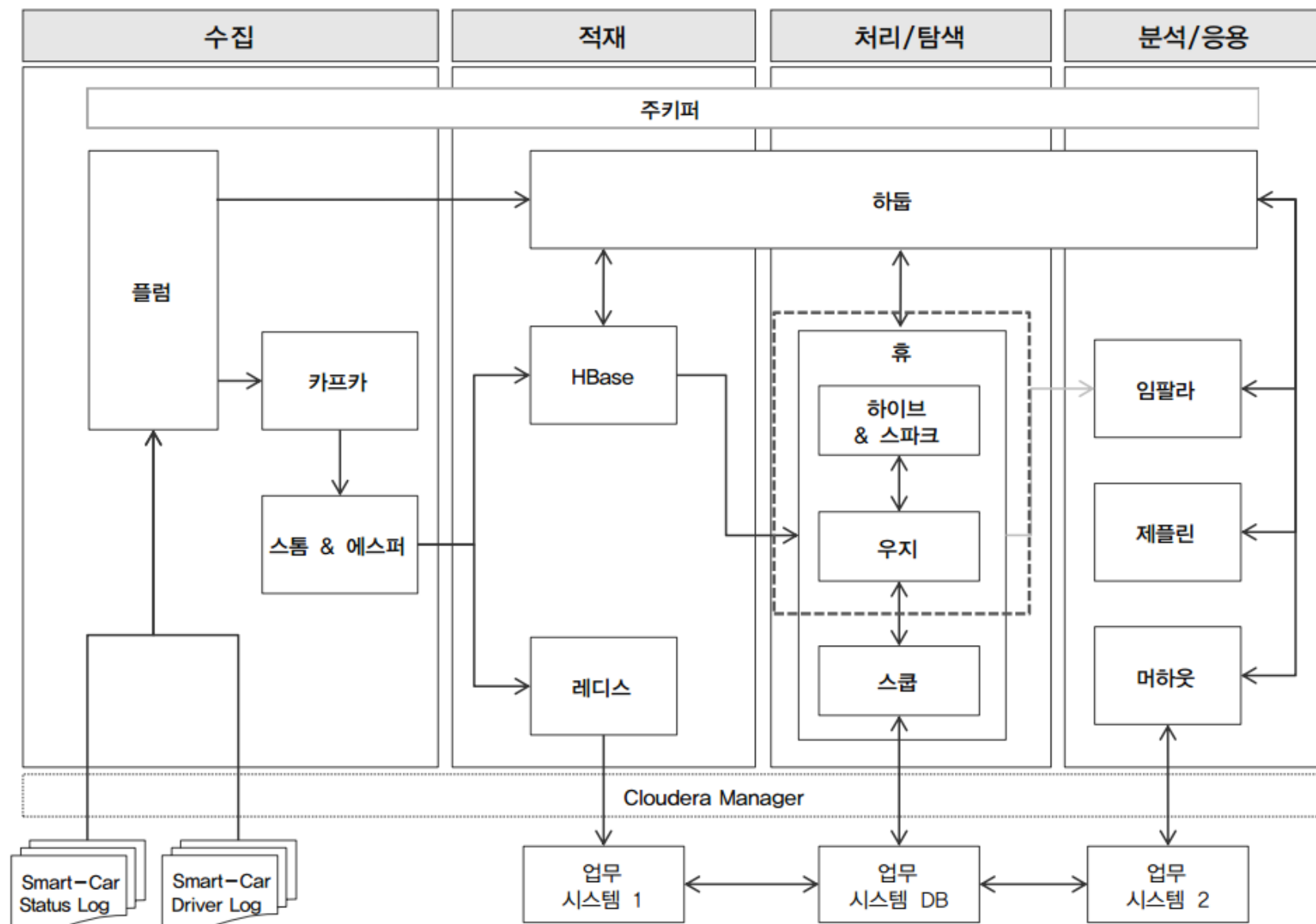


6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성



6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

01. CM의 홈에서 [서비스 추가] 화면으로 이동한다. 추가할 서비스 유형 중 하이브(Hive)를 선택하고 우측 하단의 [계속] 버튼을 클릭한다.

1 Cluster 1

CDH 6.3.2 (Pa 2 서비스 추가

Add Hosts

3개의 호스트

HDFS

YARN (MR2

ZooKeeper

Cloudera Man

Cloudera M

시작	<input type="radio"/>	Flume	Flume은 대부분의 소스에서 데이터를 수집 및 집계하여 HDFS 같은 영구 스토리지로 저장합니다.
중지	<input type="radio"/>	HBase	Apache HBase는 대규모 데이터 세트에 임의의 실시간 읽기/쓰기 액세스를 제공합니다 (HDFS와 ZooKeeper 필요).
재시작	<input type="radio"/>	HDFS	Apache HDFS(Hadoop Distributed File System)는 Hadoop 애플리케이션이 사용하는 기본 스토리지 시스템입니다. HDFS는 데이터 블록에 대한 여러 개의 복제본을 생성하고 이를 클러스터 전반에 걸쳐 컴퓨팅 호스트에 배포하여 안정적이고 매우 빠른 계산을 지원합니다.
클러스터	<input checked="" type="radio"/>	Hive	Hive는 SQL과 유사한 언어인 HiveQL을 제공하는 데이터 웨어하우스 시스템입니다.
클러스터	<input type="radio"/>	Hue	Hue는 CDH(Cloudera Distribution Including Apache Hadoop)에서 작동하는 GUI(그래픽 사용자 인터페이스)입니다(HDFS, MapReduce, Hive 필요).
클러스터	<input type="radio"/>	Impala	Impala에서는 HDFS 및 HBase에 저장된 데이터에 대해 실시간 SQL 쿼리 인터페이스를 제공합니다. Impala에는 Hive 서비스가 필요하며 Hue와 Hive Metastore를 공유합니다.

4

뒤로 계속

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

02. 하이브 작동에 필요한 의존성을 선택한다. HBase가 포함된 항목을 선택한 후 [계속] 버튼을 클릭한다.

Select Dependencies

HBase	HDFS	YARN (MR2 Included)	ZooKeeper
<input type="radio"/> 1	HDFS	YARN (MR2 Included)	ZooKeeper
<input checked="" type="radio"/> HBase	HDFS	YARN (MR2 Included)	ZooKeeper

2

뒤로 계속

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

03. 하이브와 관련된 컴포넌트 중 설치해야 할 프로그램은 4개(Gateway, MetaStore, WebHCat Server, HiveServer2)다. 이 가운데 WebHCat Server는 설치하지 않으며 WebHCat Server를 제외하고 모두 Server02에 설치한다. 그림 6.16처럼 하이브 설치에 필요한 컴포넌트를 배치하고 [계속] 버튼을 클릭한다.

역할 할당 사용자 지정

여기에서 새 서비스에 대한 역할 할당을 사용자 지정할 수 있지만 단일 호스트에 너무 많은 수의 역할을 할당하는 등 올바르지 않게 할당할 경우, 성능이 저하될 수 있습니다.

역할 할당을 호스트별로 볼 수도 있습니다. **호스트별로 보기**

The screenshot shows a configuration window for Hadoop components. It has four main sections, each with a component name and a list of hosts:

- 1 Gateway x ...**: server02.hadoop.com...
- 2 Hive Metastore Server ...**: server02.hadoop.com
- WebHCat Server**: 호스트 선택 (Host Selection)
- 3 HiveServer2 x ...**: server02.hadoop.com

At the bottom right, there are two buttons: **4 뒤로** (Back) and **계속** (Next).

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

04. 하이브의 MetaStore로 사용되는 데이터베이스를 설정한다. 기본값을 그대로 유지하고 테스트 연결이 성공적으로 끝나면 [계속] 버튼을 클릭한다.

데이터베이스 설정

데이터베이스 연결을 구성 및 테스트할 수 있습니다. 설치 가이드 [\[링크\]](#)의 **Installing and Configuring an External Database** 섹션에 설명된 대로 데이터베이스를 먼저 생성하십시오.

☐ 사용자 지정 데이터베이스 사용
 ☒ **내장된 데이터베이스 사용**

1

⚠ The embedded PostgreSQL database is not supported for use in production environments. 내장된 데이터베이스를 사용할 경우 암호가 자동으로 생성됩니다. 복사해 두십시오.

Hive

✓ 건너뛰었습니다. 이후 단계에서 Cloudera Manager가 이 데이터베이스를 생성합니다.

유형	호스트 이름 *	데이터베이스 이름 *	사용자 이름 *
PostgreS ▼	server01.hadoop.com:7432	hive	hive

2

암호 *

5myoMVk13H

3

테스트 연결

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

05. 하이브의 데이터웨어하우스 디렉터리와 메타스토어의 포트 번호를 기본값으로 유지한 상태에서 [계속] 버튼을 클릭한다.

변경 내용 검토

Hive 웨어하우스 디렉터리
hive.metastore.warehouse.dir

Hive(서비스 전체)

/user/hive/warehouse



Hive Metastore Server 포트
hive.metastore.port

Hive Metastore Server Default Group

9083



뒤로

계속

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

06. 하이브의 컴포넌트들이 설치되면서 다양한 환경 구성 정보도 설정되기 시작한다. 설치가 성공적으로 완료되면 [계속] 버튼을 눌러 진행하자.

첫 번째 실행 명령

상태 ☐ 실행 중  Mar 17, 10:48:28 PM

중단

✓ 0/1단계가 완료되었습니다.

☒ Show All Steps

☐ Show Only Failed Steps

☐ Show Only Running Steps

✓ ☐ Run a set of services for the first time
1/2단계가 완료되었습니다.

Mar 17, 10:48:28 PM

> ☐ 6 단계 순차 실행
1/2단계가 완료되었습니다.

Mar 17, 10:48:28 PM

뒤로

계속

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

하이프 설치

07. 설치 및 구성이 완료됐다. [완료] 버튼을 클릭한다.

요약

✔ 새 서비스가 클러스터에 설치 및 구성되었습니다.

i 참고: 여전히 새 서비스를 시작해야 할 수 있습니다. 시작하기 전에 오래된 구성이 포함된 모든 종속 서비스를 재시작하는 것이 좋습니다. 이러한 작업은 아래에서 완료를 클릭하면 주 페이지에서 수행할 수 있습니다.

뒤로

완료

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

미사용 수집/적재 서비스 중지

저사양 파일럿 환경: 미사용 수집/적재 서비스를 일시 중지시킨다.

- 플럼 서비스: CM 홈 → [Flume] → [정지]
- 카프카 서비스: CM 홈 → [Kafka] → [정지]
- 스톰 서비스: Server02에 SSH로 접속한 후 다음 명령을 실행

```
$ service storm-ui stop
```

```
$ service storm-supervisor stop
```


```
$ service storm-nimbus stop
```

- 레디스 서비스: Server02에 SSH로 접속한 후 다음 명령을 실행

```
$ service redis_6379 stop
```

저사양 파일럿 환경에서는 6,7장 탐색/분석 단계 실습 시 사용하지 않는 수집/적재 기능은 중지하고 진행한다.

6.4 탐색 파일럿 실행 2단계 - 탐색 환경 구성

 하이브 설치

실습