

# คู่มือติดตั้ง Hortonworks Data Platform (HDP) 3 + Hue + FreelPA

Cluster Kit Co., Ltd. January 31, 2024 เอกสารคู่มือติดตั้ง HDP+Hue+FreeIPA ฉบับนี้ เกิดขึ้นจากความตั้งใจที่จะแบ่งปันวิธีการติดตั้งระบบ Hadoop ที่มีหน้าเว็บ UI พร้อมทั้งการพิสูจน์ตัวตนกับ LDAP ตามแบบที่ใช้งานกันทั่วไป แม้ HDP จะไม่ได้ออกรุ่น ใหม่แล้ว แต่ของเก่าที่มีอยู่ยังคงใช้งานได้ ซึ่งมั่นใจว่ายังเป็นประโยชน์สำหรับผู้ที่ต้องการประมวลข้อมูลขนาดใหญ่ อีกเวลาพอสมควร เพียงพอที่จะรอให้ Ambari รุ่นใหม่และ Apache Bigtop พัฒนาจนมีระบบความปลอดภัย แบบที่ตัวเก่ามีอย่าง Ranger ที่สามารถกำหนดสิทธิ์การเข้าถึงข้อมูลของผู้ใช้งานได้

เอกสารนี้ช่วยกันเขียนโดยทีมงานของคลัสเตอร์คิท มีคุณสมภพ สุวรรณประดิษฐ์ คุณกิตติรักษ์ ม่วงมิ่งสุข เป็นผู้เขียน และคุณนันท์นภัส ม่วงมิ่งสุข เป็นผู้ตรวจทาน หวังเป็นอย่างยิ่งว่าจะเป็นประโยชน์ตามสมควร

พวกเราคลัสเตอร์คิท



1	บทน้ำ	5
2	การเตรียมระบบก่อนการติดตั้ง	5
3	FreeIPA	6
	3.1 ติดตั้ง FreelPA server	6
	3.2 ติดตั้ง FreeIPA Client	9
	3.3 LDAP binding account	9
4	Apache Ambari	
	4.1 ติดตั้ง Ambari	
	4.2 ติดตั้ง MariaDB	10
	4.3 ติดตั้ง MySQL JDBC	10
	4.4 สร้าง Database สำหรับ Ambari และผองเพื่อน	11
	4.5 คอนฟิก ambari-server	12
5	ติดตั้ง Hortonwork Data Platform (HDP)	14
	5.1 ติดตั้ง HDP ผ่าน Apache ambari	14
6	การทำ High Availability	21
	6.1 การทำ Namenode HA	21
	6.2 Enable ResourceManager HA Wizard	26
7	Hue	29
	7.1 ติดตั้ง Hue	29
	7.2 ปรับแต่ง Hue configuration file (/root/hue/conf/hue.ini)	
	7.3 ย้ายฐานข้อมูลมาใช้ MariaDB	
	7.4 กำหนด proxy user hadoop สำหรับ Hue	
	7.5 กำหนด proxy user oozie สำหรับ Hue	32
	7.6 สร้าง home directory for hue user	
	7.7 ติดตั้ง hadoop-httpfs	
8	Zeppelin authentication configuration	
9	Ranger authentication with LDAP	
	9.1 Import LDAP CA cert	
	9.2 Config Ranger authen LDAP	

# สารบัญ



10	การเพิ่มบัญชีผู้ใช้และการใช้งาน	.40
11	การใช้งาน	.43
	11.1 บริการ ssh	.43
	11.2 Hue บริการหน้าเว็บสำหรับ Hadoop Ecosystems	.43



# การติดตั้งระบบ Hortonworks HDP3+Hue+FreeIPA

### 1 บทนำ

การติดตั้ง Apache Hadoop ด้วย Hortonworks Data Platform (HDP) 3 ตามเอกสารฉบับนี้นั้น จะมี การติดตั้งเว็บสำหรับเข้าใช้งาน Hadoop ที่ชื่อว่า Hue (https://gethue.com/) และกำหนดให้มีการยืนยันตัวตน กับ LDAP เซิร์ฟเวอร์ ในที่นี้จะใช้ซอฟต์แวร์ FreeIPA ด้วยองค์ประกอบเหล่านี้ก็จะได้ระบบ Hadoop ที่พร้อมใช้ งานได้เป็นอย่างดี มีการพิสูจน์ตัวตนทั้งในส่วนของ SSH, Hue และ Zeppelin

สำหรับเครื่องที่จะติดตั้ง HDP นั้น เราใช้ระบบปฏิบัติการ CentOS-7 ส่วนเครื่องที่ติดตั้ง FreeIPA เป็น Rocky Linux 9

### 2 การเตรียมระบบก่อนการติดตั้ง

- เตรียม repository ของ HDP ในเอกสารนี้ใช้ repo ของ <u>http://makeopensourcegreatagain.com/</u> ควรจะ Mirror มาทำเป็น local repo
- สร้าง root ssh key โดยจากเครื่อง ambari server ต้องสามารถ ssh ไปยังทุกเครื่องในระบบได้โดยไม่ ติดต้องระบุรหัสผ่าน
- ปิดการทำงานของ SELinux กำหนดตัวแปร SELINUX ในไฟล์ /etc/selinux/config ให้เป็น permissive หรือ Disable
- 4) กำหนดค่าโดเมนเนมของทุกเครื่องใน DNS เพื่อให้สามารถติดต่อไปยังทุกเครื่องได้โดยใช้ชื่อเครื่อง
- 5) ตั้งชื่อเครื่องทุกเครื่องในคลัสเตอร์ให้เป็น fully qualified domain name ตามที่กำหนดใน DNS เช่น edge.example.com hostnamectl set-hostname edge.example.com
- 6) ตัวอย่างชื่อเครื่องในเอกสารนี้

Hostname	roles		
ipa.example.com	LDAP, Kerberos and DNS server		
edge.example.com	Ambari, Hue, DB		
master1.example.com	Namenode, Yarn resourcemanager		
master2.example.com	Namenode, Yarn resourcemanager		
worker1.example.com	Data node, Hbase Region server		
worker2.example.com	Data node, Hbase Region server		
worker3.example.com	Data node, Hbase Region server		

\* การตั้งโดเมนที่มีคนอื่นเป็นเจ้าของ ไม่สามารถเอาไปกับโดเมนจริง ๆ ได้ ในเอกสารนี้เราใช้ Private IP ทำเป็น ตัวอย่างใช้งานภายใน



#### 3 FreeIPA

FreeIPA เป็น LDAP และ Kerberos เซิร์ฟเวอร์ของค่าย Redhat ที่แจกจ่ายให้ใช้งานเสรี หากจะใช้งาน LDAP จำเป็นต้องตั้งชื่อเครื่องให้เป็น Fully Qualified Domain Name (FQDN) และหากจะใช้งาน Kerberos จำเป็นจะต้องกำหนดชื่อโดเมนผ่าน DNS

### 3.1 ติดตั้ง FreelPA server

ในระบบนี้ใช้งานเพียงส่วนของ LDAP ไม่ได้รวมถึง Kerberos (หากจะใช้งาน Kerberos ทุกเครื่องต้องมี hostname ใน DNS สามารถติดตั้ง ipa-server-dns เพิ่มได้) และติดตั้งบนระบบปฏิบัติการ Rocky9 การติดตั้ง FreeIPA ทำได้โดยการใช้คำสั่งต่อไปนี้

```
yum install -y ipa-server
```

หลังจากนั้นรันคำสั่งต่อไปนี้เพื่อคอนฟิก

ipa-server-install

คำสั่งข้างต้นจะไปคอนฟิกค่าต่าง ๆ ให้ มีค่าที่เราต้องกรอกหลัก ๆ คือ รหัสผ่านของ Directory Manager และ IPA Admin ผลการรันสคริปต์เป็นดังต่อไปนี้

```
[root@freeipa ~]# ipa-server-install
```

```
The log file for this installation can be found in /var/log/ipaserver-install.log

This program will set up the IPA Server.

Version 4.10.2

This includes:

* Configure a stand-alone CA (dogtag) for certificate management

* Configure the NTP client (chronyd)

* Create and configure an instance of Directory Server

* Create and configure a Kerberos Key Distribution Center (KDC)

* Configure Apache (httpd)

* Configure SID generation

* Configure the KDC to enable PKINIT

To accept the default shown in brackets, press the Enter key.

Do you want to configure integrated DNS (BIND)? [no]:

Enter the fully qualified domain name of the computer

an which you're certified up conver configure the form
```

on which you're setting up server software. Using the form <hostname>.<domainname> Example: master.example.com



#### Server host name [ipa.example.com]:

The domain name has been determined based on the host name.

#### Please confirm the domain name [example.com]:

The kerberos protocol requires a Realm name to be defined. This is typically the domain name converted to uppercase.

Please provide a realm name [EXAMPLE.COM]: Certain directory server operations require an administrative user. This user is referred to as the Directory Manager and has full access to the Directory for system management tasks and will be added to the instance of directory server created for IPA. The password must be at least 8 characters long.

#### Directory Manager password: Password (confirm):

The IPA server requires an administrative user, named 'admin'. This user is a regular system account used for IPA server administration.

#### IPA admin password: Password (confirm):

Trust is configured but no NetBIOS domain name found, setting it now. Enter the NetBIOS name for the IPA domain. Only up to 15 uppercase ASCII letters, digits and dashes are allowed. Example: EXAMPLE.

#### NetBIOS domain name [EXAMPLE]:

Do you want to configure chrony with NTP server or pool address? [no]:

The IPA Master Server will be configured with: Hostname: ipa.example.com IP address(es): 192.168.2.10 Domain name: example.com Realm name: EXAMPLE.COM

The CA will be configured with: Subject DN: CN=Certificate Authority,O=EXAMPLE.COM Subject base: O=EXAMPLE.COM Chaining: self-signed

Continue to configure the system with these values? [no]: yes

The following operations may take some minutes to complete. Please wait until the prompt is returned.

•••

...

Configuring client side components



คู่มือติดตั้ง HDP + Hue + FreelPA 7

This program will set up IPA client. Version 4.10.2 Using existing certificate '/etc/ipa/ca.crt'. Client hostname: ipa.example.com Realm: EXAMPLE.COM DNS Domain: example.com IPA Server: ipa.example.com BaseDN: dc=example,dc=com Configured /etc/sssd/sssd.conf Systemwide CA database updated. Adding SSH public key from /etc/ssh/ssh\_host\_ed25519\_key.pub Adding SSH public key from /etc/ssh/ssh\_host\_ecdsa\_key.pub Adding SSH public key from /etc/ssh/ssh\_host\_rsa\_key.pub Could not update DNS SSHFP records. SSSD enabled Configured /etc/openldap/ldap.conf Configured /etc/ssh/ssh\_config Configured /etc/ssh/sshd\_config.d/04-ipa.conf Configuring example.com as NIS domain. Client configuration complete. The ipa-client-install command was successful Please add records in this file to your DNS system: /tmp/ipa.system.records.h8lkr016.db \_\_\_\_\_ Setup complete Next steps: 1. You must make sure these network ports are open: **TCP Ports:** \* 80, 443: HTTP/HTTPS \* 389, 636: LDAP/LDAPS \* 88, 464: kerberos UDP Ports: \* 88, 464: kerberos \* 123: ntp 2. You can now obtain a kerberos ticket using the command: 'kinit admin' This ticket will allow you to use the IPA tools (e.g., ipa user-add) and the web user interface.

Be sure to back up the CA certificates stored in /root/cacert.p12 These files are required to create replicas. The password for these files is the Directory Manager password The ipa-server-install command was successful

เป็นอันเสร็จการติดตั้ง FreelPA Server



### 3.2 ติดตั้ง FreeIPA Client

FreeIPA Client จะประกอบไปด้วยส่วนของ LDAP Client และ Kerberos Client ซึ่งระบบไปคอนฟิก ให้ ต้องติดตั้งที่ทุกเครื่องในระบบ สามารถติดตั้งด้วยคำสั่งต่อไปนี้

yum install -y ipa-client

และรันคำสั่งต่อไปนี้เพื่อคอนฟิก (mkhomedir กำหนดให้สร้าง Home directory ให้ผู้ใช้เมื่อ login)

```
ipa-client-install _domain=example.com --server=ipa.example.com --mkhomedir
```

ตอบคำถามและระบุค่าดังต่อไปนี้

```
Proceed with fixed values and no DNS discovery? [no]: yes
Continue to configure the system with these values? [no]: yes
User authorized to enroll computers: admin
Password for admin@EXAMPLE.COM:
```

### 3.3 LDAP binding account

ในกรณีที่แอปพลิเคชันต้องการ Binding มายัง LDAP server ในระบบ Hadoop นี้ Ranger ต้องการ Binding บางคนอาจใช้บัญชีผู้ใช้ admin ของ LDAP แต่ในทางปฏิบัติเพื่อความปลอดภัย เราควรสร้าง user สำหรับ Binding ขึ้นมา ที่มีสิทธิ์แบบ Read only ในระบบนี้ทำที่เครื่อง edge ทำได้ตามขั้นตอนต่อไปนี้

1) ติดตั้ง openIdap-clients

yum install -y openldap-clients

2) สร้าง system account for ldap binding ให้สั่ง

```
ldapmodify -x -D 'cn=Directory Manager' -W
dn: uid=system,cn=sysaccounts,cn=etc,dc=example,dc=com
changetype: add
objectclass: account
objectclass: simplesecurityobject
uid: system
userPassword: password
passwordExpirationTime: 20380119031407Z
nsIdleTimeout: 0
<blank line>
^D
```

หมายเหตุ สองบรรทัดท้าย <blank line> ให้เว้นบรรทัดว่าง แล้วจึงกด Ctrl+D อ้างอิง

https://www.freeipa.org/page/HowTo/LDAP



#### 4 Apache Ambari

### 4.1 ติดตั้ง Ambari

4.1.1 ติดตั้ง ambari repository โดยดาวน์โหลดไฟล์ ambari.repo จาก repository ที่เตรียมไว้ ไป วางไว้ที่ /etc/yum.repos.d/

# wget -nv http://makeopensourcegreatagain.com/rpms/mosga.repo -0
/etc/yum.repos.d/ambari.repo

4.1.2 แก้ไขไฟล์ ambari.repo โดยแก้จาก https เป็น http

4.1.3 ติดตั้ง Apache Ambari

# yum install -y ambari-server

4.1.4 ถ้าทำ Local repo ก็ให้โหลด JDK และ JCE มาวางใน Local repo ด้วย แล้วแก้ไขลิงค์

ดาวน์โหลด JDK ในไฟล์ /etc/ambari-server/conf/ambari.properties จาก hortonworks repo

เป็น local repo ของเรา เช่น

# sed -i 's/public-repo-1.hortonworks.com/192.168.2.10\/hortonworks-repos/g'

### 4.2 ติดตั้ง MariaDB

MariaDB ทำหน้าที่เป็นฐานข้อมูลหลังบ้านให้กับ Ambari Server ในการเอกสารนี้เลือกใช้ MariaDB server ที่มากับ CentOS-7 ติดตั้งไว้ที่เครื่อง edge

yum install –y mariadb-server

กำหนดให้ MariaDB เริ่มการทำงานทุกครั้งที่เปิดเครื่อง และเริ่มการทำงานของ MariaDB ทันที

systemctl enable --now mariadb

เพื่อความปลอดภัยให้รันคำสั่งต่อไปนี้ เพื่อกำหนดรหัสผ่านให้ root ลบสิ่งที่ไม่จำเป็น และทำให้ระบบ

มั่นคงขึ้น /usr/bin/mysql\_secure\_installation

### 4.3 ติดตั้ง MySQL JDBC

ติดตั้งที่เครื่อง edge

```
# wget https://dev.mysql.com/get/Downloads/Connector-J/mysql-connector-java-
5.1.46.tar.gz
```

```
# tar xvf mysql-connector-java-5.1.46.tar.gz
```

```
# mkdir -p /usr/share/java
```

```
# cp mysql-connector-java-5.1.46/mysql-connector-java-5.1.46.jar /usr/share/java
# ln -sf /usr/share/java/mysql-connector-java-5.1.46.jar /usr/share/java/mysql-
connector-java.jar
```



#### 4.4 สร้าง Database สำหรับ Ambari และผองเพื่อน

อ้างอิง https://docs.hortonworks.com/HDPDocuments/Ambari-2.7.4.0/administering-

ambari/content/amb\_example\_install\_mysql-mariadb\_for\_multiple\_components.html

```
# login to mariadb
mysql -u root -p
# create database
create database ambari;
grant all privileges on ambari.* to 'ambari'@'localhost' identified by 'password';
grant all privileges on ambari.* to 'ambari'@'%' identified by 'password';
create database hive DEFAULT CHARACTER SET utf8:
grant all privileges on hive.* to 'hive'@'localhost' identified by 'password';
grant all privileges on hive.* to 'hive'@'%' identified by 'password';
create database ranger DEFAULT CHARACTER SET utf8;
grant all privileges on ranger.* to 'rangeradmin'@'localhost' identified by 'password';
grant all privileges on ranger.* to 'rangeradmin'@'%' identified by 'password';
create database rangerkms DEFAULT CHARACTER SET utf8;
grant all privileges on rangerkms.* to rangerkms@'localhost' identified by 'password';
grant all privileges on rangerkms.* to rangerkms@'%' identified by 'password';
create database oozie DEFAULT CHARACTER SET utf8:
grant all privileges on oozie.* to 'oozie'@'localhost' identified by 'password':
grant all privileges on oozie.* to 'oozie'@'%' identified by 'password';
create database superset DEFAULT CHARACTER SET utf8;
grant all privileges on superset.* to 'superset'@'localhost' identified by 'password';
grant all privileges on superset.* to 'superset'@'%' identified by 'password';
create database druid DEFAULT CHARACTER SET utf8;
grant all privileges on druid.* to 'druid'@'localhost' identified by 'password';
grant all privileges on druid.* to 'druid'@'%' identified by 'password';
create database hue DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_general_ci;
grant all ON hue.* TO 'hue'@'%' IDENTIFIED BY 'password';
grant all on hue.* to 'hue'@'localhost' identified by 'password';
```

รันสคริปต์เพื่อสร้างตารางในฐานข้อมูล ambari

```
mysql -u root -p
use ambari;
source /var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql
```



#### 4.5 คอนฟิก ambari-server

สั่งคำสั่ง ambari-server setup เพื่อเริ่ม setup ambari โดยระบบจะตรวจสอบ SELinux, ติดตั้ง JDK, config Database ตามลำดับ

# ambari-server setup Using python /usr/bin/python Setup ambari-server Checking SELinux... SELinux status is 'enabled' SELinux mode is 'permissive' WARNING: SELinux is set to 'permissive' mode and temporarily disabled. OK to continue [y/n] (y)? Customize user account for ambari-server daemon [y/n] (n)? y Enter user account for ambari-server daemon (root): Adjusting ambari-server permissions and ownership... Checking firewall status... WARNING: iptables is running. Confirm the necessary Ambari ports are accessible. Refer to the Ambari documentation for more details on ports. OK to continue [y/n] (y)? Checking JDK... [1] Oracle JDK 1.8 + Java Cryptography Extension (JCE) Policy Files 8 [2] Custom JDK Enter choice (1): To download the Oracle JDK and the Java Cryptography Extension (JCE) Policy Files you must accept the license terms found at javase/terms/license/index.html and not accepting will cancel the Ambari Server setup and you must install the JDK and JCE files manually. Do you accept the Oracle Binary Code License Agreement [y/n] (y)? Downloading JDK from -x64.tar.gz to /var/lib/ambari-server/resources/jdk-8u112-linuxx64.tar.gz jdk-8u112-linux-x64.tar.gz... 100% (174.7 MB of 174.7 MB) Successfully downloaded JDK distribution to /var/lib/ambari-server/resources/jdk-8u112linux-x64.tar.gz Installing JDK to /usr/jdk64/ Successfully installed JDK to /usr/jdk64/ Downloading JCE Policy archive from jce\_policy-8.zip to /var/lib/ambari-server/resources/jce\_policy-8.zip Successfully downloaded JCE Policy archive to /var/lib/ambari-server/resources/jce\_policy-8.zip Installing JCE policy... Check JDK version for Ambari Server... JDK version found: 8 Minimum JDK version is 8 for Ambari. Skipping to setup different JDK for Ambari Server. Checking GPL software agreement... GPL License for LZO: https://www.gnu.org/licenses/old-licenses/gpl-2.0.en.html Enable Ambari Server to download and install GPL Licensed LZO packages [y/n] (n)? Completing setup...



Configuring database... Enter advanced database configuration [y/n] (n)? y Configuring database... \_\_\_\_\_ Choose one of the following options: [1] - PostgreSQL (Embedded) [2] - Oracle [3] - MySQL / MariaDB [4] - PostgreSQL [5] - Microsoft SQL Server (Tech Preview) [6] - SQL Anywhere [7] - BDB Enter choice (1): 3 Hostname (localhost): Port (3306): Database name (ambari): Username (ambari): Enter Database Password (bigdata): Re-enter password: Configuring ambari database... Enter full path to custom jdbc driver: /usr/share/java/mysql-connector-java-5.1.46.jar Configuring remote database connection properties... WARNING: Before starting Ambari Server, you must run the following DDL against the database to create the schema: /var/lib/ambari-server/resources/Ambari-DDL-MySQL-CREATE.sql Proceed with configuring remote database connection properties [y/n] (y)? Extracting system views... ambari-admin-2.7.4.0.118.jar . . . . Ambari repo file contains latest json url hdp\_urlinfo.json, updating stacks repoinfos with it... Adjusting ambari-server permissions and ownership... Ambari Server 'setup' completed successfully.

เริ่มการทำงานของ ambari-server

```
# ambari-server setup --jdbc-db=mysql --jdbc-driver=/usr/share/java/mysql-connector-java-
5.1.46.jar
# chkconfig ambari-server on
# service ambari-server start
Using python /usr/bin/python
Starting ambari-server
Ambari Server running with administrator privileges.
Organizing resource files at /var/lib/ambari-server/resources...
Ambari database consistency check started...
Server PID at: /var/run/ambari-server/ambari-server.pid
Server out at: /var/log/ambari-server/ambari-server.out
Server log at: /var/log/ambari-server/ambari-server.log
Waiting for server start.....
Server started listening on 8080
DB configs consistency check: no errors and warnings were found.
Ambari Server 'start' completed successfully.
```



### 5 ติดตั้ง Hortonwork Data Platform (HDP)

# 5.1 ติดตั้ง HDP ผ่าน Apache ambari

5.1.1 เปิด browser ไปที่ http://<edge IPAddress>:8080 จะพบหน้า login เข้า Ambari Web UI โดยใช้ user: admin password: admin

☆ Most Visited □Internal □all □e-service □gistda @ Case Study: Fail	over us 🐵 Implementing Virtual 🕀 High Availability Virtua 🗖 B	ug #21518: mgr[zabb 🔳 Setting up a Load-Bala »
ᢙ Ambari		
	Sign in	
	Username	
	admin	
	Password	
	SIGN IN	

### 5.1.2 คลิกปุ่ม LAUNCH INSTALL WIZARD เพื่อเริ่มการติดตั้ง

ᢙ Ambari	Admin / Cluster Information	👗 admin 👻
Cluster Management  Cluster Information Remote Clusters Users	Welcome to Apache Ambari Provision a cluster, manage who can access the cluster, and customize vi	iews for Ambari users.
III Views	Create a Cluster Use the Install Wizard to select services and configure your	r cluster
	LAUNCH INSTALL WIZARD	

5.1.3 0. Get Started: ตั้งชื่อระบบคลัสเตอร์ที่จะสร้าง

Installer		💄 admin 🗸
	Get Started This wizard will walk you through the cluster Installation process. First, start by naming your new cluster.	
	Name your cluster Learn more	
	CLUSTER NAME	
CANCEL		NEXT →



#### 5.1.4 1: Select Version

หัวข้อ Repositories ให้เลือก "Use Local Repository" และ Base URL ให้ชี้ไปยัง local repository ของ redhat7 โดยระบุ URL ตามลิงก์ด้านล่าง ส่วน Distros อื่น ๆ ให้ทำการ Remove ออกจากนั้นคลิก NEXT

HDP-3.1: <a href="http://makeopensourcegreatagain.com/rpms/">http://makeopensourcegreatagain.com/rpms/</a>

HDP-UTILS:	http:/	/makeopensour	rcegreataga	in.com/rpms/	/HDP-UTILS-1	.1.0.22/

	Atlas		1.1.0				
Assign Slaves and Clients	Druid		0.12.1				
⑦ Customize Services	HBase		2.0.2				
(B) Review	Repositories Using a Public Repository requires Interne	et connectivity. Using a Local Repository requires you have	e configured the software in a repository	y available in your network.			
<ol> <li>Install, Start and Test</li> </ol>	O Use Public Repository		O Use Local Reposite	Use Local Repository			
10 Summary	Provide Base URLs for the Oper	rating Systems you are configuring.					
					+ADD -		
	OS	Name		Base URL			
	redbat7	HDP-3.1 🌶	×	http://203.150.243.109/hortonworks-repos/HDP/centos7/3.1.4.0-315/			
		HDP-UTILS-1.1.0.22 #	*	http://203.150.243.109/hortonworks-repos/HDP-UTILS/centos7/1.1.0.22/ 9			
	<ul> <li>Skip Repository Base URL valid</li> <li>Use RedHat Satellite/Spacewa</li> </ul>	dation (Advanced) 😧					
- BACK CANCEL					NEXT		

#### 5.1.5 2 : Install Options

### Target Hosts : ใส่ชื่อ FQDN เครื่องที่จะติดตั้ง Hadoop ทั้งหมด ดังนี้

- edge.example.com
- master1.example.com
- master2.example.com
- worker1.example.com
- worker2.example.com
- worker3.example.com

### Host Registration Information : ใส่ private key เพื่อให้ Ambari สามารถรีโมทไปติดตั้ง

### Agent เครื่องในระบบได้โดยไม่ติด password

Installer			👤 admin 🛩
Get Started	Install Options Enter the list of hosts to be included in the cluster and provide your SSH key.		
	Target Hosts           Enter a list of hosts using the Fully Qualified Domain Name (PGDN), one per line. Or use Pattern master, example.com master, example.com model_compared.com worker_anaple.com worker_anaple.com           Model_seargh.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           worker_anaple.com           CHOODEFILE           hodooptraining.pem           ************************************	The Expressions	
Summary	avvetykutiknileatta4stler-kkled/dueuanaaznaooafoEoxoKJREcktleSkklozeo SSH User Account	reot	
	SSH Port Number	22	
- BACK CANCEL			REGISTER AND CONFIRM



### 5.1.6 3 : Confirm Hosts จะทำการติดตั้ง Agent ไปยังเครื่องต่าง ๆ ในระบบ

	Ambari - Cluster Inst	all Wizard - Mozilla Firefox:			000
🔹 Openlandscape Cloud 🛛 🗙 203.150.	243.106/HDP/cento 🗙 🛛 🐶 Ambari - Cluster Install 🕅 🗙	+			
← → C <sup>2</sup>	<b>7</b> :8080/#/installer/step3	··· 🗵 🕁	Q Search	III 🔃 🖸 🗉	•3 © ≡
Installer					Ladmin -
Cet Started Re	nfirm Hosts gistering your hosts ase confirm the host list and remove any hosts that you do not want	to include in the cluster.			
Install Options			Show: All (6)   Installing (0)   F	legistering (0)   Success (6)	Fail (0)
Confirm Hosts	Host	Progress	Status	Action	
④ Choose Services	edge.example.com		Success	Î	
(5) Assign Masters	a master1.example.com		Success	Ĩ	
6 Assign Slaves and Clients	anaster2.example.com		Success	Î	
0	worker1.example.com		Success	<b>Î</b>	
(7) Customize Services	worker2.example.com		Success	Î	
(8) Review	worker3.example.com		Success	Î	
			Items per pag	je: 25 → 1-6of6 <>	

#### 5.1.7 4 : Choose services

เลือกบริการต่าง ๆ ที่จะติดตั้งในระบบ Hadoop ในเอกสารนี้จะเลือกติดตั้ง YARN-MapReduct2, Tez, Hive, Hbase, Pig, Sqoop, Oozie, ZooKeeper, Infra Solr, Ambari Metrics, Atlas, Kafka, Ranger, SmartSense, Spark2 และ Zeppelin Notebook จากนั้นคลิก NEXT

Installer					👤 admin 👻
Get Started	Choose File System choose which file system	0 you want to install on your cluster.			
Select Version	Service		Version	Description	
衬 Install Options	HDFS		3.1.1	Apache Hadoop Distributed File System	
Confirm Hosts					
	Choose Services				
		ou want to install on your cluster.			
(3) Assign Masters	<ul> <li>Service</li> </ul>		Version	Description	
Assign Slaves and Clients	<ul> <li>YARN + Ma</li> </ul>	apReduce2	3.3.3	Apache Hodoop HeelGen MapReduce (XANI)	
(7) Customize Services	<ul> <li>Tez</li> </ul>		0.9.1	Tez is the next generation Hadoop Query Processing framework written on top of YAPN.	
-	<ul> <li>Hive</li> </ul>		3.1.0	Data warehouse system for ad-hoc queries & analysis of large datasets and table & storage management service	
() Interest	<ul> <li>HBase</li> </ul>		2.0.2	Non-relational distributed database and centralized service for configuration management & synchronization	
Install, Start and Test	Pig		0.16.0	Scripting platform for analysing large datasets	
(1) Summary	<ul> <li>Sqoop</li> </ul>		1.4.7	Tool for transfering bulk data between Apache Hadoop and structured data stores such as relational databases	
	<ul> <li>Oozie</li> </ul>		4.3.1	System for workflow continuation and execution of Apache Hadoop jobs. This also includes the installation of the optional Oacia Web Console which relies on and will install the ExtED Library.	
	<ul> <li>ZooKeeper</li> </ul>		3.4.6	Centralized service which provides highly reliable distributed coordination	
	<ul> <li>Storm</li> </ul>		1.2.1	Apache Hadoop Bream processing framework	
	<ul> <li>Accumulo</li> </ul>		1.7.0	Robust, scalable, Nigh performance distributed key/value store.	
	<ul> <li>Infra Solr</li> </ul>		0.1.0	Core shared service used by Ambari managed components.	
	<ul> <li>Ambari Me</li> </ul>	trics	0.1.0	A system for metrics collection that provides storage and retrieval capability for metrics collected from the cluster	
	<ul> <li>Atlas</li> </ul>		1.1.0	Adas Metadata and Gevenance platform	
	<ul> <li>Kafka</li> </ul>		2.0.0	A high-thoughput distributed messaging system	
	- Knax		1.0.0	Provédes a single point of authentication and access for Apache Hadoop services in a cluster	
	<ul> <li>Log Search</li> </ul>	h	0.5.0	Log aggregation, analysis, and visualization for Ambani managed services. This service is <b>Technical Preview</b> .	
	<ul> <li>Banger</li> </ul>		1.2.0	Comprehensive security for Hadoop	
	<ul> <li>Banger KM</li> </ul>	15	1.2.0	Key Management Server	
	© SmartSens	10	1.5.1.2.7.4.0-118	SmartSense - Hostonworks SmartSense Tool (HST) helps quickly gather configuration, metrics, logs from common HCP services that aids to quickly troubleshoot support cases and receive cluster-specific recommendations.	
	<ul> <li>Spark2</li> </ul>		232	Apache Spark 2.3 is a fast and general engine for large-scale data processing	
	<ul> <li>Zeppelin N</li> </ul>	lotebook	0.8.0	A web-based notebook that enables interactive data analytics. It enables you to make beautiful data driven, interactive and collaborative documents with SQL Scala and more.	
	<ul> <li>Druid</li> </ul>		0.12.1	A fast column-oriented distributed data store.	
	<ul> <li>Superset</li> </ul>		0.23.0	Superset is a data exploration platform designed to be visual, intuitive and interactive. This service in Technical Provines.	
BACK CANCEL					NEXT



#### 5.1.8 5 : Assign Masters

เลือก Host สำหรับบริการต่าง ๆ จะติดตั้งอยู่ในเครื่องใด

Get Started	Assign Masters Assign master components to hosts you want to run them on.		
Select Version			
Install Options	SNameNode:	master1.example.com (3.7 GB, 2 cores)	edge.example.com (15.5 GB, 4 cores)     Timeline Service V1.6     Timeline Service V2.0 Reader
📀 Confirm Hosts	NameNode:	master2.example.com (3.7 GB, 2 cores)	VARN Registry DNS Hristory Server Hrive Metastore HriveServer2 Oozie Server ZoolKeeper Server Grafana Metrics Collector
Choose Services	Timeline Service V1.5:	edge.example.com (15.5 GB, 4 cores)	Activity Explorer HST Server Activity Analyzer Spark2 History Server
5 Assign Masters	ResourceManager:	master1.example.com (3.7 GB, 2 cores)	<ul> <li>master1.example.com (3.7 GB, 2 cores)</li> </ul>
Assign Slaves and Clients     Customize Services	Timeline Service V2.0 Reader.	edge.example.com (15.5 GB, 4 cores)	SkameNode ResourceManager Hilliso Master ZooKesper Server
(B) Review	YARN Registry DNS:	edge.example.com (15.5 GB, 4 cores)	<ul> <li>master2.example.com (3.7 GB, 2 cores)</li> </ul>
<ol> <li>Install, Start and Test</li> </ol>	History Server:	edge.example.com (15.5 GB, 4 cores)	Cookept Server
(1) Summary	Hive Metastore:	edge.example.com (15.5 GB, 4 cores)	*
	HiveServer2:	edge.example.com (15.5 GB, 4 cores)	•
	HBase Master.	master1.example.com (3.7 GB, 2 cores)	•
	Oozie Server.	edge.example.com (15.5 GB, 4 cores)	·
	ZooKeeper Server.	master1.example.com (3.7 GB, 2 cores)	•
	ZooKeeper Server:	edge.example.com (15.5 GB, 4 cores)	•
	ZooKeeper Server.	master2.example.com (3.7 GB, 2 cores)	- 00
	Grafana:	edge.example.com (15.5 GB, 4 cores)	*
	Metrics Collector.	edge.example.com (15.5 GB, 4 cores)	• 🖬
	Activity Explorer:	edge.example.com (15.5 GB, 4 cores)	•
	HST Server:	edge.example.com (15.5 GB, 4 cores)	•
	Activity Analyzer.	edge.example.com (15.5 GB, 4 cores)	•
	Spark2 History Server.	edge.example.com (15.5 GB, 4 cores)	•
- BACK CANCEL			NEXT

#### 5.1.9 6: Assign Slaves and Clients

เลือก Host สำหรับเป็น Datanode และ Client ตามภาพด้านล่าง จากนั้นคลิก NEXT

nstaller									💄 adı
Get Started     Select Version	Assign Slaves and Clients Assign slave and client components Hoats that are assigned master com "Client" will install HDFS Client, VARN	to hosts you want to run then ponents are shown with . Client, MapReduce2 Client, T	n on. Tez Client, Hive Client, HBase Cli	ent, Pig Client, Sqoop Client, O	ozie Client, ZooKeeper Client an	d Spark2 Client.			
Install Options	Host	all   none	all   none	all   none	all   none	all   none	all   none	all   none	ali   none
Confirm Hosts	edge.example.com @	DataNode	NFSGateway	NodeManager	RegionServer	Phoenix Query Server	<ul> <li>Livy for Spark2 Server</li> </ul>	Spark2 Thrift Server	Client
	master1.example.com	DataNode	<ul> <li>NFSGateway</li> </ul>	NodeManager	RegionServer	Phoenix Query Server	Livy for Spark2 Server	Spark2 Thrift Server	Client
Choose Services	master2.example.com e	<ul> <li>DataNode</li> </ul>	NFSGateway	NodeManager	RegionServer	<ul> <li>Phoenix Query Server</li> </ul>	Livy for Spark2 Server	<ul> <li>Spark2 Thrift Server</li> </ul>	Client
Assign Masters	worker1.example.com	DataNode	NFSGateway	NodeManager	RegionServer	Phoenix Query Server	Livy for Spark2 Server	<ul> <li>Spark2 Thrift Server</li> </ul>	Client
Assian Slaves and Clients	worker2.example.com	DataNode	NFSGateway	NodeManager	RegionServer	Phoenix Query Server	Livy for Spark2 Server	<ul> <li>Spark2 Thrift Server</li> </ul>	Client
	worker3.example.com	DataNode	NFSGateway	NodeManager	RegionServer	<ul> <li>Phoenix Query Server</li> </ul>	Livy for Spark2 Server	<ul> <li>Spark2 Thrift Server</li> </ul>	Client
) Customize Services								Items p	er page: 25 🔻 1 - 6 of 6 🔇
) Review									
) Install, Start and Test									
) Summary									
BACK CANCEL									NEXT



#### 5.1.10 7: Customize Services → CREDENTIALS

### ระบุ Username , Password ตามที่ท่านสร้างฐานข้อมูลไว้ จากนั้นคลิก NEXT

Installer					💄 admin 🗸
Get Started	CREDENTIALS EDATABASES	DIRECTORIES - LACCOUNTS -	FALL CONFIGURATIONS		
Select Version	Please provide credentials for these services				
Install Options		Username*	Password*	Confirm Password*	
Confirm Hosts	Grafana Admin	admin			
Choose Services	Hive Database	hive			
Assign Masters Assign Slaves and Clients	Oozie Database	oozie			
Customize Services	Activity Explorer's Admin	N/A			
(B) Review					
(9) Install, Start and Test					
1 Summary					
- BACK CANCEL					NEXT

5.1.11 7: Customize Services → DATABASES เชื่อมต่อ service ต่าง ๆ กับฐานข้อมูล MariaDB หัวข้อ Hive Database เลือก Existing MySQL / MariaDB คลิก TEST CONNECTION จากนั้น คลิก NEXT

	Ambari - Cluster Install Wizard - Mo	zilla Firefox					•	
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> oo	kmarks <u>T</u> ools <u>H</u> elp							
🔹 Openlandscape Cloud 🛛 🗙 🛛 20	3.150.243.106/HDP/cento × 🔗 Ambari - Cluster Install V × 🕂							
← → Ĉ û ① ½ 203.	150.107.27:8080/#/installer/step7	🖾 🗘 🔍 password	$\rightarrow$	III\ 🕅	<b>D</b>	æj	۲	≡
🥏 Get Started		ACCOUNTS 🗡 ALL CONFIGU	RATIONS					
Select Version	Please choose and configure the appropriate databases for these services							
Install Options	HIVE OOZIE RANGER							
🥏 Confirm Hosts								
Choose Services	Hive Database	Hive Database Type						
I	Existing MySQL / MariaDB 👻	mysql		C				
🥑 Assign Masters								
Assign Slaves and Clients		JDBC Driver Class						
Customiza Sarvicas	/downloads/connector/j/ from MySQL. Once downloaded to the Ambari Server /downloads/connector/j/ from MySQL. Once downloaded to the Ambari Server heat and	com.mysql.jdbc.Driver		C	•			
<ul> <li>8 Review</li> </ul>	nost, run: ambari-server setupjdbc-db=mysqljdbc-driver=/path/to/mysql /com.mysql.jdbc.Driver	Database Password						
(1) Install Start and Test	Database Name							
Constant otari and rest	hive C							
10 Summary								
https://dev.mysqt.com/downloads/co	onnector/J/	TEST CONNECTION	Connection OK					



Installer 👤 admin 👻 📀 Get Starte HDFS YARN MAPREDUCE2 TEZ HIVE HBASE OOZIE ZOOKEEPER AMBARI METRICS SPARK2 Default (6) 🔹 Filter... 👻 DATA DIRS DataNode directories o c 🔒 /hadoop/hdfs/data NameNode directories C 🔒 /hadoop/hdfs/namenode ondaryNameNode Checkpoint directories C 🔒 /hadoop/hdfs/namesecondary NFSGateway dump directory 0 C 🗎 /tmp/.hdfs-nfs NameNode Backup directory OC /tmp/upgrades

5.1.12 7: Customize Services → DIRECTORIES กำหนดเป็น Default คลิก NEXT

#### 5.1.13 7: Customize Services → ACCOUNTS กำหนดเป็น Default คลิก NEXT

Installer				👤 admin 🗸
Set Started	CREDENTIALS EDATAB/	ASES — @DIRECTORIES — LACCOUNTS — FALL CONFIGURA	TIONS	
Select Version	Please review these settings for Servi	ice Accounts		
Install Options	<ul> <li>Use Ambari to Manage Ser</li> <li>Use Ambari to Manage Gro</li> <li>Use Ambari to Manage Ser</li> </ul>	vice Accounts and Groups up Memberships vice Accounts UID's		
Confirm Hosts	Users/Groups	Usernames		
Choose Services	Smoke User	ambari-qa		
Assign Masters	Hadoop Group	hadoop		
Assign Slaves and Clients	Ambari Metrics User	ams		
Customize Services	HBase User	hbase		
(8) Review	1050 11000	hdfe		
<ol> <li>Install, Start and Test</li> </ol>	HDF5 User	Ituis		
	Proxy User Group	users		
() Summary	Hive User	hive		
	Mapreduce User	mapred		
	Oozie User	oozie		
	Livy2 Group	livy		
	Livy2 User	livy		
	Spark2 Group	spark		
	Spark2 User	spark		



Installer				💄 admin 🗸
Get Started		<u>.</u>		
Select Version	HDFS YARN MAPREDUCE2 TEZ HIVE HBASE PIG SOCOP OOZIE ZOOKEEPER AMBARI ME	TRICS SMARTSEN	NSE SPARK2 MISC	
Confirm Hosts	SETTING ADVANCED		Default (6) v Filter	•
Choose Services     Assign Masters	NameNode		DataNode	
<ul> <li>Assign Slaves and Clients</li> <li>Customize Services</li> </ul>	NameNode directories //hisdoog/hdfs/namenode	c 🔒	DataNode directories /hadoop/hdfs/data	c 🔒
(1) Review	NameNode Java heap size		A DataNode failed disk tolerance	
<ul> <li>Install, Start and Test</li> <li>Summary</li> </ul>	000 1/500 2700	C		2 <b>8</b>
	NameNode Server threads		DataNode maximum Java heap size	. •
	1 101 200 Minimum replicated blocks %	C 0	009 1.17509 3.700 DataNode max data transfer threads	
	91X 913X 101X	2 <b>0</b>	0 2.000 4000	č 💧
- BACK CANCEL				NEXT

5.1.14 7: Customize Services → ALL CONFIGURATIONS กำหนดเป็น Default คลิก NEXT

5.1.15 8: Review สรุปส่วนต่าง ๆ ที่จะติดตั้งในระบบ เมื่อคลิก next Ambari จะเริ่มติดตั้งไปยังเครื่อง ในคลัสเตอร์ คลิก DEPLOY จากนั้นรอจนติดตั้งสำเร็จ กดปุ่ม Complete

, Installer	£.	ıdmin <del>-</del>
Get Started	Review Please review the configuration before installation	
<ul> <li>Select Version</li> <li>Install Options</li> <li>Confirm Hosts</li> <li>Choose Services</li> <li>Assign Masters</li> <li>Assign Masters</li> <li>Assign Slaves and Clients</li> <li>Customize Services</li> <li>Review</li> <li>Install, Start and Test</li> </ul>	Admin Name: admin         Cluster Name: clusterkit         Total Hosts: 5 (6 new)         Repositories:         Use RedHat Satellite/Spacewalk         Services:         HDFS         Databode: 3 hosts         NameNode: master2.example.com         NSTARedWord: host         StameNode: imaster1.example.com         VRM + MapReduce2         Timeline Service V1.5 : cloge example.com         NodeManager: 3 hosts         ResourceManage: master1.example.com         NameNode: master1.example.com	
(1) Summary	Registry UNS : edge.example.com	
- BACK CANCEL	S GENERATE BLUEPRINT PRINT DEPL	DY →



### การติดตั้ง Hortonworks Data Platform (HDP) สำเร็จ

	METRICS HEATMAPS CONFIC	G HISTORY					METRIC ACTIONS -	LAST 1 HOU
rvices … 🗸								210111100
FS								
RN	NameNode Heap		HDFS Disk Usage	NameNod	e CPU WIO	1	DataNodes Live	
oReduce2					$\frown$		0.10	
	9%		8%		0.0%		3/3	
se								
	NameNode RPC		Memory Usage	Network U	sage		CPU Usage	-
qq					1 I I		100%	1
e	0 ms		4.6 GB		8			
Keeper				4.7 MB	11		50%	
ari Metrics								Marrie and
irtSense								
k2	Cluster Load		NameNode Uptime	: HBase Ma	ster Heap		HBase Ave Load	:
s	5		26m 11c				0.67	
s			2011115		2%		0.07	
ter Admin 🗸 🗸								
d Versions								
Accounts	Region In Transition	1	HBase Master Uptime	Resource	Manager Heap	1	NodeManagers Live	1

### 6 การทำ High Availability

### 6.1 การทำ Namenode HA

6.1.1 หน้า Ambari ไปที่ HDFS คลิก Actions จากนั้นเลือก Enable NameNode HA

ᢙ Ambari	ŧ	/ Services / HDFS / S	Summary			clusterkit 🔅 🕼	<b>Å</b> 0	👥 admin 🗸
🔒 Dashboard	s	SUMMARY HEATMAPS	CONFIGS METRICS					ACTIONS -
🚔 Services 🛛 🛶 🗸							► Start	
• HDFS	1	Summary					E Stop	
• YARN		Summary				40	O Resta	art All
MapReduce2			Charted	Ctarted			O Resta	art DataNodes
Tez	C C	Components	NAMENODE	SNAMENODE			→ Move	e NameNode
• Hive						•	A Move	e SNameNode
<ul> <li>HBase</li> </ul>			29m 42s NAMENODE UPTIME	6.1% 62.2 MB / 1011.3 MB		2	1 Enab	le NameNode HA
Pig				NAMENODE HEAP			Add I	New HDFS Namespace
Saoob			3/3 Started	0/0 Live	1/1 Started		D Turn	On Maintenance Mode
• Oozie			DATANODES	JOURNALNODES	NFSGATEWAYS		C Reba	lance HDFS
ZooKeener			DATANODES STATUS				Refree	esh Nodes
Amberi Matrice			3	0	0		L Dowr	nload Client Configs
<ul> <li>Amban wetrics</li> </ul>			Live	Dead	Decommissioning		Z Delet	e service
<ul> <li>SmartSense</li> </ul>								

6.1.2 จะปรากฏหน้า Wizard ทำการระบุ Nameservice ID ที่จะเป็น NameNode FQDN หลังจาก เซ็ต HA เรียบร้อยแล้ว จากนั้นคลิก NEXT



<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookr	marks <u>T</u> ools <u>H</u> elp	Ambari - test - Mozilla Firefox	008
← → ♂ ☆	50.243.106:8080/#/main/s	ervices/highAvailability/NameNode/enable/step 🛛 🏠 🔍 Search 🐘 🕱	=
Enable NameNode HA Wi	zard		admir ×
Get Started     Select Hosts	Get Started This wizard will walk you th an Active-Standby NameNo and manual steps (that you NameNode HA.	rough enabling NameNode HA on your cluster.Once enabled, you will be running a Standby NameNode in addition to your Active NameNode.Thi de configuration that automatically performs failover.The process to enable HA involves a combination of <b>automated steps</b> (that will be handled must perform in sequence as instructed by the wizard). <b>You should plan a cluster maintenance window and prepare for cluster downtime when en</b>	TIONS is allows for by the wizard) abling
	If you have HBase runnir	ng, please exit this wizard and stop HBase first.	
	Nameservice ID:	mycluster	
			_
	_		NEXT →

# 6.1.3 Select Hosts: เลือกเครื่องที่เป็น Namenode และ JournalNode

	Ambari - Lest - Mozilla Firefox							
<u>File Edit View History B</u> ookm	arks <u>T</u> ools <u>H</u> elp							
	0.243.106:8080/#/main/services/high/	Availability/NameNode/enable/step	<u></u> ତ ନ୍ଧ୍ର ସ	Search	III\ 🕅 🖸	•	۲	Ξ
Enable NameNode HA Wiza	ard					0	×	dmir
Get Started	Select Hosts Select a host that will be running the addition In addition, select the hosts to run JournalN	inal NameNode. odes, which store NameNode edit logs in a fault	tolerant manner.					IONS
③ Review	Current NameNode:	master1.example.com (3.7 GB, 2 cores)	~	edge.example.com (15	.5 GB, 4 cores)			
④ Create Checkpoint	Additional NameNode:	master2.example.com (3.7 GB, 2 cores)	•	JournalNode Resou	rceManager Reader			
5 Configure Components	JournalNode:	master2.example.com (3.7 GB, 2 cores)	-	HiveServer2 Oozie	Server			
6 Initialize JournalNodes				Accumulo Monitor	Accumulo Tracer			
(7) Start Components	JournalNode:	edge.example.com (15.5 GB, 4 cores)	-	Accumulo GC Metrie Activity Explorer HS	cs Collector Grafana ST Server	J	1	
Initialize Metadata	JournalNode:	master1.example.com (3.7 GB, 2 cores)	•	Activity Analyzer			Ц	
(9) Finalize HA Setup				master1.example.com	(3.7 GB. 2 cores)		Ц	
				NameNode Journal	Node History Server			



### 6.1.4 Review: สรุปการปรับแต่งที่จะทำ NameNode HA จากนั้นคลิก NEXT

		Ambari - test - Mozilla Firefox				• • •
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks	<u>T</u> ools <u>H</u> elp					
Ambari-test × +						
(←) → C <sup>2</sup> <sup>(1)</sup> <sup>(2)</sup> <sup>(2</sup>	106:8080/#/main/services/highAv	ailability/NameNode/enable/step	🚥 🗵 😭 🔍 Search	III\ 🔃 🖸 🗄	କ୍ଷ	. Ξ
Enable NameNode HA Wizard					×	admin
Get Started Con	riew firm your host selections.					TIONS
Select Hosts						
3 Review	Current NameNode: master1.exa Secondary NameNode: master2.exa	mple.com – TO BE DELETED				
(4) Create Checkpoint	Additional NameNode: master2.exa	mple.com 🕂 TO BE INSTALLED				
5 Configure Components	JournalNode: master2.exa edge.examp master1.exa	mple.com + TO BE INSTALLED e.com + TO BE INSTALLED mple.com + TO BE INSTALLED				
6 Initialize JournalNodes						
(7) Start Components	Review Configuration Changes. The following lists the configuration ch dfs.journalnode.edits.dir property	anges that will be made by the Wizard to enable	NameNode HA. This information is for revie	w only and is not editable except for the		Ľ
Initialize Metadata						
() Finalize HA Setup	HDFS			^		
	dfs.journalnode.edits.dir	/hadoop/hdfs/journal		C		

6.1.5 Create Checkpoint : จังหวะนี้ให้รีโมทไปที่เครื่อง master1 เพื่อเปลี่ยนเป็น safe mode ใช้ คำสั่ง <u>sudo su hdfs -l -c 'hdfs dfsadmin -safemode enter'</u> และสร้าง checkpoint ใช้คำสั่ง <u>sudo su hdfs -l -c 'hdfs dfsadmin -saveNamespace'</u> จากนั้นคลิก Next

	Ambari - Cest - Mozilla Firefox		3
mbari-test × +			
→ C 🏠 🕕 🖉 203.	150.243.106:8080/#/main/services/highAvailability/NameNode/enable/step … 🛛 🏠 🔍 Search 💷 🕅	<b>e</b> i (	9
	2/2 Ctarted 0/01 ivo 1/1 Ctarted		q
Enable NameNode HA W	fizard	×	
			ł
📀 Get Started	Manual Steps Required: Create Checkpoint on NameNode		
I			
🥑 Select Hosts	2. Put the NameNode in Safe Mode (read-only mode):		
Review	sudo su hdfs -l -c 'hdfs dfsadmin -safemode enter'		
I	3. Once in Safe Mode, create a Checkpoint:		
4 Create Checkpoint	sudo su hdís -l -c 'hdís dísadmin -saveNamespace'		
	<ol><li>You will be able to proceed once Amban detects that the NameNode is in Safe Mode and the Checkpoint has been created successfully.</li></ol>		
	If the Next button is enabled before you run the "Step 4: Create a Checkpoint" command, it means there is a recent Checkpoint already and you may proceed without		
	running uic Geop 4. Oreare a Greenpoint Contribute.		
	Checkpoint not created yet NEXT →		
<i>«</i>			d



#### Set Stard Configure Components Configur

### 6.1.6 Configure Components : รอดำเนินการคอนฟิก เมื่อสำเร็จคลิก NEXT

### 6.1.7 Manual Steps Required : Initialize JournalNodes จังหวะนี้ให้รีโมทไปที่เครื่อง master1 เริ่มการทำงาน JournalNodes โดยใช้คำสั่ง sudo su hdfs -l -c 'hdfs namenode -initializeSharedEdits' จากนั้นคลิก Next

ble NameNode HA W	rizard		
Get Started	Manual Steps Required: Initialize JournalNodes		
Select Hosts	1. Login to the NameNode host master1.example.com. 2. Initialize the JournalNodes by running:		
	sudo su hdfs ·l -c 'hdfs namenode ·initializeSharedEdits' 3. You will be able to proceed once Ambari detects that the JournalNodes have been initialized successfully.		
	JournalNode	s not initialized yet	NEVT

### 6.1.8 Start Components : เริ่มการทำงาน ZooKeeper Servers และ NameNode

Ena	ble NameNode HA Wiz	zard			^
<b></b>	Get Started	Start Components			
	Select Hosts	Please proceed to the next step.			
9	Review	Start ZooKeeper Servers			<b>E</b> x
- e	Create Checkpoint	🛹 Start NameNode			
🝦	Configure Components				
9	Initialize JournalNodes				
đ	) Start Components				
۲	) Initialize Metadata				
9	) Finalize HA Setup				



6.1.9 Manual Steps Required: Initialize NameNode HA Metadata

รีโมทไปที่เครื่อง master1 พิมพ์คำสั่ง <u>sudo su hdfs -l -c 'hdfs zkfc -formatZK'</u> และรีโมทไปที่ เครื่อง master2 พิมพ์คำสั่ง <u>sudo su hdfs -l -c 'hdfs namenode -bootstrapStandby'</u> จากนั้นคลิก Next

Anual Steps Required: initialize NameNode HA Metadata     Login to the NameNode HA Metadata	Enable NameNode HA Wi	Vizard	2
<ul> <li>Sector Loss</li> <li>Sector Loss</li> <li>Review</li> <li>Caste Checkpoint</li> <li>Configure Components</li> <li>Stat Components</li> <li>Initialize the metadata for Kaster Lesample dost.</li> <li>This is a different host from the Steps 1 and 2 above.</li> <li>Initialize the metadata for Kaster Lesample dost.</li> <li>This is a different host from the Steps 1 and 2 above.</li> <li>Stat Components</li> <li>Initialize the metadata for Kaster Lesample dost.</li> <li>This is a different host from the Steps 1 and 2 above.</li> </ul>	🥏 Get Started	Manual Steps Required: Initialize NameNode HA Metadata	
• Review       sudo su hafs - 1 - c 'haf's xkfc - formatZk'         • Create Checkpoint       c Configure Components         • Configure Components       montant Review host from the Steps 1 and 2 above.         • Initialize JournalHodes       - c 'haf's nameiode - bootstrapStandby'         • Statt Components       • Initialize Metadata         • Initialize Metadata       • Please proceed once you have completed the steps above.	Select Hosts	Login to the NameNode host master1.example.com.     Initialize the metadata for NameNode automatic failover by running:	
<ul> <li>Create Checkpoint</li> <li>Configure Components</li> <li>Components</li> <li>State Components</li> <li>Initialize Journal/Iddes</li> <li>State Components</li> <li>Initialize Metadata</li> <li>Preakze HA Setup</li> </ul>	e Review	sudo su hdfs -1 -c 'hdfs zkfc -formatZK' 3. Login to the Additional NameNode host <b>master2.example.com</b> .	
Configure C	Create Checkpoint	Important Be sure to login to the Additional NameNode host. This is a different host from the Steps 1 and 2 above.	
Start Components     Please proceed once you have completed the steps above.     Initialize Metadata     Finalize HA Setup	Initialize JournalNodes	4. initialize the metadata for the Additional NameNede by running: sudo su hdfs -1 -c 'hdfs namenode -bootstrapStandby'	
<ul> <li>Initialize Metadata</li> <li>Finalize HA Setup</li> </ul>	Start Components	Please proceed once you have completed the steps above.	
④ Finalize HA Setup			
	9 Finalize HA Setup		

6.1.10 Ambari จะปรับแต่งระบบให้ HDFS ทำงานแบบ HA ให้รอจนระบบปรับแต่งเรียบร้อย กดปุ่ม DONE เพื่อเสร็จสิ้นการทำงาน

	Amb	ari - test - Mozilla Fire	fox							
🐼 Ambari-test 🛛 🗙 🕂										
(←) → C <sup>2</sup> (1) [1] 203.150.	243.106:8080/#/main/services/highAvailability/N	ameNode/enable/s	☆ <	Search	lii\	N	<b>5</b>	) es	٢	₫ Ξ
	3/3 Started U	J/U LIVe	I/I Started							100
Enable NameNode HA Wiza	rd								×	
Cat Started	Finalize HA Setup									
Get Started	Please wait while NameNode HA is being deployed.									
Select Hosts	· · · · · · · · · · · · · · · · · · ·									
Review	Start Additional NameNode									L
Create Checkpoint	✓ Install Failover Controllers									
Configure Components	Start Failover Controllers									
Initialize JournalNodes	Reconfigure AMS									
Start Components	Reconfigure Accumulo									
	Delete Secondary NameNode									
Initialize Metadata	Stop HDFS									
9 Finalize HA Setup	Start All Services			62%						
						_		DONE		



#### 6.2 Enable ResourceManager HA Wizard

### 6.2.1 ที่หน้า Ambari ไปที่ Services > YARN คลิก Actions จากนั้นเลือก Enable

ᢙ Ambari		A / Services / YARN /	Summary				clusterkit 🛟	0 40		💄 admin 👻
<b>f</b> Dashboard		SUMMARY HEATMARS	CONFIGS METRICS							ACTIONS
🚔 Services 🛛 🗸								Start	L	, ionono
• HDFS • Yarn	1	Summary					<b>A</b> (	Stop	N Capacit	y Scheduler
<ul> <li>MapReduce2</li> </ul>		Components	<ul> <li>Started</li> </ul>	Started	<ul> <li>Started</li> </ul>	<ul> <li>Started</li> </ul>		Restart Nod	leManager	s
Tez			TIMELINE SERVICE V1.5	RESOURCEMANAGER	TIMELINE SERVICE V2.0 READER	YARN REGISTRY DNS		<ul> <li>Move Timel</li> </ul>	ine Service	2 V1.5
• Hive			2/2 Started	1 Installed	The forest			Move Resou	IrceManag	Jer
• HBase			NODEMANAGERS	YARN CLIENT				<ul> <li>Move Timel</li> <li>Move YARN</li> </ul>	Registry [	V2.0 Reader
Pig							2 [	↑ Enable Reso	ourceMana	iger HA
-			1h 21m 58s				-	C Run Service	Check	
Зцоор			RESOUNCEMAINAGEN OF TIME					Turn On Ma	Intenance	Mode
• Oozie			NODEMANACEDS STATUS					Download C	lient Confi	gs
<ul> <li>ZooKeeper</li> </ul>			3	0	0	0		× Delete Servi	ce	
• Ambari Metrics 🛛 🥹			Active	Lost	Unhealthy	Rebooted				

#### ResourceManager HA

### 6.2.2 Get Started – คลิก Next เริ่ม Wizard ResourceManager HA Wizard

		Ambari - test - M	ozilla Firefox						8
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>E</u>	ookmarks <u>T</u> ools <u>H</u> elp								
🚯 Ambari - test 🛛 🗙 🗙	÷								
← → ⊂ ☆	03.150.243.106:8080/#/main/services/highAvailabi	ility/ResourceMana	ager/e 🛛 🏠 🔍	Search	III\ 🔊	<b>2</b>	•		Ξ
-									
Enable ResourceMana	ger HA Wizard							×	.min
Cet Started     Select Host     Arrow     Review     Configure Components	<b>Get Started</b> This wizard will walk you through enabling Resource Once enabled, you will be running a Standby Resour This allows for an Active-Standby ResourceManager You should plan a cluster maintenance window and p	Manager HA on your cl ceManager in addition t configuration that auto repare for cluster down	uster, o your Active ResourceManager. omatically performs failover. time when enabling ResourceManage	r HA.					ONS
							NEXT	<b>→</b>	
Sqoop	4m 15s	_				_	_	-	
• Oozie	RESOURCEMANAGER UPTIME								
<ul> <li>ZooKeeper</li> </ul>	NODEMANAGERS STATU	IS							
Accumulo	3 Active	0	0 Linbealthy	0 Reported					
<ul> <li>Ambari Metrics</li> </ul>									
SmartSense	0 Decommissioned								
🛤 Unete									



### 6.2.3 Select host – เลือก host ที่จะทำ ResourceManager HA

		Ambari - test - Mozilla Firefox						• • •
<u>F</u> ile <u>E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookma	arks <u>T</u> ools <u>H</u> elp							
Ambari-test × +						_		
$\left( \leftarrow \right) \rightarrow \mathbb{C}$ $( \square )$ $( \square )$ 203.150	0.243.106:8080/#/main/services/high/	Availability/ResourceManager/er •••	ତ ନ୍ଥର୍ ୁ Search		III\ INI	2	• •	17 ≡
Enable ResourceManager I	HA Wizard						_	× admin
Get Started	Select Host Select a host that will be running the addition	anal ResourceManager						TIONS
<ul><li>2 Select Host</li><li>3 Review</li></ul>	Current ResourceManager.	edge.example.com (15.5 GB, 4 cores)	•	edge.example.com (15. ResourceManager	5 GB, 4 con	es)		S
④ Configure Components	Additional ResourceManager:	master2.example.com (3.7 GB, 2 cores)	•	Timeline Service V2.0 P VARN Registry DNS HiveServer2 Oozle S ZooKeeper Server A Accumulo Monitor A Accumulo BC Metrid Activity Explorer HS Activity Analyzer	eader Hive Metas server ccumulo Ma ccumulo Tra s Collector T Server	tore ster ccer Grafana	1	x
				master1.example.com ( NameNode Timeline History Server Zook	3.7 GB, 2 C Service V1. eeper Serve	ores) 5 r		

# 6.2.4 Review : หน้าสรุปการคอนฟิกก่อนการติดตั้ง

		Ambari - test - Mozilla Firefox				- 🗢 👄 😣
<u>File E</u> dit <u>V</u> iew Hi <u>s</u> tory <u>B</u> ookmarks	<u>T</u> ools <u>H</u> elp					
Ambari-test × +						_
← → C <sup>2</sup>	3.106:8080/#/main/services/highAv	ailability/ResourceManager/er •••	🗵 🗘 🔍 Search	III\ 🔃 🖸 🗉	• •	# ≡
Enable ResourceManager HA	Wizard				>	; admin
Get Started C	eview onfirm your host selections.					TIONS
Select Host  Review  Configure Companyed a	Current ResourceManager: edge.e Additional ResourceManager: maste	xample.com r2.example.com + TO BE INSTALLED				s
	Review Configuration Changes. The following lists the configuration ch	anges that will be made by the Wizard to enable R	esourceManager HA. This information is fo	r review only and is not editable.		
	YARN				^	
	yarn.resourcemanager.ha.enabled					
	yarn.resourcemanager.ha.rm-ids	rm1,rm2				
	yarn.resourcemanager.hostname. rm1	edge.example.com				
	varn resourcemanager resource-	edge example.com:8025			_	



#### Ambari - test - Mozilla Firefox 🐼 Ambari - test × + (←) → C' @ --- 🗵 🔂 🔍 Search III\ 🔃 🖸 🕄 🕲 👬 🚍 🗊 🔏 203.150.243.106:8080/#/main/services/highAvailability/ResourceManager/er 0/3 Started 1 Installed Enable ResourceManager HA Wizard Configure Components • Please wait while ResourceManager HA is being deployed. Stop Required Services Install Additional ResourceManager 🗸 Reconfigure YARN Reconfigure HDFS Start All Services 29% **n/a** Running n/a Pending n/a Submitted **n/a** Completed

# 6.2.5 เริ่มการคอนฟิก ResourceManager HA

รอจนคอนฟิกเรียบร้อย กดปุ่ม COMPLETE เพื่อจบการติดตั้ง



### 7 Hue

Hue เป็นบริการที่ช่วยให้ผู้ใช้งานสามารถใช้ Hadoop Ecosystems ผ่านหน้าเว็บได้ ไม่ได้มีมากับ Hortonworks แต่เราสามารถติดตั้งและคอนฟิกให้ใช้งานร่วมกันได้ สามารถดาวน์โหลดและดูรายละเอียดเพิ่มเติม ได้ที่ <u>https://gethue.com/</u> และแก้คอนฟิกไฟล์ที่ไฟล์โดยตรง ดังรายละเอียดต่อไปนี้

### 7.1 ติดตั้ง Hue

ในที่นี้จะใช้การติดตั้งผ่าน docker จึงจำเป็นต้องติดตั้ง docker บนลีนุกซ์ก่อน

```
wget https://download.docker.com/linux/centos/docker-ce.repo \
        -0 /etc/yum.repos.d/docker-ce.repo
yum install -y docker-ce docker-ce-cli containerd.io
systemctl enable --now docker
```

จากนั้นจึงเริ่มการทำงานของ docker hue

```
sudo docker run -d -p 8888:8888 --name hue gethue/hue
```

Hue จะเริ่มการทำงานที่พอร์ต 8888 แต่จะยังคงใช้งานไม่ได้ ต้องคอนฟิกให้รู้จักกับบริการต่าง ๆ ของ Hadoop ก่อนว่าอะไรทำงานอยู่ที่เครื่องไหน

ในระบบนี้จะใช้วิธีคัดลอกคอนฟิกูเรชันไฟล์ออกมายังลีนุกซ์เพื่อความสะดวกในการแก้ไข แล้วจึงเริ่มการ ทำงานของ Hue docker ใหม่อีกครั้ง

```
mkdir hue
cd hue
sudo docker cp hue:/usr/share/hue/desktop/conf .
```

เริ่มการทำงาน Hue ใหม่ โดยกำหนดให้ใช้คอนฟิกไฟล์ที่กำหนด

```
docker run -d -p 8888:8888 --name hue --network=host \
    -v /root/hue/conf/hue.ini:/usr/share/hue/desktop/conf/hue.ini gethue/hue
```



#### 7.2 ปรับแต่ง Hue configuration file (/root/hue/conf/hue.ini)

1) กำหนดพอร์ตและ time zone http\_port=8888

time\_zone=Asia/Bangkok

2) กำหนดให้ Hue ใช้ MySQL เป็น DBMS (จริง ๆ เราใช้ MariaDB)

[[database]]
engine=mysql
host=192.168.2.10
port=3306
user=hue
password=xxxxxx
name=hue

3) กำหนดค่า HDFS

[hadoop]

[[[default]]] # Enter the filesystem uri **fs\_defaultfs=hdfs://mycluster** #ใส่ชื่อ nameservice ที่กำหนดตอนทำ HDFS NamenodeHA

# ระบบเราใช้ httpfs (วิธีการติดตั้งอยู่ในหัวข้อถัด ๆ ไป)

webhdfs\_url=http://master1.example.com:14000/webhdfs/v1/

4) กำหนดค่าให้ Yarn

```
resourcemanager_api_url=http://master1.example.com:8088
proxy_api_url=http://master1.example.com:8088
history_server_api_url=http://master2.example.com:19888
logical_name=yarn-cluster #ค่านี้ดูจากไฟล์คอนฟิกที่ด้วแปร
yarn.resourcemanager.cluster-id
```

[[[ha]]]
# Resource Manager logical name (required for HA)
logical\_name=yarn-cluster

# Un-comment to enable
submit\_to=True

# URL of the ResourceManager API
resourcemanager\_api\_url=http://master2.example.com:8088
proxy\_api\_url=http://master2.example.com:8088
history\_server\_api\_url=http://master2.example.com:19888



5) กำหนดค่าให้ Hlve

hive\_server\_host=edge.example.com
use\_sasl=true

6) กำหนดค่าให้ OOZIE

oozie\_url=http://edge.example.com:11000/oozie

7) กำหนดค่าให้ Solr

[search]
# URL of the Solr Server
solr\_url=http://edge.example.com:8886/solr/

8) กำหนดค่าให้ Zookeeper

host\_ports=edge.example.com:2181,master1.example.com:2181,master2.example.com: 2181

9) กำหนดค่าการ Authentication กับ LDAP

```
backend=desktop.auth.backend.LdapBackend
base_dn="dc=example,dc=com"
ldap_url=ldaps://ipa.example.com
ldap_username_pattern="uid=<username>,cn=users,cn=accounts,dc=example,dc=com"
create_users_on_login = true
sync_groups_on_login=true
search_bind_authentication=true
user_filter="objectclass=person"
user_name_attr=uid
group_filter="objectclass=posixgroup"
group_name_attr=cn
group_member_attr=groups
```

10) ปรับแต่งเมนูบนหน้าเว็บ Hue ให้ uncomment ในส่วนของ interpreters ดังต่อไปนี้

[[interpreters]]

- hive
- pig
- spark2
- mapreduce
- sqoop1
- distcp



### 7.3 ย้ายฐานข้อมูลมาใช้ MariaDB

เนื่องจากเราเริ่มการทำงานของ Hue ไปแล้ว ฐานข้อมูลจะถูกสร้างบนซอฟต์แวร์ SQLite แต่เรากำหนด ในคอนฟิกให้มาใช้ MariaDB จึงต้อง sync โครงสร้างและข้อมูลต่าง ๆ มา ทำได้ดังคำสั่งต่อไปนี้

docker exec -it hue /bin/bash build/env/bin/hue syncdb build/env/bin/hue migrate exit

เริ่มการทำงานของ Hue ใหม่ เพื่อให้ใช้คอนฟิกที่เราปรับแต่งไป

docker restart hue

#### 7.4 กำหนด proxy user hadoop สำหรับ Hue

ไปที่ Ambari เมนู HDFS --> Configs --> Advanced แล้วเลื่อนลงไปหา "Custom core-site" จาก นั้นคลิก "Add Property..." และเพิ่มรายการ hadoop.proxyuser.hue.hosts และ hadoop.proxyuser.hue.groups โดยมีค่าเป็น \* ลงไป จากนั้นจึงกด Save

hadoop.proxyuser.hue.hosts	*
hadoop.proxyuser.hue.groups	*

Add Property ...

Restart services ที่เกี่ยวข้องตามที่ Ambari กำหนด

#### 7.5 กำหนด proxy user oozie สำหรับ Hue

ที่เมนู OOZIE → Configs → Advanced → Custom oozie-site เพิ่มรายการต่อไปนี้ Custom oozie-site

oozie.processing.timezone	GMT+0700
oozie.service.ProxyUserService. proxyuser.hue.groups	*
oozie.service.ProxyUserService. proxyuser.hue.hosts	*

Restart services ที่เกี่ยวข้องตามที่ Ambari กำหนด



7.6 สร้าง home directory for hue user

su - hdfs -c "hdfs dfs -mkdir /user/hue"
su - hdfs -c "hdfs dfs -chown hue:hdfs /user/hue"
su - hdfs -c "hdfs dfs -chmod 755 /user/hue"

### 7.7 ติดตั้ง hadoop-httpfs

HTTPFS เป็นบริการ API สำหรับบริการไฟล์ Hue จะ browse ไฟล์ใน HDFS ได้ก็ต้องมีบริการนี้ หรือ บริการ WebHDFS แต่เนื่องจากในระบบนี้เราทำ Active/Standby Namenode จึงเลือกใช้ HTTPFS ที่มี คุณสมบัติรองรับการ Namenode High Availability แต่ Ambari ไม่ได้ติดตั้งมาให้จึงต้องติดตั้งเอง ในระบบนี้จะ ทำที่เครื่องที่ทำหน้าที่ namenode คือเครื่อง edge, master1 มีขั้นตอนดังต่อไปนี้

yum install -y hadoop-httpfs

สร้าง httpfs-signature.secret

echo "1q2w3e4r" > /usr/hdp/3.1.4.0-315/hadoop/conf/httpfs-signature.secret

สร้าง systemd service สำหรับ httpfs ด้วยการสร้างไฟล์ /usr/lib/systemd/system/hadoophttpfs.service มีเนื้อความดังต่อไปนี้

```
[Unit]
Description = Start Hadoop HttpFS service
After = network.target
[Service]
Type=forking
Restart=always
User=hdfs
ExecStart = /bin/hdfs --daemon start httpfs
[Install]
WantedBy = multi-user.target
```

สร้างไดเรกทอรี temp และกำหนดสิทธิ์

```
mkdir /usr/hdp/3.1.4.0-315/hadoop-hdfs/temp
chmod 1777 /usr/hdp/3.1.4.0-315/hadoop-hdfs/temp
```

reload systemctl เพิ่มให้รู้จัก service ใหม่ และสั่งให้บริการ hadoop-httpfs เริ่มการทำงานทุกครั้งที่ เปิดเครื่องพร้อมทั้งเริ่มทำงานทันที

```
systemctl daemon-reload
systemctl enable --now hadoop-httpfs.service
```



กำหนด proxy user oozie สำหรับ Hue ให้สร้างไฟล์

/usr/hdp/current/hadoop-client/conf/httpfs-site.xml มีเนื้อความดังต่อไปนี้

เริ่มการทำงานของบริการ HTTPFS ใหม่อีกครั้ง

systemctl restart hadoop-httpfs

จากนั้นจึงเข้าใช้งานเข้าใช้งาน Hue ที่หน้าเว็บ <u>http://edge.example.com:8888</u>



#### 8 Zeppelin authentication configuration

การปรับแต่งให้ Zeppelin ยืนยันตัวตนผ่าน LDAP จะต้องเข้าไปคอนฟิกใน Ambari โดยให้ไปที่ Zeppelin Notebook → CONFIGS แล้วเลือกไปที่ "Advanced zeppelin-shiro-ini" แก้ไขในกรอบที่เขียนว่า "shiro\_ini\_content"

1. remark all in [users] section

```
[users]
# List of users with their password allowed to access Zeppelin.
# To use a different strategy (LDAP / Database / ...) check the shiro doc at ..
# admin = $shiro1$SHA-256$500000$p6Be9+t2hdUXJQj2D0b1fg==, admin
# user1 = $shiro1$SHA-256$500000$G2ymy/qmuZnGY6or4v2KfA==, role1, role2
# user2 = $shiro1$SHA-256$500000$aHBgiuwSgAcP3Xt5mEzeFw==, role3
# user3 = $shiro1$SHA-256$500000$nf0GzH10GbYVoxa7D010Sw==, role2
```

2. config ldap in [main] section

```
### A sample for configuring LDAP Directory Realm
ldapRealm = org.apache.zeppelin.realm.LdapGroupRealm
## search base for ldap groups (only relevant for LdapGroupRealm):
ldapRealm.contextFactory.environment[ldap.searchBase] = dc=example,dc=com
ldapRealm.userDnTemplate = uid={0},cn=users,cn=accounts,dc=example,dc=com
ldapRealm.contextFactory.authenticationMechanism = SIMPLE
```

3. Import LDAP CA Certificate to JDK

รันคำสั่งต่อไปนี้ที่เครื่องที่ Zeppelin ทำงานอยู่ เพื่อนำเข้า CA cert ของ IPA ไปยัง JDK KeyStore ด้วย คำสั่ง keytool ดังต่อไปนี้

/usr/jdk64/jdk1.8.0\_112/bin/keytool -import -trustcacerts \
 -keystore /usr/jdk64/jdk1.8.0\_112/jre/lib/security/cacerts \
 -alias IPAcert \
 -file /etc/ipa/ca.crt

ระบบจะถามรหัสผ่าน ให้ระบุว่า changeit จากนั้นให้ restart Zeppelin

Tip! ในระบบจริงควรจะเปลี่ยนรหัสผ่านของ KeyStore ตามที่เขาแนะนำว่าให้เปลี่ยนซะ (changeit) สามารถทำได้ด้วยคำสั่งต่อไปนี้

/usr/jdk64/jdk1.8.0\_112/bin/keytool -storepasswd \
 -keystore /usr/jdk64/jdk1.8.0\_112/jre/lib/security/cacerts

-ระบบจะถามรหัสผ่านเก่า แล้วให้ตั้งรหัสผ่านใหม่สองครั้ง เป็นอันเสร็จพิธี



4. ทำ Zeppelin User Impersonate โดยไปเพิ่มพารามิเตอร์เหล่านี้ โดยไปที่ Ambari เมนู HDFS --> Configs --> Advanced แล้วเลื่อนลงไปหา "Custom core-site" จากนั้นคลิก "Add Property..."

hadoop.proxyuser.zeppelin.groups	*
hadoop.proxyuser.zeppelin.users	*
hadoop.proxyuser.zeppelin.hosts	*
Add Property	

จากนั้น Restart service ที่เกี่ยวข้องตามที่ Ambari กำหนด

•	admin 🗸
About Zepp	pelin
Interpreter	Ja
Notebook F	Repos
Credential	
Helium	
Configurati	on
Logout	

5. ปรับแต่ง Zeppelin Interpreter โดยไปที่หน้าเว็บ Zeppelin แล้วล็อกอินด้วยบัญชีผู้ใช้ admin จากนั้นไปที่เมนู admin มุมบนขวา แล้วเลือกที่ Interpreter

6. ไปที่ spark2 แล้วเลือก Per User, Isolated และ User Impersonate จากนั้นจึง save



เป็นอันเสร็จพิธี



#### 9 Ranger authentication with LDAP

#### 9.1 Import LDAP CA cert

รันคำสั่งต่อไปนี้ที่เครื่องที่ Ranger ทำงานอยู่

```
/usr/jdk64/jdk1.8.0_112/bin/keytool -import -trustcacerts \
-keystore /usr/jdk64/jdk1.8.0_112/jre/lib/security/cacerts \
-alias IPAcert \
-file /etc/ipa/ca.crt
```

ระบบจะถามรหัสผ่าน ให้ระบุว่า changeit

จากนั้นให้ไปที่ Ambari → Ranger → CONFIGS แล้วค้นว่า ranger.usersync.truststore.file ให้

เปลี่ยนค่าเป็น **/usr/jdk64/jdk1.8.0\_112/jre/lib/security/cacerts** แล้วบันทึก จากนั้น restart service ตามที่ Ambari กำหนด

#### 9.2 Config Ranger authen LDAP

ที่ Ambari ให้ไปที่ Ranger → CONFIGS → RANGER USER INFO

Ranger User Info Enable User Sync Yes	กำหนดค่าต่อไปนี้ Sync Source: <b>LDAP/AD</b>
Sync Source LDAP/AD   COMMON CONFIGS USER CONFIGS GROUP CONFIGS	ที่แท็ป COMMON CONFIGS • Enable User Sync = Yes • LDAP/AD URL:
LDAP/AD URL Idap://ipa.example.com Bind User uid=system,cn=sysaccounts,cn=etc,dc=example,dc=com	<b>ldaps://ipa.example.com</b> (สามารถ ใช้ ldaps ได้)

- Bind User: **uid=system,cn=sysaccounts,cn=etc,dc=example,dc=com**
- Bind User Password:
- Incremental sync: Yes



COMMON CONFIGS USER CONFIGS GROUP CONFIGS	ที่แท็ป USER CONFIGS
I sername Attribute	• Username Attribute: <b>uid</b>
uid	<ul> <li>User Object Class: posixaccount</li> </ul>
	• User Search Base:
User Object Class	cn=accounts,dc=example,dc=com
posixaccount	• User Group Name Attribute: <b>uid</b>
User Search Base	• Group User Map Sync = <b>No</b>
cn=accounts,dc=example,dc=com	<ul> <li>Enable User Search = No</li> </ul>
COMMON CONFIGS USER CONFIGS GROUP CONFIGS	ที่แท็ป GROUP CONFIGS
Enable Group Sync	• Enable Group Sync = <b>Yes</b>
Yes	• Group Member Attribute: groups
	<ul> <li>Group Name Attribute: cn</li> </ul>
Group Member Attribute	<ul> <li>Group Object Class: posixgroup</li> </ul>
groups	Group Search Base:
Group Name Attribute	cn=groups,cn=accounts,dc=example,dc=com
cn	• Enable Group Search First = <b>No</b>
	<ul> <li>Sync Nested Groups = No</li> </ul>
Group Object Class	
posixgroup	
Group Search Base	หลังจากคอนฟิกค่าข้างต้นแล้วให้ Save แล้ว
cn=groups,cn=accounts,dc=example,dc=com	restart service ตามที่ Ambari แจ้ง

#### 1 / Services / Ranger / Summary BDUNINET 🔅 🕕 🌲 🕘 🔡 💄 admin · SUMMARY CONFIGS Summary **Quick Links** 40 Ranger Admin UI Started Started K Components RANGER ADMIN RANGER USERSYNC 1/1Started Enabled Enabled Enabled RANGER HDFS PLUGIN RANGER YARN PLUGIN RANGER HIVE PLUGIN RANGER TAGSYNCS Enabled RANGER ATLAS PLUGIN Disabled RANGER KAFKA PLUGIN Enabled RANGER HBASE PLUGIN





# จากนั้นให้ล็อกอิน

ไปที่เมนู Setting → Users/Groups

Ranger	♥Access Manager	🗅 Audit	🗿 Security Zone	Settings
				🚰 Users/Groups
Service Manag	er			🗅 Permissions 🏓

ดูที่แท็ป Users และ Groups ว่าเห็น Users และ Groups ที่มาเพิ่มไว้ใน FreeIPA หรือไม่ ถ้าเซ็ตได้ถูก ต้องจะเห็น Users และ Groups บน FreeIPA แบบนี้ก็จะสามารถกำหนดสิทธิ์การใช้งานให้กับ User และ Group เหล่านี้ได้

Range	<ul> <li>DAccess Manager</li> </ul>	🗅 Audit	🕑 Security Zone	Settings		
Users/Grou	ıps					
Users	Groups					
User List						
Q Sean	ch for your users					0
	User Name	e	Email Ad	dress	Role	User Source
	sqoop				User	External
	yarn				User	External
	mapred				User	External
	amb_ranger_admin				Admin	Internal
	demouser				User	External
	demobd				User	External



### 10 การเพิ่มบัญชีผู้ใช้และการใช้งาน

ระบบ Big Data Hadoop ชุดนี้ ได้เชื่อมโยงการตรวจสอบตัวตนผ่านโปรโตคอล LDAP กับ FreeIPA ดัง นั้นการจัดการบัญชีผู้ใช้งาน ให้ทำผ่าน FreeIPA สามารถเข้าใช้งานได้ที่ <u>https://ipa.example.com/</u>

1. การเพิ่มกลุ่มผู้ใช้งาน

การแบ่งกลุ่มผู้ใช้งานจะทำให้สะดวกในการกำหนดสิทธ์ให้กับกลุ่มผู้ใช้ ให้เข้าไปที่เมูน Groups แล้วกดปุ่ม +Add เพื่อเพิ่มกลุ่มผู้ใช้

	uthentication	Network Services	IPA Server	
ers Hosts Sen	/ices Gr	oups ID Views A	Automember 🗸	
oup categories				
er Groups	, Use	er Groups		
st Groups	Searci	h Q		2 Refresh 🗎 Delete
groups		Group name	GID	Description
		admins	1932000000	Account administrators group
		editors	1932000002	Limited admins who can edit other users
		ipausers		Default group for all users
		sysadmin	1932000054	
		trust admins		Trusts administrators group
d User Group				2. กำหนดค่า Group name แล้วกด Add
d User Group Group name *	clusterkit	1		<ul> <li>2. กำหนดค่า Group name แล้วกด Add</li> </ul>
d User Group Group name * Description	clusterkit	t		<ul> <li>2. กำหนดค่า Group name แล้วกด Adc</li> </ul>
d User Group Group name * Description	clusterkit	t		× 2. กำหนดค่า Group name แล้วกด Adc
d User Group Group name * Description	clusterkit	t		× 2. กำหนดค่า Group name แล้วกด Add
d User Group Group name * Description Group Type	Clusterkit	t OSIX 🔿 External 📀 PO	25IX	<ul> <li>× 2. กำหนดค่า Group name แล้วกด Add</li> </ul>
d User Group Group name * Description Group Type GID	Clusterkit	t OSIX () External () PO	JSIX	<ul> <li>2. กำหนดค่า Group name แล้วกด Adc</li> </ul>
Id User Group Group name * Description Group Type GID equired field	Clusterkit	t OSIX () External () PO	25IX	<ul> <li>* 2. กำหนดค่า Group name แล้วกด Add</li> </ul>

#### User Groups

Search Q			
	Group name	GID	
	admins	1932000000	
	clusterkit	1932000056	
	editors	1932000002	
	ipausers		
	sysadmin	1932000054	
	trust admins		
Sho	wing 1 to 6 of 6 entries.		

3. เพิ่มกลุ่มสำเร็จ จะพบชื่อกลุ่มพร้อม GID



	https://freeipa. <b>uni.net.th</b> /ipa/ui/#/e/user/sear	th ☆ 🛛	Q Search	luv :	🕅 🖸 🍕 🕲 🛃	<del>\$</del> ල හ ≡			
ROCKY IDENTITY MANAGEMENT									
Identity Policy Authentication Network Services IPA Server									
Users Hosts Services	Groups ID Views Automember ~	Subordinate IDs 🗸							
User categories	Active users								
Active users									
Preserved users	User login First name Last nam	ne Status	UID	Email address	Telephone Number	Job Title			
	admin Administ	rator 🗸 Enabled	1626600000						
	Showing 1 to 1 of 1 entries.								
Add User		×	5. เพิ่มข้อ แล้ากเ	อมูลผู้ใช้งาน ๑ Add	ระบุข้อมูลต่อไ	ปนี้ เสร็จ			
User login	kittirak		6661 d1 1	n Auu					
First name *	Kittirak		5.1	User log	in				
Last name *	Moungmingsuk		5.2	First nar	ne				
Class			5.3	Last nan	ne				
No private group			5.4	GID เลือก	ากลุ่มที่จะใช้ผู้ใจ	ช้งานอยู่			
GID		~ Q	5.5	New Pas	ssword				
New Password	admins		5.6	Verify Pa	assword				
Verify Password	editors	k							
* Required field	sysadmin								
	Add Add and Add Another Add	and Edit Cancel							

# เพิ่มผู้ใช้งาน ให้มาที่เมนู Users แล้วกดปุ่ม +Add

 เมื่อเพิ่มข้อมูลเสร็จจะกลับมาหน้าจอแสดงรายชื่อผู้ใช้ ให้คลิกเข้าไปที่ชื่อผู้ใช้ที่เพิ่งเพิ่มเข้าไปใหม่ เพื่อปรับ แต่ง

kittirak	Kittirak	Mo	oungmingsuk	✓ Enabled	1932000001	kittirak@example.com
Login shell	/bin/bas	sh		7	7. ที่ช่อง Log /bin/sh เช่	in shell ให้เปลี่ยนจาก ป็น /bin/bash ดังรูป
2 Refresh	් Revert	1 Save	Actions ~		8. จากนั้นให่ กระบวนเ	ห้กดปุ่ม Save เป็นอันเสร็จ การเพิ่มบัญชีผู้ใช้งาน



### 11 การใช้งาน

ผู้ใช้งานสามารถเข้าใช้บริการผ่านช่องทางดังต่อไปนี้

- ssh edge.example.com
- Hue <u>http://edge.example.com:8888</u>
- Zeppelin <u>http://edge.example.com:9995</u>

### 11.1 บริการ ssh

การสร้างบัญชีผู้ใช้ผ่าน FreeIPA นั้น ระบบจะบังคับให้ผู้ใช้งานต้องกำหนดรหัสผ่านใหม่ เมื่อล็อกอินครั้ง แรกกับบริการ ssh ดังตัวอย่างต่อไปนี้

```
sh kittirak@edge.example.com
(kittirak@edge.example.com) Password:
(kittirak@edge.example.com) Password expired. Change your password now.
Current Password:
(kittirak@edge.example.com) New password:
(kittirak@edge.example.com) Retype new password:
Creating home directory for kittirak.
Last login: Mon Dec 18 12:25:19 +07 2023
[kittirak@edge ~]$
```

หลังจากล็อกอินเข้าสู่ระบบ ระบบจะแจ้งว่ารหัสผ่านหมดอายุ แล้วจะให้เราระบุรหัสผ่านเดิมอีกครั้ง จาก นั้นจึงให้กรอกรหัสผ่านใหม่สองครั้ง รหัสผ่านนี้ถูกบันทึกที่ FreeIPA ก็จะมีผลกับทุกระบบทั้งหน้าเว็บ Hue และ Zeppelin

\*\* การพิมพ์รหัสผ่านบน ssh นั้น เคอร์เซอร์จะไม่ขยับ ให้พิมพ์ไปเลย\*\*

### 11.2 Hue บริการหน้าเว็บสำหรับ Hadoop Ecosystems

เมื่อล็อกอินบนหน้าเว็บ Hue ครั้งแรก ระบบจะสร้างโฮมไดเรอทรี (Home directory) บน HDFS (/user/USERNAME) ให้โดยอัตโนมัติ หากผู้ใช้จะใช้งาน Hadoop บน command line ก็จำเป็นต้องมาล็อกอิน บนหน้าเว็บ Hue นี้ก่อนเพื่อให้มีโฮมไดเรกทอรีของตน

