

# The State of Samba (June 2011)

# SAMBA

**Jeremy Allison**  
**Samba Team/Google Open**  
**Source Programs Office**

**[jra@samba.org](mailto:jra@samba.org)**

**[jra@google.com](mailto:jra@google.com)**

# What is Samba ?



- Provides File/Print/Authentication services to Windows clients from a UNIX/Linux server.
- First Version of Samba written in 1992
  - Initially created by Andrew Tridgell in Australia (tridge).
- Been continuously developed (by mostly the same people) since then.

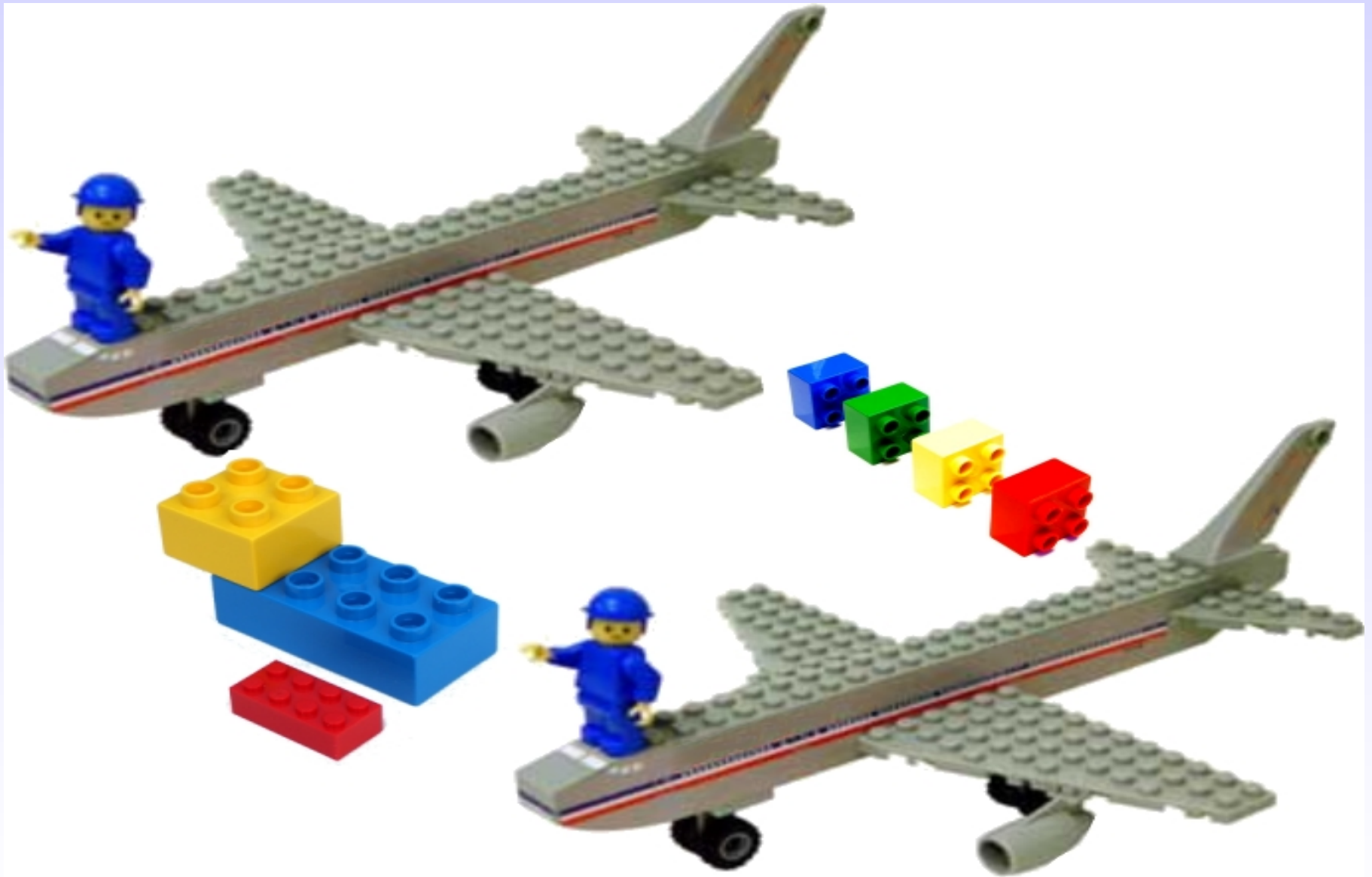
# What is Samba?

- Only fully-featured implementation of SMB/CIFS (now SMB2) outside of Microsoft.
  - Implements AD Domains, NT domains, printing, registry, performance counters, UNIX extensions and other things most other implementations don't care about :-).
- Open Source/Free Software under GPLv3 license.
- Large, old, codebase (getting larger and older).
  - Needs continuous refactoring.

# Refactoring in a nutshell..

SRMBR

Opening Windows to a  
Wider World



# Latest release 3.6.0

- Implements Microsoft's new protocol SMB2.
  - Early tests suggest 2x performance with Windows 7 clients.
- Internal code moved to asynchronous interfaces.
- Splitting up of internal functionality to allow separate daemons to implement services previously not replaceable.
- Re-written printing subsystem.

# What is SMB2 ?

- Microsoft's replacement for SMB/CIFS.
- Ships in Vista, Windows7 and Windows 2008.
- Considerably simpler than SMB/CIFS.
  - Only 19 requests vs. ~100 for CIFS/SMB.
- Much better support for asynchronous request/responses and multiple concurrent requests.
- Larger read/write sizes for more modern networks/machines.

# No longer a guessing game.

- Anti-trust cases in Europe and the US caused Microsoft to have to fully document all file sharing protocols.
  - We now have a library of complete documentation.
- Changes Samba development from network analysis to implementation of specifications.
  - Test driven development now essential.



# Samba spin-off technologies.

- Samba has had to invent several technologies that are very useful for other Free Software projects.
  - TDB (trivial database).
  - CTDB (clustered version of the above).
  - TALLOC (hierarchical memory allocator).
  - TEVENT (signal-safe asynchronous event library).
  - LDB – LDAP-like database library.



# TDB - the trivial database

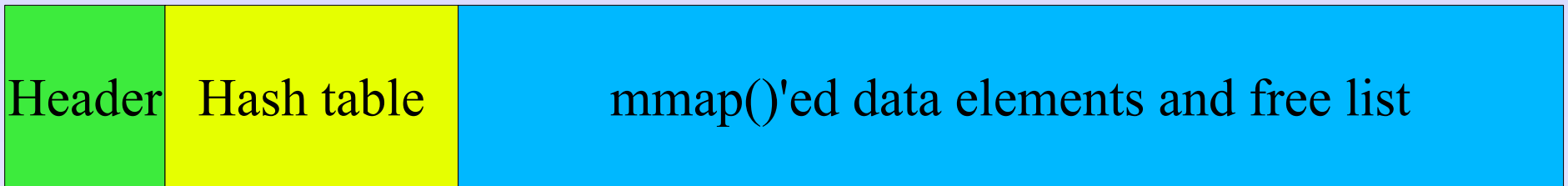
- Tdb is a very fast key/value pair database - similar to gdbm/dbm.
  - Can store blobs.
  - Multi-simultaneous readers/writers.
  - Supports transactions.
  - Core filestore for most Samba databases.
  - Available as a separate library.
- <http://tdb.samba.org>

# “Trivial” DataBase (tdb).

- Was so called because it once took less than 1000 lines of code.
  - Once transactions were added this was no longer the case.
  - mmap()'ed shared memory area arbitrated by fcntl() locks.
  - Very scalable for multi-readers/writers.
  - Dynamically expands as entries added (munmap(), ftruncate(), mmap()).
  - Depends on coherent buffer cache between pread/pwrite/mmap.

# TDB internals

Overview of a tdb file



Header contains: lock region, hash size, start of free list.

Hash table contains offset of first record in chain for each hash bucket.

Data area contains chained records (each containing their specific key, and a size, offset to the next record in the list, and the data).

Freelist contains size, and the offset to the next free area on the freelist.

# CTDB - clustered storage

- Clustered version of tdb.
  - Based on top of tdb.
  - Has high-availability features.
    - Needs shared filesystem to provide HA features - but doesn't use this filesystem for storage.
  - Contains a distributed lock manager that migrates data to nodes frequently using it.
- <http://ctdb.samba.org>

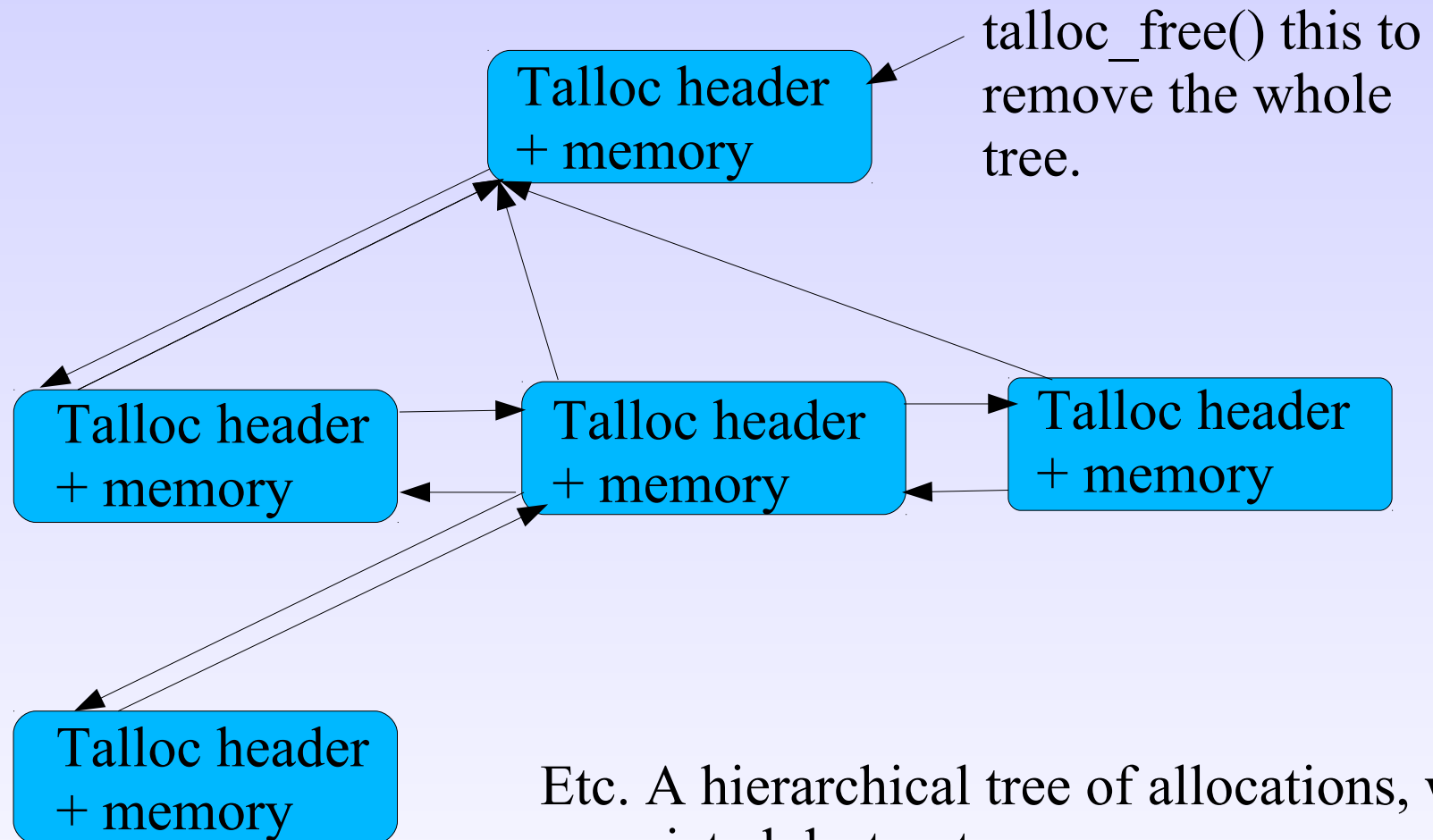
# Clustered Samba - Distributed Storage

- Samba in conjunction with ctdb and a distributed filesystem creates a highly-available, multi-node scalable network attached storage system.
  - IBM has productized with SONAS and their (proprietary) GPFS.
  - Free Software clustered filesystems:
    - Gluster.
    - GFS2.
    - OCFS.
    - CEPH.
    - Lustre.

# Talloc - C without guns

- Talloc (trivial allocator) adds hierarchical memory allocation to C programming.
- Allows destructor functions to be attached to allocated memory to deallocate resources.
  - Very much like C++ destructors.
- Allows very efficient processing of request/response protocols.
- <http://talloc.samba.org>

# talloc()



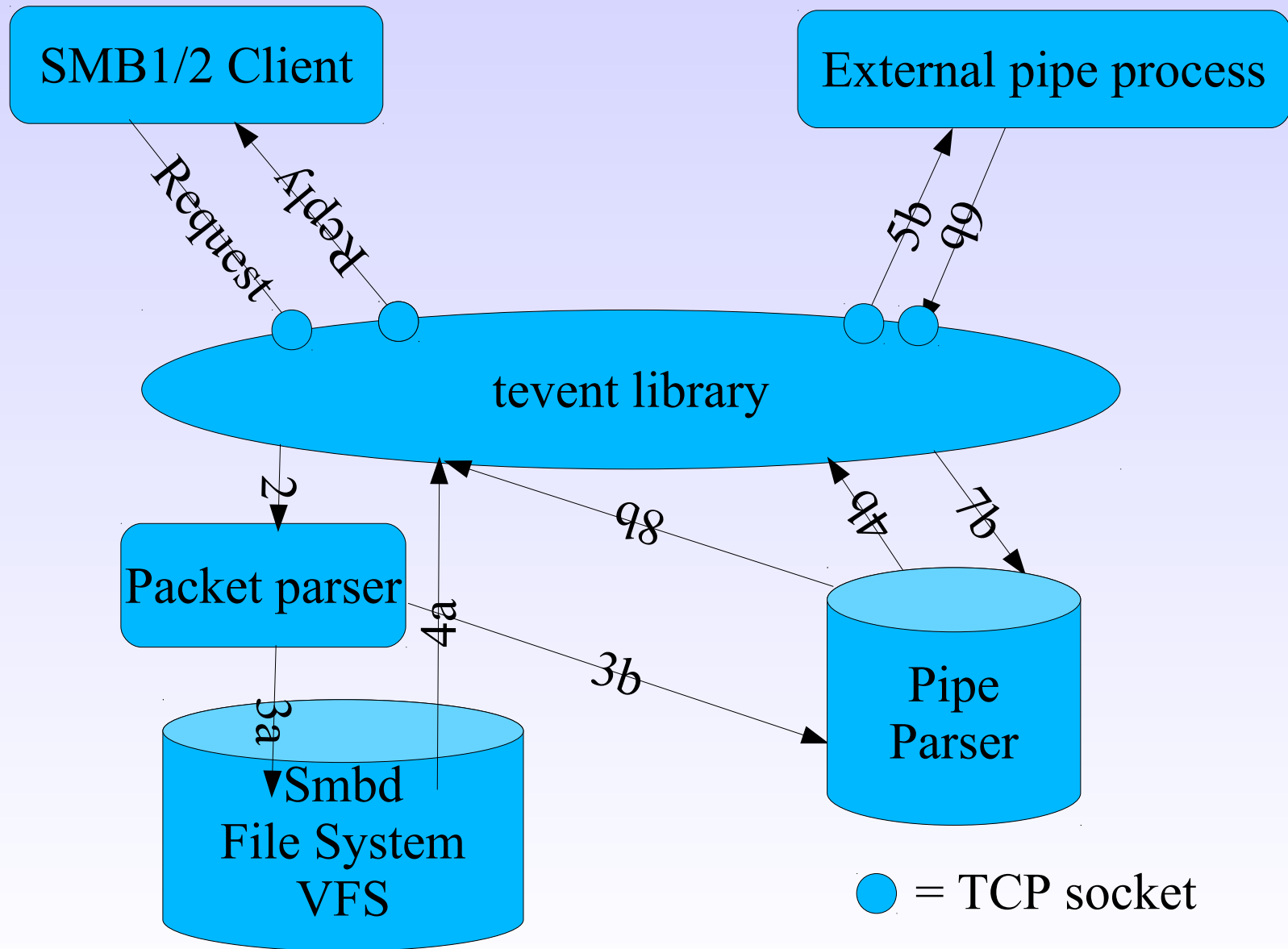
Etc. A hierarchical tree of allocations, with associated destructors.

# Tevent - Event library

- Similar to libevent/libev
  - In our defense, at the time tevent was created libevent did not handle POSIX realtime signals at all.
  - tevent heavily uses talloc(), making it much easier to integrate into the Samba event loop.
- Set all sockets into asynchronous mode, then just get callbacks when data is ready to read/write.
- <http://tevent.samba.org>



# The central role of tevent



# Ldb - LDAP made easy

- Ldb is an “LDAP-like” database built on top of local tdb files.
- Allows more complex LDAP-style queries than tdb simple key -> value lookup.
- Allows multiple indexes built on top of tdb records.
- Data imported/exported in LDIF format.
- Core database engine in Samba4.
- <http://ldb.samba.org>

# Samba 4



- The next major Samba release will be Samba4.
  - Difficult birth, conflicts between developers.
  - Finally the entire Samba Team has a ship plan and all are working towards this goal.
  - Merging of Samba3 fileserver, printserver and remote authentication (winbind) ongoing into Samba4 Active Directory code.

# Samba 4 - why so late ?

- Active Directory is way beyond the scope of file/print/authentication. It's bigger than anything we've done before.
- Active Directory includes:
  - DCE/RPC services.
  - LDAP server.
  - Kerberos server.
  - DNS server.
  - Requires integrated backend database for all the above.

# Samba 4 - getting to a release

- Full Active Directory Support, only big missing feature is trusted Domains.
  - Will probably ship without these for a first release.
  - Read-only Domains are implemented, and are much more important to get Windows network administrators used to a “second source” Active Directory.
  - Full replication is working with Microsoft Active Directory servers.
  - Several sites using this in production.

# Samba 4 - something for everyone

- New build system (waf) makes binaries much smaller.
- Will allow users to pick the parts of Samba they need.
  - Directory services, file and print.
  - Just file services (NAS vendors).
  - Just remote authentication (member services in a domain).
  - Libraries of Samba sub-components.

# The “Cloud” storage threat..



- File sharing on a LAN is obsolete. All your data belongs to the cloud..

# Is Samba obsolete ?



- File sharing on a LAN is now only for “legacy” applications.
- All modern applications are 3-tier and usually web based.
- However, the back-end servers need storage management, and NAS file storage is much easier to manage than SAN block storage.
- With SMB2, Microsoft is challenging NFSv4 in the data center.

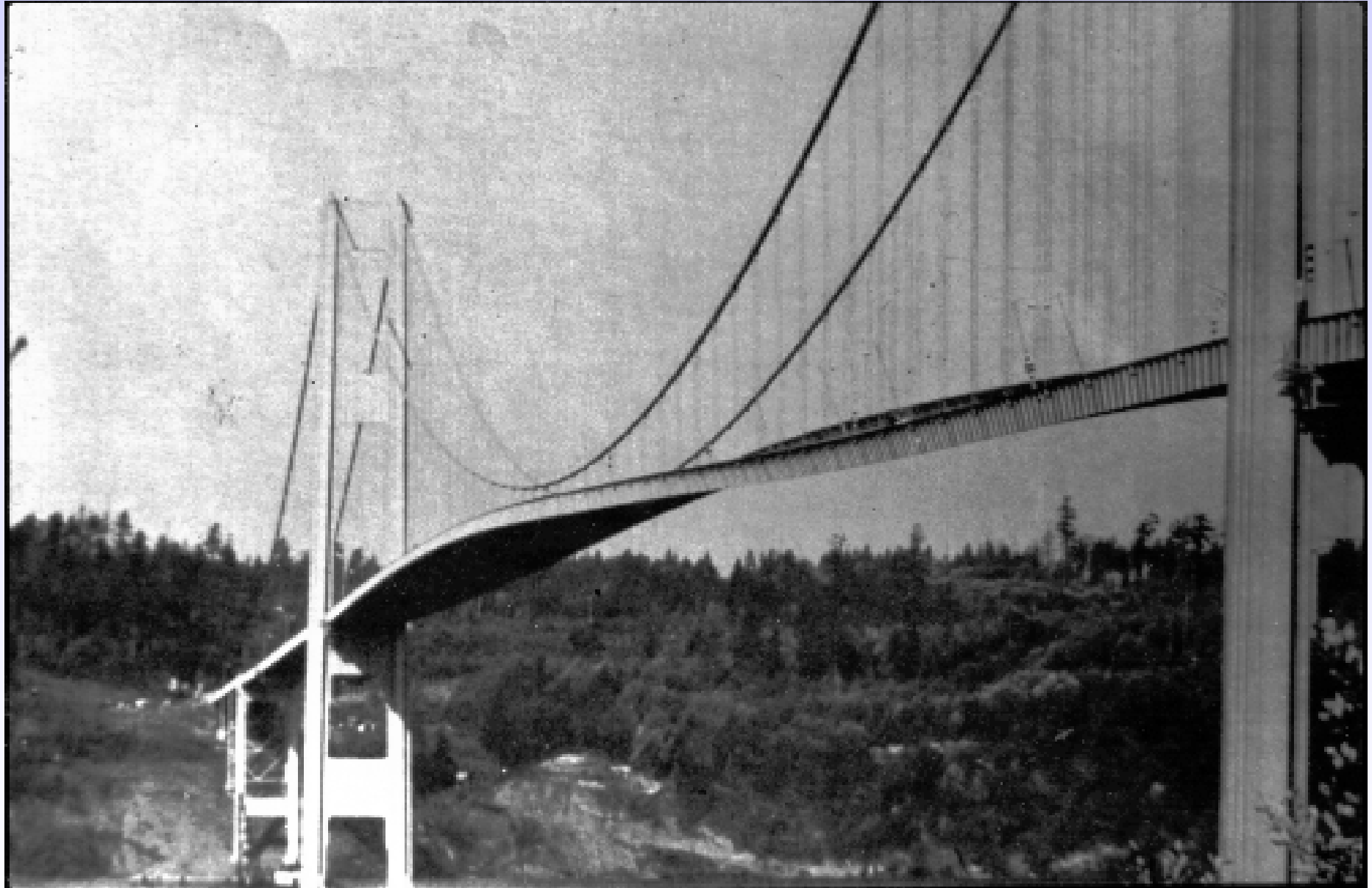


# Then there's local storage..



- Until everyone in the world has a high-bandwidth Internet connection, many people have large amounts of personal storage.
  - Cloud storage and streaming doesn't help much here.
  - NAS protocols like SMB/SMB2 are ideal here, most clients already support them.
    - Many media devices already have SMB built in.

# Can Samba bridge cloud and local storage ?



SAMBA  
Opening Windows to a  
Wider World

# Help keep Samba successful

- Everyone is welcome in the Samba ecosystem.
  - Companies who commercialize Samba help popularize it to millions of users.
  - NAS vendors and IT installations are especially welcome.
  - GPLv3 is a better license for commercial re-use of Free Software than GPLv2.
  - Programmers needed, but users can help with bug reports.
    - <http://bugzilla.samba.org>

# Questions and Comments ?

Email: [jra@samba.org](mailto:jra@samba.org)  
[jra@google.com](mailto:jra@google.com)