



How we ship

Steve Loughran

stevel at hortonworks.com

@steveloughran

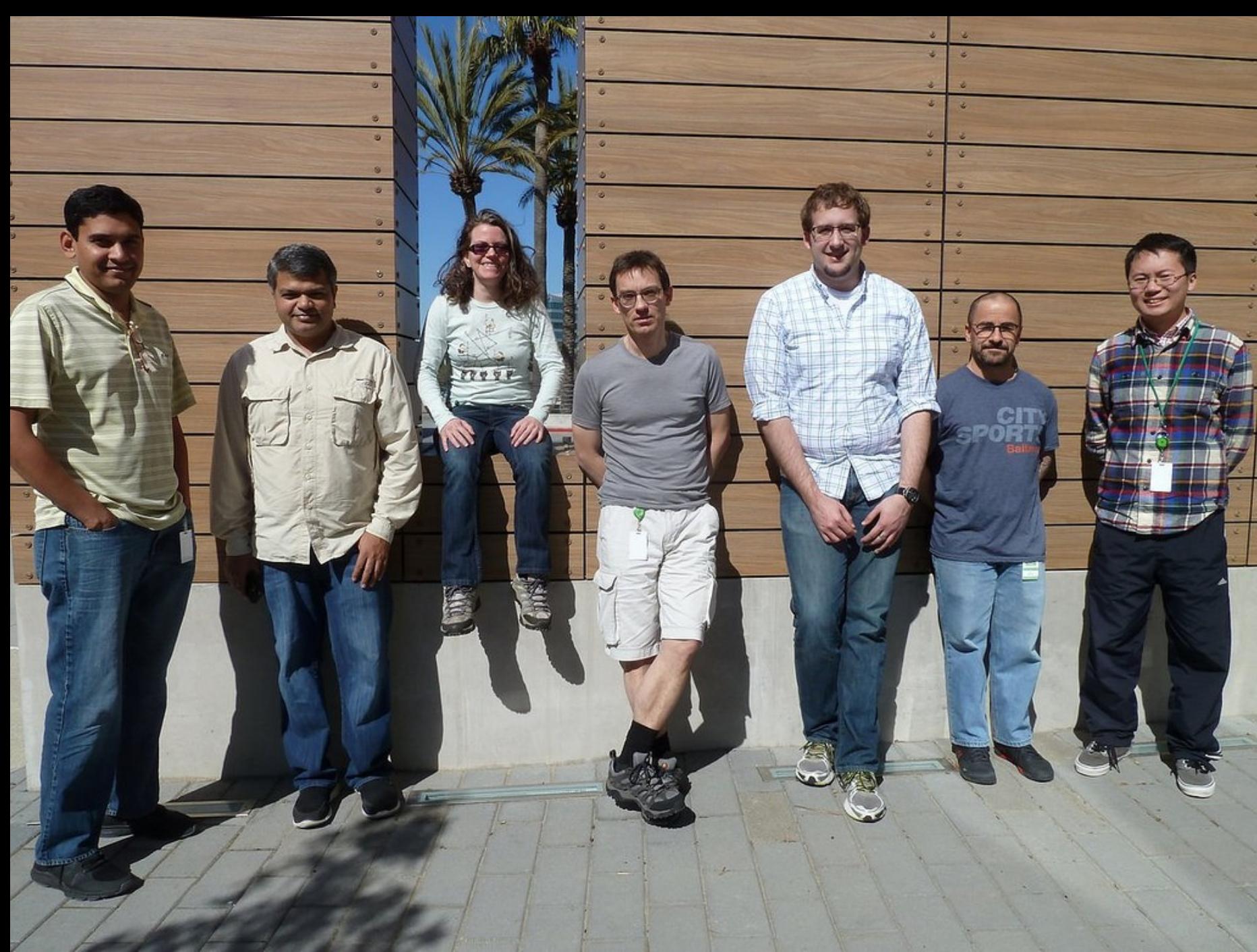
2015





Robert
Marlon
Jordi
Ivan
-06-2002

Amber
Jerico
Iris
Jordu



Some of largest filesystems ever



Some of the largest applications



Hortonworks Data Platform

GOVERNANCE

Data Workflow, Lifecycle & Governance

Falcon
Sqoop
Flume
Kafka
NFS
WebHDFS

BATCH, INTERACTIVE & REAL-TIME DATA ACCESS

Script

Pig

Tez

SQL

Hive

Tez

Java Scala

Cascading

Tez

NoSQL

HBase
Accumulo

Slider

Stream

Storm

Slider

In-Memory

Spark

Search

Solr

YARN Engines

Others

YARN READY

YARN: Data Operating System

(Cluster Resource Management)

HDFS

(Hadoop Distributed File System)

Linux

Windows

Deployment Choice

On-Premises

Cloud



SECURITY

Authentication
Authorization
Audit
Data Protection

Storage: HDFS
Resources: YARN
Access: Hive
Pipeline: Falcon
Cluster: Ranger
Cluster: Knox

OPERATIONS

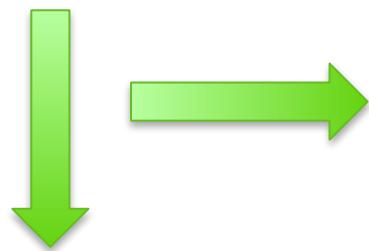
Provision, Manage & Monitor

Ambari
Zookeeper

Scheduling

Oozie

Product
Management



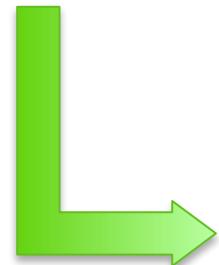
QE



Research &
Development



Jenkins



Git at Apache



Jenkins

JUnit



Git at Apache

[YARN-913] Umbrella: Add x

<https://issues.apache.org/jira/browse/YARN-913>

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

Hadoop YARN / YARN-913

Umbrella: Add a way to register long-lived services in a YARN cluster

[Edit](#) [Comment](#) [Assign](#) [More](#) [Stop Progress](#) [Submit Patch](#) [Resolve Issue](#)

Details

Type:	New Feature	Status:	IN PROGRESS	People
Priority:	Major	Resolution:	Unresolved	Assignee:
Affects Version/s:	2.5.0, 2.4.1	Fix Version/s:	None	Steve Loughran
Component/s:	api , resourcemanager			Reporter:
Labels:	None			Steve Loughran
Target Version/s:	2.8.0			Votes:

Description

In a YARN cluster you can't predict where services will come up -or on what ports. The services need to work those things out as they come up and then publish them somewhere.

Applications need to be able to find the service instance they are to bond to -and not any others in the cluster.

Some kind of service registry -in the RM, in ZK, could do this. If the RM held the write access to the ZK nodes, it would be more secure than having apps register with ZK themselves.

Attachments

2014-09-03_Proposed_YARN_Service_Registry.pdf	337 kB	03/Sep/14 09:53
2014-09-08_YARN_Service_Registry.pdf	327 kB	08/Sep/14 08:43
RegistrationServiceDetails.txt	2 kB	09/Oct/13 10:45
YARN-913-001.patch	215 kB	03/Sep/14 11:21
YARN-913-002.patch	204 kB	03/Sep/14 11:22

People

Assignee: Steve Loughran

Reporter: Steve Loughran

Votes: 4

Watchers: 59 [Stop watching this issue](#)

Dates

Created: 10/Jul/13 11:58

Updated: 01/May/15 05:53

Time Tracking

Estimated: 1.5h

Remaining: 0h

Tests

- **Unit:** classic JUnit tests
- **Minicluster:** in-VM HDFS, YARN, KDC + app
- **Installation (& Management)**
- **Integration:** applications in real clusters
- **Fault Injection:** triggering failures & verify continuity of service
- **Performance & Load:** JPerf & monitoring tools
- **Security:** auth access granted, unauthed access denied, delegation via tokens; token expiry handled...



Integrated Chaos Monkey

Hadoop [Jenkins] 2.3-maint [Jenkins]

<https://builds.apache.org/view/H-L/view/Hadoop/>

Jenkins H-L Hadoop [ENABLE AUTO REFRESH](#)

[People](#)
[Build History](#)
[Project Relationship](#)
[Check File Fingerprint](#)
[Leader board](#)
[Dependency Graph](#)

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

This is a public build and test server for [projects](#) of the [Apache Software Foundation](#). All times on this server are UTC.
See the [Jenkins wiki page](#) for more information about this service.



For more information on Hadoop, see: <http://hadoop.apache.org/>

Build Queue (116)			
kafka-trunk-git-pr	kafka-trunk-git-pr	sling-trunk-1.8	sling-trunk-1.7
S	W	Name	Last Success
		Hadoop-Hdfs-trunk-Java8	1 mo 28 days - #222
		Hadoop-branch2	3 mo 5 days - #765
		Hadoop-Common-2-Commit	1 day 4 hr - #4607
		Hadoop-Common-trunk	11 days - #1580
		Hadoop-common-trunk-Java8	11 days - #279
		Hadoop-Common-trunk-win	N/A
		Hadoop-fs-encryption-nightly	N/A
		Hadoop-HDFS-7285-Merge	6 hr 17 min - #91
		Hadoop-HDFS-7285-nightly	6 hr 12 min - #223
		Hadoop-HDFS-7285-REBASE	6 hr 13 min - #4
		Hadoop-Hdfs-7966	11 days - #42
		Hadoop-Hdfs-HBranch-build	N/A

 New Item People Build History Edit View Delete View Project Relationship Check File Fingerprint Manage Jenkins Credentials My Views Job Config History Jenkins 100K Job Priorities

Build Queue

No builds in the queue.

Build Executor Status

master

- 1 Idle
- 2 Idle
- 3 Idle
- 4 Idle
- 5 Idle

Views: openstack

S	W	Name ↓	Last Success	Last Failure	Last Duration
		analyze-launch-instances-logs	2 hr 23 min - #148	N/A	3 min 38 sec
		base-setup-instances	24 min - #60153: ambari-heavyservices-ha-rhe6-9478-split7re2	2 days 20 hr - #59083: os-u12-brdtys-devdeploy	52 sec
		create_openstack_image	4 mo 10 days - #6	4 mo 13 days - #1	6.6 sec
		create_openstack_images	3 mo 28 days - sdevineni-sp1	3 mo 25 days - sdevineni-cen5	18 min
		get-openstack-vm-list	29 min - hw-qe	2 min 0 sec - hw-qe	4.2 sec
		hw-dev-auto-terminate-expired-openstack-instances	25 min - hw-dev	10 min - hw-dev	6.4 sec
		hw-dev-cleanup-openstack-volumes	1 mo 7 days - hw-dev	N/A	4.2 sec
		hw-qe-auto-terminate-expired-openstack-instances	20 min - hw-qe	5 min 32 sec - hw-qe	21 sec
		hw-qe-cleanup-openstack-volumes	1 mo 7 days - hw-qe	N/A	6.3 sec
		Nightly-Start-EC2-Run-HDP	29 min - #65276: ambari-heavyservices-ha-rhe6-9478-split7re2	2 min 29 sec - #65282: os-r6-dbtest-ranger-22r	4 min 49 sec
		RE_smoke_test	3 mo 21 days - sdevineni-ub12	3 mo 19 days - sdevineni-2	16 min
		Start-Openstack-Instances	2 days 22 hr - #454: pt2	3 days 18 hr - #448: pramod1	8 min 15 sec
		stop-openstack-instances	38 min - os-r6-isbkws-pig-1	3 min 37 sec - ambari-heavyservices-ha-rhe6-9478-split7re2	5.5 sec

Double Click on entry to copy password

Project Start-Openstack-Instances

This build requires parameters:

CLUSTER

Give a name to your cluster!

USERNAME

Enter the username associated with this instance. This is required. THIS OPTION IS NOT IN USE YET

AMI

RHEL5
RHEL6
RHEL7
ORACLE5
ORACLE6
DEBIAN6
DEBIAN7
UBUNTU12
UBUNTU14
TDSLES11.1
SLES11.3

I that you want to use , add it here. Make sure you know what you are doing in I will be ignored.

age you want use?

want the cluster to live.

no

whether to shutdown the cluster after the tests are done running.

who to email the test results, use an individual email as group emails are having issues.

openstack

yes

qe-launch-on-openstack

Show nodes

 Back to Dashboard

 Status

 Changes

 Workspace

 Build with Parameters

 Delete Project

 Configure

 Rebuild Last

 Set Next Build Number

 Job Config History

Build History

trend 

 #454: pt2



Aug 14, 2015 12:55 PM

 #453: arpit-d7-2



Aug 14, 2015 11:41 AM

 #452: arpit-d7



Aug 14, 2015 11:34 AM

 #451: InstallDoc_Murali



Aug 14, 2015 5:08 AM

 #450: installDoc2_3



Aug 14, 2015 5:01 AM

 #449: pramod1



Aug 13, 2015 5:01 PM

 #448: pramod1



Speed



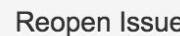
Kerberos is the path to Madness





Bug DB / BUG-38072

Spark History Server returns HTTP ERROR 401 on Ranger+TDE cluster

[Edit](#)[Comment](#)[Assign](#)[More ▾](#)[Close Issue](#)[Reopen Issue](#)

Details

Type:	<input checked="" type="checkbox"/> Bug	Status:	VERIFIED (View Workflow)
Priority:	<input checked="" type="checkbox"/> Blocker	Resolution:	Fixed
Affects Version/s:	Dal	Fix Version/s:	Dal
Component/s:	Spark		
Labels:	dal-qe-committed system_test tde		
Environment:	secure + XA + TDE		

[Field Tab](#)[Project Planning](#)

OS:

N/A

Security:

Secure

Ranger:

On

People

Assignee:



Steve Loughran

Reporter:



Yesha Vora

QE Assignee:

Yesha Vora

QE Group Lead:

Ramya Sunil

Votes:



0 Vote for this issue

Watchers:



6 Stop watching this issue

Description

Spark History server throws "HTTP ERROR 401" on cluster with secure+TDE+XA

Logs available being populated to Spark ATC web service "http://<2.52.1.54.55>/spark"



HDP: ASF + final patches

Fixes



Features

404



ORACLE®

Testing is a big data problem

