Multi-Terabyte Database Backup and Restores Over High Speed Networks

Marty Stogsdill

Discover Financial Service's Most Recognizable Brands

DISCOVER

- 1 in 4 US Households
- Leading cash rewards program

Diners Club INTERNATIONAL

- 50+ Franchises / 80+ Licensees
- Over 110K corporate clients

pulse

- #3 Market share in PIN debit
- 4,400+ Issuers

DISCOVER HOME LOANS

- Home loans, refinances, equity loans
- Real Estate Cash Rewards

DISCOVER BANK

- Online Bank
- Introducing Cashback Checking in 2014

DISCOVER STUDENT LOANS

- College and Graduate Student Loans

DISCOVER

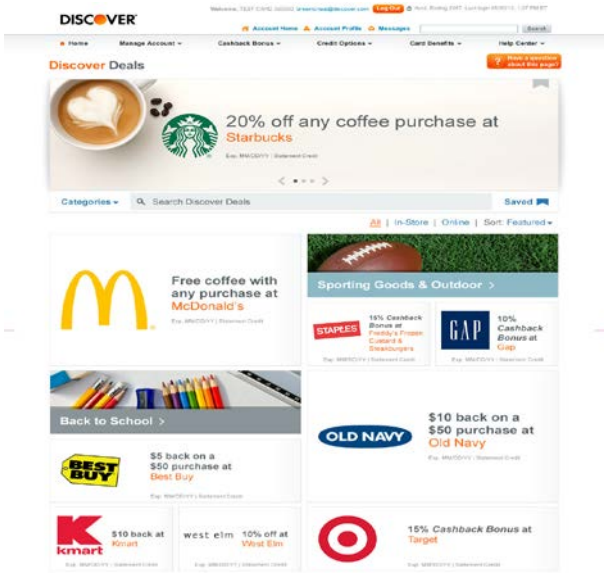
As Real Time as Possible

- DFS's business unit's interactions with customers are increasingly real time or near real time
  - SMS, Mobile App, Social Media, Website, Phone, Mail
- Faster, more intelligent responses, requires more data in OLTP systems
- Data Warehousing, BI/Analytics still important but nightly batch runs increasingly are not enough

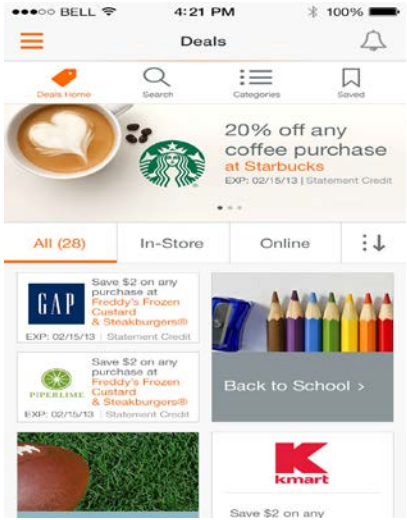
Meaningful Promotions

- ✓ Offering card holders useful promotions requires analytics
- ✓ Presenting the offer requires infrastructure and web development to execute seamlessly

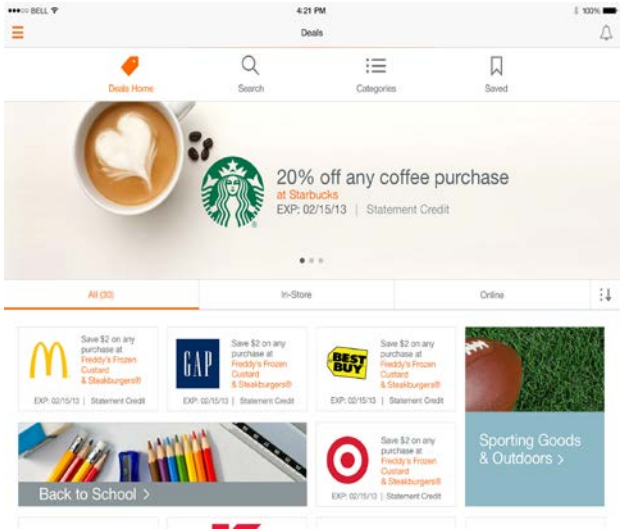
Big Browser



Mobile



IPad



DFS Enterprise IT Environment

Challenges and Opportunities

DISCOVER®

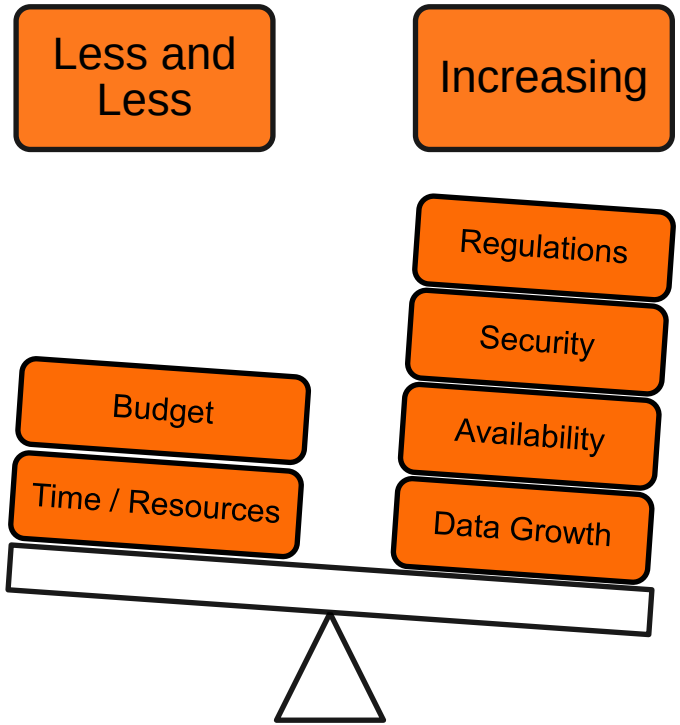
Familiar Balancing Act?

PCI DSS v3

SOX

Big Data

Internet of Things



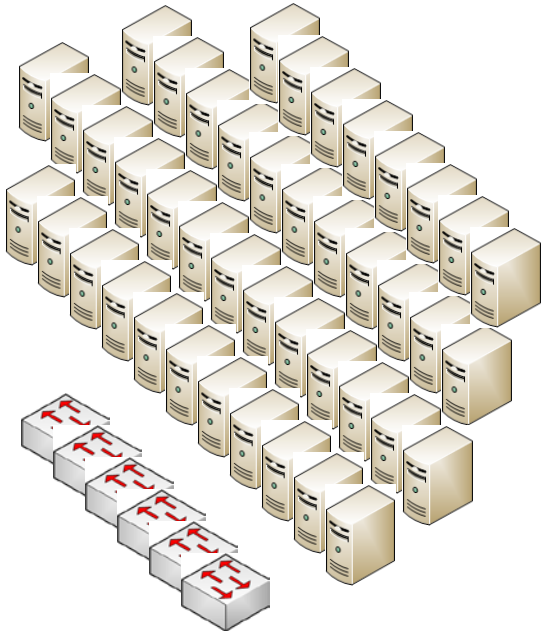
Audit

99.999%

Flash

Cloud

Traditional Enterprise IT Response



Standardize

- Unix Configuration
- x86 Configuration
- Tier 1 SAN
- Tier 2 SAN
- Switches

Virtualize

- Abstraction of Resources
- Higher Utilization Rates of Physical Assets

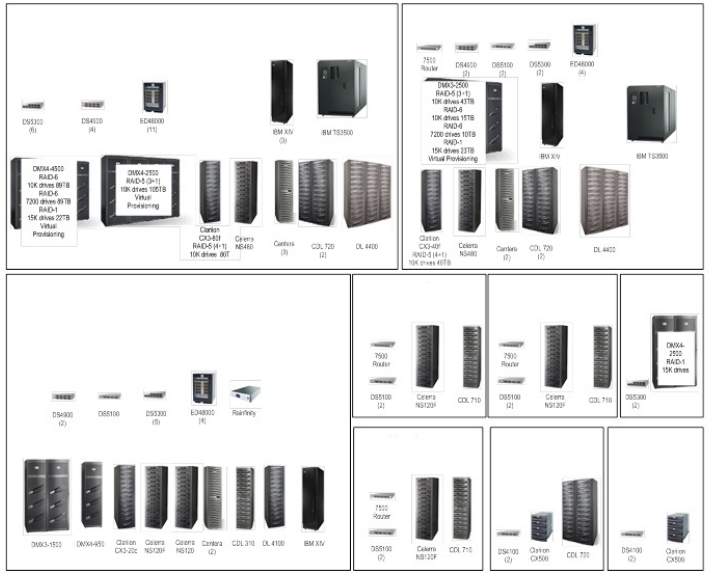
Consolidate

- Less Servers
- Less Storage Arrays
- Less Cables
- Less SW Licenses

- Lower Capital and Operating Costs
- Staffing Efficiencies
- Faster Incident Resolution

Legacy OLTP Database Environment

- Mix of Unix and Linux clusters
- FC attached Tier 1 SAN for data
- NAS attached Tier 2 for backup / restore
- Challenges
  - Deployments limited to standard offerings for Server and Storage
  - High administration cost during troubleshooting
  - Databases routinely I/O bound
  - Backup / restore not scaling with data growth



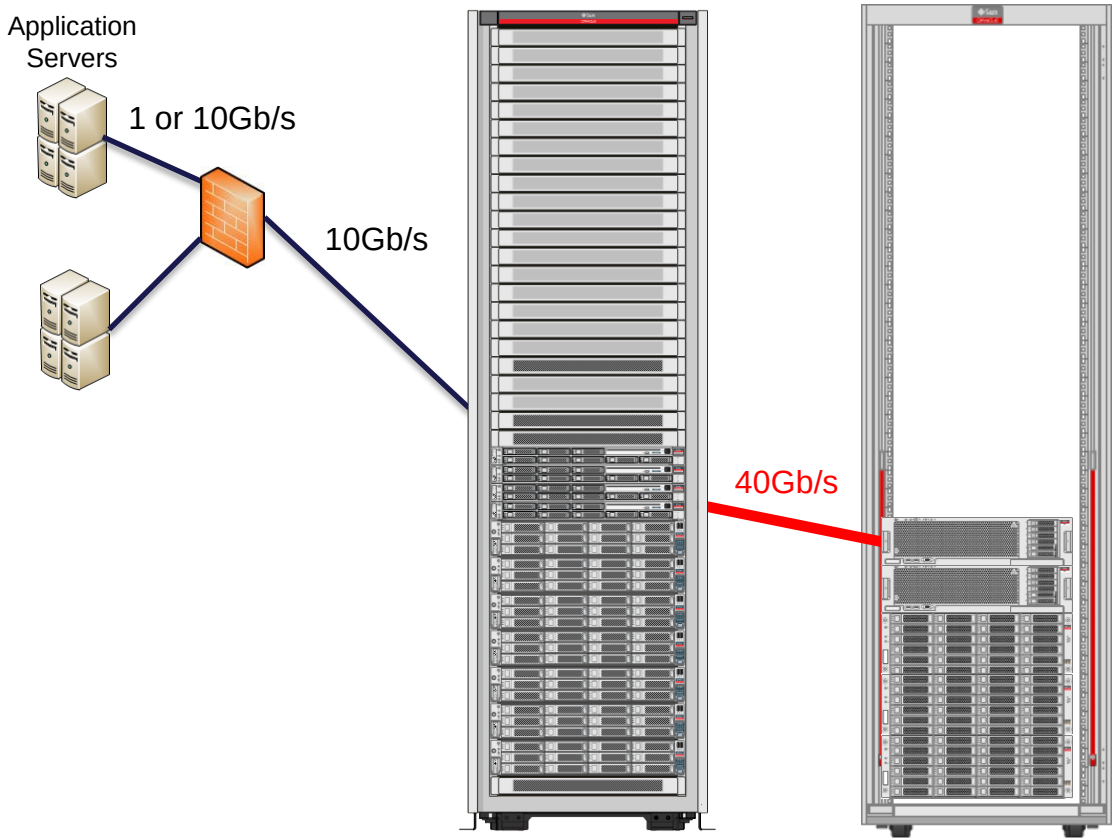
Data Growth

- 2011 capacity planning predicted that 2014 would have:
  - **23.3TB** of OLTP database production data (not counting indexes, archives, redo logs, overhead, etc.)
  - **319TB** of OLTP database disaster recover, backups, and non-production data
- 2014 so far:
  - **75+TB** of OLTP database production data
  - **1PB+** of OLTP database disaster recover, backups, and non-production data
- Growth plan only off by **3.2X**

Technical Design

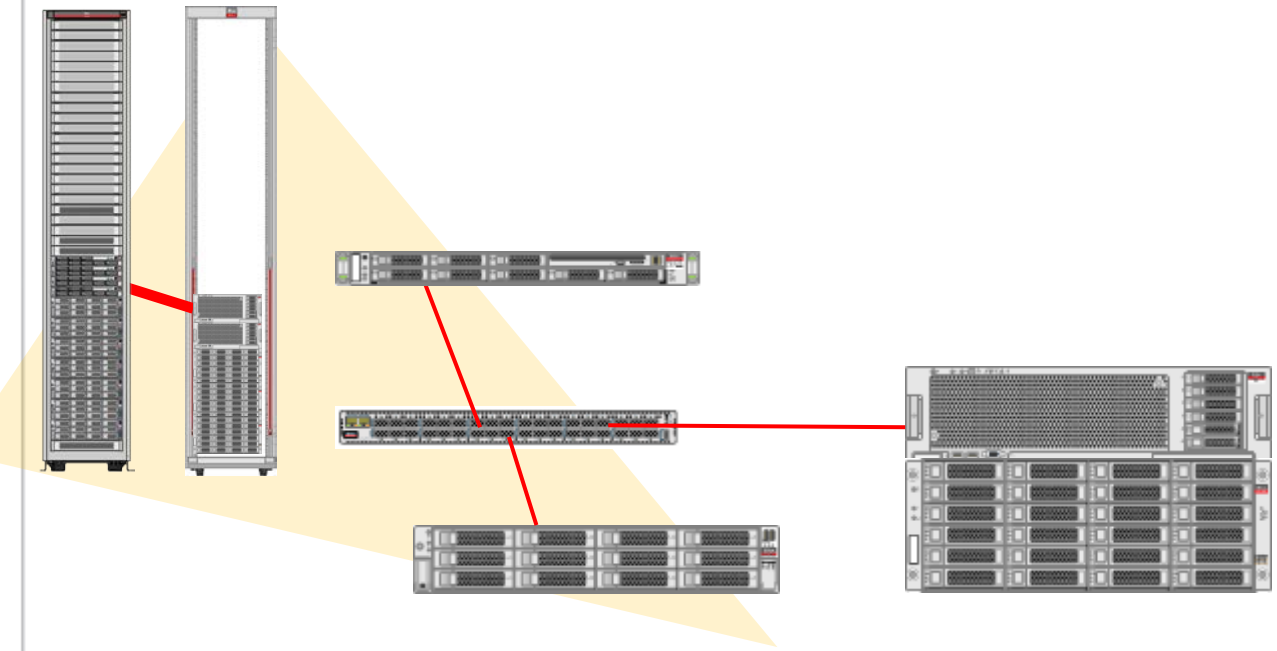
Purpose Built OLTP Database Environment

High Level OLTP Database Hosting Design



- 10GbE to database servers
- 40GbE Infiniband from database servers to x86 storage servers
- 40GbE Infiniband connectivity to external NAS array for backup / restore

Zoomed in High Level Design

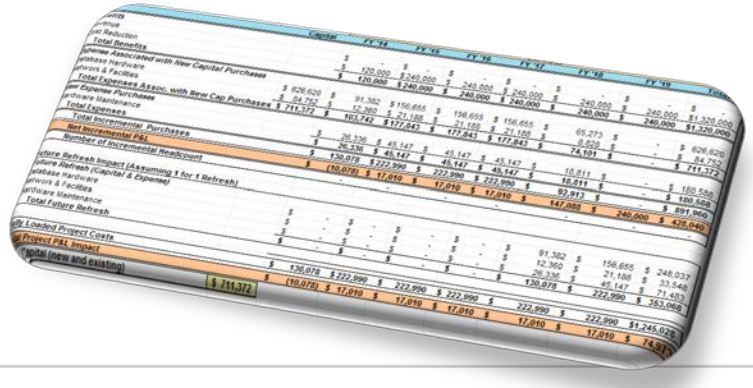


40Gb/s InfiniBand

- Database Server
  - x86 processors, DRAM
- Storage Server
  - x86 processors, flash, DRAM, commodity disk
- NAS
  - x86 processors, flash, DRAM in the controller
  - Commodity disk in drive shelves

Buy vs. Build

- DFS choose commercially available vendor solution leveraging commodity hardware
- Capital depreciation, software licensing, and support costs lower than legacy storage solution
- DFS has not evaluated opensource custom built / self supported solutions for OLTP database backups but will in the future

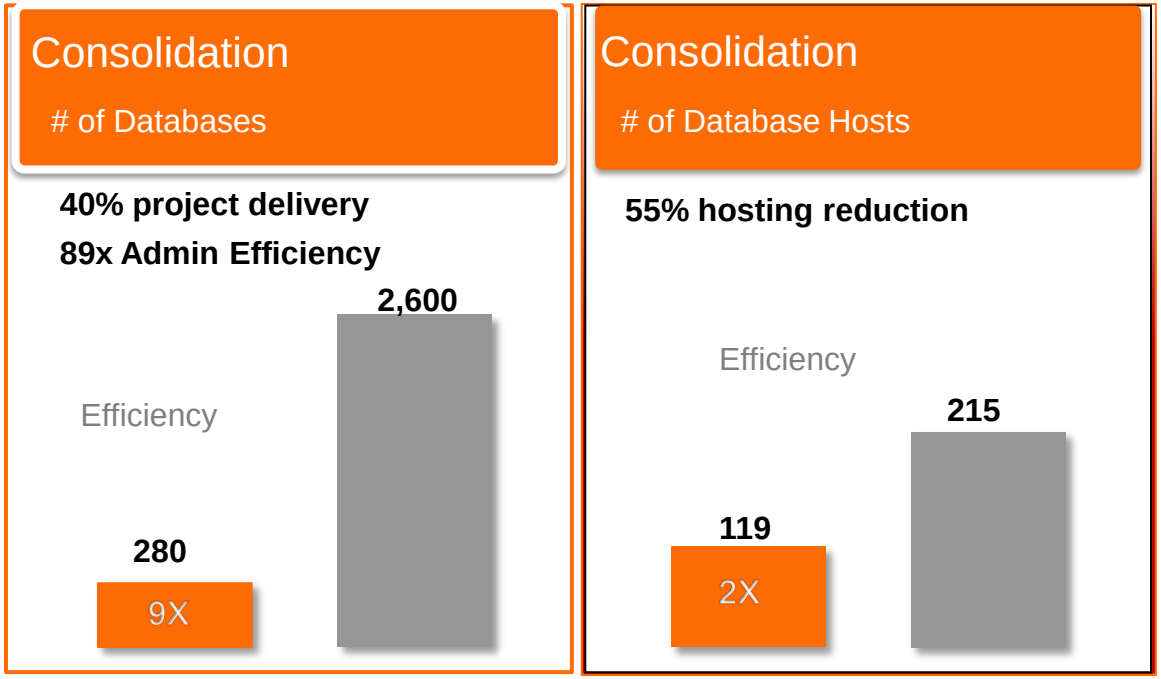


Metrics

Consolidation, Throughput, and Data Protection Strategies

DISCOVER

Database Consolidation

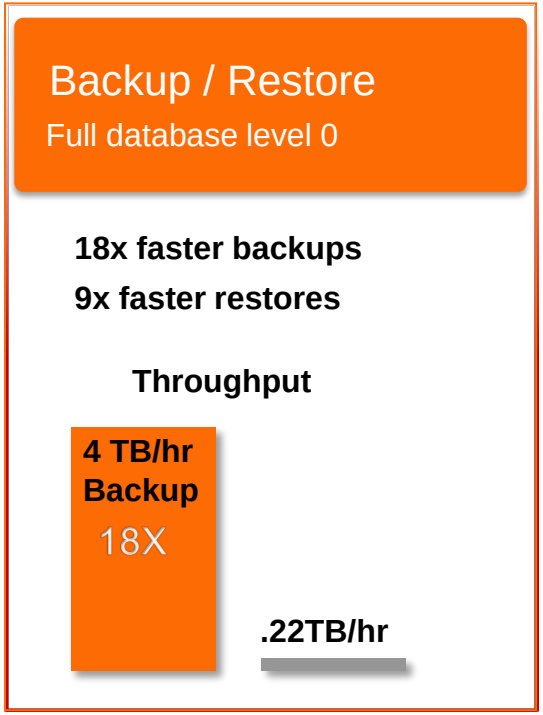


Consolidation Strategy

Databases consolidated by:

- Criticality
- Security requirements
- Change windows
- LOB channel isolation

Backup / Restore Throughput



- Limiting factor is database backup / recovery script tuning
  - Online database backups can not impact production
  - Compression and encryption takes time/CPU and reduces I/O throughput
- Currently do not have enough database servers on a single Infiniband fabric to saturate
- Testing showed similar throughput with 10GbE attached NAS but higher risk of saturation

Data Protection Strategy

Backups

- ❑ Database flashback features first line of defense
- ❑ Disk based backup primary site
- ❑ Compression level dependent on classification of application
- ❑ Mission Critical lower compression ratio = faster restore but more storage consumption
- ❑ Standby (DR) database
- ❑ Disk based backup at standby site

Restores

- ❑ Database flashback features first
- ❑ Out of place restore environment for tablespace / object level recovery
- ❑ In place recovery for issue impacting entire database

Marty Stogsdill

MartinStogsdill@discover.com