



Hanbit
RealTime

125

처음 시작하는 Apache Mesos

최규동 지음

확장 가능하고 안정적인
클러스터 환경 구축하기

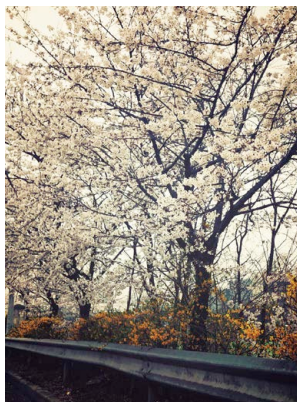


처음 시작하는 Apache Mesos

확장 가능하고 안정적인

최규동 지음

클러스터 환경 구축하기



표지 사진 김학철

이 책의 표지는 김학철님이 보내 주신 풍경사진을 담았습니다.
리얼타임은 독자의 시선을 담은 풍경사진을 책 표지로 보여주고자 합니다.

사진 보내기 ebookwriter@hanbit.co.kr

처음 시작하는 Apache Mesos 확장 가능하고 안정적인 클러스터 환경 구축하기

초판발행 2016년 3월 15일

지은이 최규동 / **펴낸이** 김태현

펴낸곳 한빛미디어(주) / **주소** 서울시 마포구 양화로 7길 83 한빛미디어(주) IT출판부

전화 02-325-5544 / **팩스** 02-336-7124

등록 1999년 9월 30일 제10-1779호

ISBN 978-89-6848-798-9 15000 / **정가** 12,000원

총괄 전태호 / **책임편집** 김창수 / **기획·편집** 정지연 / **교정** 이미연

디자인 표지/내지 여동일, 조판 최송실

마케팅 박상용, 송경석 / **영업** 김형진, 김진불, 조유미

이 책에 대한 의견이나 오타자 및 잘못된 내용에 대한 수정 정보는 한빛미디어(주)의 홈페이지나 아래 이메일로 알려주십시오.

한빛미디어 홈페이지 www.hanbit.co.kr / **이메일** ask@hanbit.co.kr

Published by HANBIT Media, Inc. Printed in Korea

Copyright © 2016 최규동 & HANBIT Media, Inc.

이 책의 저작권은 최규동과 한빛미디어(주)에 있습니다.

저작권법에 의해 보호를 받는 저작물이므로 무단 복제 및 무단 전재를 금합니다.

지금 하지 않으면 할 수 없는 일이 있습니다.

책으로 펴내고 싶은 아이디어나 원고를 메일(ebookwriter@hanbit.co.kr)로 보내주세요.

한빛미디어(주)는 여러분의 소중한 경험과 지식을 기다리고 있습니다.

지은이_ **최규동**

현재 카카오에서 Apache Mesos를 활용한 서비스를 만들고 있습니다. 프로그래밍 언어를 공부하고 새로운 기술을 배우는 것을 좋아하며, 언젠가 제주도 어느 한적한 곳에서 재택근무하는 개발자가 되기를 꿈꾸며 살고 있습니다.

국내에서는 Apache Mesos와 같은 컨테이너 기반의 클러스터 매니저가 이제 막 도입되고 있지만, 전 세계적으로는 이미 일상적으로 사용되는 기술입니다. Docker 가 점점 많이 사용됨에 따라 컨테이너의 편리함과 장점이 알려지면서 많은 개발자가 사용하게 되었고, 이와 더불어 컨테이너를 쉽게 배포하고 관리할 수 있는 Marathon(Mesos 프레임워크), Kubernetes 같은 컨테이너 오케스트레이션 툴이 주목을 받기 시작했습니다.

이 책에서 설명하고 있지만, 특히 Mesos는 계층적이고 유연한 구조 덕분에 하나의 클러스터에 다양한 목적의 여러 프레임워크를 사용할 수 있는 미래의 인프라 플랫폼이라고 할 수 있습니다.

이 책에서는 Mesos의 기본이자 핵심인 Marathon에 대한 설명을 중심으로 다양한 프레임워크를 설명했고, Mesos를 처음 접하는 독자들이 최대한 실제 업무에서 바로 활용할 수 있도록 구성했습니다. 이 책이 많은 개발자에게 새로운 영역에 대한 관심과 재미를 느끼게 해주고, 나아가 실제 서비스 운영에 적용해 좀 더 편하고 안정적으로 서비스를 운영할 수 있게 되기를 희망합니다.

늦었지만 집필을 마쳤습니다. 늦어진 원고에도 집필에 집중할 수 있게 많은 도움을 주신 한빛미디어 정지연 님께 감사의 말씀을 드립니다. 그리고 갓 태어난 아기가 있음에도 집필할 수 있는 시간을 할애해준 사랑하는 아내와 사랑하는 아들에게 감사의 말을 전합니다.



이 책은 Mesos를 처음 접하는 독자를 대상으로 합니다. Docker와 컨테이너에 대한 사전 지식이 있으면 이 책을 이해하는 데 도움이 되지만 반드시 필요한 것은 아닙니다.

이 책에서 사용한 JSON 파일과 Ansible 작성 파일은 다음에서 내려받으면 됩니다.

- <https://github.com/fpgeek/first-mesos>
- <https://github.com/fpgeek/mesos-cluster-playbook>

chapter 1 Apache Mesos — 009

- 1.1 데이터 센터와 클러스터 매니저 — 009
- 1.2 Mesos란 — 011
- 1.3 Mesos 아키텍처 — 011
 - 1.3.1 마스터와 슬레이브 — 012
 - 1.3.2 프레임워크 — 012
- 1.4 개발 환경의 Mesos 설치하기 — 013

chapter 2 Marathon — 017

- 2.1 Marathon이란 — 017
- 2.2 Marathon을 이용한 애플리케이션 실행 — 018
 - 2.2.1 일반 애플리케이션 실행 — 018
 - 2.2.2 Docker 컨테이너 실행 — 020
 - 2.2.3 인스턴스 조절 — 022
 - 2.2.4 애플리케이션 업데이트와 롤백 — 023
 - 2.2.5 패일오버 — 027
- 2.3 서비스 디스커버리와 로드밸런싱 — 028
 - 2.3.1 marathon-lb — 029
 - 2.3.2 Mesos-DNS — 033
- 2.4 REST API — 040
- 2.5 Docker 컨테이너 — 050
 - 2.5.1 동적 포트 매핑 — 052
 - 2.5.2 정적 포트 매핑 — 054
 - 2.5.3 항상 Docker 이미지 Pull하기 — 054
 - 2.5.4 cmd vs args — 054
 - 2.5.5 Privileged 모드와 추가 Docker 옵션 — 055

2.6	제약	056
2.6.1	필드	056
2.6.2	연산자	057
2.7	Health Checks	058
2.8	Event Bus	061
2.9	애플리케이션 그룹	062
2.9.1	의존 관계	063
2.9.2	그룹 단위 스케일링	063

chapter 3 Chronos — 065

3.1	Chronos란	065
3.2	예제 실행해 보기	065
3.3	REST API	067
3.3.1	주요 REST API	067
3.3.2	스케줄링	068
3.3.3	Docker 작업 실행	070
3.3.4	의존 관계 설정	072

chapter 4 Spark — 075

4.1	Spark란	075
4.2	클러스터 아키텍처	076
4.3	Spark 셀	078
4.4	Mesos에서 Spark 활용하기	080
4.4.1	준비	080
4.4.2	클라이언트 모드	080
4.4.3	클러스터 모드	083



4.4.4 실행 모드	085
4.4.5 주요 설정값	086

chapter 5 Hadoop on Mesos 089

5.1 HDFS	089
5.1.1 HDFS 네임 노드 HA	090
5.1.2 HDFS 프레임워크	092
5.2 맵리듀스	097

chapter 6 Mesos 클러스터 구축 105

6.1 Mesos와 Marathon 클러스터 구축	105
6.1.1 마스터 노드 설치	106
6.1.2 슬레이브 노드 설치	110
6.2 Ansible을 이용한 Mesos와 Marathon 클러스터 구축	112
6.2.1 Ansible 설치	113
6.2.2 Playbook 실행	113
6.2.3 프레임워크 설치	114

Apache Mesos

1.1 데이터 센터와 클러스터 매니저

IT에 종사하지 않는 사람도 ‘빅데이터’라는 말을 한 번쯤은 들어 봤을 것이다. 많은 데이터를 저장하고 가공하여 새로운 정보로 활용하는 것이 현대 IT 기업의 일반적인 모습이 되었다. 데이터는 점점 늘어나는데, 이를 처리하기 위해 더 많은 하드웨어를 구입한다면 비용과 시간이 많이 들고 자원은 비효율적으로 사용하게 된다. 그래서 많은 IT 기업에서는 더 많은 하드웨어를 구입하는 대신 기존 데이터 센터 안의 서버로 손쉽게 자원을 확보하여 안정적으로 서비스를 유지하고 싶어 한다.

우선 전통적인 방식의 데이터 센터 운영을 살펴보자. 전통적인 데이터 센터의 모습은 서버들을 정적으로 나누어 하나의 애플리케이션을 하나의 서버에 설치하여 사용하는 방식이다. 서버 자원을 효율적으로 사용하려고 가상 기술을 이용하여 하나의 물리 머신을 여러 개의 가상 머신^{VM, Virtual Machine}으로 나누어 사용한다. 여러 대의 VM에 필요한 애플리케이션을 설치하고, 더 많은 데이터를 처리해야 하면 또 다른 VM을 생성하여 애플리케이션을 설치한다. 일반적으로 하나의 VM에 하나의 애플리케이션을 설치하는데, 결국 VM을 사용하더라도 자원을 정적으로 나누어 사용하는 방식에는 변함이 없다.

이와 같이 자원을 정적으로 나누어 사용하는 데이터 센터 운영 방식에는 몇 가지 문제점이 있다.

- **복잡한 관리** 애플리케이션을 늘리거나 줄일 경우 서버에 직접 접속하여 작업해야 하고, 사용하는 서버와 여분의 서버를 항상 관리하고 있어야 한다.
- **높은 장애 발생 가능성** 애플리케이션이 설치된 서버 중 한 대의 서버에서 문제가 발생했을 때 그에 대한 대비가 없다면, 한 대의 서버에서 발생한 문제가 곧 장애로 이어지게 된다. 또한, 정적으로 나누어진 자원을 사용하는 환경에서 이러한 문제에 대비하는 것은 복잡한 면이 있다.
- **낮은 자원 효율성** 일반적으로 애플리케이션은 서버의 모든 자원을 최대한으로 활용하지 못하는데, 이 경우 낭비되는 자원으로 인하여 자원 효율성이 떨어진다. 비효율적인 자원 활용은 많은 비용이 발생하게 하므로 기업 측면에서는 굉장히 중요한 문제다.
- **개발자에 친화적 운영 환경** 서버를 관리하고 운영하는 것이 개발자에게 친화적이지 않다. 최근에는 개발자가 운영까지 하는 DevOps(데브옵스)의 시대이기 때문에 운영도 개발자 친화적으로 변경되어야 한다.

이제 전통적인 운영 방식은 많은 양의 데이터를 안정적으로 다루기에 더는 적합하지 않다. 우리가 익히 알고 있는 Twitter, Airbnb, Netflix, Apple과 같은 기업 뿐 아니라 국내외 다수의 기업이 클러스터 매니저를 활용하여 서버를 운영하는 방식으로 변화하고 있다. 클러스터 매니저는 물리적으로 나누어진 다수의 컴퓨터를 묶어 대규모 자원을 가진 하나의 컴퓨터로 취급하는 방식으로, 클러스터 매니저를 사용하면 다음과 같은 장점이 있다.

- **효율성 증가** 여러 애플리케이션이 자원을 공유하여 사용하므로 전통적인 운영 방식보다 효율적이다.
- **자원 독립성 확보** 자원을 공유하여 사용하기 위해서는 자원 간의 독립성을 확보하여야 한다. 자원 간 독립성을 확보함으로써 인하여 동일한 서버에서 여러 개의 애플리케이션을 실행하고 있어도 다른 애플리케이션이 내가 실행하고 있는 애플리케이션에 영향을 미치지 않는다.
- **확장성 용이** 필요할 때 필요한 만큼 손쉽게 자원을 확보할 수 있다.
- **더욱 견고한 시스템** 특정 서버에 발생한 장애가 서비스 장애로 이어지지 않고, 정상적으로 실행되는 다른 서버로 애플리케이션을 이동하여 계속 실행한다.

1.2 Mesos란

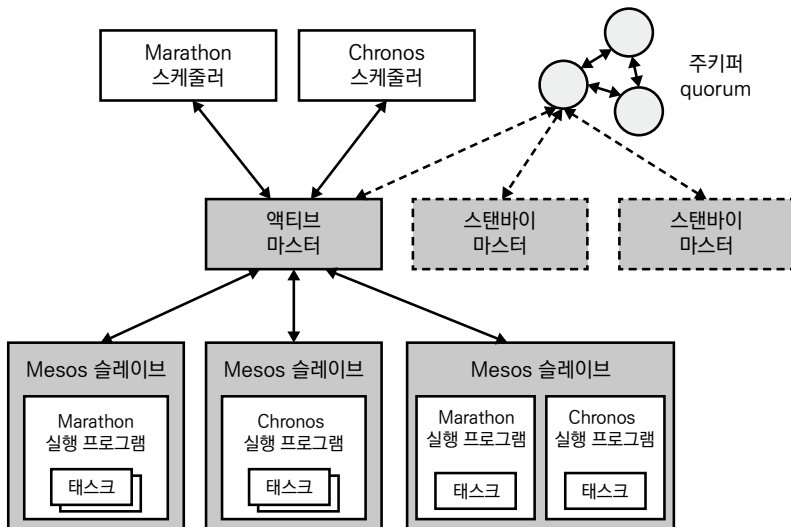
Mesos는 2009년 버클리 대학에서 시작된 오픈소스 클러스터 매니저 프로젝트로, 2011년에는 아파치 인큐베이션 프로젝트에 포함되었고 2013년에는 아파치 최상위^{Top-level} 프로젝트로 승격되었다. 이미 Twitter와 Airbnb, Apple을 포함한 많은 기업이 실 서비스 환경에서 Mesos를 사용하고 있다.

리눅스 커널이 한 대의 컴퓨터에서 다수의 애플리케이션이 자원을 공유하여 사용할 수 있게 컴퓨터 자원을 관리하는 것과 마찬가지로, Mesos는 다수의 컴퓨터를 묶어 하나의 큰 컴퓨터로 사용할 수 있도록 자원을 관리해 준다. 그래서 보통 Mesos를 ‘데이터 센터의 커널’이라고 한다.

1.3 Mesos 아키텍처

Mesos는 [그림 1-1]과 같이 마스터^{Master}, 슬레이브^{Slave}, 프레임워크^{Framework}로 구성되어 있다. 각 구성 요소에 대해 살펴보자.

그림 1-1 Mesos 아키텍처



1.3.1 마스터와 슬레이브

마스터는 슬레이브와 프레임워크를 연결하는 역할을 한다. 마스터는 각 슬레이브에서 보내는 CPU, 메모리, 디스크, 포트와 같은 자원 현황을 프레임워크에 전달하고, 프레임워크는 마스터를 통해 태스크^{Task} 실행을 요청한다. 주키퍼^{ZooKeeper}를 이용하여 1대의 액티브 마스터^{Active Master}, 2대 이상의 스탠바이 마스터^{Standby Master}를 두어 마스터 장애에 대비하는 것이 일반적이다.

슬레이브는 Mesos 클러스터에서 실제로 태스크가 실행되는 곳이다. 슬레이브는 각자의 CPU, 메모리, 디스크, 포트 등의 자원을 가지고 있고, 마스터를 통해 자원 정보를 프레임워크에 전달한다.

1.3.2 프레임워크

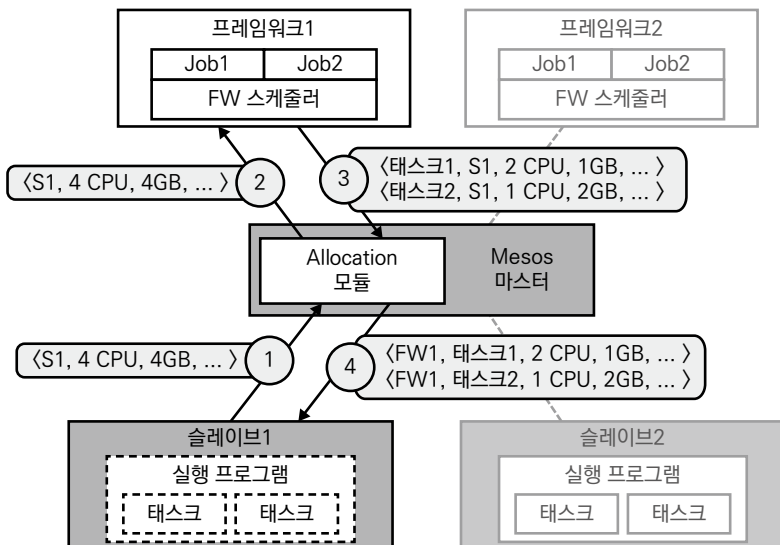
프레임워크는 Mesos 위에서 실행되는 특정 목적을 가진 애플리케이션을 말한다. 프레임워크는 각 슬레이브가 보낸 자원 현황을 바탕으로 태스크의 실행을 제어하는 스케줄러^{Scheduler}와 실제로 슬레이브에서 태스크를 실행하는 실행 프로그램^{Executor}으로 구성되어 있다.

Mesos는 리소스 오퍼^{Resource Offer}라는 정보를 주고받는데, [그림 1-2]를 참고하여 Mesos가 어떻게 동작하는지 리소스 오퍼를 중심으로 살펴보자.

1. 슬레이브1은 4 CPU와 4GB RAM을 사용할 수 있다고 마스터에게 전달한다.
마스터는 프레임워크1이 태스크를 실행하기 위해 리소스 오퍼를 요청한 사실을 Allocation 모듈에 저장하고 있다.
2. 마스터는 슬레이브1의 리소스 오퍼를 프레임워크1에 전달한다.
3. 프레임워크의 스케줄러는 슬레이브1에 2개(태스크1, 태스크2)의 태스크를 실행하겠다는 정보를 마스터에 응답으로 돌려준다. 태스크1은 2 CPU, 1GB RAM, 태스크2는 1 CPU, 2GB RAM을 사용한다.

4. 최종으로 마스터는 슬레이브1에 2개의 태스크(태스크1, 태스크2)를 전달하고, 슬레이브1에서 실행된다. 이제 슬레이브1에 남은 자원인 1 CPU와 1GB RAM 이 리소스 오퍼로 마스터에 전달된다.

그림 1-2 리소스 오퍼 예제



Mesos는 이와 같은 형태로 리소스 오퍼들이 Mesos 클러스터 내에 지속적으로 전달되면서 실행된다.

1.4 개발 환경의 Mesos 설치하기

실제로 서비스할 수 있는 Mesos 클러스터 구축은 뒤에서 다루기로 하고, 여기에서는 개발 환경의 Mesos를 손쉽게 설치해 보자. 편의를 위해 사용하고 있는 운영체제가 Mac OS X라고 가정하고 설명한다. 다른 운영체제를 사용 중이라면 VirtualBox와 Vagrant를 운영체제에 맞는 버전으로 설치한다.

VirtualBox와 Vagrant로 Ubuntu 14.04 기반의 VM을 생성하여 Mesos 클러스터 테스트 환경을 구축해 보자. 다음 방법을 따라 하면 Mesos 마스터, 슬레이브

각 1개와 이후에 살펴볼 Marathon 프레임워크와 Chronos 프레임워크가 VM 안에 설치된다.

1. **VirtualBox 설치** <https://www.virtualbox.org/wiki/Downloads> 페이지에서 OS X용 VirtualBox를 다운로드한 후 설치한다.

2. **Vagrant 설치** <http://www.vagrantup.com/downloads.html> 페이지에서 OS X용 Vagrant를 다운로드한 후 설치한다.

3. play-mesos 프로젝트를 git clone한다.

```
$ git clone https://github.com/mesosphere/playa-mesos.git
$ cd playa-mesos
```

4. 기본적인 테스트를 진행하여 정상적인 설치가 가능한 상태인지 확인한다.

```
$ bin/test
```

5. 원활한 실행을 위해 VM을 실행하기 전에 VM의 메모리를 4GB로 늘려 주자. config.json 파일 안의 vm_ram 값을 4096으로 설정하면 된다.

```
"vm_ram": "4096",
```

6. VM을 실행한다. VM이 생성되면서 Mesos 마스터, 슬레이브, Marathon 프레임워크, Chronos 프레임워크가 설치된다.

```
$ vagrant up
```

7. Mesos 마스터, Marathon 프레임워크, Chronos 프레임워크를 모두 최신 버전으로 업데이트하기 위해 다음 명령어를 실행한다(이 책을 집필하는 시점에 최신 버전은 Mesos 0.26.0, Marathon 0.15.2, Chronos 2.4.0다).

```
$ vagrant provision
```

8. VM 실행이 완료된 후 다음 주소로 접속하면 Mesos 마스터, Marathon 프레임워크, Chronos 프레임워크의 웹 화면에 접근할 수 있다.

- Mesos 마스터 <http://10.141.141.10:5050>
- Marathon 프레임워크 <http://10.141.141.10:8080>
- Chronos 프레임워크 <http://10.141.141.10:4400>

9. 복잡한 IP 대신 ‘mesos’와 ‘mesos.vm’이라는 간단한 호스트명으로 사용하기 위해 /etc/hosts 파일을 열어 다음과 같은 구문을 추가한다.

```
10.141.141.10 mesos mesos.vm
```

생성된 VM을 조작하는 명령어는 다음과 같다.

VM에 SSH로 접속

```
$ vagrant ssh
```

VM 중지

```
$ vagrant halt
```

VM 삭제

```
$ vagrant destroy
```

한빛 리얼타임은 IT 개발자를 위한 전자책입니다.

요즘 IT 업계에는 하루가 멀다 하고 수많은 기술이 나타나고 사라져 갑니다. 인터넷을 아무리 뒤져도 조금이나마 정리된 정보를 찾기도 쉽지 않습니다. 또한, 잘 정리되어 책으로 나오기까지는 오랜 시간이 걸립니다. 어떻게 하면 조금이라도 더 유용한 정보를 빠르게 얻을 수 있을까요? 어떻게 하면 남보다 조금 더 빨리 경험하고 습득한 지식을 공유하고 발전시켜 나갈 수 있을까요? 세상에는 수많은 종이책이 있습니다. 그리고 그 종이책을 그대로 옮긴 전자책도 많습니다. 전자책에는 전자책에 적합한 콘텐츠와 전자책의 특성을 살린 형식이 있다고 생각합니다.

한빛이 지금 생각하고 추구하는, 개발자를 위한 리얼타임 전자책은 이렇습니다.

1 eBook First - 빠르게 변화하는 IT 기술에 대해 핵심적인 정보를 신속하게 제공합니다

500페이지 가까운 분량의 잘 정리된 도서(종이책)가 아니라, 핵심적인 내용을 빠르게 전달하기 위해 조금은 거칠지만 100페이지 내외의 전자책 전용으로 개발한 서비스입니다. 독자에게는 새로운 정보를 빨리 얻을 기회가 되고, 자신이 먼저 경험한 지식과 정보를 책으로 펴내고 싶지만 너무 바빠서 엄두를 못 내는 선배, 전문가, 고수 분에게는 좀 더 쉽게 집필할 수 있는 기회가 될 수 있으리라 생각합니다. 또한, 새로운 정보와 지식을 빠르게 전달하기 위해 O'Reilly의 전자책 번역 서비스도 하고 있습니다.

무료로 업데이트되는 전자책 전용 서비스입니다

2 종이책으로는 기술의 변화 속도를 따라잡기가 쉽지 않습니다. 책이 일정 분량 이상으로 집필되고 정리되어 나오는 동안 기술은 이미 변해 있습니다. 전자책으로 출간된 이후에도 버전 업을 통해 중요한 기술적 변화가 있거나 저자(역자)와 독자가 소통하면서 보완하여 발전된 노하우가 정리되면 구매하신 분께 무료로 업데이트해 드립니다.

3 독자의 편의를 위해 DRM-Free로 제공합니다

구매한 전자책을 다양한 IT 기기에서 자유롭게 활용할 수 있도록 DRM-Free PDF 포맷으로 제공합니다. 이는 독자 여러분과 한빛이 생각하고 추구하는 전자책을 만들어 나가기 위해 독자 여러분이 언제 어디서 어떤 기기를 사용하더라도 편리하게 전자책을 볼 수 있도록 하기 위함입니다.

4 전자책 환경을 고려한 최적의 형태와 디자인에 담고자 노력했습니다

종이책을 그대로 옮겨 놓아 가독성이 떨어지고 읽기 어려운 전자책이 아니라, 전자책의 환경에 가능한 한 최적화하여 쾌적한 경험을 드리고자 합니다. 링크 등의 기능을 적극적으로 이용할 수 있음은 물론이고 글자 크기나 행간, 여백 등을 전자책에 가장 최적화된 형태로 새롭게 디자인하였습니다.

앞으로도 독자 여러분의 충고에 귀 기울이며 지속해서 발전시켜 나가겠습니다.

지금 보시는 전자책에 소유 권한을 표시한 문구가 없거나 타인의 소유권함을 표시한 문구가 있다면 위법하게 사용하고 있을 가능성이 큼니다. 이 경우 저작권법에 따라 불이익을 받으실 수 있습니다.

다양한 기기에 사용할 수 있습니다. 또한, 한빛미디어 사이트에서 구매하신 후에는 횡수와 관계없이 내려받을 수 있습니다.

한빛미디어 전자책은 인쇄, 검색, 복사하여 붙이기가 가능합니다.

전자책은 오타자 교정이나 내용의 수정·보완이 이뤄지면 업데이트 관련 공지를 이메일로 알려 드리며, 구매하신 전자책의 수정본은 무료로 내려받으실 수 있습니다.

이런 특별한 권한은 한빛미디어 사이트에서 구매하신 독자에게만 제공되며, 다른 사람에게 양도나 이전은 허락되지 않습니다.