

# Global File Virtualization

Charles Fan, CTO, EMC Rainfinity



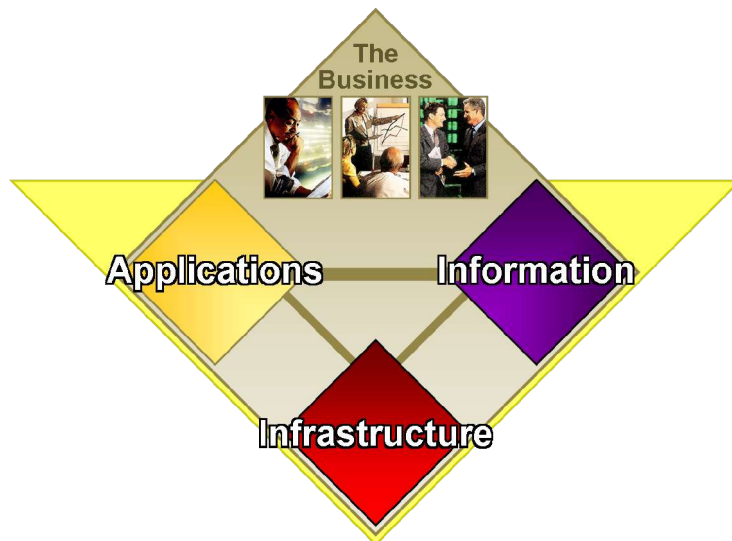
## Today's Enterprise Challenge: Managing Growth and Complexity

### Growth in Information – 60%+ per year

- Most growth is in unstructured data, in files and file systems
- NAS and file servers have proliferated
- The environment have become more difficult to manage

### Greater Complexity

- Increasingly complex resource management: more servers and storage devices to manage
- More data is in motion for protection, migration, and optimization



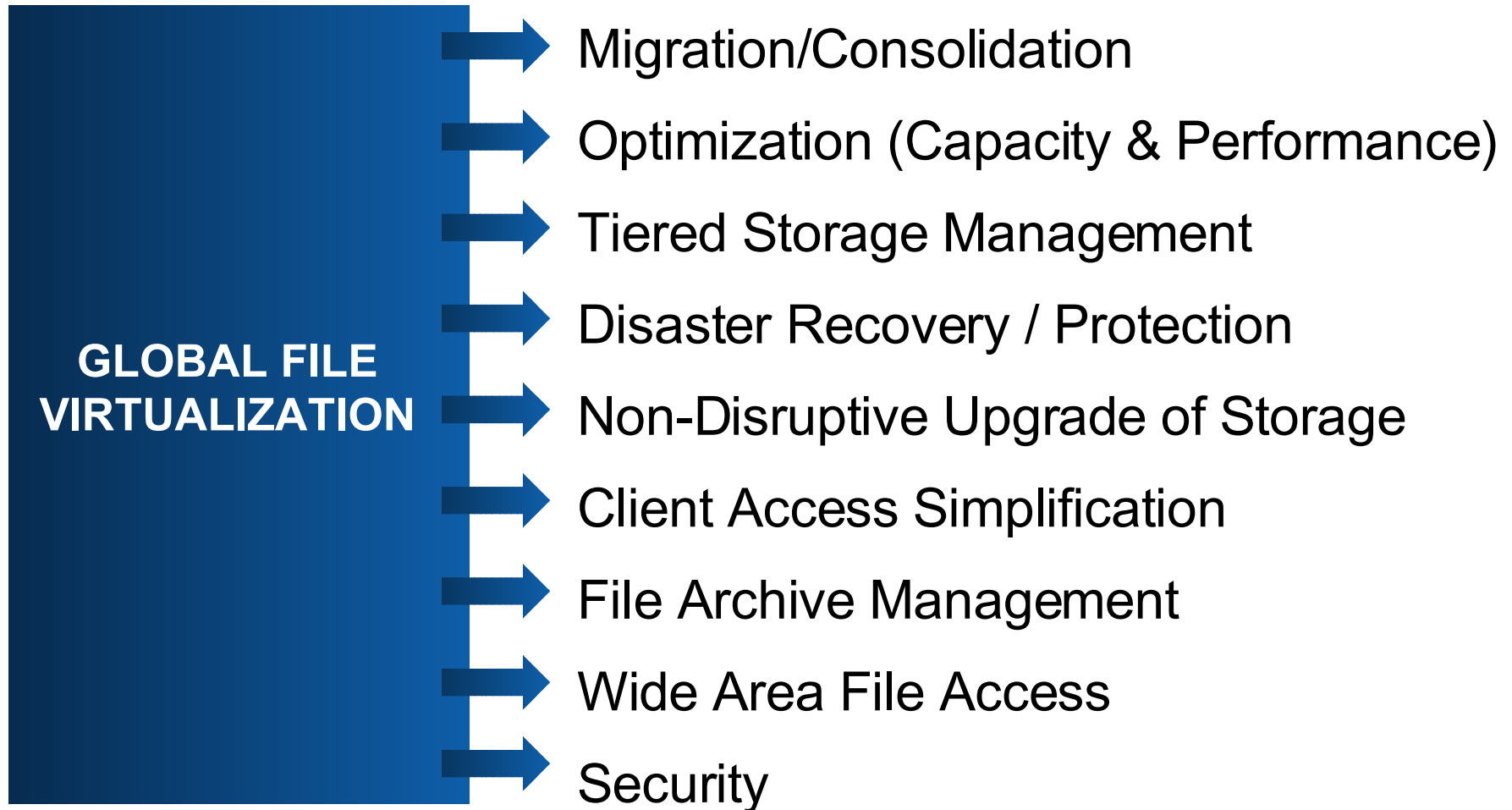
### Global File Virtualization

A key technology to

- Simplify
- Mobilize
- Optimize
- Scale Out

enterprise customers' file storage

## Key Global File Virtualization Applications



## Building Blocks

### GLOBAL FILE VIRTUALIZATION

**GLOBAL  
NAMESPACE**

**TRANSPARENT  
DATA MOBILITY**

**FILE LEVEL  
REDIRECTION**

*MANAGEMENT*

*AUTOMATION*

## Important Considerations: Safety

- **Data integrity is always priority #1**
  - Customers have zero tolerance for potential data corruption
  - No data integrity issues even when all GFV systems completely fail
  - Single authoritative copy at data any given time
- **Minimize potential points of failures**
  - Customers do not desire another hardware layer permanently on the data path
  - Stateless architecture to minimize data access exposures during failures
  - Require high availability
- **Support data protection well**
  - Must work with today's data protection solutions, including
    - Snapshot mechanisms
    - Backup/restore applications and
    - Mirroring solutions

## Important Considerations: Scalability

- **Scale out the size of data and amount of files of the data farm**
  - Petabyte is no longer a huge number
  - 10's, even 100's of billions of files
  - Expect accelerated growth
- **Scale to the growing throughput to the data farm**
  - Grid computing drives the NAS farm to the performance edge
  - Millions of ops/sec of NFS/CIFS access
- **Span multiple data centers around the globe**
  - File storage is more and more pervasive in an enterprise
  - Globalization is driving NAS to all continents of the world
  - Improved network connectivity elevates customer expectations

## Important Considerations: Transparency

- **Must non-disruptively handle active data**
  - Open files
  - Locks
  - Stale mounts
- **Require no mount point changes**
  - Applications and end-users access data same as before
  - Allow concurrent access both directly and via global namespace
  - Data location changes are masked from end users
- **No proprietary software on clients or servers**
  - Requires only standard NFS or CIFS clients
  - Support wide range of client operating systems
  - Support wide range of server vendors

## Important Considerations: Compatibility

- **Leverage existing storage investment**
  - Additive solution keep the value of existing investment
  - Manages heterogeneous multi-vendor environment
  - Facilitate easy deployment of 2<sup>nd</sup> vendor into the environment
- **Support phased deployment**
  - Integrates with customers' existing global namespace
  - Virtualized data co-exists with yet-to-virtualize data
- **File system functionalities and tools stay at file system**
  - Snapshots
  - NDMP-based backups
  - Access Control, Quota enforcement, etc.



## Building Blocks

### GLOBAL FILE VIRTUALIZATION

**GLOBAL  
NAMESPACE**

**TRANSPARENT  
DATA MOBILITY**

**FILE LEVEL  
REDIRECTION**

*MANAGEMENT*

*AUTOMATION*

## Global File Virtualization Architecture

### GLOBAL FILE VIRTUALIZATION

#### GLOBAL NAMESPACE

- Out-of-band
- Directory service
- Standards based
- Heterogeneous Support
- Span globally

#### TRANSPARENT DATA MOBILITY

- In-band file protocol processing
- Standards based (CIFS, NFS)
- Full read/write access to open files
- Security, lock compliant

#### FILE LEVEL REDIRECTION

- File system enabled
- EMC FileMover API
- NetApp fPolicy
- In-band processing when necessary
- Support NAS or archive platforms

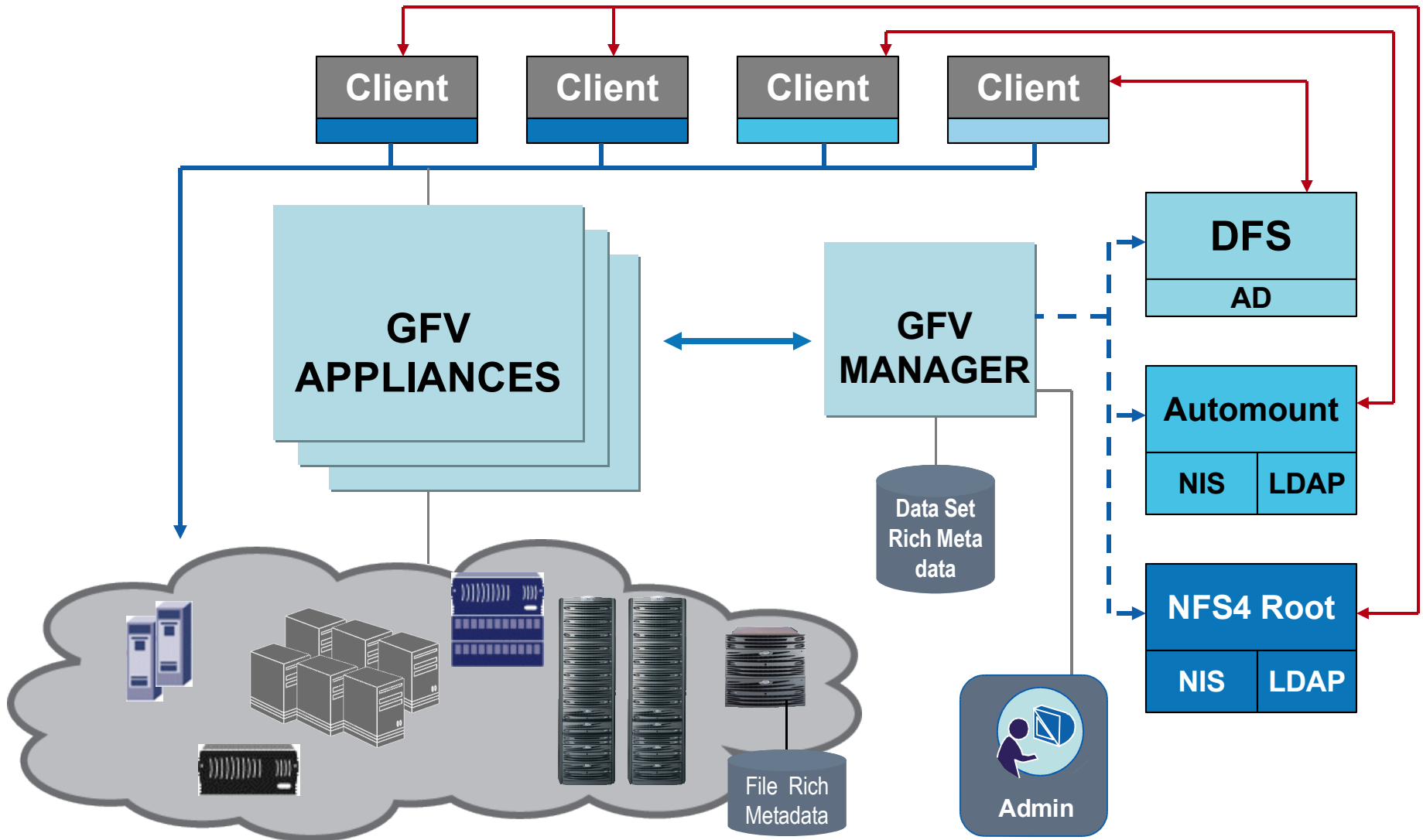
#### MANAGEMENT

- GUI, API, CLI
- Easily integrates with other solutions

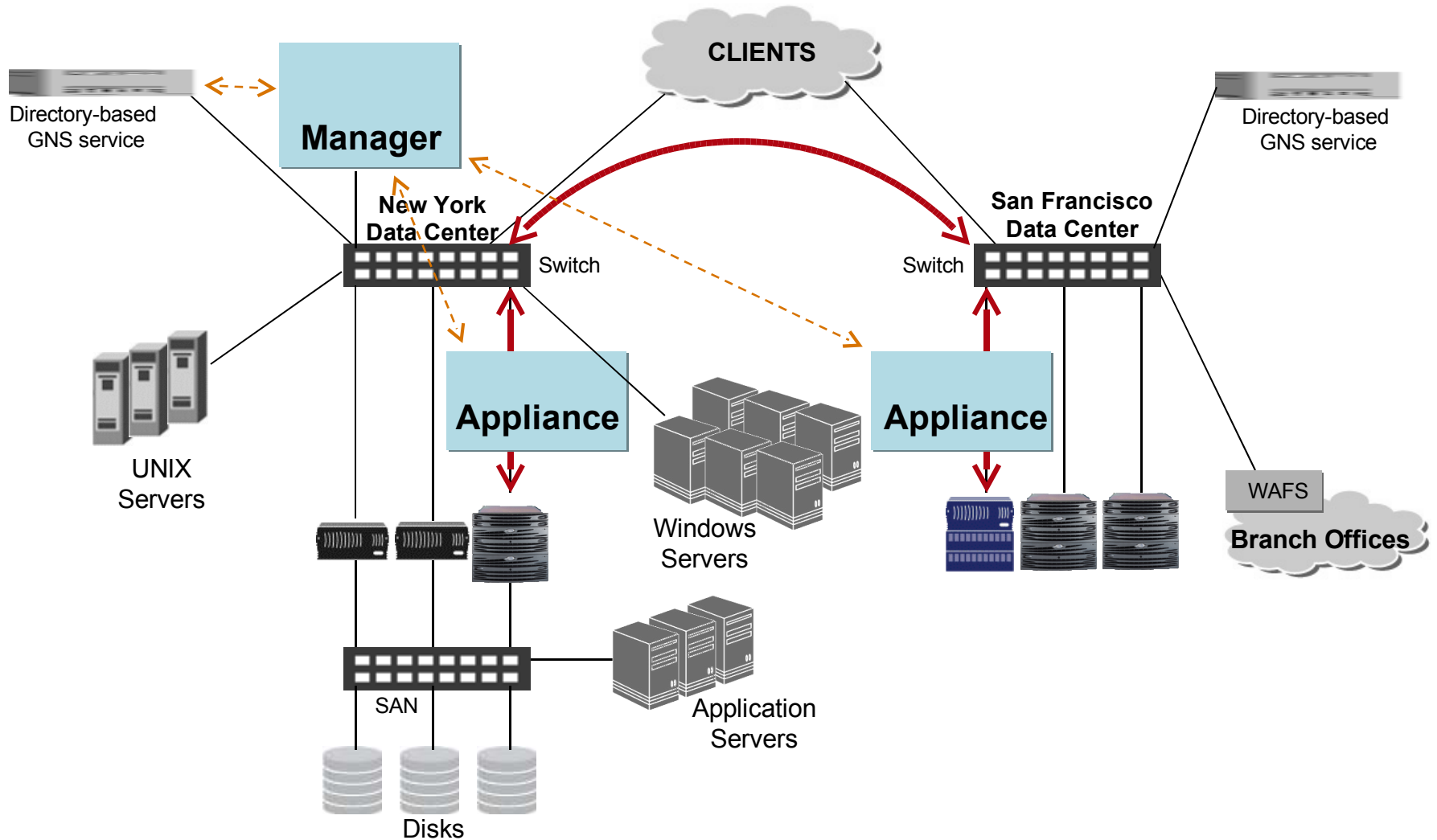
#### AUTOMATION

- Rich meta data
- Discovery
- Policy engine

# Global File Virtualization at Work



# Global Distributed Deployment



## Ongoing Standards Work

- **NFSv4 Global Namespace**

- NFSv4 referrals supports the feasibility of out-of-band namespace server
- Internet drafts in IETF NFSv4 Workgroup
  - C. Fan, D. Noveck, M. Wurzl, “NFSv4 Global Namespace Problem Statement”
  - D. Noveck, R. Burnett, “Next Steps for NFSv4 Migration/Replication”
- Referrals being implemented by client and server developers
  - UM CITI (Linux), Sun (Solaris), IBM, NetApp, EMC, etc.

- **pNFS**

- Superior architecture to address performance scaling requirement
- Ongoing work to include it in a NFSv4 minor version
- EMC participates in driving the standards effort

- **SNIA ILM TWG**

- Data and storage classification standards are key to policy-based automation
- EMC participates in driving the standards effort

## Global File Virtualization Summary

- Standards-based out-of-band global namespace with enhanced management
- Transparent in-band file system protocol processing for data in transition and in protection
- Integrated management and value-add applications address concrete customer pain points



**EMC<sup>2</sup>**

***Rainfinity***<sup>TM</sup>