Bigtop 0.4.0 hackathon May 17, 2012 agenda

Agenda

- Introduction
- Roadmap for Bigtop 0.4.0
- Packaging in Bigtop
- Validation and testing in Bigtop
- Deployment aspects of Bigtop

Things folks wanted to talk about (make sure to add to the list!)

- Anatomy of Bigtop
- How to contribute to Bigtop
- Best practices of using Git
- Best practices of using Jenkins
- Hadoop branches
- David Liu's bigtop command line launcher proposal
- Proposal of "How to add new components to Bigtop"
- can we get Bruno to talk about his Ganglia recipes? Just a demo. Don't need his code. We can make that a project for the group.
- Can we get a list of common use cases for the bigtop components? Perhaps a talk from a Solution Architect about using Hive on top of Hbase, like what StumbleUpon does? And maybe some other popular use cases we can use to duplicate and test with? Maybe duplicate the webtable Hbase example as a starting point? Hive+Hadoop or Pig+Hadoop are too simple for most deployments. Maybe a list of the most common configurations?

Biggest challenges, things we'd really like to fix:

- Building a reliable transitive closure of dependencies
  - what to do about the Maven infrastructure?
  - what to do with the ant-based projects?
- Off-line builds
- Classpath management
  - harmonization of versions (related to Maven infra)
  - construction of the classpath
  - handling of depenencies
  - local filesystem vs. HDFS (mostly Oozie's use case)
- Unification of init.d script/migration to upstart, etc.