

**IBM SPSS Analytic Server  
2.1 版**

**管理者手冊**

**IBM**

**附註**

在使用本資訊及它支援的產品之前，請閱讀第 11 頁的『注意事項』中的資訊。

**產品資訊**

此版本適用於 IBM SPSS Analytic Server 2.1.1.0 版及所有後續版本和修訂版，直到在新版本中另有聲明為止。

---

## 目錄

<b>第 1 章 概觀</b> . . . . .	<b>1</b>	版本資訊 . . . . .	9
架構 . . . . .	2	日誌收集器 . . . . .	9
2.1 版中適用於管理者的新內容 . . . . .	2	常見問題 . . . . .	9
<b>第 2 章 租戶管理</b> . . . . .	<b>3</b>	<b>注意事項</b> . . . . .	<b>11</b>
命名規則 . . . . .	4	商標 . . . . .	12
<b>第 3 章 使用者入門</b> . . . . .	<b>5</b>		
<b>第 4 章 Analytic Server 工作名稱</b> . . . . .	<b>7</b>		
<b>第 5 章 疑難排解</b> . . . . .	<b>9</b>		
記載 . . . . .	9		



---

## 第 1 章 概觀

IBM® SPSS® Analytic Server 是用於大型資料分析的解決方案，它結合 IBM SPSS 技術與大型資料系統，可容許您使用熟悉的 IBM SPSS 使用者介面來解決規模空前的問題。

### 大型資料分析為何重要

組織收集的資料量以指數增長；例如，財務與零售公司具有一年（或二年，或十年）之內的所有客戶交易，telco 提供者具有呼叫資料記錄 (CDR) 及裝置感應器讀數，而網際網路公司具有 Web 搜索的結果。

存在下列項目時，需要大型資料分析：

- 大量資料 (TB、PB、EB)，特別當結構化與非結構化資料混合時
- 快速變更/累積資料

當發生下列狀況時，大型資料分析也能予以協助：

- 建置大量（數以千計）模型
- 頻繁地建置/重新整理模型

### 盤查

因為各種原因，收集大量資料的相同組織通常難以實際予以使用：

- 傳統分析產品的架構不適合分散式計算，以及
- 現有統計演算法未設計為處理大型資料（這些演算法預期資料送入其中，但是大型資料移動成本太高），因此
- 大型資料的圖形分析執行狀態需要大型資料系統的新技能，且對這些系統非常熟悉。極少數分析師具有這些技能。
- 記憶體內解決方案適用於中等大小問題，但是不適合真正的大型資料。

### 解決方案

Analytic Server 提供：

- 以資料為中心的架構，利用大型資料系統，例如 HDFS 中含資料的「Hadoop 對映」/「Hadoop 減少」。
- 已定義介面，併入設計以跳至資料的新統計演算法。
- 熟悉 IBM SPSS 使用者介面，隱藏大型資料環境的詳細資料，以便分析師可以聚焦於分析資料。
- 可調整以適合任何大小問題的解決方案。

## 架構

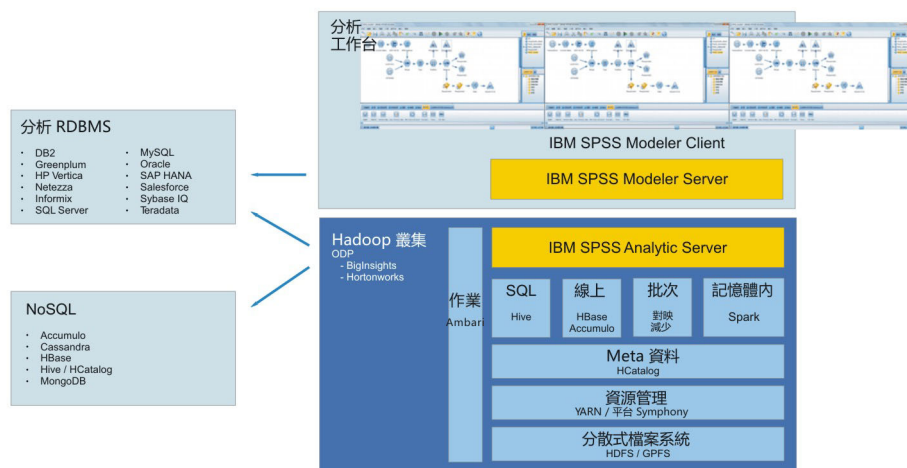


圖 1. 架構

Analytic Server 位於用戶端應用程式與 Hadoop 雲端之間。假設資料位於雲端中，使用 Analytic Server 的概要為：

1. 定義基於雲端中資料的 Analytic Server 資料來源。
2. 定義要在用戶端應用程式中執行的分析。對於現行版本，用戶端應用程式為 IBM SPSS Modeler。
3. 當您執行分析時，用戶端應用程式提交 Analytic Server 執行要求。
4. Analytic Server 編排工作以在 Hadoop 雲端中執行，並將結果報告給用戶端應用程式。
5. 您可以使用結果來定義進一步分析，並重複該循環。

## 2.1 版中適用於管理者的新內容

### Analytic Server 主控台

#### 讀者角色

在租戶內，您可以向使用者和群組指派「讀者」角色，該角色無法登入 Analytic Server，但是可以透過 Modeler 中的 Analytic Server 來源節點來讀取 Analytic Server 資料。

---

## 第 2 章 租戶管理

租戶提供使用者、專案和資料來源的高層次劃分，以便物件無法在租戶之間共用。每個使用者存取將他們指派給的租戶環境定義中的系統。

您可以在 **Analytic Server** 主控台中管理租戶，以及將使用者指派給租戶。「租戶」頁面的視圖取決於登入至主控台的使用者角色：

- 在安裝期間設定的「超級使用者」管理者是租戶管理員。只有此使用者可以建立新租戶和編輯任何租戶的內容。
- 具有管理者角色的使用者可以編輯他們登入之租戶的內容。
- 具有使用者角色的使用者不能編輯租戶內容。會向他們隱藏「租戶」頁面。
- 具有讀者角色的使用者不能編輯資料來源，甚至無法登入 **Analytic Server** 主控台。

管理者可以存取「專案」和「資料來源」頁面，以及管理任何專案或資料來源以進行清理和管理。如需相關資訊，請參閱《*IBM SPSS Analytic Server 使用手冊*》。

### 租戶清單

主要「租戶」頁面會顯示表格中的現有租戶。只有「超級使用者」管理者可以在此頁面上進行編輯。

- 按一下租戶的名稱以顯示其詳細資料，並編輯其內容。
- 按一下租戶的 URL，在該租戶的環境定義中開啓主控台。

**註：**您將登出主控台，並將需要使用該租戶的有效認證進行登入。

- 在搜尋區域中鍵入以過濾清單，從而僅顯示名稱中含有搜尋字串的租戶。
- 按一下**新建**，以您在**新增租戶**對話框中指定的名稱建立新的租戶。如需您為租戶提供之名稱的限制，請參閱第 4 頁的『命名規則』。
- 按一下**刪除**，以移除選取的租戶。
- 按一下**重新整理**，以更新清單。

### 個別租戶詳細資料

內容區劃分為數個可收合區段。

#### 詳細資料

<b>名稱</b>	一個可編輯的文字欄位，顯示租戶的名稱。
<b>說明</b>	一個可編輯的文字欄位，容許您提供關於租戶的解釋性文字。
<b>URL</b>	此 URL 提供給使用者以透過 <b>Analytic Server</b> 主控台登入，以及用來配置 <b>SPSS Modeler</b> 伺服器。如需配置 <b>SPSS Modeler</b> 的詳細資料，請參閱《 <i>IBM SPSS Analytic Server 安裝與配置手冊</i> 》。
<b>狀態</b>	<b>作用中</b> 租戶目前正在使用中。讓租戶處於 <b>非作用中</b> 會阻止使用者登入該租戶，但不會刪除任何基礎資訊。

#### 主體

主體是在安裝期間從安全提供者設定得來的使用者和群組。您可以將主體作為管理者活使用者新增至租戶。

- 在文字框中輸入內容可過濾名稱中含有搜尋字串的使用者和群組。從下拉清單中選取**管理者**、**使用者**或**讀者**，以指派其在租戶內的角色。按一下**新增參與者**以將他們新增至作者清單。
- 若要移除參與者，請在成員清單中選取使用者或群組，然後按一下**移除參與者**。

**度量** 可讓您配置租戶的資源限制。報告租戶目前使用的磁碟空間。

- 您可以設定租戶的磁碟空間配額上限；當達到此限制時，無法更多資料寫入至此租戶上的磁碟，除非清除足夠的磁碟空間，讓租戶磁碟空間用量低於配額。
- 您可以設定租戶的磁碟空間警告層次；當超出該配額時，此租戶上的主體無法提交任何分析工作，除非清除足夠的磁碟空間，讓租戶磁碟空間用量低於配額。
- 您可以設定此租戶上於單一時間執行的平行工作數目上限；當超出該配額時，此租戶上的主體無法提交任何分析工作，直到現行執行中工作完成為止。
- 您可以設定資料來源可以具有的欄位數目上限。每當建立或更新資料來源時，會檢查該限制。
- 您可以設定檔案大小上限 (MB)。上傳檔案時，會檢查該限制。

### 安全提供者配置

可讓您指定使用者鑑別提供者。**預設**使用預設租戶提供者，在安裝與配置期間設定該項。**LDAP** 可讓您向 Active Directory 或 OpenLDAP 之類的外部 LDAP 伺服器鑑別使用者。指定提供者的設定，並選擇性地指定過濾器設定，以控制「主體」區段中可用的使用者和群組。

---

## 命名規則

對於 Analytic Server 中可以提供唯一名稱的任何項目，例如資料來源與專案，下列規則適用於那些名稱。

- 在單個租戶內，名稱在相同類型的物件內必須是唯一的。例如，兩個資料來源不能都命名為 insuranceClaims，但是資料來源與專案可以都命名為 insuranceClaims。
- 名稱區分大小寫。例如，insuranceClaims 與 InsuranceClaims 都視為唯一名稱。
- 名稱會忽略前導與尾端空格。
- 下列字元在名稱中無效。

~、#、%、&、\*、{、}、\、:、<、>、?、/、\、|、"、\t、\r、\n



---

## 第 3 章 使用者入門

告知使用者導覽至 `http://<host>:<port>/<context-root>/admin/<tenant>`，並輸入其使用者名稱與密碼，從而登入至 Analytic Server 主控台。

### <host>

Analytic Server 主機的位址。

### <port>

Analytic Server 接聽的埠。依預設，為 8080。

### <context-root>

Analytic Server 的環境定義根目錄。依預設，這是分析伺服器。

### <tenant>

在多租戶環境中，表示您所屬的租戶。在單一租戶環境中，預設租戶為 **ibm**。

例如，如果主機的 IP 位址為 9.86.44.232，您已建立租戶 "mycompany" 並向其新增了使用者，且將其他設定保留為其預設值，則使用者應該導覽至 `http://9.86.44.232:8080/analyticserver/admin/mycompany`，以存取 Analytic Server 主控台。



---

## 第 4 章 Analytic Server 工作名稱

Analytic Server 產生對映化簡工作，可以透過 Hadoop 叢集的工作追蹤器進行監視。

對映化簡工作具有下列結構。

AS/{tenant name}/{user name}/{algorithm name}

### **{tenant name}**

這是執行工作所使用的租戶名稱。

### **{user name}**

這是要求工作的使用者。

### **{algorithm name}**

這是工作中的主要演算法。請注意，單一串流可能產生多個對映化簡工作；類似地，串流內的數個作業可以包含在單一對映化簡工作內。



---

## 第 5 章 疑難排解

Analytic Server 提供數個有用的工具來協助判斷問題。

---

### 記載

Analytic Server 會建立客戶日誌檔和追蹤檔案，對問題診斷會有幫助。藉由預設的 Liberty 安裝，您可以在 {AS\_ROOT}/ae\_wlpserver/usr/servers/aeserver/logs 目錄中找到日誌檔。

預設記載配置會產生兩個日誌檔（每天輪換）。

**as.log** 此檔案包含參考資訊警告及錯誤訊息的高層次摘要。當無法利用使用者介面中所顯示錯誤訊息來解決發生的伺服器錯誤時，應該首先檢查該檔案。

#### **as\_trace.log**

此檔案包含 as.log 中的所有項目，但會新增主要針對 IBM 支援中心及用於開發除錯的其他資訊。

Analytic Server 使用 Apache LOG4J 作為其基礎記載機能。藉由 LOG4J，記載可以透過編輯 {AS\_SERVER\_ROOT}/configuration/log4j.xml 配置檔進行動態調整。支援中心可能會要求您執行此動作來協助論斷問題，或要求您修改它來限制保留的日誌檔數目。系統會在數秒內自動偵測檔案變更，因此 Analytic Server 不需要重新啟動。

如需 log4j 及配置檔的相關資訊，請參閱官方 Apache 網站 <http://logging.apache.org/log4j/> 的說明文件。

---

### 版本資訊

您可以檢查 {AS\_ROOT}/properties/version 資料夾，來確定安裝哪個版本的 Analytic Server。下列檔案包含版本資訊。

#### **IBM\_SPSS\_Analytic\_Server-\*.swtag**

包含詳細產品資訊。

#### **version.txt**

已安裝產品的版本和建置號碼。

---

### 日誌收集器

當無法直接檢閱日誌檔來解決問題時，您可以組合所有日誌並將其傳送至 IBM 支援中心。提供了公用程式，以便輕鬆收集所有必要的資料。

使用指令 Shell，來執行下列指令：

```
cd {AS_ROOT}/bin
run >sh ./logcollector.sh
```

這些指令會在 {AS\_ROOT}/bin 下建立壓縮檔。壓縮檔包含所有日誌檔及產品版本資訊。

---

### 常見問題

本節說明了部分常見管理問題，以及如何對其進行修正。

## 安全

### 嘗試存取 HCatalog 資料來源時，Kerberos 鑑別失敗

如果您在日誌中看到類似下列錯誤：

```
cause:javax.security.sasl.SaslException: Failure to initialize security context
```

```
com.spss.analyticframework.api.exceptions.ComponentException: Cannot access HCatalog
```

您必須確保快取 HDFS 使用者的 Kerberos TGT，並在 Analytic Server 伺服器主機上可用。若要執行此動作：

1. 停止 Analytic Server 程序。
2. 透過 Analytic Server 主機執行 `kinit -f $hdfs.user`，其中 `$hdfs.user` 如 `config.properties` 檔案中所定義的，且具有分析根目錄的寫入權
3. 啟動 Analytic Server。

## Analytic Server 主控台

### 在 iOS 上透過 Safari 存取 Analytic Server 主控台

租戶狀態下拉清單不工作，且您無法更新檔案型資料來源的資料模型。執行這些動作時，請使用其他瀏覽器。

## 執行中串流

### PySpark 工作無法執行

1. 在 Ambari 主控台中，將 `export SPARK_HOME=/usr/iop/current/spark-client` 新增至 YARN 服務的 `yarn-env` 參數。
2. 重新啟動 YARN 服務。

---

## 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。

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  
Armonk, NY 10504-1785  
U.S.A.

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.  
1623-14, Shimotsuruma, Yamato-shi  
Kanagawa 242-8502 Japan

下列段落不適用於英國，若與任何其他國家之法律條款抵觸，亦不適用於該國：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爲了 (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 Software Group  
ATTN: Licensing

200 W. Madison St.  
Chicago, IL; 60606  
U.S.A.

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。

本文件中所含的任何效能資料是在控制環境中得出。因此，在其他作業環境中獲得的結果可能有明顯的差異。在開發層次的系統上可能有做過一些測量，但不保證這些測量在市面上普遍發行的系統上有相同的結果。再者，有些測定可能是透過推測方式來評估。實際結果可能不同。本文件的使用者應驗證其特定環境適用的資料。

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

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

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

本資訊僅作規劃目的。在產品可用前，此處的資訊可能變更。

本資訊含有日常商業運作所用之資料和報告範例。為了盡可能地加以完整說明，範例中含有個人、公司、品牌及產品的名稱。所有這些名稱全為虛構，任何與實際商場企業使用的名稱及地址類似之處，純屬巧合。

這些範例程式或任何衍生成果的每份複本或任何部分，都必須依照下列方式併入著作權聲明：

本資訊含有日常商業運作所用之資料和報告範例。為了盡可能地加以完整說明，範例中含有個人、公司、品牌及產品的名稱。所有這些名稱全為虛構，任何與實際商場企業使用的名稱及地址類似之處，純屬巧合。

這些範例程式或任何衍生成果的每份複本或任何部分，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.

若貴客戶正在閱讀本項資訊的電子檔，可能不會有照片和彩色說明。

---

## 商標

IBM、IBM 標誌及 [ibm.com](http://ibm.com) 是 International Business Machines Corp. 在世界許多管轄區註冊的商標或註冊商標。其他產品及服務名稱可能是 IBM 或其他公司的商標。IBM 商標的最新清單可在 Web 的 "Copyright and trademark information" 上找到，網址為 [www.ibm.com/legal/copytrade.shtml](http://www.ibm.com/legal/copytrade.shtml)。

Adobe、Adobe 標誌、PostScript 及 PostScript 標誌是 Adobe Systems Incorporated 在美國及（或）其他國家或地區的註冊商標或商標。

IT Infrastructure Library 是 Central Computer and Telecommunications Agency（現在是 Office of Government Commerce 的一部分）的註冊商標。

Intel、Intel 標誌、Intel Inside、Intel Inside 標誌、Intel Centrino、Intel Centrino 標誌、Celeron、Intel Xeon、Intel SpeedStep、Itanium 及 Pentium 是 Intel Corporation 或其子公司在美國及其他國家或地區的商標或註冊商標。



Linux 是 Linus Torvalds 在美國及（或）其他國家或地區的註冊商標。

Microsoft、Windows、Windows NT 及 Windows 標誌是 Microsoft Corporation 在美國及/或其他國家或地區的商標。

ITIL 是 Minister for the Cabinet Office 在美國 Patent and Trademark Office 註冊的註冊商標及註冊社群商標。

UNIX 是 The Open Group 在美國及其他國家或地區的註冊商標。

Cell Broadband Engine 是 Sony Computer Entertainment, Inc. 在美國及/或其他國家或地區的商標並在當地軟體使用權下使用。

Linear Tape-Open、LTO、LTO 標誌、Ultrium 及 Ultrium 標誌是 HP、IBM Corp. 及 Quantum 在美國及其他國家的商標。



Printed in Taiwan