2장

- 개략적 규모 추정 : 시스템에서 다루는 데이터 양
 - 。 2의 제곱수 사용해서 측정
- 응답 지연 : 통상적인 컴퓨터에서 구현된 연산들의 응답지연 값
 - 。 메모리는 빠르지만 디스크는 아직도 느림
 - 。 디스크 탐색은 가능한 피할 것
 - 。 단순한 압축 알고리즘은 빠름
 - 。 데이터를 인터넷에 전송하기 전에 가능하다면 압축할 것
 - 데이터 센터는 여러 지역에 분산되어 있고, 센터들 간에 데이터를 주고받는 데는 시 간이 걸림
- 가용성 : 시스템이 오랜 시간 동안 지속적으로 중단 없이 운영될 수 있는 능력
 - o ex) 100% ⇒ 단 한 번도 중단된 적 없음
 - 대부분의 서비스는 99%~100%의 가용성을 가짐
- QPS : Query Per Second
 - 。 가정
 - 월간 능동 사용자: 3억 명
 - 50% 사용자가 매일 이용
 - 평균적으로 각 사용자는 매일 2건의 트윗 발행
 - 미디어를 포함하는 트윗은 전체의 10%
 - 데이터는 5년간 보관
 - QPS 추정
 - 일간 능동 사용자: 3억 * 50% = 1.5억

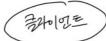
- QPS: 1.5억 * 2트윗 / 24시간 / 3600초 = 약 3500
- 최대 QPS: 2 * QPS = 약 7000
- 。 미디어 저장을 위한 저장소 요구량
 - 평균 트윗 크기
 - 트위터 아이디 : 64 byte
 - 텍스트: 140 byte
 - 미디어: 1MB
 - 미디어 저장소 요구량 : 1.5억 * 2 * 10% * 1MB = 30TB
 - 5년간 미디어를 보관하기 위한 저장소 요구량: 30TB * 365 * 5 = 약 55TB

2장

3장

- 정확한 설계를 위한 질문 목록
 - 구체적으로 어떤 기능을 만들어야 하는가?
 - 제품 사용자 수는 얼마나 되는가?
 - 회사의 규모는 얼마나 빠르게 커지리라 예상하는가? 3개월, 6개월, 1년 뒤의 규모는 얼마가 되리라 예상하는가?
 - 。 회사가 주로 사용하는 기술스택은 무엇인가?
 - 설계를 단순화하기 위해 활용할 수 있는 기존 서비스로는 어떤 것들이 있는가?
- 설계 실습?
 - 。 가정
 - 서버에 접속한 모든 유저들은 채팅 메시지를 전송할 수 있고, 전송된 채팅은 모든 유저가 볼 수 있다.
 - 서버 당 최대 접속자 수는 5000명
 - DAU는 3000명
 - 길이가 100자로 한정된 텍스트만 가능
 - 금칙어가 포함되어 있다면 해당 텍스트는 *로 대체됨
 - 24시간 동안 채팅 내용이 저장됨
 - 설계 ㅎㅎ;

〈洲罗7457



0 / 19



21

洲門得

3) J

শিন্ত সাশ

④ ↓

세팅 DB

< 对智 \$127



31

MH

0110

八张朝人