Scratching your own itch

Joey Echeverria | Solutions Architect
joey@cloudera.com
@fwiffo
About Joey

• Solutions Architect
• 6 months
• 3+ years
Why Apache Sqoop (incubating)?

• Data center integration
• Extensible
  – Sources
  – Destinations
• Company specific
• Open source
  – Yay!
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
The sqoop network graph

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

Show Help

Last updated: about an hour ago
Building the code

• Requirements
  – Ant 1.7.1
  – JDK 1.6
  – Maven
    • SQOOP-306

• Targets
  – jar
  – test
  – checkstyle
## CheckStyle Audit

Designed for use with CheckStyle and Ant.

<table>
<thead>
<tr>
<th>Summary</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Files</td>
<td>Errors</td>
</tr>
<tr>
<td>423</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Files</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Errors</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/ConnFactory.java</td>
<td>0</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/Sqoop.java</td>
<td>0</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/SqoopOptions.java</td>
<td>0</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/RelatedOptions.java</td>
<td>0</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/SqoopParser.java</td>
<td>0</td>
</tr>
<tr>
<td>/Users/joey/src/apache/sqoop/src/java/com/cloudera/sqoop/cli/ToolOptions.java</td>
<td>0</td>
</tr>
</tbody>
</table>
Testing

- JUnit
- By hand
  - ant package
  - cd build/sqoop-1.5-incubating-SNAPSHOT
  - bin/sqoop
Contributing back: Why?

• Next release
• Lower long-term cost
• For the community!
• It’s easier than you think!
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
Sqoop vs. HDFS

• Small project
  – Less contentious
  – Easier to grok the code
  – Easy deployment

• Immediate feedback

• Great place to start
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/
Posting patches

Contribution intended for inclusion in ASF products (e.g., patches, code) must be licensed to ASF under the terms of the Apache License. Other attachments (e.g., log dumps, test cases) need not be.

- Attachment not intended for inclusion
- Grant license to ASF for inclusion in ASF works (as per the Apache License §5)

Comment: Updated the patch with suggestions from review board.
The review process

• Upload your patch
• Describe your changes
• Describe your testing
• Wait for feedback
• Repeat
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
Questions?