

빅데이터

1-1 파일 시스템 방식과 관계형 데이터 베이스 시스템

File System 방식

빅데이터를 확장 가능한 분산 파일의 형태로 저장하는 방식이다.

장점 : 저 사양 서버를 이용하여 대용량, 분산, 데이터 집중형의 애플리케이션을 지원한다.

아파치(Apache)

HDFS(Hadoop Distributed File System)

구글의

GFS(Google File System)

1-1 파일 시스템 방식과 관계형 데이터 베이스 시스템

관계형 데이터 베이스 시스템

키(key)와 값(value)들의 간단한 관계를 테이블화 시킨 간단한 원칙의 데이터 베이스

관계는 테이블 간에 둘 다 존재.

오라클 DBMS(Oracle)

MS : MS SQL Server

MySQL

Altibase DB

1-1 파일 시스템 방식과 관계형 데이터 베이스 시스템

NoSQL

- A. 스키마가 없다.
- B. 분산처리를 제공한다.

컬럼 베이스 : Hbase

도큐먼트 : MongoDB

키 값 : Redis

그래프 : Neo4J

1-2 key-value vs column-oriented 데이터 베이스

Key-Value 데이터 베이스

- A. 데이터와 키와 그에 해당하는 값의 쌍으로 저장한다.
- B. RDBMS(관계형 데이터 베이스)보다 확장성이 뛰어나고 질의 응답 시간도 빠르다.

1-2 key-value vs column-oriented 데이터 베이스

Column-oriented 데이터 베이스

- A. 로우(row)가 아닌 칼럼(column)기반으로 저장하고 처리한다.
- B. 확장성을 보장하기 위해 여러 개의 노드로 분할되어 저장되고 관리된다.
- C. **Cassandra, Hbase**, HyperTable등

1-2 key-value vs column-oriented 데이터 베이스

document 데이터 베이스

- A. 문서형식의 정보를 저장, 검색, 관리하기 위한 데이터 베이스이다.
- B. **MongoDB**, SimpleDB, couchDB

1-3 인메모리(in Memory) 방식

In-Memory Database (IMDB or MMDB(Main Memory DBMS))

- A. 디스크가 아닌 주 메모리에 모든 데이터터를 보유하고 있는 데이터 베이스
- B. [장점] 디스크 검색보다 **자료 접근이 훨씬 빠르다.**
- C. [단점] DB 서버 전원이 갑자기 꺼지면 **자료들이 초고속 즉시 삭제됨.**

1-3 인메모리(in Memory) 방식

In-Memory Database (IMDB or MMDB(Main Memory DBMS))

- A. Altibase(알티베이스) – 한국
- B. SunDB – 선재소프트(한국)
- C. Redis – 오픈소스(2009)
- D. SQLite – Open Source(In memory 지원)

1-4 데이터 베이스 트랜잭션(Transaction)

Database Transaction(데이터베이스 트랜잭션)

- A. 데이터베이스 관리 시스템 또는 유사 시스템에서 상호작용의 단위이다.
- B. (명사) 처리, 처리과정이라 한다.
- C. 다수의 쿼리문으로 이루어진 논리적 작업 단위를 말한다.
- D. 논리적 작업 단위가 모두 성공적으로 실행되면 COMMIT 을 호출하여 모든 쿼리문을 한번에 데이터 베이스에 적용

2-1 RDB and NoSQL 저장 시스템

RDB(관계형 DB), NoSQL DB

A. RDB 저장 시스템

Oracle, MSSQL, MySQL, Sybase 등

⇒ 정형 데이터, 반정형 데이터

B. NoSQL 저장 시스템 (Not Only SQL)

MongoDB, Cassandra, Hbase, Redis 등

=> 반정형 데이터, 비정형 데이터

2-2 NoSQL DB 종류

가. key-value

나. column-oriented

다. document-oriented

라. graph-oriented

2-3 레코드 식별(기본키), NoSQL DB(id)

가. RDBMS 의 테이블의 레코드 식별

A. Primary key(기본키)

나. NoSQL DBMS의 데이터 식별

A. Id 이용

2-4 조인(join) 연산

가. RDBMS 의 테이블 조인(join)

A. Join 연산

나. NoSQL DBMS의 조인

A. 다른 문서와 연결된 개체(id_)

3-1 Fault tolerant system

가. 결함 감내 시스템(Fault tolerant system)

- A. 시스템을 구성하는 부품의 일부에서 결함(falut) 또는 고장(failure)이 발생하여도 정상적 혹은 부분적으로 기능을 수행하는 시스템.

3-2 scale-out(수평적 확장성)

가. Scale-out(스케일 아웃)

- A. 서버를 운영시에 갑작스런 이용자의 증가, 사업 확장 등의 이유로 더 많은 서버 용량과 성능이 필요하다. 이중에 시스템을 확장시키는 방법으로 두가지가 있다.

Scale-out, Scale-up

3-2 scale-out(수평적 확장성)

가. Scale-out(스케일 아웃)

- A. 서버를 여러 대 추가하여 시스템을 확장하는 방법

나. Scale-up(스케일 업)

- A. 1의 처리 능력을 가진 서버 한대를 '5'의 처리 능력이 있는 서버로 업그레이드 시킨다.
- B. [단점] 서버 한 대에 모든 부하가 집중된다

3-3 DBMS 인기 순위

가. DBMS 인기 순위(2016년)

<표 1-8> DBMS 인기 순위 (출처: Solid IT. (2016. 08). DB-Engines Ranking. Solid IT. <http://db-engines.com/en/ranking>)

순 위			DBMS	데이터베이스 모델
2016년 8월	2016년 7월	2015년 8월		
1.	1.	1.	Oracle	관계형 데이터베이스
2.	2.	2.	MySQL	관계형 데이터베이스
3.	3.	3.	Microsoft SQL Server	관계형 데이터베이스
4.	4.	4.	MongoDB	Document 데이터베이스
5.	5.	5.	PostgreSQL	관계형 데이터베이스
6.	6.	6.	DB2	관계형 데이터베이스
7.	7.	8.	Cassandra	Column-oriented 데이터베이스
8.	8.	7.	Microsoft Access	관계형 데이터베이스
9.	9.	9.	SQLite	관계형 데이터베이스
10.	10.	10.	Redis	Key-value 데이터베이스

3-3 DBMS 인기 순위

가. DBMS 인기 순위

342 systems in ranking, May 2018

Rank	May 2018	Apr 2018	May 2017	DBMS	Database Model	Score		
						May 2018	Apr 2018	May 2017
1.	1.	1.		Oracle +	Relational DBMS	1290.42	+0.63	-63.90
2.	2.	2.		MySQL +	Relational DBMS	1223.34	-3.06	-116.69
3.	3.	3.		Microsoft SQL Server +	Relational DBMS	1085.84	-9.67	-127.96
4.	4.	4.		PostgreSQL +	Relational DBMS	400.90	+5.43	+34.99
5.	5.	5.		MongoDB +	Document store	342.11	+0.70	+10.53
6.	6.	6.		DB2 +	Relational DBMS	185.61	-3.34	-3.23
7.	↑ 9.	↑ 9.		Redis +	Key-value store	135.35	+5.24	+17.90
8.	↓ 7.	↓ 7.		Microsoft Access	Relational DBMS	133.11	+0.89	+3.24
9.	↓ 8.	↑ 11.		Elasticsearch +	Search engine	130.44	-0.92	+21.62
10.	10.	↓ 8.		Cassandra +	Wide column store	117.83	-1.26	-5.28
11.	11.	↓ 10.		SQLite +	Relational DBMS	115.45	-0.53	-0.61
12.	12.	12.		Teradata	Relational DBMS	74.41	+0.74	-1.91
13.	13.	↑ 16.		Splunk	Search engine	65.09	+0.04	+8.40
14.	14.	↑ 18.		MariaDB +	Relational DBMS	64.99	+0.44	+14.01
15.	15.	↓ 14.		Solr	Search engine	61.51	-1.70	-2.26
16.	16.	↓ 13.		SAP Adaptive Server +	Relational DBMS	61.51	-0.12	-6.24
17.	17.	↓ 15.		HBase +	Wide column store	59.95	+0.26	+0.44
18.	18.	↑ 20.		Hive +	Relational DBMS	56.97	-0.43	+13.49
19.	19.	↓ 17.		FileMaker	Relational DBMS	54.67	-0.33	-1.81

<https://db-engines.com/en/ranking>

3-3 DBMS 인기 순위

가. DBMS 인기 순위

342 systems in ranking, May 2018

Rank	May 2018	Apr 2018	May 2017	DBMS	Database Model	Score		
						May 2018	Apr 2018	May 2017
1.	1.	1.		Oracle +	Relational DBMS	1290.42	+0.63	-63.90
2.	2.	2.		MySQL +	Relational DBMS	1223.34	-3.06	-116.69
3.	3.	3.		Microsoft SQL Server +	Relational DBMS	1085.84	-9.67	-127.96
4.	4.	4.		PostgreSQL +	Relational DBMS	400.90	+5.43	+34.99
5.	5.	5.		MongoDB +	Document store	342.11	+0.70	+10.53
6.	6.	6.		DB2 +	Relational DBMS	185.61	-3.34	-3.23
7.	↑ 9.	↑ 9.		Redis +	Key-value store	135.35	+5.24	+17.90
8.	↓ 7.	↓ 7.		Microsoft Access	Relational DBMS	133.11	+0.89	+3.24
9.	↓ 8.	↑ 11.		Elasticsearch +	Search engine	130.44	-0.92	+21.62
10.	10.	↓ 8.		Cassandra +	Wide column store	117.83	-1.26	-5.28
11.	11.	↓ 10.		SQLite +	Relational DBMS	115.45	-0.53	-0.61
12.	12.	12.		Teradata	Relational DBMS	74.41	+0.74	-1.91
13.	13.	↑ 16.		Splunk	Search engine	65.09	+0.04	+8.40
14.	14.	↑ 18.		MariaDB +	Relational DBMS	64.99	+0.44	+14.01
15.	15.	↓ 14.		Solr	Search engine	61.51	-1.70	-2.26
16.	16.	↓ 13.		SAP Adaptive Server +	Relational DBMS	61.51	-0.12	-6.24
17.	17.	↓ 15.		HBase +	Wide column store	59.95	+0.26	+0.44
18.	18.	↑ 20.		Hive +	Relational DBMS	56.97	-0.43	+13.49
19.	19.	↓ 17.		FileMaker	Relational DBMS	54.67	-0.33	-1.81

<https://db-engines.com/en/ranking>