

# OpenAFS

Improving Upon  
the Legacy of IBM AFS:  
A Development Road Map

# OpenAFS Governance

- Gatekeepers
  - Two commercial and two academic (both are CSG)
- Elders
  - Three commercial and three academic (two of three are CSG)
- Gatekeepers and Elders set long term direction, promote and fund raise
- Gatekeepers are responsible for code quality, documentation, and release management

# Major OpenAFS Improvements Since 2000

- Support for Kerberos 5 (Heimdal and MIT)
- Greater stability and performance on multi-processor systems
- Improved experience for mobile clients, NATs, and VPNs
- Large File Support
- Fast Restarts
  - --fast-restart option (manual salvage of volumes if error occurs)
  - On Demand Volume Mounting File Server and On-Line Salvager
- Audit Logging
- New platforms:
  - Microsoft Windows XP SP2, 2003 SP1, Vista, 64-bit too
  - Apple MacOS 10.3, 10.4 (Intel and PPC), (10.5 in development)
  - Solaris 9, 10 (Intel and Sparc), (11 in development)
  - Linux 2.6
  - Many more ...
- Operating System Vendors including Microsoft, Apple and Sun take OpenAFS seriously

# OpenAFS Case Studies

- Stanford Linear Accelerator Center
  - BaBar and Kavli Institute for Particle Astrophysics and Cosmology
    - >2 PB of data on tape, all tools and software in AFS
    - 5.5 TB of data in afs, 15 servers, up to 1.6TB per server, 3000 users, 18,000 volumes
  - Synchrotron Radiation Lab
    - Sequential and Parallel Gaussian Computations performed in AFS
    - Software tools in AFS
  - GEANT4 Software Toolkit
    - Distributed Development via AFS

# OpenAFS Case Studies

- ParalleIDatorCentrum, Center for Parallel Computers (KTH)
  - 5 clusters, 5 TB of data, 3000 users, 6000 volumes, 65 million files, 11 servers (1TB each), Home directories and Cluster Project storage
- [andrew.cmu.edu](http://andrew.cmu.edu)
  - 2.5TB, 29K users, ~400 peak conns/server, 46K volumes, 45M files, max of 132K files/volume
- Pictage Inc.
  - 200TB, >30,000 entries per directory, average file size 5MB, Photo editing on Windows and MacOS

# How Does OpenAFS Compare?

- Comparing AFS to solutions from Microsoft, Apple and Sun is like comparing Grapes and Oranges. They all provide basic file system but differ in implementation and capabilities
- Only AFS
  - has a caching client
    - Improves performance of clients
    - Reduces network traffic and the load on file servers
  - Has location independence and in-use volume moves
    - can load balance among servers while data is in use
    - can migrate to a new server platform without system outage
  - has working federated authentication and authorization
  - has working read-only replication and fail over

# What about Microsoft, Apple, or Sun?

- Sun NFSv4
  - A very broad and evolving file system protocol specification with few features implemented across the board by all vendors
  - Perhaps in three to five years there will be interoperable and heterogeneous support to match what AFS has today
- Microsoft Dfs
  - No location independence
  - Poorly handles failover and replication
  - Underlying file systems responsible for snapshots
  - CIFS protocol performs poorly on wide area networks
  - VPNs are required for WAN use. Vista adds TLS protection.
- Apple
  - Switching to CIFS from AFP. No Enterprise solution yet.
  - Has WebDAV but is limited by poor authentication support
    - AFS can be a back-end to WebDAV

# Development Projects Currently in Progress

- Provides Removing Reliance on DES with new security classes:
  - rxk5 (krb5) – developed at UMich; almost ready for testing
  - rxgk (GSS-API) – implementation meeting at KTH this month
  - Estimate: TBD
- Improving Performance of Namei file servers
  - Necessary for supporting AFS storage on JFS
  - Being fixed by safe removal of unnecessary fsync calls
  - Patch submitted to Gatekeepers – Testing in progress
- MacOS X Finder issues
  - To be fixed with Apple in a future release of MacOS X

# Development Projects in Search of Support

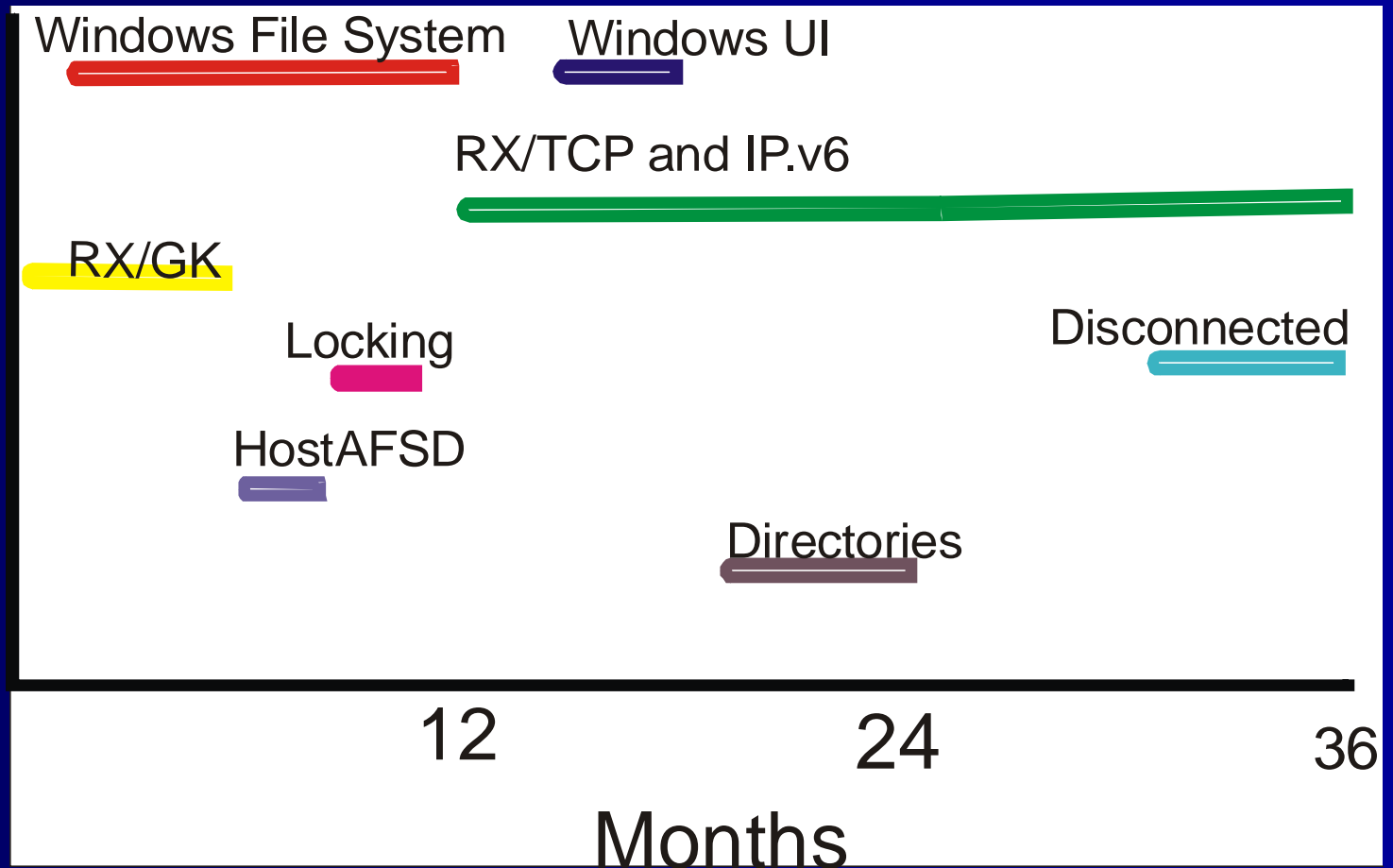
- Improving performance on 10Gbit networks and Implementing Support for IPv6 support
  - RX/TCP (NRL funded initial development)
  - Estimate: 20 to 25 developer months (\$500 - \$625K)
- Native Windows File System
  - W2K SP4, XP SP2, 2003 SP1, XP 64, 2003 64, and Vista
  - Replace SMB server w/ File System Redirector, Network Provider, and Kernel TDI module
  - Estimate: 10 to 11 developer months (\$250K - \$275K)
- Disconnected Operations
  - Read-only Estimate: 2 developer weeks (\$12K)
  - Read-write Estimate: 6 to 7 developer weeks (\$36K to \$48K)

# More Development Projects Seeking Support

- Removing Directory Limitations (poor search times, < 64,000 entries, UNICODE unaware, single stream per file)
  - New directory format and RPCs
  - Backward compatibility functionality for old clients
  - Estimate: 3 to 4 months (\$75K to \$100K)
- Mandatory Locking and Byte Range Locks
  - Estimate: 2 months (\$50K)
- Windows User Interface Improvements
  - New Explorer Shell Extensions, Control Panel and MMC plug-in
  - Estimate: 3 months (\$75K)
- HostAFSD – Local File System Exporter
  - Peer to Peer file sharing
  - Estimate: 2 months (\$50K)

# Three Year Road Map

## \$350K / year (block grant)



# Why Should CSG Help Fund OpenAFS?

- CSG already plays a significant governance role within OpenAFS
- 40% of CSG members deploy AFS today in some capacity. Many researchers use AFS from partner institutions as part of their collaboration.
- Some members are committed to AFS for at least 3 to 7 years
  - One round of major operating system and hardware revisions
- If the rest decides to migrate:
  - It will take years. OpenAFS may be able to address its weaknesses before the migration is complete.
  - Migrating to any new storage solution is hard because it takes a long time to phase out existing clients
  - Storage is complex. If the migration fails, being able to fall back on an improved OpenAFS will be desirable

# What if CSG Does Not Fund OpenAFS Development?

- OpenAFS will go on
  - CSG Members that are committed will provide what they can
  - The OpenAFS Elders will continue to solicit funding from other sources
- OpenAFS will take longer to achieve its potential
- Multi-year projects that require long term broad-based community support may never happen.
  - A native Windows file system is less pressing since there is already a client that works acceptably for most use cases
  - How many sites today have 10Gbit networks and wish to replicate or move tera byte volumes? What about five years from now?

# Q&A

See [http://www.secure-endpoints.com/private/csg\\_background.html](http://www.secure-endpoints.com/private/csg_background.html) for many more details  
Further questions can be addressed to [openafs-elders@openafs.org](mailto:openafs-elders@openafs.org)

# Related Projects

## Independent of OpenAFS

- Backup Systems
  - (but no two sites want the same thing)
- Authenticated read/write export via WebDAV, CIFS and NFSv4
- Kerberos 5 Extensions
- PTS LDAP back-end