



手把手教你玩转小米MINOS

小米科技 武泽胜 勇幸

Outline

- Minos简介
- 原理篇
 - Minos架构
 - Minos组件介绍
- 实践篇
 - 构建Minos
 - 布署Tank
 - 布署Supervisor
 - 使用Client
 - 集群布署准备
 - 布署Zookeeper
 - 布署HDFS
 - 布署Hbase
 - 布署Owl
- Minos Future

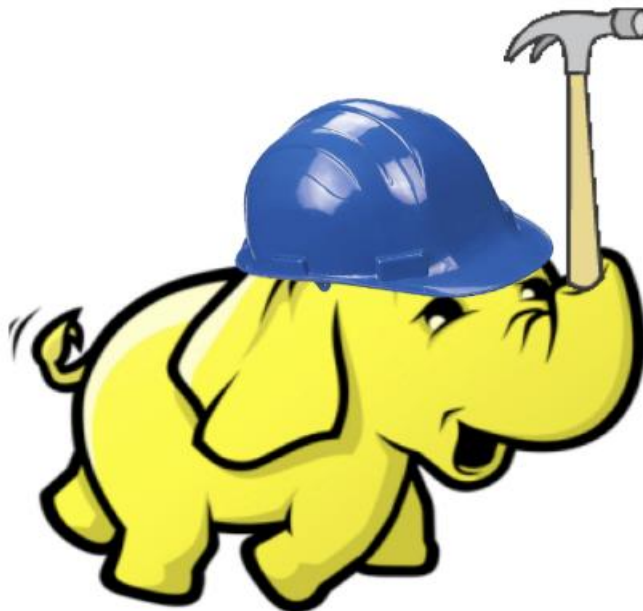


Part 1: Minos简介

Minos简介



- 目标
 - 集群级别的Hadoop配置、布署、监控系统



Minos简介



- 参考方案

- Hadoop原生脚本

- 大部分是进程级管理
 - 很多手工工作，不自动，无状态监控

- Cloudera Manager

- <http://www.cloudera.com/content/cloudera/en/products/cloudera-manager.html>
 - Cloudera商业软件，免费版: 50 nodes; 收费版: 高级feature, support
 - 安装为系统服务，难支持同机多实例
 - 比较黑盒，不方便定位错误
 - 为Hadoop Ecosystem定制，扩展支持其它服务的门槛相对比较高
 - 缺乏灵活的包版本管理，方便用官方发布包，不方便开发团队
 - 不支持metrics收集与展示

- Apache Ambari

- <http://incubator.apache.org/ambari/>
 - 由Hortonworks主导，免费、开源
 - 与Cloudera Manager 3/4/5/6类似的特点
 - 布署依赖ssh，需要用public key做无密码登录
 - Metrics通过Ganglia支持，Monitor通过Nagios支持

- 大厂通用方案

- Microsoft's Autopilot, Google's Borg, Tencent's Torca

Minos简介

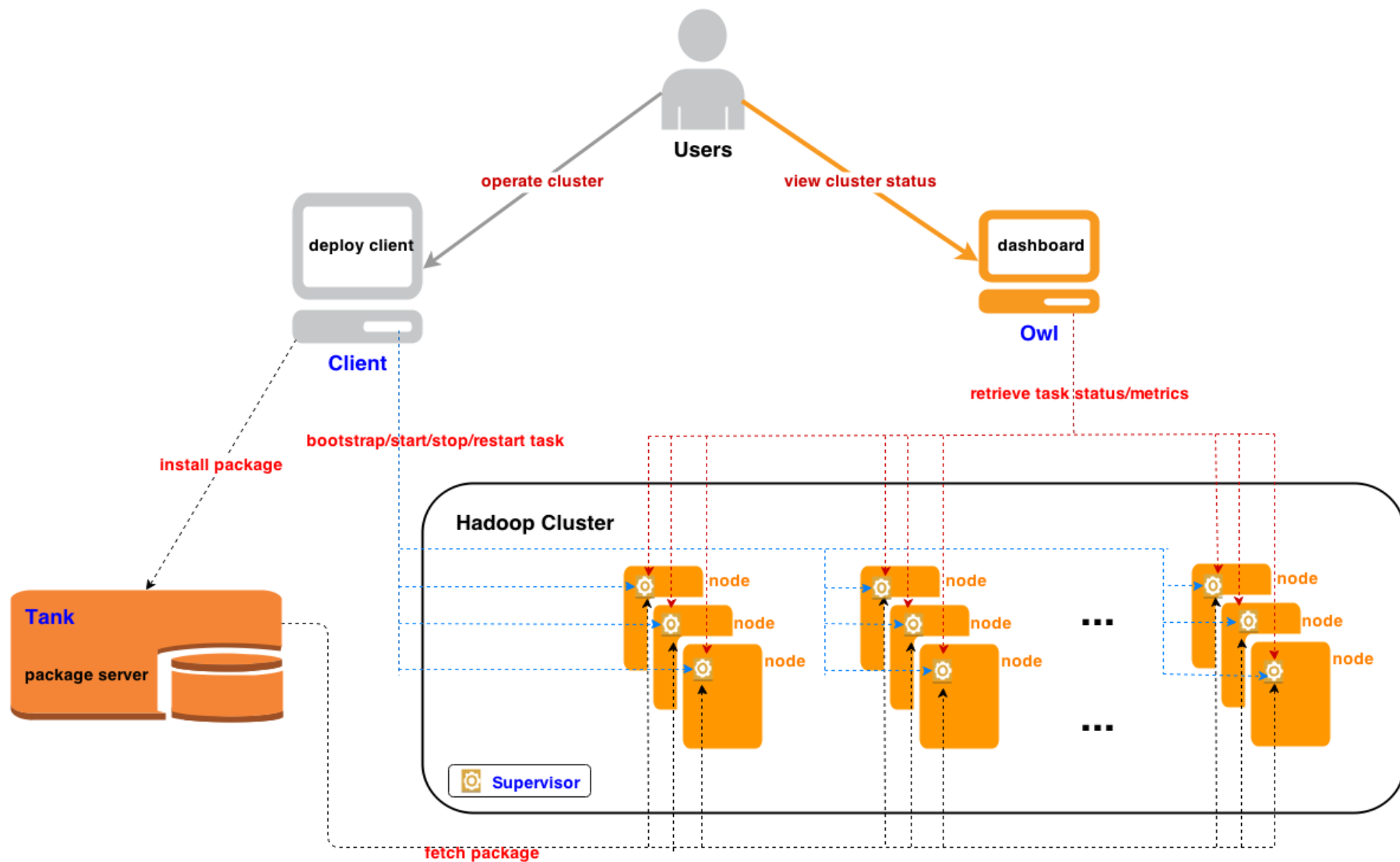


- Minos

1. <https://github.com/XiaoMi/minos>
2. 由小米大数据团队自主研发，免费、开源
3. 不强制要求布署为系统服务，能够灵活支持同机多实例
4. 灵活便捷的包管理，对开发团队更为友好
5. 直观的WebUI Dashboard，方便的Command Line Tool
6. 抽象出service/job/task的概念，直观的配置文件描述
7. 既支持集群级别的管理，也支持指定job/task级别的管理
8. 支持一键安装，方便用户使用
9. 方便扩展支持其它服务: Minos is beyond a hadoop deployment system!

Part 2: 原理篇

Minos架构



Minos组件

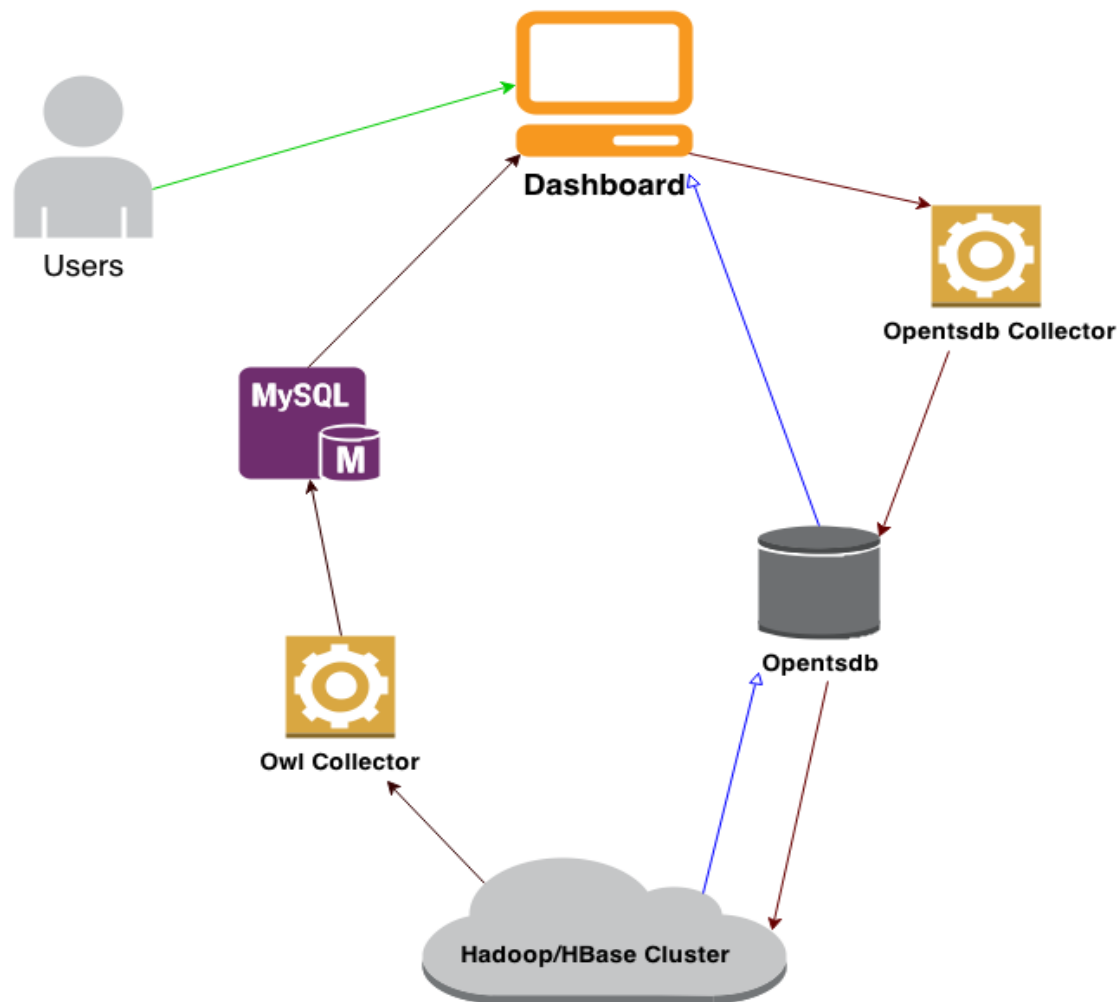


- **Client**
 - 命令行工具，集群管理入口
- **Tank**
 - 包管理服务器
 - Revision No, Timestamp, Package Name唯一标识Package
- **Supervisor**
 - 进程管理与监控服务器
 - 基于开源的Supervisor(<http://supervisord.org/>)二次开发
 - 通过xmlrpc通信，不再依赖ssh
- **Owl**
 - Metrics收集，存储，展示
 - 存储基于Opentsdb(<http://opentsdb.net/>), 具有强大的线型扩展性

Minos组件



- Owl架构



Part 3: 实践篇

构建Minos



- Prerequisites
 - Python 2.7 <http://www.python.org>
 - JDK 1.6 <http://www.oracle.com/technetwork/java/javase/downloads/index.html>
- Clone the Minos Repository
 - git clone <https://github.com/XiaoMi/minos.git>
- Build Minos

```
[work@wcc-hadoop-st01 minos]$ ./build.sh build
Creating virtual environment at /home/work/minos/build/env
New python executable in /home/work/minos/build/env/bin/python
Installing setuptools.....done.
/home/work/minos/build/env ready
2014-02-07 15:49:29 Check and install prerequisite python libraries.
2014-02-07 15:49:29 Installing configobj
Downloading/unpacking configobj==4.7.2
  Downloading configobj-4.7.2.tar.gz
  Running setup.py egg_info for package configobj

Installing collected packages: configobj
  Running setup.py install for configobj

Successfully installed configobj
Cleaning up...
2014-02-07 15:49:32 Minos client has been built.
[work@wcc-hadoop-st01 minos]$
```

部署Tank



- Start Tank

```
[work@wcc-hadoop-st01 minos]$ ./build.sh start tank --tank ip 10.237.14.236 --tank port 8000
2014-02-07 15:53:59 Building tank server.
2014-02-07 15:53:59 Check and install prerequisite python libraries.
2014-02-07 15:53:59 Installing django
Downloading/unpacking django==1.6.1
  Downloading Django-1.6.1.tar.gz (6.6MB): 6.6MB downloaded
  Running setup.py egg_info for package django

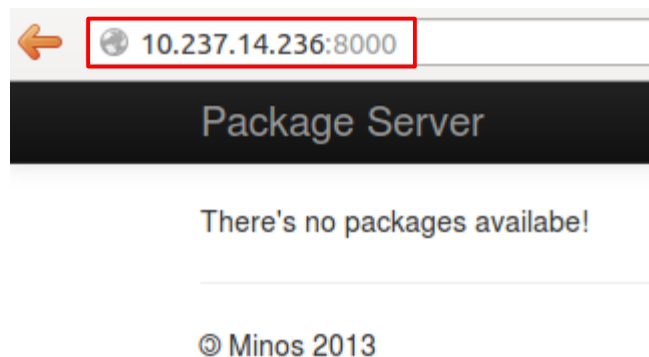
    warning: no previously-included files matching '__pycache__' found under directory '*'
    warning: no previously-included files matching '*.py[co]' found under directory '*'
Installing collected packages: django
  Running setup.py install for django
    changing mode of build/scripts-2.7/django-admin.py from 664 to 775

    warning: no previously-included files matching '__pycache__' found under directory '*'
    warning: no previously-included files matching '*.py[co]' found under directory '*'
    changing mode of /home/work/minos/build/env/bin/django-admin.py to 775
Successfully installed django
Cleaning up...
2014-02-07 15:55:58 The component tank has been built successfully.
2014-02-07 15:55:58 Starting tank server
Creating tables ...
Creating table package_server_package
Installing custom SQL ...
Installing indexes ...
Installed 0 object(s) from 0 fixture(s)
```

部署Tank



- Tank Web



- Stop Tank

```
[work@wcc-hadoop-st01 minos]$ ./build.sh stop tank  
2014-02-08 10:54:27 Stopping tank server
```

部署Supervisor

- 修改Supervisor默认配置

```
[work@wcc-hadoop-st01 ~]$ cd minos/build/template/  
[work@wcc-hadoop-st01 template]$ vi supervisord.conf.tpl
```

```
[inet_http_server]          ; inet (TCP) server disabled by default  
port=0.0.0.0:9001          ; (ip_address:port specifier, *:port for all  
username=minosuser         ; (default is no username (open server))  
password=123456            ; (default is no password (open server))
```

Supervisor webserver 默认用户名、密码及ip/port

布署Supervisor



- Prerequisites
 - 布署并启动Tank
- Start Supervisor

```
[work@wcc-hadoop-st01 minos]$ ./build.sh start supervisor --tank ip 10.237.14.236 --tank port 8006
2014-02-07 16:09:43 Building supervisor
2014-02-07 16:09:43 Check and install prerequisite python libraries.
2014-02-07 16:09:43 Installing meld3
Downloading/unpacking meld3==0.6.10
  Downloading meld3-0.6.10.tar.gz (41kB): 41kB downloaded
  Running setup.py egg_info for package meld3

Installing collected packages: meld3
  Running setup.py install for meld3

Successfully installed meld3
Cleaning up...
2014-02-07 16:09:46 Installing elementtree
Downloading/unpacking elementtree==1.2.6-20050316
  You are installing a potentially insecure and unverifiable file. Future versions of pip will default to secure files.
  Downloading elementtree-1.2.6-20050316.tar.gz (41kB): 41kB downloaded
  Running setup.py egg_info for package elementtree

Installing collected packages: elementtree
  Running setup.py install for elementtree

Successfully installed elementtree
Cleaning up...
2014-02-07 16:09:49 Installing pexpect
Downloading/unpacking pexpect==3.0
  Downloading pexpect-3.0.tar.gz (146kB): 146kB downloaded
  Running setup.py egg_info for package pexpect

Installing collected packages: pexpect
  Running setup.py install for pexpect

Successfully installed pexpect
Cleaning up...
Please input the path for deploying services:
[if you want to use the default path: /home/work, press Enter.]
:
2014-02-07 16:10:12 Creating the app root /home/work/app
2014-02-07 16:10:12 Creating the log root /home/work/log
2014-02-07 16:10:12 Creating the packages root /home/work/packages
2014-02-07 16:10:12 Deploying supervisor in /home/work/minos/supervisor
2014-02-07 16:10:12 The component supervisor has been built successfully.
2014-02-07 16:10:12 Starting supervisor
[work@wcc-hadoop-st01 minos]$
```

Hadoop cluster布署的位置
默认: 当前用户家目录

部署Supervisor



- 查看Supervisor Web

10.237.14.236:9001

Supervisor status

REFRESH

State	Description	Name	Action
running	pid 2764, uptime 15:30:14	crashmailbatch-monitor	Restart Stop Clear Log Tail -f
running	pid 2763, uptime 15:30:14	processexit-monitor	Restart Stop Clear Log Tail -f

- Stop Supervisor

```
[work@wcc-hadoop-st01 minos]$ ./build.sh stop supervisor
2014-02-08 11:20:17 Stopping supervisor
tcp      0      0 0.0.0.0:9001          0.0.0.0:*            LISTEN   2751/python
Wait for supervisor exiting...
```

Using Client

集群布署准备



- 基本约定

- 集群命名

- 集群类型: serving -> srv, processing -> prc, testing -> tst
 - Zookeeper cluster: dptst (idc + type)
 - Other cluster: dptst-example (zk + business)

- 配置文件命名

- zookeeper-dptst.cfg
 - hdfs-dptst-miliao.cfg
 - hbase-dptst-miliao.cfg

- 配置文件结构

```
hdfs/  
├─ common  
│   ├── hdfs-common.cfg  
│   ├── hdfs-prc-common.cfg  
│   ├── hdfs-srv-common.cfg  
│   └─ hdfs-tst-common.cfg  
└─ hdfs-dptst-example.cfg  
zookeeper/  
├─ common  
│   └─ zookeeper-common.cfg  
└─ zookeeper-dptst.cfg
```

集群部署准备

- 应用程序包准备
 - ZooKeeper
 - zookeeper-3.4.4
 - mvn clean package -Pdist -Dtar -DskipTests
 - HDFS
 - hadoop-2.0.0-cdh4.1.0
 - mvn clean package -Pdist -Dtar -DskipTests
 - HBase
 - hbase-0.94.3
 - mvn clean package -Dtar -DskipTests

集群布署准备

- Minos配置
 - 在作为Client端的机器上打开配置文件deploy.cfg

```
>cd minos/  
yongxing@wcc-hadoop-st01.bj ~/infra/minos $  
>vi deploy.cfg
```

- 修改supervisor和tank相应的配置

```
15 [supervisor]  
16 ; The supervisord server port  
17 server_port=9001  
18  
19 ; The supervisord username  
20 user=minosuser  
21  
22 ; The supervisord user's password  
23 password=123456  
24  
25 [tank]  
26 ; The package server host  
27 server_host=10.237.14.236  
28  
29 ; The package server port  
30 server_port=8000
```

部署Zookeeper



- Zookeeper配置

```
>cd config/conf/zookeeper/  
yongxing@wcc-hadoop-st01.bj  
>vi zookeeper-dptst.cfg
```

```
1 # deployment config for ZooKeeper  
2  
3 [cluster]  
4 # the cluster name could contain only lower case letters or numbers, and  
5 # consistent to the config file name.  
6 name=dptst  
7  
8 # version of maven artifact.  
9 version=3.4.4-mdh1.0.2  
10  
11 # all jobs below to this cluster, if more than one, separated by space.  
12 # each job has one section with the job name, as below.  
13 jobs=zookeeper  
14  
15 kerberos_username=""  
16 kerberos_realm=EXAMPLE.HADOOP  
17  
18 # the ganglia server address used to report metrics  
19 ganglia_address=10.235.2.215:8699  
20  
21 # the global log level to be configured  
22 log_level=info  
23  
24 [zookeeper]  
25 # common params for all kinds of job:  
26  
27 # zk service ports are start from base_port  
28 # Client port is base_port  
29 # Follower port is base_port + 2  
30 # Leader election port is base_port + 3  
31 # the base port must be a multiple of 100  
32 base_port=12000  
33  
34 # list all hosts here, must be IP, the id must be a non-negative integer  
35 # not necessary continuously.  
36 host.0=10.237.14.236  
37 host.1=10.237.101.56  
38 host.2=10.237.101.59  
39  
40 [[arguments]]  
41   jvm_args=''  
42     -Xmx1024m  
43     -Xms1024m  
44     -Xmn512m  
45     -XX:MaxDirectMemorySize=1024m  
46     -XX:MaxPermSize=128m  
47   ''  
48  
49 # configuration for zookeeper  
50 [configuration]  
51 # The configuration section is inherited from the base config file  
52 base=%{config_dir}/conf/zookeeper/common/zookeeper-common.cfg
```

部署 ZooKeeper



- 部署 Zookeeper 服务
 - Install

```
>./deploy.sh install zookeeper dptst
2014-02-07 16:39:38 Installing zookeeper to package server
2014-02-07 16:39:38 Uploading package: /home/yongxing/infra/zookeeper/build/zookeeper-3.4.4-mdh1.0.2.tar.gz
2014-02-07 16:39:38 Revision is: 55b7809bdca8cef23417e652364f19e025e84b79
2014-02-07 16:39:38 Generating checksum of package: /home/yongxing/infra/zookeeper/build/zookeeper-3.4.4-mdh1.0.2.tar.gz
2014-02-07 16:39:38 Checksum is: fc553735060b547c807c53343163185e646ecb8b
2014-02-07 16:39:39 Upload package /home/yongxing/infra/zookeeper/build/zookeeper-3.4.4-mdh1.0.2.tar.gz success
2014-02-07 16:39:39 Install zookeeper to package server success
{'artifact': 'zookeeper',
 'checksum': 'fc553735060b547c807c53343163185e646ecb8b',
 'package_name': 'zookeeper-3.4.4-mdh1.0.2.tar.gz',
 'revision': '55b7809bdca8cef23417e652364f19e025e84b79',
 'timestamp': '20140207-163939'}
```

← → ↻ 10.237.14.236:8000

Package Server

ID	Package Name	Revision No.	Timestamp	Checksum	Download Link
1	zookeeper-3.4.4-mdh1.0.2.tar.gz	55b7809bdca8cef23417e652364f19e025e84b79	20140207-163939	fc553735060b547c807c53343163185e646ecb8b	Download

部署ZooKeeper



- 部署Zookeeper服务
 - Bootstrap

```
>./deploy.sh bootstrap zookeeper dptst
2014-02-07 16:46:03 You should set a bootstrap password, it will be required when you do cleanup
Set a password manually? (y/n) y
Please input your password:
2014-02-07 16:46:08 Your password is: dptst-example, you should store this in a safe place, because
code used to do cleanup
2014-02-07 16:46:08 Bootstrapping task 0 of zookeeper on 10.237.14.236(0)
2014-02-07 16:46:09 Bootstrap task 0 of zookeeper on 10.237.14.236(0) success
2014-02-07 16:46:09 Starting task 0 of zookeeper on 10.237.14.236(0)
2014-02-07 16:46:09 Start task 0 of zookeeper on 10.237.14.236(0) success
2014-02-07 16:46:09 Bootstrapping task 1 of zookeeper on 10.237.101.56(0)
2014-02-07 16:46:09 Bootstrap task 1 of zookeeper on 10.237.101.56(0) success
2014-02-07 16:46:09 Starting task 1 of zookeeper on 10.237.101.56(0)
2014-02-07 16:46:09 Start task 1 of zookeeper on 10.237.101.56(0) success
2014-02-07 16:46:09 Bootstrapping task 2 of zookeeper on 10.237.101.59(0)
2014-02-07 16:46:11 Bootstrap task 2 of zookeeper on 10.237.101.59(0) success
2014-02-07 16:46:11 Starting task 2 of zookeeper on 10.237.101.59(0)
2014-02-07 16:46:11 Start task 2 of zookeeper on 10.237.101.59(0) success
```


部署ZooKeeper



- 部署Zookeeper服务
 - Show

```
>./deploy.sh show zookeeper dptst
2014-02-07 16:46:17 Showing task 0 of zookeeper on 10.237.14.236(0)
2014-02-07 16:46:17 Task 0 of zookeeper on 10.237.14.236(0) is RUNNING
2014-02-07 16:46:17 Showing task 1 of zookeeper on 10.237.101.56(0)
2014-02-07 16:46:17 Task 1 of zookeeper on 10.237.101.56(0) is RUNNING
2014-02-07 16:46:17 Showing task 2 of zookeeper on 10.237.101.59(0)
2014-02-07 16:46:17 Task 2 of zookeeper on 10.237.101.59(0) is RUNNING
```

10.237.101.59:9001

Supervisor status

REFRESH

State	Description	Name	Action
running	pid 6529, uptime 0:34:20	crashmailbatch-monitor	Restart Stop Clear Log Tail -f
running	pid 6528, uptime 0:34:20	processexit-monitor	Restart Stop Clear Log Tail -f
running	pid 6718, uptime 0:03:27	zookeeper--dptst--zookeeper	Restart Stop Clear Log Tail -f

布署ZooKeeper

- 命令使用总结
 - Pattern: `./deploy.sh $command $service $cluster`
 - Example: `./deploy.sh install zookeeper dptst`
 - Supported Commands:
 - ✓ Install: 将程序包安装到Tank服务器
 - ✓ Bootstrap: 初始化程序运行环境, 并启动程序(只需要在第一次布署时执行)
 - ✓ Show: 查看程序运行状态
 - ✓ Start: 启动程序
 - ✓ Stop: 停止程序
 - ✓ Restart: 重启程序
 - ✓ Cleanup: 清理程序运行环境, 包括集群中的数据(危险操作, 只对确定不再需要的集群做此操作)
 - ✓ Rolling_update: 逐台更新程序
 - ✓ Pack: 将程序和配置文件打包, 方便用户使用
 - ✓ Shell: 通过命令行客户端操作集群的入口

部署HDFS

- 配置HDFS
 - 修改\$minos/config/conf/hdfs/hdfs-dptst-example.cfg, 对HDFS进行配置
- 安装、启动HDFS
 - cd \$minos/client
 - ./deploy.sh install hdfs dptst-example
 - ./deploy.sh bootstrap hdfs dptst-example

部署HDFS



- 查看HDFS运行状态

```
>./deploy.sh show hdfs dptst-example
2014-02-07 17:17:52 Showing task 0 of journalnode on 10.237.14.236(0)
2014-02-07 17:17:52 Task 0 of journalnode on 10.237.14.236(0) is RUNNING
2014-02-07 17:17:52 Showing task 1 of journalnode on 10.237.101.56(0)
2014-02-07 17:17:52 Task 1 of journalnode on 10.237.101.56(0) is RUNNING
2014-02-07 17:17:52 Showing task 2 of journalnode on 10.237.101.59(0)
2014-02-07 17:17:52 Task 2 of journalnode on 10.237.101.59(0) is RUNNING
2014-02-07 17:17:52 Showing task 0 of zkfc on 10.237.14.236(0)
2014-02-07 17:17:52 Task 0 of zkfc on 10.237.14.236(0) is RUNNING
2014-02-07 17:17:52 Showing task 1 of zkfc on 10.237.101.56(0)
2014-02-07 17:17:52 Task 1 of zkfc on 10.237.101.56(0) is RUNNING
2014-02-07 17:17:52 Showing task 0 of namenode on 10.237.14.236(0)
2014-02-07 17:17:52 Task 0 of namenode on 10.237.14.236(0) is RUNNING
2014-02-07 17:17:52 Showing task 1 of namenode on 10.237.101.56(0)
2014-02-07 17:17:52 Task 1 of namenode on 10.237.101.56(0) is RUNNING
2014-02-07 17:17:52 Showing task 0 of datanode on 10.237.14.236(0)
2014-02-07 17:17:52 Task 0 of datanode on 10.237.14.236(0) is RUNNING
2014-02-07 17:17:52 Showing task 1 of datanode on 10.237.101.56(0)
2014-02-07 17:17:52 Task 1 of datanode on 10.237.101.56(0) is RUNNING
2014-02-07 17:17:52 Showing task 2 of datanode on 10.237.101.59(0)
2014-02-07 17:17:52 Task 2 of datanode on 10.237.101.59(0) is RUNNING
```

部署HDFS



- 通过Shell操作HDFS

```
>./deploy.sh shell hdfs dptst-example
help          print this help information
dfsadmin      run a DFS admin client
oev           apply the offline edits viewer to an edits file
balancer      run a cluster balancing utility
fsck          run a DFS filesystem checking utility
groups        get the groups which users belong to
fetchdt       fetch a delegation token from the NameNode
oiv           apply the offline fsimage viewer to an fsimage
jmxget        get JMX exported values from NameNode or DataNode
dfs           run a filesystem command on the file systems supported in Hadoop
haadmin       run a DFS HA admin client
getconf       get config values from configuration
yongxing@wcc-hadoop-st01.bj ~/infra/minos/client $ date:Fri 07 Feb 2014 05:20:28 PM
>./deploy.sh shell hdfs dptst-example fsck /
Connecting to namenode via http://10.237.14.236:12201
FSCK started by yongxing (auth:SIMPLE) from /10.237.14.236 for path / at Fri Feb 07
Status: HEALTHY
  Total size:      0 B
  Total dirs:      1
  Total files:     0
  Total blocks (validated):      0
  Minimally replicated blocks:  0
  Over-replicated blocks:        0
  Under-replicated blocks:       0
  Mis-replicated blocks:         0
  Default replication factor:    3
  Average block replication:     0.0
  Corrupt blocks:                0
  Missing replicas:              0
  Number of data-nodes:          3
  Number of racks:               1
FSCK ended at Fri Feb 07 17:20:38 CST 2014 in 10 milliseconds

The filesystem under path '/' is HEALTHY
```

部署HBase

- 部署Hbase

cd config/conf/hbase

vi hbase-dptst-example.cfg

```
3 [cluster]
4 # cluster name, format is [zk_service_name]-[hbase_name]
5 # The hbase name could contain only lower case letters
6 # and must be consistent to the config file name.
7 name=dptst-example
8
9 # version of maven artifact.
10 version=0.94.3-mdh1.1-SNAPSHOT
11
12 # All jobs below to this cluster, if more than one, separated by space
13 # Each job has one section with the job name, as below
14 jobs=master regionserver
15
16 kerberos_username=hbase_tst
17 kerberos_realm=EXAMPLE.HADOOP
18
19 # The HDFS cluster on which the HBase cluster depends.
20 hdfs_cluster=dptst-example
21
22 # The global log level to be configured.
23 log_level=info
24
25 [master]
26 # Client port is base_port
27 # HTTP port is base_port + 1
28 # The base port must be a multiple of 100
29 base_port=12500
30
31 # list all hosts here, must be IP, the id must be a non-negative integer
32 # not necessary continuously.
33 host.0=10.237.14.236
34 host.1=10.237.101.59
35
36 [[arguments]]
37 jvm_args='''
38 -Xmx1024m
39 -Xms1024m
40 -Xmn512m
41 -XX:MaxDirectMemorySize=1024m
42 -XX:MaxPermSize=128m
43 '''
44
45 [regionserver]
46 # Client port is base_port
47 # HTTP port is base_port + 1
48 base_port=12600
49
50 # list all hosts here, must be IP or IP/INSTANCE_NUM, the id must be a non-negative integer
51 # not necessary continuously.
52 # The regionserver can support for multiple instances on one host
53 host.0=10.237.14.236/2
54 host.1=10.237.101.59
55 host.2=10.237.101.17
```

多实例



部署HBase



- 安装、启动HBase
 - ./deploy.sh install hbase dptst-example
 - ./deploy.sh bootstrap hbase dptst-example
- 查看HBase运行状态
 - ./deploy.sh show hbase dptst-example

```
>cd client/
yongxing@wcc-hadoop-st01.bj ~/infra/minos/client $ date:Sat 08 Feb 2014 03
>./deploy.sh show hbase dptst-example
2014-02-08 15:07:12 Showing task 0 of regionserver on 10.237.14.236(0)
2014-02-08 15:07:12 Task 0 of regionserver on 10.237.14.236(0) is RUNNING
2014-02-08 15:07:12 Showing task 1 of regionserver on 10.237.14.236(1)
2014-02-08 15:07:12 Task 1 of regionserver on 10.237.14.236(1) is RUNNING
2014-02-08 15:07:12 Showing task 2 of regionserver on 10.237.101.59(0)
2014-02-08 15:07:12 Task 2 of regionserver on 10.237.101.59(0) is RUNNING
2014-02-08 15:07:12 Showing task 3 of regionserver on 10.237.101.17(0)
2014-02-08 15:07:12 Task 3 of regionserver on 10.237.101.17(0) is RUNNING
2014-02-08 15:07:12 Showing task 0 of master on 10.237.14.236(0)
2014-02-08 15:07:12 Task 0 of master on 10.237.14.236(0) is RUNNING
2014-02-08 15:07:12 Showing task 1 of master on 10.237.101.59(0)
2014-02-08 15:07:12 Task 1 of master on 10.237.101.59(0) is RUNNING
```

部署Owl

- Prerequisites

- Gnuplot

- Centos: `sudo yum install gnuplot`
 - Ubuntu: `sudo apt-get install gnuplot`

- Mysql

- Centos: `yum install mysql-server mysql mysql-devel`
 - Ubuntu:
 - `sudo apt-get install mysql-server`
 - `sudo apt-get install mysql-client`

- Attention

- Owl需要部署在作为Client端的机器上(使用同一套配置文件)

布署Owl



- 配置Owl Collector

```
[work@wcc-hadoop-st01 minos]$ cd config/owl  
[work@wcc-hadoop-st01 owl]$ vi collector.cfg
```

```
# collector config  
  
[collector]  
services=hdfs hbase 支持的服务支持  
# Period to fetch/report metrics, in seconds.  
period=10  
  
[hdfs]  
clusters=dptst-example  
jobs=journalnode namenode datanode  
# The jmx output of each bean is as following:  
# {  
#   "name" : "hadoop:service=RegionServer,name=RegionServer",  
#   "modelerType" : "org.apache.hadoop.hbase.regions",  
#   "tbl.YCSBTest.cf.test.blockCacheNumCached" : 0,  
#   "tbl.YCSBTest.cf.test.compactionBlockReadCacheHit",  
#   ...  
# }  
# Some metrics/values are from hjadoop/hbase and some from hbase environment, we specify a filter on jmx url to get  
metric_url=/jmx?qry=Hadoop:* 收集metrics的url  
  
[hbase]  
clusters=dptst-example  
jobs=master regionserver  
metric_url=/jmx?qry=hadoop:*  
  
[yarn]  
#clusters=dptst-example  
#jobs=resourcemanager nodemanager historyserver proxyserver  
#metric_url=/jmx?qry=Hadoop:*  
  
[impala]  
#clusters=dptst-example  
#jobs=statestored impalad  
#metric_url=/jmx?qry=Hadoop:*  
#need_analyze=false
```

部署Owl



- 启动Owl

```
[work@wcc-hadoop-st01 minos]$ ./build.sh start owl --local_ip 10.237.14.236
2014-02-08 16:02:59 Building owl
2014-02-08 16:02:59 Check and install prerequisite python libraries.
```

```
Please choose mysql server you intend to use:
1: Use the local mysql server
2: Use a remote mysql server
1
Please enter mysql password of the root user:root
```

选择mysql server 并输入root用户密码

```
2014-02-08 16:03:04 Configuring mysql for owl in /home/work/minos/owl/owl/settings.py
```

```
Creating tables ...
Creating table auth_permission
Creating table auth_group_permissions
Creating table auth_group
Creating table auth_user_groups
Creating table auth_user_user_permissions
Creating table auth_user
Creating table django_content_type
Creating table django_session
Creating table django_site
Creating table django_admin_log
Creating table business_business
Creating table hbase_longhaul
Creating table monitor_service
Creating table monitor_cluster
Creating table monitor_job
Creating table monitor_task
Creating table monitor_hbasecluster
Creating table monitor_regionserver
Creating table monitor_table
Creating table monitor_region
Creating table monitor_counter
Creating table monitor_quota
```

自动配置mysql、创建django数据库

```
You just installed Django's auth system, which means you don't have any superusers defined.
Would you like to create one now? (yes/no): yes
Username (leave blank to use 'work'): admin
Email address:
Password:
Password (again):
Superuser created successfully.
Installing custom SQL ...
Installing indexes ...
Installed 0 object(s) from 0 fixture(s)
```

部署Owl



- 启动Owl

```
2014-02-08 15:46:12 Setup hbase in /home/work/minos/build/download 自动部署单机hbase并启动
--2014-02-08 15:46:12-- http://www.apache.org/dist/hbase/hbase-0.94.14/hbase-0.94.14.tar.gz
Resolving www.apache.org... 140.211.11.131, 192.87.106.229, 2001:610:1:80bc:192:87:106:229
Connecting to www.apache.org|140.211.11.131|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 58436526 (56M) [application/x-gzip]
Saving to: "hbase-0.94.14.tar.gz"

100%[=====>] 58,436,526

2014-02-08 15:59:48 (70.0 KB/s) - "hbase-0.94.14.tar.gz" saved [58436526/58436526]

2014-02-08 15:59:49 Modify hbase-site.xml in /home/work/minos/build/download/hbase-0.94.14/conf
2014-02-08 15:59:49 Starting hbase
```

```
2014-02-08 16:03:18 Checkout opentsdb in /home/work/minos/build/download/opentsdb
Initialized empty Git repository in /home/work/minos/build/download/opentsdb/.git/
remote: Reusing existing pack: 4402, done. 自动部署opentsdb
remote: Total 4402 (delta 0), reused 0 (delta 0)
Receiving objects: 100% (4402/4402), 26.70 MiB | 53 KiB/s, done.
Resolving deltas: 100% (2894/2894), done.
2014-02-08 16:08:18 Compiling opentsdb in /home/work/minos/build/download/opentsdb
+ test -f configure
+ ./bootstrap
```

部署Owl



- 启动Owl

```
+ ./configure
checking for a BSD-compatible install... /usr/bin/install -c
checking whether build environment is sane... yes
checking for a thread-safe mkdir -p... /bin/mkdir -p
checking for gawk... gawk
checking whether make sets $(MAKE)... yes
checking for md5sum... /usr/bin/md5sum
checking for java... /opt/soft/jdk/bin/java
checking for javac... /opt/soft/jdk/bin/javac
checking for jar... /opt/soft/jdk/bin/jar
checking for gnuplot... /usr/bin/gnuplot
checking for javadoc... /opt/soft/jdk/bin/javadoc
checking for wget... /usr/bin/wget
checking for curl... /usr/bin/curl
configure: creating ./config.status
config.status: creating Makefile
config.status: creating opentsdb.spec
config.status: creating build-aux/fetchdep.sh
+ exec make
set dummy "http://opentsdb.googlecode.com/files" "third_party/hbase/asynchbase-1.4.1.jar"; shift;
@"
--2014-02-08 16:08:22-- http://opentsdb.googlecode.com/files/asynchbase-1.4.1.jar
Resolving opentsdb.googlecode.com... 74.125.128.82, 2404:6800:4005:c00::52
Connecting to opentsdb.googlecode.com|74.125.128.82|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 150953 (147K) [application/x-java-archive]
Saving to: "third_party/hbase/asynchbase-1.4.1.jar-t"

100%[=====] 150,953

2014-02-08 16:08:23 (292 KB/s) - "third_party/hbase/asynchbase-1.4.1.jar-t" saved [150953/150953]

set dummy "http://search.maven.org/remotecontent?filepath=com/google/guava/guava/13.0.1" "third_pa
"; shift; ./build-aux/fetchdep.sh "$@"
--2014-02-08 16:08:23-- http://search.maven.org/remotecontent?filepath=com/google/guava/guava/13.
Resolving search.maven.org... 207.223.241.72
Connecting to search.maven.org|207.223.241.72|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 1891110 (1.8M) [application/java-archive]
Saving to: "third_party/guava/guava-13.0.1.jar-t"

45% [=====] 864,228
```

自动编译opentsdb并安装第三方依赖包
由于opentsdb依赖包较多
这部分的安装需要耗费一段时间

部署Owl



- 启动Owl

```
make[1]: Leaving directory `/home/work/minos/build/download/opentsdb/build'
2014-02-08 16:20:42 Creating hbase table for opentsdb in /home/work/minos/build/download/opentsdb
HBase Shell; enter 'help<RETURN>' for list of supported commands.
Type "exit<RETURN>" to leave the HBase Shell
Version 0.94.14, r1543222, Mon Nov 18 23:23:33 UTC 2013

create 'tsdb-uid',
  {NAME => 'id', COMPRESSION => 'NONE'},
  {NAME => 'name', COMPRESSION => 'NONE'}
0 row(s) in 0.5420 seconds

create 'tsdb',
  {NAME => 't', VERSIONS => 1, COMPRESSION => 'NONE', BLOOMFILTER => 'ROW'}
0 row(s) in 1.0820 seconds

2014-02-08 16:20:47 Starting opentsdb
2014-02-08 16:20:47 Configuring opentsdb collector in /home/work/minos/config/opentsdb/metrics_colle
2014-02-08 16:20:47 Configure owl config file: /home/work/minos/config/owl/owl_config.py
2014-02-08 16:20:47 The component owl has been built successfully. owl部署成功
2014-02-08 16:20:47 Starting owl collector
nohup: appending output to `nohup.out'
2014-02-08 16:20:47 Starting opentsdb collector
nohup: appending output to `nohup.out'
2014-02-08 16:20:48 Starting owl monitor
nohup: appending output to `nohup.out'
```

自动创建hbase表

owl部署成功

owl启动

部署Owl

- 查看Owl



All clusters for all services.

name	job (running/total tasks)	cluster entry	version	description
✓ hdfs / dptst-example	✓ journalnode (3/3) ✓ namenode (2/2) ✓ datanode (3/6)	10.237.14.236:12201	2.0.0-mdh1.1-SNAPSHOT, r55b7809bdca8cef23417e652364f19e025e84b79	
✓ hbase / dptst-example	✓ master (1/2) ✓ regionserver (4/6)	10.237.101.59:12501	0.94.3-mdh1.1-SNAPSHOT, r55b7809bdca8cef23417e652364f19e025e84b79	

部署Owl



- 查看Owl

10.237.14.236:8088/monitor/cluster/1/task/

OWL- Cluster Monitor Business Longhaul admin logout

hdfs / dptst-example

Tasks Users Total

job	task id	task entry	last success	supervisor
journalnode	0	10.237.14.236:12101	Feb. 8, 2014, 4:59 p.m. ✓	10.237.14.236:9001
journalnode	1	10.237.101.17:12101	Feb. 8, 2014, 4:59 p.m. ✓	10.237.101.17:9001
journalnode	2	10.237.101.59:12101	Feb. 8, 2014, 4:59 p.m. ✓	10.237.101.59:9001
namenode	0	10.237.14.236:12201	Feb. 8, 2014, 4:59 p.m. ✓	10.237.14.236:9001
namenode	1	10.237.101.17:12201	Feb. 8, 2014, 4:59 p.m. ✓	10.237.101.17:9001
datanode	0	10.237.14.236:12401	Feb. 8, 2014, 4:59 p.m. ✓	10.237.14.236:9001
datanode	1	10.237.101.17:12401	Feb. 8, 2014, 4:59 p.m. ✓	10.237.101.17:9001
datanode	2	10.237.101.59:12401	Feb. 8, 2014, 4:59 p.m. ✓	10.237.101.59:9001

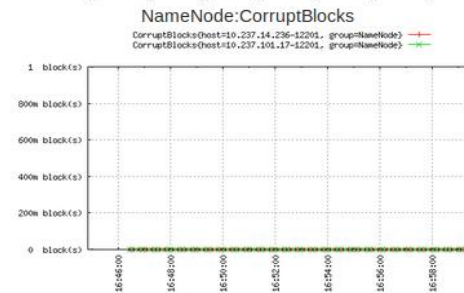
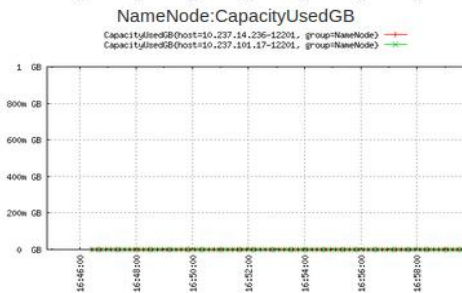
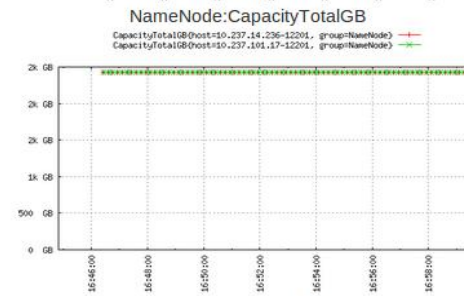
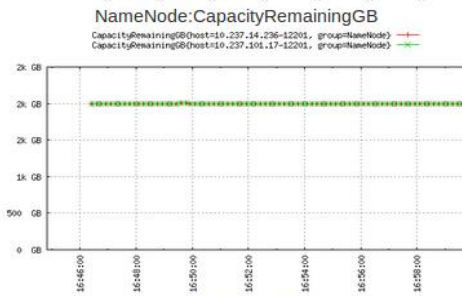
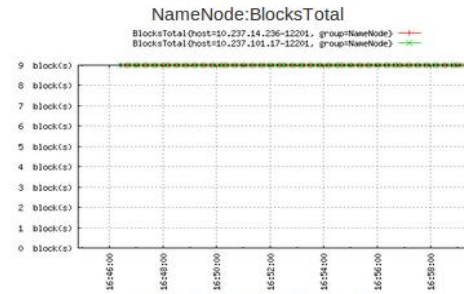
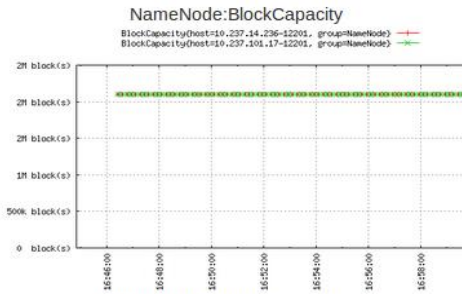
部署Owl

- 查看Owl



[hdfs](#) / [dptst-example](#) / [namenode](#) / All Tasks ▼

Overall
Operation
Rpc



部署Owl



- 查看Owl

10.237.14.236:8088/monitor/cluster/2/task/

OWL- Cluster Monitor Business Longhaul admin logout

hbase / dptst-example

Tasks Basic **Tables** RegionServers Replication

job	task id	task entry	last success	supervisor
master	0	10.237.14.236:12501	Jan. 1, 1970, 8 a.m. !	10.237.14.236:9001
master	1	10.237.101.59:12501	Feb. 8, 2014, 5:03 p.m. ✓	10.237.101.59:9001
regionserver	0	10.237.14.236:12601	Feb. 8, 2014, 5:03 p.m. ✓	10.237.14.236:9001
regionserver	1	10.237.14.236:12611	Feb. 8, 2014, 5:03 p.m. ✓	10.237.14.236:9001
regionserver	2	10.237.101.59:12601	Feb. 8, 2014, 5:03 p.m. ✓	10.237.101.59:9001
regionserver	3	10.237.101.17:12601	Feb. 8, 2014, 5:03 p.m. ✓	10.237.101.17:9001

部署Owl

- 扩展metrics

```

10.237.14.236:12201/jmx
{
  "TotalStartedThreadCount" : 382,
  "ThreadCpuTimeSupported" : true
}, {
  "name" : "Hadoop:service=NameNode,name=FSNamesystem",
  "modelerType" : "FSNamesystem",
  "tag.Context" : "dfs",
  "tag.HAState" : "active",
  "tag.Hostname" : "wcc-hadoop-st01.bj",
  "MissingBlocks" : 0,
  "ExpiredHeartbeats" : 0,
  "TransactionsSinceLastCheckpoint" : 59,
  "TransactionsSinceLastLogRoll" : 1,
  "LastWrittenTransactionId" : 256,
  "LastCheckpointTime" : 1391847389102,
  "CapacityTotalGB" : 2427.0,
  "CapacityUsedGB" : 0.0,
  "CapacityRemainingGB" : 2000.0,
  "TotalLoad" : 9,
  "BlocksTotal" : 10,
  "FilesTotal" : 25,
  "PendingReplicationBlocks" : 0,
  "UnderReplicatedBlocks" : 0,
  "CorruptBlocks" : 0,
  "ScheduledReplicationBlocks" : 0,
  "PendingDeletionBlocks" : 0,
  "ExcessBlocks" : 0,
  "PostponedMisreplicatedBlocks" : 0,
  "PendingDataNodeMessageCount" : 0,
  "MillisSinceLastLoadedEdits" : 0,
  "BlockCapacity" : 2097152,
  "TotalFiles" : 25
}, {
  "name" : "java.util.logging:type=Logging",
  "modelerType" : "java.util.logging.Logging",

```

Part 4: Minos Future

Minos Future

- 支持异构机型
- 端口动态分配
- Owl性能优化
- 多线程布署
- 支持远程操控集群机器Shell

Q & A

- Any Questions?
- Contact Me:
 - Mail: wuzesheng@xiaomi.com
 - Weibo: @wuzesheng
 - Mail: yongxing@xiaomi.com
 - Weibo: @勇幸aha