

빅데이터

Big data Framework

프레임워크



02

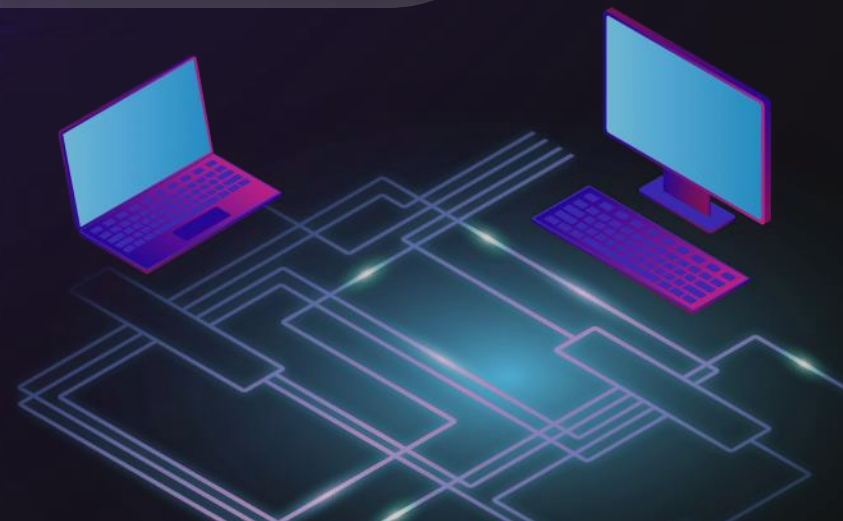
Big data Framework

빅데이터 기반 기술 - OS 및 가상화 기술






02

가상화 기술



02 | 가상화 기술

-  가상화 기술이 무엇인지 설명할 수 있다.
-  자원 가상화의 개념을 설명할 수 있다.
-  운영 체제 가상화의 개념을 설명할 수 있다.

1 가상화 기술이란?

❖ 애플리케이션, 서버, 스토리지 및 네트워크 등의 컴퓨팅 기술을 추상화하여 요청한 만큼의 리소스를 서비스 형태로 제공해 주는 기술

가상화 기술의 종류



1

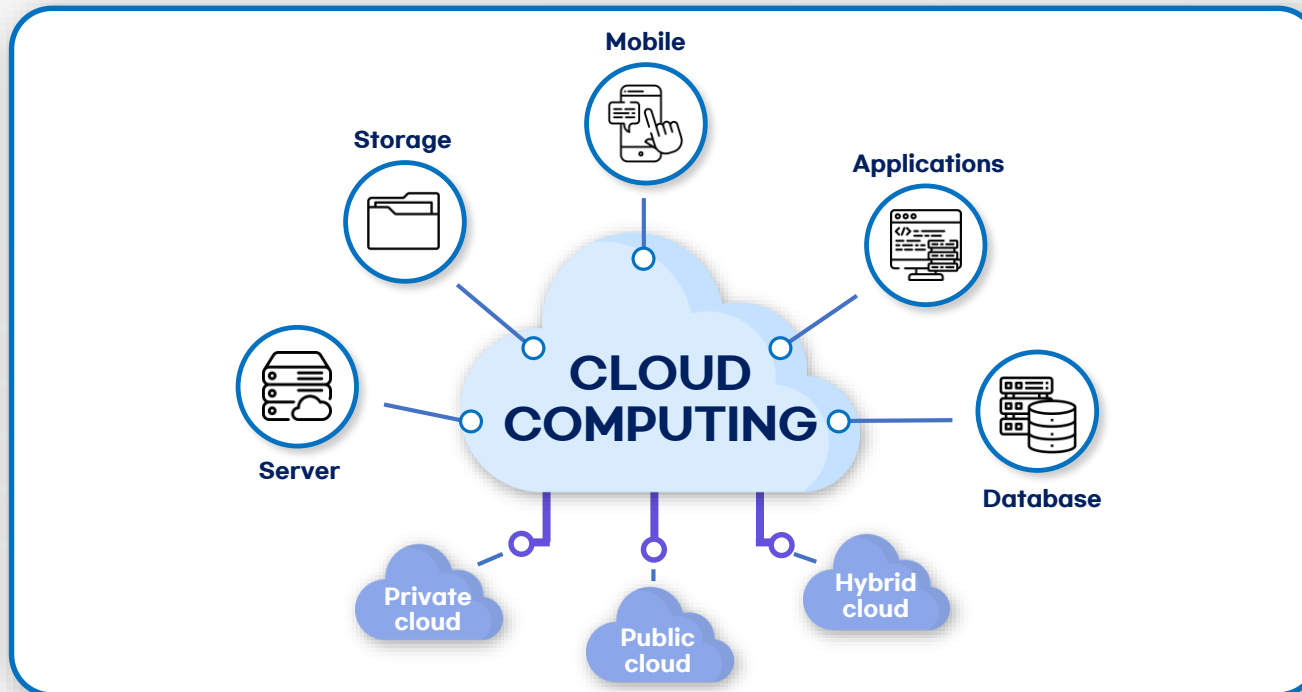
자원
가상화

2

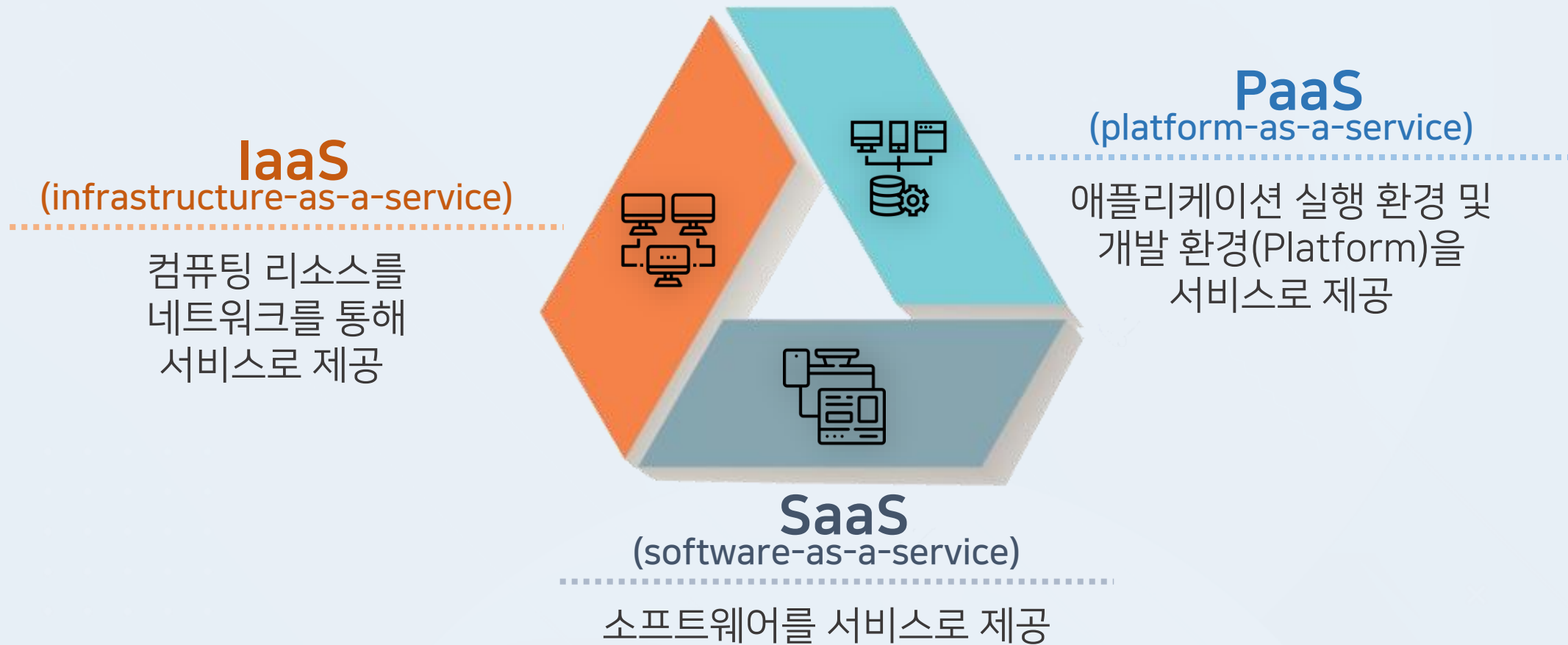
운영 체제
가상화

1 자원 가상화란?

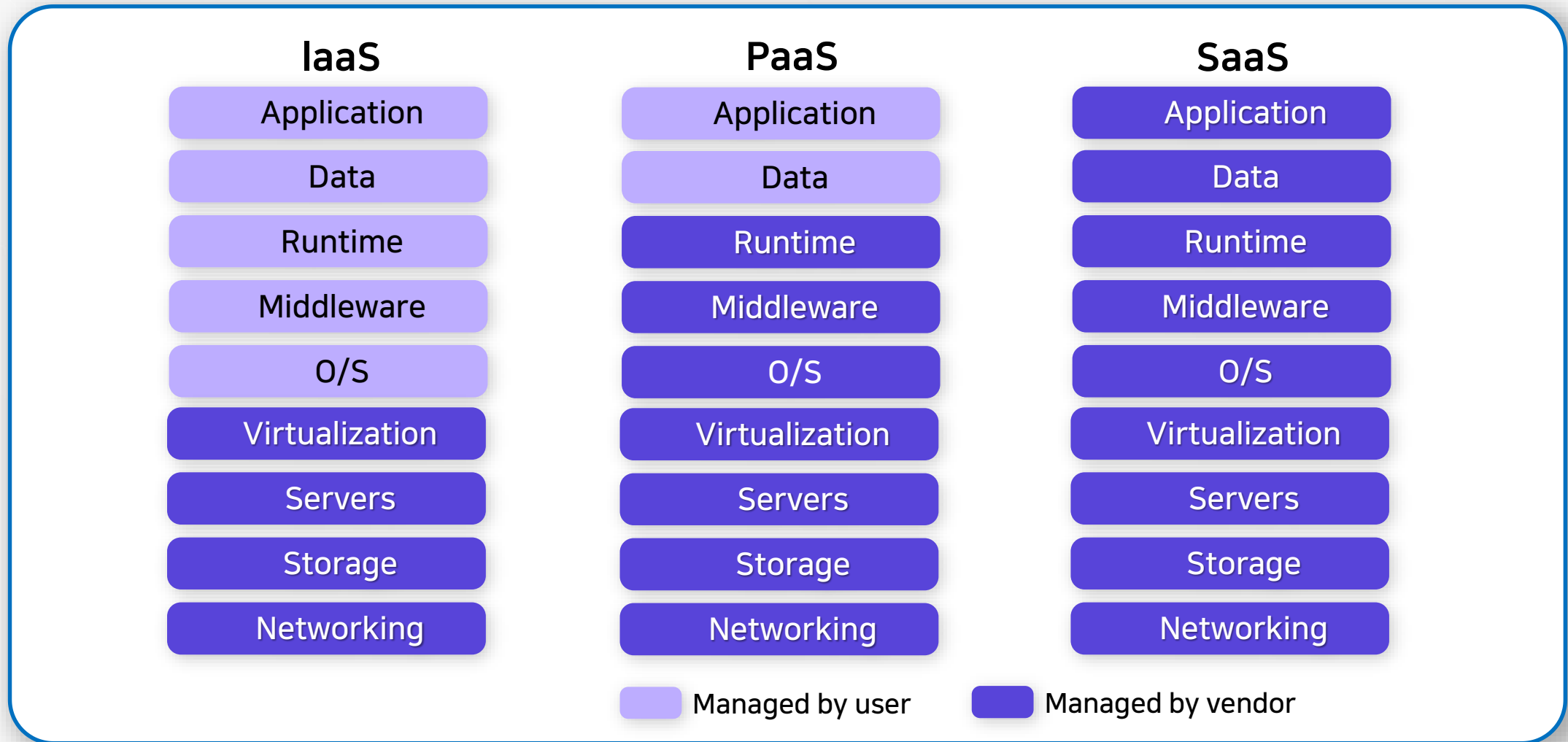
- 데이터 스토리지, 연산 장치, 네트워크 등의 컴퓨팅 리소스 자원을 가상화하여 사용자가 필요한 만큼 제공해 주는 기술
- Private Cloud, Public Cloud, Hybrid Cloud 형태가 있음



2 클라우드 컴퓨팅의 기술 요소



2 클라우드 컴퓨팅의 기술 요소



3 자원 가상화와 분산 시스템의 관계

- 빅데이터로 유효한 결과를 얻으려면 충분히 많은 데이터를 확보한 후 분석을 해야 함
- 폭발적으로 증가하는 데이터를 처리하는 시스템을 구축하고 운영하는 것은 상당히 많은 관리 포인트가 발생함



3 자원 가상화와 분산 시스템의 관계

- 클라우드 기술은 분산시스템 구축의 많은 관리 비용을 절감할 수 있음
- 데이터의 발생량에 따른 컴퓨팅 자원의 탄력적 확장이 용이함

수직적 확장(Scale Up)

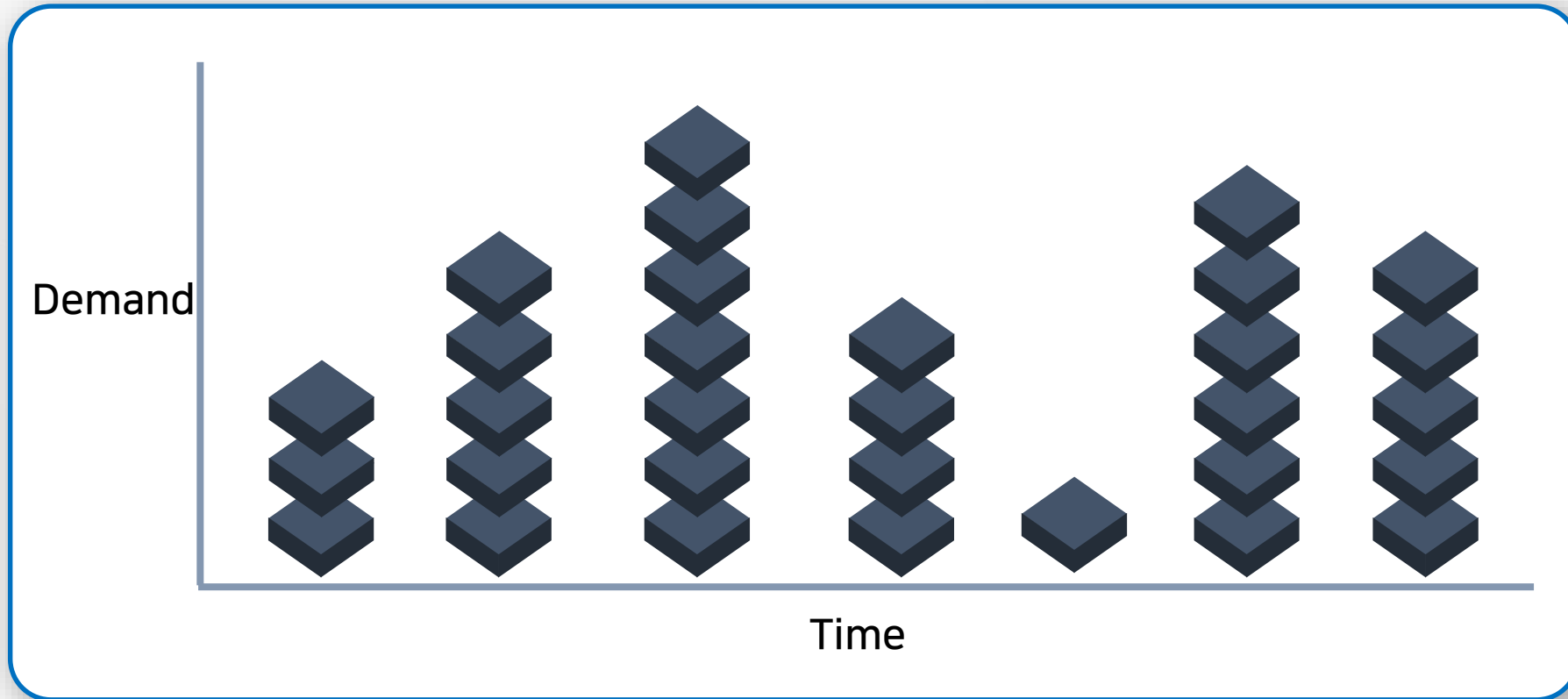


수평적 확장(Scale Out)



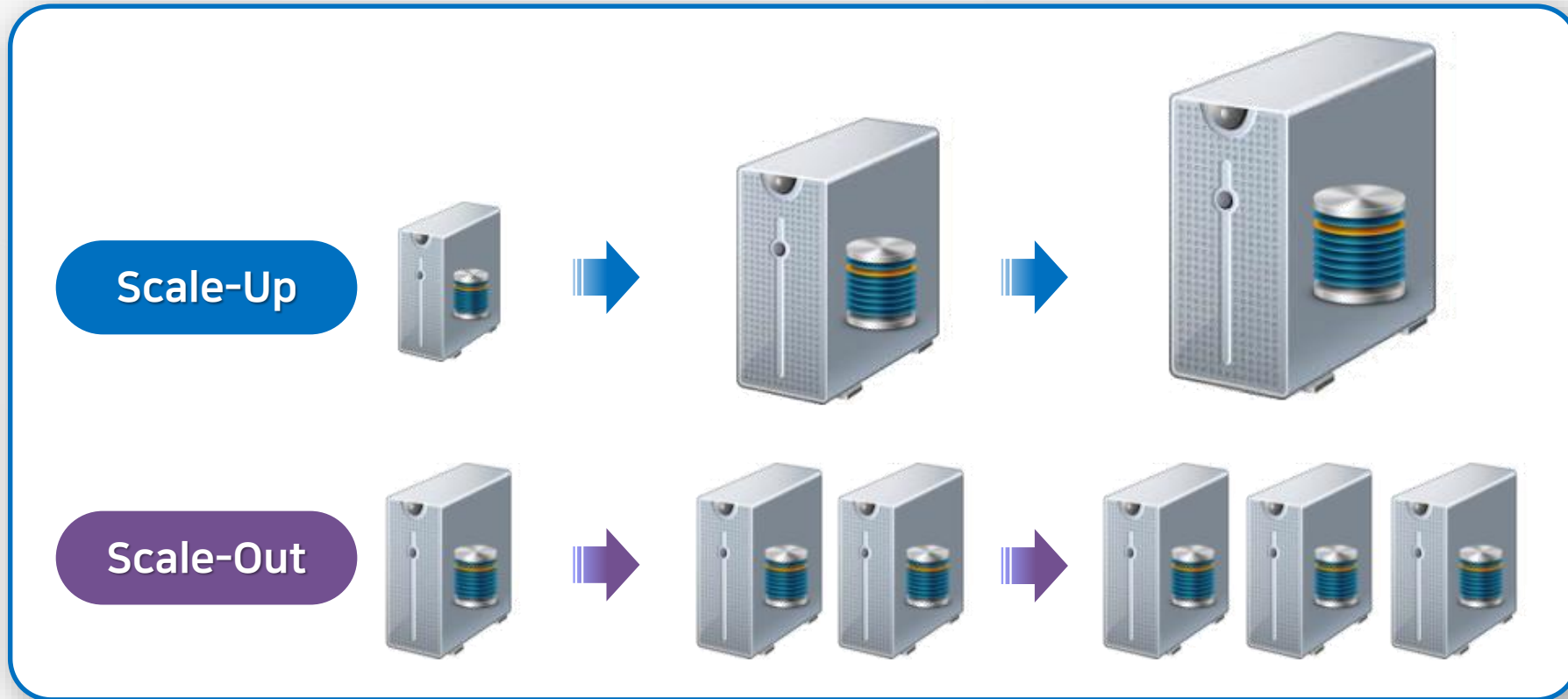
4 분산 시스템 관점에서 본 클라우드 시스템의 효율성

🏠 데이터의 규모에 따라 탄력적인 자원 사용



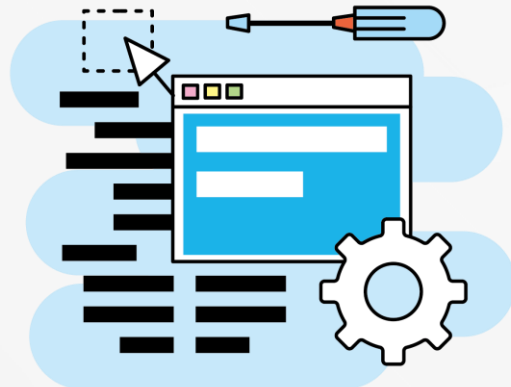
4 분산 시스템 관점에서 본 클라우드 시스템의 효율성

🏠 데이터의 규모에 따라 탄력적인 자원 사용



1 운영 체제 가상화

- ❖ 운영 체제의 커널이 하나의 사용자 공간 인스턴스가 아닌, 여러 개의 격리된 사용자 공간 인스턴스를 갖출 수 있도록 하는 가상화 기술
- ❖ 개개의 인스턴스들을 컨테이너(Container)라고도 부르며 소유자의 관점에서는 인스턴스가 실제 서버와 동일하게 보임
- ❖ 애플리케이션뿐만 아니라 이를 구동하는 모든 환경을 호스트 시스템과 분리함



2 컨테이너

❖ 애플리케이션이 동작하기 위해 필요한 모든 환경적 요소를 패키지화하여 격리하는 기술

❖ 애플리케이션과 애플리케이션의 운영 환경을 모두 묶어서 패키징 하므로 타 시스템에서 구동하기에 매우 용이함

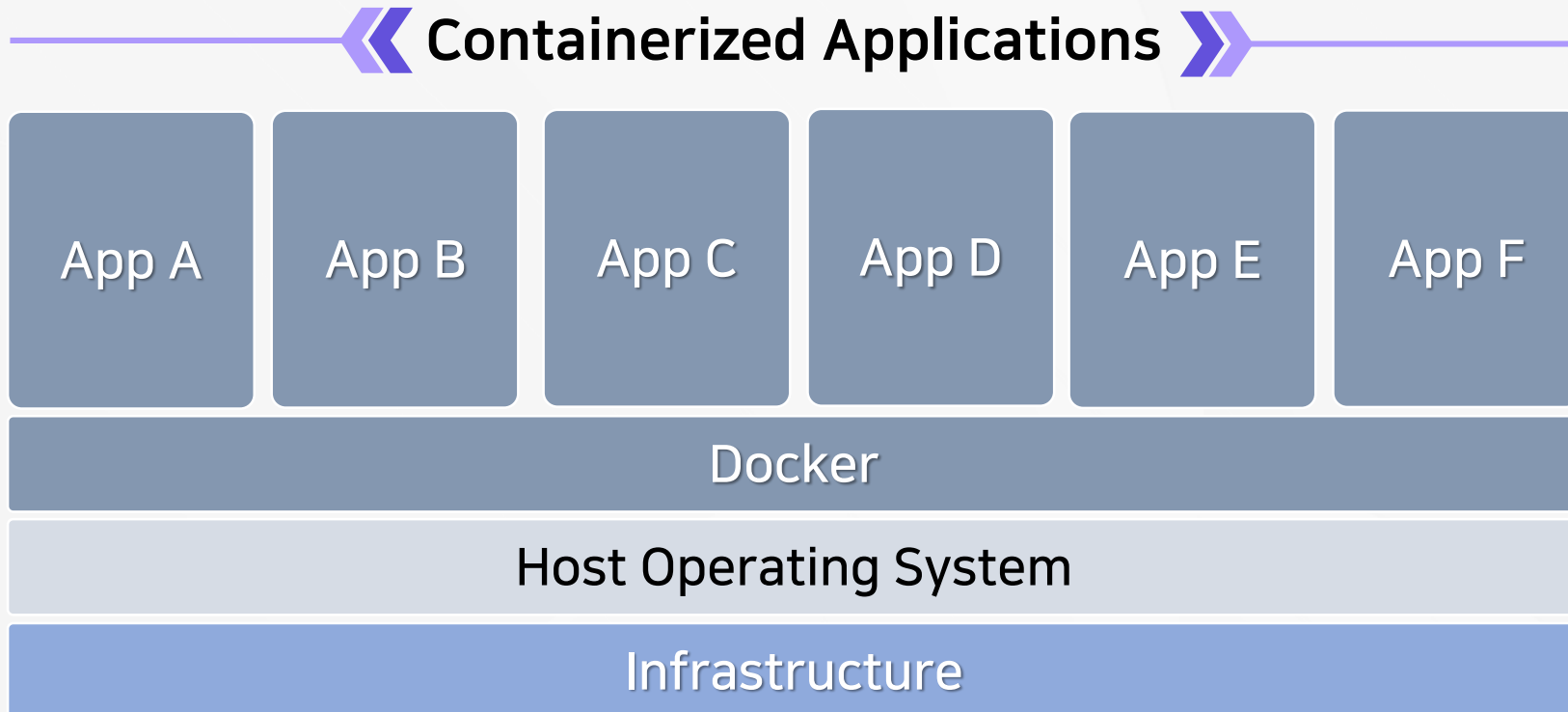
❖ 애플리케이션의 구동 시, 리소스 효율성이 높음

➔ 컨테이너 구동 시 별도의 커널을 가지고 있지 않고, 애플리케이션에 필요한 최소한의 환경과 메타데이터만 가지고 있으면 됨

❖ 컨테이너의 롤백이 가능하여, 이전의 상태를 신속히 회복함

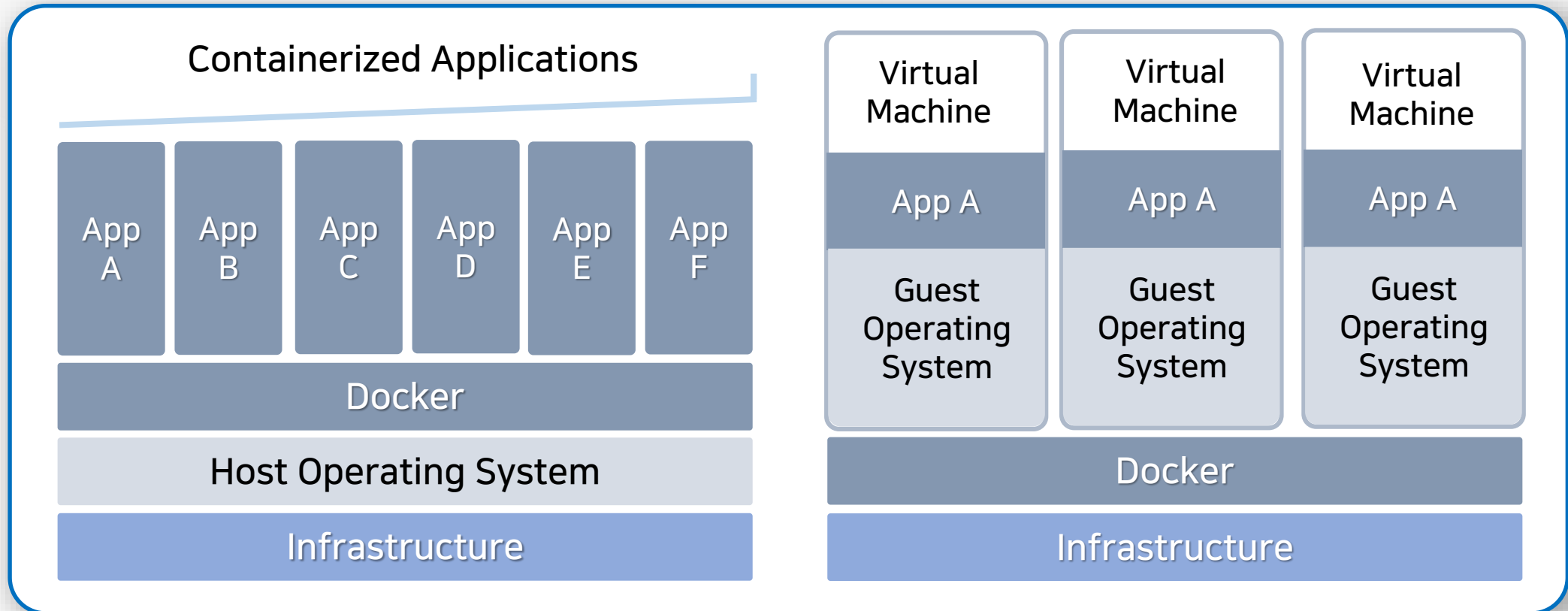
3 컨테이너의 구조

- ❖ 호스트 컴퓨터의 커널을 공유
- ❖ 각 애플리케이션은 격리된 수행환경에서 구동



4 운영 체제 가상화(컨테이너)와 자원 가상화(Virtual Machine)의 차이

- ❖ VM은 가상머신을 생성하기 위해 하이퍼바이저가 필요함
- ❖ 컨테이너는 호스트 컴퓨터의 OS 커널을 공유하지만, VM은 Guest Os를 따로 둬م



5 도커 (Docker)

📦 Go 언어로 작성된 리눅스 컨테이너 기반 오픈소스 가상화 플랫폼

📦 도커 이미지, 레지스트리, 컨테이너를 손쉽게 사용할 수 있도록 구조화되어 있음

이미지

라이브러리 및 package의 template, 읽기 전용

레지스트리

도커 이미지를 저장하는 이미지 저장소

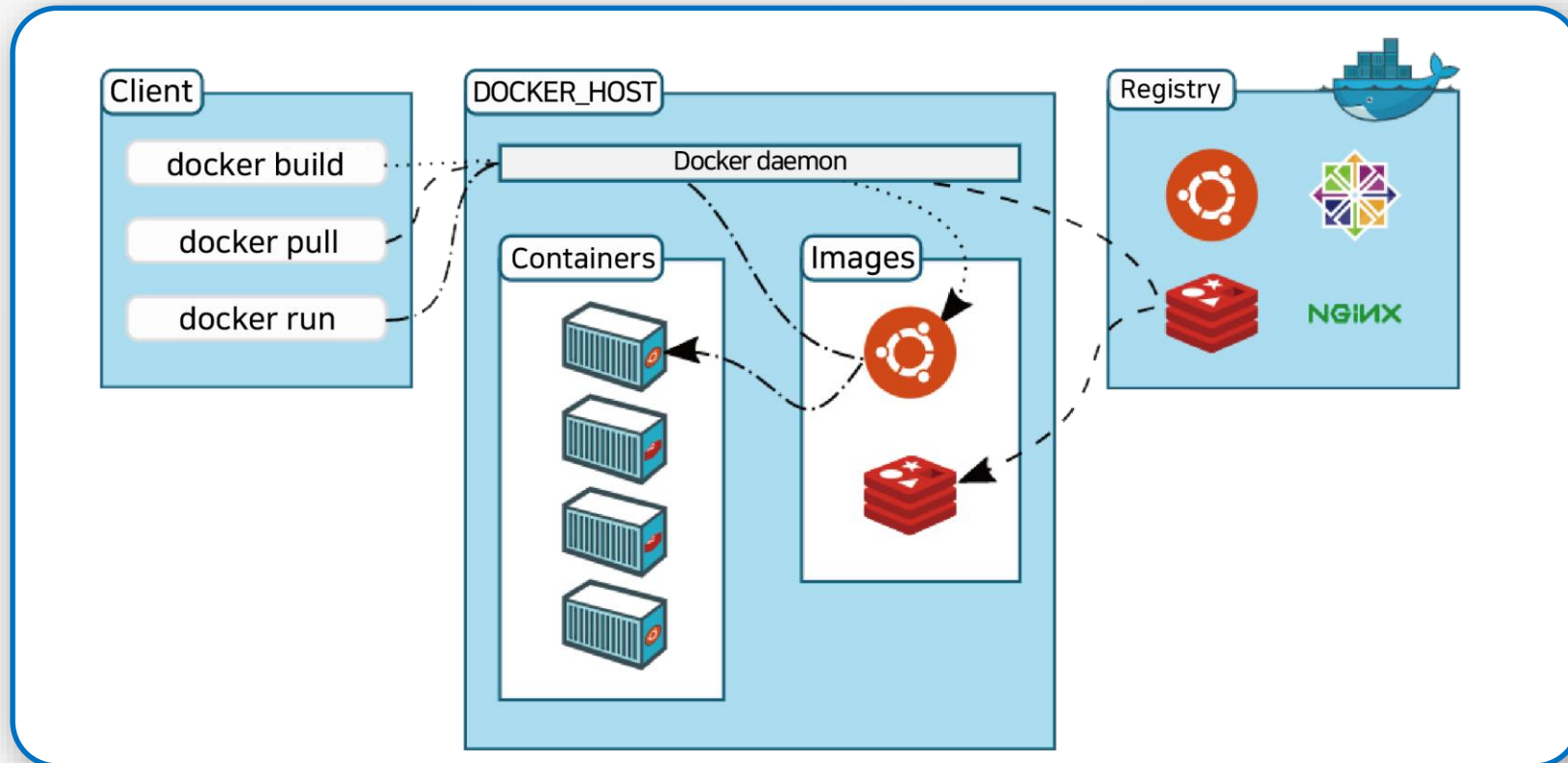
컨테이너

이미지가 로딩되어 메모리에 적재된 상태

5 도커 (Docker)

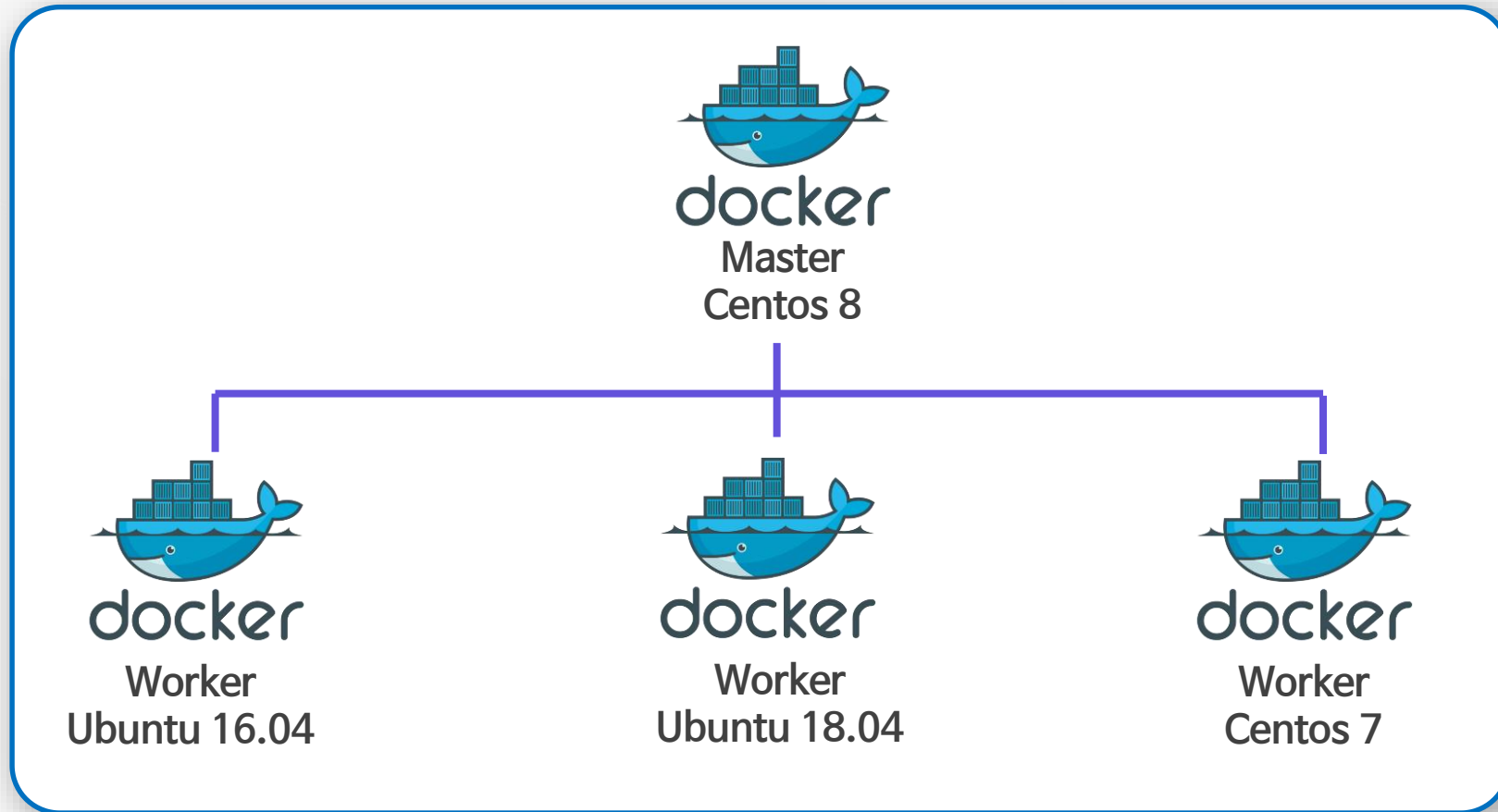
📦 Go 언어로 작성된 리눅스 컨테이너 기반 오픈소스 가상화 플랫폼

📦 도커 이미지, 레지스트리, 컨테이너를 손쉽게 사용할 수 있도록 구조화되어 있음



5 운영 체제 가상화와 분산 시스템의 관계

- 각 분산된 노드의 환경이 달라도, 운영 체제 격리를 통해 일관된 환경에서 분산 시스템 구축 및 운영이 가능함



1 가상화의 개념

- 애플리케이션, 서버, 스토리지 및 네트워크 등의 컴퓨팅 기술을 추상화하여 요청한 만큼의 리소스를 서비스 형태로 제공해 주는 기술

2 자원 가상화

- 데이터 스토리지, 연산 장치, 네트워크 등의 컴퓨팅 리소스 자원을 가상화하여 사용자가 필요한 만큼 제공해 주는 기술
- 클라우드 컴퓨팅의 기술 요소
 - IaaS, PaaS, SaaS

3 운영 체제 가상화

- 운영 체제의 커널이 하나의 사용자 공간 인스턴스가 아닌, 여러 개의 격리된 사용자 공간 인스턴스를 갖출 수 있도록 하는 가상화 기술
- 컨테이너
 - 애플리케이션이 동작하기 위해 필요한 모든 환경적 요소를 패키지화하여 격리하는 기술
- 도커
 - Go 언어로 작성된 리눅스 컨테이너 기반 오픈소스 가상화 플랫폼

✎ 빅데이터 컴퓨팅 기술, 2014, 박두순, 한빛미디어