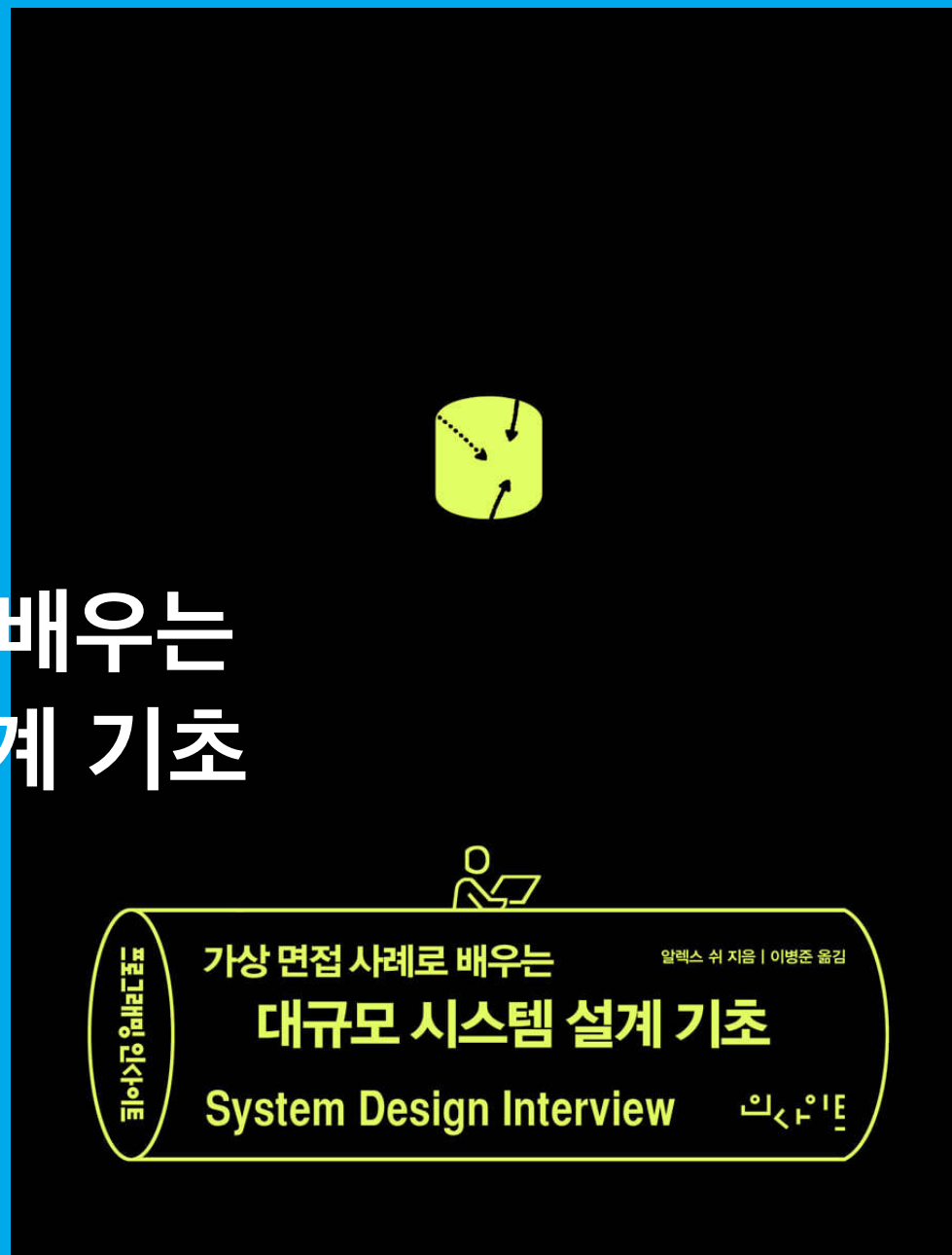


가상면접 사례로 배우는 대규모 시스템 설계 기초

PPT by 김주혁



목차

- 01 문제 이해 및 설계 범위 확정
- 02 개략적 설계안 제시 및 동의 구하기
- 03 상세 설계
- 04 마무리


8장 URL 단축기 설계

URL 단축기 ... ?



SEO.





1단계 문제 이해 및 설계 범위 확정

URL 단축키를 위한 요구사항
가정

- 쓰기 연산: 매일 1억 개의 단축 URL 생성
- 초당 쓰기: $1\text{억}(100\text{million}) / 24 / 3600 = 1160$
- 읽기 연산: 읽기 연산과 쓰기 연산의 비율은 10:1로서, 초당 11600개의 읽기 요청이 발생한다.
- 10년간 단축 서비스를 운영한다고 치면, 3650억개의 레코드를 보관할 수 있어야 한다.
- 축약 전 평균 URL의 길이는 100이라고 한다.
- 10년 동안 필요한 저장 용량은 $3650\text{억} * 100\text{바이트} \Rightarrow 36.5\text{TB}$ 이다.



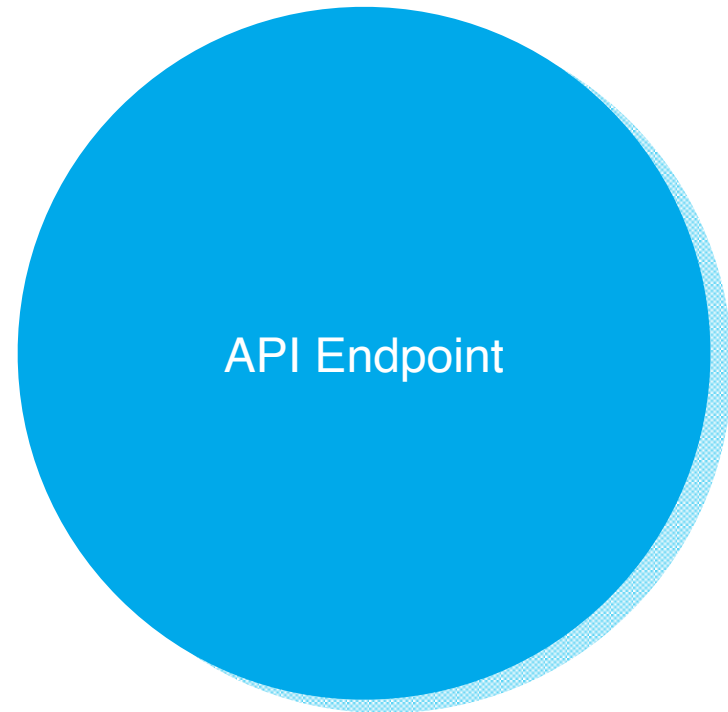
2단계 개략적 설계안 제시 및 동의 구하기



API Endpoint

URL Redirection

URL 단축



URL Redirection

Request URL: <https://tinyurl.com/qtj5opu>

Request Method: GET

Status Code: 301

Remote Address: [2606:4700:10::6814:391e]:443

Referrer Policy: no-referrer-when-downgrade

▼ Response Headers

alt-svc: h3-27=":443"; ma=86400, h3-25=":443"; ma=86400, h3-24=":443"; ma=86400, h3-23=":443"; ma=86400

cache-control: max-age=0, no-cache, private

cf-cache-status: DYNAMIC

cf-ray: 581fbd8ac986ed33-SJC

content-type: text/html; charset=UTF-8

date: Fri, 10 Apr 2020 22:00:23 GMT

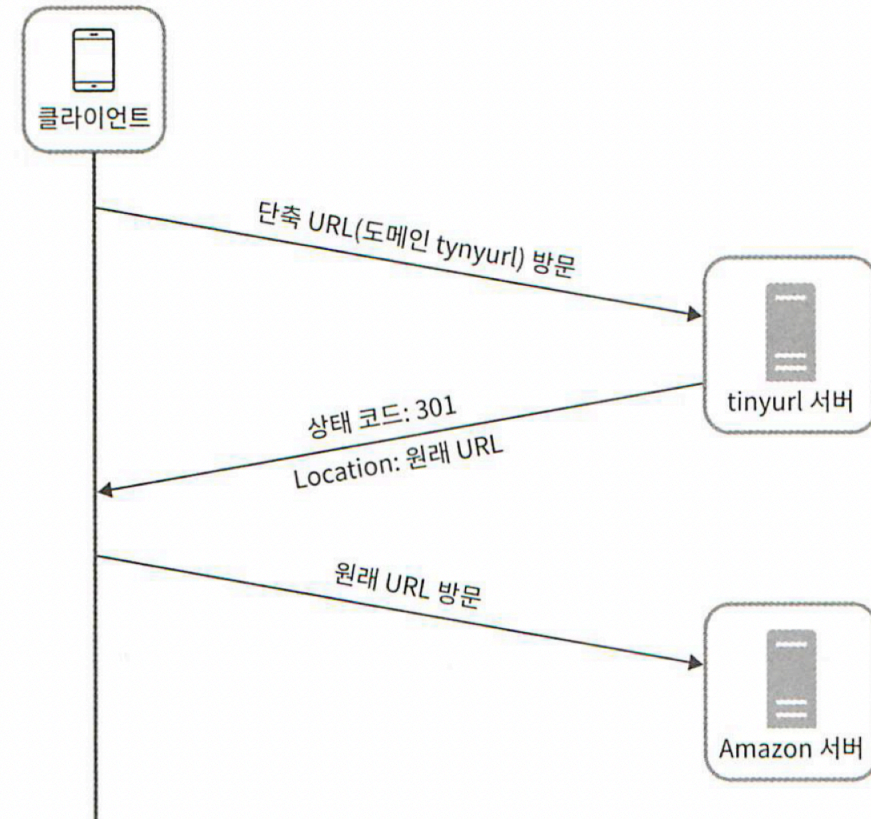
expect-ct: max-age=604800, report-uri="https://report-uri.cloudflare.com/cdn-cgi/beacon/expect-ct"

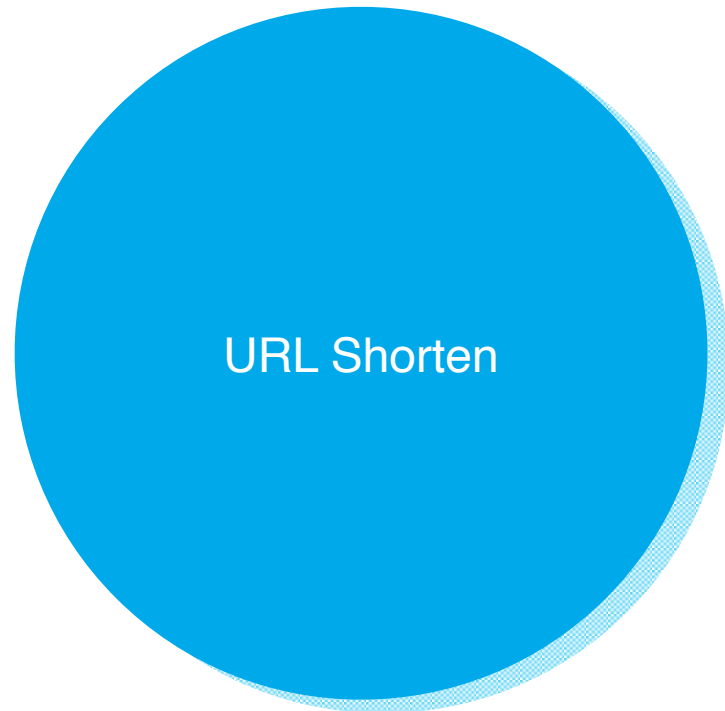
location: https://www.amazon.com/dp/B017V4NTFA?pLink=63eaeef76-979c-4d&ref=adblp13nvxx_0_2_im

URL Redirection

단축 URL: <https://tinyurl.com/qt5opu>

원래 URL: https://www.amazon.com/dp/B017V4NTFA?pLink=63eaf76-979-4d&ref=adblp13nvxx_0_2_im





URL Shorten



Hash Table



3단계 상세 설계



데이터 모델

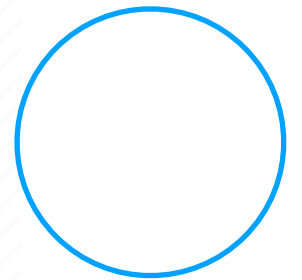


해시 함수



URL 단축기 상세 설계



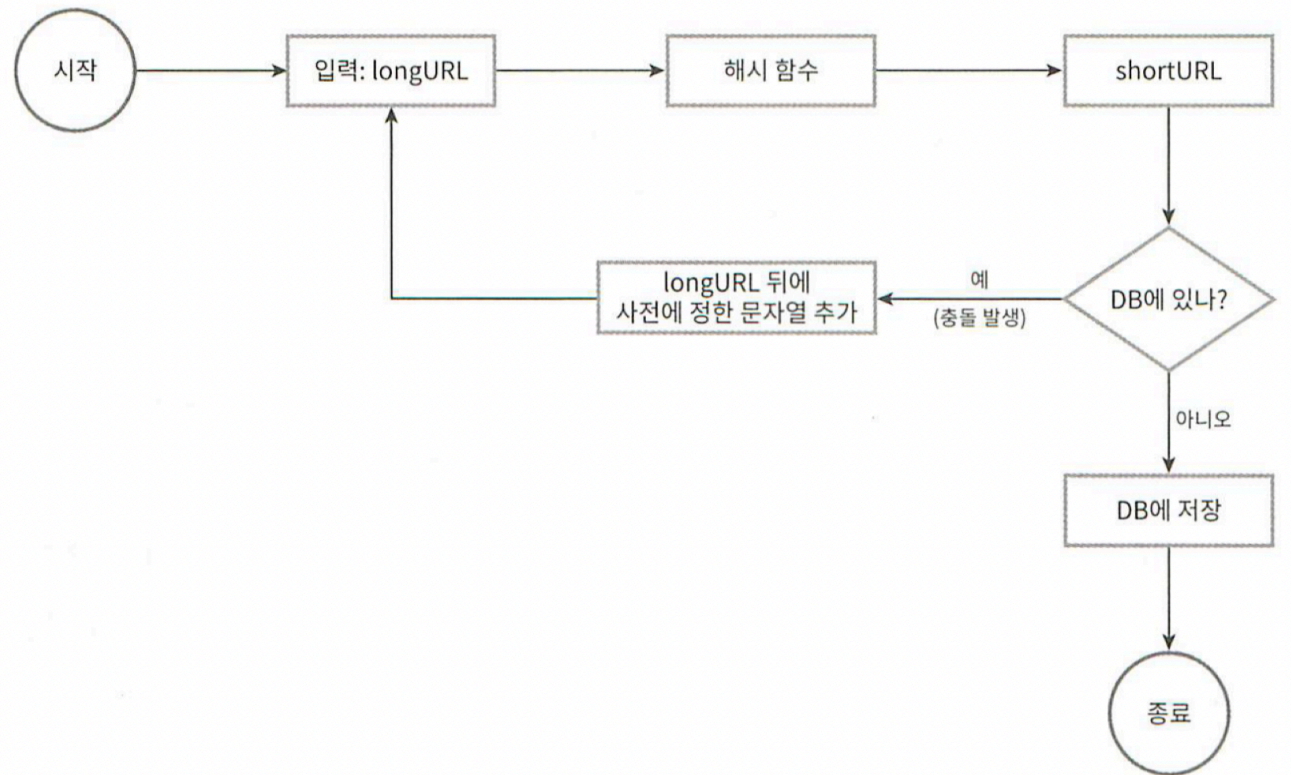


해시 함수
- 해시 충돌 후 해소

n	URL 개수
1	$62^1 = 62$
2	$62^2 = 3,844$
3	$62^3 = 238,328$
4	$62^4 = 14,776,336$
5	$62^5 = 916,132,832$
6	$62^6 = 56,800,235,584$
7	$62^7 = 3,521,614,606,208$
8	$62^8 = 218,340,105,584,896$

해시 함수	해시 결과 (16진수)
CRC 32	5cb54054
MD5	5a62509a84df9ee03fe1230b9dfb84e
SHA-1	0eeae7916c06853901d9ccbefbfcaf4de57ed85b

해시 함수 - 해시 충돌 후 해소



A large blue circle with a subtle gradient and a slight shadow, containing the text '해시 함수 - Base-62'.

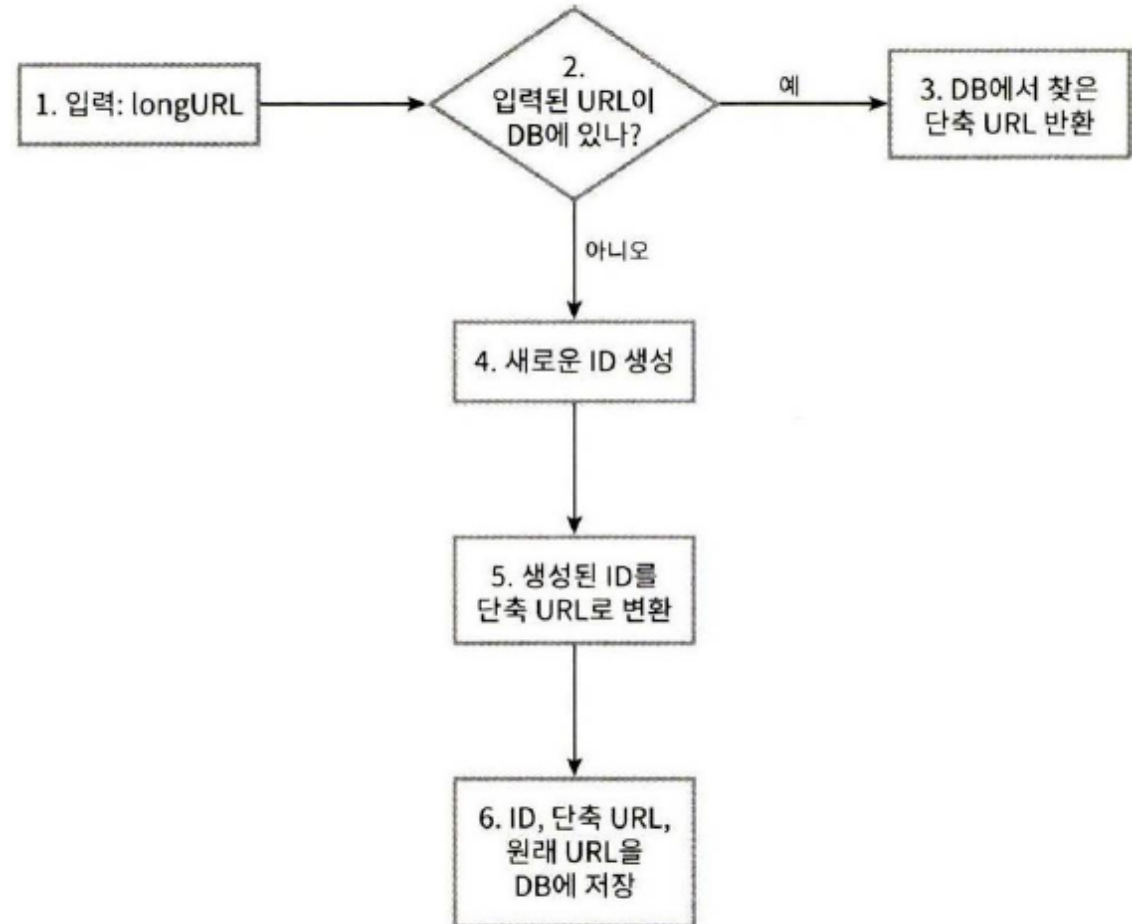
해시 함수
- Base-62

A-Z

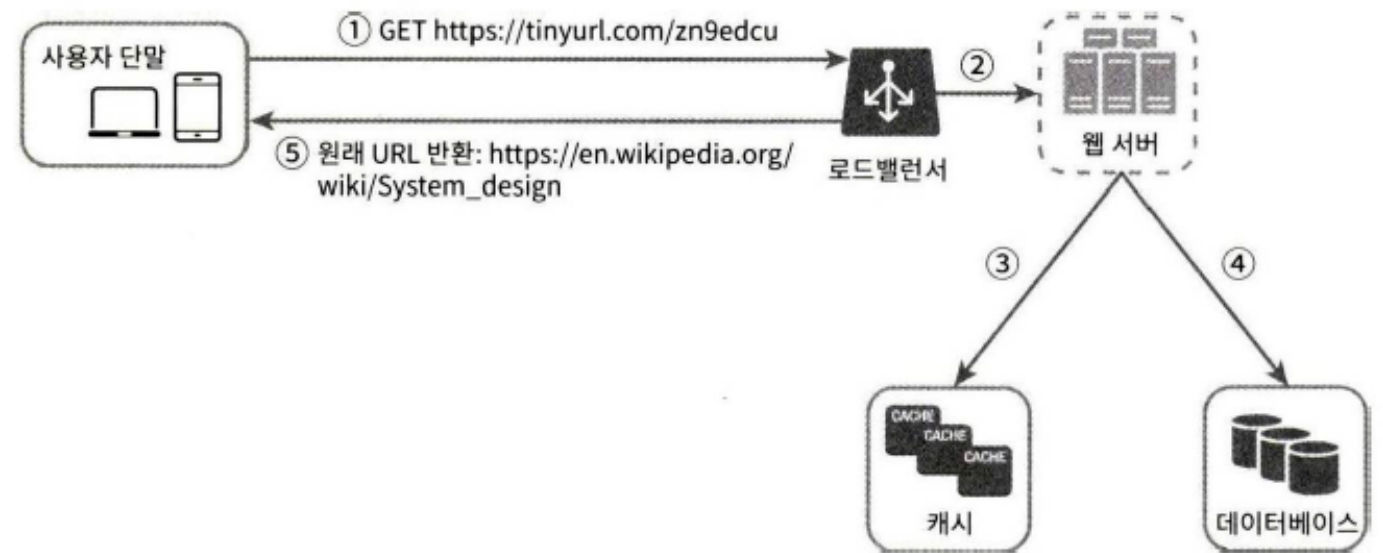
해시 함수

해시 후 충돌 해소 전략	base-62 변환
단축 URL의 길이가 고정됨	단축 URL의 길이가 가변적. ID 값이 커지면 같이 길어짐
유일성이 보장되는 ID 생성기가 필요치 않음	유일성 보장 ID 생성기가 필요
충돌이 가능해서 해소 전략이 필요	ID 유일성이 보장된 보장된 후에야 적용 가능한 전략이라 충돌은 아예 불가능
ID로부터 단축 URL을 계산하는 방식이 아니라서 다음에 쓸 수 있는 URL을 알아내는 것이 불가능	ID가 1씩 증가하는 값이라고 가정하면 다음에 쓸 수 있는 단축 URL이 무엇인지 쉽게 알아낼 수 있어서 보안상 문제가 될 소지가 있음

URL 단축기 상세 설계



URL Redirection 상세 설계



4단계 마무리

- **처리율 제한 장치(rate limiter)**
 - 지금까지 살펴본 시스템은 엄청난 양의 URL 단축 요청이 있을 경우 무력화될 수 있다는 잠재적 보안 결함을 갖고 있다.
처리율 제한 장치를 두면, IP 주소를 비롯한 필터링 규칙들을 이용해 요청을 걸러낼 수 있을 것이다.
- **웹 서버의 규모 확장**
 - 본 설계에 포함된 웹 계층은 무상태 계층이므로, 웹 서버를 자유롭게 증설하거나 삭제할 수 있다.
- **데이터베이스의 규모 확장**
 - 데이터베이스를 다중화하거나 샤딩(sharding)하여 규모 확장성을 달성할 수 있다.
- **데이터 분석 솔루션(analytics)**
 - 성공적인 비즈니스를 위해서는 데이터가 중요하다.
URL 단축기에 데이터 분석 솔루션을 통합해 두면 어떤 링크를 얼마나 많은 사용자가 클릭했는지, 언제 주로 클릭했는지 등 중요한 정보를 알아낼 수 있을 것이다.
- **가용성, 데이터 일관성, 안정성**
 - 대규모 시스템이 성공적으로 운영되기 위해서는 반드시 갖추어야 할 속성들이다.

SEO.