

# Getting Involved in World-class Software Engineering : Tips and Tricks to Join Apache Open Source Community

Evans Ye  
HadoopCon 2014  
Taiwan



# Who am I

- Evans Ye
  - Developer, SPN, Trend Micro
  - Worked on hadoop ecosystem since 2012
  - Contributor for Apache Bigtop
  - TWHUG 2013Q4 & 2014Q1



# Agenda

- Why you should join the community
- My journey to Apache Bigtop
- tips and tricks to join the community
- Apache Bigtop status quo
- Feature preview on recent development

Why you should join open  
source community?

學習上乘的coding技巧



無私奉獻的精神!



Well, Part of...



# The Real Problem

這個軟體雖然有A功能，  
卻沒有B功能耶。



超爛DER



這個軟體Bug好多，  
好難用。



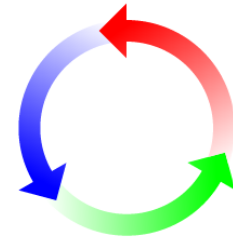
超爛DER





You can make it better  
by join the community

- Do not have the feature we need?
  - Contribute
    - Get you're desired feature keep maintained by the community(in next release)
- Too buggy?
  - improve its quality
    - less troubleshooting
    - less operational effort (apply workarounds...)

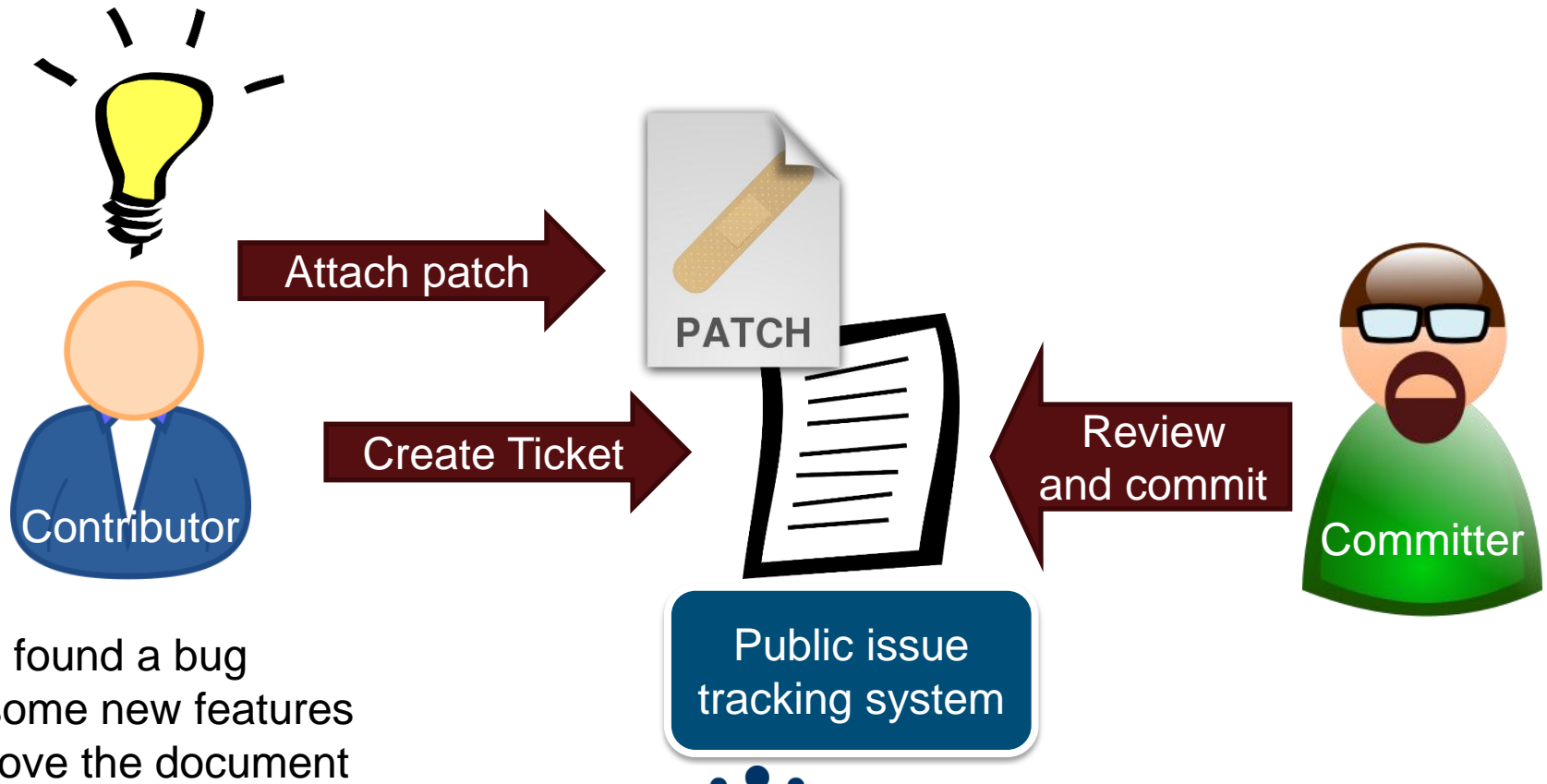


# Extra bonus for individuals

- Expert's free code review
- Embrace new technology timely just like a geek
- Personal reputation
- Improve English reading/writing skill by communicate with native speakers

# Okay, How to contribute?

# Basic working flow



The Apache  
Software Foundation  
<http://www.apache.org/>



# 真實案例： My Journey to Apache Bigtop




# Apache Bigtop is a project for...

- Packaging
  - Packages Hadoop ecosystem RPMs, DEBs
- Testing
  - integration/interoperability testing framework
- Virtualization
  - **vagrant** based hadoop provisioning
- Deployment
  - A set of **puppet** recipes for deployment covers most of the hadoop ecosystem

# Started on Dec. 2013

- Create my first ticket:

 Bigtop / BIGTOP-1163

puppet apply shows db init error if the oozie server is already running

- With a beautiful patch attached

## Attachments

 BIGTOP-1163.1.patch	0.9 kB	17/Dec/13 02:30
---	--------	-----------------

- Thinking of that will probably get committed soon😊

- And nobody gives a \_\_\_\_\_ about me



– Maybe I didn't describe the issue clearly...

# 再接再厲!

- Create another ticket:



Bigtop / BIGTOP-1167

got "Permission denied" when creating vagrant home folder in provision.sh

- Also a nice patch attached

## Attachments



BIGTOP-1167.1.patch

2 kB

22/Dec/13 02:09



BIGTOP-1167.2.patch

2 kB

22/Dec/13 02:10

# 4 days later...

- Some one jump in and tested the patch

✓  jay vyas added a comment - 26/Dec/13 09:01

2) I tested Patch one and **IT WORKS** Thanks Can a commiter weigh in and push this through 😊 I've tested it and it definetly fixes the error.. maybe [Peter Linnell](#) ?

✓  Roman Shaposhnik added a comment - 26/Dec/13 10:05

+1 on patch #1 and committed. Thanks for investigating!

COMMITTED!

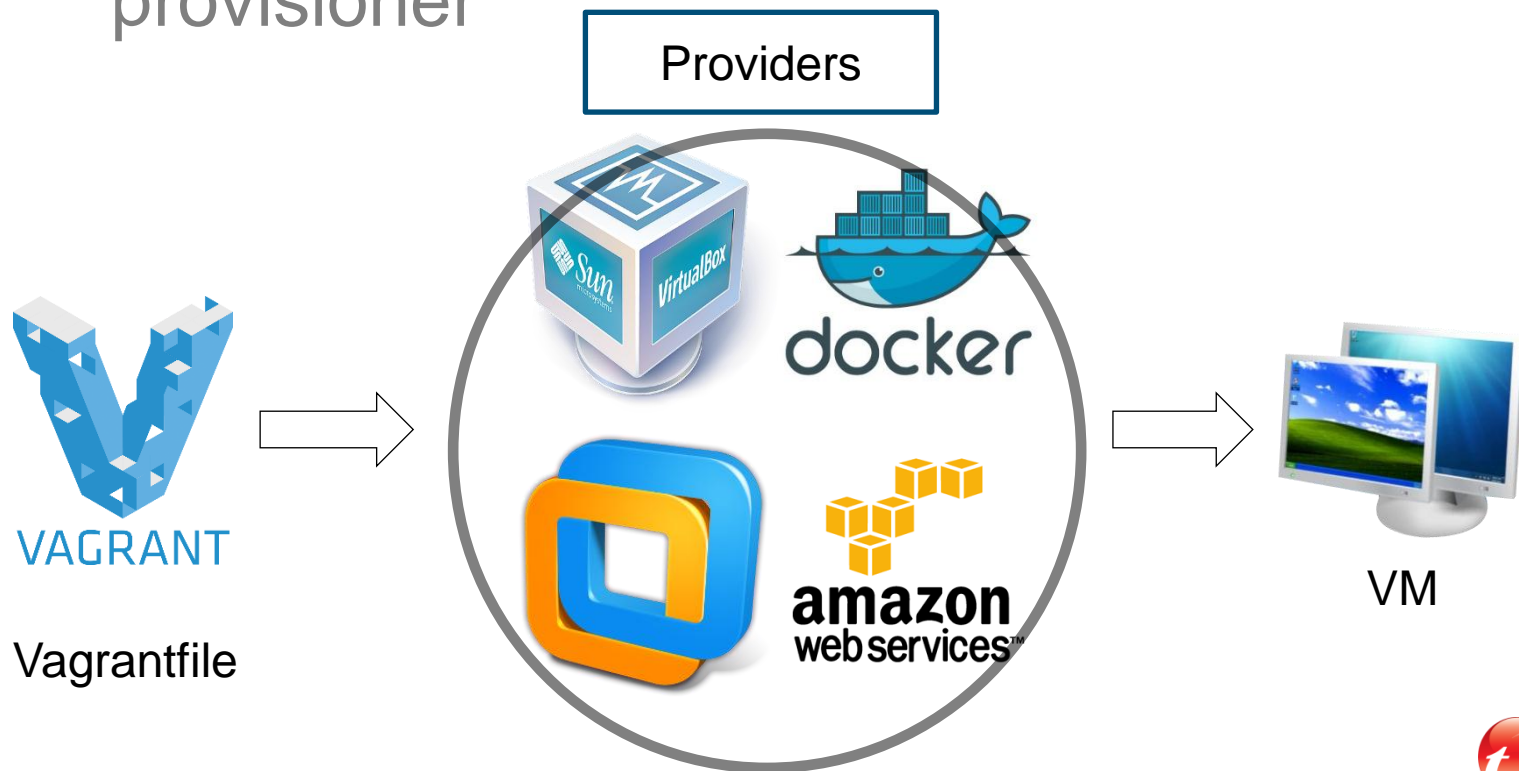
媽！  
我終於對這世界有了  
(微不足道的)貢獻



# Other Developed Features

# Quick intro. to Vagrant

- A open source **command line** VM(iaaS) provisioning tool
- Support shell, chef, puppet, ansible, etc provisioner



# Quick intro. to Puppet

- A IT automation tool to help system administrators automate the many repetitive tasks
- Suitable for deployment and configuration management
- You only need to define the desired state

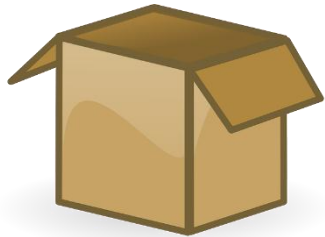
## Syntax

```
# A resource declaration:  
file { '/etc/passwd':  
    ensure => file,  
    owner  => 'root',  
    group  => 'root',  
    mode   => '0600',  
}
```

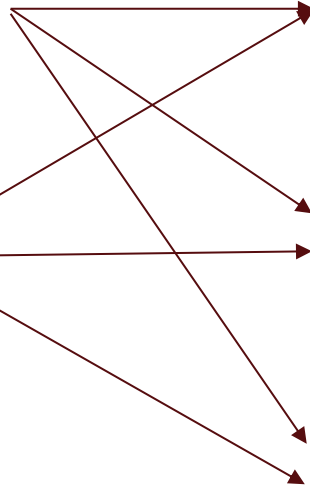
# Hadoop cluster deploying from zero



VAGRANT



Vagrant Box  
(Centos, Ubuntu,...)



Virtualbox  
VMs



APACHE  
HBASE



\$ vagrant up  
DONE!

# Usage

- System level integration test for hadoop ecosystem
- Easily to get a development environment
- Reproduce/troubleshooting hadoop cluster issues
  - Namenode, Jobtracker, ResourceManager HA
- Bigtop puppet recipes testing
- Bigtop github:
  - [bigtop/bigtop-deploy/vm/vagrant-puppet/README.md](https://github.com/bigtop/bigtop-deploy/tree/master/vm/vagrant-puppet)

# Tips and Tricks to Join Open Source Community



# Tip 1 – niche market


- Pick up a specific feature no one (or few people) interested.
  - Bigtop: vagrant based hadoop provisioning
  - HBase: canary tool
- Pros:
  - less “competitors”
  - New features are relatively easy to be improved
- Cons:
  - The feature might not be accepted by community

## Tip 2 – claim tickets

- Users bug report
- Some Sr. community members will create lots of tickets
- But they might not have time to resolve it.
- Ask politely first, don't duplicate effort

# Example – BIGTOP-1171

如果你還沒做，我可以試試看  
如果你正在做，我可以幫忙測試

✓  Evans Ye added a comment - 04/Jan/14 02:17

Hi Jay, I've done some work which is highly related to this jira, so if you're not yet started to deal with this, would you mind me to provide a concept patch for this? If you're going to do it, I can still help on the test 😊

我還沒做，但我想可以這樣做：  
step1, step2,...  
我願意幫忙測你的patch

✓  jay vyas added a comment - 04/J

Hi Evans. I dont have a patch. Herei

1) use the native vagrant puppet sup  
recipes in the bigtop/ directory can b

2) Put the configuration file defining t  
<https://issues.apache.org/jira/browse>  
the cluster configuration, and then "v

But i dont have a patch, would love t



(obviously). Maybe the puppet-deploy  
ed folder?


atched by roman here  
hared folder, so that host can update  
stem accordingly.

w 😊 thanks!

# Tip 3 – socializing with project members

- Knowing who can help is important
  - committers or active members
- Build up your personal reputation
  - your patch always has good quality
  - easy to test, easy to commit

# Examples

▼  jay vyas added a comment - 26/Dec/13 10:44

got it . thanks.

+1 for this patch ! lets get it in soon - its tested by 2 people and essential for proper provisioning .

Call for committer's help to commit this in.

▼  Roman Shaposhnik added a comment - 14/Jan/14 14:22

Peter Linnell Peter, ping! Any chance you can help commit this?

Ask other committer to help to commit the patch


# Examples

✓  Roman Shaposhnik added a comment - 23/Jun/14 11:50

+1. jay vyas can you help commit this?

✓  Roman Shaposhnik added a comment - 26/Dec/13 10:05

+1 on patch #1 and committed. Thanks for investigating!

✓  Roman Shaposhnik added a comment - 14/Jan/14 14:22

And just in case +1 from my side 😊

✓  Roman Shaposhnik added a comment - 23/Jun/14 11:51

+1. jay vyas any chance **you** can test this with the other vagrant patch and commit both? Would be very much appreciated!

Get project lead's +1 is a BIG endorsement to your patch



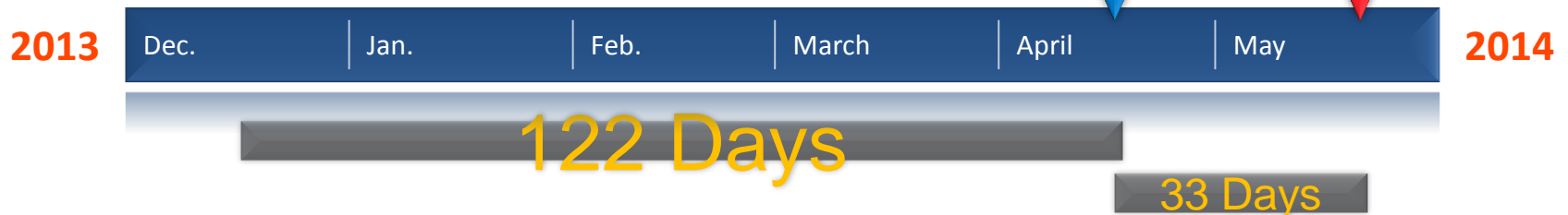
## Tip 4 – be patient

- Your patch may be ignored since they're too busy
  - Wait for a while, probably a week
  - If no one responding on your patch, you can ask for someone's opinion
  - Add more description to better describe the patch
  - Move on to other ticket, don't wait
    - Let it go~Let it go~

# Bigtop 1163, my first ticket

Add more description

Got responded by a committer and soon get committed



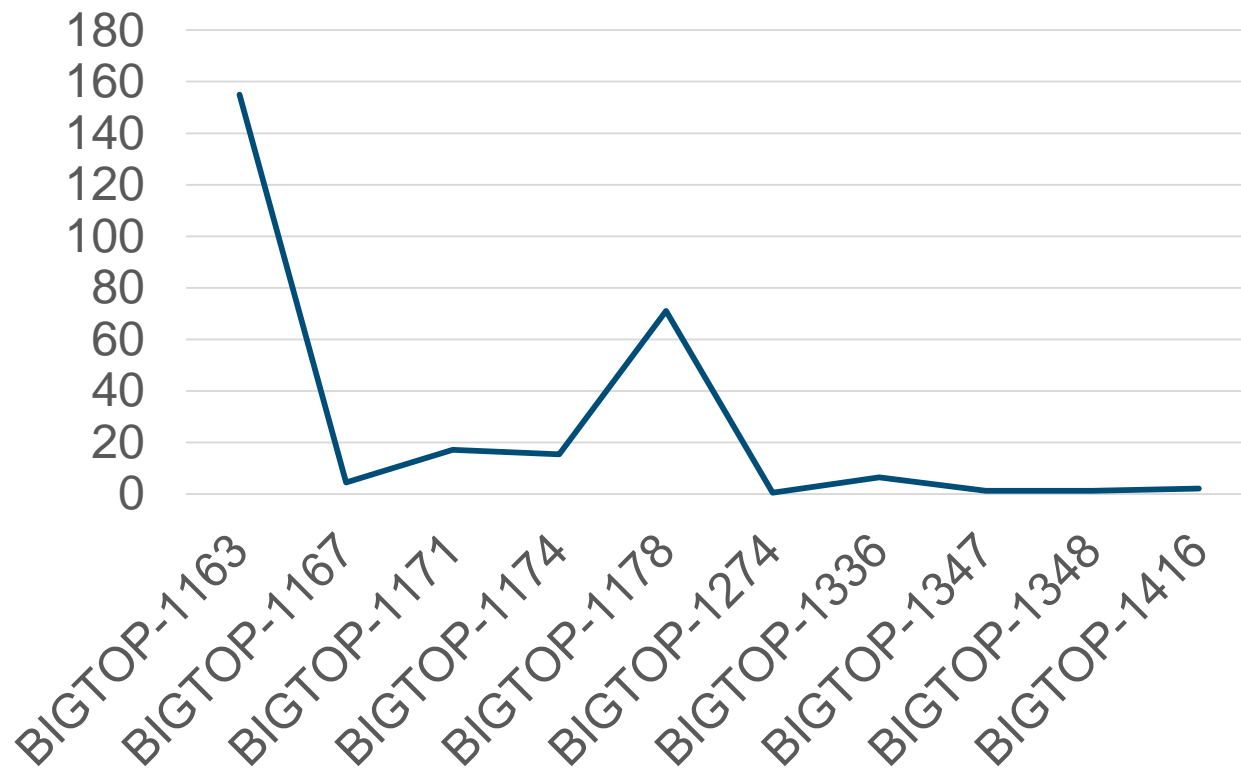
- 155 days later, finally...

✓  Konstantin Boudnik added a comment - 21/May/14 02:21

Committed, Thanks Evans!

# Average committed days

- It could be fast if community members knows well what you're doing
  - good doc, clear code, join discussion

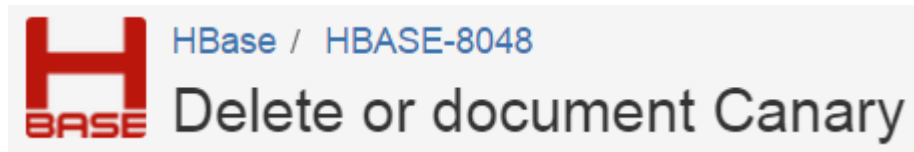


## Tip 5 – Join discussion

- Other open source members might have different opinion,  
thus you need to provide your different opinion as well.


# Example

- A community member creates a ticket:




Our HBase contributor

Delete Canary?  
No way!

▼  [takeshi.miao](#) added a comment - 09/Dec/13 16:14

I am currently working on the docs

▼  [Nick Dimiduk](#) added a comment - 10/Dec/13 06:33


Fantastic. Thanks for filling in the gaps, [takeshi.miao](#).



# Tip 6 – Join mailing list

## Mailing Lists

### Bigtop User List

If you use Bigtop, please subscribe to the Bigtop user mailing list.



The Bigtop user mailing list is : [user@bigtop.apache.org](mailto:user@bigtop.apache.org)  .

- [Subscribe to List](#) 
- [Unsubscribe from List](#) 
- [Browse archives](#) 

### Bigtop Developer List


If you'd like to contribute to Bigtop, please subscribe to the Bigtop developer mailing list.


The Bigtop developer mailing list is : [dev@bigtop.apache.org](mailto:dev@bigtop.apache.org)  .

- [Subscribe to List](#) 
- [Unsubscribe from List](#) 
- [Browse archives](#) 

### Bigtop Commit List

If you'd like to see changes made in Bigtop's version control system then subscribe to the Bigtop commit mailing list.

The Bigtop commit mailing list is : [commits@bigtop.apache.org](mailto:commits@bigtop.apache.org)  .

- [Subscribe to List](#) 
- [Unsubscribe from List](#) 
- [Browse archives](#) 

## Tip 7 – Know how to use Jira

- Carefully select the priority when creating tickets
  - Don't grant this in order to get more visibility
- remember to flip the status when patch attached



- Use Jira's mention function when you'd like to get someone's attention.
  - [`~evans_ye`], any thoughts?

# Summary



# Tips to Join Open Source Community

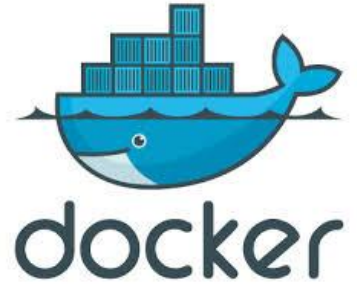
- niche market
- claim tickets
- socializing with project members
- be patient
- Join discussion
- Join mailing list
- Know how to use Jira

# Apache Bigtop Status Quo

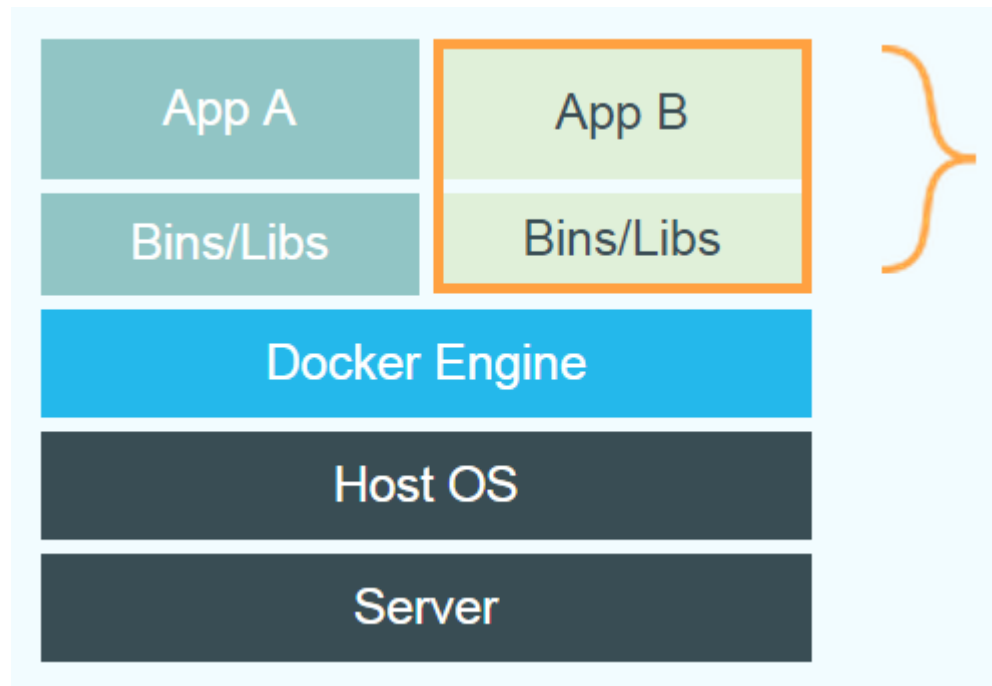
# Apache Bigtop Status Quo

- Bigtop 0.8.0 is going to release
  - Hadoop 2.4.1, HBase 0.98.5, Pig 0.12.1...
- Deprecated Makefile build system
  - switch to gradle
- Moving to JDK 7
- Docker based continuous integration
  - build, test, deploy on multiple platform
- Dropping Whirr...

# Docker?



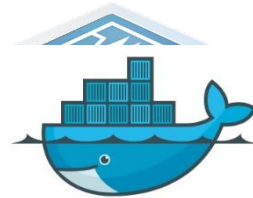
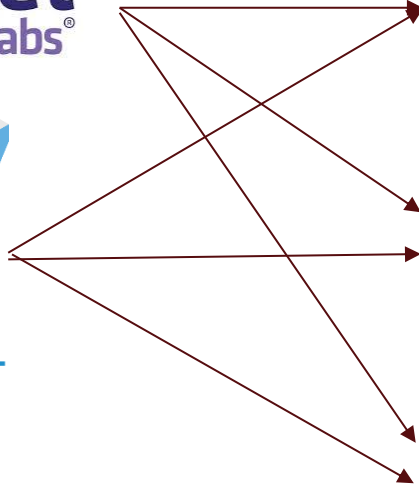
- Linux Container
  - high level view: lightweight VM
  - really fast, suitable for dev/testing



<https://www.docker.com/whatisdocker/>

# Feature preview to recent Development: Docker based hadoop provisioning

# Docker based hadoop provisioning



docker



docker



docker



APACHE  
**HBASE**



to docker image

# Lighting-fast hadoop provisioning

# LIVE DEMO

Provision a 3 node Hadoop cluster  
on Ultrabook in a minute!



# Q&A

We're hiring!