Google™

09

Monday, June 1, 2009


Google™ 09 I/O

Mercurial On Bigtable

Jacob Lee
5/28/2009

Time warp: July 2006

- 2 years since Subversion 1.0
- Monotone was 3 years old
- Git and Hg: 1 year old
- CVS was popular
- Sourceforge had just announced Subversion support

CC BY-ND http://www.flickr.com/photos/leeontheroad/89666692/

The Modern Project Hosting Ecosystem

- Google Code
  - 200,000 projects
  - Several million visitors/day
- Sourceforge
  - Shell accounts
  - DVCS
  - Trac
- GitHub, Bitbucket
  - DVCS
  - Social experience

Project Hosting Mercurial Support

Committed Changes **Branch:** default ‹ Newer May 11 - Apr 27 Older ›

Rev	Scores	Commit log message	Date	Author
bf1db50078		core: merge core-inner-animation	May 11, 2009	Mathieu Virbel <txprog>
c2edf692f4		mtwidget: add inner animation at init time	May 11, 2009	Mathieu Virbel <txprog>
b8e39e4c00		mtwidget: fix animation with list and tuple	May 11, 2009	Mathieu Virbel <txprog>
f545d57cf5		merge tip	May 11, 2009	Mathieu Virbel <txprog>
adb1374b2f		mtwidget: rework inner animation to add enable/disable/update+inner_animation()	May 11, 2009	Mathieu Virbel <txprog>
30a94a7a7a		MTKineticList: Added a "goto_head" method, which takes no arguments and take you to the top(or far	May 08, 2009	Alex Teiche <xelapond>
9ae7f162fb		Added readme file. Added exit button in upper right	May 07, 2009	Trevor Lockley <thatsyour
f6fc12c2be		Added, status post feature. Modal windows used for post error and success.	May 07, 2009	Trevor Lockley <thatsyour
403906825f		merge tip	May 07, 2009	Mathieu Virbel <txprog>
e49dcf757f		mtimagebutton: fix image reloading	May 07, 2009	Mathieu Virbel <txprog>
73ac098c43		scatter: fix division by zero + singular matrix. added scale_min to prevent scaling = 0.	May 07, 2009	Mathieu Virbel <txprog>
7731487dbb		mtwidget: hook the __setattr__ to add inner animation possibility + simplify add_animation() examples:	May 06, 2009	Mathieu Virbel <txprog>
d415fb61b3		animation: add running attribute to known if animation is running or not + add info if animation is an inn	May 06, 2009	Mathieu Virbel <txprog>
f74111f49e		core: add mercurial ignore file to prevent commit of pyc/swp/~ files	May 06, 2009	Mathieu Virbel <txprog>
3a75d99ec6		Removed twitter.pyc because I was not thinking	Apr 28, 2009	Trevor Lockley <thatsyour
574962e883		Adding a new application MTwitter	Apr 28, 2009	Trevor Lockley <thatsyour
ec7dff88d9		Kinetic, Button: Played around with it until it started working right with multiline.	Apr 29, 2009	Alex Teiche <xelapond>
7d93d9c761		MTButton: Added Multi-Line support	Apr 28, 2009	Alex Teiche <xelapond>
e3bd8cb0fb		label: add possibility to get/set attribute from pyglet label at run time.	Apr 28, 2009	Mathieu Virbel <txprog>
a5fef2e99f		label: add autoheight	Apr 28, 2009	Mathieu Virbel <txprog>

Why Mercurial

- Why Not

Comment 37 by lucky.developer, Oct 10, 2008

```
+1 for git.
```

Comment 38 by electronixtar, Oct 21, 2008

```
+1
```

Comment 39 by chris.messina, Oct 22, 2008

```
Given that Android source is hosted on GIT, I can only imagine that it's just a matter of time now.

http://tr.im/android git
```

Comment 40 by alph.pt, Oct 25, 2008

```
Yeah! :) git git git
```

Comment 41 by Bailey.D.R, Oct 25, 2008

```
Mercurial!
```

Comment 42 by patrikbeno, Oct 26, 2008

```
Bazaar, please. The only one with true rename...
```

Comment 43 by djc.ochtman, Oct 26, 2008

```
I'm sorry? Mercurial has true renames, too.
```

Why Mercurial

- Why Not

Comment 37 by lucky.developer, Oct 10, 2008

+1 for git.

Comment 38 by electronixtar, Oct 21, 2008

+1

Comment 39 by chris.messina, Oct 22, 2008

Given that Android source is hoste

http://tr.im/android_git

Comment 40 by alph.pt, Oct 25, 2008

Yeah! :) git git git

Comment 41 by Bailey.D.R, Oct 25, 2008

Mercurial!

Comment 42 by patrikbeno, Oct 26, 2008

Bazaar, please. The only one with true rename...

Comment 43 by djc.ochtman, Oct 26, 2008

I'm sorry? Mercurial has true renames, too.

Comment 58 by luckyluke56, Dec 22, 2008

Really nice, but what DVCS is google going to use? I vote for bazaar.

Comment 59 by justin.forest, Dec 22, 2008

The new "+1" preventing feature in action, hehe.

Comment 60 by pkufranky, Dec 28, 2008

git please

Comment 61 by b...@benatkin.com, Dec 31, 2008

plus won four git :p

Comment 62 by nicolas.alvarez, Dec 31, 2008

When did anyone from Google say this was a poll for which DVCS to implement??

Google is already working on it, which means they already decided on a DVCS system. When you say "<some dvcs> please" or "+1 for <some dvcs>", you're just annoying the 140 people who starred this issue. It won't change anything, because it was already decided.

The Workflow

local action

network action

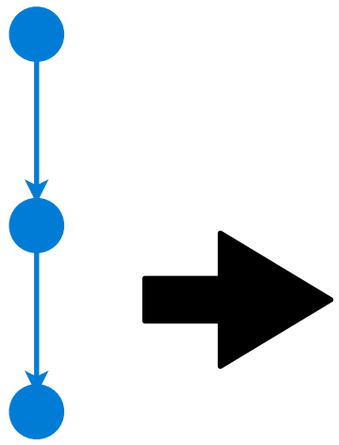
- hg init
- hack hack hack
- hg commit
- hack hack hack
- hg commit
- **hg push**

The Workflow

local action

network action

- hg init
- hack hack hack
- hg commit
- hack hack hack
- hg commit
- **hg push**



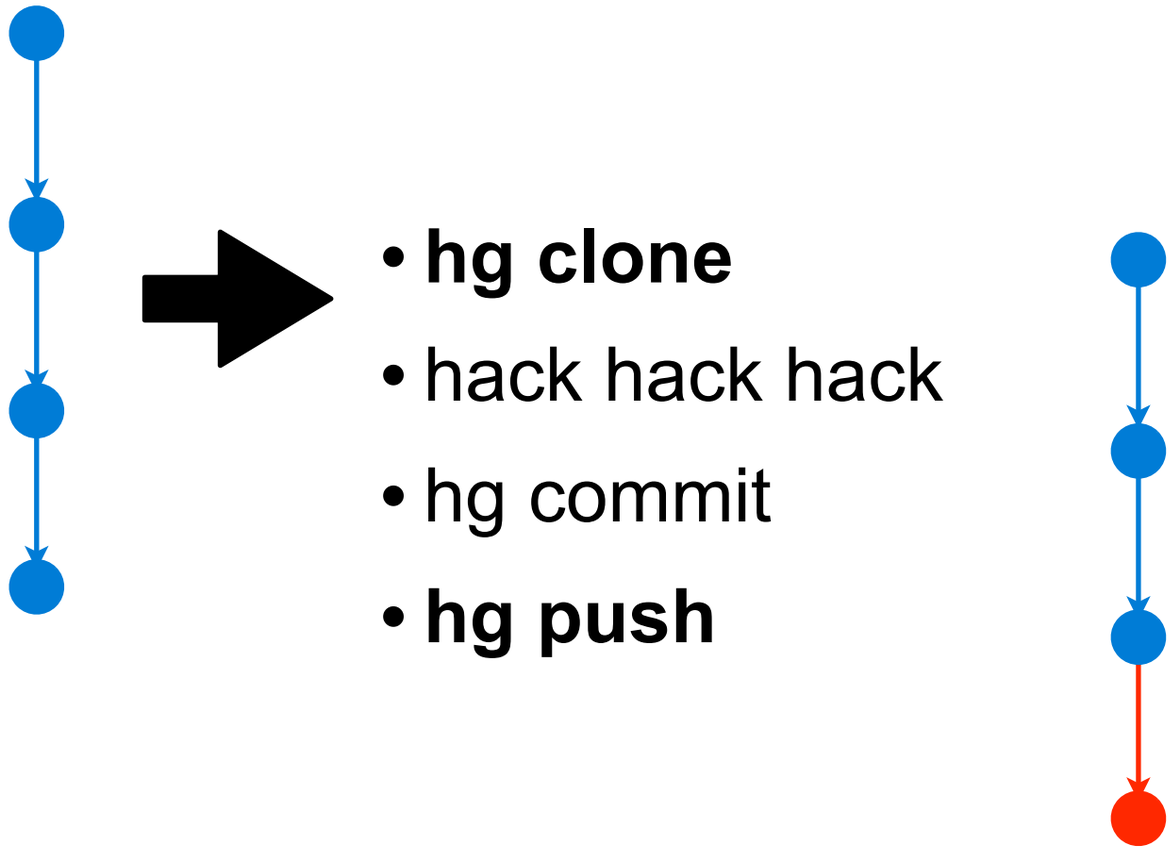
- **hg clone**
- hack hack hack
- hg commit
- **hg push**

The Workflow

local action

network action

- hg init
- hack hack hack
- hg commit
- hack hack hack
- hg commit
- **hg push**
- hack hack hack
- hg commit

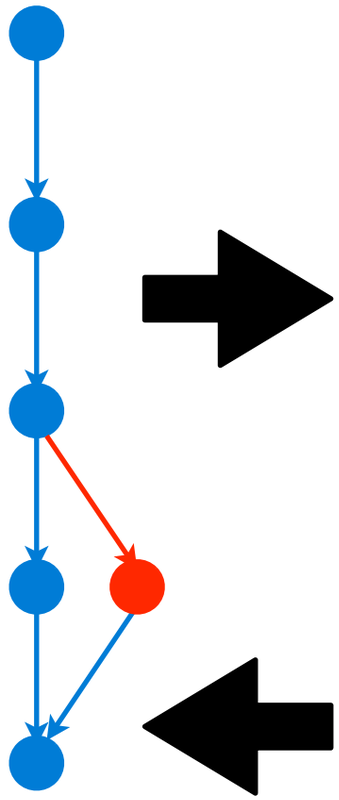


The Workflow

local action

network action

- hg init
- hack hack hack
- hg commit
- hack hack hack
- hg commit
- **hg push**
- hack hack hack
- hg commit
- **hg pull**
- hg merge

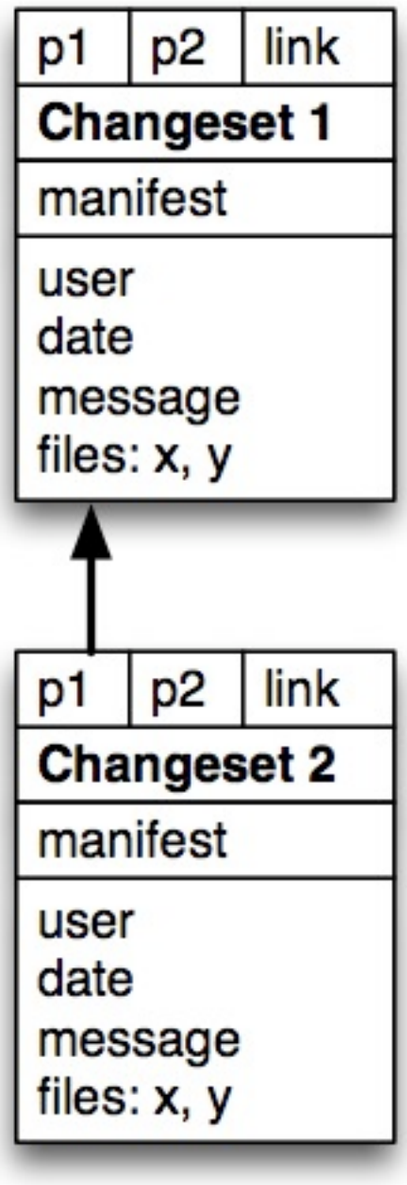


- **hg clone**
- hack hack hack
- hg commit
- **hg push**

Internals

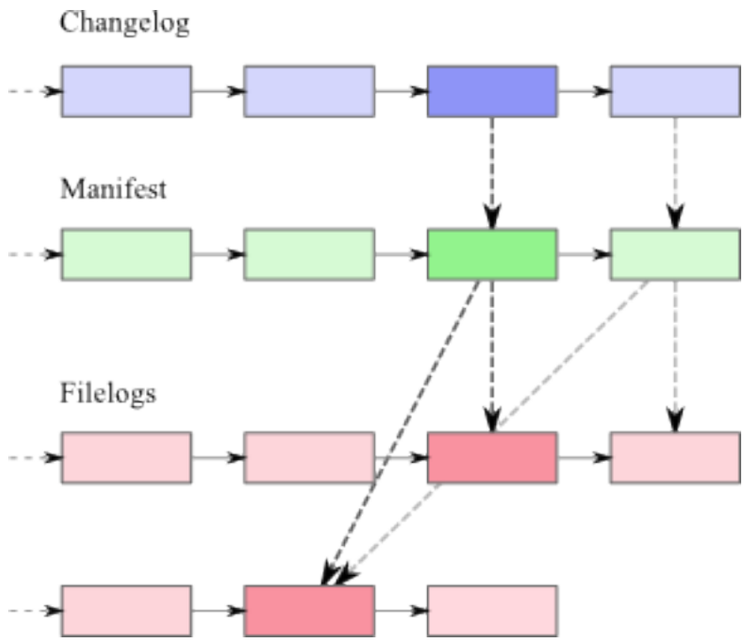
Changesets

- Hash of contents



The DAG(s)

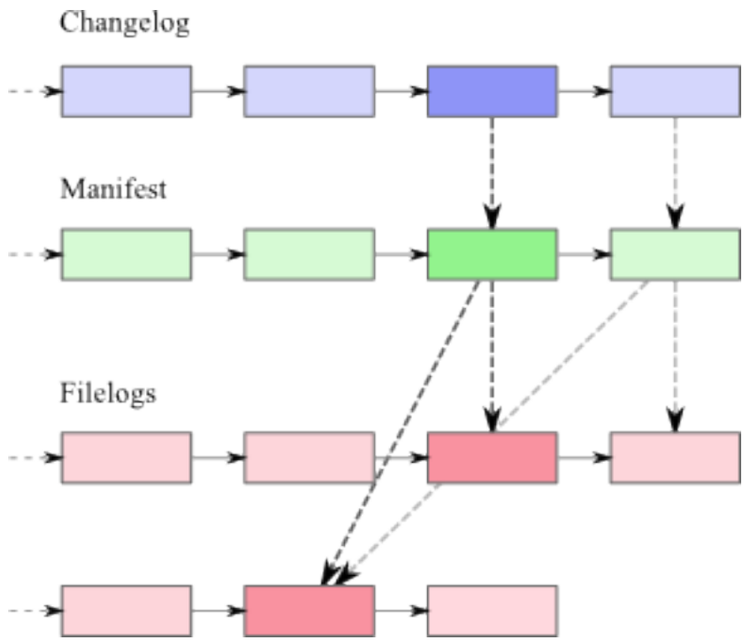
- Changesets
- Manifests
- Files



http://hgbook.red-bean.com

The Revlog

- Requirements
  - Appends
  - Random access
  - Speed
  - Compactness
- Linear Data File
  - Deltas with periodic snapshots
  - Append-only
  - Lock only for writes
- Index
  - Fixed-width records



http://hgbook.red-bean.com

Local Operations

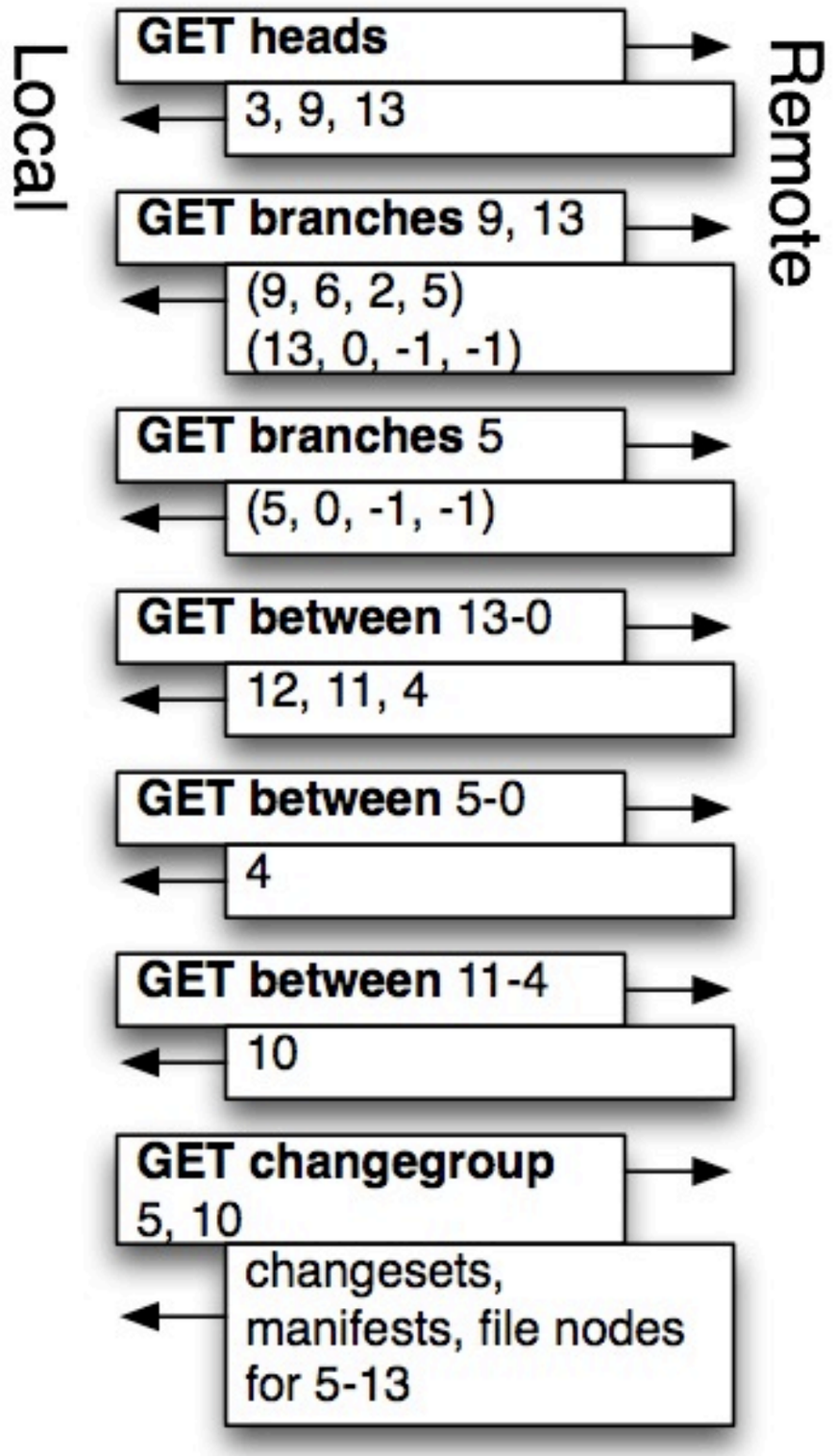
- Write
  - Always adding to repo
  - Files then manifests then changesets
  - Append to data file
  - Update index file
  - Repository is locked during entire operation
- Read
  - Index stays mostly in memory

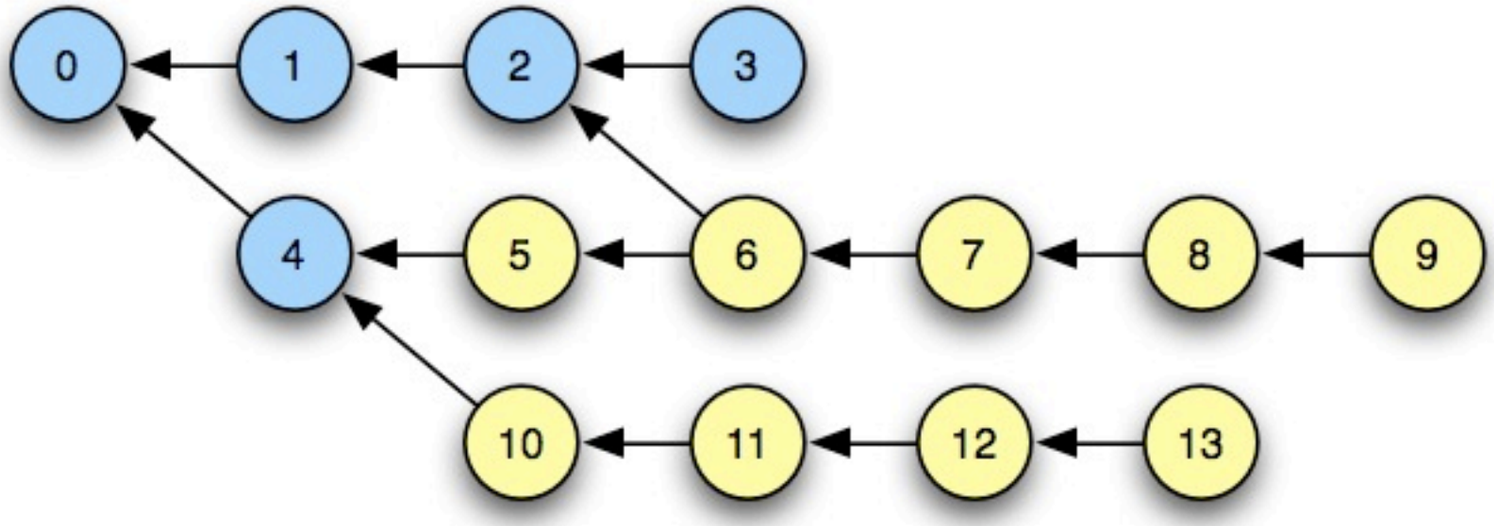
Network Operations: Push

- Get Remote Heads
- Compute
- Upload

Network Operations: Pull

- Negotiate
  - get heads
  - identify bases of new graph segments
- Fetch
  - changegroup optimized for streaming





Googlification

Good News and Bad News

- Requirements
  - No local operations
  - Push, pull
  - Source browsing
  - Simple commits

Mercurial Implementation Assumptions

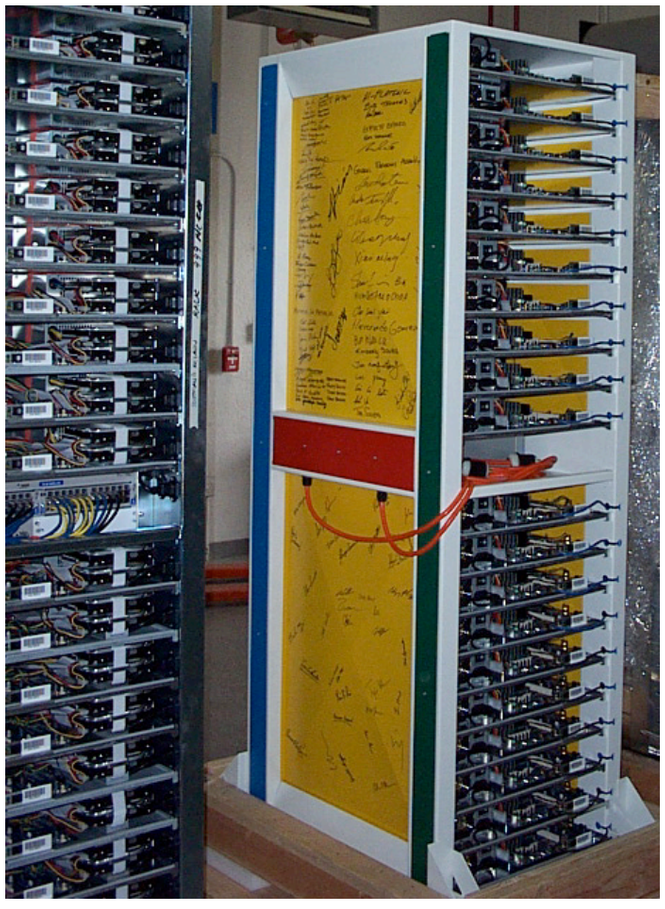
- Direct filesystem access
- Single process, single machine
- Fast in-memory random access to index files

Google Infrastructure

- Big Clusters
- Cheap Hardware
- GFS
- Bigtable

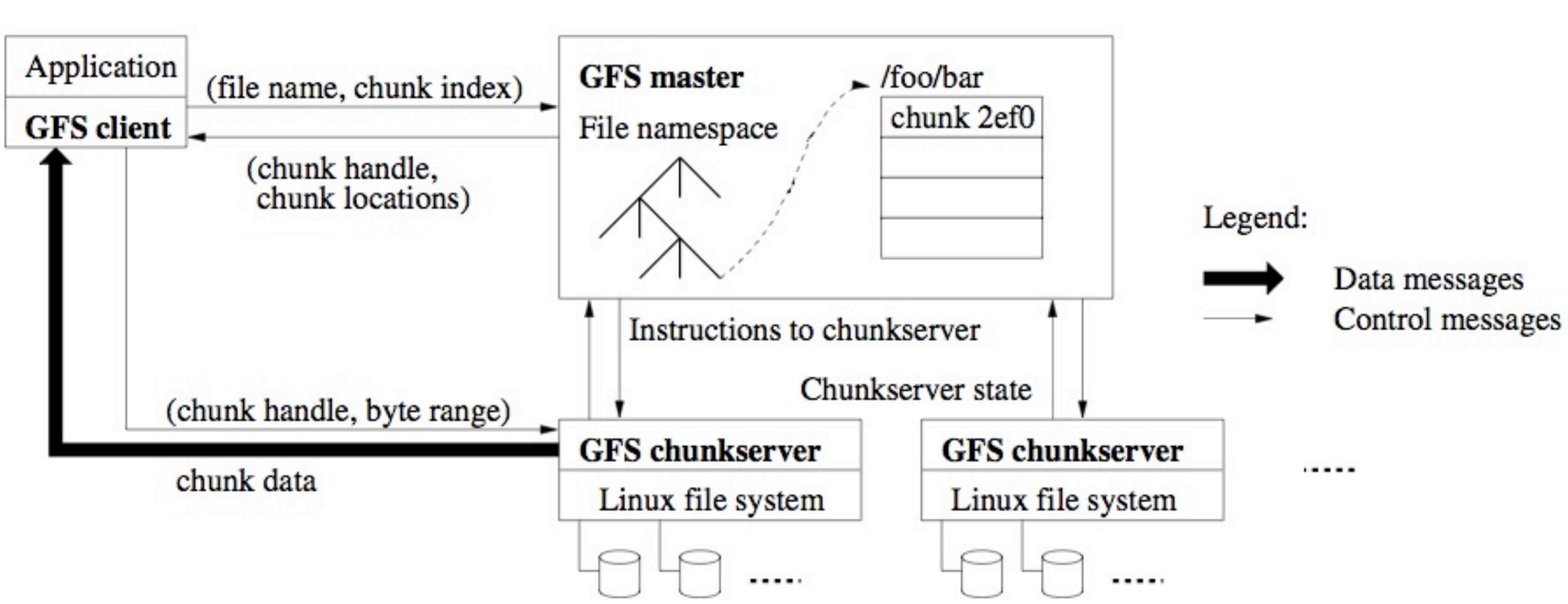
Google Infrastructure

- Big Clusters
- Cheap Hardware
- GFS
- Bigtable



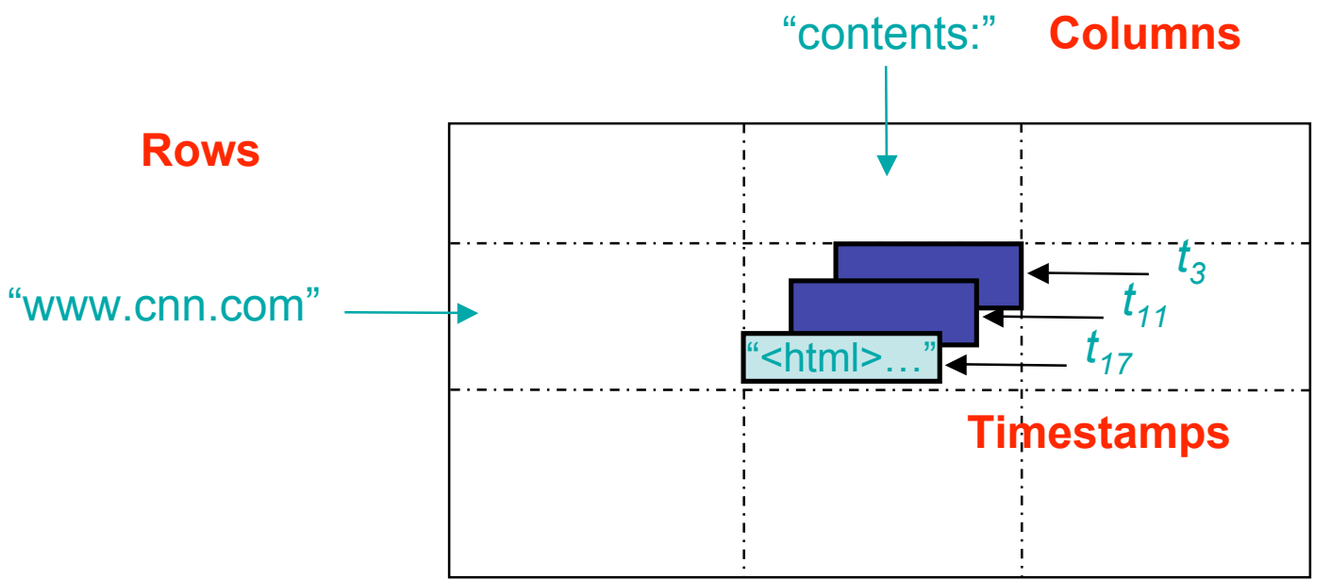
GFS

- Requirements
  - Robust
  - Hundreds of TB, Multi-GB files
  - Fast streaming reads
  - Fast, concurrent appends
- Paper Available



Bigtable

- Non-relational
- Diverse requirements
- (row, column, timestamp) -> string
- Arbitrary keys, opaque values
- Rows partitioned lexicographically into tablets
- Lookup, scan, row-atomic write
- Paper available



Jeff Dean, I/O 2008

Schema

Revision Data (one row per change per cluster)

	revfrag:	**revdata:**	**revmeta:**	**change:**
repo:C:*id*	fragment info	changeset	meta data	change info
repo:M:*id*	fragment info	manifest	meta data	
repo:F*path*:*id*	fragment info	file data (frag 0)	meta data	
repo:F*path*:*id-n*		file data (frag n)		

Repository Row (one per repository, modified atomically)

	repo:summary	**branch:**	**tag:**	**member:*id***	**children:*p*:*c***
repo:	list of heads, quota, etc	branches	tags	set if id is in repo	set if c is a child of p

Operations

- Push
  1. Shovel
  2. Finalize
- Pull
  1. Negotiate
  2. Compute Tree
  3. Stream
- Browse
  - Next and previous N revisions
  - Path history
  - File and directory contents

Results

Optimizations

- Avoid sequential reads and writes
- Let Bigtable do the concurrency
- Minimize DAG walking

Performance Characteristics

- Push
  - wicked fast
- Pull
  - fair number of synchronous reads
  - could be faster
- Browse
  - fast for certain operations

Lessons

- Certain universal truths
  - Engineering is about trade-offs
  - Local faster than remote
  - Scalability has a price

One More Cool Thing

CC BY-SA http://www.flickr.com/photos/naelyn/112322751/

Google™

09

Monday, June 1, 2009