



2장. 개략적인 규모 측정

개략적인 규모 측정

보편적으로 사용되는 성능 수치 항목들을 추정해서 계산하는 것

필요한 기본기로 **2의 제곱수**나 **응답지연(latency)** 값과 **가용성**에 관계된 수치들의 이해가 있어야 한다.

2의 제곱수

- 데이터 볼륨 단위를 계산하기 위해 사용 (양을 측정) 최소단위는 1바이트(8비트) 이다.
- 아스키 문자 하나가 차지하는 메모리의 크기가 1바이트이고, UTF-8의 경우 1에서 4바이트가 사용된다.
- 한글 또는 다국어룰 고려할 때 한 단어를 4바이트까지 고려하는 것이 좋다.
- 페타바이트 → 2의 50제곱으로 약 1000조 바이트이다.
- 대략 많이 사용되는 기가바이트는 약 10억으로 2의 30제곱이다.

데이터 볼륨 단위

2의 x 제곱	근사치	이름	축약형
10	1천(thousand)	1킬로바이트	1KB
20	1백만(million)	1메가바이트	1MB
30	10억(billion)	1기가바이트	1GB
40	1조(trillion)	1테라바이트	1TB
50	1000조(quadrillion)	1페타바이트	1PB

응답지연 값

연산명	시간
뮤텍스 락/언락	100ns
주 메모리 참조	100ns
Zippy로 1KB 압축	10us
1Gbps 네트워크로 2KB 전송	20,000ns = 10us
메모리에서 1MB 순차적으로 read	500,000ns = 250us
디스크 탐색	10ms
네트워크에서 1MB 순차적으로 read	10ms
디스크에서 1MB 순차적으로 read	30ms
한 패킷의 캘리포니아로부터 네덜란드까지의 왕복 지연시간	150ms

ns = nanosecond, us = microsecond, ms = millisecond

1ms = 1,000us = 1,000,000ns

결론

- 메모리는 빠르지만, 디스크는 아직도 느리다.
- 디스크 탐색은 가능한 피하라.
- 단순한 압축 알고리즘은 빠르다.
- 데이터를 인터넷으로 전송하기 전에 가능하면 압축하라(gzip)
- 데이터 센터는 보통 여러 지역에 분산되어 있고 센터들 간에 데이터를 주고받는데는 시간이 걸린다.

가용성에 관계된 수치들

고가용성

시스템이 오랜 시간 동안 지속적으로 중단 없이 운영될 수 있는 능력

100%는 시스템이 단 한 번도 중단된 적이 없었음을 의미한다.
퍼센테이지에 9가 많으면 많을수록 좋다.

SLA - Service Level Agreement

| 서비스 사업자가와 고객 사이에 맺어진 합의

해당 합의에는 서비스 사업자가 제공하는 서비스의 가용시간(uptime)이 공식적으로 기술되어 있다.

가용률 시간	연간 장애시간
99%	3.65일
99.9%	8.77시간
99.99%	52.60분
99.999%	5.26분

트위터 QPS와 저장소 요구량 측정

가정

- 월간 능동 사용자는 3억명이다.
- 50%의 사용자가 트위터를 매일 사용한다.
- 평균적으로 각 사용자는 매일 2건의 트윗을 올린다.
- 미디어를 포함하는 트윗은 10% 정도다
- 데이터는 5년간 보관된다.

추정

- 월간 능동 사용자는 3억명이다.
- 50%의 사용자가 트위터를 매일 사용한다.
- $3\text{억} * 50\% = 1.5\text{억명}$

QPS(Query Per Second) 추정치

- 평균적으로 각 사용자는 매일 2건의 트윗을 올린다
- $\text{QPS} = 1.5\text{억} * 2 \text{ 트윗} / 24\text{시간} / 1\text{분}(3600\text{초}) = \text{약 } 3500$
- 최대 QPS = $3500 * 2 = \text{약 } 7000$

미디어 저장을 위한 저장소 요구량

- 평균 트윗 크기(avg)
 - tweet_id 에 64바이트
 - 텍스트에 140바이트
 - 미디어에 1MB
- 미디어를 포함하는 트윗은 10% 정도다
 - $1.5\text{억 } 2(\text{개}) * 10\% * 1\text{MB} = 30\text{TB/일}$
- 데이터는 5년간 보관된다
 - $30\text{TB } 365\text{일} / 5\text{년} = \text{약 } 55\text{PB}$

팁

개략적인 규모 추정과 관계된 면접에서 가장 중요한 것은 문제를 풀어 나가는 절차다.

올바른(논리적인?) 절차를 밟느냐가 결과를 내는 것보다 중요하다. 면접자가 보고 싶어하는 것은 문제 해결 능력이다.

- 근사치를 사용한 계산을 할줄 알아야 한다. 즉 반올림, 올림, 내림에 대한 오차 수준을 생각해서 근사치로 계산시간을 줄인다.
- 가정은 적어 뒤라 위에도 적혀 있다.
- 단위를 붙이는 습관을 뒤라 이건 꼼꼼함을 챙기는 포인트로 보인다 개발자는 꼼꼼함도 중요하다.

- QPS, 최대 QPS, 저장소 요구량, 캐시 요구량, 서버 수 등 추정하는 방법을 미리 연습해 두어라.

참고 링크

<https://velog.io/@haron/가상-면접-사례로-배우는-대규모-시스템-설계-기초-2장-개략적인-규모-측정>