

IBM SPSS Analytic Server  
3.2.2 版

安裝與配置手冊



## 附註

在使用本資訊及它支援的產品之前，請閱讀第 69 頁的『注意事項』中的資訊。

## 產品資訊

除非新版中另有說明，否則此版本適用於 IBM® SPSS® Analytic Server 3 .2.2 版以及後續所有版次與修訂。

© Copyright International Business Machines Corporation .

# 目錄

<b>第 1 章必備項目</b> .....	<b>1</b>
<b>第 2 章 Ambari 安裝與配置</b> .....	<b>3</b>
Ambari 特定必備項目.....	3
安裝事先檢查與事後檢查工具 - Ambari.....	3
在 Ambari 上安裝.....	5
線上安裝.....	5
離線安裝.....	8
針對外部管理的 MySQL 環境，安裝 Analytic Server.....	14
容許非 root Ambari 代理程式.....	14
配置.....	15
安全.....	16
啟用對 Essentials for R 的支援.....	21
啟用關聯式資料庫來源.....	23
啟用 HCatalog 資料來源.....	25
變更 Analytic Server 所用的埠.....	26
高可用性 Analytic Server.....	26
最佳化小型資料的 JVM 選項.....	27
升級 Python - HDP.....	27
更新用戶端相依關係.....	28
配置 Apache Knox.....	28
為每一個 YARN 佇列配置個別動態資源配置 - HDP.....	30
在 Ambari 上移轉 IBM SPSS Analytic Server.....	32
解除安裝.....	33
解除安裝 Essentials for R.....	34
<b>第 3 章 Cloudera 安裝與配置</b> .....	<b>35</b>
Cloudera 概觀.....	35
Cloudera 特定必備項目.....	35
啟用了 Kerberos 的 Cloudera 環境.....	35
針對 Analytic Server 配置 MySQL.....	37
安裝事先檢查與事後檢查工具 - Cloudera.....	37
在 Cloudera 上安裝.....	39
配置 Cloudera.....	44
安全.....	45
啟用 Essentials for R 的支援.....	50
啟用關聯式資料庫來源.....	50
啟用 HCatalog 資料來源.....	52
配置 Apache Impala.....	53
變更 Analytic Server 所用的埠.....	54
高可用性 Analytic Server.....	54
升級 Python - CDH.....	55
最佳化小型資料的 JVM 選項.....	55
為每一個 YARN 資源儲存區配置個別動態資源配置 - Cloudera.....	55
移轉.....	56
在 Cloudera 上解除安裝 Analytic Server.....	57
<b>第 4 章配置 IBM SPSS Modeler 以與 IBM SPSS Analytic Server 搭配使用</b> .....	<b>59</b>
<b>第 5 章配置 UDF Hive 撤回</b> .....	<b>61</b>

<b>第 6 章使用 SLM 標籤以追蹤授權.....</b>	<b>63</b>
<b>第 7 章疑難排解.....</b>	<b>65</b>
<b>注意事項.....</b>	<b>69</b>
商標.....	70

# 第 1 章 必備項目

在安裝 Analytic Server 之前，請檢閱下列資訊。

## 系統需求

如需最新的系統需求資訊，請使用 IBM 技術支援網站上的詳細系統需求報告：<http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html>。在此頁面上：

1. 輸入 SPSS Analytic Server 作為產品名稱，然後按一下**搜尋**。
2. 選取報告的所需版本及範圍，然後按一下**提交**。

## WebSocket 資料流量

您必須確保用戶端與 Analytic Server 之間的 WebSocket 資料流量未受防火牆、VPN 或其他埠封鎖方法封鎖。WebSocket 埠與一般 Analytic Server 埠相同。

## SuSE Linux (SLES) 12

在 SuSE Linux 12 上安裝 Analytic Server 之前，請先執行下列作業：

1. 從下列 URL 下載公開金鑰至您的主機：<https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public>
2. 在主機上執行下列指令來匯入公開金鑰：

```
rpm --import IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
```

## Ubuntu 18.04

在 Ubuntu 18.04 上安裝 Analytic Server 之前，在所有叢集節點上執行下列作業：

1. 從下列 URL 將公開金鑰下載至您的主機：<https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public>
2. 在主機上執行下列指令來匯入公開金鑰：

```
apt-key add IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
```

## 電源系統

請確保 IBM XLC 和 XLF 編譯器已安裝且包含在叢集的所有主機上的 PATH 中。

您可以在下列網站尋找取得這些編譯器之授權的相關資訊：

- XL C for Linux：<http://www-03.ibm.com/software/products/en/xlcpp-linux>
- XL Fortran for Linux：<http://www-03.ibm.com/software/products/en/xlfortran-linux>

## Hortonworks 資料平台 (HDP)

安裝 Analytic Server 之前，您必須確保您的叢集環境中至少部署了一個 HDP 用戶端。由於管理 Ambari 管理程式的節點預期的是 `/usr/hdp` 目錄，因此 Analytic Server 在缺少 HDP 用戶端的情況下將會失敗。

## Hive/HCatalog

如果計劃使用 NoSQL 資料來源，則配置 Hive 和 HCatalog 以進行遠端存取。亦請確保 `hive-site.xml` 包含 `hive.metastore.uris` 內容，內容格式為 `thrift://<host_name>:<port>`（指向作用中的 Thrift Hive Metastore 伺服器）。請參閱 Hadoop 發行套件說明文件，以取得詳細資料。

在安裝 Analytic Server 期間，如果您要使用 Hive 2.1，則必須在 Ambari 主控台中透過啟用**互動式查詢**設定，然後輸入 2.x 作為 `hive.version` 內容。

1. 開啟 Ambari 主控台並在 **Analytic Server 的進階 `analytics.cfg` 區段**中新增下列內容。
  - 索引鍵：`hive.version`
  - 值：輸入適當的 Hive 版本（例如，2.x）
2. 儲存配置。

註：在具有 Spark 2.x 的 HDP 2.6 或更新版本上支援 Hive 2.1。對於 HDP 2.x，預設 hive.version 是 1.x；對於 HDP 3.x，預設 hive.version 是 3.x。

### meta 資料儲存庫

依預設，Analytic Server 會安裝並使用 MySQL 資料庫。此外，您可以將 Analytic Server 配置為使用現有 DB2 安裝。無論您選擇哪種資料庫類型，都必須具有 UTF-8 編碼。

### MySQL

MySQL 的預設字集取決於版本和作業系統。使用下列步驟確定 MySQL 的安裝是否設為 UTF-8。

1. 確定 MySQL 的版本。

```
mysql -V
```

2. 從 MySQL 指令行介面執行下列查詢來確定 MySQL 的預設字集。

```
mysql>show variables like 'char%';
```

如果字集已設為 UTF-8，則無需進一步變更。

3. 從 MySQL 指令行介面執行下列查詢來確定 MySQL 的預設對照。

```
mysql>show variables like 'coll%';
```

如果對照已設為 UTF-8，則無需進一步變更。

4. 如果預設字集或對照不是 UTF-8，請參閱 MySQL 說明文件，以取得有關如何編輯 /etc/my.cnf 並重新啟動 MySQL 常駐程式以將字集變更為 UTF-8 的詳細資料。

### DB2

如需配置 DB2 的相關資訊，請參閱 Knowledge Center [http://www-01.ibm.com/support/knowledgecenter/SSEPGG\\_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html](http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html)。

### 高可用性叢集

#### 負載平衡器

您的高可用性叢集應該具有支援階段作業親緣性（有時也稱為組合階段作業）的負載平衡器。

Analytic Server 使用 Cookie "request-token" 來識別階段作業。這會識別使用者登入期間的階段作業，以在應用程式控制的階段作業親緣性中使用。如需您的特定負載平衡器如何支援階段作業親緣性的詳細資料，請參閱它的說明文件。

#### Analytic Server 工作失敗

如果 Analytic Server 工作由於叢集成員失敗而失敗，則通常會在另一個叢集成員上自動重新啟動工作。如果工作未繼續執行，請檢查以確保「高可用性」叢集中至少具有 4 個叢集成員。

---

## 第 2 章 Ambari 安裝與配置

---

### Ambari 特定必備項目

---

除了一般必備項目之外，還請檢閱下列資訊。

#### 服務

Analytic Server 作為 Ambari 服務安裝。在安裝 Analytic Server 之前，您必須確保已安裝下列用戶端作為 Ambari 服務：

- HDFS/HDFS\_CLIENT
- MAPREDUCE2/MAPREDUCE2\_CLIENT
- HIVE/HIVE\_CLIENT
- SPARK2/SPARK2\_CLIENT（使用 Spark 2.x 時）
- HBASE/HBASE\_CLIENT（使用 HBASE 時）
- YARN
- Zookeeper

#### 無密碼的 SSH

為 Analytic Server 主機與叢集中的所有主機之間的 root 使用者設定無密碼的 SSH。

---

### 安裝事先檢查與事後檢查工具 - Ambari

---

#### 事先檢查工具概觀

Analytic Server 安裝事先檢查工具可在安裝 Analytic Server 之前識別潛在環境問題，以協助減少安裝問題與執行時期錯誤。

事先檢查工具可驗證：

- 本端系統上的作業系統與 Ambari 版本
- 本端系統上的作業系統 ulimit 設定
- 本端系統上可用的磁碟空間
- Hadoop 版本
- Ambari 服務可用性（HDFS、HCatalog、Spark、Hive、MapReduce、YARN、Zookeeper 等等）
- Analytic Server 特定的 Ambari 設定

註：事先檢查工具可在執行自行解壓縮的 Analytic Server 二進位檔之後使用。

#### 事後檢查工具概觀

Analytic Server 安裝事後檢查工具可在安裝 Analytic Server 之後，透過提交 REST API 要求進行處理來識別配置問題：

- HDFS 中的資料
- Hive/HCatalog 中的資料
- 壓縮資料（包括 deflate、bz2、snappy）
- PySpark 隨附的資料
- 使用原生 SPSS 元件（包括 alm、樹狀結構、neuralnet、評分、tascoring）的資料
- MapReduce 隨附的資料

- 記憶體內 MapReduce 隨附的資料

## 工具位置與必要條件

在安裝 Analytic Server 服務之前，請先在將會是 Analytic Server 服務的一部分的所有節點上執行事先檢查工具，以驗證您的 Linux 環境是否已準備安裝 Analytic Server。

在安裝過程中會自動呼叫事先檢查工具。在每一台主機上執行安裝之前，此工具都會檢查 Analytic Metastore 以及每一個 Analytic Server 節點。您也可以 Ambari Server 節點上手動呼叫事先檢查工具，這將會在安裝服務之前驗證機器。

執行自行解壓縮的 Analytic Server 二進位檔之後，事先檢查工具位於下列目錄：

### · HDP

```
/opt/ibm/spss/analyticserver-ambari/3.2/ANALYTICSERVER/package/chktool/precheck.py
[root@servername chktool]# cd /opt/ibm/spss/analyticserver-ambari/3.2/ANALYTICSERVER/package/chktool
[root@servername chktool]# ls
checkers data lib logs postcheck.py precheck.py readme.txt
```

安裝 Analytic Server 之後，事後檢查工具位於下列目錄中：

### · HDP

```
/opt/ibm/spss/analyticserver/3.2/tools/com.spss.ibm.checker.zip
```

工具必須以 root 身分執行且需要 Python 2.6.X（或更高版本）。

如果事先檢查工具報告了任何失敗，則必須先解決這些失敗，然後再繼續安裝 Analytic Server。

chktool 目錄在 Analytic Server 自行解壓縮的二進位檔執行（於第 5 頁的『在 Ambari 上安裝』一節中的步驟 2）之後才可用。如果您選擇執行第 8 頁的『離線安裝』，則 chktool 目錄會在安裝 meta 資料 RPM 之後可用。

## 執行事先檢查工具

### 自動

透過 Ambari 主控台將 Analytic Server 安裝為服務時，在安裝 Analytic Server 的過程中可以自動呼叫事先檢查工具。您必須手動輸入 Ambari 伺服器的使用者名稱和密碼：

Advanced analytics-env

Analytic_Server_UserID	<input type="text" value="3124"/>	+	-
ambari.user.name	<input type="text" value="admin"/>		
ambari.user.password	<input type="password" value="....."/>	<input type="password" value="....."/>	
as.database.type	<input type="text" value="mysql"/>	+	-

圖 1. 進階分析環境設定

### 手動

您可以在 Ambari Server 節點上手動呼叫事先檢查工具。

下列事先檢查範例會檢查在 myambarihost.ibm.com:8080 上執行的 Ambari 叢集 MyCluster（啟用了 SSL），並使用登入認證 admin:admin：

```
python ./precheck.py --target H --cluster MyCluster --username admin
--password admin --host myambarihost.ibm.com --port 8080 --ssl
```

### 附註：

- 需要引數 --target、--host、--port 及 --username。
- --host 值必須由 IP 位址或完整網域名稱提供。

- 省略密碼引數時，該工具會提示您輸入密碼。
- `precheck.py` 指令包括使用說明，使用 `--h` 引數來顯示該說明 (`python ./precheck.py --help`)。
- `--cluster` 是選用引數（未使用 `--cluster` 時會識別現行叢集）。

當事先檢查工具執行其檢查時，每個檢查的狀態都會顯示在指令視窗中。失敗發生時，會在日誌檔中記載詳細資訊（確切的日誌檔位置在指令視窗中提供）。需要更多支援時，可將日誌檔提供給 IBM 技術支援中心。

## 執行事後檢查工具

事後檢查工具會驗證 Analytic Server 是否適當地執行且是否能夠處理簡單的工作。下列事後檢查範例會檢查在 `myanalyticserverhost.ibm.com:9443` 上執行的 Analytic Server 實例（啟用了 SSL），並使用登入認證 `admin:ibmspss`：

```
python ./postcheck.py --target H --host myanalyticserverhost.ibm.com --port 9443
--username admin --password ibmspss --ssl
```

將 Knox 與 Analytic Server 搭配使用時，指令如下所示：

```
python ./postcheck.py --target H --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default
```

若要執行單一檢查，請使用下列指令：

```
python ./postcheck.py --target H --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check AS_PYSPARK_BUILDMODEL
```

### 附註：

- 需要引數 `--target`、`--host`、`--port` 及 `--username`。
- `--host` 值必須由 IP 位址或完整網域名稱提供。
- 省略密碼引數時，該工具會提示您輸入密碼。
- `postcheck.py` 指令包括使用說明，使用 `--h` 引數來顯示該說明 (`python ./postcheck.py --help`)。

當事後檢查工具執行其檢查時，每個檢查的狀態都會顯示在指令視窗中。失敗發生時，會在日誌檔中記載詳細資訊（確切的日誌檔位置在指令視窗中提供）。如果需要更多支援，可將日誌檔提供給 IBM 技術支援中心。

## 在 Ambari 上安裝

基本處理程序是在 Ambari 叢集內的主機上安裝 Analytic Server 檔案，然後將 Analytic Server 新增為 Ambari 服務。

### 第 5 頁的『線上安裝』

如果您的 Ambari 伺服器主機和叢集中的所有節點都可以存取 <https://ibm-open-platform.ibm.com>，請選擇線上安裝。

### 第 8 頁的『離線安裝』

如果您的 Ambari 伺服器主機沒有網際網路存取，請選擇離線。

## 線上安裝

如果您的 Ambari 伺服器主機和叢集中的所有節點都可以存取 <https://ibm-open-platform.ibm.com>，請選擇線上安裝。

1. 導覽至 [IBM Passport Advantage® 網站](https://ibm-passport-advantage.com)，並將特定於您堆疊、堆疊版本及硬體架構的自行解壓縮二進位檔，下載至 Ambari 管理程式節點中。可用的 Ambari 二進位檔如下所示：

表 1. Analytic Server 自行解壓縮的二進位檔	
說明	二進位檔名
IBM SPSS Analytic Server 3 .2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux x86-64 (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin
IBM SPSS Analytic Server 3 .2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux on System p LE (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin

- 執行自行解壓縮的二進位檔，然後遵循指示以檢視授權、接受授權、選擇線上安裝，以及選取 Analytic Server 使用之資料庫類型的安裝程序。為您提供了下列資料庫類型選項：

- 新的 MySQL 實例
- 預先存在的 MySQL 或 Db2 實例

- 從 `/var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts` 目錄中，執行帶有適當引數（例如，`--help` 引數）的 `update_clientdeps.sh` Script。

- 重新啟動 Ambari 伺服器。

```
ambari-server restart
```

- 登入 Ambari 伺服器，然後透過 Ambari 使用者介面將 Analytic Server 安裝為服務。

#### meta 資料儲存庫

Analytic Server 依預設會使用 MySQL 來追蹤資料來源、專案和承租人的相關資訊。在安裝期間，您需要提供 Analytic Server 與 MySQL 之間 JDBC 連線中所使用的使用者名稱 (**metadata.repository.user.name**) 及密碼 **metadata.repository.password**。安裝程式會在 MySQL 資料庫中建立使用者，但該使用者特定於 MySQL 資料庫，且不必是現有的 Linux 或 Hadoop 使用者。

註：如果您想要 Analytic Server 安裝程式建立新的 MySQL 實例，則必須在未安裝 MySQL 的機器上安裝 Analytic Server meta 儲存庫。

若要將 meta 資料儲存庫變更為 DB2，請遵循下列步驟。

註：安裝完成之後，您無法變更 meta 資料儲存庫。

- 確保 DB2 已安裝在另一台機器上。如需相關資訊，請參閱第 1 頁的『第 1 章 必備項目』主題的 meta 資料儲存庫區段。
- 在「Ambari 服務」標籤中，導覽至 Analytic Server 服務的「配置」標籤。
- 開啟進階 **analytics-env** 區段。
- 將 **as.database.type** 的值從 `mysql` 變更為 `db2`。
- 開啟進階 **analytics-meta** 區段。
- 將 **metadata.repository.driver** 的值從 `com.mysql.jdbc.Driver` 變更為 `com.ibm.db2.jcc.DB2Driver`。
- 將 **metadata.repository.url** 的值變更為 `jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName}`；其中
  - {Db2\_HOST} 是其中安裝 DB2 之伺服器的主機名稱
  - {PORT} 是 DB2 在其上接聽的埠
  - {SchemaName} 是可用的未用綱目。

如果您不確定要輸入哪些值，請與 DB2 管理者聯絡。

- 在 **metadata.repository.user.name** 及 **metadata.repository.password** 中提供有效的 DB2 認證。

i. 按一下儲存。

## LDAP 配置

Analytic Server 使用 LDAP 伺服器來儲存及鑑別使用者及群組。您在 Analytic Server 安裝期間需要提供必要的 LDAP 配置資訊。

LDAP 設定	說明
as.ldap.type	LDAP 類型。值可以是 ads、ad 或 openldap。 <ul style="list-style-type: none"><li>· ads - Apache Directory Server (預設值)</li><li>· ad - Microsoft Active Directory</li><li>· openldap - OpenLDAP</li></ul>
as.ldap.host	LDAP 主機
as.ldap.port	LDAP 埠號
as.ldap.binddn	LDAP 連結 DN
as.ldap.bindpassword	LDAP 連結 DN 密碼
as.ldap.basedn	LDAP 基本 DN
as.ldap.filter	LDAP 使用者和群組過濾器規則
as.ldap.ssl.enabled	指定是否使用 SSL 在 Analytic Server 與 LDAP 之間通訊。值可以是 true 或 false。
as.ldap.ssl.reference	LDAP SSL 參照 ID
as.ldap.ssl.content	LDAP SSL 配置

- 依預設，as.ldap.type 設為 ads，且其他相關設定包含預設值。例外情況是您必須提供 as.ldap.bindpassword 設定的密碼。Analytic Server 使用配置設定來安裝 Apache Directory Server (ADS) 並執行伺服器起始設定。預設 ADS 設定檔包括使用者 admin，其密碼為 admin。您可以透過 Analytic Server 主控台進行使用者管理，或者透過位於 <Analytic Root>/bin 資料夾中的 importUser.sh Script，從 XML 檔匯入使用者和群組資訊。
- 如果您計劃使用外部 LDAP 伺服器（例如，Microsoft Active Directory 或 OpenLDAP），您必須根據實際 LDAP 值定義配置設定。如需相關資訊，請參閱在 Liberty 中配置 LDAP 使用者登錄。
- 安裝 Analytic Server 之後，您可以變更 LDAP 配置（例如，從 Apache Directory Server 變更為 OpenLDAP）。但是，如果起始使用 Microsoft Active Directory 或 OpenLDAP 啟動，並決定稍後切換至 Apache Directory Server，則 Analytic Server 在安裝期間將不會安裝 Apache Directory Server。僅當起始 Analytic Server 安裝期間選取了 Apache Directory Server 時，才會安裝它。

▼ Advanced analytics-ldap

as.ldap.basedn	dc=ibm,dc=com
as.ldap.binddn	uid=admin,ou=system
as.ldap.bindpassword	.....
as.ldap.filter	<customFilters id="customFilters" userFilter="(&cn=%v)(objectClass=organizationalPerson))" groupFilter="(&cn=%v)(objectClass=groupOfNames))" useridMap="":cn" groupidMap="":cn"
as.ldap.host	{analytic_metastore_host}
as.ldap.port	10636
as.ldap.ssl.content	<ssl id="LDAPSSLSettings" keyStoreRef="LDAPTrustStore" trustStoreRef="LDAPTrustStore" /> <keyStore id="LDAPTrustStore" location="/opt/ibm/spss/analyticserver/{as_version} /ads/public/trustads.jks" type="JKS" password="changeit" />
as.ldap.ssl.enabled	true
as.ldap.ssl.reference	LDAPSSLSettings
as.ldap.type	ads

▶ Advanced analytics-log4j

圖 2. LDAP 配置設定範例

### 安裝之後不應變更的配置設定

安裝之後請勿變更下列設定，否則 Analytic Server 將無法工作。

- Analytic\_Server\_User
- Analytic\_Server\_UserID
- as.database.type
- metadata.repository.driver
- distrib.fs.root

6. 現在，您已具有能發揮作用的 Analytic Server 實例。可以選擇進行進一步配置。如需配置及管理 Analytic Server 的相關資訊，請參閱主題：第 15 頁的『[配置](#)』。如需將現有配置移轉至新安裝的相關資訊，請參閱主題：第 32 頁的『[在 Ambari 上移轉 IBM SPSS Analytic Server](#)』。
7. 開啟 Web 瀏覽器，然後輸入位址 `http://<host>:<port>/analyticserver/admin/ibm`，其中，<host> 是 Analytic Server 主機的位址，<port> 是 Analytic Server 接聽所在的埠。依預設，這是 9080。此 URL 會開啟 Analytic Server 主控台的登入對話框。以 Analytic Server 管理者的身分登入。依預設，此使用者 ID 為 admin，密碼為 admin。

## 離線安裝

IBM SPSS Analytic Server 離線安裝可自動執行，也可手動執行。

### 第 9 頁的『[在 HDP 上自動安裝](#)』

自動安裝程序使用 Ambari REST API，這是偏好的

安裝方法。

### 第 10 頁的『在 HDP (RHEL、SLES) 上手動安裝』

針對在 Hortonworks 資料平台上手動安裝 Analytic Server

### 第 12 頁的『在 HDP (Ubuntu) 上手動安裝』

針對在 Ubuntu Linux 上手動安裝 Analytic Server。

## 在 HDP 上自動安裝

自動安裝程序使用 Ambari REST API，這是偏好的安裝方法。

### 重要：

- 離線自動安裝程序會安裝內嵌的 Apache Directory Server (ADS)。如果您想要使用第三方 LDAP 伺服器，您可以在 IBM SPSS Analytic Server 安裝完成之後配置 LDAP 設定。
- 離線自動安裝程序只能安裝單一 Analytic Server 服務實例。在完成起始安裝之後，您可以新增更多實例。
- 離線自動安裝程序不支援在啟用了 Kerberos 的叢集上安裝 Analytic Server。

這些限制不適用於手動 HDP 或 Ubuntu 安裝。

1. 導覽至 [IBM Passport Advantage® 網站](https://ibm-open-platform.ibm.com) 並將自行解壓縮的二進位檔下載到可存取 <https://ibm-open-platform.ibm.com> 的電腦。

表 3. Analytic Server 自行解壓縮的二進位檔	
說明	二進位檔名
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux x86-64 (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-lx86.bin
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux on System p LE (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-lppc64.bin

2. 執行您在步驟 1 中下載的二進位執行檔並指定離線安裝。離線安裝會下載稍後在安裝程序中必要的 RPM 或 DEB 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的電腦上。下載的檔案位於現行二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 中。
3. 將二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 的全部內容從具備網際網路存取權的機器複製到 Ambari 管理程式節點 (受防火牆保護)。
4. 在 Ambari 管理程式節點上，使用下列指令來檢查 Ambari 伺服器是否在執行中：

```
ambari-server status
```

5. 在 Ambari 管理程式節點上，以及在您要在其中部署 Analytic Server 的所有其他節點上，安裝用於建立本端 yum 儲存庫的工具。

```
yum install createrepo (RHEL、CentOS)
```

或

```
apt-get install dpkg-dev (Ubuntu)
```

6. 在 Ambari 管理程式節點上，執行二進位執行檔 `./IBM-SPSS-AnalyticServer/packages/spss_as-ambari-offlineinstall.bin`。在安裝期間，二進位執行檔會驗證套件目錄中是否有必要的 Analytic Server RPM/DEB 檔案。您需要的 RPM 檔案取決於您的發行套件、版本與架構。

#### HDP 2.6、3.0 及 3.1 (x86\_64)

```
IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm
```

```
IBM-SPSS-AnalyticServer-3.2.2.0-1.x86_64.rpm
```

#### HDP 2.6、3.0 及 3.1 (PPC64LE)

```
IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm
```

IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

### HDP 2.6、3.0 及 3.1 (Ubuntu)

IBM-SPSS-AnalyticServer-ambari-2.x\_3.2.2.0\_amd64.deb

IBM-SPSS-AnalyticServer\_1\_amd64.deb

在安裝期間，系統會提示您輸入 Analytic Server 版本、JDBC 驅動程式、Spark 版本、Hive 版本等。

### 在 HDP (RHEL、SLES) 上手動安裝

在 HDP (RHEL、SLES) 上的一般手動離線安裝工作流程如下所示：

1. 導覽至 [IBM Passport Advantage® 網站](https://ibm-open-platform.ibm.com) 並將自行解壓縮的二進位檔下載到可存取 <https://ibm-open-platform.ibm.com> 的電腦。

說明	二進位檔名
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux x86-64 (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-1x86.bin
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux on System p LE (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-1ppc64.bin

2. 執行您在步驟 1 中下載的二進位執行檔並指定離線安裝。離線安裝會下載稍後在安裝程序中必要的 RPM 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的電腦上。下載的檔案位於現行二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 中。
3. 將二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 的全部內容從具備網際網路存取權的機器複製到 Ambari 管理程式節點的 `<AS_INSTALLABLE_HOME>` 目錄 (Ambari 管理程式節點受防火牆保護)。
4. 在 Ambari 管理程式節點上，使用下列指令來檢查 Ambari 伺服器是否在執行中：

```
ambari-server status
```

5. 安裝工具來建立本端 yum 儲存庫。

```
yum install createrepo (RHEL、CentOS)
```

或

```
zypper install createrepo (SLES)
```

6. 建立目錄以充當用來存放 Analytic Server RPM 檔的儲存庫。請參閱下列範例。

```
mkdir -p /home/root/repos/IBM-SPSS-AnalyticServer/x86_64
```

7. 將必要的 Analytic Server RPM 檔複製到新目錄。您需要的 RPM 檔案取決於您的發行套件、版本與架構。

### HDP 2.6、3.0 及 3.1 (x86\_64)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.x86\_64.rpm

### HDP 2.6、3.0 及 3.1 (PPC64LE)

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

8. 建立本端儲存庫定義。例如，使用下列內容建立名為 IBM-SPSS-AnalyticServer-3.2.2.0.repo 的檔案，建立位置是：/etc/yum.repos.d/（對於 RHEL、CentOS）或 /etc/zypp/repos.d/（對於 SLES）。

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{{本端儲存庫的路徑}}
enabled=1
gpgcheck=0
protect=1
```

9. 建立本端 yum 儲存庫。

```
createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64 ( RHEL、CentOS、SLES )
```

10. 從 root 使用者指令視窗中，cd 至 <AS\_INSTALLABLE\_HOME>/IBM-SPSS-AnalyticServer 目錄，並執行 ./offLineInstall.sh。Script 會讀取前次執行的二進位執行檔安裝指令的持續回應，並發出適當的平台指令（以安裝 RPM）。

註：僅當您使用外部管理的 MySQL 環境時才套用步驟 11。

11. 在安裝 MySQL 實例（將用來作為 AS\_MetaStore）的節點/主機上，執行 add\_mysql\_user.sh Script。

- a. 將 add\_mysql\_user.sh Script 從 <AS\_INSTALLABLE\_HOME>/IBM-SPSS-AnalyticServer 複製到 MySQL 實例安裝所在的節點/主機（將用作 AS\_MetaStore）。

· 在 MySQL 節點/主機上執行 add\_mysql\_user.sh Script。例如，./add\_mysql\_user.sh -u as\_user -p spss -d aedb。

#### 附註：

- 使用者名稱和密碼必須符合在 Ambari 配置畫面上為 AS\_Metastore 輸入的資料庫使用者名稱和密碼。
- 可以手動更新 add\_mysql\_user.sh Script，以發出指令（如有需要）。
- 當針對安全的（root 使用者存取）MySQL 資料庫執行 add\_mysql\_user.sh Script 時，使用 -r 和 -t 參數來傳入 dbuserid 和 dbuserid\_password。Script 使用 dbuserid 和 dbuserid\_password 來執行 MySQL 作業。

註：必須修改 **AS\_Configuration** 畫面上的 metadata.repository.url 設定（進階 **analytics-meta**），以指向 MySQL 資料庫主機。例如，將 JDBC 設定 mysql:// {analytic\_metastore\_host}/aedb?createDatabaseIfNotExist=true 變更為 mysql:// {MySQL\_DB}/aedb?createDatabaseIfNotExist=true

12. 更新您的 Ambari 儲存庫檔案 repoinfo.xml（一般位於 /var/lib/ambari-server/resources/stacks/\$stackName/\$stackVersion/repos/ 中），新增下列各行以使用本端 yum 儲存庫。

```
<os type="host_os">
  <repo>
    <baseurl>file:///{{path to local repository}}/</baseurl>
    <repoId>IBM-SPSS-AnalyticServer</repoId>
    <reponame>IBM-SPSS-AnalyticServer-3.2.2.0</reponame>
  </repo>
</os>
```

範例 {path to local repository} 類似於下列項目：

```
/home/root/repos/IBM-SPSS-AnalyticServer/x86_64/
```

13. 針對每一個 Ambari 非伺服器叢集節點，重複下列步驟。

- a. 將適當的 <AS\_INSTALLABLE\_HOME> 目錄的全部內容從具有網際網路存取權的機器複製到 Ambari 非伺服器叢集節點。

- b. 安裝工具來建立本端 yum 儲存庫。

```
yum install createrepo ( RHEL、CentOS )
```

或

```
zypper install createrepo ( SLES )
```

- c. 建立目錄以充當用來存放 Analytic Server RPM 檔的儲存庫。請參閱下列範例。

```
mkdir -p /home/root/repos/IBM-SPSS-AnalyticServer/x86_64
```

- d. 將必要的 Analytic Server RPM 檔複製到新目錄。您需要的 RPM 檔案取決於您的發行套件、版本與架構。

**HDP 2.6、3.0 及 3.1 (x86\_64)**

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.x86\_64.rpm

**HDP 2.6、3.0 及 3.1 (PPC64LE)**

IBM-SPSS-AnalyticServer-ambari-2.x-3.2.2.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.2.2.0-1.ppc64le.rpm

- e. 建立本端儲存庫定義。例如，使用下列內容建立名為 IBM-SPSS-AnalyticServer-3.2.2.0.repo 的檔案，建立位置是：/etc/yum.repos.d/（對於 RHEL、CentOS）或 /etc/zypp/repos.d/（對於 SLES）。

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer
baseurl=file:///{{本端儲存庫的路徑}}
enabled=1
gpgcheck=0
protect=1
```

- f. 建立本端 yum 儲存庫。

```
createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64 ( RHEL、CentOS、SLES )
```

14. 繼續執行第 5 頁的『線上安裝』小節中的步驟 3。

### 在 HDP (Ubuntu) 上手動安裝

在 HDP (Ubuntu) 上的一般手動離線安裝工作流程如下所示：

1. 導覽至 [IBM Passport Advantage® 網站](https://ibm-open-platform.ibm.com) 並將適當的 Ubuntu 自行解壓縮的二進位檔下載到可存取 <https://ibm-open-platform.ibm.com> 的電腦。

說明	二進位檔名
IBM SPSS Analytic Server 3.2.2 for Hortonworks Data Platform 2.6、3.0 及 3.1 Linux x86-64 (英文版)	spss_as-3.2.2.0-hdp2.6-3.1-1x86.bin

2. 執行您在步驟 1 中下載的二進位執行檔並指定離線安裝。離線安裝會下載稍後在安裝程序中必要的 DEB 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的電腦上。下載的檔案位於現行二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 中。
3. 將二進位執行檔目錄 `./IBM-SPSS-AnalyticServer` 的全部內容從具備網際網路存取權的機器複製到 Ambari 管理程式節點的 `<AS_INSTALLABLE_HOME>` 目錄（Ambari 管理程式節點受防火牆保護）。
4. 在 Ambari 管理程式節點上，使用下列指令來檢查 Ambari 伺服器是否在執行中：

```
ambari-server status
```

5. 建立 <local\_repo> 目錄以充當用來存放 Analytic Server DEB 檔案的儲存庫。例如：

```
mkdir -p /usr/local/mydebs
```

6. 將必要的 Analytic Server DEB 檔案複製到 <local\_repo> 目錄。

- IBM-SPSS-AnalyticServer-ambari-2.x\_3.2.2.0\_amd64.deb
- IBM-SPSS-AnalyticServer\_1\_amd64.deb

7. 建立本端儲存庫。

- a. 安裝工具來建立本端儲存庫：

```
apt-get install dpkg-dev
```

- b. 產生原始套件檔：

```
cd <local_repo>
dpkg-scanpackages . /dev/null | gzip -9c > Packages.gz
```

- c. 建立本端儲存庫的元件（主要）及架構（例如，binary-i386、binary-amd64）：

```
mkdir -p <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/
mkdir -p <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/
```

- d. 複製原始套件：

```
cp -fr <local_repo>/Packages.gz <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages
cp -fr <local_repo>/Packages.gz <local_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages
```

8. 建立本端儲存庫定義。例如，在 /etc/apt/sources.list.d 中建立包含下列內容的 IBM-SPSS-AnalyticServer-3.2.2.0.list 檔案。

```
deb file:/usr/local/mydebs ./
```

**重要：**在 Ubuntu 18.04 上，使用：deb [trusted=yes] file:/usr/local/mydebs ./

9. 執行下列指令以更新儲存庫清單：

```
apt-get update
```

10. 執行下列指令，安裝 IBM SPSS Analytic Server 3.2.2：

```
apt-get install IBM-SPSS-AnalyticServer-ambari-2.x
```

**註：**若要驗證本端儲存庫是否正確設定，請不要在 <local\_repo> 目錄中執行前一個指令。如果安裝找不到該套件，則表示本端儲存庫未正確設定（在此情況下，您必須驗證所有先前的步驟）。

11. 針對每一個 Ambari 非伺服器叢集節點，重複下列步驟。

- a. 建立 <local\_repo> 目錄以充當用來存放 Analytic Server DEB 檔案的儲存庫。例如：

```
mkdir -p /usr/local/mydebs
```

- b. 將 <local\_repo> 目錄的全部內容從 Ambari 管理程式節點複製到 Ambari 非伺服器叢集節點的 <local\_repo> 目錄。該目錄應包含下列檔案：

- <local\_repo>/IBM-SPSS-AnalyticServer-ambari-2.x\_3.2.2.0\_amd64.deb
- <local\_repo>/IBM-SPSS-AnalyticServer\_1\_amd64.deb
- <local\_repo>/Packages.gz
- <local\_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-amd64/Packages
- <local\_repo>/dists/IBM-SPSS-AnalyticServer/main/binary-i386/Packages

- c. 建立本端儲存庫定義。例如，在 /etc/apt/sources.list.d 中建立包含下列內容的 IBM-SPSS-AnalyticServer-3.2.2.0.list 檔案。

```
deb file:/usr/local/mydebs ./
```

- 重要：**在 Ubuntu 18.04 上，使用：`deb [trusted=yes] file:/usr/local/mydebs ./`
- 繼續執行第 5 頁的『線上安裝』小節中的步驟 3。

## 針對外部管理的 MySQL 環境，安裝 Analytic Server

Analytic Server 安裝程序在針對外部管理的 MySQL 環境安裝時與一般安裝不同。

下列步驟說明了針對外部管理的 MySQL 環境安裝 Analytic Server 的過程。

- 導覽至 [IBMPassport Advantage®](#) 網站，並將特定於您堆疊、堆疊版本及硬體架構的自行解壓縮二進位檔，下載至 Ambari 叢集內的主機上。
- 執行自行解壓縮的二進位檔，然後遵循指示以（選擇性地）檢視授權、接受授權。
  - 選擇線上選項。
  - 出現提示時，請選取**外部 MySQL 資料庫**選項。
- 將 `add_mysql_user.sh` Script 從 `<AS_INSTALLABLE_HOME>/IBM-SPSS-AnalyticServer` 複製到 MySQL 實例安裝所在的節點/主機（將用作 `AS_MetaStore`）。
  - 在 MySQL 節點/主機上執行 `add_mysql_user.sh` Script。例如，`./add_mysql_user.sh -u as_user -p spss -d aedb`。

### 附註：

- 使用者名稱和密碼必須符合在 Ambari 配置畫面上為 `AS_Metastore` 輸入的資料庫使用者名稱和密碼。
  - 可以手動更新 `add_mysql_user.sh` Script，以發出指令（如有需要）。
  - 當針對安全的（root 使用者存取）MySQL 資料庫執行 `add_mysql_user.sh` Script 時，使用 `-r` 和 `-t` 參數來傳入 `dbuserid` 和 `dbuserid_password`。Script 使用 `dbuserid` 和 `dbuserid_password` 來執行 MySQL 作業。
- 重新啟動 Ambari 伺服器。

```
ambari-server restart
```

- 從 Ambari 主控台，如一般作業新增 AnalyticServer 服務（輸入與第 3 步中輸入相同的資料庫使用者名稱和密碼）。

**註：**必須修改 **AS\_Configuration** 畫面上的 `metadata.repository.url` 設定（**進階 analytics-meta**），以指向 MySQL 資料庫主機。例如，將 JDBC 設定 `mysql:// {analytic_metastore_host}/aedb?createDatabaseIfNotExist=true` 變更為 `mysql://{MySQL_DB}/aedb?createDatabaseIfNotExist=true`

## 容許非 root Ambari 代理程式

以非 root 使用者身分執行 Ambari Server 及 Ambari Agent 時，Analytic Server 安裝處理程序與一般安裝不同。

### 必備項目

以 root 身分登入，在叢集中的每個主機上新增非 root 使用者，以及適用 `sudo` 存取權設定非 root 使用者。下列範例會新增非 root 使用者 `asuser` 至 `suoders` 檔（預設檔案位置是 `/etc/sudoers`）：

```
## Allow root to run any commands anywhere
asuser ALL=(ALL) ALL

## Allow root to run any commands anywhere without a password
asuser ALL=(ALL) NOPASSWD: ALL
```

如需詳細安裝資訊，請參閱第 5 頁的『線上安裝』或第 8 頁的『離線安裝』區段。

## sudo 需求

您必須在以非 root 使用者身分執行的所有指令的指令文字之前新增 sudo。

## Ambari 非 root 問題

透過 Ambari 使用者介面將 Analytic Server 新增為服務時，您可能會看到下列錯誤：

```
Error: 500 status code received on POST method for API: /api/v1/stacks/HDP/versions/2.6/recommendations
```

錯誤是 Ambari 非 root 問題的結果。您可以將資料夾 /var/run/ambari-server 的擁有者變更為非 root 使用者，然後透過 Ambari 使用者介面新增 Analytic Server 作為服務。下列範例示範將資料夾 /var/run/ambari-server 的擁有者變更為非 root 使用者 asuser 的處理程序。

```
sudo chown asuser:asuser /var/run/ambari-server/
```

## 無密碼 ssh

未設定無密碼 ssh 時，安裝期間顯示下列警告：

```
UserWarning: Failure to add as_user. This must be done manually on each node.  
warnings.warn("Failure to add as_user. This must be done manually on each node.")
```

您必須在每一個節點上手動建立 as\_user。as\_user 是具有 Analytic Server 安裝權限的使用者帳戶。例如：

```
# Create the as_user user (whatever id possible first) and note the id for use on subsequent  
nodes  
sudo useradd as_user  
  
# set the user for nologin  
sudo usermod -s /sbin/nologin as_user  
  
# Mod to as_user user id  
sudo usermod -u {as_user_id} as_user  
  
# Make primary group user_group  
sudo usermod -g hadoop as_user  
  
# Make extends group hdfs  
sudo usermod -G hdfs as_user
```

註：可以透過 id as\_user 指令在 Ambari 主要節點上找到 {as\_user\_id}。

## 配置

安裝之後，您必須在叢集作業系統上建立必要帳戶。

- 為所有使用者建立作業系統使用者帳戶，您計劃提供對每一個及每個 Analytic Server 及 Hadoop 節點上 Analytic Server 的存取權（這些使用者也配置為 LDAP 使用者登錄）。使用者群組必須設為 hadoop。
  - 請確保這些使用者的 UID 在所有機器上都相符。您可以使用 **kinit** 指令登入每一個帳戶來測試此項目。
  - 請確保 UID 遵循用於提交工作的使用者 ID 下限 YARN 設定。這是 container-executor.cfg 中的 **min.user.id** 參數。例如，如果 **min.user.id** 為 1000，則建立的每個使用者帳戶都必須具有大於或等於 1000 的 UID。
- 針對 Analytic Server 管理者使用者，在 HDFS 上建立使用者起始檔案夾。該資料夾許可權必須設為 755，擁有者必須定義為 admin，且使用者群組必須設為 hdfs。請參閱下列**粗體**範例：

```
[root@xxxxx configuration]# hadoop fs -ls /user  
drwxr-xr-x - admin hdfs 0 2017-06-08 01:33 /user/admin  
drwxr-xr-x - user1 hdfs 0 2017-06-06 01:00 /user/user1
```

3. 在 HDFS 上為所有 Analytic Server 標準使用者（例如，user1）建立使用者起始檔案夾。資料夾擁有者是實際使用者，並且使用者群組必須設為 hdfs。

安裝之後，您可以選擇性地透過 Ambari 使用者介面來配置及管理 Analytic Server。

註：下列使用慣例用於 Analytic Server 檔案路徑。

- {AS\_ROOT} 指的是 Analytic Server 的部署位置；例如，/opt/ibm/spss/analyticserver/3.2。
- {AS\_SERVER\_ROOT} 指的是配置檔、日誌檔及伺服器檔案的配置；例如，/opt/ibm/spss/analyticserver/3.2/ae\_wlpserver/usr/servers/aeserver。
- {AS\_HOME} 指的是 HDFS 上由 Analytic Server 用作根資料夾的位置；例如，/user/as\_user/analytic-root。

## 安全

### 配置 LDAP 登錄

LDAP 登錄可讓您向 Active Directory 或 OpenLDAP 之類的外部 LDAP 伺服器鑑別使用者。

**重要：**必須指定一個 LDAP 使用者作為 Analytic Server Ambari 中的管理者。

以下是 OpenLDAP 的 ldapRegistry 範例。

```
<ldapRegistry
  baseDN="ou=people,dc=aeldap,dc=org"
  ldapType="Custom"
  port="389"
  host="server"
  id="OpenLDAP"
  bindDN="cn=admin,dc=aeldap,dc=org"
  bindPassword="{xor}Dz4sLG5tbGs="
  searchTimeout="300000m"
  recursiveSearch="true">
  <customFilters
    id="customFilters"
    userFilter="( &(uid=%v)(objectClass=inetOrgPerson))"
    groupFilter="( &(cn=%v)(|(objectclass=organizationalUnit)))"
    groupMemberIdMap="posixGroup:memberUid"/>
</ldapRegistry>
```

下面的範例使用 Active Directory 提供 Analytic Server 鑑別：

```
<ldapRegistry id="Microsoft Active Directory" realm="ibm"
  host="host"
  port="389"
  baseDN="cn=users,dc=adtest,dc=mycompany,dc=com"
  bindDN="cn=administrator,cn=users,dc=adtest,dc=mycompany,dc=com"
  bindPassword="adminpassword"
  ldapType="Custom"
  <customFilters
    userFilter="( &(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="( &(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
    groupIdMap="*:cn"
    groupMemberIdMap="memberOf:member" />
</ldapRegistry>
```

註：使用 LDAP 檢視器第三方工具來驗證 LDAP 配置十分有用。

下面的範例使用 Active Directory 提供 WebSphere Liberty 設定檔鑑別：

```
<ldapRegistry id="ldap" realm="SampleLdapADRealm"
  host="ldapserver.mycity.mycompany.com" port="389" ignoreCase="true"
  baseDN="cn=users,dc=adtest,dc=mycity,dc=mycompany,dc=com"
  bindDN="cn=testuser,cn=users,dc=adtest,dc=mycity,dc=mycompany,dc=com"
  bindPassword="testuserpwd"
  ldapType="Microsoft Active Directory"
  sslEnabled="true"
  sslRef="LDAPSSLSettings">
  <activedFilters
    userFilter="( &(sAMAccountName=%v)(objectcategory=user))"
    groupFilter="( &(cn=%v)(objectcategory=group))"
    userIdMap="user:sAMAccountName"
```

```

        groupIdMap="*:cn"
        groupMemberIdMap="memberOf:member" >
    </activatedFilters>
</ldapRegistry>

<ssl id="LDAPSSLSettings" keyStoreRef="LDAPKeyStore" trustStoreRef="LDAPTrustStore" />

<keyStore id="LDAPKeyStore" location="{server.config.dir}/LdapSSLKeyStore.jks"
    type="JKS" password="{xor}CDo9HgW=" />

<keyStore id="LDAPTrustStore" location="{server.config.dir}/LdapSSLTrustStore.jks"
    type="JKS" password="{xor}CDo9HgW=" />

```

#### 附註：

- 在 Analytic Server 中的 LDAP 支援由 WebSphere Liberty 進行控制。如需相關資訊，請參閱在 [Liberty](#) 中配置 LDAP 使用者登錄。
- 使用 SSL 確保 LDAP 安全時，請遵循下列「配置從 Analytic Server 到 LDAP 的 Secure Sockets Layer (SSL) 連線」一節中的指示。

### 配置從 Analytic Server 到 LDAP 的 Secure Socket Layer (SSL) 連線

如果您在 Analytic Server 安裝期間選取 Apache Directory Server (ads) LDAP 選項，並使用預設配置，則在安裝 Apache Directory Server 時會配置並啟用 SSL（Analytic Server 將自動使用 SSL 與 Apache Directory Server 通訊）。

如果在 Analytic Server 安裝期間選取了其他某個 LDAP 選項（例如，使用外部 LDAP 伺服器），請使用下列步驟配置 SSL。

1. 以 Analytic Server 使用者的身分登入每一個 Analytic Server 機器，並為 SSL 憑證建立共用目錄。
 

註：依預設，as\_user 是 Analytic Server 使用者；請參閱 Ambari 主控台中「管理」標籤下的服務帳戶。
2. 將金鑰儲存庫和信任儲存庫檔案複製到所有 Analytic Server 機器上的部分共用目錄。另將 LDAP 用戶端 CA 憑證新增至信任儲存庫。以下是部分範例指示。

```

mkdir -p /home/as_user/security
cd /home/as_user/security
openssl s_client -connect <ldap-hostname>:636 -showcerts > client.cert
$JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore
mytrust.jks
password : changeit

```

註：JAVA\_HOME 是與用於 Analytic Server 啟動相同的 JRE。

3. 可以使用 securityUtility 工具來編碼密碼以模糊化其值，該工具位於 {AS\_ROOT}/ae\_wlpserver/bin 中。範例如下所示。

```

securityUtility encode changeit
{xor}PDC+MTg6Nis=

```

4. 登入 Ambari 主控台，並將 Analytic Server 配置設定 **ssl.keystore.config** 更新為正確的 SSL 配置設定。範例如下所示。

```

<ssl id="defaultSSLConfig" keyStoreRef="defaultKeyStore" trustStoreRef="defaultTrustStore"
    clientAuthenticationSupported="true"/>
    <keyStore id="defaultKeyStore" location="/home/as_user/security/mykey.jks"
type="JKS"
        password="{xor}0zo5PiozKxYdEgwPDaWeDG1uDz4sLCg7"/>
    <keyStore id="defaultTrustStore" location="/home/as_user/security/mytrust.jks"
type="JKS"
        password="{xor}PDC+MTg6Nis="/>

```

註：對金鑰和信任儲存庫檔案使用絕對路徑。

5. 將 Analytic Server 配置設定 **security.config** 更新為正確的 LDAP 配置設定。例如，在 **ldapRegistry** 元素中，將 **sslEnabled** 屬性設為 true，並將 **sslRef** 屬性設為 defaultSSLConfig。

## 配置 Kerberos

Analytic Server 支援使用 Ambari 的 Kerberos。

註：IBM SPSS Analytic Server 與 Apache Knox 一起使用時，不支援 Kerberos 單一登入 (SSO)。

1. 在 Kerberos 使用者儲存庫中，為您打算授與 Analytic Server 之存取權的所有使用者建立帳戶。
2. 在 LDAP 伺服器上建立相同的帳戶（依前一個步驟）。
3. 在每個 Analytic Server 節點及 Hadoop 節點上，為在前一個步驟中所建立的每一位使用者，各建立一個作業系統使用者帳戶。使用者群組必須設為 `hadoop`。
  - 請確保這些使用者的 UID 在所有機器上都相符。您可以使用 `kinit` 指令登入每一個帳戶來測試此項目。
  - 請確保 UID 遵循用於提交工作的使用者 ID 下限 YARN 設定。這是 `container-executor.cfg` 中的 `min.user.id` 參數。例如，如果 `min.user.id` 為 1000，則建立的每個使用者帳戶都必須具有大於或等於 1000 的 UID。
4. 針對 Analytic Server 管理者使用者，在 HDFS 上建立使用者起始檔案夾。該資料夾許可權必須設為 755，擁有者必須定義為 `admin`，且使用者群組必須設為 `hdfs`。請參閱下列粗體範例：

```
[root@xxxxx configuration]# hadoop fs -ls /user
Found 9 items
drwxrwxrwx - hdfs supergroup 0 2017-07-26 03:41 /user/AE
drwxr-xr-x - admin hdfs 0 2017-06-08 01:33 /user/admin
drwxr-xr-x - as_user hdfs 0 2017-06-06 01:00 /user/as_user
drwx----- - hdfs supergroup 0 2017-07-31 00:17 /user/hdfs
drwxrwxrwx - mapred hadoop 0 2017-06-05 00:28 /user/history
drwxrwxr-t - hive hive 0 2017-06-05 00:30 /user/hive
drwxrwxr-x - hue hue 0 2017-06-05 00:30 /user/hue
drwxrwxr-x - impala impala 0 2017-07-19 00:52 /user/impala
drwxr-x--x - spark spark 0 2017-06-05 01:34 /user/spark
```

5. 在 HDFS 上為所有 Analytic Server 標準使用者（例如，`user1`）建立使用者起始檔案夾。資料夾擁有者是實際使用者，並且使用者群組必須設為 `hdfs`。
6. [選用] 如果您打算使用 HCatalog 資料來源，且 Analytic Server 安裝在與 Hive Metastore 不同的機器上，則必須在 HDFS 上模擬 Hive 用戶端。
  - a. 在 Ambari 主控台中導覽至 HDFS 服務的「配置」標籤。
  - b. 編輯 `hadoop.proxyuser.hive.groups` 參數，讓它的值為 `*`，或是包含獲允許登入 Analytic Server 之所有使用者的群組。
  - c. 編輯 `hadoop.proxyuser.hive.hosts` 參數，讓它的值為 `*`，或是 Hive Metastore 及每個 Analytic Server 實例在其上被安裝為服務的主機清單。
  - d. 重新啟動 HDFS 服務。

在執行了這些步驟並安裝了 Analytic Server 之後，Analytic Server 就會無聲自動配置 Kerberos。

## 使用 Kerberos 為單一登入 (SSO) 配置 HAProxy

1. 按 HAProxy 說明文件手冊配置及啟動 HAProxy：<http://www.haproxy.org/#docs>
2. 為 HAProxy 主機建立 Kerberos 原則 (`HTTP/<proxyHostname>@<realm>`) 及 `keytab` 檔，其中 `<proxyHostname>` 是 HAProxy 主機的完整名稱，`<realm>` 是 Kerberos 領域。
3. 將 `keytab` 檔複製到每個 Analytic Server 主機作為 `/etc/security/keytabs/spnego_proxy.service.keytab`
4. 在每個 Analytic Server 主機上更新此檔案的許可權。範例如下所示。

```
chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
```

5. 開啟 Amabri 主控台，並更新 Analytic Server 「自訂 `analytics.cfg`」區段中的下列內容。

```
web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
web.authentication.kerberos.principal=HTTP/<proxy machine full name>@<realm>
```

6. 儲存配置，並從 Amabri 主控台重新啟動所有 Analytic Server 服務。

使用者現在可以使用 IBM SPSS Analytic Server 登入畫面上的單一登入選項來登入 Analytic Server。

## 啟用 Kerberos 模擬

模擬容許執行緒在與執行緒所屬程序之安全環境定義不同的安全環境定義中執行。例如，模擬可提供一種方式，讓 Hadoop 工作以標準 Analytic Server 使用者 (as\_user) 以外的使用者身分執行。若要啟用 Kerberos 模擬，請執行下列動作：

1. 在啟用了 Kerberos 的叢集中執行時，將模擬配置屬性新增至 HDFS (或 Hive 服務配置)。若為 HDFS，則下列內容必須新增至 HDFS core-site.xml 檔：

```
hadoop.proxyuser.<analytic_server_service_principal_name>.hosts = *
hadoop.proxyuser.<analytic_server_service_principal_name>.groups = *
```

其中 <analytic\_server\_service\_principal\_name> 是在 Analytic Server 配置的 Analytic\_Server\_User 欄位中指定的預設 as\_user 值。

如果是透過 Hive/HCatalog 從 HDFS 存取資料，則下列內容也必須新增至 HDFS core-site.xml 檔：

```
hadoop.proxyuser.hive.hosts = *
hadoop.proxyuser.hive.groups = *
```

2. 如果 Analytic Server 配置成使用 as\_user 以外的使用者名稱，則必須修改內容名稱以反映其他使用者名稱 (例如，hadoop.proxyuser.xxxxx.hosts，其中 xxxxx 是在 Analytic Server 配置中指定的已配置使用者名稱)。
3. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

## 啟用多個領域

定義多個領域時，需要 **as.kdc.realms** 設定。**as.kdc.realms** 值位於 Ambari 主控台的 Analytic Server 「進階 analytics.cfg」區段中。

Property Name	Value	Status	Action
admin.username	admin	OK	Refresh
as.kdc.realms	IBM.COM,SPSS.COM	OK	Refresh
distrib.fs.root	/user/{as_user}/analytic-root	OK	Refresh
hive.storagehandlers.location	/usr/share/hive	OK	Refresh
hive.version	1.x	OK	Refresh
http.port	9080	OK	Refresh
https.port	9443	OK	Refresh
jdbc.drivers.location	/usr/share/jdbc	OK	Refresh
resource.pool.enabled	false	OK	Refresh
spark.version	2.x	OK	Refresh
ssl.as.enable	false	OK	Refresh
ssl.keystore.config	None	OK	Refresh

圖 3. 進階 analytics.cfg 設定

支援多個領域名稱 (以逗點字元區隔)。指定的 Kerberos 領域名稱對應於使用者名稱並與之相關聯。例如，使用者名稱 UserOne@us.ibm.com 及 UserTwo@eu.ibm.com 對應於領域 us.ibm.com, eu.ibm.com。

在 **Kerberos 領域名稱** 中指定多個領域時，必須配置 Kerberos 交叉領域信任。在 Analytic Server 主控台登入提示期間輸入的使用者名稱是沒有領域名稱字尾的。因此，指定多個領域時，系統會向使用者呈現一個領域下拉清單，以讓使用者選取領域。

註：如果僅指定一個領域，則在使用者登入 Analytic Server 時，不會向使用者呈現領域下拉清單。

## 停用 Kerberos

1. 在 Ambari 主控台中停用 Kerberos。
2. 停止 Analytic Server 服務。
3. 按一下儲存，然後重新啟動 Analytic Server 服務。

## 啟用與 Analytic Server 主控台的 Secure Socket Layer (SSL) 連線

依預設，Analytic Server 會產生自簽憑證以啟用 Secure Socket Layer (SSL)，以便您可以透過安全埠，藉由接受自簽憑證來存取 Analytic Server 主控台。為了能更安全地進行 HTTPS 存取，您需要安裝第三方供應商憑證。

## 安裝協力廠商供應商憑證

1. 將第三方供應商金鑰儲存庫和信任儲存庫憑證複製到所有 Analytic Server 節點中的相同目錄；例如，`/home/as_user/security`。  
註：Analytic Server 使用者必須擁有此目錄的讀取權。
2. 在「Ambari 服務」標籤中，導覽至 Analytic Server 服務的「配置」標籤。
3. 編輯 `ssl.keystore.config` 參數。

```
<ssl id="defaultSSLConfig"
  keyStoreRef="defaultKeyStore"
  trustStoreRef="defaultTrustStore"
  clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
  location="<KEYSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
<keyStore id="defaultTrustStore"
  location="<TRUSTSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
```

將

- `<KEYSTORE-LOCATION>` 取代為金鑰儲存庫的絕對位置；例如：`/home/as_user/security/mykey.jks`
- `<TRUSTSTORE-LOCATION>` 取代為信任儲存庫的絕對位置；例如：`/home/as_user/security/mytrust.jks`
- `<TYPE>` 取代為憑證的類型；例如：JKS、PKCS12 等。
- `<PASSWORD>` 取代為採用 Base64 加密格式的加密密碼。如果要進行編碼，您可以使用 `securityUtility`；例如：`/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility encode <password>`

如果要產生自簽憑證，可以使用 `securityUtility`；例如：`/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,0=myorg,C=mycountry`。

附註：

- 您必須提供 CN 值的適當主機網域名稱。
- 將 `mypassword`、`myfqdnserver`、`myorg` 及 `mycountry` 取代為您的特定認證。請注意，`myfqdnserver` 是 Analytic Server 節點的完整網域名稱。

- **aeserver** 是 Liberty 伺服器的名稱（該值必須是 **aeserver**）。

如需 securityUtility 及其他 SSL 設定的相關資訊，請參閱 [WebSphere Liberty Profile](#) 及 [securityUtility 指令說明文件](#)。

4. 按一下**儲存**，然後重新啟動 Analytic Server 服務。

## 產生自簽憑證

您可以使用 securityUtility 來產生自簽憑證。例如：

```
/opt/ibm/spss/analyticserver/3.2.2/ae_wlpserver/bin/securityUtility createSSLCertificate
--server=<myserver> --password=<mypassword> --validity=365 --subject=CN=<mycompany>,O=<myOrg>,C=<myCountry>
```

### 附註：

- 您必須提供 **CN** 值的適當主機網域名稱。
- 將 key.jks 中的資訊複製到 trust.jks（兩個檔案必須相同）。
- 編輯 ssl.keystore.config 參數。例如：

```
<ssl id="defaultSSLConfig"
keyStoreRef="defaultKeyStore"
trustStoreRef="defaultTrustStore"
clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
location="/opt/ibm/spss/analyticserver/3.2.2
/ae_wlpserver/usr/servers/aeserver/resources/security/key.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
<keyStore id="defaultTrustStore"
location="/opt/ibm/spss/analyticserver/3.2.2
/ae_wlpserver/usr/servers/aeserver/resources/security/trust.jks"
type="JKS" password="{xor}Dz4sLG5tbGs="/>
```

## 透過 SSL 與 Apache Hive 通訊

您必須更新 hive.properties 檔案，才能透過 SSL 連線與 Apache Hive 通訊。或者，如果 Apache Hive 環境已啟用高可用性，您可以在主要 Analytic Server 資料來源頁面上選取高可用性參數。

### 更新 hive.properties 檔案

1. 開啟 hive.properties 檔案。該檔案位於：`/opt/ibm/spss/analyticserver/3.2/ae_wlpserver/usr/servers/aeserver/configuration/database`
2. 尋找下列行：

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasesname};user={db.username};password={db.password}
```

3. 新增下面的**粗體**資訊來更新行：

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasesname};user={db.username};password={db.password}
;
ssl=true;sslTrustStore=pathtotheirtruststorefile;trustStorePassword=xxxtheirTrustStorePassword
```

4. 儲存 hive.properties 檔案。

## 啟用對 Essentials for R 的支援

Analytic Server 支援對 R 模型評分以及執行 R Script。

若要在順利安裝 Analytic Server 之後配置對 R 的支援，請執行下列動作：

1. 供應 Essentials for R 的伺服器環境。

### RedHat Linux x86\_64

執行下列指令：

```
yum update
yum install -y zlib zlib-devel
yum install -y bzip2 bzip2-devel
yum install -y xz xz-devel
yum install -y pcre pcre-devel
yum install -y libcurl libcurl-devel
```

## Ubuntu Linux

執行下列指令：

```
apt-get update
apt-get install -y zlib1g-dev
apt-get install -y libreadline-dev
apt-get install -y libxt-dev
apt-get install -y bzip2
apt-get install -y libbz2-dev
apt-get install -y liblzma-dev
apt-get install -y libpcre3 libpcre3-dev
apt-get install -y libcurl4-openssl-dev
apt-get install -y liblzma-dev
apt-get install -y libpcre3 libpcre3-dev
apt-get install -y libcurl4-openssl-dev
```

## SUSE Linux

SUSE 上的 Essentials for R 安裝需要相容的 FORTRAN，這通常在配置的 ZYPPER 儲存庫中提供（僅位於 SUSE SDK 媒體中）。因此，由於無法安裝 FORTRAN，在 SUSE 伺服器對的 Essentials for R 進行的 Ambari 安裝將失敗。使用下列步驟以在 SUSE 上供應：

- a. 安裝 GCC C++。

```
zypper install gcc-c++
```

- b. 安裝 GCC FORTRAN。必要的 RPM 檔可以從 SUSE SDK 媒體中進行複製，且必須使用下列順序安裝。

```
zypper install libquadmath0-4.7.2_20130108-0.19.3.x86_64.rpm
zypper install libgfortran3-4.7.2_20130108-0.19.3.x86_64.rpm
zypper install gcc43-fortran-4.3.4_20091019-0.37.30.x86_64.rpm
zypper install gcc-fortran-4.3-62.200.2.x86_64.rpm
```

- c. 執行下列指令以安裝 Essentials for R 程式庫。

```
R_PREFIX=/opt/ibm/spss/R
cd $R_PREFIX
rm -fr $R_PREFIX/r_libs
mkdir -p $R_PREFIX/r_libs
cd $R_PREFIX/r_libs
wget https://zlib.net/fossils/zlib-1.2.11.tar.gz --no-check-certificate
tar xzvf zlib-1.2.11.tar.gz
cd zlib-1.2.11/
./configure
make && make install
cd $R_PREFIX/r_libs
wget http://www.bzip.org/1.0.6/bzip2-1.0.6.tar.gz
tar xzvf bzip2-1.0.6.tar.gz
cd bzip2-1.0.6
sed "s|^CC=gcc|CC=gcc -fPIC|" -i ./Makefile
make -f Makefile-libbz2_so
make clean
make
make install
cd $R_PREFIX/r_libs
wget https://tukaani.org/xz/xz-5.2.3.tar.gz
tar xzvf xz-5.2.3.tar.gz
cd xz-5.2.3
./configure
make -j3
make install
cd $R_PREFIX/r_libs
wget http://ftp.pcre.org/pub/pcre/pcre-8.38.tar.gz
tar xzvf pcre-8.38.tar.gz
cd pcre-8.38
./configure --enable-utf8
make
make install
cd $R_PREFIX/r_libs
wget https://www.openssl.org/source/openssl-1.0.2l.tar.gz --no-check-certificate
tar xzvf openssl-1.0.2l.tar.gz
cd openssl-1.0.2l/
./config shared
make
make install
echo '/usr/local/ssl/lib' >> /etc/ld.so.conf
ldconfig
cd $R_PREFIX/r_libs
wget --no-check-certificate https://curl.haxx.se/download/curl-7.50.1.tar.gz
tar xzvf curl-7.50.1.tar.gz
cd curl-7.50.1
./configure --with-ssl
make -j3
make install
cd $R_PREFIX/r_libs
wget ftp://rpmfind.net/linux/opensuse/distribution/12.3/repo/oss/suse/x86_64/
libgomp1-4.7.2_20130108-2.1.6.x86_64.rpm --no-check-certificate
rpm -ivh libgomp1-4.7.2_20130108-2.1.6.x86_64.rpm
```

2. 下載 IBM SPSS Modeler Essentials for R RPM 或 DEB 的自行解壓縮保存檔 (BIN)。Essentials for R 可供下載 (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspsp>)。選擇您的堆疊、堆疊版本及硬體架構特定的檔案。

- 執行自行解壓縮的二進位檔，然後遵循指示以（選擇性地）檢視授權、接受授權，以及選擇線上或離線安裝。

#### 線上安裝

如果您的 Ambari 伺服器主機和叢集中的所有節點都可以存取 <https://ibm-open-platform.ibm.com>，請選擇線上安裝。

#### 離線安裝

如果您的 Ambari 伺服器主機沒有網際網路存取，請選擇離線。離線安裝將會下載必要的 RPM 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的機器上。然後，可以將 RPM 檔案複製到 Ambari 伺服器主機。

- 將必要的 Essentials for R RPM 檔或 DEB 檔複製到 Ambari 伺服器主機上的任何位置。您所需的 RPM/DEB 檔案視您的發行套件、版本及架構而定，如下所示。

##### **HDP 2.6 (x86\_64)**

[IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.x86\\_64.rpm](#)

##### **HDP 3.0 及 3.1 (x86\_64)**

[IBM-SPSS-ModelerEssentialsR-ambari-2.7-HDP-3.0-9.2.0.3-1.x86\\_64.rpm](#)

##### **HDP 2.6 (PPC64LE)**

[IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.ppc64le.rpm](#)

##### **HDP 3.0 及 3.1 (PPC64LE)**

[IBM-SPSS-ModelerEssentialsR-ambari-2.7-HDP-3.0-9.2.0.3-1.ppc64le.rpm](#)

##### **HDP 2.6、3.0 及 3.1 (Ubuntu)**

[IBM-SPSS-ModelerEssentialsR-ambari-3.2.2.0\\_3.2.2.0\\_amd64.deb](#)

- 安裝 RPM 或 DEB。在下列範例中，指令會將 Essentials for R 安裝在 HDP 2.6 (x86\_64) 上。

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.6-9.2.0.3-1.x86_64.rpm
```

在下列範例中，指令會將 Essentials for R 安裝在 HDP 2.6 (Ubuntu) 上。

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.2.2.0_3.2.2.0_amd64.deb
```

- 重新啟動 Ambari 伺服器。

```
ambari-server restart
```

- 登入 Ambari 伺服器，並透過 Ambari 主控台將 SPSS Essentials for R 安裝為服務。SPSS Essentials for R 應安裝在已安裝 Analytic Server 及 Analytic Metastore 的每個主機上。

註：在安裝 R 之前，Ambari 會嘗試安裝 gcc-c++ 和 gcc-gfortran (RHEL) 以及 gcc-fortran (SUSE)。這些套件會宣告為依賴於 R 的 Ambari 服務定義。請確保要安裝及執行 R 的伺服器配置為下載 gcc-c++ 和 gcc-[g]fortran RPM，或已安裝 GCC 和 FORTRAN 編譯器。如果 Essentials for R 的安裝失敗，請在安裝 Essentials for R 之前，先手動安裝這些套件。

- 重新整理 Analytic Server 服務。
- 使用第 28 頁的『更新用戶端相依關係』中的指示，執行 update\_clientdeps Script。
- 您還必須在管理 SPSS Modeler 伺服器的機器上安裝 Essentials for R。如需詳細資料，請參閱 [SPSS Modeler 說明文件](#)。

## 啟用關聯式資料庫來源

如果您在 Analytic Server meta 儲存庫及每一個 Analytic Server 主機上的共用目錄中提供 JDBC 驅動程式，則 Analytic Server 可以使用關聯式資料庫來源。依預設，此目錄為 `/usr/share/jdbc`。

若要變更共用目錄，請遵循下列步驟。

- 在「Ambari 服務」標籤中，導覽至 Analytic Server 服務的「配置」標籤。
- 開啟進階 **analytics.cfg** 區段。
- 在 **jdbc.drivers.location** 中指定 JDBC 驅動程式之共用目錄的路徑。

4. 按一下**儲存**。
5. 停止 Analytic Server 服務。
6. 按一下**重新整理**。
7. 啟動 Analytic Server 服務。

表 6. 支援的資料庫			
資料庫	支援的版本	JDBC 驅動程式 JAR	供應商
Amazon Redshift	8.0.2 或更新版本	RedshiftJDBC41-1.1.6.1006.jar 或更新版本	Amazon
BigSQL	4.1.0.0 或更新版本	db2jcc.jar	IBM
DashDB	Bluemix 服務	db2jcc.jar	IBM
Db2 for Linux, UNIX, and Windows	11.1、10.5、10.1、9.7	db2jcc.jar	IBM
Db2 z/OS	11、10	db2jcc.jar、db2_license_cisuz.jar	IBM
Greenplum	5	postgresql.jar	Greenplum
Hive	2.1、1.2	hive-jdbc-*.jar	Apache
MySQL	5.7、5.6	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7、6.x	nzjdbc.jar	IBM
Oracle	19c、12c、11g R2 (11.2)	19c : ojdbc8.jar, orai18n.jar  12c 及 11g R2 (11.2) : ojdbc6.jar、 orai18n.jar	Oracle
SQL Server	2017、2014、2012、2008 R2	sqljdbc4.jar	Microsoft
Teradata	15.1、15	tdgssconfig.jar、 terajdbc4.jar	Teradata

#### 附註

- 如果您在安裝 Analytic Server 之前建立了 Redshift 資料來源，則需要執行下列步驟，以使用 Redshift 資料來源。
  1. 在 Analytic Server 主控台中，開啟 Redshift 資料來源。
  2. 選取 Redshift 資料庫資料來源。
  3. 輸入 Redshift 伺服器位址。
  4. 輸入資料庫名稱及使用者名稱。應該會自動移入密碼。
  5. 選取資料庫表格。
- BigSQL 是 Apache Hadoop 環境的 IBM SQL 介面。BigSQL 不是關聯式資料庫，但 Analytic Server 支援透過 JDBC（用於 DB2 的同一個 JDBC jar 檔）來存取它。

適用於 BigSQL 與 Analytic Server 的一般用法是透過 HCatalog 資料來源來存取 BigSQL Hadoop/HBase 表格。

## 啟用 HCatalog 資料來源

Analytic Server 透過 Hive/HCatalog 提供對許多資料來源的支援。部分來源需要手動配置步驟。

1. 收集必要的 JAR 檔以啟用資料來源。啟用 Apache HBase 和 Apache Accumulo 的支援不需要其他的步驟。如果是其他 NoSQL 資料來源，請與資料庫供應商聯絡，取得儲存體處理程式及相關的 JAR。如需要支援 HCatalog 資料來源的相關資訊，請參閱 [IBM SPSS Analytic Server 3.2.2 使用手冊](#) 中的「使用 HCatalog 資料來源」小節。
2. 將這些 JAR 檔新增至 Analytic Server meta 儲存庫及每一個 Analytic Server 節點上的 {HIVE\_HOME}/auxlib 目錄及 /usr/share/hive 目錄。
3. 重新啟動 Hive Metastore 服務。
4. 重新整理 Analytic Metastore 服務。
5. 重新啟動每一個 Analytic Server 服務實例。

### 附註：

- Analytic Server Metastore 無法與 Hive Metastore 安裝在同一機器上。
- 透過 Analytic Server HCatalog 資料來源存取 HBase 資料時，存取使用者必須具有 HBase 表格的讀取權。
  - 在非 Kerberos 環境中，Analytic Server 使用 as\_user 存取 HBase (as\_user 必須具有 HBase 的讀取權)。
  - 在 kerberos 環境中，as\_user 及登入使用者都必須具有 HBase 表格的讀取權。

## NoSQL 資料庫

Analytic Server 支援任何有來自供應商的 Hive 儲存體處理程式可用的 NoSQL 資料庫。

啟用 Apache HBase 和 Apache Accumulo 的支援不需要其他的步驟。

如果是其他 NoSQL 資料庫，請與資料庫供應商聯絡，取得儲存體處理程式及相關的 JAR。

## 檔案型 Hive 表格

Analytic Server 支援任何有內建或自訂 Hive SerDe (序列化程式-解除序列化程式) 可用的檔案型 Hive 表格。

用於處理 XML 檔的 Hive XML SerDe 位於 Maven Central Repository，網址為 <http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>。

## MapReduce 第 2 版工作

使用 Analytic Server 自訂 **analytic.cfg** 小節中的 **preferred.mapreduce** 設定可控制如何處理 MapReduce 工作：

內容	說明
preferred.mapreduce	控制執行 MapReduce 工作的方法。有效值包括： <ul style="list-style-type: none"><li>· spark</li><li>· m3r</li><li>· hadoop</li></ul> 例如：preferred.mapreduce=spark

## Apache Spark

如果您要使用 Spark (2.x 或更新版本)，您必須在安裝 Analytic Server 期間手動新增 `spark.version` 內容。

1. 開啟 Amabri 主控台並在 Analytic Server 的進階 **analytics.cfg** 區段中新增下列內容。

- 鍵：`spark.version`
- 值：輸入適當的 Spark 版本號碼 (例如，2.x 或 None)。

2. 儲存配置。

註：您可以透過自訂 `analytics.cfg` 設定強制 HCatalog 永不使用 Spark。

1. 開啟 Amabri 主控台並在 Analytic Server 的自訂 **analytic.cfg** 區段中新增下列內容。

- 鍵：`spark.hive.compatible`
- 值：`false`

## 已啟用 Kerberos 的 HDP 3.0 (或更新版本) 環境

已啟用 Kerberos 的 HDP 3.0 (或更新版本) 環境可能需要其他安全配置設定。在 HDFS 中，`filesystem` `facl's` 使用在 `/warehouse/tablespace/managed/hive` 目錄中。`messages.log` 或 `as_trace.log` 檔中出現下列異常狀況時，您可以識別需求以在 Hive meta 儲存庫中設定 `facls`：

```
原因：org.apache.hadoop.hive.q1.metadata.HiveException: java.security.AccessControlException:
許可權遭拒：user=xxxx, access=READ, inode="/warehouse/tablespace/managed/hive/hcat_primitives":hive:hadoop:drwxrwx---
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:399)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:261)
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkPermission(FSPermissionChecker.java:193)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1850)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPermission(FSDirectory.java:1834)
at org.apache.hadoop.hdfs.server.namenode.FSDirectory.checkPathAccess(FSDirectory.java:1784)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkAccess(FSNamesystem.java:7767)
```

下列範例顯示 `setfacl` 指令，提供對 Hive warehouse 目錄的廣泛存取 (在此範例中，對 hadoop 群組的所有成員)：

```
hadoop fs -setfacl -R -m group:hadoop:rxw /warehouse/tablespace/managed/hive/
```

此外，需要更精細的存取控制時，應該使用更多限制變異。

下列網站提供其他參照資訊。

[https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/hdfs-acls/content/acl\\_examples.html](https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/hdfs-acls/content/acl_examples.html)

[https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/securing-hive/content/hive\\_sba\\_permissions\\_model.html](https://docs.hortonworks.com/HDPDocuments/HDP3/HDP-3.1.0/securing-hive/content/hive_sba_permissions_model.html)

## 變更 Analytic Server 所用的埠

依預設，Analytic Server 對 HTTP 使用 9080 埠，並對 HTTPS 使用 9443 埠。若要變更埠設定，請遵循下列步驟。

1. 在「Ambari 服務」標籤中，導覽至 Analytic Server 服務的「配置」標籤。
2. 開啟進階 **analytics.cfg** 區段。
3. 分別在 **http.port** 及 **https.port** 中指定所要的 HTTP 和 HTTPS 埠。
4. 按一下儲存。
5. 重新啟動 Analytic Server 服務。

## 高可用性 Analytic Server

您可以使 Analytic Server 高度可用，方法為將它作為服務新增至叢集中的多個節點。

1. 在 Ambari 主控台中，導覽至「主機」標籤。
2. 選取尚未將 Analytic Server 作為服務執行的主機。
3. 在「摘要」標籤上，按一下**新增**，然後選取 Analytic Server。

## 4. 按一下確認新增

### 多叢集支援

多叢集特性是 IBM SPSS Analytic Server 的高可用性功能的加強功能，並在多租戶環境中提供改良的隔離。依預設，安裝 Analytic Server 服務（在 Ambari 或 Cloudera Manager 中）會定義單一分析伺服器叢集。

叢集規格會定義 Analytic Server 叢集成員資格。使用 XML 內容達成修改叢集規格（在 Ambari Analytic Server 配置的 `analytics-cluster` 欄位中，或透過編輯 Cloudera Manager 的 `configuration/analytics-cluster.xml` 檔）。當配置多個 Analytic Server 叢集時，需要使用其自己的負載平衡器將要求饋送至每一個 Analytic Server 叢集。

使用多叢集特性會確保一個租戶的工作不能對另一個租戶叢集中執行工作產生負面影響。對於高可用性工作，僅在起始工作所在的 Analytic Server 叢集範圍內發生工作失效接手。下列範例提供多叢集 XML 規格。

**註：**可以透過將 Analytic Server 作為服務新增至叢集中的多個節點，讓其成為高度可用。

```
<analyticServerClusterSpec>
  <cardinality>1+</cardinality>
  <cluster name="cluster1">
    <memberName>one.cluster</memberName>
    <memberName>two.cluster</memberName>
  </cluster>
  <cluster name="cluster2">
    <memberName>three.cluster</memberName>
    <memberName>four.cluster</memberName>
  </cluster>
</analyticServerClusterSpec>
```

在前一個範例中，需要兩個負載平衡器。一個負載平衡器將要求傳送至 `cluster1` 成員（`one.cluster` 及 `two.cluster`），另一個將要求傳送至 `cluster2` 成員（`three.cluster` 及 `four.cluster`）。

下列範例提供單一叢集 XML 規格（預設配置）。

```
<analyticServerClusterSpec>
  <cardinality>1</cardinality>
  <cluster name="cluster1">
    <memberName>*</memberName>
  </cluster>
</analyticServerClusterSpec>
```

在前一個範例中，需要單一負載平衡器來處理已配置多個叢集成員的情況。

### 附註

- 只有單態叢集支援在 `memberName` 元素中使用萬用字元（例如，叢集基數 = "1"）。基數元素的有效值是 1 與 1+。
- `memberName` 的指定方式，必須與為其指派 Analytic Server 角色之主機名稱的指定方式相同。
- 套用叢集配置變更之後，必須重新啟動叢集中的所有所有叢集中的所有伺服器。
- 在 Cloudera Manager 中，您必須在所有 Analytic Server 節點上修改維護 `analytics-cluster.xml` 檔。必須維護所有節點，以確保它們包含相同的內容。

## 最佳化小型資料的 JVM 選項

您可以編輯 JVM 內容，以最佳化執行小型 (M3R) 工作時的系統。

在 Ambari 主控台中，請參閱 Analytic Server 服務中「配置」標籤的進階 `analytics-jvm-options` 區段。修改下列參數，會設定伺服器（用來管理 Analytic Server）上所執行工作的資料堆大小；亦即，不是 Hadoop。這對於執行較小 (M3R) 工作而言非常重要，而且可能需要使用這些值進行試驗以最佳化系統。

```
-Xms512M
-Xmx2048M
```

## 升級 Python - HDP

本節說明從 Python 2.x 手動升級至 Python 3.7 的處理程序

1. 在每一個叢集節點上安裝 Python 3.7。如需相關資訊，請參閱 [Python 網站](#)。
2. 在每一個叢集節點上安裝 NumPy。如需相關資訊，請參閱 [NumPy 安裝指示](#)。

3. 在每一個叢集節點上安裝 pandas。如需相關資訊，請參閱 [pandas 安裝指示](#)。
4. 將 `spark.driver.python=<python3.7 executable path>` 新增至 Ambari 配置的自訂 **analytics.cfg** 區段。例如：

```
spark.driver.python=/opt/python3/bin/python3.7
```

## 更新用戶端相依關係

本節說明如何使用 `update_clientdeps` Script，來更新 Analytic Server 服務的相依關係。

1. 以 root 身分登入 Ambari 伺服器。
2. 切換至 `/var/lib/ambari-server/resources/stacks/<stack-name>/<stack-version>/services/ANALYTICSERVER/package/scripts` 目錄；請參閱範例下列。

```
cd "/var/lib/ambari-server/resources/stacks/HDP/2.6/services/ANALYTICSERVER/package/scripts"
```

3. 使用下列引數執行 `update_clientdeps` Script。

- u <ambari-user>**  
Ambari 帳戶使用者名稱
- p <ambari-password>**  
Ambari 帳戶使用者的密碼。
- h <ambari-host>**  
Ambari 伺服器的主機名稱。
- x <ambari-port>**  
Ambari 在其上接聽的埠。

請參閱下列範例。

```
./update_clientdeps.sh -u admin -p admin -h host.domain -x 8080
```

4. 使用下列指令重新啟動 Ambari 伺服器。

```
ambari-server restart
```

## 配置 Apache Knox

Apache Knox 閘道是可以為 Apache Hadoop 服務提供單點安全存取的系統。該系統為使用者（存取叢集資料及執行工作）和操作員（控制存取及管理叢集）簡化了 Hadoop 安全。閘道可以作為伺服器（或伺服器叢集）來執行，以為一或多個 Hadoop 叢集提供服務。

註：IBM SPSS Analytic Server 與 Kerberos 單一登入 (SSO) 一起使用時，不支援 Apache Knox。

Apache Knox 閘道能夠有效地隱藏 Hadoop 叢集拓撲的詳細資料，並與 Enterprise LDAP 及 Kerberos 相整合。下列各節提供了必要的 Apache Knox 及 Analytic Server 配置作業的相關資訊。

### 必備項目

- 已知 Apache Knox 問題不會傳送 HTTP Cookie 及標頭中包含的安全資訊（如需相關資訊，請參閱 <https://issues.apache.org/jira/browse/KNOX-895>）。此問題在 Knox 0.14.0（或更新版本）中已解決。您必須取得包含 Knox 0.14.0（或更新版本）的 Hortonworks 發行套件，然後才能將 Knox 與 Analytic Server 搭配使用。請聯絡 Hortonworks 提供者，以取得相關資訊。
- Analytic Server 節點必須透過無密碼的 SSH 連線與 Knox 伺服器連接。無密碼的 SSH 連線將從 Analytic Server 移至 Knox (**Analytic Server > Knox**)。
- Analytic Server 必須在 Knox 服務安裝之後進行安裝。

在某些情況下，非預期的問題會導致無法自動複製配置檔。在這些情況下，您必須手動複製下列配置檔：

- `com.ibm.spss.knox_0.6-3.2.2.0.jar`：該檔案必須從 Analytic Server 位置複製：

```
<Analytic_Server_Installation_Path>/ae_wlpserver/usr/servers/aeserver/apps/  
AE_BOOT.war/WEB-INF/lib
```

至 Knox 伺服器節點：

```
/KnoxServicePath/ext
```

例如：/usr/iop/4.1.0.0/knox/ext

· rewrite.xml 和 service.xml：這兩個檔案必須從 Analytic Server 位置複製：

```
<Analytic_Server_Installation_Path>/ae_wlpserver/usr/servers/aeserver/  
configuration/knox
```

至 Knox 伺服器節點：

```
/KnoxServicePath/data/services
```

例如：/usr/iop/4.1.0.0/knox/data/services

註：有兩組 rewrite.xml 及 service.xml 檔案（一組用於 http://rest 資料流量，另一組用於 ws://websocket 資料流量）。將 analyticserver 及 analyticserver\_ws 的所有 rewrite.xml 及 service.xml 檔案複製到 Knox 伺服器節點中。

## 配置 Ambari

必須在 Ambari 使用者介面中配置 Analytic Server 服務：

1. 在 Ambari 使用者介面中，導覽至 **Knox > 配置 > 進階拓撲**。現行 Knox 配置設定顯示在內容視窗中。
2. 將下列兩個服務新增至 Knox 配置的進階拓撲區段中：

```
<service>  
  <role>ANALYTICSERVER</role>  
  <url>http://{analyticserver-host}:{analyticserver-port}/analyticserver</url>  
</service>  
<service>  
  <role>ANALYTICSERVER_WS</role>  
  <url>ws://{analyticserver-host}:{analyticserver-port}/analyticserver</url>  
</service>
```

{analyticserver-host} 和 {analyticserver-port} 必須取代為適當的 Analytic Server 伺服器名稱及埠號：

- {analyticserver-host} URL 可在 Ambari 使用者介面中找到（**SPSS Analytic Server > 摘要 > Analytic Server**）。
- {analyticserver-port} 編號可在 Ambari 使用者介面中找到（**SPSS Analytic Server > 配置 > 進階 analytics.cfg > http.port**）。

註：當 Analytic Server 部署至多個節點，並使用 LoadBalancer 時，{analyticserver-host} 和 {analyticserver-port} 必須對應於 LoadBalancer URL 和埠號。

3. 重新啟動 Knox 服務。

當使用 LDAP 時，Knox 預設為提供的「展示」LDAP。您可以變更為企業 LDAP 伺服器（例如：Microsoft LDAP 或 OpenLDAP）。

## 配置 Analytic Server

若要對 Analytic Server 使用 LDAP，必須將 Analytic Server 配置為使用和 Apache Knox 所用相同的 LDAP 伺服器。必須更新下列 Ambari 設定的 <value> 項目，以反映適當的 Knox LDAP 伺服器設定：

- main.ldapRealm.userDnTemplate
- main.ldapRealm.contextFactory.url

這些值在 Ambari 使用者介面中可用：**Knox > 配置 > 進階拓撲**。例如：

```
<param>  
  <name>main.ldapRealm.userDnTemplate</name>  
  <value>uid={0},ou=people,dc=hadoop,dc=apache,dc=org</value>  
</param>  
<param>
```

```
<name>main.ldapRealm.contextFactory.url</name>
<value>ldap://{{knox_host_name}}:33389</value>
</param>
```

更新 LDAP 設定之後，重新啟動 Knox 服務。

**重要：** Analytic Server 管理者密碼必須與 Knox 管理者密碼相同。

## 配置 Apache Knox

1. 重新整理 Knox gateway.jks 檔案：

- a. 在 Knox 伺服器上，停止 Knox 服務。
- b. 將 gateway.jks 從 /var/lib/knox/data-2.6.2.0-205/security/keystores 中刪除。
- c. 重新啟動 Knox 服務。

2. 在 Knox 伺服器上，建立子目錄 <knox\_server>/data/service/analyticserver/3.2.2.0，然後將 service.xml 及 rewrite.xml 檔案上傳至新目錄。這兩個檔案位於 Analytic Server 的 <analytic\_server>/configuration/knox/analyticserver/ 上（例如，/opt/ibm/spss/analyticserver/3.2/ae\_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/\*.xml）

3. 在 <knox\_server>/bin 中，執行 Script ./knoxcli.sh redeploy --cluster default

4. 將 com.ibm.spss.knoxservice\_0.6-\*.jar 檔案上傳至 <knox\_server>/ext。該檔案位於 Analytic Server 的 <analytic\_server>/apps/AE\_BOOT.war/WEB-INF/lib/com.ibm.spss.knox\_0.6-3.2.2.0.jar 上（例如，/opt/ibm/spss/analyticserver/3.2/ae\_wlpserver/usr/servers/aeserver/apps/AE\_BOOT.war/WEB-INF/lib/com.ibm.spss.knox\_0.6-3.2.2.0.jar）。

5. 在 Ambari 使用者介面中，在 **Knox > 配置 > 進階拓撲** 中新增下列元素：

```
<service>
  <role>ANALYTICSERVER</role>
  <url>http://{analyticserver-host}:{analyticserver-port}/analyticserver</url>
  <role>ANALYTICSERVER_WS</role>
  <url>ws://{analyticserver-host}:{analyticserver-port}/analyticserver</url>
</service>
```

**註：** 依預設，WebSocket 功能已停用。透過將 /conf/gateway-site.xml 檔案中的 gateway.websocket.feature.enabled 內容變更為 true 可啟用。

6. 在 Ambari 使用者介面中，於 **Knox > 配置 > 進階 users-ldif** 內新增或更新使用者（例如，admin、qauser1、qauser2）。
7. 從 **Knox > 服務動作 > 啟動展示 LDAP** 重新啟動 LDAP。
8. 重新啟動 Knox 服務。

## 已啟用 Apache Knox 之 Analytic Server 的 URL 結構

已啟用 Knox 之 Analytic Server 使用者介面的 URL 為 https://{knox-host}:{knox-port}/gateway/default/analyticserver/admin

- https 通訊協定 - 使用者必須接受憑證才能繼續在 Web 瀏覽器中進行操作。
- knox-host 是 Knox 的主機。
- knox-port 是 Knox 的埠號。
- URI 為 gateway/default/analyticserver。

## 為每一個 YARN 佇列配置個別動態資源配置 - HDP

您可以為每一個 YARN 佇列配置個別「動態資源配置」。

## 使用者及租戶模式對映 - Hortonworks 資料平台

使用者及租戶可以提交至不同的 YARN 佇列，並且每一個使用者或租戶都會對映至不同的 YARN 佇列（以利用「動態資源配置」）。可以定義 **user** 模式或 **tenant** 模式，以對映至 YARN 佇列。在 Analytic Server 3.2.1 Fix Pack 1 之前，所有 Spark 工作都限制為單一 YARN 佇列。

從 IBM SPSS Analytic Server 3.2.1 Fix Pack 1 開始，當使用者/租戶的串流導致 Spark 工作在系統上執行時，個別 YARN 佇列將以提交串流至 Analytic Server 的使用者/租戶身分執行。可以為不同的使用者/租戶作業同時執行多個 YARN 佇列。

只要使用者登入 Analytic Server（有時，在使用者登出且沒有其他作用中使用者工作之後），每一個 YARN 佇列就會繼續執行。配置變數可以控制登出之後的時間量：**as.spark.driver.cleanup.delay**。

為提交 Spark 工作的每一個使用者建立 **SparkDriver** 處理程序。在使用者沒有作用中工作約 2 分鐘（預設值）且沒有 **HTTPSession** 活動之後，每一個使用者的 **SparkDriver** 處理程序都會終止。

註：Analytic Server 關閉時，所有 **SparkDriver** 處理程序都會終止。

使用下列步驟，以將 Analytic Server 新增至現有叢集：

1. 在 Ambari 使用者介面中，導覽至 **SPSS Analytic Server 服務 > 配置 > 進階 analytics.cfg** 標籤。
2. 將 **resource.pool.enabled** 值變更為 **true**。
3. 在 **Custom analytics.cfg** 標籤上新增下列內容：

```
//Using user/tenant mapping to YARN pool
yarn.queue.mode=<user/tenant>
yarn.queue.mapping=<user1:test,user2:production>/<tenant1:test,tenant2:production>
yarn.queue.default=default
as.spark.driver.cleanup.delay=2
as.sparkdriver.max.memory=1G
```

表 8. 自訂 *analytics.cfg* 內容

內容	說明
yarn.queue.mode	設定 YARN 佇列的對映模式。yarn.queue.mode=user 時，未提交工作/串流至 Analytic Server 的每一個使用者執行個別 YARN 佇列。可以為不同的使用者工作/串流同時執行多個 YARN 佇列。  yarn.queue.mode=tenant 時，未提交工作/串流至 Analytic Server 的每一個租戶執行個別 YARN 佇列。可以為不同的租戶工作/串流同時執行多個 YARN 佇列。
yarn.queue.mapping	將使用者或租戶配對對映至在「YARN 佇列管理程式」中定義的 YARN 佇列。配對必須以逗點區隔（例如，tenant1:test,tenant2:production 用於租戶，或者 user1:test,user2:production 用於使用者）。
yarn.queue.default	將應用程式提交到的預設 YARN 佇列的名稱。您可以在「YARN 佇列管理程式」中指定自訂的 YARN 佇列名稱。
as.spark.driver.cleanup.delay	代表登出之後終止使用者 YARN 佇列之前分鐘數目的整數。預設值是 <b>2</b> 。此內容是選用項目。
as.sparkdriver.max.memory	設定每一個 <b>SparkDriver</b> 處理程序使用的記憶體數量。預設值是 <b>1G</b> 。此內容是選用項目。

4. 儲存配置，然後重新啟動 Analytic Server 服務。

## 參照

如需相關資訊，請參閱下列網站：

- <https://spark.apache.org/docs/latest/running-on-yarn.html>

## 在 Ambari 上移轉 IBM SPSS Analytic Server

Analytic Server 可以將資料及配置設定從現有 Analytic Server 安裝移轉至新的安裝。移轉可在相同叢集環境或新的叢集環境上發生。

### 將 Analytic Server 3.2.1.1 移轉至相同伺服器叢集上的 3.2.2

如果您已安裝 Analytic Server 3.2.1.1，您可以將 3.2.1.1 配置設定移轉至相同伺服器叢集上的 3.2.2 安裝。

1. 從 Analytic Server 舊版本 (Analytic Server 3.2.1.1) 收集配置設定。
  - a. 展開 `{AS_ROOT}\tools\unzip configcollector.zip` 保存檔 (其將建立名為 `configcollector` 的新資料夾)。
  - b. 執行 `configcollector` 資料夾中的 `configcollector.sh` Script。將產生的 (ZIP) `ASConfiguration_3.2.1.1.xxx.zip` 壓縮檔複製至不同資料夾位置 (作為備份)。
2. 將舊的 Analytic Server 3.2.1.1 版安裝中的分析根目錄備份到新位置中。
  - a. 如果您不確定分析根目錄的位置，請執行 `hadoop fs -ls` 指令。分析根目錄的路徑類似於 `/user/as_user/analytic-root/analytic-workspace`，其中 `as_user` 是擁有分析根目錄的使用者 ID。
  - b. 使用 `hadoop fs -copyToLocal` 及 `hadoop fs -copyFromLocal` 指令，將舊的 Analytic Server 版本 `analytic-workspace` 資料夾複製到新位置中 (例如，`/user/as_user/analytic-root/AS3211Location`)。
3. 如果您使用內嵌的 Apache Directory Server，請使用第三方 LDAP 用戶端工具備份現行使用者/群組配置。安裝 Analytic Server 3.2.2 之後，將備份使用者/群組配置匯入至 Apache Directory Server。

**註：**如果您使用外部 LDAP 伺服器，則可跳過此步驟。
4. 開啟 Ambari 主控台，並停止 **Analytic Server 服務**。
5. 解除安裝 Analytic Server 舊版本 (Analytic Server 3.2.1.1)，然後安裝 Analytic Server 3.2.2。如需安裝指示，請參閱 [第 3 頁的『第 2 章 Ambari 安裝與配置』](#)。
6. 開啟 Ambari 主控台，並停止 **Analytic Server 服務** (在 Ambari 上，確保 **Analytic Metastore 服務** 在執行中)。
7. 將步驟 2 中備份的 Analytic Server 3.2.1.1 分析根目錄複製到新的 Analytic Server 版本位置。
  - a. 將 `analytic-workspace` 從新安裝的 Analytic Server 版本中移除。
  - b. 將備份的 Analytic Server 3.2.1.1 分析工作區資料夾 (`/user/as_user/analytic-root/AS3211Location`) 複製到新版本位置 (例如，`/user/as_user/analytic-root/analytic-workspace`)。您必須確保分析工作區擁有者定義為 `as_user`。
8. 清除 Zookeeper 狀態。在 Zookeeper bin 目錄中 (例如，Hortonworks 上的 `/usr/hdp/current/zookeeper-client`)，執行下列指令：

```
./zkCli.sh rmr /AnalyticServer
```

9. 將步驟 1 中的備份保存檔 `ASConfiguration_3.2.1.1.xxx.zip` 複製到新的 Analytic Server 版本位置 (例如，`/opt/ibm/spss/analyticserver/3.2/`)。
10. 透過執行 `migrationtool.sh` Script 並傳遞 `ASConfiguration_3.2.1.1.xxx.zip` 保存檔 (由配置收集程式所建立) 的路徑作為引數，來執行移轉工具。例如：

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

11. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

12. 在 Ambari 主控台中，啟動 **Analytic Server 服務**。

## 將 Analytic Server 3.2.1.1 移轉至新伺服器叢集上的 3.2.2

如果您已安裝 Analytic Server 3.2.1.1，您可以將 3.2.1.1 配置設定移轉至新的伺服器叢集上的 3.2.2 安裝。

1. 根據第 5 頁的『在 Ambari 上安裝』中的指示，安裝新的 Analytic Server 版本。
2. 將舊安裝中的分析工作區複製到新安裝中。
  - a. 如果您不確定分析工作區的位置，請執行 `hadoop fs -ls`。分析工作區的路徑類似於 `/user/as_user/analytic-root/analytic-workspace`，其中 `as_user` 是擁有分析工作區的使用者 ID。
  - b. 移除新伺服器上的 `analytic-workspace`。
  - c. 使用 `hadoop fs -copyToLocal` 與 `hadoop fs -copyFromLocal`，將舊伺服器的分析工作區複製到新伺服器的 `/user/as_user/analytic-root/analytic-workspace/` 資料夾（確保擁有者設為 `as_user`）。
3. 如果您使用內嵌的 Apache Directory Server，請使用第三方 LDAP 用戶端工具備份現行使用者/群組配置。安裝 Analytic Server 3.2.2 之後，將備份使用者/群組配置匯入至 Apache Directory Server。  
**註：**如果您使用外部 LDAP 伺服器，則可跳過此步驟。
4. 在新伺服器上，開啟 Ambari 主控台，並停止 Analytic Server 服務（在 Ambari 上，確保 Analytic Metastore 服務在執行中）。
5. 從舊安裝收集配置設定。
  - a. 將新安裝中的 `configcollector.zip` 保存檔複製到舊安裝中的 `{AS_ROOT}\tools`。
  - b. 擷取 `configcollector.zip` 的副本，這會在您的舊安裝中建立一個新的 `configcollector` 子目錄。
  - c. 在舊安裝中執行配置收集器工具，方式是執行 `{AS_ROOT}\tools\configcollector` 中的 **configcollector** Script。將產生的壓縮 (ZIP) 檔複製到管理新安裝的伺服器。

**重要：**提供的 **configcollector** Script 與最近的 Analytic Server 版本不相容。如果您的 **configcollector** Script 發生問題，請聯絡 IBM 技術支援代表。

6. 清除 Zookeeper 狀態。在 Zookeeper bin 目錄中（例如，Hortonworks 上的 `/usr/hdp/current/zookeeper-client`），執行下列指令。

```
./zkCli.sh rmr /AnalyticServer
```

7. 透過執行 **migrationtool** Script 並傳遞配置收集器建立的壓縮檔路徑作為引數，來執行移轉工具。範例如下所示。

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

8. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

9. 在 Ambari 主控台中，啟動 Analytic Server 服務。

**註：**如果您已配置 R 以與現有 Analytic Server 安裝搭配使用，請遵循其配置步驟來配置新的 Analytic Server 安裝。

## 解除安裝

**重要：**當已安裝 Essentials for R 時，必須首先執行 `remove_R.sh` Script。在解除安裝 Analytic Server 之前，若無法解除安裝 Essentials for R，則會導致稍後無法解除安裝 Essentials for R。當解除安裝 Analytic Server 時，會移除 `remove_R.sh` Script。如需解除安裝 Essentials for R 的相關資訊，請參閱第 34 頁的『解除安裝 Essentials for R』。

1. 在 Analytic Metastore 主機上，以下列參數執行 `{AS_ROOT}/bin` 目錄中的 `remove_as.sh` Script。

- u** 必要。Ambari 伺服器管理者的使用者 ID。
- p** 必要。Ambari 伺服器管理者的密碼。
- h** 必要。Ambari 伺服器主機名稱。
- x** 必要。Ambari 伺服器埠。
- l** 可選。僱用安全模式。

範例如下。

```
remove_as.sh -u admin -p admin -h one.cluster -x 8081
```

從具有 Ambari 主機 one.cluster 的叢集移除 Analytic Server。

```
remove_as.sh -u admin -p admin -h one.cluster -x 8081 -l
```

在安全模式下，從具有 Ambari 主機 one.cluster 的叢集移除 Analytic Server。

註：此作業會移除 HDFS 上的 Analytic Server 資料夾。

註：此作業不會移除與 Analytic Server 相關聯的任何 DB2 綱目。請參閱 DB2 文件，以取得手動移除綱目的相關資訊

## 解除安裝 Essentials for R

1. 在 Essentials for R 主機上，以下列參數執行 {AS\_ROOT}/bin 目錄中的 remove\_R.sh Script。

- u** 必要項目。Ambari 伺服器管理者的使用者 ID。
- p** 必要項目。Ambari 伺服器管理者的密碼。
- h** 必要項目。Ambari 伺服器主機名稱。
- x** 必要項目。Ambari 伺服器埠。
- l** 選用項目。僱用安全模式。

範例如下。

```
remove_R.sh -u admin -p admin -h one.cluster -x 8081
```

從具有 Ambari 主機 one.cluster 的叢集移除 Essentials for R。

```
remove_R.sh -u admin -p admin -h one.cluster -x 8081 -l
```

在安全模式下，從具有 Ambari 主機 one.cluster 的叢集移除 Essentials for R。

2. 從 Ambari 伺服器服務目錄移除 R 服務目錄。例如，在 HDP 2.6 中，ESSENTIALR 目錄位於 /var/lib/ambari-server/resources/stacks/HDP/2.6/services 中。
3. 在 Ambari 主控台中，驗證 Essentials for R 服務是否不再存在。

## 第 3 章 Cloudera 安裝與配置

### Cloudera 概觀

Cloudera 是開放程式碼 Apache Hadoop 發行套件。包括 Apache Hadoop 的 Cloudera 發行套件 (CDH) 以該技術的企業級部署為目標。

Analytic Server 可以在 CDH 平台上執行。CDH 包含 Hadoop 的主要核心元素（其會對大型資料集（主要是 MapReduce 及 HDFS）提供可靠的可調式分散式資料處理），以及其他企業導向的元件（其會提供安全性、高可靠性及與硬體及其他軟體的整合）。

### Cloudera 特定必備項目

除了一般必備項目之外，還請檢閱下列資訊。

#### 服務

確保在每一個 Analytic Server 主機上已安裝下列實例。

- HDFS：Gateway、DataNode 或 NameNode
- Hive：Gateway、Hive Metastore Server 或 HiveServer2
- YARN：Gateway、ResourceManager 或 NodeManager

僅當使用它們的特性時才需要下列實例。

- Accumulo：Gateway
- HBase：Gateway、Master 或 RegionServer
- Spark 2：Gateway

#### meta 資料儲存庫

您可以使用 Db2 及 MySQL 作為 Analytic Server meta 資料儲存庫。如果您計劃使用 MySQL 作為 Analytic Server meta 資料儲存庫，請遵循第 37 頁的『針對 Analytic Server 配置 MySQL』的指示。

#### 無密碼的 SSH

為 Analytic Server 主機與叢集中的所有主機之間的 root 使用者設定無密碼的 SSH。

#### 預設 Umask

「預設 Umask」設定必須設為 022。例如：

<b>Default Umask</b> dfs.umaskmode, fs.permissions.umask-mode	<b>HDFS (Service-Wide)</b> <div style="border: 1px solid black; padding: 10px; display: inline-block;">022</div>
---	---

022 設定是最受限的 Umask，容許 Analytic Server 運作。

### 啟用了 Kerberos 的 Cloudera 環境

如果您計劃在啟用了 Kerberos 的 Cloudera 環境中安裝 Analytic Server，則必須驗證已使用與 Analytic Server 相容的方式正確配置 Kerberos。

下列各節適用於已安裝 Kerberos 的 Cloudera 環境。在 Cloudera 中安裝 Analytic Server 之前必須遵循下列各節。假定您具有基本的 Kerberos 鑑別知識，因為這些小節中包含 Kerberos 特定術語（例如，**kinit**、**kadmin** 等等）。

註：Analytic Server 會檢查 HDFS 配置，以使用 Kerberos 相關值進行鑑別。

## Kerberos 鑑別

安裝 Analytic Server 之前，驗證已在每一個 Cloudera 叢集節點上配置 Kerberos 鑑別。如需相關資訊，請參閱 Cloudera 產品說明文件中的在 Cloudera Manager 中配置鑑別。

註：在每一個 Cloudera 叢集節點上配置 Kerberos 鑑別之後，必須先重新啟動 **cloudera-scm-server** 及 **cloudera-scm-agent** 服務，然後再安裝 Analytic Server。必須在所有叢集節點上重新啟動 **cloudera-scm-agent** 服務。

## 在 Kerberos 中建立必要的帳戶

1. 在 Kerberos 使用者儲存庫中，為您打算授與 Analytic Server 之存取權的所有使用者建立帳戶。
2. 在 LDAP 伺服器上建立相同的帳戶（依前一個步驟）。
3. 在每個 Analytic Server 節點及 Hadoop 節點上，為在前一個步驟中所建立的每一位使用者，各建立一個作業系統使用者帳戶。使用者群組必須設為 **hadoop**。
  - 請確保這些使用者的 UID 在所有機器上都相符。您可以使用 **kinit** 指令登入每一個帳戶來測試此項目。
  - 請確保 UID 遵循用於提交工作的使用者 ID 下限 YARN 設定。這是 **container-executor.cfg** 中的 **min.user.id** 設定。例如，如果 **min.user.id** 為 1000，則建立的每個使用者帳戶都必須具有大於或等於 1000 的 UID。
4. 針對 Analytic Server 管理者使用者，在 HDFS 上建立使用者起始檔案夾。該資料夾許可權必須設為 755，擁有者必須定義為 **admin**，且使用者群組必須設為 **hdfs**。請參閱下列粗體範例：

```
[root@xxxxx configuration]# hadoop fs -ls /user
Found 9 items
drwxIwxIwx - hdfs supergroup 0 2017-07-26 03:41 /user/AE
drwxI-xI-x - admin hdfs 0 2017-06-08 01:33 /user/admin
drwxI-xI-x - as_user hdfs 0 2017-06-06 01:00 /user/as_user
drwx----- - hdfs supergroup 0 2017-07-31 00:17 /user/hdfs
drwxIwxIwx - mapred hadoop 0 2017-06-05 00:28 /user/history
drwxIwxI-t - hive hive 0 2017-06-05 00:30 /user/hive
drwxIwxI-x - hue hue 0 2017-06-05 00:30 /user/hue
drwxIwxI-x - impala impala 0 2017-07-19 00:52 /user/impala
drwxI-x--x - spark spark 0 2017-06-05 01:34 /user/spark
```

5. 在 HDFS 上為所有 Analytic Server 標準使用者（例如，**user1**）建立使用者起始檔案夾。資料夾擁有者是實際使用者，並且使用者群組必須設為 **hdfs**。
6. 如果您打算使用 HCatalog 資料來源，且 Analytic Server 安裝在與 Hive metastore 不同的機器上，則必須在 HDFS 上模擬 Hive 用戶端。
  - a. 在 Cloudera Manager 中導覽至 HDFS 服務的「配置」標籤。

註：如果尚未設定下列設定，則在配置標籤上，可能不會顯示這些設定。在此情況下，請執行搜尋來尋找它們。
  - b. 編輯 **hadoop.proxyuser.hive.groups** 設定，讓它的值為 **\***，或是包含獲允許登入 Analytic Server 之所有使用者的群組。
  - c. 編輯 **hadoop.proxyuser.hive.hosts** 設定，讓它的值為 **\***，或是 Hive Metastore 及每個 Analytic Server 實例在其上被安裝為服務的主機清單。
  - d. 重新啟動 HDFS 服務。

在執行了這些步驟並安裝了 Analytic Server 之後，Analytic Server 就會無聲自動配置 Kerberos。

## 啟用 Kerberos 模擬

模擬容許執行緒在與執行緒所屬程序之安全環境定義不同的安全環境定義中執行。例如，模擬可提供一種方式，讓 Hadoop 工作以標準 Analytic Server 使用者 (as\_user) 以外的使用者身分執行。若要啟用 Kerberos 模擬，請執行下列動作：

1. 開啟 Cloudera Manager，並在 **core-site.xml** 的叢集層面進階配置 **Snippet (安全閘)** 區域中 (位於 **HDFS (服務層面) > 配置標籤**上)，新增或更新下列內容。

- 名稱：hadoop.proxyuser.as\_user.hosts
- 值：\*
- 名稱：hadoop.proxyuser.as\_user.groups
- 值：\*

註：**core-site.xml** 設定適用於 Hadoop 配置 (而非 Analytic Server)。

2. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

## 針對 Analytic Server 配置 MySQL

在 Cloudera Manager 中配置 IBM SPSS Analytic Server 需要安裝及配置 MySQL 伺服器資料庫。

1. 在儲存 MySQL 資料庫的節點上，透過指令視窗，執行下列指令：

```
yum install mysql-server
```

註：針對 SuSE Linux，使用 `zypper install mysql`。

2. 在每一個 Cloudera 叢集節點上，透過指令視窗，執行下列指令：

```
yum install mysql-connector-java
```

註：針對 SuSE Linux，使用 `sudo zypper install mysql-connector-java`。

3. 決定並記下 Analytic Server 存取 MySQL 資料庫時所使用的 Analytic Server 資料庫名稱、資料庫使用者名稱及資料庫密碼。
4. 根據第 39 頁的『在 Cloudera 上安裝』中的指示，安裝 Analytic Server。
5. 將 `/opt/cloudera/parcels/AnalyticServer/bin/add_mysql_user.sh` Script 從 Cloudera 管理的其中一個伺服器，複製到 MySQL 資料庫安裝所在的節點。使用適合於特定配置的參數，執行 Script。例如：

```
./add_mysql_user.sh -u <database_user_name> -p <database_password> -d <database_name>
```

附註：當資料庫在安全模式 (已設定 root 使用者密碼) 下執行時，需要 `a -r <dbRootPassword>` 參數。

當資料庫在安全模式下執行但使用者名稱不是 root 時，需要 `-r <dbUserPassword>` 和 `-t <dbUserName>` 參數。

## 安裝事先檢查與事後檢查工具 - Cloudera

### 工具位置與必要條件

在安裝 Analytic Server 服務之前，請先在將會是 Analytic Server 服務的一部分的所有節點上執行事先檢查工具，以驗證您的 Linux 環境是否已準備安裝 Analytic Server。

在安裝過程中會自動呼叫事先檢查工具。在每一台主機上執行安裝之前，此工具都會檢查每一個 Analytic Server 節點。您也可以將在每一個節點上手動呼叫事先檢查工具，這可以在安裝服務之前驗證機器。

執行自行解壓縮的 Analytic Server 二進位檔之後，事先檢查工具位於下列目錄：

## · Cloudera

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/com.spss.ibm.checker.zip  
[root@servername ~]# cd /opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/  
[root@servername tools]# ls  
com.spss.ibm.checker.zip configcollector.zip regex-files
```

註：在您執行二進位執行檔、配送（**下載 > 配送**）並在 Cloudera Manager 的「包裹」頁面中啟動 Analytic Server 之前，tools 目錄中的事先檢查工具無法使用。

安裝 Analytic Server 之後，事後檢查工具位於下列目錄中：

## · Cloudera

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/tools/com.spss.ibm.checker.zip
```

工具必須以 root 身分執行且需要 Python 2.6.X（或更高版本）。

如果事先檢查工具報告了任何失敗，則必須先解決這些失敗，然後再繼續安裝 Analytic Server。

## 執行事先檢查工具

### 自動

透過 Cloudera Manager 主控台將 Analytic Server 安裝為服務時，在安裝 Analytic Server 的過程中可以自動呼叫事先檢查工具。您必須手動輸入 Cloudera Manager 管理者的使用者名稱和密碼：

### Add SPSS Analytic Server Service to Cluster 1

#### Review Changes

<b>Cloudera Manager</b> Administrator account username <small>cm.admin.username</small>	<b>Analytic Server Default Group</b> <input type="text" value="admin"/> <small>Missing required value: Cloudera Manager Administrator account username</small>
<b>Cloudera Manager</b> Administrator account password <small>cm.admin.password</small>	<b>Analytic Server Default Group</b> <input type="password" value="....."/> <small>Missing required value: Cloudera Manager Administrator account password</small>

圖 4. Cloudera Manager 管理者設定

### 手動

您可以在每一個叢集節點上手動呼叫事先檢查工具。

下列事先檢查範例會檢查在 myclouderahost.ibm.com:7180 上執行的 Cloudera 叢集 MyCluster)，並使用登入認證 admin:admin：

```
python ./precheck.py --target C --cluster MyCluster --username admin  
--password admin --host myclouderahost.ibm.com --port 7180 --ssl
```

### 附註：

- 需要引數 --target、--host、--port 及 --username。
- --host 值必須由 IP 位址或完整網域名稱提供。
- 省略密碼引數時，該工具會提示您輸入密碼。
- precheck.py 指令包括使用說明，使用 --h 引數來顯示該說明 (python ./precheck.py --help)。
- --cluster 是選用引數（未使用 --cluster 時會識別現行叢集）。

當事先檢查工具執行其檢查時，每個檢查的狀態都會顯示在指令視窗中。失敗發生時，會在日誌檔中記載詳細資訊（確切的日誌檔位置在指令視窗中提供）。需要更多支援時，可將日誌檔提供給 IBM 技術支援中心。

## 執行事後檢查工具

事後檢查工具會驗證 Analytic Server 是否適當地執行且是否能夠處理簡單的工作。下列事後檢查範例會檢查在 `myanalyticserverhost.ibm.com:9443` 上執行的 Analytic Server 實例（啟用了 SSL），並使用登入認證 `admin:ibmspss`：

```
python ./postcheck.py --target C --host myanalyticserverhost.ibm.com --port 9443
--username admin --password ibmspss --ssl
```

將 Knox 與 Analytic Server 搭配使用時，指令如下所示：

```
python ./postcheck.py --target C --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default
```

若要執行單一檢查，請使用下列指令：

```
python ./postcheck.py --target C --host myknoxserverhost.ibm.com --port 8443
--username admin --password ibmspss --ssl --gateway_url /gateway/default --check AS_PYSPARK_BUILDMODEL
```

### 附註：

- 需要引數 `--target`、`--host`、`--port` 及 `--username`。
- `--host` 值必須由 IP 位址或完整網域名稱提供。
- 省略密碼引數時，該工具會提示您輸入密碼。
- `postcheck.py` 指令包括使用說明，使用 `--h` 引數來顯示該說明 (`python ./postcheck.py --help`)。

當事後檢查工具執行其檢查時，每個檢查的狀態都會顯示在指令視窗中。失敗發生時，會在日誌檔中記載詳細資訊（確切的日誌檔位置在指令視窗中提供）。如果需要更多支援，可將日誌檔提供給 IBM 技術支援中心。

## 在 Cloudera 上安裝

下列步驟說明在 Cloudera Manager 上手動安裝 IBM SPSS Analytic Server 的處理程序。

### Analytic Server 3.2.2

#### 線上安裝

1. 導覽至 [IBM Passport Advantage® 網站](#)，並將特定於您的堆疊、堆疊版本以及硬體架構的自行解壓縮二進位檔下載至 Cloudera 叢集內的主機中。可用的 Cloudera 二進位檔如下所示：

說明	二進位檔名
IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11、5.12、5.13、5.14、5.15、6.0、6.1、6.2 及 6.3 Ubuntu（英文版）	<code>spss_as-3.2.2.0-cdh5.11-6.3-ubun.bin</code>
IBM SPSS Analytic Server 3.2.2 for Cloudera 5.11、5.12、5.13、5.14、5.15、6.0、6.1、6.2 及 6.3 Linux x86-64（英文版）	<code>spss_as-3.2.2.0-cdh5.11-6.3-lx86.bin</code>

2. 在 Cloudera Manager 主要叢集節點上，執行 Cloudera 自行解壓縮的 `*.bin` 安裝程式。遵循安裝提示，接受授權合約，並保留預設 CSD 安裝目錄。

註：如果該目錄與預設位置不同，您必須指定不同的 CSD 目錄。

3. 在安裝完成之後，使用下列指令來重新啟動 Cloudera Manager：

```
service cloudera-scm-server restart
```

- 開啟 Cloudera Manager 介面（例如，[http://\\${CM\\_HOST}:7180/cm/manager/login](http://${CM_HOST}:7180/cm/manager/login)，其使用預設登入認證 admin/admin），重新整理遠端包裹儲存庫 URL（位於主機 > 包裹 > 配置中），並驗證此 URL 正確。例如：

```
https://ibm-open-platform.ibm.com
```

註：可以更新包裹更新頻率及遠端包裹儲存庫 URL，以滿足您的特定需要。

- 在 Cloudera Manager 重新整理包裹檔案之後（您可以透過按一下**檢查新包裹**，以手動重新整理包裹檔案），您會看到 **AnalyticServer** 包裹狀態設為**遠端可用**。
- 選取**下載 > 配送 > 啟動**。**AnalyticServer** 包裹狀態會更新為**已配送、已啟動**。
- 在 Cloudera Manager 中，將 Analytic Server 新增為服務，並決定在何處放置 Analytic Server。您需要在「**新增服務精靈**」中提供下列資訊：

註：「**新增服務精靈**」會顯示服務建立處理程序各階段期間的整體進度，並在服務於叢集上順利安裝及配置時提供最終確認訊息。

- Analytic Server meta 儲存庫主機名稱
- Analytic Server meta 儲存庫資料庫名稱
- Analytic Server meta 儲存庫使用者名稱
- Analytic Server meta 儲存庫密碼

#### MySQL 用作 Analytic Server meta 資料儲存庫

- Analytic Server meta 儲存庫驅動程式類別：`com.mysql.jdbc.Driver`
- Analytic Server meta 儲存庫 URL：`jdbc:mysql://${MySQL_DB}/${DBName}?createDatabaseIfNotExist=true`

{MySQL\_DB} 是其中安裝 MySQL 之伺服器的主機名稱

#### Db2 用作 Analytic Server meta 資料儲存庫

- Analytic Server meta 儲存庫驅動程式類別：`com.ibm.db2.jcc.DB2Driver`
- Analytic Server meta 儲存庫 URL：`jdbc:db2://{Db2_HOST}:{PORT}/{DBName}:currentSchema={SchemaName};`

{Db2\_HOST} 是其中安裝 Db2 之伺服器的主機名稱。

{PORT} 是 Db2 在其上接聽的埠。

{SchemaName} 是可用的未用綱目。

如果您不確定要輸入哪些值，請與 Db2 管理者聯絡。

#### LDAP 配置

Analytic Server 使用 LDAP 伺服器來儲存及鑑別使用者及群組。您在 Analytic Server 安裝期間需要提供必要的 LDAP 配置資訊。

表 10. LDAP 配置設定	
LDAP 設定	說明
as.ldap.type	LDAP 類型。值可以是 ads、ad 或 openldap。 <ul style="list-style-type: none"> <li>· ads - Apache Directory Server（預設值）</li> <li>· ad - Microsoft Active Directory</li> <li>· openldap - OpenLDAP</li> </ul>
as.ldap.host	LDAP 主機
as.ldap.port	LDAP 埠號

表 10. LDAP 配置設定 (繼續)	
LDAP 設定	說明
as.ldap.binddn	LDAP 連結 DN
as.ldap.bindpassword	LDAP 連結 DN 密碼
as.ldap.basedn	LDAP 基本 DN
as.ldap.filter	LDAP 使用者和群組過濾器規則 註：當此值包含垂直線   字元時，字元必須使用反斜線 (例如，\ ) 跳出。
as.ldap.ssl.enabled	指定是否使用 SSL 在 Analytic Server 與 LDAP 之間通訊。值可以是 true 或 false。
as.ldap.ssl.reference	LDAP SSL 參照 ID
as.ldap.ssl.content	LDAP SSL 配置

- 依預設，as.ldap.type 設為 ads，且其他相關設定包含預設值。例外情況是您必須提供 as.ldap.bindpassword 設定的密碼。Analytic Server 使用配置設定來安裝 Apache Directory Server (ADS) 並執行伺服器起始設定。預設 ADS 設定檔包括使用者 admin，其密碼為 admin。您可以透過 Analytic Server 主控台進行使用者管理，或者透過位於 <Analytic Root>/bin 資料夾中的 importUser.sh Script，從 XML 檔匯入使用者和群組資訊。
- 如果您計劃使用外部 LDAP 伺服器（例如，Microsoft Active Directory 或 OpenLDAP），您必須根據實際 LDAP 值定義配置設定。如需相關資訊，請參閱在 Liberty 中配置 LDAP 使用者登錄。
- 安裝 Analytic Server 之後，您可以變更 LDAP 配置（例如，從 Apache Directory Server 變更為 OpenLDAP）。但是，如果起始使用 Microsoft Active Directory 或 OpenLDAP 啟動，並決定稍後切換至 Apache Directory Server，則 Analytic Server 在安裝期間將不會安裝 Apache Directory Server。僅當起始 Analytic Server 安裝期間選取了 Apache Directory Server 時，才會安裝它。

LDAP type as ldap.type	Analytic Server Default Group <input type="radio"/> openldap <input type="radio"/> ad <input checked="" type="radio"/> ads	?
LDAP host as ldap.host	Analytic Server Default Group <input type="text"/>	?
Bind DN as ldap.binddn	Analytic Server Default Group <input type="text" value="uid=admin,ou=system"/>	?
Bind password as ldap.bindpassword	Analytic Server Default Group <input type="text"/>	?
Base DN as ldap.basedn	Analytic Server Default Group <input type="text" value="dc=ibm,dc=com"/>	?
Enable SSL as ldap.ssl.enabled	<input checked="" type="checkbox"/> Analytic Server Default Group	?
SSL settings id as ldap.ssl.reference	Analytic Server Default Group <input type="text" value="LDAPSSLSettings"/>	?
SSL configuration as ldap.ssl.content	Analytic Server Default Group <input type="text" value="&lt;ssl id='LDAPSSLSettings' keyStoreRef='LDAPTrustStore' trustStoreRef='LDAPTrustStore' /&gt; &lt;-keyStore id='LDAPTrustStore' location='/opt,'"/>	?
LDAP user and group filter as ldap.filter	Analytic Server Default Group <input type="text" value="&lt;customFilters id='customFilters' userFilter='(&amp;amp;(cn=%v)(objectClass=organizationalPerson))' groupFilter='(&amp;amp;(cn=%v)(objectClass='"/>	?
LDAP Port as ldap.port	Analytic Server Default Group <input type="text" value="10636"/>	?

圖 5. LDAP 配置設定範例

8. 在啟用了 Kerberos 的 Cloudera 環境中安裝 Analytic Server 時，必須還在「新增服務精靈」中配置下列設定：

註：Analytic Server 會檢查 HDFS 配置，以使用 Kerberos 相關值進行鑑別。

- 如果想要在登入 Analytic Server 主控台時啟用 Kerberos 鑑別，請選取 Kerberos 作為 **Analytic Server 安全設定**。選取 **Kerberos** 作為 **Analytic Server 安全設定** 時，Analytic Server 主控台預設為 Kerberos 登入模式。
- 當您想要連接至啟用了 Kerberos 的資料庫時，請選取 Kerberos 作為 **Analytic Server 資料庫資料來源連線方法**。選取 **Kerberos** 作為 **Analytic Server 資料庫資料來源連線方法** 設定時，Analytic Server 主控台會使用 Kerberos 模式連接至資料庫。
- 需要 **Kerberos 領域名稱** 及 **KDC 主機** 設定。**Kerberos 領域名稱 (as.kdc.realms)** 和 **KDC 主機 (kdcserver)** 值位於 Kerberos Key Distribution Center (KDC) 伺服器上的 `krb5.conf` 檔中。

支援多個領域名稱（以逗點字元區隔）。指定的 Kerberos 領域名稱對應於使用者名稱並與之相關聯。例如，使用者名稱 `UserOne@us.ibm.com` 及 `UserTwo@eu.ibm.com` 對應於領域 `us.ibm.com`, `eu.ibm.com`。

在 **Kerberos 領域名稱** 中指定多個領域時，必須配置 Kerberos 交叉領域信任。在 Analytic Server 主控台登入提示期間輸入的使用者名稱是沒有領域名稱字尾的。因此，指定多個領域時，系統會向使用者呈現一個 **領域** 下拉清單，以讓使用者選取領域。

註：如果僅指定一個領域，則在使用者登入 Analytic Server 時，不會向使用者呈現 **領域** 下拉清單。

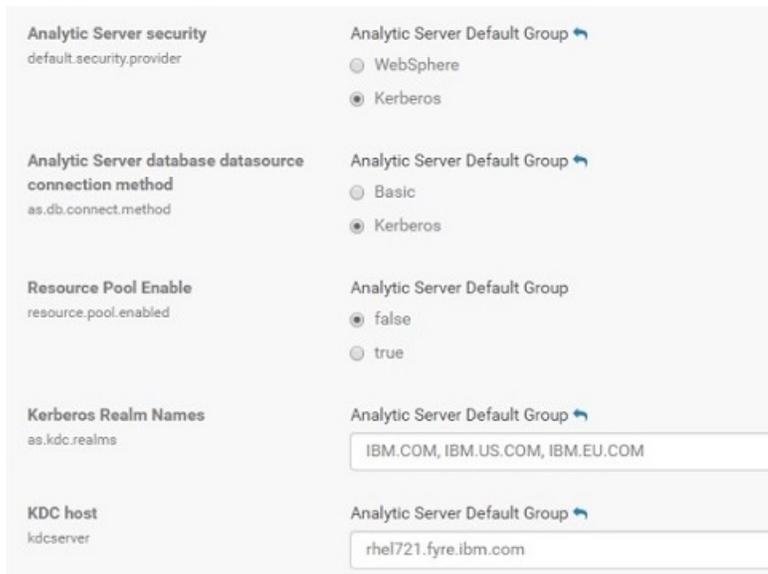


圖 6. Kerberos 設定範例

**附註：**

- **Analytic Server 安全及 Analytic Server 資料庫資料來源連線方法**設定適用於 IBM SPSS Modeler 用戶端及 Analytic Server 主控台鑑別。
- 當 **Analytic Server 資料庫資料來源連線方法**設為 Kerberos 時，您必須確保目標資料庫也已啟用 Kerberos。
- **Analytic Server 安全及 Analytic Server 資料庫資料來源連線方法**設定不會在 Hadoop 叢集上配置 Kerberos 鑑別。如需相關資訊，請參閱「啟用 Kerberos 模擬」小節。
- 如果您要在登入時啟用 Kerberos 鑑別，則必須將 IBM SPSS Modeler 用戶端部署為有效的 Kerberos 用戶端。使用「Kerberos 金鑰配送中心 (KDC)」伺服器中的 **addprinc** 指令即可達成。如需相關資訊，請參閱 IBM SPSS Modeler 說明文件。

在啟用了 Kerberos 的 Cloudera 環境中安裝 Analytic Server 時，您必須還在 Kerberos 中建立必要帳戶並啟用 Kerberos 模擬。如需詳細資料，請參閱第 45 頁的『配置 Kerberos』。



**警告：**順利安裝 Analytic Server 之後，請不要按一下 Cloudera Manager 中 Analytic Server 服務頁面之「動作」清單中的**建立 Analytic Server Metastore**。建立 meta 儲存庫會改寫現有 meta 資料儲存庫。

**離線安裝**

離線安裝步驟與線上步驟相同，只是您必須手動下載適用於您特定作業系統的包裹檔案及 meta 資料。

RedHat Linux 需要下列檔案：

- [AnalyticServer-3.2.2.0-el7.parcel](#)
- [AnalyticServer-3.2.2.0-el7.parcel.sha](#)
- [manifest.json](#)

SuSE Linux 需要下列檔案：

- [AnalyticServer-3.2.2.0-sles12.parcel](#)
- [AnalyticServer-3.2.2.0-sles12.parcel.sha](#)
- [manifest.json](#)

Ubuntu Linux 16.04 需要下列檔案：

- [AnalyticServer-3.2.2.0-xenial.parcel](#)
- [AnalyticServer-3.2.2.0-xenial.parcel.sha](#)

Ubuntu Linux 18 需要下列檔案：

- [AnalyticServer-3.2.2.0-bionic.parcel](#)
- [AnalyticServer-3.2.2.0-bionic.parcel.sha](#)

1. 在 Cloudera Manager 主要叢集節點上，下載並執行 Cloudera 自行解壓縮的 \*.bin 安裝程式。遵循安裝提示，接受授權合約，並保留預設 CSD 安裝目錄。

註：如果該目錄與預設位置不同，您必須指定不同的 CSD 目錄。

2. 將必要的包裹和 meta 資料檔案複製到 Cloudera Manager 主要叢集節點上的本端 Cloudera repo 路徑。預設路徑為 /opt/cloudera/parcel-repo（該路徑可在 Cloudera Manager 使用者介面中進行配置）。

3. 使用下列指令來重新啟動 Cloudera Manager：

```
service cloudera-scm-server restart
```

在 Cloudera Manager 重新整理 **AnalyticServer** 包裹之後，該包裹會顯示為**已下載**。您可以按一下**檢查新包裹**，以強制重新整理。

4. 按一下**配送 > 啟動**。

**AnalyticServer** 包裹會顯示為已配送及已啟動。

5. 在 Cloudera Manager 中，新增 Analytic Server 作為服務。請參閱「線上安裝」小節中的步驟 7 和 8，以取得相關資訊。

## 配置 Cloudera

安裝之後，您必須在叢集作業系統上建立必要帳戶。

1. 為所有使用者建立作業系統使用者帳戶，您計劃提供對每一個及每個 Analytic Server 及 Hadoop 節點上 Analytic Server 的存取權（這些使用者也配置為 LDAP 使用者登錄）。使用者群組必須設為 hadoop。

· 請確保這些使用者的 UID 在所有機器上都相符。您可以使用 **kinit** 指令登入每一個帳戶來測試此項目。

· 請確保 UID 遵循用於提交工作的使用者 ID 下限 YARN 設定。這是 container-executor.cfg 中的 **min.user.id** 參數。例如，如果 **min.user.id** 為 1000，則建立的每個使用者帳戶都必須具有大於或等於 1000 的 UID。

2. 針對 Analytic Server 管理者使用者，在 HDFS 上建立使用者起始檔案夾。該資料夾許可權必須設為 755，擁有者必須定義為 admin，且使用者群組必須設為 hdfs。請參閱下列**粗體**範例：

```
[root@xxxx configuration]# hadoop fs -ls /user
drwxr-xr-x - admin hdfs 0 2017-06-08 01:33 /user/admin
drwxr-xr-x - user1 hdfs 0 2017-06-06 01:00 /user/user1
```

3. 在 HDFS 上為所有 Analytic Server 標準使用者（例如，user1）建立使用者起始檔案夾。資料夾擁有者是實際使用者，並且使用者群組必須設為 hdfs。

安裝之後，您可以選擇性地透過 Cloudera Manager 來配置及管理 Analytic Server。

註：下列使用慣例用於 Analytic Server 檔案路徑。

· {AS\_ROOT} 指的是 Analytic Server 部署所在的位置；例如，/opt/cloudera/parcels/AnalyticServer。

· {AS\_SERVER\_ROOT} 指的是配置、日誌及伺服器檔案的位置；例如，/opt/cloudera/parcels/AnalyticServer/ae\_wlpserver/usr/servers/aeserver。

· {AS\_HOME} 指的是 HDFS 上由 Analytic Server 用作根資料夾的位置；例如，/user/as\_user/analytic-root。

## 安全

IBM SPSS Modeler `options.cfg` 檔案中的預設 `tenant_id` 值為 `ibm`。您可以在 Analytic Server 主控台中檢視「租戶」。如需有關承租人管理的詳細資料，請參閱《IBM SPSS Analytic Server 管理手冊》。

### 配置 LDAP 登錄

Analytic Server 安裝期間會配置 LDAP。在 Analytic Server 安裝之後，您可以變更為另一種 LDAP 伺服器方法。

註：在 Analytic Server 中的 LDAP 支援由 WebSphere Liberty 進行控制。如需相關資訊，請參閱在 Liberty 中配置 LDAP 使用者登錄。

### 配置從 Analytic Server 到 LDAP 的 Secure Socket Layer (SSL) 連線

1. 以 Analytic Server 使用者的身分登入每一個 Analytic Server 機器，並為 SSL 憑證建立共用目錄。

註：在 Cloudera 上，Analytic Server 使用者一律是 `as_user`，而且無法變更。

2. 將金鑰儲存庫和信任儲存庫檔案複製到所有 Analytic Server 機器上的部分共用目錄。另將 LDAP 用戶端 CA 憑證新增至信任儲存庫。以下是部分範例指示。

```
mkdir -p /home/as_user/security
cd /home/as_user/security
openssl s_client -connect <ldap-hostname>:636 -showcerts > client.cert
$JAVA_HOME/bin/keytool -import -file ./client.cert -alias ldapCA -keystore
mytrust.jks
password : changeit
```

註：`JAVA_HOME` 是與用於 Analytic Server 啟動相同的 JRE。

3. 可以使用 `securityUtility` 工具來編碼密碼以模糊化其值，該工具位於 `{AS_ROOT}/ae_wlpserver/bin` 中。範例如下所示。

```
securityUtility encode changeit
{xor}Pdc+MTg6Nis=
```

4. 登入 Cloudera Manager，並將 Analytic Server 配置設定 `ssl_cfg` 更新為正確的 SSL 配置設定。範例如下所示。

```
<ssl id="defaultSSLConfig" keyStoreRef="defaultKeyStore" trustStoreRef="defaultTrustStore"
  clientAuthenticationSupported="true"/>
  <keyStore id="defaultKeyStore" location="/home/as_user/security/mykey.jks"
type="JKS"
  password="{xor}0zo5PiozKxYdEgwPDAweDG1uDz4sLCg7"/>
  <keyStore id="defaultTrustStore" location="/home/as_user/security/mytrust.jks"
type="JKS"
  password="{xor}Pdc+MTg6Nis="/>
```

註：對金鑰和信任儲存庫檔案使用絕對路徑。

5. 將 Analytic Server 配置設定 `security_cfg` 更新為正確的 LDAP 配置設定。例如，在 `ldapRegistry` 元素中，將 `sslEnabled` 屬性設為 `true`，並將 `sslRef` 屬性設為 `defaultSSLConfig`。

### 配置 Kerberos

Analytic Server 在 Cloudera 中支援 Kerberos。下列各節提供了一些配置設定，以確保使用與 Analytic Server 相容的方式正確配置 Kerberos。

註：Analytic Server 會檢查 HDFS 配置，以使用 Kerberos 相關值進行鑑別。

### Analytic Server 及 Kerberos 設定

在啟用了 Kerberos 的 Cloudera 環境中安裝 Analytic Server 時，請記住下列設定。

- 如果想要在登入 Analytic Server 主控台時啟用 Kerberos 鑑別，請選取 Kerberos 作為 **Analytic Server 安全設定**。選取 **Kerberos** 作為 **Analytic Server 安全設定** 時，Analytic Server 主控台預設為 Kerberos 登入模式。

- 當您想要連接至啟用了 Kerberos 的資料庫時，請選取 Kerberos 作為 **Analytic Server 資料庫資料來源連線方法**。選取 **Kerberos** 作為 **Analytic Server 資料庫資料來源連線方法** 設定時，Analytic Server 主控台會使用 Kerberos 模式連接至資料庫
- 需要 **Kerberos 領域名稱** 及 **KDC 主機** 設定。**Kerberos 領域名稱 (as.kdc.realms)** 和 **KDC 主機 (kdcserver)** 值位於 Kerberos Key Distribution Center (KDC) 伺服器上的 krb5.conf 檔中。

支援多個領域名稱（以逗點字元區隔）。指定的 Kerberos 領域名稱對應於使用者名稱並與之相關聯。例如，使用者名稱 UserOne@us.ibm.com 及 UserTwo@eu.ibm.com 對應於領域 us.ibm.com, eu.ibm.com。

在 **Kerberos 領域名稱** 中指定多個領域時，必須配置 Kerberos 交叉領域信任。在 Analytic Server 主控台登入提示期間輸入的使用者名稱是沒有領域名稱字尾的。因此，指定多個領域時，系統會向使用者呈現一個 **領域** 下拉清單，以讓使用者選取領域。

**註：**如果僅指定一個領域，則在使用者登入 Analytic Server 時，不會向使用者呈現 **領域** 下拉清單。

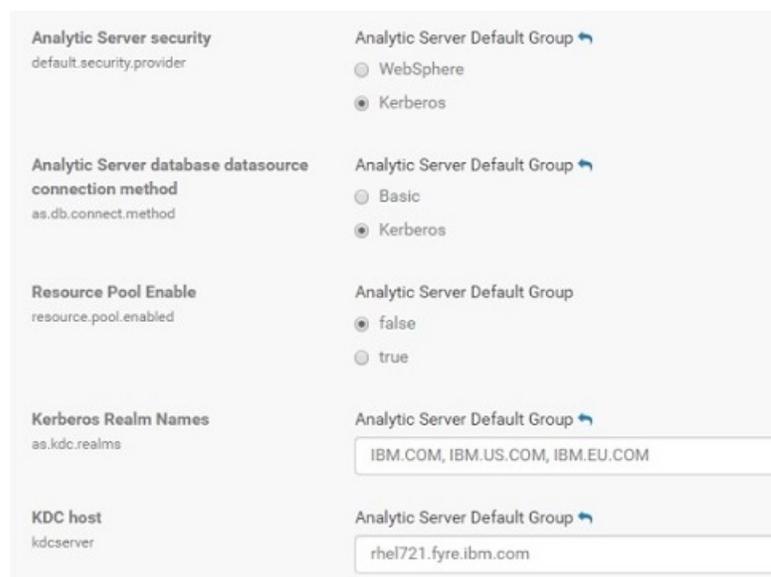


圖 7. Kerberos 設定範例

**附註：**

- **Analytic Server 安全及 Analytic Server 資料庫資料來源連線方法** 設定適用於 IBM SPSS Modeler 用戶端及 Analytic Server 主控台鑑別。
- 當 **Analytic Server 資料庫資料來源連線方法** 設為 Kerberos 時，您必須確保目標資料庫也已啟用 Kerberos。
- **Analytic Server 安全及 Analytic Server 資料庫資料來源連線方法** 設定不會在 Hadoop 叢集上配置 Kerberos 鑑別。如需相關資訊，請參閱「啟用 Kerberos 模擬」小節。
- 如果您要在登入時啟用 Kerberos 鑑別，則必須將 IBM SPSS Modeler 用戶端部署為有效的 Kerberos 用戶端。使用「Kerberos 金鑰配送中心 (KDC)」伺服器中的 **addprinc** 指令即可達成。如需相關資訊，請參閱 IBM SPSS Modeler 說明文件。

### 在 Kerberos 中建立必要的帳戶

1. 在 Kerberos 使用者儲存庫中，為您打算授與 Analytic Server 之存取權的所有使用者建立帳戶。
2. 在 LDAP 伺服器上建立相同的帳戶（依前一個步驟）。
3. 在每個 Analytic Server 節點及 Hadoop 節點上，為在前一個步驟中所建立的每一位使用者，各建立一個作業系統使用者帳戶。使用者群組必須設為 hadoop。
  - 請確保這些使用者的 UID 在所有機器上都相符。您可以使用 kinit 指令登入每一個帳戶來測試此項目。

· 請確保 UID 遵循用於提交工作的使用者 ID 下限 YARN 設定。這是 `container-executor.cfg` 中的 `min.user.id` 設定。例如，如果 `min.user.id` 為 1000，則建立的每個使用者帳戶都必須具有大於或等於 1000 的 UID。

4. 針對 Analytic Server 管理者使用者，在 HDFS 上建立使用者起始檔案夾。該資料夾許可權必須設為 755，擁有者必須定義為 `admin`，且使用者群組必須設為 `hdfs`。請參閱下列粗體範例：

```
[root@xxxxx configuration]# hadoop fs -ls /user
Found 9 items
drwxrwxrwx - hdfs supergroup 0 2017-07-26 03:41 /user/AE
drwxr-xr-x - admin hdfs 0 2017-06-08 01:33 /user/admin
drwxr-xr-x - as_user hdfs 0 2017-06-06 01:00 /user/as_user
drwx----- - hdfs supergroup 0 2017-07-31 00:17 /user/hdfs
drwxrwxrwx - mapred hadoop 0 2017-06-05 00:28 /user/history
drwxrwxr-t - hive hive 0 2017-06-05 00:30 /user/hive
drwxrwxr-x - hue hue 0 2017-06-05 00:30 /user/hue
drwxrwxr-x - impala impala 0 2017-07-19 00:52 /user/impala
drwxr-x--x - spark spark 0 2017-06-05 01:34 /user/spark
```

5. 在 HDFS 上為所有 Analytic Server 標準使用者（例如，`user1`）建立使用者起始檔案夾。資料夾擁有者是實際使用者，並且使用者群組必須設為 `hdfs`。
6. 如果您打算使用 HCatalog 資料來源，且 Analytic Server 安裝在與 Hive metastore 不同的機器上，則必須在 HDFS 上模擬 Hive 用戶端。
  - a. 在 Cloudera Manager 中導覽至 HDFS 服務的「配置」標籤。

註：如果尚未設定下列設定，則在配置標籤上，可能不會顯示這些設定。在此情況下，請執行搜尋來尋找它們。
  - b. 編輯 `hadoop.proxyuser.hive.groups` 設定，讓它的值為 `*`，或是包含獲允許登入 Analytic Server 之所有使用者的群組。
  - c. 編輯 `hadoop.proxyuser.hive.hosts` 設定，讓它的值為 `*`，或是 Hive Metastore 及每個 Analytic Server 實例在其上被安裝為服務的主機清單。
  - d. 重新啟動 HDFS 服務。

在執行了這些步驟並安裝了 Analytic Server 之後，Analytic Server 就會無聲自動配置 Kerberos。

## 啟用 Kerberos 模擬

模擬容許執行緒在與執行緒所屬程序之安全環境定義不同的安全環境定義中執行。例如，模擬可提供一種方式，讓 Hadoop 工作以標準 Analytic Server 使用者 (`as_user`) 以外的使用者身分執行。若要啟用 Kerberos 模擬，請執行下列動作：

1. 開啟 Cloudera Manager，並在 `core-site.xml` 的叢集層面進階配置 **Snippet**（安全閘）區域中（位於 **HDFS**（服務層面） > 配置標籤上），新增或更新下列內容。

- 名稱：`hadoop.proxyuser.as_user.hosts`
- 值：`*`
- 名稱：`hadoop.proxyuser.as_user.groups`
- 值：`*`

註：`core-site.xml` 設定適用於 Hadoop 配置（而非 Analytic Server）。

2. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

## 使用 Kerberos 為單一登入 (SSO) 配置 HAProxy

1. 按 HAProxy 說明文件手冊配置及啟動 HAProxy：<http://www.haproxy.org/#docs>
2. 為 HAProxy 主機建立 Kerberos 原則 (`HTTP/<proxyHostname>@<realm>`) 及 `keytab` 檔，其中 `<proxyHostname>` 是 HAProxy 主機的完整名稱，`<realm>` 是 Kerberos 領域。
3. 將 `keytab` 檔複製到每個 Analytic Server 主機作為 `/etc/security/keytabs/spnego_proxy.service.keytab`

4. 在每個 Analytic Server 主機上更新此檔案的許可權。範例如下所示。

```
chown root:hadoop /etc/security/keytabs/spnego_proxy.service.keytab
chmod 440 /etc/security/keytabs/spnego_proxy.service.keytab
```

5. 開啟 Cloudera Manager 並在 Analytic Server **analyticserver-conf/config.properties** 的 **Analytic Server 進階配置 Snippet (安全閘)** 區域中，新增或更新下列內容。

```
web.authentication.kerberos.keytab=/etc/security/keytabs/spnego_proxy.service.keytab
web.authentication.kerberos.principal=HTTP/<proxy machine full name>@<realm>
```

6. 儲存配置，並從 Cloudera Manager 重新啟動所有 Analytic Server 服務。
7. 指示使用者配置其瀏覽器，以使用 Kerberos。

使用者現在可以使用 IBM SPSS Analytic Server 登入畫面上的**單一登入**選項來登入 Analytic Server。

## 停用 Kerberos

1. 在 Cloudera Manager 主控台中停用 Kerberos。
2. 停止 Analytic Server 服務。
3. 在 **analyticserver-conf/config.properties** 的 **Analytic Server 進階配置 Snippet (安全閘)** 區域中修改下列設定：

**Analytic Server 安全 (default.security.provider) > WebSphere**

**Analytic Server 資料庫資料來源連線方法 (as.db.connect.method) > 基本**

4. 按一下**儲存變更**，然後重新啟動 Analytic Server 服務。

## 啟用與 Analytic Server 主控台的 Secure Socket Layer (SSL) 連線

依預設，Analytic Server 會產生自簽憑證以啟用 Secure Socket Layer (SSL)，以便您可以透過安全埠，藉由接受自簽憑證來存取 Analytic Server 主控台。為了能更安全地進行 HTTPS 存取，您需要安裝第三方供應商憑證。

## 安裝協力廠商供應商憑證

1. 將第三方供應商金鑰儲存庫和信任儲存庫憑證複製到所有 Analytic Server 節點中的相同目錄；例如，`/home/as_user/security`。

註：Analytic Server 使用者必須擁有此目錄的讀取權。

2. 在 Cloudera Manager 中，導覽至 Analytic Server 服務的「配置」標籤。
3. 編輯 **ssl\_cfg** 參數。

```
<ssl id="defaultSSLConfig"
  keyStoreRef="defaultKeyStore"
  trustStoreRef="defaultTrustStore"
  clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
  location="<KEYSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
<keyStore id="defaultTrustStore"
  location="<TRUSTSTORE-LOCATION>"
  type="<TYPE>"
  password="<PASSWORD>"/>
```

將

- `<KEYSTORE-LOCATION>` 取代為金鑰儲存庫的絕對位置；例如：`/home/as_user/security/mykey.jks`
- `<TRUSTSTORE-LOCATION>` 取代為信任儲存庫的絕對位置；例如：`/home/as_user/security/mytrust.jks`
- `<TYPE>` 取代為憑證的類型；例如：JKS、PKCS12 等。

- <PASSWORD> 取代為採用 Base64 加密格式的加密密碼。如果要進行編碼，您可以使用 securityUtility；例如：{AS\_ROOT}/ae\_wlpserver/bin/securityUtility encode <password>

如果要產生自簽憑證，可以使用 securityUtility；例如：{AS\_ROOT}/ae\_wlpserver/bin/securityUtility createSSLCertificate --server=aeserver --password=mypassword --validity=365 --subject=CN=myfqdnserver,O=myorg,C=mycountry。如需 securityUtility 及其他 SSL 設定的相關資訊，請參閱 WebSphere Liberty Profile 文件。

#### 附註：

- 您必須提供 CN 值的適當主機網域名稱。
- 將 **mypassword**、**myfqdnserver**、**myorg** 及 **mycountry** 取代為您的特定認證。請注意，**myfqdnserver** 是 Analytic Server 節點的完整網域名稱。
- **aeserver** 是 Liberty 伺服器名稱（該值必須是 **aeserver**）。

如需 **securityUtility** 及其他 SSL 設定的相關資訊，請參閱 [WebSphere Liberty Profile](#) 及 [securityUtility](#) 指令說明文件。

4. 按一下**儲存變更**，然後重新啟動 Analytic Server 服務。

### 產生自簽憑證

您可以使用 securityUtility 來產生自簽憑證。例如：

```
/opt/cloudera/parcels/AnalyticServer-3.2.2.0/ae_wlpserver/bin/securityUtility createSSLCertificate
--server=<myserver> --password=<mypassword> --validity=365 --subject=CN=<mycompany>,O=<myOrg>,C=<myCountry>
```

#### 附註：

- 您必須提供 **CN** 值的適當主機網域名稱。
- 將 key.jks 中的資訊複製到 trust.jks（兩個檔案必須相同）。
- 編輯 ssl.keystore.config 參數。例如：

```
<ssl id="defaultSSLConfig"
keyStoreRef="defaultKeyStore"
trustStoreRef="defaultTrustStore"
clientAuthenticationSupported="true"/>
<keyStore id="defaultKeyStore"
location="/opt/cloudera/parcels/AnalyticServer-3.2.2.0
/ae_wlpserver/usr/servers/aeserver/resources/security/key.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
<keyStore id="defaultTrustStore"
location="/opt/cloudera/parcels/AnalyticServer-3.2.2.0
/ae_wlpserver/usr/servers/aeserver/resources/security/trust.jks"
type="JKS"
password="{xor}Dz4sLG5tbGs="/>
```

### 透過 SSL 與 Apache Hive 通訊

您必須更新 hive.properties 檔案，才能透過 SSL 連線與 Apache Hive 通訊。或者，如果 Apache Hive 環境已啟用高可用性，您可以在主要 Analytic Server 資料來源頁面上選取高可用性參數。

#### 更新 hive.properties 檔案

1. 開啟 hive.properties 檔案。該檔案位於：/opt/cloudera/parcels/AnalyticServer/ae\_wlpserver/usr/servers/aeserver/configuration/database
2. 尋找下列行：

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password}
```

3. 新增下面的**粗體**資訊來更新行：

```
jdbcurl = jdbc:hive2://{db.servername}:{db.serverport}/{db.databasename};user={db.username};password={db.password}
;
ssl=true;sslTrustStore=pathtotheirtruststorefile;trustStorePassword=xxxtheirTrustStorePassword
```

4. 儲存 hive.properties 檔案。

## 啟用 Essentials for R 的支援

Analytic Server 支援對 R 模型評分以及執行 R Script。

若要在 Cloudera Manager 中順利安裝 Analytic Server 之後安裝 Essentials for R：

1. 供應 Essentials for R 的伺服器環境。如需相關資訊，請參閱第 21 頁的『啟用對 Essentials for R 的支援』中的步驟 1。
2. 下載 IBM SPSS Modeler Essentials for R RPM 的自行解壓縮保存檔 (BIN)。Essentials for R 可供下載 (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>)。選擇您的堆疊、堆疊版本及硬體架構特定的檔案。
3. 在 Cloudera Manager 伺服器主機上，以 root 或 sudo 使用者身分，執行自行解壓縮的保存檔。透過已配置的儲存庫，必須安裝或提供下列套件：
  - Red Hat Linux：gcc-gfortran、zip、gcc-c++
  - SUSE Linux：gcc-fortran、zip、gcc-c++
  - Ubuntu Linux：gcc-fortran、zip、gcc-c++
4. 自行解壓縮的安裝程式執行下列作業：
  - a. 顯示必要授權，並提示安裝程式接受它們。
  - b. 提示安裝程式輸入 R 來源位置，或者使用預設位置繼續。所安裝的預設 R 版本為 3.5.1。若要安裝不同的版本：
    - 線上安裝：提供所需 R 版本保存檔的 URL。例如，對於 R 3.4.4，為 <https://cran.r-project.org/src/base/R-3/R-3.4.4.tar.gz>。
    - 離線安裝：下載所需 R 版本保存檔，並將其複製到 Cloudera Manager 伺服器主機上。請勿重新命名保存檔（依預設，其會命名為 R-x.x.x.tar.gz）。請提供所複製 R 保存檔的 URL，如下所示：`file://<R_archive_directory>/R-x.x.x.tar.gz`。如果已下載 R-3.4.4.tar.gz 保存檔，並將其複製到 /root，則 URL 為 `file:///root/R-3.4.4.tar.gz`。

註：其他 R 版本可在 <https://cran.r-project.org/src/base/> 中找到。
  - c. 安裝 R 所需的套件。
  - d. 下載並安裝 R 及 Essentials for R 外掛程式。
  - e. 建立包裹及 `parcel.sha` 檔案，並將它們複製到 `/opt/cloudera/parcel-repo`。如果位置已變更，請輸入正確的位置。
5. 安裝完成之後，在 Cloudera Manager 中配送並啟動 **Essentials for R** 包裹（按一下**檢查新包裹**，以重新整理包裹清單）。
6. 如果已安裝 Analytic Server 服務：
  - a. 停止該服務。
  - b. 重新整理 Analytic Server 二進位檔。
  - c. 啟動該服務，以完成 Essentials for R 安裝。
7. 如果未安裝 Analytic Server 服務，則繼續其安裝。

註：所有 Analytic Server 主機都必須已安裝適當的保存檔（zip 及 unzip）套件。

## 啟用關聯式資料庫來源

如果您在 Analytic Server meta 儲存庫及每一個 Analytic Server 主機上的共用目錄中提供 JDBC 驅動程式，則 Analytic Server 可以使用關聯式資料庫來源。依預設，此目錄為 `/usr/share/jdbc`。

若要變更共用目錄，請遵循下列步驟。

1. 在 Cloudera Manager 中，導覽至 Analytic Server 服務的「配置」標籤。
2. 在 **jdbc.drivers.location** 中指定 JDBC 驅動程式之共用目錄的路徑。
3. 按一下**儲存變更**。
4. 從**動作**下拉清單中選取**停止**，以停止 Analytic Server 服務。

5. 從**動作**下拉清單中選取**重新整理 Analytic Server 二進位檔**。

6. 從**動作**下拉清單中選取**啟動**，以啟動 Analytic Server 服務。

表 11. 支援的資料庫			
資料庫	支援的版本	JDBC 驅動程式 JAR	供應商
Amazon Redshift	8.0.2 或更新版本	RedshiftJDBC41-1.1.6.1006.jar 或更新版本	Amazon
Apache Impala	使用 2.5.5 或更新版本的 JDBC 4	ImpalaJDBC4.jar、commons-codec-*.jar、commons-logging-*.jar、httpclient-*.jar、httpcore-*.jar、log4j-*.jar、libthrift-*.jar、libfb303-*.jar、slf4j-api-*.jar、ql.jar、zookeeper-*.jar、TCLIServiceClient.jar	Apache
DashDB	Bluemix 服務	db2jcc.jar	IBM
Db2 for Linux, UNIX, and Windows	11.1、10.5、10.1、9.7	db2jcc.jar	IBM
Db2 z/OS	11、10	db2jcc.jar、db2_license_cisuz.jar	IBM
Greenplum	5.x	postgresql.jar	Greenplum
Hive	1.1	hive-jdbc-*.jar	Apache
MySQL	5.7、5.6	mysql-connector-java-commercial-5.1.25-bin.jar	MySQL
Netezza	7、6.x	nzjdbc.jar	IBM
Oracle	19c、12c、11g R2 (11.2)	19c : ojdbc8.jar, orai18n.jar 12c 及 11g R2 (11.2) : ojdbc6.jar、 orai18n.jar	Oracle
SQL Server	2017、2014、2012、2008 R2	sqljdbc4.jar	Microsoft
Teradata	15.1、15	tdgssconfig.jar、 terajdbc4.jar	Teradata

## 附註

· 如果您在安裝 Analytic Server 之前建立了 Redshift 資料來源，則需要執行下列步驟，以使用 Redshift 資料來源。

1. 在 Analytic Server 主控台中，開啟 Redshift 資料來源。
2. 選取 Redshift 資料庫資料來源。
3. 輸入 Redshift 伺服器位址。
4. 輸入資料庫名稱及使用者名稱。應該會自動移入密碼。
5. 選取資料庫表格。

## 啟用 HCatalog 資料來源

Analytic Server 透過 Hive/HCatalog 提供對許多資料來源的支援。部分來源需要手動配置步驟。

1. 收集必要的 JAR 檔以啟用資料來源。如需詳細資料，請參閱下列各節。
2. 將這些 JAR 檔新增至 Analytic Server meta 儲存庫及每一個 Analytic Server 節點上的 {HIVE\_HOME}/auxlib 目錄及 /usr/share/hive 目錄。
3. 重新啟動 Hive Metastore 服務。
4. 重新啟動每一個 Analytic Server 服務實例。

### 註：

透過 Analytic Server HCatalog 資料來源存取 HBase 資料時，存取使用者必須具有 HBase 表格的讀取權。

- 在非 Kerberos 環境中，Analytic Server 使用 `as_user` 存取 HBase (`as_user` 必須具有 HBase 的讀取權)。
- 在 kerberos 環境中，`as_user` 及登入使用者都必須具有 HBase 表格的讀取權。

## NoSQL 資料庫

Analytic Server 支援任何有來自供應商的 Hive 儲存體處理程式可用的 NoSQL 資料庫。

啟用 Apache HBase 和 Apache Accumulo 的支援不需要其他的步驟。

如果是其他 NoSQL 資料庫，請與資料庫供應商聯絡，取得儲存體處理程式及相關的 JAR。

## 檔案型 Hive 表格

Analytic Server 支援任何有內建或自訂 Hive SerDe (序列化程式-解除序列化程式) 可用的檔案型 Hive 表格。

用於處理 XML 檔的 Hive XML SerDe 位於 Maven Central Repository，網址為 <http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>。

## MapReduce 第 2 版工作

使用 `analyticserver-conf/config.properties` 的 Analytic Server 進階配置 Snippet (安全閥) 區域中的 `preferred.mapreduce` 設定可控制如何處理 MapReduce 工作：

內容	說明
<code>preferred.mapreduce</code>	控制執行 MapReduce 工作的方法。有效值包括： <ul style="list-style-type: none"><li>· spark</li><li>· m3r</li><li>· hadoop</li></ul> 例如： <code>preferred.mapreduce=spark</code>

## Apache Spark

如果您要使用 Spark (2.x 或更新版本)，您必須在安裝 Analytic Server 期間選取 spark.version。

1. 開啟 Cloudera Manager 並在 **Analytic Server Spark 版本** 區域中選取適當的 spark.version (例如，None 或 2.x)。
2. 儲存配置。

## 配置 Apache Impala

在 Cloudera 上針對 Analytic Server 資料庫資料來源或 HCatalog 資料來源執行時，支援 Apache Impala (無論 Impala 是否啟用 SSL)。

### 為 Apache Impala 資料建立資料庫資料來源

1. 在主要 Analytic Server 的資料來源頁面上，按一下**新建**以建立新的資料來源。即會顯示「新資料來源」對話框。
2. 在**新資料來源**欄位中輸入適當的名稱，選取 Database 作為**內容類型**值，然後按一下**確定**。
3. 開啟**資料庫選項**區段，輸入下列資訊。

**資料庫：**

從下拉功能表選取 **Impala**。

**伺服器位址：**

輸入管理 Impala 常駐程式的伺服器 URL。當針對 Analytic Server 啟用 Kerberos 時，需要完整網域名稱。

**伺服器埠：**

輸入 Impala 資料庫接聽所在的埠號。

**資料庫名稱：**

輸入您要連接的資料庫名稱。

**使用者名稱：**

輸入有權登入 Impala 資料庫的使用者名稱。

**密碼：**

輸入適當的使用者名稱密碼。

**表格名稱：**

輸入來自您要使用之資料庫中的表格名稱。按一下**選取**以手動選取檔案。

**並行讀取數目上限：**

輸入可從 Analytic Server 傳送至資料庫，以讀取資料來源中的指定表格的平行查詢數目限制。

4. 輸入所需資訊之後按一下**儲存**。

### 為 Apache Impala 資料建立 HCatalog 資料來源

1. 在主要 Analytic Server 的資料來源頁面上，按一下**新建**以建立新的資料來源。即會顯示「新資料來源」對話框。
2. 在**新資料來源**欄位中輸入適當的名稱，選取 HCatalog 作為**內容類型**值，然後按一下**確定**。
3. 開啟**資料庫選項**區段，輸入下列資訊。

**資料庫：**

從下拉功能表選取**預設值**。

**表格名稱：**

輸入來自您要使用之資料庫中的表格名稱。

**HCatalog 綱目**

選取 **HCatalog** 元素選項，然後選取適當的 **HCatalog** 欄位對映選項。

4. 輸入所需資訊之後按一下**儲存**。

## 連接啟用了 SSL 的 Apache Impala 資料

1. 在 Cloudera Manager 主控台中定義下列 Impala SSL 設定。

### 為 Impala 啟用 TLS/SSL (client\_services\_ssl\_enabled)

選取 **Impala (服務範圍)** 選項。

### Impala TLS/SSL 伺服器憑證檔案 (PEM 格式) (ssl\_server\_certificate)

輸入 PEM 格式的自簽憑證位置及檔名 (例如: /tmp/<user\_name>/ssl/114200v21.crt)。

### Impala TLS/SSL 伺服器私密金鑰檔案 (PEM 格式) (ssl\_private\_key)

輸入 PEM 格式的私密金鑰位置及檔名 (例如: /tmp/<user\_name>/ssl/114200v21.key)。

2. 在 Analytic Server 主機上，將 \*.crf 檔 (用來啟用 Impala SSL) 匯入 \*.jks 檔。該檔案可以是 cacert 檔案 (例如， /etc/pki/java/cacerts) 或任何其他 \*.jks 檔。
3. 在 Analytic Server 主機上，更新 Impala 配置檔 (impala.properties)，即附加下列 jdbcurl 金鑰值：

```
SSL=1;AllowSelfSignedCerts=1;CAIssuedCertNamesMismatch=1;
```

註：使用 \*.jks 檔 (cacert 檔案除外) 時，您還需要指定下列內容：

```
SSLTrustStore=<your_pks_file>;SSLTrustStorePwd=<password_for_pks_file>;
```

4. 在 Cloudera Manager 主控台中重新啟動 Analytic Server。

## 變更 Analytic Server 所用的埠

依預設，Analytic Server 對 HTTP 使用 9080 埠，並對 HTTPS 使用 9443 埠。若要變更埠設定，請遵循下列步驟。

1. 在 Cloudera Manager 中，導覽至 Analytic Server 服務的「配置」標籤。
2. 分別在 **http.port** 及 **https.port** 參數中指定所要的 HTTP 和 HTTPS 埠。

註：您可能需要在「過濾器」區段中選取埠及位址種類，以查看這些參數。

3. 按一下**儲存變更**。
4. 重新啟動 Analytic Server 服務。

## 高可用性 Analytic Server

您可以使 Analytic Server 高度可用，方法為將它作為服務新增至叢集中的多個節點。

1. 在 Cloudera Manager 中，導覽至 Analytic Server 服務的「實例」標籤。
2. 按一下**新增角色實例**，並選取要將 Analytic Server 新增為服務的主機。

### 多叢集支援

多叢集特性是 IBM SPSS Analytic Server 的高可用性功能的加強功能，並在多租戶環境中提供改良的隔離。依預設，安裝 Analytic Server 服務 (在 Ambari 或 Cloudera Manager 中) 會定義單一分析伺服器叢集。

叢集規格會定義 Analytic Server 叢集成員資格。使用 XML 內容達成修改叢集規格 (在 Ambari Analytic Server 配置的 analytics-cluster 欄位中，或透過編輯 Cloudera Manager 的 configuration/analytics-cluster.xml 檔)。當配置多個 Analytic Server 叢集時，需要使用其自己的負載平衡器將要求饋送至每一個 Analytic Server 叢集。

使用多叢集特性會確保一個租戶的工作不能對另一個租戶叢集中執行的工作產生負面影響。對於高可用性工作，僅在起始工作所在的 Analytic Server 叢集範圍內發生工作失效接手。下列範例提供多叢集 XML 規格。

註：可以透過將 Analytic Server 作為服務新增至叢集中的多個節點，讓其成為高度可用。

```
<analyticServerClusterSpec>
  <cardinality>1+</cardinality>
  <cluster name="cluster1">
    <memberName>one.cluster</memberName>
    <memberName>two.cluster</memberName>
  </cluster>
  <cluster name="cluster2">
    <memberName>three.cluster</memberName>
```

```
<memberName>four.cluster</memberName>
</cluster>
</analyticServerClusterSpec>
```

在前一個範例中，需要兩個負載平衡器。一個負載平衡器將要求傳送至 `cluster1` 成員 (`one.cluster` 及 `two.cluster`)，另一個將要求傳送至 `cluster2` 成員 (`three.cluster` 及 `four.cluster`)。

下列範例提供單一叢集 XML 規格 (預設配置)。

```
<analyticServerClusterSpec>
  <cardinality>1</cardinality>
  <cluster name="cluster1">
    <memberName>*</memberName>
  </cluster>
</analyticServerClusterSpec>
```

在前一個範例中，需要單一負載平衡器來處理已配置多個叢集成員的情況。

## 附註

- 只有單態叢集支援在 `memberName` 元素中使用萬用字元 (例如，叢集基數 = "1")。基數元素的有效值是 1 與 1+。
- `memberName` 的指定方式，必須與為其指派 Analytic Server 角色之主機名稱的指定方式相同。
- 套用叢集配置變更之後，必須重新啟動叢集中的所有伺服器。
- 在 Cloudera Manager 中，您必須在所有 Analytic Server 節點上修改維護 `analytics-cluster.xml` 檔。必須維護所有節點，以確保它們包含相同的內容。

## 升級 Python - CDH

本節說明從 Python 2.x 手動升級至 Python 3.7 的處理程序

1. 在每一個叢集節點上安裝 Python 3.7。如需相關資訊，請參閱 [Python 網站](#)。
2. 在每一個叢集節點上安裝 NumPy。如需相關資訊，請參閱 [NumPy 安裝指示](#)。
3. 在每一個叢集節點上安裝 pandas。如需相關資訊，請參閱 [pandas 安裝指示](#)。
4. 在 Cloudera Manager 中，更新 `analyticserver-conf/config.properties` 的 **Analytic Server 進階配置 Snippet (安全閥)** 區段，以包括 Python 3.7 可執行路徑。例如：

```
spark.driver.python=/opt/python3/bin/python3.7
```

## 最佳化小型資料的 JVM 選項

您可以編輯 JVM 內容，以最佳化執行小型 (M3R) 工作時的系統。

在 Cloudera Manager 中，查看 Analytic Server 服務中「配置」標籤上的 **Jvm 選項 (jvm.options)** 控制項。修改下列參數，會設定伺服器 (用來管理 Analytic Server) 上所執行工作的資料堆大小；亦即，不是 Hadoop。這對於執行較小 (M3R) 工作而言非常重要，而且可能需要使用這些值進行試驗以最佳化系統。

```
-Xms512M
-Xmx2048M
```

## 為每一個 YARN 資源儲存區配置個別動態資源配置 - Cloudera

您可以為每一個 YARN 資源儲存區配置個別「動態資源配置」。

### 使用者及租戶模式對映 - Cloudera

使用者及租戶可以提交至不同的 YARN 資源儲存區，並且每一個使用者或租戶都會對映至不同的 YARN 資源儲存區 (以利用「動態資源配置」)。可以定義 **user** 模式或 **tenant** 模式，以對映至 YARN 資源儲存區。在 Analytic Server 3.2.1 Fix Pack 1 之前，所有 Spark 工作都限制為單一 YARN 資源儲存區。

從 IBM SPSS Analytic Server 3.2.1 Fix Pack 1 開始，當使用者/租戶的串流導致 Spark 工作在系統上執行時，個別 YARN 資源儲存區將以提交串流至 Analytic Server 的使用者/租戶身分執行。可以為不同的使用者/租戶作業同時執行多個 YARN 資源儲存區。

只要使用者登入 Analytic Server（有時，在使用者登出且沒有其他作用中使用者工作之後），每一個 YARN 資源儲存區就會繼續執行。配置變數可以控制登出之後的時間量：**as.spark.driver.cleanup.delay**。

為提交 Spark 工作的每一個使用者建立 **SparkDriver** 處理程序。在使用者沒有作用中工作約 2 分鐘（預設值）且沒有 **HTTPSession** 活動之後，每一個使用者的 **SparkDriver** 處理程序都會終止。

註：Analytic Server 關閉時，所有 **SparkDriver** 處理程序都會終止。

使用下列步驟，以將 Analytic Server 新增至現有叢集：

1. 在 Cloudera Manager 中，導覽至 **SPSS Analytic Server 服務 > 配置**。
2. 將資源儲存區啟用：**resource.pool.enabled** 值變更為 **true**。
3. 將下列內容新增至 **Analytic Server 進階配置 Snippet（安全閘） > analyticsserver-conf.config.properties**：

```
//Using user/tenant mapping to YARN pool
yarn.queue.mode=<user/tenant>
yarn.queue.mapping=<user1:test,user2:production>/<tenant1:test,tenant2:production>
yarn.queue.default=default
as.spark.driver.cleanup.delay=2
as.sparkdriver.max.memory=1G
```

內容	說明
yarn.queue.mode	設定 YARN 資源儲存區的對映模式。yarn.queue.mode=user 時，未提交工作/串流至 Analytic Server 的每一個使用者執行個別 YARN 應用程式。可以為不同的使用者工作/串流同時執行多個 YARN 應用程式。  yarn.queue.mode=tenant 時，未提交工作/串流至 Analytic Server 的每一個租戶執行個別 YARN 應用程式。可以為不同的租戶工作/串流同時執行多個 YARN 應用程式。
yarn.queue.mapping	將使用者或租戶配對對映至在「Cloudera 管理程式動態資源儲存區」中定義的 YARN 資源儲存區。配對必須以逗點區隔（例如，tenant1:test,tenant2:production 用於租戶，或者 user1:test,user2:production 用於使用者）。
yarn.queue.default	將應用程式提交到的預設 YARN 資源儲存區的名稱。您可以在「動態資源儲存區配置」中指定自訂的 YARN 資源儲存區。
as.spark.driver.cleanup.delay	代表登出之後終止使用者 YARN 應用程式之前分鐘數目的整數。預設值是 <b>2</b> 。此內容是選用項目。
as.sparkdriver.max.memory	設定每一個 <b>SparkDriver</b> 處理程序使用的記憶體數量。預設值是 <b>1G</b> 。此內容是選用項目。

## 參照

如需相關資訊，請參閱下列網站：

· <https://spark.apache.org/docs/latest/running-on-yarn.html>

## 移轉

Analytic Server 容許您將資料及配置設定從現有 Analytic Server 安裝移轉至新的安裝。

### 升級至新版本的 Analytic Server

如果您已安裝 Analytic Server 3.2.1.1，而且購買了更新的版本，您可以將 3.2.1.1 配置設定移轉至新安裝。

**限制：**3.2.1.1 及新安裝無法同時存在於同一個 Hadoop 叢集上。如果將新安裝配置為與 3.2.1.1 安裝使用相同的 Hadoop 叢集，則 3.2.1.1 安裝就不再起作用。

### 移轉步驟：3.2.1.1 至更新版本

1. 根據第 39 頁的『在 Cloudera 上安裝』中的指示，進行 Analytic Server 的新安裝。
2. 將舊安裝中的分析工作區複製到新安裝中。
  - a. 如果您不確定分析工作區的位置，請執行 `hadoop -fs ls`。分析工作區的路徑格式將為 `/user/as_user/analytic-root/analytic-workspace`，其中 `as_user` 是擁有分析工作區的使用者 ID。
  - b. 使用 `as_user` 身分登入新 Analytic Server 安裝的主機。如果 `/user/as_user/analytic-root/analytic-workspace` 目錄已存在，將其刪除。
  - c. 使用 `hadoop fs -copyToLocal` 與 `hadoop fs -copyFromLocal`，將舊伺服器的分析工作區複製到新伺服器的 `/user/as_user/analytic-root/analytic-workspace/` 資料夾（確保擁有者設為 `as_user`）。
3. 如果您使用內嵌的 Apache Directory Server，請使用第三方 LDAP 用戶端工具備份現行使用者/群組配置。安裝 Analytic Server 3.2.2 之後，將備份使用者/群組配置匯入至 Apache Directory Server。

註：如果您使用外部 LDAP 伺服器，則可跳過此步驟。
4. 在 Cloudera Manager 中，停止 Analytic Server 服務。
5. 從舊安裝收集配置設定。
  - a. 將新安裝中的 `configcollector.zip` 保存檔複製到舊安裝中的 `{AS_ROOT}\tools`。
  - b. 擷取 `configcollector.zip` 的副本。這會在舊安裝中建立新的 `configcollector` 子目錄。
  - c. 在舊安裝中執行配置收集器工具，方式是執行 `{AS_ROOT}\tools\configcollector` 中的 **configcollector** Script。將產生的壓縮 (ZIP) 檔複製到管理新安裝的伺服器。

**重要：**提供的 **configcollector** Script 與最近的 Analytic Server 版本不相容。如果您的 **configcollector** Script 發生問題，請聯絡 IBM 技術支援代表。

6. 清除 Zookeeper 狀態。在 Zookeeper bin 目錄中（例如，Cloudera 上的 `/opt/cloudera/parcels/CDH-5.4..../lib/zookeeper/bin`），執行下列指令。

```
./zkCli.sh rmr /AnalyticServer
```

7. 透過執行 **migrationtool** Script 並傳遞配置收集器建立的壓縮檔路徑作為引數，來執行移轉工具。範例如下所示。

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.2/ASConfiguration_3.2.1.1.xxx.zip
```

8. 在 Analytic Server 節點上，透過指令 Shell 執行下列指令：

```
hadoop fs -chmod -R 755 /user/as_user/analytic-root/analytic-workspace
```

9. 在 Cloudera Manager 中，啟動 Analytic Server 服務。

註：如果已將 R 配置成與現有 Analytic Server 安裝搭配使用，則將需要遵循下列步驟來配置成搭配使用新的 Analytic Server 安裝。

## 在 Cloudera 上解除安裝 Analytic Server

Cloudera 會自動處理解除安裝 Analytic Server 服務及包裹所需的大部分步驟。

從 Cloudera 環境中清除 Analytic Server 需要下列步驟：

1. 停止並刪除 Analytic Server 服務。
2. 取消啟動、從主機移除及刪除 Analytic Server 包裹。
3. 刪除 HDFS 中的 Analytic Server 使用者目錄。預設位置為 `/user/as_user/analytic-root`。
4. 刪除 Analytic Server 所使用的資料庫或綱目。
5. 清除 Analytic Server 安裝套件的任何殘餘。透過刪除下列項目即可達成：

- csd 資料夾
- 位於 parcels、parcel-cache 及 parcel-repo 資料夾中的任何現有 3.2.2 檔案。

## 第 4 章 配置 IBM SPSS Modeler 以與 IBM SPSS Analytic Server 搭配使用

若要讓 SPSS Modeler 搭配 Analytic Server 使用，您需要對 SPSS Modeler 伺服器安裝執行一些更新。

1. 配置 SPSS Modeler 伺服器以使其與 Analytic Server 安裝相關聯。

a. 編輯主要伺服器安裝目錄之 config 子目錄中的 options.cfg 檔案，並新增或編輯下列行：

```
as_ssl_enabled, {Y|N}  
as_host, "{AS_SERVER}"  
as_port, PORT  
as_context_root, "{CONTEXT-ROOT}"  
as_tenant, "{TENANT}"  
as_prompt_for_password, {Y|N}  
as_kerberos_auth_mode, {Y|N}  
as_kerberos_krb5_conf, {CONF-PATH}  
as_kerberos_krb5_spn, {AS-SPN}
```

### **as\_ssl\_enabled**

如果在 Analytic Server 上配置了安全通訊，請指定 Y；否則，請指定 N。

### **as\_host**

管理 Analytic Server 之伺服器的 IP 位址/主機名稱。

註：為 Analytic Server 啟用 SSL 時，您必須提供適當的 IP 位址/主機網域名稱。

### **as\_port**

Analytic Server 接聽所在的埠（依預設，這是 9080）。

### **as\_context\_root**

Analytic Server 環境定義根目錄（依預設，這是 analyticserver）。

### **as\_tenant**

SPSS Modeler 伺服器安裝屬於其中一個成員的承租人（預設承租人是 ibm）。

### **as\_prompt\_for\_password**

如果用在 Analytic Server 上所用的相同使用者與密碼的鑑別系統配置 SPSS Modeler 伺服器，請指定 N；例如，當使用 Kerberos 鑑別時。否則，請指定 Y。

以批次模式執行 SPSS Modeler 時，將 `-analytic_server_username {ASusername} - analytic_server_password {ASpassword}` 作為引數新增至 `clemb` 指令。

### **as\_kerberos\_auth\_mode**

指定 Y 可從 SPSS Modeler 啟用 Kerberos SSO。

### **as\_kerberos\_krb5\_conf**

指定 Analytic Server 應使用之 Kerberos 配置檔的路徑；例如，`\etc\krb5.conf`。

### **as\_kerberos\_krb5\_spn**

指定 Analytic Server Kerberos SPN；例如，`HTTP/ashost.mydomain.com@MYDOMAIN.COM`。

b. 重新啟動 SPSS Modeler 伺服器服務。

為了連接至已啟用 SSL/TLS 的 Analytic Server 安裝，配置 SPSS Modeler 伺服器與用戶端安裝需要執行一些其他步驟。

a. 導覽至 `http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT}` 並登入 Analytic Server 主控台。

b. 從瀏覽器下載憑證檔案並將它儲存到檔案系統中。

c. 將憑證檔案新增至 SPSS Modeler 伺服器與 SPSS Modeler 用戶端安裝的 JRE。要更新的位置可在 SPSS Modeler 安裝路徑的 `/jre/lib/security/cacerts` 子目錄中找到。

1) 請確保 `cacerts` 檔不是唯讀的。

- 2) 使用 Modeler 隨附的 keytool 程式 – 此程式可在 SPSS Modeler 安裝路徑的 /jre/bin/keytool 子目錄中找到。

執行下列指令

```
keytool -import -alias <as-alias> -file <cert-file> -keystore "<cacerts-file>"
```

請注意，<as-alias> 是 cacerts 檔案的別名。您可以使用您喜歡的任何名稱，只要它在 cacerts 檔案中是唯一的即可。

因此，範例指令類似如下：

```
keytool -import -alias MySSLCertAlias -file C:\Download\as.cer  
-keystore "c:\Program Files\IBM\SPSS\Modeler\{ModelerVersion}\jre\lib\security  
\cacerts"
```

- d. 重新啟動 SPSS Modeler 伺服器與 SPSS Modeler 用戶端。
2. [選用] 如果計劃使用 Analytic Server 資料來源對串流中的 R 模型進行評分，請安裝 IBM SPSS Modeler - Essentials for R。IBM SPSS Modeler - Essentials for R 可供下載 (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>)。

## 第 5 章 配置 UDF Hive 撤回

Hive UDF 登錄至 HiveDB 之後，Analytic Server 可以使用 UDF 功能以執行撤回。

依預設停用 UDF Hive 撤回，並且必須透過 ASModules.xml 檔中的 **udfmodule** 設定手動啟用（將 **disabled** 值變更為 **enabled**）。啟用設定之後，您必須重新啟動 Analytic Server，並手動將 UDF 登錄至 Hive。

### 附註：

- 在 HDP 3.x 上使用 HDP 3.x 資料來源時，您可能遇到下列錯誤：

錯誤：您嘗試載入的檔案不符合目的地表格的檔案格式。

1. 請開啟 Ambari 主控台，並在 **Hive > 配置 > 進階 > 進階 hive-site** 區段中變更下列內容。

```
Key: hive.default.fileformat.managed  
Value: TextFile (change the default value from ORC to TextFile)
```

2. 儲存配置。

- 在非 Kerberos 環境中使用 Hive 資料來源時，確保您在 **資料庫選項** 區段中輸入的使用者名稱與登入 Analytic Server 使用者相同。

下列範例示範如何在 HDP 及 Cloudera 環境中將 UDF 登錄/取消登錄至 Hive。

### 在 HDP 上登錄/取消登錄 UDF

#### 登錄 UDF

```
su - hive -c 'hive -f /opt/ibm/spss/analyticserver/3.2/bin/udfRegister.sql'
```

#### 取消登錄 UDF

```
su - hive -c 'hive -f /opt/ibm/spss/analyticserver/3.2/bin/udfUnregister.sql'
```

### 在 Cloudera 上登錄/取消登錄 UDF

#### 登錄 UDF

```
sudo -u hive kinit -k -t hive.keytab hive/bosperf5.fyre.ibm.com@IBM.COM  
sudo -u hive hive -f /opt/cloudera/parcels/AnalyticServer/bin/udfRegister.sql
```

#### 取消登錄 UDF

```
sudo -u hive hive -f /opt/cloudera/parcels/AnalyticServer/bin/udfUnregister.sql
```



## 第 6 章 使用 SLM 標籤以追蹤授權

SLM 標籤基於「資源使用率度量」的 ISO/IEC 19770-4 標準草稿版。SLM 標籤提供產品的標準化功能，以報告其授權標準（與軟體資產使用相關的）的耗用情資源況。在產品中啟用 SLM 之後，在執行時期，會產生 XML 檔以自行報告其授權使用。

啟動 Analytic Server 時，會在 `<as_installation_path>/logs/slmtag` 資料夾中建立 `slmtag` 檔案。

由於有兩個授權類型，因此會定期兩個不同的度量值：

- 對於現行 Analytic Server 版本，授權基於 Hadoop 叢集中的資料節點總數（根據虛擬伺服器）。節點數目記錄在下列 `slmtag` 檔案區段中。

```
<Type>VIRTUAL_SERVER</Type>
<SubType>Number of Data Nodes in Hadoop</SubType>
<Value>2</Value>
...
```

- 對於 Analytic Server 3.1 版之前的版本，授權基於 Hadoop 叢集中的 HDFS 儲存體大小（根據 RVU）。例如，儲存體大小（兆位元組）記錄在下列 `slmtag` 檔案區段中。

```
<Type>RESOURCE_VALUE_UNIT</Type>
<SubType>HDFS storage (Unit: Tega byte)</SubType>
<Value>0.21</Value>
```

SLM 標籤輸出在執行緒中啟動，並受 `SlmTagOutput.properties` 檔中定義的內容影響。該檔案位於 `<as_installation_path>/configuration` 資料夾中。

內容	說明
<code>license.metric.logger.output.enabled</code>	控制 SLM 日誌檔產生。預設值為 <code>False</code> 。
<code>license.metric.logger.output.dir</code>	用來儲存 SLM 標籤檔之目錄的相對路徑。預設目錄為 <code>&lt;as_installation_path&gt;/ logs</code> 。
<code>license.metric.logger.output.SLMLogFrequency</code>	用來收集 SLM 日誌的時間間隔 (unit:milliseconds)。
<code>icense.metric.logger.file.size</code>	SML 標籤檔案大小上限（以位元組為單位）。
<code>license.metric.logger.file.number</code>	一個軟體身分實例的 SLM 標籤檔數目上限。



## 第 7 章 疑難排解

本小節說明部分常見的安裝與配置問題，及如何修正這些問題。

### 一般問題

安裝成功但出現警告，不過使用者無法建立資料來源並出現錯誤「無法完成要求。原因：許可權遭拒」將 **distrib.fs.root** 參數設定為 Analytic Server 使用者（預設為 `as_user`）沒有存取權的目錄會導致錯誤。請確定已授權 Analytic Server 使用者讀取、寫入及執行 **distrib.fs.root** 目錄。

**Analytic Server 效能逐漸變差。**

當 Analytic Server 效能不符合預期時，請從 Knox 服務部署路徑移除所有 `*.war` 檔：  
<KnoxServicePath>/data/deployments。例如：`/usr/hdp/3.1.0.0-78/knox/data/deployments`。

### 在 Ambari 上解除安裝 Analytic Server or Essentials for R

在某些情況下，在 Ambari 上解除安裝 Analytic Server 或 Essentials for R 時，解除安裝程序會當掉。發生問題之後，您必須手動停止 Ambari 伺服器的程序 ID。

### Analytic Server 安裝在使用 OpenJDK 的 POWER System 上時

在使用 OpenJDK 的 POWER System 上執行 Analytic Server 時，您必須手動執行下列配置步驟，以確保座標系統 API 如預期運作

註：如果您未使用座標系統 API，則可以捨棄配置需求。

1. 在 Ambari 主控台中，導覽至 **Analytic Server 服務 > Configs 標籤 > 進階 analytics-jvm-options**，並將下列行新增至內容區域：

```
-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*
```

2. 在 Ambari 主控台中，導覽至自訂 **analytics.cfg** 區段，然後新增下列 3 個配置：

#### **spark.executor.extraJavaOptions**

將值設為：`-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

#### **spark.driver.extraJavaOptions**

將值設為：`-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

#### **mapred.child.java.opts**

將值設為：`-XX:CompileCommand=exclude,com/esri/sde/sdk/pe/engine/PeHznTwoPointEquidistant$GCSHorizon.*`

### 在 SuSE Linux 12 上安裝 Analytic Server 時發生錯誤

在 SuSE Linux 12 上安裝 Analytic Server 時，可能會發生下列錯誤：

```
簽章驗證失敗 [有 4 個簽章的公開金鑰無法使用]
```

在 SuSE Linux 12 上安裝 Analytic Server 之前，先執行下列作業即可解決該問題：

1. 從下列 URL 下載公開金鑰至您的主機：

```
https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.2.2.0/IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
```

2. 在主機上執行下列指令來匯入公開金鑰：

```
rpm --import IBM-SPSS-ANALYTICSERVER-GPG-KEY.public
```

### 特定 Hadoop 發行套件問題

#### Hortonworks 2.3-2.6 上已停用 Analytic Server 服務的重新整理動作

若要在 Hortonworks 2.3-2.6 上手動重新整理 Analytic Server 程式庫，請執行下列步驟。

1. 以 Analytic Server 使用者（預設為 `as_user`）身分登入執行 Analytic Metastore 的主機。

註：您可以從 Ambari 主控台尋找此主機名稱。

2. 執行目錄 `{AS_ROOT}/bin` 中的 **refresh** Script；例如：

```
cd /opt/ibm/spss/analyticserver/3.2/bin
./refresh
```

3. 在 Ambari 主控台中重新啟動 Analytic Server 服務。

### 從外部網站下載的套件無法通過 Cloudera Manager 中的雜湊檢查

雜湊驗證錯誤會顯示在包裹清單中。可以容許下載程序完成，然後透過 `cloudera-scm-server` 服務重新啟動 Cloudera 來解決此問題。當服務重新啟動之後，就不會發生此錯誤。

### HDFS supergroup 內容

如果 `as_user` 不是下列 HDFS 群組內容的成員，則 Analytic Server 在啟動期間將記載異常狀況：

**dfs.permissions.supergroup/dfs.permissions.superusergroup**。例如：

```
[11/15/17 7:32:35:510 PST] 000000bf SystemOut
0 2017-11-15 07:32:35,510 : | | | | | ERROR | slmtagoutput.SlmOutputAgent | SLM Logger => Error in performing callback function when
calculating number
of nodes in kerberos environment: org.apache.hadoop.ipc.RemoteException(org.apache.hadoop.security.AccessControlException): Access denied for user
as_user.
Superuser privilege is required
at org.apache.hadoop.hdfs.server.namenode.FSPermissionChecker.checkSuperuserPrivilege(FSPermissionChecker.java:93)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.checkSuperuserPrivilege(FSNamesystem.java:6606)
at org.apache.hadoop.hdfs.server.namenode.FSNamesystem.datanodeReport(FSNamesystem.java:5595)
at org.apache.hadoop.hdfs.server.namenode.NameNodeRpcServer.getDatanodeReport(NameNodeRpcServer.java:928)
at
org.apache.hadoop.hdfs.server.namenode.AuthorizationProviderProxyClientProtocol.getDatanodeReport(AuthorizationProviderProxyClientProtocol.java:390)
at
org.apache.hadoop.hdfs.protocolPB.ClientNamenodeProtocolServerSideTranslatorPB.getDatanodeReport(ClientNamenodeProtocolServerSideTranslatorPB.java:6
94)
at org.apache.hadoop.hdfs.protocol.proto.ClientNamenodeProtocolProtos$ClientNamenodeProtocol$2.callBlockingMethod(ClientNamenodeProtocolProtos.java)
at org.apache.hadoop.ipc.ProtobufRpcEngine$Server$ProtoBufRpcInvoker.call(ProtobufRpcEngine.java:617)
at org.apache.hadoop.ipc.RPC$Server.call(RPC.java:1073)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2141)
at org.apache.hadoop.ipc.Server$Handler$1.run(Server.java:2137)
at java.security.AccessController.doPrivileged(Native Method)
at javax.security.auth.Subject.doAs(Subject.java:415)
at org.apache.hadoop.security.UserGroupInformation.doAs(UserGroupInformation.java:1912)
at org.apache.hadoop.ipc.Server$Handler.run(Server.java:2135)
```

您必須將 `as_user` 手動新增至 `hdfs-site` 配置內容中定義的作業系統群組：

**dfs.permissions.supergroup/dfs.permissions.superusergroup**。

- 對於 Cloudera，預設內容值為 **supergroup**，且必須變更為實際存在的作業系統群組。如需 Cloudera 中 `supergroup` 設定的相關資訊，請參閱 [Cloudera 說明文件](#)。
- 對於 Ambari，預設內容值為 **hdfs**。依預設，在 Ambari 安裝期間，Analytic Server 會將 `as_user` 新增至 HDFS 及 Hadoop 群組。

在 Linux 上，使用 **usermod** 指令，來將 `as_user` 新增至 HDFS **superusergroup**（如果尚未存在）。

如需 HDFS 許可權的一般資訊，請參閱 [HDFS 許可權手冊](#)。

### HDP 3.0 上的 MapReduce 工作失敗

您可能在 HDP 3.0 上發現 MapReduce 工作具有下列錯誤：

無法完成要求。原因：java.lang.IllegalStateException：工作處於 DEFINE 狀態，而非 RUNNING (as\_trace.log)

我們可以據以解決的錯誤狀態：

1. 將下列配置新增至 Custom `analytics.cfg` 檔：

```
exclude.mapreduce.jars=icu4j-
```

2. 重新啟動 Analytic Server。

重新啟動 Analytic Server 之後，MapReduce 工作將正常執行。

### 由於 Cloudera 問題，將日期或時間戳記值寫入至 Hive 表格失敗

當 Analytic Server 嘗試將日期或時間戳記值寫入至 Cloudera 環境中的 Hive 表格時，處理程序由於已知的 Cloudera 問題而失敗 (<https://issues.apache.org/jira/browse/HIVE-11024>)。

註：日期值問題不會影響 Hive 1.3.0 或 2.0.0 (<https://issues.apache.org/jira/browse/HIVE-11024>)；時間戳記值問題不會影響 Hive 2.0.0 (<https://issues.apache.org/jira/browse/HIVE-11748?jql=project%20%3D%20HIVE%20AND%20text%20~%20%22jdbc%20timestamp%22>)。您必須確保受支援的 Hive 版本（1.3.0 或 2.0.0）呈現在您的 Cloudera 環境中。

## meta 資料儲存庫的問題

### 執行 add\_mysql\_user Script 時，CREATE USER 作業失敗

執行 `add_mysql_user` Script 之前，您需要先手動從 mysql 資料庫中，移除嘗試新增的使用者。您可以透過 MySQL 工作台使用者介面或 MySQL 指令，來移除使用者。例如：

```
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'localhost';"  
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'$METASTORE_HOST';"  
mysql -u root -e "DROP USER '$AEDB_USERNAME_VALUE'@'%';"
```

在上述指令中，將 `$AEDB_USERNAME_VALUE` 取代為您要移除的使用者名稱，並將 `$METASTORE_HOST` 取代為資料庫安裝所在的主機名稱。

## Apache Spark 的問題

### Spark 處理程序中執行串流的問題

當在 Spark 程序內強制執行時，SPSS Modeler 串流無法完成。失敗的 SPSS Modeler 串流是使用 Analytic Server 來源節點（HDFS 檔）建置的，它鏈結至排序節點，然後設定為匯出至另一個 Analytic Server 資料來源。執行串流之後，「資源管理程式」使用者介面指出應用程式正在執行，但串流永不完全且保留執行中狀態。在 Analytic Server 日誌、YARN 日誌或 Spark 日誌中沒有訊息指出串流無法完成的原因。

可以將 `spark.executor.memory` 設定新增至 Analytic Server 配置中的自訂 `analytics.cfg` 檔案，來解決此問題。將記憶體值設定為 4GB，可容許先前失敗的 SPSS Modeler 串流在 2 分鐘以內完成（位於單節點叢集環境中）。

當執行 SparkML 案例時遇到「在 `HdfsAuthcom.spss.utilities.i18n.Loc` 期間發生異常狀況。異常狀況：執行失敗。原因：`com.spss.ae.filesystem.exception.FileSystemException`：無法起始設定檔案系統存取。」。

當 Spark 找不到世系日誌目錄時會產生該錯誤。該問題的暫行解決方法是將 `spark.lineage.log.dir` 重新導向至 `/ae_wlpserver/usr/servers/aeserver/logs/spark`。

## 高可用性叢集

由於相依關係中的變更，Analytic Server 無法新增至更多主機。

使用第 28 頁的『更新用戶端相依關係』中的指示，執行 `update_clientdeps` Script。

「Analytic 叢集服務非預期地失去了與 Zookeeper 聯絡，此 JVM 將終止以維護叢集完整性。」

可能導致此問題的一個原因是要寫入 Zookeeper 的資料量太大。如果在 Zookeeper 中，日誌為如下所示的異常狀況：

```
java.io.IOException: Unreasonable length = 2054758
```

或者在 Analytic Server 中，日誌為如下所示的訊息：

```
Caused by: java.io.UTFDataFormatException: encoded string too long: 2054758 bytes  
at java.io.DataOutputStream.writeUTF(DataOutputStream.java:375)
```

1. 在 Ambari 主控台中，導覽至 Zookeeper 服務「配置」標籤，並將下列行新增至 `env` 範本，然後重新啟動 Zookeeper 服務。

```
export JVMFLAGS="-Xmx2048m -Djute.maxbuffer=2097152"
```

2. 在 Ambari 主控台中，導覽至 Analytic Server 服務「配置」標籤，並將下列內容新增至進階 `analytics-jvm-options`，然後重新啟動 Analytic Cluster 服務。

```
-Djute.maxbuffer=2097152
```

針對 `jute.maxbuffer` 設定指定的數字應該高於異常狀況訊息中指出的數字。

## Zookeeper 交易資料變得不可管理

將 zoo.cfg 中的 **autopurge.purgeInterval** 參數設為 1，以啟用 Zookeeper 交易日誌自動清除作業。

## Analytic 叢集服務失去了與 Zookeeper 的聯絡

在 zoo.cfg 中檢閱並修改 **tickTime**、**initLimit** 和 **syncLimit** 參數。例如：

```
# The number of milliseconds of each tick
tickTime=2000
# The number of ticks that the initial
# synchronization phase can take
initLimit=30
# The number of ticks that can pass between
# sending a request and getting an acknowledgement
syncLimit=15
```

請參閱 Zookeeper 說明文件以取得詳細資料：<https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html>

## Analytic Server 工作未繼續執行

Analytic Server 工作未繼續執行有一種常見情況。

- 如果 Analytic Server 工作由於叢集成員失敗而失敗，則通常會在另一個叢集成員上自動重新啟動工作。如果工作未繼續執行，請檢查以確保「高可用性」叢集中至少具有 4 個叢集成員。

## Hive 後推

啟用後推時，您可能收到下列錯誤訊息：

```
(AEQAE2103E) SQL 執行失敗 - 編譯陳述式時發生錯誤：
失敗：語意異常狀況 [錯誤 10014]：明細行 3:47 引數錯誤 '9223372036854775808'：
出於安全原因，已停用不同類型之間的不安全比較。如果您知道正在做什麼，
請將 hive.strict.checks.type.safety 設為 false，並確保 hive.mapred.mode
未設為 'strict' 以繼續。請注意，如果您在使用部分不安全的特性
時發生錯誤，則可能會取得錯誤或不正確的結果。(as_trace.log)
```

可以透過採用下列其中一個方法來解決錯誤。

- 將 **hive.sql.check=true** 新增至 Analytic Server config.properties 檔。
- 在 Hive 配置中將 **hive.strict.checks.type.safety** 設定變更為 **false**。

## 注意事項

本資訊係針對 IBM 在美國所提供之產品與服務所開發。IBM 可能以其他語言提供本資訊。不過，您可能需要擁有一份該語言的產品或產品版本的副本，才能存取該產品或產品版本。

在其他國家，IBM 不見得有提供本文件所提及之各項產品、服務或功能。請洽詢當地的 IBM 業務代表，以取得當地目前提供的產品和服務之相關資訊。本文件在提及 IBM 的產品、程式或服務時，不表示或暗示只能使用 IBM 的產品、程式或服務。只要未侵犯 IBM 之智慧財產權，任何功能相當之產品、程式或服務皆可取代 IBM 之產品、程式或服務。不過，任何非 IBM 之產品、程式或服務，使用者必須自行負責作業之評估和驗證責任。

本文件所說明之主題內容，IBM 可能擁有其專利或專利申請案。提供本文件不代表提供這些專利的授權。您可以書面提出授權查詢，來函請寄到：

*IBM Director of Licensing*  
*IBM Corporation*  
*North Castle Drive, MD-NC119*  
*Armonk, NY 10504-1785*  
*US*

如果是有關雙位元組 (DBCS) 資訊的授權查詢，請洽詢所在國的 IBM 智慧財產部門，或書面提出授權查詢，來函請寄到：

*Intellectual Property Licensing*  
*Legal and Intellectual Property Law*  
*IBM Japan Ltd.*  
*19-21, Nihonbashi-Hakozakicho, Chuo-ku*  
*Tokyo 103-8510, Japan*

International Business Machines Corporation 只依「現況」提供本出版品，不提供任何明示或默示之保證，其中包括且不限於不侵權、可商用性或特定目的之適用性的隱含保證。有些地區在特定交易上，不允許排除明示或暗示的保證，因此，這項聲明不一定適合您。

本資訊中可能會有技術上或排版印刷上的訛誤。因此，IBM 會定期修訂；並將修訂後的內容納入新版中。IBM 隨時會改進及/或變更本出版品所提及的產品及/或程式，不另行通知。

本資訊中任何對非 IBM 網站的敘述僅供參考，IBM 對該網站並不提供任何保證。這些網站所提供的資料不是 IBM 本產品的資料內容，如果要使用這些網站的資料，您必須自行承擔風險。

IBM 得以各種 IBM 認為適當的方式使用或散布 貴客戶提供的任何資訊，而無需對 貴客戶負責。

如果本程式之獲授權人為了 (i) 在個別建立的程式和其他程式（包括本程式）之間交換資訊，以及 (ii) 相互使用所交換的資訊，因而需要相關的資訊，請洽詢：

*IBM Director of Licensing*  
*IBM Corporation*  
*North Castle Drive, MD-NC119*  
*Armonk, NY 10504-1785*  
*US*

上述資料之取得有其特殊要件，在某些情況下必須付費方得使用。

IBM 基於 IBM 客戶合約、IBM 國際程式授權合約或雙方之任何同等合約的條款，提供本文件所提及的授權程式與其所有適用的授權資料。

所述的效能資料及客戶範例僅供示範之用。實際的效能結果可能會因特定的配置及作業狀況而異。

本文件所提及之非 IBM 產品資訊，取自產品的供應商，或其發佈的聲明或其他公開管道。IBM 並未測試過這些產品，也無法確認這些非 IBM 產品的執行效能、相容性或任何對產品的其他主張是否完全無誤。有關非 IBM 產品的性能問題應直接洽詢該產品供應商。

所有關於 IBM 未來方針或目的之聲明，隨時可能更改或撤銷，不必另行通知，且僅代表目標與主旨。

所有 IBM 價格都是 IBM 建議的零售價格，可隨時變更而不另行通知。經銷商價格可不同。

本資訊僅作規劃目的。在產品可用前，此處的資訊可能變更。

本資訊含有日常商業運作所用之資料和報告範例。為了盡可能地加以完整說明，範例中含有個人、公司、品牌及產品的名稱。此等名稱皆屬虛構，凡有類似實際個人或企業所用之名稱及地址者，皆屬巧合。

著作權：

本資訊含有日常商業運作所用之資料和報告範例。為了盡可能地加以完整說明，範例中含有個人、公司、品牌及產品的名稱。此等名稱皆屬虛構，凡有類似實際個人或企業所用之名稱及地址者，皆屬巧合。

這些範例程式或任何衍生著作的每份副本或任何部分，都必須依照下列方式併入著作權聲明：

© IBM 2020. 本程式之若干部分係衍生自 IBM 公司的範例程式。

© Copyright IBM Corp. 1989 - 2020. All rights reserved.

## 商標

---

IBM、IBM 標誌及 [ibm.com](http://ibm.com) 是 International Business Machines Corp. 在世界許多管轄區註冊的商標或註冊商標。其他產品及服務名稱可能是 IBM 或其他公司的商標。IBM 商標的最新清單可在 Web 的 "Copyright and trademark information" 中找到，網址為 [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)。

Adobe、Adobe 標誌、PostScript 及 PostScript 標誌是 Adobe Systems Incorporated 在美國及（或）其他國家或地區的註冊商標或商標。

IT Infrastructure Library 是 Central Computer and Telecommunications Agency（現在是 Office of Government Commerce 的一部分）的註冊商標。

Intel、Intel 標誌、Intel Inside、Intel Inside 標誌、Intel Centrino、Intel Centrino 標誌、Celeron、Intel Xeon、Intel SpeedStep、Itanium 及 Pentium 是 Intel Corporation 或其子公司在美國及其他國家或地區的商標或註冊商標。

Linux 是 Linus Torvalds 在美國及（或）其他國家或地區的註冊商標。

Microsoft、Windows、Windows NT 及 Windows 標誌是 Microsoft Corporation 在美國及/或其他國家或地區的商標。

ITIL 是 Minister for the Cabinet Office 在美國 Patent and Trademark Office 註冊的註冊商標及註冊社群商標。

UNIX 是 The Open Group 在美國及其他國家或地區的註冊商標。

Cell Broadband Engine 是 Sony Computer Entertainment, Inc. 在美國及/或其他國家或地區的商標並在當地軟體使用權下使用。

Linear Tape-Open、LTO、LTO 標誌、Ultrium 及 Ultrium 標誌是 HP、IBM Corp. 及 Quantum 在美國及其他國家的商標。



