

**IBM SPSS Analytic Server
3.1.0 版**

安裝與配置手冊

IBM

附註

在使用本資訊及它支援的產品之前，請閱讀第 65 頁的『注意事項』中的資訊。

產品資訊

此版本適用於 IBM SPSS Analytic Server 3.1.0 版及所有後續版本與修訂，但新版本另有說明者不在此限。

目錄

第 1 章 必備項目	1
第 2 章 Ambari 安裝與配置	3
Ambari 特定必備項目	3
IBM SPSS Analytic Server 安裝事先檢查與事後檢查 工具	3
在 Ambari 上安裝	5
離線安裝	7
針對外部管理的 MySQL 環境，安裝 Analytic Server	10
配置	11
安全	11
啟用對 Essentials for R 的支援	16
啟用關聯式資料庫來源	17
啟用 HCatalog 資料來源	18
變更 Analytic Server 所用的埠	19
高可用性 Analytic Server	19
最佳化小型資料的 JVM 選項	20
更新用戶端相依關係	20
配置 Apache Knox	20
升級和移轉	24
解除安裝	27
解除安裝 Essentials for R	27
第 3 章 Cloudera 安裝與配置	29
Cloudera 概觀	29
Cloudera 特定必備項目	29
針對 Analytic Server 配置 MySQL	29
在 Cloudera 上安裝	30
配置 Cloudera	31
安全	32
啟用 Essentials for R 的支援	36
啟用關聯式資料庫來源	37
啟用 HCatalog 資料來源	38
配置 Apache Impala	39
變更 Analytic Server 所用的埠	40
高可用性 Analytic Server	40
最佳化小型資料的 JVM 選項	40
移轉	41
在 Cloudera 上解除安裝 Analytic Server	42
第 4 章 MapR 安裝與配置	43
MapR 概觀	43
在 MapR 上安裝 Analytic Server	43
配置 MapR	46
啟用資料庫推回	46
啟用 Apache Hive	47
啟用 Apache HBase	48
啟用 Apache SPARK	49
啟用功能旗標	51
啟用 R	51
啟用 LZ0	51
針對 MapR 設定 IBM SPSS Analytic Server 叢 集	52
解除安裝 MapR	52
在 MapR 上移轉 IBM SPSS Analytic Server	53
MapR 疑難排解	54
第 5 章 Huawei FusionInsight HD 安裝 與配置	55
FusionInsight HD 概觀	55
在 Huawei FusionInsight HD 上安裝	55
第 6 章 配置 IBM SPSS Modeler 以與 IBM SPSS Analytic Server 搭配使用	59
第 7 章 疑難排解	61
注意事項	65
商標	66

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。

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>

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，其格式為 `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。

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。

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。

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，您應該確保將 HDFS、YARN、MapReduce2、Hive 及 Zookeeper 新增為 Ambari 服務。

無密碼的 SSH

為 Analytic meta 儲存庫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。

IBM SPSS Analytic Server 安裝事先檢查與事後檢查工具

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：事先檢查工具可在安裝 Analytic Server 之後執行。

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、cmx)

註：cmx 僅在 BigInsights 中受支援。

- PySpark 隨附的資料
- 使用原生 SPSS 元件 (包括 alm、樹狀結構、neuralnet、評分、tascoreing) 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

事先檢查與事後檢查工具位於下列目錄中：

- **BigInsights**

```
/var/lib/ambari-server/resources/stacks/BigInsights/4.X/services/ANALYTICSERVER/package/chktool
```

- **HDP**

```
/var/lib/ambari-server/resources/stacks/HDP/2.X/services/ANALYTICSERVER/package/chktool
```

- **Cloudera**

工具包裝在 `AnalyticServer-*.jar` 檔中（該檔案位於 `/opt/cloudera/csd`）。

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

安裝 Analytic Server 之前，事先檢查工具應該在將要管理 Analytic Server 服務的所有 Ambari 節點上執行。在不同節點上執行工具需要將整個 `chktool` 目錄複製到節點。

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 7 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

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 Ambari -cluster MyCluster -username admin  
-password admin -host myambarihost.ibm.com -port 8080 -as_host myashost.ibm.com -ssl
```

附註：

- `as_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 -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 -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 -host myknosserverhost.ibm.com -port 8443
-username admin -password ibmspss -ssl -gateway_url /gateway/default -check AS_PYSPARK_BUILDMODEL
```

附註：

- 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。更多詳細步驟如下所示。

重要：Analytic Server 不支援以非 root 使用者身分在執行 Ambari 伺服器的環境中安裝。

1. 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。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 3.1 for BigInsights 4.1 和 4.3 Linux on System p LE (英文)	spss_as-3.1-bi4.1-4.3-1ppc64le_en.bin
IBM SPSS Analytic Server 3.1 for BigInsights 4.1、4.2 和 4.3 Linux x86-64 (英文)	spss_as-3.1-bi4.1-4.2-4.3-1x86_en.bin
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.3 和 2.4 Ubuntu (英文)	spss_as-3.1-hdp2.3-2.4-ubun_en.bin
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.4、2.5 和 2.6 Linux x86-64 (英文)	spss_as-3.1-hdp2.4-2.6-1x86_en.bin
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.6 Linux on System p LE (英文)	spss_as-3.1-hdp2.6-1ppc64le_en.bin

2. 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。

[僅限 *GFPS (Spectrum Scale)*] 下載檔案 <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo> (x86 和 ppc64le) 或 <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/Ubuntu/IBM-SPSS-AnalyticServer.list> (Ubuntu)，並將其移至您在其中新增 Analytic Server Metastore 作為服務的每個節點上的資料夾 `/etc/yum/repos.d` (RHEL、CentOS)、`/etc/zypp/repos.d` (SLES) 或 `/etc/apt/sources.list.d` (Ubuntu)。

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。如需詳細資料，請參閱第 7 頁的『離線安裝』。

3. 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，使用 `--help` 引數）的 `update_clientdeps.sh` Script。

4. 重新啟動 Ambari 伺服器。

```
ambari-server restart
```

5. 如果您正在執行離線安裝，請驗證 `reponfo.xml` 和 `IBM-SPSS-AnalyticServer-3.1.0.0.repo` 變更是否仍然存在。如果您發現 `reponfo.xml` 及 `IBM-SPSS-AnalyticServer-3.1.0.0.repo` 變更未持續存在，請與 IBM 支援中心代表聯絡。
6. 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。

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。

- a. 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。
- b. 在「Ambari 服務」標籤中，導覽至 Analytic Server 服務的「配置」標籤。
- c. 開啟進階 **analytics-env** 區段。
- d. 將 `as.database.type` 的值從 `mysql` 變更為 `db2`。
- e. 開啟進階 **analytics-meta** 區段。
- f. 將 `metadata.repository.driver` 的值從 `com.mysql.jdbc.Driver` 變更為 `com.ibm.db2.jcc.DB2Driver`。
- g. 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。
- h. 在 `metadata.repository.user.name` 及 `metadata.repository.password` 中提供有效的 DB2 認證。
- i. 按一下儲存。

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`

7. 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 11 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 24 頁的『升級和移轉』。

- 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 在其上接聽的埠。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

離線安裝的一般工作流程如下所示：

- 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。可用的 Ambari 二進位檔如下所示：

表 2. *Analytic Server* 自行解壓縮的二進位檔

說明	二進位檔名
IBM SPSS Analytic Server 3.1 for BigInsights 4.1 和 4.3 Linux on System p LE (英文)	<code>spss_as-3.1-bi4.1-4.3-lppc64le_en.bin</code>
IBM SPSS Analytic Server 3.1 for BigInsights 4.1、4.2 和 4.3 Linux x86-64 (英文)	<code>spss_as-3.1-bi4.1-4.2-4.3-lx86_en.bin</code>
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.3 和 2.4 Ubuntu (英文)	<code>spss_as-3.1-hdp2.3-2.4-ubun_en.bin</code>
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.4、2.5 和 2.6 Linux x86-64 (英文)	<code>spss_as-3.1-hdp2.4-2.6-lx86_en.bin</code>
IBM SPSS Analytic Server 3.1 for Hortonworks Data Platform 2.6 Linux on System p LE (英文)	<code>spss_as-3.1-hdp2.6-lppc64le_en.bin</code>

- 執行二進位執行檔並指定離線安裝。離線安裝會下載必要的 RPM 檔案或 DEB 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的電腦上。執行檔二進位檔位於可用的 Ambari `<AS_INSTALLABLE_HOME>` 配送目錄中。
- 將適當的 `<AS_INSTALLABLE_HOME>` 目錄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）。
- 繼續在適當的 Ambari 管理程式節點（受防火牆保護）上執行適用於您的特定發行套件（BigInsights/HDP 或 Ubuntu）的指示步驟。

BigInsights 和 HDP 指示 - 3.1.0

重要：Analytic Server 不支援以非 root 使用者身分在執行 Ambari 伺服器的環境中安裝。

- 使用下列指令來檢查 Ambari 伺服器是否在執行中：

```
ambari-server status
```

關閉 Ambari 伺服器節點（如果目前正在執行中）：

```
ambari-server stop
```

- 安裝工具來建立本端 yum 儲存庫。

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

或

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

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

- 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。

BigInsights 4.1、4.2 及 4.3 (x86_64)

SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

BigInsights 4.1 和 4.3 (PPC64LE)

SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.1.0.0-1.ppc64le.rpm

HDP 2.3、2.4 和 2.5 (x86_64)

SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm

- 建立本端儲存庫定義。例如，在 `/etc/yum.repos.d/` (RHEL、CentOS) 或 `/etc/zypp/repos.d/` (SLES) 中建立 `IBM-SPSS-AnalyticServer-3.1.0.0.repo` 檔，其中包括下列內容。

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

- 建立本端 yum 儲存庫。
`createrepo /home/root/repos/IBM-SPSS-AnalyticServer/x86_64` (RHEL、CentOS、SLES)
- 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 8 和步驟 9。

- 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。
 - 將 `add_mysql_user.sh` Script 從 `/opt/AS_Installable/IBM-SPSS-AnalyticServer` 複製到將用作 `AS_MetaStore` 之 MySQL 實例安裝所在的節點/主機。例如，`/opt/AS_InstallTools`。
 - 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`

- 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:///{{本端儲存庫的路徑}}/</baseurl>
    <repoid>IBM-SPSS-AnalyticServer</repoid>
    <reponame>IBM-SPSS-AnalyticServer-3.1.0.0</reponame>
  </repo>
</os>

```

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

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

10. 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> 目錄的整體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 /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。

BigInsights 4.1、4.2 及 4.3 (x86_64)

```
SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm
```

```
IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
```

BigInsights 4.1 和 4.3 (PPC64LE)

```
SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm
```

```
IBM-SPSS-AnalyticServer-3.1.0.0-1.ppc64le.rpm
```

HDP 2.3、2.4 和 2.5 (x86_64)

```
SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm
```

```
IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
```

e. 建立本端儲存庫定義。例如，在 /etc/yum.repos.d/ (RHEL、CentOS) 或 /etc/zypp/repos.d/ (SLES) 中建立 IBM-SPSS-AnalyticServer-3.1.0.0.repo 檔，其中包括下列內容。

```

[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)
```

11. 繼續執行第 5 頁的『在 Ambari 上安裝』主題中的步驟 5。

Ubuntu 指示 - 3.1.0

1. 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適當 Ubuntu 自行解壓縮二進位檔，下載至 Ambari 叢集內的主機上。可用的二進位檔如下所示：

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 3.1.0 for Hortonworks Data Platform 2.3 和 2.5 Ubuntu (英文)	spss_as-3.1.0-hdp2.3-2.5-ubun_en.bin

2. 在具有網際網路存取權的機器上，執行二進位執行檔並指定離線安裝。離線安裝會下載必要的 DEB 檔案，且應該執行於可存取 <https://ibm-open-platform.ibm.com> 的電腦上。二進位執行檔位於可用的 Ambari <AS_INSTALLABLE_HOME> 配送目錄中。
3. 必要的 Analytic Server DEB 檔位於下列目錄。
IBM-SPSS-AnalyticServer/packages
4. 使用下列指令來安裝 Analytic Server 3.1.0：

```
dpkg -i ./IBM-SPSS-AnalyticServer-ambari-HDP-2.5_3.1.0.0_amd64.deb (or IBM-SPSS-AnalyticServer-ambari-HDP-2.3_3.1.0.0_amd64.deb)
dpkg -i ./IBM-SPSS-AnalyticServer_1_amd64.deb
```
5. 重新啟動 Ambari 伺服器。
ambari-server restart
6. 登入 Ambari 伺服器，然後透過 Ambari 使用者介面將 Analytic Server 安裝為服務。

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。

1. 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。
2. 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。
 - a. 選擇線上選項。
 - b. 出現提示時，請選取外部 **MySQL** 資料庫選項。
3. 將 `add_mysql_user.sh` Script 從 `/opt/AS_Installable/IBM-SPSS-AnalyticServer` 複製到將用作 `AS_MetaStore` 之 MySQL 實例安裝所在的節點/主機。例如，`/opt/AS_InstallTools`。
 - 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。
4. 重新啟動 Ambari 伺服器。
 5. 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`

配置

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} 是指在其中部署 Analytic Server 的位置；例如，/opt/IBM/SPSS/AnalyticServer/{version}。
- {AS_SERVER_ROOT} 是指配置檔、日誌檔和伺服器檔案的位置；例如，/opt/IBM/SPSS/AnalyticServer/{version}/ae_wlpserver/usr/servers/aeserver。
- {AS_HOME} 是指 HDFS 上由 Analytic Server 用作根資料夾的位置。

安全

security.config 參數定義可以作為主體新增至 Analytic Server 系統之使用者和群組的登錄。

依預設，基本登錄定義為使用單一使用者 admin 和密碼 admin。您可以透過編輯 **security.config** 或配置 Kerberos 來變更登錄。在 Analytic Server 服務之「配置」標籤的 **Analytic Server 進階配置 Snippet 區** 段中，可以找到 **security_cfg** 參數。

註：如果您編輯 **security.config** 參數以變更登錄，則需要將任何新使用者作為主體新增至 Analytic Server 系統。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》。

對基本登錄進行變更

基本登錄可讓您在 **security.config** 參數中定義使用者和群組的資料庫。

預設基本登錄類似如下。

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="admin"/>
</basicRegistry>
```

變更後的基本登錄範例如下所示。

```
<basicRegistry id="basic" realm="ibm">
  <user name="user1" password="{xor}Dz4sLG5tbGs="/>
  <user name="user2" password="Pass"/>
  <user name="user3" password="Pass"/>
  <user name="user4" password="Pass"/>
  <user name="admin" password="{xor}KzosKw="/>
  <group name="Development">
    <member name="user1"/>
    <member name="user2"/>
  </group>
  <group name="QA">
    <member name="user3"/>
    <member name="user4"/>
  </group>
  <group name="ADMIN">
    <member name="user1"/>
    <member name="admin"/>
  </group>
</basicRegistry>
```

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。

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

註：如需 securityUtility 工具的詳細資料，請參閱 http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html。

註：在沙盤推演環境中基本登錄非常有用，但不建議在正式作業環境中使用。

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="ldapserverserver.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">
  <activatedFilters
    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

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 /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。

註：如果 Analytic Server 安裝使用基本登錄，則它必須包含 Kerberos 使用者帳戶（使用 "-" 作為密碼）。範例如下所示。

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="-"/>
  <user name="user1" password="-"/>
  <user name="user2" password="-"/>
  <group name="group1">
    <member name="admin"/>
    <member name="user1"/>
    <member name="user2"/>
  </group>
  <group name="group2">
    <member name="admin"/>
    <member name="user1"/>
  </group>
</basicRegistry>
```

2. 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。
 - 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。
3. 針對 Analytic Server 中的所有主體，在 HDFS 上建立使用者起始檔案夾。例如，如果您將 `testuser1` 新增至 Analytic Server 系統，則在 HDFS 上建立 `/user/testuser1` 之類的起始檔案夾，並確保 `testuser1` 擁有此資料夾的讀取權和寫入權。
4. [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。

現在，使用者可以使用 Kerberos SSO 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)。

停用 Kerberos

1. 在 Ambari 主控台中停用 Kerberos。
2. 停止 Analytic Server 服務。
3. 從自訂 analytics.cfg 中移除下列參數。

```
default.security.provider
hdfs.keytab
hdfs.user
java.security.krb5.conf
as.db.connect.method
web.authentication.kerberos.keytab
web.authentication.kerberos.principal
```

4. 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。

若要安裝第三方供應商憑證，請遵循下列步驟。

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/analyticsserver/3.0/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/analyticsserver/3.0/ae_wlpserver/bin/securityUtility createSSLCertificate --server=myserver --password=mypassword --validity=365 --subject=CN=mycompany,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。

4. 按一下儲存，然後重新啟動 Analytic Server 服務。

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. 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-tpssp>)。選擇您的堆疊、堆疊版本及硬體架構特定的檔案。
2. 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。

[僅限 *GPFS (Spectrum Scale)*] 下載檔案 https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/x86_64/IBM-SPSS-AnalyticServer-3.1.0.0.repo (x86)、<https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/ppc64le/IBM-SPSS-AnalyticServer-3.1.0.0.repo> (ppc64le) 或 <https://ibm-open-platform.ibm.com/repos/IBM-SPSS-ModelerEssentialsR/3.1.0.0/Ubuntu/IBM-SPSS-ModelerEssentialsR-3.1.0.0.list> (Ubuntu)，並將其移至您在其上新增 Analytic Server Meta 資料儲存庫作為服務的每個節點上的 /etc/yum.repos.d (RHEL、CentOS)、/etc/zypp/repos.d (SLES) 或 /etc/apt/sources.list.d (Ubuntu)。

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。

- a. 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。

BigInsights 4.1 和 4.2 (x86_64)

```
IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.x86_64.rpm
```

BigInsights 4.1 (PPC64LE)

```
IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.ppc64le.rpm
```

HDP 2.3 和 2.4 (x86_64)

```
IBM-SPSS-ModelerEssentialsR-ambari-2.1-HDP-2.3-8.4.1.0-1.x86_64.rpm
```

HDP 2.4 (Ubuntu)

```
IBM-SPSS-ModelerEssentialsR-ambari-3.1.0.0_3.1.0.0_amd64.deb
```

- b. 安裝 RPM 或 DEB。在下列範例中，指令會將 Essentials for R 安裝在 BigInsights 4.2 上。

```
rpm -i IBM-SPSS-ModelerEssentialsR-ambari-2.1-BI-4.1-8.4.1.0-1.x86_64.rpm
```

在下列範例中，指令會將 Essentials for R 安裝在 HDP 2.4 (Ubuntu) 上。

```
dpkg -i IBM-SPSS-ModelerEssentialsR-ambari-3.1.0.0_3.1.0.0_amd64.deb
```

3. 重新啟動 Ambari 伺服器。

```
ambari-server restart
```

4. 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。

5. 重新整理 Analytic Server 服務。
6. 使用第 20 頁的『更新用戶端相依關係』中的指示，執行 update_clientdeps Script。
7. 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

如果您在每個 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「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. 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 4. 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	10.5、10.1、9.7	db2jcc.jar	IBM
DB2 z/OS	11、10	db2jcc.jar、db2_license_cisuz.jar	IBM
Greenplum	5、4.2.x	postgresql.jar	Greenplum
Hive	1.1、1.2	hive-jdbc-*.jar	Apache
Netezza	7、6.x	nzjdbc.jar	IBM
Oracle	12c、11g R2 (11.2)	ojdbc6.jar、orai18n.jar	Oracle
SQL Server	2014、2012、2008 R2	sqljdbc4.jar	Microsoft
Sybase IQ	16.x、15.4、15.2	jconnect70.jar	Sybase
Teradata	14、14.1、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。

BigSQL 是 BigInsights 增值服務。因此，其版本與 BigInsights 版本相同。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。如需詳細資料，請參閱下列各節。
2. 將這些 JAR 檔新增至 {HIVE_HOME}/auxlib 目錄以及每一個 Analytic Server 節點上的 /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。

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>。

Apache Spark

如果您要使用 Spark（1.5 版或更新版本）與 HCatalog 輸入資料來源，則必須將 `spark.version` 內容手動新增至自訂的 `analytics.cfg` 檔。

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，`1.x`、`2.x` 或 `None`）。
2. 儲存配置，並從 Amabri 主控台重新啟動所有 Analytic Server 服務。

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`

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**

最佳化小型資料的 JVM 選項

您可以編輯 JVM 內容，以最佳化執行較小 (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，而且可能需要嘗試使用這些值以最佳化系統。

```
-Xms512M  
-Xmx2048M
```

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；請參閱下列範例。

```
cd "/var/lib/ambari-server/resources/stacks/HDP/2.4/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 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

- 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_0.6-3.1.0.0.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

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 <service> 新增至 Knox 配置：

```
<service>
  <role>ANALYTICSERVER</role>
  <url>http://{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，以反映對應的 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. 在 Knox 伺服器上，建立子目錄 `<knox_server>/data/service/analyticserver/3.1`，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/3.1` 上（例如，`/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/knox/analyticserver/3.1/*.xml`）
2. 在 `<knox_server>/bin` 中，執行 Script `./knoxcli.sh redeploy --cluster default`
3. 將 `com.ibm.spss.knoxservice_0.6-*.jar` 檔上傳至 `<knox_server>/ext`。該檔案位於 Analytic Server 的 `<analytic_server>/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.1.0.0.jar`（例如，`/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/com.ibm.spss.knox_0.6-3.1.0.0.jar`）。
4. 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://{AS-Host}:{AS-port}/analyticserver</url>
</service>

```

5. 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`）。
6. 從 **Knox > 服務動作 > 啟動展示 LDAP** 重新啟動 LDAP。
7. 重新啟動 Knox 服務。

在 Hortonworks 資料平台 (HDP) 上安裝 Apache Knox

下列步驟輪廓了在 HDP 叢集中安裝 Apache Knox 的過程。

1. 驗證 HDP 叢集上是否存在 Knox 使用者。如果 Knox 使用者不存在，則必須建立一個。
2. 下載 Apache Knox 並將其解壓縮至 `/home/knox` 下的資料夾中。
3. 在 HDP 中，切換至 Knox 使用者並跳至 `knox` 資料夾。Knox 使用者必須具有所有 `knox` 子資料夾的許可權(RWX)。
4. 為 Analytic Server 配置 Apache Knox。如需相關資訊，請參閱配置 **Apache Knox** 一節。
 - a. 在 `{knox}/data/services` 下建立 `analyticserver/3.0` 資料夾階層。
 - b. 從 Analytic Server 位置複製 `rewrite.xml` 和 `service.xml` 檔案：

```

/opt/ibm/spss/analyticserver/3.0/ae_wlpserver/usr/servers/aeserver/configuration/knox/
analyticserver/3.1

```

至 Knox 伺服器節點：

```

{knox}/data/services/analyticserver/3.1

```

- c. 從 Analytic Server 主機複製 Knox `*.jar` 檔：

```
/opt/ibm/spss/analyticserver/3.1/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/
lib/com.ibm.spss.knox_0.6-*.jar
```

至 Knox ext 目錄：

```
{knox}/ext
```

- d. 更新 {knox}/conf/topologies 中的 default.xml 檔案，以符合下列範例：

註：如果該檔案不存在，則必須建立它。

```
<topology>
  <gateway>
    <provider>
      <role>authentication</role>
      <name>ShiroProvider</name>
      <enabled>true</enabled>
      <param>
        <name>sessionTimeout</name>
        <value>30</value>
      </param>
      <param>
        <name>main.ldapRealm</name>
        <value>org.apache.hadoop.gateway.shiro.realm.KnoxLdapRealm</value>
      </param>
      <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://localhost:33389</value>
      </param>
      <param>
        <name>main.ldapRealm.contextFactory.authenticationMechanism</name>
        <value>simple</value>
      </param>
      <param>
        <name>urls./*</name>
        <value>authcBasic</value>
      </param>
    </provider>
    <provider>
      <role>identity-assertion</role>
      <name>Default</name>
      <enabled>true</enabled>
    </provider>
    <provider>
      <role>authorization</role>
      <name>AclsAuthz</name>
      <enabled>true</enabled>
    </provider>
  </gateway>

  <!--other service-->
  <service>
    <role>ANALYTICSERVER</role>
    <!--replace the {AS-host}nas {AS-port} with real value-->
    <url>http://{AS-host}:{AS-port}/analyticserver</url>
  </service>
</topology>
```

5. 執行 {knox}/bin/knoxcli.sh。

6. 執行 {knox}/bin/ldap.sh start。

註：Script 使用埠 33389。請確定目前沒有使用該埠。

7. 執行 {knox}/bin/gateway.sh start。

註：Script 使用埠 8443。請確定目前沒有使用該埠。

8. 驗證安裝。

- a. 針對 Knox URL 上的 Analytic Server 執行 curl 指令：

```
curl -ikvu {username}:{password} https://{knox-host}:8443/gateway/default/analyticserver/admin
```

疑難排解

問題：Analytic Server 安裝之後在 Knox 中不工作。

解決方案：停止 Knox，移除 {knox}/data/deployments/* 下的所有檔案，然後重新啟動 Knox。

問題：無法透過 Knox 登入 Analytic Server。

解決方案：驗證 {knox}/conf/users.ldif 中的使用者。更新現有使用者，或新增 Analytic Server 使用者。Knox 使用者主體和認證必須符合 Analytic Server 使用者。

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/analyticsserver/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/analyticsserver。

升級和移轉

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

從 3.0.1 版升級至 3.1.0 版 - BigInsights 和 Hortonworks

如果您已安裝 Analytic Server 3.0.1，則可以就地將安裝升級至 3.1.0 版。

1. 在 Ambari 主控台中，停止 Analytic Server 服務。
2. 根據您的安裝類型，遵循下列步驟。

線上升級

- a. 請確保您的 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>。
- b. 在每個 Analytic Server 主機上，從 `https://ibm-open-platform.ibm.com/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms/IBM-SPSS-AnalyticServer-3.1.0.0.repo` 下載 IBM-SPSS-AnalyticServer-3.1.0.0.repo 檔 (x86 和 ppc64le)，並將其移至資料夾 /etc/yum.repos.d (RHEL 或 CentOS) 或 /etc/zypp/repos.d (SLES)。

離線升級

- a. 離線升級將會下載必要的 RPM 檔案，且應該執行於可存取 `https://ibm-open-platform.ibm.com` 的機器上。
- b. 建立新目錄，以充當 Analytic Server RPM 檔案的儲存庫。請參閱下列範例：

```
mkdir /home/root/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms
```
- c. 將必要的 Analytic Server RPM 檔案複製到此目錄中。您所需的 RPM 檔案視您的發行套件、版本及架構而定。若為 BigInsights 4.2，必要檔案如下所示。

表 5. *BigInsights 4.2 RPM*

BigInsights 4.2 (x86_64)
IBM-SPSS-AnalyticServer-ambari-2.x-3.1.0.0-1.noarch.rpm

- d. 建立本端儲存庫定義。例如，在 `/etc/yum.repos.d/`（針對 RHEL、CentOS）或 `/etc/zypp/repos.d/`（針對 SLES）中建立名為 `analyticserver.repo` 的檔案，且包含下列內容。

```
[IBM-SPSS-AnalyticServer]
name=IBM-SPSS-AnalyticServer-3.1.0.0
baseurl=file:///{{本端儲存庫的路徑}}
enabled=1
gpgcheck=0
protect=1
```

- e. 建立本端 yum 儲存庫。請參閱下列範例：

```
createrepo /home/root/repos/IBM-SPSS-AnalyticServer/3.1.0.0/rpms
```

3. 從本端快取清除 Ambari meta 資料。例如，若要在 RHEL 或 CentOS 上清除快取，請執行下列指令：

```
sudo yum clean all
```

註：當列出兩個 Analytic Server 儲存庫時，yum 指令不工作。結果，必須重新命名或移除與 Analytic Server 相關的原始 *.repo 檔案。在 SLES 上，如下變更該指令：

```
sudo zypper refresh
```

4. 在每一個 Analytic Server 主機上，升級 RPM。例如，若要在 RHEL 或 CentOS 上進行升級，請執行下列指令：

```
chown -R as_user:hadoop /opt/ibm/spss/analyticserver/3.0
sudo yum upgrade IBM-SPSS-AnalyticServer
```

在 SLES 上，如下變更該指令：

```
sudo zypper up IBM-SPSS-AnalyticServer
```

5. 重新整理堆疊。

BigInsights

- a. 在 Ambari 主控台中，啟動然後停止 Analytic Server 服務。
- b. 執行自訂重新整理動作。

Hortonworks

導覽至其中一個 Analytic Server 節點，並執行下列指令：

```
sudo -u as_user /opt/ibm/spss/analyticserver/3.1/bin/refresh.sh
```

6. 僅限離線安裝。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:///{{本端儲存庫的路徑}}</baseurl>
    <repoid>IBM-SPSS-AnalyticServer</repoid>
    <reponame>IBM-SPSS-AnalyticServer-3.1.0.0</reponame>
  </repo>
</os>
```

7. 清除 Zookeeper 狀態。在 Zookeeper bin 目錄中執行下列指令（例如，`/usr/iop/current/zookeeper-server/bin`）：

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

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

移轉至新版本的 **Analytic Server**

如果您已安裝 Analytic Server 2.0 或 2.1，而且購買了 3.1.0，則可以將 2.0/2.1 配置設定移轉至 3.1.0 安裝。

限制：

- 如果安裝的版本比 2.0 版更舊，則必須先從舊版移轉至 2.0/2.1 版，然後再從 2.0/2.1 版移轉至 3.1.0 版。
- 2.0/2.1 版及 3.1.0 版安裝無法同時存在於同一個 Hadoop 叢集上。如果將 3.1.0 安裝配置成使用和 2.0/2.1 安裝相同的 Hadoop 叢集，則 2.0/2.1 安裝就不再起作用。

移轉步驟：2.0/2.1 至 3.1.0

1. 根據第 5 頁的『在 Ambari 上安裝』中的指示，進行 Analytic Server 的新安裝。
2. 將舊安裝中的分析根目錄複製到新安裝中。
 - a. 如果您不確定分析根目錄的位置，請執行 `hadoop -fs ls`。分析根目錄的路徑將為 `/user/aeuser/analytic-root` 格式，其中 `aeuser` 是擁有分析根目錄的使用者 ID。
 - b. 將分析根目錄的所有權從 `aeuser` 變更為 `as_user`

```
hadoop dfs -chown -R {as_user:{group}} {path to 2.0/2.1 analytic-root}
```
- 註：如果您計劃在移轉之後使用現有的 Analytic Server 安裝，請在 HDFS 中製作 `analytic-root` 目錄的副本，然後變更該目錄副本的所有權。
- c. 使用 `as_user` 身分登入新 Analytic Server 安裝的主機。如果 `/user/as_user/analytic-root` 目錄已存在，將其刪除。
- d. 執行下列副本 Script。

```
hadoop distcp hftp://{host of 2.0/2.1 namenode}:50070/{path to 2.0/2.1 analytic-root}
hdfs://{host of 3.1.0 namenode}/user/as_user/analytic-root
```
3. 在 Ambari 主控台中，停止 Analytic Server 服務。
4. 確保 Analytic Metastore 服務正在執行。
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。
6. 清除 Zookeeper 狀態。在 Zookeeper bin 目錄中（例如，Hortonworks 上的 `/usr/hdp/current/zookeeper-client` 或 BigInsights 上的 `/usr/iop/current/zookeeper-server`），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/3.1/ASConfiguration_2.1.0.0.xxx.zip
```
8. 在 Ambari 主控台中，啟動 Analytic Server 服務。

註：如果配置了 R 以與現有 Analytic Server 安裝搭配使用，則將需要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『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。啟用安全模式。

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 BigInsights 4.2 中，ESSENTIALR 目錄位於 `/var/lib/ambari-server/resources/stacks/BigInsights/4.2/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 包含 Hadoop 的主要核心元素，以對大型資料集（主要是 MapReduce 及 HDFS）提供可靠的可調式分散式資料處理功能，還包含其他企業導向的元件，以提供安全性、高可用性及與硬體和其他軟體的整合。

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

meta 資料儲存庫

如果您計劃使用 MySQL 作為 Analytic Server 的 meta 資料儲存庫，請遵循『針對 Analytic Server 配置 MySQL』的指示。

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 30 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> 及 -t <dbUserName> 參數。

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 3.1.0

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：

表 6. Analytic Server 自行解壓縮的二進位檔

說明	二進位檔名
IBM SPSS Analytic Server 3.1 for Cloudera 5.8、5.9 和 5.10 Ubuntu (英文)	spss_as-3.1-cdh5.8-5.10-ubun_en.bin
IBM SPSS Analytic Server 3.1 for Cloudera 5.8、5.9 和 5.10 Linux x86-64 (英文)	spss_as-3.1-cdh5.8-5.10-lx86-en.bin

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. 在安裝完成之後重新啟動 Cloudera Manager。
4. 開啟 Cloudera Manager 介面（例如，[http://\\${CM_HOST}:7180/cm/login](http://${CM_HOST}:7180/cm/login)，其使用預設登入認證 admin/admin），重新整理遠端包裹儲存庫 URL，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。

5. 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。
6. 選取下載 > 配送 > 啟動。 **AnalyticServer** 包裹狀態會更新為已配送、已啟動。
7. 針對 Analytic Server 配置 MySQL。
8. 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：
 - 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

「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。

註：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-3.1.0.0-el6.parcel
- AnalyticServer-3.1.0.0-el6.parcel.sha
- manifest.json

或

- AnalyticServer-3.1.0.0-el7.parcel
- AnalyticServer-3.1.0.0-el7.parcel.sha

SuSE Linux 需要下列檔案：

- AnalyticServer-3.1.0.0-sles11.parcel
- AnalyticServer-3.1.0.0-sles11.parcel.sha
- manifest.json

Ubuntu Linux 需要下列檔案：

- AnalyticServer-3.1.0.0-trusty.parcel
- AnalyticServer-3.1.0.0-trusty.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）。

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。

3. 按一下配送 > 啟動。

AnalyticServer 包裹會顯示為已配送及已啟動。

在 Cloudera 上升級至 Analytic Server 3.1.0

如果您已安裝 Analytic Server 3.0/3.0.1，則可以就地將安裝升級至 3.1.0 版。

1. 在 Cloudera Manager 中，停止然後刪除 Analytic Server 服務。
2. 在 Cloudera Manager 中，停用舊版 Analytic Server。
3. 如需安裝 Analytic Server 3.1.0 的指示，請參閱第 24 頁的『升級和移轉』中的「線上」或「離線」一節。
4. 在 Cloudera Manager 中安裝和新增 Analytic Server 服務之後，請執行重新整理 **Analytic Server** 二進位檔。Analytic Server 3.1.0 現已備妥使用。

配置 Cloudera

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

security_cfg 參數定義可以作為主體新增至 Analytic Server 系統之使用者和群組的登錄。

依預設，基本登錄定義為使用單一使用者 admin 和密碼 admin。您可以透過編輯 **security_cfg** 或配置 Kerberos 作為安全提供者，來變更登錄。在 Analytic Server 服務之「配置」標籤的 **Analytic Server 進階配置 Snippet** 區段中，可以找到 **security_cfg** 參數。

註：如果您編輯 **security_cfg** 參數以變更登錄，則需要將任何新使用者作為主體新增至 Analytic Server 系統。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》。

對基本登錄進行變更

基本登錄可讓您在 **security_cfg** 參數中定義使用者和群組的資料庫。

預設基本登錄類似如下。

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="admin"/>
</basicRegistry>
```

變更後的基本登錄範例如下所示。

```
<basicRegistry id="basic" realm="ibm">
  <user name="user1" password="{xor}Dz4sLG5tbGs="/>
  <user name="user2" password="Pass"/>
  <user name="user3" password="Pass"/>
  <user name="user4" password="Pass"/>
  <user name="admin" password="{xor}KzosKw="/>
  <group name="Development">
    <member name="user1"/>
    <member name="user2"/>
  </group>
  <group name="QA">
    <member name="user3"/>
    <member name="user4"/>
  </group>
  <group name="ADMIN">
    <member name="user1"/>
    <member name="admin"/>
  </group>
</basicRegistry>
```

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。

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

註：如需 securityUtility 工具的詳細資料，請參閱 http://www-01.ibm.com/support/knowledgecenter/SSD28V_8.5.5/com.ibm.websphere.wlp.core.doc/ae/rwlp_command_securityutil.html。

註：在沙盤推演環境中基本登錄非常有用，但不建議在正式作業環境中使用。

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。

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>
```

如需更多配置範例，請參閱範本資料夾 {AS_ROOT}/ae_wlpserver/templates/config。

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 /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。

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。

註：如果 Analytic Server 安裝使用基本登錄，則它必須包含 Kerberos 使用者帳戶（使用 "-" 作為密碼）。範例如下所示。

```
<basicRegistry id="basic" realm="ibm">
  <user name="admin" password="-"/>
  <user name="user1" password="-"/>
  <user name="user2" password="-"/>
  <group name="group1">
    <member name="admin"/>
    <member name="user1"/>
    <member name="user2"/>
  </group>
  <group name="group2">
    <member name="admin"/>
    <member name="user1"/>
  </group>
</basicRegistry>
```

2. 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。
 - 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。
3. 針對 Analytic Server 中的所有主體，在 HDFS 上建立使用者起始檔案夾。例如，如果您將 `testuser1` 新增至 Analytic Server 系統，則在 HDFS 上建立 `/user/testuser1` 之類的起始檔案夾，並確保 `testuser1` 擁有此資料夾的讀取權和寫入權。
4. 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：如果尚未設定下列參數，則在「配置」標籤上，可能不會顯示這些參數。在此情況下，請執行搜尋來尋找它們。
 - 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 Cloudera Manager 並在 Analytic Server **analyticserver-conf/config.properties** 的 **Analytic Server 進階配置 Snippet (安全閥)** 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。

現在，使用者可以使用 Kerberos SSO 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. 開啟 Cloudera Manager 並在 Analytic Server **core-site.xml** 的分析伺服器進階配置 **Snippet (安全閥)** 區域中，新增或更新下列內容。

```
- hadoop.proxyuser.as_user.hosts = *
- hadoop.proxyuser.as_user.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）。

註：Ambari 中會自動新增內容（根據 Analytic Server 配置中的值）。

停用 Kerberos

1. 在 Ambari 主控台中停用 Kerberos。
2. 停止 Analytic Server 服務。
3. 從 **analyticserver-conf/config.properties** 的 **Analytic Server 進階配置 Snippet (安全閥)** 區域中移除下列參數。

```
default.security.provider
hdfs.keytab
hdfs.user
java.security.krb5.conf
as.db.connect.method
web.authentication.kerberos.keytab
web.authentication.kerberos.principal
```

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。

若要安裝第三方供應商憑證，請遵循下列步驟。

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=myserver --password=mypassword --validity=365 --subject=CN=mycompany,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。

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

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. 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。
2. 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++
3. 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 3.1.0。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 2.15.3，為 <https://cran.r-project.org/src/base/R-2/R-2.15.3.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）。請提供指向已複製 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-2.15.3.tar.gz 保存檔，並將其複製到 /root，則 URL 為 `file:///root/R-2.15.3.tar.gz`。

註：其他 R 版本可在 <https://cran.r-project.org/src/base/> 中找到。

- 安裝 R 所需的套件。
 - 下載並安裝 R 及 Essentials for R 外掛程式。
 - 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。
- 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）。
 - 如果已安裝 Analytic Server 服務：
 - 停止該服務。
 - 重新整理 Analytic Server 二進位檔。
 - 啟動該服務，以完成 Essentials for R 安裝。
 - 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

如果您在每個 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 Cloudera Manager 中，導覽至 Analytic Server 服務的「配置」標籤。
- 在 `jdbc.drivers.location` 中指定 JDBC 驅動程式之共用目錄的路徑。
- 按一下儲存變更。
- 從動作下拉清單中選取停止，以停止 Analytic Server 服務。
- 從動作下拉清單中選取重新整理 **Analytic Server** 二進位檔。
- 從動作下拉清單中選取啟動，以啟動 Analytic Server 服務。

表 7. 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
DashDB	Bluemix 服務	db2jcc.jar	IBM
DB2 for Linux, UNIX, and Windows	10.5、10.1、9.7	db2jcc.jar	IBM
DB2 z/OS	11、10	db2jcc.jar、db2_license_cisuz.jar	IBM
Greenplum	5、4.2.x	postgresql.jar	Greenplum
Hive	1.1、1.2	hive-jdbc-*.jar	Apache

表 7. 支援的資料庫 (繼續)

資料庫	支援的版本	JDBC 驅動程式 JAR	供應商
Netezza	7、6.x	nzjdbc.jar	IBM
Oracle	12c、11g R2 (11.2)	ojdbc6.jar、orai18n.jar	Oracle
SQL Server	2014、2012、2008 R2	sqljdbc4.jar	Microsoft
Sybase IQ	16.x、15.4、15.2	jconnect70.jar	Sybase
Teradata	14、14.1、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 {HIVE_HOME}/auxlib 目錄以及每一個 Analytic Server 節點上的 /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。

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>。

Apache Spark

如果您要搭配使用 Spark (1.5 版或更新版本) 與 HCatalog 輸入資料來源，則必須手動新增 spark.version=X.X.0 內容 (例如，spark.version=2.0.0)

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

spark.version=2.0.0

2. 儲存配置，並從 Cloudera Manager 重新啟動所有 Analytic Server 服務。

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 Apache Impala 的資料

1. 在 Analytic Server 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)

輸入自簽的 PEM 格式的憑證位置及檔名 (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。

最佳化小型資料的 JVM 選項

您可以編輯 JVM 內容，以最佳化執行較小 (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，而且可能需要嘗試使用這些值以最佳化系統。

```
-Xms512M  
-Xmx2048M
```

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

如果您具有 Analytic Server 2.0/2.1 的現有安裝，而且購買了更新的版本，您可以將 2.0/2.1 配置設定移轉至新安裝。

限制：如果安裝的版本比 2.0 版更舊，則必須先從較舊的版本移轉至 2.0/2.1 版，然後再從 2.0/2.1 版移轉至較新的版本。

限制：2.0/2.1 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 2.0/2.1 安裝使用相同的 Hadoop 叢集，則 2.0/2.1 安裝就不再起作用。

移轉步驟：2.1 至較新版本

1. 根據第 30 頁的『在 Cloudera 上安裝』中的指示，進行 Analytic Server 的新安裝。
2. 將舊安裝中的分析根目錄複製到新安裝中。
 - a. 如果您不確定分析根目錄的位置，請執行 `hadoop -fs ls`。分析根目錄的路徑將為 `/user/aeuser/analytic-root` 格式，其中 `aeuser` 是擁有分析根目錄的使用者 ID。
 - b. 將分析根目錄的所有權從 `aeuser` 變更為 `as_user`

```
hadoop dfs -chown -R {as_user:{group}} {path to 2.1 analytic-root}
```

註：如果您計劃在移轉之後使用現有的 Analytic Server 安裝，請在 HDFS 中製作 `analytic-root` 目錄的副本，然後變更該目錄副本的所有權。
 - c. 使用 `as_user` 身分登入新 Analytic Server 安裝的主機。如果 `/user/as_user/analytic-root` 目錄已存在，將其刪除。
 - d. 執行下列副本 Script。

```
hadoop distcp http://{host of 2.1 namenode}:50070/{path to 2.1 analytic-root}
hdfs://{host of 3.1 namenode}/user/as_user/analytic-root
```
3. 在 Cloudera Manager 中，停止 Analytic Server 服務。
4. 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。
5. 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/3.1/ASConfiguration_2.1.0.0.xxx.zip
```
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. 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. 取消啟動並從主機移除 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。

第 4 章 MapR 安裝與配置

MapR 概觀

MapR 是 Apache Hadoop 的完整發行套件，其透過 Hadoop 生態系統，包裝超過十多個專案，以提供廣泛的海量資料功能集。

在伺服器叢集之外，無法存取 MapR 檔案系統。因此，必須在 MapR 叢集節點中部署 IBM SPSS Analytic Server。在此部署實務範例中，Analytic Server 必須由擁有下列權限的使用者執行：有權存取 MapR 檔案系統且有權將工作提交至 Yarn 以部署至 Analytic Server（以 <as_user> 身分）。

在 MapR 上安裝 Analytic Server

下列步驟詳細說明在 MapR 叢集上手動安裝 IBM SPSS Analytic Server 的處理程序。

在 MapR 5.0 或 5.1 上安裝 Analytic Server 3.1.0

1. 導覽至 [IBM Passport Advantage® 網站](#) 並下載 MapR 自行解壓縮的二進位檔。

表 8. MapR 自行解壓縮的二進位檔

說明	二進位檔名
IBM SPSS Analytic Server 3.1 for MapR 5.0 和 5.1 Linux x86-64 (英文)	spss_as-3.1.0-mapr5.0-5.1-lx86_en.bin

2. 以 root 使用者身分或 sudo 使用者身分執行 Analytic Server 安裝程式。遵循安裝提示，接受授權，並選擇線上或離線安裝 Analytic Server。
 - a. 當管理 Analytic Server 的伺服器具有 <https://ibm-open-platform.ibm.com> 的網際網路連線時，請選取線上選項。安裝程式會自動安裝 Analytic Server。
 - b. 當管理 Analytic Server 的伺服器沒有 <https://ibm-open-platform.ibm.com> 的網際網路連線時，請選取離線選項。在具有該 URL 存取權的另一部伺服器上執行安裝程式，並選擇離線安裝 Analytic Server。安裝程式會自動下載 RPM 或 DEB 套件。
3. 針對 Analytic Server，尋找並執行 RPM 或 DEB：

- RedHat 或 SuSe Linux：

```
rpm -ivh IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
```

- Ubuntu Linux：

```
dpkg -i IBM-SPSS-AnalyticServer_1_amd64.deb
```

無論是線上還是離線安裝模式，Analytic Server 都會安裝至 /opt/ibm/spss/analyticserver/3.1（作為 <as_installation_path>）。

4. 將安裝路徑中的所有檔案變更為執行 Analytic Server 的使用者：

```
chown -R <as_user> <as_installation_path>
```

將使用者切換至 <as_user>；所有先前的步驟都使用 <as_user>。

5. 配置 HTTP 內容。在路徑 <as_installation_path>/ae_wlpserver/usr/servers/aeserver 中建立名為 http_endpoint.xml 的檔案，並新增下列行至該檔案：

```
<server>
  <httpEndpoint host="*" id="defaultHttpEndpoint" httpPort="<http_port>" httpsPort="<https_port>" onError="FAIL"/>
</server>
```

<http_port> 及 <https_port> 是 Analytic Server 透過 HTTP 及 HTTPS 通訊協定使用的埠。請將它們取代為任何可用的埠。

6. 新增使用者及群組。在路徑 <as_installation_path>/ae_wlpserver/usr/servers/aeserver 中建立名為 security_cfg.xml 的檔案，並新增下列行至該檔案：

```
<server>
  <basicRegistry id="basic" realm="ibm">
    <user name="admin" password="test"/>
  </basicRegistry>
</server>
```

在預設狀態中，XML 檔案僅包含管理使用者。您必須在 <basicRegistry> 設定中手動新增其他使用者和群組，或者將設定變更為 ldapRegistry。

7. 設定 meta 資料資料庫。Analytic Server 支援 DB2 及 MySQL 資料庫。

- a. 配置資料庫使用者。使用 MySQL 資料庫時，請在 MySQL Shell 中執行下列 SQL Script：

```
DROP DATABASE IF EXISTS <db_name>;
CREATE DATABASE <db_name> DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_bin;
CREATE USER '<db_username>@%' IDENTIFIED BY '<db_password>';
CREATE USER '<db_username>@localhost' IDENTIFIED BY '<db_password>';
GRANT ALL PRIVILEGES ON *.* TO '<db_username>@%';
GRANT ALL PRIVILEGES ON *.* TO '<db_username>@localhost';
```

- b. 加密密碼。資料庫使用者的密碼必須已加密，然後才能傳遞至 Analytic Server。執行下列指令：

```
java -Duser.language=en -cp <as_installation_path>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*
com.spss.ae.encryption.provider.EncryptKeystorePassword <db_password>
```

註：當指令直接在 Linux Shell 中執行時，字元 * 可能需要跳出為 *。

指令輸出讀作：The encrypted password is '<encrypted_db_password>'。請記錄已加密的資料庫密碼。

- c. 刪除 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties 檔案（如果存在的話），並使用相同名稱建立一個新檔案。使用 DB2 資料庫時，變更下列內容：

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:db2://<db_host>:<db_port>/<db_name>:currentSchema=<db_schema_name>;
jndi.aedb.driver=com.ibm.db2.jcc.DB2Driver
jndi.aedb.username=<db_username>
jndi.aedb.password=<encrypted_db_password>
```

如果 <db_schema_name> 綱目不存在，則使用者 <db_username> 必須具有建立該綱目的隱含許可權。使用 MySQL 資料庫時，變更下列內容：

```
jndi.aedb=jdbc/aeds
jndi.aedb.url=jdbc:mysql://<db_host>:<db_port>/<db_name>?createDatabaseIfNotExist=true
jndi.aedb.driver=com.mysql.jdbc.Driver
jndi.aedb.username=<db_username>
jndi.aedb.password=<encrypted_db_password>
```

- d. 使用 MySQL 資料庫時，必須安裝 MySQL JDBC 驅動程式。執行下列指令：

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

- e. 執行下列指令，以建立所需表格：

```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/sql/<db_type>
java -Xmx128m -Xms128m -cp <as_installation_path>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*:/usr/share/java/*
com.spss.ae.dbscript.ScriptRunner ../../configuration/config.properties schema.sql true
```

<db_type> 是 db2 或 mysql，具體取決於所使用的資料庫。

註：當搭配使用 MySQL 與 MYISAM 引擎時，第二個指令會報告下列錯誤訊息，您可放心地忽略它們：


```

Error executing: set global innodb_large_prefix=ON
java.sql.SQLException: Unknown system variable 'innodb_large_prefix'
Error executing: set global innodb_file_format=BARRACUDA
java.sql.SQLException: Unknown system variable 'innodb_file_format'
Error executing: set global innodb_file_format_max=BARRACUDA
java.sql.SQLException: Unknown system variable 'innodb_file_format_max'
Error executing: set global innodb_file_per_table=TRUE
java.sql.SQLException: Variable 'innodb_file_per_table' is a read only variable

```

- 執行下列指令，以解壓縮 cf 程式庫。

```

cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration
unzip cf.zip

```

- 透過在路徑 <as_installation_path>/ae_wlpserver/usr/servers/aeserver 中建立名為 private_library.xml 的檔案，以配置 JAAS 登入模組類別路徑，並在該檔案中，輸入下列資訊：

```

<server>
<library id="maprLib">
<fileset dir="{wlp.install.dir}/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib" includes="*.jar"/>
<fileset dir="/usr/share/java" includes="*.jar"/>
<folder dir="/opt/mapr/hadoop/hadoop-2.7.0/etc/hadoop"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common/lib" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/hdfs" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/hdfs/lib" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/yarn" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/yarn/lib" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce" includes="*.jar"/>
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce/lib" includes="*.jar"/>
</library>
<jaasLoginModule id="maprLoginModule1" className="org.apache.hadoop.security.login.GenericOSLoginModule"
controlFlag="REQUIRED" libraryRef="maprLib"></jaasLoginModule>
<jaasLoginModule id="maprLoginModule2" className="org.apache.hadoop.security.login.HadoopLoginModule"
controlFlag="REQUIRED" libraryRef="maprLib"></jaasLoginModule>
<jaasLoginContextEntry id="hadoop_simple" name="hadoop_simple" loginModuleRef="maprLoginModule1,maprLoginModule2" />
<application context-root="/analyticsserver" id="AS_BOOT" location="AE_BOOT.war" name="AS_BOOT" type="war">
<classloader commonLibraryRef="maprLib"></classloader>
</application>
<application id="help" location="help.war" name="help" type="war" context-root="/analyticsserver/help"/>
</server>

```

註：前一個範例用於配置 hadoop_simple login 模組。當 MapR 使用其他登入模組時，必須變更配置。

- 驗證檔案 ASModules.xml 在路徑 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/ 中是否存在。如果檔案不存在，請將檔案 ASModules.xml.template（在相同的路徑中）重新命名為 ASModules.xml
- 透過在檔案 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties 中新增下列內容，以配置叢集資訊。

```

ae.cluster.zookeeper.connect.string=
ae.cluster.member.name=
ae.cluster.collective.name=mapr_5.1

```

ae.cluster.zookeeper.connect.string 內容是以逗點區隔的 zookeeper 節點清單。該內容可以共用 MapR 所使用的 zookeeper 叢集。ae.cluster.member.name 是管理 Analytic Server 之節點的主機名稱。

下列範例示範 ae.cluster.zookeeper.connect.string 格式：

```

ae.cluster.zookeeper.connect.string=<zookeeper host 1>:<zookeeper port 1>,<zookeeper host 2>:<zookeeper port 2>,<zookeeper host 3>:<zookeeper port 3>...

```

當 Analytic Server 與 MapR 共用同一 zookeeper 叢集時，ae.cluster.zookeeper.connect.string 值必須與 MapR warden.conf 檔（該檔案的預設位置為 /opt/mapr/conf）中的 zookeeper.servers 內容相同。

- 開啟檔案 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/server.env，並新增下列行至該檔案：

```

JAVA_HOME=<java_home>
PATH=<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<java_home>/jre/lib/amd64:/usr/sbin:/usr/bin:/sbin:/bin
IBM_SPSS_AS_NATIVE_PATH=<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64
LD_LIBRARY_PATH=<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<java_home>/jre/lib/amd64:/opt/mapr/hadoop/hadoop-2.7.0/lib/native

```

將 `<as_installation_path>` 及 `<java_home>` 取代為實際的安裝路徑及 Java 起始目錄路徑。

13. 透過開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`，並新增下列行，以編輯 Analytic 根目錄：

```
distrib.fs.root=<analytic_root>
```

`<analytic_root>` 是管理必要 Analytic Server 遠端檔案之 MapR 檔案系統中的路徑。建議的路徑為 `/user/<as_user>/analytic-root`。

14. 透過開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`，並新增下列行，來設定管理者使用者：

```
admin.username=admin
```

該值必須是 Analytic Server 管理者使用者名稱，且必須是在 `security_cfg.xml` 檔案中配置的其中一個使用者。

15. 透過在檔案 `<as_installation_path>/bin/hdfsUpdate.sh` 第 69 行中新增下列行，將 Analytic Server 相依關係上傳至 MapR 檔案系統：

```
JAVA_CLASS_PATH=`hadoop classpath`:$JAVA_CLASS_PATH
```

執行下列指令，以建立 `<analytic_root>`：

```
cd <as_installation_path>/bin
./hdfsUpdate.sh
```

`<as_user>` 必須具有 `<analytic_root>` 上層目錄的寫入權。

16. 啟動和停止 Analytic Server。
 - a. 執行下列指令，以啟動 Analytic Server：

```
cd <as_installation_path>/ae_wlpserver/bin
./server start aeserver
```

- b. 執行下列指令，以停止 Analytic Server：

```
cd <as_installation_path>/ae_wlpserver/bin
./server stop aeserver
```

在 MapR 5.2 上安裝 Analytic Server 3.1.0

1. 導覽至 [IBM Passport Advantage® 網站](#) 並下載 MapR 自行解壓縮的二進位檔。

表 9. MapR 自行解壓縮的二進位檔

說明	二進位檔名
IBM SPSS Analytic Server 3.1.0.0 for MapR 5.0, 5.1 和 5.2 Linux x86-64 (英文)	spss_as3.1.0.0-mapr5-5.2-lx86_en.bin

2. 用於安裝 Analytic Server 的剩餘步驟大部分與用於在 MapR 5.0 或 5.1 上安裝 Analytic Server 3.1.0 的那些步驟相同。但是，MapR 5.1 與 5.2 之間的第 48 頁的『啟用 Apache HBase』和 第 49 頁的『啟用 Apache SPARK』資訊不同。請參閱這些主題以瞭解有關在 MapR 5.2 上安裝的資訊。

配置 MapR

安裝之後，您可以選擇性地配置及管理 Analytic Server MapR 功能。

啟用資料庫推回

資料庫推回實際上是從資料庫讀取資料，並直接對資料進行處理。

IBM SPSS Analytic Server 支援下列資料庫的推回：

- DashDB
- DB2
- DB2 for Z
- Hive
- MySQL
- Netezza
- Oracle
- PostgreSQL
- Redshift
- SQL Server
- Sybase IQ
- Teradata

使用下列步驟，以啟用資料庫推回。

1. 將適當的 JDBC 驅動程式 JAR 檔複製到 `<as_installation_path>/jdbc`。
2. 開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml`，找到 ID 為 `maprLib` 的標籤庫，並在標籤中新增下列行：

```
<fileset dir="<as_installation_path>/jdbc" includes="*.jar"/>
```

3. 執行下列指令：

```
cd <as_installation_path>/jdbc
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

4. 重新啟動 Analytic Server。

啟用 Apache Hive

Apache Hive 是在 Hadoop 基礎之上建置的資料倉儲基礎架構，用於提供資料摘要、查詢及分析。

註：Hive 必須配置為使用 MySQL 作為 meta 儲存庫。在管理 IBM SPSS Analytic Server 之節點上存在的 `hive-site.xml` 檔案應該與執行 Hive Metastore 之節點中的檔案相同。

若要在順利安裝 MapR 之後啟用 Apache Hive 支援：

1. 透過執行下列指令，將 Hive 及 hcatalog 相依關係上傳至 MapR 檔案系統：

```
cd /opt/mapr/hive/hive-1.2/lib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
cd /opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

`<as_analytic_root>` 是第 43 頁的『在 MapR 上安裝 Analytic Server』中定義的 Analytic 根路徑。

2. 開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml`，找到 ID 為 `maprLib` 的標籤庫，並在標籤中新增下列行：

```
<fileset dir="/opt/mapr/hive/hive-1.2/lib" includes="*.jar"/>
<fileset dir="/opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog" includes="*.jar"/>
```

3. 執行下列指令，以建立 Hive 及 hcatalog 配置檔鏈結：

```
mkdir <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hive-conf
ln -s /opt/mapr/hive/hive-1.2/conf/* <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hive-conf
```

4. 當 Hive 的 `auxlib` 中存在額外的 Jar 檔時，將下列行新增至 `private_library.xml` 檔案：

```
<fileset dir="/opt/mapr/hive/hive-1.2/auxlib" includes="*.jar"/>
```

新增上一行之後，執行下列指令：

```
cd /opt/mapr/hive/hive-1.2/auxlib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

5. 重新啟動 Analytic Server。

以 HTTP 模式執行 Hive

依預設，以二進位模式（TCP 模式）執行 Hive。若要以 HTTP 模式執行 Hive，您必須更新下列 Hive 配置內容（特別是 `hive.server2.transport.mode` 內容）。

註：如需每一個內容的相關資訊，請參閱 Hive 配置內容。

表 10. HTTP 模式的 Hive 內容

內容名稱	預設值	說明
<code>hive.server2.transport.mode</code>	binary	伺服器的傳輸模式。值可以是 binary 或 http。設為 http 以啟用 HTTP 傳輸模式。
<code>hive.server2.thrift.http.port</code>	10001	HTTP 模式中的埠號。
<code>hive.server2.thrift.http.path</code>	cliservice	處於 HTTP 模式時，URL 端點的路徑元件。
<code>hive.server2.thrift.http.min.worker.threads</code>	5	處於 HTTP 模式時，伺服器儲存區中工作程式執行緒的數目下限。
<code>hive.server2.thrift.http.max.worker.threads</code>	500	處於 HTTP 模式時，伺服器儲存區中工作程式執行緒的數目上限。

註：更新內容之後，必須重新啟動 Hive。

啟用 Apache HBase

Apache HBase 是以 Java 撰寫的開放程式碼、非關聯式分散式資料庫。它是作為 Apache Software Foundation Apache Hadoop 專案的一部分進行開發的，並在 HDFS（Hadoop 分散式檔案系統）基礎之上執行。

若要在順利安裝 MapR 之後啟用 Apache HBase 支援：

MapR 5.0/5.1 上的 IBM SPSS Analytic Server 3.1.0

1. 將 HBase 相依關係上傳至 MapR 檔案系統，並執行下列指令：

```
cd /opt/mapr/hbase/hbase-0.98.12/lib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

<as_analytic_root> 是第 43 頁的『在 MapR 上安裝 Analytic Server』中定義的 Analytic 根路徑。

2. 開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml`，找到 ID 為 `maprLib` 的標籤庫，並在標籤中新增下列行：

```
<fileset dir="/opt/mapr/hbase/hbase-0.98.12/lib" includes="*.jar"/>
```

3. 執行下列指令，以建立 HBase 及 hcatalog 配置檔鏈結：

```
mkdir <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
ln -s /opt/mapr/hbase/hbase-0.98.12/conf/* <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
```

4. 重新啟動 IBM SPSS Analytic Server。

MapR 5.2 上的 IBM SPSS Analytic Server 3.1.0

1. 透過執行下列指令將 HBase 相依關係上傳至 MapR 檔案系統：

```
cd /opt/mapr/hbase/hbase-1.1.1/lib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

<as_analytic_root> 是在第 43 頁的『在 MapR 上安裝 Analytic Server』的第 12 步設定的路徑。

- 開啟 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml` 並尋找 ID 為 `maprLib` 的標籤庫。將下列行新增至標籤：


```
<fileset dir="/opt/mapr/hbase/hbase-1.1.1/lib" includes="*.jar"/>
```
- 執行下列指令以建立 Hive 與 HCatalog 配置檔案的鏈結：


```
mkdir <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
ln -s /opt/mapr/hbase/hbase-1.1.1/conf/* <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/hbase-conf
```
- 將下列行新增至 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties`

```
spark.executor.extraClassPath=/opt/mapr/hbase/hbase-1.1.1/lib/*
```
- 重新啟動 Analytic Server。

啟用 Apache SPARK

Apache Spark 是一種開放式標準，用於對記憶體內資料進行靈活的處理，以供進行批次、即時的進階分析。

若要在順利安裝 MapR 之後啟用 Apache Spark 支援：

MapR 5.0/5.1 上的 IBM SPSS Analytic Server 3.1.0

- 將檔案 `spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar` 從 `/opt/mapr/spark/spark-1.4.1/lib` 複製到 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/`。
- 將 Spark 相依關係上傳至 MapR 檔案系統，並執行下列指令：


```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

`<as_analytic_root>` 是第 43 頁的『在 MapR 上安裝 Analytic Server』中定義的 Analytic 根路徑。
- 開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml`，找到 ID 為 `maprLib` 的標籤庫，並在標籤中新增下列行：


```
<fileset dir="/opt/mapr/spark/spark-1.4.1/lib" includes="spark-assembly-*.jar"/>
```
- 執行下列指令，以建立 Spark 配置檔鏈結：


```
mkdir <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
ln -s /opt/mapr/spark/spark-1.4.1/conf/* <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
```
- 在檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/server.env` 中新增下列行：


```
SPARK_HOME=/opt/mapr/spark/spark-1.4.1
```
- 在檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties` 中新增下列行：


```
spark.executor.extraLibraryPath=/opt/mapr/hadoop/hadoop-2.7.0/lib/native
```
- 重新啟動 IBM SPSS Analytic Server。
- 若要啟用 PySpark 功能，請在 `yarn-env.sh` 檔案中新增下列行，然後重新啟動 ResourceManagers 及 NodeManagers：


```
export SPARK_HOME=/opt/mapr/spark/spark-1.4.1
```

MapR 5.2 上的 IBM SPSS Analytic Server 3.1.0

步驟根據 Spark 版本不同而異。

Spark 1.x

- 將 `spark-assembly-1.4.1-hadoop2.5.1-mapr-1501.jar` 檔從 `/opt/mapr/spark/spark-1.4.1/lib`

複製到

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/
```

2. 開啟下列檔案：

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml
```

並尋找 ID 為 maprLib 的標籤庫。將下列行新增至標籤：

```
<fileset dir="/opt/mapr/spark/spark-1.4.1/lib" includes="spark-assembly-*.jar"/>
```

3. 刪除下列檔案：

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.sparkmapreduce_2-3.1.0.0.jar
```

Spark 2.x

1. 刪除下列檔案：

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/com.ibm.spss.sparkmapreduce-3.1.0.0.jar
```

2. 開啟下列檔案：

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/private_library.xml
```

並尋找 ID 為 maprLib 的標籤庫。將下列行新增至標籤：

```
<fileset dir="/opt/mapr/spark/spark-2.0.1/jars" includes="*.jar"/>
```

```
<fileset dir="/opt/mapr/spark/spark-2.0.1/scala/lib" includes="*.jar"/>
```

```
<fileset dir="<as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark" includes="*.jar"/>
```

3. 將行 spark.version=2.0 新增至下列檔案：

```
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties
```

下列步驟通用於 Spark 1.x 和 2.x，必須在之前的 Spark 1.x 或 2.x 步驟之後完成。

註：對 <spark_version> 的所有參照都必須取代為實際的 Spark 版本（1.4.1 或 2.0.1）。

1. 透過執行下列指令將 Spark 相依關係上傳至 MapR 檔案系統：

```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/  
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

<as_analytic_root> 是在第 43 頁的『在 MapR 上安裝 Analytic Server』的第 12 步設定的路徑。

2. 執行下列指令以建立 Spark 配置檔案的鏈結：

```
mkdir <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf  
ln -s /opt/mapr/spark/spark-<spark_version>/conf/*  
<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/spark-conf
```

3. 將下列行新增至檔案 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/server.env

```
SPARK_HOME=/opt/mapr/spark/spark-<spark_version>
```

4. 將下列行新增至檔案 <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties

```
spark.executor.extraLibraryPath=/opt/mapr/hadoop/hadoop-2.7.0/lib/native
```

5. 重新啟動 Analytic Server。

6. 如果需要啟用 PySpark 功能，請在 yarn-env.sh 檔中新增下列行：

```
export SPARK_HOME=/opt/mapr/spark/spark-<spark_version>
```

重新啟動 ResourceManagers 和 NodeManagers。

啟用功能旗標

功能旗標會提供啟用及停用特定應用程式功能的能力。

若要在順利安裝 MapR 之後啟用功能旗標支援：

1. 在檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties` 中新增下列行：

```
load.feature.flags.on.msg=true
```

2. 重新啟動 IBM SPSS Analytic Server。

啟用 R

R 是用於統計計算與圖形的語言及環境。

若要在順利安裝 MapR 之後啟用 R 支援：

註：必須安裝下列套件，您才可以在管理節點管理程式及 IBM SPSS Analytic Server 的所有叢集節點上，執行安裝程式：

```
gcc-gfortran  
libgfortran  
gcc-c++
```

1. 在管理節點管理程式及 Analytic Server 的所有叢集節點上，執行安裝程式 `spss_er-8.4.0.0-mapr5-1x86_64_en.bin`。執行安裝程式的使用者必須具有 R 及 Analytic Server 安裝路徑的寫入權。
2. 遵循安裝提示，接受授權合約，並輸入必要資訊。如果 Analytic Server 安裝在安裝伺服器上，請在提示時選擇是，並輸入 `<as_installation_path>`。如果 Analytic Server 未安裝在安裝伺服器上，請在提示時選擇否。
3. 如果已安裝 Analytic Server，則 Essentials for R 會自動安裝在 Analytic Server 安裝路徑中。
 - 如果未安裝 Analytic Server，則 Essentials for R 會安裝在 `<installer_path>/IBM_SPSS_ModelerEssentialsR/linux` 路徑中。
 - 如果稍後安裝 Analytic Server，請使用下列指令，將 Essentials for R 複製到安裝 Analytic Server 的 Analytic Server 配置路徑中。

```
cp -r <installer_path>/IBM_SPSS_ModelerEssentialsR/linux <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration
```

4. 在路徑 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration` 中刪除 `cf.zip` 檔案，並使用下列指令，產生新檔案：

```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration  
zip -r cf.zip linux
```

5. 執行下列指令：

```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration  
hadoop fs -rm <as_analytic_root>/cluster1/configuration/cf.zip  
hadoop fs -put cf.zip <as_analytic_root>/cluster1/configuration/
```

6. 重新啟動 Analytic Server。

啟用 LZO

LZO 是不失真資料壓縮程式庫，相較於壓縮比例，其更重視速度。必須對 MapR 進行手動配置，以提供 LZO 支援。

下列網站提供 LZO 安裝及配置指示：<https://github.com/twitter/hadoop-lzo>。

下列步驟詳細說明將 LZO 程式庫匯入 MapR 的處理程序。

1. 將 `hadoop-lzo-<version>.jar` 檔案複製到 Hadoop 類別路徑。建議的路徑為 `/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/mapreduce/lib`。

- 將原生檔案 `libgplcompression.so` 及 `liblzo2.so.2` 複製到 `/opt/mapr/hadoop/hadoop-2.7.0/lib/native`，並將下列內容新增至 `core-site.xml` 檔案：

```
<property>
  <name>io.compression.codecs</name>
  <value>org.apache.hadoop.io.compress.GzipCodec,org.apache.hadoop.io.compress.DefaultCodec,com.hadoop.compression.lzo.LzoCodec,com.hadoop.compression.lzo.LzopCodec,org.apache.hadoop.io.compress.BZip2Codec</value>
</property>
<property>
  <name>io.compression.codec.lzo.class</name>
  <value>com.hadoop.compression.lzo.LzoCodec</value>
</property>
```

- 開啟檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/server.env`，並將 `<lzo_native_path>` 新增至 `LD_LIBRARY_PATH` 參數。`<lzo_native_path>` 是包含 Hadoop-LZO 原生程式庫的資料夾。

```
LD_LIBRARY_PATH=<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:<java_home>/jre/lib/amd64:/opt/mapr/hadoop/hadoop-2.7.0/lib/native:<lzo_native_path>
```

- 重新啟動 IBM SPSS Analytic Server。

針對 MapR 設定 IBM SPSS Analytic Server 叢集

使用下列步驟，以設定 MapR 支援的 IBM SPSS Analytic Server 叢集環境。

- 在檔案 `<as_installation_path>/ae_wlpserver/usr/servers/aeserver/configuration/config.properties` 中新增下列行。

```
enable.resume=true
```

- 將安裝路徑複製到其他叢集節點，並將 `config.properties` 檔案中的 `ae.cluster.member.name` 內容變更為正確的主機名稱。
- 啟動所有叢集節點。

解除安裝 MapR

下列步驟說明解除安裝 MapR 的處理程序：

- 停止 IBM SPSS Analytic Server。
- 刪除 meta 資料資料庫。

- 執行下列指令：

```
cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/sql/<db_type>
java -Xmx128m -Xms128m -cp <as_installation_path>/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib/*:/usr/share/java/*
com.spss.ae.dbscript.ScriptRunner ../../configuration/config.properties drop.sql true
```

- 執行下列 SQL 陳述式，以捨棄資料庫：

```
drop database <db_name>
```

- 解除安裝 RPM 套件：

```
rpm -e IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64
```

- 刪除安裝路徑：

```
rm -r <as_installation_path>
```

- 刪除 Analytic 根目錄：

```
hadoop fs -rm -r <analytic-root>
```

- 刪除 zookeeper 資料：

```
/opt/mapr/zookeeper/zookeeper-3.4.5/bin/zkCli.sh -server <zookeeper_host>:<zookeeper_port>
rmr /AnalyticServer
```


在 MapR 上移轉 IBM SPSS Analytic Server

可以在 MapR 上移轉 IBM SPSS Analytic Server。

使用下列步驟，在 MapR 上將 IBM SPSS Analytic Server 2.0 或 2.1 版移轉至 3.1.0 版。

1. 透過遵循第 43 頁的『在 MapR 上安裝 Analytic Server』中的安裝指示，在 MapR 叢集上安裝 Analytic Server 3.1.0。
2. 複製分析根目錄。

註：如果 Analytic 根目錄未變更，則可忽略此步驟。

- 如果 Analytic Server 2.0/2.1 版及 3.1.0 版的 Analytic 根目錄位於相同的 MapR 叢集上，請在其中一個資料節點上，執行下列指令：

```
hadoop fs -cp <old_analytic_root>/analytic-workspace/* <new_analytic_root>/analytic-workspace
```

- 已安裝的 WEBHDFS 或 NFS 服務會控制 Analytic Server 2.0/2.1 版及 3.1.0 版的 Analytic 根目錄何時位於不同的 MapR 叢集上。由於無法在叢集外部直接存取 MapR 檔案系統，因此若要複製分析根目錄資料，需要 WEBHDFS 或 NFS。

- a. 當舊 Analytic Server 2.0/2.1 叢集包括 WEBHDFS 服務時，請在其中一個新 Analytic Server 3.1.0 叢集節點上執行下列指令：

```
hadoop distcp webhdfs://<webhdfs_server>:<webhdfs_port>/<old_analytic_root>/analytic-workspace/*  
maprfs://<new_analytic_root>/analytic-workspace
```

- b. 當新 Analytic Server 3.1.0 叢集包括 WEBHDFS 服務時，請在其中一個舊 Analytic Server 2.0/2.1 叢集節點上執行下列指令：

```
hadoop distcp maprfs://<old_analytic_root>/analytic-workspace/*  
webhdfs://<webhdfs_server>:<webhdfs_port>/<new_analytic_root>/analytic-workspace
```

- c. 當舊叢集包括 NFS 且 NFS 也裝載在其中一個新 Analytic Server 3.1.0 叢集節點上時，請在其中一個舊 Analytic Server 2.0/2.1 叢集節點上執行下列指令：

```
hadoop distcp file:///<mount_path>/<old_analytic_root>/analytic-workspace/* maprfs://<new_analytic_root>/analytic-workspace
```

- d. 當新叢集包括 NFS 且 NFS 也裝載在其中一個舊 Analytic Server 2.0/2.1 叢集節點上時，請在其中一個新 Analytic Server 3.0.1 叢集節點上執行下列指令：

```
hadoop discp maprfs://<old_analytic_root>/analytic-workspace/* file:///<mount_path>/<new_analytic_root>/analytic-workspace
```

檢閱 MapR 資料移轉網站，以取得在不同 MapR 叢集之間移轉資料的相關資訊。

3. 執行下列指令，以變更新 Analytic 根目錄的擁有者及許可權：

```
hadoop fs -chown -R <as_user> <analytic_root>  
hadoop fs -chmod -R 755 <>
```

4. 停止 Analytic Server 3.1.0，但請確保 meta 資料資料庫仍在執行中。

5. 從舊 Analytic Server 2.0/2.1 叢集安裝收集配置設定。

- a. 將 configcollector.zip 保存檔從新 Analytic Server 3.1.0 叢集安裝複製到舊 Analytic Server 2.0/2.1 叢集安裝的 <old_as_installation_path>/tools 中。
- b. 在舊 Analytic Server 2.0/2.1 叢集安裝上解壓縮 configcollector.zip 內容。即會在舊 Analytic Server 2.0/2.1 叢集安裝中建立新 configcollector 子目錄。
- c. 在舊 Analytic Server 2.0/2.1 叢集安裝中執行配置收集器工具，方法是從 <old_as_installation_path>/tools/configcollector 執行 configcollector Script。將所產生的壓縮檔 (ZIP) 複製到新 Analytic Server 3.1.0 叢集安裝中。

6. 在新 Analytic Server 3.1.0 叢集上執行移轉工具，方法是執行 migrationtool Script，並將配置收集器所建立的壓縮檔路徑，作為引數進行傳遞。例如：

```
migrationtool.sh /opt/ibm/spss/analyticserver/3.1/ASConfiguration_2.1.0.0.xxx.zip
```

7. 啟動 Analytic Server 3.1.0。

MapR 疑難排解

本節說明部分常見的 MapR 安裝與配置問題，以及如何修正這些問題。

hdfsUpdate.sh Script 的問題

hdfsUpdate.sh Script 只能執行一次，因為該 Script 會在上傳新檔案之前，先移除 analytic-root 中的所有檔案。當 Script 執行多次時，您必須重新上傳資料庫推回、Hive、HBase 及 Spark 的相依關係。執行下列指令，以重新上傳必要相依關係：

```
cd <as_installation_path>/jdbc
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath

cd /opt/mapr/hive/hive-1.2/lib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
cd /opt/mapr/hive/hive-1.2/hcatalog/share/hcatalog
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath

cd /opt/mapr/hbase/hbase-0.98.12/lib
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath

cd <as_installation_path>/ae_wlpserver/usr/servers/aeserver/modules/spark/
hadoop fs -put *.jar <as_analytic_root>/cluster1/classpath
```

MapR 與 Spark 版本之間發生衝突會造成 Spark 工作執行失敗

當 MapR 版本為 5.1 或更新版本時，在 MapR 與 Spark (1.6.1) 之間發生類別衝突問題。該衝突會導致 Spark 工作執行失敗。可透過修改 <as_installation_path>/ae_wlpserver/usr/servers/aeserver 中的 private_library.xml 檔來解決該問題。下列範例可識別所需要的變更：

```
.....
<fileset dir="/opt/mapr/hadoop/hadoop-2.7.0/share/hadoop/common/lib" includes="*.jar" excludes="jackson-databind-*.jar" />
.....
```

第 5 章 Huawei FusionInsight HD 安裝與配置

FusionInsight HD 概觀

Huawei FusionInsight HD 提供一個綜合性的海量資料軟體平台，以供利用開放程式碼 Hadoop 及 Spark 技術進行批次與即時分析。系統利用 HDFS、HBase、MapReduce 和 YARN/Zookeeper 來形成 Hadoop 的叢集，並利用 Apache Spark 進行快速即時分析與互動式查詢。

Analytic Server 可在 FusionInsight HD 平台上執行。FusionInsight HD 包含 Hadoop 的主要核心元素，以對大型資料集（主要是 MapReduce 及 HDFS）提供可靠的可調式分散式資料處理功能，還包含其他企業導向的元件，以提供安全性、高可用性與硬體和其他軟體的整合。

在 Huawei FusionInsight HD 上安裝

下列步驟說明在 Huawei FusionInsight HD 中手動安裝 IBM SPSS Analytic Server 的程序。

Analytic Server 3.1.0

1. 導覽至 [IBM Passport Advantage® 網站](#) 並將下列自行解壓縮的二進位檔下載到 FusionInsight HD 叢集中的主機。

表 11. Analytic Server 自行解壓縮的二進位檔

說明	二進位檔名稱
IBM SPSS Analytic Server 3.1 for FusionInsight HD 2.6 Linux x86-64 (英文)	spss_as-3.1-fhd2.6-1x86_en.bin

2. 在 FusionInsight Manager 主要叢集節點上執行自行解壓縮的 *.bin 安裝程式。遵循安裝提示，接受授權合約，並保留預設安裝目錄。安裝程式會下載必要的 RPM 檔案，且必須在可以存取 <https://ibm-open-platform.ibm.com> 的電腦上執行。可執行二進位檔位於可用的 FusionInsight HD <AS_INSTALLABLE_HOME> 配送目錄中。
3. 使用下列指令來安裝 Analytic Server 3.1.0：

```
# yum install -y IBM-SPSS-AnalyticServer-3.1.0.0-1.x86_64.rpm
```

4. 使用 omm 來登入，並建立 analyticserver.keytab：

```
# su omm
# source /opt/huawei/Bigdata/om-0.0.1/meta-0.0.1-SNAPSHOT/kerberos/scripts/component_env
# kadmin -p kadmin/admin
```

預設 kadmin 密碼為 Admin@123。第一次使用時必須變更密碼。在下列指令中，將 _HOST 取代為主機名稱。

```
kadmin > addprinc -randkey omm/_HOST@HADOOP.COM
kadmin > ktadd -k /opt/ibm/spss/analyticserver/3.1/analyticserver.keytab HTTP/_HOST@HADOOP.COM
kadmin > ktadd -k /opt/ibm/AnalyticServer/analyticserver.keytab omm/_HOST@HADOOP.COM
```

5. 安裝 MYSQL 並手動建立 aedb。例如：

```
# cd /etc/yum.repos.d
# wget http://dev.mysql.com/get/mysql157-community-release-e17-9.noarch.rpm
# yum -y install mysql157-community-release-e17-9.noarch.rpm
# yum repolist all | grep mysql
# yum -y install mysql-community-server
# yum install -y mysql-connector-java
# systemctl enable mysqld.service
# systemctl start mysqld.service
```

擷取 MYSQL root 使用者密碼：

```
# grep 'temporary password' /var/log/mysql.log
# mysql -uroot -p
# MySQL> set global validate_password_policy=0;
# MySQL> DROP DATABASE IF EXISTS aedb;
# MySQL> CREATE DATABASE aedb DEFAULT CHARACTER SET utf8 DEFAULT COLLATE utf8_bin;
# MySQL> CREATE USER 'aeuser'@%' IDENTIFIED BY 'Pass1234';
# MySQL> CREATE USER 'aeuser'@'localhost' IDENTIFIED BY 'Pass1234';
# MySQL> GRANT ALL PRIVILEGES ON *.* TO 'aeuser'@%;
# MySQL> GRANT ALL PRIVILEGES ON *.* TO 'aeuser'@'localhost';
```

6. 在 /opt 中建立 install_as.sh Script，使用 omm 使用者來執行它：

```
# chown -R omm:wheel /opt/ibm/*
```

```
# su omm
# /opt/install_as.sh
```

```
install_as.sh
```

install_as.sh 包含下列 Script：

```
cd /opt/ibm/spss/analyticsserver/3.1
mkdir hadoop
mkdir zookeeper
mkdir spark-client
```

```
cd /opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/
echo "JAVA_HOME=/opt/huawei/Bigdata/jdk/jre" > server.env
echo "PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/huawei/Bigdata/jdk/jre/lib/amd64:/usr/sbin:/usr/bin:/usr/local/bin:/usr/local/sbin" > server.env
echo "IBM_SPSS_AS_NATIVE_PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64" > server.env
echo "LD_LIBRARY_PATH=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64:/opt/huawei/Bigdata/jdk/jre/lib/amd64:/usr/local/lib:/usr/lib:/usr/lib64:/usr/local/lib64:/usr/lib32:/usr/lib64:/usr/libx32:/usr/libx86_64:/usr/libx86_32:/usr/libx86_64:/usr/libx86_32:/usr/libx86_64:/usr/libx86_32" > server.env
echo "SPARK_HOME=/opt/ibm/spss/analyticsserver/3.1/spark-client" >> server.env
```

```
echo "<server description=\"new server\">" > server.xml
echo "<!-- Enable features -->" >> server.xml
echo " <featureManager>" >> server.xml
echo " <feature>Servlet-3.1</feature>" >> server.xml
echo " <feature>jsp-2.3</feature>" >> server.xml
echo " <feature>jdbc-4.0</feature>" >> server.xml
echo " <feature>jndi-1.0</feature>" >> server.xml
echo " <feature>localConnector-1.0</feature>" >> server.xml
echo " <feature>jaxrs-2.0</feature>" >> server.xml
echo " <feature>json-1.0</feature>" >> server.xml
echo " <feature>appSecurity-2.0</feature>" >> server.xml
echo " <feature>ldapRegistry-3.0</feature>" >> server.xml
echo " <feature>restConnector-1.0</feature>" >> server.xml
echo " <feature>monitor-1.0</feature>" >> server.xml
echo " <feature>ssl-1.0</feature>" >> server.xml
echo " </featureManager>" >> server.xml
echo " <applicationManager startTimeout=\"120s\" />" >> server.xml
echo " <executor name=\"LargeThreadPool\" id=\"default\" coreThreads=\"100\" keepAlive=\"60s\" stealPolicy=\"STRICT\" rejectedWorkPolicy=\"CALLER_RUNS\" />" >> server.xml
echo " <webContainer deferServletLoad=\"false\" disallowAllFileServing=\"false\" fileServingEnabled=\"true\" trusted=\"false\" directoryBrowsingEnabled=\"false\" />" >> server.xml
echo " <classloading useJarUrls=\"true\" />" >> server.xml
echo " <applicationMonitor updateTrigger=\"mbean\" />" >> server.xml
echo " <mimeTypes>" >> server.xml
echo " <type>svg=image/svg+xml</type>" >> server.xml
echo " </mimeTypes>" >> server.xml
echo " <variable name=\"AE_DATABASE\" value=\"\${wlp.install.dir}/usr/servers/aeserver/aedb\" />" >> server.xml
echo " <administrator-role>" >> server.xml
echo " <user>admin</user>" >> server.xml
echo " </administrator-role>" >> server.xml
echo " <include optional=\"true\" location=\"\${server.config.dir}/private_library.xml\" />" >> server.xml
echo " <include optional=\"true\" location=\"\${server.config.dir}/http_endpoint.xml\" />" >> server.xml
echo " <include optional=\"true\" location=\"\${server.config.dir}/security_cfg.xml\" />" >> server.xml
echo " <include optional=\"true\" location=\"\${server.config.dir}/ssl_cfg.xml\" />" >> server.xml
echo " <include optional=\"true\" location=\"\${server.config.dir}/configuration/key.xml\" />" >> server.xml
echo "</server>" >> server.xml
```

```
touch http_endpoint.xml
echo "<server>" > http_endpoint.xml
echo " <httpEndpoint host=\"*\\" id=\"defaultHttpEndpoint\" httpPort=\"9080\" httpsPort=\"9443\" onError=\"FAIL\" />" >> http_endpoint.xml
echo "</server>" >> http_endpoint.xml
```

```
touch private_library.xml
echo "<server>" > private_library.xml
echo " <application context-root=\"/analyticsserver\" id=\"AS_B00T\" location=\"/AE_B00T.war\" name=\"/AS_B00T\" type=\"war\">" >> private_library.xml
echo " <classloader>" >> private_library.xml
echo " <privateLibrary>" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/usr/servers/aeserver/apps/AE_B00T.war/WEB-INF/lib\" includes=\"*.jar\" />" >> private_library.xml
echo " <fileset dir=\"/usr/share/java\" includes=\"*.jar\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../lib\" includes=\"*.jar\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../spark-client/lib\" includes=\"spark-assembly-*.jar\" />" >> private_library.xml
echo " <folder dir=\"\${wlp.install.dir}/usr/servers/aeserver/configuration/hadoop-conf\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../jdbc\" includes=\"postgresql-*.jar\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../jdbc\" includes=\"*.jar\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../hive\" includes=\"*.jar\" />" >> private_library.xml
echo " <fileset dir=\"\${wlp.install.dir}/../zookeeper\" includes=\"*.jar\" />" >> private_library.xml
```

```

echo "      <fileset dir=\"\${wlp.install.dir}/../hadoop\" includes=\"*.jar\"/>" >> private_library.xml
echo "    </privatelibrary>" >> private_library.xml
echo "  </classloader>" >> private_library.xml
echo " </application>" >> private_library.xml
echo " <application id=\"help\" location=\"help.war\" name=\"help\" type=\"war\" context-root=\"/analyticsserver/help\"/>" >> private_library.xml
echo "</server>" >> private_library.xml

touch security_cfg.xml
echo "<server>" > security_cfg.xml
echo "  <basicRegistry id=\"basic\" realm=\"ibm\">" >> security_cfg.xml
echo "    <user name=\"admin\" password=\"admin\"/>" >> security_cfg.xml
echo "  </basicRegistry>" >> security_cfg.xml
echo "</server>" >> security_cfg.xml

touch jaas.conf
echo "Client {" > jaas.conf
echo "  com.sun.security.auth.module.Krb5LoginModule required" >> jaas.conf
echo "  keyTab=\"/opt/ibm/spss/analyticsserver/3.1/analyticsserver.keytab\"" >> jaas.conf
echo "  principal=\"omm/huawei-1@HADOOP.COM\"" >> jaas.conf
echo "  useKeyTab=true" >> jaas.conf
echo "  useTicketCache=true" >> jaas.conf
echo "  storeKey=true" >> jaas.conf
echo "  debug=true;" >> jaas.conf
echo "};" >> jaas.conf

cd /opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration
echo "spark.version=1.x" > config.properties
echo "http.port=9080" >> config.properties
echo "https.port=9443" >> config.properties
echo "ae.cluster.zookeeper.connect.string=172.16.155.123:24002,172.16.155.212:24002,172.16.186.208:24002" >> config.properties
echo "ae.cluster.member.name=huawei-1" >> config.properties
echo "ae.cluster.collective.name=Test_01" >> config.properties
echo "jndi.aedb=jdbc/aeds" >> config.properties
echo "jndi.aedb.url=jdbc:mysql://huawei-1/aedb?createDatabaseIfNotExist=true" >> config.properties
echo "jndi.aedb.username=aeuser" >> config.properties
echo "jndi.aedb.driver=com.mysql.jdbc.Driver" >> config.properties
echo "distrib.fs.root=/user/as_user/analytic-root" >> config.properties
echo "admin.username=admin" >> config.properties
echo "enable.resume=true" >> config.properties
echo "load.feature.flags.on.msg=true" >> config.properties
echo "jndi.aedb.password=FEFFUy9FQ0IvUEtDUzVQYWRkaW5nAGk3bIuya2BzXyYFc0rxo=" >> config.properties
echo "ae.kerberos.principal=omm/huawei-1@HADOOP.COM" >> config.properties
echo "hdfs.user=omm/huawei-1@HADOOP.COM" >> config.properties
echo "web.authentication.kerberos.principal=HTTP/huawei-1@HADOOP.COM" >> config.properties
echo "java.security.krb5.conf=/home/omm/kerberos/var/krb5kdc/krb5.conf" >> config.properties
echo "web.authentication.kerberos.keytab=/opt/ibm/spss/analyticsserver/3.1/analyticsserver.keytab" >> config.properties
echo "hdfs.keytab=/opt/ibm/spss/analyticsserver/3.1/analyticsserver.keytab" >> config.properties
echo "ae.db.connect.method=Kerberos" >> config.properties
echo "kdcrealm=HADOOP.COM" >> config.properties
echo "kdcserver=172.16.155.212:21732" >> config.properties
echo "encryption.keystore.password=FEFFUy9FQ0IvUEtDUzVQYWRkaW5nAMDJul7PVsvdIyLlZjeS8ws=" >> config.properties
echo "encryption.keystore.base64=zs70zgAAAAIAAAABAAAAAwA6Y29tLnNwc3MuYUuZw5jc1wdG1vb15wcm92aWR1ci51bmNyeXB0aW9ucHJvdmlkZXJpbXBsLmFlcwAAAATg2Ahyr00ABX" >> config.properties
echo "jdbc.drivers.location=/usr/share/jdbc" >> config.properties
echo "default.security.provider=Kerberos" >> config.properties
echo "load.feature.flags.on.msg=true" >> config.properties
echo "spark.serializer=org.apache.spark.serializer.JavaSerializer" >> config.properties
echo "spark.executor.extraLibraryPath=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64" >> config.properties
echo "zookeeper.server.jaas.config=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/jaas.conf" >> config.properties

chmod 644 *.xml

unzip cf.zip

mkdir hadoop-conf

```

7. 下載 FusionInsight SPARK 和 Zookeeper 用戶端，解壓縮內容，然後將 Spark 配置檔複製到 /opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/hadoop-conf folder。
8. 將原生 Hadoop 檔複製到下列資料夾：/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64
9. 將下列行新增至檔案 /opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/jvm.option
-Dconfig.folder.path=/opt/ibm/spss/analyticsserver/3.1/ae_wlpserver/usr/servers/aeserver/configuration
10. 執行 hdfsUpdate.sh Script (/opt/ibm/spss/analyticsserver/3.1/bin/hdfsUpdate.sh)。
11. 執行 start.sh script (/opt/ibm/spss/analyticsserver/3.1/bin/start.sh) 以啟動 Analytic Server 服務；執行 stop.sh Script (/opt/ibm/spss/analyticsserver/3.1/bin/stop.sh) 以停止 Analytic Server 服務。

第 6 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。

as_port

Analytic Server 接聽的埠（依預設，這是 8080）。

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 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/iop/4.1.0.0/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。

特定 Hadoop 發行套件問題

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

若要在 Hortonworks 2.3 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/analyticsserver/3.0/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。

LZO 壓縮程式不適用於 BigInsights 4.2.X

您必須執行 `yum install lzo* hadoop-lzo*` 以在 BigInsights 4.2.X 中安裝 LZO 元件。由於安裝行為已在 BigInsights 4.2 環境中變更，現在，LZO 元件會安裝在 `/usr/lib/hadoop-lzo/lib` 目錄之下（此目錄不包括 Hadoop 執行環境）。請使用下列步驟在 BigInsights 4.2.X 上手動配置 LZO 元件。

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

註：可使用 Ambari 主控台來判定主機名稱。

2. 將 Hadoop LZO jar 檔複製到 Analytic Server 檔案庫目錄。例如：

```
cp /usr/lib/hadoop-lzo/lib/hadoop-lzo*.jar {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF
```

3. 將 LZO 原生檔案庫複製到 Analytic Server 原生檔案庫目錄。例如：

```
cp /usr/lib/hadoop-lzo/lib/native/* {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/linux/lib_64
```

4. 執行位於 `{AS_ROOT}/bin` 目錄中的 `refresh` Script。例如：

```
cd /opt/ibm/spss/analyticsserver/3.0/bin
./refresh
```

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

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 取代之為您要移除的使用者名稱，並將 \$METASTORE_HOST 取代之為資料庫安裝所在的主機名稱。

Spark 處理程序中執行 IBM SPSS Modeler 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，但串流永不完成且保留執行中狀態。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）。

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 20 頁的『更新用戶端相依關係』中的指示，執行 update_clientdeps Script。

java.net.SocketTimeoutException：讀取逾時

變更 Liberty ND 逾時環境變數，如下所示：

```
export LIBERTYND_READ_TIMEOUT=<milliseconds>
```

其中，<milliseconds> 是用於 JMX 讀取逾時的秒數。

java.io.IOException：CWWKX7202E：指令 ./server 啟動的逾時值 60（秒）已過期

將下列行新增至「控制器伺服器」的 server.xml

```
<!-- Increase start and stop server timeout to accommodate slow hardware -->  
<serverCommands startServerTimeout="120" stopServerTimeout="120"/>
```

java.lang.OutOfMemoryError：Java 資料堆空間

在每個「高可用性」叢集成員的 jvm.options 中，新增下列行。

```
-Xms512M  
-Xmx2048M
```

「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。

有時，當伺服器關機時，**Analytic Server** 伺服器會當掉
手動結束伺服器。

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：

© (您的公司名稱) (年份)。本程式之若干部分係衍生自 IBM 公司的範例程式。

© Copyright IBM Corp. (輸入年份). All rights reserved.

商標

IBM、IBM 標誌及 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。

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。



Printed in Taiwan