The Death of File Sharing Protocols

Jeremy Allison
Samba Team/Google Open Source Programs Office
jra@samba.org
jra@google.com

Opening Windows to a Wider World

Not *that* file sharing..

This file sharing

~~AT&T - RFS~~

Sun - NFS

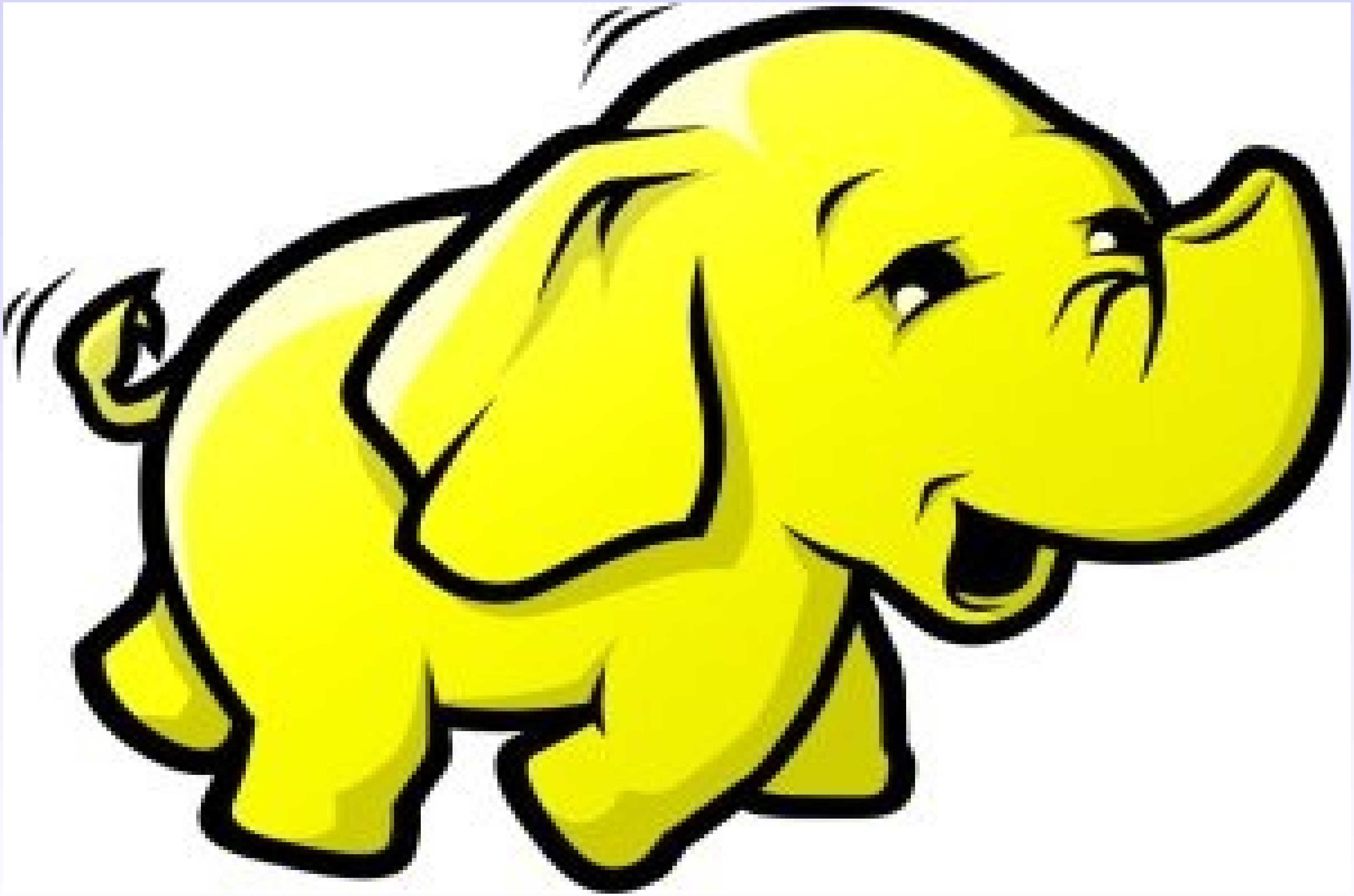
IBM – SMB

Microsoft SMB2

Remote file systems that attempt to completely emulate all the semantics of a local file system is an idea whose time has come..

..And gone.

The high end is pushing down



Microsoft/Yahoo are using Hadoop.

Google changed from using MapReduce to “Colossus”.

Amazon has S3.

High end users create application specific file systems for their specific needs.

Clustered file systems create their own protocols either proprietary or Open Source.

No widely adopted standards.

Opening Windows to a Wider World

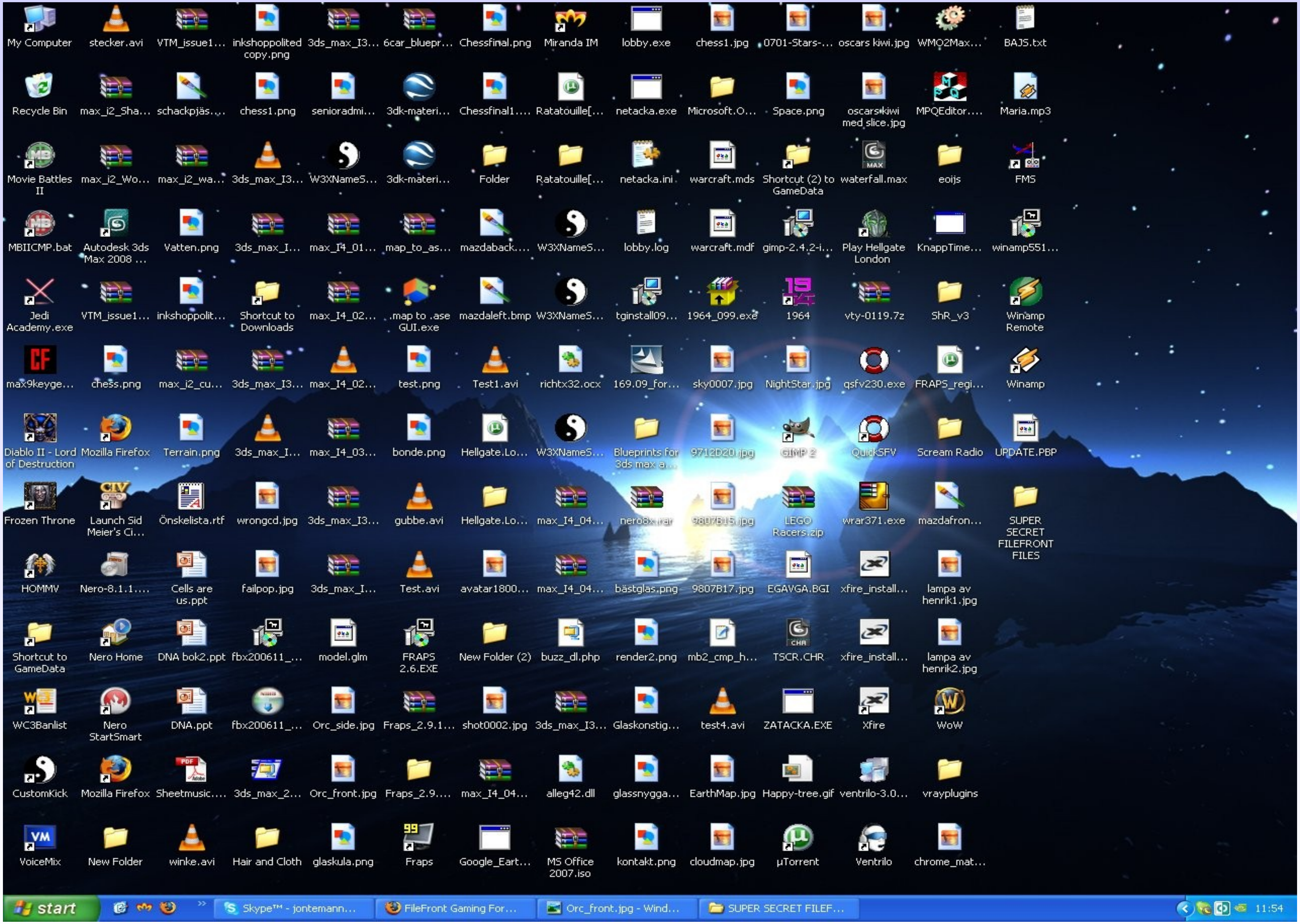
File systems like Lustre, GFS, GPFS, Glusterfs, OCFS2, live in back end data centers. They don't need to provide full local filesystem semantics.

Access to cloud-based storage over the Internet allows genius-level file system engineering to be available to people who don't understand what a file is.

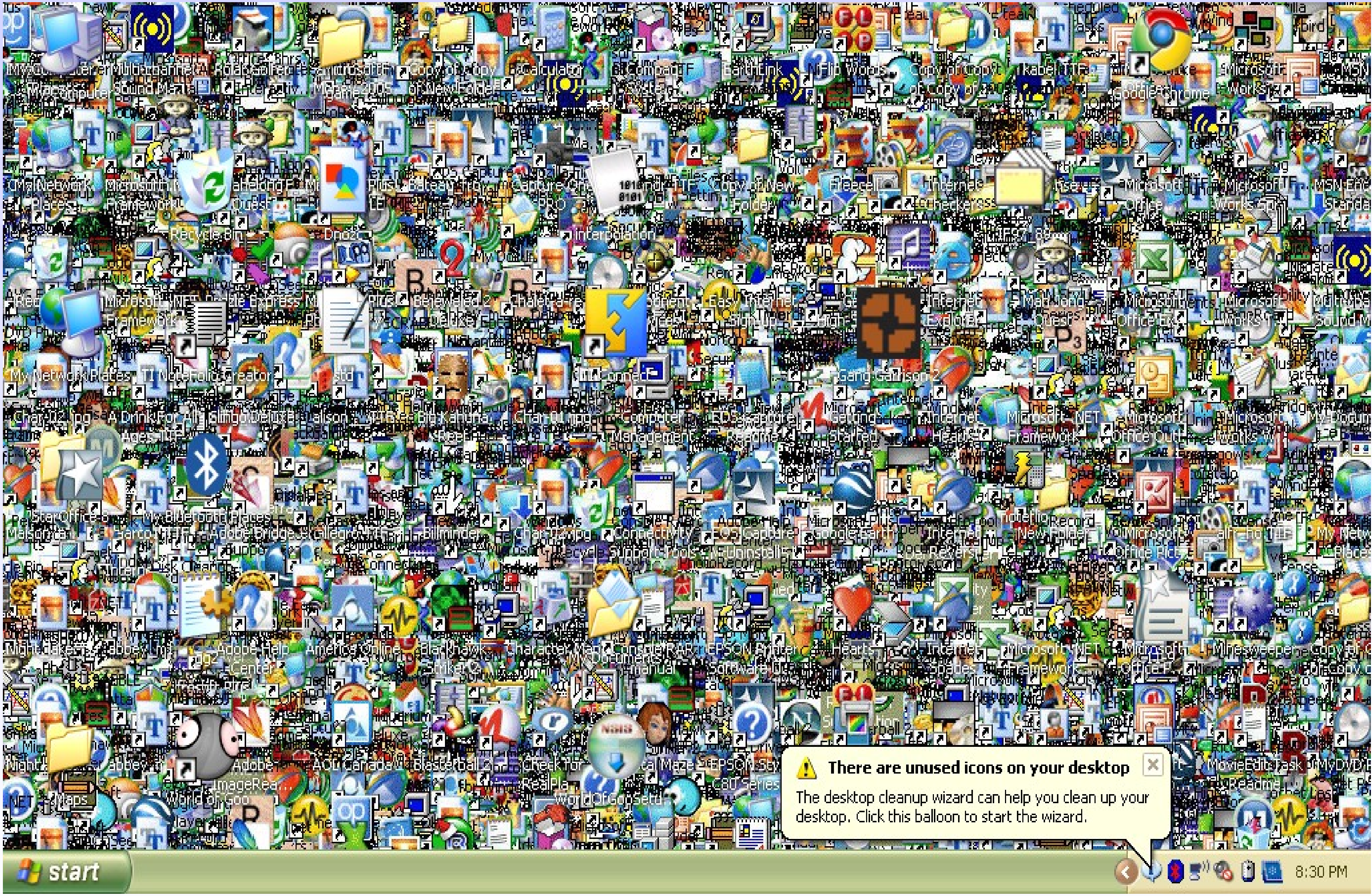
Opening Windows to a Wider World

The low end is pushing up

Non-geek desktops look like this



Or eventually even like this



There are unused icons on your desktop

The desktop cleanup wizard can help you clean up your desktop. Click this balloon to start the wizard.

Opening Windows to a Wider World

No one can find anything on *local* multi-terabye drives.

Adding a remote file system makes it worse.

Normal users fundamentally don't understand files and directories.

Modern applications have to incorporate a database in order to allow users to find material.

Local or remote file systems simply become a cache for recently accessed content.

Modern distributed applications that depend on local file system locking behavior are considered broken.

What kind of applications do people now use to collaborate ?

Microsoft
Office SharePo
Server 2007
Linked in
facebook
oogle
BETA
Docs &
Spreadsheets
Google wave

But what about my home media server ?

Media Servers (should) provide streaming services, not a generic file system view.

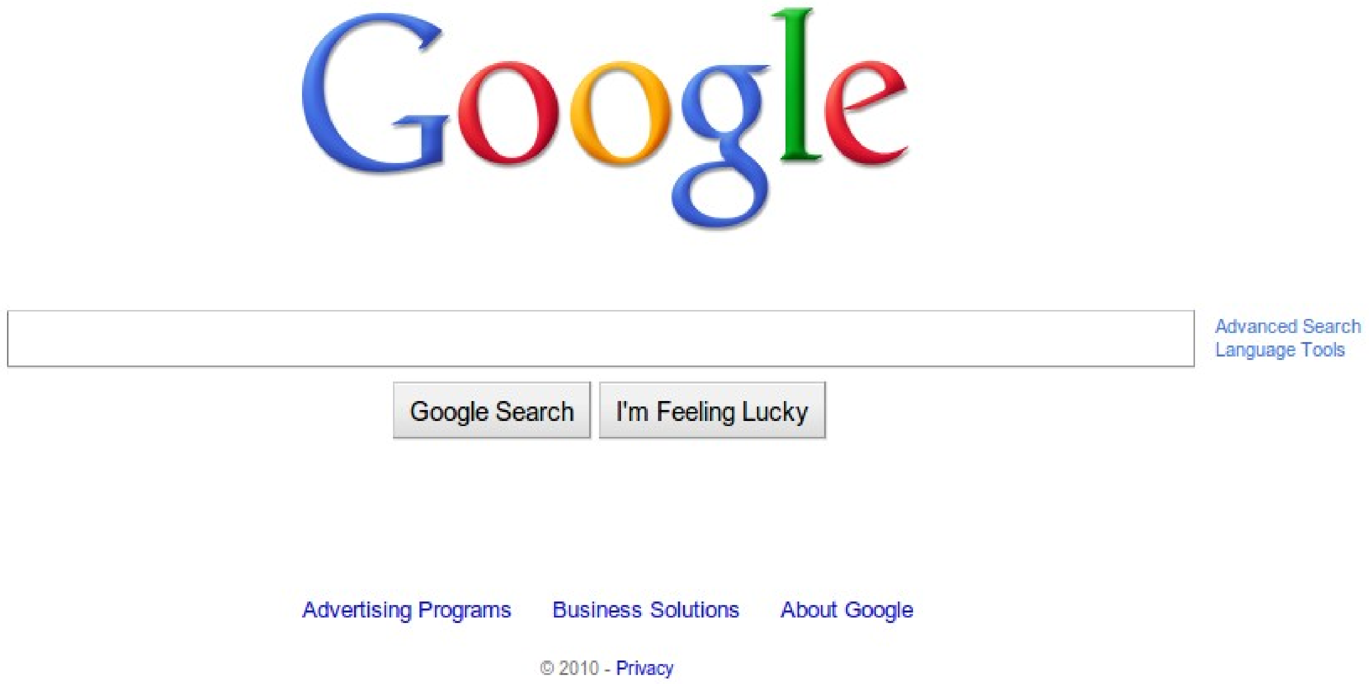
“*Eventually, all media streaming problems become file transfer protocol*”: Vint Cerf.

Televisions and media servers shouldn't present file/directory based views.

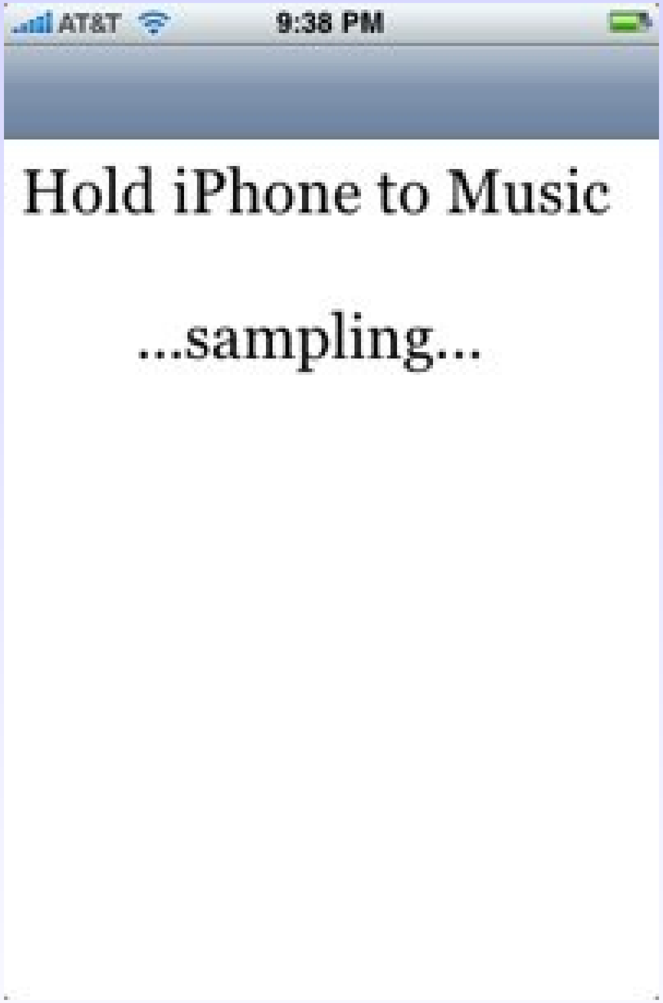
They don't need share modes, ACLs or file locking.

They need good search capabilities.

So what does a good search interface look like ?

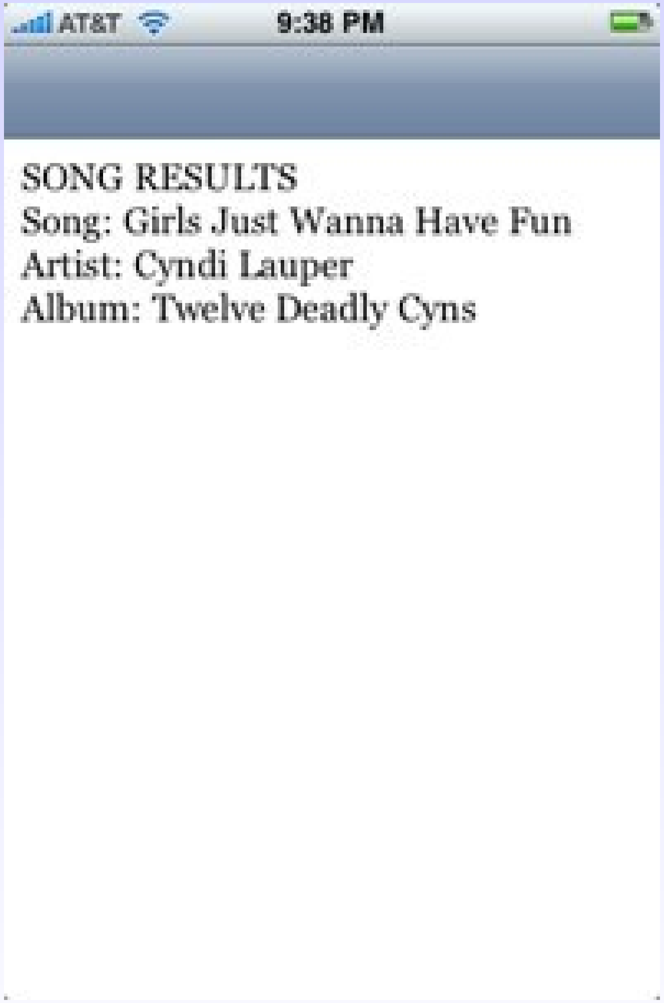


Or this..

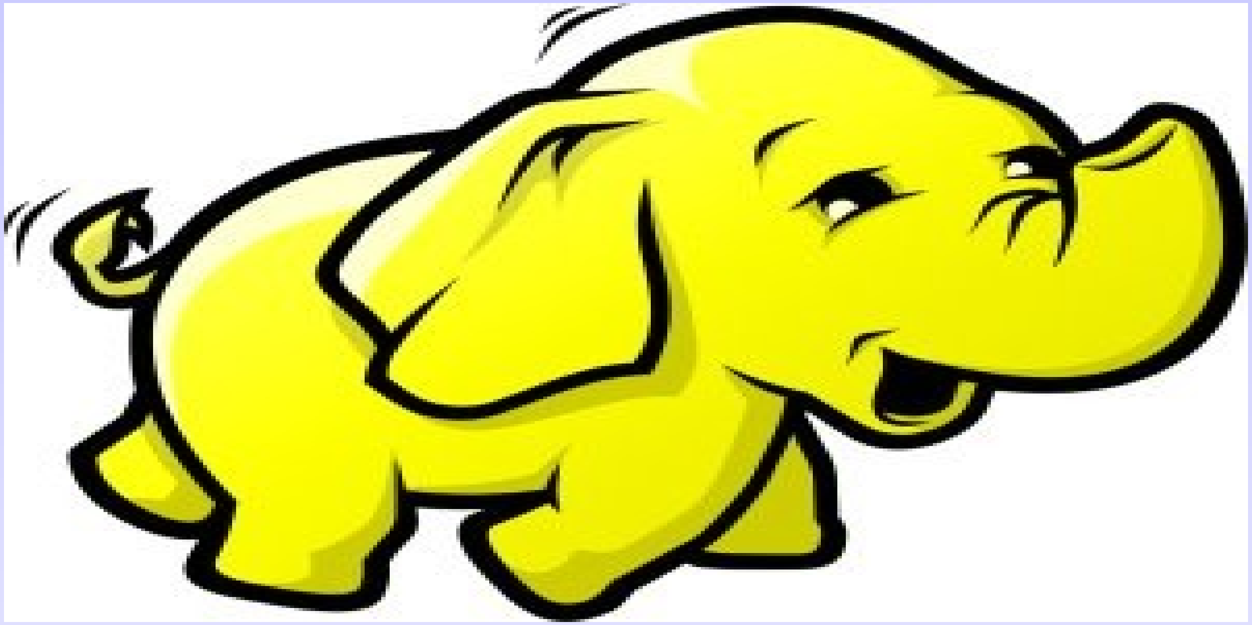


Hold iPhone to Music

...sampling...



SONG RESULTS
Song: Girls Just Wanna Have Fun
Artist: Cyndi Lauper
Album: Twelve Deadly Cyns



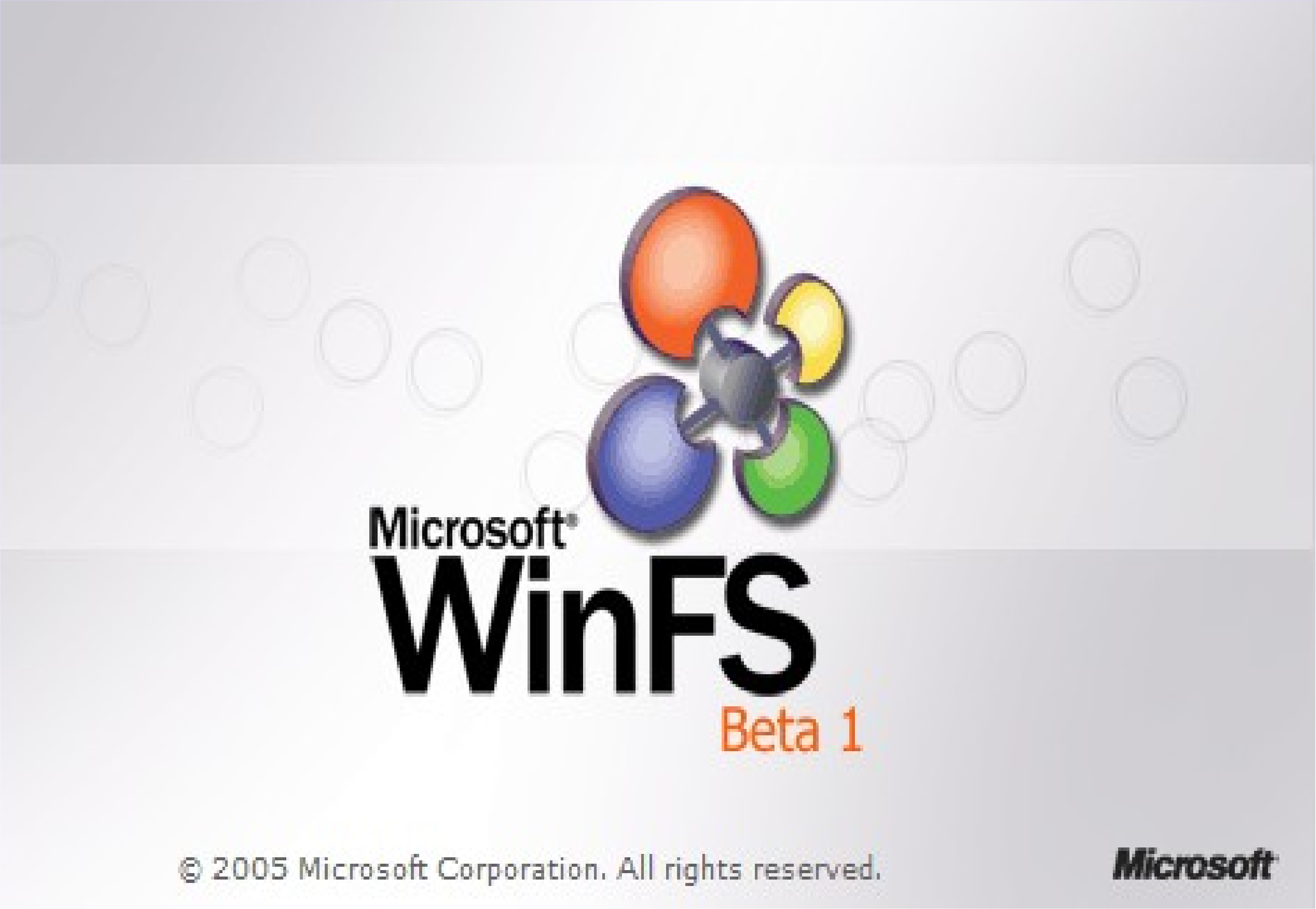
CIFS/NFS/SMB2 File sharing protocols.

How to survive the crunch ?

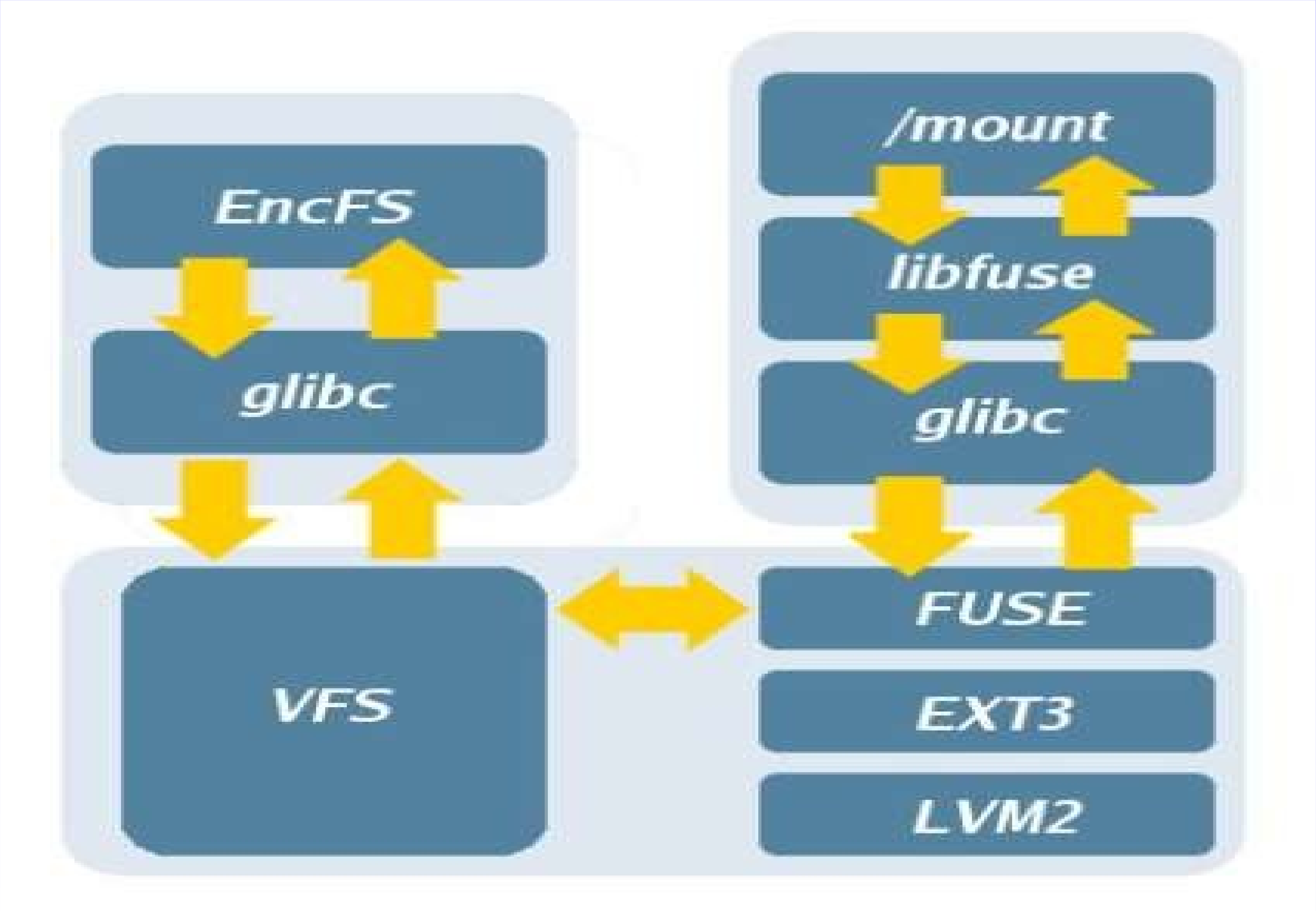
- We can keep hiding in the data center, providing high availability to applications that need local file system semantics.
  - There will be fewer and fewer of these in the future.
  - It's a living – but not a thriving one.
    - Like becoming a backup vendor..

What if we embraced search as a core part of file sharing protocols ?

But didn't someone already try this ?



And FUSE in the Open Source world..



What might this look like in a file sharing protocol ?

- opendir(“pathname”,....);
  - or
- FindFirstFile(“pathname”, ...);
- Might become:
- opendir(&list_of_search_attributes,...);
  - Or
- FindFirstFile(&list_of_search_attributes,...);

Would need a change in the wire protocol

- This is easy in SMB2 as the “open” call already takes attributes.
- NFSv4 would need extensions for this.

Might be easier to do at mount time

- Specify a list of attributes to match when creating the mount point/share.
- Could even be encoded in the share/mount point name ?
- Any normal opendir() then returns a list of paths matching the requested attributes. A “virtual” directory.

Is it worth it ?

- Should we leave all this to individual application developers ?
  - Windows explorer/Gnome tracker will do most of this based on top of existing file systems.
  - Still leaves the “what is a file” problem for normal users.
    - Save on the desktop and let “tracker” sort it out ?

There is a product hiding here somewhere. One that might be worth building..

Hopefully this made everyone think: Let's do a poll.

- 1). Do you agree there is a problem with the protocols we're currently creating and using ?
- 2). Is the problem I described the real issue ?
- 3). Does anyone agree with my conclusions ?

Questions and Comments ?

Email: jra@samba.org
jra@google.com