### Distributed Version Control with Mercurial

Martin Geisler (mg@aragost.com)

ClearCase in the 21st Century, Zurich February 3rd, 2012



## About the Speaker

Martin Geisler:

- core Mercurial developer:
  - reviews patches from the community
  - helps users in our IRC channel
- works at aragost Trifork, Zurich:
  - offers professional Mercurial support
  - customization, migration, training
  - advice on best practices

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

#### Frontends

Comparison with Git

### Wrapping Up

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

Frontends

Comparison with Git



### Introduction Mercurial

Centralized vs Distributed Key Mercurial Concepts

### Using Mercurial

Workflows Branches

### Frontends

### Comparison with Git

### Wrapping Up

# What is Mercurial?

Main features:

- fast, distributed revision control system
  - robust support for branching and merging
  - free and open source
  - commercial support available: aragost Trifork and others

# What is Mercurial?

Main features:

- fast, distributed revision control system
  - robust support for branching and merging
  - free and open source
  - commercial support available: aragost Trifork and others
- ▶ installers for Windows, Mac OS X, Linux, ...
  - TortoiseHg is a cross-platform graphical frontend
  - MacHg and SourceTree are fast native Mac OS X frontends
  - ▶ plugins for MS Visual Studio, Eclipse, ...



# What is Mercurial?

Main features:

- fast, distributed revision control system
  - robust support for branching and merging
  - free and open source
  - commercial support available: aragost Trifork and others
- ▶ installers for Windows, Mac OS X, Linux, ...
  - TortoiseHg is a cross-platform graphical frontend
  - MacHg and SourceTree are fast native Mac OS X frontends
  - ▶ plugins for MS Visual Studio, Eclipse, ...
- very user-friendly
  - extensive built-in help for all commands
  - command set resembles CVS and SVN
  - destructive commands delegated to extensions

# Who is Using it?

Mercurial is used by:

- ► Oracle for Java, OpenSolaris, NetBeans, OpenOffice, ...
- Mozilla for Firefox, Thunderbird, ...



#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

### Using Mercurial

Workflows Branches

### Frontends

### Comparison with Git

### Wrapping Up

## Centralized Revision Control

Single repository, multiple working copies:



## Centralized Revision Control

Single repository, multiple working copies:



## Distributed Revision Control

Mercurial duplicates the history on many servers:



## Distributed Revision Control

Mercurial duplicates the history on many servers:



# Why Distributed?

Distributed revision control gives you:

- offline commits
- rich set of fast local operations
- great flexibility



# Why Distributed?

Distributed revision control gives you:

- offline commits
- rich set of fast local operations
- great flexibility

Derived effects:

- fine-grained commits
- searchable history
- branching and merging become a natural task (not something to be feared)
- enables better workflows



### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

### Using Mercurial

Workflows Branches

### Frontends

### Comparison with Git

### Wrapping Up

Pull and merge:

Alice Bob 1

Pull and merge: Alice Bob  $A_1$ 1 commit









Pull and merge:





Changesets:

- atomic repository-wide snapshot
- very similar to a baseline in ClearCase UCM



Pull and merge:







Changesets:

- atomic repository-wide snapshot
- very similar to a baseline in ClearCase UCM

Merging:

- find common ancestor of  $A_2$  and  $B_1$ : I
- do three-way merge between I,  $A_2$ , and  $B_1$
- merge often  $\Rightarrow$  common ancestor is close to tip

# Key Mercurial Commands



Local commands:

- ▶ hg commit: save a snapshot into the current repository
- ▶ hg update: checkout revision into working directory
- hg merge: join different lines of history

# Key Mercurial Commands



Local commands:

- hg commit: save a snapshot into the current repository
- ▶ hg update: checkout revision into working directory
- hg merge: join different lines of history

Network commands:

- ▶ hg pull: retrieve changesets from another repository
- hg push: send your changesets to another repository

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

Frontends

Comparison with Git



### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

Frontends

Comparison with Git

### Wrapping Up

# Workflow in a Team

Mercurial scales from a single team...:



# Workflow in a Team

Mercurial scales from a single team...:



## Workflow Between Company Divisions

... to enterprise-wide development...:



# Workflow Between Two Computers

... to working with yourself:



### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

Frontends

Comparison with Git

### Wrapping Up





## Release Branches


















Named branches:

- heavy-weight branches, integral part of each changeset
- allow tracking and auditing of old history

# Outline

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

### Frontends

### Comparison with Git



# Third-Party Tools

Mercurial is mature and has wide-spread tool support:

- ► Graphical frontends: TortoiseHg, MacHg, SourceTree, ...
- ► IDEs: Eclipse, NetBeans, IntelliJ, Visual Studio, ...
- ▶ Project Support: Trac, JIRA, Maven, Hudson, ...



Context menu in Windows Explorer:







#### Browsing history:





#### Interactive commit tool:







Update with shelve option:



#### Detecting renames based on file similarity:



tortoisehg/util/cache.py  Source  Dest  % Match    tortoisehg/util/cache.py  tortoisehg/util/cache.py  100%    Min Similarity: 50%	tortoisehg/util/cache.py  Source  Dest  % Match    tortoisehg/util/cache.thg.py  tortoisehg/util/cache.thg.py  tortoisehg/util/cache.py  100%    Min Similarity: 50%	Unrevisioned Files	Candidate Match	ies	
Itertoisehg/util/cachethg.py  tortoisehg/util/cache.py  100%    Min Similarity:  50%	Itertoisehg/util/cachethg.py  tertoisehg/util/cache.py  100%    Min Similarity: 50%  Itertoisehg/util/cachethg.py  tertoisehg/util/cachethg.py    Im Similarity: 50%  Itertoisehg/util/cachethg.py	tortoisehg/util/cache.py	Source	Dest	% Match
Min Similarity: 50%	Min Similarity: 50% Accept Selected Mate		tortoisehg/util/c	achethg.py tortoisehg/util/c	ache.py 100%
Differences from Source to Dest	Differences from Source to Dest	Min Similarity: 50%	d Rename		Accept Selected Mat
		Differences from Source to Dest			

### Cloning a repository:



	Destinati	on:		
	Sh	ortName:	hgcollapse	
	Lo	cal Path:	/Users/jason/hgcollapse	Brows
hgcollapse				
dvanced Options: 🔻				
	Use SSH comma	und 🗌		
Clone only Revi	sion. Tag or Bran	ch 🗆		
Class and	Caraifia Daviaia	C		
Cione only	specific Revision	(5)		
	Remote Comma	ind 🔄		

### Updating to a revision:



ect Revision to L	Jpdate to: author date	labels	description	changeset
0303	FUJIWARA F 3 mont		i18n-ja: synchronized with e7727a545c48	35bcb3152e67
0304	Georg Bran 3 mont		i18n: fix typo in German translation	0824aedd0079
0305	Martin Geis 2 mont		Merge with stable.	81211cf5703b
0306	Wagner Bru 2 mont		i18n-pt_BR: synchronized with b08ffd27dfc8	e7639acd54f6
0307	Wagner Bru 2 mont		merge with i18n stable	b0298eaaeaf7
308	Matt Macka 2 mont	1.4.3	Merge with i18n	4aa619c4c2c0
309	Matt Macka 2 mont		Merge with i18n	0df9e63f0c07
0310	Matt Macka 2 mont		Added tag 1.4.3 for changeset 4aa619c4c2c0	61750d3aa00c
311	Matt Macka 2 mont		Added signature for changeset 4aa619c4c2c0	7c5eb0988e7a
1312	Matt Marka 2 mont		Merge with stable	131a012aa878
Changeset: 10 Parents: 10	0310 : 61750d3aa00c 0308:4aa619c4c2c0			
Changeset: 10 Parents: 10 Author: M Date: 2 Modified: h Description: Ac	0310 : 61750d3aa00c 0308:4aa619c4c2c0 att Mackall months ago (2010-02- gtags dded tag 1.4.3 for chang	01 09:34 -060 geset 4aa619	00) 040200	



### Collapsing changesets:



rision 🔺 gra	ph	author	date	labels	description	changeset
7		jfh	8 weeks		- change for CD	6a45701c259c
8	+	jfh	8 weeks		- change for CC	Øfe4cd5dfØef
9	+	jfh	8 weeks		- change for CB	6ed1596d9af7
10	<b>–</b>	jfh	8 weeks		- change for CA	a2fd8700743f
11		jfh	8 weeks		- BugFix for CH	c53e5e8a6c8e
12	11	jfh	8 weeks		- BugFix for CI	d2d1f3c8ba70
	+					
15	+	jfh	8 weeks		- change for CE	dbf40643867d
16		jfh	31 hour	default, tip	- A simple change	3adf10560070
Changeset: Tags: Branch: Parents: Author:	16 : 3a tip default 7:6a45 jfh	df1056007 701c259c	0			



#### Removing changesets from a repository:

Hg

ision 🔺	graph	author	date	labels	description	changeset
3	1	Jin	o weeks		- mira change	56243884038
4	1	Jfn	8 weeks		- change for four	80919088010
5	1	Jfn	8 weeks		- change for five	114924380110
6	1	jfn	8 weeks		- change for six	67305027800
<i>′</i>		jfn :e-	8 weeks		- change for CD	68457010259
8		Jrn	8 weeks		- change for CC	øre4cd3drøe
	1 I *					
15		jm 10-	8 weeks		- change for CE	db148043807
16	1	Jrn	31 nour	defends sim	- A simple change	38071056007
17	-	Jtn	31 nour	default, tip	- Commit the merge	1200240304
Descript Commit: Descript Commit: Descript	ion: - char 10 (iff ion: - char 11 (iff ion: - Bugi	nge for CB n), 8 weeks nge for CA n), 8 weeks Fix for CH	ago ago			
at Will H	lappen: Th	e selected re writes histor	visions withir . Therefore y	9 through t ou should ne you are doir	to 17 will be removed from the repositor ever strip any revisions that have already re.	y. This strip operation is destructiv been pushed to other repositories



## Outline

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

#### Frontends

### Comparison with Git



# Git

Git was created by Linus Torvalds for Linux Kernel development:

- ► focus on speed
- builds on the same concepts as Mercurial:
  - fully distributed with push/pull between clones
  - commits are organized in a changeset graph

# Git

Git was created by Linus Torvalds for Linux Kernel development:

- focus on speed
- builds on the same concepts as Mercurial:
  - fully distributed with push/pull between clones
  - commits are organized in a changeset graph

Git is extremely flexible, but flexibility comes at a price:

- ▶ git log has 150 different flags, log help text is 1496 lines
- default command set includes destructive commands
- steep learning curve must learn about the "staging area"

# Git

Git was created by Linus Torvalds for Linux Kernel development:

- focus on speed
- builds on the same concepts as Mercurial:
  - fully distributed with push/pull between clones
  - commits are organized in a changeset graph

Git is extremely flexible, but flexibility comes at a price:

- ▶ git log has 150 different flags, log help text is 1496 lines
- default command set includes destructive commands
- ► steep learning curve must learn about the "staging area" Git is still somewhat Linux-centric:
  - started as a tool for Linux Kernel hackers, made by Linux Kernel hackers
  - Git has since then been ported to Windows
  - Windows performance is still reported to be sub-par

Similar to using bookmarks in Mercurial:



Similar to using bookmarks in Mercurial:





Similar to using bookmarks in Mercurial:





Similar to using bookmarks in Mercurial:

branch pointer advances on commit:



pull and merge of changesets from others:



Similar to using bookmarks in Mercurial:

branch pointer advances on commit:



pull and merge of changesets from others:



Similar to using bookmarks in Mercurial:



# Deleting a Git Branch

Git can garbage collect changesets:

repository with experimental branch:



# Deleting a Git Branch

Git can garbage collect changesets:

repository with experimental branch:



- delete branch with git branch -D experiments
- faint commits may disappear on next git gc



# Deleting a Git Branch

Git can garbage collect changesets:

repository with experimental branch:



- delete branch with git branch -D experiments
- faint commits may disappear on next git gc
- delete on server with git push origin :experiments

### Outline

#### Introduction

Mercurial Centralized vs Distributed Key Mercurial Concepts

Using Mercurial Workflows Branches

Frontends

#### Comparison with Git

### Wrapping Up

### Mercurial in a Nutshell

Mercurial changes the way you develop:

- simple yet strong model for both branching and merging
- power tool instead of necessary evil
- light-weight and snappy

### More Information

- Mercurial homepage: http://mercurial.selenic.com/
- Mercurial: The Definitive Guide: http://hgbook.red-bean.com/
- Getting Started:

http://mercurial.aragost.com/kick-start/ http://mercurial.ch/ http://hginit.com/

Some free Mercurial hosting sites: http://bitbucket.org/ http://code.google.com/ http://sourceforge.net/ http://www.codeplex.com/

### Mercurial Contributors

From http://ohloh.net/p/mercurial/map:


#### Mercurial Contributors

From http://ohloh.net/p/mercurial/map:



#### Outline

#### Performance Study: OpenOffice

Subversion and Branches

The Underlying Model

Using History

Changing History



# OpenOffice

Fairly large repository:

- ▶ 70,000 files, 2,0 GB of data
- 270,000 changesets, 2,3 GB of history

Mercurial is fast on a repository of this size:

```
$ time hg status
0.45s user 0.15s system 99% cpu 0.605 total
$ time hg tip
0.28s user 0.03s system 99% cpu 0.309 total
$ time hg log -r DEV300_m50
0.30s user 0.04s system 99% cpu 0.334 total
$ time hg diff
0.74s user 0.16s system 88% cpu 1.006 total
$ time hg commit -m 'Small change'
1.77s user 0.25s system 98% cpu 2.053 total
```



#### Outline

Performance Study: OpenOffice

#### Subversion and Branches

The Underlying Model

Using History

Changing History



Subversion knows nothing about branches!

- but SVN has a cheap copy mechanism
- used for tags and branches



Subversion knows nothing about branches!

- but SVN has a cheap copy mechanism
- used for tags and branches

r10

trunk/ hello.c Makefile branches/ tags/

Subversion knows nothing about branches!

- but SVN has a cheap copy mechanism
- used for tags and branches



Subversion knows nothing about branches!

- but SVN has a cheap copy mechanism
- used for tags and branches



# Merging Branches in SVN

The support is incomplete and fragile:

- renamed files are not merged correctly
- old clients will not update the merge info

# Merging Branches in SVN

The support is incomplete and fragile:

- renamed files are not merged correctly
- old clients will not update the merge info

From the SVN Book:

The bottom line is that Subversion's merge-tracking feature has an extremely complex internal implementation, and the svn:mergeinfo property is the only window the user has into the machinery. Because the feature is relatively new, a numbers of edge cases and possible unexpected behaviors may pop up. —Version Control with Subversion

(Mercurial has robust built-in support for merging branches.)



#### Outline

Performance Study: OpenOffice

Subversion and Branches

The Underlying Model

Using History

Changing History



# The Underlying Model

- A Mercurial changeset conceptually consist of:
  - ▶ 0-2 parent changeset IDs:
    - root changeset has no parents
    - normal changesets have one parent
    - merge changesets have two parents
  - date, username, commit message
  - difference from first parent changeset
  - changeset ID is computed as SHA-1 hash of the above
  - makes it impossible to inject malicious code on server



#### Immutable History

SHA-1 hashes as changeset IDs have some consequences:

- ► a changeset ID is a hash of the entire history
- changing history changes subsequent changesets
- history is immutable, you can only make new history:

# $A \longrightarrow B \longrightarrow C \longrightarrow D \longrightarrow E$



#### Immutable History

SHA-1 hashes as changeset IDs have some consequences:

- a changeset ID is a hash of the entire history
- changing history changes subsequent changesets
- history is immutable, you can only make new history:



#### Immutable History

SHA-1 hashes as changeset IDs have some consequences:

- a changeset ID is a hash of the entire history
- changing history changes subsequent changesets
- history is immutable, you can only make new history:





#### Outline

Performance Study: OpenOffice

Subversion and Branches

The Underlying Model

#### Using History

Changing History



### Browsing the History of a File

The hg annotate command is invaluable:

- you see when each line was introduced
- you can quickly jump back to earlier versions

History of Mercurial's README file:

```
3942: Basic install:
445:
3942: $ make  # see install targets
3942: $ make install  # do a system-wide install
3942: $ hg debuginstall # sanity-check setup
3942: $ hg  # see help
0:
# ...
```

#### Better interface in hg serve

# Searching File Content

Ever wondered when a function was introduced?

hg grep can help you!

Example: When was hg forget introduced?

\$ hg grep --all 'def forget' commands.py commands.py:8902:+:def forget(ui, repo, \*pats, \*\*opts): commands.py:3522:-:def forget(ui, repo, \*pats, \*\*opts): commands.py:814:-:def forget(ui, repo, file1, \*files): commands.py:814:+:def forget(ui, repo, \*pats, \*\*opts): # ...





















#### Outline

Performance Study: OpenOffice

Subversion and Branches

The Underlying Model

Using History

#### Changing History



Tired of all those merges? Use the rebase extension!

► Revision graph:





Tired of all those merges? Use the rebase extension!

► Revision graph:





В

Tired of all those merges? Use the rebase extension!

D

Ε

Ζ

► Revision graph:

A



Tired of all those merges? Use the rebase extension!

► Revision graph:







Tired of all those merges? Use the rebase extension!

► Revision graph:





 $\blacktriangleright \text{ Rebase:}$   $A \longrightarrow B \longrightarrow C \longrightarrow D \longrightarrow E$   $X \longrightarrow Y \longrightarrow Z \longrightarrow D' \longrightarrow E'$ 

Tired of all those merges? Use the rebase extension!

► Revision graph:





 $\blacktriangleright \text{ Rebase:}$   $A \longrightarrow B \longrightarrow C \longrightarrow D \longrightarrow E$   $X \longrightarrow Y \longrightarrow Z \longrightarrow D' \longrightarrow E'$ 

Beware: public changes should never be rebased.

#### Maintaining Patch Series

The mq extension makes it easy to maintain a patch series:

Works nicely for local modification for upstream sources.

#### Maintaining Patch Series

The mq extension makes it easy to maintain a patch series:



Works nicely for local modification for upstream sources.
### Maintaining Patch Series

The mq extension makes it easy to maintain a patch series:



Works nicely for local modification for upstream sources.

### Maintaining Patch Series

The mq extension makes it easy to maintain a patch series:



Works nicely for local modification for upstream sources.

### Maintaining Patch Series

The mq extension makes it easy to maintain a patch series:



Works nicely for local modification for upstream sources.

Inspired by git rebase -i, histedit lets you

► reorder changesets:

$$A \longrightarrow B \longrightarrow C \quad \rightsquigarrow \quad A \longrightarrow C' \longrightarrow B'$$



Inspired by git rebase -i, histedit lets you

B'

ВС

- ► reorder changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A \rightarrow C'$
- ► fold changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A$



Inspired by git rebase -i, histedit lets you

- ► reorder changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A \rightarrow C'$
- ► fold changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow$

$$A \longrightarrow BC$$

B'

► drop changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A \rightarrow C'$ 

Inspired by git rebase -i, histedit lets you

- ► reorder changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A \rightarrow C' \rightarrow B'$
- ► fold changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow$

$$A \longrightarrow BC$$

• drop changesets:  $A \rightarrow B \rightarrow C$ 

$$A \rightarrow C'$$

► edit changesets:  $A \rightarrow B \rightarrow C \quad \rightsquigarrow \quad A \rightarrow X \rightarrow B' \rightarrow C'$ 

#### Outline

Performance Study: OpenOffice

Subversion and Branches

The Underlying Model

Using History

Changing History

#### Talking to Other Systems

# Migrating History

The convert extension can import history:

- ► CVS, SVN, Git, Bazaar, Darcs, ...
- incremental conversion
- many options for fiddling with branches, authors, ...



# Migrating History

The convert extension can import history:

- ► CVS, SVN, Git, Bazaar, Darcs, ...
- incremental conversion

many options for fiddling with branches, authors, ...

Interestingly, **convert** can import from Mercurial:

- --filemap lets you exclude and rename files
- --branchmap lets you rename branches

## Interfacing with Subversion

The hgsubversion extension let's you:

- use hg clone on a SVN URL
- use hg pull to convert new SVN revisions
- use hg push to commit changesets to SVN server

Goal: make hg a better Subversion client than svn!





# Interfacing with Git

Need to work on a Git repository? Try hg-git!

- ► Mercurial extension: you get the nice hg command line
- round-tripping: changeset hashes are preserved

