SOSCON

삼성 오픈소스 컨퍼런스

SAMSUNG OPEN SOURCE CONFERENCE

OPEN YOUR UNIVERSE WITH SOSCON

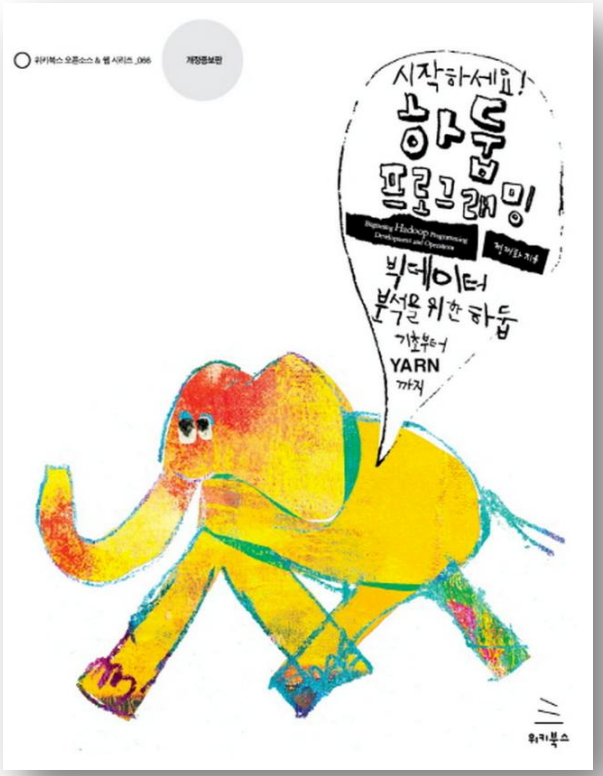
Expanding Your Data Warehouse with Tajo

그루터 / 정재화

2015.10.27

About me

- Gruter Corp / BigData Engineer (jhjung@gruter.com)
- Committer and PMC member of Apache Tajo
- The author of Hadoop book
- Home Page: http://blrunner.com
- Twitter: @blrunner78



Agenda

1. 데이터 웨어하우스란?
2. 데이터 웨어하우스의 문제점
3. 스토리지 엔진을 확장하는 방법
4. SQL 엔진을 확장하는 방법
5. 적용 사례

데이터 웨어하우스란?

출처 http://www.internetguncatalog.com/Portals/0/Warehouse%2006-08-07%20(26).jpg

정보(data)
+ 창고(warehouse)

A data warehouse is a system used for reporting and data analysis. DWs are central repositories of integrated data from one or more disparate sources.

- Wikipedia -

= data warehouse

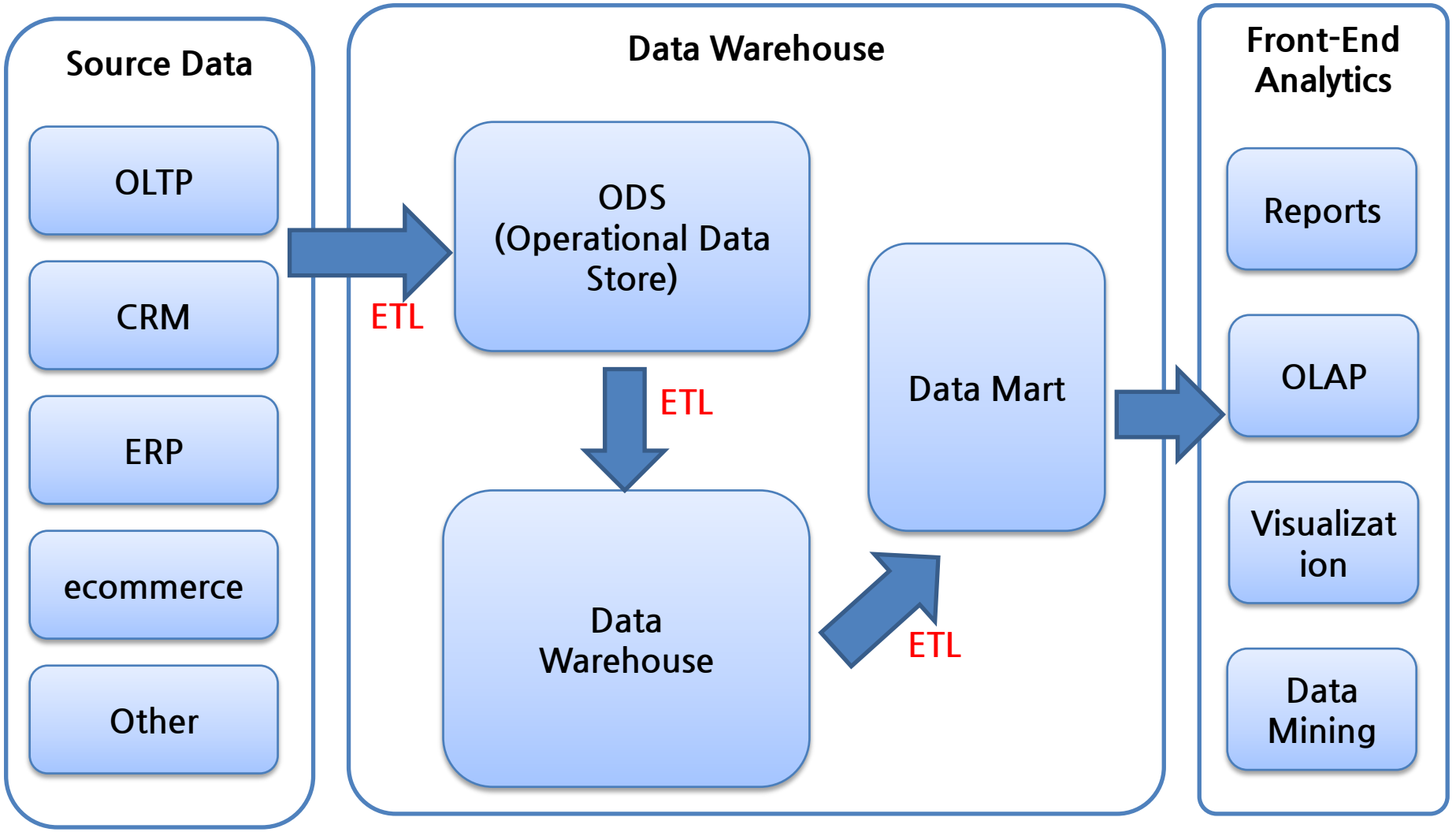
= DW or DWH

= Enterprise data warehouse

= EDW

DW의 아키텍처는 어떻게 구성될까?

전통적인 DW 아키텍처


Source Data
OLTP
CRM
ERP
ecommerce
Other
ETL
Data Warehouse
ODS
(Operational Data
Store)
ETL
Data
Warehouse
ETL
Data Mart
Front-End
Analytics
Reports
OLAP
Visualizat
ion
Data
Mining

기존 데이터 웨어하우스의 문제점

데이터 볼륨과 복잡도가 급격히 증가하고 있습니다.

- Web logs / Click Stream
- User Generated Contents
- Sensors / RFID / Devices
- Spatial & GPS
- Speech to Text
- Etc …

그냥 버리고 싶지만

분석 요건은
더 많고, 더 다양한
데이터를 필요로 합니다.

- 서비스/제품 품질 개선
- 고객 Retention / 추천
- 경영 의사 결정
- Etc …

그래서 DW 확장을 시도하지만…

출처: https://thinkbannedthoughts.files.wordpress.com/2012/05/banging-head-on-wall.jpg

용량 단위로 증가되는
라이선스 비용

-1 테라 바이트: 2만 ~ 5만 달러 (2013년 기준)

→2천만 ~ 5천만원

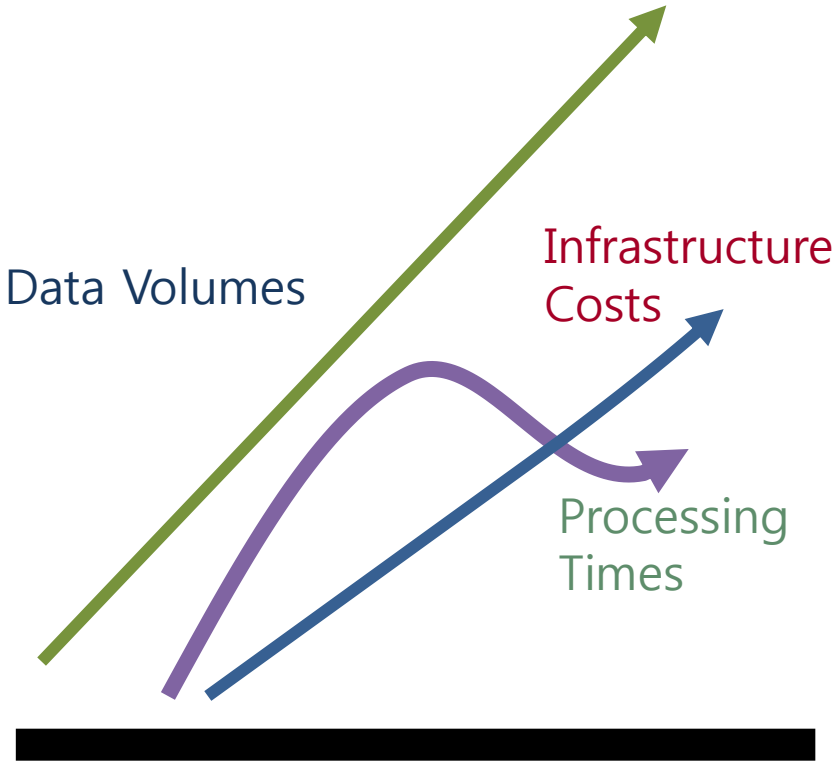
-1 페타 바이트: 2천만 ~ 5천만 달러

→ 200 ~ 500억원

출처: http://www.slideshare.net/cloudera/hadoop-extending-your-data-warehouse

데이터 용량을
증가시키더라도…

늘어나는 데이터 볼륨만큼,
처리 속도가 증가하지 않습니다.



출처: http://www.slideshare.net/cloudera/hadoop-extending-your-data-warehouse

또한 다양한 데이터 포맷을 정형화된 포맷으로 변경해야 합니다.

어떻게 하면 이런 DW를 구축할 수 있을까?

- 낮은 TCO 보장
- 선형적인 용량 및 성능 확장
- 다양한 데이터 형태의 데이터 수용
- 빠른 프로세싱 및 손쉬운 데이터 접근

GRUTER

SOSCON

스토리지 엔진을 확장하는 방법

스토리지 엔진으로 어떤 엔진이 적절할까?

Hadoop이란
대용량 데이터를 분산 처리할 수
있는 자바 기반의
오픈소스 프레임워크입니다.

분산 파일 시스템 (HDFS)
+
분산 처리 시스템 (MapReduce)

Hadoop 의 주요 특징은 무엇이 있을까?

Data Locality

- 데이터가 있는 곳에서 로직을 수행합니다.

Fault Tolerant

- Hadoop은 x86 서버에 설치할 수 있습니다. (vs. 유닉스 서버)
- 하드웨어 장애는 피할 수 없다는 가정하에 설계됐습니다.

Scalable

- 서버(노드)를 추가하면, 용량과 컴퓨팅 성능의 선형적인 확장이 가능합니다.

Hadoop으로 어떻게 스토리지 엔진을 구성해야 할까?

모든 데이터를 Hadoop으로 모아줍니다!

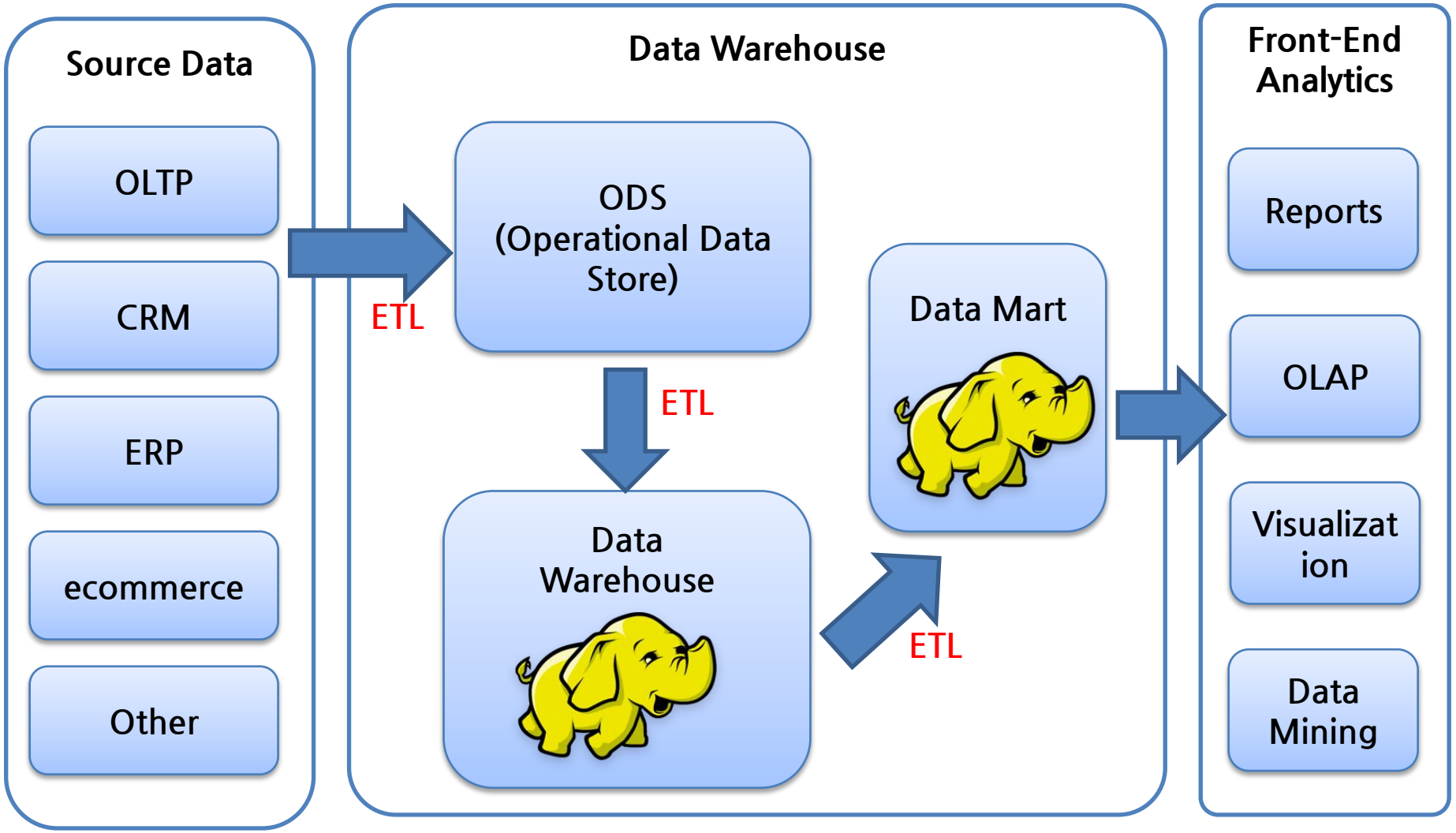
어떻게 모아줄까?

- DB 데이터 : Sqoop, SQL-on-hadoop
- Raw 파일 : `cp` 명령어 한 줄이면 끝!
- 로그 수집 : Flume, Scribe, Chuckwa, …

Case by Case

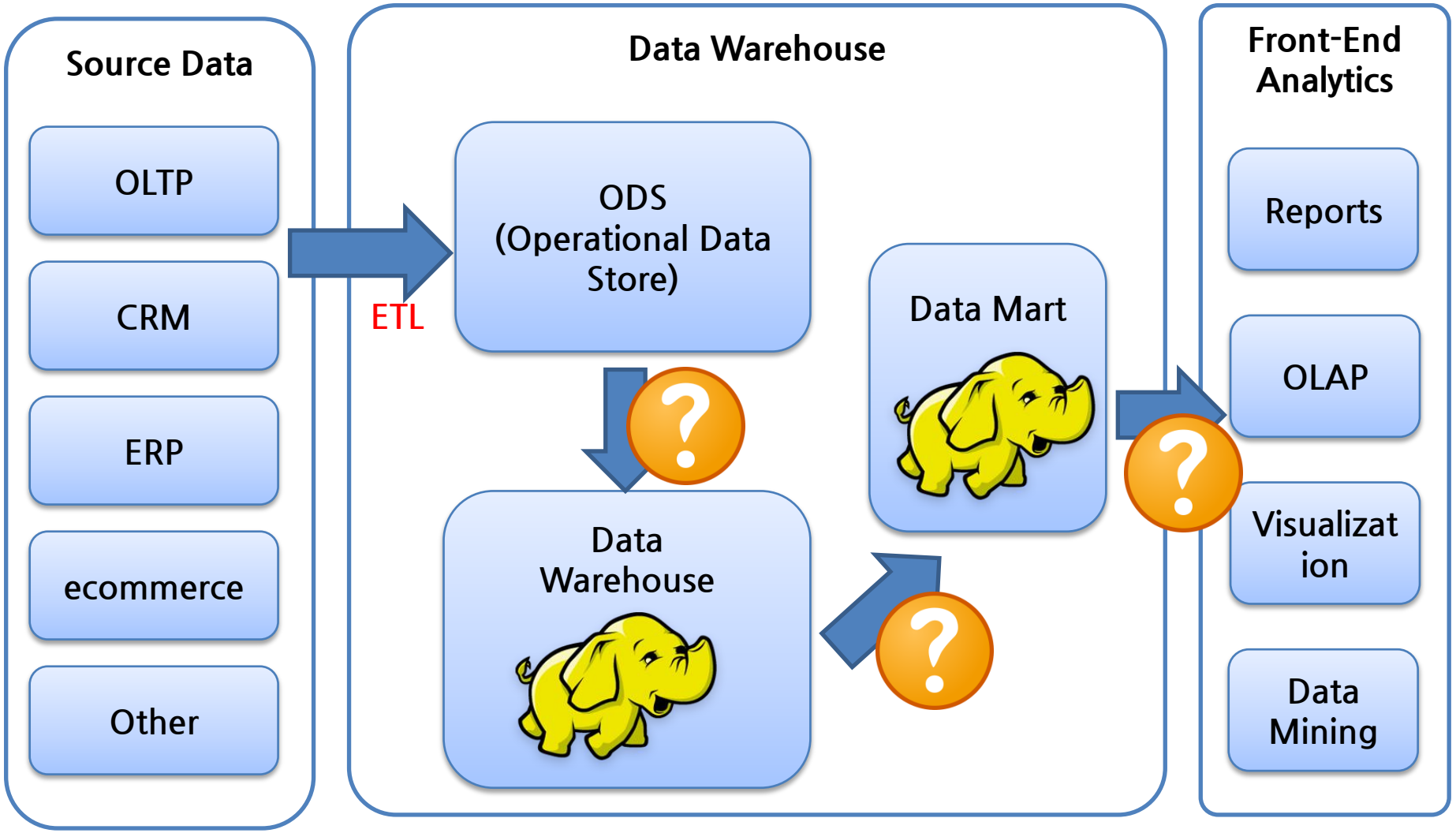
최종적으로 이런 그림이 나옵니다.

Hadoop 기반 DW 아키텍처



그런데 …

ETL과 분석질의는 어떻게 처리해야할까?

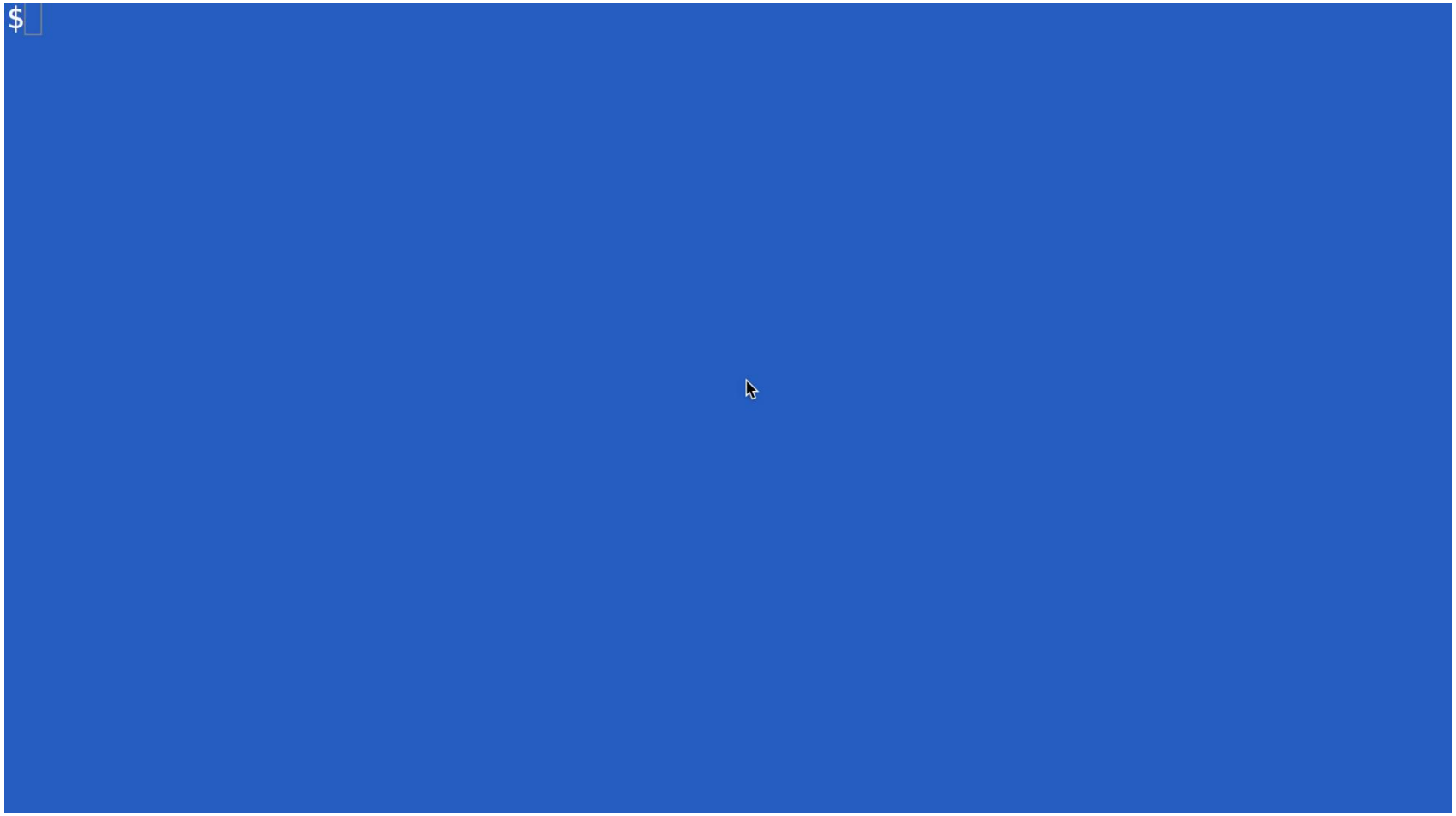


GRUTER

SOSCON

SQL 엔진을 확장하는 방법

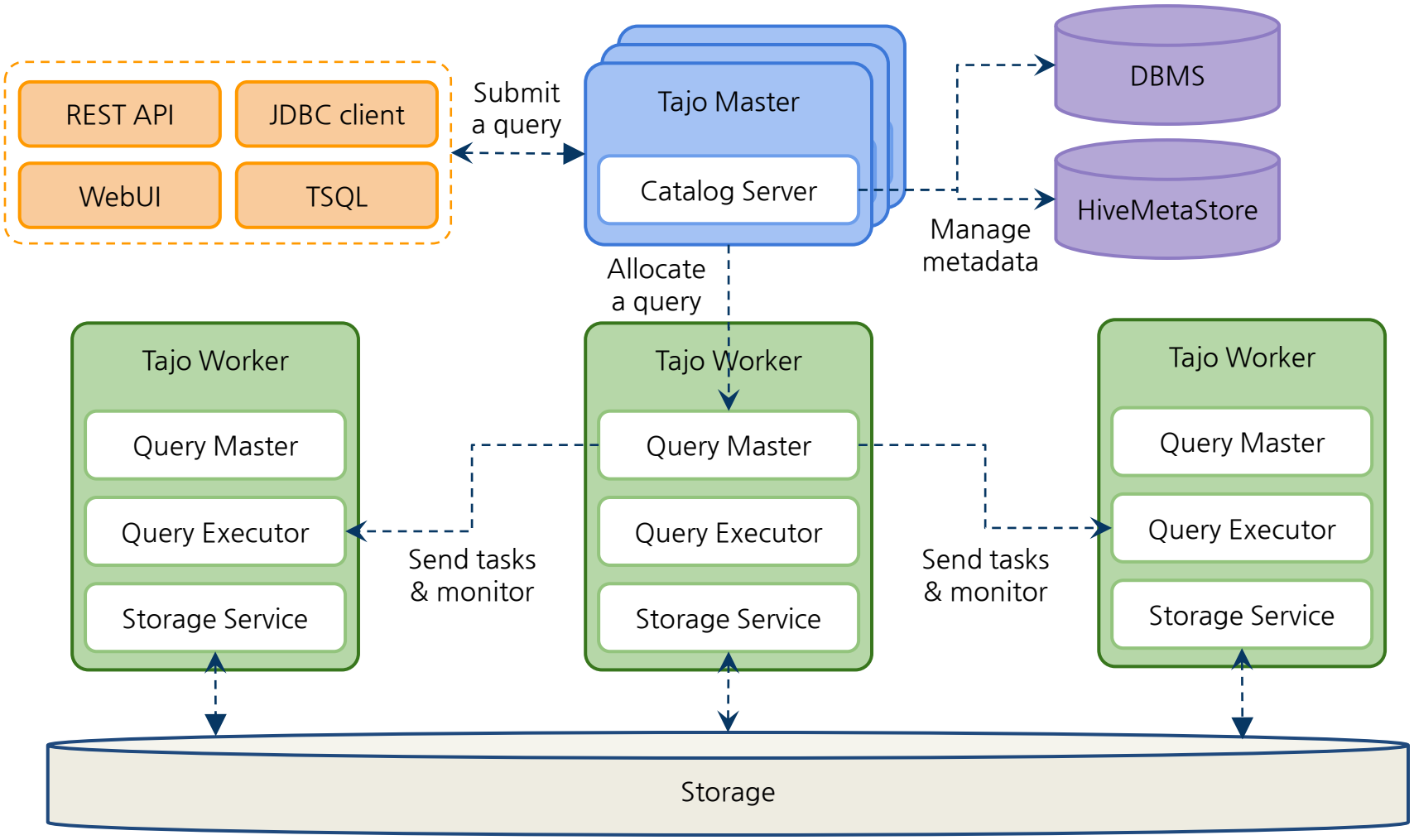
SQL 엔진으로 어떤 엔진이 적절할까?



Tajo란 Hadoop 기반의
빅데이터 웨어하우스
시스템입니다.

Tajo는 아파치 탑 레벨 프로젝트이며, ANSI SQL을 지원하며, 자체 분산 처리 엔진을 제공합니다.

Tajo 아키텍처


REST API
JDBC client
WebUI
TSQL
Submit
a query
Tajo Master
Catalog Server
DBMS
HiveMetaStore
Manage
metadata
Allocate
a query
Tajo Worker
Query Master
Query Executor
Storage Service
Send tasks
& monitor
Tajo Worker
Query Master
Query Executor
Storage Service
Send tasks
& monitor
Tajo Worker
Query Master
Query Executor
Storage Service
Storage

Tajo의 비교우위는?

ANSI SQL 지원 = 새로운 엔진에 대한 학습 비용을 최소화하고, 기존 시스템을 쉽게 전환할 수 있습니다.

클러스터 확장성 = 노드를 추가할 수록 선형적인 용량 및 성능 확장이 가능합니다.

고성능 분산 처리 엔진 = 수시간 이상 소요되는 ETL 질의부터, 수백 밀리 세컨드내에 처리되는 Interactive 질의까지 모두 지원합니다.

구체적인 질의 처리 속도는?

- 스캔속도: 물리적 디스크당 100MB/sec (SATA 기준)
- 1TB를 10 여대의 노드로 처리
  - 노드당 10 여개의 디스크가 설치되어 있다고 가정함
  - 간단한 aggregation 쿼리: 30초 ~ 1분 내외
  - 간단한 join 쿼리: 1 ~ 2 분 내외
  - 복잡한 join 및 distinct aggregation : 수 분에서 10여분

Tajo 의 주요 특징은 무엇이 있을까?

풍부한 SQL 지원

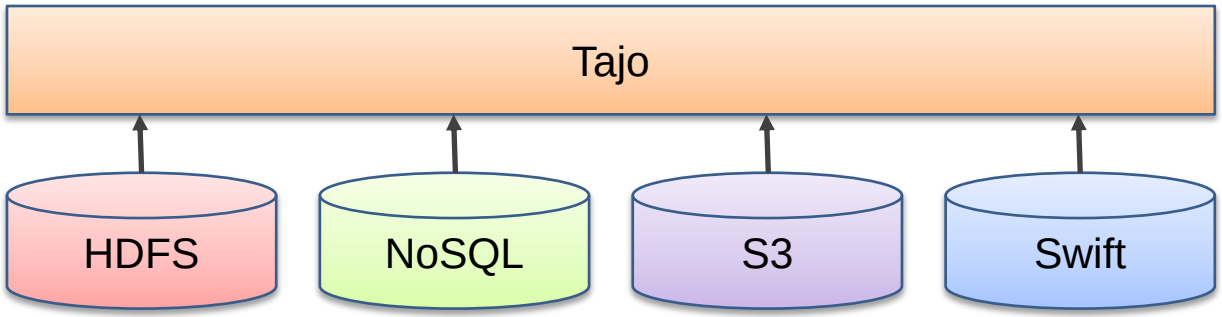
- **질의 분산 처리**
  - Inner join, and left/right/full outer join
  - GroupBy, sort, multiple distinct aggregation, window function
- **SQL 데이터 타입**
  - CHAR, BOOL, INT, DOUBLE, TEXT, DATE, Etc
- **다양한 파일 포맷**
  - Text file, SequenceFile, RCFile, ORC, Parquet, Avro

쿼리 최적화

- **Cost-based Join Optimization (Greedy Heuristic)**
  - 사용자가 최선의 Join 순서를 추측하는 수고 제거
- **확장 가능한 rewrite rule 엔진**
  - rewrite rule 인터페이스 제공과 다양한 유틸리티 제공
- **쿼리 최적화 (Progressive Query Optimization)**
  - 실행 시간 통계 수집
  - 분산 정렬을 위한 범위 분할 (range partitioning)의 적절한 파티션 범위, 개수 등을 런타임에 조정
  - 분산 Join, 그룹바이를 위한 파티션 개수를 런타임에 조정

쿼리 Federation 및 TableSpace 지원

- 다양한 데이터 소스간의 Join 및 Union 쿼리를 지원합니다.
- 장점
  - 데이터 마이그레이션 : RDBMS → 하둡
  - 기존 RBMS 데이터와 하둡 데이터의 Join 쿼리 처리
  - SQL를 이용한 NoSQL 및 다양한 스토리지 (S3, Swift, HBase, ElasticSearch, Kafka)
  - SQL 도구를 이용한 인터페이스 표준화



Nested 및 JSON 포맷 지원

별도 가공없이 Nested 및 JSON 포맷 파일의
SQL 처리 가능합니다.

SQL 문

```
SELECT
  table1.title,
  table1.name.first_name,
  name.last_name,
FROM
  table1
WHERE
  name.last_name LIKE '%Stark%';
```

테이블 정의

```
CREATE EXTERNAL TABLE table1 (
  title TEXT,
  name RECORD (
    first_name TEXT,
    last_name TEXT
  )
) USING JSON LOCATION '/xxx/json_Table';
```

```
{ "title" : "Hand of the King", "name" : { "first_name": "Eddard", "last_name": "Stark"}}
{ "title" : "Assassin", "name" : { "first_name": "Arya", "last_name": "Stark"}}
{ "title" : "Dancing Master", "name" : { "first_name": "Syrio", "last_name": "Forel"}}
```

입력 데이터

파티션 테이블

Hive 호환성 지원을 위한 Column Value 방식의 테이블 파티션 지원합니다.

- 테이블 생성 쿼리문

```
CREATE TABLE student (
  id      INT,
  name    TEXT,
  grade   TEXT
) USING PARQUET
PARTITION BY COLUMN (country TEXT, city TEXT);
```

- 파티션 디렉터리 구성

```
/tajo/warehouse/student/country=KOREA/city=SEOUL/
/tajo/warehouse/student/country=KOREA/city=PUSAN/
/tajo/warehouse/student/country=KOREA/city=INCHEON/
/tajo/warehouse/student/country=USA/city=NEWYORK/
/tajo/warehouse/student/country=USA/city=BOSTON/
 . . .
```

0.11.0 버전에서는

- **2015년 10월 릴리즈**
- **주요 기능**
  - 다중 쿼리 동시 실행 지원
  - Tablespace 및 JDBC Storage 지원
  - Nested Record 타입 지원
  - JSON 등 self-describing 데이터에 대한 schemaless 지원
  - ORC 파일 지원
  - JDBC 및 Client의 ResultSet fetch의 성능 향상
  - Python UDF/UDAF 지원
  - 향상된 조인 최적화, 쿼리 처리 성능
  - 반응속도 향상 및 버그 수정

0.12.0 버전에서는

- **주요 기능**
  - YARN 지원
  - 사용자 인증 지원
  - Scalar 및 Exist 서브쿼리 지원
  - ALTER TABLE ADD/DROP 파티션 지원
  - 하이브 UDF 호환
  - WITH 절 지원

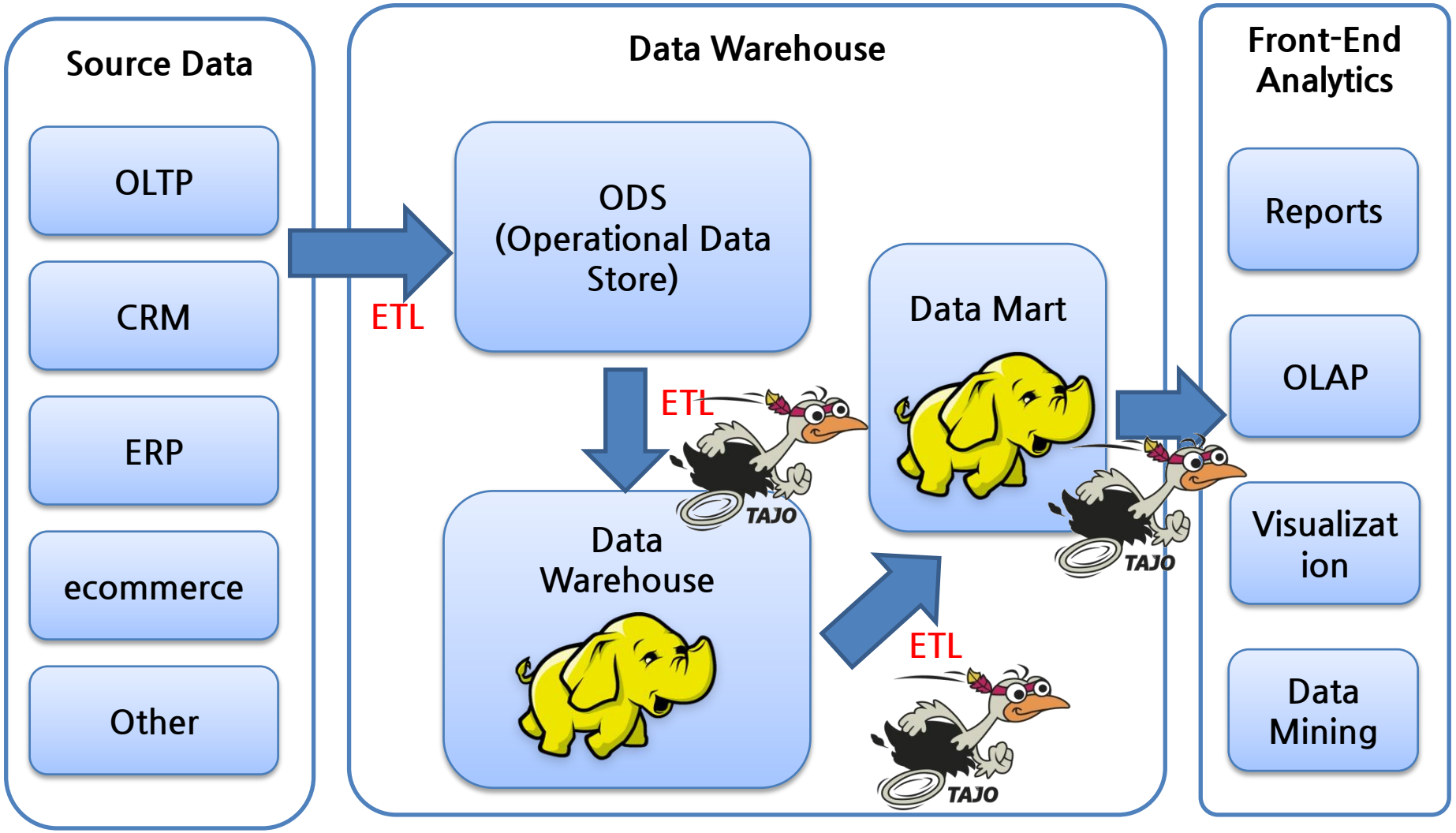
Tajo로 어떻게 SQL 엔진을 구성해야할까?

- Hadoop 경로를 설정하고,
- DW의 Root 경로를 설정하고,
- 어떤 DB를 카탈로그로 쓸지 경정하고,
- 얼마나 많은 Task를 동시에 실행할지

결정하면…

최종적으로 이런 그림이 나옵니다.

Tajo 기반 DW 아키텍처



실제 적용 사례를 알아봅시다.

상용 DW 대체

- **국내 이동통신 점유율 1위 회사**
  - ETL 작업 대체 : 일일 4TB, 120개 이상 질의 처리
  - OLAP 분석 대체: 500개 이상 질의 처리
- **도입 효과**
  - 데이터 분석을 위한 아키텍처 간소화

  → DW ETL, OLAP, Hadoop ETL을 위한 시스템 통합
  - 저비용으로 상용 수준의 SLA와 데이터 볼륨 확장

  → 상용 DW 라이선스 비용 절감

데이터 Discovery

- 국내 음원시장 점유율 1위 회사
- 2,800만 고객의 소비 이력과 활동 내역 분석
- 도입 효과
  - Hive 에서 Tajo로 분석 작업 전환

  → 최소 1.5배에서 최대 수십배 성능 향상

  - 대용량 데이터에 대한 Interactive 질의 수행

Cohort 분석

- 스마트폰 잠금화면 리워드형 광고 서비스
- S3에 저장된 원본 로그에 대한 Cohort 분석
- 분석 결과는 RDS에 저장
- 도입 효과
  - EC2 인스턴스 스펙 : c3.2xlarge
    → vCPU : 8, 메모리: 15GB, HDD: 2 x 80GB
  - EC2 인스턴스 10개로 수십GB 로그를 약 40초에 처리

  → 총 비용 : 0.420 * 10 = 4.20 달러 (한화: 4756.08원)

Welcome to Tajo

- Homepage: http://tajo.apache.org
- 한국 타조 사용자 그룹
    - 구글 그룹: https://groups.google.com/forum/#!forum/tajo-user-kr
    - 페이스북: https://www.facebook.com/groups/tajokorea
- 타조 한글 문서화 프로젝트: http://bit.ly/1Ir417T
- 기타 참고 사이트
    - http://www.gruter.com/blog/tag/apache-tajo
    - http://teamblog.gruter.com/tag/apache-tajo

```
tajo> select question from you;
```

SOSCON
삼성 오픈소스 컨퍼런스
SAMSUNG OPEN SOURCE CONFERENCE
OPEN YOUR UNIVERSE WITH SOSCON

THANK YOU!