Data and AI

AI시대 효과적 데이터 관리를 위한 데이터 현대화

— Db2 : Cloud Native Database

소가영 실장

Data & AI
IBM Korea

© 2020 IBM Corporation

“2020 년 금융 기업은 AI 와 Data 의 확장 으로 비즈니스 프로세스 혁신을 가속화하고 있습니다.”

1

AI 및 데이터 확장을 위한 적절한 **운영 모델** 설계 및 구축

: Cognitive Factory / Garage

2

적절한 **데이터 기반** 으로 변환 하고 기술 채택

: 데이터 플랫폼을 재설계하고 데이터 현대화

3

신뢰할 수 있고 투명하며 설명 가능

ADOPTION

75% 의 기업(EU)이 AI를 구현하고 있습니다.
47% 는 비즈니스 전반에 AI를 배포 할 계획을 가지고 있습니다.

BARRIERS FOR AI

37% - 기술 부족
31% - 데이터 복잡성 및 단편화
26% - AI 모델 개발을 위한 도구 부족

CRITICAL TRUST

78% - AI의 결과가 공정하고 안전하며 신뢰할 수 있다고 믿을 수있는 것이 '매우'중요합니다.
83% - AI가 어떻게 결정에 도달했는지 설명할 수 있어야합니다.

**IBM Study, From Roadblock to Scale, 4514 online interviews w/ Executive from US, Europe, China, October 2019*

© 2020 IBM Corporation

데이터 현대화
클라우드로의 전환

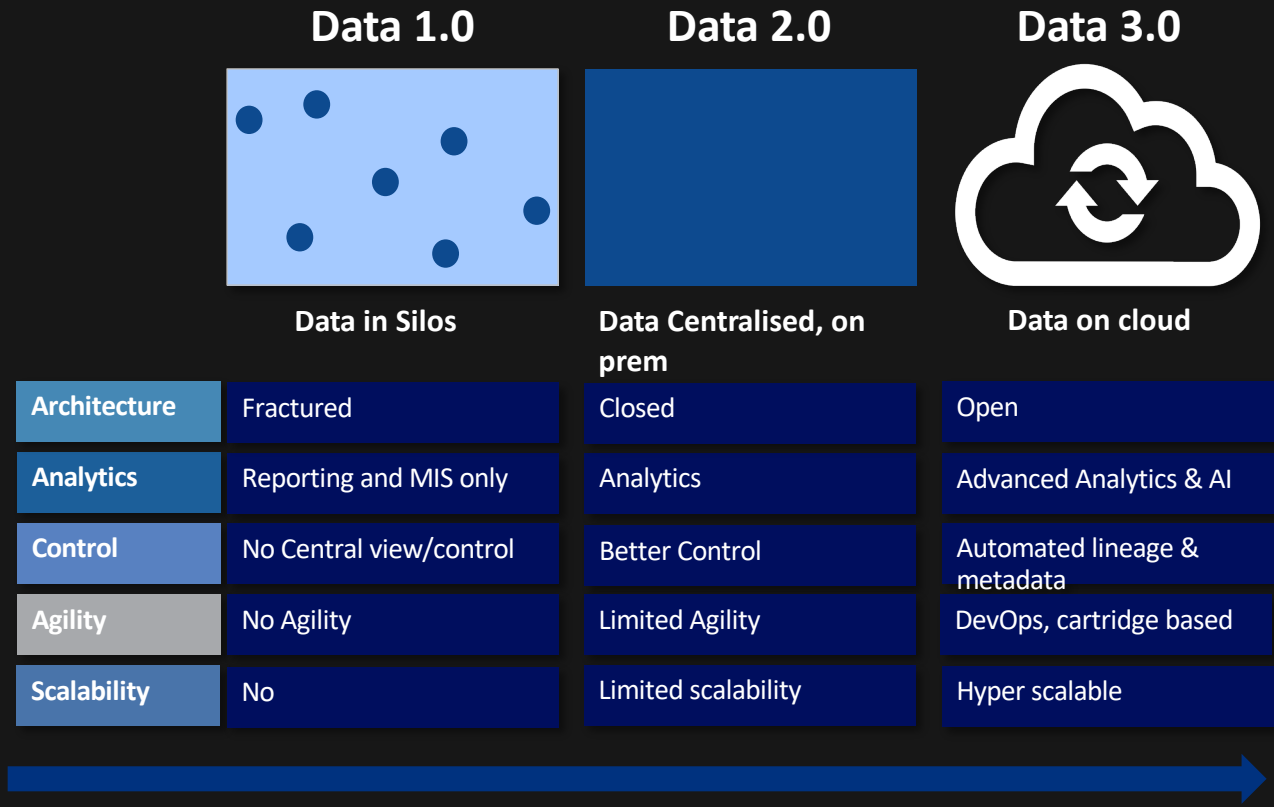
- 분리된 아키텍처에서 개방형 아키텍처로 변화

- 기본 리포팅에서 고급 분석 및 AI 로 확장

- 자동화된 데이터 계보 및 메타데이터 관리로 데이터에 대한 중앙 집중적 관리

- 데이터에 대한 민첩성 확보

- 확장성이 뛰어난 플랫폼



© 2020 IBM Corporation

데이터 Anywhere

Public, Private, On-premise간 이식과 연계가 용이

Cloud Ready

- Public Cloud, Private Cloud, 전용 Appliance 이 동일한 S/W Stack으로 구성되어 이식과 연계가 용이

Common SQL

- 공통 SQL 사용으로 개발 생산성 향상 및 이관 용이

Managed public Cloud DBaaS	Software defined warehouse On-Premises or in Cloud	Dedicated analytics appliance	Custom deployable database	Open source Hadoop with Hortonworks
Db2 on Cloud Db2 Warehouse on Cloud	Db2 Warehouse	Integrated Analytics System	Db2 Database	Db2 Big SQL

- **신속한 Application 개발**
  Write once, run anywhere
  One ISV product certification for all platforms
- **운영 호환성**
  Reuse operational and housekeeping procedures
- **유연한 라이센스**
  Flexible entitlements for business agility and cost-optimization
- **통합**
  Common Data Virtualization capabilities for query federation and data movement
- **표준화된 분석**
  Common programming model for in-DB analytics
- **공통 스킬**
  One skill set for all deployments
  Drive higher efficiencies and portfolio rationalization

© 2020 IBM Corporation

데이터 관리 현대화의 핵심 요건

Omni-Experience, Customer Engagement

- 엣지-코어-클라우드 : 끊김없는 데이터 연결 및 가용성
- 데이터 품질 및 프로파일링
- 데이터 동기화 (내부 및 외부 데이터 소스)
- 메타 데이터 관리
- 데이터 위치 최적화
- 분석 / AI 워크로드의 가속화 및 최적화
- 일관되고 균형 잡힌 성능

Digital Trust & Stewardship

- 데이터 손실 방지
- 암호화, 감사, 로깅, 역할 기반 액세스 및 제어
- 국가 산업 규정 준수
- 데이터 배포 위치에서 일관되고 세분화된 백업, 복원, 재해 복구 및 연속성
- 리스크 완화

Efficiency and Agility

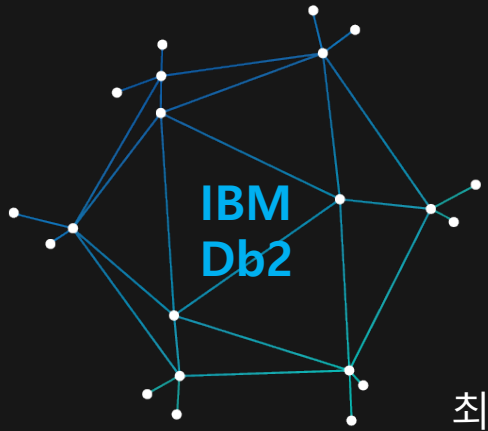
- 기존 투자 활용
- 레거시 인프라 현대화
- 클라우드 인프라 채택 및 통합과 최적화
- IT 인프라, DevOps, 컨테이너, 하이퍼 컨버지드 인프라 현대화
- 비즈니스 민첩성 지원

© 2020 IBM Corporation

Db2 11.5 : AI ready database

우수한 쿼리 성능
- 머신러닝 Optimizer, 인메모리 분석

성능 저하 없는 실시간 데이터 압축
- 적응형 압축 기술



이기종 데이터 실시간 액세스
- 데이터 가상화 기술

무중단 아키텍처(pureScale)
- 온라인 유지보수, 무중단 DBMS패치

최신 AI 애플리케이션 개발
- Python, GO , JSON , Jupyter notebooks

Cloud Ready
- 컨테이너 기반, Continuous Delivery

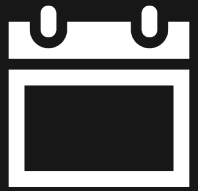
© 2020 IBM Corporation

Db2 in Container (via OpenShift 시너지)

클라우드 네이티브 데이터베이스

손쉽고 빠른 HA 구성

보안성 높은 컨테이너 환경

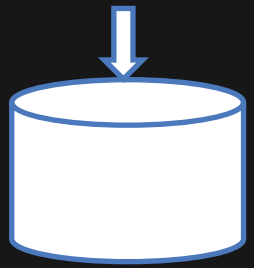


소프트웨어 정의 스토리지

© 2020 IBM Corporation

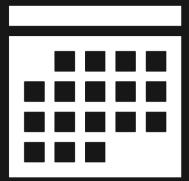
Cloud Native 데이터베이스

민첩하고
탄력적이며
모듈화된
데이터베이스를
신속히 배포



배포 신속성

단순화 된 라이프 사이클 관리
(Helm 통한 업그레이드, Operator 백업 / 복원)



독립적 개발/업데이트/확장 가능한 마이크로 서비스 형태로 아키텍처 및 배포

Continuous Delivery 를 통한 새로운 기능의 신속한 제공

온 프레미스, 모든 클라우드에 배포 가능한 유연성

© 2020 IBM Corporation

손쉽고 빠른 HA 구성

높은 가용성으로 시스템을 연중 무휴 24 시간 운영

성능에 거의 영향없이 무정지 상태로 시스템을 유지하고 손쉽게 관리

- ❖ 워크로드를 분산하고 장애 시 즉각 복구
- ❖ Kubernetes 기반 높은 가용성 (Standby 통한 자동 복구)

다양한 토폴로지 구성

- ❖ 동일 OpenShift/Kubernetes cluster 내 Primary 와 Standby (same or different namespaces)
- ❖ 다른 OpenShift/Kubernetes cluster 내 Primary 와 Standby

© 2020 IBM Corporation

높은 보안성

Db2 보안기능
+
RHOS 컨테이너
보안 기능

Db2 보안기능과 RedHat OpenShift Container 보안 기능과 통합되어 데이터를 보다 완벽하게 보호

Db2 보안

- **클라이언트-서버간 네트워크 전송중인 데이터를 암호화하는 TLS (Transport Layer Security)**
- **디스크 수준 데이터 암호화 (Native Encryption)**

OpenShift Container 보안

- **Non-root 사용자로 실행되고 root 액세스가 필요한 프로세스는 SUDO 마스킹**
- **OpenShift의 SELinux 모드 (액세스 제어 보안 정책을 지원하는 Linux 커널 보안 모듈)**
- **Db2 컨테이너 이미지 검증 (컨테이너가 OpenShift와 호환되고 상업적으로 지원되며 이미지에 플랫폼 구성 요소 또는 계층의 알려진 취약성이 없는지 확인)**

© 2020 IBM Corporation

기존 스토리지 + 소프트웨어 정의 스토리지

성능 및 확장 요구에 따라 스토리지 옵션을 선택하고 필요에 따라 스토리지를 확장하여 비용을 절감

- 기존 스토리지 지원
  (성능 위한 SAN, NFS, GPFS 등 클러스터 파일 시스템, 로컬 디스크)
- 소프트웨어 정의 스토리지 지원
  - RedHat OpenShift 및 Kubernetes에 기본 제공되며 동적 프로비저닝, 복제, 스냅샷, 확장 등 스토리지 계층 관리
  - 하드웨어에서 소프트웨어를 분리한 스토리지 가상화로 값 비싼 특수 하드웨어의 필요성 제거

  (Portworx, Ceph, OCS, IBM Cloud File Storage)

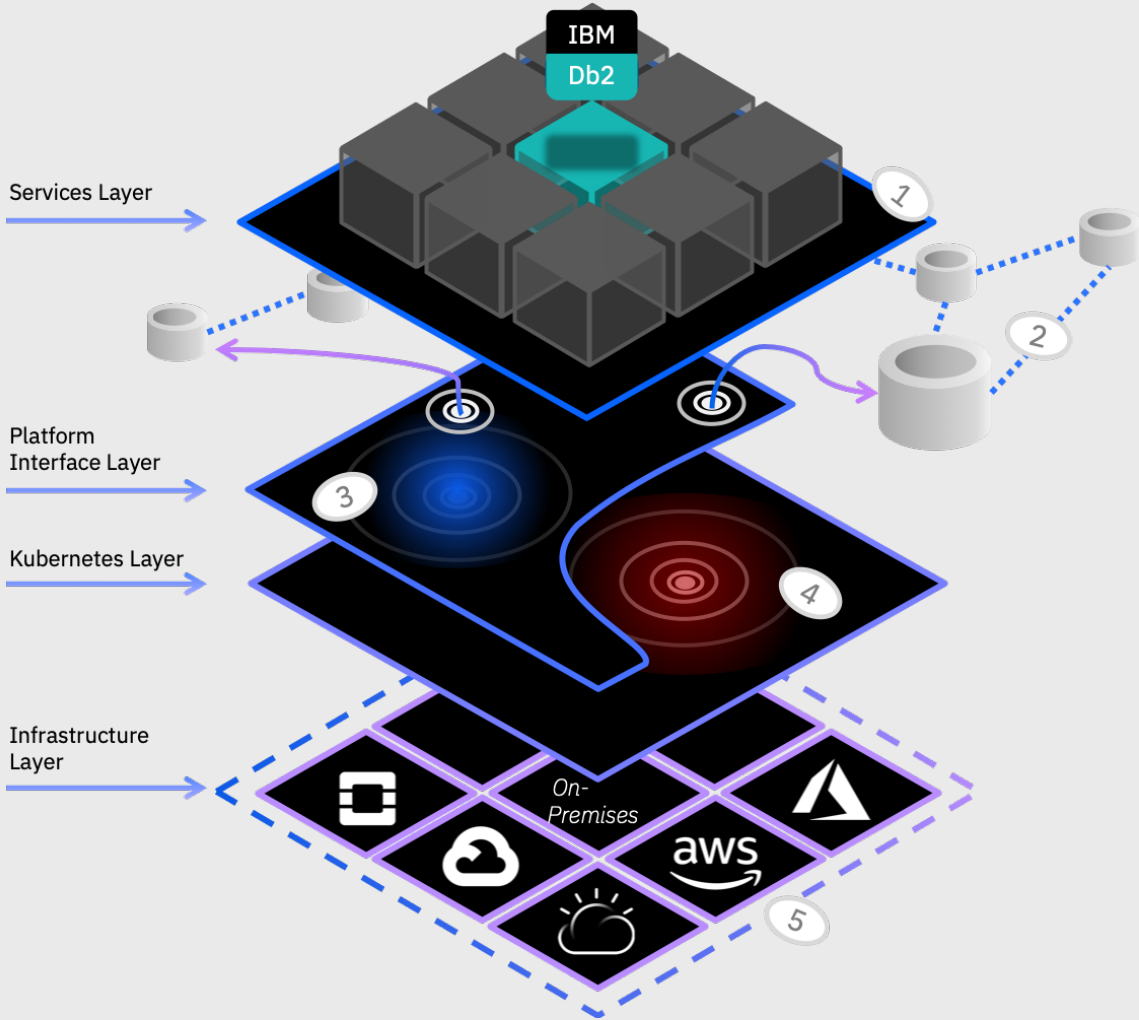
※ Internal Test에서 Portworx와 Ceph중 Ceph가 더 성능 우수

© 2020 IBM Corporation

IBM Cloud Pak for Data Db2

- **Ecosystem integration**
- **Platform integration**
- **Data&AI integration**
- **Administration integration**

- Data virtualization
- Data warehousing
- Databases on-demand
- Data source ingestion
- Distributed processing



© 2020 IBM Corporation

Cloud Pak for Data

1. 서비스 에코시스템

한번의 클릭으로 IBM 및 3’rd party의 45개 이상의 분석 서비스와 템플릿으로 구성된 에코시스템에 액세스

2. 데이터 가상화

데이터 이동없이 여러 데이터 소스에 걸쳐 빠르고 쉽게 쿼리.

3. 플랫폼 인터페이스

데이터 관리, 데이터 거버넌스 및 분석을 통합하여 효율성을 높이고 리소스 사용을 개선하는 단일 사용자 환경으로 가치 창출 시간 단축

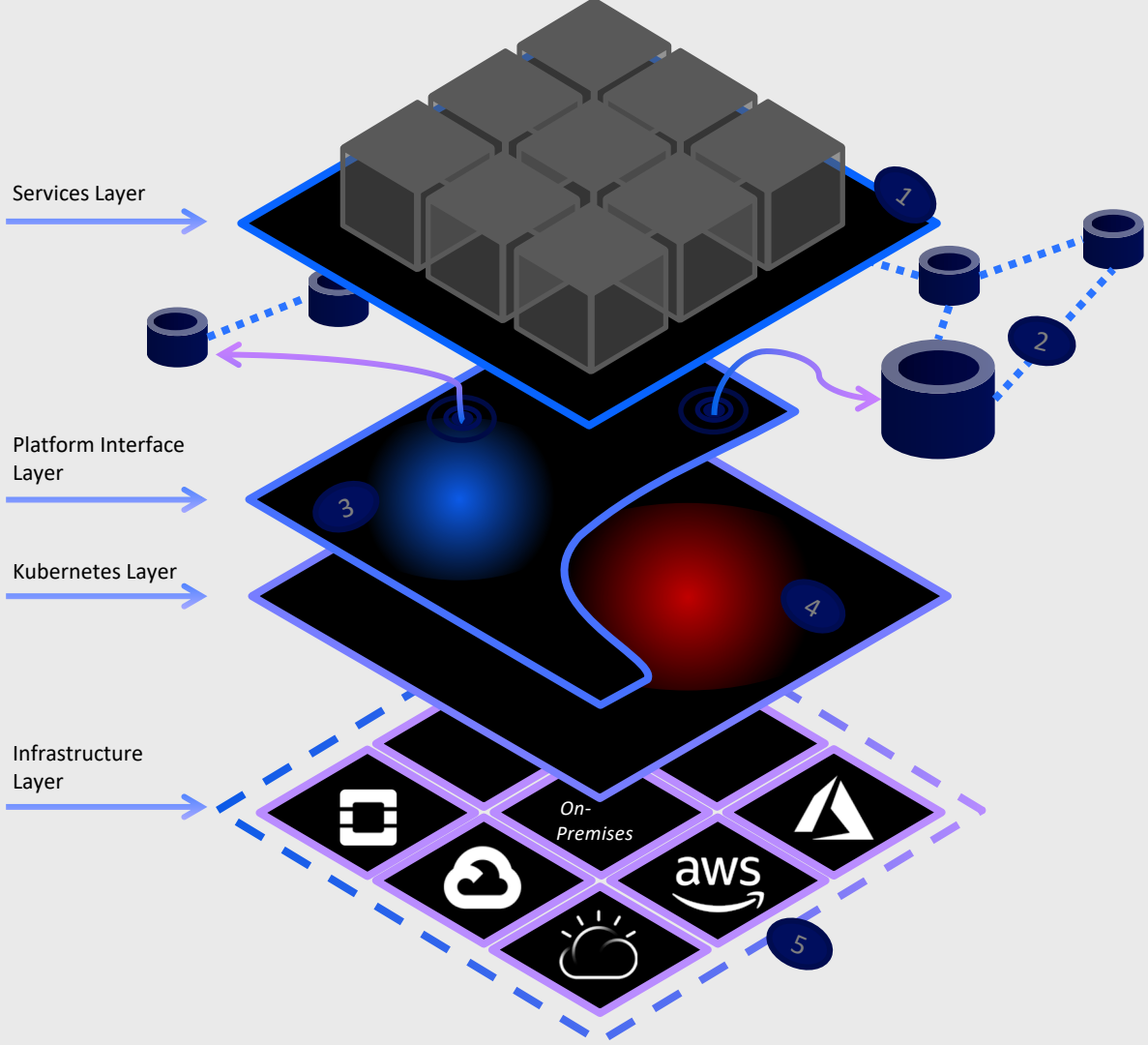
4. Red Hat OPENSHIFT®

업계 최고의 하이브리드 클라우드 엔터프라이즈 컨테이너 플랫폼으로 혁신적이고 신속한 구현 전략 수립

5. Any Cloud

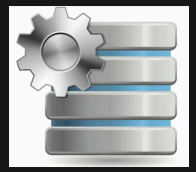
멀티 클라우드로 모든 클라우드 인프라 활용

.



© 2020 IBM Corporation

Db2 in Cloud Pak for Data 장점



클라우드 및AI 여정을 가속화하고 복잡성 감소

Easy to Deploy

- **쿠키 커터 배포**
  - 몇 분 내에 배포가 가능하고 유연하게 표준화된 배포 및 관리 가능
  - 프로비저닝시 사전 맞춤형 DB 생성
- **자동 장애조치**
  - 즉시 사용 가능한 향상된 Kubernetes 가용성
  - HA / DR 지원
- **자동 업데이트**
  - 버튼 클릭만으로 서비스 팩, 버전 및 모드 배포 가능

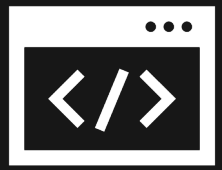
Easy to Manage

- **"애플리케이션 그룹"에 의한 자동화 된 관리**
  - 네임 스페이스를 사용하여 액세스 제어 및 프로비저닝 옵션 관리
- **하나의 애플리케이션 레벨로 모니터링 및 관리**
  - 플랫폼 및 서비스 수준 모니터링 및 관리

© 2020 IBM Corporation

플랫폼 통합 : Unified Db2 Console

통합 된 뷰를 사용하여 Cloud Pak for data 내의 Db2의 빠르고 쉬운 모니터링 및 성능 관리



다양한 폼팩터의 100개 이상의 Db2 데이터베이스의 관리 및 모니터링

모니터링 및 경고로 영향을 미치기 전에 문제 파악

리소스 사용 및 데이터 경합 관리

개방형 API로 커스터마이징 가능

팀 협업기능 내장

© 2020 IBM Corporation

Data and AI 통합 : Db2 ML Capabilities

AI 기능을 활용하여 비즈니스에 파급 효과가있는 데이터베이스 성능 최적화

Db2가 CP4D Marketplace에서 배포되는 즉시 Watson Knowledge Catalog (Organize) 및 Watson Studio (Analyze)와 같은 CP4D의 다른 AI 서비스에서 사용할 수 있으며 고객은 내장 된 Db2 AI 기능을 활용할 수 있습니다.

머신 러닝 쿼리 Optimizer:
기계 학습을 사용하여 쿼리 성능 향상

통합 데이터 관리 기능이있는 다중 모델 데이터베이스

Visual Studio 추가 기능을 사용하여 Visual Studio 개발 환경에 원활하게 통합

Support for:

Popular Languages for AI
Python
Go

JSON Open source file format

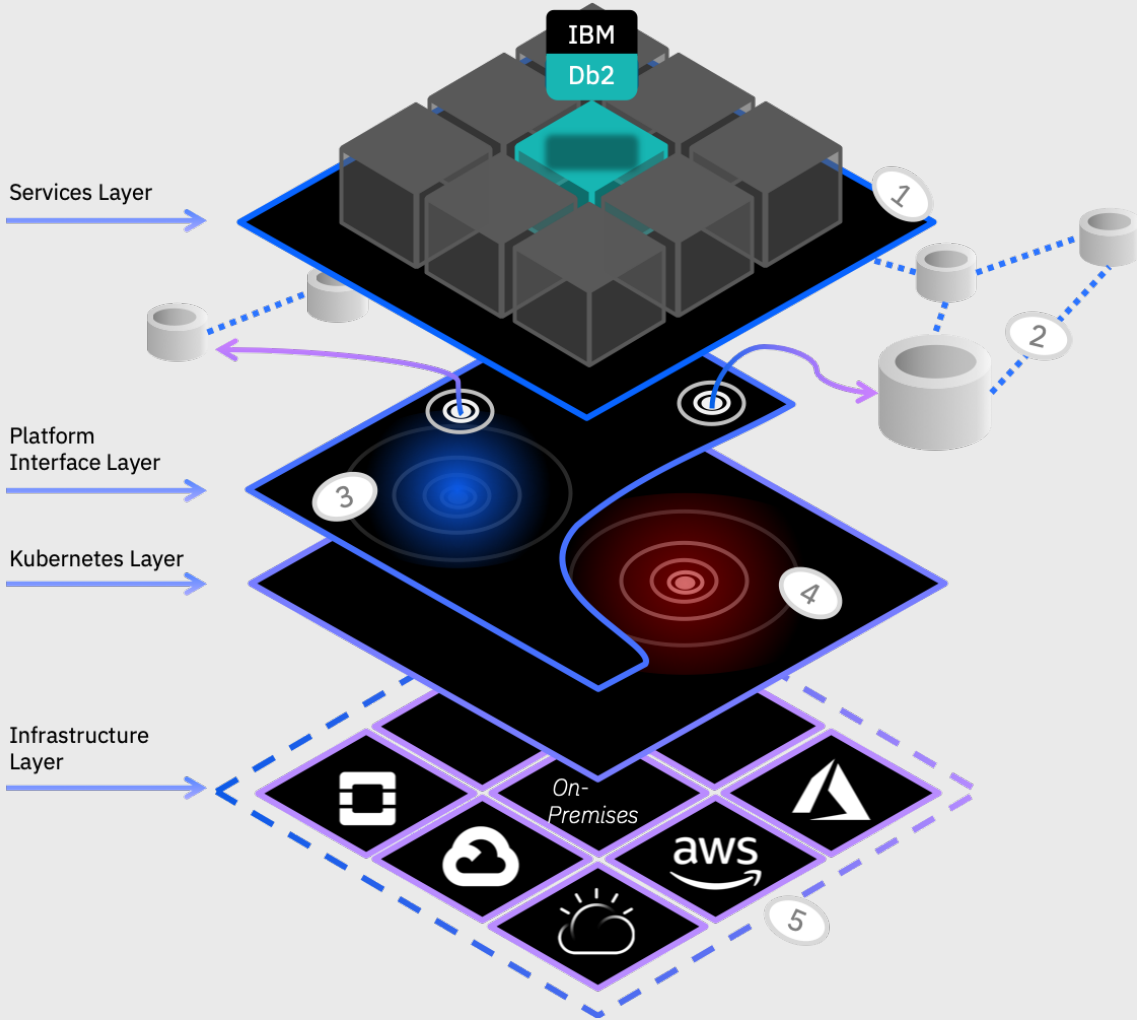
Jupyter Notebooks
Collaborative Dev Environment

© 2020 IBM Corporation

Trade-up to
IBM Cloud Pak for Data Db2

- Db2 라이센스 업그레이드로 Cloud Pak for Data 와 Db2에 대한 사용권 동시 확보

- 고객은 Bare Metal 환경 Db2에서 점진적으로 컨테이너화 가능하며 동시에 Cloud Pak for Data의 서비스를 추가 이용



© 2020 IBM Corporation

Thank you

© 2020 IBM Corporation