

弹性 MapReduce

最佳实践

产品文档





【版权声明】

©2013-2024 腾讯云版权所有

本文档著作权归腾讯云单独所有,未经腾讯云事先书面许可,任何主体不得以任何形式复制、修改、抄袭、传播全部或部分本文档内容。

【商标声明】



及其它腾讯云服务相关的商标均为腾讯云计算(北京)有限责任公司及其关联公司所有。本文档涉及的第三方主体的商标,依法由权利人所有。

【服务声明】

本文档意在向客户介绍腾讯云全部或部分产品、服务的当时的整体概况,部分产品、服务的内容可能有所调整。您 所购买的腾讯云产品、服务的种类、服务标准等应由您与腾讯云之间的商业合同约定,除非双方另有约定,否则, 腾讯云对本文档内容不做任何明示或默示的承诺或保证。



文档目录

最佳实践

数据迁移

开始使用

通过 COS 迁移数据

如何查看 COS 信息

DistCp 迁移数据

基本说明

网络打通

执行拷贝

注意事项

Hive 迁移指导手册

Kudu 数据迁移指导手册

数据迁移实践

HDFS 通过对象存储数据迁移实践

HDFS 通过 DistCp 数据迁移实践



最佳实践 数据迁移 开始使用

最近更新时间:2022-04-18 12:18:28

腾讯云 弹性MapReduce 数据迁移目前支持两种数据迁移,第一种是通过 COS (腾讯云对象存储)中转,第二种是 自建 HDFS 通过 DistCp 方式迁移,此模式下需要实现自建集群和 EMR 集群的网络互通。



通过 COS 迁移数据

最近更新时间:2023-02-23 14:17:19

原始数据非 HDFS 数据

如果您的原始数据不是 HDFS 数据而是其他形式的文件数据,可以通过 COS 的 web 控制台或者 COS 提供的 API 来把数据传入到 COS,然后在 EMR 集群中进行分析, COS 传输数据请查看资料。

原始数据在 HDFS 的数据迁移

1. 获取 COS 迁移工具

获取迁移工具,更多迁移工具请参考工具概览。

2. 工具配置

配置文件统一放在工具目录里的 conf 目录,将需要同步的 HDFS 集群的 core-site.xml 拷贝到 conf 中,其中包含

了 NameNode 的配置信息,编辑配置文件 cos_info.conf,包括 appid、bucket、region 以及密钥信息。

3. 命令参数说明

注意:

- 建议用户使用子账号密钥,遵循最小权限原则说明,防止泄漏目标存储桶或对象之外的资源。
- 如果您一定要使用永久密钥, 建议遵循 最小权限原则说明 对永久密钥的权限范围进行限制。

```
-ak <ak> the cos secret id //用户的 SecretId, 建议使用子账号密钥, 授权遵循最小权限指引,
降低使用风险。子账号密钥获取可参考:https://www.tencentcloud.com/document/product/598/3
2675
-appid, --appid <appid> the cos appid
-bucket, --bucket <bucket_name> the cos bucket name
-cos_info_file, --cos_info_file <arg> the cos user info config default is ./conf/c
os_info.conf
-cos_path, --cos_path <cos_path> the absolute cos folder path
-h, --help print help message
-hdfs_conf_file, --hdfs_conf_file <arg> the hdfs info config default is ./conf/cor
e-site.xml
-hdfs_path, --hdfs_path <hdfs_path> the hdfs path
-region, --region <region> the cos region. legal value cn-south, cn-east, cn-nort
h, sq
-sk <sk> the cos secret key //用户的 SecretKey, 建议使用子账号密钥, 授权遵循最小权限指
引,降低使用风险。子账号密钥获取可参考:https://www.tencentcloud.com/document/product/5
98/32675
-skip_if_len_match, --skip_if_len_match skip upload if hadoop file length match co
```

S



4. 执行迁移

所有操作都要在工具目录下。如果同时设置了配置文件和命令行参数,以命令行参数为准 ./hdfs_to_cos_cmd -h # 从 HDFS 拷贝到 COS (如果 COS 上已存在文件,则会覆盖) ./hdfs_to_cos_cmd --hdfs_path=/tmp/hive --cos_path=/hdfs/20170224/ # 从 HDFS 拷贝到 COS,同时要拷贝的文件和 COS 的长度一致,则忽略上传(适用于拷贝一次后,重新 拷贝) # 这里只做长度的判断,因为如果将 Hadoop 上的文件摘要算出,开销较大 ./hdfs_to_cos_cmd --hdfs_path=/tmp/hive --cos_path=/hdfs/20170224/ -skip_if_len _match # 完全通过命令行设置参数 ./hdfs_to_cos_cmd -appid 1252xxxxx -ak AKIDVt55xxxxxxxxxxxxxxxxxxxxxxx -sk KS08jDVbVElxxxxxxxxxxxxxxxxxxxxx -bucket test -cos_path /hdfs -hdfs_path /data/data -region cn-south -hdfs_conf_file /home/hadoop/hadoop-2.8.1/etc/hadoop/core-site.xml

5. 验证运行命令后,输出如下日志

```
[Folder Operation Result : [ 53(sum) / 53(ok) / 0(fail)]]
[File Operation Result: [22(sum) / 22(ok) / 0(fail) / 0(skip)]]
[Used Time: 3 s]
```

- sum 表示总共需要迁移的文件数。
- ok 表示成功迁移的文件数。
- fail 表示迁移失败的文件数。
- skip 表示在添加 skip_if_len_match 参数后,由于上传文件和同名文件具有相同长度的文件,则跳过的数量。

您也可以登录 COS 控制台查看数据是否已经正确迁移过来。

常见问题

- 请确保填写的配置信息,包括 appID、密钥信息、bucket 和 region 信息正确,以及机器的时间和北京时间一致 (如相差1分钟左右是正常的),如果相差较大,请设置机器时间。
- 请保证对于 DateNode, 拷贝程序所在的机器也可以连接。因 NameNode 有外网 IP 可以连接, 但获取的 block 所 在的 DateNode 机器是内网 IP, 无法连接上,因此建议同步程序放在 Hadoop 的某个节点上执行,保证对 NameNode 和 DateNode 皆可访问。
- 权限问题,用当前账户使用 Hadoop 命令下载文件,看是否正常,再使用同步工具同步 Hadoop 上的数据。



- 对于 COS 上已存在的文件,默认进行重传覆盖,除非用户明确的指定 -skip_if_len_match,当文件长度一致时则 跳过上传。
- cos path 都认为是目录,最终从 HDFS 上拷贝的文件都会存放在该目录下。



如何查看 COS 信息

最近更新时间:2022-05-16 12:52:26

登录腾讯云 对象存储 控制台,选择存储桶列表。

• 若没有存储桶可自行创建,选择创建存储桶。

cos	Bucket List			Console Reference 🗾 SDK	Documentation 😰 API Documentation 😰
Overview	If root account grants permissions to a sub-accoun	t for buckets opera	ations, but not grant for the bucket list operations, sub-ac	count can go to Access Path List to ad	d authorized bucket access paths.
P Access Path List	Create Bucket			Bucket Name	Please enter bucket na Q
	Bucket Name	Monitoring	Region	Time Created	Actions
		di	(u)	2019-07-17 19:47:48	Delete

• 若已有对应存储桶,可选择对应桶,并查看对应存储桶的访问域名。

8	← Back to Bucke	et					Documentation Guide 🗷			
	Overview		The usage overview data is not used as billing metering d	lata, and the non-real-time data (with a de	a delay of about 2 hours) is for reference only. For billing metering data, please go to Billing CenterDownload and view usage details.					
C.	File List Basic Configurations	v	Usage Overview STANDARD *							
	Security Management	~	Number of Objects v Storage		Month-to-Date	e Total Traffic 💌	Total Requests for This Month 👻			
Ð	Permission Management	~	O individual VS Yesterday: ↑ 0%	O _B VS Yesterday: ↑ 0%	O B last monthTota	al Traffic 0 B	O time(s) last monthTotal requests 0 per			
⊕	Domains and Transfer	~	VS Last Month: 1 U%	VS Last Month: † 0%						
®	Fault Tolerance and Disaster Recovery	×	Information		Domain Information					
® 0	Logging	v	Bucket Name		Endpoint Use the access domain name for Intra					
۲	Data Processing		Region Guangzhou (China) (ap-guang Creation Time 2022-04-24 16:11:35	znou)	Default CDN Acceleration Domain					
	Workflow NEW		Access Permission Private Read/Write		Custom Endpoint	 0 per				
	Function Service	~	Alarm Configuration	Configure Alarm Policy	Global Acceleration Endpoint Static Website Endpoint	Disabled				
Ľ			Current Alarms	0	Note: COS domains use smart DNS. If address. Cross-region access cannot Creation OverviewMore.	vour other Tencent Cloud services access CC use a private network and thus will be resolved	OS, intra-region access will be resolved to a private d to a public address. For details, please see Request			



• 查询密钥, 单击左侧菜单栏密钥管理。

Clock of Virtual Machine Pelkey Centre Ref of Person AP Decomment of Virtual Machine 12 Nation Image: Person of Virtual Machine I				
I Instances I Instances I Resonant Group I Statistic I Sta	Cloud Virtual Machine	API key	Console Reference 🗾	API Documentation 🗾
Pleasent Group Image: Statistic Image: Statistic <	E Instances			
State <	Placement Group			
Surger Che ha Processing Workflow Che ha Processing W	🔿 Statistic 🗸 🗸	Please turn to [Access Key]		
Data Processing Workflow Deta Processing Workflow Ecological Service Application Integration Integration S DK Download	Storage+	Get the Appld, SecretId, SecretKey and other information you need to process Tencent Cloud object storage COS resources. You can also find the "Access Key" in the "Management and Audit" category of the "Cloud Products" in the upper menu and go to the management console to view		
 ≥ Batch Operation Evological Service No Tools > SDK Download Imagement	● E Data Processing Workflow			
Ecological Saruto Application integration Colos	Batch Operation			
 Application Integration Tools → SDK Download Application → Applicati	Ecological Service			
 ♀ Tools ► > SDK Download ● key management 	Application Integration			
SK Download Sk pmanagement				
Se key management	③ SDK Download			
	key management			
	Ξ			

• 单击云 API 密钥。

Cloud Access Management	Manage API Key											
Dashboard Users • • • • • • • • • • • • • • • • • • •	 Safety Warning Your API key represents your account identify and permissions. You can operate all the Tencent Cloud resources under your account with TencentCloud API. For your property and service security, please keep the key property and change it regularly. Please do not upload or share your key information by any means (such as GitHub). For details, see Security Setting Policy IZ 											
	Create Key APPID 1307265578	Key Secretik SecretKey: *****Show	1	Creation Time 2022-04-24 16:17:16	Last Access Time	Status On	Operation Diaable					
3												



DistCp 迁移数据 基本说明

最近更新时间:2021-06-30 15:01:50

DistCp(distributed copy)是用于大型内部/集群内复制的工具。它使用 MapReduce 来实现其分发、错误处理和恢 复、报告的功能。它将文件和目录的列表扩展为映射任务的输入,每个任务将复制源列表中指定的文件的分区。 DistCp 是 Hadoop 自带的文件迁移工具。



网络打通

最近更新时间:2021-10-29 10:07:28

本地自建 HDFS 文件迁移到 EMR

本地自建 HDFS 文件迁移到 EMR 集群需要有专线打通网络,可以联系开发人员协助解决。

CVM 上的自建 HDFS 文件迁移到 EMR

- CVM 的所属网络和 EMR 集群的所属网络在同一 VPC 下,则可以自由传送文件。
- CVM 的所属网络和 EMR 集群的所属网络在不同 VPC 下,需要使用对等连接将网络打通。

使用对等连接

网段1:广州的 VPC1 中的子网 A 192.168.1.0/24。 网段2:北京的 VPC2 中的子网 B 10.0.1.0/24。

1. 登录 私有网络控制台,选择对等连接页,在列表上方选择地域广州,选择私有网络 VPC1,然后单击+新建。

Virtual Private Cloud «	Peering Connec	tions	South China	a (Guangzhou) 🔻	All VPCs 🔻						Help	of Peering Connection 🗈
Network Topology Virtual Private Cloud	To get notified a CCNProvides sir	about ab Igle-poir	onormal peer o nt access and	connection behaviors full-mesh interconne	instantly, please Co ction between Tence	nfigure Alarms。 ent Cloud VPCs, and I	between VPC and ci	ustomer data center ,	Apply Nov	N>>		
Subnets Route Tables	+ New										Enter	peering connecti Q 🌣
ENI	ID/Name	м	Status	Local Region	Local VPC	Peer Region	Peer account	Peer VPC	Ban	Serv	Billing Mode	Operation
EIP NAT Gateway	pcx-finhtmkno	di	Connect	South China (vpc-e0krgxj7 farley_test_vpc(South China (My Account	unc-gasw6eh9	Unlimi ted	Gold	Free	Delete
Peering Connections												
VPN Connection *												
Direct Connect Gateway												
Cloud Connect Network												
Security -												

2. 进入建立对等连接页。

- 。名称:对等连接的名称,例如 PeerConn。
- 本端地域:填写本地端地域,例如广州。
- •本端网络:填写本端网络,例如 VPC1。
- 对端账户类型:填写对端网络所属账户,如果广州和北京两个网络在同一账户下,选择**我的账户**,如果不在同一账户,则要选择**其它账户**。

说明:

腾讯云

如果本端网络和对端网络都在同一地域,例如广州,通信是免费的,也不需要选择**带宽上线**;如果不在 同一地域,就要进行收费,同时带宽上限可选。

- 对端地域:填写对端地域,例如北京。
- 对端网络:填写对端网络,例如 VPC2。

Create a peering conn	ection		×
Name			
Local Region	South China (Guangzhou)	Ŧ	
Local network	vpc-s0xvxze5 (EMR-Glenn-Test 10.0.0	Ŧ	
Peer account type	O My Account Other accounts		
Peer Region	South China (Guangzhou)	٣	
Peer network	Please select	٣	
Bandwidth Cap	No restriction		
Billing method	Free		
	Create Cancel		

- 3. 同账户内私有网络进行连接,新建后对等连接立即生效;与其它账户私有网络创建对等连接,需要对端接受此对 等连接后生效。参见 同账号创建对等连接通信 和 跨账号创建对等连接通信。
- 4. 为对等连接配置本端和对端路由表。
 - 登录 腾讯云控制台,选择云产品>网络>私有网络,进入私有网络控制台。单击左侧目录中的子网,进入管理 页面。单击对等连接本端指定子网(例如广州的子网 VPC1)的关联路由表 ID,进入路由表详情页。



Su	bnet South Chi	ina (Guangzhou)	▼ All VPCs ▼							VPC and S	Subnets Documentation E
	+ New								Q ¢		
	ID/Name \$	Network	IPv4 CIDR	IPv6 CIDR	Avail 🛈	Associated ro	Subnet broad	CVM	Available IP	Default Subnet	Operation
	A	VPC1		-	Guangzhou Zo			0 👩	253	No	Delete Change route tab le

。 单击+新增路由策略。

Routing Rules + New routing	policies				
Destination	Next hop type	Next hop	Notes	Enable routing	Operation
Local	Local	Local	Released by the system by default,		(j)

• 目的端中填入对端 CIDR(例如北京的 VPC2 的 CIDR 是10.0.1.0/24),下一跳类型选择**对等连接**,下一跳选择已建立的对等连接(PeerConn)。

Add a route					×
Destination	Next hop type	Next hop	Notes	Operation	
10.0.1.0/24	Peering Connections v	pcx-		0	- -
+ Add a line					
Routing policies cor	ntrols the traffic flow in the subnet. For details	, please see <u>Configuring Routing Policie</u>	<u>25</u> .		
		Create Close			

- 以上步骤是配置广州 VPC1 到北京 VPC2 的路由表,还需要配置北京 VPC2 到广州 VPC1 的配置,配置过程同上。
- 路由表配置完成后,不同私有网络的网段之间即可进行通信。



执行拷贝

最近更新时间:2020-10-10 17:34:55

集群间的拷贝, 将一个文件夹拷贝到另一个集群 hadoop distcp hdfs://nn1:9820/foo/bar hdfs://nn2:9820/bar/foo

指定文件拷贝 hadoop distcp hdfs://nn1:9820/foo/a hdfs://nn1:9820/foo/b hdfs://nn2:9820/bar/foo

如果指定的文件太多, 可使用 -f 参数。



注意事项

最近更新时间:2021-06-30 11:11:34

- 1. 对于上述命令,必须要求源和目的版本相同。
- 2. 如果另一个客户端仍然在写入源文件,则该拷贝可能会失败;如果一个文件正在被拷贝到目的端,试图重写该文件的操作会失败;如果源文件在被复制之前被移动,那么拷贝将失败,报错信息为 FileNotFoundException。



Hive 迁移指导手册

最近更新时间:2023-03-15 10:12:39

Hive 迁移涉及两部分,数据迁移和元数据迁移。Hive 表数据主要存储在 HDFS 上,故数据的迁移主要在 HDFS 层。 Hive 的元数据主要存储在关系型数据库,可平滑迁移到云上 TencentDB,并可保障高可用。

Hive 元数据迁移

1. Dump 源 Hive 元数据库。

```
mysqldump -hX.X.X.X -uroot -pXXXX --single-transaction --set-gtid-purged=OFF hi
vemetastore > hivemetastore-src.sql
# 如果 mysql 数据没有开启 GTID, 请删除命令行中的 --set-gtid-purged=OFF
# X.X.X.X为数据库服务器地址
# XXXX为数据库密码
# 如果数据库用户不是 root, 请用正确的用户名
# hivemetastore 是 Hive 元数据库名
```

2. 确认目标集群 Hive 表数据在 HDFS 中的默认存储路径。

Hive 表数据在 HDFS 中的默认存储路径由 hive-site.xml 中的 hive.metastore.warehouse.dir 配 置项指定。如果目标集群 Hive 表在 HDFS 的存储路径需要与源集群 Hive 表路径一致,可以参考以下示例对配置 文件进行修改。例如,源集群 hive-site.xml 中 hive.metastore.warehouse.dir 为下面的值。

```
<property>
<name>hive.metastore.warehouse.dir</name>
<value>/apps/hive/warehouse</value>
</property>
```

目标集群 hive-site.xml 中 hive.metastore.warehouse.dir 为下面的值。

```
<property>
<name>hive.metastore.warehouse.dir</name>
<value>/usr/hive/warehouse</value>
</property>
```

如果目标集群 Hive 表在 HDFS 的存储位置依然保持与源集群 Hive 一致,那么修改目标 hive-site.xml 中的 hive.metastore.warehouse.dir ,即为:



<property>

```
<name>hive.metastore.warehouse.dir</name>
<value>/apps/hive/warehouse</value>
</property>
```

3. 确认目标 Hive 元数据 SDS.LOCATION 和 DBS.DB_LOCATION_URI 字段。 通过下面的查询获取当前 SDS.LOCATION 和 DBS.DB_LOCATION_URI 字段。

```
SELECT DB_LOCATION_URI from DBS;
SELECT LOCATION from SDS;
```

查询出的结果类似如下:

```
mysql> SELECT LOCATION from SDS;
+------+
| LOCATION |
+-----+
| hdfs://HDFS2648/usr/hive/warehouse/hitest.db/t1 |
| hdfs://HDFS2648/usr/hive/warehouse/wyp |
+-----+
mysql> SELECT DB_LOCATION_URI from DBS;
+-----+
| DB_LOCATION_URI |
+-----+
| hdfs://HDFS2648/usr/hive/warehouse |
| hdfs://HDFS2648/usr/hive/warehouse |
+-----+
```

其中 hdfs://HDFS2648 是 HDFS 默认文件系统名,由 core-site.xml 中的 fs.defaultFS 指定。

```
<property>
<name>fs.defaultFS</name>
<value>hdfs://HDFS2648</value>
</property>
```

/usr/hive/warehouse 为 Hive 表在 HDFS 中的默认存储路径,也是 hive-site.xml 中
 hive.metastore.warehouse.dir 指定的值。所以我们需要修改源 hive 元数据 sql 文件中的
 SDS.LOCATION 和 DBS.DB_LOCATION_URI 两个字段。确保被导入的 Hive 元数据库中的这两个字段使用的是正确的路径。可使用如下 sed 命令批量修改 sql 文件。

```
替换ip:sed -i 's/oldcluster-ip:4007/newcluster-ip:4007/g' hivemetastore-src.sql
替换defaultFS:sed -i 's/old-defaultFS/new-defaultFS/g' hivemetastore-src.sql
```



说明:

如果使用了 Kudu、Hbase 等部分组件,用 Metastore 作为元数据服务,也需更改目标 Hive 元数据中对应 location 字段。

4. 停止目标 Hive 服务 MetaStore、HiveServer2、WebHcataLog。

5. 备份目标 Hive 元数据库。

mysqldump -hX.X.X.X -uroot -pXXXX --single-transaction --set-gtid-purged=OFF hi
vemetastore > hivemetastore-target.sql
如果 mysql 数据没有开启 GTID, 请删除命令行中的 --set-gtid-purged=OFF
x.X.X.X为数据库服务器地址
xXXX为数据库密码
如果数据库用户不是 root, 请用正确的用户名
hivemetastor 是 Hive 元数据库名

6. Drop/Create 目标 Hive 元数据。

mysql> drop database hivemetastore; mysql> create database hivemetastore;

7. 导入源 Hive 元数据库到目标数据库。

mysql -hX.X.X.X -uroot -pXXXX hivemetastore < hivemetastore-src.sql # x.x.x.x为数据库服务器地址

xxxx为数据库密码

- # 如果数据库用户不是 root, 请用正确的用户名
- # hivemetastor 是 Hive 元数据库名

8. Hive 元数据升级。

如果目标和源 Hive 版本一致,则可直接跳过该步骤;否则,分别在源集群和目标集群查询 Hive 版本。

hive --service **version**

 hive 的升级脚本存放在
 /usr/local/service/hive/scripts/metastore/upgrade/mysql/
 目录下。

 hive 不支持跨版本升级,例如 hive 从1.2升级到2.3.0需要依次执行:



```
upgrade-1.2.0-to-2.0.0.mysql.sql -> upgrade-2.0.0-to-2.1.0.mysql.sql -> upgrade-2
.1.0-to-2.2.0.mysql.sql -> upgrade-2.2.0-to-2.3.0.mysql.sql
```

升级脚本主要操作为建表、加字段、改内容。如果表或字段已经存在,则升级过程中字段已存在的异常可以忽略。 例如 hive 从2.3.3升级至3.1.1。

```
mysql> source upgrade-2.3.0-to-3.0.0.mysql.sql;
mysql> source upgrade-3.0.0-to-3.1.0.mysql.sql;
```

9. 如果源 Hive 中有 phoneix 表, 修改目标 Hive 元数据中 phoneix 表的 zookeeper 地址。 通过下面的查询获取 phoenix 表的 phoenix.zookeeper.quorum 配置。

```
mysql> SELECT PARAM_VALUE from TABLE_PARAMS where PARAM_KEY = 'phoenix.zookeepe
r.quorum';
+-----+
| PARAM_VALUE |
+-----+
| 172.17.64.57,172.17.64.78,172.17.64.54 |
+-----+
```

查看目标集群的 zookeeper 地址,即 hive-site.xml 配置文件中 hbase.zookeeper.quorum 指定的值。

```
<name>hbase.zookeeper.quorum</name>
<value>172.17.64.98:2181,172.17.64.112:2181,172.17.64.223:2181</value>
</property>
```

将目标 Hive 元数据中的 phoenix 表的 zookeeper 地址改为目标集群的 zookeeper 地址。

```
mysql> UPDATE TABLE_PARAMS set PARAM_VALUE = '172.17.64.98,172.17.64.112,172.17.6
4.223' where PARAM_KEY = 'phoenix.zookeeper.quorum';
```

0. 检查目标 Hive 元数据中表名的大小写格式,参考以下示例将所有小写表名改为大写:

alter table metastore_db_properties rename to METASTORE_DB_PROPERTIES;

1. 启动目标 Hive 服务 MetaStore、HiveServer2、WebHcataLog。

2. 最后可通过简单的 Hive sql 查询进行验证。



Kudu 数据迁移指导手册

最近更新时间:2022-05-16 12:52:26

Kudu 可以通过 rebalance tool 做数据迁移。

注意:

一次只能下线一台 tserver,如果下线多台,需重复执行下述步骤。

Kudu 基于 rebalance tool 迁移

1. 确保集群状态ok。

```
/usr/local/service/kudu/bin/kudu cluster ksck 10.0.1.29:7051,10.0.1.16:7051,10.
0.1.36:7051
```

Warnings:

```
Some masters have unsafe, experimental, or hidden flags set
Some tablet servers have unsafe, experimental, or hidden flags set
```

OK [hadoop@10 bin]\$ []

2. 使用步骤1的 ksck 命令, 获取下线的节点 uid。

Tablet Server Summary					
UUID	Address	Status	Location	Tablet Leaders	Active Scanners
	t	t	++	+·	+
20681b1d6b9942cbab95dded905406ec	10.0.1.37:7050	HEALTHY	<none></none>	9	0
6929daf14f8647c89fb8cc51db5d70b6	10.0.1.15:7050	I HEALTHY	<none></none>	5	i o
b53b28bfad2c41d38d6f08a261ceb486	10.0.1.40:7050	I HEALTHY	<none></none>	2	i o
be018287364d4443a48ad1bba248c87f	10.0.1.9:7050	I HEALTHY	<none></none>	i 0	i o
fb9afb1b2989456cac5800bf6990dfea	10.0.1.45:7050	HEALTHY	<none></none>	Ő	i 0
				-	i

以 fb9afb1b2989456cac5800bf6990dfea 节点为例子。

3. 将 fb9afb1b2989456cac5800bf6990dfea 节点进入维护模式。



/usr/local/service/kudu/bin/kudu tserver state enter_maintenance 10.0.1.29:7051
,10.0.1.16:7051,10.0.1.36:7051 fb9afb1b2989456cac5800bf6990dfea

4. 执行rebalance命令。

```
/usr/local/service/kudu/bin/kudu cluster rebalance 10.0.1.29:7051,10.0.1.16:705
1,10.0.1.36:7051 --ignored_tservers fb9afb1b2989456cac5800bf6990dfea --move_rep
licas_from_ignored_tservers
```

等待命令执行结束,再次用 ksck 检查,状态为 ok,继续后面步骤。

5. 暂停 fb9afb1b2989456cac5800bf6990dfea 对应节点10.0.1.45的 tserver 进程。注意此时,使用 ksck 命令,集群 状态不健康,需要重启 tmaster。

Tablet Server Summary UUID	Address	Status	Location	Tablet Leaders	Active Scanners
20681b1d6b9942cbab95dded905406ec 6929daf14f8647c89fb8cc51db5d70b6 b53b28bfad2c41d38d6f08a261ceb486 be018287364d4443a48ad1bba248c87f fb9afb1b2989456cac5800bf6990dfea	10.0.1.37:7050 10.0.1.15:7050 10.0.1.40:7050 10.0.1.9:7050 10.0.1.45:7050	HEALTHY HEALTHY HEALTHY HEALTHY HEALTHY UNAVAILABLE	<none> <none> <none> <none></none></none></none></none>	9 5 2 0 n/a	0 0 0 0 n/a

6. 在 EMR 控制台重启 master。注意需要手动一台一台地重启(不建议使用控制台的滚动重启)。重启结束后,使 用 ksck 命令,确保集群状态健康。

Cluster	Service /	KUDU -						Help Documentation 🗷 😐
Service	e Status	Role Management	Configuration Ma	nagement Co	Configuration Record			
Restar	rt Service	Enter Maintenance	Exit Maintenance	Start Pau	se	Enter a node IP to search	Q	Nodes pending restart $~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~$
F	Role T	Health Status	Operation Status	Configura T	Node Type 🔻	Maintenance	Node IP	Last Restart \$
~ •	KuduMaster	🕑 Good	Started	kudu-defaultGroup	Master	RtatHa mode	10.0.0.144	
- H	KuduMaster	Good Good	Started	kudu-defaultGroup	Master	Normal mode	10.0.0.143	
- H	KuduMaster	Good Good	Started	kudu-defaultGroup	Common	Normal mode	10.0.0.5	
- H	KuduServer	Good Good	Started	kudu-defaultGroup	Core	Normal mode	10.0.0.60	
- H	KuduServer	🕑 Good	Started	kudu-defaultGroup	Core	Normal mode	10.0.0.138	
- H	KuduServer	Good Good	Started	kudu-defaultGroup	Core	Normal mode	10.0.0.59	
Total 6 it	tems					Lines per page 20 🔻		1 /1 page 🕨 🕨



数据迁移实践 HDFS 通过对象存储数据迁移实践

最近更新时间:2023-07-14 10:57:58

如果您需要将自有 HDFS 的原始数据迁移至腾讯云 EMR,可以通过两种方式进行数据迁移,第一种是通过腾讯云对象存储(COS)进行数据中转迁移,第二种是通过 Hadoop 自带文件迁移工具 DistCp 进行数据迁移。本文主要介绍通过腾讯云对象存储(COS)进行数据中转迁移。

原始数据非 HDFS 数据

如果您的原始数据不是 HDFS 数据而是其他形式的文件数据,可以通过 COS 的 web 控制台或者 COS 提供的 API 来把数据传入到 COS,然后在 EMR 集群中进行分析, COS 传输数据请查看资料。

原始数据在 HDFS 的数据迁移

1. 获取 COS 迁移工具。

获取迁移工具,更多迁移工具请参考工具概览。

2. 工具配置。

配置文件统一放在工具目录里的 conf 目录,将需要同步的 HDFS 集群的 core-site.xml 拷贝到 conf 中,其中包含 了 NameNode 的配置信息,编辑配置文件 cos_info.conf,包括 appid、bucket、region 以及密钥信息。

注意:

- 建议用户使用子账号密钥,遵循最小权限原则说明,防止泄漏目标存储桶或对象之外的资源。
- 如果您一定要使用永久密钥,建议遵循最小权限原则说明对永久密钥的权限范围进行限制。

命令参数说明:

```
-ak <ak> the cos secret id //用户的 SecretId, 建议使用子账号密钥, 授权遵循最小权限指
引, 降低使用风险。子账号密钥获取可参考:https://cloud.tencent.com/document/product/59
8/37140
-appid,--appid <appid> the cos appid
-bucket,--bucket <bucket_name> the cos bucket name
-cos_info_file,--cos_info_file <arg> the cos user info config default is ./conf
/cos_info.conf
-cos_path,--cos_path <cos_path> the absolute cos folder path
-h,--help print help message
-hdfs_conf_file,--hdfs_conf_file <arg> the hdfs info config default is ./conf/c
```



```
ore-site.xml
-hdfs_path,--hdfs_path <hdfs_path> the hdfs path
-region,--region <region> the cos region. legal value cn-south, cn-east, cn-nor
th, sg
-sk <sk> the cos secret key //用户的 SecretKey, 建议使用子账号密钥, 授权遵循最小权限指
引, 降低使用风险。子账号密钥获取可参考:https://cloud.tencent.com/document/product/59
8/37140
-skip_if_len_match,--skip_if_len_match skip upload if hadoop file length match
cos
```

3. 执行迁移:

```
# 所有操作都要在工具目录下。如果同时设置了配置文件和命令行参数,以命令行参数为准
./hdfs_to_cos_cmd -h
# 从 HDFS 拷贝到 cos (如果 cos 上已存在文件,则会覆盖)
./hdfs_to_cos_cmd --hdfs_path=/tmp/hive --cos_path=/hdfs/20170224/
# 从 HDFS 拷贝到 cos,同时要拷贝的文件和 cos 的长度一致,则忽略上传(适用于拷贝一次后,重新
拷贝)
# 这里只做长度的判断,因为如果将 Hadoop 上的文件摘要算出,开销较大
./hdfs_to_cos_cmd --hdfs_path=/tmp/hive --cos_path=/hdfs/20170224/ -skip_if_len
_match
# 完全通过命令行设置参数
./hdfs_to_cos_cmd -appid 1252xxxxx -ak
AKIDVt55xxxxxxxxxxxxxxxxxxxxxxx -sk
KS08jDVbVElxxxxxxxxxxxxxxxxxxxxxxx -bucket test -cos_path /hdfs
-hdfs_path /data/data -region cn-south -hdfs_conf_file
/home/hadoop/hadoop-2.8.1/etc/hadoop/core-site.xml
```

4. 验证运行命令后,输出如下日志::

```
[Folder Operation Result : [ 53(sum) / 53(ok) / 0(fail)]]
[File Operation Result: [22(sum) / 22(ok) / 0(fail) / 0(skip)]]
[Used Time: 3 s]
```

- sum 表示总共需要迁移的文件数。
- ok 表示成功迁移的文件数。
- fail 表示迁移失败的文件数。
- skip 表示在添加 skip_if_len_match 参数后,由于上传文件和同名文件具有相同长度的文件,则跳过的数量。

您也可以登录 COS 控制台查看数据是否已经正确迁移过来。对象存储使用指引请参见快速入门。

常见问题



- 请确保填写的配置信息,包括 appID、密钥信息、bucket 和 region 信息正确,以及机器的时间和北京时间一致 (如相差1分钟左右是正常的),如果相差较大,请设置机器时间。
- 请保证对于 DateNode,拷贝程序所在的机器也可以连接。因 NameNode 有外网 IP 可以连接,但获取的 block 所 在的 DateNode 机器是内网 IP,无法连接上,因此建议同步程序放在 Hadoop 的某个节点上执行,保证对 NameNode 和 DateNode 皆可访问。
- 权限问题,用当前账户使用 Hadoop 命令下载文件,看是否正常,再使用同步工具同步 Hadoop 上的数据。
- 对于 COS 上已存在的文件,默认进行重传覆盖,除非用户明确的指定 -skip_if_len_match,当文件长度一致时则 跳过上传。
- cos path 都认为是目录,最终从 HDFS 上拷贝的文件都会存放在该目录下。



HDFS 通过 DistCp 数据迁移实践

最近更新时间:2023-07-14 10:57:58

如果您需要将自有 HDFS 的原始数据迁移至腾讯云 EMR,可以通过两种方式进行数据迁移,第一种是通过腾讯云对象存储(COS)进行数据中转迁移,第二种是通过 Hadoop 自带文件迁移工具 DistCp 进行数据迁移。本文主要介绍通过 DistCp 进行数据迁移。

DistCp(distributed copy)是 Hadoop 自带的文件迁移工具。它使用 MapReduce 来实现其分发、错误处理和恢复、 报告的功能。它将文件和目录的列表扩展为映射任务的输入,每个任务将复制源列表中指定的文件的分区。使用 DistCp 需要实现自建集群和 EMR 集群的网络互通。 使用 DistCP 数据迁移步骤如下:

步骤1:网络打通

本地自建 HDFS 文件迁移到 EMR

本地自建 HDFS 文件迁移到 EMR 集群需要有专线打通网络,可以联系开发人员协助解决。

CVM 上的自建 HDFS 文件迁移到 EMR

- CVM 的所属网络和 EMR 集群的所属网络在同一 VPC 下,则可以自由传送文件。
- CVM 的所属网络和 EMR 集群的所属网络在不同 VPC 下,需要使用对等连接将网络打通。

使用对等连接

网段1:广州的 VPC1 中的子网 A 192.168.1.0/24。 网段2:北京的 VPC2 中的子网 B 10.0.1.0/24。

1. 登录 私有网络控制台-对等连接, 在列表上方选择地域广州, 选择私有网络 VPC1, 然后单击**+新建**。

Virtual Private Cloud	Peering connection	South China (C	Suangzho	-) -	All VPCs 🔻									Help of peering connection \mathbb{E}
l目 Network Topology Map		To get notified abou	ut abnorm	al peer conne	ction behaviors instantly	r, please Configure alarms,								
Virtual Private Cloud	[+ Create											Search by peering connect $ {\bf Q}_{\!$	\$
Subnet		ID/Name	Mo	Status	Local region	Local VPC	Peer region	Peer account	Peer VPC	Band	Servi	Billing mode	Operation	
Route Tables							No	results found						
묘 IP and Interface 💉														
Shared Bandwidth Pack														
🛱 NAT Gateway														
A Peering Connections														

2. 进入建立对等连接页。

• 名称:对等连接的名称,例如 PeerConn。



- 本端地域:填写本地端地域,例如广州。
- •本端网络:填写本端网络,例如 VPC1。
- 对端账户类型:填写对端网络所属账户,如果广州和北京两个网络在同一账户下,选择我的账户,如果不在同一
 账户,则要选择其它账户。

说明:

如果本端网络和对端网络都在同一地域,例如广州,通信是免费的,也不需要选择**带宽上线**;如果不在同 一地域,就要进行收费,同时带宽上限可选。

• 对端地域:填写对端地域,例如北京。



• 对端网络:填写对端网络,例如 VPC2。

Create peering connectio	n	×
Name		
Local region	South China (Guangzhou)	
Local network	vpc-3dio4ezl (Roy-001 10.0.0.0/16) 🔹	
Destination account type	O My Account Other accounts	
Peer region	South China (Guangzhou)	
Peer network	Please select	
Bandwidth cap	Unlimited	
Billing method	A 5 Gbps free tier is applied for intra-region bandwidth. And for cross-region bandwidth, No free tier , Billing description	or
	Create Cancel	

- 同账户内私有网络进行连接,新建后对等连接立即生效;与其它账户私有网络创建对等连接,需要对端接受此对 等连接后生效。参见 同账号创建对等连接通信 和 跨账号创建对等连接通信。
- 5. 为对等连接配置本端和对端路由表。
- 登录 私有网络控制台,单击左侧目录中的**子网**,进入管理页面。单击对等连接本端指定子网(例如广州的子网 VPC1)的关联路由表 ID,进入路由表详情页。



Subnet	Subnet Subnet Guangchou 3 * Al VPCs *										
Create										Please enter the Subnet	Q ¢ ‡ ₺
ID/Name		Network	CIDR	Availability zone T	Associated route table	CVM	Available IPs	Default subnet	Creation time	Tags T	Operation
						0 🍞	253	No	2022-07-01 10:37:10		Delete More 🔻

• 单击**+新增路由策略**。

Add route policy	Export Enable Disab	e				Destination address	Q
Destination	Next hop type 🔻	Next hop	Remark	Enable routing	Route status in CCN	Operation	
1	LOCAL	Local	Delivered by default, indicates that CVMs in the VPC are interconnected.		-	①Publish to CCN	
					20 /		ь ы

• 目的端中填入对端 CIDR(例如北京的 VPC2 的 CIDR 是10.0.1.0/24),下一跳类型选择**对等连接**,下一跳选择已 建立的对等连接(PeerConn)。

Add route policy	xport Enable Disab	e				Destination address	Q
Destination	Next hop type T	Next hop	Remark	Enable routing	Route status in CCN	Operation	
-	LOCAL	Local	Delivered by default, indicates that CVMs in the VPC are interconnected.		-	Publish to CCN	
+					30 /		ь ы

- 以上步骤是配置广州 VPC1 到北京 VPC2 的路由表,还需要配置北京 VPC2 到广州 VPC1 的配置,配置过程同上。
- 路由表配置完成后,不同私有网络的网段之间即可进行通信。

步骤2:执行拷贝

集群间的拷贝, 将一个文件夹拷贝到另一个集群 hadoop distcp hdfs://nn1:9820/foo/bar hdfs://nn2:9820/bar/foo

指定文件拷贝 hadoop distcp hdfs://nn1:9820/foo/a hdfs://nn1:9820/foo/b hdfs://nn2:9820/bar/foo

如果指定的文件太多, 可使用 -f 参数。

注意:



- 对于上述命令,必须要求源和目的版本相同。
- 如果另一个客户端仍然在写入源文件,则该拷贝可能会失败;如果一个文件正在被拷贝到目的端,试图重 写该文件的操作会失败;如果源文件在被复制之前被移动,那么拷贝将失败,报错信息为 FileNotFoundException。