

**IBM SPSS Analytic Server
V1**

安装指示信息

IBM

目录

IBM SPSS Analytic Server 安装指示信息 1

受支持的平台	1
先决条件	1
使用安装程序安装 Analytic Server	1
在 Hadoop 上进行手动安装	2
config.properties	4
添加 JDBC 驱动程序	5
WebSphere Liberty 配置	5
基本注册表	5
LDAP 注册表配置	6
使用 Kerberos 安全提供程序进行认证	6

启用 Hadoop 2.0 中的高可用性 (HA) 方式	8
启用 HCatalog 数据源	8
启用对 R 的必备软件的支持	9
用户入门	10
更新 Analytic Server	10
卸载 Analytic Server 服务器	11
问题确定	11
日志记录	11
版本信息	12
日志收集程序	12

IBM SPSS Analytic Server 安装指示信息

受支持的平台

表 1. 环境

产品	发行版	处理器	文字大小
Red Hat Enterprise Linux	6.x	x64	64 位
Red Hat Enterprise Linux	5.x	x64	64 位
SuSE Linux Enterprise Server	11	x64	64 位
SuSE Linux Enterprise Server	10	x64	64 位

先决条件

- 具有以下某个发行版：
 - Apache Hadoop 0.20.2
 - Apache Hadoop 1.0.3
 - IBM BigInsights 2.0
 - Cloudera CDH3
 - Cloudera CDH4 MRv1
 - Hortonworks Data Platform 1.x
- 为了使安装程序能够复制 Hadoop 发行版中的必要文件，您必须已安装了 SSH 和 SCP，并且能够使用公用密钥认证进行连接，而无需使用密码登录。否则，必须手动复制这些文件。为了验证 SSH 连接是否起作用，请执行以下操作：
 - 使用公用密钥认证登录到 Hadoop 名称节点服务器。
 - 运行以下代码，并将 {HADOOP_SERVER} 替换为您的 Hadoop 服务器。

```
export HADOOP_SERVER_USERNAME='whoami'  
export HADOOP_SERVER={HADOOP_SERVER}  
ssh -q -o "BatchMode=yes" -o "ConnectTimeout 5"  
    $HADOOP_SERVER_USERNAME@$HADOOP_SERVER "echo 2>&1"  
    && echo HOST_SSH_OK || echo HOST_SSH_NOK
```
 - 如果返回“HOST_SSH_OK”，表示已成功。否则，表示未正确配置。

使用安装程序安装 Analytic Server

安装程序会自动操作若干步骤，或者手动执行步骤。安装程序必须对 HDFS 环境有访问权。安装 Analytic Server 的步骤如下所示：

- 运行安装程序文件 `install.bin`。
 - 审阅许可协议，然后接受该协议以继续。
 - 选择安装路径。
 - 根据提示将文件安装到系统上。

- d. 配置 Analytic Server 选项。管理员帐户名称（缺省：admin）是对控制台上的 Analytic Server 应用程序具有“超级用户”管理员权限的用户名。
- e. 密钥库密码。这是一个密钥库，用于存储对存储在 Analytic Server 数据库中的数据库数据源密码进行加密时所使用的加密密钥。
- f. 自动检测 Hadoop 版本

是 指定 Hadoop Namenode Server（缺省：localhost）。如果 Hadoop 位于安装了此产品的同一机器上，那么使用 localhost 作为服务器名称。然后，安装程序将确定安装了哪个版本的 Hadoop，并从该 Hadoop 发行版中收集所需的文件。注意：必须已安装了 SSH 和 SCP，并且能够使用公用密钥认证进行连接，而无需使用密码登录。

否

- 1) 在列表中选择 Hadoop 服务器版本。
- 2) 指定 Hadoop 发行版中所需文件的位置，或者稍后手动复制这些文件。
- 3) 如果发行版为 Cloudera CDH3，请指定 Oracle Java Runtime Environment 1.6 的位置，或者稍后手动配置。

- g. 配置 Hadoop 参数

Hadoop 用户名

运行 Analytic Server 的 Unix 用户名。

密码 Hadoop 用户名的密码。

Analytic Server 根目录

此路径是 Analytic Server 设置其工作空间以及存储分发到 Hadoop 节点的可执行文件的位置。

名称节点主机

HDFS 名称节点主机的名称或 IP 地址。

名称节点端口

名称节点主机的端口号。

作业跟踪程序

Hadoop 作业跟踪程序主机的名称或 IP 地址。

作业跟踪程序端口号

作业跟踪程序的端口号。

- 2. 配置 WebSphere Liberty 以注册用户。有关更多信息，请参阅第 5 页的『WebSphere Liberty 配置』。
- 3. 通过运行 {AS_ROOT}/bin/start.sh 启动服务器，其中 {AS_ROOT} 指的是在其中部署了 Analytic Server 的文件夹；例如，/opt/IBM/SPSS/AnalyticServer/1.0。
- 4. 打开 Web 浏览器，然后输入地址：http://localhost:8080。此 URL 将打开 Analytic Server 控制台的登录对话框。以上述定义的“超级用户”身份登录。请参阅《IBM® SPSS® Analytic Server 管理员指南》，以获取有关租户管理的详细信息。

如果安装程序未能连接到 Hadoop 服务器，或未找到有效的 Hadoop 版本，请遵循手动步骤。

在 Hadoop 上进行手动安装

- 1. 运行 install.bin。这将创建 Analytic Server 数据库并设置 admin 用户凭证。
 - a. {AS_ROOT} 指的是部署 Analytic Server 的文件夹；例如 /opt/IBM/SPSS/AnalyticServer/1.0。
 - b. {AS_SERVER_ROOT} 指的是配置文件、日志文件和服务器文件的位置；例如，/opt/IBM/SPSS/AnalyticServer/1.0/ae_wlpserver/usr/servers/aeserver。

- c. {AS_HOME} 指的是 HDFS 上 Analytic Server 用作根文件夹的文件夹。
2. 编辑配置文件 {AS_SERVER_ROOT}/configuration/config.properties。有关更多信息，请参阅第 4 页的『config.properties』。根据系统的特定 Hadoop 配置来配置 Hadoop 设置 (**hdfs.namenode.url**, **hdfs.user**, **hdfs.password**)。
 3. Analytic Server 在 IBM WebSphere Liberty 的顶部运行，缺省情况下，已将其配置为在端口 8080 上运行。
 4. 将适用的 Hadoop JAR 文件复制到 {AS_SERVER_ROOT}/apps/AE_BOOT.war/WEB-INF/lib 中。这些 JAR 基于安装，并且可以在安装了 Hadoop 的 Hadoop 集群名称节点上找到；例如，在 /usr/lib/hadoop 和 /usr/lib/hadoop/lib 文件夹中。

- Apache Hadoop 0.20.2

hadoop-0.2*-core.jar

- Apache Hadoop 1.0.3

commons-configuration-*.jar

commons-io-*.jar

commons-lang-*.jar

hadoop-core-1.*.jar

jackson-core-asl-*.jar

jackson-mapper-asl-*.jar

jersey-core-*.jar

jersey-server-*.jar

- Cloudera CDH3

aspectjrt-1.6.5.jar

aspectjtools-1.6.5.jar

core-3.1.1.jar

guava-r09*.jar

hadoop-fairscheduler-0.20.2-*.jar

hadoop-core-0.20.2-*.jar

hadoop-tools-0.20.2-*.jar

jackson-core-asl-1.5.2.jar

jackson-mapper-asl-1.5.2.jar

jasper-compiler-5.5.12.jar

jasper-runtime-5.5.12.jar

Cloudera CDH3 update 5 和更高版本需要附加的 JAR 文件

jersey-core-1.8.jar

jersey-json-1.8.jar

jersey-server-1.8.jar

- Cloudera 4.0.1

aspectjrt-1.6.5.jar

aspectjtools-1.6.5.jar

commons-io-2*.jar

guava-*.jar

hadoop-fairscheduler-2*.jar

hadoop-core-2*.jar

hadoop-tools-2*.jar

jackson-core-asl-1*.jar

jackson-mapper-asl-1*.jar

jasper-compiler-5*.jar

jasper-runtime-5*.jar

- Cloudera 4.1.*

avro-*.jar

avro-compiler-*.jar

commons-cli-*.jar

commons-configuration-*.jar

commons-io-2*.jar

commons-lang-*.jar

guava-*.jar

hadoop-auth-*.jar

hadoop-common-2*.jar

hadoop-core-2*.jar

hadoop-hdfs-2.0.0-cdh4.1.2.jar

jackson-core-asl-*.jar

jackson-mapper-asl-*.jar

jersey-core-*.jar

jersey-server-*.jar

protobuf-java-*.jar

slf4j-api-*.jar

slf4j-log4j12-*.jar

- IBM BigInsights 2.0

commons-configuration-*.jar

commons-lang-*.jar

commons-io-2*.jar

hadoop-core-1.*.jar

jackson-core-asl-*.jar

jackson-mapper-asl-*.jar

jersey-core-*.jar

jersey-server-*.jar

- Hortonworks Data Platform 1.x

commons-configuration-*.jar	jackson-core-asl-*.jar
commons-lang-*.jar	jackson-mapper-asl-*.jar
commons-io-2*.jar	jersey-core-*.jar
hadoop-core-1*.jar	jersey-server-*.jar

5. Cloudera CDH3 需要 Oracle JVM 1.6。通过更新 JAVA_HOME 并将缺省路径更改为 Oracle JVM 来更新 {AE_ROOT}/bin 中的 setenv.sh 脚本。
6. 如果使用的是 Apache Hadoop 1.0.3 或更高版本，请将 {AS_SERVER_ROOT}/apps/AE_BOOT.war/WEB-INF/lib 中的 commons-io-1.4.jar 替换为 commons-io-2.x.jar。
7. 如果 Hadoop 版本为 2.0 或更高版本，请将 {AS_ROOT}/cloudera/cloudera4/hadoopmapreduce_1.1-1.0.0.0.jar 复制到 {AS_SERVER_ROOT}/apps/AE_BOOT.war/WEB-INF/lib 并移除 hadoopmapreduce-1.0.0.0.jar。
8. 请确保 HDFS 正在运行且可访问。将 Analytic Server 可执行文件复制到 HDFS，并通过运行 {AS_ROOT}/bin/hdfsUpdate.sh 来初始化 Analytic Server 工作空间。
9. 通过运行 {AS_ROOT}/bin/run.sh 来启动服务器
10. 在浏览器中输入 http://localhost:8080 来访问 Analytic Server 控制台。以 config.properties 中定义的 admin.username 身份登录。请参阅《IBM SPSS Analytic Server 管理员指南》，以获取有关租户管理的详细信息。
11. 要停止服务器，请运行 {AS_ROOT}/bin/stop.sh
12. 可以在 {AS_SERVER_ROOT}/logs/ 中找到服务器日志

config.properties

config.properties 文件中包含若干 Analytic Server 设置；下列设置相当于通过安装程序进行指定的设置；手动安装该产品之前必须确定这些设置是可接受的。

root.folder.path

此路径是存储 Analytic Server 用户工件（输入和输出）的位置。

admin.consumer=ibm

缺省管理员帐户租户。租户提供了角色、许可权、项目和数据源的高级划分。

admin.username=admin

缺省管理员帐户名称。此用户名对控制台的 Analytic Server 应用程序具有初始管理员权限。

hdfs.namenode.url=hdfs://hadoop:54310/user/hdpadding

Hadoop 名称节点的 URL。

hdfs.user=hdpadding

Hadoop 用户名。运行 Analytic Server 的 Unix 用户名。

hdfs.password=<ENTER PASSWORD>

hdfs.user 设置中指定的用户的密码。

hdfs.classpath.folder=/user/hdpadding/classpath

Analytic Server 根目录。此路径是 Analytic Server 设置其工作空间以及存储分发到 Hadoop 节点的可执行文件的位置。

mapred.job.tracker=hadoop:54311

Hadoop 作业跟踪程序主机的名称或 IP 地址和端口号。

encryption.keystore.password

这是一个密钥库，用于存储对存储在 Analytic Server 数据库中的数据库数据源密码进行加密时所使用的加密密钥。

添加 JDBC 驱动程序

要支持数据库数据源，必须将 JDBC 驱动程序添加到 Analytic Server 中。

1. 通过运行 {AS_ROOT}/bin/stop.sh 来停止 Analytic Server
2. 将所需的 JDBC 驱动程序 JAR 复制到 {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib 中
3. 通过运行 {AS_ROOT}/bin/hdfsUpdate.sh 来更新 Analytic Server
4. 通过运行 {AS_ROOT}/bin/start.sh 来启动 Analytic Server

表 2. 受支持的数据库

数据库	受支持的版本	JDBC 驱动程序 JAR	供应商
DB2 for Linux、UNIX 和 Windows	9.5、9.7 和 10.0	db2jcc.jar	IBM
DB2 z/OS	10	db2jcc.jar 和 db2_license_cisuz.jar	IBM
Teradata	13.1、14	tdgssconfig.jar 和 terajdbc4.jar	Teradata
SQL Server	2012、2008 R2	sqljdbc4.jar	Microsoft
Netezza	6.x、7	nzjdbc.jar	IBM
Oracle	12g、11g R2	ojdbc6.jar 和 orai18n.jar	Oracle

WebSphere Liberty 配置

WebSphere Liberty 概要文件是 IBM WebSphere 的轻量级实现。Analytic Server 可以使用 WebSphere 应用程序安全性来对用户进行认证。由于部署了服务器 Analytic Server，因此这是 server.xml 中的设置。要启用 Liberty 中的应用程序安全性，appSecurity-1.0 功能部件必须包含在功能管理器中：

```
<featureManager onError="FAIL">
...
<feature>appSecurity-1.0</feature>
...
</featureManager>
```

要在服务器上启动 SSL，那么必须将 SSL 功能包含到 server.xml 文件中：

```
<featureManager>
  <feature>ssl-1.0</feature>
</featureManager>
```

您可以在 ftp://ftp.software.ibm.com/software/websphere/appserv/library/v85/was85base_security.pdf 中找到有关 WebSphere 安全性的详细信息。

基本注册表

基本注册表允许管理员在 {AS_SERVER_ROOT}/server.xml 文件中定义用户和组的数据库。可以使用位于 {AS_ROOT}/ae_wlpserver/bin 的 securityUtil 工具对密码进行编码，从而对它们的值进行加密。

基本注册表在沙盒环境中有用，但不建议在生产环境中使用。

```
<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>

```

LDAP 注册表配置

LDAP 注册表为管理员提供了一种使用外部 LDAP 服务器（如 Active Directory 或 OpenLDAP）对用户进行认证的方法。此处是针对 OpenLDAP 的 ldapRegistry 示例。

```

<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>

```

有关配置的更多示例，请参阅模板文件夹 {AS_ROOT}/ae_wlpserver/templates/config。

使用 Kerberos 安全提供程序进行认证

可以配置 Kerberos 之前，必须先从 Hadoop 管理员处获取以下信息：

1. Kerberos 域；例如，ASSSO.COM
2. Kerberos 密钥分发中心 (KDC) 主机名；例如，kdc.assso.com
3. “名称”节点 Kerberos 主体；例如，hdfs/namenode.assso.com@ASSSO.COM。
4. MapReduce 节点 Kerberos 主体；例如，mapred/jobtracker.assso.com@ASSSO.COM。

然后，必须配置 /etc/krb5.conf 中的 krb5.conf 文件；例如：

```

[libdefaults]
default_realm = ASSSO.COM
default_tkt_enctypes = rc4-hmac des-cbc-md5
default_tgs_enctypes = rc4-hmac des-cbc-md5
dns_lookup_realm = false
dns_lookup_kdc = false
ticket_lifetime = 24h
forwardable = yes

[realms]
ASSSO.COM = {
  kdc = kdc.assso.com:88

```

```
default_domain = asso.com  
}
```

```
[dmain_realm]  
.asso.com = ASSSO.COM  
asso.com = ASSSO.COM
```

创建 `krb5.conf` 文件后，修改 `config.properties` 文件，如下所示：

1. 在 `Analytic Server` 模块部分：

- 添加 `hdfsauth` 和 `kerberossecurityprovider` 模块
- 移除 `wssecurityprovider` 模块。

2. 添加以下配置属性：

```
#Kerberos authentication parameters  
hadoop.security.authentication=kerberos  
dfs.namenode.kerberos.principal=hdfs/namenode.asso.com@ASSSO.COM  
mapreduce.jobtracker.kerberos.principal=mapred/jobtracker.asso.com@ASSSO.COM  
java.security.krb5.conf=/etc/krb5.conf
```

其中

hadoop.security.authentication

是 Hadoop 安全认证。指定 `kerberos` 来启用 Kerberos 安全提供程序。

dfs.namenode.kerberos.principal

用于 `keytab` 文件的 Kerberos 主体，`keytab` 文件用于启动“名称”节点。

mapreduce.jobtracker.kerberos.principal

用于 `keytab` 文件的 Kerberos 主体，`keytab` 文件用于启动作业跟踪程序。

java.security.krb5.conf

是 Kerberos 配置文件位置。

3. 配置 `service.xml` 中的 Liberty LDAP 用户存储库。请参阅第 5 页的『WebSphere Liberty 配置』。Liberty 用户存储库中指定的所有用户必须与 Kerberos 用户帐户相匹配，并且必须在 Kerberos 服务器中使用同一 LDAP 设置。
4. 缺省情况下，`Analytic Server` 使用用户的主目录下的 `.temp` 目录作为临时目录，但如果您想将临时目录配置在不同的位置，请按照以下指示信息进行操作。
- a. 编辑 `config.properties` 并取消注释以下配置设置。

```
#as.temp.folder=/.temp
```

根据需要设置更改为临时目录的绝对路径。如果使用 `/.temp` 目录，则不需要进行任何更改。

- b. 更改此文件夹的许可权以允许所有用户访问；例如，`hadoop fs -chmod 777 /.temp`。
5. 为所有 Kerberos 用户授予 `classpath` 和 `configuration` 文件夹的读许可权。
- a. 打开 `config.properties` 文件并记录下参数 `hdfs.classpath.folder` 和 `component.framework.bin.path` 的设置。
- b. 设置读许可权，如下所示。

如果 `hdfs.classpath.folder=/user/hdadmin/classpath` 并且 `component.framework.bin.path=/user/hdadmin/configuration`，请运行

```
hadoop fs -chmod -R 755 /user/hdadmin  
hadoop fs -chmod -R 755 /user/hdadmin/classpath  
hadoop fs -chmod -R 755 /user/hdadmin/configuration
```

启用 Hadoop 2.0 中的高可用性 (HA) 方式

Analytic Server 支持以 Hadoop 定制方式（如高可用性 (HA) 方式）运行，方法是向 Analytic Server 配置文件提供集群客户机 API 配置。仅当存在定制设置（如 HA）时，才强制要求向 Analytic Server 提供 Hadoop 集群客户机 API 配置。只有具有了这些设置，Analytic Server API 才能在集群名称节点失败时均衡请求。

为了配置定制方式，在成功安装 Analytic Server 之后，请执行以下步骤：

1. 获取集群客户机配置文件（`hdfs-site.xml` and `core-sites.xml`），这些文件通常位于名称节点机器上的 `/etc/hadoop/conf` 中。
2. 将这些文件复制到 `{AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/hadoop-conf`。
3. 如果 Hadoop 集群配置为 HA 方式，请确保 `config.properties` 文件中的 Analytic Server 配置属性 `hdfs.namenode.url` 指向 HDFS 服务名称和 `/user/Username`；例如，`hdfs://nameservice1/user/hdadmin`。
4. 如果已经为作业跟踪程序服务设置 HA，请更新 `config.properties` 中的 `mapred.job.tracker` 属性以指向 HDFS 服务名称。
5. 通过执行以下命令来更新 Hadoop 文件系统：
`{AS_ROOT}/bin/hdfsUpdate.sh`
6. 通过执行以下命令来启动 Analytic Server：
`{AS_ROOT}/bin/run.sh`

启用 HCatalog 数据源

为了将 Analytic Server 配置为与 HCatalog 数据库一起使用，在成功安装 Analytic Server 之后，请执行以下步骤：

1. 将 `hcataloginput` 条目添加到 `{AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/config.properties` 文件的 Analytic Server 模块列表中。例如：

```
ae.modules=securityprovidermanager,\
localClient,\
componentframework,\
...
hcataloginput
```
2. 取消注释或向 `config.properties` 添加下列行

```
hive.metastore.local=false
hive.metastore.uris=thrift://hostname:portnum
```

其中

hostname

是用于托管 Thrift 服务器的机器名

portnum

HCatalog 安装脚本中所使用的端口号

3. 请确保 `{AS_ROOT}/ae_wlpserver/usr/servers/aeserver/apps/AE_BOOT.war/WEB-INF/lib` 目录中提供了以下文件，并将这些文件复制到 HDFS 目录 `/user/{ae_admin}/classpath` 中。

下列 JAR 文件适用于 HCatalog 0.4.0 和 Hive 0.9.0。必须采用相应的 HCatalog 和 Hive，以及适用于其他版本的从属 JAR 文件。

- 可以从 HCatalog 服务器的安装目录复制以下文件。
`hcatalog-0.4.0.jar`

- 可以从 Hive 服务器的安装目录复制以下文件。

```
hive-exec-0.9.0.jar
hive-metastore-0.9.0.jar
libfb303-0.7.0.jar
slf4j-api-1.6.1.jar
slf4j-log4j12-1.6.1.jar
```

- 可以从 Hadoop 1.X 的安装目录或通过 <http://jackson.codehaus.org/> 复制以下文件。

```
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar
```

下列 JAR 文件适用于 HCatalog 0.5.0。

- 可以从 <http://code.google.com/p/guava-libraries/wiki/Release13> 复制下列文件。

```
guava-13.0.1.jar
```

4. 如果打算使用 Hive 复杂类型（图、数组和结构），请将以下 JAR 文件复制到 Hive /usr/local/hive/hive-0.9.0/auxlib/ 目录中：

```
hcatalog-0.4.0.jar
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar
```

5. 基于压缩序列文件的 HCatalog 数据源要求 Hadoop 本机库在 Analytic Server 机器上可用。将在 {HADOOP}/lib/native/Linux-amd64-64 中找到的 Hadoop 本机库复制到 Analytic Server 机器上的目录中，然后编辑 {AS_ROOT}/bin/start.sh 中的 **LIB_PATH** 变量以包含该目录的路径。例如：

```
export LIB_PATH=$AE_BASE/ae_wlpserver/usr/servers/aeserver/configuration/lib_32:
          $AE_BASE/ae_wlpserver/usr/servers/aeserver/configuration/lib_64:
          <hadoop_native_libraries_directory>:
```

注：如果 Analytic Server 安装已经有这些 JAR 文件的更新版本，那么您不应复制较早的版本。例如，如果 Analytic Server 安装程序已从更新的 Hadoop 发行版复制 jackson-core-asl-1.8.0，那么您不应复制 jackson-core-asl.1.7.3 JAR 文件。

启用对 R 的必备软件的支持

Analytic Server 支持对 R 进行评分，并支持运行 R 脚本。

为了配置对 R 的支持，成功安装 Analytic Server 后，请执行以下步骤：

1. 使用以下步骤在托管 Analytic Server 的服务器上安装 R 引擎：

```
mkdir /home/hdpadmin/APPS/R
cd /home/hdpadmin/APPS/R
wget http://cran.r-project.org/src/base/R-2/R-2.15.2.tar.gz
tar -xzf R-2.15.2.tar.gz
cd R-2.15.2
./configure --enable-R-shlib
make
```

2. 通过运行安装程序文件 install.bin 并遵循屏幕上的指示信息来在托管了 Analytic Server 的服务器上安装 R 的必备软件。安装程序：

- a. 在 Analytic Server 上更新 R 引擎安装；它将添加“R 插件”，并
- b. 更新 {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/ext_64/bin 目录以向 pasw.rstask 模块添加本机库和配置文件。

3. 为 Hadoop 部署 R 引擎和 R 组件。

- a. 如果 Analytic Server 与所有 Hadoop 节点有相同的操作系统版本和相同的处理器体系结构，那么：
 - 1) 使用以下脚本为 R 引擎创建归档

```
#!/usr/bin/env bash
echo Creating R.zip...
cd /tmp
rm -r -f R
rm -f R.zip
mkdir R
cp -r $R_HOME/* ./R/
cp $(ldd ./R/bin/exec/R ./R/bin/Rscript|cut -d\ -f3|grep \.so\.|sort -u) ./R/lib/
cp -P /usr/lib64/libgfortran.so.3 ./R/lib
cp /usr/lib64/libgfortran.so.3.0.0 ./R/lib
rm -r ./R/doc ./R/src ./R/include ./R/tests
zip -r R.zip R
```

2) 将归档 R.zip 复制到 {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/app_64 目录中。

b. 运行 {AS_ROOT}/bin/hdfsUpdate.sh, 从而将更改传送给 HDFS。

c. 如果 Analytic Server 的操作系统版本与 Hadoop 节点的不同, 请在每个 Hadoop 节点上安装 R 引擎和 R 的必备软件, 并且安装目录名称与 Analytic Server 中 R 引擎的安装目录名称相同。在 Hadoop 节点上安装 R 的必备软件时, 只需指定 R 引擎的位置, 可以跳过指定 ../ext_64/bin 位置的步骤。

注: R 引擎应安装在所有用户都可以访问的位置, 因为 R 将以不同于 Analytic Server 用户的用户身份在 Hadoop 集群中运行。

还必须在托管 SPSS Modeler 的机器上安装 R 的必备软件。请参阅《IBM SPSS Analytic Server 安装指示信息的 IBM SPSS Modeler 更新和扩展》, 以获取详细信息。

用户入门

告诉用户导航至 <http://<host>:<port>/<tenant>> 并输入用户名和密码以登录到 Analytic Server 控制台。

<host>

Analytic Server 主机的地址

<port>

Analytic Server 正在侦听的端口

<tenant>

在多租户环境中, 指您所属的租户。在单租户环境中, 缺省租户为 **ibm**。

要访问 IBM SPSS Analytic Catalyst, 请导航至 <http://<host>:<port>/catalyst.html> 并输入用户名和密码进行登录。

更新 Analytic Server

使用安装程序进行重新安装时, 您可以更新 Analytic Server 选项、“密钥库密码”参数和“配置 Hadoop”参数。

更新 Hadoop 服务器版本

有时可能有必要更新运行 Analytic Server 的 Hadoop 服务器的版本, 例如, 将 Cloudera 3.x 升级到 Cloudera 4.x。在这种情况下, 有必要使用新的 Hadoop 服务器版本中的客户机 JAR 来更新 Analytic Server。

注: 这些步骤假设与 Hadoop 有关的配置属性均未更改, 例如: 名称节点主机/端口、作业跟踪程序主机/端口或 Hadoop 用户名/密码。

1. 通过执行以下命令来停止 Analytic Server。

```
{AS_ROOT}/bin/stop.sh
```

2. 按照供应商提供的用于升级 Hadoop 服务器的指示信息进行操作，并确保该服务器正在运行。
3. 将必需客户机 JAR 复制到 Analytic Server 的本地目录中；该目录将称为 {HADOOP_SRC_FOLDER}。必需 JAR 的列表取决于 Hadoop 服务器的新版本，可以在第 2 页的『在 Hadoop 上进行手动安装』中找到。
4. 执行在 {AS_ROOT}/bin/hadoop_setup 中找到的 install_hadoop_files shell 脚本，在此过程中提供 {HADOOP_SRC_FOLDER} 作为参数。例如：./install_hadoop_files.sh /opt/hadoop_src。
 - a. 选择是否要移除任何已安装的 Hadoop 文件（建议这样做）。
 - b. 选择要使用的 Hadoop 版本。
5. 通过执行以下命令使用 Analytic Server 更新 Hadoop 文件系统。

```
{AS_ROOT}/bin/hdfsUpdate.sh
```
6. 通过执行以下命令重新启动 Analytic Server。

```
{AS_ROOT}/bin/run.sh
```

卸载 Analytic Server 服务器

1. 运行 {AS_ROOT}/_uninstall/uninstaller.bin。

注：该操作不会移除 HDFS 上的 Analytic Server 文件夹中的任何数据。要移除整个 Analytic Server 主文件夹，请移除存储 Analytic Server 用户工件（输入和输出）的 HDFS 目录。这是 {AS_SERVER_ROOT}/configuration/config.properties 中指定作为 root.folder.path 的目录。

注：如果卸载程序由于 Java OutOfMemoryException 无法启动，那么应增加最大堆大小。{AS_ROOT}/_uninstall/uninstaller.bin.lax 中的 **lax.nl.java.option.java.heap.size.max** 属性定义将调用的安装程序的最大堆大小（以字节为单位）。该值始终按字节而不是千字节或兆字节来指定，类似于 VM 参数 -mx 或 Xmx。缺省值为 50331648 (48 MB)。我们建议将该值更改为 134217728 (128 MB)，然后再启动卸载程序。

问题确定

Analytic Server 提供了多种用于问题确定的有用工具。

日志记录

Analytic Server 会创建对诊断问题有用的客户日志文件和跟踪文件。在 Liberty 进行缺省安装的情况下，您可以在 {LIBERTY_PROFILE_FOLDER}/logs 日志目录中找到日志文件。

缺省的日志记录配置会生成两个每日滚动的日志文件。

ae.log

此文件包含了信息警告和错误消息的高级摘要。出现无法使用用户界面中所显示的错误消息来解决的服务器错误时，请首先检查此文件。

ae_trace.log

此文件汇总包含了 ae.log 中的所有项，不过，出于调试目的，还添加了主要针对 IBM 支持和开发的其他信息。

Analytic Server 将 Apache LOG4J 作为其底层的日志记录工具。使用 LOG4J 可通过编辑 {AS_SERVER_ROOT}/configuration/log4j.xml 配置文件来动态调整日志记录。“支持中心”可能会要求您执行此操作来帮助诊断问题，或者要求您进行修改以限制保留的日志文件数。数秒中内就会自动检测到该文件的更改，因此 Analytic Server 无需进行重新启动。

有关 log4j 及配置文件的更多信息，请参阅 Apache 官方 Web 站点 <http://logging.apache.org/log4j/> 上的文档。

版本信息

可以通过检查 {AS_ROOT}/properties/version 文件夹来确定安装了哪个版本的 Analytic Server。以下文件中包含版本信息。

analytic_engine.swtag

包含详细的产品信息。

version.txt

已安装产品的版本和构建号。

日志收集程序

如果无法通过直接查看日志文件来解决问题，那么可以捆绑所有的日志，并将其发送到 IBM 支持。此处提供了一个实用程序，使您能够方便地收集所有必要的信息。

使用命令 shell 来运行下列命令

```
cd <AS_ROOT>/tools/support/logcollector  
run >sh ./logcollector.sh
```

这些命令将在 <AS_ROOT>/tools/support/logcollector 下创建一个压缩文件。该压缩文件中包含所有日志文件和产品版本信息。



Printed in China