

 굿모닝아이텍(주)

 joinusbiz
조인어소비스

 디지털아이텍(주)
DIGITAL INFORMATION TECHNOLOGY

 디지털비즈니스(주)
Digital Business

굿모닝아이텍 무빙 클래스

The Lakehouse CLASS

2022년 11월 14일(월)

DX 사업본부 클라우드 사업부

김정희 부장



Contents

1. Data Lakehouse의 탄생
2. Modern Data Stack
3. Data Lakehouse Sample
4. Use-Cases(Apache Iceberg)
5. 클래스 맺음말



Chapter 0. ICE BREAKING

Chapter 0. ICE BREAKING



<https://www.linkedin.com/feed/update/urn:li:activity:6988181955502632960>



Chapter 1. Data Lakehouse의 탄생

Chapter 1. Data Lakehouse의 탄생

Data Lakehouse (이솜이 아닌) 우화.



<https://twitter.com/GaryStafford/status/1553844632097882116>

Chapter 1. Data Lakehouse의 탄생

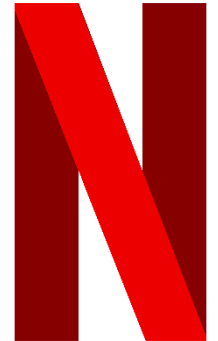
우리 모두가 마케팅 용어의 홍수 속에서 혼란에 빠져 있다.



<https://twitter.com/GaryStafford/status/1553844632097882116>

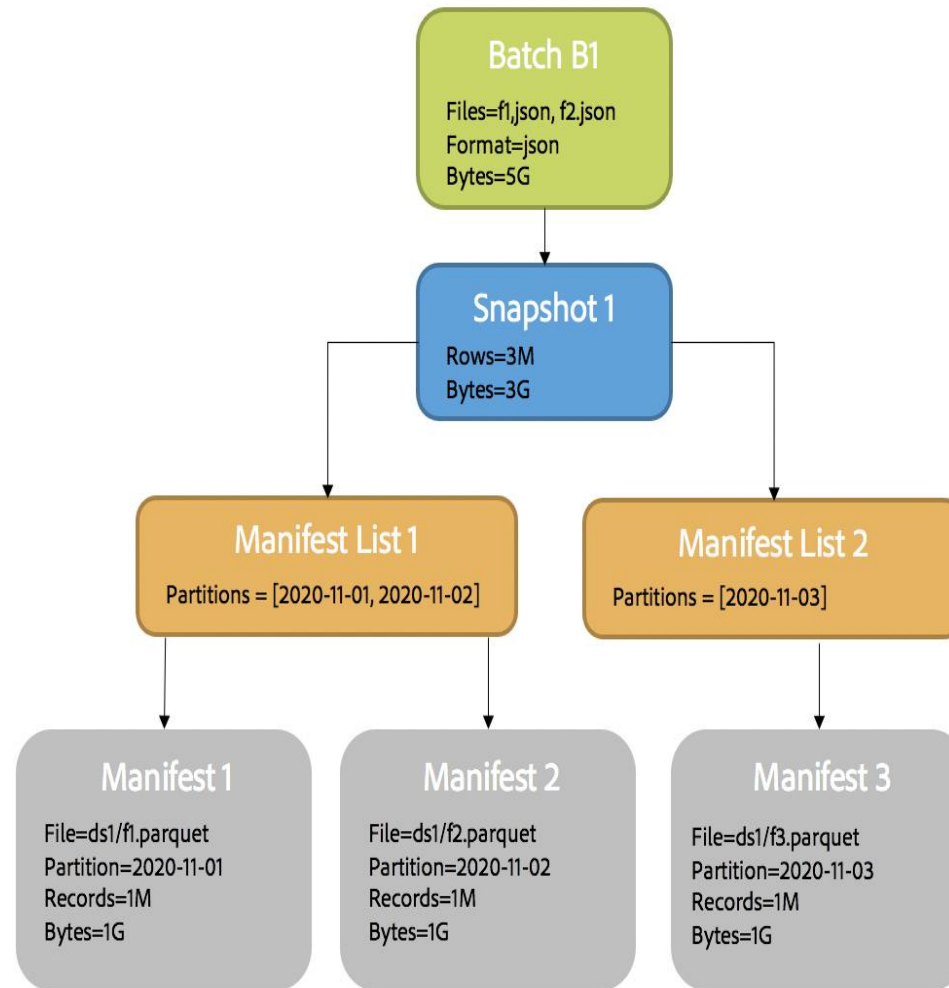
Chapter 1. Data Lakehouse의 탄생

2018년 Lakehouse의 태동.



Chapter 1. Data Lakehouse의 탄생

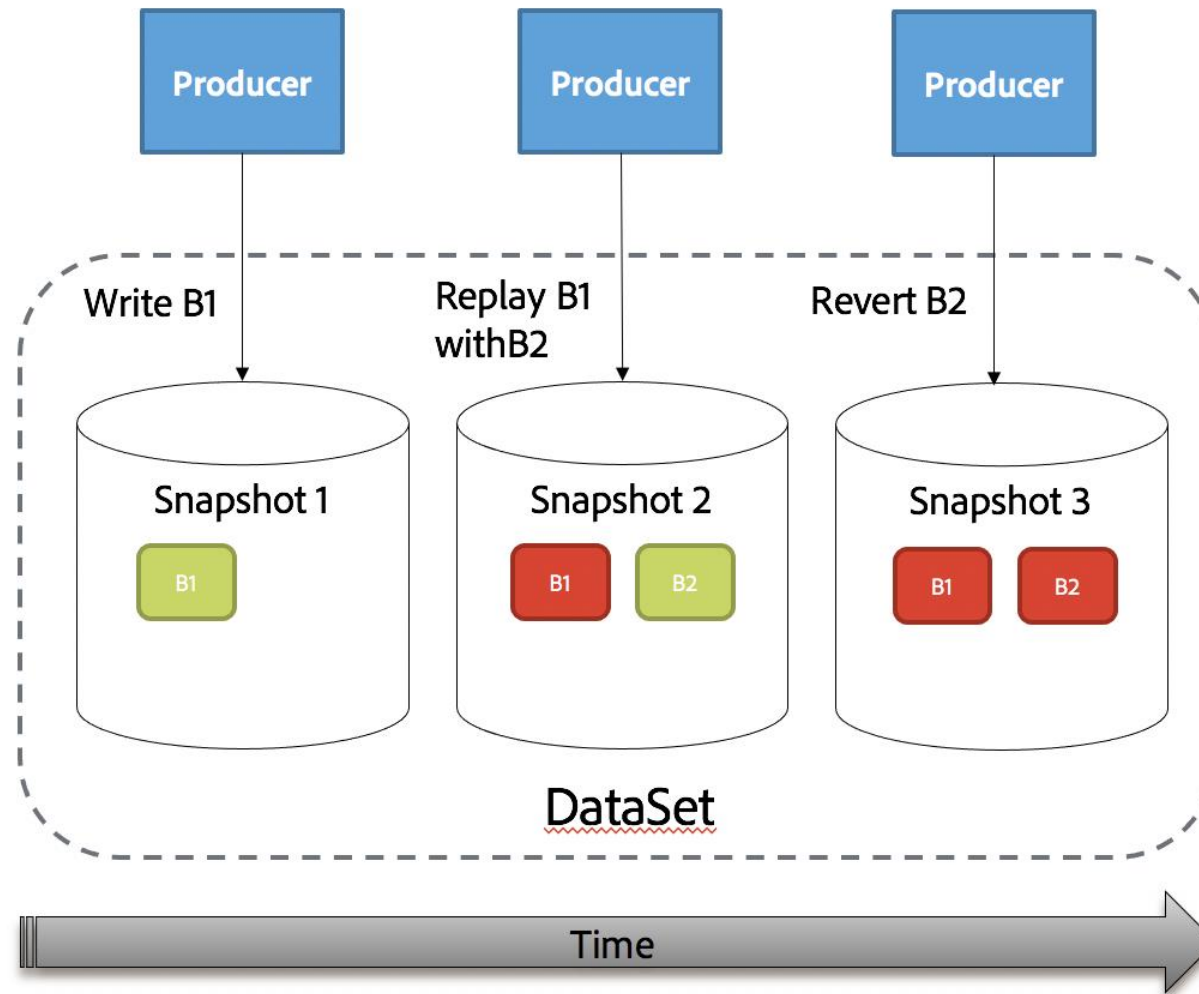
Apache Iceberg가 각광 받는 이유 (1) Manifest로 데이터 관리 최적화



<https://blog.developer.adobe.com/iceberg-at-adobe-88cf1950e866>

Chapter 1. Data Lakehouse의 탄생

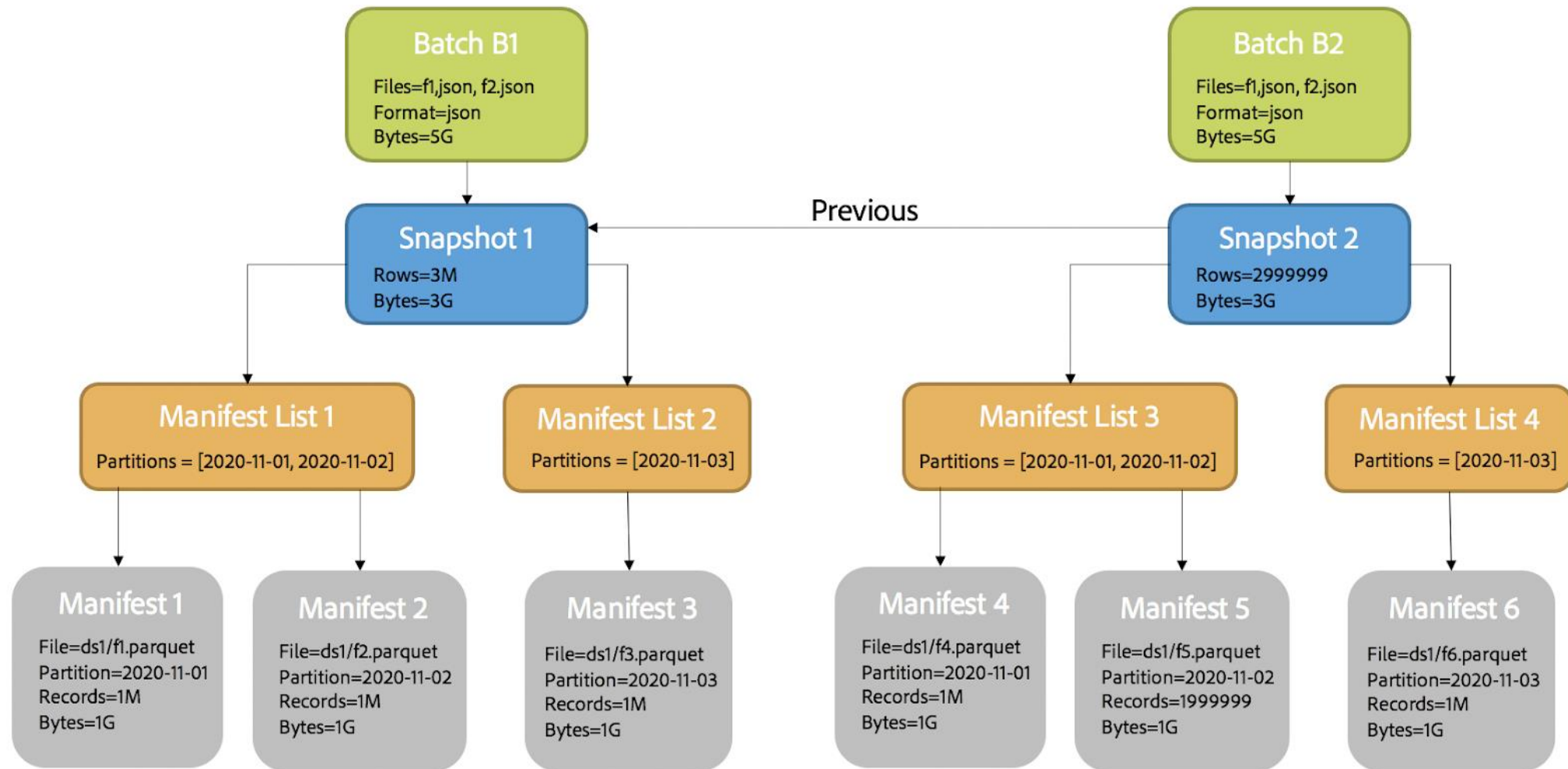
Apache Iceberg가 각광 받는 이유 (2) Snapshot을 통한 데이터 수정의 편리함



<https://blog.developer.adobe.com/iceberg-at-adobe-88cf1950e866>

Chapter 1. Data Lakehouse의 탄생

Apache Iceberg가 각광 받는 이유 (3) 시간여행



<https://blog.developer.adobe.com/iceberg-at-adobe-88cf1950e866>

Chapter 1. Data Lakehouse의 탄생

Apache Iceberg 핵심 정리

User experience	Schema evolution
	Hidden partitioning
	Partition layout evolution
	Time travel
Reliability and performance	Scan planning is fast
	Advanced filtering
	Works with any cloud store
	Serializable isolation
	Multiple concurrent writers
Open standard	Designed and developed to be an open community

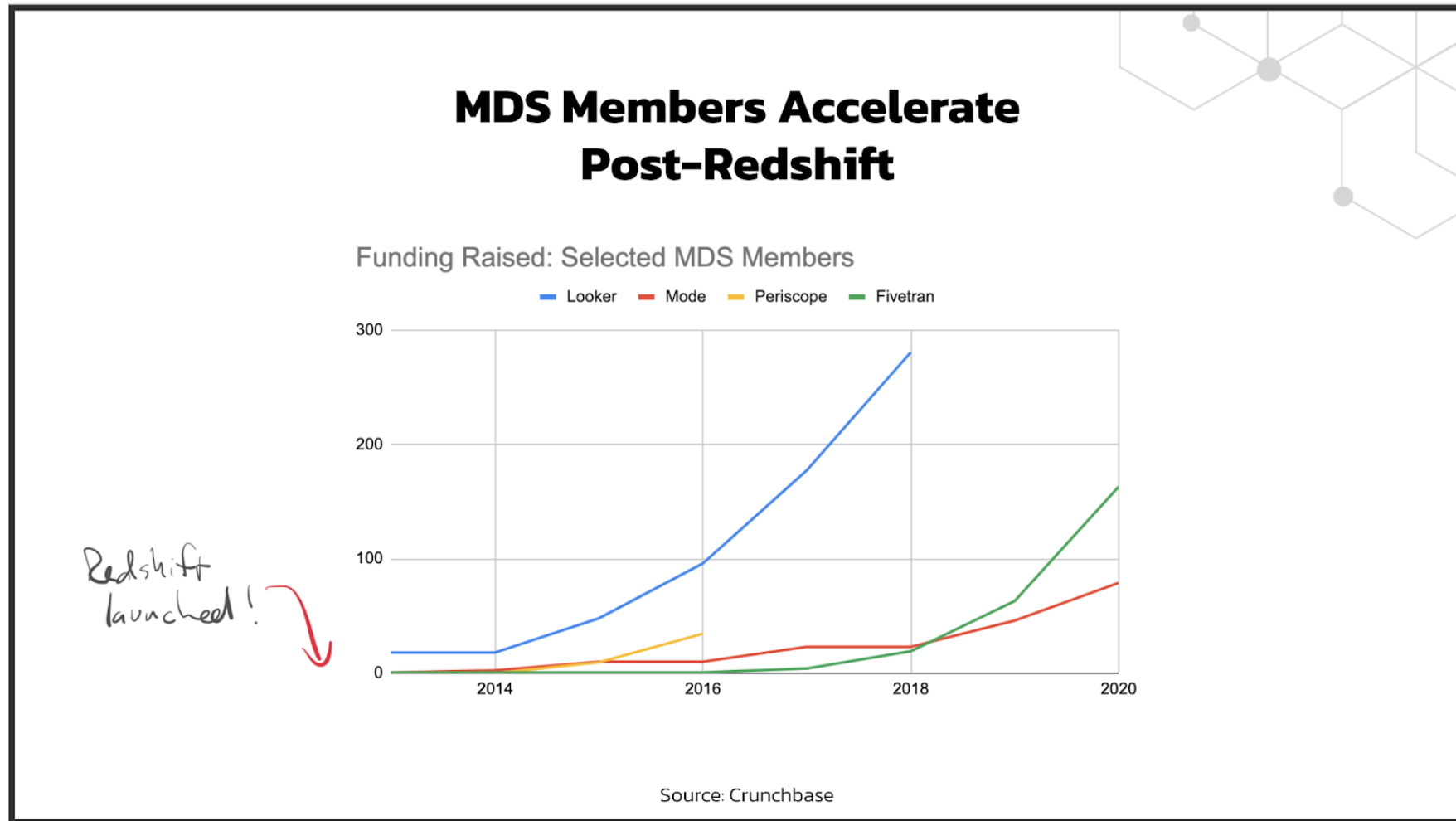
<https://iceberg.apache.org/docs/latest/>



Chapter 2. Modern Data Stack

Chapter 2. Modern Data Stack

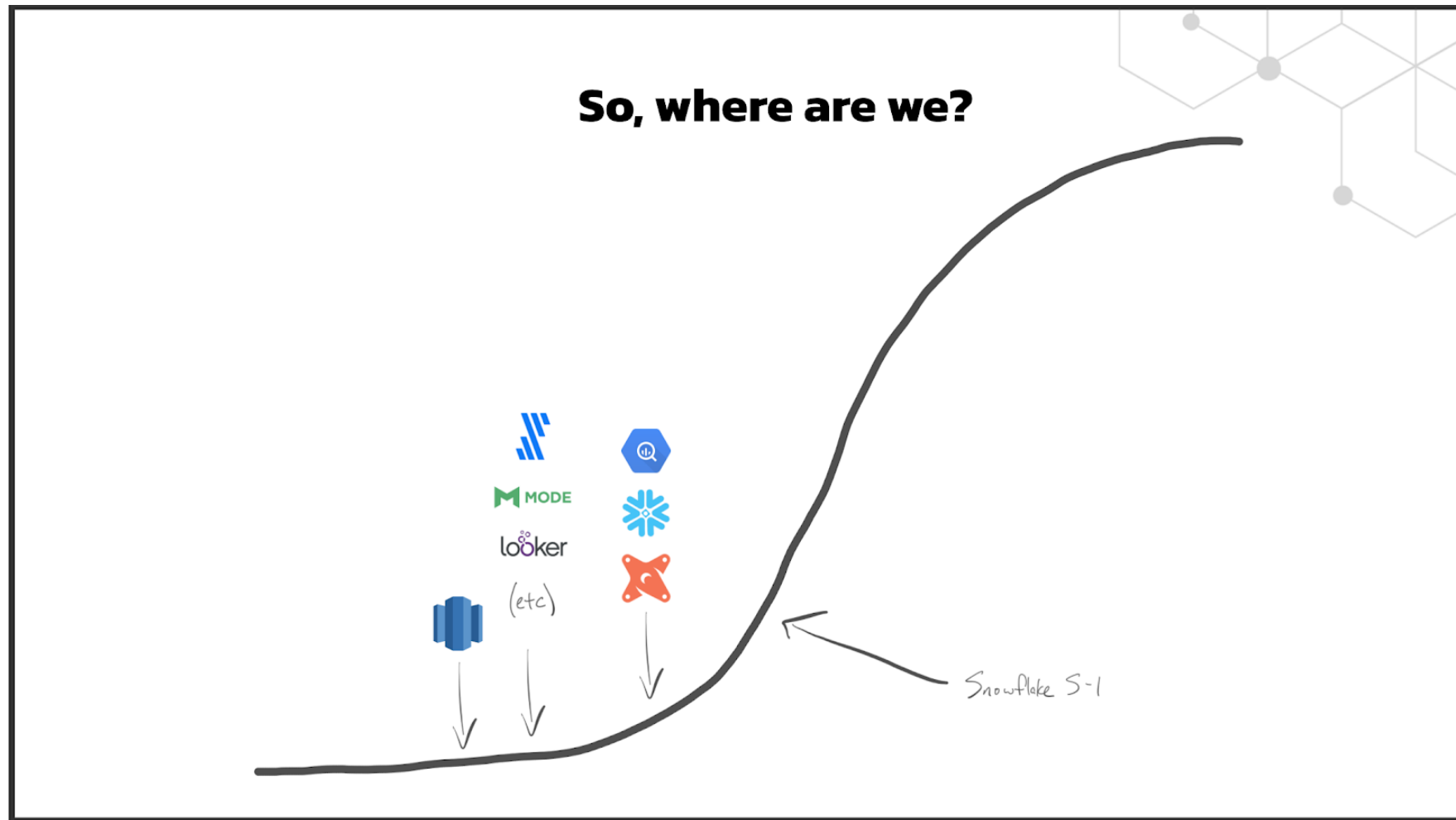
Modern Data Stack(2012-2016)은 AWS Redshift 서비스 런칭(2012) 이후 본격화



<https://www.getdbt.com/blog/future-of-the-modern-data-stack/>

Chapter 2. Modern Data Stack

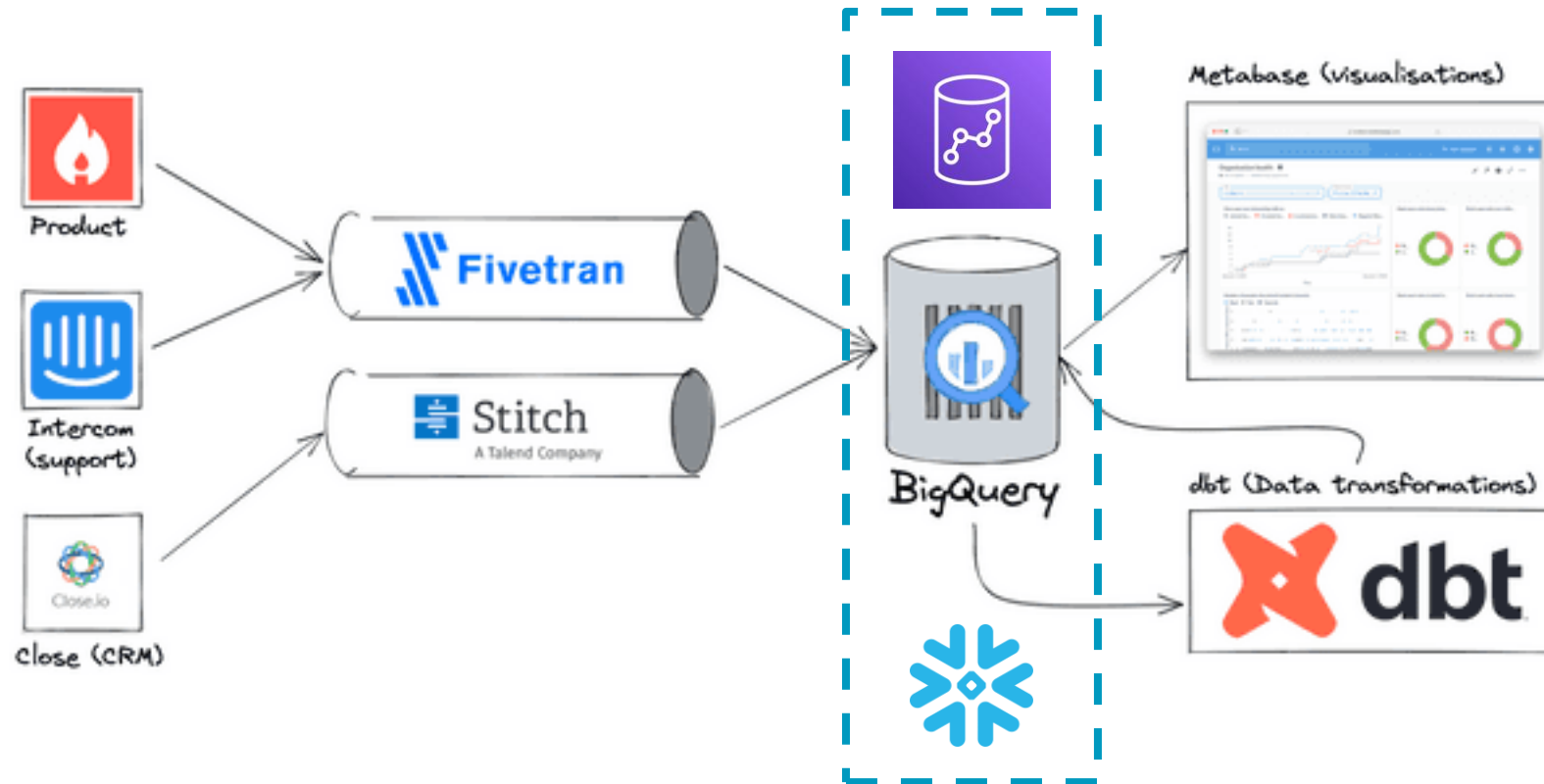
Modern Data Stack(2016-2020) 의 기술 성숙도는 어디까지 왔나?



<https://www.getdbt.com/blog/future-of-the-modern-data-stack/>

Chapter 2. Modern Data Stack

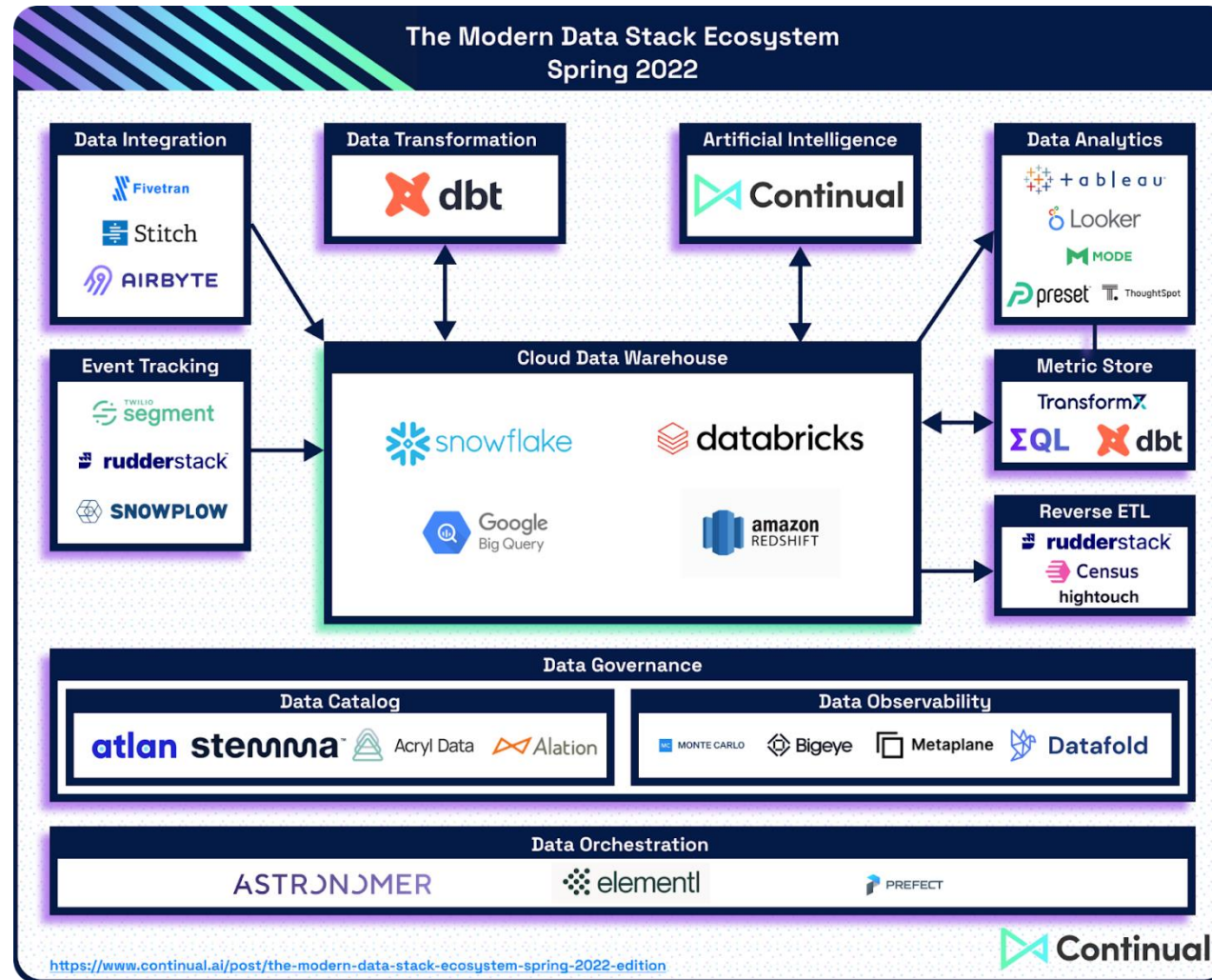
스타트업 기업을 위한 Modern Data Stack.



<https://incident.io/blog/data-stack>

Chapter 2. Modern Data Stack

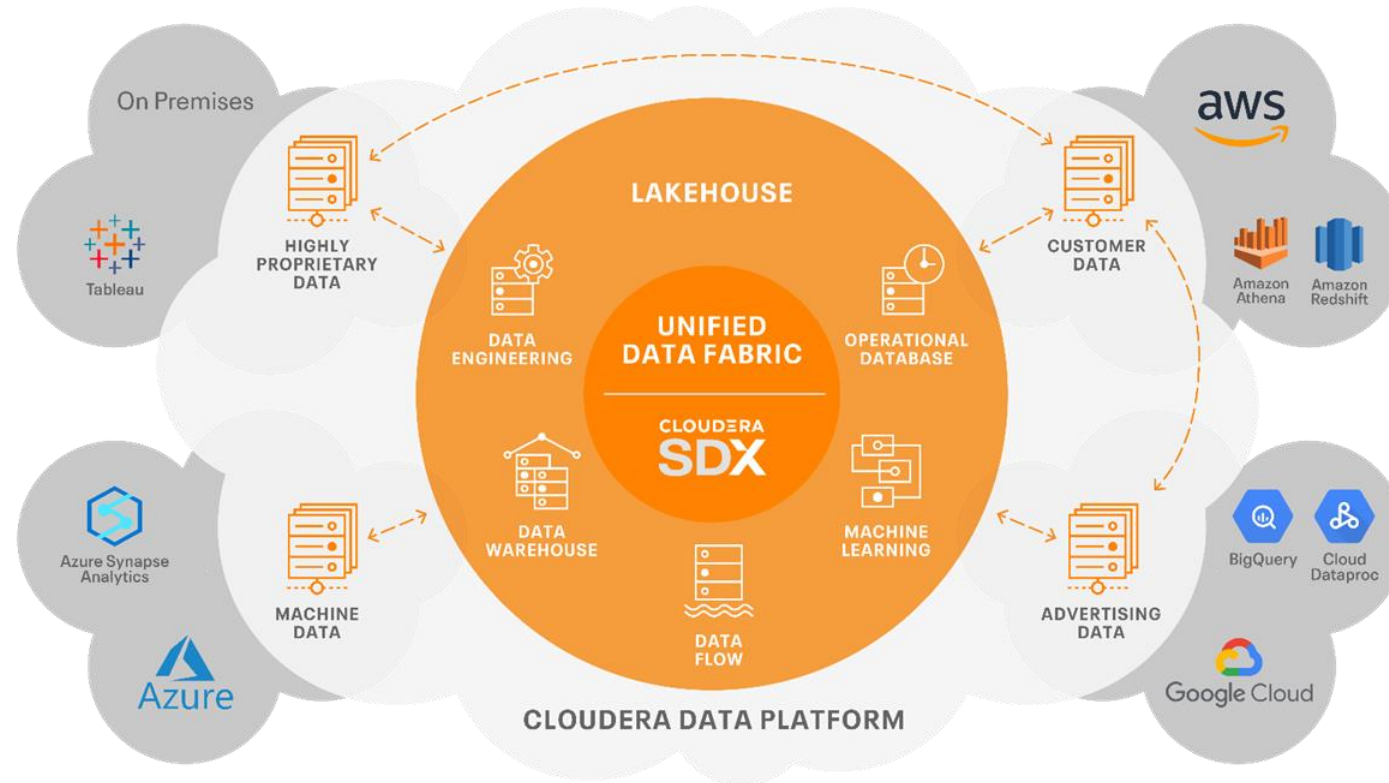
AI 회사 Continual이 제시한 Modern Data Stack 생태계.



<https://continual.ai/post/the-modern-data-stack-ecosystem-spring-2022-edition>

Chapter 2. Modern Data Stack

CLUDERA의 Modern Data Stack 소개



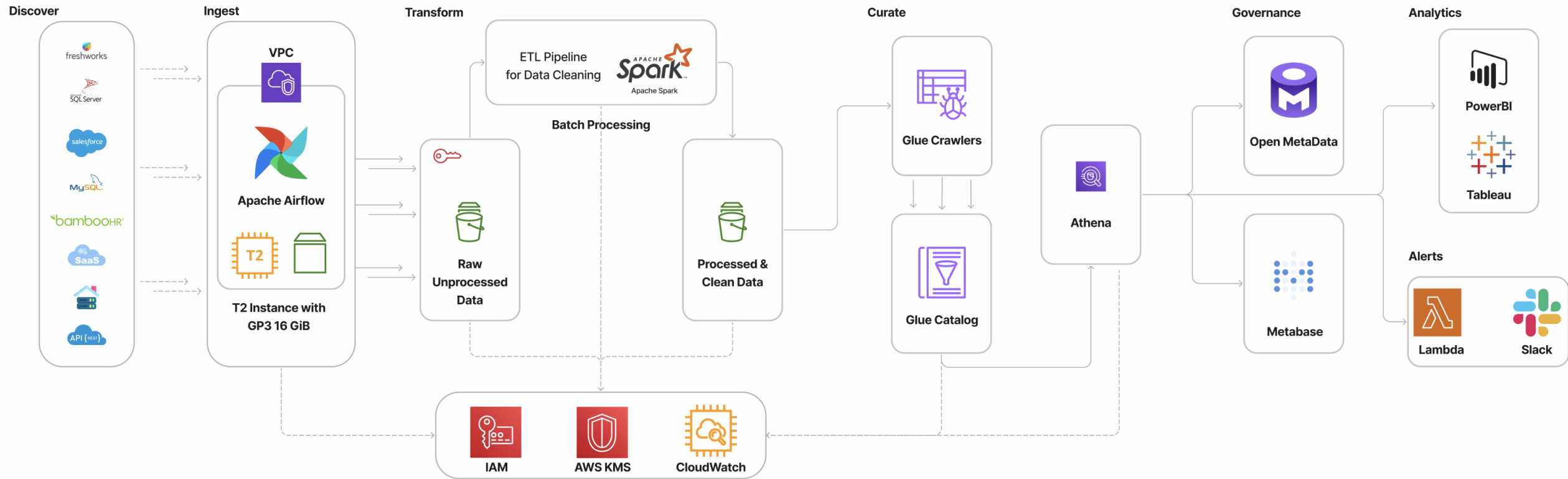
<https://www.cloudera.com/products/open-data-lakehouse.html>



Chapter 3. Data Lakehouse Sample

Chapter 3. Data Lakehouse Sample

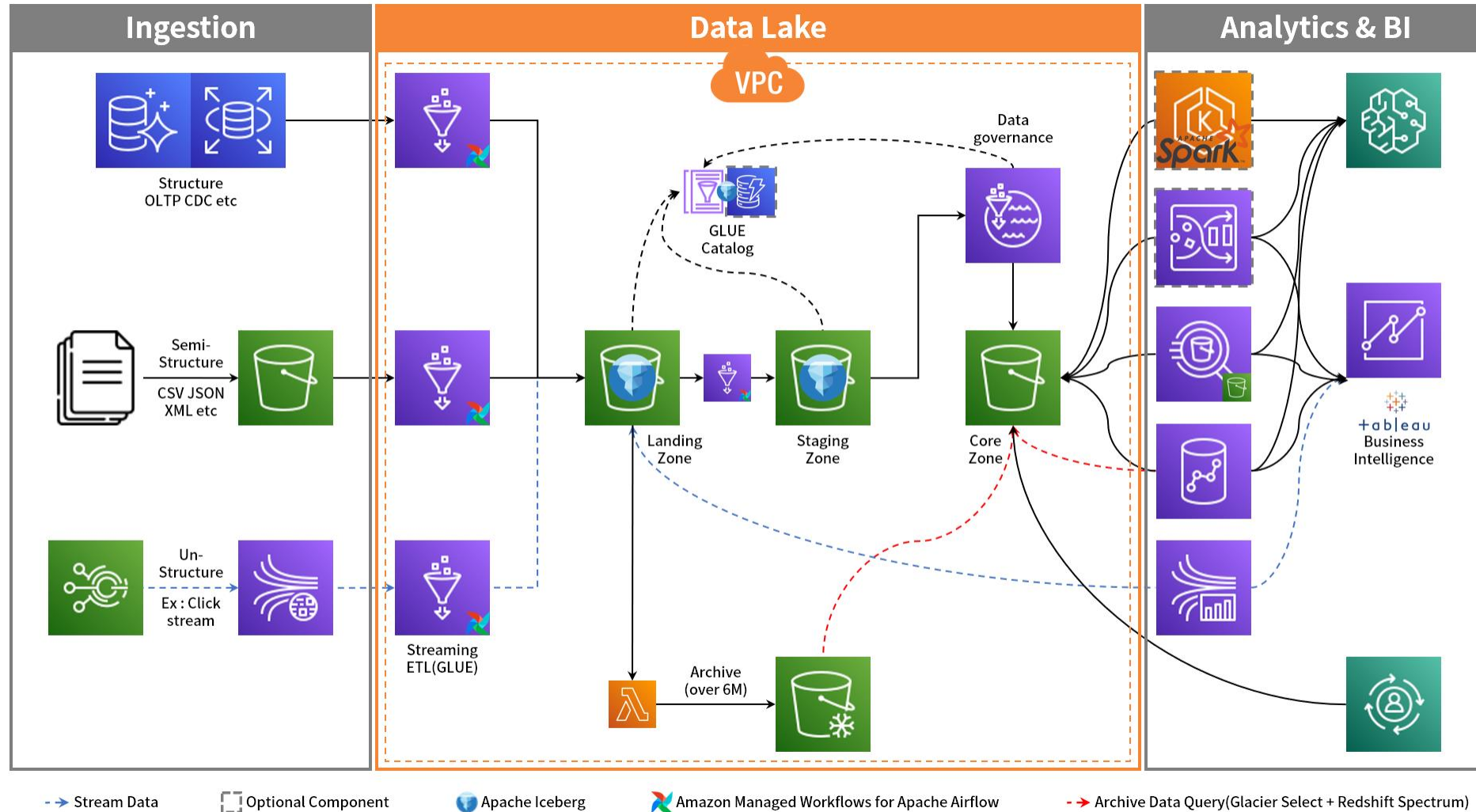
어느 회사의 데이터 과학자의 Data Lakehouse 아키텍처 on AWS



<https://abhijit.work/how-weve-increased-analytics-productivity-by-almost-100-with-data-lakehouse-with-modern-data-6b7afecab14c>

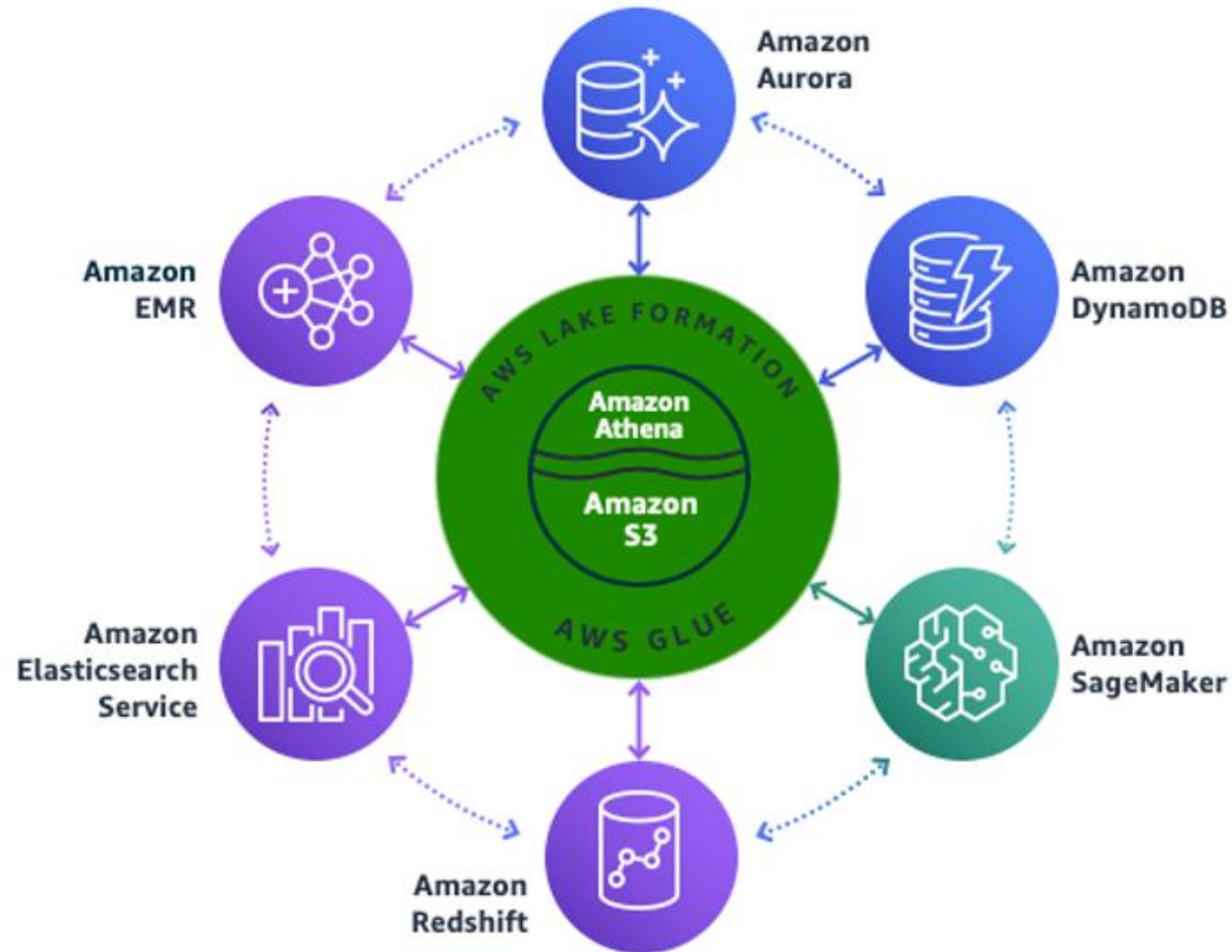
Chapter 3. Data Lakehouse Sample

Data Lakehouse 기반 Modern Data Stack on AWS



Chapter 3. Data Lakehouse Sample

Lake House Architecture on AWS



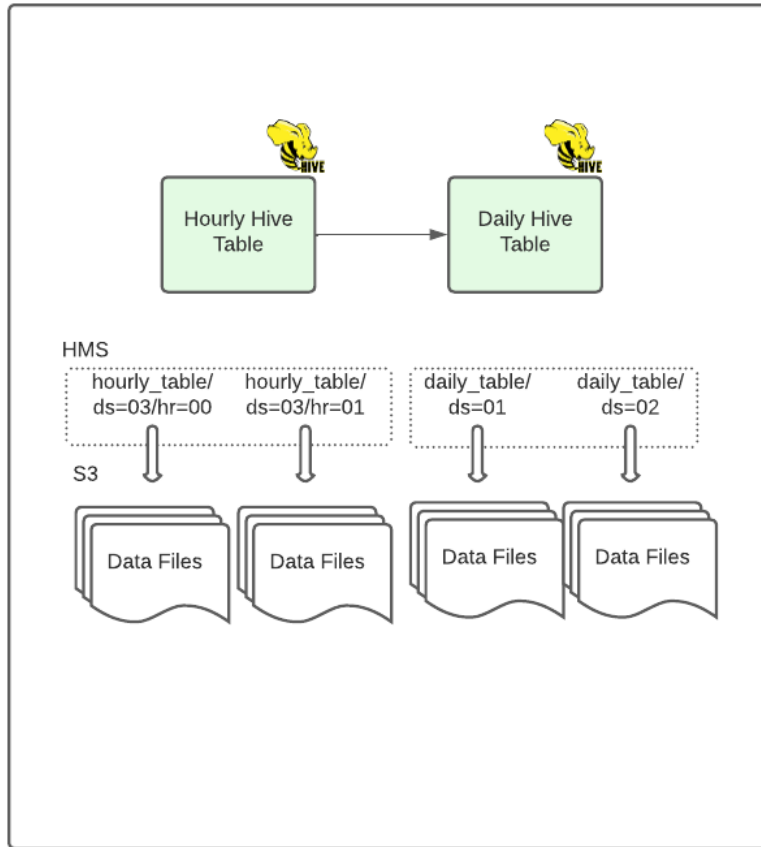
<https://aws.amazon.com/ko/blogs/architecture/address-modernization-tradeoffs-with-lake-house-architecture/>

A man with glasses and a beard is pointing at a whiteboard. The image is overlaid with a large diagonal graphic consisting of a green-to-blue gradient bar and a solid blue bar. The background is a blurred office setting.

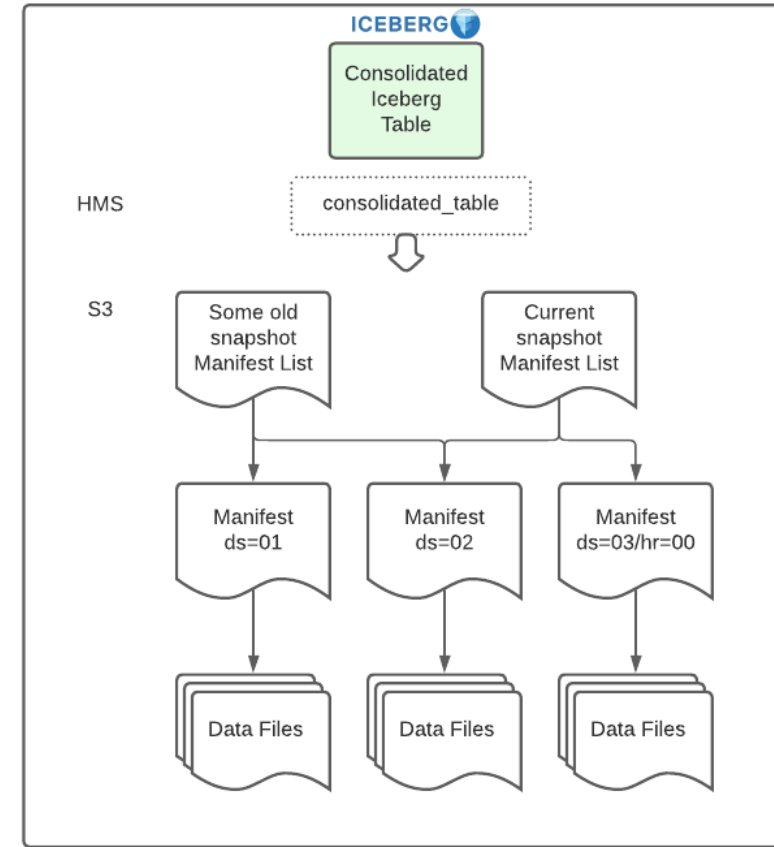
Chapter 4. Use-Cases(Iceberg)

Chapter 4. Use-cases(Apache Iceberg)

Upgrading Data Warehouse Infrastructure at Airbnb



Before

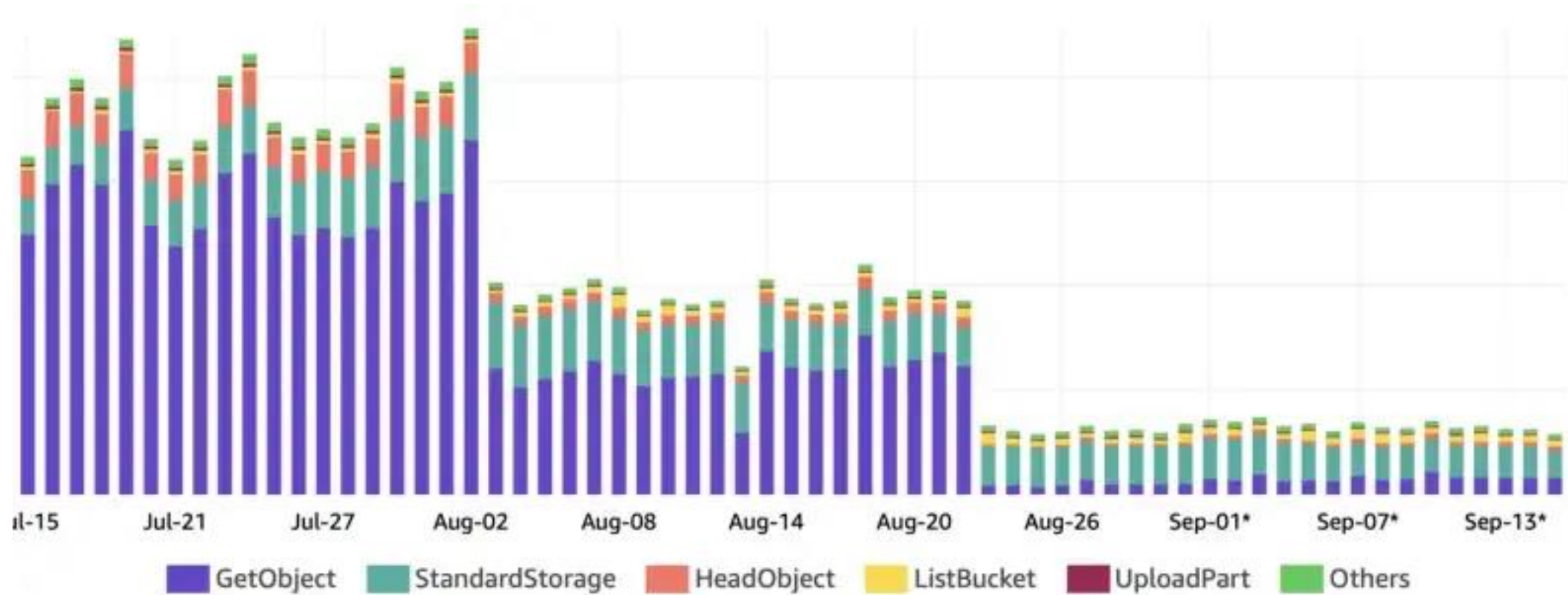


After

<https://medium.com/airbnb-engineering/upgrading-data-warehouse-infrastructure-at-airbnb-a4e18f09b6d5>

Chapter 4. Use-cases(Apache Iceberg)

Apache Iceberg Reduced Our Amazon S3 Cost by 90%



<https://medium.com/insiderengineering/apache-iceberg-reduced-our-amazon-s3-cost-by-90-997cde5ce931>

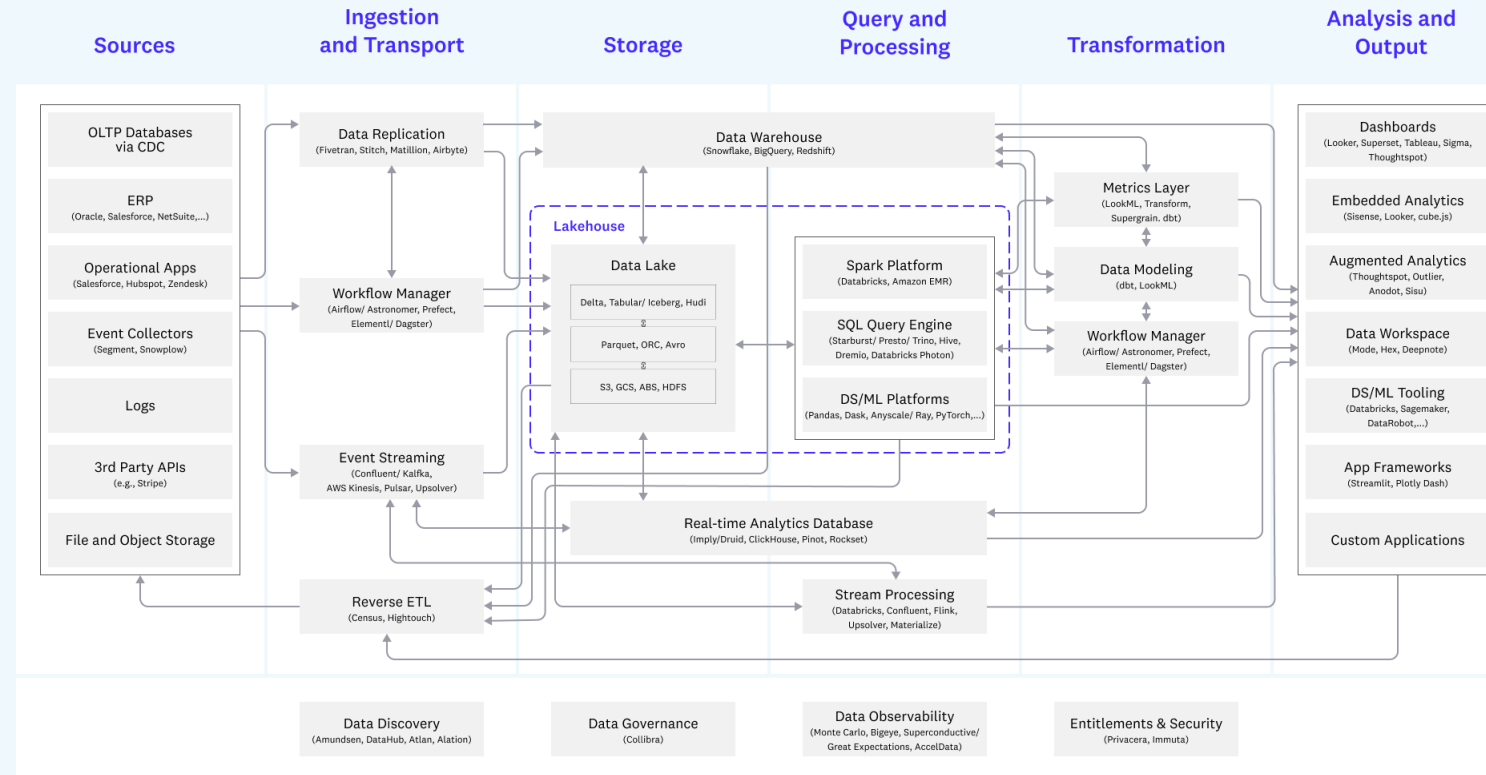


Chapter 5. 클래스 맷음말

Chapter 5. 클래스 맺음말

Data Infrastructure의 진화는 현재진행형.

Unified Data Infrastructure (2.0)



<https://future.com/emerging-architectures-modern-data-infrastructure/>

Chapter 5. 클래스 맺음말

굿모닝아이텍 무빙 클래스를 마치며... 질문 받습니다.



 굿모닝아이텍(주)

 joinusbiz
조인어소비스

 디지털아이텍(주)
DIGITAL INFORMATION TECHNOLOGY

 디지털비즈니스(주)
Digital Business

감 사 합 니 다