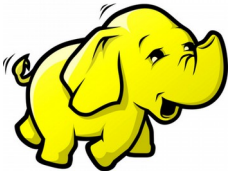


Cloudera Administrator Apache Hadoop

Parte 04-2 Manutenção, Monitoramento e Troubleshooting



Marco Reis
<http://marcoreis.net>

Agenda

- Manutenção
- Monitoramento
- Troubleshooting

Manutenção



Adicionando discos - preparação



- Adicionar ou remover discos é uma tarefa comum em um cluster
 - A sugestão é usar ao menos 2 discos em cada host
- O padrão sugerido pela Cloudera é `/data/n/`, onde `n` é o número do disco
 - No DataNode a sugestão é usar o diretório `/data/n/dfs/dn/`
 - No NameNode a sugestão é usar `/data/n/dfs/nn/`
- Vamos adicionar um novo disco (`/data/2/`). Os passos são:
 - Criar o mapeamento no Linux e configurar as permissões
 - Configurar o novo diretório no CDH
 - Verificar a integridade e a nova capacidade do HDFS
- Exemplo:
 - `$ mkdir -p /data/2`
 - `$ chown -R hdfs:hdfs /data/2`
- Monte o novo disco no diretório `/data/2` e não se esqueça do `/etc/fstab`
- Realize a manutenção em um disco de cada vez!!!

Adicionar o novo disco

- Para adicionar um novo disco no datanode01:
 - Vá no serviço HDFS → Instances → datanode01 → Configuration
- Procure pelo parâmetro dfs.datanode.data.dir e adicione o novo disco
- Salve as alterações e clique no botão Actions → Refresh DataNode Configuration
- Verifique a capacidade do cluster na NameNode Web UI
- É interessante verificar o estado do HDFS com o comando:
 - `$ hdfs fsck /`
 - Verifique se não há nenhum bloco com estado de under replication, ou seja, um bloco sem o nível mínimo de replicação
 - Não altere os discos enquanto houver blocos neste estado!

The screenshot shows the Hadoop DataNode Configuration page for datanode01. The top navigation bar includes 'Status', 'Configuration' (selected), 'Processes', 'Commands', 'Charts Library', 'Audits', 'Log Files', 'Stacks Logs', 'DataNode Web UI', and 'Quick Links'. A search bar contains 'dfs.datanode.data.dir'. The main content area is divided into two sections: 'DataNode Data Directory' and 'DataNode Data Directory Permissions'. The 'DataNode Data Directory' section shows the current value '/dfs/dn' and a new value '/data/2/dfs/dn' being entered. The 'DataNode Data Directory Permissions' section shows the current value '755'. A sidebar on the left lists various categories and their counts: Advanced (0), Logs (0), Main (1), Monitoring (0), Performance (0), Plugins (0), Ports and Addresses (0), Resource Management (0), and Security (1). A 'Switch to the classic layout' button is visible in the top right corner.

DataNode (Cluster 1, HDFS, datanode01) Actions Sep 4, 9:43 PM BRT

Status Configuration Processes Commands Charts Library Audits Log Files Stacks Logs DataNode Web UI Quick Links

dfs.datanode.data.dir Switch to the classic layout

Filters

CATEGORY

Advanced	0
Logs	0
Main	1
Monitoring	0
Performance	0
Plugins	0
Ports and Addresses	0
Resource Management	0
Security	1

DataNode Data Directory

DataNode (datanode01) ✕

dfs.data.dir /dfs/dn

dfs.datanode.data.dir /data/2/dfs/dn

DataNode Data Directory Permissions

DataNode (datanode01) ✕

dfs.datanode.data.dir.perm 755

Show All Descriptions

Opcional: remover/alterar um disco

- Remover o disco /dfs e substituí-lo por /data/1 (imagem)
- Salve as configurações e execute o balancer para recuperar os blocos perdidos

DataNode Data Directory
dfs.data.dir,
dfs.datanode.data.dir

DataNode (datanode01) ✕



Balanceamento intra-datanode



- Esta é uma funcionalidade recente, na qual um DataNode com múltiplos discos pode balancear os blocos internamente
- Especialmente importante quando um novo disco é adicionado, pois a capacidade do host é medida pelo total disponível e não pela quantidade de cada disco individualmente
- Para habilitar o recurso adicione a configuração indicada no HDFS (imagem)
 - `dfs.disk.balancer.enabled=true`
- Verifique a utilização dos discos antes do balanceamento:
 - `$ df -h`
- Vamos começar o procedimento. Primeiro temos de nos autenticar como `hdfs` no `datanode01`, no qual foi adicionado um novo disco
 - `$ hdfs diskbalancer -plan datanode01.lab`
- O plano de execução foi criado no diretório do HDFS `/system/diskbalancer`. Para encontrá-lo use o comando:
 - `$ hdfs dfs -ls -R /system/diskbalancer/`
- Nosso plano está disponível em `/system/diskbalancer/2018-Sep-13-18-20-00/datanode01.lab.plan.json`. Vamos executá-lo com o comando:
 - `$ hdfs diskbalancer -execute /system/diskbalancer/2016-Aug-17-17-03-56/172.26.10.16.plan.json`
- Enquanto o balanceamento é executado podemos verificar o andamento assim:
 - `$ hdfs diskbalancer -query datanode01.lab`
- Depois de concluído, podemos verificar que os discos estão com a utilização mais equilibrada:
 - `$ df -h`
- Outro comando útil é o `report`, que mostra quais os DataNodes que precisam de balanceamento interno
 - `$ hdfs diskbalancer -report datanode01.lab`
- O resultado abaixo mostra que o `datanode02` está com alta densidade em um dos discos, enquanto que o `datanode01` já está com baixa densidade, o que era esperado depois do balanceamento. O `datanode03` só tem 1 disco e não pode ser balanceado
 - Reporting top 3 DataNode(s) benefiting from running DiskBalancer.
 - 1/3 datanode02.lab[192.168.122.13:50020] - <527890dd-6663-400d-aa7d-201a14578342>: 2 volumes with node data density 0.74.
 - 2/3 datanode01.lab[192.168.122.12:50020] - <a1b7d24b-0708-415b-9a60-17935174cf93>: 2 volumes with node data density 0.04.
 - 3/3 datanode03.lab[192.168.122.14:50020] - <7cebd4f4-0df3-4fec-bec0-8bde90c8b3be>: 1 volumes with node data density 0.00.

HDFS Service Advanced
Configuration Snippet
(Safety Valve) for hdfs-
site.xml

HDFS (Service-Wide) 🔗

Name	dfs.disk.balancer.enabled
Value	true
Description	Description
<input type="checkbox"/> Final	











Host template

- O Host Template é um recurso usado para facilitar a adição de novos nós no cluster
- Vamos criar um template para novos DataNodes, com os serviços que precisamos neste tipo de host
- Selecione o menu superior Hosts → Host Template → Create
 - Template Name: datanode-host
 - Serviços deste template:
 - HDFS: DataNode
 - Hive: Gateway
 - Impala: Impala Daemon
 - Spark: Gateway
 - YARN: NodeManager

Create New Host Template For Cluster 1

Template Name

Select Role Groups to Include:

>  Flume	
>  HDFS	DataNode Group 1
>  Hive	Gateway Default Group
>  Hue	
>  Impala	Impala Daemon Default Group
>  Oozie	
>  Spark	Gateway Default Group
>  Sqoop 1 Client	
>  YARN (MR2 Include...	NodeManager Group 1
>  ZooKeeper	



Preparação do SO

- O host precisa das mesmas configurações dos outros nós, incluindo o Kerberos
- Atenção para a configuração da rede e do SO
- A seguir estão listados os comandos necessários para todos os hosts do cluster, na forma de um guia rápido
 - `$ sudo echo 10 > /proc/sys/vm/swappiness`
 - `$ sudo sysctl -w vm.swappiness=10`
 - `$ wget 'https://archive.cloudera.com/cm5/ubuntu/xenial/amd64/cm/cloudera.list' -O /etc/apt/sources.list.d/cloudera.list`
 - `$ wget http://archive.cloudera.com/cdh5/ubuntu/xenial/amd64/cdh/archive.key`
 - `$ apt-key add archive.key`
 - `$ apt-get update`
 - `$ apt-get -y install oracle-j2sdk1.7 libmysql-java ntp`
 - `$ apt-get install krb5-user -y`
- Copiar a chave a partir do HeadNode:
 - `$ ssh-copy-id -i ~/.ssh/id_rsa.pub datanode04.lab`

Observações sobre os hosts



- Preferencialmente, os hosts devem ter a mesma configuração
 - Memória, CPU e Disco
- Também é importante que as versões do CDH sejam iguais em todos os hosts, ou seja, deve-se evitar de ter versões diferentes dentro do mesmo cluster
- Caso seja necessário usar um host com configuração diferente, crie um grupo de funções (Role Groups) para este tipo de máquina
 - Exemplo: criar um Role Group para hosts com 4 discos de 1 TB e 128 GB
 - Os Role Groups estão disponíveis em todos os serviços, na aba Configuration → botão Role Groups

Adicionar novo host no cluster

- A adição de novos hosts é uma operação natural dentro do cluster
- O Cloudera Manager permite a administração dos hosts e de seus serviços de uma forma bastante simples por meio de wizards
- O wizard para adicionar um host está disponível no menu superior Hosts → All Hosts → botão Add New Hosts to Cluster

Add Hosts Wizard

Use this installer to install CDH and the Cloudera Manager Agent on a host so that it can be added to your cluster.

If you are using a different mechanism to install CDH and the Cloudera Manager Agent, then you do not need to use this wizard. The host will be listed in the Cloudera Manager Admin Console after the Cloudera Manager Agent contacts the Cloudera Manager Server.

If the cluster is using Kerberos authentication, follow the steps documented here: [Enabling Kerberos Authentication Using the Wizard](#) to configure the new host for Kerberos. Failure to do so will prevent services on the new host from functioning.

Especifique o host

- Este wizard é similar à instalação inicial do cluster, assim, informe o nome do host para continuar a operação

Specify hosts for your CDH cluster installation.

Hint: Search for hostnames and IP addresses using [patterns](#) 

1 hosts scanned, 1 running SSH. [New Search](#)

<input checked="" type="checkbox"/> Expanded Query	Hostname (FQDN)	IP Address	Currently Managed	Result
<input checked="" type="checkbox"/> datanode04.lab	datanode04.lab	192.168.122.16	No	✔ Host ready: 1 ms response time.

- Selecione o repositório, como na imagem abaixo

Add New Hosts to Cluster

Select Repository

Select the specific release of the Cloudera Manager Agent you want to install on your hosts.

☒ Matched release for this Cloudera Manager Server

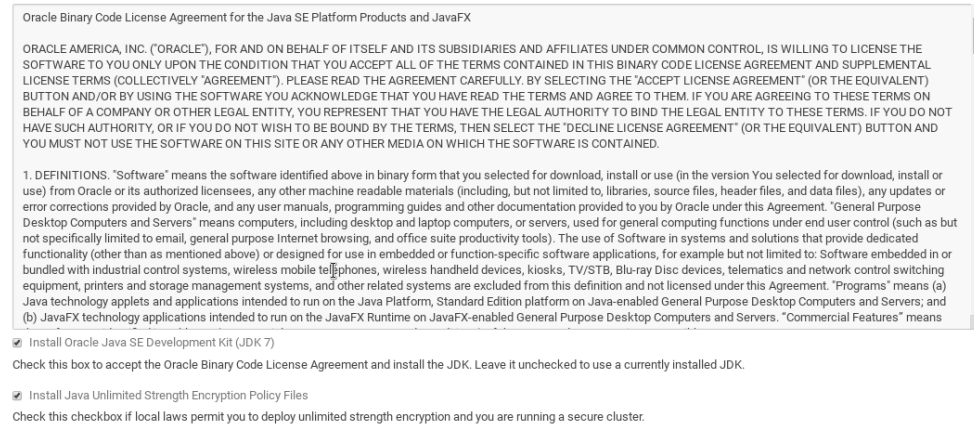
☐ Custom Repository

JDK

- O JDK a ser instalado é o padrão sugerido pelo wizard
 - Veja que é possível selecionar o JCE para o Kerberos com a opção Install Java Unlimited Strength Policy Files

Add New Hosts to Cluster

Accept JDK License



Oracle Binary Code License Agreement for the Java SE Platform Products and JavaFX

ORACLE AMERICA, INC. ("ORACLE"), FOR AND ON BEHALF OF ITSELF AND ITS SUBSIDIARIES AND AFFILIATES UNDER COMMON CONTROL, IS WILLING TO LICENSE THE SOFTWARE TO YOU ONLY UPON THE CONDITION THAT YOU ACCEPT ALL OF THE TERMS CONTAINED IN THIS BINARY CODE LICENSE AGREEMENT AND SUPPLEMENTAL LICENSE TERMS (COLLECTIVELY "AGREEMENT"). PLEASE READ THE AGREEMENT CAREFULLY. BY SELECTING THE "ACCEPT LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND/OR BY USING THE SOFTWARE YOU ACKNOWLEDGE THAT YOU HAVE READ THE TERMS AND AGREE TO THEM. IF YOU ARE AGREEING TO THESE TERMS ON BEHALF OF A COMPANY OR OTHER LEGAL ENTITY, YOU REPRESENT THAT YOU HAVE THE LEGAL AUTHORITY TO BIND THE LEGAL ENTITY TO THESE TERMS. IF YOU DO NOT HAVE SUCH AUTHORITY, OR IF YOU DO NOT WISH TO BE BOUND BY THE TERMS, THEN SELECT THE "DECLINE LICENSE AGREEMENT" (OR THE EQUIVALENT) BUTTON AND YOU MUST NOT USE THE SOFTWARE ON THIS SITE OR ANY OTHER MEDIA ON WHICH THE SOFTWARE IS CONTAINED.

1. DEFINITIONS. "Software" means the software identified above in binary form that you selected for download, install or use (in the version You selected for download, install or use) from Oracle or its authorized licensees, any other machine readable materials (including, but not limited to, libraries, source files, header files, and data files), any updates or error corrections provided by Oracle, and any user manuals, programming guides and other documentation provided to you by Oracle under this Agreement. "General Purpose Desktop Computers and Servers" means computers, including desktop and laptop computers, or servers, used for general computing functions under end user control (such as but not specifically limited to email, general purpose Internet browsing, and office suite productivity tools). The use of Software in systems and solutions that provide dedicated functionality (other than as mentioned above) or designed for use in embedded or function-specific software applications, for example but not limited to: Software embedded in or bundled with industrial control systems, wireless mobile telephones, wireless handheld devices, kiosks, TV/STB, Blu-ray Disc devices, telematics and network control switching equipment, printers and storage management systems, and other related systems are excluded from this definition and not licensed under this Agreement. "Programs" means (a) Java technology applets and applications intended to run on the Java Platform, Standard Edition platform on Java-enabled General Purpose Desktop Computers and Servers; and (b) JavaFX technology applications intended to run on the JavaFX Runtime on JavaFX-enabled General Purpose Desktop Computers and Servers. "Commercial Features" means

☒ Install Oracle Java SE Development Kit (JDK 7)

Check this box to accept the Oracle Binary Code License Agreement and install the JDK. Leave it unchecked to use a currently installed JDK.

☒ Install Java Unlimited Strength Encryption Policy Files

Check this checkbox if local laws permit you to deploy unlimited strength encryption and you are running a secure cluster.

Processo de instalação

- O wizard vai baixar os arquivos necessários de acordo com a opção de instalação selecionada, no caso, a Parcels

Add New Hosts to Cluster

Install Parcels

The selected parcels are being downloaded and installed on all the hosts in the cluster.



- O novo host precisa configurar o Kerberos para acesso seguro ao cluster e iniciar os serviços

Add New Hosts to Cluster

Deploy Kerberos Client Configuration Command

Status Running Context [Cluster 1](#) Sep 11, 2:17:19 PM [Abort](#)

Completed 1 of 2 step(s).

[Show All Steps](#) [Show Only Failed Steps](#) [Show Running Steps](#)

Execute global command Wait for configuration staleness computation	Sep 11, 2:17:19 PM	38ms
Deploy Kerberos Client Configuration 0/6 steps completed.	Sep 11, 2:17:20 PM	

Seleção do template

- Selecione o template datanode-host, uma vez que este será um novo DataNode
- Caso não tenha nenhum template, a seleção dos serviços do host pode ser feita posteriormente, sem problema

Add New Hosts to Cluster

Select Host Template

Select a host template to apply to the new hosts in order to populate them with role instances.

☐ None

☒ datanode-host

[Create...](#)

☒ Start newly created roles after applying the host template.

- Os serviços selecionados pelo template serão iniciados

Add New Hosts to Cluster

Start Roles on Hosts When Free Command

Status Running Sep 11, 10:25:41 PM [Abort](#)

Completed 3 of 5 step(s).

☒ Show All Steps ☐ Show Only Failed Steps ☐ Show Running Steps

Wait for Service Commands	Hive	Sep 11, 10:25:41 PM	1.84s
Wait for Service Commands	Impala	Sep 11, 10:25:41 PM	1.57s
Wait for Service Commands	Spark	Sep 11, 10:25:41 PM	1.66s
Wait for Service Commands	YARN (MR2 Included)	Sep 11, 10:25:41 PM	Abort
Wait for Service Commands	HDFS	Sep 11, 10:25:41 PM	Abort

Novo host no cluster

- O novo host foi incluído com sucesso no cluster

Actions for Selected ▾ Columns: 10 Selected ▾

Status	Name	IP	Roles	Commission State	Last Heartbeat	Load Average	Disk Usage	Physical Memory	Swap Space
	datanode01.lab	192.168.122.12	5 Role(s)	Commissioned	521ms ago	0.32 0.12 0.08	<div><div>25 GiB / 77.5 GiB</div></div>	<div><div>762.7 MiB / 3.9 GiB</div></div>	<div><div>0 B / 980 MiB</div></div>
	datanode02.lab	192.168.122.13	5 Role(s)	Commissioned	481ms ago	0.08 0.07 0.08	<div><div>26.7 GiB / 77.5 GiB</div></div>	<div><div>752.8 MiB / 3.9 GiB</div></div>	<div><div>0 B / 980 MiB</div></div>
	datanode03.lab	192.168.122.14	6 Role(s)	Commissioned	435ms ago	0.09 0.08 0.09	<div><div>21.3 GiB / 38.3 GiB</div></div>	<div><div>905.7 MiB / 3.9 GiB</div></div>	<div><div>0 B / 980 MiB</div></div>
	datanode04.lab	192.168.122.16	5 Role(s)	Commissioned	8.84s ago	0.03 0.09 0.06	<div><div>10.6 GiB / 38.3 GiB</div></div>	<div><div>1 GiB / 3.9 GiB</div></div>	<div><div>908 KiB / 980 MiB</div></div>
	edgenode.lab	192.168.122.11	16 Role(s)	Commissioned	453ms ago	0.23 0.29 0.50	<div><div>8.3 GiB / 38.3 GiB</div></div>	<div><div>2.6 GiB / 3.9 GiB</div></div>	<div><div>0 B / 980 MiB</div></div>
	headnode.lab	192.168.122.10	9 Role(s)	Commissioned	469ms ago	0.86 1.12 1.16	<div><div>16.7 GiB / 38.3 GiB</div></div>	<div><div>5.8 GiB / 11.7 GiB</div></div>	<div><div>0 B / 980 MiB</div></div>

- Verifique na NameNode Web UI a nova configuração
 - Opcionalmente, você pode executar o balanceamento do cluster

Node	Last contact	Capacity	Blocks	Block pool used	Version
datanode01.lab (192.168.122.12:1004)	Tue Sep 11 15:07:47 -0300 2018	34.6 GB <div><div></div></div>	826	12.99 GB (37.55%)	2.6.0-cdh5.15.1
datanode02.lab (192.168.122.13:1004)	Tue Sep 11 15:07:47 -0300 2018	70.84 GB <div><div></div></div>	1021	16.48 GB (23.27%)	2.6.0-cdh5.15.1
datanode03.lab (192.168.122.14:1004)	Tue Sep 11 15:07:47 -0300 2018	34.6 GB <div><div></div></div>	390	13.3 GB (38.43%)	2.6.0-cdh5.15.1
datanode04.lab (192.168.122.16:1004)	Tue Sep 11 15:07:49 -0300 2018	34.6 GB <div><div></div></div>	0	24 KB (0%)	2.6.0-cdh5.15.1

Showing 1 to 4 of 4 entries

Previous 1 Next

Gerenciamento de funções dos hosts

- Cada host tem uma lista de funções (roles) que pode sofrer alterações durante a operação do cluster
- A configuração do cluster é dinâmica e muda de acordo com a carga de trabalho
 - Exemplo: um DataNode pode ter adicionada uma nova função, como Gateway do Hive, ZooKeeper, HBase etc.
- Para selecionar as funções do host deve-se considerar sua capacidade, tanto CPU quanto memória
- A seleção das funções, tanto para adicionar quanto para remover, é feita na aba Instances do serviço (HDFS, YARN, Impala) → Add Role Instances → selecione as novas funções e os respectivos hosts

Add Role Instances to HDFS

Assign Roles

You can specify the role assignments for your new roles here.

You can also view the role assignments by host.

[View By Host](#)

 Failover Controller

[Select hosts](#)

 SecondaryNameNode x 1

[Select hosts](#)

 NFS Gateway

[Select hosts](#)

 HttpFS x 1

[Select hosts](#)

 NameNode x 1

[Select hosts](#)

 Gateway

[Select hosts](#)

 JournalNode

[Select hosts](#)

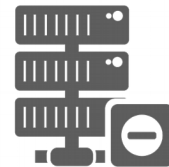
 DataNode x 1

[Select hosts](#)

Considerações sobre os hosts



- A seleção das funções do host deve considerar a capacidade total do cluster, a carga de trabalho e as expectativas dos usuários em relação à performance
- Não deve-se esperar que um sistema de big data resolva problemas de performance sozinho. Na verdade, um Data Warehouse pode ser tão eficiente em termos de performance quando um cluster de big data. A diferença está na relação entre custo e escala, porque quando a escala aumenta, o custo do cluster tende a ser (muito) menor, considerando licenças, treinamento, storage e servidores
 - A escolha de ferramentas é essencial para cada tipo de processamento
 - Exemplo: ferramentas para streaming não devem ser usadas para batch
- A carga de trabalho pode ser intensiva em relação ao I/O, memória, rede e CPU, desta forma, os hosts devem ser especializados em uma função, evitando misturar tipos de processamento diferentes para que a concorrência não prejudique a performance
 - Exemplo 1: um DataNode é intensivo em relação ao I/O e não deveria ser usado para streaming, que é intenso em memória e rede
 - Exemplo 2: um HeadNode é intensivo em CPU e memória, assim, não deveria ser usado como DataNode
- Algumas sugestões de funções para o host:
 - HeadNode: serviços master do Hadoop
 - EdgeNode: serviços redundantes
 - WorkerNode (HBase Region Server, DataNode, Impala Daemon, NodeManager etc.)
 - Flume Agent: entrada de dados no HDFS
 - Kafka Broker: processamento em tempo real
 - Gateway: hosts clientes



Decommission

- A decommission é uma remoção temporária ou permanente de um host do cluster
- Para remover um datanode com segurança do cluster é preciso desabilitá-lo com a opção Decommission
 - Serviço HDFS → aba Instances → seleção do DataNode → botão Actions for Selected → Decommission
- Todos os serviços do host serão parados
- A decommission pode ser feita em qualquer serviço (HDFS, YARN, etc.)
- Se a decommission for no HDFS, o fator de replicação deve suportar a quantidade de hosts após a desabilitação
 - Após a desabilitação, o número de hosts deve ser igual ou maior ao fator de replicação
- Essa operação deve ser feita em um host de cada vez
- Se for apenas para manutenção, o host pode ser habilitado no mesmo menu com a opção Recommission

Snapshot

- O Snapshot é uma funcionalidade usada em conjunto com a replicação para proteção dos dados
- Podem ocorrer erros durante replicações demoradas e para ajudar neste caso podemos criar snapshots, que é na prática a cópia do diretório
- Os snapshots podem ajudar a melhorar a performance, porque o HDFS sempre cria um snapshot antes da replicação
- Diretórios importantes, como as bases do Hive, devem ser protegidos com este recurso
 - No File Browser, selecione o diretório /user/hive/warehouse e clique no botão Enable Snapshots
 - O snapshot é gravado no mesmo caminho, no subdiretório .snapshot, por exemplo /user/hive/warehouse/.snapshot

HDFS (Cluster 1) Actions Sep 7, 1:36 AM BRT

Status Instances Configuration Commands **File Browser** Charts Library Cache Statistics Audits NameNode Web UI Quick Links

/ / user / hive / warehouse / Edit Showing 1 to 4

Name	Owner	Group	Last Modified	Size	Mode
..					-
avisaocivil.db	hive	hive	08/21/2018 11:47 PM		- drwxrwxrwx
datalake.db	dataengineer	hive	08/31/2018 9:32 AM		- drwxrwxrwx
dbdataengineer.db	dataengineer	hive	08/25/2018 1:07 AM		- drwxrwxrwx
social.db	hive	hive	08/22/2018 12:26 AM		- drwxrwxrwx

/user/hive/warehouse

Parent	/user/hive
Owner	hive
Group	hive
Mode	drwxrwxrwx
Last Modified	August 31, 2018 9:26 AM

Quota Management

Edit Quota

Snapshots

Snapshots are not enabled for this directory.

Enable Snapshots

Snapshot Policy

- Para automatizar o snapshot vamos criar uma política de criação diária
 - Menu Backup → Snapshot Policies → Create Snapshot Policy
- Preencha as informações conforme a imagem ao lado, principalmente em relação aos alertas

The image shows a web form for configuring a Snapshot Policy. The form includes the following fields and options:

- Service:** A dropdown menu with "HDFS (Cluster 1)" selected.
- Name:** A text input field containing "snapshot-hive-diario".
- Description:** A large, empty text area.
- Paths:** A dropdown menu with "/user/hive/warehouse" selected.
- Schedule:** A section with a red asterisk and a help icon. It contains radio buttons for "Hourly", "Daily", "Weekly", "Monthly", and "Yearly". The "Daily" option is selected. Below the radio buttons, there is a label "Take snapshots every day at" followed by two dropdown menus for "hour(s)" and "minute(s)", both set to "00". Below that, there is a label "Daily snapshots to keep" followed by a text input field containing "1".
- Alerts:** A section with a label "Alerts" and four checked checkboxes: "On failure", "On start", "On success", and "On abort".

Below the Paths field, there is a note: "The paths listed above are snapshottable. Navigate to the 'File Browser' tab of the relevant HDFS service page to make other paths snapshottable."

Snapshot - restauração

- O snapshot é uma cópia dos arquivos, então é possível recuperar o conteúdo do diretório com a opção Restore Snapshot
- É possível usar os comandos copy e DistCp para a restauração
 - Copy: operação de cópia sequencial dos dados, é indicada para volumes pequenos de dados dentro do próprio cluster
 - DistCp: operação de cópia distribuída usando MapReduce, é indicada para cópia de grandes volumes de dados entre clusters, como no caso de backups

Restore Snapshot

×

Select a snapshot to restore `/user/hive/warehouse` to:

Snapshot

September 8, 2018 6:30 PM — snap: ▼

Select Restore Method

☒ Use the HDFS 'copy' command.
Executes more slowly. Does not require credentials in a secure cluster. Copies the contents of the snapshot as a subdirectory or as files in the target directory.

☐ Use DistCp / MapReduce.
Executes more quickly. Requires credentials (Run As) in secure clusters. Merges the target directory with the contents of the source snapshot.

Cancel

Restore

Manutenção do host

- Paradas programadas para manutenção podem ser feitas de duas formas
 - Decommission: pode demorar bastante
 - DataNode Offline: opção para manutenção programada rápida
- A opção está disponível no menu superior Hosts → All Hosts → selecione o host → Actions for Selected → Begin Maintenance (Suppress Alerts/Decommission)

Begin Maintenance (Suppress Alerts/Decommission)

Host	Role Instances
datanode03.lab	DN G ID G NM S

Suppress Alerts
The selected host(s) will enter Maintenance Mode and all alerts will be suppressed.

☒ **Decommission Host(s)**
All roles on the selected host(s) will be decommissioned (where applicable) and stopped.
For HDFS DataNodes on the selected host(s):

☒ **Decommission DataNodes**
This action replicates all HDFS data to other hosts before stopping and **could take significant time**.
Example: Disk replacement, repurposing the hosts, ...

☐ **Take DataNode offline**
HDFS will not re-replicate data to other DataNodes for:

Minutes
This action is **the least disruptive to workloads** and is optimal for **short-term maintenance** in which all storage volumes used by HDFS will remain intact.
Example: Quick reboot, CPU/RAM upgrade, switching network cables, ...

Taking multiple DataNodes offline increases the chance that some HDFS data may become unavailable.

Manutenção com DN decommissioned

- O HDFS é rebalanceado e os blocos são copiados para os hosts ativos
- Verifique o estado do cluster com o comando:
 - `$ hdfs fsck /`
 - Se estiver com a segurança ativada tem de usar o superuser (hdfs)
- Verifique o estado do datanode na NameNode Web UI







Actions for Selected ▾						
<input type="checkbox"/>	Status	Name	IP	Roles	Commission State	
<input type="checkbox"/>	✓	datanode01.lab	192.168.122.12	5 Role(s)	Commissioned	
<input type="checkbox"/>	✓	datanode02.lab	192.168.122.13	5 Role(s)	Commissioned	
<input type="checkbox"/>	✓	datanode03.lab	192.168.122.14	6 Role(s)	Decommissioned	
<input type="checkbox"/>	✓	edgenode.lab	192.168.122.11	16 Role(s)	Commissioned	
<input type="checkbox"/>	✓	headnode.lab	192.168.122.10	9 Role(s)	Commissioned	

Node	Last contact	Capacity	Blocks	Block pool used	Version
✓ datanode01.lab (192.168.122.12:1004)	Thu Sep 06 13:20:19 -0300 2018	70.84 GB <div><div></div></div>	923	19.65 GB (27.74%)	2.6.0-cdh5.15.1
✓ datanode02.lab (192.168.122.13:1004)	Thu Sep 06 13:20:19 -0300 2018	70.84 GB <div><div></div></div>	923	19.65 GB (27.74%)	2.6.0-cdh5.15.1
⚠ datanode03.lab (192.168.122.14:1004)	Thu Sep 06 13:15:07 -0300 2018	34.6 GB <div><div></div></div>	99	6.61 GB (19.12%)	2.6.0-cdh5.15.1

Finalizar manutenção

- Para finalizar a manutenção e voltar o datanode para operação
 - Menu superior Hosts → All Hosts → seleção do datanode → Actions for Selected → Exit Maintenance Mode

End Maintenance (Enable Alerts/Recommission)

Host	Role Instances
datanode03.lab	 DN  G  ID  G  NM  S

☒ Recommission Host(s)

☐ Bring hosts online and start all roles

All decommissioned roles will be recommissioned and started.

Note: HDFS DataNodes will be started first and brought online to avoid excess replication.

☐ Bring hosts online

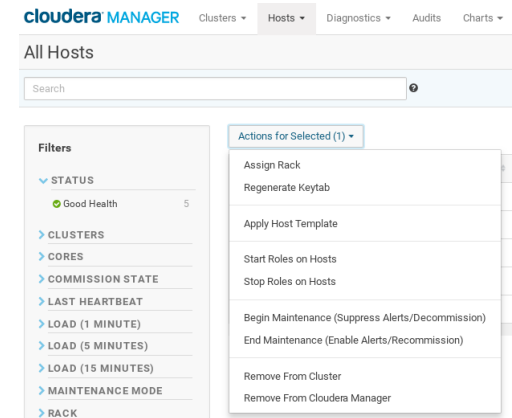
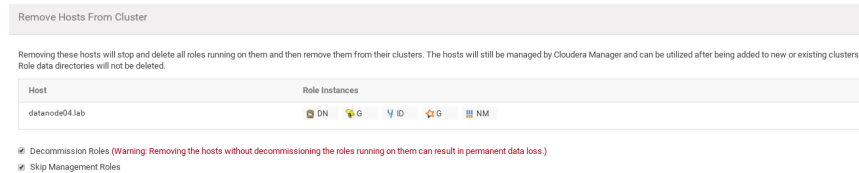
All decommissioned roles will be recommissioned but remain stopped.

☒ Re-enable All Alerts

The selected host(s) will exit Maintenance Mode and all alerts will be enabled.

Remoção do host

- Para remover definitivamente o host do cluster (depois da desabilitação)
 - Menu superior Hosts → All Hosts
 - Seleção do host
 - Remove From Cluster



Configurações do banco de dados

- As configurações do cluster estão gravadas em um banco de dados fora do cluster, como foi definido na instalação do CDH
- Os dados de acesso estão na página inicial do Cloudera Manager → Configuration → Database Settings

Database Settings

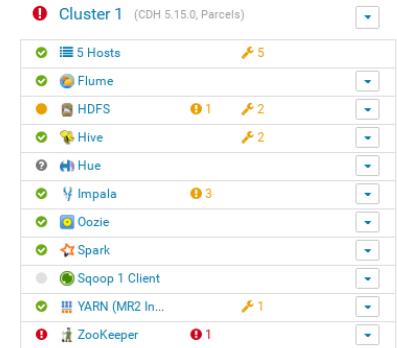
Hue Database Type engine	Cluster 1 > Hue (Service-Wide) ↩ <input checked="" type="radio"/> MySQL <input type="radio"/> SQLite3 <input type="radio"/> PostgreSQL <input type="radio"/> Oracle
Hue Database Directory dir	Cluster 1 > Hue (Service-Wide) <input type="text" value="/var/lib/hue/desktop.db"/>
Hue Database Hostname host	Cluster 1 > Hue (Service-Wide) ↩ <input type="text" value="mariadbserver"/>
Hue Database Port port	Cluster 1 > Hue (Service-Wide) <input type="text" value="3306"/>
Hue Database Username user	Cluster 1 > Hue (Service-Wide) <input type="text" value="hue"/>
Hue Database Password password	Cluster 1 > Hue (Service-Wide) ↩ <input type="password" value="*****"/>
Hue Database Name name	Cluster 1 > Hue (Service-Wide) <input type="text" value="hue"/>
Database Dump File	Cluster 1 > Hue (Service-Wide) <input type="text" value="/tmp/hue_database_dump.json"/> Cluster 1 > Oozie (Service-Wide) <input type="text" value="/tmp/oozie_database_dump.zip"/>

Monitoramento e Solução de Problemas



Monitoramento

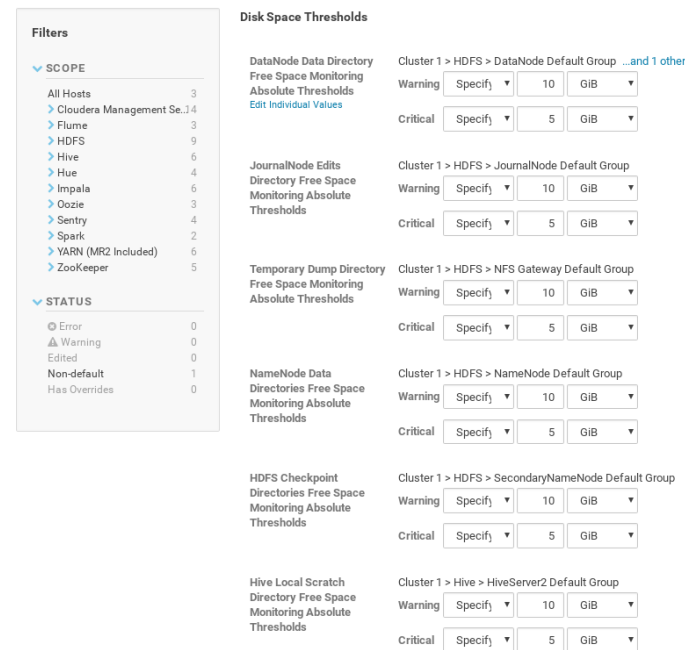
- Cada serviço mostra o estado de sua saúde na página inicial do CM
 - Good Health
 - Concerning Health
 - Bad Health



Cluster 1 (CDH 5.15.0, Parcels)		
5 Hosts	5	
Flume		
HDFS	1	2
Hive		2
Hue		
Impala	3	
Oozie		
Spark		
Sqoop 1 Client		
YARN (MR2 In...)		1
ZooKeeper	1	

Limites de espaço em disco

- A configuração do espaço deve ser monitorada frequentemente, pois a operação do cluster gera muita gravação de log em disco
- Está disponível na página inicial → Configuration → Disk Space Threshold
- Na imagem, a configuração indica que os alertas são disparados quando o disco está abaixo de 10 GB de espaço livre e o estado crítico é atingido quando há menos de 5 GB
- O swap também deve ser verificado, uma vez que não é desejável ter swap acionado durante o processamento



Component	Cluster Path	Warning Threshold	Critical Threshold
DataNode Data Directory Free Space Monitoring Absolute Thresholds	Cluster 1 > HDFS > DataNode Default Group ...and 1 other	10 GiB	5 GiB
JournalNode Edits Directory Free Space Monitoring Absolute Thresholds	Cluster 1 > HDFS > JournalNode Default Group	10 GiB	5 GiB
Temporary Dump Directory Free Space Monitoring Absolute Thresholds	Cluster 1 > HDFS > NFS Gateway Default Group	10 GiB	5 GiB
NameNode Data Directories Free Space Monitoring Absolute Thresholds	Cluster 1 > HDFS > NameNode Default Group	10 GiB	5 GiB
HDFS Checkpoint Directories Free Space Monitoring Absolute Thresholds	Cluster 1 > HDFS > SecondaryNameNode Default Group	10 GiB	5 GiB
Hive Local Scratch Directory Free Space Monitoring Absolute Thresholds	Cluster 1 > Hive > HiveServer2 Default Group	10 GiB	5 GiB

Exemplos de alerta

- Os alertas de nível crítico (vermelho) e de atenção (laranja) são disparados quando algum dos eventos configurados é atingido
 - Exemplo: quando o espaço livre no disco está abaixo de 10 GB
- O administrador tem duas opções ao detectar o problema
 - Proceder com a correção, movendo arquivos antigos ou adicionando novos discos
 - Alterar o limite de monitoramento (não é o recomendado)

Status	Configuration	Processes	Commands	Charts Library	Audit
Health Tests					
Create Trigger					
<div><div><div><div><div></div><div>Log Directory Free Space</div></div><div>This role's Log Directory is on a filesystem with less than 5.0 GiB of its space free. /var/log/hadoop-hdfs (free: 2.4 GiB (4.40%), capacity: 54.9 GiB)</div></div><div>Suppress...</div></div><div><div><div><div></div><div>Host Health</div></div><div>The health of this role's host is concerning. The following health tests are concerning: agent parcel directory.</div></div><div>Suppress...</div></div><div><div><div><div></div><div>Show 13 Good</div></div></div><div><div><div><div></div><div>Show 2 Disabled</div></div></div></div></div></div>					

Log Directory Free Space (Cluster 1, Hive, Hive Metastore Server, hulk)

Sep 12, 10:36 AM BFT

Test of whether this role's log directory has enough free space.

Concerning : This role's Log Directory is on a filesystem with less than 10.0 GiB of its space free. /var/log/hive (free: 5.6 GiB (10.22%), capacity: 54.9 GiB)

Actions

- Change **Log Directory Free Space Monitoring Absolute Thresholds** for all roles in **Hive Metastore Server Default Group** role group
- Change **Log Directory Free Space Monitoring Percentage Thresholds** for all roles in **Hive Metastore Server Default Group** role group
- Change **Log Directory Free Space Monitoring Absolute Thresholds** for this role instance
- Change **Log Directory Free Space Monitoring Percentage Thresholds** for this role instance
- View the log for this role instance at the time of the health test

Advice

This Hive Metastore Server health test checks that the filesystem containing the log directory of this Hive Metastore Server has sufficient free space.

This test can be configured using the **Log Directory Free Space Monitoring Absolute Thresholds** and **Log Directory Free Space Monitoring Percentage Thresholds** Hive Metastore Server monitoring settings.

Diretórios de dados locais

- Os serviços gravam dados no sistema de arquivos do SO e não apenas no HDFS
- Uma grande parte das configurações locais está definida em `/var/lib/` ou `/var/run`
- Esses diretórios podem ser utilizados para identificação de problemas no cluster
- A opção está disponível na página inicial → Configuration → Local Data Directories and Files

The screenshot shows the 'Local Data Directories and Files' configuration page. On the left is a sidebar with filters categorized into SCOPE, CATEGORY, and STATUS. The main area on the right is titled 'Local Data Directories and Files' and contains several configuration sections, each with a text input field and a help icon.

Filters

- SCOPE**
 - All Hosts: 2
 - Settings: 3
 - Cloudera Management Serv...: 6
 - HBase: 4
 - HDFS: 10
 - Hive: 8
 - Hue: 10
 - Impala: 13
 - Kafka: 5
 - Oozie: 3
 - Spark: 1
 - Spark 2: 1
 - YARN (MR2 Included): 7
 - ZooKeeper: 2
- CATEGORY**
 - Advanced: 3
 - Database: 0
 - Hive Metastore Database: 0
 - Kerberos: 0
 - Main: 2
 - Parcels: 0
 - Performance: 0
 - Security: 2
- STATUS**

Local Data Directories and Files

System User's Home Directory

- Cluster 1 > HDFS (Service-Wide):
- Cluster 1 > YARN (MR2 Included) (Service-Wide):

Hadoop TLS/SSL Server Keystore File Location
all.servers.keystore.location
[Edit Individual Values](#)

- Cluster 1 > HDFS (Service-Wide) ...and 1 other:

Deploy Directory
[Edit Individual Values](#)

- Cluster 1 > HBase > Gateway Default Group:
- Cluster 1 > HDFS > Gateway Default Group ...and 1 other:
- Cluster 1 > Hive > Gateway Default Group:
- Cluster 1 > Kafka > Gateway Default Group:
- Cluster 1 > Spark > Gateway Default Group:

Diretórios de logs

- Um cluster de big data demanda monitoramento constante, uma vez que é um sistema complexo com centenas ou até mesmo milhares de elementos interconectados
- Além dos alertas, que são características do Cloudera Manager, os logs são uma fonte de informação importante para a identificação de problemas na operação do cluster
- Por padrão, os logs estão disponíveis no diretório `/var/log`, onde cada serviço tem um diretório específico
 - O HDFS grava o log em `/var/log/hadoop-hdfs/`
- O menu desta funcionalidade está disponível na página inicial → Configuration → Log Directories

Filters	
Clear All	
▼ SCOPE	Clear
▶ Cloudera Management Se...	9
▶ Flume	1
▶ HDFS	7
▶ Hive	3
▶ Hue	3
▶ Impala	4
▶ Oozie	1
▶ Sentry	1
▶ Spark	1
▶ YARN (MR2 Included)	3
▶ ZooKeeper	1
▼ STATUS	
⊙ Error	0
⚠ Warning	0
Edited	0
Non-default	0
Has Overrides	0

Log Directories	
Activity Monitor Log Directory	Cloudera Management Service > Activity Monitor Default Group <code>/var/log/cloudera-scm-firehose</code>
Alert Publisher Log Directory	Cloudera Management Service > Alert Publisher Default Group <code>/var/log/cloudera-scm-alertpublisher</code>
Event Server Log Directory	Cloudera Management Service > Event Server Default Group <code>/var/log/cloudera-scm-eventserver</code>
Host Monitor Log Directory	Cloudera Management Service > Host Monitor Default Group <code>/var/log/cloudera-scm-firehose</code>
Navigator Audit Server Log Directory	Cloudera Management Service > Navigator Audit Server Default Group <code>/var/log/cloudera-scm-navigator</code>
Navigator Metadata Server Log Directory	Cloudera Management Service > Navigator Metadata Server Default Group <code>/var/log/cloudera-scm-navigator</code>

Portas

- O Cloudera Manager mostra as portas utilizadas no menu está disponível na página inicial → Configuration → Ports
- Assim como os logs, o número da porta HTTP é importante para a identificação de problemas
- Além disso, as portas são importantes para configuração de redes quando é preciso algum tipo de permissão no firewall

Filters		Clear All	
▼ SCOPE		Clear	
All Hosts	1		
Settings	5		
▶ Cloudera Management Se.22			
▶ Fume	1		
▶ HDFS	20		
▶ Hive	5		
▶ Hue	3		
▶ Impala	14		
▶ Oozie	6		
▶ Sentry	3		
▶ Spark	3		
▶ YARN (MR2 Included)	14		
▶ ZooKeeper	5		
▼ CATEGORY			
Advanced	0		
Database	0		
Hive Metastore Database	0		
Kerberos	0		
Network	0		
Ports and Addresses	20		
SNMP	0		
Sentry Server Database	0		
▼ STATUS			
🔴 Error	0		
⚠ Warning	0		
🟢 Edited	0		

Ports	
DataNode Protocol Port dfs.datanode.ipc.address Edit Individual Values	Cluster 1 > HDFS > DataNode Default Group ...and 2 others 50020
DataNode Transceiver Port dfs.datanode.address Edit Individual Values	Cluster 1 > HDFS > DataNode Default Group ...and 2 others 1004
DataNode HTTP Web UI Port dfs.datanode.http.address Edit Individual Values	Cluster 1 > HDFS > DataNode Default Group ...and 2 others 1006
Secure DataNode Web UI Port (TLS/SSL) dfs.datanode.https.address Edit Individual Values	Cluster 1 > HDFS > DataNode Default Group ...and 2 others 50475
REST Port hdfs.https.http.port	Cluster 1 > HDFS > HttpFS Default Group 14000
Administration Port hdfs.https.admin.port	Cluster 1 > HDFS > HttpFS Default Group 14001
JournalNode RPC Port dfs.journalnode.rpc.address	Cluster 1 > HDFS > JournalNode Default Group 8485

Verificação das portas
































- O Hadoop e o CM usam diversas portas para publicar os serviços e não é necessário decorar seus números
- Algumas portas acabam sendo mais conhecidas, como a 8042 (NodeManager) e a 7190 (Agent)
- O comando Linux para mostrar as portas utilizadas é:
 - `$ netstat -ant |grep LISTEN`
- Outra opção é testar se a porta está aberta em outro servidor, quando o netstat não está disponível, como na imagem abaixo:
 - `$ telnet ip porta`
- Caso a porta esteja indisponível, o computador remoto vai responder desta forma:
 - telnet: Unable to connect to remote host: Connection refused

```
dataengineer@datanode04:~$ netstat -vatn |grep LISTEN
tcp        0      0 192.168.122.16:8041 0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:7337        0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:8042 0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:1004 0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:1006 0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:22000       0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:80          0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:4433        0.0.0.0:*           LISTEN
tcp        0      0 127.0.0.1:7190      0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:22          0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:7191        0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:27000 0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:23000       0.0.0.0:*           LISTEN
tcp        0      0 127.0.0.1:19001     0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:21050       0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:13562       0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:50020 0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:21000       0.0.0.0:*           LISTEN
tcp        0      0 0.0.0.0:25000       0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:8040 0.0.0.0:*           LISTEN
tcp        0      0 127.0.0.1:43336     0.0.0.0:*           LISTEN
tcp        0      0 192.168.122.16:9000 0.0.0.0:*           LISTEN
tcp6       0      0 :::4434             :::*                LISTEN
tcp6       0      0 :::22               :::*                LISTEN
tcp6       0      0 :::7191             :::*                LISTEN
```

```
dataengineer@datanode04:~$ telnet 192.168.122.16 8042
Trying 192.168.122.16...
Connected to 192.168.122.16.
Escape character is '^['.
```

Resolução de problemas

- Um sistema distribuído como o Hadoop precisa de manutenção constante porque apresenta com frequência falhas referentes a:
 - Rede
 - Sistema Operacional
 - Configurações dos serviços (HDFS, YARN, Impala, Hive etc.)
 - Autenticação e autorização de usuários e grupos
 - Capacidade e utilização dos discos
- Um problema geralmente está relacionado a falhas em ao menos serviço de um host
- O ponto de partida para identificação de problemas é a página inicial do CM, na qual podemos verificar os alertas sobre a saúde dos serviços
 - Os alertas estão indicados pela interrogação com o número de ocorrências que aparece ao lado do serviço

 Hosts	 4	
 HBase	 2	
 HDFS	 3  1	
 Hive		
 Hue	 1	
 Impala		
 Kafka		
 Oozie	 1	
 Spark		
 Sqoop 2		
 YARN (MR2 Incl...)	 3	
 ZooKeeper	 3	

Inspeção dos hosts

- O CM tem uma facilidade para inspeção dos hosts disponível no menu superior → Hosts → All Hosts → botão Inspect All Hosts
- Os resultados da inspeção mostram as validações executadas em todos os componentes instalados nos hosts e, em caso de problemas, o inspetor vai indicar possíveis ações que podem ajudar na solução
- É possível fazer o download dos resultados ou apenas mostrar na tela
- Na imagem abaixo podemos ver um alerta de configuração

✔	No conflicts detected between packages and parcels.
✔	No kernel versions that are known to be bad are running.
⚠	Cloudera recommends setting <code>/proc/sys/vm/swappiness</code> to a maximum of 10. Current setting is 60. Use the <code>sysctl</code> command to change this setting at run time and edit <code>/etc/sysctl.conf</code> for this setting to be saved after a reboot. You can continue with installation, but Cloudera Manager might report that your hosts are unhealthy because they are swapping. The following hosts are affected: ▼ View Details datanode[01-03].lab; edgenode.lab; headnode.lab
✔	No performance concerns with Transparent Huge Pages settings.
✔	CDH 5 Hue Python version dependency is satisfied.
✔	0 hosts are running CDH 4 and 5 hosts are running CDH 5.

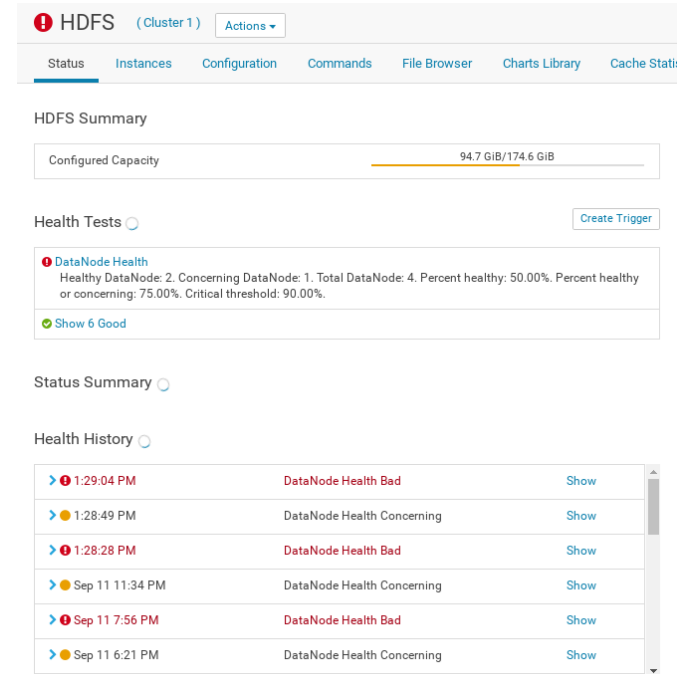
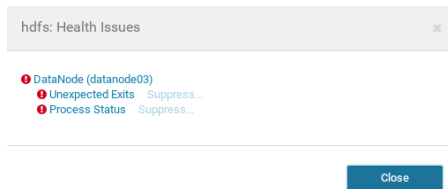
Diagnostics e Audit



- Diagnostics: o CDH coleta os eventos dos hosts do cluster para ajudar na solução de problemas
 - Events: mostra os eventos e os detalhes, com opção de filtros
 - Logs: registros dos hosts, com opção de consulta e filtros
 - Server Logs: registros do SCM (Cloudera Manager Server)
- Audit: histórico de operações para auditoria
 - Operações realizadas no cluster, como iniciar e parar serviços, alteração de configuração, alteração de usuários etc.

Exemplo de Problema Crítico

- Neste exemplo o processo do DataNode foi parado, o que gerou um alerta crítico no sistema
 - A capacidade do HDFS diminuiu e o DataNode Health está crítico
- Ao detectar o problema na página inicial do CM (imagem abaixo) selecione o problema e proceda para a identificação das causas



Hosts com serviço Down

- Verifique na página dos hosts que um deles está down (datanode03.lab), mesmo que seu estado seja Started

Filters

STATUS

None

1

Down

1

Concerning Health

1

Good Health

5

COMMISSION STATE

MAINTENANCE MODE

RACK

ROLE GROUP

ROLE TYPE

STATE

HEALTH TESTS

Actions for Selected

Migrate Roles

Add Role Instances

Role Groups

<input type="checkbox"/>	Role Type	State	Host
<input type="checkbox"/>	Balancer	N/A	edgenode.lab
<input type="checkbox"/>	DataNode	Started	datanode04.lab
<input type="checkbox"/>	DataNode	Started	datanode03.lab
<input type="checkbox"/>	DataNode	Started	datanode01.lab
<input type="checkbox"/>	DataNode	Started	datanode02.lab
<input type="checkbox"/>	HttpFS	Started	edgenode.lab
<input type="checkbox"/>	NameNode (Active)	Started	headnode.lab
<input type="checkbox"/>	SecondaryNameNode	Started	edgenode.lab

Atividades Opcionais



Atividades

- Sentry
- HDFS High Availability
- YARN High Availability

Sentry

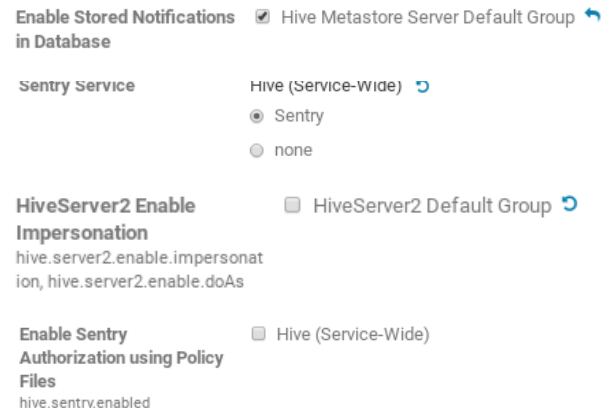
- O Sentry é um serviço para autorização usado no Hive/Impala/Solr, garantindo controle a nível de usuário nas tabelas do cluster (GRANT/REVOKE)
- Para criar o usuário e a base de dados (MariaDB):
 - > grant all on *.* to 'sentry'@'%' identified by 'sentry' with grant option;
 - > CREATE DATABASE sentry DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_general_ci;
- Adicione o serviço na página inicial → Add Service → Sentry
- O servidor Sentry será instalado no EdgeNode e os DataNodes serão os gateways



- A base de dados vai usar as configurações já conhecidas:
 - Servidor MySQL/MariaDB: mariadbserver
 - Database: sentry
 - User: sentry
 - Password: sentry

Configuração do Sentry no Hive/Impala/Hue

- Configurações do Hive
 - Selecione a opção Enable Stored Notifications in Database
 - Defina o parâmetro Sentry Service com o valor Sentry
 - Desabilite a opção hive.server2.enable.doAs
 - Desabilite a opção Enable Sentry Authorization using Policy Files
- Configurações do Impala
 - Defina o parâmetro Sentry Service com o valor Sentry
- Configuração do Hue
 - Defina o parâmetro Sentry Service com o valor Sentry
- Reinicie os serviços Impala, Hive e Hue



Acesso Hive/Impala/Hue com Kerberos



- Com o Kerberos ativado, nenhum usuário tem acesso aos dados pelos clientes do Hive/Impala/Hue
- Para habilitar os usuários temos de criar uma ROLE e associá-la ao grupo de usuários
- Para usar o cliente Hive/Impala a sintaxe é:
 - `$ beeline -u "jdbc:hive2://edgenode.lab:10000;principal=hive/_HOST@LAB.BIGDATA.COM"`
- A primeira ROLE é a dos administradores e deve ser criada com o usuário hive autenticado no kerberos:
 - `> CREATE ROLE admin;`
 - `> GRANT ALL ON SERVER server1 TO ROLE admin;`
 - `> GRANT ALL ON DATABASE datalake TO ROLE admin;`
 - `> GRANT ROLE admin TO GROUP hive;`
- Vamos adicionar uma ROLE para cada grupo com acesso a sua respectiva base de dados:
 - `> CREATE ROLE dataengineer_role;`
 - `> GRANT ALL ON DATABASE datalake TO ROLE dataengineer_role;`
 - `> GRANT ROLE dataengineer_role TO GROUP dataengineer;`
- Para criar tabelas externas no Hive temos mais um passo, que é permitir o acesso ao diretório de origem
 - `GRANT ALL ON URI 'hdfs://NameServiceHA/user/dataengineer/dados/municipio' TO ROLE dataengineer_role;`
- Agora é possível que o dataengineer acesse normalmente os dados da base, até mesmo criar tabelas
- Por fim, para acessar o Impala com Kerberos
 - `$ impala-shell -i datanode01.lab -k`

HDFS High Availability

- As falhas são consideradas inevitáveis em sistemas distribuídos, especialmente no Hadoop
- A resiliência é a capacidade de suportar falhas sem downtime, ou seja, mantendo o serviço disponível para o usuário mesmo quando um servidor ou serviço para inesperadamente
- As versões iniciais do Hadoop tinham pontos de falha críticos (Single Point of Failure - SPoF): o NameNode e o ResourceManager. Como são os serviços principais, tanto o HDFS quando o YARN ficariam indisponíveis em caso de falha
- Esta característica é chamada de alta disponibilidade (High Availability - HA) e, no HDFS, é implementada por meio da replicação do NameNode e de JournalNodes
 - As alterações no HDFS são replicadas nos NameNodes, um deles fica em estado Active e o outro em Stand By
 - Os JournalNodes são usados para sincronizar os NameNodes
 - Em caso de queda do NN Active, o NN Stand By verifica as atualizações nos JournalNodes e se torna o NN Active
- Para ativar o HDFS HA selecione o serviço HDFS → Actions → Enable High Availability

Enable High Availability for HDFS

Getting Started

This wizard leads you through adding a standby NameNode, restarting this HDFS service and any dependent services, and then re-deploying client configurations.

Nameservice Name *

Enabling High Availability creates a new nameservice. Accept the default name **nameservice1** or provide another name in **Nameservice Name**.

Roles para o NameServiceHA

- Os hosts para NameNode e JournalNode devem ter capacidades semelhantes
 - No nosso cluster estamos usando os DataNodes, o que não seria a melhor opção, mas é apenas para demonstração, e não para máxima performance
- Indique os diretórios dos JournalNodes conforme a imagem
 - A quantidade de JNs deve ser ímpar, com o mínimo de 3

Enable High Availability for HDFS

Assign Roles

NameNode Hosts	<input type="text" value="headnode.lab (Current)"/>
	<input type="text" value="edgenode.lab"/>
JournalNode Hosts	<input type="text" value="datanode[01-03].lab"/>

We recommend that JournalNodes be hosted on machines of similar hardware specifications as the NameNodes. The hosts of NameNodes and the ResourceManager are generally good options. You must have a minimum of three and an odd number of JournalNodes.

Enable High Availability for HDFS

Review Changes

Set the following configuration values for your new role(s). Required values are marked with *.

Parameter	Group	Value	Description
Service HDFS			
NameNode Data Directories*	edgenode	/dfs/n	Determines where on the local file system the NameNode should store the name table (fsimage). For redundancy, enter a comma-delimited list of directories to replicate the name table in all of the directories. Typical values are /data/N/dfs/n where N=1..3.
dfs.namenode.name.dir		Inherited from: NameNode Default Group	
	headnode	/dfs/n	
		Inherited from: NameNode Default Group	
JournalNode Edits Directory*	datanode01	/dfs/jn	Directory on the local file system where NameNode edits are written.
dfs.journalnode.edits.dir		<input type="text" value="Reset to empty default value"/>	
	datanode02	<input type="text" value="Reset to empty default value"/>	
	datanode03	<input type="text" value="Reset to empty default value"/>	

Extra Options

- ❗ Force initialize the Zookeeper ZNode for autolover. Any previous ZNode used for this nameservice will be overwritten.
- ❗ Clear any existing data present in name directories of Standby NameNode.
Make sure you have backed up any existing data in the name directories of Standby NameNode.
- ❗ Clear any existing data present in the JournalNode edits directory for this nameservice.
Make sure you have backed up any existing data in the edits directory on all hosts running JournalNodes.

Concluindo o procedimento

- A inicialização pode demorar um pouco, então, paciência!

Enable High Availability for HDFS

Enable High Availability Command

Status Finished Context [HDFS](#) Sep 12, 9:21:18 PM 9:1m

Successfully enabled High Availability and Automatic Failover

Completed 20 of 21 step(s).

Show All Steps	Show Only Failed Steps	Show Running Steps
1	2	3
Check that name directories for the new Standby NameNode either do not exist or are writable and empty. Can optionally clear directories.	edgecode lab	Sep 12, 9:21:18 PM 22.25s
Check that edit directories for the nameservice either do not exist or are writable and empty. Can optionally clear directories.		Sep 12, 9:21:40 PM 13.85s
Stop hdfs and its dependent services	Cluster 1	Sep 12, 9:21:56 PM 38.5s
Creating roles to enable High Availability		Sep 12, 9:22:27 PM 372ms
Waiting for credentials to be generated for the new roles.		Sep 12, 9:22:27 PM 16.63s
Deleting the SecondaryNameNode role. The checkpoint directories of the SecondaryNameNode will not be deleted.		Sep 12, 9:22:45 PM 1.56s
Configuring NameNodes and the HDFS service to enable High Availability		Sep 12, 9:22:46 PM 20ms

- O passo seguinte é atualizar a MetaStore do Hive para usar a HA. Neste caso, a recomendação da Cloudera é fazer um backup do banco de dados antes de continuar
 - `$ mysqldump -hmariadbserver -uroot -proot hive > /tmp/hive-backup.sql`

Enable High Availability for HDFS

Congratulations!

Successfully enabled High Availability.

The following manual steps must be performed after completing this wizard:

- For each of the Hive service(s) **Hive**, stop the Hive service, back up the Hive Metastore Database to a persistent store, run the service command "Update Hive Metastore NameNodes", then restart the Hive services.

YARN High Availability

- O YARN também era um ponto de falha no Hadoop (SPoF), entretanto, já está disponível o recurso de alta disponibilidade, ou High Availability (HA)
- Chamado de ResourceManager HA, o recurso usa a redundância do RM, ou seja, a implementação é feita com 2 RMs com a estratégia de Active/Stand by, assim como no HDFS. Desta forma, as aplicações ficam protegidas e continuam rodando mesmo em caso de falhas no RM

Enable High Availability for YARN (MR2 Included)

Getting Started

This wizard leads you through adding a standby ResourceManager, restarting this YARN (MR2 Included) service and any dependent services, and then re-deploying client configurations.





ResourceManager Hosts	<input type="text" value="headnode.lab (Current)"/>
	<input type="text" value="edgenode.lab"/>

- Verifique o estado das novas funções selecionando o serviço YARN → Instances

<input type="checkbox"/>	<input checked="" type="checkbox"/>	ResourceManager (Standby)	Started	edgenode.lab	Commissioned	ResourceManager Default Group
<input type="checkbox"/>	<input checked="" type="checkbox"/>	ResourceManager (Active)	Started	headnode.lab	Commissioned	ResourceManager Default Group

Teste da HA (HDFS e YARN)

- Com a HA habilitada o cluster suporta a perda de um dos NameNodes sem downtime
- Para confirmar, vamos matar o processo do NameNode ativo (HeadNode) e verificar o que acontece
 - `$ ps aux | grep namenode`
 - `$ kill xxx`

<input type="checkbox"/>	 JournalNode	Started	datanode01.lab	Commissioned	JournalNode Default Group
<input type="checkbox"/>	 JournalNode	Started	datanode03.lab	Commissioned	JournalNode Default Group
<input type="checkbox"/>	 JournalNode	Started	datanode02.lab	Commissioned	JournalNode Default Group
<input type="checkbox"/>	 NameNode (Standby)	Started	edgenode.lab	Commissioned	NameNode Default Group
<input type="checkbox"/>	 NameNode (Active)	Started	headnode.lab	Commissioned	NameNode Default Group

Dúvidas?

Marco Reis
<http://marcoreis.net>