

Jenkins at Apache

Andrew Bayer

abayer@apache.org



APACHECON
EUROPE

CORINTHIA HOTEL
BUDAPEST, HUNGARY
— NOVEMBER 17-21, 2014 —



What We Provide

- builds.apache.org
- Open to any ASF projects
- 1,126 jobs (as of 14 November 2014)
- 20 always-on and functional slaves (2 Windows), total of 35 executors
- 4,297 builds run from ~7 Nov - ~14 Nov, from 518 different jobs.
- Very little queuing other than deliberate configuration and multiple jobs locked to single executor



Where We Were In Spring 2014

- Not in a good place
- Less and smaller slaves - overwhelmingly VMs or ancient Hadoop-dedicated slaves hosted by Yahoo!
- Inconsistent slave environments
- Things just generally sucked.
 - builds.a.o crashed/hung All. The Time.



builds.apache.org OW OW OW OW

- Brutally unstable master
 - Major performance and reliability problems due to disk issues - no one was watching, so no one found the real problem for months.
 - Unstable private build of Jenkins
 - Combination uncovered lots and lots of edge-case performance problems in Jenkins at scale



How We Got To Where We Are Now

- Fixed the zfs volume issues causing brutal disk IO issues on builds.a.o master.
- Trimmed problematic plugins and obsolete data, speeding up load time significantly.
- Moved to the Jenkins LTS train - stable, tested releases with critical backports added, plus a little curation of our own.
 - Some of the edge-case IO issues still cropped up due to ZFS, so we have backported fixes for them.



builds.apache.org Is Aliiiiiiiiive

- End result:
 - builds.a.o much, much, much more stable and much speedier.
 - Master host better monitored, with more people keeping track of it.
 - I personally check builds.a.o load/perf daily to see if anything looks wrong.



Jenkins Enterprise!

- Biiiiig thanks to CloudBees for this!
- We are running the Jenkins Enterprise plugins, with a generously donated license from CloudBees.
- Initial value is in some backend plugins
 - Faster archiving based on diffs
 - CloudBees Support plugin for regular system health information
 - Improved SSH slaves plugin
- Also the GitHub Pull Request Builder plugin!



Slaves!

- Yahoo! provided 17 modern physical hosts running Ubuntu 14.04, replacing the 6-8 much smaller Debian 32-bit slaves for Hadoop - 10 used by Hadoop ecosystem, 8 for everyone.
- All Ubuntu always-on slaves now run on bare metal, generally with multiple executors.
- More stuff installed!
 - Most notably, Docker is installed on all slaves with the “Docker” label.



Slave configuration!

- Slave configs managed with Ansible (based on initial config set up by Y!/Hortonworks on the new machines).
 - Playbook on GitHub at <https://github.com/apache/toolchain>
- We intend to move to Puppet to unify with the rest of ASF Infra, but just having *something* has made a huge difference.



Burst Capacity!

- Thanks to Rackspace for this!
- When all Ubuntu nodes are at capacity, up to 5 single node slave VMs are spun up at Rackspace and added to Jenkins.
- Specs are a bit in flux - initially at 4gb RAM, but now trying out 8gb RAM.
- Images built using Packer and Puppet - that puppet will also be merged into the general ASF Infra puppet.



More bursting! MORE BURSTING!

- We want more!
- Burst capacity is very very good to have - less need for dedicated hardware for jobs that can run just as happily on dynamically provision slaves.
- To grow more, we need more providers to donate us credits! =)



BUILDS JIRA project

- Build/Jenkins-related JIRAs can get lost in INFRA.
- So we created BUILDS.
- Use it! =)



Where We Plan To Go From Here

- Address remaining queuing issues
 - Get functional/reliable FreeBSD and Solaris slaves back up.
 - Get at least one more slave (or on-demand image) that the Lucene/Solr jobs can use.
- More on-demand!
 - Like I said. =)



Windows. Ow. Windows.

- Two Windows slaves currently.
- Working with Microsoft Azure to add more.
- Need a config management solution for setting up these slaves.
- If anyone's got experience in Windows config management and would like to help, please let us know! (builds@apache.org) =)



One Config Management

- Multiple config management tools/repos etc makes things confusing.
- Unify on just one - ASF Infra Puppet.
- Move Ansible-powered slaves at Y! to ASF Infra Puppet.
- Move the Packer on-demand slave image creation to use ASF Infra Puppet as well.



Better Performance on builds.apache.org master

- ZFS on Linux is kind of a pain and eats all the RAM.
- And more importantly, it's not actually useful for Jenkins, since Jenkins generally reads files just once to load them into memory.
- So we'd like to:
 - Get more RAM on crius.apache.org (builds.apache.org's host) - can't ever have too much RAM!
 - Move builds.apache.org's storage from ZFS to XFS or ext4.



Where *I'd* Like To Go

- builds.apache.org is one of the biggest public Jenkins instances. That's cool.
- But it should be *bigger*!
 - A fair number of projects don't use builds.a.o - they use their own Jenkins masters.
 - I'd like to bring them home!



More Dogfooding!

- Hmm...dynamically provisioned slaves...what Apache projects could be relevant to that...oh yeah!
 - Apache jclouds (cheating - we already use this for slave provisioning)
 - Apache CloudStack - forget commercial providers, we should be using CloudStack for slaves!
 - Apache Mesos - Docker, Mesos, Jenkins - they all go verrrrry well together.
- How to do this? Well, get hardware, to start.



Parallel Testing Infrastructure!

- A few ASF projects are already working on this - Hive, Hadoop, for example.
- Run build locally without test execution, then spin up a bunch of VMs and distribute the actual test execution across them.
- It works! It's nice! It's *fast*!
- Work underway to standardize the tooling for use by any Java project at ASF.
- ...then we just need the VMs. =)



Reviewbot!

- Jenkins Reviewbot plugin is...not bad.
- Polls Reviewboard for pending changes, kicks off builds to report back to RB - like the JIRA Patch build setup.
- Could be more like GitHub PR builder - web hooks calling to Jenkins, triggering builds directly.
- Investigation continuing.



Test Reporting Database!

- Jenkins is great at running builds and reporting on them.
- Jenkins is not great at storing lots of historical data and being a source for analytics.
- Working on a plugin to store JUnit-style test reports in an RDBMS
 - Results stay forever, even once build is removed from Jenkins
 - Queryable with SQL, obviously, and a REST API exposed by Jenkins.
 - Find flaky tests programmatically!
- More to come in the future on this.



Where Do *You* Want builds.apache.org to go?

- The roadmap on previous slides comes from my observations on the admin side, and talking with a few projects.
- But I'm not talking with all of them!
- So - what are your problems on builds.a.o?
- What do you want to do on builds.a.o that you can't now?
- And how can I help make all that happen?



Thanks!