November 2011

# Scratching your own itch

**Joey Echeverria** | Solutions Architect

joey@cloudera.com

@fwiffo

**cloudera**

# About Joey

- Solutions Architect
- 6 months
- 3+ years

cloudera

# Why Apache Sqoop (incubating)?

- Data center integration
- Extensible
  - Sources
  - Destinations
- Company specific
- Open source
  - Yay!

cloudera

# Getting the code

- Apache SVN
  - https://svn.apache.org/repos/asf/incubator/sqoop/

- Apache GIT
  - http://git.apache.org/sqoop.git
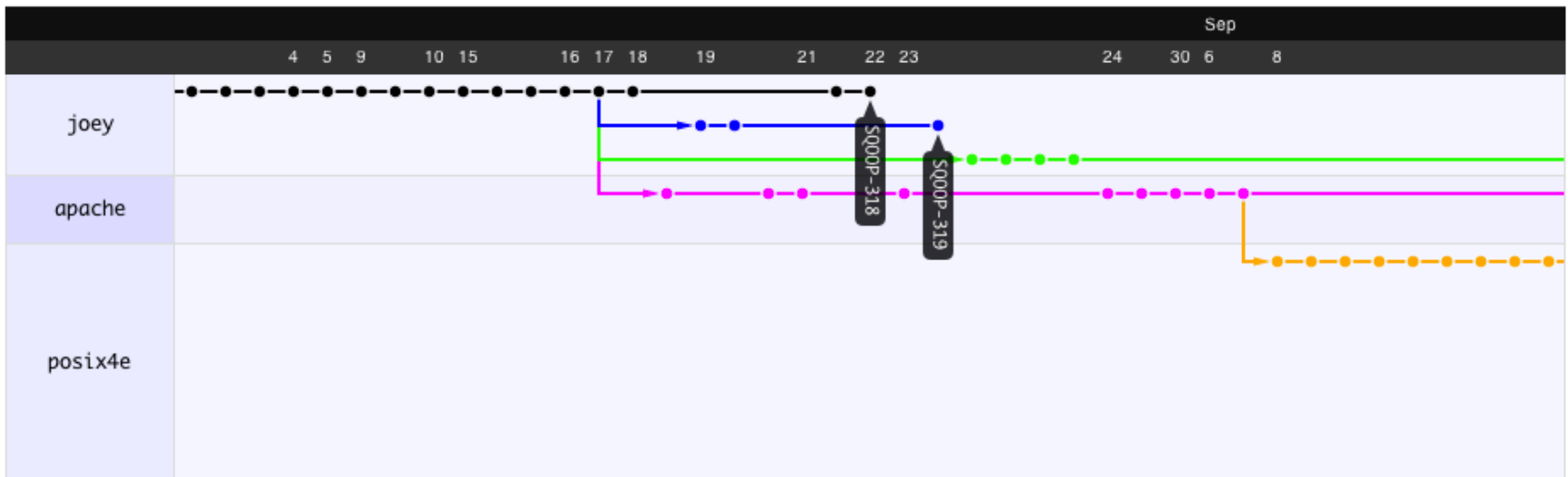
- GitHub
  - https://github.com/apache/sqoop.git

cloudera

# Fork you



The sqoop network graph

All branches in the network using *joey/sqoop* as the reference point. *Read our blog post about how it works.*

cloudera

# Building the code

- Requirements
  - Ant 1.7.1
  - JDK 1.6
  - Maven
    - SQOOP-306
- Targets
  - jar
  - test
  - checkstyle

cloudera

# Checkstyle

## CheckStyle Audit

Designed for use with CheckStyle and Ant.

| Summary | |
|---------|---|
| **Files** | **Errors** |
| 423 | 0 |

| Files | |
|-------|---|
| **Name** | **Errors** |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/ConnFactory.java | 0 |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/Sqoop.java | 0 |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/SqoopOptions.java | 0 |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/RelatedOptions.java | 0 |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/SqoopParser.java | 0 |
| /Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/ToolOptions.java | 0 |

cloudera

# IDE

cloudera

# Testing

- JUnit
- By hand
  - ant package
  - cd build/sqoop-1.5-incubating-SNAPSHOT
  - bin/sqoop

cloudera

# Contributing back: Why?

- Next release
- Lower long-term cost
- For the community!
- It's easier than you think!

cloudera

# Fears

- People won't want my patch
  - Rare
  - Small projects
- Only I need this feature
  - That's what everyone says…
- Too much process!
  - A little time now, a massive headache later

cloudera

# Sqoop vs. HDFS

- Small project
  - Less contentious
  - Easier to grok the code
  - Easy deployment
- Immediate feedback
- Great place to start

cloudera

# Contributing back: How?

- File a JIRA
  - https://issues.apache.org/jira/browse/SQOOP
- Post your patch on the issue
- File a review request
  - https://reviews.apache.org/groups/Sqoop/

cloudera

# Posting patches

cloudera

# The review process

- Upload your patch
- Describe your changes
- Describe your testing
- Wait for feedback
- Repeat

cloudera

# What else?

- Subscribe to the mailing list
  - sqoop-dev@incubator.apache.org
- Look for open issues
  - https://issues.apache.org/jira/browse/SQOOP
- Get hacking

cloudera

# Questions?

cloudera