Disk-Based Backups: The Next Step in Storage Management Data Protection

Eric Carter

Product Manager
LEGATO Software – EMC Software Group

Agenda

- Why we're here
- Why Backup-to-disk
- Backup-to-disk considerations
- Justifying the Value
- List of resources

Why We're Here

What Backup-to-disk, but we're really talking about restore

When Now, new technology has changed the economics and the market acceptance is high

Why Tape is no longer meeting recovery needs

Where Not just the data center but all over

Who Those with the most to lose/gain

- CIO
- Data Center Managers
- Backup Admins

Recovery Factoids

- Reliance on tape alone for data recovery is no longer a best practice
- More than 80% of restore requests are made within 48 hours of data loss over 90% are made within 30 days
- 60-70% of storage management effort is devoted to B/R
- 15% of a storage administrator's time is spent on recovery operations
- Over 5% of B/R jobs fail nightly

—*META Group, April 1, 2004*

Traditional Backup and Recovery

Challenges:

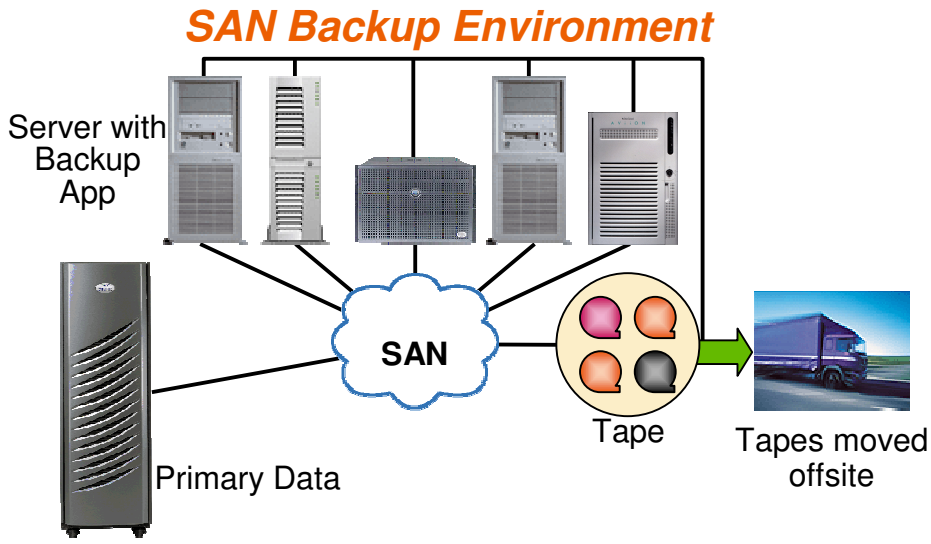
Pressure to meet the changing needs of the business

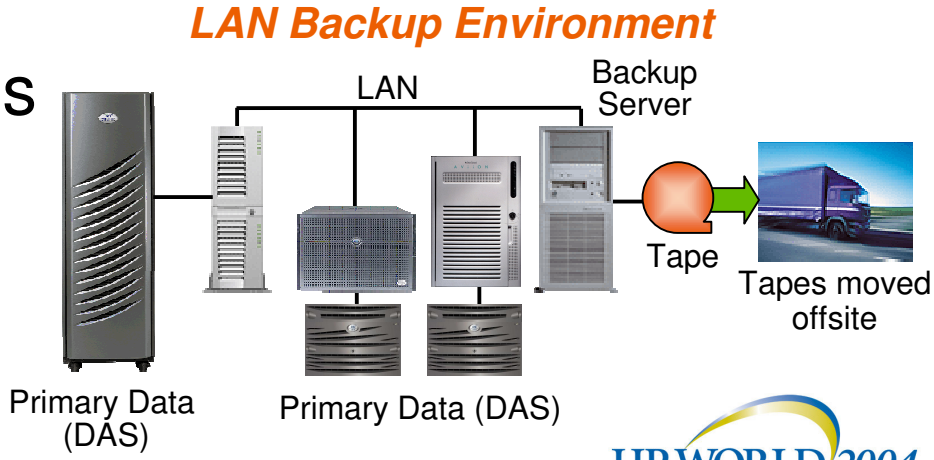
- Availability = Revenue
- Traditional methods are unreliable
- Operations are highly manual and inefficient
- Backup and restore environments continue to grow in complexity
- Costs of backup and restore continue to rise

HP WORLD 2004
Solutions and Technology Conference & Expo

Why Tape Doesn't Always Meet the Challenges

- **Performance** limitations
  - Limited throughput
  - No random access to data
- **Availability** limitations
  - Tape media does not support data protection, like RAID
  - Tape media is "expired" after certain number of uses
- **Management** complexities
  - Restoring from multiple incremental copies is painful
  - Media replacement costs
  - Vaulting costs
  - Labor costs





Today's Backup and Recovery

Solution:

Backup-to-disk: moves data to disk rather than tape to improve operational backup and restore

Why Disk Can Meet the Challenges

Faster, more predictable backup and recovery

Performance

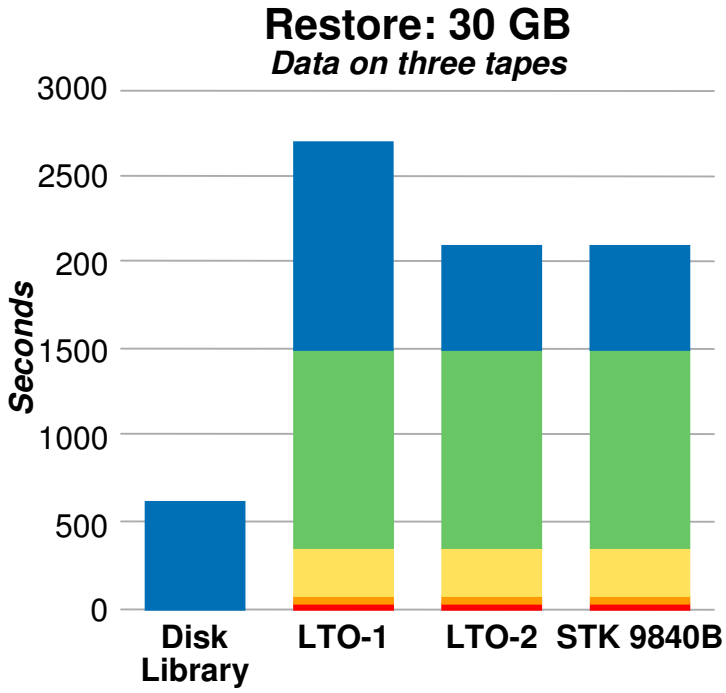
- 80MB/sec single stream throughput
- Immediate, random access to data
- Backup AND recover simultaneously on the same media

Availability

- High-availability built in
- Supports RAID
- Advanced data protection

Management

- Reduced number of drives and libraries
- More efficient use of staff
- Eliminate pain of incremental restores



70–77% faster restore!

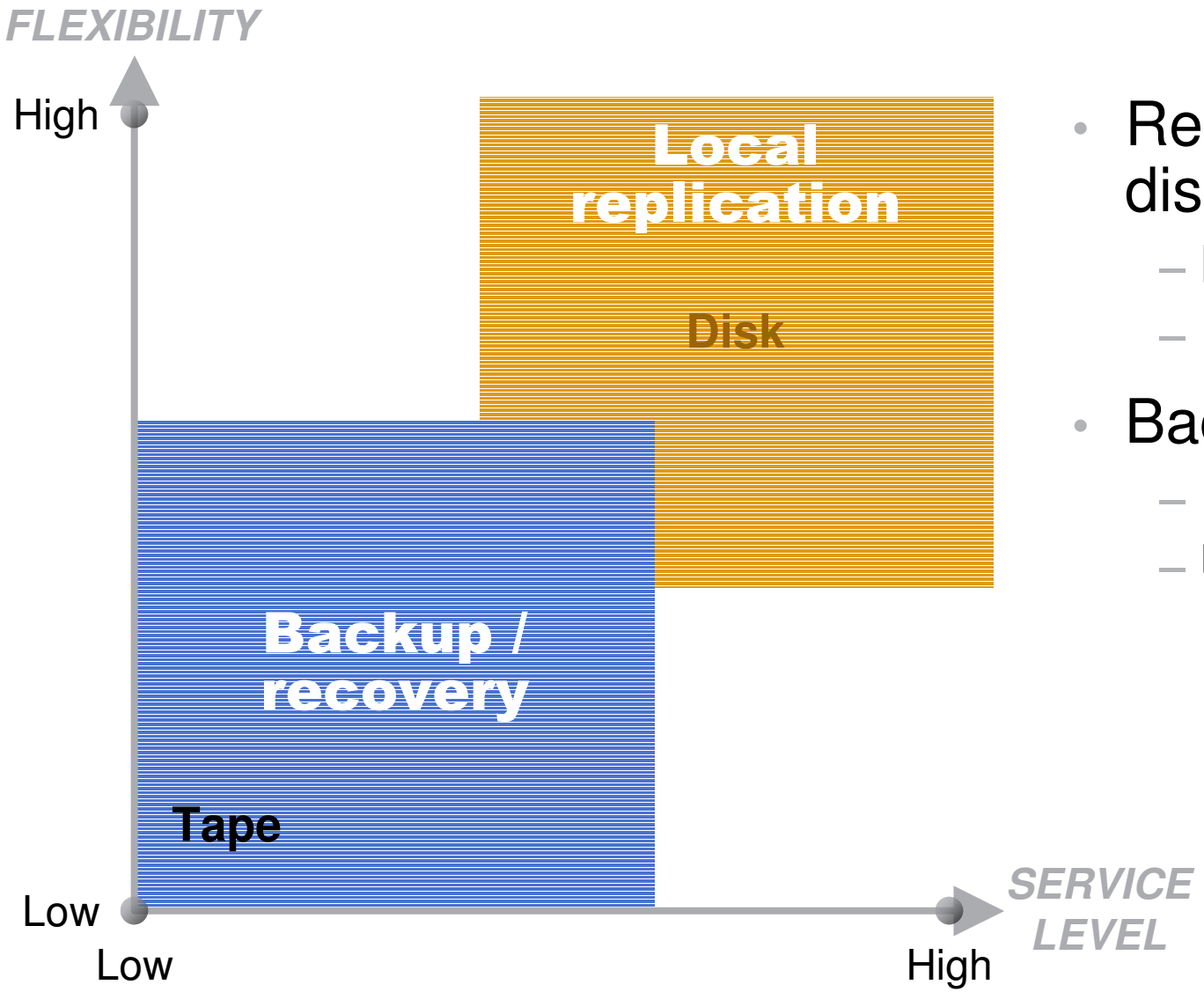
NOTE:
LTO1, LTO2 9840B assume 2:1 compression

HP WORLD 2004
Solutions and Technology Conference & Expo

Targets for Backup-to-disk

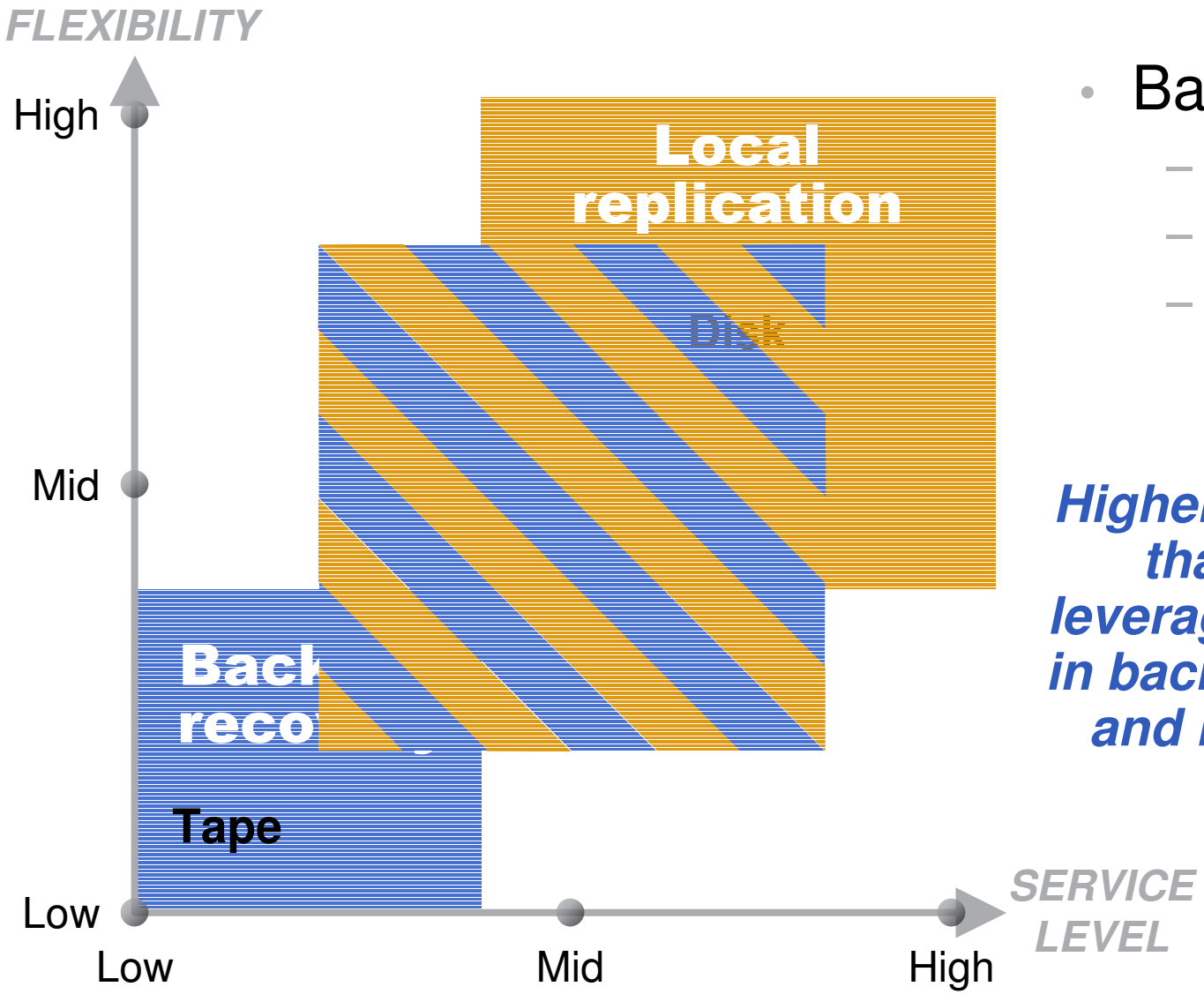
- **Disk Arrays**
  - Typically leveraging ATA drives for lower cost, higher capacity
  - Direct Attached, SAN Attached, NAS Attached
  - Good flexibility – support for point-in-time copies
  - Can support both transactional and backup data
- **Virtual Tape Libraries (VTLs)**
  - Disk arrays that appear as a tape library
    - Emulates a variety of libraries and drives
    - Typically leverages ATA drives
  - Disks are invisible to the backup application
    - It only sees a tape library
  - A dedicated appliance
  - Because it emulates tape – not all Backup-to-disk advantages apply

Typical Information Protection Choices



- Replication to disk
  - Fast
  - Reliable
- Backup to tape
  - Integrated
  - Ubiquitous

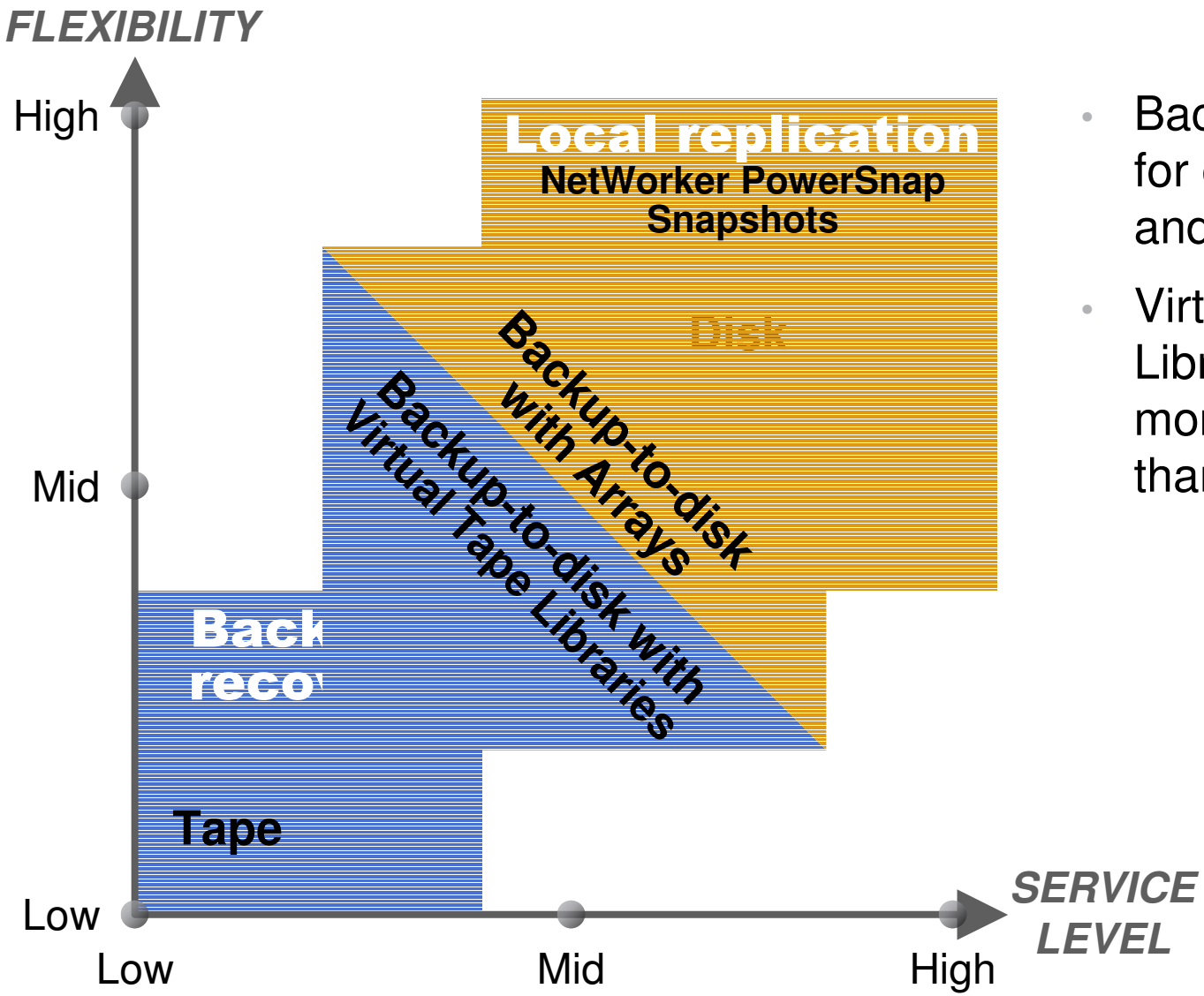
Today's Information Protection Options



- Backup to disk
  - Fast
  - Reliable
  - Integrated

Higher service levels than tape and leverages investment in backup application and infrastructure

Tiered Data Protection Offering



- Backup-to-disk—best for operational backup and restore
- Virtual Tape Libraries—faster and more reliable backup than tape

Information Protection Requirements

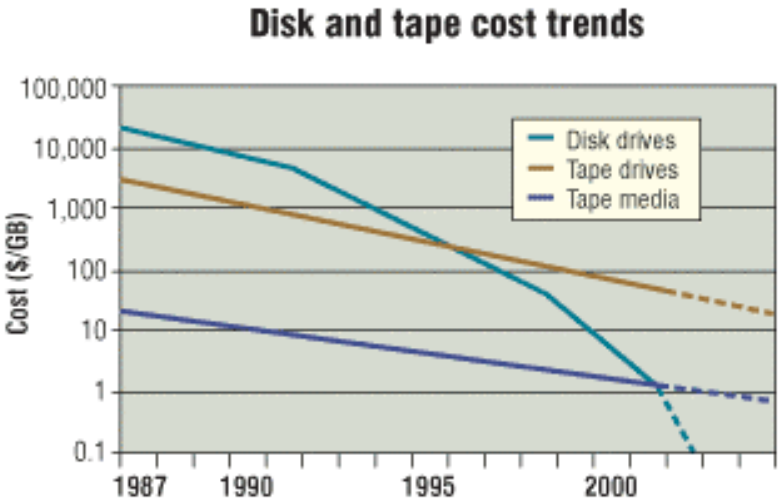
Transactional ⟷ **Referential**

Business Need	Continuous Availability	Data & Operational Recovery	Disaster & Long-term Protection	Compliance
Backup Window	*Instant/Seconds*	*Seconds/Minutes*	*Minutes/ Hours*	*Hours*
Recovery **Time Objective**	*Instant*	*Seconds/Minutes*	*Hours*	*Hours/Days*
Retrieval **Time Objective**	*Instant*	*Seconds*	*Seconds/Minutes*	*Minutes*
Retention **Requirement**	*Hours/Days*	*Days/Weeks*	*Months*	*Years*
Solution Options	Snapshot, Mirroring	**Backup-to-disk**, Replication	Tape Backup, HSM	HSM, Archive, Vaulting

SERVICE-LEVEL REQUIREMENTS

Disk-based Solutions – Why now?

- It's Affordable
  - ATA drives have lowered the cost per megabyte more than any other technology
- It Offers Reliability
  - Greater confidence due to redundancy, RAID, and other high-availability features
- Technology Supports It
  - Disk-based backup can be architected with most backup/recovery solutions
- Customer Environments can Support It
  - Storage networks: SAN and LAN
  - More servers can take advantage of the infrastructure
- Business Requirements Demand a "Better Way"
  - Better, Faster, Cheaper…



"75% of storage management is backup and recovery yet 30% of all data recovery instances fail due to botched backups."

— Anders Lofgren, Forrester Research, 2003

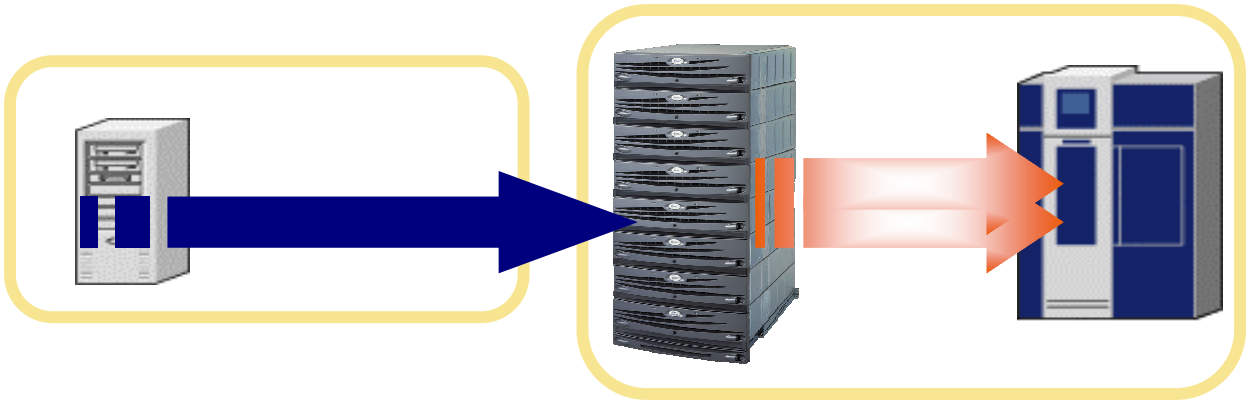
Backup-to-Disk: What It Isn't

- Not an array-based replica
  - BCV, clone, snapshot, etc.
- Not an archiving solution
  - B2D is a second copy—"operational restore"
- Not a replacement for all tape
  - Offsite requirements
  - Transportability (intra / inter-platform)
  - Legal and regulatory issues
  - Heritage

What about Physical Tape?

Including Tape in the Backup-to-disk Paradigm

- Most Companies Still Require Tape!
- Tape remains a best practice and the long time archival vehicle of choice (mobility)
- Older backups need to be available – don't expect to replace tape completely overnight!
- Tape still much cheaper, allowing for multiple, stable copies to be offline or offsite

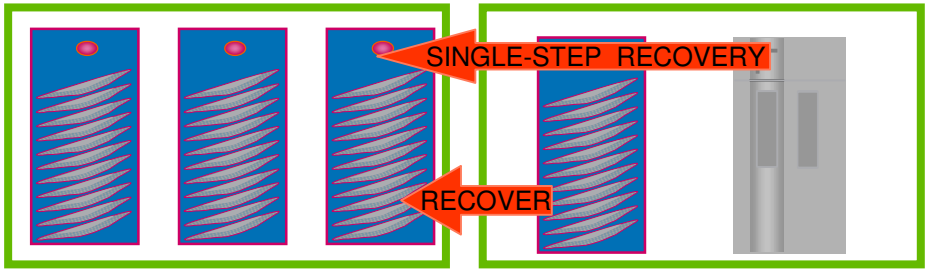


Recover Data Faster and More Efficiently

Restore-from-Disk Characteristics

- No tape headers, file marks, catalog info – faster, less capacity consumed than tape
- Leverage characteristics of disk – high-speed random access read
- Simultaneous parallel recovery of multiple data sets from disk
- Eliminates time consumed by overhead of mounting/un-mounting/positioning tapes

Recover from Disk, Single-Stage Recovery from Tape



Production Storage

Backup/Recovery Storage

Single-Step Recovery from Tape

- Recoveries from tape are returned directly to clients in an optimized sequential manner
- Fewer clients wait for recoveries from tape to complete, minimizing downtime and costs – optimized for disaster recovery
- Backup software database maintains a record of where data resides, by client

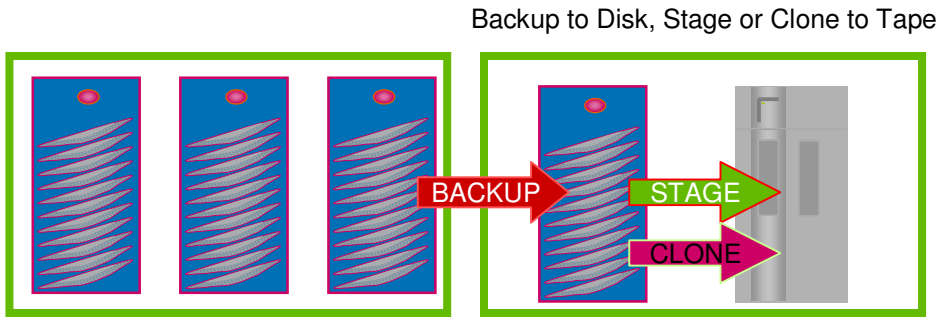
Disk Device Managed Automatically

- Automatically sense and remove incomplete backups from disk; purge expired data
- Reduce admin overhead and fully leverage disk resources

Protect Data Faster and More Efficiently

Backup-to-Disk Characteristics

- No tape headers, file marks, catalog info – faster, less capacity consumed than tape
- Leverage the high-speed random access write characteristics of disk – faster
- Simultaneous parallel multi-client backups are multiplexed (interleaved) to file type devices on demand – faster
- Eliminates the time consumed by overhead of mounting/un-mounting/positioning tapes



Read from and write to disk simultaneously

- Perform staging/cloning and backups *concurrently*
- Recover data contiguously from devices engaged in backups – faster

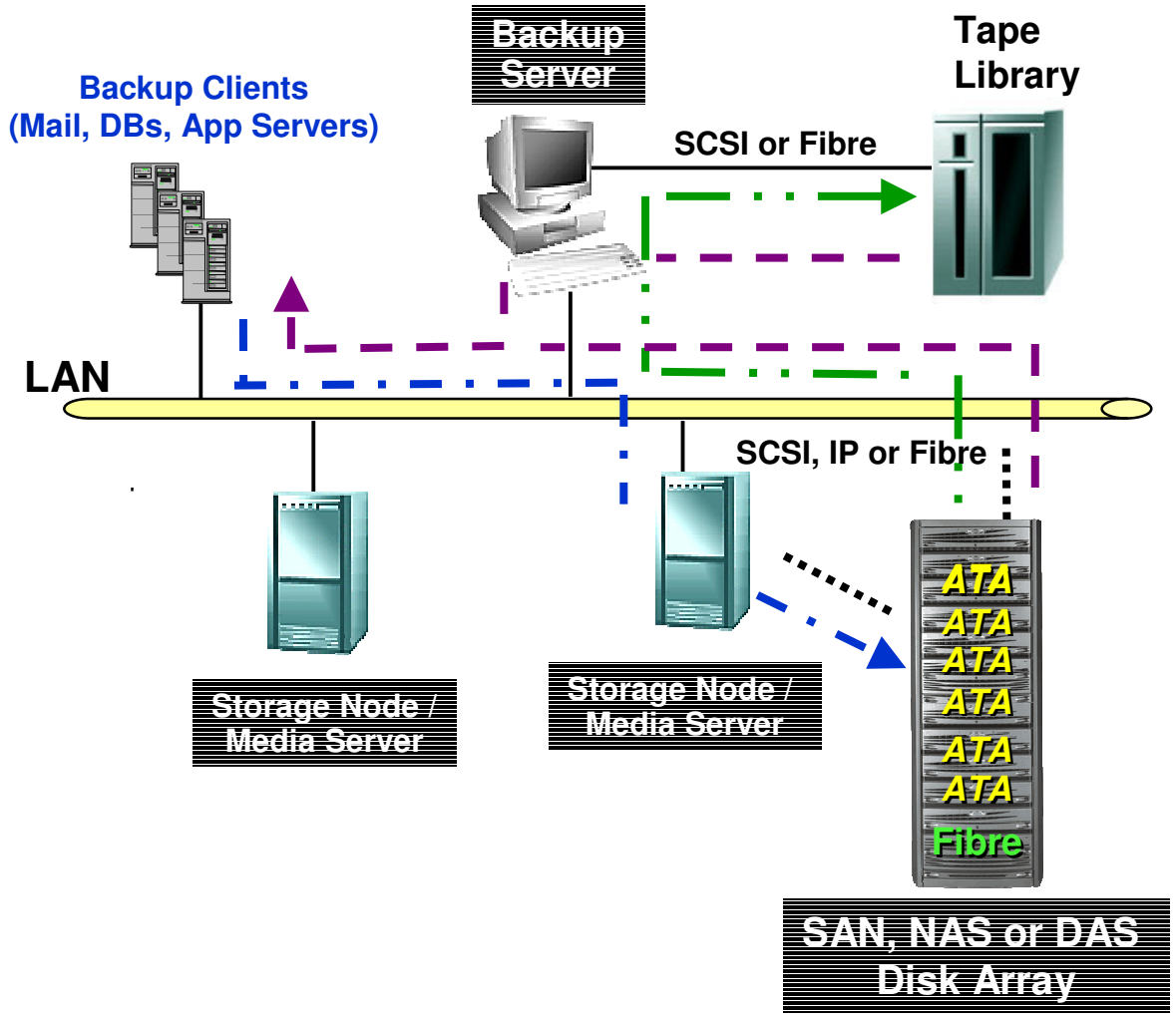
Fast, Efficient Staging to Tape

- High-speed serial streaming to tape
- Policy-based staging automatically frees space on disk, optimizing full backups

Faster Tape Cloning

- Completion lags only behind the time of the last *client* backed up (not the entire completed backup) – faster, efficient

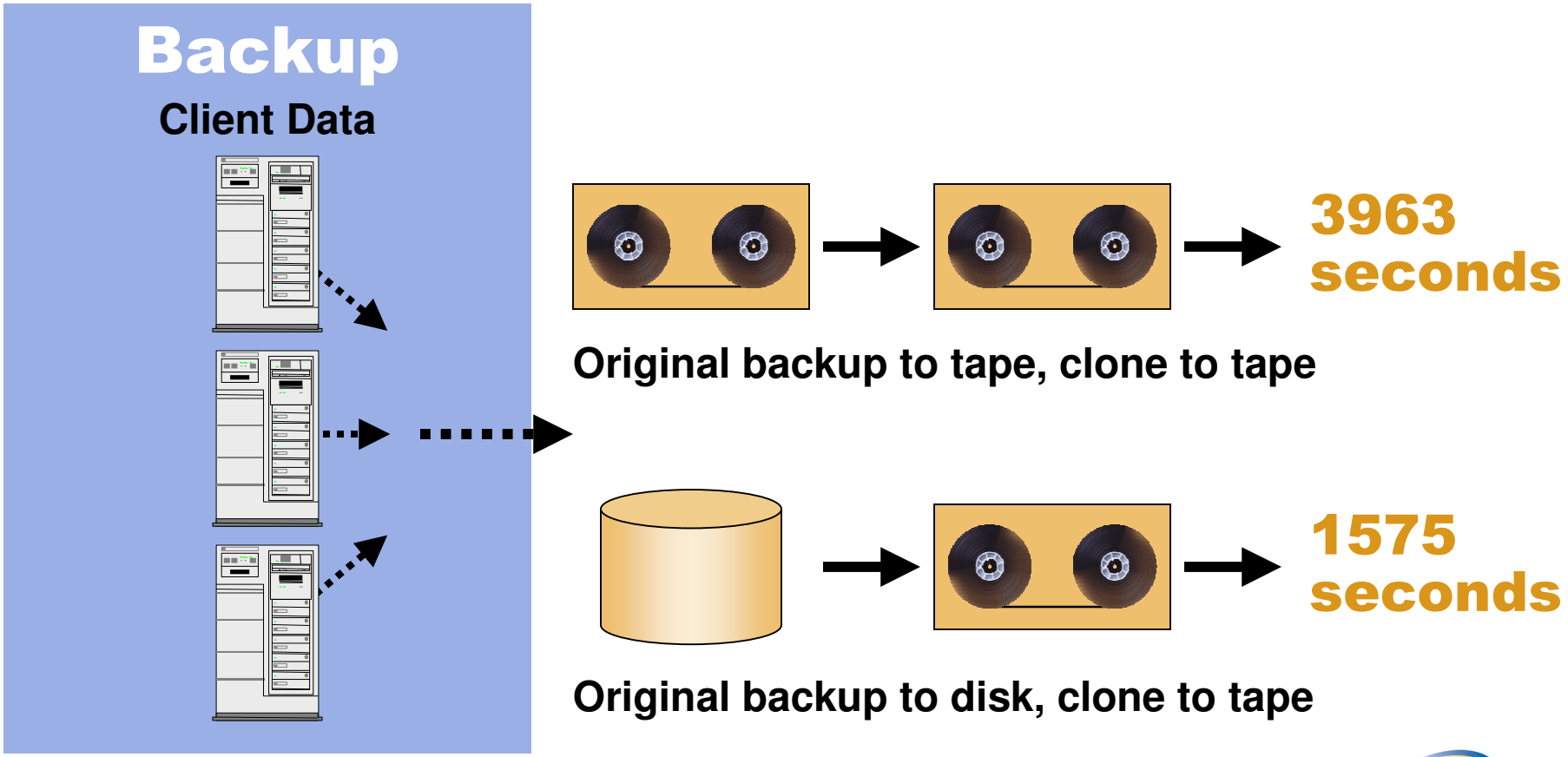
Backup-to-disk Implementation



- **Primary Protection**
  - **Backup** unlimited clients directly to disk
- **Secondary Protection**
  - Automatically **move** data to tape via Staging, or **copy** to tape with Cloning
- **Single-step Recovery**
  - **Recover** files directly from tape (or disk) to client

Performance Advantages

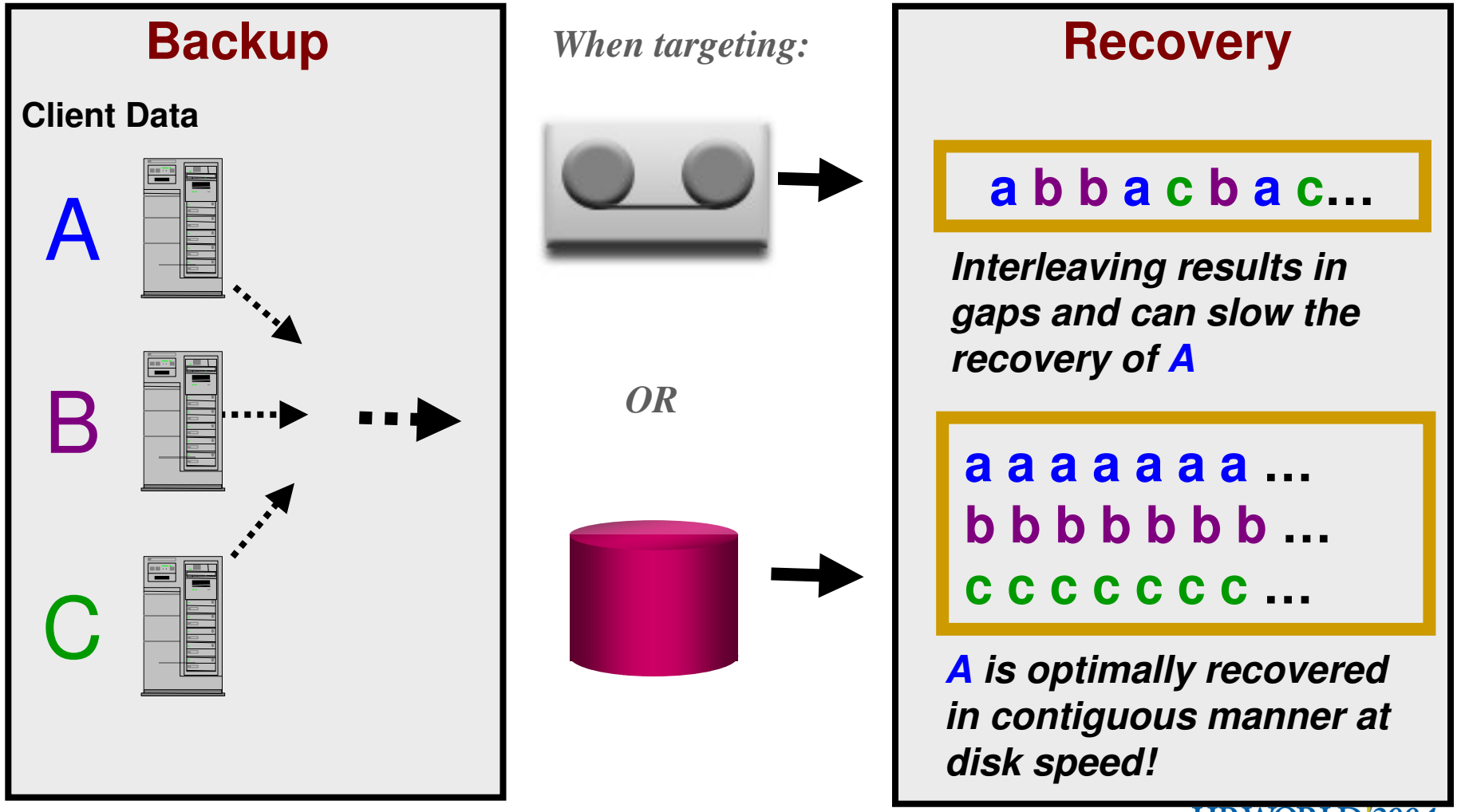
Comparing Time to Create an Offsite Copy—Traditional vs. DiskBackup



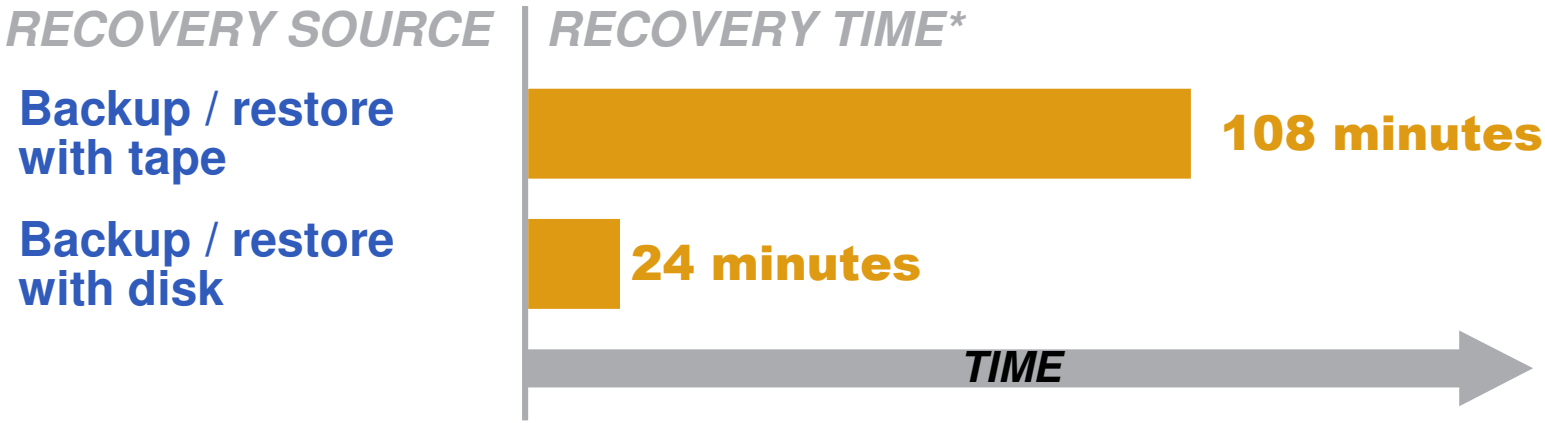
More than two-times faster!

HP WORLD 2004
Solutions and Technology Conference & Expo

Backup-to-disk Contiguous Recovery Advantage



Restore Comparison



**Total time from point of failure to return of service to e-mail users*

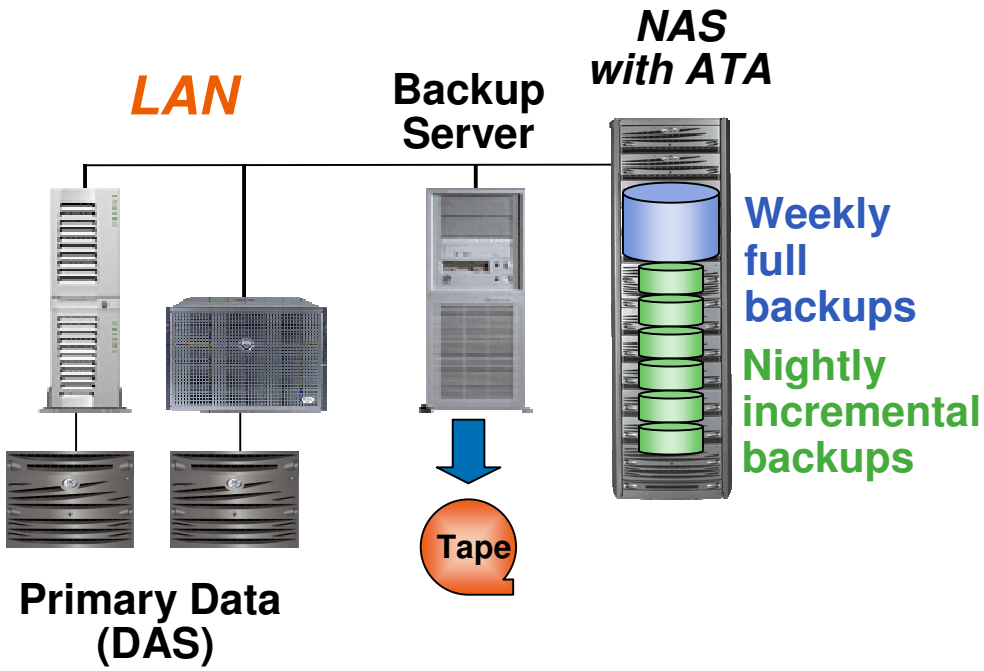
Typical scenario:

- 800 users, 75 MB mailbox
- 60 GB database

Source: EMC Engineering and EMC IT

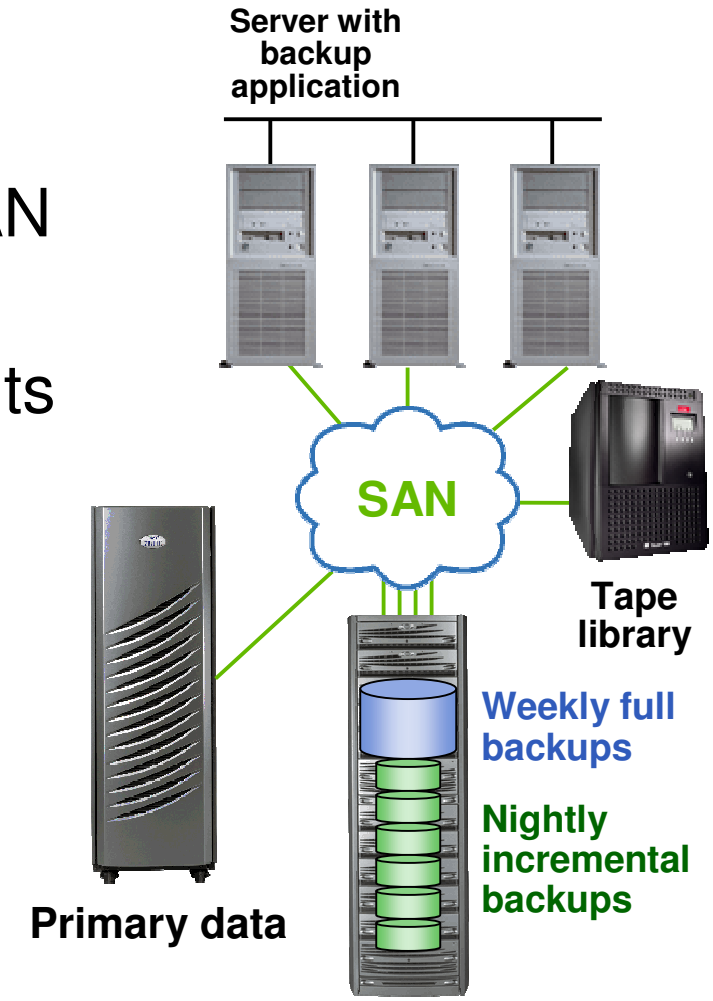
LAN Environment

- Uses existing LAN infrastructure
- Ideal for distributed deployments
- Attractive option where no SAN infrastructure exists
- Offloads storage processing from backup server



SAN Environment

- Leverage high-performance SAN infrastructure
- Ideal for centralized deployments



SAN Storage with ATA

Operational Benefits of Disk in this Environment

Cost Savings...

- Tape libraries and drives
- Physical media
- Staffing and floor space

Performance...

- Up to 5x performance increase on restore
- Up to 3x performance increase on backup
- Single-stream performance up to 80 MB/s

Data Access...

- Read only what you need to
- No waiting for a cartridge to load
- No waiting as the tape fast forwards

Disk Backup Case Study – TV2 Norway

- The largest commercial TV station in Norway
  - In operation 22 hours a day, 365 days a year
  - Significant investment in tape
  - Need for improved service-levels
- The Challenges
  - An increasing amount of data - 5TB backup data
  - Complex environment
    - Solaris, SGI IRIX, Linux and Windows
    - 3 offices across Norway
  - A very small backup window
  - STK library with DLT drives
    - Could not deliver backup within small window
    - Could not recover quickly
    - Did not keep up with data growth

Disk Backup Case Study – TV2 Norway

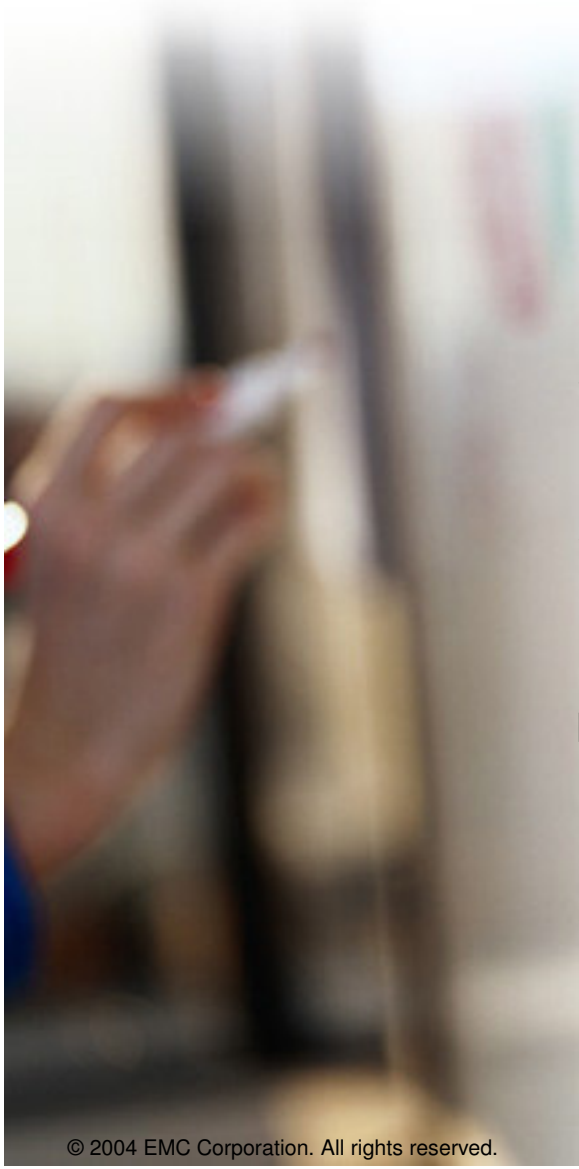
Disk Backup Helps Meet Service Levels

- LEGATO NetWorker DiskBackup Option with an ATA-based disk array
  - Fast, reliable backups fit within 2 hour window
  - Improved restore performance by 10x – responsive, “instant” file-level restores in seconds rather than minutes!
  - Tapes optimized for recovery
  - Less wear and tear on tape library – extending asset utilization = continued ROI
  - Lower administrative overhead
  - User down-time reduced
  - Future scalability

“If a user calls because he has lost and important file, we can pretty much recovery before the phone conversation is over!”

Kare Teigland
System Engineer, TV2 Norway

Getting Started



1. Defining the success criteria for backup to disk
2. Discovering the content
3. Sizing capacity and performance – leveraging RPO and RTO
4. Developing the implementation plan
5. Justification of the solution

HP WORLD 2004
Solutions and Technology Conference & Expo

Implementing Backup to Disk

1. **Defining the success criteria for backup to disk**
2. Discovering the content
3. Sizing capacity and performance – leveraging RPO and RTO
4. Developing the implementation plan
5. Justification of the solution

HP WORLD 2004
Solutions and Technology Conference & Expo

Step 1 – Defining the Success Criteria

Basic operational concerns are addressed when implementing backup to disk:

- Restores must be completed in a certain amount of time
- Backups have to fit within a given backup window
- Reliability of recovery has to be improved
- Costs have to be efficiently managed
- Management of the solution has to be simplified

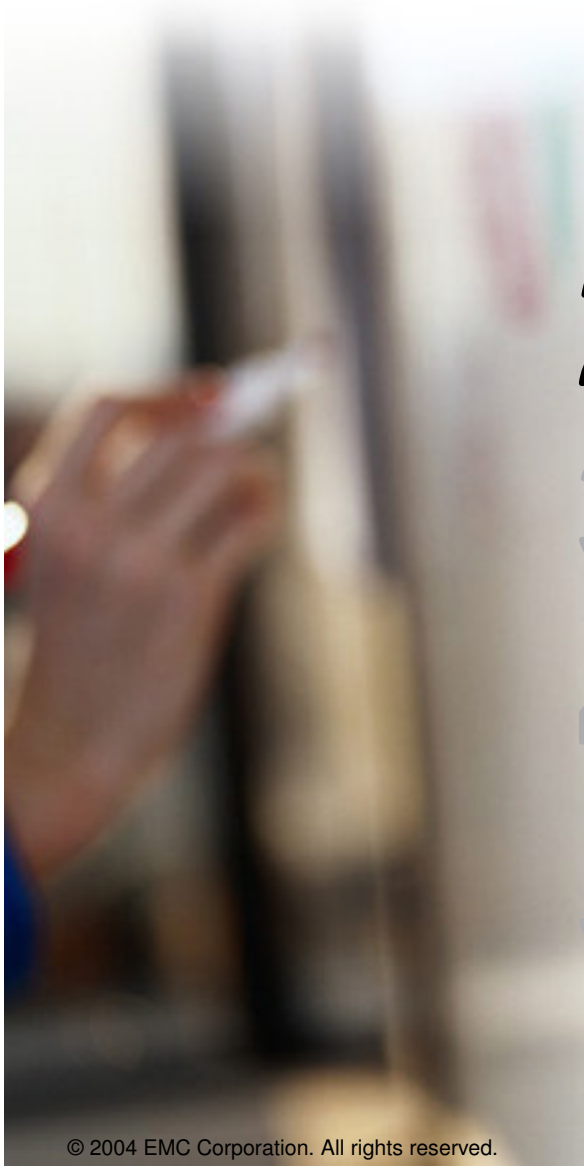
Developing the Success Criteria…

- Requires understanding of:
  - Each application's criticality to the business
  - Application capacity to address
  - Recovery point objectives
    - Ties to backup frequency and retention timelines
  - Recovery time objectives
    - Ties to service level requirements
- Choice of connectivity
  - SAN, LAN, or combination

Application Mix Example

	Uptime	*RTO*	*RPO*	*Backup Window*
Tier 1 applications	24x7x365	Seconds	Last transaction	None
E-mail	24x7x365	Minutes	Full restore	Minutes
Tier 2 applications	Business hours	Minutes to hours	Minimal loss	Minutes to hours
File servers	Business hours	Minutes to hours	Minimal loss	Hours
Business records and archived data	Business hours	Hours to days	Best effort (unless regulated)	Days

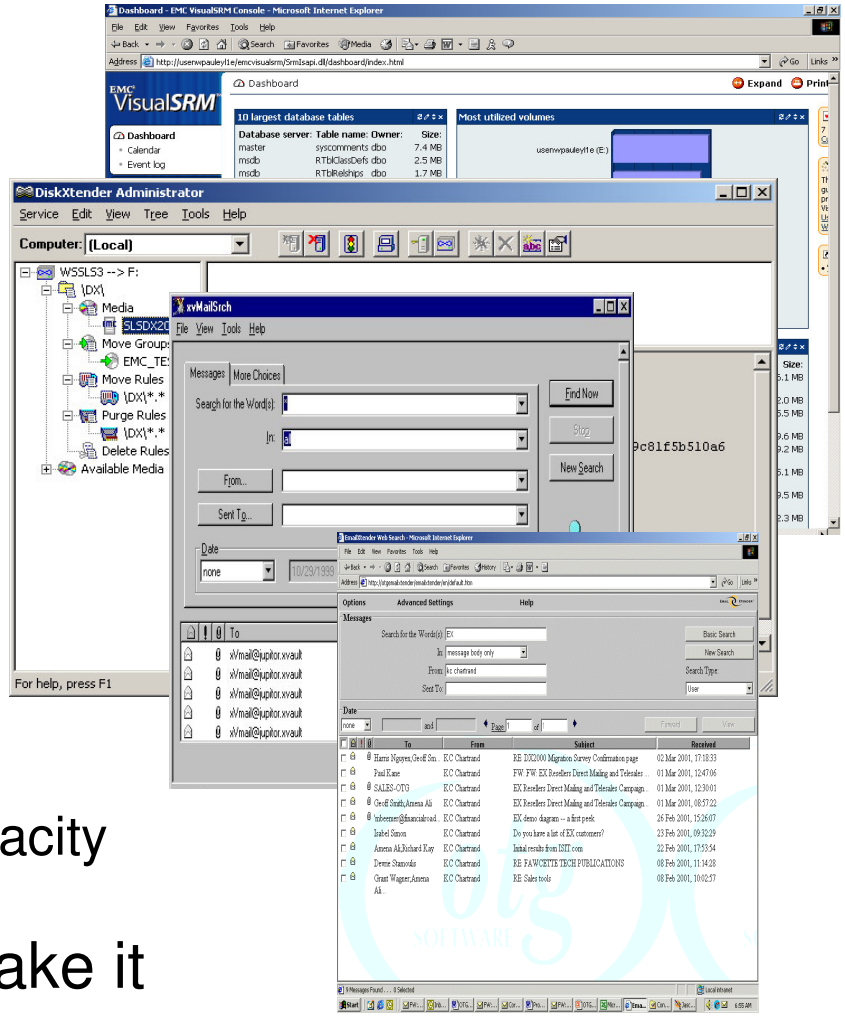
Implementing Backup to Disk



1. Defining the success criteria for backup to disk
2. **Discovering the content**
3. Sizing capacity and performance – leveraging RPO and RTO
4. Developing the implementation plan
5. Justification of the solution

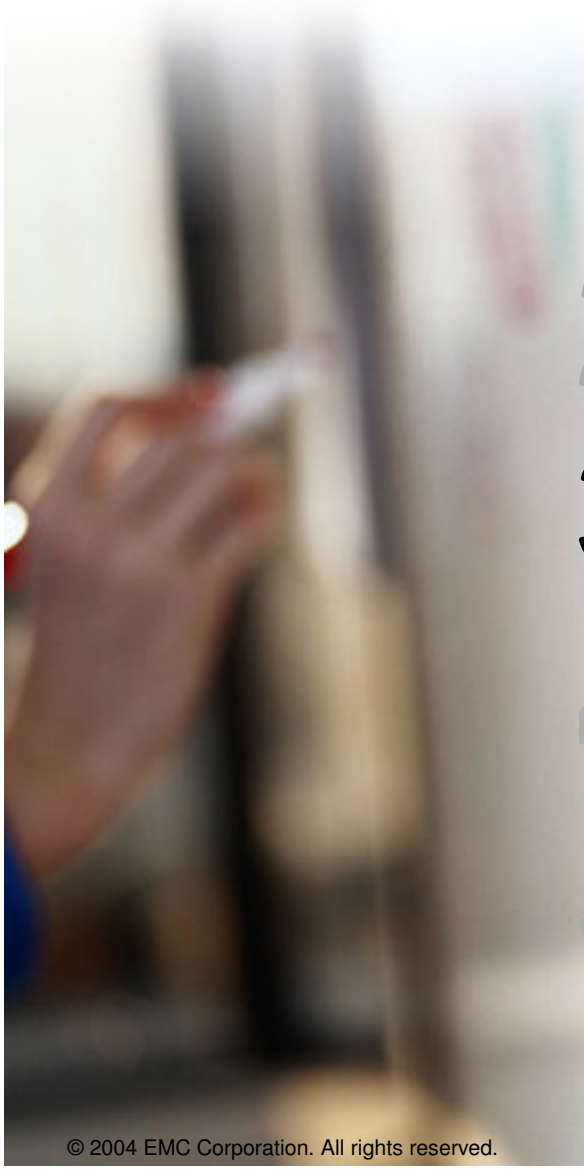
Step 2 – Inventory and Gather Data

- Backup content
  - How much is backed up?
  - How often is it backed up?
  - How long is it retained for?
- Clean house!
  - Stale data, duplicate data
  - Non-corporate data like MP3s
  - Extinct user data
- Removing the inactive data…
  - Accelerates backups
  - Accelerates restores
  - Relieves pressure on Tier 1 capacity growth
- Formalize the process, and make it repeatable
  - Tier – archive – delete



HP WORLD 2004
Solutions and Technology Conference & Expo

Implementing Backup to Disk



1. Defining the success criteria for backup to disk
2. Discovering the content
3. **Sizing capacity and performance – leveraging RPO and RTO**
4. Developing the implementation plan
5. Justification of the solution

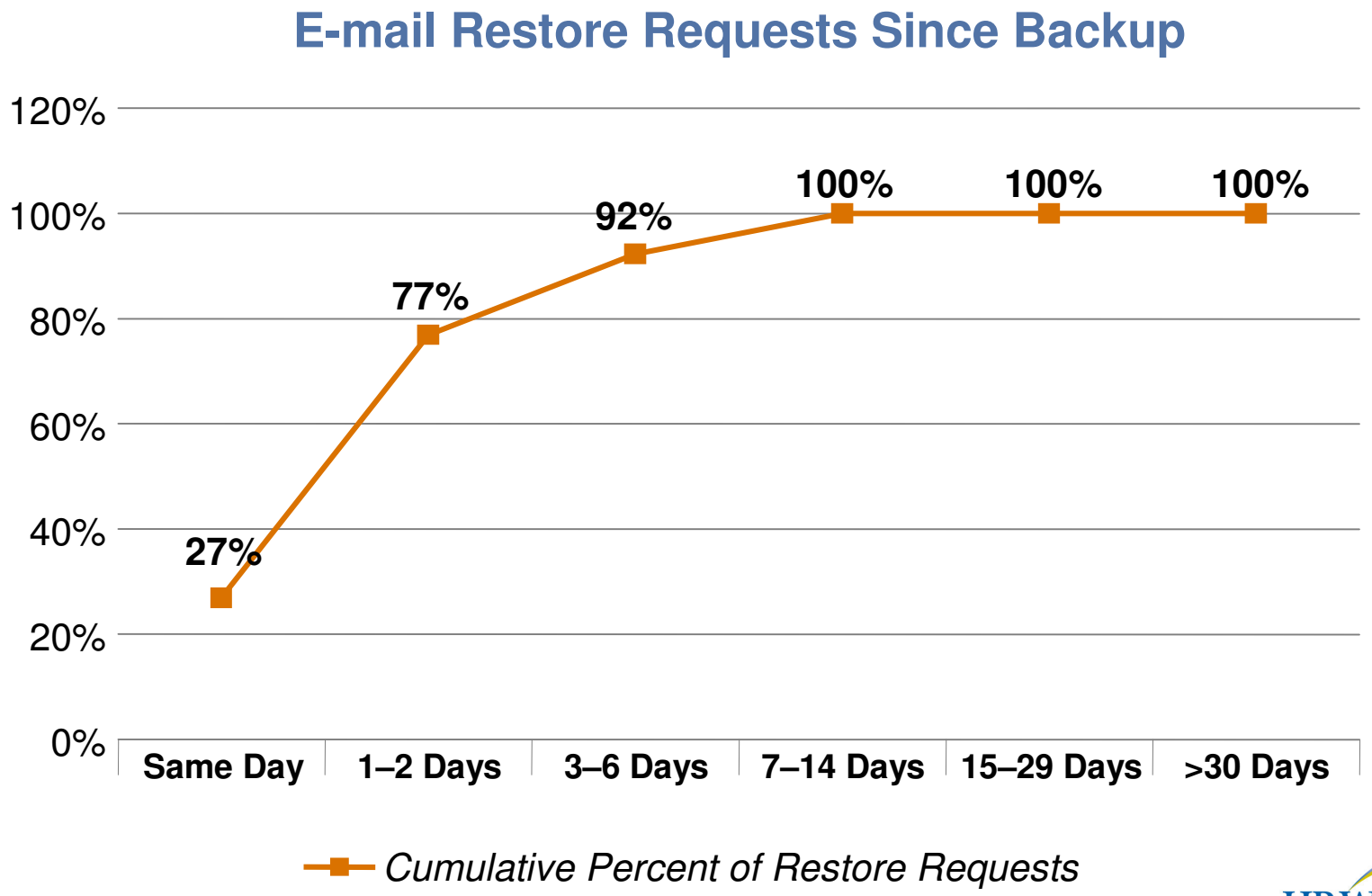
HP WORLD 2004
Solutions and Technology Conference & Expo

Step 3 – Sizing

- Capture recovery / retention data for each of your major applications
  - Understanding restore frequency relative to age of data can make the picture clearer
- Make informed choices in terms of your RPO service levels (information required to be retained on disk for a given period)
  - Number of days
  - Frequency of backup

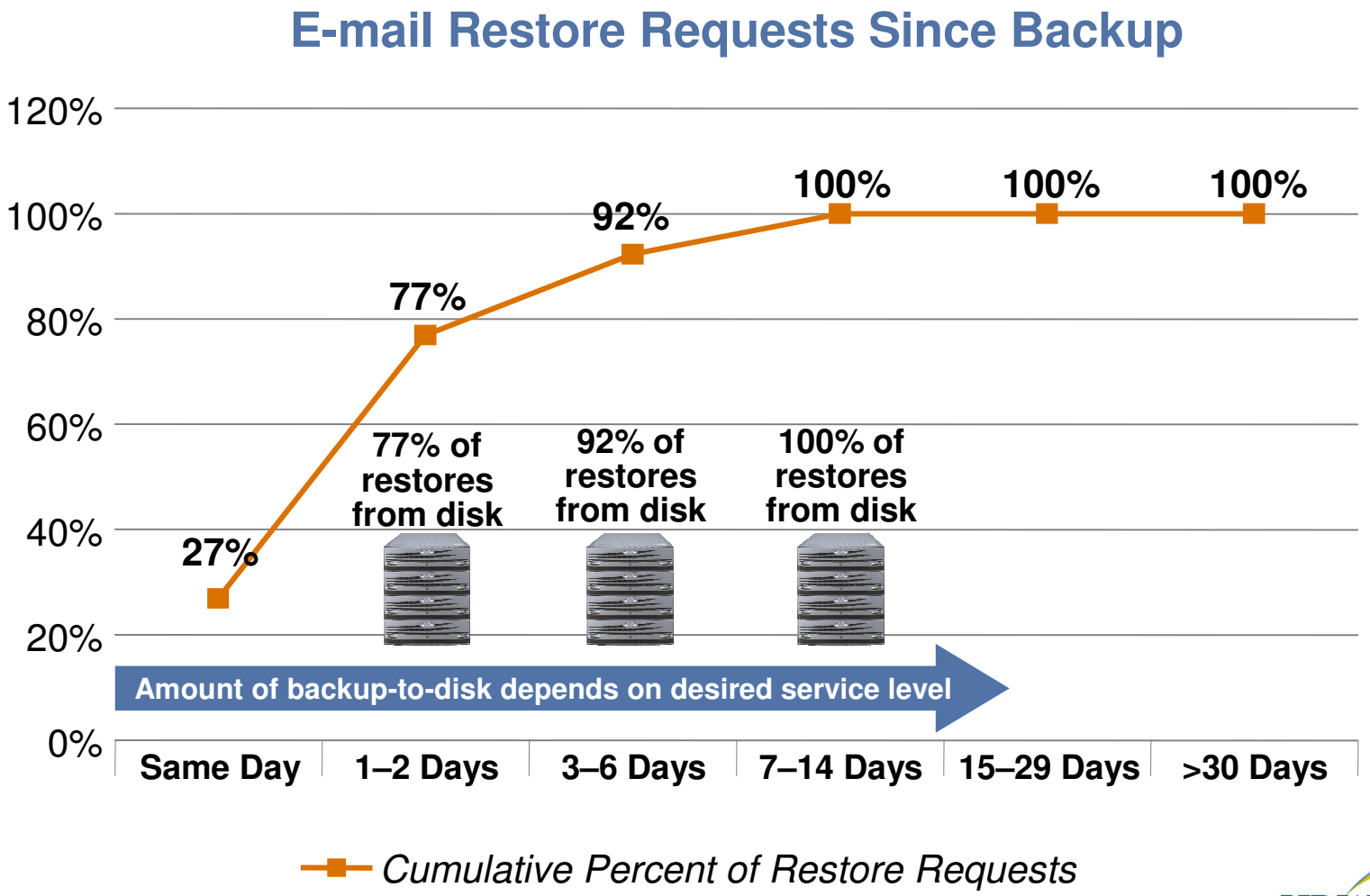
Sizing the Requirement—Backup Capacity

Internal Case Study



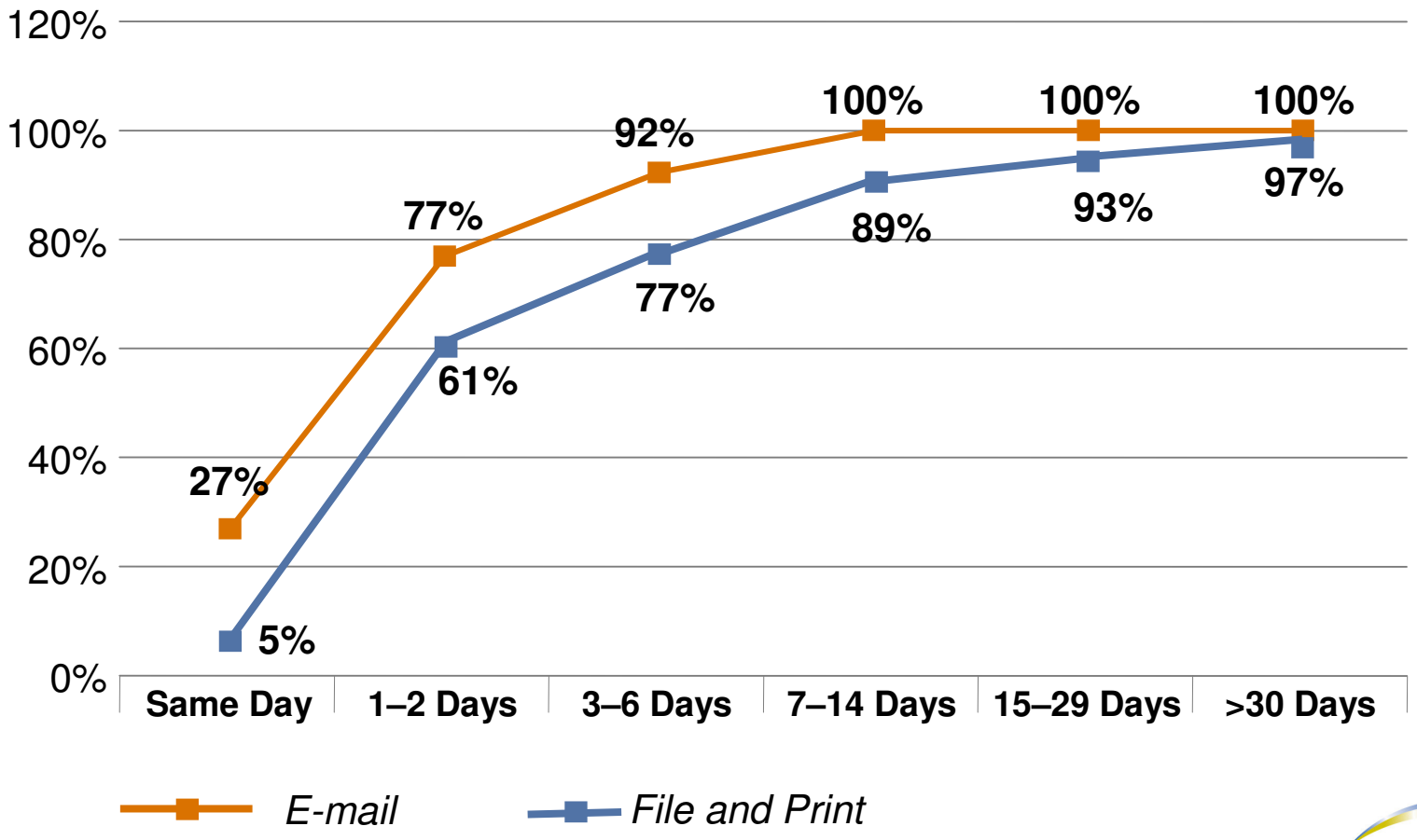
Sizing the Requirement—Backup Capacity

Internal Case Study



Sizing the Requirement—Backup Capacity

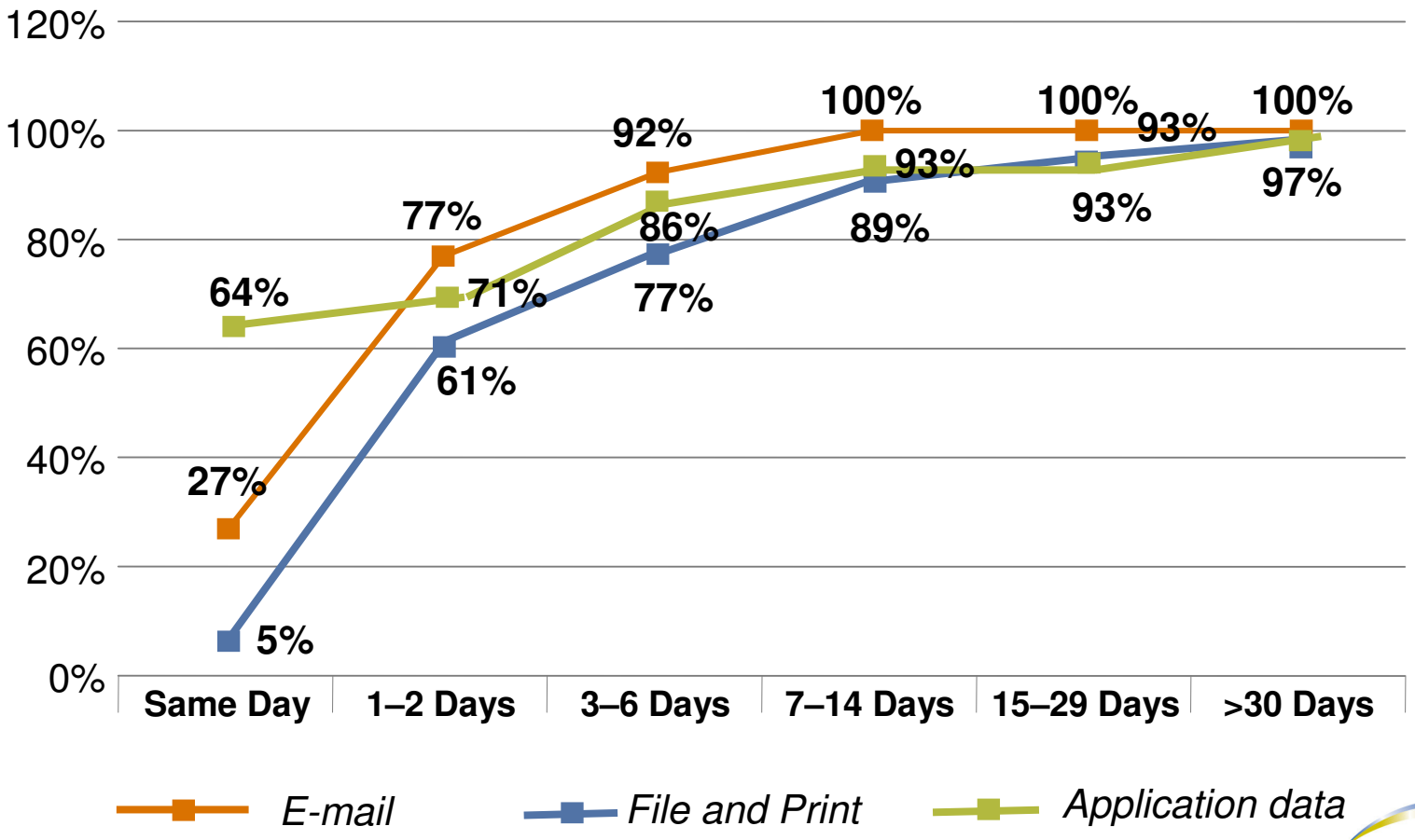
Enterprise-wide Application Recovery Request Graphs



HP WORLD 2004
Solutions and Technology Conference & Expo

Sizing the Requirement—Backup Capacity

Enterprise-wide Application Recovery Request Graphs



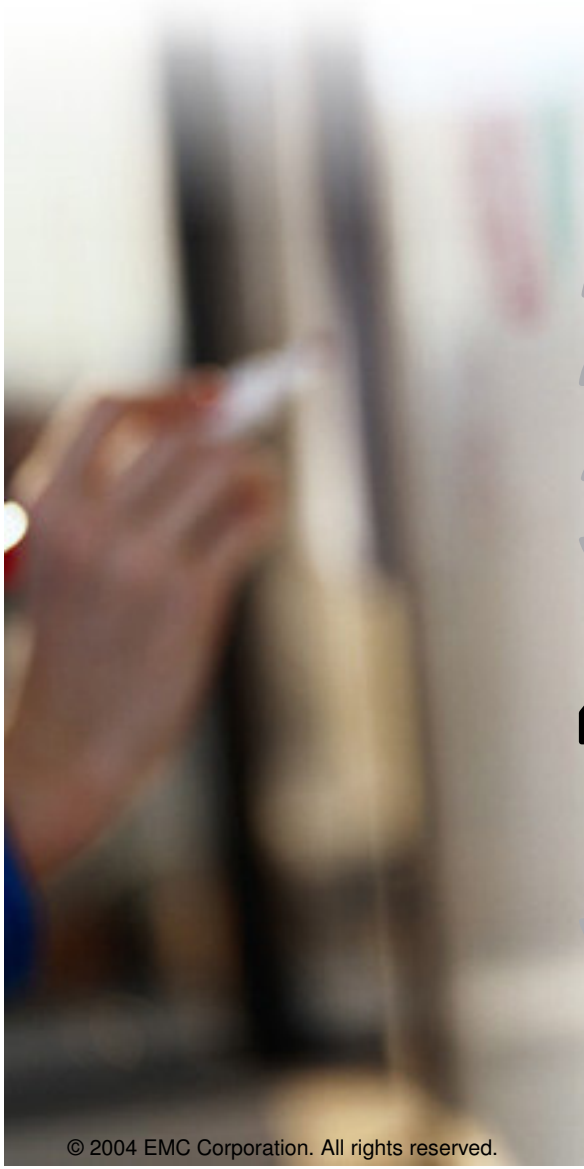
Sizing the Requirement—Performance

- Define the backup and recovery time objectives for each of the major applications
  - Applications need to be tiered to reflect mission criticality
- Size the solution to match the recovery goals
  - Number of LUNs
  - RAID 3
  - Number of RAID groups
  - Model and Number of arrays
- Need to understand external factors also
  - Server, network, etc.

Performance Characteristics to Think About

- Your mileage may vary…
  - 5X performance enhancement on restore
  - 3X performance enhancement on backup
  - Single stream performance up to 80 MB/s
- De-multiplexing removed from the restore process
  - Read only what you need to
- Immediate, random access to ALL data
  - No waiting for a cartridge to load
  - No waiting for tape positioning
- Disk-based reclamations, saveset consolidation/ synthetic full backups
- Backup AND recover simultaneously to same media
- Add a few ATA drives = adding a tape drive

Implementing Backup to Disk



1. Defining the success criteria for backup to disk
2. Discovering the content
3. Sizing capacity and performance – leveraging RPO and RTO
4. **Developing the implementation plan**
5. Justification of the solution

HP WORLD 2004
Solutions and Technology Conference & Expo

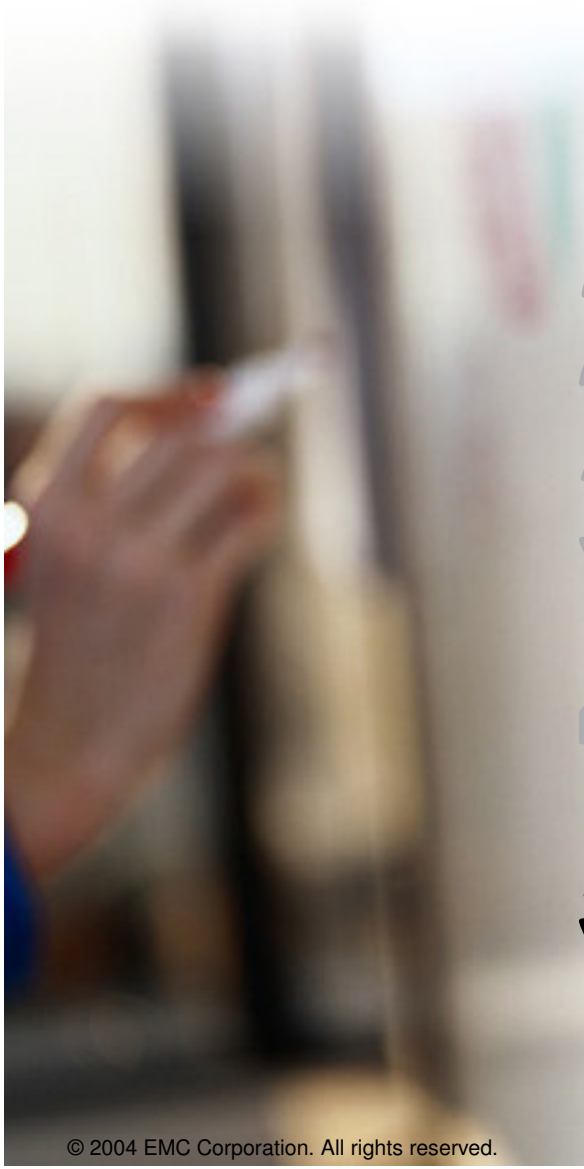
Step 4 – Develop the Implementation Plan

- Combine all data and make decisions based on:
  - Restore frequency per application
  - Restore time per application
  - Backup window per application
  - Retention period per application
- Assign specific technologies to address service levels

Developing the Plan—Example

	Uptime	*RTO*	*RPO*	*Solution*
Tier 1 Applications	24x7x365	Seconds	Last transaction	Remote mirror → B2D → tape
E-mail	24x7x365	Minutes	Full restore	Snapshot → B2D → tape → CAS
Tier 2 applications	Business hours	Minutes to hours	Minimal loss	Snapshot → B2D → tape
File servers	Business hours	Minutes to hours	Minimal loss	B2D → tape
Business records and long-term data	Business hours	Hours to days	Best effort (unless regulated)	ATA, CAS, or tape

Implementing Backup to Disk



1. Defining the success criteria for backup to disk
2. Discovering the content
3. Sizing capacity and performance – leveraging RPO and RTO
4. Developing the implementation plan
5. **Justification of the solution**

HP WORLD 2004
Solutions and Technology Conference & Expo

Step 5 – Justifying the Solution

Savings from:

- Deferral of tape equipment purchases
- Removal or re-deployment of existing drives and libraries
  - Maintenance, personnel etc.
- Reduced volume of tape media
- Vaulting expenses
- Floor space and power

Revenue from:

- Higher system availability
- Accelerated restore / recovery
- Staff re-allocated to more productive work

Justifying the Value

Business Impact

- Application availability
  - Faster backup and restores

Operational Impact

- Reduce numbers of cartridges, drives and libraries to manage
- Staffing
  - More efficient use of Librarian, vaulting, and other staff

Financial Impact

- Minimize future tape infrastructure investments
  - Floor space and power
  - Reduced media replacement costs
- Vaulting expenses
  - Reduced media into and out of Vault site

Implementation Steps

- Enable backup software for backup-to-disk functionality (may require application upgrade or enhancement)
  - Not necessary with Disk Library
- Re-assign chosen application backups destination to disk
  - Set frequency and retention periods
- Configure movement of aged data from disk to tape
  - Set frequency and retention periods
- Update vaulting procedures accordingly

LEGATO NetWorker
Disk-based Backup Solutions
Faster, More Predictable Backup and Recovery

- Shorter backup times
  - Meet more aggressive backup windows and improve availability
- Faster and consistent restore times
  - Deliver higher and more predictable service levels
- Improved reliability
  - Greater confidence in the recoverability of critical business data
  - Redundancy, RAID, and high-availability features built in
- Easier management
  - Reduce complexity and administration of tape infrastructure
- Investment protection
  - Leverages existing connectivity (SAN or NAS)

Is Backup to Disk Right for You?

- Do you require faster recovery to meet stringent service level requirements and reduce downtime costs?
- Do you require a faster backup solution to meet the demands of increasing data volumes and a shrinking backup window?
- Do you seek to lower tape storage TCO and increase ROI?

For More Information

Come by LEGATO or EMC booths TODAY!

Contact:
Eric Carter, Product Manager - ecarter@legato.com

On the Web:
LEGATO NetWorker DiskBackup Option
http://legato.com

EMC Backup-to-Disk Solutions
http://www.emc.com

Hewlett Packard
http://www.hp.com

HP WORLD 2004
Solutions and Technology Conference & Expo

HP WORLD 2004

Solutions and Technology Conference & Expo

Co-produced by:

interex

shared knowledge • shared power

encompass

AN HP USER GROUP