



Design, Care, and Feeding of an Enterprise Backup System

UUASC/LA, March 2006

Michael Barrow

<michael at michaelbarrow dot name>



Who is this guy?

- Technical lead in Amgen's Enterprise Backup Group
- Define enterprise-wide standards and practices for backup, archive, & restore
- Working in the industry for 16 years
 - Industries: ISPs, academia, consulting, dot-coms, and enterprise
 - Domains: sysadmin, programming, network design & engineering, network security
- Spent the last 2½ years eating, sleeping, and breathing backups
 - Currently implementing backup redesign for Amgen



Agenda

- **Business Side**
 - Disaster recovery & business continuity
 - Reasons for performing backups
 - Designing backup solution
- **Technology Side**
 - Overview of backup technologies
 - An enterprise backup environment
 - NetBackup



DR versus BC

- Disaster Recovery (DR)
 - How to restore the business quickly after a disaster
- Business Continuity (BC)
 - How to keep the business running after the disaster until environment returns to normal
- Key focus areas
 - Business processes
 - Resources (technical, real estate, human, etc.)
 - Communication



Case Study from 9/11

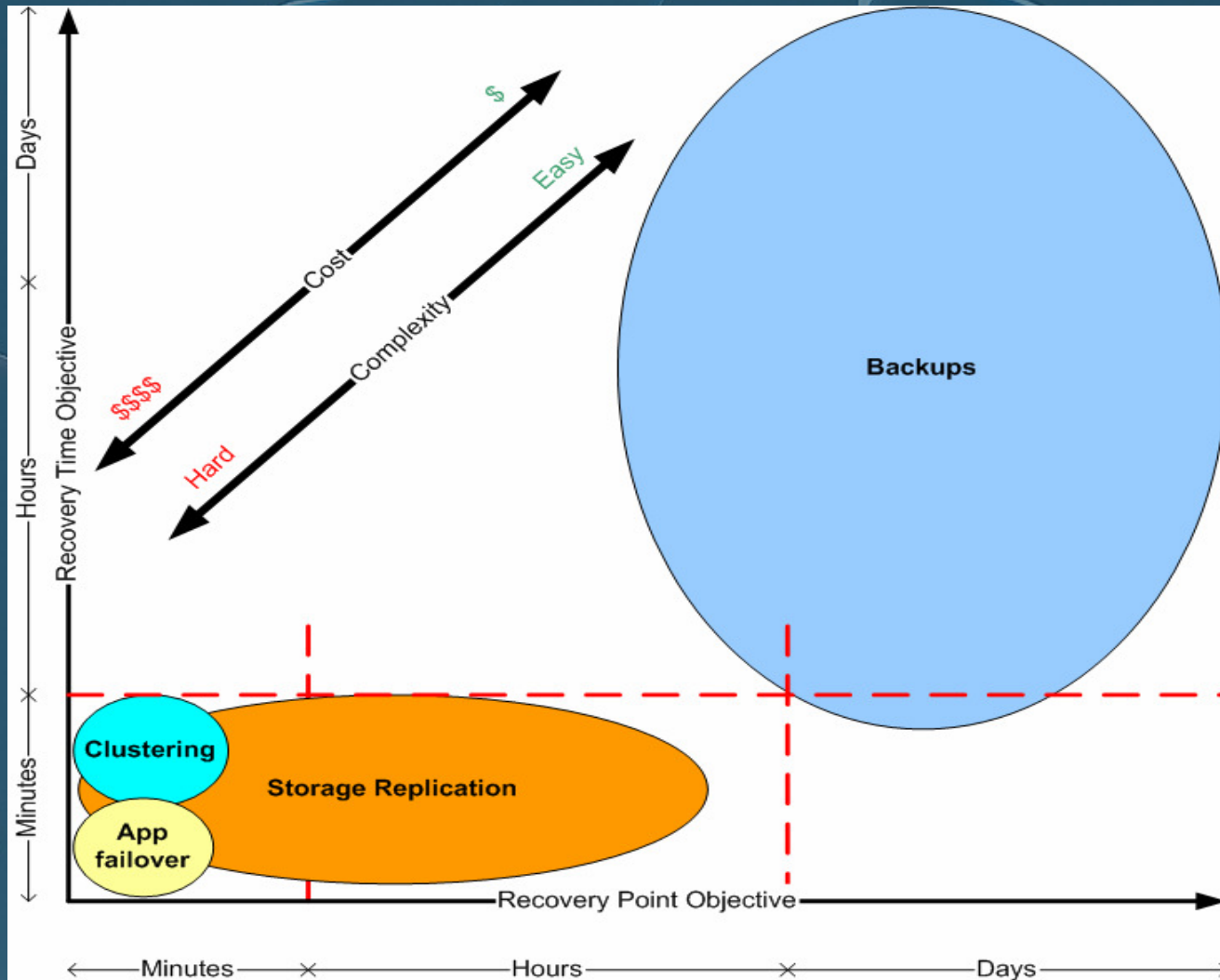
- **Impact**
 - Main NYC data center was 6 blocks from WTC
 - Closest CO was severely damaged
 - Loss of power, telecommunications, access
- **Disaster recovery activities**
 - Accounted for all employees
 - Damage assessment of infrastructure
 - "Hijacked DNS" with minimal failover to CA
- **Business continuity activities**
 - Acquired generator for data center
 - Contacted customers to make arrangements
 - Brought up replica systems in midtown, CA, and Long Island
 - Baby-sitting plan for downtown data center



Key Factors of DR Solution

- **Recovery Time Objective (RTO)**
 - How long it takes to recover a system after a disaster
- **Recovery Point Objective (RPO)**
 - How far behind a system is once it has been recovered

Comparison of DR Solutions





Why Backups?

- Data protection/disaster recovery
 - “Oopsies”
 - Data corruption – hardware or software
 - Catastrophic failures
 - Natural: earthquake, fire, flood, tornado, ...
 - Man made: theft, terrorism, ID-10-T, etc.
- Data preservation, a/k/a archiving
 - Point in time snapshot of system
 - Migrate data off-line
- Support larger ILM strategy



Business Design for Backups

- Data protection (backups)
 - Define different classes of service based on RPO/RTO -- it's all about recovery!
 - Emphasize purpose of backups
 - Establish generations with finite cycle times
 - Focus: system/application admin
- Data preservation (archives & ILM)
 - Determine data retention period – work with compliance/lawyers
 - Establish good cataloging and indexing
 - Keep datasets isolated from each other
 - Focus: business/data owner



Technology Side

- Components of an enterprise backup environment
 - High speed networking
 - Disk-based backup
 - Encryption service
 - Removable media
- Real-world example
- NetBackup overview & demo



High Speed Networking

- **Fibre channel SANs are required**
 - All high-end backup gear has FC ports
 - Parallel SCSI is fast enough, but can be painful to administer
 - iSCSI used for some low-end/mid-range systems
 - Client-less & server-less backups need FC SANs
- **Backup LANs**
 - Isolate backup traffic from production traffic for performance & security
 - Backup networks may use different technology (fast Ethernet v. gigabit, jumbo frames, etc.)



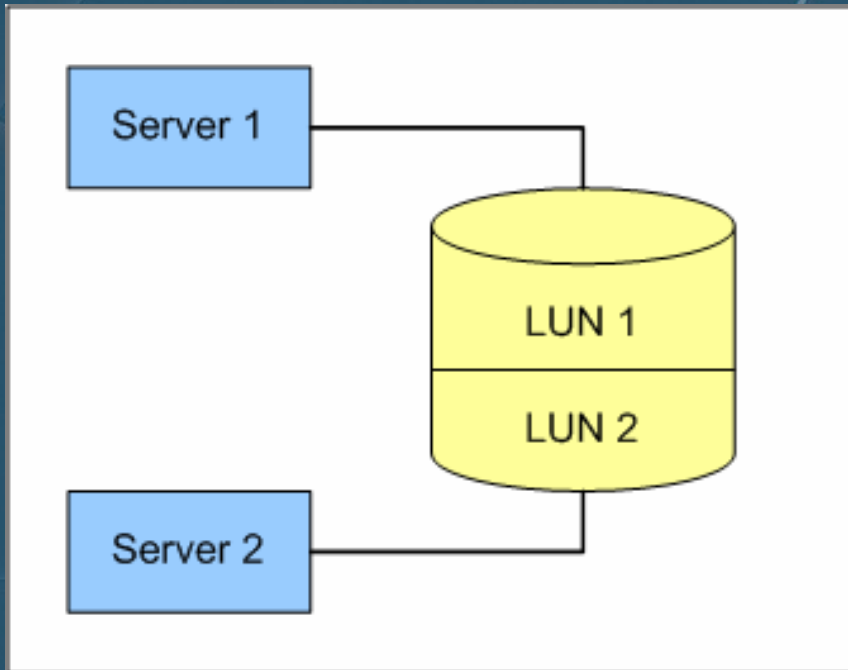
Disk-Based Backup

- Methodologies
 - D2T: disk to tape
 - D2D: disk to disk
 - D2D2T: disk to disk to tape, disk staging
- Disk as disk
- Disk as tape (VTL)
- Hybrid appliances
 - Massive array of idle disks (MAID)
 - Data reduction

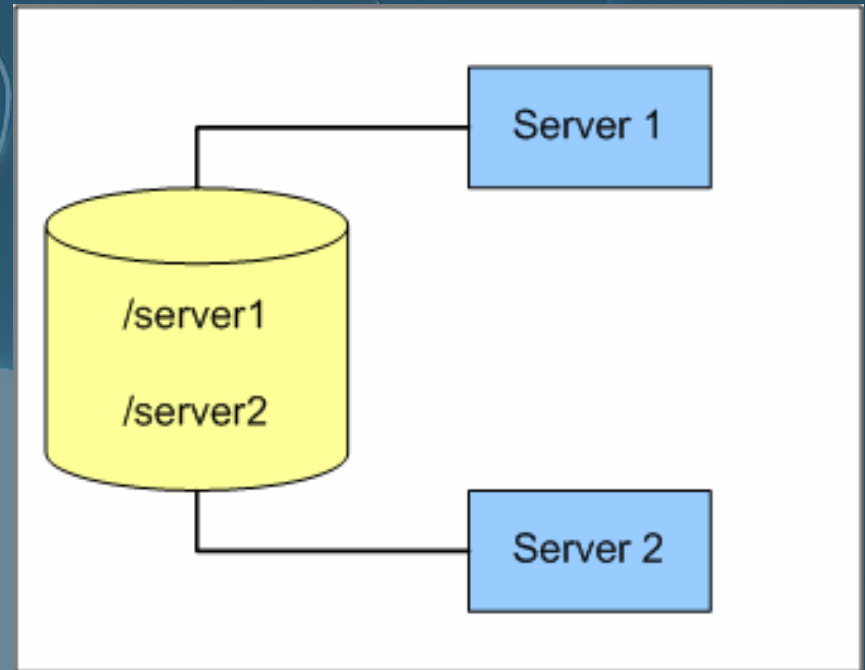


Disk as Disk

Block/SAN-based



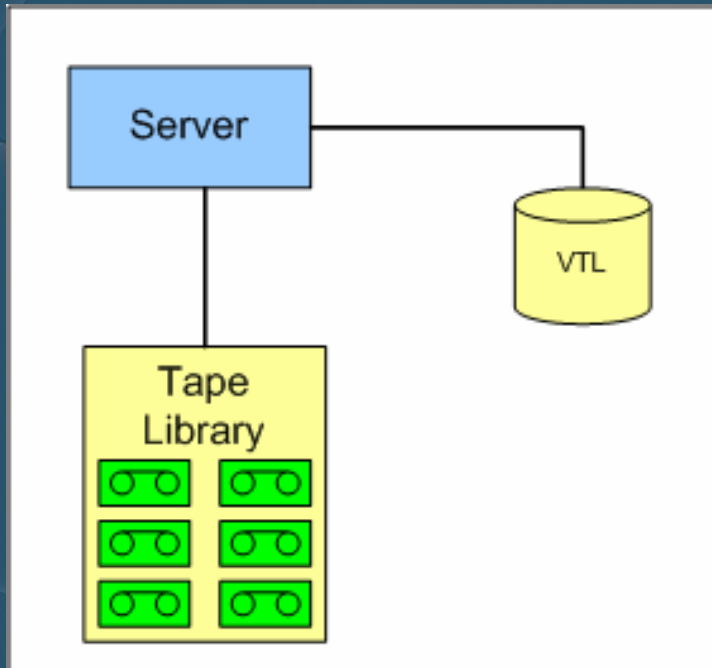
NAS-based



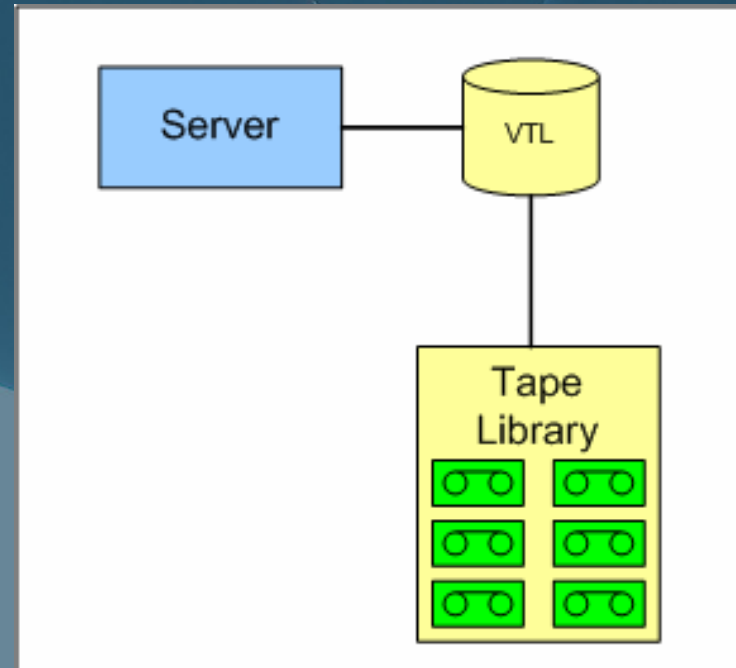


Virtual Tape Libraries

Stand alone



Integrated





Tape's Demise is Exaggerated

Format	Capacity	Data Rate
LTO2	400 GB	70 MB/s
LTO3 (W)	800 GB	160 MB/s
LTO4	1.6 TB	240 MB/s
LTO5	3.2 TB	360 MB/s
LTO6	6.4 TB	540 MB/s
SDLT 320 (W)	320 GB	32 MB/s
SDLT 600 (W)	600 GB	72 MB/s
DLT-S4	1.6 TB	120 MB/s
DLT-S5	3 – 3.5 TB	200 – 250 MB/s
DLT-S6	6 – 7 TB	400 – 500 MB/s
DLT-S7	12 – 14 TB	800 – 1000 MB/s

Format	Capacity	Data Rate
AIT-3	260 GB	31.2 MB/s
AIT-4	520 GB	?
AIT-5	1 TB	?
AIT-6	2 TB	?
SAIT-1 (W)	1.3 TB	78 MB/s
SAIT-2	2.8 TB	?
SAIT-3	5.2 TB	?
SAIT-4	10.4 TB	?
T9940B (W)	400 GB	70 MB/s
T10000 (W)	1000 GB	240 MB/s

Red: Currently available

(W): WORM available



Encrypt this, buddy!

- Tapes get lost or stolen
 - BofA: 1.2M victims, including members of Congress
 - AOL/TW: 600K victims
 - ABN Amro: 2M potential, but tape found
 - Marriott Vacation Club: 206K victims
- California 2002 SB 1386
 - "Any agency that owns or licenses computerized data that includes personal information shall disclose any breach of the security of the system following discovery or notification of the breach in the security of the data to any resident of California whose unencrypted personal information was, or is reasonably believed to have been, acquired by an unauthorized person."
 - Personal information: first name or first initial and last name in combination with any one or more of:
 - SSN
 - Driver's license or California ID card number
 - Account number, credit or debit card number, in combination with PIN or password that would permit access to financial account

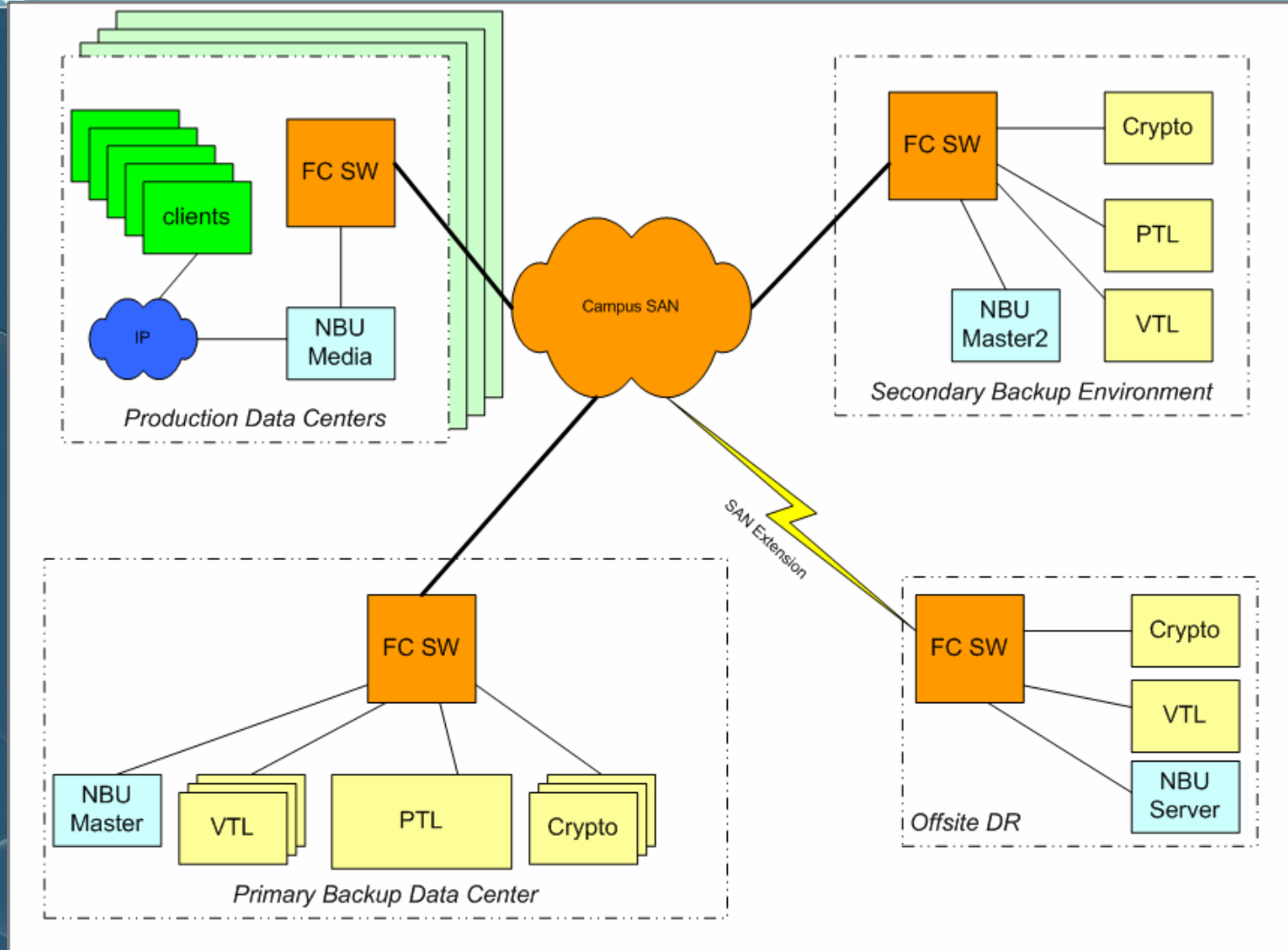


Encryption Solutions

- Encrypt within application (e.g. Oracle Advanced Security option)
- Methods to encrypt data at rest
 - PGP/GPG
 - Backup software
 - Encryption appliances



Real World Example





NetBackup 6.0

- Veritas purchased by Symantec
- Major features
 - Updated database – Sybase engine
 - NetBackup Operations Manager
 - Disk-based backups
 - Advanced client
 - NetApp integration
 - Synthetic backups