Using Jenkins to Build Apache VHosts for GitHub Pull Requests on PHP Apps

Sam Moffatt
AccountBouncer
http://www.accountbouncer.com

October 23, 2014

#jenkinsconf

Overview

- About Me
- Background
- Setup
- Demo

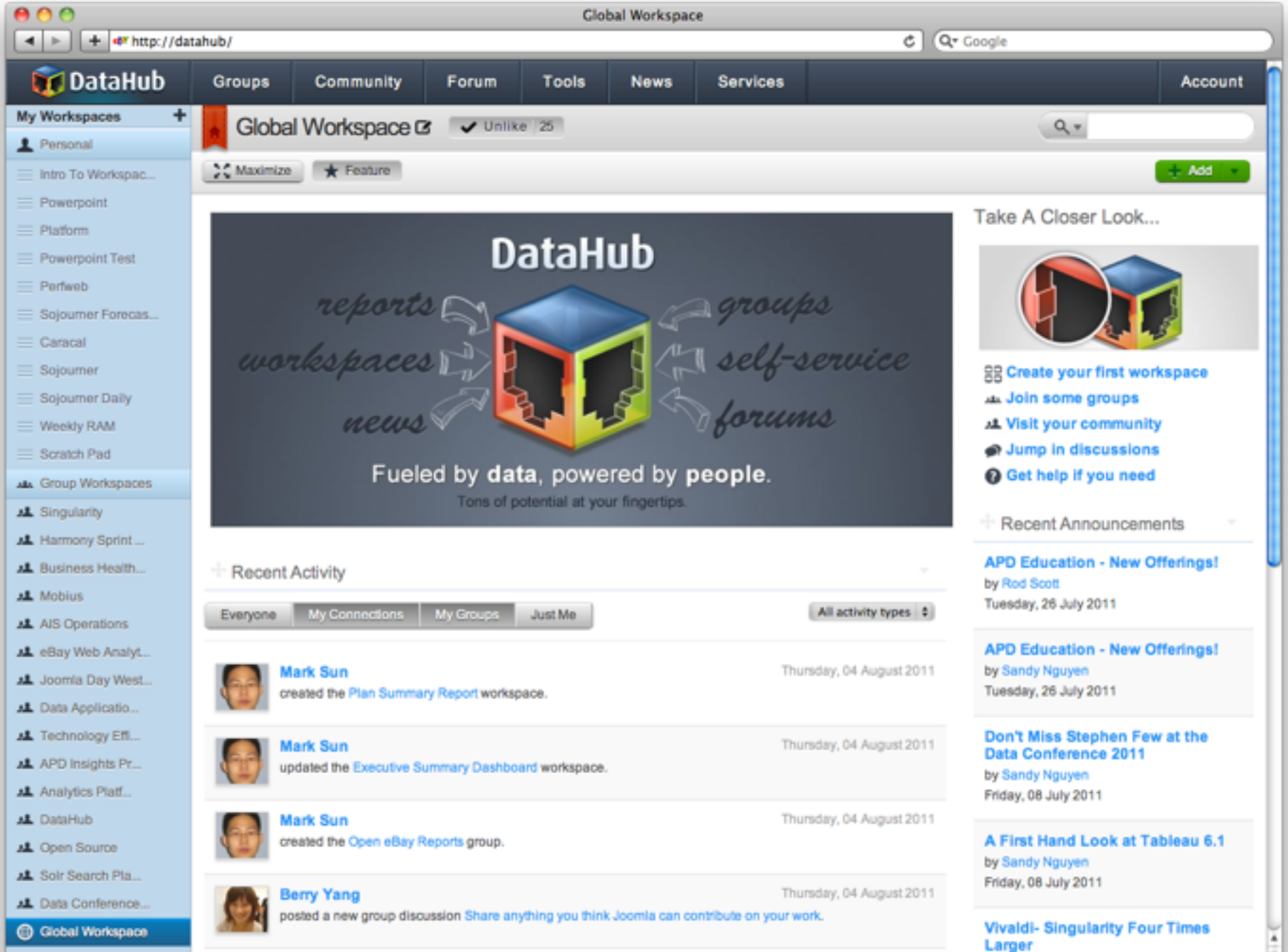
About Me

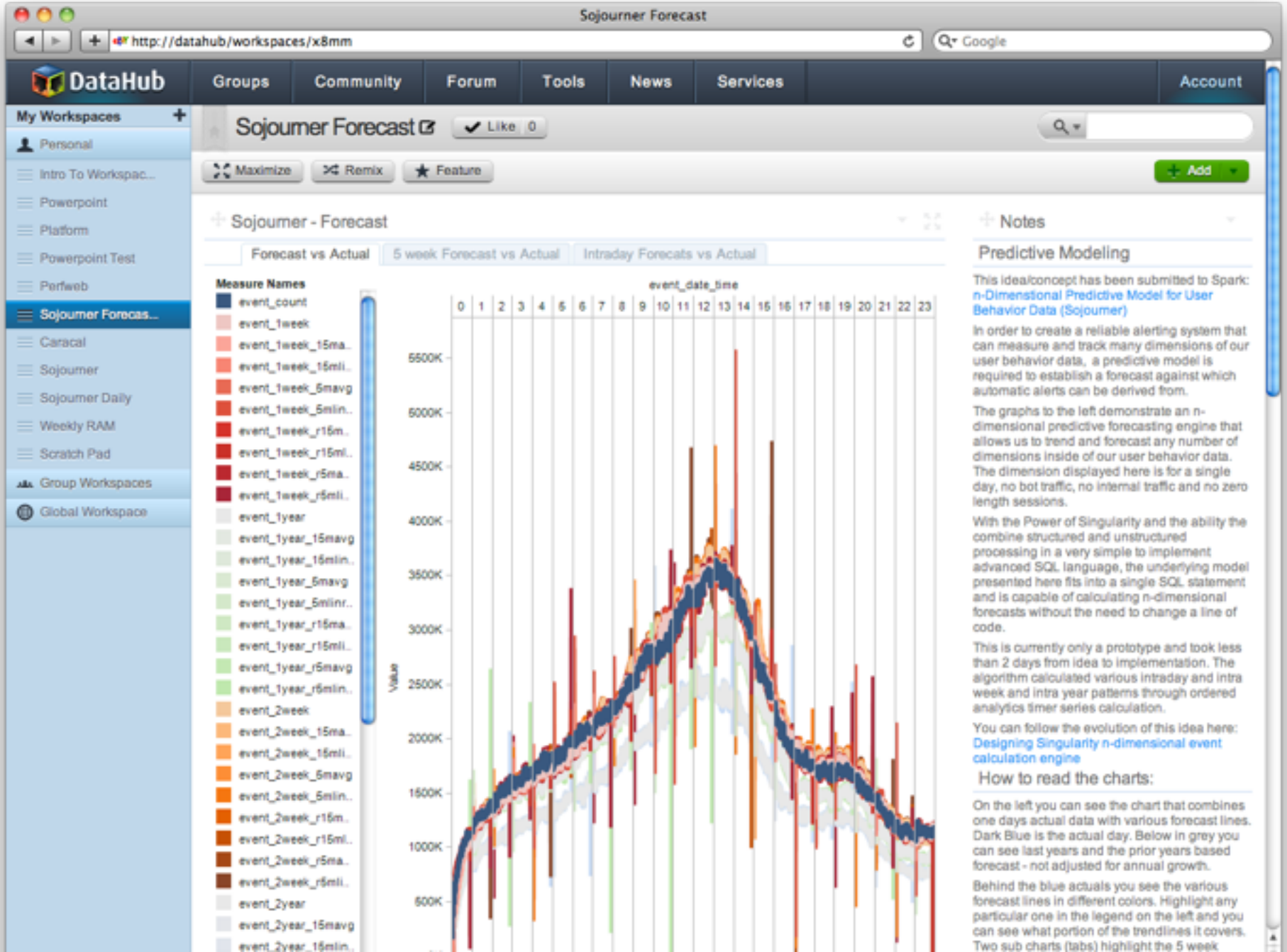
- Founder of AccountBouncer
- Varied roles in life
  - LinkedIn as Site Reliability Engineer
  - Sears as Systems Engineer
  - eBay as Software Engineer
  - USQ as Library Systems Co-ordinator
  - USQ as Analyst Programmer
  - TRC as Web Developer

Background

Joomla!™

Background


Global Workspace
http://datahub/
Google
DataHub
Groups
Community
Forum
Tools
News
Services
Account
My Workspaces
Personal
Intro To Workspac...
Powerpoint
Platform
Powerpoint Test
Perfweb
Sojourner Forecas...
Caracal
Sojourner
Sojourner Daily
Weekly RAM
Scratch Pad
Group Workspaces
Singularity
Harmony Sprint ...
Business Health...
Mobius
AIS Operations
eBay Web Analyt...
Joomla Day West...
Data Applicatio...
Technology Effi...
APD Insights Pr...
Analytics Platf...
DataHub
Open Source
Solr Search Pla...
Data Conference...
Global Workspace
Global Workspace
Unlike 25
Maximize
Feature
Add
DataHub
reports
workspaces
news
groups
self-service
forums
Fueled by data, powered by people.
Tons of potential at your fingertips.
Take A Closer Look...
Create your first workspace
Join some groups
Visit your community
Jump in discussions
Get help if you need
Recent Activity
Everyone
My Connections
My Groups
Just Me
All activity types
Mark Sun
created the Plan Summary Report workspace.
Thursday, 04 August 2011
Mark Sun
updated the Executive Summary Dashboard workspace.
Thursday, 04 August 2011
Mark Sun
created the Open eBay Reports group.
Thursday, 04 August 2011
Berry Yang
posted a new group discussion Share anything you think Joomla can contribute on your work.
Thursday, 04 August 2011
Recent Announcements
APD Education - New Offerings!
by Rod Scott
Tuesday, 26 July 2011
APD Education - New Offerings!
by Sandy Nguyen
Tuesday, 26 July 2011
Don't Miss Stephen Few at the Data Conference 2011
by Sandy Nguyen
Friday, 08 July 2011
A First Hand Look at Tableau 6.1
by Sandy Nguyen
Friday, 08 July 2011
Vivaldi- Singularity Four Times Larger


Sojourner Forecast
http://datahub/workspaces/x8mm
Google
DataHub
Groups
Community
Forum
Tools
News
Services
Account
My Workspaces
Personal
Intro To Workspac...
Powerpoint
Platform
Powerpoint Test
Perfweb
Sojourner Forecas...
Caracal
Sojourner
Sojourner Daily
Weekly RAM
Scratch Pad
Group Workspaces
Global Workspace
Sojourner Forecast
Like 0
Maximize
Remix
Feature
Add
Sojourner - Forecast
Forecast vs Actual
5 week Forecast vs Actual
Intraday Forecats vs Actual
Measure Names
event_count
event_1week
event_1year
event_2week
event_2year
event_date_time
Value
5500K
5000K
4500K
4000K
3500K
3000K
2500K
2000K
1500K
1000K
500K
Notes
Predictive Modeling
This idea/concept has been submitted to Spark: n-Dimensional Predictive Model for User Behavior Data (Sojourner)
In order to create a reliable alerting system that can measure and track many dimensions of our user behavior data, a predictive model is required to establish a forecast against which automatic alerts can be derived from.
The graphs to the left demonstrate an n-dimensional predictive forecasting engine that allows us to trend and forecast any number of dimensions inside of our user behavior data. The dimension displayed here is for a single day, no bot traffic, no internal traffic and no zero length sessions.
With the Power of Singularity and the ability the combine structured and unstructured processing in a very simple to implement advanced SQL language, the underlying model presented here fits into a single SQL statement and is capable of calculating n-dimensional forecasts without the need to change a line of code.
This is currently only a prototype and took less than 2 days from idea to implementation. The algorithm calculated various intraday and intra week and intra year patterns through ordered analytics timer series calculation.
You can follow the evolution of this idea here: Designing Singularity n-dimensional event calculation engine
How to read the charts:
On the left you can see the chart that combines one days actual data with various forecast lines. Dark Blue is the actual day. Below in grey you can see last years and the prior years based forecast - not adjusted for annual growth.
Behind the blue actuals you see the various forecast lines in different colors. Highlight any particular one in the legend on the left and you can see what portion of the trendlines it covers. Two sub charts (tabs) highlight the 5 week

Forum

http://datahub/forum/26-singularity-forum/1655-designing-singularity-n-dimensional-event-calculation-engine

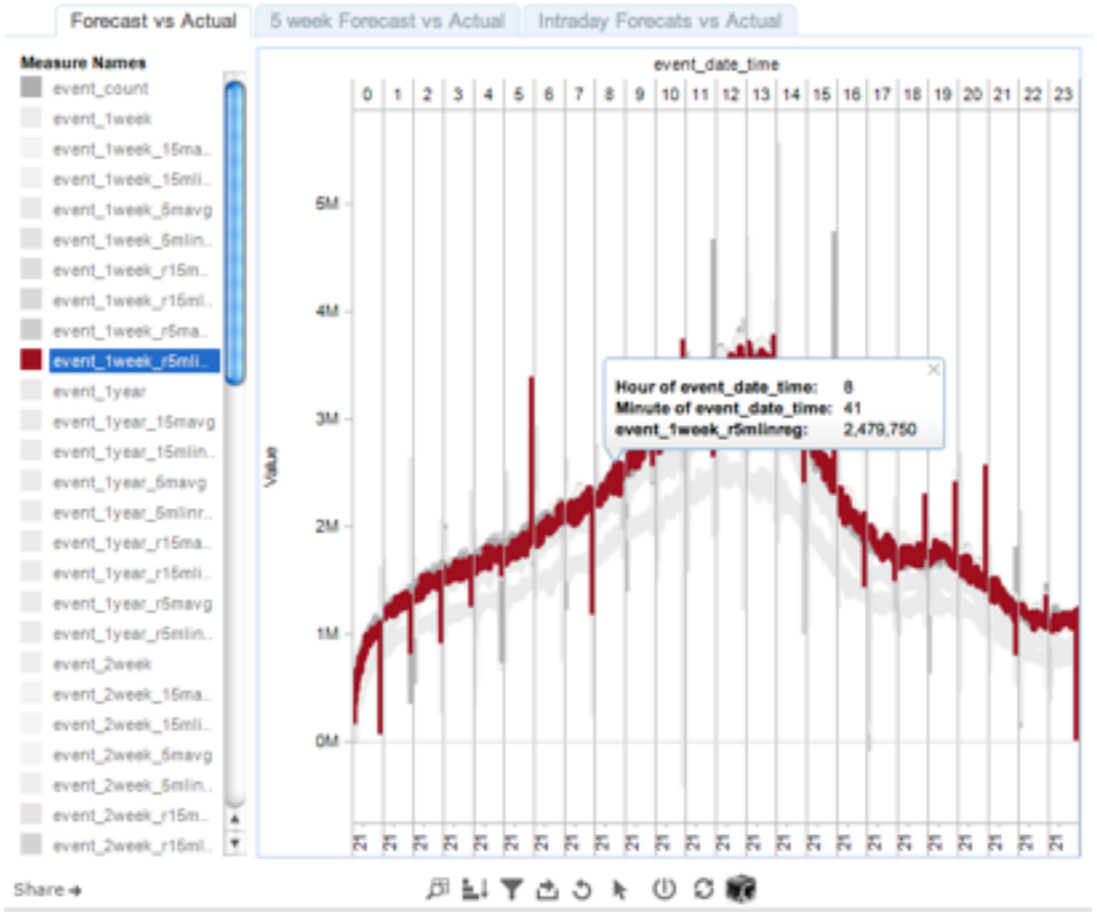
Twitter: @ratzesberger

Quick Reply | Reply | Quote | Edit | Moderate | Delete

Re: Designing Singularity n-dimensional event calculation engine 5 months, 1 week ago #1674

Oliver Ratzesberger

And finally here are some graphs demonstrating the basic algorithm. Only one dimension being shown at the moment: No bot, No zero length sessions, No internal traffic (bot=0&zero=0&internal=N) Report to moderator



You can see some of the linear regression artifacts at the top of the hour, caused by peaks in traffic. For the final model these

There were challenges…

- Significant UI changes caused surprises.
- Building out new user interaction flows was hard to do well.
- Team spread out between Austin, San Jose and Shanghai.
- Contention for QA review meant sometimes premature code check-ins.
- Hard to evolve ideas and get feedback.

…but also solutions!

- GitHub pull requests provided
  - a place for automated testing
  - communication about code changes
  - a better way for review

Solution

- GitHub Pull Request triggers:
  - automated tests (unit test, code quality)
  - builds dedicated environment
- Two extra features:
  - QA can run UI tests against pull requests
  - Product can review UX changes
- Also compatible with feature toggles!

SETUP

Setup

- Debian 7.6 “Wheezy”
  - bind9 (DNS Server)
  - Apache 2.2 + PHP 5.4 (module)
  - mod_vhost_alias OR
  - mod_vhost_hash_alias
- Jenkins 1.583
- GitHub

Setup: Bind 9

- Append to /etc/bind/named.local.conf:

```
zone "test" {
        type master;
        file "/etc/bind/test.zone";
};
```

Setup: Bind 9

- Create /etc/bind/test.zone:

```
test. IN      SOA     dns.test. admin.test. (
                      1130659718
                10800
                3600
                604800
                38400 )
test.           IN   NS  test.
                A    10.0.1.11
*.test.         IN   A    10.0.1.11
```

Setup: Apache

- Beginner way: mod_vhost_alias
  - Pros: standard module, easy to configure
  - Cons: doesn't play well with others vhost
- Intermediate way: mod_vhost_hash_alias
  - Pros: works with other vhost settings on same IP (needs to be last!)
  - Cons: 3rd party, requires directory structure including a hash

Setup: Apache

- Alternatives: mod_vhost_ldap/ mod_vhost_dbd
  - Pros: more options/flexibility
  - Cons: requires external server, DBD does query on every request, more dependencies

Setup: Apache mod_vhost_alias

```
UseCanonicalName Off
NameVirtualHost 10.0.1.11:80
LogFormat "%V %h %l %u %t \"%r\" %s %b" vcommon

<Directory /var/www/prtester>
        Options FollowSymLinks
        AllowOverride All
</Directory>

<VirtualHost 10.0.1.11:80>
        VirtualDocumentRoot /var/www/prtester/%1/web
        ErrorLog ${APACHE_LOG_DIR}/prtest.error.log
        CustomLog ${APACHE_LOG_DIR}/prtest.access.log vcommon
</VirtualHost>
```

Setup: Apache mod_vhost_hash_alias (Part 1)

HashType md5

HashEncoding hexa

HashSplit 2 2

HashLimit 4

HashDocumentRootPrefix /var/www/hash

HashDocumentRootSuffix htdocs

HashAddAliasPrefix www

Setup: Apache mod_vhost_hash_alias (Part 2)

```
<Directory /var/www/hash>
      Options FollowSymLinks
      AllowOverride All
</Directory>

<VirtualHost 10.0.1.10:80>
       DocumentRoot /var/www/hash
       HashEnable On
       ErrorLog ${APACHE_LOG_DIR}/prtest-hash.error.log
       LogFormat "%V %h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%
{User-agent}i\"" hashlog
       CustomLog ${APACHE_LOG_DIR}/prtest-hash.access.log hashlog
</VirtualHost>
```

Setup: Apache mod_vhost_hash_alias (All Together)

```
HashType md5
HashEncoding hexa
HashSplit 2 2
HashLimit 4
HashDocumentRootPrefix /var/www/hash
HashDocumentRootSuffix htdocs
HashAddAliasPrefix www

<Directory /var/www/hash>
        Options FollowSymLinks
        AllowOverride All
</Directory>

<VirtualHost 10.0.1.10:80>
      DocumentRoot /var/www/hash
      HashEnable On
      ErrorLog ${APACHE_LOG_DIR}/prtest-hash.error.log
      LogFormat "%V %h %l %u %t \"%r\" %>s %b \"%{Referer}i\" \"%{User-agent}i\"" hashlog
      CustomLog ${APACHE_LOG_DIR}/prtest-hash.access.log hashlog
</VirtualHost>
```

Setup: Jenkins System
GitHub Pull Request Builder

- Install "GitHub pull request builder plugin" via "Manage Jenkins" → "Manage Plugins"
- Navigate to "Manage Jenkins" → "Configure System" → "Github Pull Request Builder"

Setup: Jenkins System GitHub Pull Request Builder

- If you have GitHub Enterprise, change "GitHub server api URL" to your internal server.
- If you have an access token for GitHub you can put it in here, or click "Advanced" and fill in your username and password to have Jenkins create a token for you.
- Add your GitHub username to the "admin" list.

Setup: Jenkins Job GitHub Pull Request Builder

- Create a new “freestyle project”
- Under “SCM” select “Git” and then “Advanced…”, set:
  - Repository URL: Your GitHub repo
  - Refspec: “+refs/pull/*:refs/remotes/origin/pr/*”
  - Branches to build: “${sha1}”

Setup: Jenkins Job
GitHub Pull Request Builder

- Under "Build Triggers", select "GitHub Pull Request Builder" and click "Advanced":
  - **White list:** List of users who will trigger builds without admin approval.
  - **List of organisations:** List of organisations whose members will trigger builds without admin approval.
  - **Allow members of whitelisted organisations as admins:** All members can trigger builds for other users.

Setup: Jenkins Job
GitHub Pull Request Builder

- Under "Build Triggers", select "GitHub Pull Request Builder" and click "Advanced":
  - **Build every pull request automatically without asking (Dangerous!).**

Setup: Jenkins Job mod_vhost_alias

- We configured the path to be: /var/www/prtester/**%1**
- **%1** is the first part of the domain, so **pr1**.test will use /var/www/prtester/**pr1**
- So we need to create this directory structure.

Setup: Jenkins Job mod_vhost_alias

- Simple Shell Script:

```
TARGET=/var/www/prtester/pr${ghprbPullId} &&
rm -rf $TARGET &&
mkdir -p $TARGET/web &&
cp * $TARGET/web &&
chmod -R u+rX,g+rX,o+rX $TARGET
```

Setup: Jenkins Job mod_vhost_hash_alias

- We configured the path to be: **/var/www/hash**
- We set there to be hash limit of **4** and the split as **2 2**
- We set it to use **MD5** and **hexa**decimal encoding
- We have a suffix of **htdocs**

Jenkins: Job Setup
mod_vhost_hash_alias

- For **pr1.hashtest** we need the MD5: cb6c1a353e84ea0d3edd326e957aaafa
- Then we get the first four characters and split into two: **/cb/6c**
- We now need the following path structure: /var/www/hash/cb/6c/**pr1.hashtest**/htdocs
- Note: www.pr1.hashtest will work too!

Jenkins: Job Setup
mod_vhost_hash_alias

```
DOMAIN=pr${ghprbPullId}.hashtest
HASH=$(echo -n $DOMAIN | md5sum |
    sed -e's#^\(..\)\(..\).*#/\1/\2/#')
ROOT=/var/www/hash
TARGET=$ROOT/$HASH/$DOMAIN
rm -rf $TARGET &&
mkdir -p $TARGET/htdocs &&
cp * $TARGET/htdocs &&
chmod -R u+rX,g+rX,o+rX $ROOT
```

DEMO

Pitfalls

- With GitHub Enterprise:
  - if you have custom SSL certificates, make sure that Jenkins trusts them
  - make sure that the API is set up properly!
- If you have more than one pull request builder job for the same repo then GitHub will likely report the wrong status!

Further Reading

- GitHub Pull Request Builder plugin page: https://wiki.jenkins-ci.org/display/JENKINS/GitHub+pull+request+builder+plugin

Thank You To Our Sponsors

Platinum

CloudBees

Gold

XebiaLabs
keynote
zend
BlazeMeter
THE LOAD TESTING CLOUD
ravello systems
elasticbox
JFrog
perfecto mobile
black diamond SOFTWARE

Silver

SOASTA
YAHOO!
Pivotal
SERENA
choose digital
sauce LABS
opensourcearchitect

Corporate

cloudera