다양한 농업 데이터가 쌓이고 있다. 이러한 데이
터들을 이용하여 농민과 농업 관련기업에 가치
있고 유용한 정보를 찾아낼 수 있다. 그러나 이
러한 데이터들을 사용하기 위해서는 우선 엄청
난 양의 데이터들을 수집하고 저장하는 것이 중
요하다. 일부 시스템은 저장할 수 있는 데이터
유형이 제한되어 여러 종류의 농업데이터를 가
공 및 분석하는데 적합하지 않기 때문에, 수집되
는 데이터의 특성을 고려하여 수집 및 저장 구성
요소를 잘 설계해야 한다. 또한, 농업 빅데이터
분석 시 자주 사용되는 기능들에 대한 고려를 통
해, 보다 효과적인 분석을 수행할 수 있도록 플
랫폼의 주요 구성요소들을 설계해야 한다.

레임워크인 HBase와 많은 양의 고속 랜덤 액세

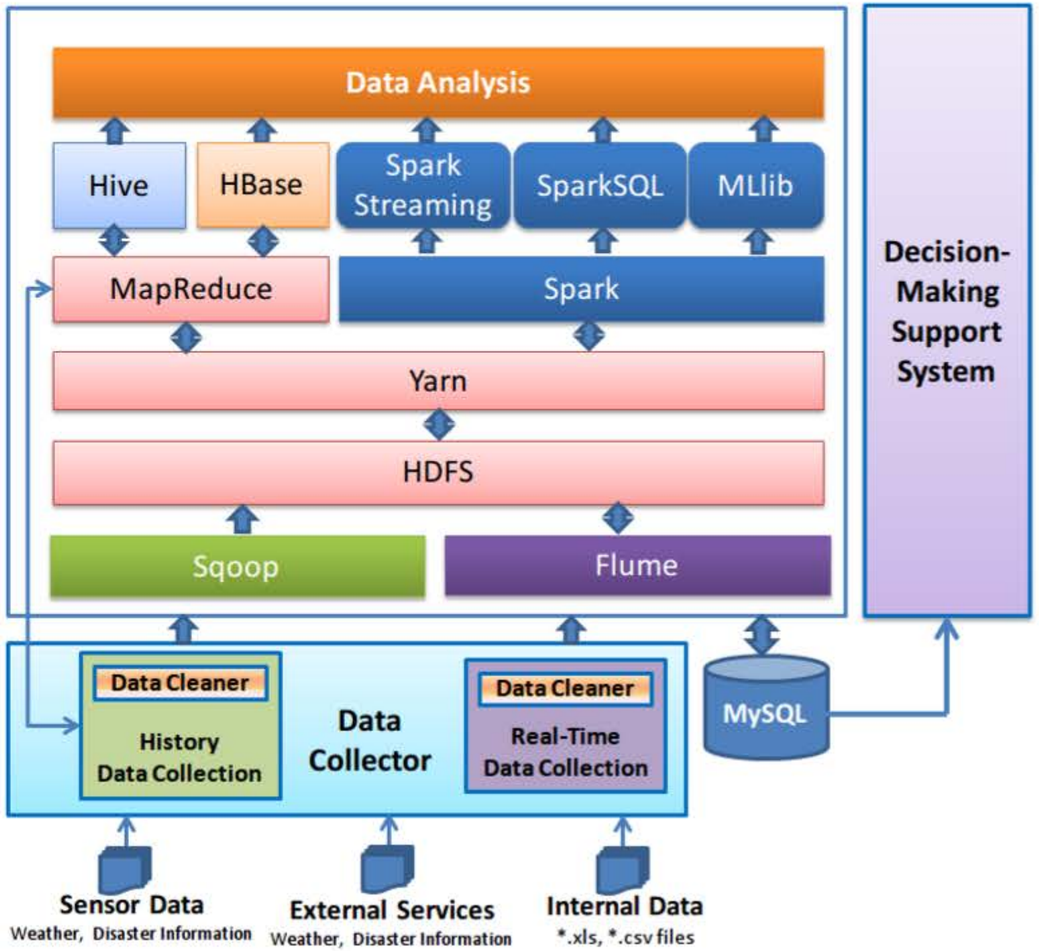
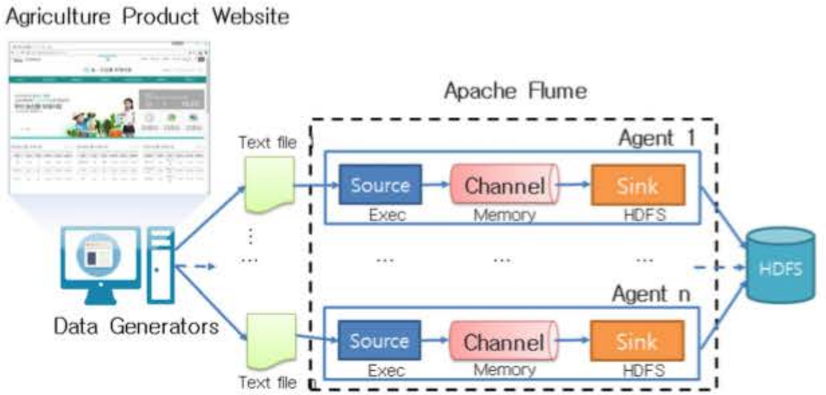
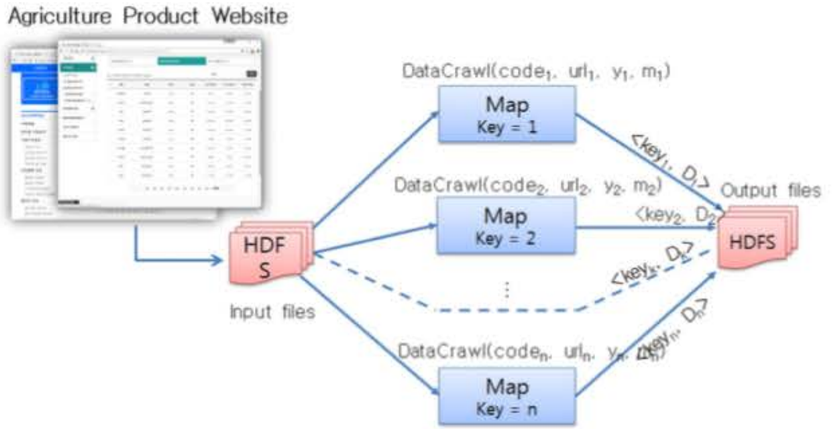


그림 1. 설계된 농업 빅데이터 플랫폼 구조의 개요[3]



Flume-based for real-time collecting agriculture data from the web

그림 2. Apache flume을 이용한 실시간 데이터 수집 및 저장



MapReduce-based for collecting agriculture data from the web

그림 3. 하둡 맵리듀스를 이용한 데이터 수집 모듈

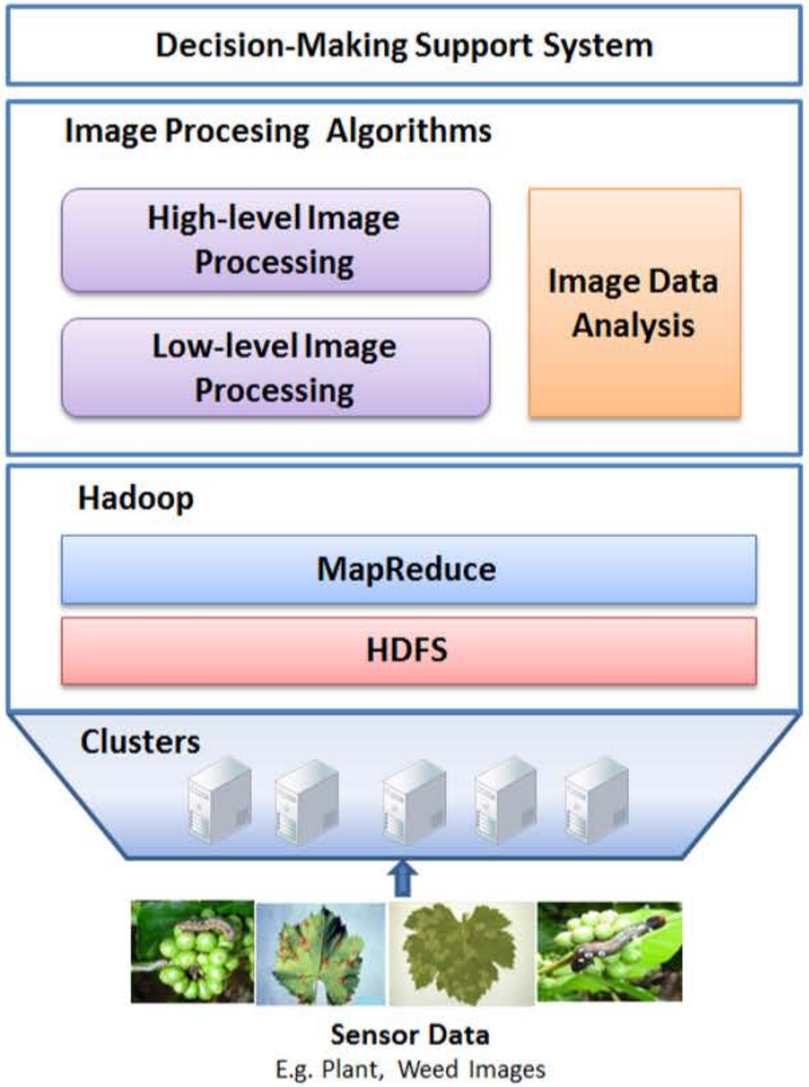


그림 5. 이미지프로세싱을 위한 빅데이터 플랫폼[7]

리즘이다.

위 설명 중 '~를 구매한다면' 에 해당하는 부분을 조건절(Antecedent)라고 한다. 그리고 그 뒷부분을 결과절(Consequent)라고 한다. 그리고 조건절과 결과절을 구성하는 아이템들의 집합을 아이템 집합(Item Set)이라 한다. 예를 들

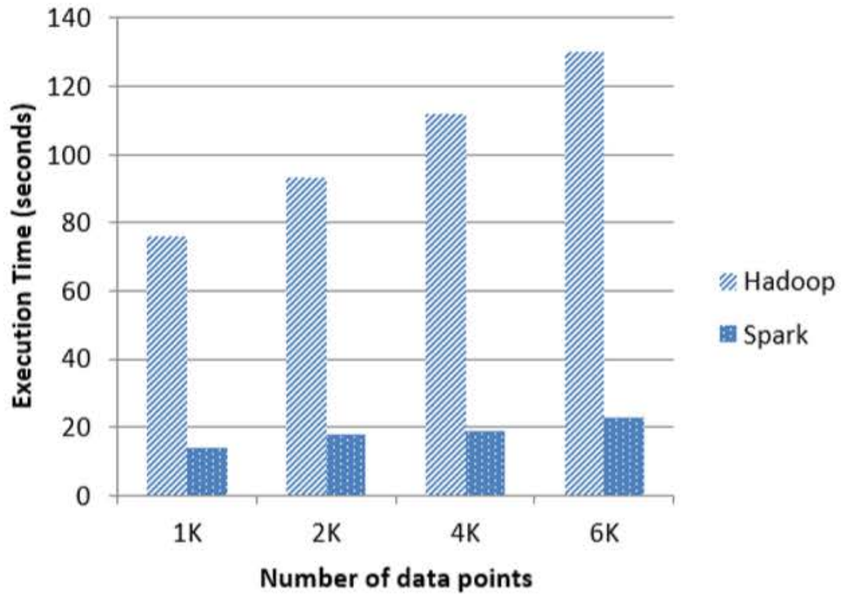
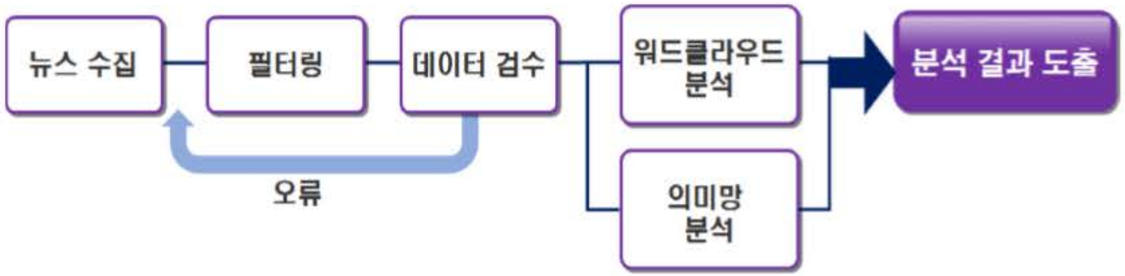


그림 6. 데이터 포인트 개수에 따른 K-mean 알고리즘 수행시간



[그림 2] 빅데이터 분석절차

[그림 3] 데이터 마이닝 절차

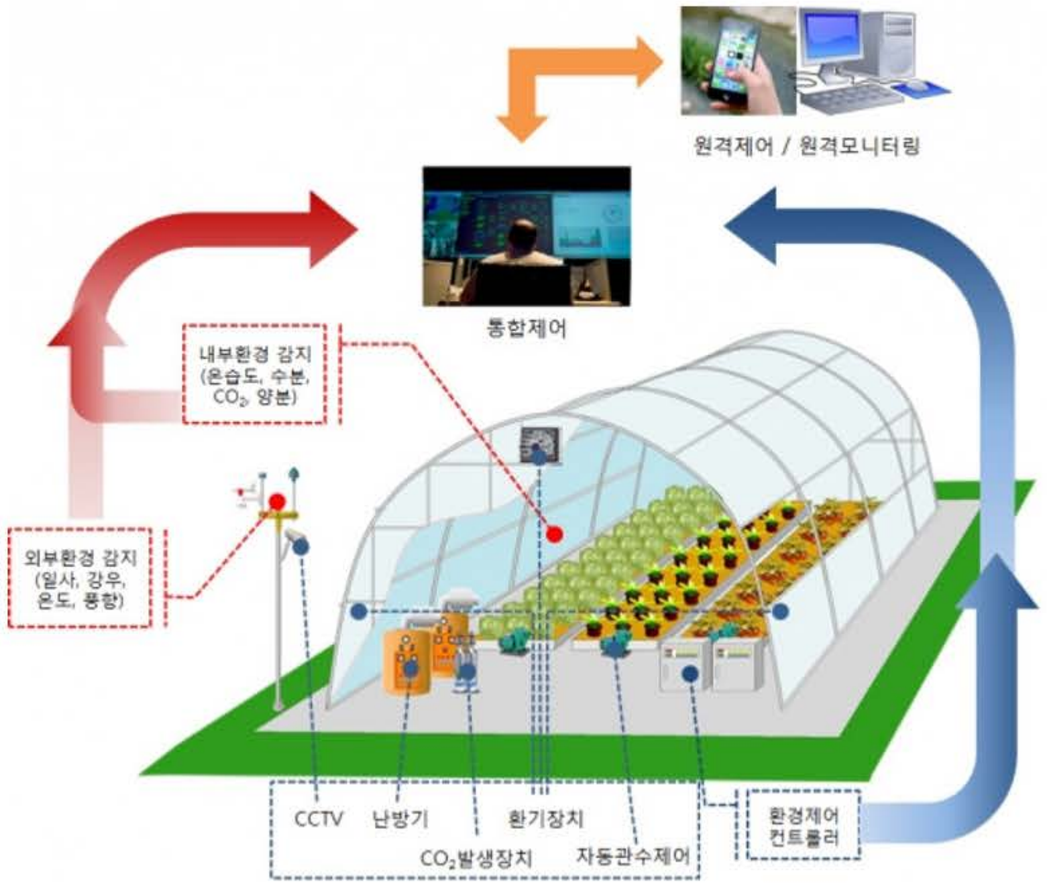


그림 6. 딸기 스마트팜 구성도

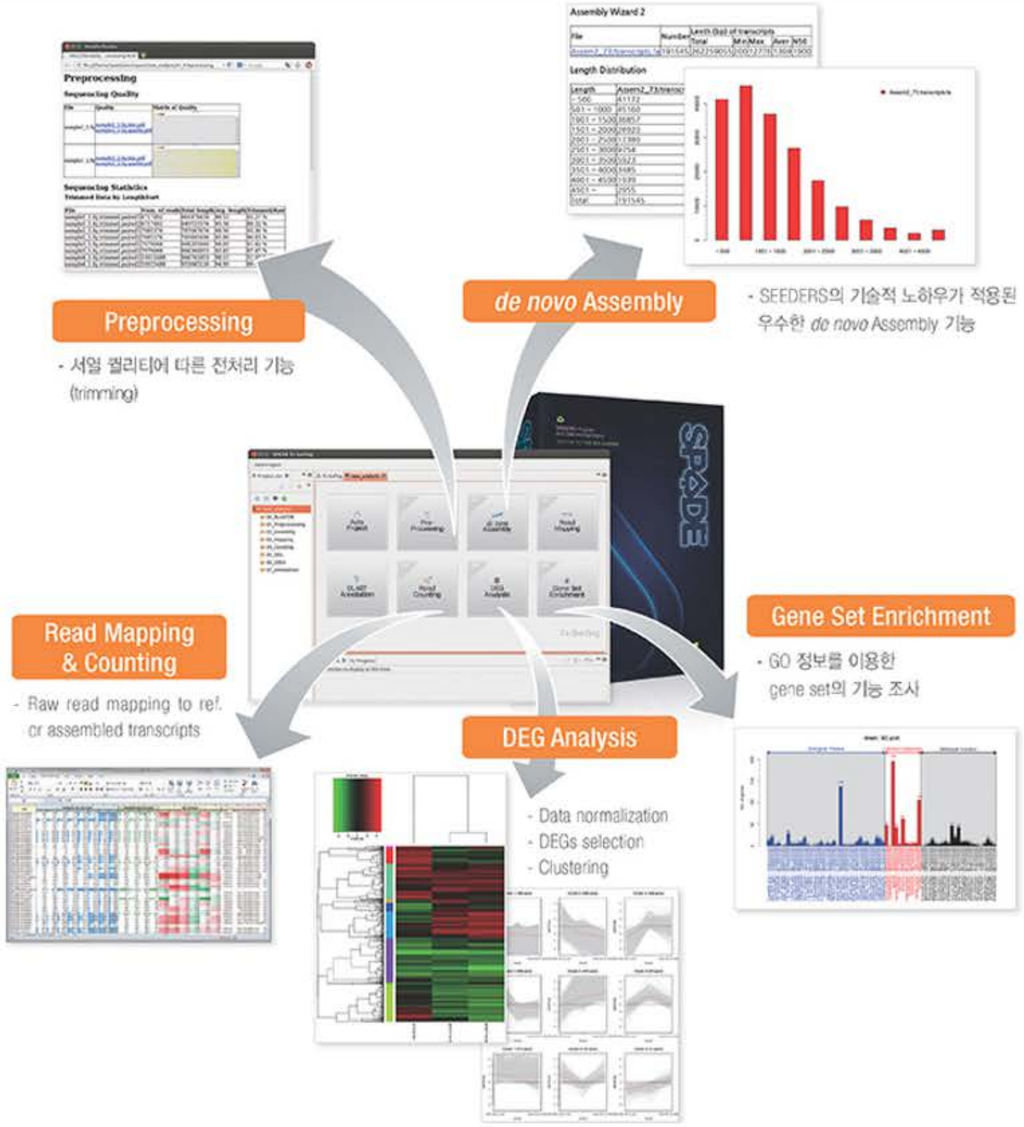


그림 5. 씨더스의 소프트웨어 제품 기능 소개

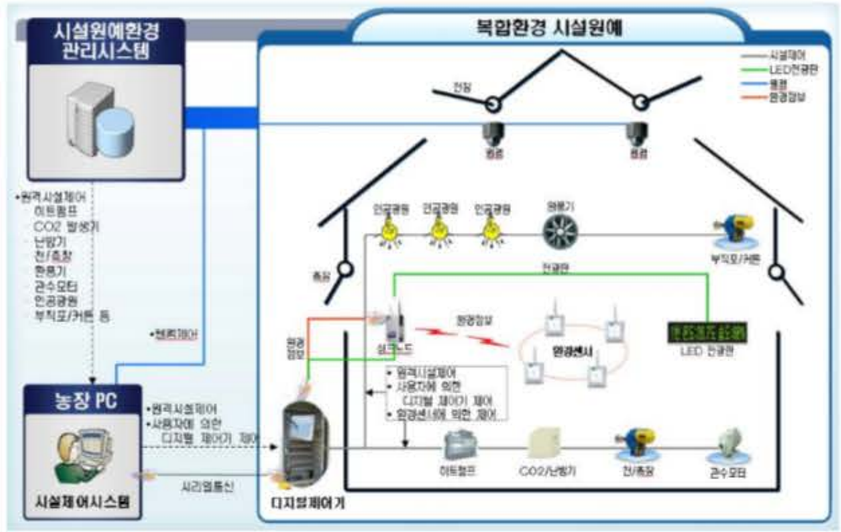


그림 3. 스마트 그린 하우스 시스템 프로세스

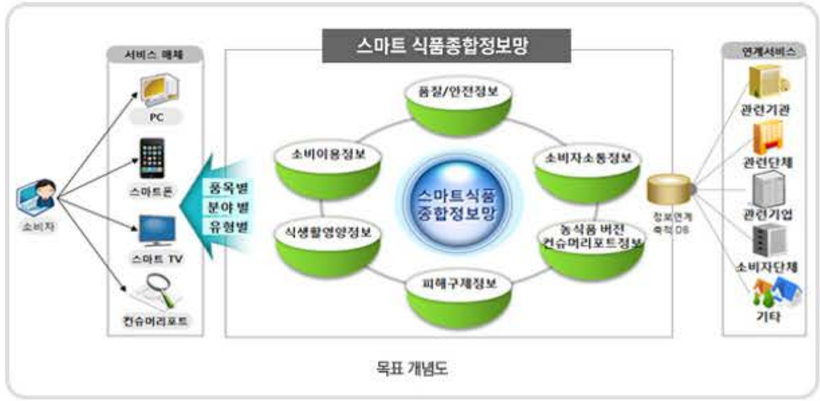


그림 4. 스마트 식품종합정보망 프로세스

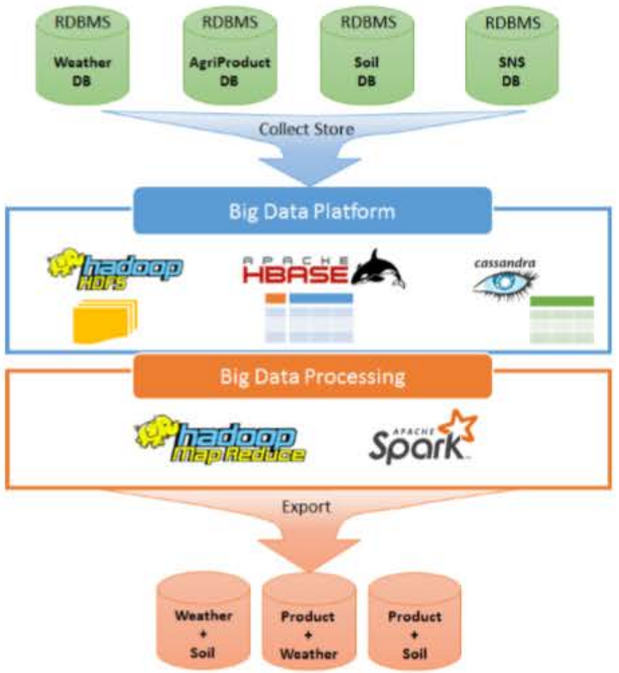


그림 15. 농업 빅데이터 수집저장처리 플랫폼 구성안

그림 16. 농업경영 통합정보 시스템 기능 및 구성안

Learning Goals

In this activity, you will:

- Download and Install VirtualBox.
- Download and Install Cloudera Virtual Machine (VM) Image.
- Launch the Cloudera VM.

Hardware Requirements: (A) Quad Core Processor (VT-x or AMD-V support recommended), 64-bit; (B) 8 GB RAM; (C) 20 GB disk free. How to find your hardware information: Open System by clicking the Start button, right-clicking Computer, and then clicking Properties. Most computers with 8 GB RAM purchased in the last 3 years will meet the minimum requirements.You will need a high speed internet connection because you will be downloading files up to 4 Gb in size.

Instructions

Please use the following instructions to download and install the Cloudera Quickstart VM with VirutalBox before proceeding to the Getting Started with the Cloudera VM Environment video. The screenshots are from a Mac but the instructions should be the same for Windows. Please see the discussion boards if you have any issues.

1. **Install VirtualBox. Go to https://www.virtualbox.org/wiki/Downloads to download and install VirtualBox for your computer. The course uses Virtualbox 5.1.X, so we recommend clicking VirtualBox 5.1 builds on that page and downloading the older package for ease of following instructions and screenshots. However, it shouldn't be too different if you choose to use or upgrade to VirtualBox 5.2.X. For Windows, select the link "VirtualBox 5.1.X for Windows hosts x86/amd64" where 'X' is the latest version.**

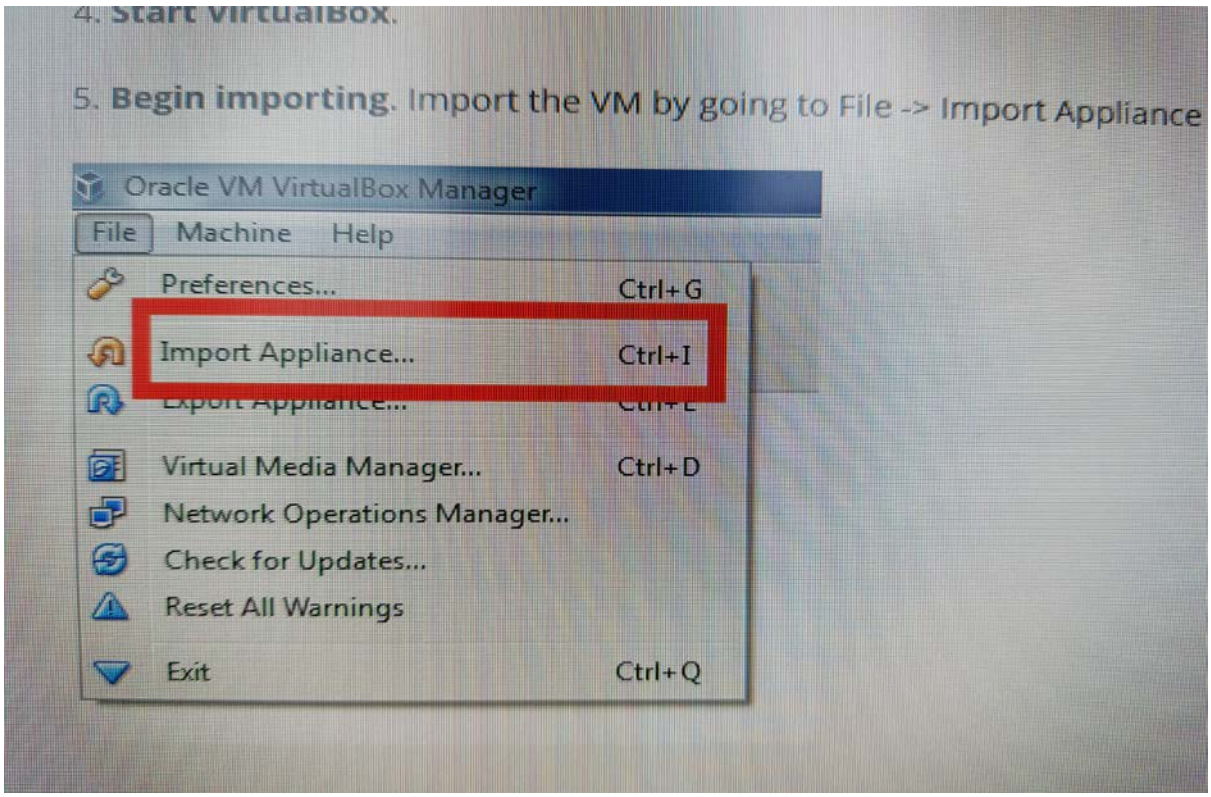
2. **Download the Cloudera VM. Download the Cloudera VM fromhttps://downloads.cloudera.com/demo_vm/virtualbox/cloudera-quickstart-vm-5.4.2-0-virtualbox.zip. The VM is over 4GB, so will take some time to download.**

3. **Unzip the Cloudera VM**:

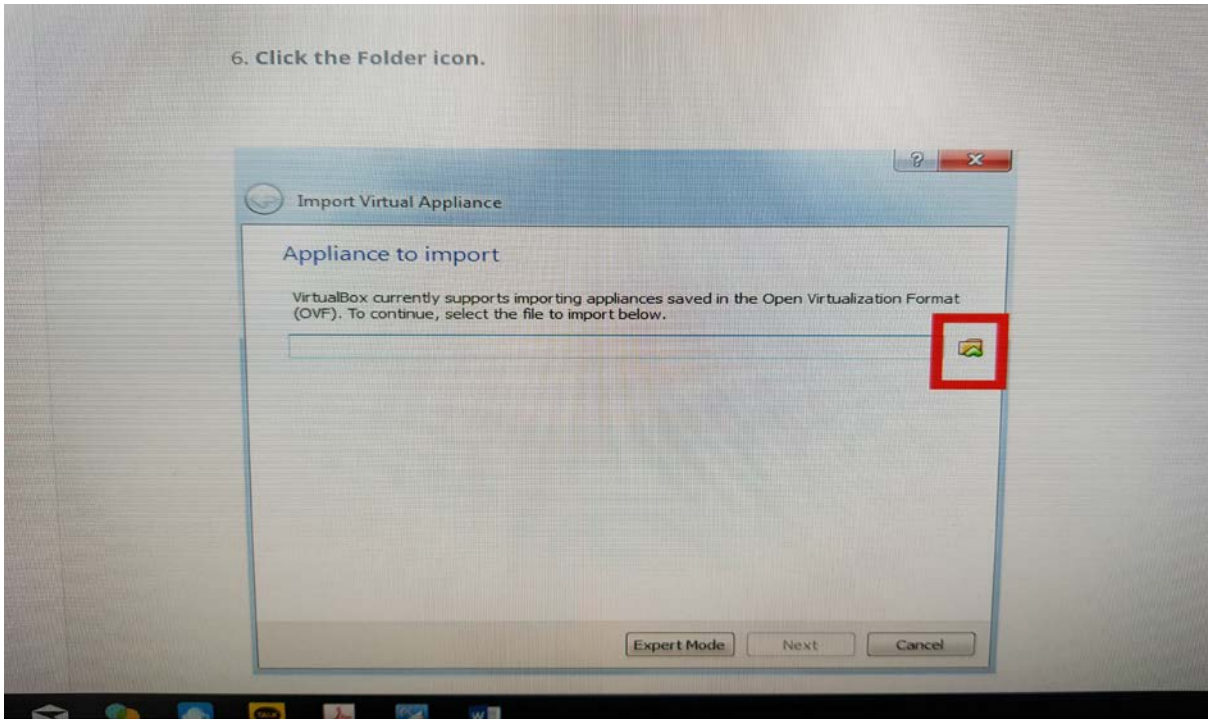
Right-click cloudera-quickstart-vm-5.4.2-0-virtualbox.zip and select “Extract All…”

4. **Start VirtualBox**.

5. **Begin importing**. Import the VM by going to File -> Import Appliance

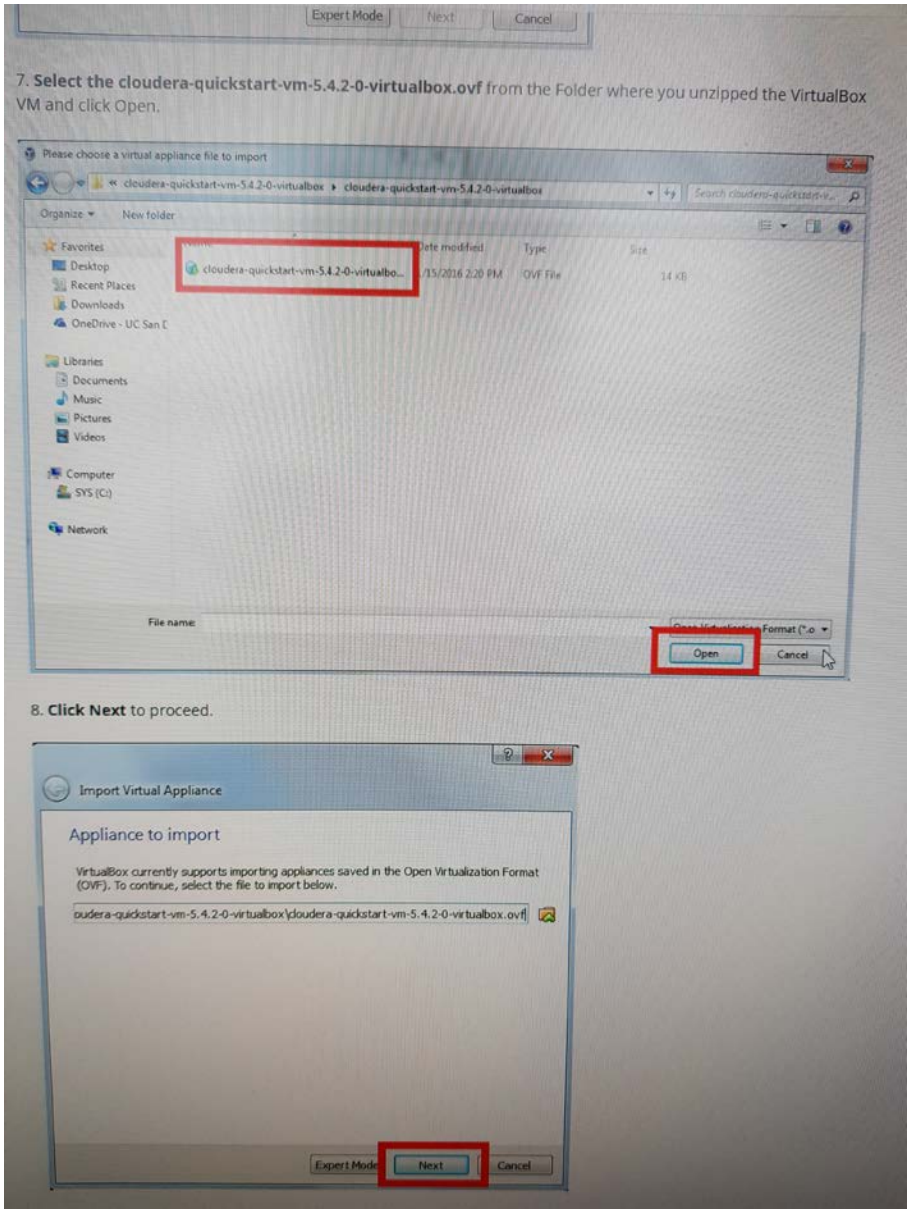


6. **Click the Folder icon.**



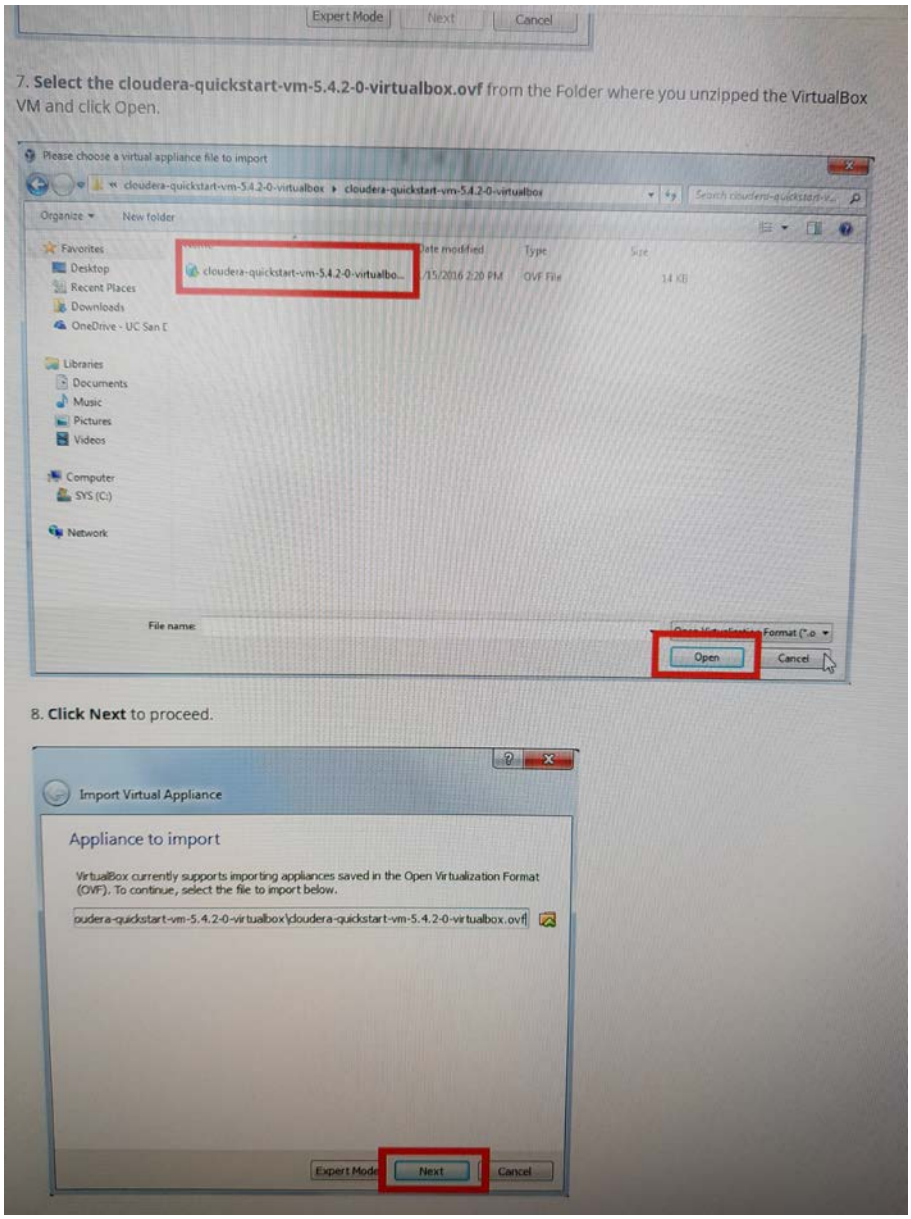
7. **Select the cloudera-quickstart-vm-5.4.2-0-virtualbox.ovf** from the Folder where you unzipped the VirtualBox VM and click Open.

8. **Click Next** to proceed.

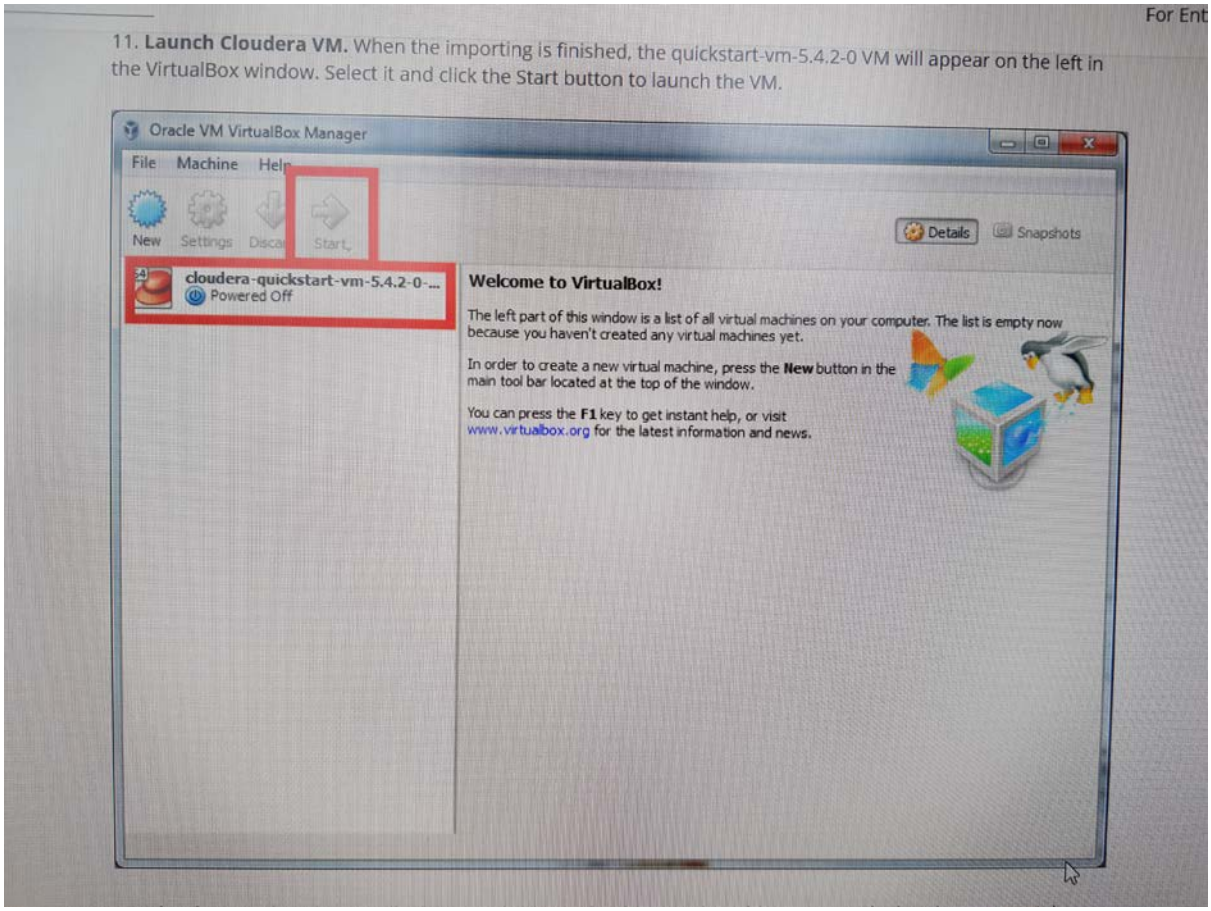


9. **Click Import.**

10. **The virtual machine image will be imported.** This can take several minutes.

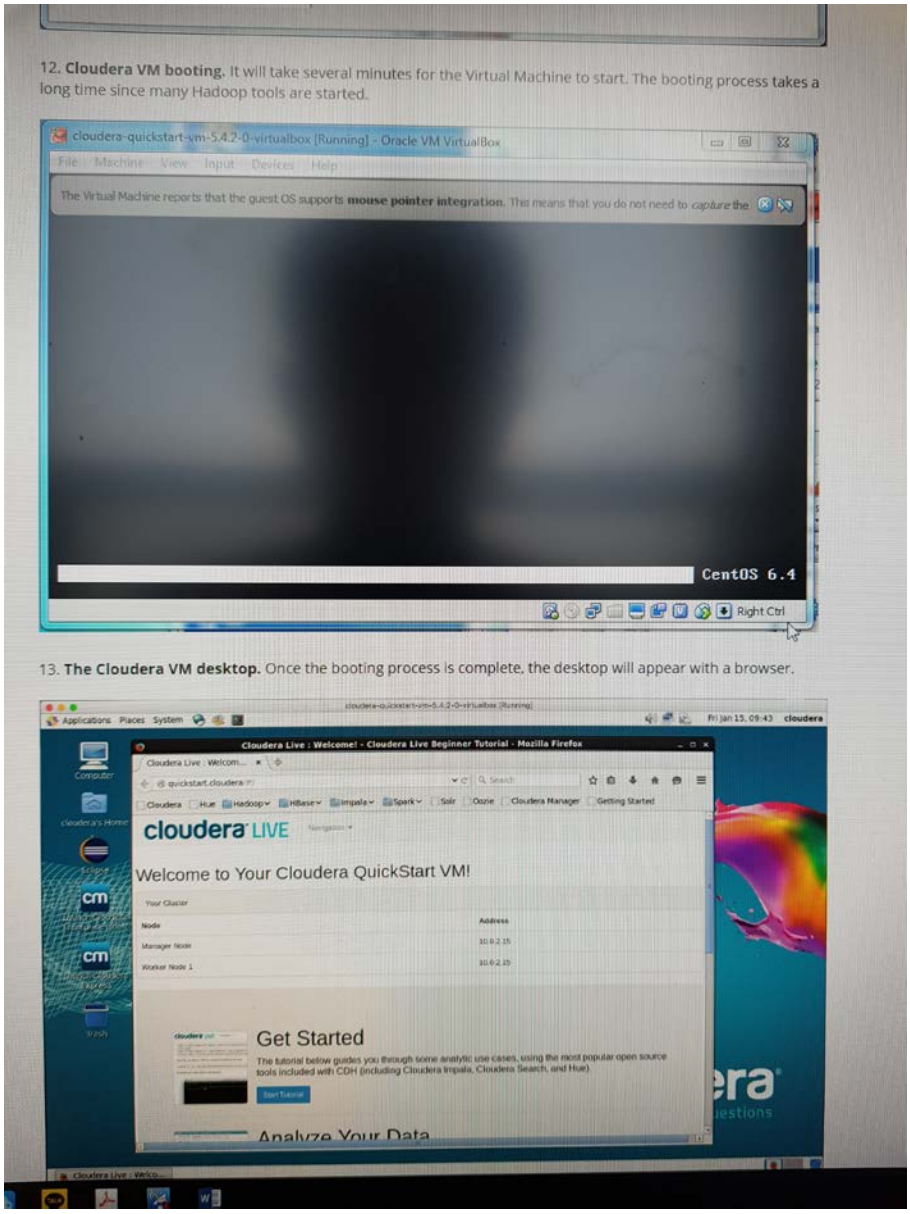


11. **Launch Cloudera VM.** When the importing is finished, the quickstart-vm-5.4.2-0 VM will appear on the left in the VirtualBox window. Select it and click the Start button to launch the VM.



12. **Cloudera VM booting.** It will take several minutes for the Virtual Machine to start. The booting process takes a long time since many Hadoop tools are started.

13. **The Cloudera VM desktop.** Once the booting process is complete, the desktop will appear with a browser.



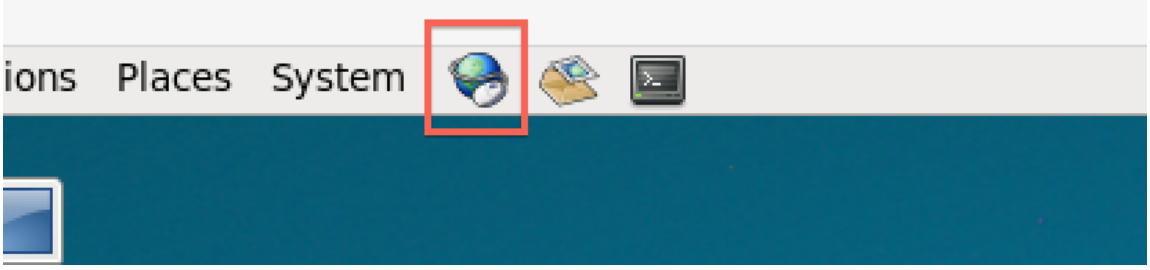
Learning Goals

By the end of this activity, you will be able to:

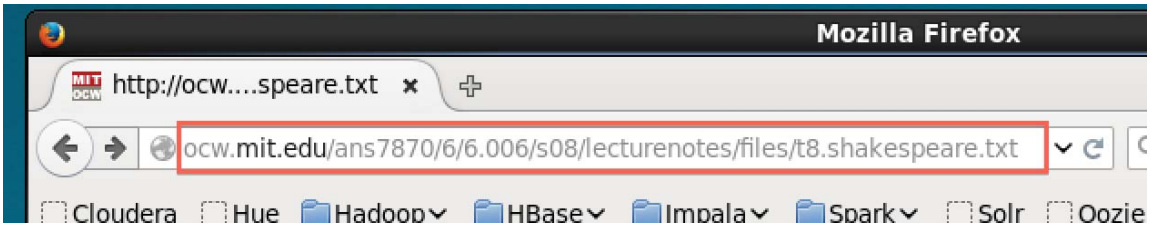
- Interact with Hadoop using the command-line application.
- Copy files into and out of the Hadoop Distributed File System (HDFS).

Instructions

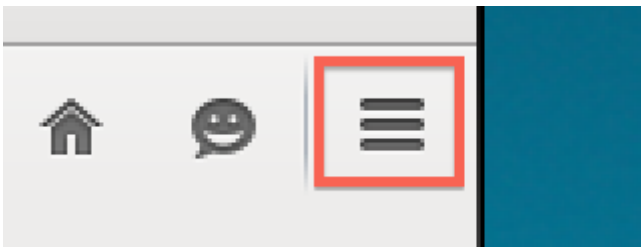
1. **Open a browser.** Open the browser by click on the browser icon on the top left of the screen.



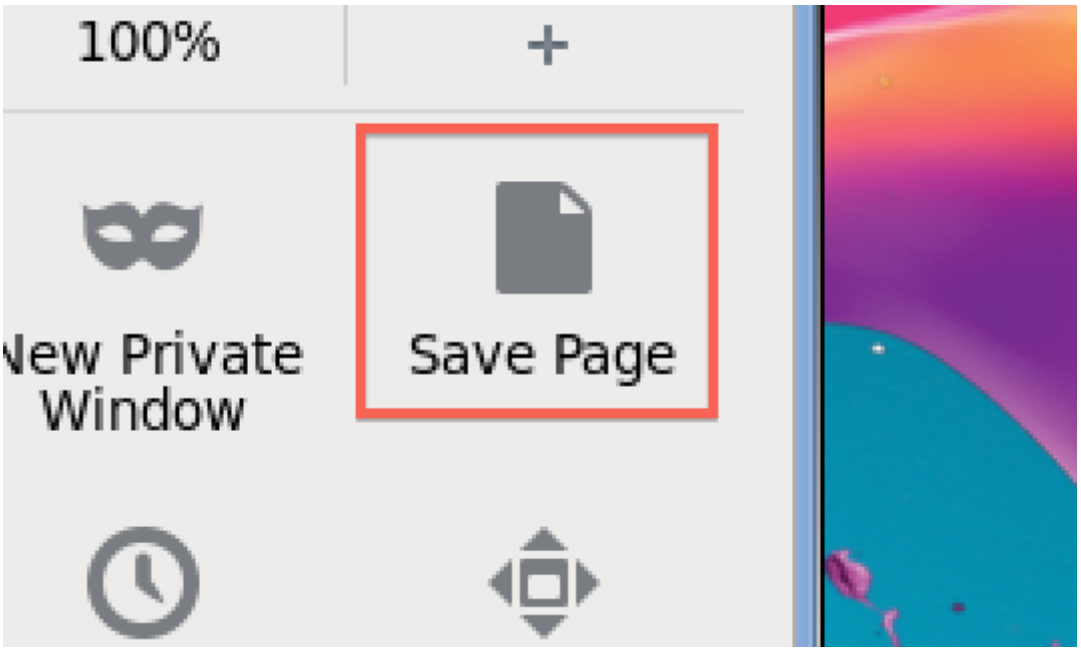
2. Download the Shakespeare. We are going to download a text file to copy into HDFS. Enter the following link in the browser: http://ocw.mit.edu/ans7870/6/6.006/s08/lecturenotes/files/t8.shakespeare.txt



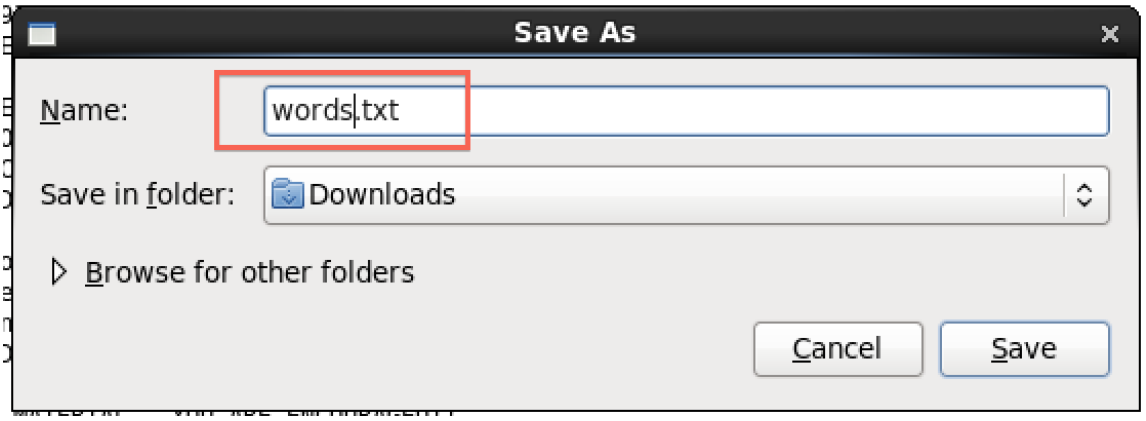
Once the page is loaded, click on the Open menu button.



Click on Save Page



Change the output to words.txt and click Save.



2. Open a terminal shell. Open a terminal shell by clicking on the square black box on the top left of the screen.

Run *cd Downloads* to change to the Downloads directory.

```
[cloudera@quickstart ~]$ cd Downloads/
[cloudera@quickstart Downloads]$
```

Run *ls* to see that words.txt was saved.

```
[cloudera@quickstart Downloads]$ ls
words.txt
```

3. **Copy file to HDFS.** Run *hadoop fs –copyFromLocal words.txt* to copy the text file to HDFS.

```
[cloudera@quickstart Downloads]$ hadoop fs -copyFromLocal words.txt
[cloudera@quickstart Downloads]$
```

4. **Verify file was copied to HDFS.** Run *hadoop fs –ls* to verify the file was copied to HDFS.

```
[cloudera@quickstart Downloads]$ hadoop fs -ls
Found 1 items
-rw-r--r--   1 cloudera cloudera    5458199 2016-02-12 15:14 words.txt
[cloudera@quickstart Downloads]$
```

5. **Copy a file within HDFS**. You can make a copy of a file in HDFS. Run *hadoop fs -cp words.txt words2.txt* to make a copy of words.txt called words2.txt

```
[cloudera@quickstart Downloads]$ hadoop fs -cp words.txt words2.txt
[cloudera@quickstart Downloads]$
```

We can see the new file by running *hadoop fs -ls*

```
[cloudera@quickstart Downloads]$ hadoop fs -ls
Found 2 items
-rw-r--r--   1 cloudera cloudera    5458199 2016-02-12 15:14 words.txt
-rw-r--r--   1 cloudera cloudera    5458199 2016-02-12 15:15 words2.txt
[cloudera@quickstart Downloads]$
```

6. **Copy a file from HDFS.** We can also copy a file from HDFS to the local file system. Run *hadoop fs -copyToLocal words2.txt .* to copy words2.txt to the local directory.

```
[cloudera@quickstart Downloads]$ hadoop fs -copyToLocal words2.txt
[cloudera@quickstart Downloads]$
```

Let's run *ls* to see that the file was copied to see that words2.txt is there.

```
[cloudera@quickstart Downloads]$ ls
words2.txt  words.txt
[cloudera@quickstart Downloads]$
```

7. **Delete a file in HDFS.** Let's the delete words2.txt in HDFS. Run *hadoop fs -rm words2.txt*

```
[cloudera@quickstart Downloads]$ hadoop fs -rm words2.txt
16/02/12 15:17:01 INFO fs.TrashPolicyDefault: Namenode trash configuration: Dele
tion interval = 0 minutes, Emptier interval = 0 minutes.
Deleted words2.txt
[cloudera@quickstart Downloads]$
```

Run *hadoop fs -ls* to see that the file is gone.

```
[cloudera@quickstart Downloads]$ hadoop fs -ls
Found 1 items
-rw-r--r--   1 cloudera cloudera    5458199 2016-02-12 15:14 words.txt
[cloudera@quickstart Downloads]$
```

Mark as completed

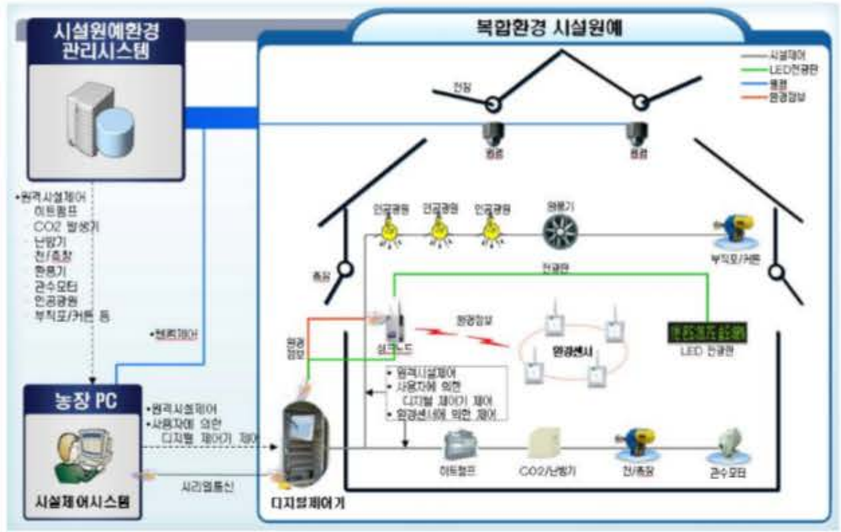


그림 3. 스마트 그린 하우스 시스템 프로세스

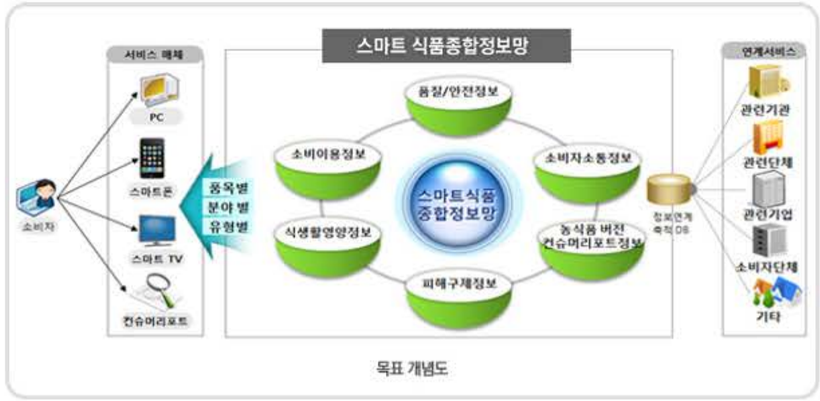


그림 4. 스마트 식품종합정보망 프로세스

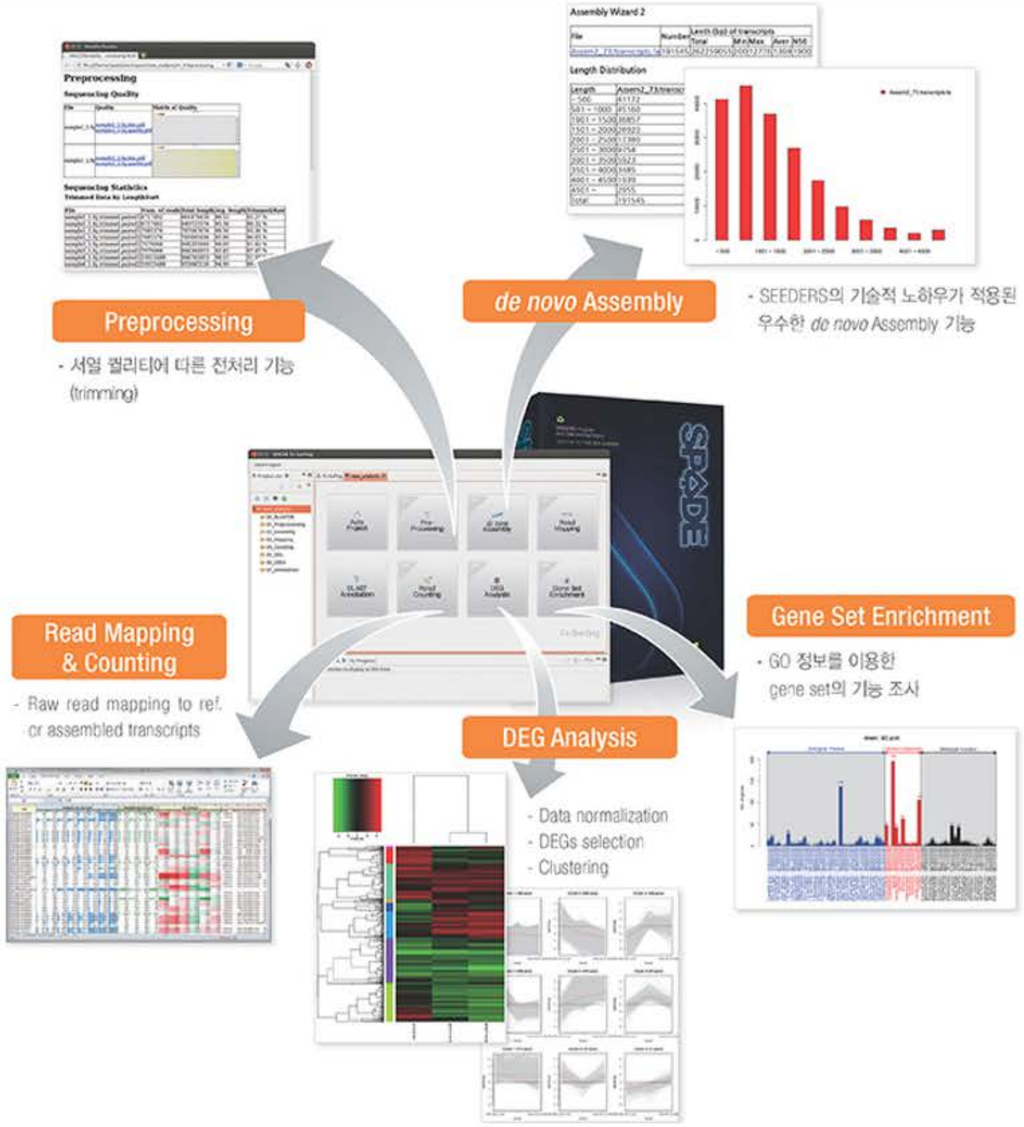


그림 5. 씨더스의 소프트웨어 제품 기능 소개

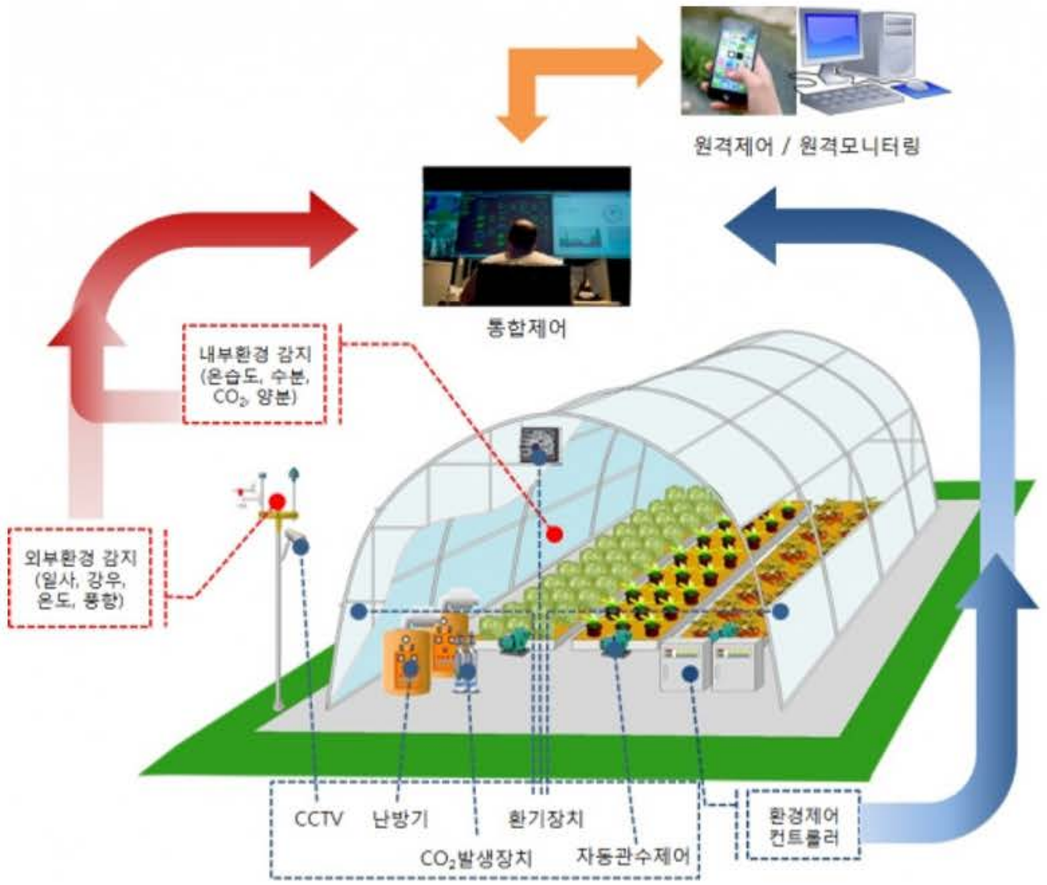


그림 6. 딸기 스마트팜 구성도

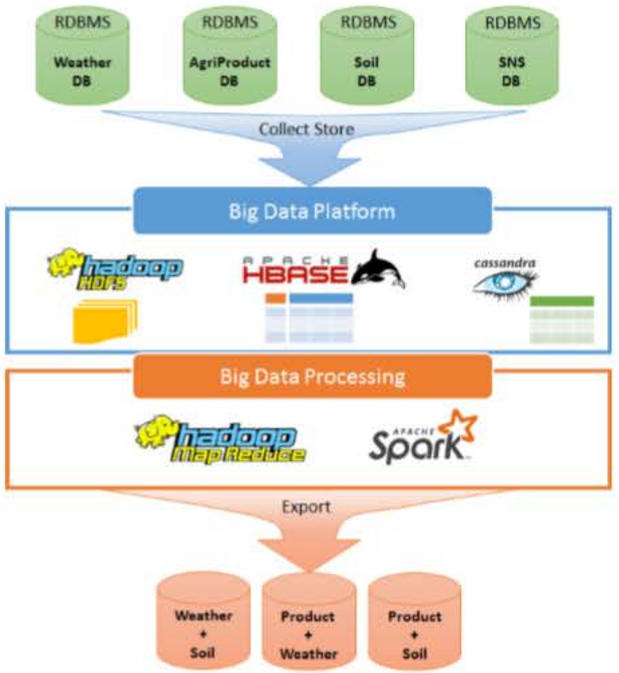


그림 15. 농업 빅데이터 수집저장처리 플랫폼 구성안

그림 16. 농업경영 통합정보 시스템 기능 및 구성안