**산학협력프로젝트 – 알티베이스 과제명 사전조사**

**WEB WAS DB 스케일아웃 솔루션과 알티베이스 연계 기반 성능모니터링 솔루션 개발**

키워드 : WEB, WAS, DB, 스케일아웃 솔루션, 알티베이스 연계~~ 솔루션

**1.WEB**

Web의 본 의미는 거미줄을 뜻하는 영단어 web으로부터 파생된 것으로, 하나의 사이트나 다른 사이트와의 관계가 거미줄처럼 복잡하게 얽혀 있는 것을 뜻한다. WWW(World Wide Web)라고도 불리며, 제공하는 정보 검색 서비스로 그림, 동화상, 소리 등을 지원한다. 하이퍼텍스트 개념을 도입, 쉽게 원하는 정보를 찾을 수 있다. 아마 광범위한 WEB이라는 키워드를 사용한 것으로 보아 WAS와 연동하여 ‘WEB 서버’라는 의미로 사용한 것 같다. 링크1 참조.

**2.WAS**

웹 애플리케이션 서버 즉 웹 응용 서버(Web Application Server)의 약자로, 인터넷 상에서 HTTP를 통해 사용자 컴퓨터나 장치에 애플리케이션을 수행해주는 미들웨어(소프트웨어 엔진)이다. Server 에서 애플리케이션을 동작할 수 있도록 지원한다. 동적 서버 컨텐츠를 수행하여 웹 서버와 구별되며, 주로 데이터베이스 서버와 같이 수행된다. 주로 자바 기반으로 JAVA EE 표준을 수용하나, 자바 기반이지만 JAVA EE를 따르지 않거나 .NET이나 Citrix 기반인 계열도 존재한다. WAS는 안정된 시스템을 구성하며 자동적인 어플리케이션 복구를 제공한다. 또한 업무 로직이 중간 어플리케이션 서버에 존재하여 쉽고 빠르게 구축 가능하다.

**3.DB**

데이터베이스, 여러 사람에 의해 공유되어 사용될 목적으로 통합하여 관리되는 데이터의 집합을 뜻한다. 자료의 중복을 없애고 구조화하여 저장함으로써 검색과 갱신의 효율을 높인다. 데이터베이스 자체로는 아무것도 못하기 때문에, 그걸 관리하는 시스템과 통합되어 사용하며, 정확한 명칭으로는 데이터베이스 관리 시스템(DBMS)가 된다. 데이터베이스만 제공되는 건 CSV와 같이 아주 단순한 데이터에 국한되는데, 이걸 주로 사용하지는 않고 RAW데이터로 간주해 다른 DBMS 시스템에 적재하는게 일반적이다.

관계형 데이터 베이스(RDBMS)가 가장 널리 쓰인다. 이를 이용하기 위한 표준 언어가 만들어져 있는데, 그것이 SQL(Structured Query Language)이다. 비관계형 DB는 NoSQL이라 불린다. DB의 대부분은 SQL을 사용한다.

설계 단계 : 요구 조건 분석>개념적 설계>논리적 설계>물리적 설계>구현>운영 및 개선

데이터베이스의 종류에는 5가지가 있다.

**3.1 관계형(Relational)**

데이터를 Column과 Row라는 일종의 표 형태로 저장한다. 데이터의 종속성은 Relation이다. 테이블의 모든 로우는 같은 길이의 칼럼을 가진다. 이 테이블의 칼럼의 구조와 데이터의 관계는 테이블 Schema로 정의된다. 가장 오랫동안 사용된 것인 만큼 데이터의 분류, 정렬, 탐색 속도가 빠르다.

다만, 데이터가 2차원적으로만 표현되기 때문에, 트리 구조로 조직화되는 객체들과 궁합이 잘 맞지 않는다. ORM기법으로 커버하고 있으나 객체 친화적인 DB를 사용하는 게 좋다.

**3.2 키-값형(KV Store)**

모든 데이터를 Key와 Value의 쌍으로 매핑한다. Hash와 비슷한 것 같다. 특정 값 하나를 찾아내는 데에는 가장 뛰어난 성능을 보인다. 하지만 데이터를 그룹화하고 정렬하는 기능은 거의 없다. 대신 RDBMS에 비해 가볍고, 빠르고, 다루기 쉽다. 2차원 이상 데이터 형식은 KV Store말고 다른 형식을 추천.

**3.3 객체형(Object)**

프로그래밍 언어에서 객체지향의 개념이 포함되듯, DB에서도 객체지향을 구현하였는데 이것이 바로 객체형 DB이다. ODBMS라고 불린다. 다만 DB분야에서 주류는 아닌데, 그 이유는 쿼리 사용이 복잡해서이다. 따라서 아직 RDBMS 방식이 널리 쓰인다.

**3.4 문서형(Document)**

객체형과 비슷하다. Schema가 필수가 아닌 것이 특징. Index Field를 제외한 필드를 넣고 뺄 수 있다. 데이터의 Schema가 자주 바뀌고 데이터가 다계층의 객체 형태이면서 업무가 갱신, 삽입(문서단위), 조회에 집중되어있고 통계 연산에 쓰일 일이 많지 않은 곳에서 사용하기 좋다. 통계 연산이 필요할 경우엔 필요한 데이터를 추출해 RDBMS로 옮기고 SQL로 처리한다. 혹은 DBMS에서 자체 제공하는 map/reduce 연산을 사용한다.

**3.5 컬럼형(Column)**

디스크에 조차 다 담지 못할 정도의 거대한 데이터를 다루는 게 특징이다. 다만 데이터 검색용으로 사용하기보다는 데이터의 통계 처리용으로 주로 사용된다. 사용하기가 더럽게 어렵다. 최소 12대 이상의 컴퓨터로 클러스터 머신을 구성해야 한다. DB 뿐 아니라 네트워크와 하드웨어에 대한 심도깊은 이해가 필요하여 지금은 인기가 많이 식었다.

**4.스케일아웃 솔루션**

서버 확장에는 스케일 아웃과 스케일 업 두가지 방식이 있다.

스케일 아웃이란 서버를 여러 대 추가하여 시스템을 확장하는 방식이다. 서버가 여러 대가 되므로 각 서버에 걸리는 부하를 균등하게 해주는 ‘로드밸런싱’이 필수적으로 동반되어야 한다. 스케일 아웃의 경우, 서버 한 대가 장애로 다운되어도 다른 서버로 서비스 제공이 가능하다. 서버의 가상화 기능을 사용하고 하나의 케이스 내에서 가상으로 복수 서버를 구축하여 스케일 아웃과 같은 효과를 볼 수 있는데, 이를 가상 스케일 아웃이라고 한다. 수평 확장, 지속성 확장이 가능하다. 비용 부담이 적다. 다만 서버 대수가 늘어나면 관리 편의성이 떨어지고, 서버의 상면 비용을 포함한 운영 비용이 증가한다.

읽기/쓰기가 여러 대의 서버에 분산되어 처리됨으로 장애 시 전면 장애의 가능성이 적다.

Sharding, Query-off Loading, Queue, In Memory Cache, NoSQL, Object Storage, Distributed Storage 등의 기술이 사용된다.

**5.알티베이스 관련 성능 모니터링**

우선, 알티베이스는 알티베이스사에서 제작한 한국산 데이터베이스 관리 시스템, 즉 DBMS이다. 빠른 응답 처리가 가능한 인메모리(In-Memory) DBMS와 대용량 데이터 처리가 가능한 온디스크(On-Disk) DBMS를 동시에 사용할 수 있다는 장점을 지니고 있다.

성능 모니터링 솔루션은 말 그대로 해당 시스템의 성능을 모니터링할 수 있는 솔루션, 툴을 뜻한다. 링크2 참조. 컨테이너 관리 및 시스템 모니터링, 유형별 모니터링 환경을 제공한다.

**6.LINK**

**1. WEB Server와 Web Application Server의 연동**

<https://dodo4513.github.io/2017/06/18/server_architecture/>

**2. Tmax Soft – 성능 모니터링 솔루션**

<https://kr.tmaxsoft.com/pdsDownload.do?board_file_seq=425>