

IBM SPSS Analytic Server
Versão 2

Guia de Instalação e Configuração

IBM

Nota

Antes de usar estas informações e o produto suportado por elas, leia as informações em “Avisos” na página 31.

Informações do produto

Esta edição aplica-se à versão 2, liberação 0, modificação do 0 do IBM SPSS Analytic Server e a todas as liberações e modificações subsequentes até que seja indicado de outra forma em novas edições.

Índice

Capítulo 1. Visão geral	1
Arquitetura	2
O que há de novo para instaladores na versão 2	2
Capítulo 2. Instalação	5
Pré-requisitos	5
Pré-requisitos de cluster de alta disponibilidade	5
Instalação	6
Capítulo 3. Configuração	9
Usando a ferramenta de configuração	9
Crie um registro do usuário	12
Ativação das Origem de Dados HCatalog	13
Ativação do Suporte Para Essentials for R	15
Configurando o IBM SPSS Modeler para Utilização com o IBM SPSS Analytic Server	16
Capítulo 4. Migração	19
Usando a ferramenta de migração	20

Capítulo 5. Fazendo mudanças em seu sistema	23
Usando a ferramenta de administração	23
Editando manualmente propriedades de configuração	24
Implementar para novos membros de cluster de alta disponibilidade	26
Capítulo 6. Desinstalando	27
Capítulo 7. Resolução de problemas	29
Avisos	31
Marcas comerciais	33

Capítulo 1. Visão geral

O IBM® SPSS Analytic Server é uma solução para analítica de Big Data que combina a tecnologia do IBM SPSS com sistemas Big Data e permite trabalhar com interfaces com o usuário do IBM SPSS familiares para resolver problemas em uma escala anteriormente inalcançável.

Porque a analítica de big data é importante

Os volumes de dados coletados pelas organizações estão crescendo exponencialmente; por exemplo, negócios financeiros e varejistas têm todas as transações dos clientes por um ano (ou dois anos, ou dez anos), provedores telco têm Call Data Records (CDR) e leituras de sensor de dispositivo e as empresas de Internet têm os resultados de crawls da web.

A analítica de big data é necessária onde existe:

- Um grande volume de dados (terabytes, petabytes, exabytes), principalmente quando há uma mistura de dados estruturados & dados não estruturados
- Dados que são alterados/acumulados rapidamente

A analítica de big data também ajuda quando:

- Um grande número (milhares) de modelos está sendo construído
- Modelos são construídos/atualizados frequentemente

Desafios

As mesmas organizações que coletam grandes volumes de dados costumam ter dificuldades ao fazerem uso real deles por vários motivos:

- A arquitetura dos produtos de analítica tradicionais não é adequada para cálculo distribuído, e
- Algoritmos estatísticos existentes não são projetados para trabalharem com Big Data (esses algoritmos esperam que os dados cheguem até a eles, mas big data são muito caros para se mover); dessa forma
- Executar o estado da analítica moderna em Big Data requer novas qualificações e profundo conhecimento de sistemas de big data. Poucos analistas têm essas qualificações.
- As soluções em memória trabalham para solucionar problemas de médio porte, mas não se adaptam muito bem a dados realmente grandes.

Solução

Servidor Analítico fornece:

- Uma arquitetura centrada em dados que alavanca sistemas de big data, como Hadoop Map/Reduce com dados em HDFS.
- Uma interface definida para incorporar novos algoritmos estatísticos projetados para acessar dados.
- Interfaces com o usuário do IBM SPSS familiares que ocultam os detalhes dos ambientes de big data para que os analistas possam se concentrar na análise dos dados.
- Uma solução escalável para problemas de qualquer porte.

Arquitetura

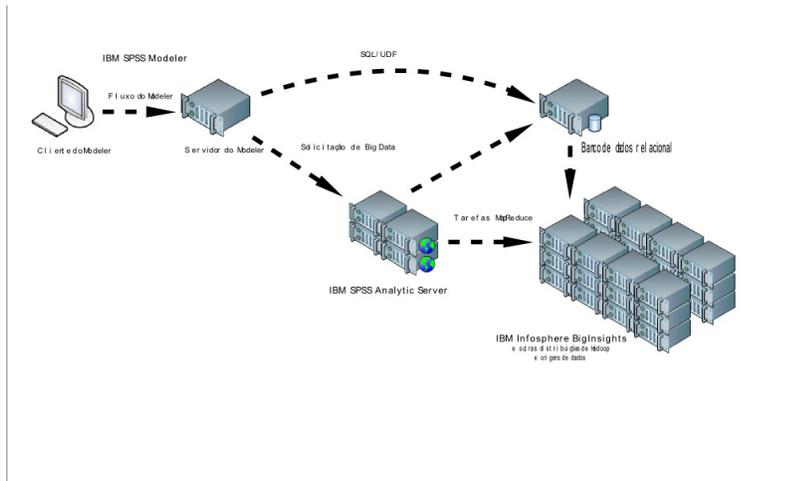


Figura 1. Arquitetura

O Servidor Analítico fica entre um aplicativo cliente e uma nuvem Hadoop. Supondo que os dados residam na nuvem, a estrutura de tópicos geral para se trabalhar com o Servidor Analítico é:

1. Defina as origens de dados do Servidor Analítico sobre os dados na nuvem.
2. Defina a análise que você deseja executar no aplicativo cliente. Para a liberação atual, o aplicativo cliente é IBM SPSS Modeler.
3. Ao executar a análise, o aplicativo cliente envia uma solicitação de execução do Servidor Analítico.
4. O Servidor Analítico orquestra a tarefa a ser executada na nuvem Hadoop e relata os resultados para o aplicativo cliente.
5. É possível usar os resultados para definir outras análises e as repetições de ciclo.

O que há de novo para instaladores na versão 2

Instalação e configuração

A instalação do produto agora usa o IBM Installation Manager para fornecer uma experiência perfeita ao instalar vários produtos IBM.

As etapas para configurar o Servidor Analítico são manipuladas em uma nova ferramenta de configuração executada depois de todas as extensões do produto serem instaladas.

Plataforma

Suporte para novos sistemas operacionais e distribuições do Hadoop.

Sistemas operacionais

O Servidor Analítico agora é executado no Microsoft Windows.

Distribuições do Hadoop

O Servidor Analítico pode ser configurado para executar o Hortonworks Data Platform no Microsoft Windows.

Adaptadores HCatalog

As origens de dados podem ser definidas para Apache Cassandra, MongoDB e Oracle NoSQL. Há mais etapas de configuração para ativar essas origens de dados.

Clusters de alta disponibilidade

Servidor Analítico pode ser instalado e configurado em um cluster de alta disponibilidade usando o IBM WebSphere Liberty Network Deployment.

Capítulo 2. Instalação

Pré-requisitos

Antes de instalar o Servidor Analítico, revise as informações a seguir.

Requisitos do sistema

Para obter informações mais atualizadas sobre os requisitos do sistema, use os relatórios de requisitos do sistema Detalhados no site de Suporte Técnico da IBM: <http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html>. Nesta página:

1. Digite SPSS Servidor Analítico como o nome do produto e clique em **Procurar**.
2. Selecione a versão e o escopo do relatório desejados e, em seguida, clique em **Enviar**.

Privilégios administrativos

O usuário que instala o Servidor Analítico deve ter acesso de gravação ao Hadoop Distributed File System.

IBM Installation Manager

Uma versão atual do IBM Installation Manager deve ser instalada e configurada para utilizar um repositório que contém os arquivos de instalação do Servidor Analítico. Se você tiver uma versão antiga do IBM Installation Manager, será necessário atualizá-la para uma versão suportada como parte da instalação.

Se o IBM Installation Manager não estiver presente no sistema, faça o download e instale o IBM Installation Manager a partir do site de suporte da IBM Corp.: <http://www.ibm.com/support>.

Importante: No Microsoft Windows, instale o IBM Installation Manager de uma conta do Administrador. É possível fazer uma instalação de grupo como outro usuário com: `./groupinstc -acceptLicense`. Consulte a documentação do IBM Installation Manager para obter detalhes. http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html

Hive/HCatalog

Se você planeja usar origens de dados NoSQL, configure Hive and HCatalog para acesso remoto. Além disso, assegure que o `hive-site.xml` contenha a propriedade `hive.metastore.uris` no formato `thrift://<host_name>:<port>` que aponta para o servidor Thrift Hive Metastore ativo. Consulte a documentação da distribuição do Hadoop para obter detalhes.

Repositório de metadados

Por padrão, o Servidor Analítico instala e utiliza um banco de dados Apache Derby. Como alternativa, será possível configurar o Servidor Analítico para utilizar uma instalação existente do DB2. O banco de dados deve ter uma codificação de UTF-8. IBM DB2 10.5 Workgroup Server Edition está disponível na mídia do Servidor Analítico. Para obter mais informações: http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html.

Cluster de alta disponibilidade

Consulte “Pré-requisitos de cluster de alta disponibilidade” para obter os pré-requisitos para a instalação do Servidor Analítico em um cluster de alta disponibilidade.

Pré-requisitos de cluster de alta disponibilidade

Há pré-requisitos extras para a instalação do Servidor Analítico em um cluster de alta disponibilidade.

Configuração de Sistemas e DNS

Os sistemas que são utilizados no ambiente precisam ser capazes de se comunicar uns com os outros. Apenas os sistemas que pertencem a um domínio deveriam ter o sufixo do domínio. Por exemplo, somente sistemas registrados com o domínio "IBM.COM" deve ter o sufixo "ibm.com" para o Nome Completo do Domínio (FQDN).

Se um sistema não estiver no domínio IBM.COM terá um sufixo FQDN do "ibm.com", em seguida, o FQDN para o sistema deverá ser inserido no DNS com os sistemas que têm o servidor DNS na configuração TCP. Caso contrário, todos os FQDNs de todos os sistemas que estão envolvidos deverão estar nos arquivos hosts de todos os sistemas envolvidos.

Repositório de metadados

O Servidor Analítico deve utilizar o DB2 quando instalado em um cluster de alta disponibilidade.

Zookeeper 3.4.5

Um mínimo de três instâncias Zookeeper são necessárias. Consulte a instalação Zookeeper e a documentação de configuração : http://zookeeper.apache.org/doc/r3.4.5/zookeeperStarted.html#sc_RunningReplicatedZooKeeper.

Nota: Essas etapas devem ser executadas em cada sistema onde o Zookeeper está instalado.

Liberty Network Deployment 8.5.5.4

O ambiente de alta disponibilidade consiste em um coletivo Liberty com um controlador, duas réplicas e, pelo menos, dois membros de cluster. Consulte a documentação do WebSphere Application Server Network Deployment para obter detalhes sobre a configuração de um coletivo Liberty: http://www-01.ibm.com/support/knowledgecenter/SSAW57/mapfiles/product_welcome_wasnd.html.

Balancedador de carga

Seu cluster de alta disponibilidade deve ter um balanceador de carga que suporta afinidade de sessão, conhecida também às vezes como sessões persistentes. O Servidor Analítico identifica as sessões com o cookie "request-token". Isso identifica uma sessão para a duração de um login de usuário para uso em afinidade de sessão controlada pelo aplicativo. Consulte a documentação do seu balanceador de carga particular para obter os detalhes de como ela suporta afinidade de sessão.

Nota: O nó líder não deve ser incluído no balanceador de carga para o cluster. A maneira de se identificar o líder é que ele é o primeiro nó iniciado.

Visual C++ 2008 (somente Windows)

Se você tiver um membro de cluster executando o Servidor Analítico no Windows, será necessário instalar o Visual C 2008 Runtime antes de instalar o Servidor Analítico.

Instalação

Os arquivos de aplicativos do Servidor Analítico são instalados no sistema host com o IBM Installation Manager. Os arquivos de instalação podem ser localizados na mídia física ou transferidos por download do IBM Passport Advantage.

1. Inicie o IBM Installation Manager:

Modo da GUI

```
{IBM Installation Manager installation directory}/eclipse/IBMIM
```

Modo de linha de comandos

```
{IBM Installation Manager installation directory}/eclipse/tools/imcl -c
```

Nota: O usuário que executa o IBM Installation Manager é o usuário que executa o processo Servidor Analítico. Este usuário requer acesso ao ambiente do Hadoop.

2. Se o repositório de instalação não estiver configurado, especifique o caminho do repositório. Por exemplo, especifique o disco de instalação montado ou um local no sistema de arquivos do host, a rede ou um endereço HTTP.
3. No menu principal, selecione **Instalar**.
4. Selecione o pacote para instalação. A opção disponível depende da versão do Servidor Analítico adquirida.

IBM SPSS Analytic Server

Este pacote é adequado para a maioria dos sistemas.

IBM SPSS Analytic Server Alta Disponibilidade

Esse pacote será requerido se você planejar executar o Servidor Analítico em um cluster de alta disponibilidade. Diferenças na configuração do IBM SPSS Analytic Server Alta Disponibilidade estão documentadas no local.

5. Leia o contrato de licença e aceite seus termos.
6. Especifique o grupo de pacotes e o diretório de instalação.
 - Um novo grupo de pacotes é necessário para instalação do IBM SPSS Analytic Server.
 - Especifique o diretório de instalação para recursos compartilhados. É possível especificar o diretório de recursos compartilhados somente a primeira vez que um pacote for instalado.
7. Selecione os recursos a serem instalados.
8. Revise as informações de resumo e continue com a instalação. Os arquivos do aplicativo estão instalados no diretório especificado.
9. Uma configuração adicional é necessária. Para obter mais informações sobre uma configuração nova, consulte o tópico: “Usando a ferramenta de configuração” na página 9. Para obter informações sobre como migrar uma configuração existente para uma nova instalação, consulte o tópico: Capítulo 4, “Migração”, na página 19. Se estiver executando o IBM Installation Manager no modo GUI, a opção para iniciar a ferramenta de configuração ou migração será apresentada no último passo.

Se ocorrerem problemas durante a instalação, será possível solucioná-los usando os logs do IBM Installation Manager. Os arquivos de log podem ser acessadas a partir do menu principal do IBM Installation Manager.

Para automatizar a instalação, o IBM Installation Manager pode ser executado no modo silencioso. Para obter mais informações, consulte a documentação do IBM Installation Manager: http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

Capítulo 3. Configuração

Após a instalação, deve-se configurar o Servidor Analítico antes de poder utilizar o produto.

Nota: As convenções a seguir são utilizadas para os caminhos de arquivo do Servidor Analítico.

- {AS_ROOT} refere-se à localização em que Servidor Analítico é implementado; por exemplo, /opt/IBM/SPSS/AnalyticServer/{version}.
- {AS_SERVER_ROOT} refere-se ao local dos arquivos de configuração, de log e do servidor; por exemplo, /opt/IBM/SPSS/AnalyticServer/{version}/ae_wlpserver/usr/servers/aeserver.
- {AS_HOME} refere-se à localização no HDFS usada pelo Servidor Analítico como uma pasta-raiz.

Etapas de configuração

1. [Opcional] Siga os passos para ativar as origens de dados HCatalog se desejar executar o Servidor Analítico em origens de dados NoSQL ou XML.

Nota: É possível executar uma configuração simples do Servidor Analítico agora e ativar origens de dados HCatalog depois usando a ferramenta de administração.

2. Execute a ferramenta de configuração A ferramenta automatiza inúmeras outras etapas manuais. Se você usou a versão da GUI do IBM Installation Manager, a ferramenta de configuração será ativada automaticamente, caso contrário será necessário executar o script **configtool** no {AS_ROOT}/bin.
3. Crie um registro de usuários que podem ser incluídos no sistema do Servidor Analítico.
4. Configure o IBM SPSS Modeler para uso com o Servidor Analítico.
5. [Opcional] Ative o Support for Para Essentials for R se quiser que o Servidor Analítico execute o código R.
6. Certifique-se de que o Servidor Analítico esteja em execução. É possível iniciar o processo executando o script **start** em {AS_ROOT}/bin.
7. Abra um navegador da web e insira o endereço <http://localhost:8080/analyticserver/admin/ibm>. Essa URL abre o diálogo de login para o console do Servidor Analítico. Efetue login como o administrador do Servidor Analítico definido na ferramenta de configuração.

Usando a ferramenta de configuração

A ferramenta de configuração automatiza inúmeras outras etapas manuais. A ferramenta de configuração deve ter acesso ao ambiente do HDFS.

Se você usou a versão da GUI do IBM Installation Manager, a ferramenta de configuração será ativada automaticamente. Siga os prompts para configurar seu sistema.

Também é possível ativar a ferramenta de configuração executando o script **configtool** no {AS_ROOT}\bin. Por padrão, a ferramenta detecta se a ativação está no modo de GUI ou console. É possível forçar o modo do console, transmitindo a ele o argumento **-c**, ou o modo silencioso, transmitindo a ele o caminho para um arquivo que especifica a configuração desejada. Consulte o arquivo **configtool.response** no {AS_ROOT}\bin para obter uma amostra.

Boas-vindas

A etapa de Boas Vindas determina qual caminho tomar através da Ferramenta de Configuração.

Típico O caminho Típico automatiza muitas decisões e simplifica suposições sobre seu sistema. Selecione

esta opção se desejar configurar rapidamente um ambiente não produção. Sempre é possível alterar sua configuração na ferramenta de Administração ou migrar sua configuração para outro ambiente posteriormente.

Em uma configuração Típica, a ferramenta solicitará a localização do nó de nome Hadoop e progressivamente pedirá mais informações quando necessário.

Em uma configuração Típica de uma instalação de Alta Disponibilidade, a ferramenta de configuração solicitará informações sobre o seu ambiente do WebSphere® Application Server Liberty Network Deployment e seu conjunto Zookeeper.

Customizado

Isso permite que você execute tarefas de configuração mais avançadas. Essa opção é requerida se você desejar configurar NoSQL ou bancos de dados relacionais para uso com o Servidor Analítico agora.

As seções abaixo fornecem mais detalhes sobre o caminho customizado.

Cluster do Hadoop

As etapas do cluster do Hadoop configuram o Servidor Analítico para ser usado com sua distribuição de Hadoop. A ferramenta de configuração pode descobrir automaticamente sua distribuição, ou é possível ainda selecionar manualmente os arquivos necessários.

Coletar arquivos automaticamente a partir de sua distribuição de Hadoop

A ferramenta tentará detectar a distribuição de Hadoop e progressivamente pedirá mais informações quando necessário.

1. Será sempre necessário fornecer o nome do host do cluster do Hadoop.
2. Se o cluster não estiver configurado com SSH e com a autenticação de chave pública para o usuário que está executando a ferramenta de configuração, será necessário fornecer credenciais válidas de modo que a ferramenta de configuração possa reunir arquivos do cluster do Hadoop.

Restrição: Somente a criptografia RSA SSH é suportada.

3. Se o cluster do Hadoop foi instalado em um local não padrão, pode ser necessário fornecer o local de instalação.

Coletar arquivos manualmente de sua distribuição Hadoop

Identifique a distribuição Hadoop que está sendo usada. A ferramenta de configuração identificará os arquivos que precisam ser reunidos no cluster do Hadoop. Copie esses arquivos para o local especificado e será possível continuar as etapas de configuração.

Nota: O arquivo JAR `hadoop-lzo-*.jar` só é necessário quando o valor `"com.hadoop.compression.lzo.LzoCodec"` é especificado para o parâmetro `io.compression.codecs` no arquivo `core-site.xml`.

Hcatalog/Hive

As etapas Hcatalog/Hive são mostradas apenas como parte de uma configuração Customizada.

Conectar automaticamente e detectar sua configuração Hive / HCatalog

A ferramenta tentará detectar sua versão do HCatalog / Hive e coletar os arquivos necessários para configurar o Servidor Analítico.

Selecionar manualmente sua configuração do Hive / HCatalog

Identifique os bancos de dados NoSQL que estão sendo usados. A ferramenta de configuração identificará os arquivos que precisam ser reunidos. Copie esses arquivos para o local especificado e será possível continuar as etapas de configuração.

Ignorar configuração do HCatalog / Hive

É possível configurar o HCatalog / Hive posteriormente.

JDBC

Os passos JDBC só são mostrados como parte de uma configuração Customizada. Copie os drivers para os bancos de dados que estão sendo usados para o local especificado e continue as etapas de configuração.

Tabela 1. Bancos de Dados Suportados

Database	Versões suportadas	Jars de driver JDBC	Vendor
DB2 para Linux, UNIX e 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, ora118n.jar	Oracle
Servidor SQL	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

Application Server

As etapas do Application Server são dependentes do tipo de instalação.

- Em uma instalação que não seja de Alta Disponibilidade, é possível configurar comunicação segura (SSL/TLS) e configurar as portas nas quais o Servidor Analítico atende.
- Em uma instalação de Alta Disponibilidade, a ferramenta de configuração solicitará informações sobre o seu ambiente do WebSphere® Application Server Liberty Network Deployment.

Zookeeper

A etapa zookeeper será necessária somente como parte da configuração para uma instalação de Alta Disponibilidade. Forneça a lista de servidores em sua combinação Zookeeper.

Repositório de metadados

As etapas do repositório de metadados são dependentes do tipo de instalação.

- Em uma instalação não Alta Disponibilidade, você tem a escolha Derby ou DB2. Se você escolher o DB2, ele deverá ser instalado e estar disponível antes da execução da ferramenta de configuração. Consulte “Pré-requisitos” na página 5 para obter informações adicionais.
- Uma configuração de uma instalação de Alta Disponibilidade deve sempre usar o DB2 como o repositório de metadados e a instância do DB2 deve ser instalada e estar disponível antes da execução da ferramenta de configuração..

Servidor analítico

Identifique o principal que atende como o administrador do Servidor Analítico. Esse principal será capaz de criar, excluir e gerenciar locatários.

Resumo

A etapa de Resumo permite revisar as etapas de configuração antes de concluir a configuração.

Pós-configuração

A etapa de Pós configuração fornecerá informações que a configuração foi bem sucedida. Caso ocorra uma falha, revise o arquivo de log no diretório logs.

Crie um registro do usuário

O locatário padrão criado durante a instalação e a configuração usa o perfil do WebSphere Liberty para autenticar usuários, que, por sua vez, podem fazer referência a um servidor LDAP externo ou definir um registro básico de usuários e grupos.

Nota: Depois de editar o arquivo `{AS_SERVER_ROOT}/server.xml` para configurar o WebSphere Liberty Profile para autenticar usuários, inclua esses usuários ao sistema do Servidor Analítico. Consulte o *IBM SPSS Analytic Server Guia do Administrador* para obter detalhes sobre o gerenciamento de locatário.

LDAP

O registro LDAP permite autenticar usuários com um servidor LDAP externo, como o Active Directory ou OpenLDAP. Veja aqui um exemplo 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>
```

Para obter mais exemplos de configurações, consulte a pasta de modelos `{AS_ROOT}/ae_wlpserver/templates/config`.

Básico

O Registro Básico permite definir um banco de dados de usuários e grupos dentro do arquivo `{AS_SERVER_ROOT}/server.xml`. As senhas podem ser codificadas para ofuscar seus valores com a ferramenta `securityUtility`, que está em `{AS_ROOT}/ae_wlpserver/bin`.

O Registro Básico é útil em um ambiente de simulação, mas não é recomendado para um ambiente de produção.

```
<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>

```

Ativação das Origem de Dados HCatalog

O Servidor Analítico fornece suporte para dados XML e bancos de dados NoSQL por meio do Hive/HCatalog.

A ferramenta de configuração prepara sua instalação do Servidor Analítico para uso com Hive/HCatalog, mas existem alguns passos manuais. É necessário copiar alguns arquivos JAR para o diretório do Hive {HIVE_HOME}/auxlib/, pois o Servidor Analítico implementa algumas interfaces requeridas pelo HCatalog. O Servidor Analítico não pode usar manipuladores de armazenamento para o Hive diretamente no HCatalog porque o HCatalog tem uma API um pouco diferente.

As seções a seguir detalham os arquivos necessários, com base em quais origens de dados você deseja utilizar.

Importante: Esses arquivos devem ser copiados no diretório {HIVE_HOME}/auxlib/ antes da execução da ferramenta de configuração ou ferramenta de administração para que o Servidor Analítico faça uso deles.

Importante: Hortonworks Data Platform on Windows não reconhece o diretório {HIVE_HOME}/auxlib/. Para ativar o HCatalog, é possível:

1. Copiar os arquivos necessários, conforme descrito abaixo, no diretório {HIVE_HOME}/lib/.
2. Remover o diretório {HIVE_HOME}/auxlib/.
3. Atualizar os arquivos hive-site.xml e hive-config.cmd para remover quaisquer referências ao diretório {HIVE_HOME}/auxlib/.
4. Reiniciar o cluster do Hortonworks Data Platform para que as mudanças entrem em vigor.

Consulte <http://hortonworks.com/community/forums/topic/hive-auxlib-issues/> para obter informações adicionais.

Todas as instalações do Hive/HCatalog

1. Os arquivos a seguir podem ser localizados no archive hive_auxlib no diretório {AS_ROOT}/add-ons.

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

2. Copie a versão mais recente do hivexmlserde-{version}.jar no Maven Central Repository em <http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>.
3. [Opcional] Se planeja usar os tipos complexos do Hive (mapas, matrizes, estruturas), copie estes arquivos JAR. Eles podem ser localizados na instalação do HCatalog. Antes do Hive 0.11.0, isto está em {HCATALOG_HOME}/share/hcatalog; iniciando com Hive 0.11.0, isto está em {HIVE_HOME}/hcatalog/share/hcatalog.

```

hcatalog-core-{version}.jar (for hadoop 1.2)
hive-hcatalog-{version}.jar (for hadoop 2.x)
jackson-core-asl-1.8.8.jar
jackson-mapper-asl-1.8.8.jar

```

4. Se você estiver usando o Hive versão 0.14.0, faça o download do seguinte arquivo JAR.
<https://github.com/dvasilen/Hive-HCatalog-Compatibility/raw/master/release/hive-hcatalog-compatibility-0.14.0.jar>

Dados XML

1. [Opcional] Se você pretende trabalhar com arquivos XML compactados, copie os seguintes arquivos JAR localizados no archive hive_auxlib no diretório {AS_ROOT}/add-ons.
com.ibm.spss.com.ibm.biginsights.compression_2.1.jar (somente distribuição BigInsights Hadoop)
com.ibm.spss.hivexmlserdecomp-2.0.0.0.jar

Apache Accumulo

1. Os arquivos a seguir podem ser localizados na instalação do HCatalog. Antes do Hive 0.11.0, isto está em {HCATALOG_HOME}/share/hcatalog; iniciando com Hive 0.11.0, isto está em {HIVE_HOME}/hcatalog/share/hcatalog.
commons-io-{version}.jar
2. O manipulador de armazenamento Apache Accumulo pode ser localizado em <https://github.com/dvasilen/accumulo-hive-storage-manager/tree/master>. Esta é a ramificação principal; a partir daqui, deve-se mudar para a ramificação para a sua versão do Accumulo e do Hadoop. O arquivo JAR pode ser localizado no diretório /release e ele tem o nome nesse formato:
accumulo-hive-storage-handler_{hadoopVersion}-{accumuloVersion}.jar
3. Os arquivos a seguir podem ser localizados na instalação do 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. Os arquivos a seguir podem ser localizados na liberação do Apache Cassandra 1.2.9, em <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. O manipulador de armazenamento Apache Cassandra pode ser localizado em <https://github.com/dvasilen/Hive-Cassandra>. Essa é a ramificação principal; aqui você deve alternar para a ramificação para sua versão do Cassandra e Hadoop. O arquivo JAR pode ser localizado no diretório /release e ele tem o nome nesse formato:
hive-
<hiveVersion>-hadoop-
<hadoopVersion>-cassandra-
<cassandraVersion>.jar

Nota: Para versões de Hadoop anteriores à 2.0, não há uma parte hadoop no nome ou ramificação do arquivo.

Nota: O Servidor Analítico pode se comunicar com Cassandra 2.0, mas seus arquivos JAR são compilados para Java 7. Portanto, é necessário copiar os arquivos JAR do Cassandra 1.2.9 porque eles são compilados para o Java 6.

HBase

Não há arquivos extras para copiar para o HBase.

MongoDB

1. Os arquivos a seguir devem ser construídos localmente para sua distribuição do Hadoop a partir da ramificação "principal" do projeto "mongodb/mongo-hadoop". Consulte <https://github.com/mongodb/mongo-hadoop> para obter mais detalhes.

```
mongo-hadoop-core-{version}.jar  
mongo-hadoop-hive-{version}.jar
```

2. O arquivo a seguir pode ser localizado em <https://github.com/downloads/mongodb/mongo-java-driver/mongo-2.10.1.jar>
mongo-2.10.1.jar

Oracle NoSQL

1. O arquivo a seguir pode ser localizado em <https://github.com/vilcek/HiveKVStorageHandler2/blob/master/dist/HiveKVStorageHandler.jar>

```
HiveKVStorageHandler.jar
```

2. O arquivo a seguir pode ser localizados na instalação do Apache Hive.
avro-*<version>.jar

Ativação do Suporte Para Essentials for R

Servidor Analítico suporta modelos R de pontuação e scripts R de execução.

Restrição: Atualmente, o suporte para R é limitado a instalações do Servidor Analítico no Linux.

Para configurar o suporte para R após uma instalação bem-sucedida do Servidor Analítico:

1. Instale o R Engine no servidor que hospeda o Servidor Analítico.

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

Importante: Instale o Mecanismo R em um local acessível a todos os usuários, pois o R é executado no no cluster do Hadoop como um usuário diferente do usuário do Servidor Analítico.

Importante: Em uma instalação de Alta Disponibilidade, o R Engine deve ser instalado no mesmo local em cada membro de cluster.

2. Instale o Essentials for R no servidor que hospeda o Servidor Analítico executando o arquivo do instalador, seguindo as instruções na tela. O Essentials for R está disponível para download (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>).
3. Implementa o mecanismo R e o Componente R para o Hadoop.
 - a. Se o Servidor Analítico e todos os nós do Hadoop tiverem a mesma versão de sistema operacional e a mesma arquitetura do processador:

- 1) Crie um archive para o Mecanismo R usando o seguinte script

```
#!/usr/bin/env bash  
repetir Criando 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: Há dois espaços entre -d\ e -f3 no script. O documento PDF dá a aparência de dois espaços, mas é um único espaço grande. Ao copiar e colar este script do PDF, insira o segundo espaço antes de executar o script.

Nota: As linhas que copiam as bibliotecas Fortran são dependentes da versão do Fortran instalado. Se, por exemplo, a versão 1.0 está instalada, essas linhas são lidas como se segue. Edite conforme a necessidade de seu sistema.

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

- 2) Copie o archive R.zip no diretório {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/linux/app_64.
- b. Se as versões dos sistemas operacionais do Servidor Analítico e os nós do Hadoop forem diferentes, instale o mecanismo R e o Essentials for R de cada nó do Hadoop. Instale-os em um diretório com o mesmo caminho e nome em que o mecanismo R está instalado no Servidor Analítico. Se o Essentials for R for instalado no nó do Hadoop, especifique o diretório no qual o R está instalado e ignore a etapa de especificação do local de ../ext_64/bin.
- c. Execute o **hdfsUpdate** script no diretório {AS_ROOT}/bin para propagar as mudanças para o HDFS.
4. Em uma instalação de Alta Disponibilidade, também se deve executar o script **deployRemote** no diretório {AS_ROOT}/bin para propagar as mudanças para as instâncias do Servidor Analítico em cada membro de cluster.
5. Deve-se também instalar o Essentials for R na máquina que hospeda o SPSS Modeler Server.

Configurando o IBM SPSS Modeler para Utilização com o IBM SPSS Analytic Server

Para ativar o SPSS Modeler para uso com o Servidor Analítico, é necessário fazer algumas atualizações para a instalação do SPSS Modeler Server.

1. Configure o SPSS Modeler Server para associá-lo a uma instalação do Servidor Analítico.
 - a. Edite o arquivo **options.cfg** no subdiretório **config** do diretório de instalação de servidor principal e inclua ou edite as linhas a seguir:

```
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 se a comunicação segura estiver configurada no Servidor Analítico; caso contrário, N.

as_host

O endereço IP do servidor que hospeda o Servidor Analítico.

as_port

A porta na qual o Servidor Analítico está atendendo (por esse padrão é 8080).

as_context_root

O contexto raiz Servidor Analítico (por padrão, esse é analyticserver).

as_tenant

O locatário da instalação do SPSS Modeler Server é um membro de (o locatário padrão é ibm).

as_prompt_for_password

Especifique N se o SPSS Modeler Server estiver configurado com o mesmo sistema de autenticação para usuários e senhas que aquele utilizado no Servidor Analítico; caso contrário, Y.

Ao executar o SPSS Modeler em modo em lote, inclua `-analytic_server_username {ASusername} -analytic_server_password {ASpassword}` como argumentos para o comando `c1emb`.

- b. Reinicie o serviço do SPSS Modeler Server.

Para se conectar a uma instalação do Servidor Analítico que tem SSL/TLS ativado, há alguns passos adicionais para a configuração de suas instalações de cliente e SPSS Modeler Server.

- a. Navegue para `http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT}` e efetue logon no console do Servidor Analítico.
- b. Faça download do arquivo de certificação do navegador e salve-o em seu sistema de arquivos.
- c. Inclua o arquivo de certificação para o JRE das instalações do SPSS Modeler Server e SPSS Modeler Client. O local para atualizar pode ser localizado no subdiretório `/jre/lib/security/cacerts` do caminho da instalação do SPSS Modeler.
 - 1) Certifique-se de que o arquivo `cacerts` não seja de somente leitura.
 - 2) Use o programa `keytool` Modeler fornecido com – isso pode ser localizado no subdiretório `/jre/bin/keytool` do caminho da instalação do SPSS Modeler.

Execute o comando a seguir

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

Observe que `<as-alias>` é um alias para o arquivo `cacerts`. É possível usar qualquer nome que você gostaria, contanto que seja exclusivo para o arquivo `cacerts`.

Portanto, um exemplo de comando seria semelhante ao seguinte.

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

- d. Reinicie seu SPSS Modeler Server e SPSS Modeler Client .

2. [opcional] Instale o IBM SPSS Modeler - Essentials for R se planeja pontuar modelos R em fluxos com origens de dados do Servidor Analítico. O IBM SPSS Modeler - Essentials for R está disponível para download (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>).

Capítulo 4. Migração

Servidor Analítico suporta inúmeros cenários nos quais você migra definições de configuração de uma instalação do Servidor Analítico existente para uma nova.

Fazer upgrade para uma nova versão do Servidor Analítico

Se você tiver uma instalação existente do Servidor Analítico 1.0.1 e tiver comprado a 2.0, é possível migrar suas definições de configuração da 1.0.1 para a instalação 2.0. Nesse cenário, a ferramenta de migração lhe guiará pelo caminho apropriado.

Restrição: Se você tiver a versão 1.0 instalada, será preciso migrar primeiro da versão 1.0 para a 1.0.1 e depois da versão 1.0.1 para a 2.0.

Restrição: Suas instalações 1.0.1 e 2.0 não podem coexistir no mesmo cluster do Hadoop. Se você configurar instalações 2.0 para usar o mesmo cluster Hadoop que a instalação 1.0.1, a instalação 1.0.1 não funcionará mais.

Criar uma cópia do Servidor Analítico existente

Se você tiver uma instalação existente do Servidor Analítico 2.0, é possível replicar definições de configuração em uma nova instalação 2.0. Isso é útil, por exemplo, quando se tem um ambiente de desenvolvimento existente e é preciso configurar ambientes de teste e produção.

Passos de migração, 1.0.1 para 2.0

1. Instale a nova instalação do Servidor Analítico de acordo com as instruções em “Instalação” na página 6.
2. No final do processo de instalação, escolha migrar uma configuração. Isso ativará automaticamente a ferramenta de migração. Conclua os passos a seguir antes de retornar para a ferramenta de migração.
 - a. Copie o diretório `{AS_ROOT}\tools\configcollector` em sua nova instalação para o local correspondente em sua instalação antiga.
 - b. Execute a ferramenta coletora de configuração em sua instalação existente executando o script **configcollector** em `{AS_ROOT}\tools\configcollector`. Anote o local do arquivo compactado (.zip) resultante.
 - c. Copie a raiz analítica. Por exemplo, se você instalou a versão 1.0.1 com um valor de raiz analítica de `/user/aeuser/analytic-root` e desejar migrar para um local de raiz analítica de `/user/aeuser/analytic-root_new`, será necessário executar o seguinte script de cópia.

```
hadoop distcp /user/aeuser/analytic-root /user/aeuser/analytic-root_new
```
3. Você precisará fornecer à ferramenta de migração o local da nova raiz analítica.
4. No final da ferramenta de migração, siga quaisquer passos da ferramenta pós-configuração de acordo com as instruções em Capítulo 3, “Configuração”, na página 9.

Nota: Se você configurou R para usar com a instalação do Servidor Analítico existente, será necessário seguir os passos para configurá-lo com a nova instalação do Servidor Analítico.

Passos de migração, 2.0 para 2.0

1. Execute a ferramenta coletora de configuração em sua instalação existente executando o script **configcollector** em `{AS_ROOT}\tools\configcollector`. Anote o local do arquivo compactado (.zip) resultante.
2. Instale a nova instalação do Servidor Analítico de acordo com as instruções em “Instalação” na página 6.
3. No final do processo de instalação, escolha migrar uma configuração. Isso ativará automaticamente a ferramenta de migração.

4. No final da ferramenta de migração, siga quaisquer passos da ferramenta pós-configuração de acordo com as instruções em Capítulo 3, "Configuração", na página 9.

Nota: Se você configurou R para usar com a instalação do Servidor Analítico existente, será necessário seguir os passos para configurá-lo com a nova instalação do Servidor Analítico.

Usando a ferramenta de migração

A ferramenta de migração automatiza inúmeras outras etapas manuais para migrar sua configuração existente para uma nova instalação. A ferramenta de migração deve ter acesso ao ambiente do HDFS.

Se você usou a versão da GUI do IBM Installation Manager, a ferramenta de migração será ativada automaticamente. Siga os prompts para configurar seu sistema.

Também é possível ativar a ferramenta de migração executando o script **migrationtool** em {AS_ROOT}\bin. Por padrão, a ferramenta detecta se a ativação está no modo de GUI ou console. É possível forçar o modo de console passando para ele o argumento **-c**.

Boas-vindas

O passo Bem-vindo requer que você forneça o arquivo compactado (.zip) contendo a configuração do Servidor Analítico sendo migrada.

Cluster do Hadoop

As etapas do cluster do Hadoop configuram o Servidor Analítico para ser usado com sua distribuição de Hadoop.

Mesmo cluster do Hadoop que a instalação original

Essa é uma cópia "destrutiva" na qual a nova instalação assume a propriedade do cluster do Hadoop da instalação antiga. Isso será apropriado se você não tiver um uso para a antiga instalação do Servidor Analítico.

Cluster do Hadoop diferente

A nova instalação é conectada a um cluster do Hadoop diferente da instalação antiga.

As telas que seguem essa opção inicial são semelhantes às de um caminho Típico por meio da ferramenta de configuração.

Hcatalog/Hive

Os passos Hcatalog/Hive serão mostrados somente se sua instalação original tiver origens HCatalog / Hive.

- Em uma migração de Cópia para o mesmo cluster do Hadoop, você tem a opção de desativar o Hive.
- Em uma migração de Cópia para um cluster do Hadoop diferente, você recebe um conjunto completo de opções do caminho Customizado à ferramenta de configuração para configurar o HCatalog / Hive.

JDBC

Em uma migração 1.0.1 para 2.0, se você planeja usar origens de dados do banco de dados, será necessário incluir drivers JDBC na nova instalação por meio da ferramenta de administração.

Em uma migração 2.0 para 2.0, quaisquer drivers JDBC usados na instalação original serão migrados para a nova instalação.

Application Server

Os passos do Application Server dependem do tipo da nova instalação.

- Em uma instalação que não seja de Alta Disponibilidade, esses passos são ignorados. É possível configurar comunicação segura (SSL) e configurar as portas nas quais o Servidor Analítico atende por meio da ferramenta de administração.
- Em uma instalação de Alta Disponibilidade, a ferramenta de migração solicitará informações sobre o seu ambiente do WebSphere® Application Server Liberty Network Deployment.

Nota: Se a instalação original era de Alta Disponibilidade, essas etapas serão pré-preenchidas com os valores dessa instalação original.

Zookeeper

A etapa zookeeper será necessária somente como parte da configuração para uma instalação de Alta Disponibilidade. Forneça a lista de servidores em sua combinação Zookeeper.

Nota: Se a instalação original era de Alta Disponibilidade, essas etapas serão pré-preenchidas com os valores dessa instalação original.

Repositório de metadados

Os passos do repositório de metadados são dependentes do tipo da nova instalação.

- Em uma instalação não Alta Disponibilidade, você tem a escolha Derby ou DB2. Se você escolher o DB2, ele deverá ser instalado e estar disponível antes da execução da ferramenta de migração. Consulte “Pré-requisitos” na página 5 para obter informações adicionais.
- Uma configuração de uma instalação de Alta Disponibilidade deve sempre usar o DB2 como o repositório de metadados, e a instância do DB2 deve ser instalada e estar disponível antes da execução da ferramenta de configuração.

Nota: Se a instalação original era de Alta Disponibilidade, essas etapas serão pré-preenchidas com os valores dessa instalação original.

Servidor analítico

A ferramenta de migração usa a configuração original para determinar o principal que serve de administrador do Servidor Analítico. Esse principal será capaz de criar, excluir e gerenciar locatários. É possível alterar o principal por meio da ferramenta de administração.

Resumo

O passo Resumo permite revisar os passos antes da conclusão da migração.

Pós-migração

O passo Pós-migração fornece informações de que a migração foi bem-sucedida. Caso ocorra uma falha, revise o arquivo de log no diretório logs.

Capítulo 5. Fazendo mudanças em seu sistema

Usando a ferramenta de administração

A ferramenta de administração automatiza inúmeros outros passos manuais para atualizar sua configuração. A ferramenta de administração deve ter acesso ao ambiente do HDFS.

É possível ativar a ferramenta de administração executando o script `admintool` em `{AS_ROOT}\bin`. Por padrão, a ferramenta detecta se a ativação está no modo de GUI ou console. É possível forçar o modo de console passando para ele o argumento `-c`.

Resumo

O painel Resumo é sempre mostrado antes de qualquer mudança permanente no sistema. Quando você conclui um conjunto de passos sob uma área Administrativa, o botão final o leva até o painel Resumo. Lá, é possível selecionar outra área Administrativa para modificar ou clicar em Aplicar para aplicar mudanças no sistema.

Cluster do Hadoop

Os passos do cluster do Hadoop permitem reconfigurar o Servidor Analítico para uso com sua distribuição do Hadoop.

Nota: A ferramenta de administração permitirá a atualização da versão da distribuição do Hadoop, caso o fornecedor do Hadoop suporte uma atualização de versão no local. Se você estiver alterando fornecedores do Hadoop, é possível migrar as configurações para sua instalação existente do Servidor Analítico para uma nova instalação configurada para usar a nova distribuição do Hadoop. Para obter mais informações, consulte Capítulo 4, “Migração”, na página 19.

Atualização Simples

Isso será útil se a versão da distribuição do Hadoop tiver mudado e você simplesmente precisar atualizar os arquivos JAR usados pelo Servidor Analítico.

Atualização Orientada

Isso será útil se outras propriedades de seu cluster do Hadoop tiverem mudado. As telas que seguem a atualização Orientada são semelhantes àquelas na ferramenta de configuração.

Hcatalog/Hive

Os passos Hcatalog/Hive permitem reconfigurar o Servidor Analítico para uso com sua instalação do Hive ou ativar a instalação existente do Servidor Analítico com uma nova instalação do Hive.

Se você tiver uma instalação existente do Hive, haverá duas opções para a atualização da configuração.

Atualização Simples

Isso será útil se a versão do Hive tiver mudado e você simplesmente precisar atualizar os arquivos JAR usados pelo Servidor Analítico.

Atualização Orientada

Isso será útil se outras propriedades da instalação do Hive tiverem mudado. As telas que seguem a atualização Orientada são semelhantes àquelas na ferramenta de configuração.

JDBC

Copie os drivers para os bancos de dados que estão sendo usados para o local especificado e continue os passos de administração.

Application Server

Os passos do Application Server são dependentes do tipo de instalação.

- Em uma instalação que não seja de Alta Disponibilidade, é possível configurar comunicação segura (SSL/TLS) e configurar as portas nas quais o Servidor Analítico atende.
- Em uma instalação de Alta Disponibilidade, a ferramenta de administração permitirá que você observe quaisquer mudanças no ambiente do WebSphere® Application Server Liberty Network Deployment.

Nota: Para incluir ou remover um servidor de quarentena, você precisa editar a propriedade de configuração `ae.cluster.quarantine.server.name`. Para obter mais informações, consulte “Editando manualmente propriedades de configuração”.

Zookeeper

A ferramenta de administração permite modificar a lista de servidor na combinação do Zookeeper.

Repositório de metadados

Os passos do Repositório de Metadados permitem modificar sua instância do DB2. Se você alterar o esquema, o novo esquema especificado já deverá existir no banco de dados.

Se você configurou Servidor Analítico para usar Derby como repositório de metadados e desejar alternar para o DB2, use a ferramenta de migração.

Servidor analítico

O servidor analítico permite alterar o principal que serve de administrador do Servidor Analítico. Esse principal será capaz de criar, excluir e gerenciar locatários.

Pós-administração

O passo Pós-administração fornece informações de que as mudanças na configuração foram bem-sucedidas. Caso ocorra uma falha, revise o arquivo de log no diretório logs.

Editando manualmente propriedades de configuração

Há algumas propriedades de configuração que podem ser alteradas editando-as manualmente.

`config.properties`

O arquivo `config.properties` pode ser localizado no diretório `{AS_SERVER_ROOT}/configuration`. Há duas maneiras de se editar esse arquivo.

- A ferramenta de administração, quando ativada no modo GUI, permite fazer mudanças nesse arquivo.
- Caso contrário, é possível editar esse arquivo manualmente.

Em uma instalação de Alta Disponibilidade, é possível enviar mudanças por push para `config.properties` por meio da ferramenta de administração. Em uma instalação que não seja de Alta Disponibilidade, você precisará reiniciar o Servidor Analítico.

join.small.data.size

A quantidade máxima de dados (em bytes) que o Servidor Analítico tentará juntar em um algoritmo Map side.

mapred.child.java.opts

Controla o tamanho de heap da JVM para as tarefas mapear e reduzir que são executadas no Hadoop. Configure esse valor para o máximo que os nós no cluster podem manipular.

mapred.job.reuse.jvm.num.tasks

Reutilize JVMs para tarefas mapear e reduzir. Quando a sobrecarga para iniciar a JVM estiver alta, aumente o valor dessa propriedade. Por exemplo, quando o valor estiver configurado para 1, cada tarefa mapear e reduzir será executada em uma nova JVM. Quando o valor estiver configurado para 5, uma nova JVM será criada para cada 5 tarefas.

mmr.threads

O número de encadeamentos a ser usado para tarefas de redução de mapa na memória (M3R).

Nota: As propriedades prefixadas por "mmr" controlam a tarefa mapear reduzir na memória e são importantes para qualquer instalação que faça muito uso de pequenos arquivos. Talvez seja necessário experimentar esses valores para otimizar seu sistema.

mmr.upper.bound.threshold

Essa é a quantidade máxima de dados que serão processados pela M3R. Grandes quantidades de dados serão processadas pelo Hadoop.

mmr.synthetic.block.size

O lado do bloco padrão se o arquivo estiver no sistema de arquivos local.

mmr.diskspill.threshold

Dados de M3R intermediários são gravados em disco quando excedem esse limite.

mmr.diskspill.output.buffer

O buffer de saída do mapeador. Ele deve ser menor que **mmr.diskspill.threshold**.

mmr.taskparallel.targets.threshold

Se a razão de destinos/núcleos for menor que esse limite, as tarefas serão processadas pela M3R.

node.max.jobs

Em uma instalação de Alta Disponibilidade, o número máximo de tarefas do Servidor Analítico que um membro de cluster pode executar simultaneamente em um determinado momento.

ae.cluster.quarantine.server.name

Em uma instalação de Alta Disponibilidade, o nome do servidor designado como servidor de quarentena.

jvm.options

O arquivo `jvm.options` pode ser localizado no diretório `{AS_SERVER_ROOT}`. A inclusão ou modificação das seguintes linhas para `jvm.options` configura o tamanho de heap para tarefas executadas no servidor hospedando Servidor Analítico; ou seja, não Hadoop. Isso será importante se você estiver executando pequenas tarefas (M3R), e talvez seja necessário experimentar esses valores para otimizar seu sistema.

```
-Xms512M  
-Xmx2048M
```

Em uma instalação que não seja de Alta Disponibilidade, você precisará reiniciar o Servidor Analítico. Em uma instalação de Alta Disponibilidade, não há uma maneira simples de enviar mudanças por push para o `jvm.options` para o conjunto; é necessário atualizar manualmente o `jvm.options` em cada membro de cluster.

Implementar para novos membros de cluster de alta disponibilidade

Após juntar um novo membro do servidor no cluster no Liberty, será possível implementar o Servidor Analítico em novos membros executando o comando a seguir.

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

Sem a opção **addClusterMembers**, o Servidor Analítico é reimplementado para cada membro do cluster em alta disponibilidade.

Capítulo 6. Desinstalando

1. Execute o IBM Installation Manager (GUI ou linha de comandos), selecione a opção para desinstalar o IBM SPSS Analytic Server, e siga os prompts. O IBM Installation Manager também pode ser executado no modo silencioso. Para obter mais informações, consulte a documentação do IBM Installation Manager: http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

Nota: Essa operação não remove nenhum dado da pasta Servidor Analítico no HDFS. Para remover a pasta inicial inteira do Servidor Analítico, remova o diretório do HDFS no qual os artefatos do usuário do Servidor Analítico (entradas e saídas) estão armazenados. Esse é o diretório especificado como `root.folder.path` em `{AS_SERVER_ROOT}/configuration/config.properties`.

Capítulo 7. Resolução de problemas

Esta seção descreve alguns problemas comuns de instalação e configuração e como corrigi-los.

Clusters de alta disponibilidade

java.net.SocketTimeoutException: Read timed out

Altere a variável de ambiente Liberty ND timeout como a seguir:

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

em que <milliseconds> é o número de segundos usado para o tempo limite de leitura do JMX.

java.io.IOException: CWWKX7202E: The timeout value 60 (seconds) for command ./server start expired

Inclua o seguinte no server.xml do Controller Server

```
<!-- Aumente os tempos limite de início e parada do servidor para acomodar o hardware lento -->
<serverCommands startServerTimeout="120" stopServerTimeout="120"/>
```

java.lang.OutOfMemoryError: Java heap space

Inclua as seguintes linhas no jvm.options em cada membro do cluster de AD.

```
-Xms512M
-Xmx2048M
```

Dados de transação do Zookeeper não podem ser gerenciados

Configure o parâmetro **autopurge.purgeInterval** em zoo.cfg para 1 para ativar limpezas automáticas do log de transação do Zookeeper.

Serviço de cluster analítico perde contato com Zookeeper

Revise e modifique os parâmetros **tickTime**, **initLimit** e **syncLimit** no zoo.cfg. Por exemplo:

```
# 0 número de milissegundos de cada marcação
tickTime=2000
# 0 número de marcações que a fase de sincronização # inicial pode obter
initLimit=30
# 0 número de marcações que podem ser passadas entre
# o envio de uma solicitação e a obtenção de reconhecimento
syncLimit=15
```

Consulte a documentação do Zookeeper para obter detalhes: <https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html>

Tarefas do Servidor Analítico não continuam

Há duas situações comuns em que tarefas do Servidor Analítico não continuam.

1. Quando uma tarefa do Servidor Analítico falha porque um membro de cluster falha, normalmente a tarefa é reiniciada automaticamente em outro membro de cluster. Se a tarefa não continuar, verifique se há pelo menos 4 membros de cluster no cluster de Alta Disponibilidade.
2. Quando você coloca um membro de cluster em modo quiesce, todas as tarefas do Servidor Analítico nesse servidor normalmente continuam em outro membro de cluster. Para assegurar que as tarefas continuem, configure `-Dcom.spss.ae.remoteclient.failover.threshold=100` e use o modo remoto.

Avisos

Estas informações foram desenvolvidas para produtos e serviços oferecidos nos Estados Unidos.

É possível que a IBM não ofereça os produtos, serviços ou recursos discutidos neste documento em outros países. Consulte um representante IBM local para obter informações sobre produtos e serviços disponíveis atualmente em sua área. Qualquer referência a produtos, programas ou serviços IBM não significa que apenas produtos, programas ou serviços IBM possam ser utilizados. Qualquer produto, programa ou serviço funcionalmente equivalente, que não infrinja nenhum direito de propriedade intelectual da IBM poderá ser usado em substituição a este produto, programa ou serviço. Entretanto, a avaliação e verificação da operação de qualquer produto, programa ou serviço não IBM são de responsabilidade do usuário.

A IBM pode ter patentes ou solicitações de patentes pendentes relativas a assuntos tratados neste documento. O fornecimento desse documento não garante a você nenhum direito sobre tais patentes. Pedidos de licença devem ser enviados, por escrito, para:

Gerência de Relações
Industriais e Comerciais da IBM Brasil
Av. Pasteur, 138-146 Botafogo
Botafogo
Rio de Janeiro, RJ
CEP 22290-240

Para consultas sobre licença relacionadas a informações de DBCS (Conjunto de Caracteres de Byte Duplo), entre em contato com o Departamento de Propriedade Intelectual da IBM em seu país ou envie pedidos de licença, por escrito, para:

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

O parágrafo a seguir não se aplica ao Reino Unido ou qualquer outro país em que tais disposições estejam inconsistentes com a lei local: A INTERNATIONAL BUSINESS MACHINES CORPORATION FORNECE ESTA PUBLICAÇÃO "NO ESTADO EM QUE SE ENCONTRA", SEM GARANTIA DE NENHUM TIPO, SEJA EXPRESSA OU IMPLÍCITA, INCLUINDO, MAS A ELAS NÃO SE LIMITANDO, AS GARANTIAS IMPLÍCITAS DE NÃO INFRAÇÃO, COMERCIALIZAÇÃO OU ADEQUAÇÃO A UM DETERMINADO PROPÓSITO. Alguns estados não permitem a renúncia de responsabilidade de garantias expressas ou implícitas em determinadas transações, portanto, esta declaração pode não se aplicar a você.

Essas informações podem incluir imprecisões técnicas ou erros tipográficos. Periodicamente, são feitas mudanças nas informações aqui contidas; tais mudanças serão incorporadas em novas edições da publicação. A IBM pode, a qualquer momento, aperfeiçoar e/ou alterar os produtos e/ou programas descritos nesta publicação, sem aviso prévio.

Quaisquer referências nestas informações ou a websites não IBM são fornecidas apenas por conveniência e não representam de forma alguma um endosso a esses documentos ou websites. Os materiais contidos nesses websites não fazem parte dos materiais desse produto IBM e a utilização desses websites é de inteira responsabilidade do Cliente.

A IBM pode utilizar ou distribuir informações fornecidas da forma que julgar apropriada sem incorrer em qualquer obrigação para com o Cliente.

Licenciados deste programa que pretendam obter informações adicionais com o objetivo de permitir: (i) a troca de informações entre programas criados independentemente e outros programas (incluindo este) e (ii) o uso mútuo das informações trocadas, devem entrar em contato com:

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

Tais informações podem estar disponíveis, sujeitas a termos e condições apropriadas, incluindo em alguns casos o pagamento de uma taxa.

O programa licenciado descrito neste documento e todo o material licenciado disponível são fornecidos pela IBM sob os termos do IBM Customer Agreement, do Contrato de Licença do Programa Internacional da IBM ou de qualquer outro contrato equivalente.

Todos os dados de desempenho aqui contidos foram determinados em um ambiente controlado. Portanto, os resultados obtidos em outros ambientes operacionais podem variar significativamente. Algumas medidas podem ter sido tomadas em sistemas em nível de desenvolvimento e não há garantia de que estas medidas serão iguais em sistemas geralmente disponíveis. Além disso, algumas medidas podem ter sido estimadas por meio de extrapolação. Os resultados reais podem variar. Os usuários deste documento devem verificar os dados aplicáveis para seu ambiente específico.

As informações relativas a produtos não IBM foram obtidas junto aos fornecedores dos respectivos produtos, de seus anúncios publicados ou de outras fontes disponíveis publicamente. A IBM não testou estes produtos e não pode confirmar a precisão de seu desempenho, compatibilidade, nem qualquer outra reivindicação relacionada a produtos não IBM. Perguntas sobre os recursos de produtos não IBM devem ser endereçadas aos fornecedores desses produtos.

Todas as instruções relativas às direções ou intenções futuras da IBM estão sujeitas a mudanças ou retirada sem aviso prévio, e apenas representam metas e objetivos.

Todos os preços da IBM mostrados são preços de varejo sugeridos pela IBM, são atuais e estão sujeitos a mudanças sem aviso prévio. Os preços para o revendedor podem variar.

Estas informações têm apenas o propósito de planejamento. As informações aqui contidas estão sujeitas a mudanças antes que os produtos descritos estejam disponíveis.

Estas informações contêm exemplos de dados e relatórios utilizados em operações comerciais diárias. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos estes nomes são fictícios e qualquer semelhança com nomes e endereços utilizados por uma empresa real é mera coincidência.

Cada cópia ou qualquer parte desses programas de amostra ou qualquer trabalho derivado deve incluir um aviso de copyright como a seguir:

Estas informações contêm exemplos de dados e relatórios utilizados em operações comerciais diárias. Para ilustrá-los da forma mais completa possível, os exemplos podem incluir nomes de indivíduos, empresas, marcas e produtos. Todos estes nomes são fictícios e qualquer semelhança com nomes e endereços utilizados por uma empresa real é mera coincidência.

Cada cópia ou qualquer parte desses programas de amostra ou qualquer trabalho derivado deve incluir um aviso de copyright como a seguir:

© (nome da empresa) (ano). Partes desse código são derivadas dos Programas de Amostra da IBM Corp.

© Copyright IBM Corp. _insira o(s) ano(s). All rights reserved.

Se estiver vendo estas informações em formato eletrônico, as fotografias e ilustrações coloridas podem não aparecer.

Marcas comerciais

IBM, o logotipo IBM e ibm.com são marcas comerciais ou marcas registradas da International Business Machines Corp., registradas em vários países no mundo todo. Outros nomes de produtos e serviços podem ser marcas comerciais da IBM ou de outras empresas. Uma lista atual de marcas comerciais da IBM está disponível na Web em "Copyright and trademark information", em www.ibm.com/legal/copytrade.shtml.

Adobe, o logotipo Adobe, PostScript e o logotipo PostScript são marcas comerciais ou marcas registradas da Adobe Systems Incorporated nos Estados Unidos, e/ou outros países.

IT Infrastructure Library é uma marca registrada da Central Computer and Telecommunications Agency, a qual agora é parte do departamento de comércio do governo.

Intel, logotipo Intel, Intel Inside, logotipo Intel Inside, Intel Centrino, logotipo Intel Centrino, Celeron, Intel Xeon, Intel SpeedStep, Itanium e Pentium são marcas comerciais ou marcas registradas da Intel Corporation ou de suas subsidiárias nos Estados Unidos e em outros países.

Linux é uma marca registrada de Linus Torvalds nos Estados Unidos e/ou em outros países.

Microsoft, Windows, Windows NT e o logotipo Windows são marcas comerciais da Microsoft Corporation nos Estados Unidos, e/ou em outros países.

ITIL é uma marca registrada e uma marca comercial de comunidade registrada do The Minister for the Cabinet Office e está registrada no U.S. Patent and Trademark Office.

UNIX é uma marca registrada da The Open Group nos Estados Unidos e em outros países.

Java e todas as marcas comerciais e logotipos baseados em Java são marcas comerciais ou marcas registradas da Oracle e/ou suas afiliadas.

Cell Broadband Engine é uma marca comercial da Sony Computer Entertainment, Inc. nos Estados Unidos e/ou em outros países e é usada sob licença desta empresa.

Linear Tape-Open, LTO, o logotipo LTO, Ultrium e o logotipo Ultrium são marcas comerciais da HP, IBM Corp. e Quantum nos Estados Unidos e em outros países.



Impresso no Brasil