Steven Kleiman

SVP & CTO

Network Appliance, Inc.

The Data Management Challenge

Steven Kleiman

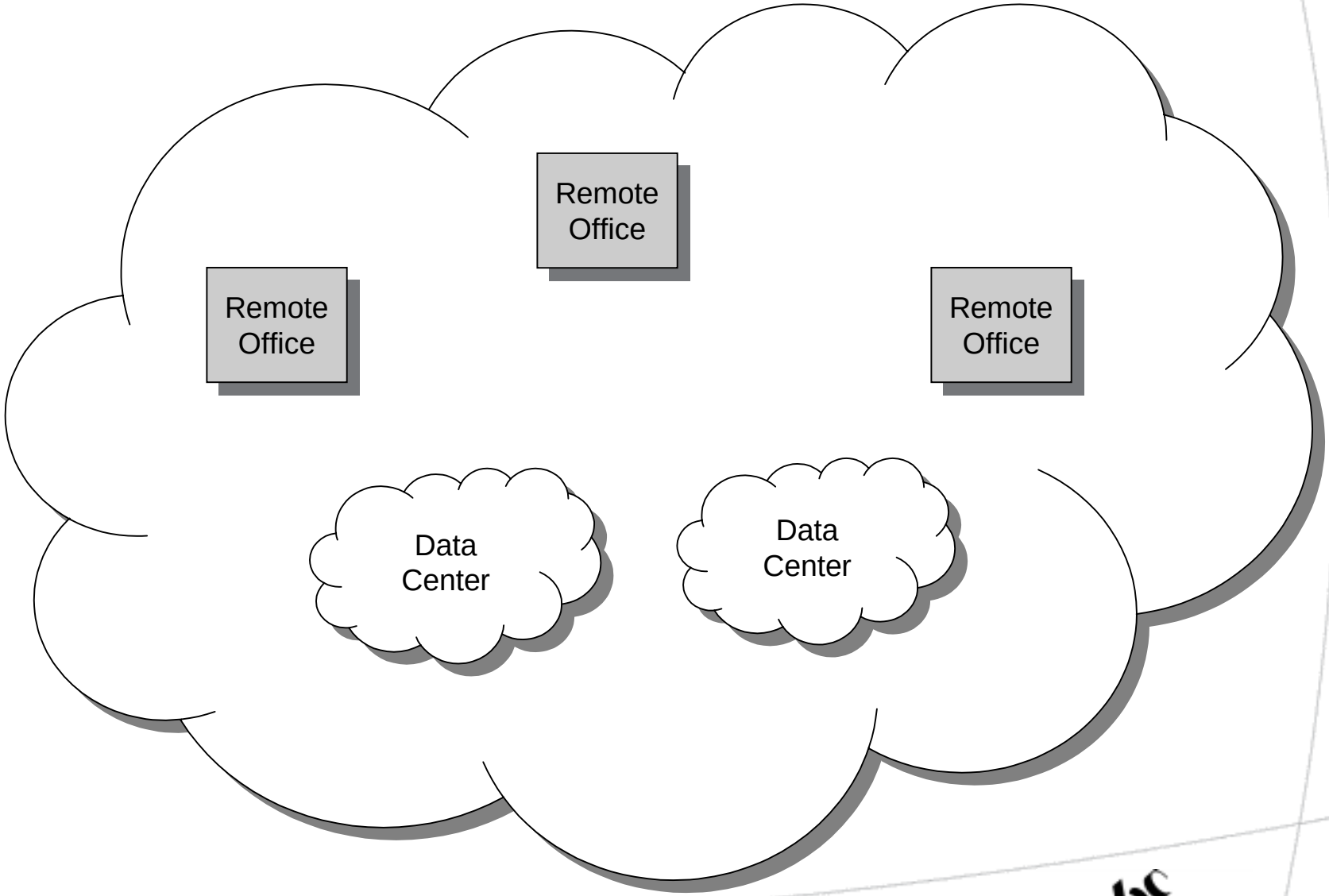
SVP & CTO

Network Appliance, Inc.

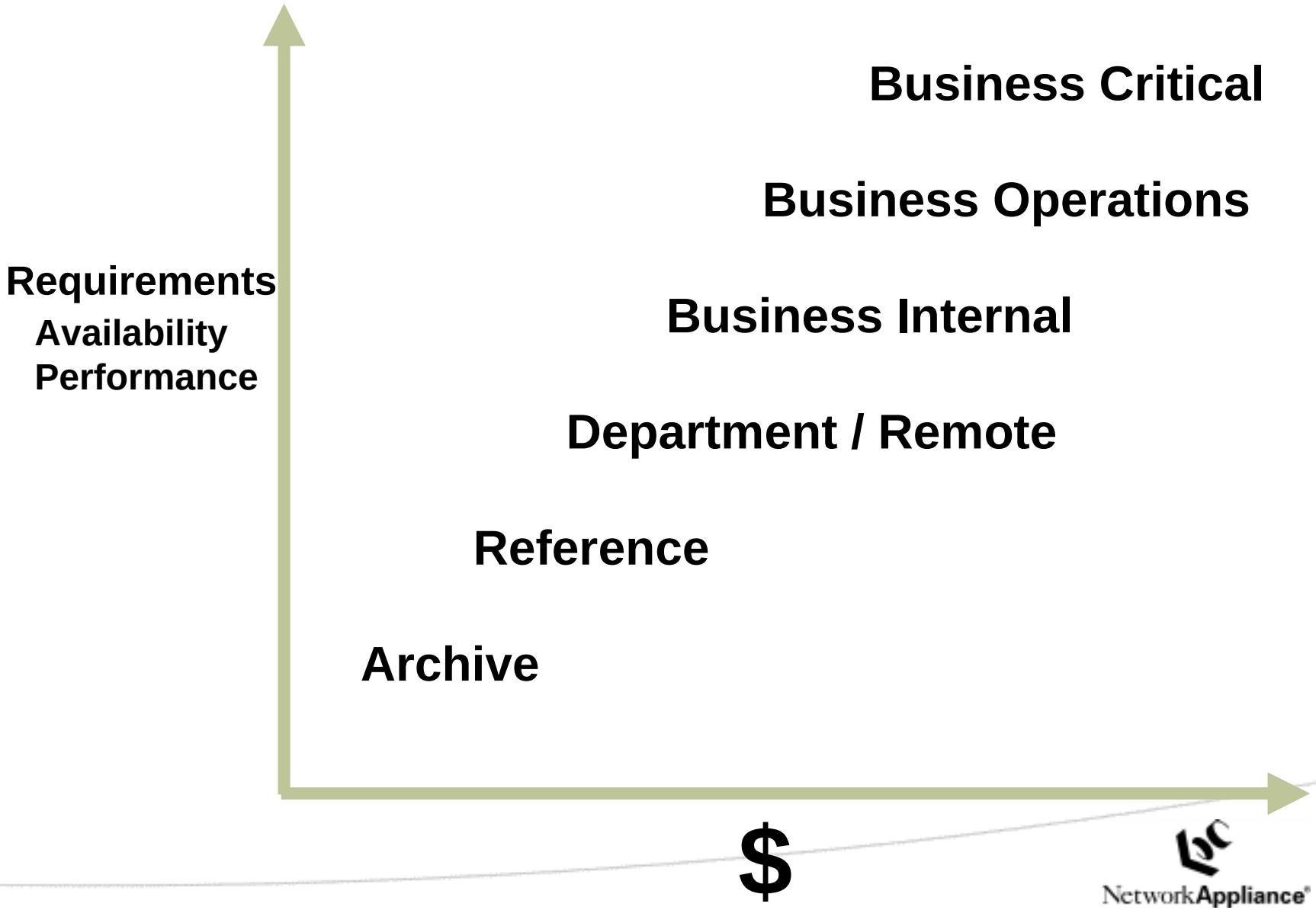
Data storage problems have evolved

- Petabytes of data
- Several geographically dispersed data centers
- Many remote locations
- Thousands of users and nodes
- Diverse applications
  - Database, collaboration, web, streaming, backup, home directory, simulation, …
- Different data availability and recovery requirements
- Different types of access
  - Web: HTTP, FTP
  - Streaming: Real, Windows Media, Quicktime, MPEG4
  - File access: NFS, CIFS, DAFS
  - Block access: FCP, iSCSI

Big Cloud/Little Cloud



Not All Data Is Created Equal



Evolution of Enterprise Architectures

- **Proprietary Networks**
  - Single vendor solutions
- **Open Protocol LANs**
  - Within workgroups or buildings
- **Open Protocol WANs**
  - Connect locations
- **Enterprise Network Infrastructure**
  - Enterprise-wide, multi-protocol, network architecture
- **Proprietary Storage**
  - Single vendor solutions
- **Open protocol NAS and SAN**
  - Within data centers, workgroups or buildings
- **Long distance data**
  - Disaster recovery
- **Enterprise Data Infrastructure**
  - Enterprise-wide, multi-protocol, data architecture

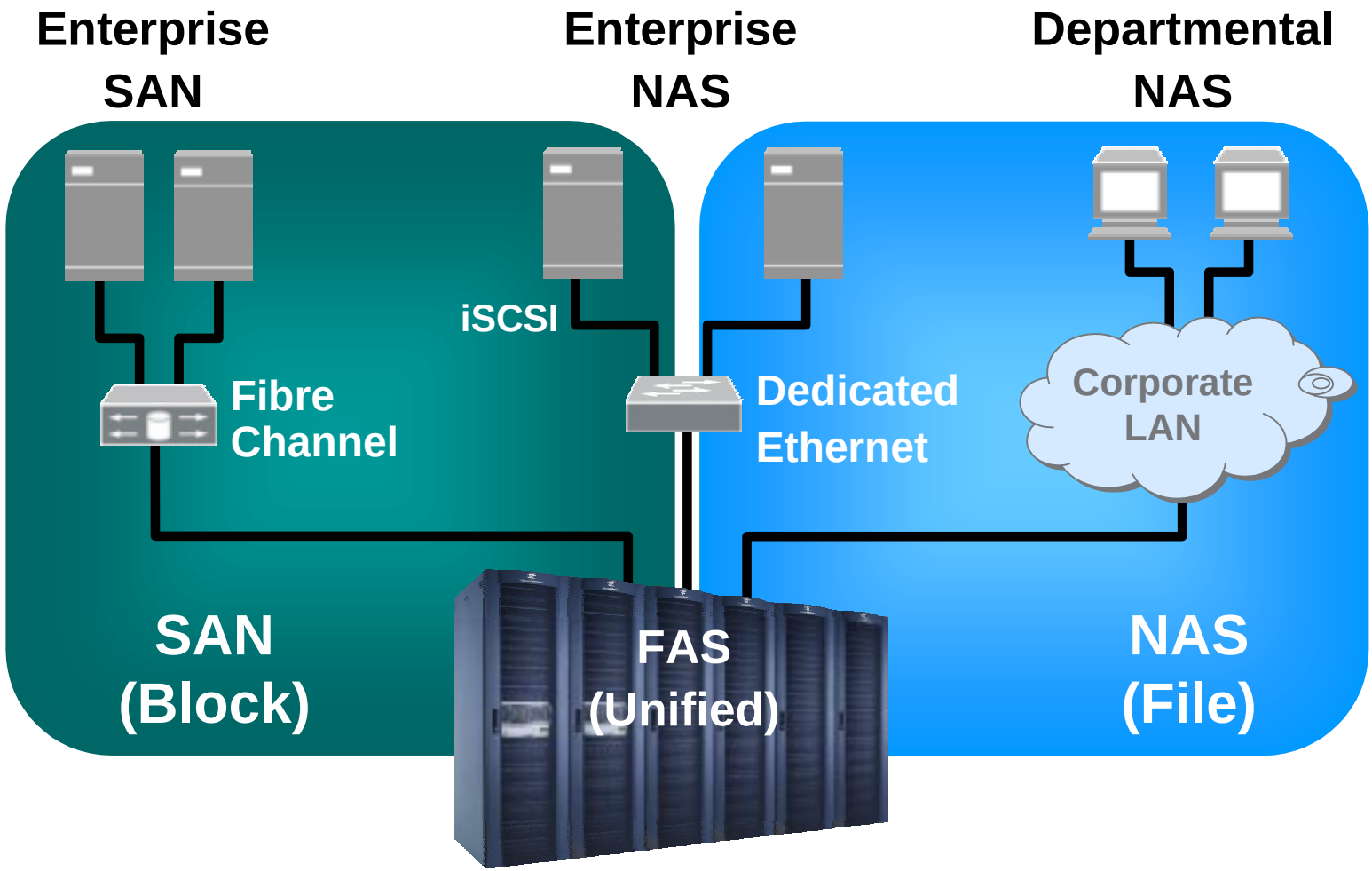
Technology has evolved

- Network links are as fast as storage links
  - Started 100:1, now 1:1 (or better)
- Low cost, high volume, reliable disks
  - Raw disk $/GB can be as cheap as tape
    - Disk is getting cheaper faster
  - High throughput *and* random access
- High speed, long distance links are getting cheaper
  - MAN technology
- Commodity high performance data center fabrics
- Direct application->storage access
- Volume servers are best price/performance

Data Management

- **The solution is not to have lots of things that are each easy to manage**
  - **Drag and drop GUIs won't solve it**
- **Need to eliminate or drastically simplify some existing paradigms**
  - **SAN/NAS separation, Backup/restore , disaster recovery, data distribution**
- **Virtualization is important!**
  - **Unhitch the client model from physical layout**
  - **Both file and block virtualization are required**
  - **Beware of trying to hide "1,000 chickens"**

Unified Storage



Files vs. Blocks

- **Things a block storage device knows:**
  - **All blocks may have data**
  - **The geometry of underlying RAID devices**
- **Things a file storage device knows:**
  - **The set of blocks that have allocated data**
  - **The set of blocks that comprise a file**
  - **The set of blocks that comprise a file system**
  - **The set of users that have permission to access each block**
  - **Which application has locks on each block**
  - **The geometry of underlying RAID devices**

Lower Cost Storage

Redundant Array of Inexpensive Disks

Redundant Array of Independent Disks

Redundant Array of Inexpensive Disks

Lower Cost Storage

- **Volume drives**
  - **Much lower $/GB**
    - **Raw disk approaching tape media**
  - **Lower $/op then enterprise drives**
  - **Slower than enterprise drives, a lot faster than tape**
  - **Less reliable than enterprise drives (?), more reliable than tape**

Tape Backup Issues

- Expensive robots
- Tape is a poor recovery mechanism
  - Typical tape drive: ~14-27 hours/Terabyte
- Long latency to restore individual files
- Not accessible or searchable
- Lots of staff to manage
- Not practical for small, remote sites

Simplified Data Recovery Models

- **Self-recovery, file restore**
  - **User access of Snapshot data**
- **Recover from inconsistency**
  - **Fast restore of large data from Snapshots**
- **Easy Hierarchical Storage**
  - **SnapVault: long term Snapshot storage on another filer**
  - **NearStore as backing storage**
- **File caching …**

File Caching

- Local data performance at remote sites
  - CIFS/NFS/iSCSI/FCP
- Hierarchical storage
- File server acceleration
- File and block virtualization
- Replace on failure

NFSv4 Improvements

- **Security**
- **Integrated locking & delegations**
- **Transparent data migration hooks**

- **However**
  - **Clients and servers need to exploit these features**

NFSv4.X Challenges

- **Transport**
  - **Overhead <= blocks: RDMA, user I/O**
  - **Sessions**
    - **Failover, trunking**
- **Data management**
  - **Data movement protocols**

NFSv4.X Challenges

- **Reliable grid computing**
  - **Fencing**
  - **Lock break indication**
  - **Rollback locking**
- **Scalability**
  - **Parallel files (e.g. pDAFS)**
  - **Separate control and data transfer (e.g. NASD)**

NFSv4.X Challenges

- **Application I/O**
  - **Traditionally constrained by LCD OS semantics**
  - **Examples**
    - **Direct application access**
    - **Bulk I/O**
    - **Cache control**

Beyond File Semantics

- Simple directory trees are probably inadequate for petabytes of data
- Indexing
- Stronger transactional semantics
- Query protocols

Summary

- Today's data management is too complex
  - Need to change or eliminate some paradigms
- New technology enables new strategies
  - Rethink:
    - Backup/restore techniques
    - Tape
    - Disaster recovery
    - Hierarchical storage
- Focus next generation file access protocols on Big Cloud and Little Cloud issues


NetApp FAS960