

IBM SPSS Analytic Server
Versión 2

Guía de instalación y configuración

IBM

Aviso

Antes de utilizar esta información y el producto al que da soporte, lea la información de "Avisos" en la página 31.

Información sobre el producto

Esta edición se aplica a la versión 2, release 0, modificación 0 de IBM SPSS Analytic Server y a todos los releases y modificaciones posteriores hasta que se indique lo contrario en nuevas ediciones.

Contenido

Capítulo 1. Visión general	1	Capítulo 4. Migración	19
Arquitectura	2	Utilización de la herramienta de migración	20
Novedades para los instaladores en la versión 2	2	Capítulo 5. Realización de cambios en el sistema	23
Capítulo 2. Instalación	5	Utilización de la herramienta de administración	23
Requisitos previos	5	Edición manual de las propiedades de configuración	24
Requisitos previos del clúster de alta disponibilidad	5	Despliegue en nuevos miembros del clúster de alta disponibilidad	26
Instalación	6	Capítulo 6. Desinstalación	27
Capítulo 3. Configuración	9	Capítulo 7. Resolución de problemas	29
Utilización de la herramienta de configuración	9	Avisos	31
Crear un registro de usuarios	12	Marcas registradas	33
Habilitación de orígenes de datos de HCatalog	13		
Habilitación del soporte para Essentials for R	15		
Configuración de IBM SPSS Modeler para su uso con IBM SPSS Analytic Server	16		

Capítulo 1. Visión general

IBM® SPSS Analytic Server es una solución de análisis masivo de datos que combina tecnología de IBM SPSS con sistemas de datos masivos, y que permite trabajar con interfaces de usuario de IBM SPSS conocidas para resolver problemas a una escala antes impensable.

Por qué es importante el análisis masivo de datos

El volumen de datos recopilados por las organizaciones crece de forma exponencial; por ejemplo, las empresas financieras y de venta al por menor tienen todas las transacciones de clientes de un año (o dos años, o diez), los proveedores de telecomunicaciones tienen los registros de datos de llamadas (CDR) y lecturas de sensores de dispositivos, y las empresas de internet tienen los resultados de los rastreos web.

Un análisis masivo de datos es necesario cuando existe:

- Un gran volumen de datos (terabytes, petabytes o exabytes), sobre todo cuando es una mezcla de datos estructurados y no estructurados.
- Datos que cambian/se acumulan con rapidez.

El análisis masivo de datos también es de ayuda cuando:

- Se construye un gran número de modelos (del orden de miles).
- Los modelos se construyen/renuevan con frecuencia.

Retos

Las mismas organizaciones que recopilan grandes volúmenes de datos suelen tener dificultades a la hora de utilizarlos, por una serie de razones:

- la arquitectura de los productos analíticos tradicionales no está pensada para la computación distribuida, y
- Los algoritmos estadísticos existentes no están diseñados para trabajar con cantidades masivas de datos (tales algoritmos esperan que los datos les lleguen, pero cuesta mucho mover datos masivos), por tanto
- el análisis de datos masivos con tecnología puntera requiere nuevas habilidades y un conocimiento a fondo de los sistemas de datos masivos. Muy pocos analistas poseen estas habilidades.
- Las soluciones residentes en memoria son aptas para problemas de tamaño medio, pero no escalan bien a datos realmente masivos.

Solución

Analytic Server proporciona:

- Una arquitectura centrada en datos que saca partido de sistemas de datos masivos tales como Hadoop Map/Reduce con datos en HDFS.
- Una interfaz definida para incorporar nuevos algoritmos estadísticos diseñados para ir a los datos.
- Conocidas interfaces de usuario de IBM SPSS que ocultan los detalles de los entornos de datos masivos, de modo que el analista pueda centrarse en el análisis de los datos.
- Una solución escalable a problemas de cualquier tamaño.

Arquitectura

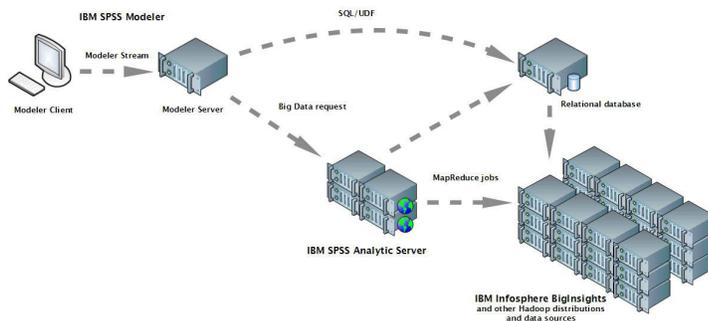


Figura 1. Arquitectura

Analytic Server se sitúa entre una aplicación cliente y una nube Hadoop. Suponiendo que los datos residan en la nube, la forma de trabajar con Analytic Server sería a grandes rasgos:

1. Se definen los orígenes de datos de Analytic Server que dan acceso a los datos de la nube.
2. Se define el análisis que se desea realizar en la aplicación cliente. En el release actual, la aplicación cliente es IBM SPSS Modeler.
3. Cuando se ejecuta el análisis, la aplicación cliente envía una solicitud de ejecución de Analytic Server.
4. Analytic Server organiza el trabajo para que ejecute en la nube de Hadoop e informa de los resultados a la aplicación cliente.
5. Los resultados pueden utilizarse para definir análisis adicionales, con lo que se repetiría el ciclo.

Novedades para los instaladores en la versión 2

Instalación y configuración

La instalación del producto ahora utiliza IBM Installation Manager para proporcionar una experiencia sin fisuras al instalar varios productos de IBM.

Los pasos para configurar Analytic Server se manejan en una nueva herramienta de configuración que se ejecuta después de instalar todas las ampliaciones del producto.

Plataforma

Soporte para nuevos sistemas operativos y distribuciones de Hadoop.

Sistemas operativos

Analytic Server se ejecuta ahora en Microsoft Windows.

Distribuciones de Hadoop

Analytic Server puede configurarse para que se ejecute con Hortonworks Data Platform en Microsoft Windows.

Adaptadores de HCatalog

Pueden definirse orígenes de datos para Apache Cassandra, MongoDB, y Oracle NoSQL. Hay pasos de configuración adicionales para habilitar estos orígenes de datos.

Clústeres de alta disponibilidad

Analytic Server puede instalarse y configurarse en un clúster de alta disponibilidad utilizando IBM WebSphere Liberty Network Deployment.

Capítulo 2. Instalación

Requisitos previos

Antes de instalar Analytic Server, revise la información siguiente.

Requisitos del sistema

Para obtener la información más actualizada sobre los requisitos del sistema, utilice los informes detallados de requisitos del sistema en el sitio de soporte técnico de IBM: <http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html>. En esta página:

1. Especifique SPSS Analytic Server como nombre de producto y pulse **Search**.
2. Seleccione la versión deseada y el ámbito del informe y, a continuación, haga clic en **Submit**.

Privilegios de administración

El usuario que instala Analytic Server debe tener acceso de grabación al sistema de archivos distribuido de Hadoop.

IBM Installation Manager

Debe instalarse y configurarse una versión actual de IBM Installation Manager para utilizar un repositorio que contenga los archivos de instalación de Analytic Server. Si tiene una versión anterior de IBM Installation Manager, se le solicitará que la actualice a una versión admitida como parte de la instalación.

Si IBM Installation Manager no está presente en el sistema, descargue e instale IBM Installation Manager del sitio de soporte de IBM Corp.: <http://www.ibm.com/support>.

Importante: En Microsoft Windows, instale IBM Installation Manager desde una cuenta de administrador. Es posible realizar una instalación de grupo con otro usuario con: `./groupinstc -acceptLicense`. Consulte la documentación de IBM Installation Manager para obtener más detalles. http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html

Hive/HCatalog

Si tiene previsto utilizar los orígenes de datos NoSQL, configure Hive y HCatalog para el acceso remoto. Asegúrese también de que `hive-site.xml` contiene una propiedad `hive.metastore.uris` de la forma `thrift://<nombre_host>:<puerto>` que señala al servidor Thrift Hive Metastore activo. Consulte la documentación de distribución de Hadoop para obtener detalles.

Repositorio de metadatos

De forma predeterminada, Analytic Server instala y utiliza una base de datos Apache Derby. De forma alternativa, puede configurar Analytic Server para que utilice una instalación existente de DB2. La base de datos debe tener una codificación UTF-8. IBM DB2 10.5 Workgroup Server Edition está disponible en el soporte de Analytic Server. Para obtener más información: http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html.

Clúster de alta disponibilidad

Consulte “Requisitos previos del clúster de alta disponibilidad” para conocer los requisitos previos para la instalación de Analytic Server en un clúster de alta disponibilidad.

Requisitos previos del clúster de alta disponibilidad

Existen requisitos previos adicionales para la instalación de Analytic Server en un clúster de alta disponibilidad.

Configuración de sistemas y DNS

Los sistemas que se utilizan en el entorno deben poder comunicarse entre sí. Sólo los sistemas que pertenecen a un dominio deben tener el sufijo de dominio. Por ejemplo, sólo los sistemas registrados en el dominio "IBM.COM" deben tener el sufijo "ibm.com" para el nombre de dominio completo (FQDN).

Si un sistema que no se encuentran en el dominio IBM.COM tiene un sufijo de FQDN "ibm.com", el FQDN del sistema debe especificarse en DNS con los sistemas que tienen el servidor DNS en la configuración de TCP. De lo contrario, todos los FQDN de todos los sistemas que participan deben estar en los archivos hosts de todos los sistemas implicados.

Repositorio de metadatos

Analytic Server debe utilizar DB2 cuando se instala en un clúster de alta disponibilidad.

Zookeeper 3.4.5

Son necesarias tres instancias de Zookeeper como mínimo. Consulte la documentación de instalación y configuración de Zookeeper: http://zookeeper.apache.org/doc/r3.4.5/zookeeperStarted.html#sc_RunningReplicatedZooKeeper.

Nota: estos pasos deben llevarse a cabo en cada sistema donde está instalado Zookeeper.

Liberty Network Deployment 8.5.5.4

El entorno de alta disponibilidad consta de un colectivo Liberty con un controlador, dos réplicas y al menos dos miembros del clúster. Consulte la documentación de WebSphere Application Server Network Deployment para obtener detalles acerca de la configuración de un colectivo Liberty: http://www-01.ibm.com/support/knowledgecenter/SSAW57/mapfiles/product_welcome_wasnd.html.

Equilibrador de carga

El clúster de alta disponibilidad debe tener un equilibrador de carga que dé soporte a la afinidad de sesiones, a veces también conocida como sesiones permanentes. Analytic Server identifica las sesiones con la cookie "request-token". Ésta identifica una sesión durante la duración de un inicio de sesión del usuario para su uso en la afinidad de sesiones controlada por aplicación. Consulte la documentación de su equilibrador de carga para conocer los detalles de soporte a la afinidad de sesiones.

Nota: El nodo líder no se debe incluir en el equilibrador de carga del clúster. El modo de identificar el líder es que se trata del primer nodo que inicia.

Visual C++ 2008 (solo Windows)

Si tiene un miembro de clúster que ejecuta Analytic Server en Windows, tendrá que instalar Visual C 2008 Runtime antes de instalar Analytic Server.

Instalación

Los archivos de la aplicación Analytic Server se instalan en el sistema host con el Gestor de instalación de IBM. Los archivos de instalación se pueden encontrar en el soporte físico o descargarse de IBM Passport Advantage.

1. Inicie IBM Installation Manager:

Modalidad de GUI

{directorio de instalación de IBM Installation Manager}/eclipse/IBMIM

Modalidad de línea de mandatos

{directorio de instalación de IBM Installation Manager}/eclipse/tools/imcl -c

Nota: El usuario que ejecuta el Gestor de instalación de IBM es el usuario que ejecuta el proceso de Analytic Server. Este usuario necesita acceso al entorno Hadoop.

2. Si el repositorio de instalación no está configurado, especifique la vía de acceso del repositorio. Por ejemplo, especifique el disco de instalación montado o una ubicación del sistema de archivos del host, la red o una dirección HTTP.
3. En el menú principal, seleccione **Instalar**.
4. Seleccione el paquete que debe instalarse. La opción disponible depende de la versión de Analytic Server que haya adquirido.

IBM SPSS Analytic Server

Este paquete es adecuado para la mayoría de los sistemas.

IBM SPSS Analytic Server Alta disponibilidad

Este paquete es necesario si tiene previsto ejecutar Analytic Server en un clúster de alta disponibilidad. Las diferencias en la configuración de IBM SPSS Analytic Server Alta disponibilidad se documentan en el lugar pertinente.

5. Lea el acuerdo de licencia y acepte los términos.
6. Especifique el grupo de paquetes y el directorio de instalación.
 - Se requiere un grupo de paquetes nuevo para instalar en IBM SPSS Analytic Server.
 - Especifique el directorio de instalación para los recursos compartidos. Puede especificar el directorio de recursos compartidos sólo la primera vez que instale un paquete.
7. Seleccione las características que se instalarán.
8. Revise la información de resumen y continúe con la instalación. Los archivos de la aplicación se instalarán en el directorio especificado.
9. Es necesario realizar una configuración adicional. Para obtener más información sobre una nueva configuración, consulte el tema: "Utilización de la herramienta de configuración" en la página 9. Para obtener información sobre la migración de una configuración existente a una nueva instalación, consulte el tema: Capítulo 4, "Migración", en la página 19. Si está ejecutando IBM Installation Manager en la modalidad de GUI, la opción para iniciar la herramienta de migración o de configuración se presenta en el último paso.

Si se producen problemas durante la instalación, puede resolverlos utilizando los registros de IBM Installation Manager. Puede acceder a los archivos de registro desde el menú principal de IBM Installation Manager.

Para automatizar la instalación, IBM Installation Manager puede ejecutarse en modalidad silenciosa. Para obtener más información, consulte la documentación de IBM Installation Manager:
http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

Capítulo 3. Configuración

Después de la instalación, debe configurar Analytic Server antes de poder utilizar el producto.

Nota: Para las vías de acceso de archivo de Analytic Server se utilizan las convenciones siguientes.

- {RAÍZ_AS} hace referencia a la ubicación en la que está desplegado Analytic Server; por ejemplo, >/opt/IBM/SPSS/AnalyticServer/{version}.
- {RAÍZ_SERVIDOR_AS} hace referencia a la ubicación de los archivos de configuración, registro y servidor; por ejemplo, /opt/IBM/SPSS/AnalyticServer/{version}/ae_wlpserver/usr/servers/aeserver.
- {INICIO_AS} hace referencia a la ubicación de HDFS que utiliza Analytic Server como carpeta raíz.

Pasos de la configuración

1. [Opcional] Siga los pasos para habilitar orígenes de datos HCatalog si desea ejecutar Analytic Server en orígenes de datos XML o NoSQL.

Nota: Puede realizar un Analytic Server de configuración simple ahora y habilitar los orígenes de datos de HCatalog más tarde mediante la herramienta de administración.

2. Ejecute la herramienta de configuración. La herramienta automatiza una serie de pasos que de otra forma tendrían que hacerse manualmente. Si utilizó la versión de la GUI de IBM Installation Manager, la herramientas de configuración se inicia automáticamente, en caso contrario necesitará ejecutar el script **configtool** en {RAÍZ_AS}/bin.
3. Cree un registro de usuarios que pueden añadirse al sistema de Analytic Server .
4. Configure IBM SPSS Modeler para su uso con Analytic Server.
5. [Opcional] Habilite el soporte para Essentials for R si desea que Analytic Server ejecute código R.
6. Asegúrese de que Analytic Server se esté ejecutando. Puede iniciar el proceso ejecutando el script **start** en {RAÍZ_AS}/bin.
7. Abra un navegador web y escriba la dirección <http://localhost:8080/analyticserver/admin/ibm>. Este URL abre el diálogo de inicio de sesión de la consola de Analytic Server. Inicie la sesión como el administrador de Analytic Server definido en la herramienta de configuración.

Utilización de la herramienta de configuración

La herramienta de configuración automatiza una serie de pasos que de lo contrario deberían realizarse manualmente. La herramienta de configuración debe tener acceso al entorno HDFS.

Si utilizó la versión de la GUI de IBM Installation Manager, la herramientas de configuración se inicia automáticamente. Siga las pantallas de solicitud para configurar su sistema.

También puede iniciar la herramienta de configuración ejecutando el script **configtool** en {RAÍZ_AS}\bin. De forma predeterminada, la herramienta detecta si se va a lanzar en modalidad de GUI o de consola. Puede forzar la modalidad de consola pasando el argumento **-c**, o la modalidad silenciosa pasando la vía de acceso a un archivo que especifique la configuración deseada. Consulte el archivo **configtool.response** en {RAÍZ_AS}\bin para obtener un ejemplo.

Bienvenida

El paso de bienvenida determina qué ruta tomar a través de la herramienta de configuración.

Típica La ruta Típica automatiza muchas de las decisiones y realiza suposiciones de simplificación acerca del sistema. Seleccione esta opción si desea configurar rápidamente un entorno no de

producción. Siempre puede cambiar la configuración en la herramienta de administración o migrar la configuración a otro entorno más adelante.

En una configuración típica, la herramienta le solicitará la ubicación del nombre de nodo de Hadoop, y progresivamente solicitará más información cuando sea necesario.

En una instalación típica de alta disponibilidad, la herramienta de configuración le solicitará información sobre el entorno WebSphere® Application Server Liberty Network Deployment y el conjunto Zookeeper.

Personalizada

Esta opción permite realizar tareas de configuración más avanzadas. Esta opción es necesaria si desea configurar NoSQL o bases de datos relacionales para su uso con Analytic Server de modo inmediato.

Las secciones siguientes proporcionan más detalles sobre la ruta de acceso personalizada.

Clúster de Hadoop

Los pasos del clúster de Hadoop configuran Analytic Server para su uso con la distribución de Hadoop. La herramienta de configuración puede descubrir automáticamente la distribución, o puede seleccionar manualmente los archivos necesarios.

Recopilar automáticamente los archivos de la distribución de Hadoop

La herramienta intentará detectar la distribución de Hadoop y progresivamente solicitará más información cuando sea necesario.

1. Siempre será necesario proporcionar el nombre de host del clúster de Hadoop.
2. Si el clúster no está configurado con SSH y autenticación de clave pública para el usuario que ejecuta la herramienta de configuración, tendrá que proporcionar credenciales válidas para que la herramienta de configuración pueda recopilar los archivos del clúster de Hadoop.

Restricción: Solo se admite el cifrado RSA SSH.

3. Si el clúster de Hadoop se ha instalado en una ubicación no estándar, puede que tenga que proporcionar la ubicación de instalación.

Recopilar manualmente los archivos de la distribución de Hadoop

Identifique la distribución de Hadoop que está utilizando. La herramienta de configuración identificará los archivos que debe recopilar del clúster de Hadoop. Copie estos archivos en la ubicación especificada y podrá continuar los pasos de configuración.

Nota: El archivo JAR `hadoop-lzo-*.jar` es necesario solo si se especifica el valor `"com.hadoop.compression.lzo.LzoCodec"` para el parámetro `io.compression.codecs` en el archivo `core-site.xml`.

Hcatalog/Hive

Los pasos para Hcatalog/Hive sólo se muestran como parte de una configuración Personalizada.

Conectarse automáticamente y detectar la configuración de HCatalog/Hive

La herramienta intentará detectar la versión de HCatalog/Hive y recopilará los archivos necesarios para configurar Analytic Server.

Seleccionar manualmente la configuración de Hive/HCatalog

Identifique las bases de datos NoSQL que está utilizando. La herramienta de configuración identificará los archivos que necesita recopilar. Copie estos archivos en la ubicación especificada y podrá continuar los pasos de configuración.

Omitir la configuración de HCatalog/Hive

Puede configurar HCatalog/Hive más adelante.

JDBC

Los pasos para JDBC sólo se muestran como parte de una configuración Personalizada. Copie los controladores para las bases de datos que está utilizando en la ubicación especificada y siga los pasos de configuración.

Tabla 1. Bases de datos soportadas

Base de datos	Versiones soportadas	Archivos JAR del controlador JDBC	Distribuidor
DB2 for Linux, UNIX y Windows	11, 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
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

Servidor de aplicaciones

Los pasos de servidor de aplicaciones dependen del tipo de instalación.

- En una instalación que no sea de alta disponibilidad, puede configurar la comunicación segura (SSL/TLS) y establecer los puertos en los que Analytic Server está a la escucha.
- En una instalación de alta disponibilidad, la herramienta de configuración le solicitará información sobre el entorno WebSphere® Application Server Liberty Network Deployment.

Zookeeper

El paso para Zookeeper sólo es necesario como parte de la configuración de una instalación de alta disponibilidad. Especifique la lista de servidores del conjunto Zookeeper.

Repositorio de metadatos

Los pasos para el repositorio de metadatos dependen del tipo de instalación.

- En una instalación que no sea de alta disponibilidad, tiene la opción de elegir Derby o DB2. Si elige DB2, debe estar instalada y disponible antes de ejecutar la herramienta de configuración. Consulte “Requisitos previos” en la página 5 para obtener información adicional.
- Una configuración de una instalación de alta disponibilidad siempre debe utilizar DB2 como repositorio de metadatos, y la instancia de DB2 debe estar instalada y disponible antes de ejecutar la herramienta de configuración.

Analytic Server

Identifique el sujeto principal que actúa como administrador de Analytic Server. Este sujeto principal podrá crear, suprimir y gestionar arrendatarios.

Resumen

El paso de Resumen permite revisar los pasos de configuración antes de completarla.

Postconfiguración

El paso de postconfiguración informará de que la configuración fue satisfactoria. En caso de error, revise el archivo de registro en el directorio logs.

Crear un registro de usuarios

El arrendatario predeterminado creado durante la instalación y configuración utiliza el perfil WebSphere Liberty para autenticar usuarios, que a su vez pueden hacer referencia a un servidor LDAP externo o definir un registro básico de usuarios y grupos.

Nota: Después de editar el archivo `{RAÍZ_SERVIDOR_AS}/server.xml` para configurar el perfil WebSphere Liberty para autenticar a los usuarios, añada esos usuarios al sistema de Analytic Server. Consulte la *IBM SPSS Analytic Server Guía del administrador* para obtener detalles sobre la gestión del arrendatario.

LDAP

El registro LDAP permite autenticar a los usuarios con un servidor LDAP externo como Active Directory u OpenLDAP. A continuación se muestra un ejemplo de `ldapRegistry` para OpenLDAP.

```
<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>
```

Puede obtener ejemplos adicionales de configuraciones consultando la carpeta de plantillas `{RAÍZ_AS}/ae_wlpserver/templates/config`.

Básico

El registro básico permite definir una base de datos de usuarios y grupos en el archivo `{RAÍZ_SERVIDOR_AS}/server.xml`. Las contraseñas pueden codificarse para ocultar sus valores con la herramienta `securityUtility`, ubicada en `{RAÍZ_AS}/ae_wlpserver/bin`.

El registro básico es útil en un entorno de recinto de seguridad (sandbox), pero no se recomienda en un entorno de producción.

```
<basicRegistry id="basic" realm="ibm">
  <user name="usuario1" password="{xor}Dz4sLG5tbGs="/>
  <user name="usuario2" password="Pass"/>
  <user name="usuario3" password="Pass"/>
  <user name="usuario4" password="Pass"/>
  <user name="admin" password="{xor}KzosKw=""/>
  <group name="Desarrollo">
    <member name="usuario1"/>
    <member name="usuario2"/>
  </group>
  <group name="QA">
    <member name="usuario3"/>
    <member name="usuario4"/>
  </group>
</basicRegistry>
```

```

</group>
<group name="ADMIN">
  <member name="usuario1"/>
  <member name="admin"/>
</group>
</basicRegistry>

```

Habilitación de orígenes de datos de HCatalog

Analytic Server proporciona soporte para datos XML y bases de datos NoSQL a través de HCatalog y Hive.

La herramienta de configuración prepara la instalación de Analytic Server para utilizarla con Hive/HCatalog, pero hay algunos pasos manuales. Es necesario copiar algunos archivos JAR en el directorio {INICIO_HIVE}/auxlib/ de Hive, ya que Analytic Server implementa algunas interfaces necesarias para HCatalog. Analytic Server no puede utilizar manejadores de almacenamiento de Hive directamente en HCatalog debido a que HCatalog tiene una API ligeramente diferente.

Las secciones siguientes detallan los archivos necesarios, en función de los orígenes de datos que desee utilizar.

Importante: Se deben copiar estos archivos en el directorio {INICIO_HIVE}/auxlib/ antes de ejecutar la herramienta de configuración o la herramienta de administración para que Analytic Server las utilice.

Importante: Hortonworks Data Platform en Windows no reconoce el directorio {INICIO_HIVE}/auxlib/. Para habilitar HCatalog, puede:

1. Copiar los archivos necesarios, como se describe a continuación, en el directorio {INICIO_HIVE}/lib/.
2. Eliminar el directorio {INICIO_HIVE}/auxlib/.
3. Actualizar los archivos hive-site.xml y hive-config.cmd para eliminar las referencias al directorio {INICIO_HIVE}/auxlib/.
4. Reiniciar el clúster de Hortonworks Data Platform para que entren en vigor los cambios.

Consulte <http://hortonworks.com/community/forums/topic/hive-auxlib-issues/> para obtener información adicional.

Todas las instalaciones de Hive/HCatalog

1. Los archivos siguientes se encuentran en el archivado hive_auxlib del directorio {RAÍZ_AS}/add-ons.

```
com.ibm.spss.com.hcatalogstoragehandler-{version}.jar
```

2. Copie la última versión de hivexmlserde-{versión}.jar del repositorio central de Maven en <http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>.
3. [Opcional] Si tiene previsto utilizar tipos complejos Hive (correlaciones, matrices, estructuras), copie estos archivos JAR. Pueden encontrarse en la instalación de HCatalog. Antes de Hive 0.11.0, es {INICIO_HCATALOG}/share/hcatalog; a partir de Hive 0.11.0, es {INICIO_HIVE}/hcatalog/share/hcatalog.

```
hcatalog-core-{version}.jar (para hadoop 1.2)
```

```
hive-hcatalog-{version}.jar (para hadoop 2.x)
```

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

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

4. Si utiliza Hive versión 0.14.0, bájese el archivo JAR siguiente. <https://github.com/dvasilen/Hive-HCatalog-Compatibility/raw/master/release/hive-hcatalog-compatibility-0.14.0.jar>

Data XML

1. [Opcional] Si tiene previsto trabajar con archivos XML comprimidos, copie los archivos JAR siguientes que se encuentran en el archivado hive_auxlib en el directorio {RAÍZ_AS}/add-ons.

com.ibm.spss.com.ibm.biginsights.compression_2.1.jar (sólo distribución BigInsights Hadoop)
com.ibm.spss.hivexmlserdecomp-2.0.0.0.jar

Apache Accumulo

1. Los archivos siguientes pueden encontrarse en la instalación de HCatalog. Antes de Hive 0.11.0, es {INICIO_HCATALOG}/share/hcatalog; a partir de Hive 0.11.0, es {INICIO_HIVE}/hcatalog/share/hcatalog.

commons-io-{version}.jar

2. El manejador de almacenamiento de Apache Accumulo puede encontrarse en <https://github.com/dvasilen/accumulo-hive-storage-manager/tree/master>. Esta es la ramificación maestra; desde aquí, debe conmutar a la rama correspondiente a su versión de Accumulo y de Hadoop. El archivo JAR puede encontrarse en el directorio /release y tiene un nombre con el formato:

accumulo-hive-storage-handler_{versiónHadoop}-<versiónAccumulo>.jar

- 3.

Los archivos siguientes pueden encontrarse en la instalación de Apache Accumulo.

accumulo-trace-{version}.jar

accumulo-fate-{version}.jar

accumulo-core-{version}.jar

accumulo-server-{version}.jar

accumulo-start-{version}.jar

Apache Cassandra

1. Los archivos siguientes pueden encontrarse en el release de Apache Cassandra 1.2.9 en <http://www.apache.org/dyn/closer.cgi?path=/cassandra/1.2.9/apache-cassandra-1.2.9-bin.tar.gz>

apache-cassandra-1.2.9.jar

apache-cassandra-clientutil-1.2.9.jar

apache-cassandra-thrift-1.2.9.jar

2. El controlador de almacenamiento Apache Cassandra se encuentra en <https://github.com/dvasilen/Hive-Cassandra>. Se trata de la ramificación maestra; desde aquí debe cambiar a la ramificación de la versión de Cassandra y Hadoop. El archivo JAR puede encontrarse en el directorio /release y tiene un nombre con el formato:

hive-<versiónHive>-hadoop-<versiónHadoop>-cassandra-<versiónCassandra>.jar

Nota: Para las versiones de Hadoop anteriores a la 2.0, no hay ninguna parte hadoop del nombre de archivo o ramificación.

Nota: Analytic Server puede comunicarse con Cassandra 2.0, pero sus archivos JAR se compilan para Java 7. Por lo tanto, tendrá que copiar los archivos JAR de Cassandra 1.2.9 porque se compilan para Java 6.

HBase

No hay archivos adicionales que deban copiarse para HBase.

MongoDB

1. Los archivos siguientes deben compilarse localmente para la distribución de Hadoop desde la rama "maestra" del proyecto "mongodb/mongo-hadoop". Consulte <https://github.com/mongodb/mongo-hadoop> para obtener detalles.

mongo-hadoop-core-{version}.jar

mongo-hadoop-hive-{version}.jar

2. El archivo siguiente puede encontrarse en <https://github.com/downloads/mongodb/mongo-java-driver/mongo-2.10.1.jar>

mongo-2.10.1.jar

Oracle NoSQL

1. El archivo siguiente puede encontrarse en <https://github.com/vilcek/HiveKVStorageHandler2/blob/master/dist/HiveKVStorageHandler.jar>
HiveKVStorageHandler.jar
2. El archivo siguiente puede encontrarse en la instalación de Apache Hive
avro-*<version>.jar

Habilitación del soporte para Essentials for R

Analytic Server soporta la puntuación de modelos R y la ejecución de scripts R.

Restricción: El soporte de R está limitado actualmente a las instalaciones de Analytic Server en Linux.

Para configurar el soporte de R tras una instalación satisfactoria de Analytic Server:

1. Instale R Engine en el servidor que aloja Analytic Server.

```
wget http://cran.r-project.org/src/base/R-3/R-3.1.0.tar.gz
tar -xzf R-3.1.0.tar.gz
cd R-3.1.0
./configure --enable-R-shlib
make
```

Importante: instale el motor R en una ubicación accesible para todos los usuarios, ya que R se ejecuta en el clúster de Hadoop como un usuario distinto del usuario de Analytic Server.

Importante: En una instalación de alta disponibilidad, R Engine debe estar instalado en la misma ubicación en todos los miembros de clúster.

2. Para instalar Essentials for R en el servidor que aloja Analytic Server ejecute el archivo de instalador, siguiendo las instrucciones que aparecen en pantalla. Essentials for R está disponible para descarga (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspsp>).
3. Despliega el motor R y el componente R en Hadoop.

a. Si Analytic Server y todos los nodos de Hadoop tienen la misma versión del sistema operativo y la misma arquitectura de procesador:

- 1) Crea un archivo para el motor R utilizando el script siguiente

```
#!/usr/bin/env bash
echo Creando R.zip...
cd /tmp
rm -r -f R
rm -f R.zip
mkdir R
export R_HOME=/home/hdpadmin/APPS/R/R-3.1.0
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
```

Nota: Hay dos espacios entre -d\ y -f3 en el script. El documento PDF ofrece el aspecto de dos espacios, pero se trata de un único espacio amplio. Al copiar y pegar este script desde el PDF, inserte el segundo espacio antes de ejecutar el script.

Nota: Las líneas que copian bibliotecas Fortran dependen de la versión de Fortran instalada. Por ejemplo, si está instalada la versión 1.0, esas líneas son las siguientes. Edítelas según sea necesario para el sistema.

```
cp -P /usr/lib64/libgfortran.so.1 ./R/lib
cp /usr/lib64/libgfortran.so.1.0.0 ./R/lib
```

- 2) Copie el archivo R.zip en el directorio {RAÍZ_AS}/ae_wlpserver/usr/servers/aeserver/configuration/linux/app_64.
- b. Si las versiones de los sistemas operativos para los nodos de Analytic Server y Hadoop son diferentes, instale el motor R y Essentials for R en cada nodo Hadoop. Instálelos en un directorio con la misma vía de acceso y nombre con los que se ha instalado el motor de R en Analytic Server. Si instala Essentials for R en el nodo de Hadoop, especifique el directorio en el que está instalado R y omita el paso de especificar la ubicación de ../ext_64/bin.
- c. Ejecute el script **hdfsUpdate** en el directorio {RAÍZ_AS}/bin para propagar los cambios a HDFS.
4. En una instalación de alta disponibilidad, debe ejecutar también el script **deployRemote** en el directorio {RAÍZ_AS}/bin para propagar los cambios a las instancias de Analytic Server en todos los miembros de clúster.
5. También deberá instalar Essentials for R en la máquina que aloja SPSS Modeler Server.

Configuración de IBM SPSS Modeler para su uso con IBM SPSS Analytic Server

Para habilitar SPSS Modeler a fin de utilizarlo con Analytic Server, debe realizar algunas actualizaciones en la instalación de SPSS Modeler Server.

1. Configure SPSS Modeler Server para asociarlo con una instalación de Analytic Server.
 - a. Edite el archivo `options.cfg` en el subdirectorio `config` del directorio de instalación del servidor principal, y añada o edite las líneas siguientes:

```
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_ssl_enabled

Especifique Y si la comunicación segura está configurada en Analytic Server; de lo contrario, especifique N.

as_host

La dirección IP del servidor que aloja Analytic Server.

as_port

El puerto en el que Analytic Server está a la escucha (el valor predeterminado es 8080).

as_context_root

La raíz de contexto de Analytic Server (el valor predeterminado es `analyticserver`).

as_tenant

El arrendatario del que la instalación de SPSS Modeler Server forma parte (el arrendatario predeterminado es `ibm`).

as_prompt_for_password

Especifique N si SPSS Modeler Server está configurado con el mismo sistema de autenticación de usuarios y contraseñas que el utilizado en Analytic Server; de lo contrario, especifique Y.

Al ejecutar SPSS Modeler en modalidad de proceso por lotes, añada `-analytic_server_username {ASusername} -analytic_server_password {ASpassword}` como argumentos al mandato `clemb`.

- b. Reinicie el servicio de SPSS Modeler Server.

Para poder conectarse a una instalación de Analytic Server que tiene habilitado SSL/TLS, deben realizarse algunas tareas adicionales para configurar las instalaciones de SPSS Modeler Server y de cliente.

- a. Navegue a `http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT}` e inicie la sesión en la consola de Analytic Server.
 - b. Descargue el archivo de certificación del navegador y guárdelo en su sistema de archivos.
 - c. Añada el archivo de certificación en el JRE de las instalaciones de SPSS Modeler Server y de SPSS Modeler Client. La ubicación de actualizaciones puede encontrarse en el subdirectorio `/jre/lib/security/cacerts` de la vía de instalación de SPSS Modeler.
 - 1) Asegúrese de que el archivo `cacerts` no sea de sólo lectura.
 - 2) Utilice el programa `keytool` que se suministra con Modeler - que puede encontrarse en el subdirectorio `/jre/bin/keytool` de la vía de instalación de SPSS Modeler.
Ejecute el siguiente mandato


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

 Tenga en cuenta que `<alias-as>` es un alias para el archivo `cacerts`. Puede utilizar cualquier nombre que desee, siempre y cuando sea exclusivo para el archivo `cacerts`.
Un mandato de ejemplo podría ser parecido al siguiente.


```
keytool -import -alias MySSLCertAlias -file C:\Download\as.cer
-keystore "c:\Program Files\IBM\SPSS\Modeler\{ModelerVersion}\jre\lib\security\cacerts"
```
 - d. Reinicie SPSS Modeler Server y SPSS Modeler Client.
2. [opcional] Instale IBM SPSS Modeler - Essentials for R si tiene previsto puntuar modelos R en secuencias con orígenes de datos de Analytic Server. IBM SPSS Modeler - Essentials for R está disponible para descarga (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>).

Capítulo 4. Migración

Analytic Server admite varios escenarios en los que puede migrar los valores de configuración de una instalación de Analytic Server existente a una nueva instalación.

Actualice a una versión nueva de Analytic Server

Si tiene una instalación existente de Analytic Server 1.0.1 y ha adquirido la versión 2.0, puede migrar los valores de configuración de 1.0.1 a la instalación de 2.0. En este escenario, la herramienta de migración le guiará por la ruta adecuada.

Restricción: Si tiene instalada la versión 1.0, debe migrar primero de la versión 1.0 a la 1.0.1 y luego de la versión 1.0.1 a la 2.0.

Restricción: Las instalaciones de 1.0.1 y 2.0 no pueden coexistir en el mismo clúster de Hadoop. Si configura la instalación de 2.0 para utilizar el mismo clúster de Hadoop que la instalación de 1.0.1, la instalación de 1.0.1 ya no funcionará.

Cree una copia del Analytic Server existente

Si tiene una instalación existente de Analytic Server 2.0, puede replicar los valores de configuración en una nueva instalación de 2.0. Esto resulta de utilidad, por ejemplo, si tiene un entorno de desarrollo existente y tiene que configurar los entornos de prueba y producción.

Pasos de migración, 1.0.1 a 2.0

1. Instale la nueva instalación de Analytic Server de acuerdo a las instrucciones de “Instalación” en la página 6.
2. Al final del proceso de instalación, elija migrar una configuración. Esto iniciará automáticamente la herramienta de migración. Complete los pasos siguientes antes de volver a la herramienta de migración.
 - a. Copie el directorio `{RAÍZ_AS}\tools\configcollector` de la nueva instalación en la ubicación correspondiente de la instalación anterior.
 - b. Ejecute la herramienta de recopilador de configuración en la instalación existente ejecutando el script **configcollector** en `{RAÍZ_AS}\tools\configcollector`. Anote la ubicación del archivo (.zip) comprimido producido.
 - c. Copie el `analytic-root`. Por ejemplo, si ha instalado la versión 1.0.1 con un valor de `analytic-root` de `/user/aeuser/analytic-root` y desea migrar a una ubicación de `analytic-root` de `/user/aeuser/analytic-root_new`, debe ejecutar el script de copia siguiente.

```
hadoop distcp /user/aeuser/analytic-root /user/aeuser/analytic-root_new
```
3. Tendrá que proporcionar a la herramienta de migración la ubicación del nuevo raíz de análisis.
4. Al final de la herramienta de migración, siga los pasos de la herramienta de postconfiguración según las instrucciones de Capítulo 3, “Configuración”, en la página 9.

Nota: Si ha configurado R para utilizarlo con la instalación de Analytic Server existente, tendrá que seguir los pasos para configurarlo con la nueva instalación de Analytic Server.

Pasos de migración, 2.0 a 2.0

1. Ejecute la herramienta de recopilador de configuración en la instalación existente ejecutando el script **configcollector** en `{RAÍZ_AS}\tools\configcollector`. Anote la ubicación del archivo (.zip) comprimido producido.
2. Instale la nueva instalación de Analytic Server de acuerdo a las instrucciones de “Instalación” en la página 6.

3. Al final del proceso de instalación, elija migrar una configuración. Esto iniciará automáticamente la herramienta de migración.
4. Al final de la herramienta de migración, siga los pasos de la herramienta de postconfiguración según las instrucciones de Capítulo 3, "Configuración", en la página 9.

Nota: Si ha configurado R para utilizarlo con la instalación de Analytic Server existente, tendrá que seguir los pasos para configurarlo con la nueva instalación de Analytic Server.

Utilización de la herramienta de migración

La herramienta de migración automatiza una serie de pasos que de otra forma debería realizarse manualmente para migrar la configuración existente a una nueva instalación. La herramienta de migración debe tener acceso al entorno HDFS.

Si utilizó la versión de la GUI de IBM Installation Manager, la herramienta de migración se inicia automáticamente. Siga las pantallas de solicitud para configurar su sistema.

También puede iniciar la herramienta de migración ejecutando el script **migrationtool** en {RAÍZ_AS}\bin. De forma predeterminada, la herramienta detecta si se va a lanzar en modalidad de GUI o de consola. Puede forzar la modalidad de consola pasando el argumento **-c**.

Bienvenida

El paso Bienvenida requiere que proporcione el archivo (.zip) comprimido que contiene la configuración de Analytic Server que va a migrar.

Clúster de Hadoop

Los pasos del clúster de Hadoop configuran Analytic Server para su uso con la distribución de Hadoop.

Igual clúster de Hadoop que la instalación original

Se trata de una copia "destruktiva" en la que la nueva instalación toma el control de la propiedad del clúster de Hadoop de la instalación anterior. Esto resulta adecuado si no utiliza la instalación de Analytic Server anterior.

Clúster Hadoop distinto

La nueva instalación se conecta a un clúster de Hadoop distinto de la instalación anterior.

Las pantallas que siguen a esta opción inicial son similares a las de la ruta típica a través de la herramienta de configuración.

Hcatalog/Hive

Los pasos de Hcatalog/Hive se muestran solo si la instalación original tiene orígenes de datos de HCatalog / Hive.

- En una migración de copia al mismo clúster de Hadoop, tiene la opción de desactivar Hive.
- En una migración de copia a un clúster de Hadoop distinto, se le proporciona el conjunto completo de opciones de la ruta de acceso personalizada a través de la herramienta de configuración para configurar HCatalog / Hive.

JDBC

En una migración de 1.0.1 a 2.0, si tiene previsto utilizar los orígenes de datos de base de datos, tendrá que añadir los controladores JDBC a la nueva instalación a través de la herramienta de administración.

En una migración de 2.0 a 2.0, todos los controladores JDBC utilizados en la instalación original se migran a la nueva instalación.

Servidor de aplicaciones

Los pasos de servidor de aplicaciones dependen del tipo de la instalación nueva.

- En una instalación no de alta disponibilidad, se omiten estos pasos. Puede configurar la comunicación segura (SSL) y establecer los puertos en los que Analytic Server escucha a través de la herramienta de administración.
- En una instalación de alta disponibilidad, la herramienta de migración solicitará información sobre el entorno Liberty Network Deployment de WebSphere® Application Server.

Nota: Si la instalación original era de alta disponibilidad, estos pasos se llenan previamente con los valores de la instalación original.

Zookeeper

El paso para Zookeeper sólo es necesario como parte de la configuración de una instalación de alta disponibilidad. Especifique la lista de servidores del conjunto Zookeeper.

Nota: Si la instalación original era alta disponibilidad, estos pasos se llenan previamente con los valores de la instalación original.

Repositorio de metadatos

Los pasos para el repositorio de metadatos dependen del tipo de la nueva instalación.

- En una instalación que no sea de alta disponibilidad, tiene la opción de elegir Derby o DB2. Si elige DB2, debe estar instalada y disponible antes de ejecutar la herramienta de migración. Consulte “Requisitos previos” en la página 5 para obtener información adicional.
- Una configuración de una instalación de alta disponibilidad siempre debe utilizar DB2 como repositorio de metadatos y la instancia de DB2 debe estar instalada y disponible antes de ejecutar la herramienta de configuración.

Nota: Si la instalación original era alta disponibilidad, estos pasos se llenan previamente con los valores de la instalación original.

Analytic Server

La herramienta de migración utiliza la configuración original para determinar el principal que actúa como administrador de Analytic Server. Este sujeto principal podrá crear, suprimir y gestionar arrendatarios. Puede cambiar el principal a través de la herramienta de administración.

Resumen

El paso Resumen permite revisar los pasos antes de completar la migración.

Postmigración

El paso postmigración informa de que la migración ha sido satisfactoria. En caso de error, revise el archivo de registro en el directorio logs.

Capítulo 5. Realización de cambios en el sistema

Utilización de la herramienta de administración

La herramienta de administración automatiza una serie de pasos que de otra forma tendrían que hacerse manualmente para actualizar la configuración. La herramienta de administración debe tener acceso al entorno HDFS.

Puede iniciar la herramienta de administración ejecutando el script `admintool` en `{RAÍZ_AS}\bin`. De forma predeterminada, la herramienta detecta si se va a lanzar en modalidad de GUI o de consola. Puede forzar la modalidad de consola pasando el argumento `-c`.

Resumen

Siempre se muestra el panel Resumen antes de realizar cualquier cambio permanente en el sistema. Cuando completa un conjunto de pasos bajo un área Administrativa, el botón final le lleva al panel Resumen. Desde aquí, puede seleccionar otra área Administrativa o pulsar Aplicar para aplicar los cambios en el sistema.

Clúster de Hadoop

Los pasos del clúster de Hadoop vuelven a configurar Analytic Server para su uso con la distribución de Hadoop.

Nota: La herramienta de administración permitirá actualizar la versión de la distribución de Hadoop, si el proveedor de Hadoop admite una actualización de versión en su lugar. Si va a cambiar de proveedor de Hadoop, puede migrar los valores de la instalación de Analytic Server existente a una nueva instalación configurada para utilizar la nueva distribución de Hadoop. Para obtener más información, consulte Capítulo 4, “Migración”, en la página 19.

Renovación simple

Esto resulta útil si la versión de distribución de Hadoop ha cambiado y simplemente necesita renovar los archivos .jar utilizados por Analytic Server.

Renovación guiada

Esto resulta de utilidad si otras propiedades del clúster de Hadoop han cambiado. Las pantallas que siguen la renovación guiada son similares a las de la herramienta de configuración.

Hcatalog/Hive

Los pasos de Hcatalog/Hive permiten volver a configurar Analytic Server para utilizarlo con la instalación de Hive o habilitan la instalación de Analytic Server existente con una nueva instalación de Hive.

Si tiene una instalación de Hive existente, hay dos opciones para actualizar la configuración.

Renovación simple

Esto resulta útil si la versión de Hive ha cambiado y simplemente necesita renovar los archivos .jar utilizados por Analytic Server.

Renovación guiada

Esto resulta de utilidad si otras propiedades de la instalación de Hive han cambiado. Las pantallas que siguen la renovación guiada son similares a las de la herramienta de configuración.

JDBC

Copie los controladores para las bases de datos que está utilizando en la ubicación especificada y siga los pasos de administración.

Servidor de aplicaciones

Los pasos de servidor de aplicaciones dependen del tipo de instalación.

- En una instalación que no sea de alta disponibilidad, puede configurar la comunicación segura (SSL/TLS) y establecer los puertos en los que Analytic Server está a la escucha.
- En una instalación de alta disponibilidad, la herramienta de administración permite darse cuenta de los cambios en el entorno Liberty Network Deployment de WebSphere® Application Server.

Nota: Para añadir o eliminar un servidor en cuarentena, tiene que editar la propiedad de configuración **ae.cluster.quarantine.server.name**. Para obtener más información, consulte “Edición manual de las propiedades de configuración”.

Zookeeper

La herramienta de administración permite modificar la lista de servidores en el conjunto Zookeeper.

Repositorio de metadatos

El repositorio de metadatos permite modificar la instancia de DB2. Si cambia el esquema, el nuevo esquema que especifique debe ya existir en la base de datos.

Si ha configurado Analytic Server para utilizar Derby como el repositorio de metadatos y desea cambiar a DB2, utilice la herramienta de migración.

Analytic Server

Los pasos de Analytic Server permiten cambiar el principal que sirve como el administrador de Analytic Server. Este sujeto principal podrá crear, suprimir y gestionar arrendatarios.

Postadministración

El paso postadministración informa de que los cambios de configuración han sido satisfactorios. En caso de error, revise el archivo de registro en el directorio logs.

Edición manual de las propiedades de configuración

Hay algunas propiedades de configuración que se pueden cambiar editando manualmente las propiedades de configuración.

config.properties

El archivo `config.properties` se encuentra en el directorio `{AS_SERVER_ROOT}/configuration`. Hay dos modos de editar este archivo.

- La herramienta de administración, cuando se inicia en modo de GUI, permite realizar cambios en este archivo.
- De lo contrario, puede editar manualmente el archivo.

En una instalación de alta disponibilidad, puede enviar los cambios a `config.properties` a través de la herramienta de administración. En una instalación no de alta disponibilidad, tendrá que reiniciar Analytic Server.

join.small.data.size

La cantidad máximo de datos (en bytes) que Analytic Server intentará unir en un algoritmo de lado de correlación.

mapred.child.java.opts

Controla los tamaños de almacenamiento dinámico de JVM para las tareas de correlación y reducción que se ejecutan en Hadoop. Establezca este valor en un tamaño tan grande como puedan gestionarlo los nodos del clúster.

mapred.job.reuse.jvm.num.tasks

Reutilice las JVM para las tareas de correlación y reducción. Cuando la sobrecarga para iniciar una JVM es elevada, aumente el valor de esta propiedad. Por ejemplo, si este valor se establece en 1, cada tarea de correlación y de reducción se ejecuta en una nueva JVM. Si el valor se establece en 5, se crea una nueva JVM para las 5 tareas.

mmr.threads

El número de hebras que se van a utilizar para los trabajos de reducción de correlación en memoria (M3R).

Nota: Las propiedades con el prefijo "mmr" controlan la reducción de correlación en memoria y son importantes para cualquier instalación que hace un uso intenso de archivos pequeños. Puede que tenga que experimentar con estos valores para optimizar el sistema.

mmr.upper.bound.threshold

Se trata de la cantidad máxima de datos que M3R procesará. Hadoop procesará cantidad más grandes de datos.

mmr.synthetic.block.size

El lado de bloque predeterminado si el archivo se encuentra en el sistema de archivos local.

mmr.diskspill.threshold

Los datos de M3R intermedios se graban en disco cuando supera su umbral.

mmr.diskspill.output.buffer

El almacenamiento intermedio de salida del correlacionador. Debe ser menor que **mmr.diskspill.threshold**.

mmr.taskparallel.targets.threshold

Si la proporción de destinos/núcleos es inferior a este umbral, M3R procesa los trabajos.

node.max.jobs

En una instalación de alta disponibilidad, el número máximo de trabajos de Analytic Server que un miembro de clúster puede ejecutar en un momento determinado.

ae.cluster.quarantine.server.name

En una instalación de alta disponibilidad, el nombre del servidor designado como el servidor de cuarentena.

jvm.options

El archivo `jvm.options` se encuentra en el directorio `{AS_SERVER_ROOT}`. Al añadir o modificar las líneas siguientes en `jvm.options` se establece el tamaño de almacenamiento dinámico para los trabajos que se ejecutan en el servidor que aloja Analytic Server; es decir, no Hadoop. Esto es importante si se ejecutan trabajos (M3R) pequeños y es posible que tenga que experimentar con estos valores para optimizar el sistema.

```
-Xms512M  
-Xmx2048M
```

En una instalación no de alta disponibilidad, tendrá que reiniciar Analytic Server. En una instalación de alta disponibilidad, no hay un modo sencillo de enviar los cambios en `jvm.options` al colectivo; tiene que actualizar manualmente las `jvm.options` en cada miembro de clúster.

Despliegue en nuevos miembros del clúster de alta disponibilidad

Una vez que se una a un nuevo miembro servidor del clúster de Liberty, puede desplegar Analytic Server a los nuevos miembros ejecutando el mandato siguiente.

```
{RAÍZ_AS}/bin/deployRemote.sh -addClusterMembers true
```

Sin la opción de **addClusterMembers**, Analytic Server vuelve a desplegarse en cada miembro del clúster de alta disponibilidad.

Capítulo 6. Desinstalación

1. Ejecute IBM Installation Manager (GUI o línea de mandatos), seleccione la opción de desinstalación de IBM SPSS Analytic Server y siga las solicitudes. IBM Installation Manager también se puede ejecutar en modo silencioso. Para obtener más información, consulte, consulte la documentación de IBM Installation Manager: http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

Nota: Esta operación no elimina ningún dato de la carpeta de Analytic Server en HDFS. Para eliminar la carpeta de inicio de Analytic Server, elimine el directorio HDFS donde están almacenados los artefactos de usuario (entradas y salidas) de Analytic Server. Es el directorio especificado en `root.folder.path` en `{RAÍZ_SERVIDOR_AS}/configuration/config.properties`.

Capítulo 7. Resolución de problemas

En este apartado se describen algunos problemas comunes de instalación y configuración y cómo corregirlos.

Clústeres de alta disponibilidad

java.net.SocketTimeoutException: se ha agotado el tiempo de espera de lectura

Cambie la variable de entorno de tiempo de espera de Liberty ND como se indica a continuación:
export LIBERTYND_READ_TIMEOUT=<milisegundos>

Donde <milisegundos> es el número de segundos para utilizar el tiempo de espera de lectura de JMX.

java.io.IOException: CWWKX7202E: El valor de tiempo de espera de 60 (segundos) para el mandato ./server start ha caducado

Añada lo siguiente al archivo server.xml de Controller Server

```
<!-- Aumente el tiempo de espera de inicio y parada de servidor para adaptar el hardware lento -->  
<serverCommands startServerTimeout="120" stopServerTimeout="120"/>
```

java.lang.OutOfMemoryError: espacio de almacenamiento dinámico de Java

Añada las líneas siguientes a jvm.options en cada miembro del clúster de alta disponibilidad.

```
-Xms512M  
-Xmx2048M
```

Los datos de transacciones de Zookeeper dejan de ser gestionables

Establezca el parámetro **autopurge.purgeInterval** de zoo.cfg en 1 para habilitar las depuraciones automáticas del registro de transacciones de Zookeeper.

El servicio de clúster de análisis ha perdido contacto con Zookeeper

Revise y modifique los parámetros **tickTime**, **initLimit** y **syncLimit** de zoo.cfg. Por ejemplo:

```
# El número de milisegundos de cada marca  
tickTime=2000  
# El número de marcas que la  
# fase de sincronización inicial puede aceptar  
initLimit=30  
# El número de marcas que pueden pasar entre  
# el envío de una solicitud y la obtención de un acuse de recibo  
syncLimit=15
```

Consulte la documentación de Zookeeper para obtener detalles: <https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html>

Los trabajos de Analytic Server no se reanudan

Hay dos situaciones comunes en las que los trabajos de Analytic Server no se reanudan.

1. Cuando un trabajo de Analytic Server falla porque falla un miembro de clúster, el trabajo se suele reiniciar automáticamente en otro miembro de clúster. Si no se reanuda el trabajo, compruebe para asegurarse de que hay por lo menos 4 miembros de clúster en el clúster de alta disponibilidad.
2. Cuando desactiva temporalmente un miembro de clúster, todos los trabajos de Analytic Server en ese servidor se reanudan normalmente en otro miembro de clúster. Para asegurarse de que se reanudan los trabajos, establezca `-Dcom.spss.ae.remoteclient.failover.threshold=100` y utilice la modalidad remota.

Avisos

Esta información se ha desarrollado para productos y servicios que se comercializan en los EE.UU.

Es posible que IBM no ofrezca en otros países los productos, servicios o características que se describen en este documento. Póngase en contacto con el representante local de IBM, que le informará sobre los productos y servicios disponibles actualmente en su área. Las referencias a programas, productos o servicios de IBM no pretenden establecer ni implicar que sólo puedan utilizarse dichos productos, programas o servicios de IBM. En su lugar, se puede utilizar cualquier producto, programa o servicio equivalente que no infrinja ninguno de los derechos de propiedad intelectual de IBM. No obstante, es responsabilidad del usuario evaluar y verificar el funcionamiento de cualquier producto, programa o servicio que no sea de IBM.

IBM puede tener patentes o solicitudes de patente pendientes que cubran la materia descrita en este documento. El suministro de este documento no le otorga ninguna licencia sobre dichas patentes. Puede enviar consultas sobre licencias, por escrito, a:

IBM Director of Licensing
IBM Corporation
North Castle Drive
Armonk, NY 10504-1785
U.S.A.

Si tiene consultas sobre licencias relacionadas con información DBCS (de doble byte), póngase en contacto con el Departamento de propiedad intelectual de IBM en su país o envíelas, por escrito, a:

Intellectual Property Licensing
Legal and Intellectual Property Law
IBM Japan Ltd.
1623-14, Shimotsuruma, Yamato-shi
Kanagawa 242-8502 Japan

El párrafo siguiente no se aplica al Reino Unido ni a ningún otro país donde estas disposiciones sean incompatibles con la legislación local: INTERNATIONAL BUSINESS MACHINES CORPORATION PROPORCIONA ESTA PUBLICACIÓN TAL CUAL, SIN NINGÚN TIPO DE GARANTÍA, EXPLÍCITAS NI IMPLÍCITAS, INCLUYENDO PERO NO LIMITÁNDOSE A ELLAS, LAS GARANTÍAS IMPLÍCITAS DE NO VULNERACIÓN, COMERCIALIZACIÓN O IDONEIDAD PARA UN PROPÓSITO DETERMINADO. Algunas legislaciones no contemplan la declaración de limitación de responsabilidad, ni implícita ni explícita, en determinadas transacciones, por lo que cabe la posibilidad de que esta declaración no sea aplicable en su caso.

Es posible que esta información contenga imprecisiones técnicas o errores tipográficos. Periódicamente se realizan cambios en la información que aquí se presenta; estos cambios se incorporarán en las nuevas ediciones de la publicación. IBM puede realizar en cualquier momento mejoras o cambios en los productos o programas descritos en esta publicación sin previo aviso.

Cualquier referencia en esta información a sitios Web que no sean de IBM se proporciona sólo para su comodidad y no constituyen una recomendación de dichos sitios Web. Los materiales de estos sitios Web no forman parte de los materiales destinados a este producto de IBM, y el usuario será responsable del uso que se haga de estos sitios Web.

IBM puede utilizar o distribuir la información que se le proporcione del modo que considere adecuado sin incurrir por ello en ninguna obligación con el remitente.

Los poseedores de licencias de este programa que deseen obtener información sobre éste a efectos de permitir: (i) el intercambio de información entre programas creados de forma independiente y otros programas (incluido éste) y (ii) el uso mutuo de la información intercambiada, deben ponerse en contacto con:

IBM Software Group
ATTN: Licensing
200 W. Madison St.
Chicago, IL; 60606
U.S.A.

Dicha información puede estar disponible, sujeta a los términos y condiciones correspondientes, incluyendo, en algunos casos, el pago de una tarifa.

El programa bajo licencia que se describe en este documento y todo el material bajo licencia disponible los proporciona IBM bajo los términos de las Condiciones Generales de IBM, Acuerdo Internacional de Programas Bajo Licencia de IBM o cualquier acuerdo equivalente entre las partes.

Cualquier dato que se encuentre en este documento se ha determinado en un ambiente controlado. Por lo tanto, los resultados obtenidos en otros entornos operativos pueden variar considerablemente. Algunas mediciones podrían haberse realizado en sistemas en desarrollo y, por lo tanto, no existe ningún tipo de garantía de que dichas mediciones sean las mismas en los sistemas con disponibilidad general. Además, algunas medidas se podrían haber estimado en extrapolación. Los resultados reales pueden variar. Los usuarios de este documento deben verificar los datos aplicables a su entorno específico.

La información respecto a productos que no son de IBM se obtuvo de los proveedores de estos productos, sus anuncios publicados u otras fuentes disponibles de forma pública. IBM no ha probado dichos productos y no puede confirmar la precisión en cuanto al rendimiento, la compatibilidad u otras características relacionadas con productos que no son de IBM. Las consultas sobre las prestaciones de productos que no sean de IBM se deben dirigir a los proveedores de esos productos.

Todas las declaraciones relativas a la dirección o intenciones futuras de IBM pueden cambiar o ser retiradas sin aviso, y representan sólo propósitos y objetivos.

Todos los precios de IBM que se muestran son precios actuales recomendados por IBM de venta al público y están sujetos a cambios sin notificación previa. Los precios en los distribuidores pueden variar.

Esta información es sólo para fines de planificación. Dicha información está sujeta a cambios antes de que los productos descritos estén disponibles.

Esta información contiene ejemplos de datos e informes utilizados en operaciones empresariales diarias. Para ilustrarlas lo mejor posible, los ejemplos contienen nombres de personas, compañías, marcas y productos. Todos estos nombres son ficticios y cualquier similitud a los nombres y direcciones que haya utilizado una empresa real es pura coincidencia.

Cada copia o cada parte de estos programas de ejemplo, o trabajos derivados, debe incluir un aviso de copyright como se indica a continuación:

Esta información contiene ejemplos de datos e informes utilizados en operaciones empresariales diarias. Para ilustrarlas lo mejor posible, los ejemplos contienen nombres de personas, compañías, marcas y productos. Todos estos nombres son ficticios y cualquier similitud a los nombres y direcciones que haya utilizado una empresa real es pura coincidencia.

Cada copia o cada parte de estos programas de ejemplo, o trabajos derivados, debe incluir un aviso de copyright como se indica a continuación:

© el nombre de su empresa) (año). Partes de este código se derivan de IBM Corp. Sample Programs.

© Copyright IBM Corp. _especifique el año o años_. Reservados todos los derechos.

Si está viendo esta información en copia software, es posible que las fotografías y las ilustraciones en color no aparezcan.

Marcas registradas

IBM, el logotipo de IBM e ibm.com son marcas registradas o marcas comerciales registradas de International Business Machines Corp., registrada en muchas jurisdicciones en todo el mundo. Otros nombres de productos y servicios podrían ser marcas registradas de IBM u otras compañías. Encontrará una lista de marcas actuales de IBM en el sitio Web "Copyright and trademark information" en www.ibm.com/legal/copytrade.shtml.

Adobe, el logotipo de Adobe, PostScript y el logotipo de PostScript son marcas registradas o marcas comerciales de Adobe Systems Incorporated en los Estados Unidos y/o en otros países.

IT Infrastructure Library es una marca registrada de la Agencia central de informática y telecomunicaciones que ahora es parte de la Cámara de Comercio.

Intel, el logotipo de Intel, Intel Inside, el logotipo de Intel Inside, Intel Centrino, el logotipo de Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium y Pentium son marcas registradas de Intel Corporation o de sus subsidiarias en EE.UU. y en otros países.

Linux es una marca registrada de Linus Torvalds en Estados Unidos y/o en otros países.

Microsoft, Windows, Windows NT y el logotipo de Windows son marcas registradas de Microsoft Corporation en los Estados Unidos, otros países o ambos.

ITIL es una marca registrada, y una marca de comunidad registrada de The Minister for the Cabinet Office, y está registrada en U.S. Patent and Trademark Office.

UNIX es una marca registrada de The Open Group en Estados Unidos y en otros países.

Java y los logotipos y marcas registradas basadas en Java son marcas registradas o marcas comerciales registradas de Oracle o de sus compañías asociadas.

Cell Broadband Engine es una marca comercial de Sony Computer Entertainment, Inc. en Estados Unidos, otros países o ambos y se utiliza bajo licencia.

Linear Tape-Open, LTO, el logotipo de LTO, Ultrium y el logotipo de Ultrium son marcas comerciales de HP, IBM Corp. y Quantum en Estados Unidos y otros países.



Impreso en España