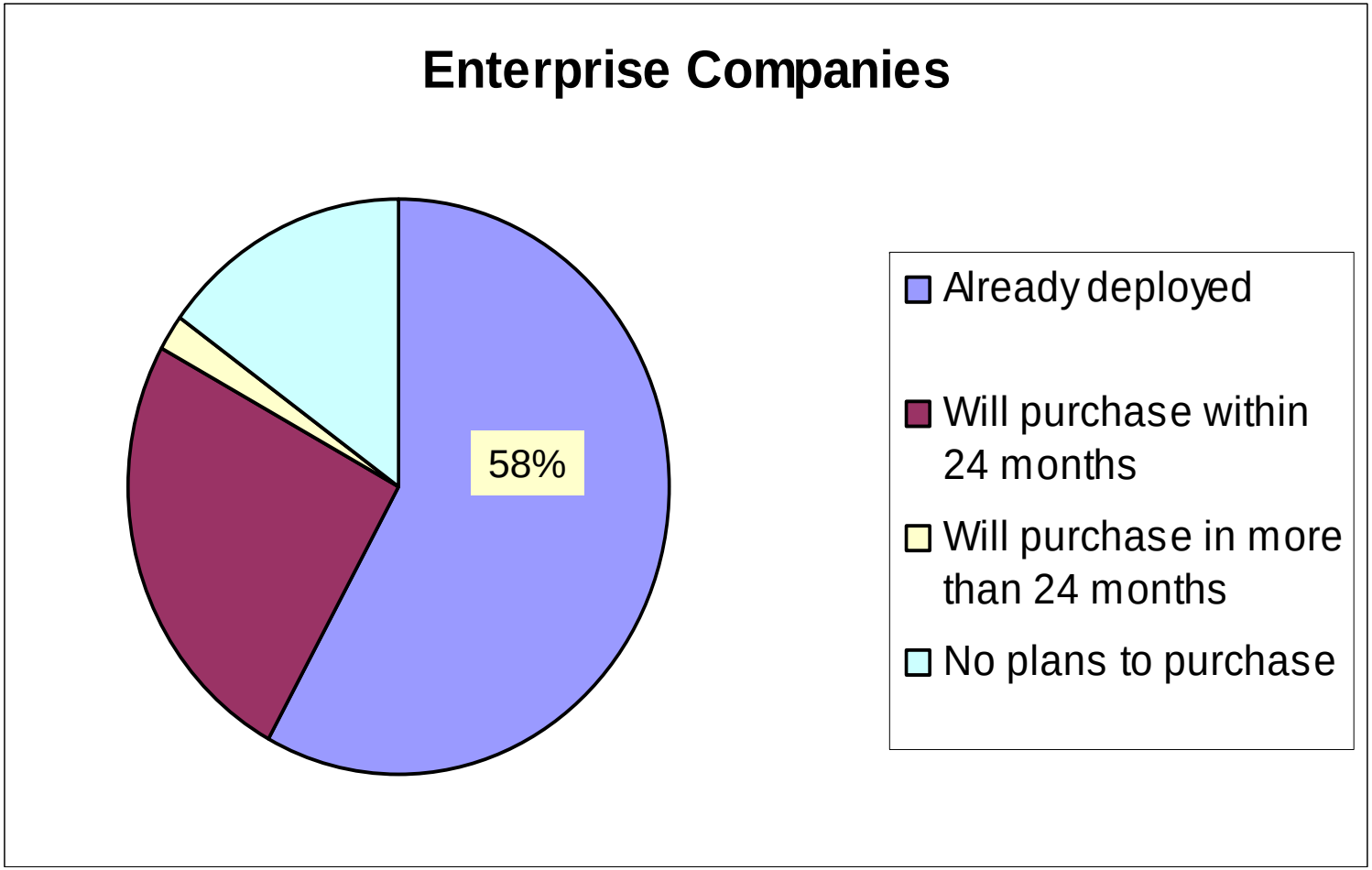
The NDMP v4 File Service Extension

Hugo Patterson

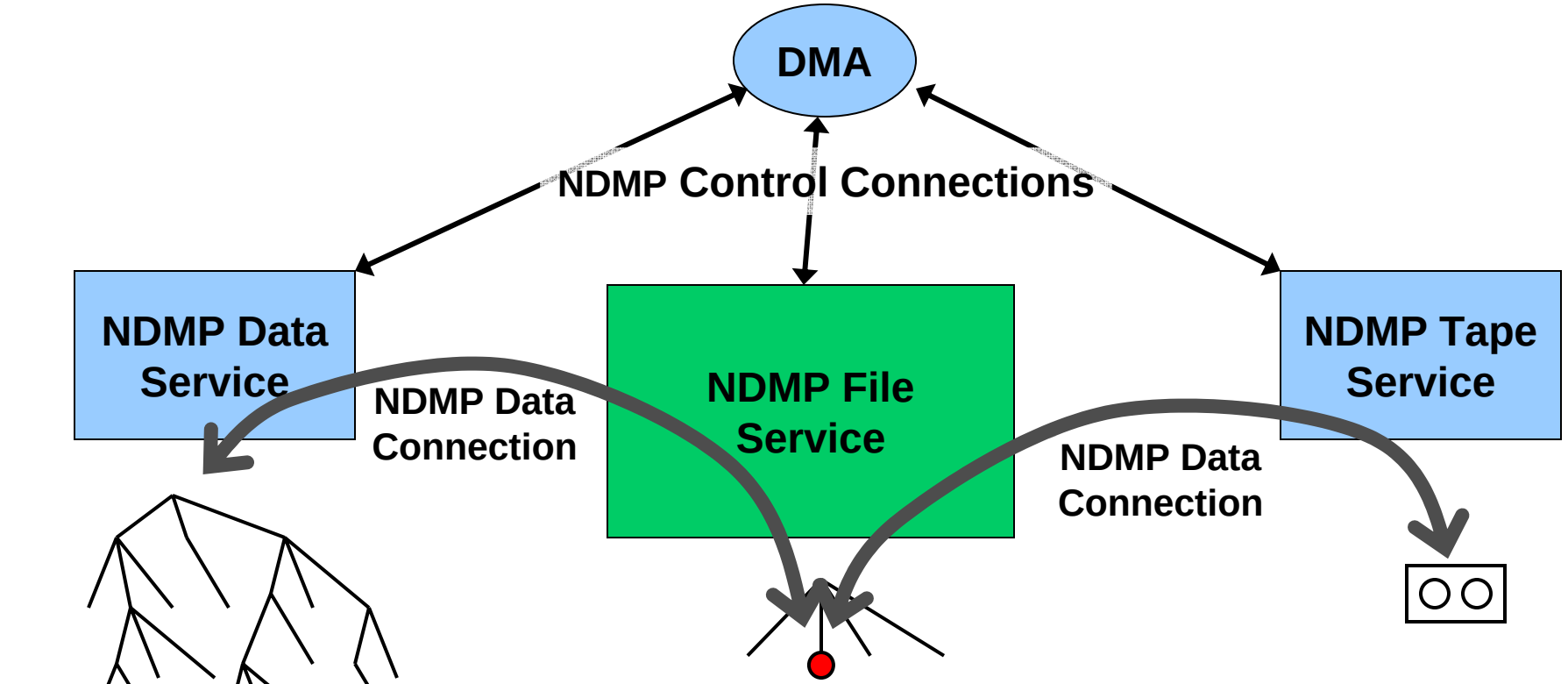
Bob Fozard

Backup to disk is now standard



Source: ESG Research 2004

NDMP File Service for D2D Backup/Restore and Staging



- Store NDMP stream in a file
- Restore from file to NDMP Data Server
- Copy to an NDMP Tape Server later

Outline

- File Service definition and use
- API overview
- Advanced applications
- Summary

File Service highlights

- An NDMP service like the Data and Tape Services
- Controlled by a DMA via an NDMP Control Connection
- Interoperates with existing NDMP Data and Tape Services
- Accepts data from an NDMP Data Connection and stores it in a file
- Reads data from a file and sends it out an NDMP Data Connection
- Can transfer data at a byte granularity from one file to another within the file service
- Supports the Mover interface for familiar operation
- Supports Advanced File Service commands for new applications (more on this later)

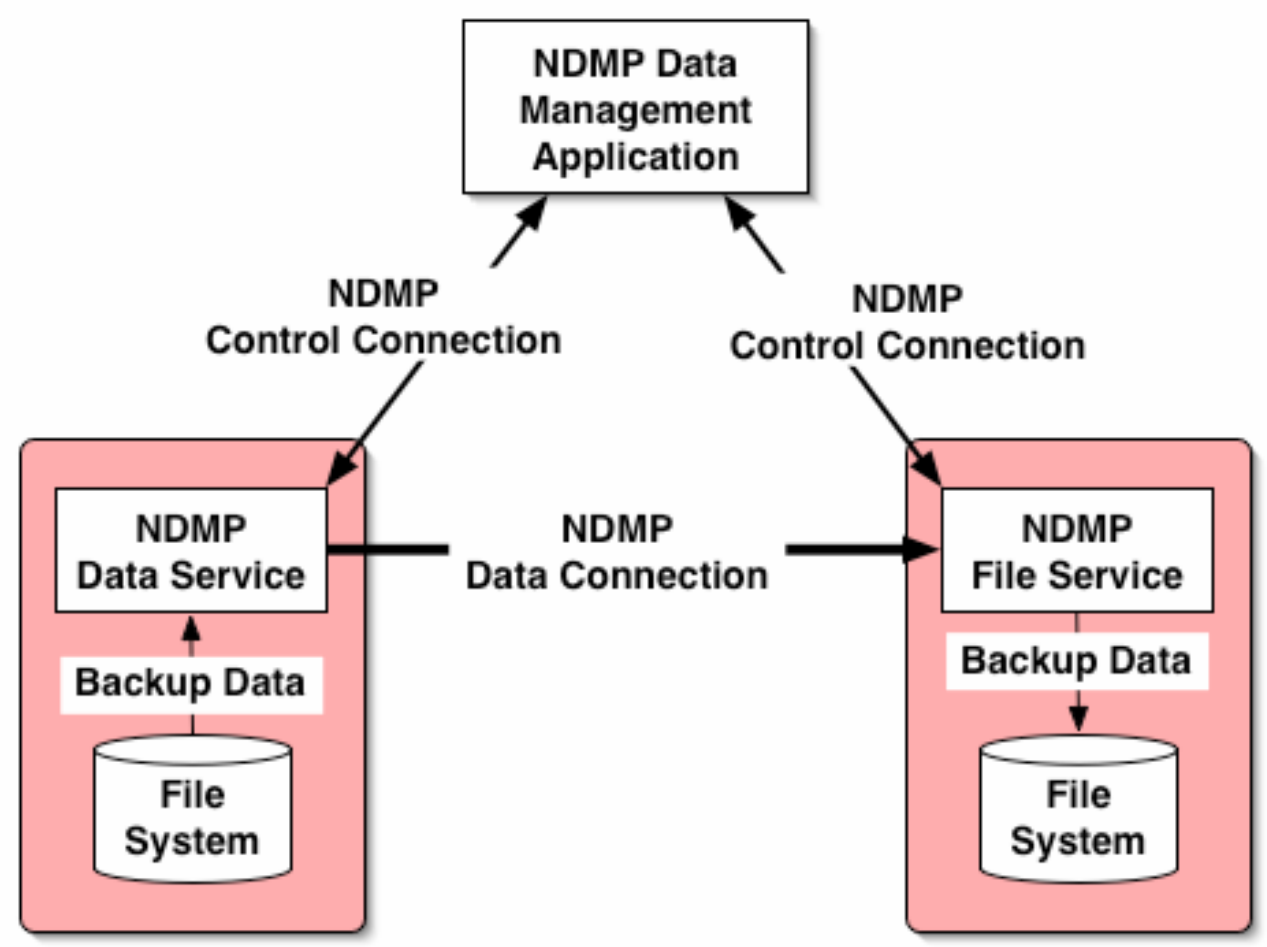
Basic use cases

- **Backup/restore a filer with disk via NDMP**
  - Delivers all the benefits of disk-based backup to NDMP
    - More reliable
    - Less manual intervention ==> lower TCO
    - Verifiable recoverability
- **Stage from filer to disk and then to tape**
  - Filer backups can't always stream the fastest tape
  - Stage to a disk file first, then stream to tape
  - Good for backing up SnapVault servers

Outline

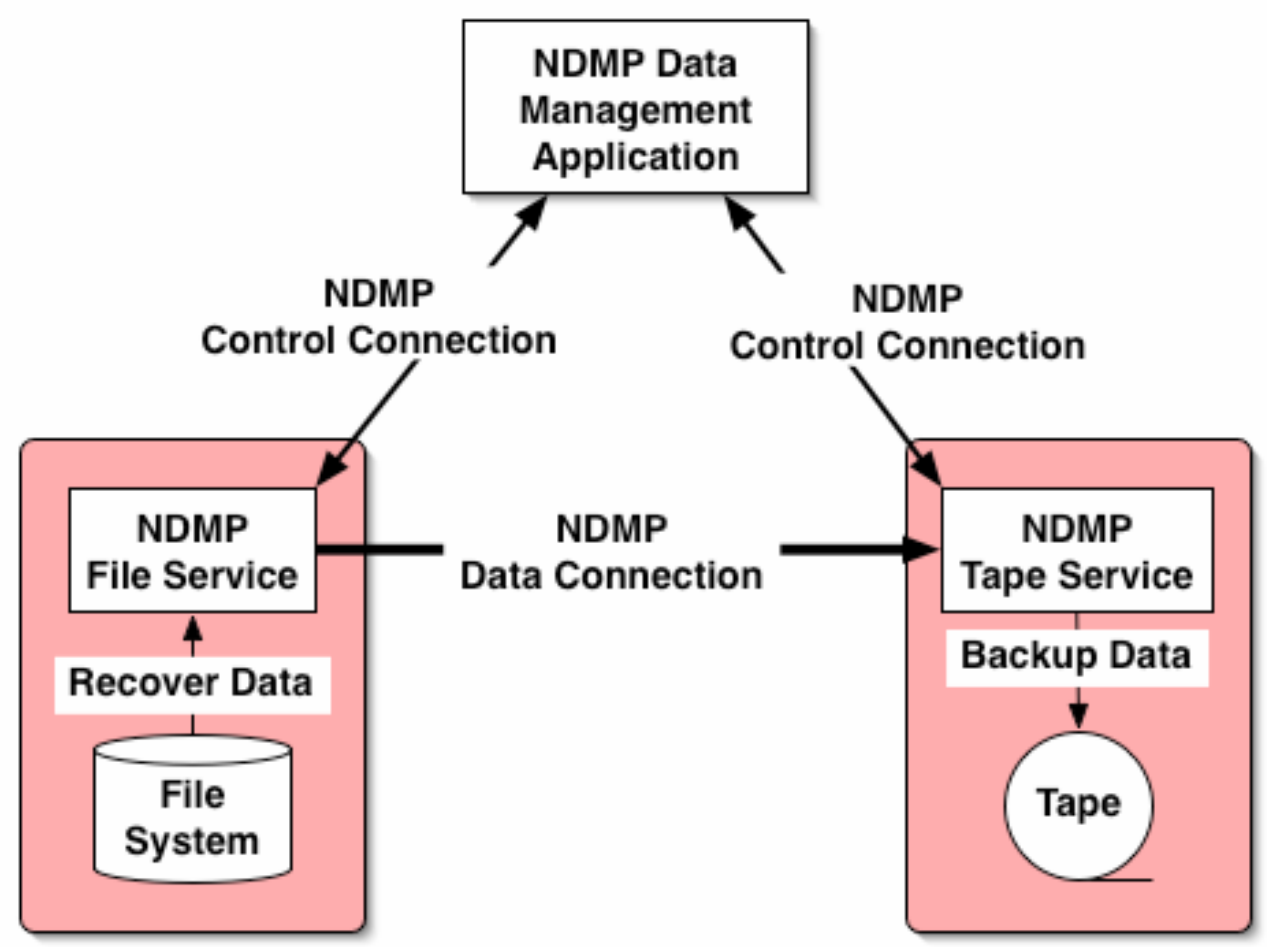
- File Service definition and use
- **API overview**
- Advanced applications
- Summary

Topology: 3-way Data Service to File Service



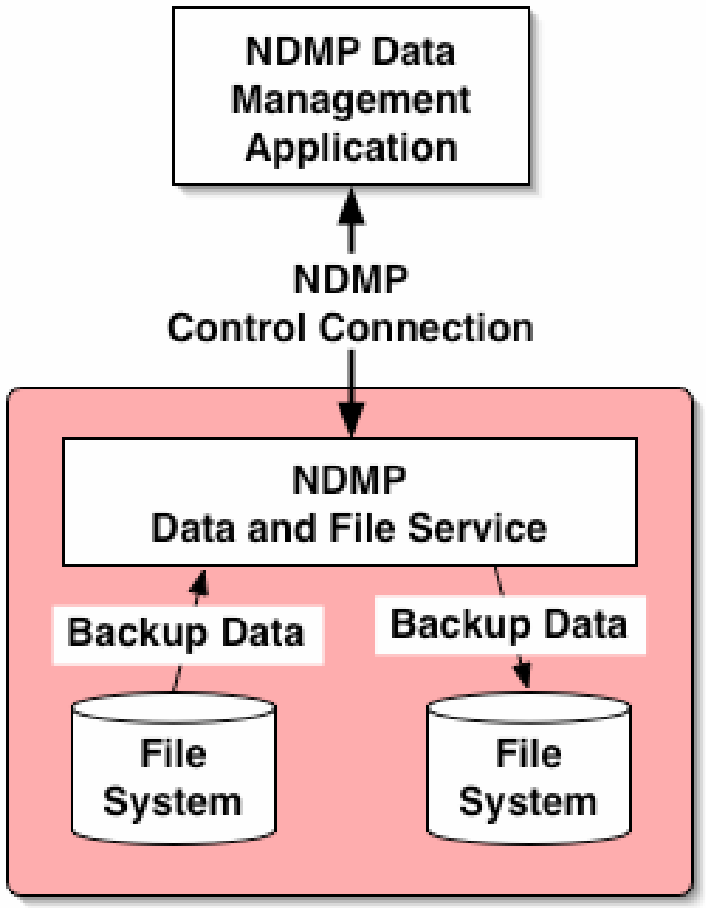
- **Staging or backing up to disk**

Topology: 3-way File Service to Tape Service



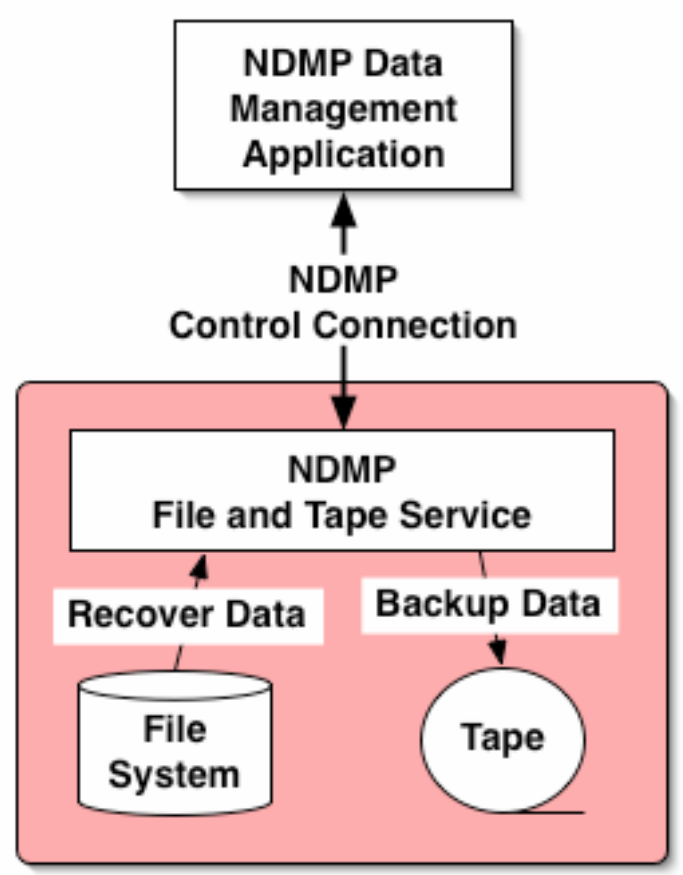
- **Copying to tape**

Topology: Simple Data Service to File Service



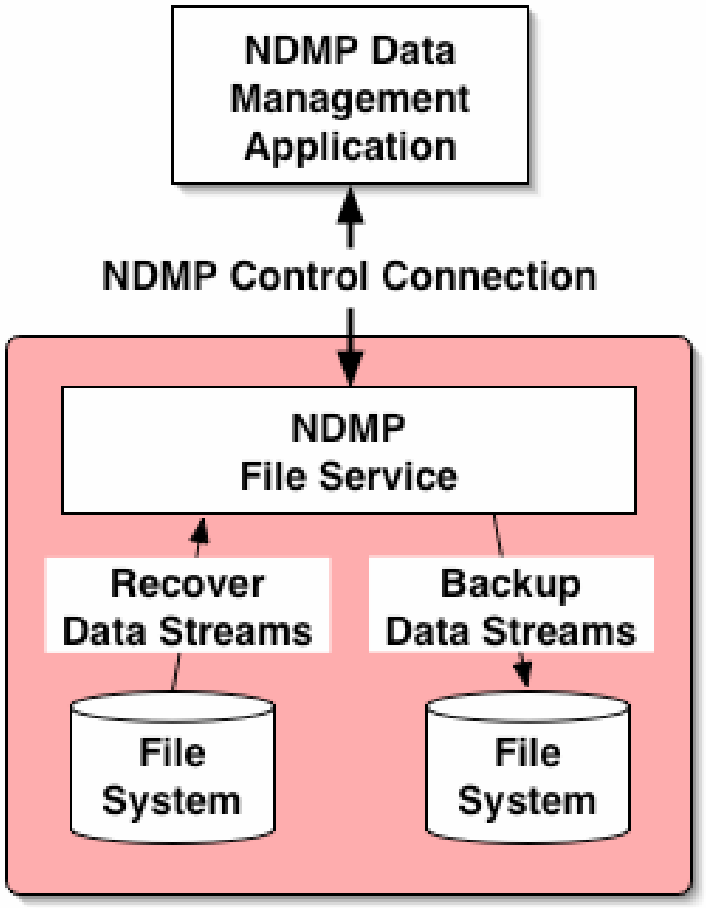
- **Staging to disk within same filer**

Topology: Simple File Service to Tape Service



- Copying to tape within same filer
- Requires use of Advanced File commands

Topology: Simple File to File Service



- Copying, Multiplexing, Demultiplexing, Synthesizing
- Requires use of Advanced File commands

Tape vs. File Backup Workflow Comparison

Backup Phase	Tape Service	Basic File Service	Adv. File Service
Connect/Config	CONNECT_* CONFIG_*	CONNECT_* CONFIG_*	CONNECT_* CONFIG_*
Target Setup	TAPE_OPEN TAPE_MTIO TAPE_WRITE MOVER: _SET_RECORD_SIZE _SET_WINDOW _LISTEN	FILE_OPEN FILE_WRITE / META FILE_BIND_MOVER MOVER: _SET_RECORD_SIZE _SET_WINDOW _LISTEN	FILE_OPEN FILE_WRITE / META FILE_ADV_LISTEN
Data Setup	DATA_CONNECT DATA_START_BACKUP	DATA_CONNECT DATA_START_BACKUP	DATA_CONNECT DATA_START_BACKUP

Tape vs. File Backup Workflow Comparison (cont'd)

Backup Phase	Tape Service	Basic File Service	Adv. File Service
Transfer	DATA_START_BACKUP Additional Tapes: MOVER_PAUSED TAPE_MTIO TAPE_WRITE MOVER_CONTINUE	DATA_START_BACKUP	FILE_ADV_RECEIVE DATA_START_BACKUP
Completion	NOTIFY_DATA_HALTED MOVER_STOP TAPE_CLOSE	NOTIFY_DATA_HALTED MOVER_STOP FILE_CLOSE	NOTIFY_DATA_HALTED FILE_ADV_STOP FILE_CLOSE

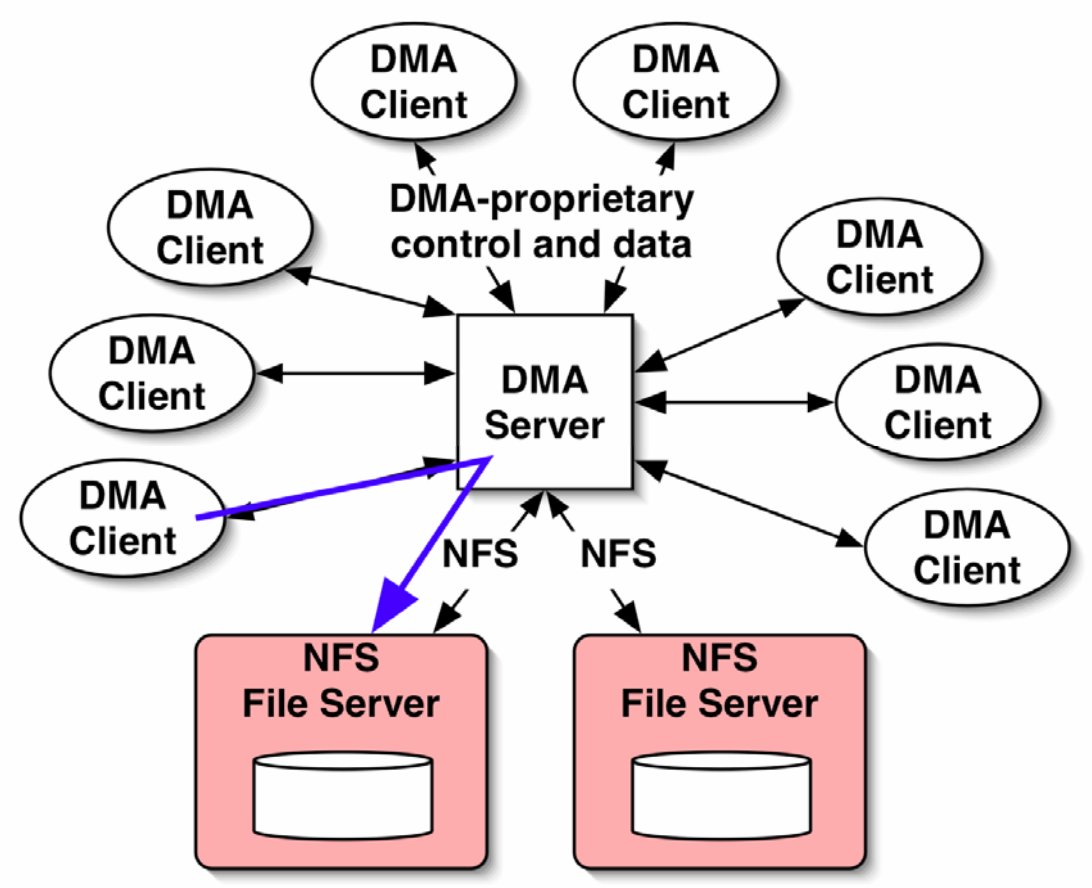
Outline

- File Service definition and use
- API overview
- **Advanced applications**
- Summary

Advanced (unconventional) applications

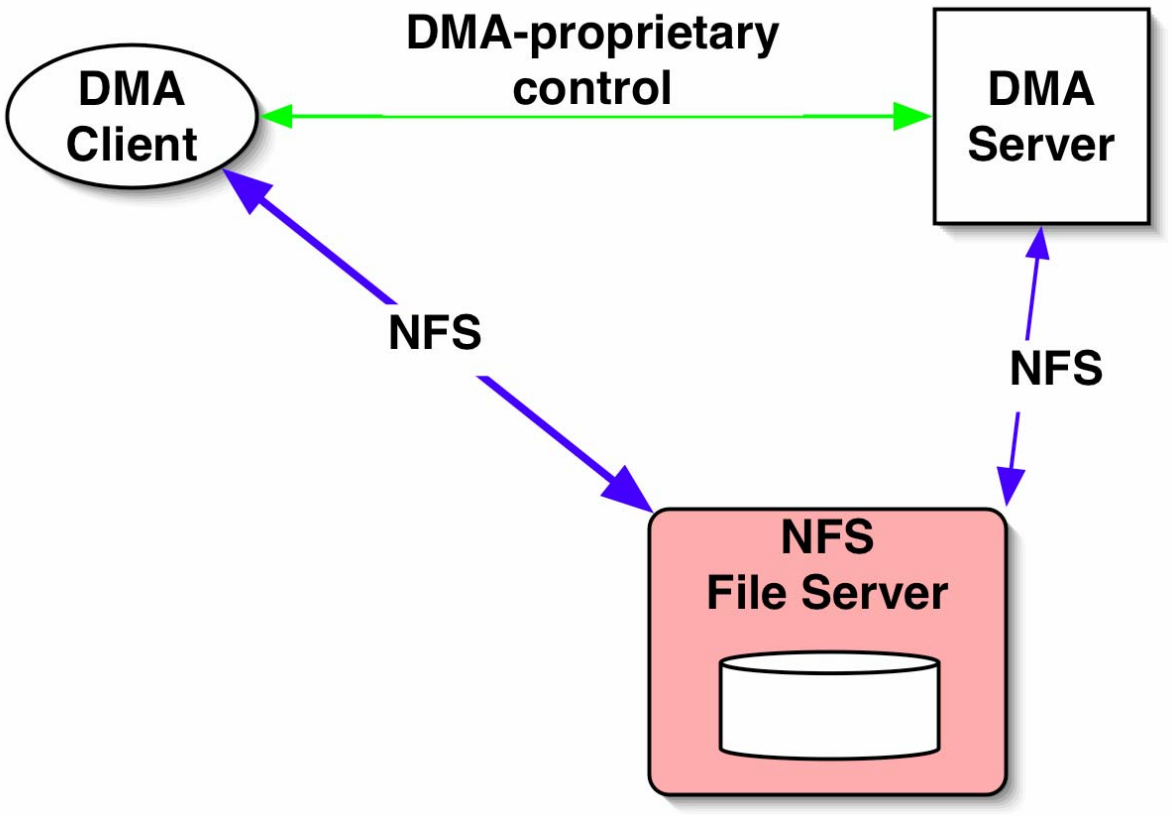
- Use NDMP File Service for DMA-proprietary data streams
- Direct transfers from non-NDMP DMA agents to NAS
- Serverless creation of synthetic fulls

Standard D2D backup architecture



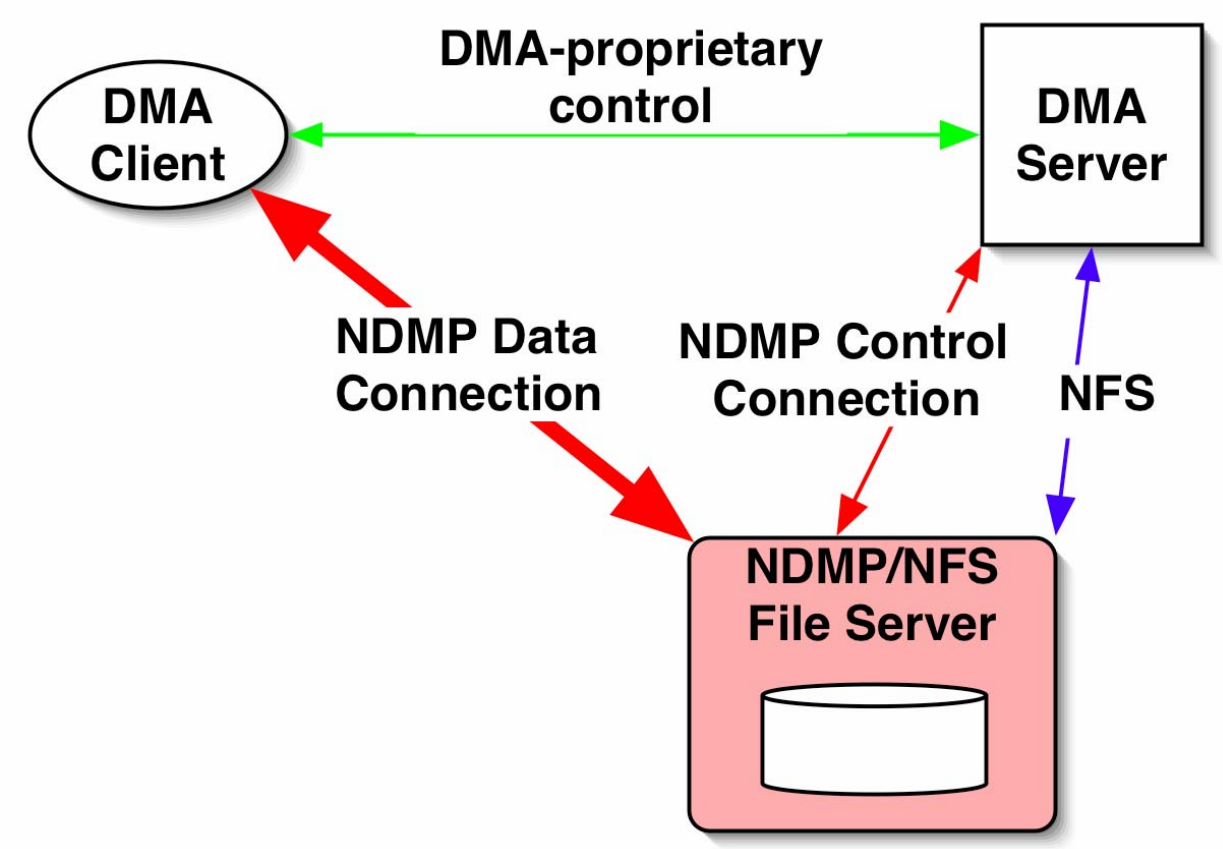
- Data goes from client through server to disk
- DMA backup server is the bottleneck

Better: Direct transfers via NFS



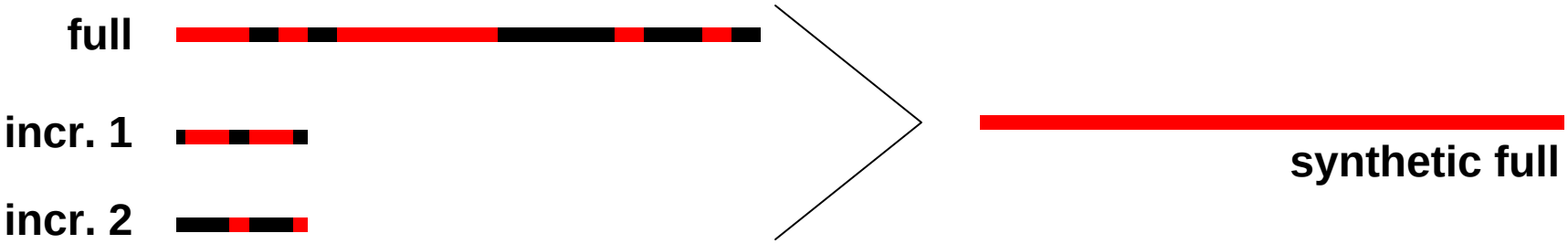
- Direct NFS transfer avoids DMA backup server bottleneck
- Trashes client file cache
- Creates security problems; backup data put at risk

Best: Direct transfers via NDMP



- Avoids DMA backup server bottleneck
- Leaves client cache intact
- Lets DMA control access to files to keep backups secure

Synthetic fulls



- **Merge several full and incremental backups into one new synthetic backup image**
- **Requires knowledge of backup image format**
- **Saves clients from doing a full**
- **Heavy load on backup servers to create the synthetic full**
  - Must read and write a full set of data ==> 2x data transfer

Serverless synthetic fulls

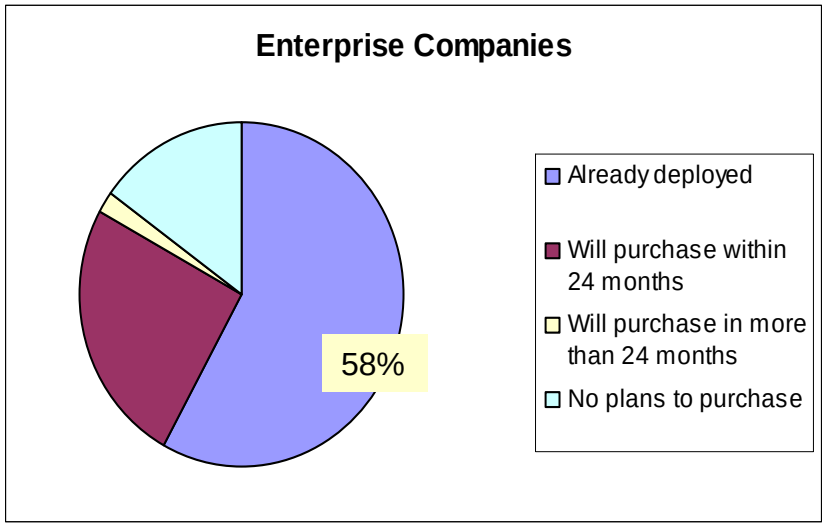
- **Save DMA-proprietary backups on networked storage**
  - Use standard NFS, direct transfer NFS, or direct transfer NDMP File Service to get backups to storage server
- **Use NDMP Advanced File Commands to create the synthetic full**
  - DMA gives NDMP File Service a recipe for creating the full
  - Commands similar to an IO vector
- **Data format remains opaque to the NDMP File Service**
- **Data format not opaque to the DMA since it created the backup images in the first place**
- **No user data goes through the backup server**
- **Some NDMP File Services can avoid any data movement**
- **Synthetic fulls become very easy to create**
  - Make one every day

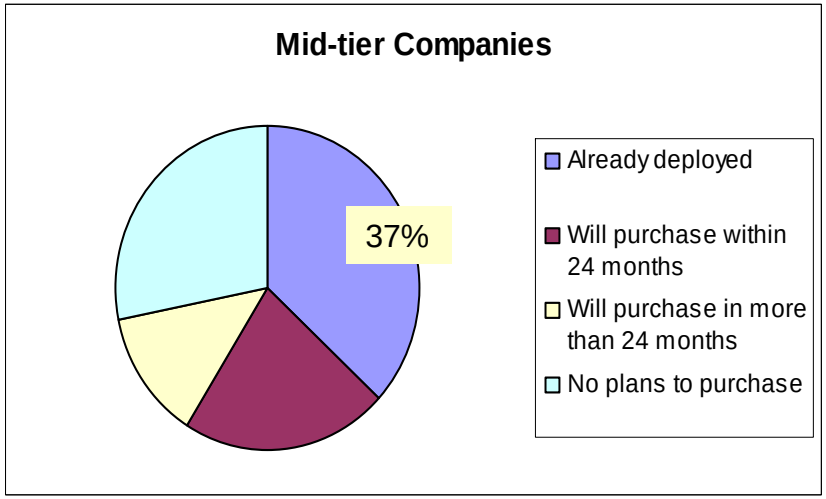
Summary

- **File Service adds support for D2D and D2D2T backup to NDMP**
- **Backward compatible with existing Data and Tape Service implementations**
- **Supports familiar Mover interface**
- **Advanced commands add flexibility and simplicity**
  - File to tape transfers with one NDMP Control Connection
  - No mover window or record size manipulations
- **File Service has big advantages for DMA data streams**
  - More secure, lower overhead than NFS for direct transfers from DMA clients to networked storage
  - DMA can give recipes for serverless synthetic full generation
    - Data format is opaque to File Service, but not DMA
- **Specification and reference implementation available:**

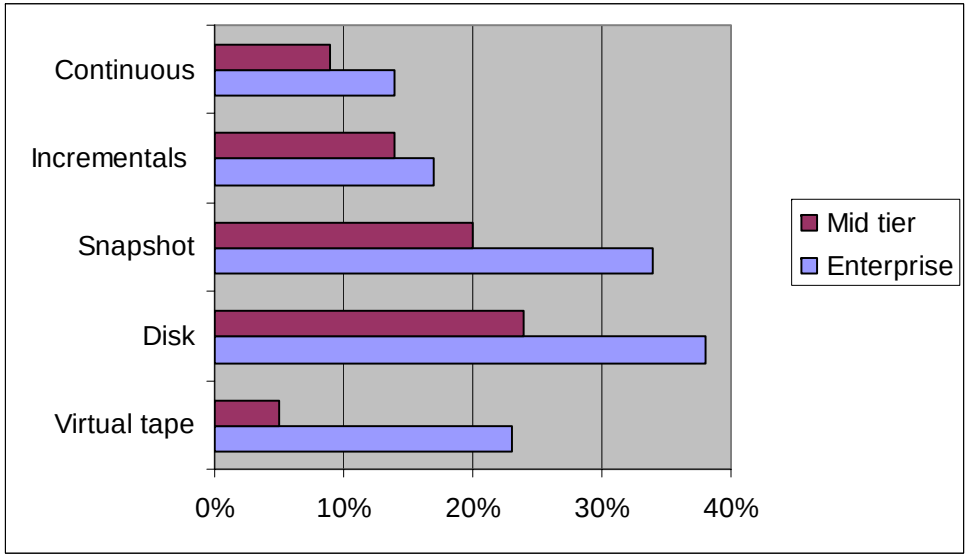
  http://sourceforge.net/projects/ndmfs/

Where is the market with disk backup?





Source: ESG Research 2004



(percentage of users who have already deployed some form of disk based data protection)