep focused on the actual goal -- availability of the data -- and to balance how we achieve that goal by using the right set of tools for the specific job. Held in the balance are concepts like data importance or business criticality, critical to the business operations team and budget, speed, and cost of downtime, usually the concerns of backup administrators.

DDRR - The Process

Detection

- Corruption or failure noted

Diagnosis / Decision

- What went wrong?
- What recovery point should be used?
- What method of recovery -- overall strategy for the recovery?

Restoration

- This phase involves moving data -- from tape to disk or disk to disk -- from the backup or archive (source) to the primary or production disks.

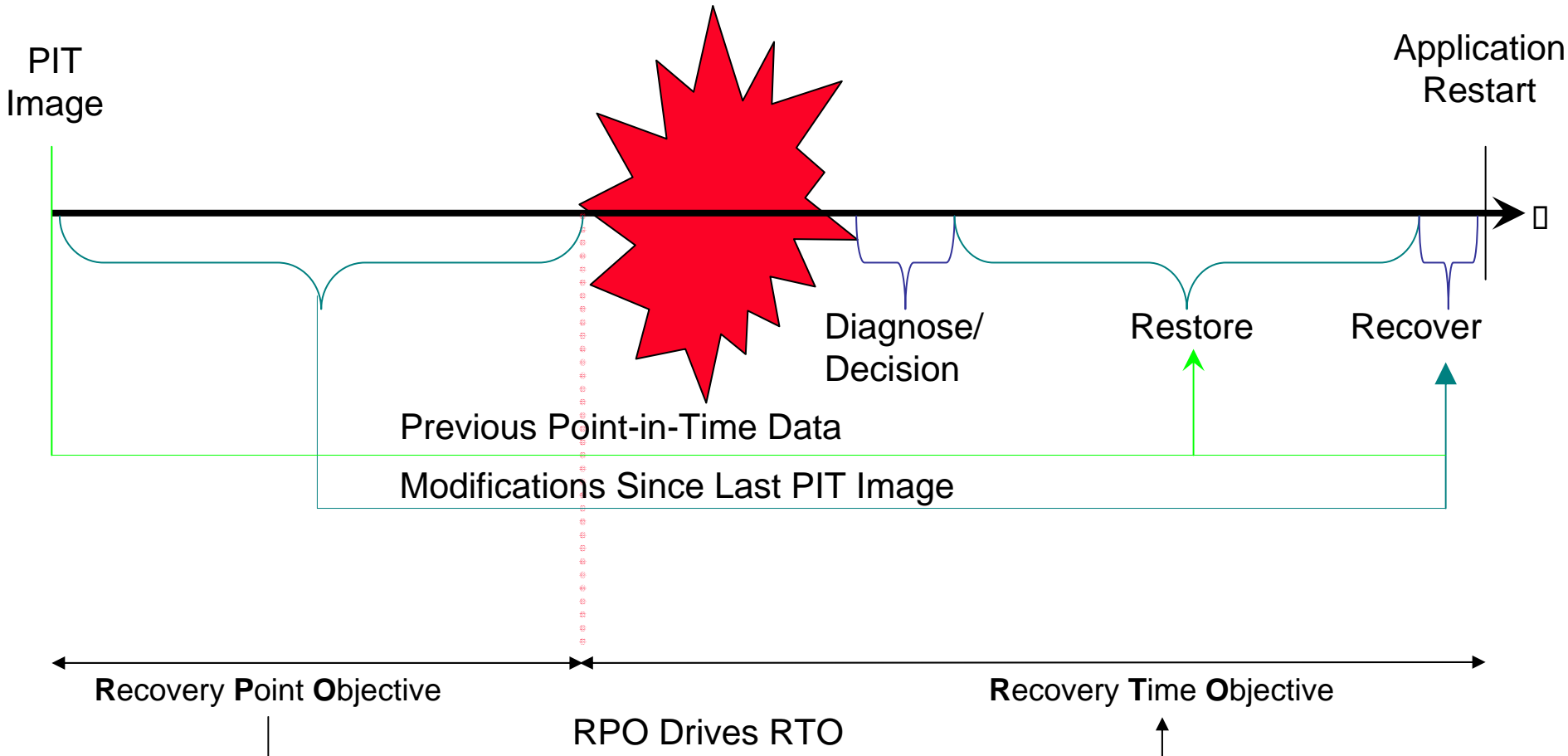
Recovery

- Final phase: application environment perform standard recovery and startup operations -- log replays for a database or journals replays for a file system.

- ***RPO - Recovery Point Objective***
 - The maximum desired time period prior to a failure or disaster during which changes to data may be lost as a consequence of recovery. Data changes preceding the failure or disaster by at least this time period are preserved by recovery. Zero is a valid value and is equivalent to a "zero data loss" requirement

- ***RTO - Recovery Time Objective***
 - The maximum desired time period required to bring one or more applications and associated data back to a correct operational state.

RPO & RTO



Methodologies of Backup

➤ Cold

- ◆ Systems are offline for the period of time it takes to capture an image of the data in total. As backup window shrinks and data size expands, cold backup becomes untenable. If possible, this is the easiest and cheapest form of protection.

➤ Application Consistent

- ◆ Application supports ability to take pieces of overall data set offline for a period of time to protect it - application knows how to recover from a collection of individual consistent pieces. No downtime for backup window.

➤ Crash Consistent or Atomic

- ◆ Data can be copied or frozen at the exact same moment across the entire dataset. Application recovery from an atomic backup performs like a high availability failover. No backup window.

Protection Design Trade-off

- What's most important:
 - ◆ Backup Window (speed of backup)
 - ◆ Recovery: RTO (speed of recovery, cost of downtime)
 - ◆ Recovery: RPO (amount of data loss)
- There are trade-offs everywhere
 - ◆ Newer technology minimizes trade-offs
 - But they still exist!!!
 - ◆ Need to identify the priority order, and establish SLA targets for each piece (Backup, RTO, RPO)
 - ◆ Ease of use, audit-ability, and error recovery need to be considered too...

“Pay me now, or pay me later!”

Backup versus Disk-based Replication

➤ Backup

- ◆ Protecting data by making copies or allowing copies to be generated from saved data
- ◆ Examples: snapshots, split mirrors, VTL, tape, CDP
- ◆ When?
 - > **Multiple Recovery Points needed**
 - > **Recovery from data corruption**
 - > **Archival and indexing**

➤ Mirroring/Replication

- ◆ Protecting data by moving the data, usually as it changes, to a remote copy. Synchronous or Asynchronous mirroring
- ◆ When?
 - ***Disaster Recovery Time Objective (DR/RTO centric usually)***
 - ***Data Migration***
 - ***Content Distribution***

- Fundamental concepts for modern, high-performance, scalable systems for Data Protection using networked storage
- **An Overview of Backup Mechanisms for Storage Networks**
- Conclusions
- Appendix

B/R Topology Components

Runs the **backup engine** to get the data from the platform to backup

Collects the **data** and **meta-data** accordingly to the requested level of abstraction

Collects the data from the agent to redirect them on a secondary storage device

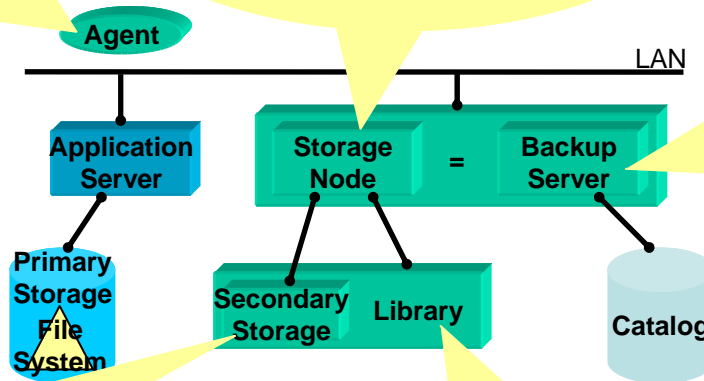
May offer some **caching** functionality for stream manipulation like :

- Multiplexing (aka fan-in)
- Multiple streams handling (aka fan-out)

Central point of **administration** and management for the backup application

Usually concentrates all the meta-data information in a repository (aka **catalog**) that may be built from a collection of flat files to a real database

May offer some **redundancy** or disaster recovery methods for securing the catalog



Device on which the **data** are **written**

If plugged on a SAN, it is **shareable**. Thus the backup application must :

- authorize multiple attachments of this same physical device
- handle dynamic bus address change (aka serialization)

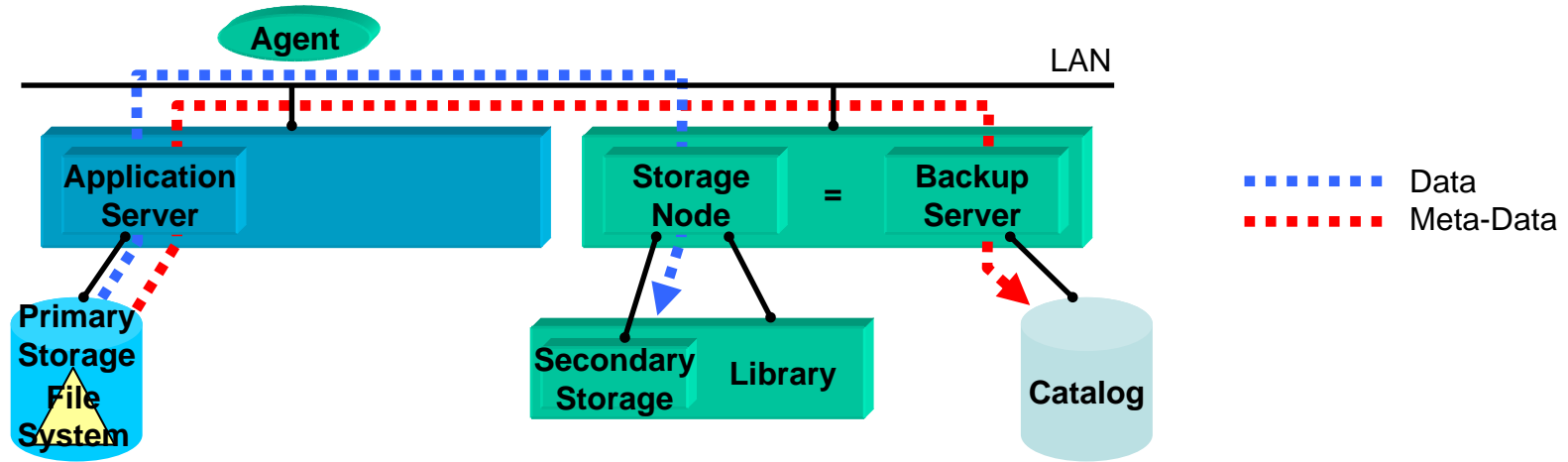
Mechanical device **moving removable media** in secondary storage devices

Made up of **slots, mailbox, picker** and locations for secondary storage devices

Usually attached only once and driven from a single host

The backup software may enable a logical split to create as many different logical smaller Libraries

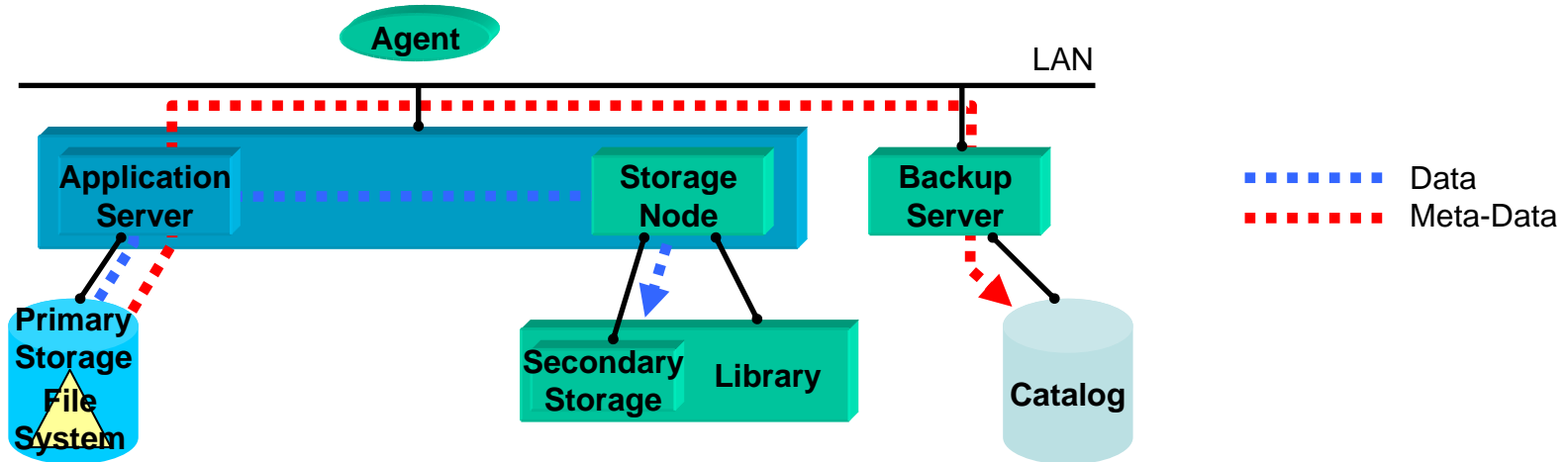
Traditional LAN Backup



In this mode, the application server reads and sends the data over the LAN

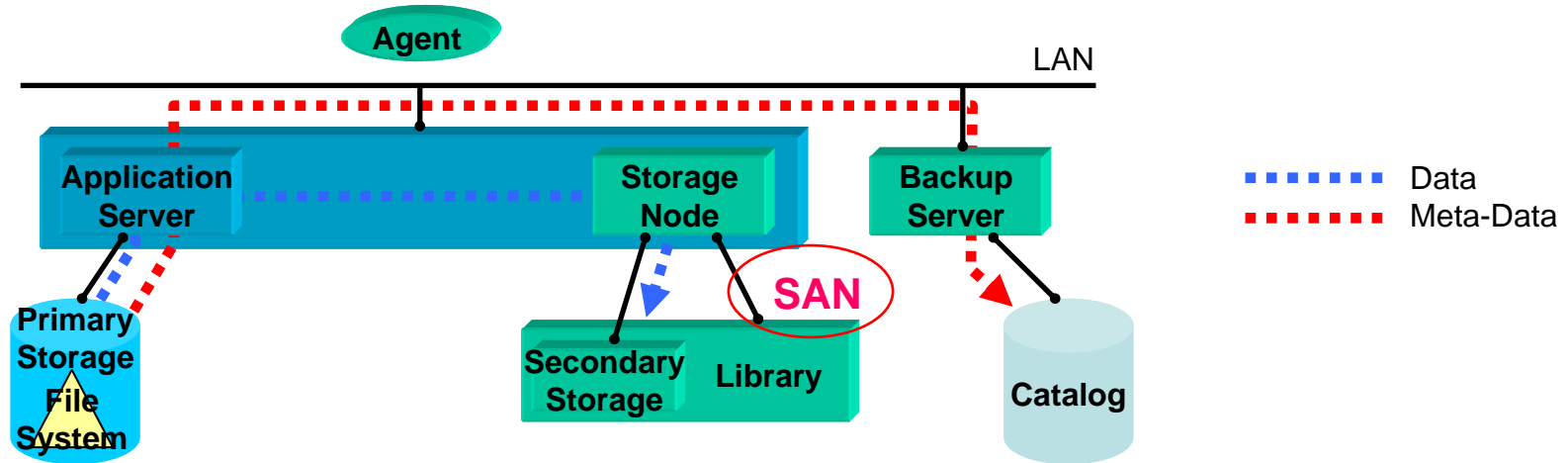
- Data and meta-data go to the backup server
- **LAN impacted by NAS, iSCSI, Backup & Mirroring/Replication!**
- Application server impacted by primary storage I/O
- Various network protocols: NFS, CIFS, iSCSI, proprietary protocol, NDMP

Traditional Local Backup



In this mode, the application server reads and writes the data locally

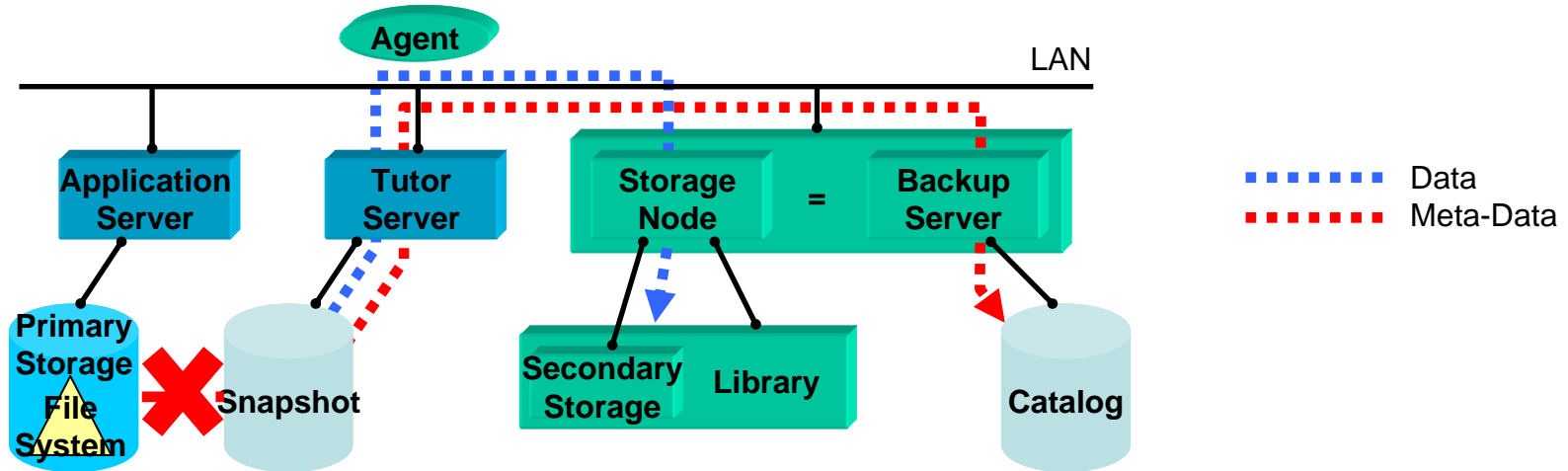
- Secondary storage statically attached to the application server
- Meta-data go to the backup server
- **No LAN impact – except meta-data (significant for millions of files)**
- Application server impacted by both primary and secondary storage I/O
- May be done by proprietary protocol or NDMP



In this mode, the application server reads and writes the data locally through the SAN

- Secondary storage is:
 - Attached to the application server via the SAN
 - Resources attached and shared among multiple application servers
 - Dynamically allocated by the backup application
- Meta-data go to the backup server
- **No LAN impact – except meta-data (significant for millions of files)**
- Application server impacted by both primary and secondary storage I/O

Application Server-free Backup

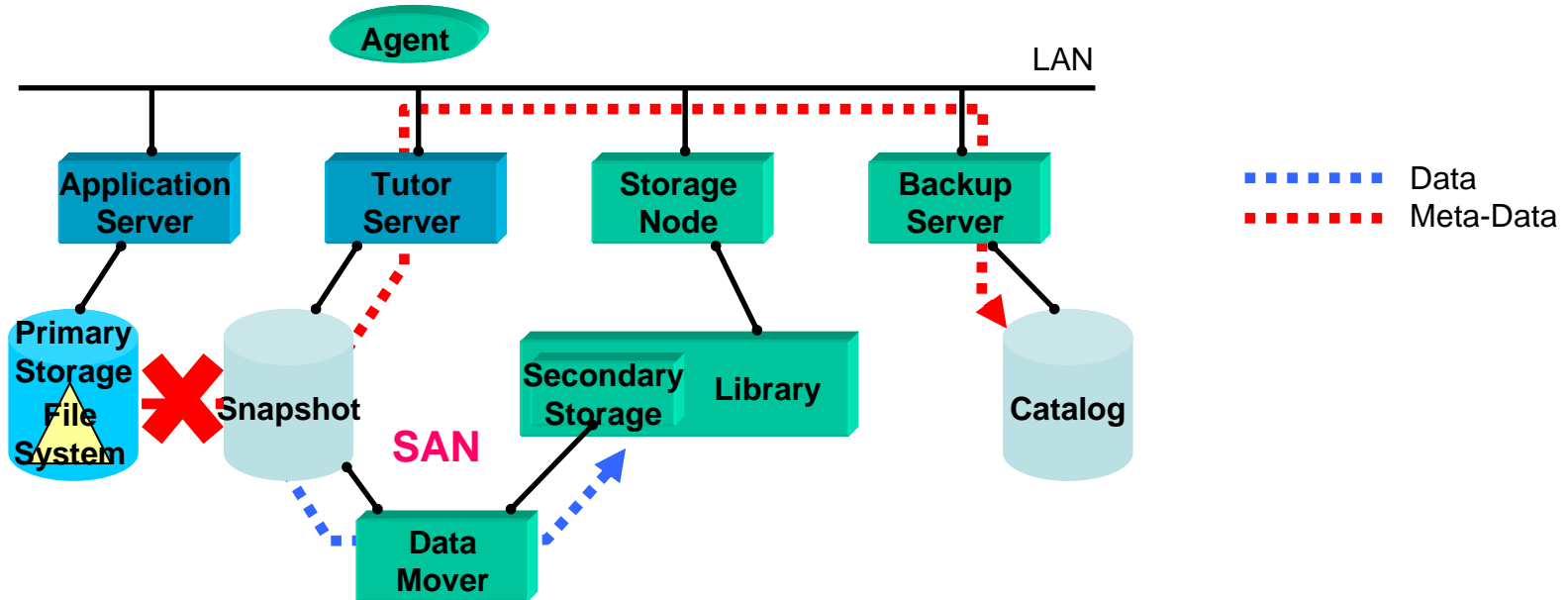


In this mode, the application server gives a **snapshot** of the primary storage volume to a **tutor server** that reads and sends the **data over the LAN or SAN**

- **Tutor server (Slave server) must understand the volume structure**
- Snapshot volume may be mounted on tutor server or extent list map created
- Meta-data (and perhaps data) go to the backup server
- **Mirror** : Application server impacted when re-silvering the mirror

Or

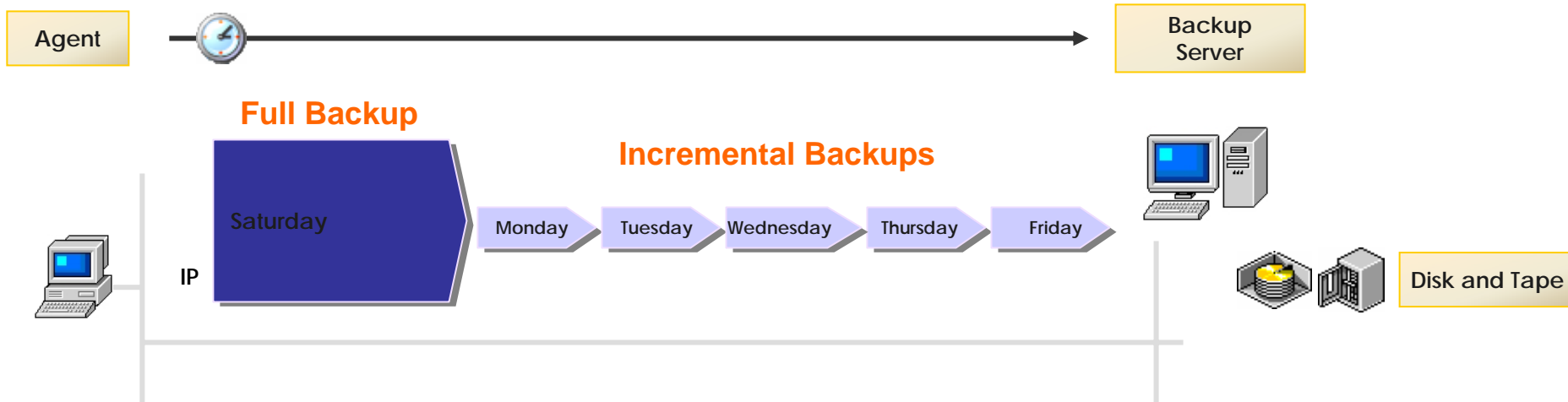
- **Snapshot** (shared blocks) : Application server impacted by volume access



In this mode, the server **delegates** the primary storage I/O processing to a “**Data-mover**” device on the SAN, using SCSI **Extended Copy** (XCOPY or “Third-Party Copy”) command. Stable image (snapshot) backup and extent list meta-data are necessary components of this.

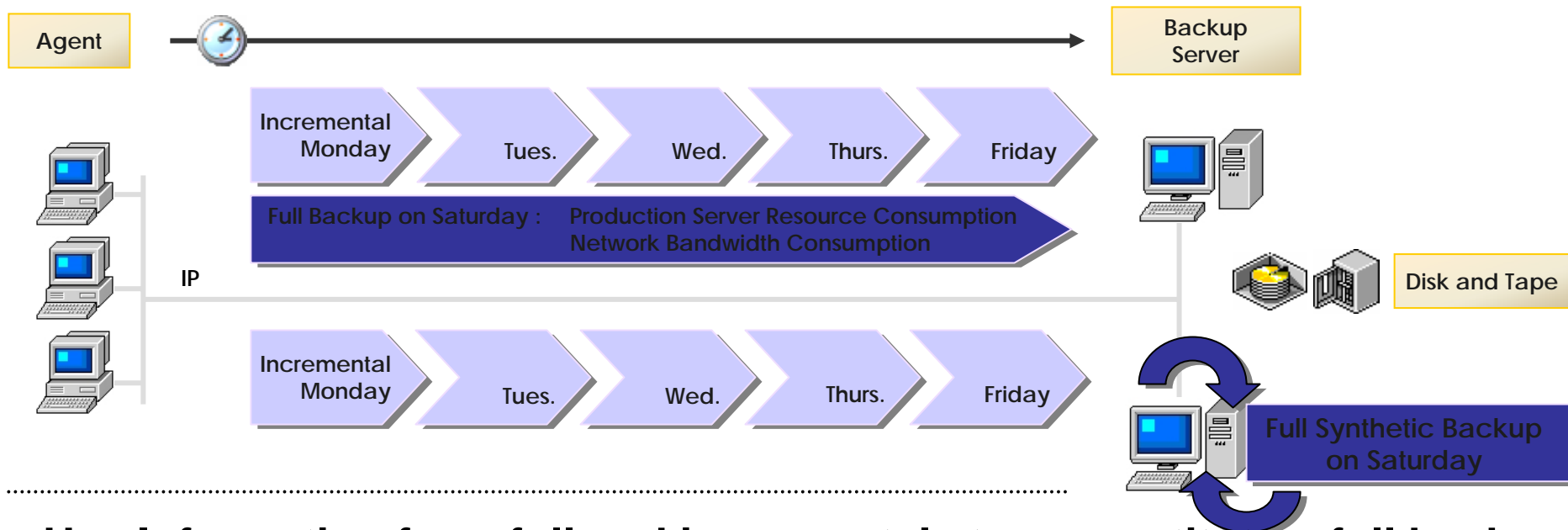
- **Tutor server must understand the volume structure**
- Meta-data go to the backup server but with less granularity
- No LAN impact – there are fewer meta-data
- Application server impacted when re-silvering the mirror or by snapshot volume access
- No storage node impact

Full and Incremental Backup



- **Full Backup --**
 - **Everything copied to backup (cold or hot backup)**
 - **Restoration straight-forward for cold backup (and, typically, hot backup)**
 - **Huge resource consumption (server, network, tapes)**
- **Incremental Backup --**
 - **Only the data that changed since last full/incremental copied**
 - **Massive reduction in data moved/copied**
 - **Can lead to nightmares in restoration and recovery**

Synthetic Full Backup

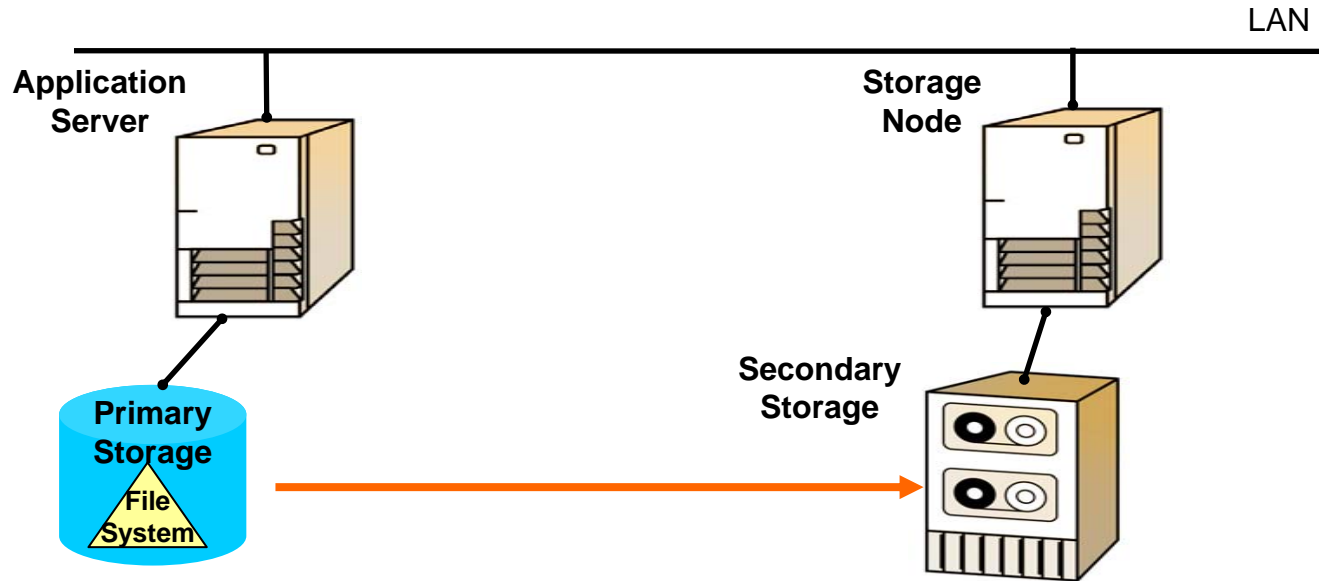


- **Use Information from full and incrementals to reconstitute a full backup**
 - No server involvement
 - Minimize network resource consumption
 - Not without it's challenges like long processing times and metadata consistency

Tape Technology

- Sequential technology
 - ◆ Versus random access
- Can be sequestered
 - ◆ Away from prying hands, DR
 - ◆ Legal - regulated industry?
- Media replacement costs
 - ◆ Tape life, reusability
 - ◆ Reliability
- Performance and Utilization
 - ◆ Streaming and multiplexing
- Libraries and Robotics
- Removability

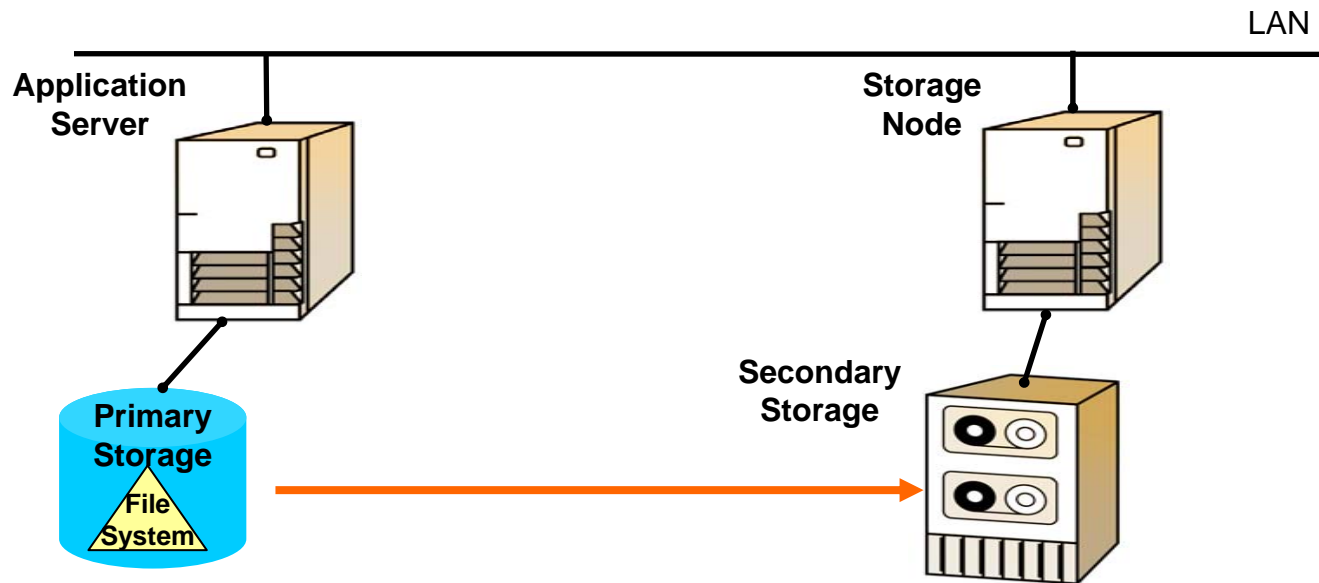
Tape Backup



Traditional approach:

- Data go **over LAN** from the Primary Storage to the Secondary Storage
- The Secondary Storage is composed of a **Media Manager** and **Tape Drives**
- Typically managed by a B/R application

Tape Backup: Pros and Cons



Pros:

- Lower costs
- Established process
- Off-site media storage

Cons:

- Performance
- Device sharing
- Reliability of devices and media
- Risk of Tape loss

Tape challenges for Enterprise customers

- **Speed and Reliability**
 - Need to reduce backup window
 - Need faster, more reliable restores
- **Management**
 - Media management of tape-based B/R program can be complex
- **Cost**
 - ROI analysis of hard and soft costs
(Equipment costs, maintenance, media, down-time)
- **Device sharing**
 - Complexity, schedule risk, resource conflicts

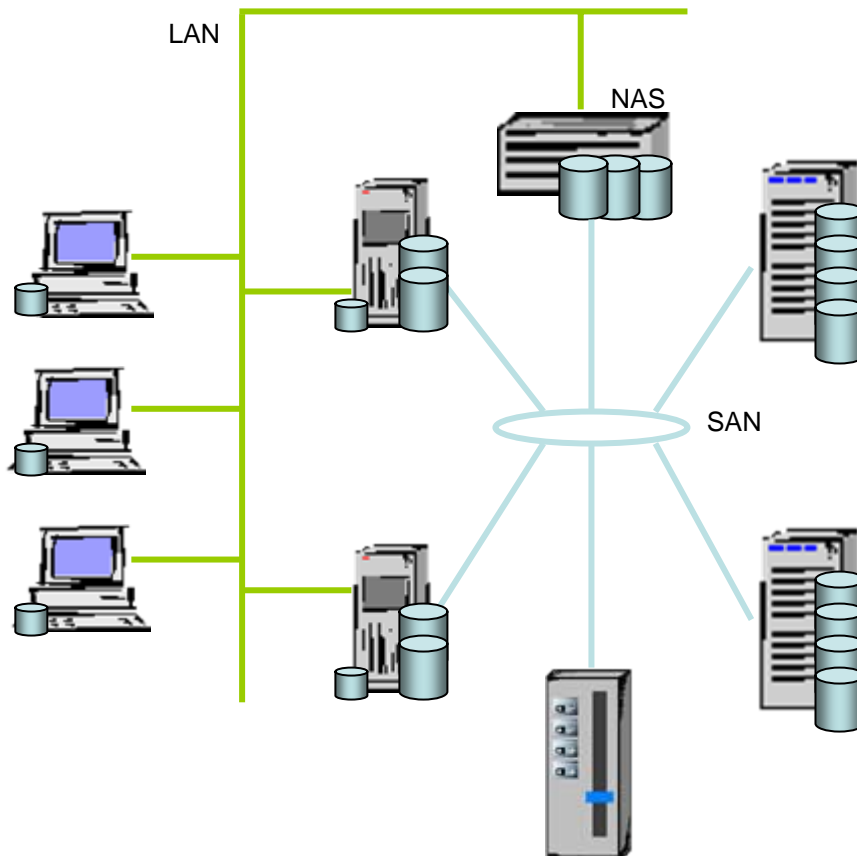
Backup to Disk (B2D)

- Replaces primary backup target with disks
- Advantages:
 - ◆ Speed (Backup and Restore)
 - ◆ Scale
 - ◆ Fewer shared devices
- Fibre Channel Disks versus ATA versus SAS
 - ◆ I/O per second random access, MB/s sequential
 - ◆ Large # of small I/O's, Small numbers of large I/O's
- Backup to disk may require updates to backup software or extra modules

- Originated in mainframe environments
 - Multiple applications
 - Many data sets with long data life cycle times
- Integration into Open Systems
 - Designed to avoid changes to backup software
 - Packaged, formatted, optimized
 - Eliminate media handling, improves BU and Restore

Virtual Tape Advantages

- **Speed and reliability**
 - › Single stream performance exceeding tape drives
 - › Aggregate performance without multiplexing
 - › No mechanical failures, No robots
 - › Inherent RAID protection
- **Management**
 - › Plug and play in existing back-up environment
 - › Appears as open systems tape cartridges, drives, libraries
 - › Compatible with existing backup & restore operations
 - › Easily integrates with today's off-site processes
- ◆ **Cost**
 - › Cost effective & scalable
 - › Leverage lower cost disks
- ◆ **Reduced device sharing**
 - › Reduces complexity, schedule risk, resource conflicts

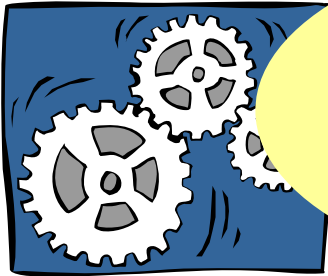


Constant change and **heterogeneity** in technologies

- Operating Systems
- Disk Storage Appliances
- Network Architectures / Topologies
- Tape Storage Devices

Challenge

- **Protect** mission-critical data
- **Timely** backup and restore
- **Administration** overhead
- **Optimize** storage resources

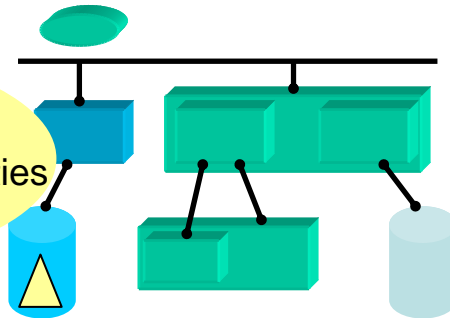


Heterogeneity

- Multiple platforms (HW / OS)
- Multiple tape drives & libraries
- Multiple applications
- NAS and SAN

Leverage architecture

- **Snapshot** facilities
- **SAN, LAN**



Open

- Use of **standards**
- API, command line

Security

- **Authentication**
- DMZ
- **Firewall** support



Advanced tape management

- **Tape mirroring**
- **Off-site storage**

Advanced library management

- **Sharing, splitting**
- Port handling



Easy Installation
and
deployment


Centralized Administration

- GUI & smart interface
- Backup strategies
- Scheduling
- Media management



Centralized Supervision

- **Real-time** monitoring
- Alarms
- Event log
- SNMP compliant, integration with **Frameworks**

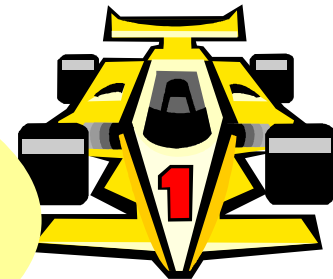


Accounting

- **Reporting** about Protection services
- Allows internal or external **billing**

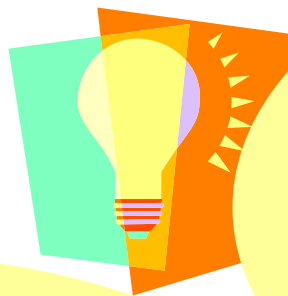
Performance

- **Caching** on storage node for **multiplexing** and **stream** management
- Synthetic full backup – **offload backup engine**
 - faster backup = Incrementals
 - faster restore = restore from full session
 - LAN-free full backup from incremental



Manageable by user

- **User profile** and authentication
- Delegate: not only an administrator tool



Smart recovery functions

- Most recent image
- **Consistent view** = true image
- **Minimize downtime** :
 - **time to diagnose**
 - **time to restore**
- Lost time + downtime = total loss time

Please send any comments on this tutorial to
SNIA: trackdatamgmt@snia.org

**Many thanks to the following individuals for
their contributions to this tutorial:**

**SNIA Data protection Initiative
SNIA “Data Management Forum”
SNIA Tech Council**

**Nancy Clay
Mike Rowan
SW Worth
Mike Fishman
Jason lehl**

Get Involved !

www.snia-dmf.org

- Find a passion**
- Join a committee**
- Gain knowledge & influence**
- Make a difference**

Thank you for your feedback

Questions and Answers



Education

APPENDIX

www.snia.org/education/tutorials

« A **collection of data** stored on (usually **removable**) non-volatile storage media for purposes of recovery in case the original copy of data is lost or becomes inaccessible. Also called *backup copy*. To be useful for recovery, a backup must be made by copying the source data image when it is in a **consistent state**...

or contains elements and information enabling a consistent state to be recovered. »

Source: SNIA Dictionary

Secondary Storage

Standard Data Formats

Posix 1003.1 Archive/Interchange File Format also well known as **tar** or **cpio**

.Pros : implemented and available on **all Unix** systems

.Cons : meta-data **limited to posix** system implementations and designed for sequential **unaltered file level data** (no streams, no compression, no multiplexing, etc.)

SIDF, **S**ystem **I**ndependent **D**ata **F**ormat, focuses on representing system data and file data and meta-data for **all types of operating systems** and platforms in a media independent common format.

▪Pros : handle **all kinds of stream** manipulation and meta-data forms with tags

▪Cons : used privately in backup products, i.e. **no native command** on systems yet

<http://www.cs.wisc.edu/~jgast/sidf/>

Secondary Storage

Proprietary Data Formats

Server vendor based (usually at volume or lower file system level)

- .Pros : maximum support of **vendor specific features** (volume specificities, unique meta-data, etc.)
- .Cons : **no possible interchange** with other platforms

Backup software vendor based

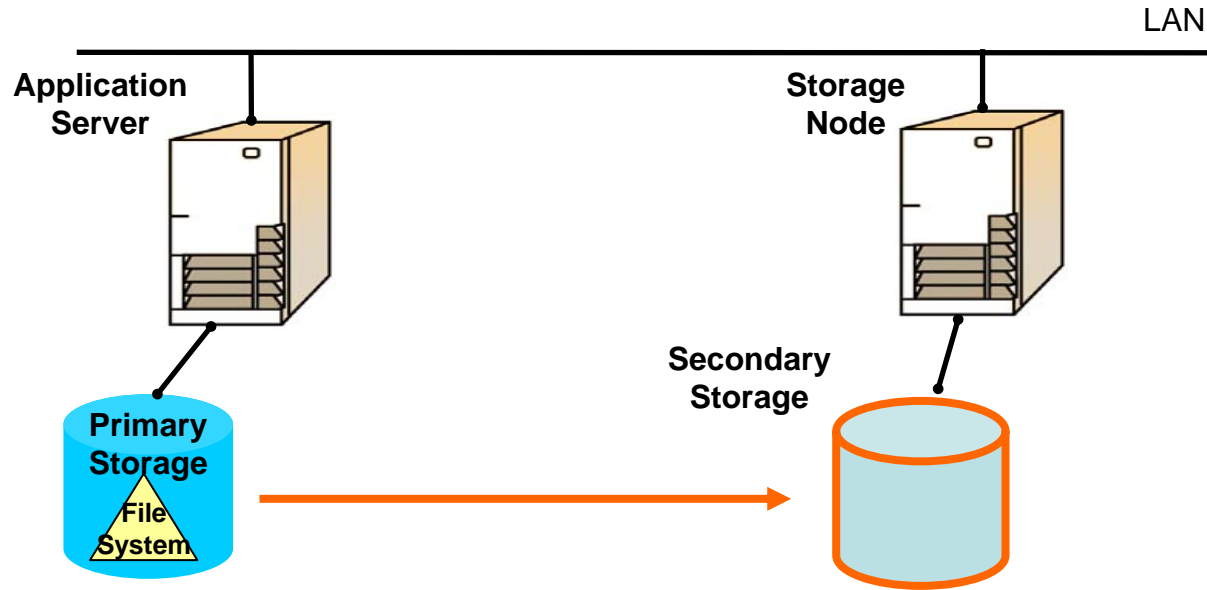
- .Pros : maximum support of vendor **specific features** like multiplexing, compression, etc. – all kinds of implemented **stream manipulation**
- .Cons : **no possible interchange** with other products and must be read via the backup software (may be « processed » sometimes to be read in standard format or standalone command available)

Network Data Management Protocol

- **NDMP** is a general open network protocol for controlling the exchange of data between two parties
- Partition the problem between vendors
- Each vendor implements solutions in compliance with this protocol: Data Server / Tape Server / Client
- Enable « best of breed » **combinations of multi-vendor products** in customer solutions
- Enhanced **interoperability**
- Vendor
 - focus on core competencies
 - Improve user offering
 - reduce time-to-market

<http://www.ndmp.org>

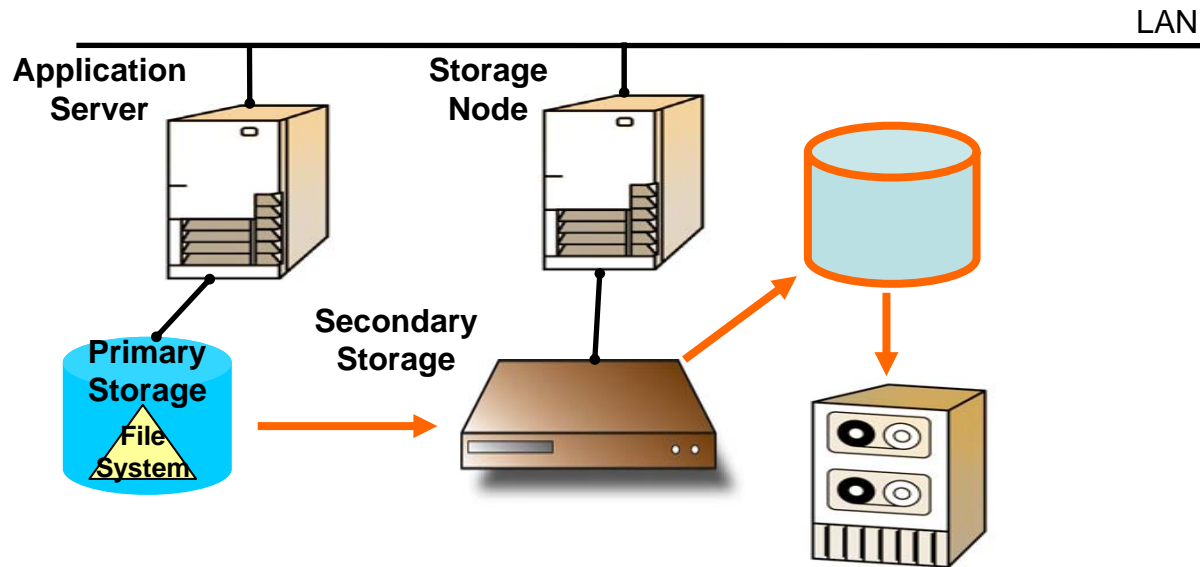
B/R Application Disk Support



First Backup-to-Disk approach:

- Data go **over LAN** from the Primary Storage to the Secondary Storage
- The Secondary Storage is composed of **Disk Drives managed by the B/R Application**
- Storage maybe **block** or **file** based with **raw**, **filesystem** or **sequential format**

Virtualization Approach



Virtualization approach

- Data sent **over LAN** from the Primary Storage to the Secondary Storage
- The Secondary Storage is an appliance filled with **Disk Drives**
- It behaves like a regular **Media Manager** and **Tape Drives**