in the Age of Microsoft
Glasnost

Christopher R. Hertel
Storage Architect, CIFS Geek
Founder and CTO

TCCC 2011
Octoberly, 2011
Copyright 2011, ubiqx Consulting, Inc.
Introductions
Introducing: Me

Your Friendly Neighborhood CIFS Geek

- **Samba Team** member (since 1998-ish)
- **jCIFS** Project co-founder
- **CIFS Author** *(shameless plug)*
- Network Storage Geek
- Incurable Idealist
- Etc., etc., ad nauseum

A ruminant mammal (Geekus geekus) with long legs, humped shoulders, and broadly palmated antlers.
Introducing: You

Code Monkeys!

This is Code Camp, after all.
Introducing: The Samba Team

Members of the Samba Team gather at the 10th annual Samba eXPerience conference in Göttingen, Germany.
SMB = Server Message Block protocol
A stateful network file system protocol originally created by IBM in the early 1980s for use with the PC-DOS operating system.

CIFS = Common Internet File System
A “marketing upgrade” to SMB. This new name for SMB was coined in the mid 1990's. The term “CIFS” is now often used as a name for the complete suite of protocols that include and provide support for SMB. Often written “SMB/CIFS”.

SMB2 = Server Message Block protocol version 2
A complete rewrite of the SMB protocol, introduced with Windows Vista. SMB2 reduces the top-level command set from 75 commands to 19.
Introducing: More About Me

Java SMB/CIFS Client Toolkit

The very best developer's guide to SMB/CIFS.

The very worst developer's guide to SMB/CIFS.

The very only developer's guide to SMB/CIFS.

Some Trouble I've Caused...
Open Source Credentials Notwithstanding...

*Microsoft* asked a member of the *Samba Team* to document SMB/CIFS!
Thus, SMB/CIFS is covered in two documents:

[MS-CIFS]
- Provides the base specification of the “NT LM 0.12” dialect.
- A “snapshot in time”.
- Most of this stuff is still there in current Windows versions. Really.

[MS-SMB]
- “Extends” [MS-CIFS].
- Documents changes made to SMB starting in W2K.
- Still the same “NT LM 0.12” dialect.

Note: The naming is backwards!
Where are we going?

- SMB/CIFS is Dead
  - Long live SMB2?

- The Entourage
  - 400 documents, give or take

- BITS Upload Protocol
  - An easy start

- PeerDist Protocol
  - Three-part harmony

- SMB2.2
  - The future’s so bright, I gotta wear shades
CIFS is Dead
The current dialect is “NT LM 0.12”. It was introduced with Windows NT.
The current dialect is “NT LM 0.12”. It was introduced with Windows NT.

New dialects of SMB/CIFS are unlikely.
CIFS is Dead

- The current dialect is “NT LM 0.12”. It was introduced with Windows NT.
- New dialects of SMB/CIFS are unlikely.
- Microsoft is focused on SMB2.
SMB/CIFS is still the most widely used network file system on the planet, by a wide margin.

Supported in all versions of MS-Windows.

Supported by 3rd-party NAS Vendors.
SMB/CIFS code examples:

- Samba, of course
- jCIFS – Smaller, simpler, client-only
- Sample code in *Implementing CIFS*
- See also: [http://www.ubiqx.org/libcifs/](http://www.ubiqx.org/libcifs/)
CIFS is Undead and Living at:

CIFS.ORG
The Docs
The Docs

Go here:
http://www.microsoft.com/openspecifications/

Over 400 documents have been published, covering:
- Authentication
- Windows Internals
- File Formats
- Client-Server Protocols
- Server-Server Protocols

Overview docs provide starting points for understanding groups of docs.
“We should implement them all.”

- There is an opportunity here to leverage both the technology and the installed base.
- The documentation includes previews of SMB2.2, and other features of Windows 8.
- This will feed the software engineering ecosystem for years.
The PFIF provides subcontractors with access to Microsoft documentation under the terms of the Work Group Server Protocol Program (WSPP).

See the PFIF website for more information.

See also the “Open Specification Promise”.
Okay...

...so what can you *do* with all of this?
“BITS is Earth’s most widely used file transfer service, with more than 600 million unique users across the planet.”
– Vipul Bansal, Microsoft WMI Blog, Jan 2009.
“BITS is Earth’s most widely used file transfer service, with more than 600 million unique users across the planet.”

– Vipul Bansal, Microsoft WMI Blog, Jan 2009.

Note Well:
“BITS is Earth’s most widely used file transfer service, with more than 600 million unique users across the planet.”
– Vipul Bansal, Microsoft WMI Blog, Jan 2009.

Note Well: nobody cares.
BITS: Background “Intelligent” Transfer Service

“BITS is Earth’s most widely used file transfer service, with more than 600 million unique users across the planet.”

– Vipul Bansal, Microsoft WMI Blog, Jan 2009.

What does that mean anyway?

- It does not say “protocol”, it says “file transfer service”.
- BITS is the Windows system service used by Windows Update to download patches.
- Most users don't even know it's there.
BITS: Background “Intelligent” Transfer Service

BITS Features

- Built into Windows
- Restartable Transfers
  - ...but only linearly; does not “patch”.
- Both Download and Upload
  - ...and “Upload Reply”.
- Job priority levels
- Senses network traffic to manage impact
**BITS**: Background “Intelligent” Transfer Service

**BITS Download Jobs**

- The overwhelming majority of BITS jobs are probably Windows Update downloads.
- BITS Downloads use HTTP/HTTPS.
- Sort of like uucp?
  - `wget + batch + nice + diffserv`?

The “special sauce” is the use of network traffic monitoring to limit BITS data transfer rates.
BITS: Background “Intelligent” Transfer Service

BITS Upload Jobs

- Much less common.
- Proprietary extensions to HTTP/HTTPS.
- Only between Windows BITS clients and Windows HTTP[S] servers.
BITS: Background “Intelligent” Transfer Service

BITS Upload Jobs

- Much less common.
- Proprietary extensions to HTTP/HTTPS.
- Only between Windows BITS clients and Windows HTTP[S] servers – Until now!

STiB
**BITS:** Background “Intelligent” Transfer Service

**STiB means:**
- Slow Transfer in Background?
- Silly Technology is Boring?
- Sipping Tea in Belgium?
- BITS spelled sdrawkcab with a small ‘i’?

**STiB: It Is what It Is.**
- ...a toolkit for testing BITS Uploads.
- ...example code for others to read / use.

A CGI script could be written to accept BITS Uploads.
BITS: Background “Intelligent” Transfer Service

BITS Upload Extensions:

- HTTP Extension Method: BITS_POST
- BITS Packet Types
  - Ping
  - Create-Session
  - Fragment
  - Cancel-Session
  - Close-Session
  - Ack

BITS Documentation:

- MSDN: BITS Upload Protocol
- WSPP: [MC-BUP]
BITS: Background “Intelligent” Transfer Service

BITS Upload Extensions:
- HTTP Extension Method: BITS_POST
- BITS Packet Types
  - Ping
  - Create-Session
  - Fragment
  - Cancel-Session
  - Close-Session
  - Ack

BITS Documentation:
- MSDN: BITS Upload Protocol
- WSPP: [MC-BUP]
BITS: Background “Intelligent” Transfer Service

Do we care?

YAWP (Yet Another Windows Protocol)

 BITS Upload is supported in IIS,
   ✷ and in Microsoft's “lightweight” HTTP server.
 BITS It's convenient when working with Windows,
   ✷ Not nearly as powerful as, eg., rsync.
   ✷ Not as secure as sftp, scp, or sshfs.
Do we care?

YAWP (Yet Another Windows Protocol)

BITS Upload is supported in IIS,
and in Microsoft's “lightweight” HTTP server.

It's convenient when working with Windows.
Not nearly as powerful as, eg., rsync.
Not as secure as sftp, scp, or sshfs.

MS-BITS, however, also supports BranchCache™, which suggests some very useful testing scenarios.

GET support added to STiB,
PeerDist included in the header,
It works!

TCCC Oct 2011
BITS: Background “Intelligent” Transfer Service

http://www.ubiqx.org/proj/STiB/

STiB is at version 0.2

It contains stibtest, which can:

- Send files using MS-BUP protocol,
- Get files using HTTP1.1,
- Get a subrange of a requested file,
- Specify “peerdist” encoding when requesting all or part of a file.

Please download and test it.
Send patches.
“Someone” should write an Apache Module that to handle BITS Upload Protocol.

http://www.ubiqx.org/proj/STiB/
Pay Attention!

This is where it gets interesting.
What the heck is Prequel?
**Prequel**

**Prequel**: A project to build an Open Source Implementation of Microsoft's BranchCache™.

So what the heck is BranchCache™?
Prequel

Prequel: A project to build an Open Source Implementation of Microsoft's BranchCache™.

BranchCache™ is a distributed content caching system
- supported in W2K8r2 servers,
- and Windows7 clients.

Cheap, effective WAN acceleration for SMB2, HTTP, and BITS.
Prequel

BranchCache Architecture
A quick overview

Content Servers
Have content to share with multiple clients.

Clients (peers)
Request & receive content from content servers.

The Cache
A copy of the original content, divided into segments and blocks, accessed via hash tags.

TCCC Oct 2011
Content Servers:

- Web Servers (HTTP, BITS)
- File Servers (SMB2.1)

The client must know to ask for content tags instead of actual content.

- If the tags are already calculated, they are returned by the BranchCache™-enabled server.
- Otherwise the content is returned, and the server (W2K8r2) calculates the tags for next time.
This is IE 8 indicating support for BranchCache™ by listing "peerdist" as an acceptable encoding.

Accept-Encoding: gzip, deflate, peerdist
Client-side PeerDist Caching

There are two modes of operation:

- Distributed Mode
- Hosted Mode
Distributed Mode
Distributed Mode

- Each client keeps a local cache.
- A client requests PeerDist tags from the server, then broadcasts to find the cached content.
- If the content is not cached,
  - The client requests the content from the content server,
  - The client stores both content and tags in its own cache.

Reminiscent of the CIFS Browse Service.
Hosted Mode

Content Server

WAN Link/VPN

Cache
Hosted Mode

- A client request tags from the content server.
- The client then asks the local cache server for the content.
- If the content is not cached, the client requests content from the content server.
- The client sends both content and tags to the cache server.
- Content can now be retrieved from the cache server using only tags.

TCCC Oct 2011
Content Tags

Blocks
- Are a unit of download
  (from either the content server or cache server)
- Are 64K
  (or less, for the last block in a file only)

The block tag is an SHA{256, 384, 512} hash of the block.

Segments
- Are a unit of discovery
- One segment is 32M ≈ 512 blocks
  (or less, if the last block is short)

Segments are identified by a hash of the block hashes.
Prequel does not have a release number yet.

**pq_cgi** – CGI program to generate PeerDist Content Information.
- Tested with Apache.

**pdDump** – Pretty-print Content Information.

**_*key_dx** – Extract W2K8r2 Server Passphrase and Server Secret
Prequel Dæmon

Conceptual Overview
SMB2.2

The Future's So Bright, I Gotta Wear Shades
SMB2.2 is a Game-Changer

- Multipath
- Cluster capabilities
- Distributed caching
- SMB over RDMA

SMB2.2 is aimed at supporting Application Servers in Enterprise Datacenters.

It presents a strong challenge to NFSv4.x.
Stocking the SMB/CIFS Pond

- The SMB/CIFS/SMB2 Talent Pool is very shallow.
- The SMB/CIFS/SMB2-talented are big fish.
- The number of organizations in the SNIA CIFS Plugfest has grown from 14 to 27 in 4 years.
- SMB2.2 is likely to generate a lot of work.
Whither Samba?

Now that Microsoft is sharing, do we need the Rebel Alliance?

When will we see Samba 4?

What have the Samba Team been doing for the past few years?

- Directory Services
- Python Bindings
- CTDB Clustering
- Restructure the Code
Dive in!
The End