

MongoDB 개요



2016년도 2학기 2학년 방과후학교

MongoDB 소개

- 10gen사에서 개발한 솔루션 (C++)
- Key-value와 다르게 여러 용도로 사용이 가능 (범용적)
- 스키마를 고정하지 않는 형태
- 스키마 변경으로 오는 문제 없음
- 데이터를 구조화해서 JSON 형태로 저장 (데이터를 key-value화 저장)
- Join이 불가능하기 때문에 join이 필요 없도록 데이터 설계 필요

MongoDB 특징

- 메모리 맵 형태의 파일엔진 DB이기 때문에 메모리에 의존적
- 메모리 크기가 성능을 좌우
- 메모리를 넘어서는 경우 성능이 급격히 저하됨
- 쌓아놓고 삭제가 없는 경우가 적합
ex) “로그 데이터”, “이벤트 참여 내역”, “세션”
- 트랙잭션 처리가 필요한, 금융, 결제, 빌링, 회원정보 등에는 부적합

도큐먼트 데이터 모델

- 속성의 이름과 값으로 이루어진 쌍의 집합
- 속성은 문자열이나 숫자, 날짜 가능
- 배열 또는 다른 도큐먼트를 지정하는 것도 가능함
- 하나의 document에 필요한 정보를 모두 담아야 함
- one query로 모두 해결이 되게끔 collection model 설계를 해야 함
- Join이 불가능하므로 미리 embedding시켜야 함

도큐먼트 형태 구조



```
{  
  "_id" : objectId("4c03e856e528c2701930c091"),  
  "title" : "Welcome to MongoDB",  
  "body" : "Today, we're gonna totally rock your world...",  
  "published" : true,  
  "created" : "Mon May 31 2010 12:48:22 GMT-0400 (EDT)",  
  "updated" : "Mon May 31 2010 12:48:22 GMT-0400 (EDT)",  
  "tags" : [ "databases", "MongoDB", "awesome" ],  
  "comments" : [ {  
    "comment_id" : "9023091210",  
    "author" : "Bob",  
    "email" : "bob@example.com",  
    "body" : "My mind has been totally blown!",  
    "created" : "Mon May 31 2010 12:48:22 GMT-0400 (EDT)"  
  } ]  
}
```

MongoDB 장점

- Schema-less 구조
 - 다양한 형태의 데이터 저장 가능
 - 데이터 모델의 유연한 변화 가능 (데이터 모델 변경, 필드 확장 용이)
- Read/Write 성능 뛰어남
- 장비 확장이 간단함
- 방대한 데이터를 빠르게 처리 가능
 - Memory Mapped(데이터 쓰기 시에 OS의 가상 메모리에 데이터를 넣은 후 비동기로 디스크에 기록하는 방식)를 사용
- JSON 구조 : 데이터 직관적 이해 가능
- 사용 방법 쉽고, 개발이 편리함

MongoDB 단점

- 데이터 업데이트 중 장애 발생 시, 데이터 손실 가능
- 많은 인덱스 사용 시, 충분한 메모리 확보 필요
- 데이터 공간 소모가 RDBMS에 비해 많음 (비효율적인 Key 중복 입력)
- 복잡한 JOIN 사용시 성능 제약이 따름
- Transactions 지원이 RDBMS 대비 미약함
- 제공되는 MapReduce 작업이 Hadoop에 비해 성능이 떨어짐

MongoDB vs MSSQL Server 2008

- 쓰기 작업 성능 비교 : MongoDB가 100배 이상 빠름
- 읽기 작업 성능 비교 : MongoDB가 평균 3배 이상 빠름
- 읽기/쓰기 동시 작업 성능 비교 : 역시 3배 이상 빠름

Number of Parallel Clients 5			Time in seconds			
Basic Insert	Total Rows	Rows / client	SQL Time	Mongo Time	Sql Ops/sec	Mongo Ops/sec
several columns 600 bytes per row	100	20	0.19	0.011	526	9,091
	1,000	200	1.8	0.02	556	50,000
	5,000	1,000	9	0.25	556	20,000
	25,000	5,000	100	1.5	250	16,667
	50,000	10,000	270	2.5	185	20,000

Number of Parallel Clients 5			Time in seconds			
Basic Query	Total Rows	Rows / client	SQL Time	Mongo Time	Sql Ops/sec	Mongo Ops/sec
with index	50	10	0.1	0.08	500	625
	500	100	0.38	0.1	1,316	5,000
	5,000	1,000	2.8	1.2	1,786	4,167
	25,000	5,000	14	4	1,786	6,250
	50,000	10,000	28	10.4	1,786	4,808