

Alibaba Cloud E-MapReduce

集群規劃與配置

檔案版本：20200121

目錄

| | |
|----------------------------|-----------|
| 1 叢集規劃 | 1 |
| 1.1 角色授權..... | 1 |
| 1.2 使用者管理..... | 2 |
| 1.3 執行個體類型..... | 3 |
| 1.4 Gateway執行個體..... | 4 |
| 1.5 ECS執行個體說明..... | 4 |
| 1.6 儲存說明..... | 5 |
| 1.7 本地碟機型支援..... | 7 |
| 1.8 傳統網路與VPC互訪..... | 8 |
| 1.9 EMR叢集容災能力..... | 8 |
| 2 叢集 | 10 |
| 2.1 建立叢集..... | 10 |
| 2.2 擴容叢集..... | 14 |
| 2.3 釋放叢集..... | 15 |
| 2.4 叢集續約..... | 15 |
| 2.5 SSH 登入叢集..... | 17 |
| 2.6 服務列表..... | 21 |
| 2.7 叢集指令碼..... | 22 |
| 2.8 安全性群組..... | 23 |
| 2.9 建立Gateway..... | 24 |
| 2.10 Auto Scaling..... | 29 |
| 2.10.1 Auto Scaling概述..... | 29 |
| 2.10.2 按時間伸縮規則配置..... | 30 |
| 2.10.3 Auto Scaling記錄..... | 31 |
| 2.11 專用網路..... | 32 |
| 2.12 MetaService..... | 33 |

1 叢集規劃

1.1 角色授權

在使用者開通 E-MapReduce 服務時，需要授予一個名稱為“AliyunEMRDefaultRole”的系統預設角色給 E-MapReduce 的服務帳號，若且唯若該角色被正確授予後，E-MapReduce 才能正常地調用相關服務（ECS，OSS 等），建立叢集以及儲存日誌等。

角色授權流程

1. 當使用者建立叢集或建立按需執行計畫時，如果沒有正確地給 E-MapReduce 的服務帳號授予預設角色，則會看到如下提示，此時需單擊前往 RAM 進行授權，進行角色授權。
2. 單擊同意授權，將預設角色 AliyunE-MapReduceDefaultRole 授予給 E-MapReduce 的服務帳號。
3. 當完成以上授權步驟後，使用者需重新整理 E-MapReduce 的控制台，然後就可以進行了。如果您想查看 AliyunE-MapReduceDefaultRole 相關的詳細策略資訊，可以登入 RAM 的控制台查看，也可以單擊[查看連結](#)。

預設角色包含的許可權內容

預設角色 AliyunEMRDefaultRole 包含的許可權資訊如下：

- ECS 相關許可權：

| 許可權名稱 (Action) | 許可權說明 |
|-----------------------------|-------------|
| ecs:CreateInstance | 建立 ECS 執行個體 |
| ecs:RenewInstance | ECS 執行個體續約 |
| ecs:DescribeRegions | 查詢 ECS 地區資訊 |
| ecs:DescribeZones | 查詢 Zone 資訊 |
| ecs:DescribeImages | 查詢鏡像資訊 |
| ecs:CreateSecurityGroup | 建立安全性群組 |
| ecs:AllocatePublicIpAddress | 分配公網 IP |
| ecs>DeleteInstance | 刪除機器執行個體 |
| ecs:StartInstance | 啟動機器執行個體 |

| 許可權名稱 (Action) | 許可權說明 |
|------------------------------------|-------------|
| ecs:StopInstance | 停止機器執行個體 |
| ecs:DescribeInstances | 查詢機器執行個體 |
| ecs:DescribeDisks | 查詢機器相關磁碟資訊 |
| ecs:AuthorizeSecurityGroup | 設定安全性群組入規則 |
| ecs:AuthorizeSecurityGroupEgress | 設定安全性群組出規則 |
| ecs:DescribeSecurityGroupAttribute | 查詢安全性群組詳情 |
| ecs:DescribeSecurityGroups | 查詢安全性群組列表資訊 |

• OSS 相關許可權：

| 許可權名稱 (Action) | 許可權說明 |
|-----------------|------------|
| oss:PutObject | 上傳檔案或檔案夾對象 |
| oss:GetObject | 擷取檔案或檔案夾對象 |
| oss:ListObjects | 查詢檔案清單資訊 |

1.2 使用者管理

使用者管理是在指定叢集上建立相關服務所需的帳號，目前支援建立Knox帳號和Kerberos帳號。

建立RAM帳號

1. 登入[阿里雲 E-MapReduce 控制台](#)，進入叢集列表頁面。
2. 單擊叢集ID右側的組態管理。
3. 在左側導覽列中單擊使用者管理。
4. 單擊右側頁面建立RAM子帳號。

添加Knox帳號

1. 進入使用者管理頁面，在頁面上選擇需要添加到叢集的帳號，單擊操作列的設定Knox密碼。
2. 在添加Knox使用者的對話方塊中設定密碼，然後單擊確認。
3. 重新整理使用者管理頁面，當Knox帳號列的狀態顯示已同步時，表示Knox帳號添加成功。

添加成功後，您即可使用帳號名稱及**第2步**中設定的密碼登入Knox。

刪除Knox帳號

1. 進入使用者管理頁面，在頁面上選擇需要添加到叢集的帳號，單擊操作列的刪除Knox密碼。
2. 重新整理使用者管理頁面，當Knox帳號列的狀態顯示未同步時，表示Knox帳號刪除成功。

常見問題

- 不同叢集不能共用Knox帳號，即：在cluster-1上添加Knox帳號A之後，並不會共用給cluster-2，若想在cluster-2上使用Knox帳號A，需要在cluster-2上重新添加帳號A。原因是Knox帳號是建立在叢集中的，每個叢集的Knox帳號不互連。
- 添加Knox帳號過程中，若顯示同步失敗，請單擊重試重新添加。
- 添加Knox帳號過程中，多次重試仍然失敗，單擊左側叢集與服務管理，檢查ApacheDS是否已停止，若已停止則啟動ApacheDS後，再回到使用者管理進行重試。

1.3 執行個體類型

EMR叢集中由多個不同的節點執行個體類型組成，他們分別是主執行個體節點（Master），核心執行個體節點（Core）和計算執行個體節點（Task），每一種不同的執行個體在部署的時候會部署完全不同的服務進程，以完成完全不同的任務。舉例來說，我們會在主執行個體節點（Master）上部署 Hadoop HDFS 的 Namenode 服務，Hadoop YARN 的 ResourceManager 服務，而在核心執行個體節點（Core）上部署 Datanode 服務，Hadoop YARN 的 NodeManager 服務，在計算執行個體節點（Task）顧名思義，只進行計算，部署 Hadoop YARN 的 NodeManager 服務，不部署任何 HDFS 相關的服務。

在建立叢集的時候需要確定對應的三種執行個體類型的 ECS 規格，相同執行個體類型的 ECS 在同一個執行個體組內。並且可以在後期通過擴容來擴容對應執行個體組內的機器數量（主執行個體組除外）。



说明:

計算執行個體節點（Task）從3.2.0及以後版本開始支援。

Master 主執行個體

主執行個體是叢集服務的管控等組件的部署的節點，舉例來說，Hadoop YARN 的 ResourceManager 就部署在主執行個體節點上。使用者可以通過 SSH 的方式串連到主執行個體上，通過軟體的 Web UI 來查看叢集上的服務的運行情況。同時當需要進行快速的測試或者運行作業的時候，也可以登入到主執行個體上，通過命令列來直接提交作業。當叢集開啟了高可用的時候會有2個主執行個體節點，預設只有1個。

Core 核心執行個體

核心執行個體是被主執行個體管理的執行個體節點。上面會運行 Hadoop HDFS 的 Datanode 服務，並儲存所有的資料。同時也會部署計算服務，比如 Hadoop YARN 的 NodeManager 服務，來執行計算任務。為滿足儲存資料量或者是計算量上升的需要，核心執行個體可以隨時的擴容，不影響當前叢集的正常運行。核心使用可以使用多種不同的儲存介質來儲存資料。參考磁碟介紹。

Task 計算執行個體

計算執行個體是專門負責計算的執行個體節點，是一個可選的執行個體類型。如果核心執行個體的計算能力足夠的情況下，可以不使用計算執行個體。計算執行個體可以在任何時候快速的為叢集增加額外的計算能力，如 Hadoop 的 MapReduce tasks, Spark executors 等。在計算執行個體上不會儲存 HDFS 的資料，因此在計算執行個體上不運行 Hadoop HDFS 的 Datanode 服務。計算執行個體可以隨時的新增和減少，都不會影響到現有叢集的運行。計算執行個體節點的減少可能會引起 MapReduce 和 Spark 的作業的失敗，能否成功取決於該計算服務的重試容錯能力。

1.4 Gateway 執行個體

Gateway 一般為獨立的一個叢集，由多台相同配置的節點群組成。

在建立 Gateway 叢集時，可以關聯到一個已經存在的 Hadoop 叢集上，該叢集上會部署 Hadoop (HDFS+YARN)、Hive、Spark、Sqoop、Pig 等用戶端，方便對叢集進行操作。這樣做的好處是：它可以作為一個獨立的提交點，不會佔用叢集的資源，尤其是在 Master 提交的方式，可以提高 Master 節點的穩定性。如果作業太多，可以動態增加節點。

您可以建立多個不同的 Gateway 叢集，來給不同的使用者使用，讓他們可以使用各自專屬的環境配置來滿足不同的業務需求。

1.5 ECS 執行個體說明

EMR 目前支援的 ECS 執行個體類型

- 通用型

VCPU 與 Memory 比為 1: 4，如 32 核 128G，使用雲端硬碟作為儲存。

- 計算型

VCPU 與 Memory 比為 1: 2，如 32 核 64G，使用雲端硬碟作為儲存，提供了更多的計算資源。

- 記憶體型

VCPU與Memory比為1: 8，如32核256G，使用雲端硬碟作為儲存，提供了更多的記憶體資源。

- 大資料型

使用本地SATA盤作為資料存放區，擁有很高的儲存性價比，是大資料量（T層級的資料量）情境下的推薦機型。

- 本地SSD型

使用本地SSD盤，擁有極高的本地iops和吞吐能力。總體的儲存

- 共用型（入門級）

共用CPU的機型，在大計算量的情境下，穩定性不夠。入門級學習使用，不推薦企業客戶使用。

- GPU

使用GPU的異構機型，可以用來運行機器學習服務等情境。

執行個體類型適用情境

- Master 主執行個體

適合通用型或記憶體型執行個體，資料直接使用阿里雲的雲端硬碟來儲存，有三個備份的保證，資料高可靠。

- Core 核心執行個體

小資料量（T以下）或者是使用OSS作為主要的資料存放區時，可以使用通用型，計算型或記憶體型。當資料量較大的情況下，如10T或以上，推薦使用大資料機型，會獲得極高的性價比。使用本地碟會有一個資料可靠性的挑戰，會由EMR平台來進行維護和保證。

- Task 計算執行個體

作為叢集的計算能力的補充，可以使用除大資料型以外的所有的機型。目前本地SSD型尚未支援，後續會加入到Task中。

1.6 儲存說明

雲端硬碟與本地碟

在節點上存在2種角色的磁碟，一類是系統硬碟，用來安裝作業系統。一類是資料盤，用來儲存資料。系統硬碟預設都是一塊，必須使用雲端硬碟。而資料盤可以有許多塊，目前上限可以一個節點掛16塊。每一塊都可以有不同的配置，類型和容量都可以不同。EMR預設使用SSD雲端硬碟作為

叢集的系統硬碟。EMR 預設掛載4塊雲端硬碟，目前的內網頻寬的情況下4塊雲端硬碟是比較合理的配置。

有2種類型的磁碟可以用作資料的儲存：

- 雲端硬碟

包括，SSD 雲端硬碟，高效雲端硬碟，普通雲端硬碟。

特點是，磁碟並不直接掛載在本地的計算節點上，通過網路訪問遠端的一個儲存節點。每一份資料在後端都有2個即時備份，一共三份資料。所以當一份資料損毀的時候（磁碟損壞，不是使用者自己的業務上的破壞），會自動的使用備份資料恢復。

- 本地碟

包括，大資料型的 SATA 本地碟，和本地 SSD 的本地 SSD 盤。

直接掛載在計算節點上的磁碟，擁有超過雲端硬碟的效能表現。使用本地碟的時候不能選擇數量，只能使用預設配置好的數量，和線下物理機一樣，資料沒有後端的備份機制，需要上層的軟體來保證資料可靠性。

適用的情境

在EMR中，所有雲端硬碟和本地碟都會在節點釋放的時候清除資料，磁碟無法獨立的儲存下來，並再次使用。Hadoop HDFS 會使用所有的資料盤作為資料存放區。Hadoop YARN 也會使用所有的資料作為計算的臨時儲存。

當業務資料量並不太大（T層級以下）的時候，可以使用雲端硬碟，IOPS和吞吐相比本地碟都會小些。資料量大的時候，推薦都使用本地碟，EMR 會來維護本地碟的資料可靠性。如果發現在使用中明顯的輸送量不夠用，可以切換到本地碟的儲存上。

OSS

在 EMR 中可以將 OSS 作為 HDFS 使用。使用者可以非常方便的讀寫OSS，所有使用 HDFS 的代碼也可以簡單的修改就能訪問 OSS 上的資料了。

比如：

Spark中讀取資料

```
sc.textfile("hdfs://user/path")
```

替換儲存類型 hdfs -> oss

```
sc.textfile("oss://user/path")
```

對於MR或者Hive作業也是一樣

HDFS命令直接操作OSS資料

```
hadoop fs -ls oss://bucket/path  
hadoop fs -cp hdfs://user/path oss://bucket/path
```

這個過程，不需要輸入AK和endpoint，EMR都會自動替使用者使用當前叢集所有者的資訊補全。

但OSS的iops不高，在一些需要高IOPS的情境，不適合使用，比如流式計算Spark Streaming或HBase。

1.7 本地碟機型支援

阿里雲為了滿足大資料情境下的儲存的需求，在雲上推出了本地碟的機型：D1系列。這個系列提供了本地碟而非雲端硬碟作為儲存。解決了之前使用雲端硬碟的多份冗餘資料導致的成本高問題，同時資料的傳輸不再需要全部通過網路，從而提高了磁碟的吞吐能力。同時還能發揮Hadoop的就近計算的優勢。

相比於使用雲端硬碟的方式，極大的提高了儲存的效能，並降低了儲存的單價，達到和線下物理機幾乎相同的成本。

本地碟機型在提供了大量的優勢的情況下，也帶來了一個問題：資料可靠性。對於雲端硬碟來說，由於有阿里雲預設的磁碟多備份策略，所以使用者可以說完全感知不到磁碟的損壞，由雲端硬碟自動保證資料可靠，當使用了本地碟以後這個就需要由上層的軟體來保證。同時如果有磁碟與節點的故障情況，也需要進行人工的營運處理。

EMR + D1 方案

EMR產品針對本地碟機型，如D1，推出了一整套的自動化營運方案，協助阿里雲使用者方便可靠的使用本地碟機型，不需要關心整個營運的過程的同時，做到資料高可靠，服務高可用。

主要的一些點如：

- 強制節點的高可靠分布
- 本地碟與節點的故障監控
- 資料移轉時機自動決策
- 自動的故障節點遷移與資料平衡
- 自動的HDFS資料檢測
- 網路拓撲調優

EMR通過整個背景管控系統的自動化營運，協助使用者更好的使用本地碟機型，實現高性價比的大資料系統。



说明:

如需使用 D1 機型搭建 Hadoop 叢集，請工單聯絡我們協助操作。

1.8 傳統網路與VPC互訪

目前阿里雲存在兩種網路類型，一個是傳統網路，一個是VPC。很多使用者的業務系統因為曆史的原因還會在傳統網路中，而EMR叢集是在VPC中，本節將會介紹如何使傳統網路的ECS可以和VPC網路下的EMR叢集網路互訪。

ClassicLink

為瞭解決這個問題，阿里雲推出了 [ClassLink](#) 方案。

大致步驟如下，詳細的請參考上面的ClassLink文檔：

1. 首先按照上面文檔中指定網段建立vswitch。
2. 在建立叢集的時候，請使用該網段的vswitch來部署叢集。
3. 在ECS控制台將對應的傳統網路節點串連到這個VPC。
4. 設定安全性群組訪問規則。

然後就可以讓傳統網路的ECS和VPC的叢集互訪了。

1.9 EMR叢集容災能力

資料容災

Hadoop分散式檔案系統(HDFS)將每一個檔案的資料進行分區塊存放裝置，同時每一個資料區塊又儲存有多個副本(系統預設為每一個資料區塊存放3個副本)，盡量保證這些資料區塊副本分布在不同的機架之上（在大多數情況下，副本係數是3，HDFS的存放策略是將一個副本存放在本地機架節點上，一個副本存放在同一個機架的另一個節點上，最後一個副本放在不同機架的節點上）。

HDFS會定期掃描資料副本，若探索資料副本發生丟失，則會快速的進行資料的複製以保證副本的數量。若發現節點丟失，則節點上的所有資料也會快速的進行複製恢復。在阿里雲上，如果是使用雲端硬碟的技術，則在後台每一個雲端硬碟都會對應三個資料副本，當其中的任何一個出現問題時，副本資料都會自動進行切換並恢復，以保證資料的可靠性。

Hadoop HDFS是一個經歷了長時間考驗且具有高可靠性的資料存放區系統，已經能夠實現海量資料的高可靠性儲存。同時雲端式上的特性，也可以在OSS等服務上進行資料的額外備份，來達到更高的資料可靠性。

服務容災

Hadoop的核心組件都會進行HA的部署，即有至少2個節點的服務互備，如YARN，HDFS，Hive Server，Hive Meta，以保證在任何時候，其中任何一個服務節點掛掉時，當前的服務節點都能自動的進行切換，保證服務不會受到影響。

2 叢集

2.1 建立叢集

進入建立叢集頁面

1. 登入[阿里雲 E-MapReduce 控制台叢集列表](#)。
2. 完成 RAM 授權，操作步驟請參見[角色授權](#)。
3. 在上方選擇所在的地區（Region），所建立叢集將會在對應的地區內，一旦建立後不能修改。
4. 單擊右上方的建立叢集，進行建立。

建立叢集流程



注意：

叢集除了名字以外，一旦建立完成就無法被修改。所以在建立時請仔細確認需要的配置。

要建立叢集，您需要繼續完成以下 3 個步驟：

1. 軟體配置

配置項說明：

- **產品版本：**E-MapReduce 產品的主要版本，代表了一整套的開源軟體環境，它會定時的根據內部組成軟體的升級進行升級。一般如果 Hadoop 相關的軟體有進行升級，E-MapReduce 也會升級，這個時候就會升級這個主要版本號。低版本的叢集無法自動的升級到一個高版本上。
- **叢集類型：**目前的EMR提供了
 - Hadoop標準的 Hadoop 叢集，包含了大部分的 Hadoop 相關的組件，具體的組件資訊可以在選擇介面的列表中查看。
 - Kafka獨立的 Kafka 叢集，提供Message Service。
- **包含配置：**展示選擇的叢集類型下的所有的軟體組件列表，包括名稱和版本號碼。根據需求，您可選擇不同的組件，被選中的組件會預設啟動相關的服務進程。



說明：

您選擇的組件越多，對您機器的配置要求就越高，否則很可能無法有足夠的資源來運行這些服務。

- 安全模式：是否開啟叢集的 Kerberos 認證功能。
- 軟體配置（可選）：可以對叢集中的基礎軟體例如 Hadoop、Spark、Hive 等進行配置，詳細使用方法請參見[#unique_14](#)。

2. 硬體設定

配置項說明：

· 付費配置

- 付費類型：訂用帳戶是一次性支付一個長期的費用，價格相對來說會比較便宜，特別是包三年的時候折扣會很大。隨用隨付是根據實際使用的小時數來支付費用，每個小時計一次費用。適合與短期的測試或者是靈活的動態任務，價格相對來說會貴一些。
- 購買時間長度：您可選擇購買 1 個月、2 個月、3 個月、6 個月、9 個月、1 年、2 年、3 年。

· 叢集網路設定

- 叢集可用性區域：選擇叢集所在的可用性區域（Zone），不同的可用性區域會有不同的機型和磁碟。在每個 Region 記憶體在多個可用性區域。可用性區域在物理上屬於不同的地區，一般來說如果需要較好的網路，推薦您選擇相同的可用性區域，但是這樣也會使建立叢集失敗的風險增大，因為單個可用性區域的存庫不一定那麼充足。如果需要大量的機器可以工單諮詢我們。
- 網路類型：可以選擇傳統網路和專用網路（VPC），專用網路需要額外提供所屬 VPC 以及子網（交換器），若還未建立，可前往[VPC控制台](#)進行建立。E-MapReduce 專用網路詳細使用說明查看[專用網路](#)。

注意：傳統網路與專用網路不互連，購買後不能更換網路類型。

- ECS 執行個體系列：不同的可用性區域有不同的執行個體系列，系列 I、II、III 等。盡量使用最新的系列。
- VPC：選擇在該地區的 VPC。
- 交換器：選擇在對應的 VPC 下的在對應可用性區域的交換器，如果在這個可用性區域沒有可用的交換器，那麼就需要前往去建立一個新的使用。
- 建立安全性群組：一般使用者初次來到這裡還沒有安全性群組，開啟“建立安全性群組”開關，在“安全性群組名稱”裡面填上新的安全性群組的名字。
- 選擇安全性群組：叢集所屬的安全性群組。這裡只展示使用者在 E-MapReduce 產品中建立的安全性群組，目前尚不支援選擇在 E-MapReduce 外建立的安全性群組。如果需要建立安全性群組，可以選擇“建立安全性群組”選項，同時輸入安全性群組的名字完

成建立。長度限制為 2-64 個字元，以大小寫字母或中文開頭，可使用中文、字母、數字、“-”和“_”。

· 叢集節點配置

- 高可用叢集：開啟後，Hadoop 叢集會有 2 個 master 來支援 Resource Manager 和 Name Node 的高可用。HBase 叢集原來就支援高可用，只是另一個節點用其中一個 core 節點來充當，如果開啟高可用，會獨立使用一個 master 節點來支援高可用，更加的安全可靠。預設為非高可用模式，master 節點數量為 1。
- 節點類型：
 - Master 主執行個體節點，主要負責 Resource Manager, Name node 等控制進程的部署
 - Core 核心執行個體節點，主要負責叢集所有資料的儲存，可以按照需要進行擴容
 - Task 純計算節點，不儲存資料。調整叢集的計算力使用。
- 節點配置：不同規格的機型的選擇。各個機型有各自比較適用的情境，可以根據需要選擇。
- 資料盤類型：叢集的節點使用的資料盤類型，資料盤有 3 種類型，普通雲端硬碟、高效雲端硬碟和 SSD 雲端硬碟，根據不同機型和不同的 Region，會有不同。當使用者選擇不同的區的時候，該區支援什麼盤，下拉框就會展示什麼類型的盤。資料盤預設設定為隨著叢集的釋放而釋放。本地碟的計算節點，磁碟是預設選定的，無法修改。
- 資料盤容量：目前推薦的叢集容量最小是 40G 單機，最大可以到 32T 單節點。本地碟的容量是預設的，無法調整。
- 執行個體數量：需要的總的節點的台數。一個叢集至少需要 3 台執行個體（高可用叢集需增加 1 個 Master 節點，至少 4 台）。按量叢集目前最大台數是 50 台，如果需要超過 50 台，請工單聯絡我們。包月叢集最大 100 台，超過 50 台請工單聯絡我們。

3. 基礎配置

配置項說明：

· 基本資料

叢集名稱：叢集的名字，長度限制為 1-64 個字元，僅可使用中文、字母、數字、“-”和“_”。

· 作業記錄

- 作業記錄：是否儲存作業的日誌，日誌儲存預設是開啟的。開啟後會需要您選擇用來儲存日誌的 OSS 目錄位置，會將您的作業的日誌儲存到該 OSS 儲存目錄上。當然，您要使用

這個功能必須先開通 OSS，同時上傳的檔案會按照使用的量來計算使用者的費用。強烈建議您開啟 OSS 日誌儲存功能，這會對您的作業調試和錯誤排查有極大的協助。

- 日誌路徑：儲存日誌的 OSS 路徑。
- 統一Meta資料庫：將你所有的 Hive 的元資料都儲存到叢集外部的資料庫上，由EMR產品提供。推薦當叢集使用 OSS 作為主要的儲存的時候，使用這個功能。
- 使用權限設定
 - 服務角色：這個是使用者將許可權授予EMR服務，允許 EMR 代表使用者調用其他阿里雲的服務，例如 ECS 和 OSS
 - ECS應用角色：這個是當使用者的程式在 EMR 計算節點上啟動並執行時候，可以不填寫阿里雲的 AK 來訪問相關的雲端服務，比如OSS。EMR 會自動的申請一個臨時 AK 來授權這次訪問。而這個 AK 的許可權控制將由這個角色來控制。
- 登入設定

登入密碼：設定 master 節點的登入密碼。8 - 30 個字元，且必須同時包含大寫字母、小寫字母、數字和特殊字元!@#%\$%^&*。
- 引導操作（可選）：您可以在叢集啟動 Hadoop 前執行您自訂的指令碼，詳細使用說明請參見[#unique_16](#)。

配置清單和叢集費用

頁面右邊會顯示您所建立叢集的配置清單以及叢集費用。根據付費類型的不同，會展示不同的價格資訊。隨用隨付叢集顯示每小時費用，訂用帳戶顯示總費用。

確認建立

當所有的資訊都有效填寫以後，建立按鈕會亮起，確認無誤後單擊建立將會建立叢集。



说明:

- 若是隨用隨付叢集，叢集會立刻開始建立。頁面會返回叢集列表頁，就能看到在列表中有一個叢集建立中的叢集。請耐心等待，叢集建立會需要幾分鐘時間。完成之後叢集的狀態會切換為叢集空閑。
- 若是訂用帳戶叢集，則會先產生訂單，在支付完成訂單以後叢集才會開始建立。

登入Core節點

登入Core節點，請按照以下步驟操作：

1. 在Master節點上切換到hadoop帳號。

```
su hadoop
```

2. 免密碼SSH登入到對應的Core節點。

```
ssh emr-worker-1
```

3. 通過sudo命令可以獲得root許可權。

```
sudo vi /etc/hosts
```

建立失敗

如果建立失敗，在叢集列表頁上會顯示叢集建立失敗，將滑鼠移動到紅色的驚嘆號上會看到失敗原因。

建立失敗的叢集可以不用處理，對應的計算資源並沒有真正的建立出來。這個叢集會在停留3天以後自動隱藏。

2.2 擴容叢集

當您的叢集資源（計算資源、儲存資源）不足的時候，您可以將您的叢集進行水平擴充。目前只能擴充您的 Core 節點，且使用的配置預設與您之前購買的 ECS 配置一致。

擴容入口

在叢集列表頁上，找到需要擴充的叢集條目，單擊調整規模就會進入叢集擴容頁面。也可以單擊查看詳情，然後在詳情頁的 Core 節點資訊位置單擊調整叢集規模。

擴容介面



說明：

目前，只支援擴容，不支援縮容。

- 當前 Master 節點數量：不可以調整，當前預設值是 1 個。
- 當前 Core 台數：預設顯示的是您當前的所有 Core 節點數量。
- 增加 Core 台數：輸入您實際需要增加的量，右側會顯示擴容後的叢集總費用，然後單擊確定就會進行擴容。安全起見，強烈推薦您在擴容的時候小批量的進行，比如 1 到 2 台，不要一次擴容太大的數量。

擴容狀態

您可以在詳情頁的 Core 節點資訊上看到叢集的擴容情況，正在擴容的節點，其狀態會顯示為擴容中。當這台 ECS 的狀態轉為正常後，該 ECS 即已經加入該叢集，並可正常提供服務了。

2.3 釋放叢集

在叢集列表頁面，您可以單擊叢集條目右側的釋放按鈕對叢集進行釋放操作。只有以下狀態的隨用隨付的叢集才可以被釋放。

- 建立中
- 運行中
- 空閑中

普通釋放

釋放前會提示您再次確認，一旦確認釋放，會發生以下的操作：

1. 所有在叢集上的作業都會被強制終止。
2. 如果您選擇了儲存日誌到 OSS，那麼當前作業的日誌會被儲存到 OSS，所需時間取決於日誌大小。日誌的上傳和作業的運行是並行的，作業組建記錄檔的同時，就會進行日誌的上傳。所以最終作業停止時，需要上傳的日誌一般不會特別多，正常在幾分鐘內都會完成。
3. 終止並釋放所有的ECS。這個過程取決於叢集的大小，越小的叢集會越快，正常都在幾秒內完成，至多不會超過5分鐘。最遲釋放的ECS在等待釋放時仍然計費。



警告：

如果您想要省錢而控制在整點前釋放，請務必留出一定的釋放時間保證確實在整點前釋放。

4. 釋放叢集結束。

強制釋放

如果您不需要任何日誌，只是想快速結束叢集的運行，那麼可以開啟強制釋放。釋放過程就會跳過日誌收集（如果有開啟日誌收集的話），直接進入ECS釋放階段。

釋放失敗的叢集

由於系統錯誤等原因，叢集有可能會在確認釋放後，釋放失敗。您單擊釋放是一個非同步過程，可能等一會兒會發生叢集釋放失敗的情況，但是不用擔心，E-MapReduce 會啟動後台保護，自動重試釋放該叢集，直到叢集被成功釋放為止。

2.4 叢集續約

當您的訂用帳戶叢集服務即將到期的時候，您需要執行叢集的續約操作，以繼續您的 E-MapReduce 叢集服務。叢集續約包括 E-MapReduce 服務費的續約以及叢集中 ECS 的續約。

續約入口

1. 登入 [阿里雲 E-MapReduce 叢集列表頁](#)。

2. 找到需要續約的叢集條目。
3. 單擊相應叢集條目操作欄下的續約，進入叢集續約頁面。

續約介面

- 續約：勾選需要續約的機器。
- ECS 執行個體 ID：叢集中這台機器的 ECS 執行個體 ID。
- ECS 當前到期時間：這台 ECS 到期的時間。
- E-MapReduce 當前到期時間：E-MapReduce 服務到期的時間。
- ECS 續約時間長度：需要對這個節點續約的時間長度，目前支援 1-9 個月、1 年、2 年和 3 年的續約時間長度。
- E-MapReduce 續約時間長度：對應的這個節點的 E-MapReduce 服務費的時間長度，推薦和 ECS 的設定一致。
- ECS 續約價格：ECS 節點對應的續約價格。
- E-MapReduce 續約價格：E-MapReduce 服務對應節點的續約價格。

支付訂單



注意：

叢集續約的費用為 ECS 續約價格與 E-MapReduce 服務供應項目價格的總和。當叢集列表中存在未支付的訂單時，您將無法執行叢集的擴容和續約操作。

1. 單擊確定後，會彈出下單成功提示的提示框（該提示資訊可能有較長的時間延遲，請耐心等待）。
2. 單擊支付訂單，跳轉到訂單支付頁面。支付頁面會顯示您應付的總金額以及各訂單的詳情，其中一個為 E-MapReduce 產品費用訂單，其它的為叢集續約的 ECS 訂單。
3. 單擊確認支付，完成付款。
4. 完成支付之後，單擊已完成，返回叢集列表頁面。

此時，在叢集列表頁上續約成功叢集的到期時間會更新為續約之後的到期時間，對應的 ECS 的到期時間更新目前存在一定的延遲，一般約 3-5 分鐘之後會更新成續約之後的到期時間。

若您只是確認了續約訂單但並未進行支付，您可在叢集列表頁面中找到該叢集條目，在其右側的操作欄中會出現前往支付和取消訂單的按鈕。您可單擊前往支付，完成對應的訂單支付和叢集擴容流程；或單擊取消訂單，以取消本次續約操作。

2.5 SSH 登入叢集

登入叢集內部的用途

若您覺得在網頁上的作業和執行計畫無法滿足您更加複雜的應用需求，您可以登入到 E-MapReduce 叢集的主機上，找到叢集的詳情頁，其中就有叢集 master 機器的公網 IP 位址，您可以直接 SSH 登入到這台機器上，查看各種設定與狀態。

機器上已為您設定好相關的環境變數，其中包括以下常用的環境變數：

- JAVA_HOME
- HADOOP_HOME
- HADOOP_CONF_DIR
- HADOOP_LOG_DIR
- YARN_LOG_DIR
- HIVE_HOME
- HIVE_CONF_DIR
- PIG_HOME
- PIG_CONF_DIR

您可以在指令碼中直接引用這些變數，但請不要去修改這些變數的值，以免造成 E-MapReduce 的意外錯誤。

登入 master 主機步驟

1. 使用如下命令 SSH 登入到 master 主機。請在[#unique_21](#)的硬體資訊列中擷取叢集 master 機器的公網 IP。

```
ssh root@ip.of.master
```

2. 輸入建立時設定的密碼。

打通本地機器與叢集 master 機器的 SSH 無密碼登入

通常，您需要登入到叢集上進行一些管理和操作。為便於登入叢集 master 機器，您可打通與 master 機器的 SSH 無密碼登入（叢集 master 機器預設開通了公網 IP）。操作步驟如下：

1. 通過上面提到的 root + 密碼的方式登入到 master 主機。
2. 切換到到Hadoop 使用者或者 hdfs 使用者。

Linux 的 SSH 方式

1. 複製私密金鑰到本地。

```
sz ~/.ssh/id_rsa
```

2. 回到您的本地機器，嘗試重新登入 master 機器。

```
ssh -i 私密金鑰存放路徑/id_rsa hadoop@120.26.221.130
```

當然如果你只有這一個私密金鑰，也可以直接放到你的 `~/.ssh/` 下，預設使用這個私密金鑰，就不需要 `-i` 指定了。

Windows 的 SSH 方式

在 Windows 下你可以有多種方式來使用 SSH 免密碼登入 master 機器。

- 方式一：使用 PuTTY

1. 下載 [PuTTY](#)。
2. 在同樣的位置下載 PuTTYgen。
3. 開啟 PuTTYgen，並 Load 您的私密金鑰。



注意：

請妥善保管這個私密金鑰，保證該私密金鑰的安全。若私密金鑰不幸洩漏了，請立刻重建一個新的取代。

4. 使用預設的配置，並 Save private key。會儲存出一個尾碼為 `ppk` 的 PuTTY 使用的密鑰檔案。
5. 運行 PuTTY，並在配置頁面選擇 Session。
6. 輸入您要串連的目標機器公網 IP 位址，要加上登入使用的使用者名，類似 `hadoop@MasterNodeIP`。
7. 在配置頁面，選擇 Connection > SSH > Auth。
8. 選擇之前產生好的 `ppk` 檔案。
9. 最後單擊 Open，就會自動登入到 master 節點了。

- 方式二：使用 Cygwin | MinGW

這是在 Windows 上類比 Linux 的非常方便的工具，使用起來也非常簡單。

如果採用這種方式，串連過程就可以參考上面的 Linux 的 SSH 方式了。

推薦採用 MinGW 的方式，這個是最小巧的一種方式。如果官網打不開，可以下載 git 的用戶端，預設帶的 Git Bash 就可以滿足。

查看 Hadoop、Spark、Ganglia 等系統的 webui



说明:

在進行本步驟前，請確認您已經完成了上面的 [SSH 無密碼登入](#) 流程。

由於安全的緣故，E-MapReduce 叢集的 Hadoop、Spark 和 Ganglia 等系統的 webui 監控系統的連接埠都沒有對外開放。如果使用者想要訪問這些 webui，需要建立一個 SSH 隧道，通過連接埠轉寄的方式來達到目的。有如下兩種方式：



注意:

下面的操作是在您本地機器上完成的，不是叢集內部機器。

· 方式一：連接埠動態轉寄

建立一個 SSH 隧道，該隧道可打通您本地機器跟 E-MapReduce 叢集的 master 機器的某個動態連接埠的串連。

```
ssh -i /path/id_xxx -ND 8157 hadoop@masterNodeIP
```

8157 是您本地機器沒有被使用過的任何一個連接埠，使用者可以自定義。

完成動態轉寄以後，您可以選擇如下兩種方式來查看。

- 推薦方式

推薦使用 Chrome 瀏覽器，可以使用如下的方式來訪問 Web UI：

```
chrome --proxy-server="socks5://localhost:8157" --host-resolver-rules="MAP * 0.0.0.0 , EXCLUDE localhost" --user-data-dir=/tmp/
```

若是 Windows 系統，這裡的 *tmppath* 可以寫成類似 *d:/tmppath*，若是 Linux 或者 OSX，可以直接寫成 */tmp/*。

在不同的作業系統中，Chrome 的位置不同，請參見下表：

| 作業系統 | Chrome 位置 |
|----------|--|
| Mac OS X | /Applications/Google Chrome.app/Contents/MacOS/Google Chrome |
| Linux | /usr/bin/google-chrome |

| 作業系統 | Chrome 位置 |
|---------|---|
| Windows | C:\Program Files (x86)\Google\Chrome\Application\chrome.exe |

- 外掛程式方式

此時，您本地機器跟 E-MapReduce 叢集的 master 主機的 SSH 通道已經打通，要在瀏覽器中查看 Hadoop、Spark、Ganglia 的 webui，您還需要配置一個本地代理。操作步驟如下：

1. 假設您使用的是 Chrome 或者 Firefox 瀏覽器，請點擊[下載 FoxyProxy Standard 代理軟體](#)。
2. 安裝完成並重啟瀏覽器後，開啟一個文字編輯器，編輯如下內容：

```
<?xml version="1.0" encoding="UTF-8"?>
<foxyproxy>
<proxies>
<proxy name="aliyun-emr-socks-proxy" id="2322596116" notes="" fromSubscription="false" enabled="true" mode="manual" selectedTabIndex="2" lastresort="false" animatedIcons="true" includeInCycle="true" color="#0055E5" proxyDNS="true" noInternalIPs="false" autoconfMode="pac" clearCacheBeforeUse="false" disableCache="false" clearCookiesBeforeUse="false" rejectCookies="false">
<matches>
<match enabled="true" name="120.*" pattern="http://120.*" isRegex="false" isBlackList="false" isMultiLine="false" caseSensitive="false" fromSubscription="false" ></match>
</matches>
<manualconf host="localhost" port="8157" socksVersion="5" isSocks="true" username="" password="" domain="" ></manualconf>
</proxy>
</proxies>
```

```
</foxyproxy>
```

其中：

- Port 8157 是您本地用來建立與叢集 master 機器 SSH 串連的連接埠，這個需要跟您之前執行的在終端中執行的 SSH 命令中使用的連接埠匹配。
- 120.* 這個匹配是用來匹配 master 主機的 IP 位址，請根據 master 的 IP 位址的情況來定。

3. 在瀏覽器中單擊Foxyproxy按鈕，選擇 Options。
4. 選擇 Import/Export。
5. 選擇剛才您編輯的 xml 檔案，單擊 Open。
6. 在 Import FoxyProxy Setting 對話方塊中，單擊 Add。
7. 點擊瀏覽器中的 Foxyproxy 按鈕，選擇 Use Proxy aliyun-emr-socks-proxy for all URLs。
8. 在瀏覽器中輸入 localhost:8088，就可以開啟遠端的 Hadoop 介面了。

· 方式二：本地連接埠轉寄



注意：

這個方式的缺陷是只能看到最外層的介面，一旦要看詳細的作業資訊，就會出錯。

```
ssh -i /path/id_rsa -N -L 8157:masterNodeIP:8088 hadoop@masterNodeIP
```

參數說明：

- path：私密金鑰存放路徑。
- masterNodeIP：要串連的 master 節點 IP。
- 8088：是 master 節點上 ResourceManager 的訪問連接埠號碼。

2.6 服務列表

叢集詳情頁的tab下面增加了服務列表頁，服務列表展示了HDFS，YARN等服務在叢集節點上的運行狀態。

只有叢集狀態是空閑、運行中的叢集會展示服務的運行狀態，如果點擊的服務在建立叢集時未勾選，例如Storm，則會提示沒有記錄。

列表包括節點內網ip和狀態，狀態有運行中和未運行兩種。如果某個節點上的服務狀態是未運行，您可以通過master節點跳轉，登入到對應節點上查看服務進程情況。

2.7 叢集指令碼

叢集指令碼的作用

叢集，特別是訂用帳戶叢集，在使用過程中，可能會有新的安裝第三方軟體，修改叢集運行環境的需求。叢集指令碼功能可以在叢集建立好後批量選擇節點，運行您指定的指令碼，以實現個人化的需求。

叢集指令碼類似引導操作，您可以在叢集建立好後安裝很多目前叢集尚未支援的軟體到您的叢集上，例如：

- 使用 yum 安裝已經提供的軟體。
- 直接下載公網上的一些公開的軟體。
- 讀取 OSS 中您的自有資料。
- 安裝並運行一個服務，例如 Flink 或者 Impala，但需要編寫的指令碼會複雜些。

強烈建議您先用單個節點進行叢集指令碼的測試，測試都正確以後再在整個叢集上操作。

如何使用

1. 叢集狀態是空閑或者運行中的叢集可以運行叢集指令碼，叢集列表頁面點擊對應叢集的查看詳情按鈕
2. 左側菜單單擊叢集指令碼，即會進入該叢集的叢集指令碼執行介面，右側是已經執行過的叢集指令碼列表。
3. 單擊右上方建立並執行，進入建立介面。
4. 填寫建立介面上的配置項，選擇執行的節點，點擊執行，完成添加並執行操作。
5. 叢集指令碼列表可以看到新建立的叢集指令碼，點擊重新整理可以更新叢集指令碼的狀態。
6. 單擊查看詳情可以看到指令碼在各個節點上的運行情況，單擊重新整理可以更新指令碼在各個節點上的運行狀態。

只有空閑或者運行中的可用叢集才能使用叢集指令碼功能。叢集指令碼適用於長期存在的叢集，對按需建立的臨時叢集，應使用引導操作來完成叢集初始化工作。

叢集指令碼會在您指定的節點上下載oss上的指令碼並運行，根據傳回值是否為0判斷執行成功還是失敗。如果運行狀態是失敗，您可以登入到各個節點上查看作業記錄，作業記錄記錄在每個節點的/var/log/cluster-scripts/clusterScriptId目錄下。如果叢集配置了oss日誌目錄，作業記錄也會上傳到osslogpath/clusterId/ip/cluster-scripts/clusterScriptId目錄下方便查看。

預設會使用 root 賬戶執行您指定的指令碼，您可以在指令碼中使用 su hadoop 切換到 Hadoop 賬戶。

叢集指令碼可能在部分節點上運行成功，部分節點上運行失敗，例如節點重啟導致的指令碼運行失敗。您可以在解決異常問題後，單獨指定失敗的節點再次運行。當叢集擴容後，您也可以指定擴容的節點單獨運行叢集指令碼。

1個叢集同一時間只能運行一個叢集指令碼，如果有正在啟動並執行叢集指令碼，無法提交執行新的叢集指令碼。每個叢集最多保留10個叢集指令碼記錄，超過10個需要將之前的記錄刪除才能建立新的叢集指令碼。

指令碼的例子

類似引導操作的指令碼，您可以在指令碼中指定從 OSS 下載需要的檔案，下面的例子會將 `oss://yourbucket/myfile.tar.gz` 這個檔案下載到本地，並解壓到 `/yourdir` 目錄下：

```
#!/bin/bash
osscmd --id=<yourid> --key=<yourkey> --host=oss-cn-hangzhou-internal.
aliyuncs.com get oss://<yourbucket>/<myfile>.tar.gz ./<myfile>.tar.gz
mkdir -p /<yourdir>
tar -zxvf <myfile>.tar.gz -C /<yourdir>
```

osscmd 已預先安裝在節點上，可以直接調用來下載檔案。



说明:

OSS 地址 host 有內網地址、外網地址和 VPC 網路地址之分。如果用傳統網路，需要指定內網地址，杭州是 `oss-cn-hangzhou-internal.aliyuncs.com`。如果用 VPC 網路，要指定 VPC 內網可訪問的網域名稱，杭州是 `vpc100-oss-cn-hangzhou.aliyuncs.com`。

指令碼也可以通過 yum 安裝額外的系統軟體包，下面的例子會安裝 `ld-linux.so.2`：

```
#!/bin/bash
yum install -y ld-linux.so.2
```

2.8 安全性群組

目前 E-MapReduce 建立叢集的時候需要使用在 E-MapReduce 中建立的安全性群組。

主要是因為 E-MapReduce 建立的叢集只開放了 22 連接埠，推薦您將 ECS 執行個體也按照功能劃分，放於不同的使用者安全性群組中。例如，E-MapReduce 的安全性群組為 E-MapReduce 安全性群組，而您已有的安全性群組為使用者-安全性群組，每個安全性群組按照不同的需要開啟不同的存取控制。

如果需要和已有叢集進行聯動，請參考如下做法。

將 E-MapReduce 叢集加入現有安全性群組

1. 登入 [阿里雲 E-MapReduce 控制台叢集列表](#)。

2. 找到需加入安全性群組的叢集條目，單擊其操作中的查看詳情，即會進入叢集詳情頁。
3. 在叢集詳情頁上，找到“網路資訊”下的所屬安全性群組，查看叢集所有 ECS 執行個體所在的安全性群組名稱。
4. 前往[阿里雲ECS的管理主控台](#)，單擊頁面左側的安全性群組，在列表中找到步驟 3 中查看到的安全性群組條目。
5. 單擊該安全性群組操作中的管理執行個體，您會看到很多名字以 emr-xxx 開頭的 ECS 執行個體，這些就是對應的 E-MapReduce 叢集中的 ECS。
6. 全選這些執行個體並單擊移入安全性群組，選擇想要加入的新的安全性群組即可。

將現有叢集加入 E-MapReduce 安全性群組

和上面的操作一樣，先找到現有叢集所在安全性群組，重複如上的操作，移入 E-MapReduce 的安全性群組即可。如果是一些零散的機器，也可以直接在 ECS 的控制台界面上選擇機器，然後通過下方的批量操作將叢集移入 E-MapReduce 的安全性群組。

安全性群組的規則

一個 ECS 執行個體在多個不同的安全性群組的時候，安全性群組的規則是或的關係。舉例來說就是，E-MapReduce 的安全性群組只開放了 22 連接埠，而使用者-安全性群組開放了所有的連接埠。當 E-MapReduce 的叢集加入使用者-安全性群組以後，E-MapReduce 中的機器也會開放所有連接埠，所以在使用上請特別注意。

2.9 建立Gateway

Gateway是與EMR叢集處於同一個內網中的ECS伺服器，使用者可以使用Gateway實現負載平衡和安全隔離，也可以通過Gateway向E-MapReduce叢集提交作業。

您可以通過以下兩種方式建立Gateway：

- （推薦）通過[EMR控制台](#)建立。
- 手動搭建。

通過E-MapReduce控制台建立Gateway

在建立Gateway前，請確保您已經建立了E-MapReduce叢集。建立Gateway，請按照如下步驟進行操作：

1. 登入[阿里雲E-MapReduce控制台](#)，進入叢集列表頁面。
2. 單擊頁面右上方的建立Gateway按鈕。

3. 在建立Gateway頁面中進行配置。

- 付費類型：訂用帳戶是一次性支付一段時間的費用，價格相對來說會比較便宜，特別是包三年的時候折扣會比較大。隨用隨付是根據實際使用的小時數來支付費用，每個小時計一次費用。
- 我的叢集：為該叢集建立Gateway，即建立的Gateway可以向哪個叢集提交作業。Gateway將會自動設定與該叢集一致的Hadoop環境。
- 配置：該地區內可選擇的ECS執行個體規格。
- 系統硬碟配置：Gateway節點使用的系統硬碟類型。系統硬碟有2種類型：SSD雲端硬碟和高效雲端硬碟，根據不同機型和不同的Region，系統硬碟顯示類型會有不同。系統硬碟預設隨著叢集的釋放而釋放。
- 系統硬碟大小：最小為40GB，最大為500GB。預設值為300GB。
- 資料盤配置：Gateway節點使用的資料盤類型。資料盤有2種類型：SSD雲端硬碟和高效雲端硬碟，根據不同機型和不同的Region，資料盤顯示類型會有不同。資料盤預設隨著叢集的釋放而釋放。
- 資料盤大小：最小為200GB，最大為4000GB。預設值為300GB。
- 數量：資料盤的數量，最小設定為1台，最大設定為10台。
- 叢集名稱：建立的Gateway的名稱，長度限制為1~64個字元，只允許包含中文、字母、數字、串連號 (-)、底線 (_)。
- 密碼/金鑰組：
 - Cipher 模式：在文字框中輸入登入Gateway的密碼。
 - 金鑰組模式：在下拉式功能表中選擇登入Gateway的金鑰組名稱。如果還未建立過金鑰組，單擊右側的建立金鑰組連結，進入ECS控制台建立。請妥善保管好金鑰組所對應的私密金鑰檔案 (.pem檔案)。在Gateway建立成功後，該金鑰組的公開金鑰部分會自動綁定到Gateway所在的Elastic Compute Service上。在您通過SSH登入Gateway時，請輸入該私密金鑰檔案中的私密金鑰。

4. 單擊建立完成Gateway的建立。

新建立的Gateway會顯示在叢集列表中，建立完成後，狀態列會變為空閑狀態。

手動搭建Gateway

· 網路環境

首先要保證Gateway節點在EMR對應叢集的安全性群組中，Gateway節點可以順利的訪問EMR叢集。設定節點的安全性群組請參考[建立安全性群組](#)。

- 軟體環境
 - 系統內容：推薦使用CentOS 7.2及以上版本。
 - Java環境：安裝JDK 1.7及以上版本，推薦使用OpenJDK version 1.8.0版本。
- 搭建步驟
 - EMR 2.7及以上版本，3.2及以上版本

這些版本推薦直接使用EMR控制台來建立Gateway。

如果您選擇手動搭建，請先建立一個指令碼，指令碼內容如下所示，然後在Gateway節點上執行。執行命令為：`sh deploy.sh <master_ip> master_password_file`。

■ `deploy.sh`：指令碼名稱，內容見下面代碼。

■ `master_ip`：叢集的master節點的IP，請確保可以訪問。

■ `master_password_file`：儲存master節點的密碼檔案，將master節點的密碼直接寫在檔案內即可。

```
#!/usr/bin/bash
if [ $# != 2 ]
then
    echo "Usage: $0 master_ip master_password_file"
    exit 1;
fi
masterip=$1
masterpwdfile=$2
if ! type sshpass >/dev/null 2>&1; then
    yum install -y sshpass
fi
if ! type java >/dev/null 2>&1; then
    yum install -y java-1.8.0-openjdk
fi
mkdir -p /opt/apps
mkdir -p /etc/ecm
echo "Start to copy package from $masterip to local gateway(/opt/apps)"
echo " -copying hadoop-2.7.2"
sshpass -f $masterpwdfile scp -r -o 'StrictHostKeyChecking no'
root@$masterip:/usr/lib/hadoop-current /opt/apps/
echo " -copying hive-2.0.1"
sshpass -f $masterpwdfile scp -r root@$masterip:/usr/lib/hive-current /opt/apps/
echo " -copying spark-2.1.1"
sshpass -f $masterpwdfile scp -r root@$masterip:/usr/lib/spark-current /opt/apps/
echo "Start to link /usr/lib/\${app}-current to /opt/apps/\${app}"
if [ -L /usr/lib/hadoop-current ]
then
    unlink /usr/lib/hadoop-current
fi
ln -s /opt/apps/hadoop-current /usr/lib/hadoop-current
if [ -L /usr/lib/hive-current ]
then
    unlink /usr/lib/hive-current
fi
```

```

ln -s /opt/apps/hive-current /usr/lib/hive-current
if [ -L /usr/lib/spark-current ]
then
    unlink /usr/lib/spark-current
fi
ln -s /opt/apps/spark-current /usr/lib/spark-current
echo "Start to copy conf from $masterip to local gateway(/etc/ecm
)"
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/ecm/hadoop-
conf /etc/ecm/hadoop-conf
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/ecm/hive-conf
/etc/ecm/hive-conf
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/ecm/spark-
conf /etc/ecm/spark-conf
echo "Start to copy environment from $masterip to local gateway(/
etc/profile.d)"
sshpass -f $masterpwdfile scp root@$masterip:/etc/profile.d/hdfs.
sh /etc/profile.d/
sshpass -f $masterpwdfile scp root@$masterip:/etc/profile.d/yarn.
sh /etc/profile.d/
sshpass -f $masterpwdfile scp root@$masterip:/etc/profile.d/hive.
sh /etc/profile.d/
sshpass -f $masterpwdfile scp root@$masterip:/etc/profile.d/spark.
sh /etc/profile.d/
if [ -L /usr/lib/jvm/java ]
then
    unlink /usr/lib/jvm/java
fi
echo "" >>/etc/profile.d/hdfs.sh
echo export JAVA_HOME=/usr/lib/jvm/jre-1.8.0 >>/etc/profile.d/hdfs
.sh
echo "Start to copy host info from $masterip to local gateway(/etc
/hosts)"
sshpass -f $masterpwdfile scp root@$masterip:/etc/hosts /etc/
hosts_bak
cat /etc/hosts_bak | grep emr | grep cluster >>/etc/hosts
if ! id hadoop >& /dev/null
then
    useradd hadoop
fi

```

- EMR 2.7以下版本， 3.2以下版本

建立一個指令碼，指令碼內容如下所示，然後在Gateway節點上執行。執行命令為：`sh deploy.sh <masterip> master_password_file`。

■ `deploy.sh`: 指令碼名稱，內容見下面代碼。

■ `masterip`: 叢集的master節點的IP，請確保可以訪問。

■ `master_password_file`: 儲存master節點的密碼檔案，將master節點的密碼直接寫在檔案內即可。

```

#!/usr/bin/bash
if [ $# != 2 ]
then
    echo "Usage: $0 master_ip master_password_file"
    exit 1;
fi
masterip=$1

```

```

masterpwdfile=$2
if ! type sshpass >/dev/null 2>&1; then
    yum install -y sshpass
fi
if ! type java >/dev/null 2>&1; then
    yum install -y java-1.8.0-openjdk
fi
mkdir -p /opt/apps
mkdir -p /etc/emr
echo "Start to copy package from $masterip to local gateway(/opt/
apps)"
echo " -copying hadoop-2.7.2"
sshpass -f $masterpwdfile scp -r -o 'StrictHostKeyChecking no'
root@$masterip:/usr/lib/hadoop-current /opt/apps/
echo " -copying hive-2.0.1"
sshpass -f $masterpwdfile scp -r root@$masterip:/usr/lib/hive-
current /opt/apps/
echo " -copying spark-2.1.1"
sshpass -f $masterpwdfile scp -r root@$masterip:/usr/lib/spark-
current /opt/apps/
echo "Start to link /usr/lib/\${app}-current to /opt/apps/\${app}"
if [ -L /usr/lib/hadoop-current ]
then
    unlink /usr/lib/hadoop-current
fi
ln -s /opt/apps/hadoop-current /usr/lib/hadoop-current
if [ -L /usr/lib/hive-current ]
then
    unlink /usr/lib/hive-current
fi
ln -s /opt/apps/hive-current /usr/lib/hive-current
if [ -L /usr/lib/spark-current ]
then
    unlink /usr/lib/spark-current
fi
ln -s /opt/apps/spark-current /usr/lib/spark-current
echo "Start to copy conf from $masterip to local gateway(/etc/emr
)"
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/emr/hadoop-
conf /etc/emr/hadoop-conf
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/emr/hive-conf
/etc/emr/hive-conf
sshpass -f $masterpwdfile scp -r root@$masterip:/etc/emr/spark-
conf /etc/emr/spark-conf
echo "Start to copy environment from $masterip to local gateway(/
etc/profile.d)"
sshpass -f $masterpwdfile scp root@$masterip:/etc/profile.d/hadoop
.sh /etc/profile.d/
if [ -L /usr/lib/jvm/java ]
then
    unlink /usr/lib/jvm/java
fi
ln -s /usr/lib/jvm/java-1.8.0-openjdk-1.8.0.131-3.b12.el7_3.x86_64
/jre /usr/lib/jvm/java
echo "Start to copy host info from $masterip to local gateway(/etc
/hosts)"
sshpass -f $masterpwdfile scp root@$masterip:/etc/hosts /etc/
hosts_bak
cat /etc/hosts_bak | grep emr | grep cluster >>/etc/hosts
if ! id hadoop >& /dev/null
then
    useradd hadoop

```

```
fi
```

- 測試

- Hive

```
[hadoop@iZ23bc05hrvZ ~]$ hive
hive> show databases;
OK
default
Time taken: 1.124 seconds, Fetched: 1 row(s)
hive> create database school;
OK
Time taken: 0.362 seconds
hive>
```

- 運行Hadoop作業

```
[hadoop@iZ23bc05hrvZ ~]$ hadoop jar /usr/lib/hadoop-current/share
/hadoop/mapreduce/hadoop-mapreduce-examples-2.6.0.jar pi 10 10
Number of Maps = 10
Samples per Map = 10
Wrote input for Map #0
Wrote input for Map #1
Wrote input for Map #2
Wrote input for Map #3
Wrote input for Map #4
Wrote input for Map #5
Wrote input for Map #6
Wrote input for Map #7
Wrote input for Map #8
Wrote input for Map #9
File Input Format Counters
  Bytes Read=1180
File Output Format Counters
  Bytes Written=97
Job Finished in 29.798 seconds
Estimated value of Pi is 3.20000000000000000000
```

2.10 Auto Scaling

2.10.1 Auto Scaling概述

在以下情境中，您可以通過E-MapReduce Auto Scaling功能節省成本，提高執行效率。

- 臨時需要按照時間段添加計算節點，補充計算能力。
- 確保重要作業按時完成，按照某些叢集指標擴充計算節點。



说明:

Auto Scaling功能僅能對Task節點進行擴容或縮容。

開啟Auto Scaling

1. 登入[阿里雲 E-MapReduce 控制台](#)，進入叢集列表頁面。

2. 單擊叢集ID右側的組態管理。
3. 在左側導覽列中單擊Auto Scaling。
4. 單擊頁面右上方的開啟Auto Scaling按鈕。

如果當前帳號是首次使用Auto Scaling功能，需對E-MapReduce授予Auto Scaling（ESS）的預設角色。在ESS的授權介面中單擊確定即可。

關閉Auto Scaling

單擊關閉Auto Scaling，當前已經通過Auto Scaling功能擴充的Task節點將會被全部釋放，HDFS儲存的資料位元於Core節點，不會受影響。

2.10.2 按時間伸縮規則配置

如果Hadoop叢集計算量在一定的周期記憶體在明顯的波峰、波穀，您可以設定在每天、每周或每月的固定時間段擴出一定量的Task節點來補充計算能力，這樣在保證作業完成的同時，可以節省您的成本。

由於Auto Scaling節點均為隨用隨付的購買方式，並且隨用隨付和訂用帳戶的同等計算能力價格大概在3:1左右，所以需要根據您的Auto Scaling時間來設計訂用帳戶計算能力和隨用隨付計算能力的比例。例如業務波峰時間段每天持續8小時，訂用帳戶和隨用隨付的價格大致相同，當大於8小時時，訂用帳戶比Auto Scaling的購買方式更優惠。

配置伸縮執行個體數

- **最大節點數：**Auto Scaling的Task節點上限。一旦達到上限，即使滿足Auto Scaling的規則，也不會繼續進行Auto Scaling的動作。目前可設定的Auto Scaling最大上限為1000。
- **最小節點數：**Auto Scaling的Task節點下限。如果Auto Scaling規則中設定的增加或減少Task節點數小於此處的最小節點數，那麼在首次執行時，叢集會以最小節點數為準進行伸縮。

例如，設定彈性擴容規則為每天零點動態添加1個節點，但最小節點數為3。那麼系統在第一天的零點時會添加3個節點，以滿足最小節點數的要求。

配置伸縮規則

伸縮規則分為擴容規則和縮容規則。叢集關閉Auto Scaling功能後，所有規則會被清空，再次開啟Auto Scaling功能時，需要重新設定伸縮規則。

- **規則名稱：**在同一個叢集中，伸縮規則名稱（包括擴容規則和縮容規則）不允許重複。

- 規則執行循環：
 - 只執行一次：叢集在指定的時間點執行一次Auto Scaling動作。
 - 重複執行：使用者可以選擇每天、每周或每月的某一特定時間點執行一次Auto Scaling動作。
- 重試到期時間：Auto Scaling在到達指定時間時可能由於各種原因不能執行，通過設定重試到期時間，系統會在該時間範圍內每隔30秒一直檢測可以執行伸縮的時機，直到在滿足條件時執行伸縮。設定範圍為0到21600秒。

假設在指定時間段需要進行Auto Scaling動作A，如果有其他Auto Scaling動作B或正處在冷卻期，則動作A無法執行。在您設定的重試到期時間內，每隔30秒會重試一次，嘗試執行A，一旦條件滿足，叢集會立刻執行Auto Scaling。

- 增加或減少Task節點數：規則被觸發時，叢集每次執行增加或減少的Task節點數量。
- 冷卻時間：每次Auto Scaling動作執行完成，到可以再次進行Auto Scaling的時間間隔。在冷卻時間內，不會發生Auto Scaling動作。

配置伸縮規格

Auto Scaling配置可以指定伸縮的節點的硬體規格。使用者只能在開啟Auto Scaling功能時配置，儲存後不能更改。如出於特殊情況確實需要修改，可以關閉Auto Scaling功能後，再次開啟。

- 選擇vCPU和記憶體規格時，系統會根據您的選擇自動匹配出滿足條件的執行個體，顯示在下面的備選執行個體列表中。您需要添加備選的執行個體到右側列表中，以便叢集按照已選的執行個體規格進行伸縮。
- 為避免由於ECS庫存不足造成的Auto Scaling失敗，您最多可以選擇3種ECS執行個體。
- 無論是選擇高效雲端硬碟還是SSD雲端硬碟，資料盤最小設定為40G。

2.10.3 Auto Scaling記錄

Auto Scaling執行完成後，您可以單擊Auto Scaling頁面上方的Auto Scaling記錄頁籤，查看Auto Scaling活動的執行記錄，以及Auto Scaling活動執行完成後的節點數量等資訊。

Auto Scaling的執行狀態包括以下4類：

- 執行中：Auto Scaling活動正在執行。
- 成功：根據伸縮規則，所有Auto Scaling中的所有節點被加入或移出叢集。
- 部分成功：根據伸縮規則，有部分節點成功被加入或移出叢集，但是受磁碟配額管理或ECS庫存的影響，部分節點執行失敗。
- 全部失敗：根據伸縮規則，沒有一個節點被加入或移出叢集。

2.11 專用網路

專用網路 (Virtual Private Cloud, VPC) 為使用者建立一個隔離的網路環境，使用者可以選擇自有的 IP 位址範圍、劃分網路、配置路由表、網關等。

詳見[專用網路產品簡介](#)。另外通過[Express Connect](#) 可以實現跨地區或跨使用者的 VPC 內網互連、VPC 與物理 IDC 機房互連。

建立專用網路叢集

E-MapReduce 在建立叢集的時候可以選擇網路類型，即傳統網路/專用網路，若選擇專用網路，需要如下額外操作：

- 所屬 VPC：選擇將當前建立的 E-MapReduce 叢集放在哪個 VPC 中，如果還沒建立可以進入[VPC 控制台](#)進行建立，一般一個帳號最多建立 2 個 VPC 網路，超過 2 個需要提工單。
- 子網（交換器）：E-MapReduce 叢集內的 ECS 執行個體通過交換器進行通訊，如果還沒建立可以進入[VPC 控制台](#)進行建立，因為交換器有可用性區域的屬性，所以在 E-MapReduce 建立叢集時選定了可用性區域後，建立的交換器也必須屬於該可用性區域。
- 建立安全性群組：如果開啟，那麼就需要輸入建立的安全性群組的名字。
- 所屬安全性群組：叢集所屬的安全性群組，傳統網路的安全性群組不能在 VPC 中使用，VPC 的安全性群組只能在當前 VPC 中使用。這裡只展示使用者在 E-MapReduce 產品中建立的安全性群組。因為一些安全的原因目前尚不支援選擇在 E-MapReduce 外建立的安全性群組。如果需要建立安全性群組，可以選擇“建立安全性群組”選項，同時輸入安全性群組的名字完成建立。

樣本

不同 VPC 中的 EMR 叢集通訊 (Hive 訪問 HBase)

1. 建立叢集。

在 E-MapReduce 上面建立兩個叢集，Hive 叢集 C1 處於 VPC1 中，HBase 叢集 C2 處於 VPC2 中，兩個叢集都在杭州地區。

2. 配置 Express Connect。

配置詳見同地區下的[VPC 私網互連](#)。

3. ssh 登入 HBase 叢集，通過 HBase Shell 建立表。

```
hbase(main):001:0> create 'testfromHbase','cf'
```

4. ssh 登入 Hive。

a. 修改 hosts，增加如下一行：

```
$zk_ip emr-cluster //$zk_ip為Hbase叢集的zk節點IP
```

b. 通過 Hive Shell 訪問 HBase。

```
hive> set hbase.zookeeper.quorum=172.16.126.111,172.16.126.112,172.16.126.113;
hive> CREATE EXTERNAL TABLE IF NOT EXISTS testfromHive (rowkey STRING, pageviews Int, bytes STRING) STORED BY 'org.apache.hadoop.hive.hbase.HBaseStorageHandler' WITH SERDEPROPERTIES ('hbase.columns.mapping' = ':key,cf:c1,cf:c2') TBLPROPERTIES ('hbase.table.name' = 'testfromHbase');
```

此時命令會卡住，然後會報 `java.net.SocketTimeoutException` 的異常，原因是 HBase 叢集的 ECS 所在的安全性群組限制了相關連接埠的訪問（E-MapReduce 建立的安全性群組預設只開放 22 連接埠）E-MapReduce，所以需要給 HBase 叢集的安全性群組增加安全性群組規則開放連接埠給 Hive 叢集，如下圖所示：

2.12 MetaService

E-MapReduce環境下提供MetaService服務。基於此服務，您可以在E-MapReduce叢集中以
免AK的方式訪問阿里雲資源。

預設應用角色

預設地，您在建立叢集時將需要向E-MapReduce服務授權一個應用角色

(`AliyunEmrEcsDefaultRole`)。授權之後，您在E-MapReduce上的作業將可以無需顯式輸入AK來訪問阿里雲資源。`AliyunEmrEcsDefaultRole`預設授予以下權限原則：

```
{
  "Version": "1",
  "Statement": [
    {
      "Action": [
        "oss:GetObject",
        "oss:ListObjects",
        "oss:PutObject",
        "oss:DeleteObject",
        "oss:ListBuckets",
        "oss:AbortMultipartUpload"
      ],
      "Resource": "*",
      "Effect": "Allow"
    }
  ]
}
```

```
    ]
  }
```

所以預設情況下，基於MetaService的作業將只能訪問OSS資料。如果您想基於MetaService訪問其他阿里雲資源，例如LogService等等，則需要給AliyunEmrEcsDefaultRole補充授予相應的許可權。以上操作需要登入[RAM控制台](#)完成。



注意：

當前metaservice服務只支援OSS，LogService和MNS資料的免AK操作。請謹慎編輯，刪除預設角色，否則會造成叢集建立失敗或者作業運行失敗。

自訂應用角色

大多數情況下，您只需要使用預設應用角色或者修改預設應用角色即可。E-MapReduce同時支援您使用自訂的應用角色。在建立叢集時，您既可以使用預設應用角色，也可以選擇自訂應用角色。如何建立角色並授權給服務，請參考[RAM](#)的相關文檔。

訪問MetaService

MetaService是一個HTTP服務，您可以直接存取這個HTTP服務來擷取相關Meta資訊：例如 “`curl http://localhost:10011/cluster-region`” 可以獲得當前叢集所在Region。

當前MetaService支援以下幾類資訊：

- Region: `/cluster-region`
- 角色名稱: `/cluster-role-name`
- AccessKeyId: `/role-access-key-id`
- AccessKeySecret: `/role-access-key-secret`
- SecurityToken: `/role-security-token`
- 網路類型: `/cluster-network-type`

使用MetaService

基於MetaService服務，我們可以在作業中免AK地訪問阿里雲資源，這樣可以帶來兩個優勢：

- 降低AK泄漏的風險。基於RAM的使用方式，可以將安全風險降到最低。需要什麼許可權就給角色授予什麼許可權，做到許可權最小化。
- 提高使用者體驗。尤其在互動式訪問OSS資源時，可以避免寫一長串的OSS路徑。

下面樣本幾種使用方式：

```
I. Hadoop命令列查看OSS資料
  舊方式: hadoop fs -ls oss://ZaH*****As1s:Ba23N*****
sdaBj2@bucket.oss-cn-hangzhou-internal.aliyuncs.com/a/b/c
  新方式: hadoop fs -ls oss://bucket/a/b/c
II. Hive建表
```

舊方式:

```
CREATE EXTERNAL TABLE test_table(id INT, name string)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY '/t'
LOCATION 'oss://ZaH*****As1s:Ba23N*****sdaBj2@bucket
.oss-cn-hangzhou-internal.aliyuncs.com/a/b/c';
```

新方式:

```
CREATE EXTERNAL TABLE test_table(id INT, name string)
ROW FORMAT DELIMITED
FIELDS TERMINATED BY '/t'
LOCATION 'oss://bucket/a/b/c';
```

III. Spark

舊方式: val data = sc.textFile("oss://ZaH*****As1s:Ba23N
*****sdaBj2@bucket.oss-cn-hangzhou-internal.aliyuncs.com/a/b/
c")

新方式: val data = sc.textFile("oss://bucket/a/b/c")