

IBM SPSS Analytic Server
버전 2

설치 및 구성 안내서

IBM

참고

이 정보와 이 정보가 지원하는 제품을 사용하기 전에, 31 페이지의 『주의사항』의 정보를 읽으십시오.

제품 정보

이 개정판은 새 개정판에 별도로 명시하지 않는 한 IBM SPSS Analytic Server의 버전 2, 릴리스 0, 수정사항 0 및 모든 후속 릴리스와 수정사항에 적용됩니다.

목차

제 1 장 개요	1	제 4 장 마이그레이션	19
아키텍처	2	마이그레이션 도구 사용	20
버전 2 설치 프로그램의 새로운 사항	2	제 5 장 시스템 변경하기	23
제 2 장 설치	5	관리 도구 사용	23
필수 소프트웨어	5	구성 특성 수동 편집	25
고가용성 클러스터 필수조건	6	새 고가용성 클러스터 멤버에 배치.	26
설치	7	제 6 장 설치 제거	27
제 3 장 구성	9	제 7 장 문제 해결	29
구성 도구 사용	9	주의사항	31
사용자 레지스트리 작성	12	상표.	33
HCatalog 데이터 소스 사용.	13		
Essentials for R에 대한 지원	16		
IBM SPSS Analytic Server와 함께 사용하도록			
IBM SPSS Modeler 구성	17		

제 1 장 개요

IBM® SPSS® Analytic Server는 IBM SPSS 기술을 빅 데이터 시스템과 통합하여 사용자가 친숙한 IBM SPSS 사용자 인터페이스로 작업하면서 기존에 다룰 수 없던 문제까지 해결할 수 있도록 지원하는 솔루션입니다.

빅 데이터 분석이 중요한 이유

조직에서 수집하는 데이터 볼륨이 기하급수적으로 증가하고 있습니다. 예를 들어, 금융 및 소매 업계는 모든 고객의 1년(또는 2년이나 10년까지) 치 거래를 가지고 있고 통신 제공업체의 경우 CDR(Call Data Record) 및 디바이스 센서 자료를 인터넷 회사는 웹 크롤링 결과를 보관합니다.

다음과 같은 경우 빅 데이터 분석이 필요합니다.

- 대형 데이터(테라바이트, 페타바이트, 엑사바이트)의 경우. 특히 구조화 및 비구조화 데이터가 혼합된 경우.
- 빠르게 변하거나 누적되는 데이터

빅 데이터 분석은 다음과 같은 경우에도 도움이 됩니다.

- 수천 개의 모델을 빌드하는 경우
- 모델을 자주 빌드하거나 갱신하는 경우

해결 과제

대용량의 데이터를 수집하는 같은 조직에서도 여러 이유로 이러한 데이터를 실제로 활용하지 못하는 경우가 있습니다.

- 기존 분석 제품의 아키텍처는 분산 계산에 적합하지 않습니다. 또한
- 기존 통계 알고리즘은 빅 데이터 작업에 맞게 설계되어 있지 않습니다(이러한 알고리즘은 데이터를 가져와야 하지만 빅 데이터를 이동하는 데에는 많은 비용이 듭). 따라서
- 빅 데이터에 대한 최신 분석을 수행하려면 새로운 기술과 빅 데이터 시스템에 대한 상세한 지식이 필요합니다. 이러한 기술이 있는 분석가는 많지 않습니다.
- 인메모리 솔루션은 중소형 문제를 다룰 수 있지만 빅 데이터를 취급할 수는 없습니다.

해결책

Analytic Server에서는 다음 기능을 제공합니다.

- HDFS 데이터가 있는 빅 데이터 시스템(예: Hadoop Map/Reduce)을 활용하는 데이터 중심 아키텍처.
- 데이터로 이동하는 새 통계 알고리즘을 통합하는 정의된 인터페이스.
- 빅 데이터 환경의 세부사항을 숨겨서 분석가가 데이터 분석에 집중할 수 있는 친숙한 IBM SPSS 사용자 인터페이스.
- 모든 규모의 문제를 해결할 수 있는 확장 가능한 솔루션.

아키텍처

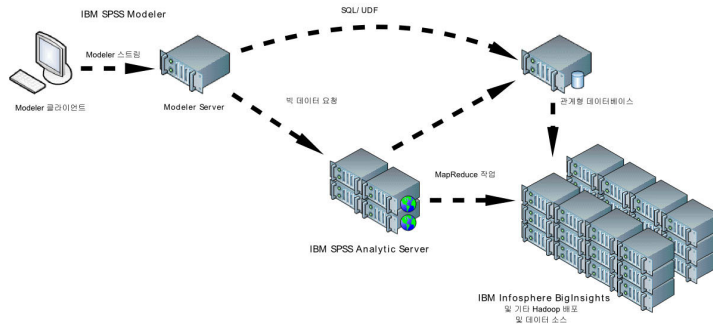


그림 1. 아키텍처

Analytic Server는 클라이언트 애플리케이션과 Hadoop 클라우드 사이에 있습니다. 데이터가 클라우드에 있다고 가정할 때 Analytic Server 작업의 일반 개요에서 다음 내용을 다룹니다.

1. 클라우드의 데이터에 대해 Analytic Server 데이터 소스를 정의합니다.
2. 클라이언트 애플리케이션에서 수행할 분석을 정의합니다. 현재 릴리스의 경우, 클라이언트 애플리케이션은 IBM SPSS Modeler입니다.
3. 분석을 실행할 때 클라이언트 애플리케이션에서 Analytic Server 실행 요청을 제출합니다.
4. Analytic Server에서 작업을 조정하여 Hadoop 클라우드에서 실행하도록 하고 결과를 클라이언트 애플리케이션에 보고합니다.
5. 결과를 사용하여 추가로 분석을 정의하고 주기를 반복할 수 있습니다.

버전 2 설치 프로그램의 새로운 사항

설치 및 구성

제품 설치 시 여러 IBM 제품을 설치하는 경우 이제 IBM Installation Manager를 사용하여 중단 없는 설치 경험을 제공합니다.

Analytic Server를 구성하는 단계는 모든 제품 확장을 설치한 후에 실행되는 새 구성 도구에서 처리됩니다.

플랫폼

새 운영 체제와 Hadoop 배포를 지원합니다.

운영 체제

Analytic Server가 이제 Microsoft Windows에서 실행됩니다.

Hadoop 배포

Analytic Server를 Hortonworks 데이터 플랫폼과 함께 Microsoft Windows에서 실행하도록 구성할 수 있습니다.

HCatalog 어댑터

Apache Cassandra, MongoDB 및 Oracle NoSQL에 대한 데이터 소스를 정의할 수 있습니다. 이러한 데이터 소스를 사용하도록 설정하는 추가 구성 단계가 있습니다.

고가용성 클러스터

IBM WebSphere Liberty Network Deployment를 사용하여 고가용성 클러스터에 Analytic Server를 설치하여 구성할 수 있습니다.

제 2 장 설치

필수 소프트웨어

Analytic Server를 설치하기 전에 다음 정보를 검토하십시오.

시스템 요구 사항

최신 시스템 요구 사항 정보를 보려면 IBM 기술 지원 사이트의 세부 시스템 요구 사항 보고서를 확인하십시오. <http://publib.boulder.ibm.com/infocenter/prodguid/v1r0/clarity/softwareReqsForProduct.html>. 이 페이지에서 다음을 수행하십시오.

1. SPSS Analytic Server를 제품 이름으로 입력하고 검색을 클릭하십시오.
2. 원하는 보고서 버전과 범위를 선택하고 제출을 클릭하십시오.

관리 권한

Analytic Server를 설치한 사용자는 Hadoop 배포 파일 시스템에 대한 쓰기 액세스 권한을 갖고 있어야 합니다.

IBM Installation Manager

Analytic Server 설치 파일이 포함된 리포지토리를 사용하려면 현재 버전의 IBM Installation Manager를 설치하여 구성해야 합니다. IBM Installation Manager가 이전 버전인 경우, 설치 과정에서 지원되는 버전으로 업데이트해야 합니다.

IBM Installation Manager가 시스템에 없으면 IBM Corp. 지원 사이트(<http://www.ibm.com/support>)에서 IBM Installation Manager를 다운로드하여 설치하십시오.

중요사항: Microsoft Windows의 경우, 관리자 계정에서 IBM Installation Manager를 설치하십시오. `./groupinstc -acceptLicense`를 사용하여 다른 사용자로 그룹 설치를 수행할 수 있습니다. 자세한 정보는 IBM Installation Manager 문서를 참조하십시오. http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html

Hive/HCatalog

NoSQL 데이터 소스를 사용할 계획이라면 Hive 및 HCatalog를 원격 액세스용으로 구성하십시오. 또한 `hive-site.xml`에 활성 Thrift Hive Metastore 서버를 가리키는 `thrift://<host_name>:<port>` 양식으로 `hive.metastore.uris` 특성이 포함되어 있는지 확인하십시오. 세부사항은 Hadoop 배포 문서를 참조하십시오.

메타데이터 리포지토리

기본적으로 Analytic Server는 Apache Derby 데이터베이스를 설치하여 사용합니다. 또는 Analytic Server에서 기존 DB2 설치를 사용하도록 구성할 수 있습니다. 데이터베이스의 인코딩은 UTF-8이어야 합니다. IBM DB2 10.5 Workgroup Server Edition은 Analytic Server 매체에서 제공됩니다. 자세한 정보는 http://www-01.ibm.com/support/knowledgecenter/SSEPGG_10.5.0/com.ibm.db2.luw.kc.doc/welcome.html의 내용을 참조하십시오.

고가용성 클러스터

고가용성 클러스터에 Analytic Server를 설치하기 위한 필수조건은 『고가용성 클러스터 필수조건』의 내용을 참조하십시오.

고가용성 클러스터 필수조건

Analytic Server를 고가용성 클러스터에 설치하는 데 필요한 추가 필수조건이 있습니다.

시스템 및 DNS 구성

환경에서 사용되는 시스템들은 서로 통신할 수 있어야 합니다. 도메인에 속한 시스템만 도메인 접미문자를 가질 수 있습니다. 예를 들어, "IBM.COM" 도메인에 등록된 시스템만 FQDN(Fully Qualified Domain Name)에 접미문자 "ibm.com"을 사용할 수 있습니다.

IBM.COM 도메인에 있지 않은 시스템에 FQDN 접미문자 "ibm.com"이 있으면 시스템의 FQDN을 DNS에 입력할 때 TCP 구성에 DNS 서버가 있는 시스템을 사용해야 합니다. 그렇지 않으면 연관된 모든 시스템의 모든 FQDN이 연관된 모든 시스템의 호스트 파일에 있어야 합니다.

메타데이터 리포지토리

Analytic Server가 고가용성 클러스터에 설치된 경우 DB2를 사용해야 합니다.

Zookeeper 3.4.5

세 개 이상의 Zookeeper 인스턴스가 필요합니다. Zookeeper 설치 및 구성 문서를 참조하십시오. http://zookeeper.apache.org/doc/r3.4.5/zookeeperStarted.html#sc_RunningReplicatedZooKeeper.

참고: 이 단계를 Zookeeper가 설치된 각 시스템에서 수행해야 합니다.

Liberty Network Deployment 8.5.5.4

고가용성 환경은 한 개의 제어기, 두 개의 복제본 및 두 개 이상의 클러스터 멤버가 있는 Liberty 집단으로 구성됩니다. Liberty 집단을 구성하는 데 관한 세부사항은 WebSphere Application Server Network Deployment 문서를 참조하십시오. http://www-01.ibm.com/support/knowledgecenter/SSAW57/mapfiles/product_welcome_wasnd.html.

로드 밸런서

고가용성 클러스터에는 세션 선호도(스티키(sticky) 세션이라고도 함)를 지원하는 로드 밸런서가 있어야 합니다. Analytic Server는 쿠키 "request-token"이 있는 세션을 식별합니다. 이를 통해 애플리케이션 제어 세션 선호도에 사용할 사용자 로그인 기간에 대한 세션을 식별합니다. 자세한 세션 선호도 지원 방법을 보려면 특정 로드 밸런서의 문서를 확인하십시오.

참고: 리더 노드는 클러스터의 로드 밸런서에 포함될 수 없습니다. 사용자가 시작한 첫 노드가 리더 노드입니다.

Visual C++ 2008(Windows 전용)

Windows에서 Analytic Server를 실행하는 클러스터 멤버가 있으면 Analytic Server를 설치하기 전에 Visual C 2008 Runtime을 설치해야 합니다.

설치

Analytic Server 애플리케이션 파일은 IBM Installation Manager를 사용하여 호스트 시스템에 설치됩니다. 설치 파일은 실제 매체에 있거나 IBM Passport Advantage에서 다운로드할 수 있습니다.

1. IBM Installation Manager를 시작하십시오.

GUI 모드

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

명령행 모드

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

참고: IBM Installation Manager를 실행하는 사용자는 Analytic Server 프로세스를 실행하는 사용자입니다. 이 사용자는 Hadoop 환경에 액세스할 수 있어야 합니다.

2. 설치 리포지토리가 구성되지 않은 경우, 리포지토리 경로를 지정하십시오. 예를 들어, 마운트된 설치 디스크나 호스트 파일 시스템의 위치, 네트워크 또는 HTTP 주소를 지정하십시오.
3. 기본 메뉴에서 설치를 선택하십시오.
4. 설치할 패키지를 선택하십시오. 사용 가능한 옵션은 구입한 Analytic Server 버전에 따라 달라집니다.

IBM SPSS Analytic Server

이 패키지는 대부분의 시스템에 적합합니다.

IBM SPSS Analytic Server High Availability

Analytic Server를 고가용성 클러스터에서 실행하려는 경우 이 패키지가 필요합니다. IBM SPSS Analytic Server High Availability의 구성 차이점이 문서화되어 있습니다.

5. 라이선스 계약을 읽고 계약을 승인하십시오.
6. 패키지 그룹과 설치 디렉토리를 지정하십시오.
 - IBM SPSS Analytic Server 설치에 새 패키지 그룹이 필요합니다.
 - 공유 자원의 설치 디렉토리를 지정하십시오. 패키지를 처음 설치할 때 만 공유 자원 디렉토리를 지정할 수 있습니다.
7. 설치할 기능을 선택하십시오.
8. 요약 정보를 검토하고 설치를 계속하십시오. 애플리케이션 파일이 지정된 디렉토리에 설치됩니다.
9. 추가 구성이 필요합니다. 신규 구성에 대한 자세한 정보는 9 페이지의 『구성 도구 사용』 주제를 참조하십시오. 기존 구성을 새 설치로 마이그레이션하는 방법에 대한 자세한 정보는 19 페이지의 제 4 장 『마이그레이션』 주제를 참조하십시오. IBM Installation Manager를 GUI 모드로 실행 중인 경우, 구성 또는 마이그레이션 도구를 시작하는 옵션이 마지막 단계에 표시됩니다.

설치 중에 문제가 발생하면 IBM Installation Manager 로그를 사용하여 해결할 수 있습니다. 로그 파일은 IBM Installation Manager의 기본 메뉴에서 액세스할 수 있습니다.

IBM Installation Manager를 자동 모드로 실행하여 설치를 자동화할 수 있습니다. 자세한 정보는 IBM Installation Manager 문서를 참조하십시오. http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

제 3 장 구성

설치 후 제품을 사용하기 전에 Analytic Server를 구성해야 합니다.

참고: Analytic Server 파일 경로에는 다음 규칙을 사용합니다.

- {AS_ROOT}는 Analytic Server가 배포된 위치를 나타냅니다. 예: /opt/IBM/SPSS/AnalyticServer/{version}.
- {AS_SERVER_ROOT}는 구성, 로그 및 서버 파일의 위치를 나타냅니다. 예: /opt/IBM/SPSS/AnalyticServer/{version}/ae_wlpserver/usr/servers/aeserver.
- {AS_HOME}은 Analytic Server에서 루트 폴더로 사용되는 HDFS의 위치를 나타냅니다.

구성 단계

1. [선택사항] Analytic Server를 NoSQL 또는 XML 데이터 소스에서 실행하려면 HCatalog 데이터 소스 사용에 대한 단계를 따르십시오.

참고: Analytic Server 단순 구성을 지금 수행하고 관리 도구를 사용하여 나중에 HCatalog 데이터 소스를 사용하도록 설정할 수 있습니다.

2. 구성 도구를 실행하십시오. 해당 도구는 다른 여러 수동 단계를 자동화합니다. IBM Installation Manager의 GUI 버전을 사용한 경우 구성 도구가 자동으로 시작되며 그렇지 않으면 {AS_ROOT}/bin에서 **configtool** 스크립트를 실행해야 합니다.
3. Analytic Server 시스템에 추가할 수 있는 사용자의 레지스트리를 작성하십시오.
4. Analytic Server와 사용할 IBM SPSS Modeler를 구성하십시오.
5. [선택사항] Analytic Server에서 R 코드를 실행하도록 하려면 Essentials for R에 대한 지원을 사용하십시오.
6. Analytic Server가 실행 중인지 확인하십시오. {AS_ROOT}/bin의 **start** 스크립트를 실행하여 프로세스를 시작할 수 있습니다.
7. 웹 브라우저를 열고 주소 <http://localhost:8080/analyticserver/admin/ibm>을 입력하십시오. 이 URL은 Analytic Server 콘솔의 로그인 대화 상자를 엽니다. 구성 도구에 정의된 Analytic Server 관리자로 로그인하십시오.

구성 도구 사용

구성 도구는 다른 여러 수동 단계를 자동화합니다. 구성 도구에는 HDFS 환경에 대한 액세스 권한이 있어야 합니다.

IBM Installation Manager의 GUI 버전을 사용한 경우 구성 도구가 자동으로 시작됩니다. 프롬프트에 따라서 시스템을 구성하십시오.

{AS_ROOT}\bin에서 **configtool** 스크립트를 실행하여 구성 도구를 실행할 수도 있습니다. 기본적으로 이 도구는 GUI 모드나 콘솔 모드의 실행 여부를 감지합니다. 인수 **-c**를 전달하여 콘솔 모드를 강제 실행하거나 원하는 구성을 지정하는 파일 경로를 전달하여 자동 모드를 강제 실행할 수 있습니다. 샘플은 {AS_ROOT}\bin의 configtool.response 파일을 참조하십시오.

시작

시작 단계에서는 구성 도구를 통해 수행하는 경로를 판별합니다.

일반 일반 경로에서는 여러 의사결정을 자동화하고 시스템에 대한 가정을 단순화합니다. 비프로덕션 환경을 빠르게 구성하려는 경우 이 옵션을 선택하십시오. 항상 관리 도구에서 구성을 변경하거나 구성을 다른 환경에 마이그레이션할 수 있습니다.

일반 구성에서 이 도구는 Hadoop 이름 노드의 위치를 입력하도록 프롬프트를 표시하며 필요한 경우 계속해서 추가 정보를 요청합니다.

고가용성 설치의 일반 구성에서는 구성 도구가 WebSphere® Application Server Liberty Network Deployment 환경과 Zookeeper 앙상블에 대한 정보를 입력하도록 프롬프트를 표시합니다.

사용자 정의

고급 구성 작업을 수행할 수 있습니다. Analytic Server와 사용할 관계형 데이터베이스 또는 NoSQL을 설정하는 경우 이 옵션이 필요합니다.

아래 섹션에서는 사용자 정의 경로에 대한 세부사항을 제공합니다.

Hadoop 클러스터

Hadoop 클러스터 단계에서는 Hadoop 배포와 함께 사용할 Analytic Server를 구성합니다. 구성 도구에서 사용자의 배포를 자동으로 발견하거나 사용자가 필요한 파일을 수동으로 선택할 수 있습니다.

Hadoop 배포에서 자동으로 파일 수집

해당 도구는 Hadoop 배포를 발견하고자 하며 필요한 경우 계속해서 추가 정보를 요청합니다.

1. Hadoop 클러스터의 호스트 이름은 항상 제공해야 합니다.
2. 구성 도구를 실행 중인 사용자의 공용 키 인증 및 SSH를 사용하여 클러스터를 설정하지 않은 경우 유효한 신임 정보를 입력해야만 구성 도구에서 Hadoop 클러스터의 파일을 수집할 수 있습니다.

제한사항: RSA SSH 암호화만 지원됩니다.

3. Hadoop 클러스터가 비표준 위치에 설치된 경우, 설치 위치를 제공해야 할 수 있습니다.

Hadoop 배포에서 수동으로 파일 수집

사용 중인 Hadoop 배포를 식별하십시오. 구성 도구에서 Hadoop 클러스터로부터 수집해야 하는 파일을 식별합니다. 이러한 파일을 지정된 위치로 복사하면 구성 단계를 계속할 수 있습니다.

참고: core-site.xml 파일의 **io.compression.codecs** 매개변수에 대해 "com.hadoop.compression.lzo.LzoCodec" 값이 지정된 경우에만 hadoop-lzo-*.jar JAR 파일이 필요합니다.

Hcatalog/Hive

Hcatalog/Hive 단계도 사용자 정의 구성의 일부로 표시됩니다.

HCatalog/Hive 구성 자동 연결 및 발견

이 도구는 사용자의 HCatalog / Hive 버전을 발견하고 필요한 파일을 수집하여 Analytic Server를 구성합니다.

HCatalog/Hive 구성 수동 선택

사용 중인 NoSQL 데이터베이스를 식별하십시오. 구성 도구에서 사용자가 수집해야 하는 파일을 식별합니다. 이러한 파일을 지정된 위치로 복사하면 구성 단계를 계속할 수 있습니다.

HCatalog/Hive 구성 건너뛰기

HCatalog/Hive를 나중에 구성할 수 있습니다.

JDBC

JDBC 단계도 사용자 정의 구성의 일부로 표시됩니다. 사용 중인 데이터베이스의 드라이버를 지정된 위치에 복사하고 구성 단계를 계속하십시오.

표 1. 지원되는 데이터베이스

데이터베이스	지원되는 버전	JDBC 드라이버 jar	벤더
Linux, UNIX 및 Windows용 DB2	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, ora18n.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

애플리케이션 서버

애플리케이션 서버 단계는 설치 유형에 따라 달라집니다.

- 고가용성이 아닌 설치에서는 보안 통신(SSL/TLS)을 구성하고 Analytic Server에서 청취하는 포트를 설정할 수 있습니다.
- 고가용성 설치에서는 구성 도구에서 사용자의 WebSphere® Application Server Liberty Network Deployment 환경에 대한 정보를 입력하도록 프롬프트를 표시합니다.

Zookeeper

Zookeeper 단계는 고가용성 설치를 위한 구성의 일부로만 필요합니다. Zookeeper 앙상블에 서버 목록을 제공하십시오.

메타데이터 리포지토리

메타데이터 리포지토리 단계는 설치 유형에 따라 달라집니다.

- 고가용성이 아닌 설치에서는 Derby 또는 DB2를 선택할 수 있습니다. DB2를 선택하는 경우, 구성 도구를 실행하기 전에 설치하여 가용케 해야 합니다. 자세한 정보는 5 페이지의 『필수 소프트웨어』의 내용을 참조하십시오.
- 고가용성 설치의 구성에서는 DB2를 메타데이터 리포지토리로 항상 사용해야 하며 DB2 인스턴스는 구성 도구를 실행하기 전에 설치하여 가용케 해야 합니다.

Analytic Server

Analytic Server 관리자 역할을 하는 프린시펄을 식별하십시오. 이 프린시펄은 테넌트를 작성, 삭제 및 관리할 수 있습니다.

요약

요약 단계에서는 구성을 완료하기 전에 구성 단계를 검토할 수 있습니다.

사후 구성

구성 후 단계에서는 구성 완료에 대한 정보를 제공합니다. 구성이 실패한 경우에는 logs 디렉토리에서 로그 파일을 검토하십시오.

사용자 레지스트리 작성

설치 및 구성 중에 작성된 기본 테넌트는 WebSphere Liberty 프로파일을 사용하여 사용자를 인증하며 이 경우 외부 LDAP 서버를 차례로 참조하거나 사용자 및 그룹의 기본 레지스트리를 정의할 수 있습니다.

참고: {AS_SERVER_ROOT}/server.xml 파일을 편집한 후에 WebSphere Liberty 프로파일을 구성하여 사용자를 인증하고 이러한 사용자를 Analytic Server 시스템에 추가하십시오. 테넌트 관리에 대한 세부사항은 *IBM SPSS Analytic Server 관리자 안내서*를 참조하십시오.

LDAP

LDAP 레지스트리를 사용하면 외부 LDAP 서버(예: Active Directory 또는 OpenLDAP)로 사용자를 인증할 수 있습니다. 다음은 OpenLDAP용 ldapRegistry의 예제입니다.

```
<ldapRegistry
  baseDN="ou=people,dc=aeldap,dc=org"
  ldapType="Custom"
  port="389"
  host="server"
```



```

id="OpenLDAP"
bindDN="cn=admin,dc=ae1dap,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>

```

추가 구성 예제에 대해서는 템플릿 폴더 {AS_ROOT}/ae_wlpserver/templates/config를 참조하십시오.

기본

기본 레지스트리를 사용하면 {AS_SERVER_ROOT}/server.xml 파일 내에서 사용자 및 그룹의 데이터베이스를 정의할 수 있습니다. {AS_ROOT}/ae_wlpserver/bin에 있는 securityUtility 도구를 사용하여 비밀번호를 명확하지 않은 값으로 인코딩할 수 있습니다.

기본 레지스트리는 샌드박스 환경에서는 유용하지만 프로덕션 환경에는 사용하지 않는 것이 좋습니다.

```

<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>

```

HCatalog 데이터 소스 사용

Analytic Server에서는 Hive/HCatalog를 통해 XML 데이터 및 NoSQL 데이터베이스의 지원을 제공합니다.

구성 도구에서 Hive/HCatalog와 함께 사용할 Analytic Server 설치를 준비하지만 수동으로 수행해야 할 단계가 있습니다. Analytic Server에서 HCatalog에 필요한 일부 인터페이스를 구현하므로 일부 JAR 파일을 Hive {HIVE_HOME}/auxlib/ 디렉토리로 복사해야 합니다. HCatalog는 조금 다른 API를 사용하므로 Analytic Server는 HCatalog에서 직접 Hive용 저장 공간 핸들러를 사용할 수 없습니다.

다음 섹션에서는 사용할 데이터 소스를 기반으로 필요한 파일을 자세하게 보여줍니다.

중요사항: Analytic Server에서 이러한 파일을 사용하려면 구성 도구 또는 관리 도구를 실행하기 전에 해당 파일을 {HIVE_HOME}/auxlib/ 디렉토리로 복사해야 합니다.

중요사항: Windows에서는 Hortonworks 데이터 플랫폼이 {HIVE_HOME}/auxlib/ 디렉토리를 인식하지 못합니다. HCatalog를 사용하려면 다음과 같이 수행하십시오.

1. 아래에서 설명하는 대로 필수 파일을 {HIVE_HOME}/lib/ 디렉토리로 복사하십시오.
2. {HIVE_HOME}/auxlib/ 디렉토리를 제거하십시오.
3. hive-site.xml 및 hive-config.cmd 파일을 업데이트하여 {HIVE_HOME}/auxlib/ 디렉토리에 대한 모든 참조를 제거하십시오.
4. Hortonworks 데이터 플랫폼 클러스터를 다시 시작하여 변경사항을 적용하십시오.

자세한 정보는 <http://hortonworks.com/community/forums/topic/hive-auxlib-issues/>의 내용을 참조하십시오.

모든 Hive/HCatalog 설치

1. 다음 파일은 {AS_ROOT}/add-ons 디렉토리의 hive_auxlib 아카이브에 있습니다.

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

2. Maven Central Repository(<http://search.maven.org/#search%7Cga%7C1%7Chivexmlserde>)에서 최신 버전의 hivexmlserde-`{version}`.jar을 복사하십시오.
3. [선택사항] Hive 복합 유형(맵, 배열, 구조)을 사용하려면 다음 JAR 파일을 복사하십시오. 해당 파일은 HCatalog 설치에서 찾을 수 있습니다. Hive 0.11.0 이전 버전에서는 {HCATALOG_HOME}/share/hcatalog에 있고 Hive 0.11.0부터는 {HIVE_HOME}/hcatalog/share/hcatalog에 있습니다.

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. Hive 버전 0.14.0을 사용 중인 경우, 다음 JAR 파일을 다운로드하십시오. <https://github.com/dvasilen/Hive-HCatalog-Compatibility/raw/master/release/hive-hcatalog-compatibility-0.14.0.jar>

XML 데이터

1. [선택사항] Optional] 압축된 XML 파일로 작업하려는 경우, {AS_ROOT}/add-ons 디렉토리의 hive_auxlib 아카이브에 있는 다음 JAR 파일을 복사하십시오.

com.ibm.spss.com.ibm.biginsights.compression_2.1.jar (BigInsights Hadoop distribution only)
com.ibm.spss.hivexmlserdecomp-2.0.0.0.jar

Apache Accumulo

1. 다음 파일은 HCatalog 설치에 있습니다. Hive 0.11.0 이전 버전에서는 {HCATALOG_HOME}/share/hcatalog에 있고 Hive 0.11.0부터는 {HIVE_HOME}/hcatalog/share/hcatalog에 있습니다.

commons-io-`{version}`.jar

2. Apache Accumulo 저장 공간 핸들러는 <https://github.com/dvasilen/accumulo-hive-storage-manager/tree/master>에서 찾을 수 있습니다. 이는 마스터 분기이며 여기에서 사용자 버전의 Accumulo 및 Hadoop의 분기로 전환해야 합니다. JAR 파일은 /release 디렉토리에 있으며 다음 양식의 이름을 사용합니다.:
accumulo-hive-storage-handler_<hadoopVersion>-<accumuloVersion>.jar

3.

다음 파일은 Apache Accumulo 설치에 있습니다.

```
accumulo-trace-{version}.jar
accumulo-fate-{version}.jar
accumulo-core-{version}.jar
accumulo-server-{version}.jar
accumulo-start-{version}.jar
```

Apache Cassandra

1. 다음 파일은 Apache Cassandra 1.2.9 릴리스(<http://www.apache.org/dyn/closer.cgi?path=/cassandra/1.2.9/apache-cassandra-1.2.9-bin.tar.gz>)에 있습니다.

```
apache-cassandra-1.2.9.jar
apache-cassandra-clientutil-1.2.9.jar
apache-cassandra-thrift-1.2.9.jar
```

2. Apache Cassandra 저장 공간 핸들러는 <https://github.com/dvasilen/Hive-Cassandra>에서 찾을 수 있습니다. 이는 마스터 분기이며 여기에서 사용자 버전의 Cassandra 및 Hadoop의 분기로 전환해야 합니다. JAR 파일은 /release 디렉토리에 있으며 다음 양식의 이름을 사용합니다.:

```
hive-<hiveVersion>-hadoop-<hadoopVersion>-cassandra-<cassandraVersion>.jar
```

참고: 2.0 이전 버전의 Hadoop의 경우 파일 또는 분기 이름의 hadoop 부분이 없습니다.

참고: Analytic Server는 Cassandra 2.0과 통신할 수 있지만 해당 JAR 파일은 Java 7용으로 컴파일됩니다. 따라서 Java 6용으로 컴파일되는 Cassandra 1.2.9 JAR 파일을 복사해야 합니다.

HBase

HBase용으로 복사할 추가 파일이 없습니다.

MongoDB

1. "mongodb/mongo-hadoop" 프로젝트의 "master" 분기에서 사용자의 Hadoop 배포를 위해 다음 파일을 로컬로 빌드해야 합니다. 자세한 정보는 <https://github.com/mongodb/mongo-hadoop>의 내용을 참조하십시오.

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

2. 다음 파일은 <https://github.com/downloads/mongodb/mongo-java-driver/mongo-2.10.1.jar>에 있습니다.

```
mongo-2.10.1.jar
```

Oracle NoSQL

1. 다음 파일은 <https://github.com/vilcek/HiveKVStorageHandler2/blob/master/dist/HiveKVStorageHandler.jar>에 있습니다.

HiveKVStorageHandler.jar

2. 다음 파일은 Apache Hive 설치에 있습니다.

avro-*<version>.jar

Essentials for R에 대한 지원

Analytic Server는 R 모델 스코어링 및 R 스크립트 실행을 지원합니다.

제한사항: R에 대한 지원은 현재 Linux에서의 Analytic Server 설치로 제한됩니다.

Analytic Server 설치를 완료한 후 R에 대한 지원을 구성하려면 다음을 수행하십시오.

1. Analytic Server를 호스팅하는 서버에 R Engine을 설치하십시오.

```
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
```

중요사항: R은 Hadoop 클러스터에서 Analytic Server 사용자와 다른 사용자로 실행되므로 R Engine을 모든 사용자가 액세스할 수 있는 위치에 설치하십시오.

중요사항: 고가용성 설치에서는 R Engine이 모든 클러스터 멤버와 동일한 위치에 설치되어야 합니다.

2. 화면의 지시사항에 따라 설치 프로그램 파일을 실행하여 Analytic Server를 호스팅하는 서버에 Essentials for R을 설치하십시오. Essentials for R을 다운로드할 수 있습니다(<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>).

3. R 엔진 및 R 구성요소를 Hadoop에 배포하십시오.

- a. Analytic Server 및 모든 Hadoop 노드에 동일한 버전의 운영 체제와 동일한 프로세서 아키텍처가 있는 경우:

- 1) 다음 스크립트를 사용하여 R 엔진에 대한 아카이브를 작성하십시오.

```
#!/usr/bin/env bash
echo Creating 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
```

참고: 스크립트의 -d\ 및 -f3 사이에 두 칸의 공백이 있습니다. PDF 문서에서는 공백이 두 칸으로 보이지만 이는 한 칸의 공백입니다. 이 스크립트를 PDF에서 복사하여 붙여넣는 경우 스크립트를 실행하기 전에 두 번째 공백을 삽입하십시오.

참고: Fortran 라이브러리를 복사하는 행은 설치된 Fortran 버전에 따라 달라집니다. 예를 들어, 1.0 버전이 설치된 경우, 해당 행이 다음과 같이 표시됩니다. 시스템에 따라 이를 편집하십시오.

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

- 2) R.zip 아카이브를 {AS_ROOT}/ae_wlpserver/usr/servers/aeserver/configuration/linux/app_64 디렉토리에 복사하십시오.
 - b. Analytic Server 및 Hadoop 노드의 운영 체제 버전이 다른 경우 R 엔진과 Essentials for R을 각 Hadoop 노드에 설치하십시오. Analytic Server에 설치된 R 엔진의 경로와 이름을 사용하여 디렉토리에 설치하십시오. Essentials for R을 Hadoop 노드에 설치하는 경우, R이 설치된 디렉토리를 지정하고 ../ext_64/bin의 위치를 지정하는 단계를 건너뛰십시오.
 - c. {AS_ROOT}/bin 디렉토리의 **hdfsUpdate** 스크립트를 실행하여 HDFS에 대한 변경사항을 전파하십시오.
4. 고가용성 설치에서는 {AS_ROOT}/bin 디렉토리의 **deployRemote** 스크립트도 실행하여 각 클러스터 멤버의 Analytic Server 인스턴스에 변경사항을 전파해야 합니다.
 5. SPSS Modeler 서버를 호스팅하는 시스템에도 Essentials for R을 설치해야 합니다.

IBM SPSS Analytic Server와 함께 사용하도록 IBM SPSS Modeler 구성

SPSS Modeler를 Analytic Server와 함께 사용하려면 SPSS Modeler 서버 설치를 업데이트해야 합니다.

1. SPSS Modeler 서버를 구성하여 Analytic Server 설치와 연관시키십시오.
 - a. 기본 서버 설치 디렉토리의 config 서브디렉토리에서 options.cfg 파일을 편집하고 다음 행을 추가하거나 편집하십시오.

```
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

Analytic Server에 보안 통신이 구성된 경우 Y를 지정하고 그렇지 않으면 N을 지정하십시오.

as_host

Analytic Server를 호스팅하는 서버의 IP 주소입니다.

as_port

Analytic Server가 청취하는 포트(기본값: 8080).

as_context_root

Analytic Server 컨텍스트 루트(기본값: analyticserver).

as_tenant

SPSS Modeler 서버 설치가 멤버인 테넌트(기본 테넌트는 ibm).

as_prompt_for_password

SPSS Modeler 서버가 Analytic Server에서 사용하는 사용자 및 비밀번호 인증 시스템과 같은 것으로 구성된 경우 N을 그렇지 않으면 Y를 지정하십시오.

SPSS Modeler를 일괄처리 모드로 실행 중이면 `-analytic_server_username {ASusername}`
`-analytic_server_password {ASpassword}`를 `clemb` 명령의 인수로 추가합니다.

b. SPSS Modeler 서버 서비스를 다시 시작하십시오.

SSL/TLS를 사용하는 Analytic Server 설치에 연결하려면 SPSS Modeler 서버와 클라이언트 설치를 구성하는 추가 단계를 수행해야 합니다.

- a. `http{s}://{HOST}:{PORT}/{CONTEXT-ROOT}/admin/{TENANT}`로 이동하여 Analytic Server 콘솔에 로그인하십시오.
- b. 브라우저에서 인증 파일을 다운로드하여 사용자의 파일 시스템에 저장하십시오.
- c. SPSS Modeler 서버와 SPSS Modeler 클라이언트 설치에서 각각 JRE에 인증 파일을 추가하십시오. 업데이트 위치는 SPSS Modeler 설치 경로의 `/jre/lib/security/cacerts` 서브디렉토리에 있습니다.
 - 1) `cacerts` 파일이 읽기 전용 파일이면 안됩니다.
 - 2) Modeler와 함께 제공된 `keytool` 프로그램을 사용하십시오. 이 프로그램은 SPSS Modeler 설치 경로의 `/jre/bin/keytool` 서브디렉토리에 있습니다.

다음 명령을 수행하십시오.

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

<as-alias>는 `cacerts` 파일의 별명입니다. `cacerts` 파일에 대해 고유한 이름을 임의로 사용할 수 있습니다.

따라서 예제 명령은 다음과 비슷합니다.

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

d. SPSS Modeler 서버와 SPSS Modeler 클라이언트를 다시 시작하십시오.

2. [선택사항] 스트림에서 R 모델을 Analytic Server 데이터 소스와 함께 스코어링하려면 IBM SPSS Modeler - Essentials for R을 설치하십시오. IBM SPSS Modeler - Essentials for R을 다운로드할 수 있습니다 (<https://www14.software.ibm.com/webapp/iwm/web/preLogin.do?source=swg-tspssp>).

제 4 장 마이그레이션

Analytic Server는 사용자가 기존 Analytic Server 설치에서 새 설치로 구성 설정을 마이그레이션하는 여러 시나리오를 지원합니다.

Analytic Server의 새 버전으로 업그레이드

Analytic Server 1.0.1이 이미 설치되어 있으며 2.0을 구매한 경우에는 1.0.1 구성 설정을 2.0 설치로 마이그레이션할 수 있습니다. 이 시나리오에서는 마이그레이션 도구가 적절한 경로를 안내합니다.

제한사항: 1.0 버전이 설치되어 있는 경우, 먼저 버전 1.0에서 1.0.1로 마이그레이션한 다음 버전 1.0.1에서 2.0으로 마이그레이션해야 합니다.

제한사항: 1.0.1 및 2.0 설치는 동일한 Hadoop 클러스터에 공존할 수 없습니다. 1.0.1 설치와 동일한 Hadoop 클러스터를 사용하도록 2.0 설치를 구성한 경우, 1.0.1 설치가 더 이상 작동하지 않습니다.

기존 Analytic Server의 사본 작성

Analytic Server 2.0이 이미 설치되어 있으면 새 2.0 설치에서 구성 설정을 복제할 수 있습니다. 예를 들어, 기존 개발 환경이 있으며 테스트 및 프로덕션 환경을 설정해야 하는 경우에 유용합니다.

1.0.1에서 2.0으로 마이그레이션하는 단계

1. 7 페이지의 『설치』의 지시사항에 따라 Analytic Server의 새 설치를 수행합니다.
2. 설치 프로세스의 마지막에 구성을 마이그레이션하도록 선택합니다. 그러면 자동으로 마이그레이션 도구가 실행됩니다. 마이그레이션 도구로 돌아가기 전에 다음 단계를 완료하십시오.
 - a. 새 설치의 {AS_ROOT}\tools\configcollector 디렉토리를 이전 설치의 해당 위치로 복사합니다.
 - b. {AS_ROOT}\tools\configcollector의 **configcollector** 스크립트를 실행하여 기존 설치에서 구성 수집기 도구를 실행합니다. 그 결과로 생성된 압축(.zip) 파일의 위치를 기록합니다.
 - c. analytic root를 복사합니다. 예를 들어, /user/aeuser/analytic-root의 analytic-root 값을 사용하여 버전 1.0.1을 설치하였으며 /user/aeuser/analytic-root_new의 analytic-root 위치로 마이그레이션하려는 경우, 다음 복사 스크립트를 실행해야 합니다.

```
hadoop distcp /user/aeuser/analytic-root /user/aeuser/analytic-root_new
```
3. 마이그레이션 도구에 새 analytic root의 위치를 제공해야 합니다.
4. 마이그레이션 도구 실행 마지막에 9 페이지의 제 3 장 『구성』의 지시사항에 따라 모든 구성 후 도구 단계를 수행합니다.

참고: 기존 Analytic Server 설치와 함께 사용하도록 R을 구성한 경우, 새 Analytic Server 설치를 사용하여 이를 구성하는 단계에 따라야 합니다.

2.0에서 2.0으로 마이그레이션하는 단계

1. {AS_ROOT}\tools\configcollector의 **configcollector** 스크립트를 실행하여 기존 설치에서 구성 수집 도구를 실행합니다. 그 결과로 생성된 압축(.zip) 파일의 위치를 기록합니다.
2. 7 페이지의 『설치』의 지시사항에 따라 Analytic Server의 새 설치를 수행합니다.
3. 설치 프로세스의 마지막에 구성을 마이그레이션하도록 선택합니다. 그러면 자동으로 마이그레이션 도구가 실행됩니다.
4. 마이그레이션 도구 실행 마지막에 9 페이지의 제 3 장 『구성』의 지시사항에 따라 모든 구성 후 도구 단계를 수행합니다.

참고: 기존 Analytic Server 설치와 함께 사용하도록 R을 구성한 경우, 새 Analytic Server 설치를 사용하여 이를 구성하는 단계에 따라야 합니다.

마이그레이션 도구 사용

마이그레이션 도구는 기존 구성을 새 설치로 마이그레이션하는 기타 여러 수동 단계를 자동화합니다. 마이그레이션 도구에는 HDFS 환경에 대한 액세스 권한이 있어야 합니다.

IBM Installation Manager의 GUI 버전을 사용한 경우 마이그레이션 도구가 자동으로 시작됩니다. 프롬프트에 따라서 시스템을 구성하십시오.

{AS_ROOT}\bin에서 **migrationtool** 스크립트를 실행하여 마이그레이션 도구를 실행할 수도 있습니다. 기본적으로 이 도구는 GUI 모드나 콘솔 모드의 실행 여부를 감지합니다. 인수 **-c**를 전달하여 콘솔 모드를 강제 실행할 수 있습니다.

시작

시작 단계에서는 사용자가 마이그레이션 중인 Analytic Server 구성을 포함하는 압축(.zip) 파일을 제공해야 합니다.

Hadoop 클러스터

Hadoop 클러스터 단계에서는 Hadoop 배포와 함께 사용할 Analytic Server를 구성합니다.

원래 설치와 동일한 Hadoop 클러스터

새 설치가 이전 설치로부터 Hadoop 클러스터의 소유권을 인수하는 "소거식" 복사입니다. 이 방법은 이전 Analytic Server 설치가 필요하지 않은 경우에 사용하는 방법입니다.

다른 Hadoop 클러스터

새 설치가 이전 설치로부터 다른 Hadoop 클러스터로 연결됩니다.

이 초기 선택을 따르는 화면은 구성 도구를 통한 일반적인 경로의 화면과 유사합니다.

Hcatalog/Hive

Hcatalog/Hive 단계는 원래 설치에 HCatalog/Hive 데이터 소스가 있는 경우에만 표시됩니다.

- 동일한 Hadoop 클러스터에 대한 복사 마이그레이션에서는 Hive를 끌 수 있는 옵션이 있습니다.
- 다른 Hadoop 클러스터에 대한 복사 마이그레이션에서는 HCatalog/Hive를 구성하기 위한 구성 도구의 사용자 정의 경로에서 전체 옵션 세트가 제공됩니다.

JDBC

1.0.1에서 2.0으로 마이그레이션하는 경우, 데이터베이스 데이터 소스를 사용할 계획이라면 관리 도구에서 JDBC 드라이버를 새 설치에 추가해야 합니다.

2.0에서 2.0으로 마이그레이션하는 경우, 원래 설치에 사용된 모든 JDBC 드라이버가 새 설치로 마이그레이션됩니다.

애플리케이션 서버

애플리케이션 서버 단계는 새 설치 유형에 따라 달라집니다.

- 고가용성이 아닌 설치에서는 다음 단계를 건너뛵니다. 관리 도구에서 보안 통신(SSL)을 구성하고 Analytic Server에서 청취하는 포트를 설정할 수 있습니다.
- 고가용성 설치에서는 마이그레이션 도구에서 사용자의 WebSphere® Application Server Liberty Network Deployment 환경에 대한 정보를 입력하도록 프롬프트를 표시합니다.

참고: 원래 설치가 고가용성인 경우, 이러한 단계가 원래 설치의 값으로 미리 채워져 있습니다.

Zookeeper

Zookeeper 단계는 고가용성 설치를 위한 구성의 일부로만 필요합니다. Zookeeper 양상블에 서버 목록을 제공하십시오.

참고: 원래 설치가 고가용성인 경우, 이러한 단계가 원래 설치의 값으로 미리 채워져 있습니다.

메타데이터 리포지토리

메타데이터 리포지토리 단계는 새 설치 유형에 따라 달라집니다.

- 고가용성이 아닌 설치에서는 Derby 또는 DB2를 선택할 수 있습니다. DB2를 선택하는 경우, 마이그레이션 도구를 실행하기 전에 설치하여 가용케 해야 합니다. 자세한 정보는 5 페이지의 『필수 소프트웨어』의 내용을 참조하십시오.
- 고가용성 설치의 구성에서는 DB2를 메타데이터 리포지토리로 항상 사용해야 하며 DB2 인스턴스는 구성 도구를 실행하기 전에 설치하여 가용케 해야 합니다.

참고: 원래 설치가 고가용성인 경우, 이러한 단계가 원래 설치의 값으로 미리 채워져 있습니다.

Analytic Server

마이그레이션 도구는 원래 구성을 사용하여 Analytic Server 관리자 역할을 하는 프린시펄을 판별할 수 있습니다. 이 프린시펄은 테넌트를 작성, 삭제 및 관리할 수 있습니다. 관리 도구에서 프린시펄을 변경할 수 있습니다.

요약

요약 단계에서는 마이그레이션을 완료하기 전에 단계를 검토할 수 있습니다.

마이그레이션 후

마이그레이션 후 단계에서는 마이그레이션 성공에 대한 정보를 제공합니다. 구성이 실패한 경우에는 logs 디렉토리에서 로그 파일을 검토하십시오.

제 5 장 시스템 변경하기

관리 도구 사용

관리 도구는 구성 업데이트의 일부 수동 단계를 자동화합니다. 관리자 도구에는 HDFS 환경에 대한 액세스 권한이 있어야 합니다.

{AS_ROOT}\bin에서 **admintool** 스크립트를 실행하여 관리 도구를 실행할 수 있습니다. 기본적으로 이 도구는 GUI 모드나 콘솔 모드의 실행 여부를 감지합니다. 인수 **-c**를 전달하여 콘솔 모드를 강제 실행할 수 있습니다.

요약

시스템에 대한 영구적인 변경을 수행하기 전에는 항상 요약 패널이 표시됩니다. 관리 영역에서 일련의 단계를 완료하고 최종 단추를 누르면 요약 패널이 표시됩니다. 여기서 다른 관리 영역을 선택하여 수정하거나 적용을 클릭하여 변경사항을 시스템에 적용할 수 있습니다.

Hadoop 클러스터

Hadoop 클러스터 단계에서는 Hadoop 배포와 함께 사용할 Analytic Server를 재구성할 수 있습니다.

참고: 관리 도구를 사용하면 Hadoop 벤더가 내부 버전 업데이트를 지원하는 경우에 Hadoop 배포의 버전을 업데이트할 수 있습니다. Hadoop 벤더를 변경하는 경우, 기존 Analytic Server 설치에 대한 설정을 새 Hadoop 배포와 함께 사용하도록 구성된 새 설치로 마이그레이션할 수 있습니다. 자세한 정보는 19 페이지의 제 4 장 『마이그레이션』의 내용을 참조하십시오.

단순 새로 고치기

Hadoop 배포 버전이 변경되었으며 Analytic Server에서 사용되는 JAR 파일만 새로 고치는 경우에 유용합니다.

안내에 따라 새로 고치기

Hadoop 클러스터의 기타 특성이 변경된 경우에 유용합니다. 안내에 따라 새로 고치기를 수행하는 화면은 구성 도구의 화면과 유사합니다.

Hcatalog/Hive

Hcatalog/Hive 단계에서는 Hive 설치와 함께 사용할 Analytic Server를 재구성하거나 기존 Analytic Server 설치를 새 Hive 설치와 함께 사용할 수 있습니다.

기존 Hive 설치가 있는 경우, 구성을 업데이트할 수 있는 두 가지 선택 사항이 있습니다.

단순 새로 고치기

Hive 버전이 변경되었으며 Analytic Server에서 사용되는 JAR 파일만 새로 고치는 경우에 유용합니다.

안내에 따라 새로 고치기

Hive 설치의 기타 특성이 변경된 경우에 유용합니다. 안내에 따라 새로 고치기를 수행하는 화면은 구성 도구의 화면과 유사합니다.

JDBC

사용 중인 데이터베이스의 드라이버를 지정된 위치에 복사하고 관리 단계를 계속하십시오.

애플리케이션 서버

애플리케이션 서버 단계는 설치 유형에 따라 달라집니다.

- 고가용성이 아닌 설치에서는 보안 통신(SSL/TLS)을 구성하고 Analytic Server에서 청취하는 포트를 설정할 수 있습니다.
- 고가용성 설치에서는 관리 도구를 사용하여 사용자의 WebSphere® Application Server Liberty Network Deployment 환경에 대한 모든 변경사항을 기록할 수 있습니다.

참고: 격리 서버에 대해 추가하거나 제거하려면 `ae.cluster.quarantine.server.name` 구성 특성을 편집해야 합니다. 자세한 정보는 25 페이지의 『구성 특성 수동 편집』의 내용을 참조하십시오.

Zookeeper

관리 도구를 사용하면 Zookeeper 앙상블에서 서버 목록을 수정할 수 있습니다.

메타데이터 리포지토리

메타데이터 리포지토리 단계에서는 DB2 인스턴스를 수정할 수 있습니다. 스키마를 변경하는 경우, 사용자가 지정하는 새 스키마가 데이터베이스에 이미 있어야 합니다.

Analytic Server가 메타데이터 리포지토리로 Derby를 사용하도록 구성된 상태에서 DB2로 변경하려면 마이그레이션 도구를 사용하십시오.

Analytic Server

Analytic Server 단계에서는 Analytic Server 관리자 역할을 하는 프린시펄을 변경할 수 있습니다. 이 프린시펄은 테넌트를 작성, 삭제 및 관리할 수 있습니다.

관리 후

관리 후 단계에서는 구성 변경 성공에 대한 정보를 제공합니다. 구성이 실패한 경우에는 logs 디렉토리에서 로그 파일을 검토하십시오.

구성 특성 수동 편집

일부 구성 특성은 수동으로 편집하여 변경할 수 있습니다.

config.properties

config.properties 파일은 {AS_SERVER_ROOT}/configuration 디렉토리에 있습니다. 이 파일을 편집하는 방법은 두 가지입니다.

- GUI 모드로 실행한 경우, 관리 도구를 사용하면 이 파일을 변경할 수 있습니다.
- 그렇지 않으면 파일을 수동으로 편집할 수 있습니다.

고가용성 설치에서는 관리 도구를 통해 변경사항을 config.properties에 푸시할 수 있습니다. 고가용성이 아닌 설치에서는 Analytic Server를 다시 시작해야 합니다.

join.small.data.size

Analytic Server가 맵 방향 알고리즘에 결합하려고 시도할 수 있는 최대 데이터 양(바이트 단위)입니다.

mapred.child.java.opts

Hadoop에서 실행되는 reduce 태스크 및 맵에 대한 JVM 힙 크기를 제어합니다. 클러스터 내의 노드가 처리할 수 있는 크기만큼 이 값을 설정하십시오.

mapred.job.reuse.jvm.num.tasks

맵 및 reduce 태스크에 대해 JVM을 재사용합니다. JVM을 시작하기 위한 오버헤드가 높은 경우, 이 특성의 값을 늘리십시오. 예를 들어, 값이 1로 설정되면 각 맵 및 각 reduce 태스크가 새 JVM에서 실행됩니다. 값이 5로 설정되면 다섯 개의 태스크마다 새 JVM이 작성됩니다.

mmr.threads

인메모리 맵 감소(M3R) 작업에 사용할 스레드의 수입니다.

참고: "mmr" 접두부가 있는 특성은 인메모리 맵 감소를 제어하며 작은 파일을 많이 사용하는 모든 설치에 중요합니다. 시스템을 최적화하기 위해 해당 값을 사용하여 시험해야 합니다.

mmr.upper.bound.threshold

M3R에 의해 처리될 최대 데이터 양입니다. Hadoop에 의해 더 많은 양의 데이터가 처리됩니다.

mmr.synthetic.block.size

파일이 로컬 파일 시스템에 있는 경우 기본 블록 방향입니다.

mmr.diskspill.threshold

이 임계값을 초과하면 중간 M3R 데이터가 디스크에 기록됩니다.

mmr.diskspill.output.buffer

맵퍼 결과 버퍼입니다. `mmr.diskspill.threshold` 미만이어야 합니다.

mmr.taskparallel.targets.threshold

대상/core의 비율이 이 임계값 미만이면 작업이 M3R에 의해 처리됩니다.

node.max.jobs

고가용성 설치에서 클러스터 멤버가 지정된 시간에 실행할 수 있는 Analytic Server 작업의 최대 수입니다.

ae.cluster.quarantine.server.name

고가용성 설치에서 격리 서버로 지정된 서버의 이름입니다.

jvm.options

jvm.options 파일은 {AS_SERVER_ROOT} 디렉토리에 있습니다. 다음 행을 jvm.options에 추가하거나 수정하여 Hadoop이 아니라 Analytic Server를 호스팅하는 서버에서 실행되는 작업에 대한 힙 크기를 설정하십시오. 이는 작은(M3R) 작업을 실행하는 경우에 중요하며 시스템을 최적화하기 위해 해당 값을 사용하여 시험해야 합니다.

```
-Xms512M  
-Xmx2048M
```

고가용성이 아닌 설치에서는 Analytic Server를 다시 시작해야 합니다. 고가용성 설치에서는 jvm.options에 대한 변경사항을 집단에 푸시하는 간단한 방법이 없습니다. 각 클러스터 멤버에서 수동으로 jvm.options를 업데이트해야 합니다.

새 고가용성 클러스터 멤버에 배치

Liberty의 클러스터에 새 서버 멤버를 결합한 후, 다음 명령을 실행하여 새 멤버에 Analytic Server를 배포할 수 있습니다.

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

addClusterMembers 옵션 없이 Analytic Server가 고가용성 클러스터의 모든 멤버에게 재배포됩니다.

제 6 장 설치 제거

1. IBM Installation Manager(GUI 또는 명령행)를 실행하고 IBM SPSS Analytic Server를 설치 제거하는 옵션을 선택한 후 프롬프트의 지시에 따르십시오. IBM Installation Manager는 자동 모드로도 실행할 수 있습니다. 자세한 정보는 IBM Installation Manager 문서를 참조하십시오. http://www-01.ibm.com/support/knowledgecenter/SSDV2W/im_family_welcome.html.

참고: 이 작업은 HDFS의 Analytic Server 폴더에서 데이터를 제거하지 않습니다. 전체 Analytic Server 홈 폴더를 제거하려면 Analytic Server 사용자 아티팩트(입력 및 출력)가 저장된 HDFS 디렉토리를 제거하십시오. 이 디렉토리는 {AS_SERVER_ROOT}/configuration/config.properties에서 `root.folder.path`로 지정됩니다.

제 7 장 문제 해결

이 절에서는 공통 설치 및 구성 문제와 수정 방법에 대해 설명합니다.

고가용성 클러스터

java.net.SocketTimeoutException: 읽기 제한시간

Liberty ND 제한시간 환경 변수를 다음과 같이 변경하십시오.

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

여기서, <milliseconds>는 JMX 읽기 제한시간에 사용할 초 수입니다.

java.io.IOException: CWWKX7202E: ./server start 명령이 만료되는 제한시간 값 60(초)

제어기 서버 server.xml에 다음을 추가하십시오.

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

java.lang.OutOfMemoryError: Java 힙 공간

HA 클러스터의 모든 멤버에서 다음 행을 jvm.options에 추가하십시오.

```
-Xms512M  
-Xmx2048M
```

Zookeeper 트랜잭션 데이터를 관리할 수 없음

Zookeeper 트랜잭션 로그를 자동으로 제거할 수 있도록 설정하려면 zoo.cfg에서 **autopurge.purgeInterval** 매개변수를 1로 설정하십시오.

Analytic 클러스터 서비스와 Zookeeper의 연결이 끊김

zoo.cfg에서 **tickTime**, **initLimit** 및 **syncLimit** 매개변수를 검토하고 수정하십시오. 예를 들면, 다음과 같습니다.

```
# The number of milliseconds of each tick  
tickTime=2000  
# The number of ticks that the initial  
# synchronization phase can take  
initLimit=30  
# The number of ticks that can pass between  
# sending a request and getting an acknowledgement  
syncLimit=15
```

세부사항은 <https://zookeeper.apache.org/doc/r3.3.3/zookeeperAdmin.html>에서 Zookeeper 문서를 참조하십시오.

Analytic Server 작업이 다시 실행되지 않음

Analytic Server 작업이 다시 실행되지 않는 일반적인 상황이 두 가지 있습니다.

1. 클러스터 멤버가 실패하여 Analytic Server 작업이 실패하는 경우에는 일반적으로 다른 클러스터 멤버에서 작업이 자동으로 다시 시작됩니다. 작업이 다시 실행되지 않으면고가용성 클러스터에 네 개 이상의 클러스터 멤버가 있는지 확인하십시오.
2. 클러스터 멤버를 비활성화하면 일반적으로 해당 서버의 모든 Analytic Server 작업이 다른 클러스터 멤버에서 다시 실행됩니다. 작업이 다시 실행되었는지 확인하려면 `-Dcom.spss.ae.remoteclient.failover.threshold=100`을 설정하고 원격 모드를 사용하십시오.

주의사항

이 정보는 미국에서 제공되는 제품 및 서비스용으로 작성된 것입니다.

IBM은 다른 국가에서 이 책에 기술된 제품, 서비스 또는 기능을 제공하지 않을 수도 있습니다. 현재 사용할 수 있는 제품 및 서비스에 대한 정보는 한국 IBM 담당자에게 문의하십시오. 이 책에서 IBM 제품, 프로그램 또는 서비스를 언급했다고 해서 해당 IBM 제품, 프로그램 또는 서비스만을 사용할 수 있다는 것을 의미하지는 않습니다. IBM의 지적 재산을 침해하지 않는 한, 기능상으로 동등한 제품, 프로그램 또는 서비스를 대신 사용할 수도 있습니다. 그러나 비IBM 제품, 프로그램 또는 서비스의 운영에 대한 평가 및 검증은 사용자의 책임입니다.

IBM은 이 책에서 다루고 있는 특정 내용에 대해 특허를 보유하고 있거나 현재 특허 출원 중일 수 있습니다. 이 책을 제공한다고 해서 특허에 대한 라이선스까지 부여하는 것은 아닙니다. 라이선스에 대한 의문사항은 다음으로 문의하십시오.

135-700

서울특별시 강남구 도곡동 467-12, 군인공제회관빌딩

한국 아이.비.엠 주식회사

고객만족센터

전화번호: 080-023-8080

2바이트(DBCS) 정보에 관한 라이선스 문의는 한국 IBM 고객만족센터에 문의하거나 다음 주소로 서면 문의하시기 바랍니다.

Intellectual Property Licensing

Legal and Intellectual Property Law

IBM Japan Ltd.

1623-14, Shimotsuruma, Yamato-shi

Kanagawa 242-8502 Japan

다음 단락은 현지법과 상충하는 영국이나 기타 국가에서는 적용되지 않습니다. IBM은 타인의 권리 비침해, 상품성 및 특정 목적에의 적합성에 대한 묵시적 보증을 포함하여(단, 이에 한하지 않음) 묵시적이든 명시적이든 어떠한 종류의 보증 없이 이 책을 "현상태대로" 제공합니다. 일부 국가에서는 특정 거래에서 명시적 또는 묵시적 보증의 면책사항을 허용하지 않으므로, 이 사항이 적용되지 않을 수도 있습니다.

이 정보에는 기술적으로 부정확한 내용이나 인쇄상의 오류가 있을 수 있습니다. 이 정보는 주기적으로 변경되며, 변경된 사항은 최신판에 통합됩니다. IBM은 이 책에서 설명한 제품 및/또는 프로그램을 사전 통지 없이 언제든지 개선 및/또는 변경할 수 있습니다.

이 정보에서 언급되는 비IBM의 웹 사이트는 단지 편의상 제공된 것으로, 어떤 방식으로든 이들 웹 사이트를 옹호하고자 하는 것은 아닙니다. 해당 웹 사이트의 자료는 본 IBM 제품 자료의 일부가 아니므로 해당 웹 사이트 사용으로 인한 위험은 사용자 본인이 감수해야 합니다.

IBM은 귀하의 권리를 침해하지 않는 범위 내에서 적절하다고 생각하는 방식으로 귀하가 제공한 정보를 사용하거나 배포할 수 있습니다.

(1) 독립적으로 작성된 프로그램과 기타 프로그램(본 프로그램 포함)간의 정보 교환 및 (2) 교환된 정보의 상호 이용을 목적으로 본 프로그램에 관한 정보를 얻고자 하는 라이선스 사용자는 다음 주소로 문의하십시오.

135-700

서울특별시 강남구 도곡동 467-12, 군인공제회관빌딩

한국 아이.비.엠 주식회사

고객만족센터

이러한 정보는 해당 조건(예를 들면, 사용료 지불 등)하에서 사용될 수 있습니다.

이 정보에 기술된 라이선스가 부여된 프로그램 및 프로그램에 대해 사용 가능한 모든 라이선스가 부여된 자료는 IBM이 IBM 기본 계약, IBM 국제 프로그램 라이선스 계약(IPLA) 또는 이와 동등한 계약에 따라 제공한 것입니다.

본 문서에 포함된 모든 성능 데이터는 제한된 환경에서 산출된 것입니다. 따라서 다른 운영 환경에서 얻어진 결과는 상당히 다를 수 있습니다. 일부 성능은 개발 단계의 시스템에서 측정되었을 수 있으므로 이러한 측정치가 일반적으로 사용되고 있는 시스템에서도 동일하게 나타날 것이라고는 보증할 수 없습니다. 또한 일부 성능은 추정을 통해 추측되었을 수도 있으므로 실제 결과는 다를 수 있습니다. 이 책의 사용자는 해당 데이터를 본인의 특정 환경에서 검증해야 합니다.

비IBM 제품에 관한 정보는 해당 제품의 공급업체, 공개 자료 또는 기타 범용 소스로부터 얻은 것입니다. IBM에서는 이러한 비IBM 제품을 반드시 테스트하지 않았으므로, 이들 제품과 관련된 성능의 정확성, 호환성 또는 기타 주장에 대해서는 확인할 수 없습니다. 비IBM 제품의 성능에 대한 의문사항은 해당 제품의 공급업체에 문의하십시오.

IBM이 제시하는 방향 또는 의도에 관한 모든 언급은 특별한 통지 없이 변경될 수 있습니다.

여기에 나오는 모든 IBM의 가격은 IBM이 제시하는 현 소매가이며 통지 없이 변경될 수 있습니다. 실제 판매가는 다를 수 있습니다.

이 정보는 계획 수립 목적으로만 사용됩니다. 이 정보는 기술된 제품이 GA(General Availability)되기 전에 변경될 수 있습니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제에는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 기업의 이름 및 주소와 유사하더라도 이는 전적으로 우연입니다.

이러한 샘플 프로그램 또는 파생 제품의 각 사본이나 그 일부에는 반드시 다음과 같은 저작권 표시가 포함되어야 합니다.

이 정보에는 일상의 비즈니스 운영에서 사용되는 자료 및 보고서에 대한 예제가 들어 있습니다. 이들 예제에는 개념을 가능한 완벽하게 설명하기 위하여 개인, 회사, 상표 및 제품의 이름이 사용될 수 있습니다. 이들 이름은 모두 가공의 것이며 실제 기업의 이름 및 주소와 유사하더라도 이는 전적으로 우연입니다.

이러한 샘플 프로그램 또는 파생 제품의 각 사본이나 그 일부에는 반드시 다음과 같은 저작권 표시가 포함되어야 합니다.

© (귀하의 회사명) (연도). 이 코드의 일부는 IBM Corp.의 샘플 프로그램에서 파생됩니다.

© Copyright IBM Corp. _연도_. All rights reserved.

이 정보를 소프트카피로 확인하는 경우에는 사진과 컬러 삽화가 제대로 나타나지 않을 수도 있습니다.

상표

IBM, IBM 로고 및 ibm.com은 전세계 여러 국가에 등록된 International Business Machines Corp.의 상표 또는 등록상표입니다. 기타 제품 및 서비스 이름은 IBM 또는 타사의 상표입니다. 현재 IBM 상표 목록은 웹 (www.ibm.com/legal/copytrade.shtml)의 "저작권 및 상표 정보"에 있습니다.

Adobe, Adobe 로고, PostScript 및 PostScript 로고는 미국 또는 기타 국가에서 사용되는 Adobe Systems Incorporated의 등록상표 또는 상표입니다.

IT Infrastructure Library는 현재 Office of Government Commerce의 일부인 Central Computer and Telecommunications Agency의 등록상표입니다.

Intel, Intel 로고, Intel Inside, Intel Inside 로고, Intel Centrino, Intel Centrino 로고, Celeron, Intel Xeon, Intel SpeedStep, Itanium 및 Pentium은 미국 또는 기타 국가에서 사용되는 Intel Corporation 또는 그 계열사의 상표 또는 등록상표입니다.

Linux는 미국 또는 기타 국가에서 사용되는 Linus Torvalds의 등록상표입니다.

Microsoft, Windows, Windows NT 및 Windows 로고는 미국 또는 기타 국가에서 사용되는 Microsoft Corporation의 상표입니다.

ITIL은 미국 특허청(U.S. Patent and Trademark Office)에 등록된 The Minister for the Cabinet Office의 등록상표 및 등록 공동체 상표입니다.

UNIX는 미국 및 기타 국가에서 사용되는 The Open Group의 등록상표입니다.

Java 및 모든 Java 기반 상표와 로고는 Oracle 및/또는 그 계열사의 상표 또는 등록상표입니다.

Cell Broadband Engine은 미국 또는 기타 국가에서 사용되는 Sony Computer Entertainment, Inc.의 상표이며 이에 따른 라이선스가 적용됩니다.

Linear Tape-Open, LTO 및 LTO 로고, Ultrium 및 Ultrium 로고는 미국 또는 기타 국가에서 사용되는 HP, IBM Corp. 및 Quantum의 상표입니다.

