



# (주)베스트텍시스템 Big data 플랫폼 제안서

Big data 플랫폼, Storage, Server, Solution IT 컨설팅 전문기업



# 목차

## Big data 플랫폼 제안서



- 회사 소개
- 회사연혁
- (주)베스트텍시스템 파트너
- 빅데이터 구축의 어려움
- (주)베스트텍시스템 제안
- (주)베스트텍시스템과 Microsoft Azure
- (주)베스트텍시스템 빅데이터 요소 기술
- (주)베스트텍시스템 시각화 플랫폼
- 빅데이터 플랫폼 구축 Case Sturdy - 편의점
- 빅데이터 플랫폼 구축 Case Sturdy - 보안, 백신 회사
- 빅데이터 플랫폼 구축 Case Sturdy - 게임사

# 회사 소개



<b>Company</b>	(주)베스트텍시스템	<b>Represented by</b>	서우승
<b>Established</b>	2006년 8월 ~ 현재		
<b>Business Area</b>	Storage / Big data / Server / Solution		
<b>Items</b>	<ul style="list-style-type: none"> <li>• 빅데이터 플랫폼</li> <li>• 스토리지 (NAS, DAS, All flash array )</li> <li>• 서버</li> <li>• Solution (네트워크, 백업 솔루션, 복구 솔루션, 웹 방화벽)</li> </ul>		
<b>Address</b>	본사 : 서울시 성동구 성수일로8길 53 7층 부산지사 : 부산 남구 신선로 428	<b>TEL</b>	82-2-497-0090~2
		<b>FAX</b>	82-2-497-0093
		<b>Website</b>	<a href="http://bestts.co.kr">http://bestts.co.kr</a>

**빅데이터**

빅데이터 플랫폼

- Microsoft Azure
- Biz Flash 빅데이터 플랫폼

**스토리지**

**Storage**

- Thecus (NAS) **Distributor**
- EMC
- FlashMAXII (DAS)

**All flash array**

- Kaminario

**서버**

**Server**

- HP
- Lenovo
- Fujitsu

**솔루션**

Fortinet firewall  
 CA 백업 솔루션  
 Kaspersky 보안 솔루션

## 회사 연혁

# (주)베스트텍시스템

- 2006년 회사 설립
- 2006년 IBM 벨류 파트너 인증
- 2007년 벤처기업인증(기술신용보증기금)
- **2007년 네트워크 스토리지(NAS) "Thecus" 한국 공식 수입원**
- 2008년 "BestStor 상표출원 (출원번호:40-2006-0048201)"
- 2008년 기업 전담부서 등록
- 2011년 Sans Digital사 Elite-NAS 한국 공식 총판 인증
- 2011년 HP 공식 파트너 인증
- 2012년 베스트텍시스템 부산지사 사세확장
- 2013년 Fortinet 협력업체 지정
- 2014년 EMC, Fujitsu, Nimble Storage 파트너쉽 체결
- 2015년 Lenovo 비즈니스 파트너 체결
- 2015년 Kaminario 공급파트너 체결
- **2016년 MS Azure 공급파트너 체결 (Azure 유일 기술 파트너)**



## (주)베스트텍시스템 파트너



### Microsoft Azure 유일 기술 파트너



(주)베스트텍시스템은 마이크로소프트(MS)와 클라우드 컴퓨팅 플랫폼 애저(Azure)의 공급 계약을 체결하였으며, **국내 유일의 MS 애저 오픈 소스 플랫폼 기술 파트너**로 빅데이터 플랫폼 구축 기업의 입지를 다져 나가고 있습니다.

### 다양한 협력사

(주)베스트텍시스템은 2006년 설립 이후, 비즈니스의 경쟁력을 높이기 위하여 다양한 사업분야의 파트너와 함께 최적의 서비스를 제공할 수 있도록 국, 내외 업체와 긴밀한 협력관계를 이어가고 있습니다.

- Big Data
- Storage
- Server
- Solution

# 빅데이터 플랫폼 구축의 어려움



## RDBMS

- Oracle,
- MS-SQL,
- MySQL,
- Postgresql



## NoSQL

- MongoDB,
- Cassandra



## Files

- Text,
- JSON,
- XML



## Log

- WAS,
- Application,
- System



### 사용자들이 겪는 일반적인 문제점

- 다양한 종류의 데이터가 존재.
- 일반적으로 한가지 종류의 데이터만 분석하지는 않음.
- 종류별로 분석 방법도 다름.



### 너무나도 다양한 분석 방법

- 데이터 종류별로 별도의 분석 방법을 사용.
- 다양한 분석 도구를 사용하기 쉽지 않음.

# 빅데이터 플랫폼 구축의 어려움



Standard

표준화된 방법이 없음.  
상황에 따라 다양한 한정된  
방법을 사용해야 함.



Sqoop

테이블 구조에 따라 성능이 천차만별.  
테이블 별로 커스텀 SQL 작성이 필요.



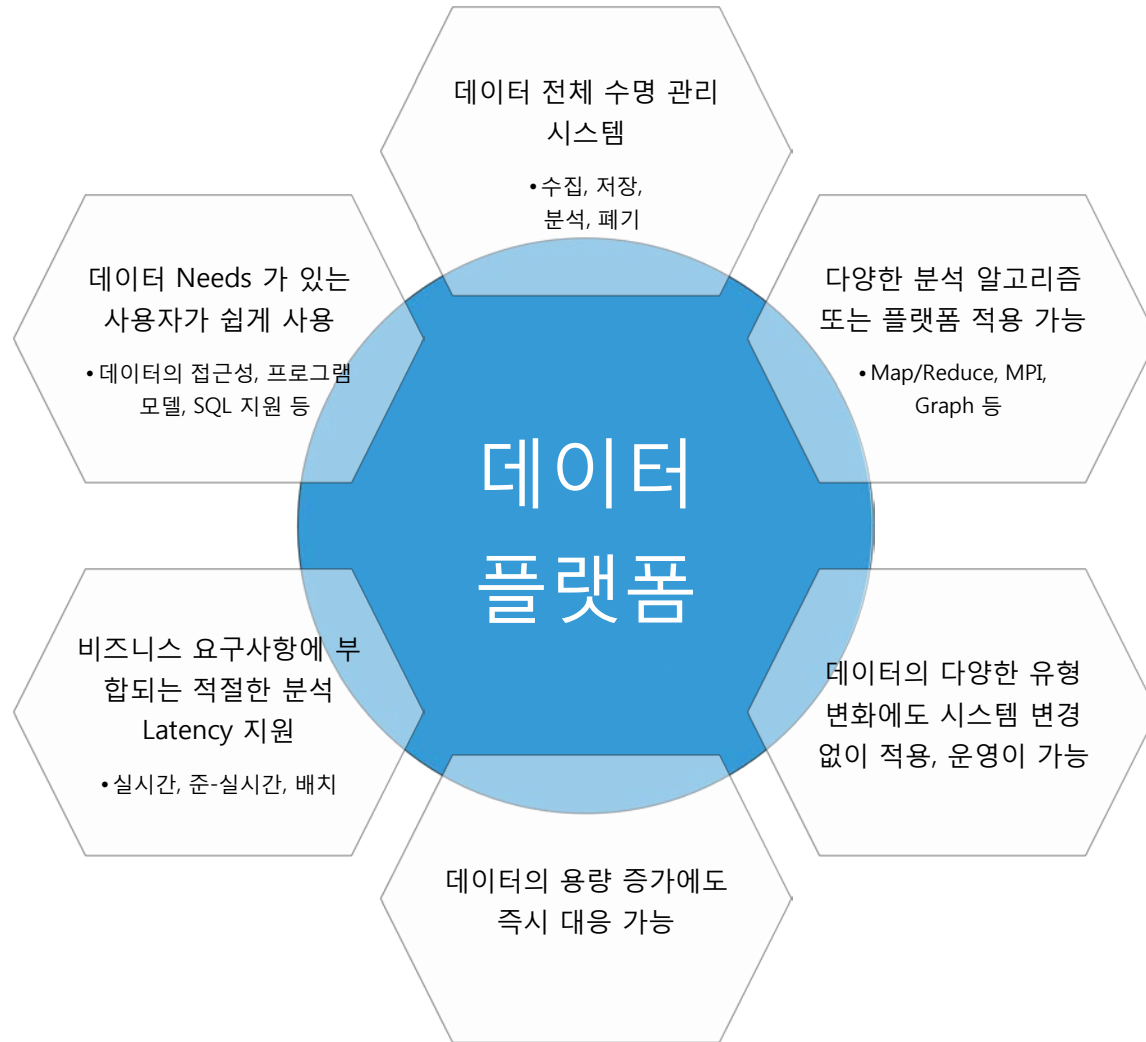
Python

필요에 따라 전처리를 위한 스크립트 필요.  
(특수문자 처리, 데이터 포맷 변경등)

## 데이터 적재의 문제

- 일반적인 ETL 도구를 사용할 수 없는 경우도 많음.
- 상황에 따라 주어진 도구(Python, Bash)를 활용하여 작업을 수행 해야함.

# (Big)Data플랫폼이란?





## (주)베스트텍시스템 BigData 구축 전략



- **ROI [Return on Investment]**

- 전통적인 데이터 분석 솔루션을 이용하여 원하는 데이터를 분석할 수 있어야 함
- 투자금액대비 이익 및 성과 필요



@ **해결 방안**

오픈 소스 중심으로 발생 비용을 절감

- **플랫폼**

- 데이터 통합 관리 및 분석 체계 구축 필요
- 분석가는 쉽게 데이터에 접근할 수 있어야 함



@ **해결 방안**

사용자 선택에 따라  
Cloud 서비스 이용 가능 ( MS Azure )

- **내재화**

- 데이터 분석은 고객이 직접 할 수 있어야 한다
- 데이터 플랫폼의 운영 및 진화도 고객이 직접 수행할 수 있어야 함

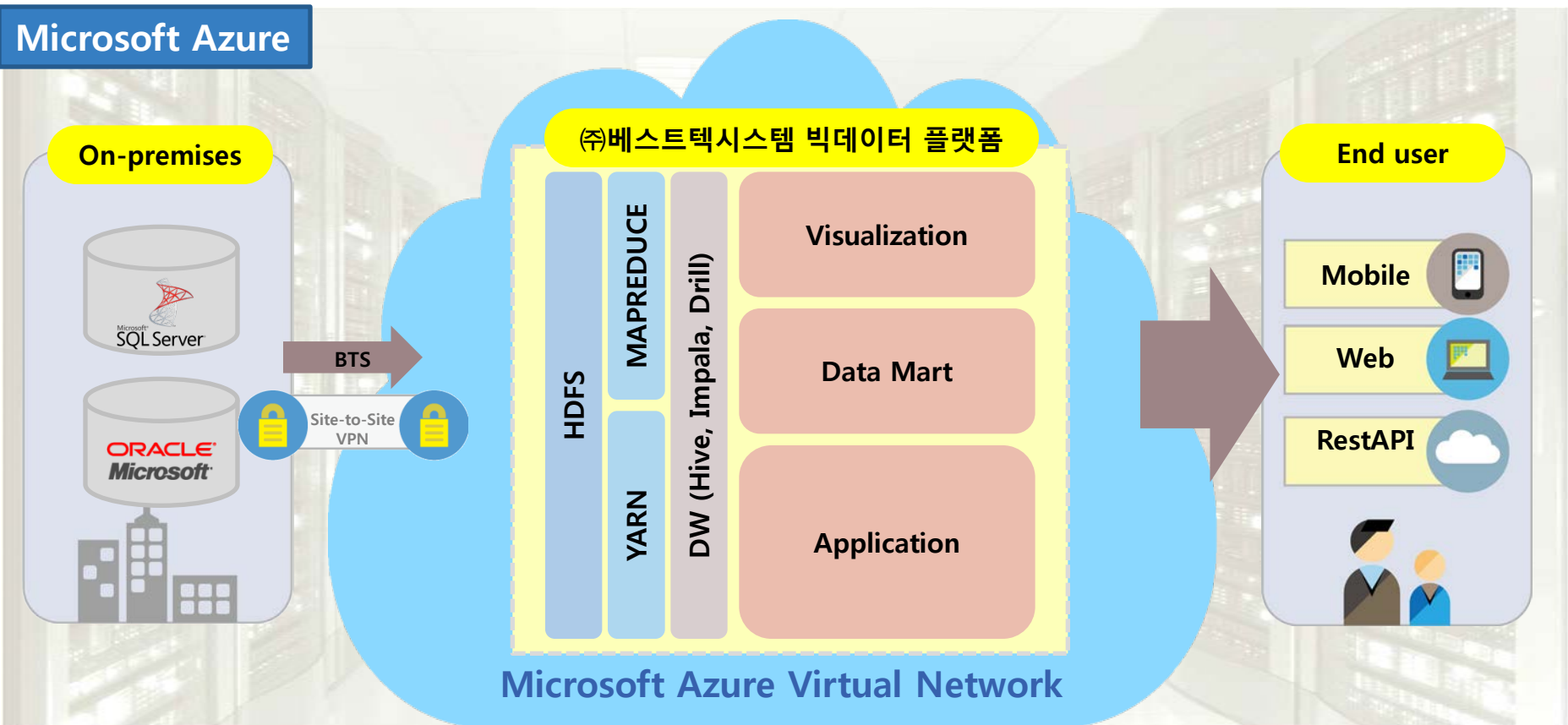


@ **해결 방안**

BizFlash를 이용한 시각화 및  
(주)베스트텍시스템 기술 지원


## (주)베스트텍시스템 빅데이터 플랫폼 (Azure를 이용한 구성)

- MS Azure 클라우드를 이용해 오픈소스 플랫폼을 구축할 수 있는 애저 플랫폼을 제공
- 일련의 쿼리를 통해 데이터를 분석하고 웹 브라우저 기반으로 데이터를 시각화하여 사용자에게 가장 효과적인 결과를 도출, 공유
- 별도의 소프트웨어 설치 없이 클라우드 기반에서 MS의 데이터 분석도구를 함께 이용



## 플랫폼 주요 컴포넌트



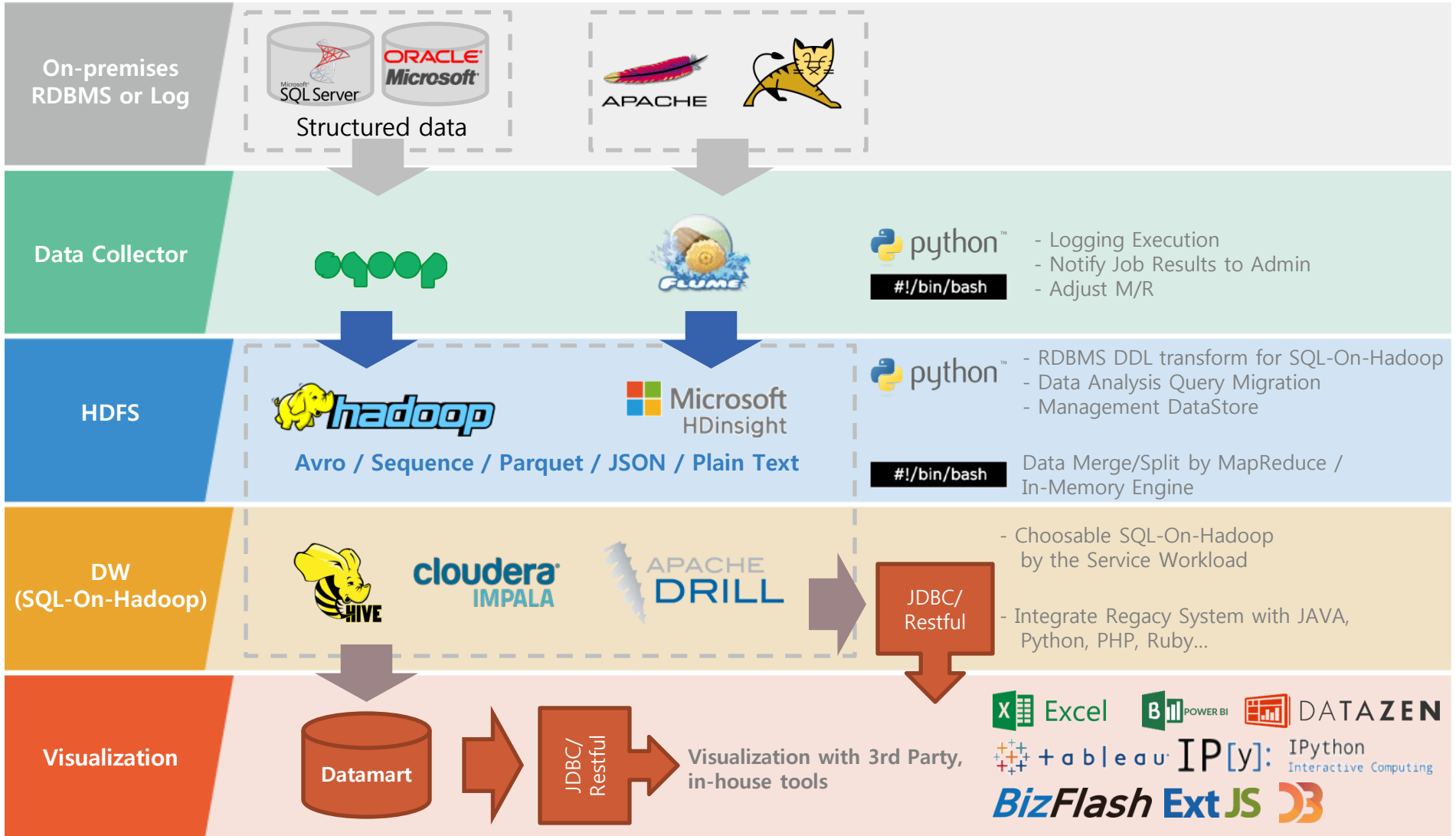
구성요소	역할	솔루션
데이터 수집	데이터 소스(설비, Legacy)로부터 안전한 데이터 저장소까지 데이터를 수집 저장	 sqoop, flume 등
데이터 저장	다양한 형태의 데이터를 안정적으로 저장 (원본데이터, 테이블 데이터, 메타 데이터, 실행 결과 데이터, 임시 데이터 등)	Hadoop File System MS Azure Blob
분석 모델 개발 엔진	수집된 데이터를 이용하여 다양한 분석 모델을 쉽고 빠르게 생성할 수 있는 기능 제공	Hadoop MapRedude SQL-On Hadoop R, Jupyter, Spark
데이터 인터페이스	기존 BI 도구와 Hadoop에 저장된 데이터의 연동을 지원하는 인터페이스 데이터 연동은 SQL 기반의 JDBC/ODBC를 이용한 연동	JDBC/ODBC
분석 모델 엔진	개발된 모델을 실제 운영환경에서 실행하는 엔진으로 기본 컴포넌트 구성은 분석 모델 개발 엔진과 동일하지만 실행 성능 QoS 보장을 위해 별도 클러스터로 구성	Hadoop MapRedude Spark
검색 인덱스 및 엔진	사용자가 요구한 다양한 조건에 맞춰 빠르게 응답하는 검색 기능 수행	ElasticSearch Hadoop Plugin
플랫폼 운영/관리 툴	시스템 전체 하드웨어, 소프트웨어에 대한 모니터링, 운영 및 관리 도구	Ambari

## 시스템 장점

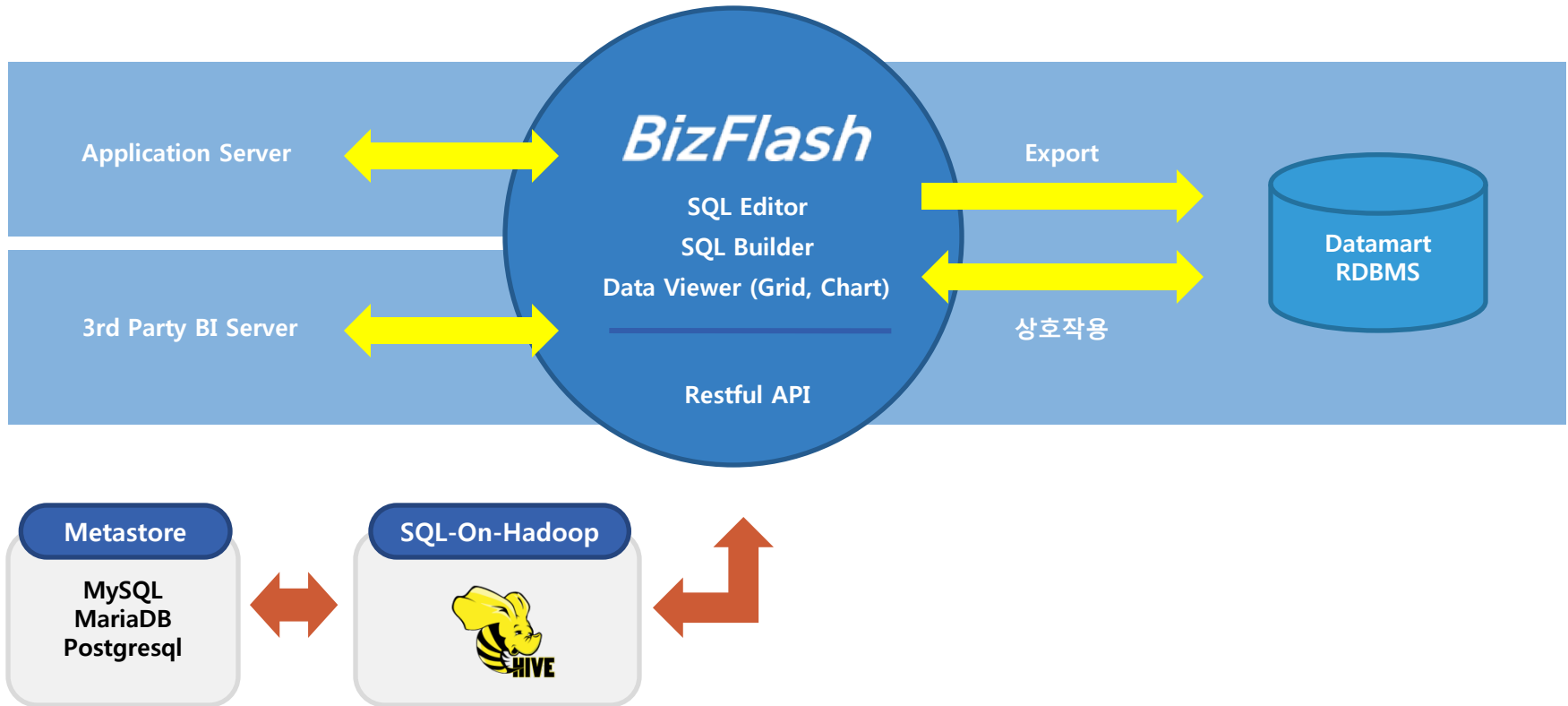


- **오픈 소스 중심의 아키텍처**
  - 대부분 컴포넌트는 Hadoop과 Hadoop eco 시스템 기반 오픈 소스로 구성
  - 오픈 소스이지만 많은 기업에서 사용하는 빅데이터 산업 표준 솔루션 중심 도입
- **쉬운 분석 모델 개발 지원**
  - 사용자가 익숙한 SQL을 이용한 데이터 가공 및 단순한 분석 수행
- **초기 도입 비용 절감**
  - 솔루션 도입 비용 최소화로 분석 모델 개발에 집중 가능
  - 고객 선택에 따라 Cloud 구축 가능
- **쉬운 확장**
  - 데이터 증가 및 분석 대상 증가 시 프로그램 수정 없이 하드웨어 증설만으로 성능 확보 가능(Cloud)
  - 물리적 확장 시 하드웨어 비용만으로 확장 가능, 별도 소프트웨어 비용 미발생(On-Premise)
- **다양한 요구사항 대응**
  - 어떤 형태의 데이터도 저장 가능하며, 필요 시 원본 데이터로부터 필요한 데이터 모델로 빠르게 변환 가능
  - 요구사항 변경 시 적절한 오픈 소스 도입 및 적용 가능

# (주)베스트텍시스템 빅데이터 요소 기술

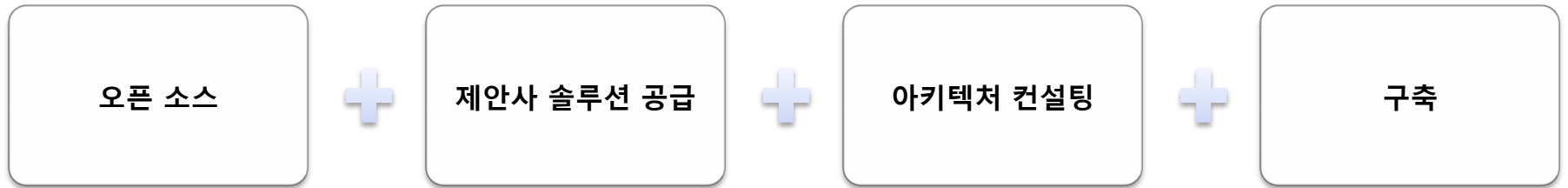


# (주)베스트텍시스템 빅데이터 시각화



# 납품 기간

## • 빅데이터 플랫폼 Delivery 전략

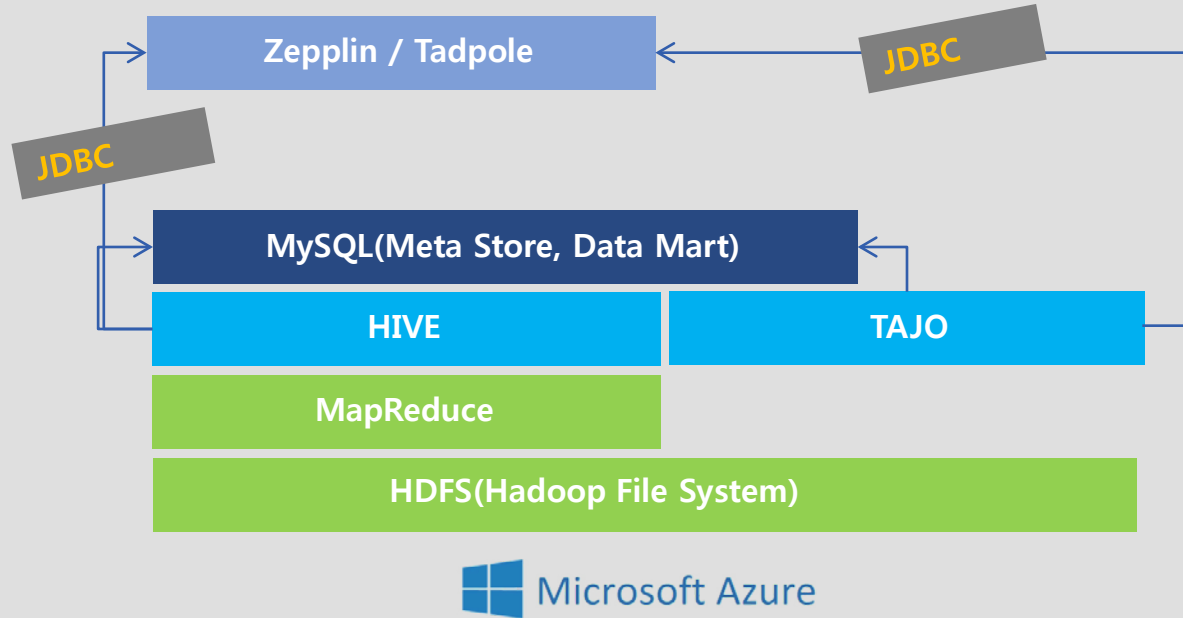


## • 주요 Delivery 내용

<p><b>고객의 요구사항에 적합한 시스템 아키텍처 수립</b></p>	<p><b>플랫폼 구축</b></p>	<p><b>분석 모듈 개발</b></p>
<ul style="list-style-type: none"> <li>• 사용자 요구사항에 부합되는 아키텍처 및 솔루션 선정</li> <li>• 하드웨어 스펙 및 필요 대수 산정</li> <li>• 솔루션은 대부분 오픈 소스 중심으로 구성</li> </ul>	<ul style="list-style-type: none"> <li>• 선정된 솔루션 설치 및 실행 환경 구성</li> <li>• 일부 컴포넌트 개발</li> </ul>	<ul style="list-style-type: none"> <li>• 고객과 협의된 업무 및 데이터를 고객 요청에 따라 물리적 또는 클라우드 플랫폼에 ETL</li> <li>• 분석 프로그램 연동 및 설정</li> </ul>

# 빅데이터 플랫폼 구축 Case Sturdy

## M사 : 편의점



**@ 고객 환경 및 니즈**

1. 일 250만 건의 판매 로그 저장
2. 분석대상은 6개월 ~ 1년
3. 데이터 저장은 Oracle
4. 전체 데이터는 4억 5천만건(6개월)

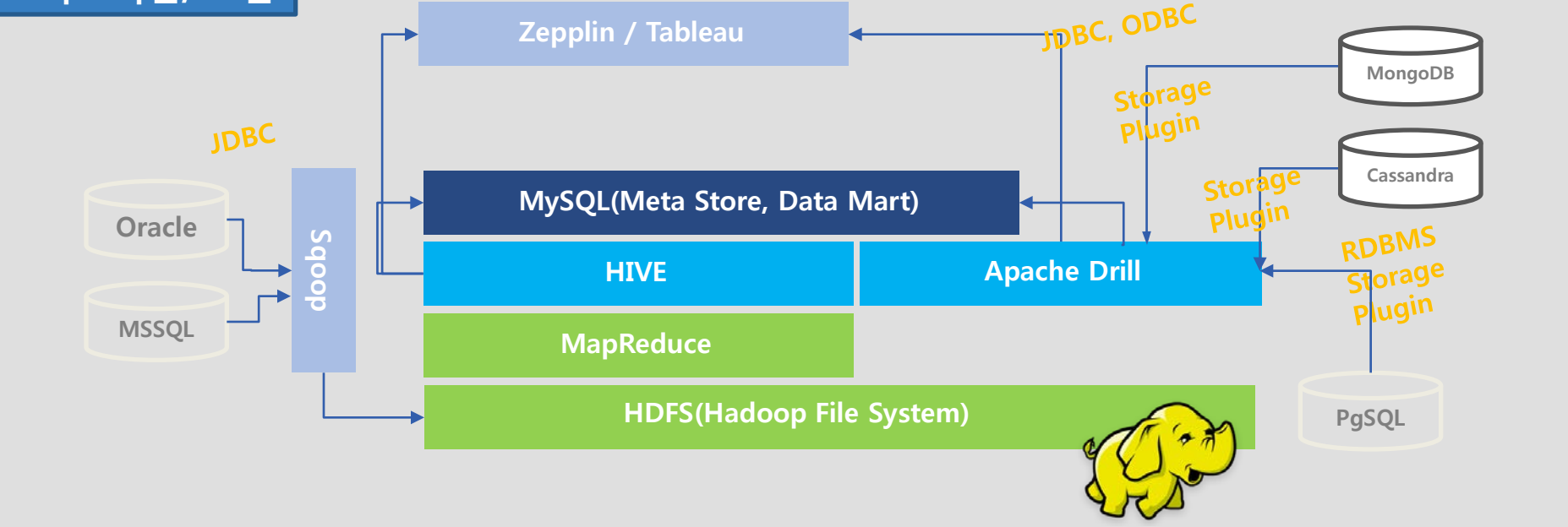
**@ 해결 방안**

- 데이터 ETL(일 단위, CHARACTER SET 변환 등)
- 오라클 분석 쿼리 Migration ( SQL-On-Hadoop)
- DBA 와 마케터등 실무자들을 위한 SQL GUI Client 제공
- SQL 기반의 분석 및 시각화 도구 제공



# 빅데이터 플랫폼 구축 Case Study

## A사 : 백신, 보안



### @ 고객 환경 및 니즈

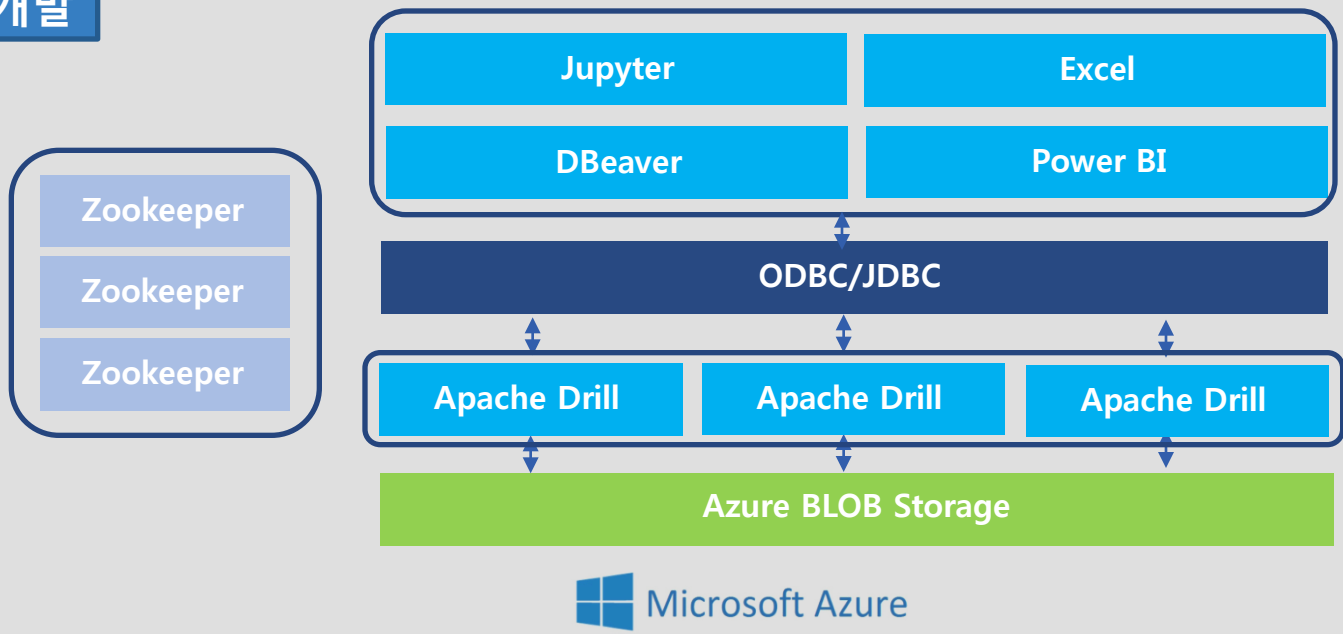
- 다양한 데이터 소스 산재
- ANSI SQL을 이용한 분석 도구 필요
- 3rd Party 분석 도구 연계 필요
- 목표시간 내 데이터 ETL 완료

### @ 해결 방안

- Apache Drill 을 통해 다양한 데이터 소스 통합.
- Client 분석 도구 사용을 위한 인터페이스 제공(ODBC/JDBC).
- SQL 을 통해 RDBMS, NoSQL, Files(CSV, JSON등) 을 모두 분석.

# 빅데이터 플랫폼 구축 Case Sturdy

W사 : 게임,개발



**@ 고객 환경 및 니즈**

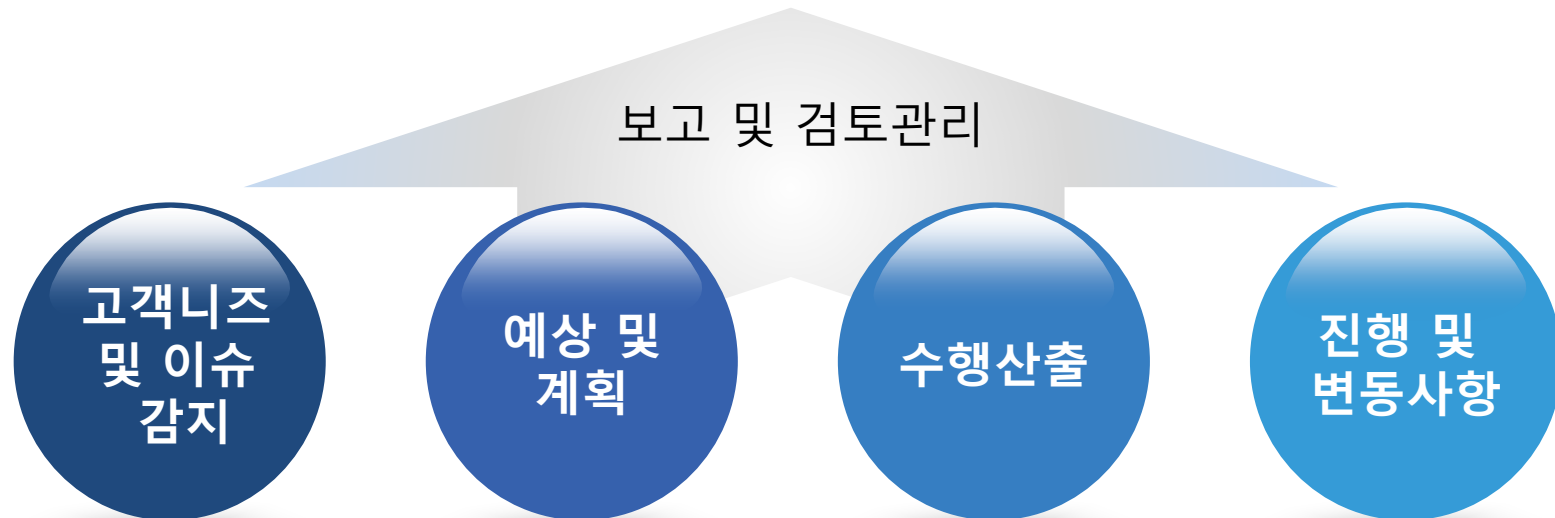
- 대량의 게임 로그 분석 플랫폼 필요
- MapReduce의 성능개선을 위한 대안 필요
- Jupyter 와의 연동 필요
- 마케터 또는 DBA를 위한 분석 도구 제공
- 대화형 데이터 시각화 BI 도구 제공

**@ 해결 방안**

- Parquet 변환을 통해 수배 ~ 수십배 의 성능 향상(Plain Text Gzipped 압축 대비)
- Apache Drill User Defined Function 개발(날짜 관련)
- Client 분석 도구 사용을 위한 인터페이스 제공(ODBC)
- Jupyter 에서 MR 코드의 일부를 SQL 로 대체

## (주)베스트텍시스템 제안

비즈니스 산업에 따라 IT 산업은 매해 엄청난 대용량 데이터베이스를 요구하고 있습니다. (주)베스트텍시스템은 고가용성 시스템 제공을 통해 고객이 필요로 하는 요구사항에 맞추어 가장 적합한 솔루션을 제안합니다. 또한, 축적된 빅 데이터를 지속해서 유지 관리가 가능하도록 컨설팅을 통해 서비스 중단 및 피해 최소화를 위하여 최적의 방안을 모색, 제안합니다.





# 감사합니다.

Big data 플랫폼, Storage, Server, Solution IT 컨설팅 전문기업

(주)베스트텍시스템

