Design a Unique ID generator.

     ↳ At Scale.

⇒ Auto Increment Ids.

Count = 0 1 2 3 4
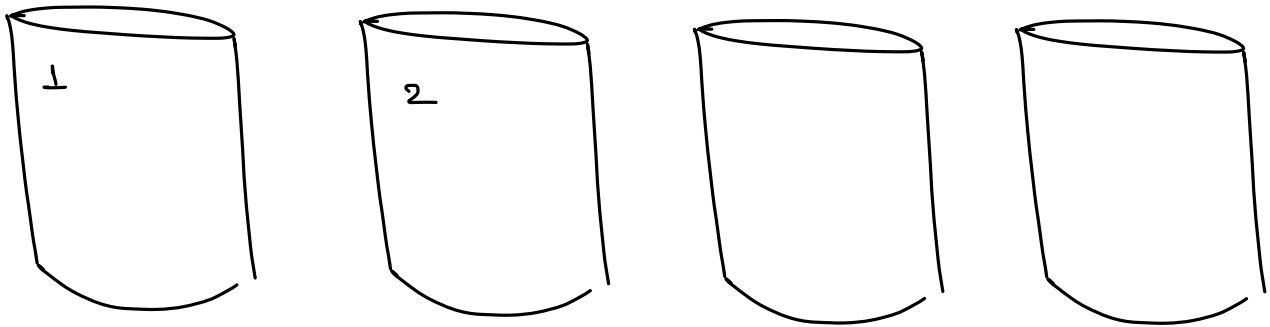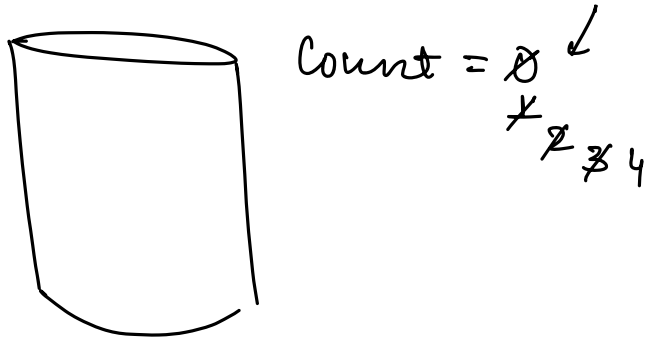
⇒ Sparse

Id's : 1, 2, 3, 4, 5, 6, - - - -] Dense

Youtube : youtube.com/videos/✗

                                  axY249zmgNAX

Amazon Pay Voucher. => xxxx - xxxx . . . . .

Aadhar Card : _ _ _ _ _ _ _ _ _ _ _ _

aadhar.gov.in | details | 1234-5678-1234

_ _ _ _ _ _ _ _ _ _ _ _

↑

$[0-9] = 10$ possibilities.

Unique aadhar No's possible $= 10^{12}$

No. of users / Population $= 1.5 B$

$$\approx 1.5 \times 10^{9}$$

$$= \frac{1.5 \times 10^{9}}{10^{12} \quad 10^{3}}$$

$$= \frac{1.5}{10^{3}} \times 100 \ \%$$

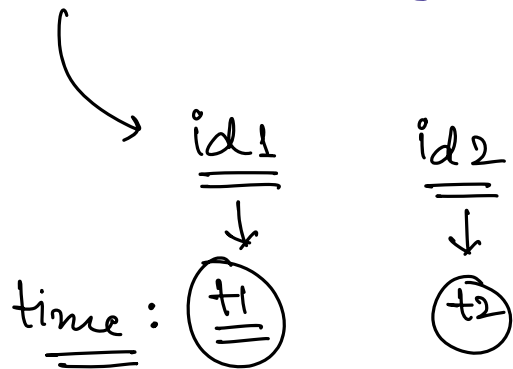$$= \frac{1.5}{10} = \underline{0.15 \%}$$

ID generation can happen on client or backend side.

Server Side ID generation

I) ID should be unique across multiple servers.

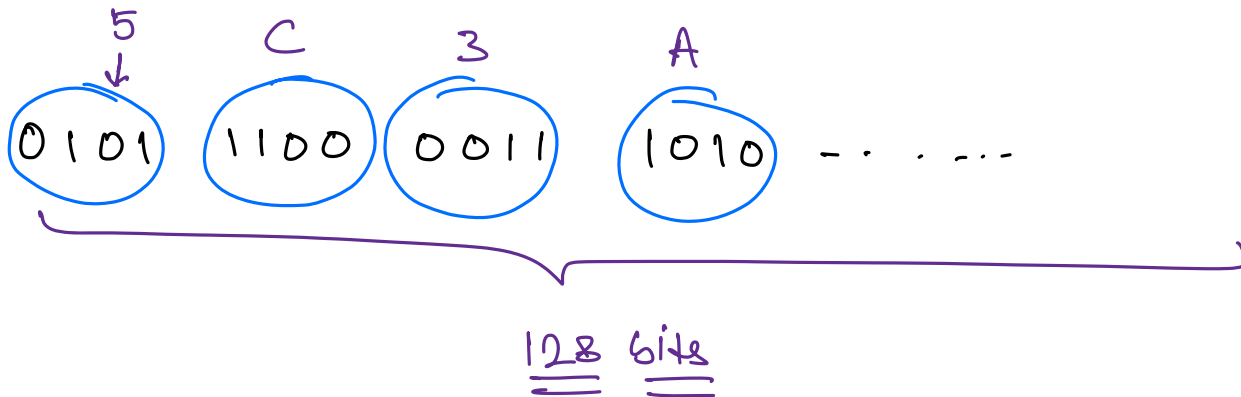II) ID's should be sparse.

III) ID's should be incremental.

$$id_1 \qquad id_2$$

$$\downarrow \qquad \downarrow$$

time : $\boxed{t_1} \qquad \boxed{t_2}$

if $t_2 > t_1 \Rightarrow id_2 > id_1$

PK

Post_id	
—	
—	
—	
—	

sorted

UUID.

↳ 128 bit Numbers.

⇒ 2^{128}

5   C   3   A

0101   1100   0011   1010 — · · · ···

128 bits

$$\frac{128}{4} = 32$$

4 bits : 0-15

↓

0000 to 1111

Hexadecimal.

5C3A — ☐ — ☐

0
⋮
⋮
9

10	A
11	B
12	C
13	D
14	E
15	f

\# timestamp + device_id

Client info

$t1$ $t1$

SQL
DB

Count = $\emptyset$
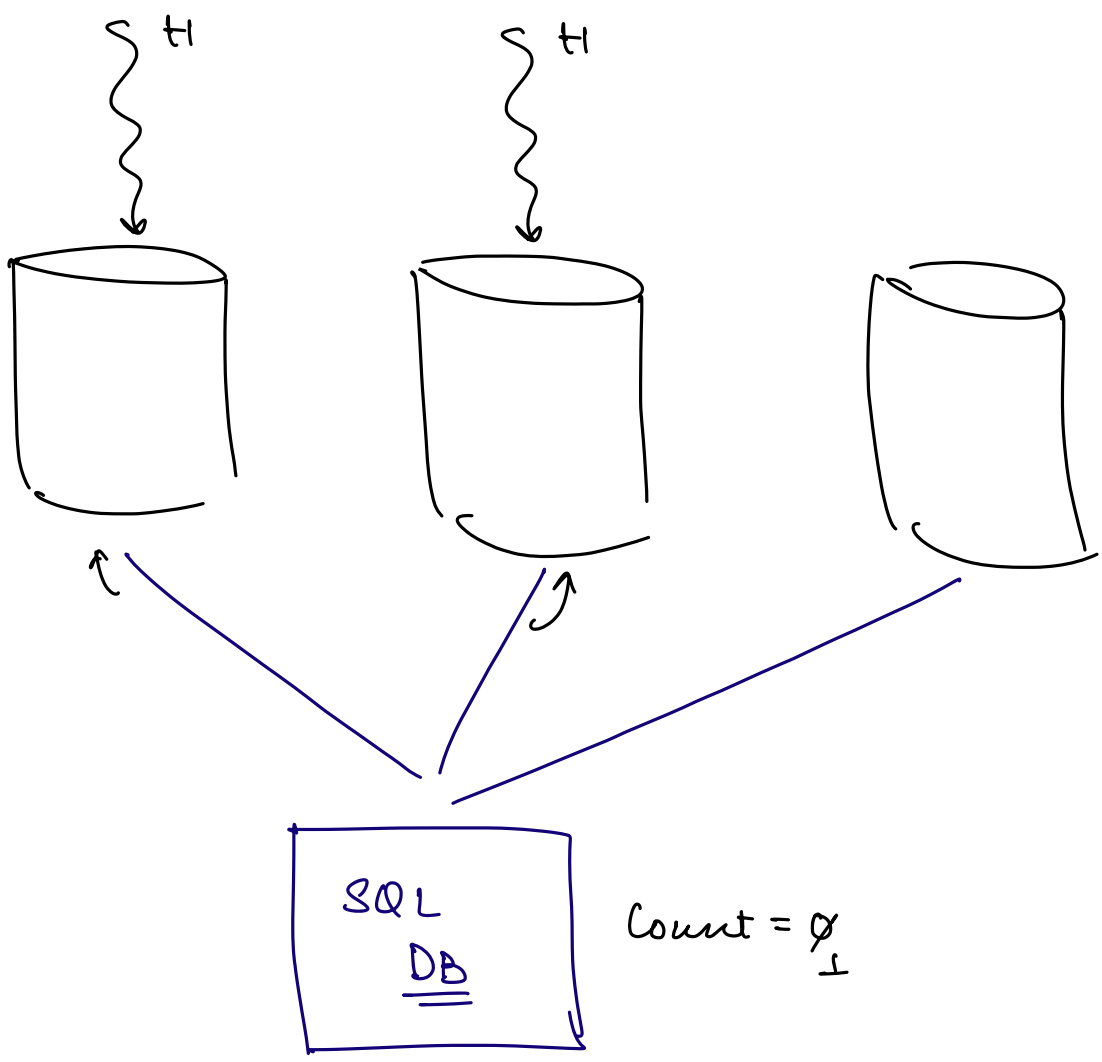
→ Extra N/w hop
    ↳ Slow.

→ Bottleneck.

Twitter's Snowflake Algorithm.

→ 2010.

⇒ 64-bits.

1	41 Bits	5 Bits	5 Bits	12 Bits

MSB → Sign Bit. → Soft Delete

Timestamp.

datacenter id

machine id

Sequence #

LSB.
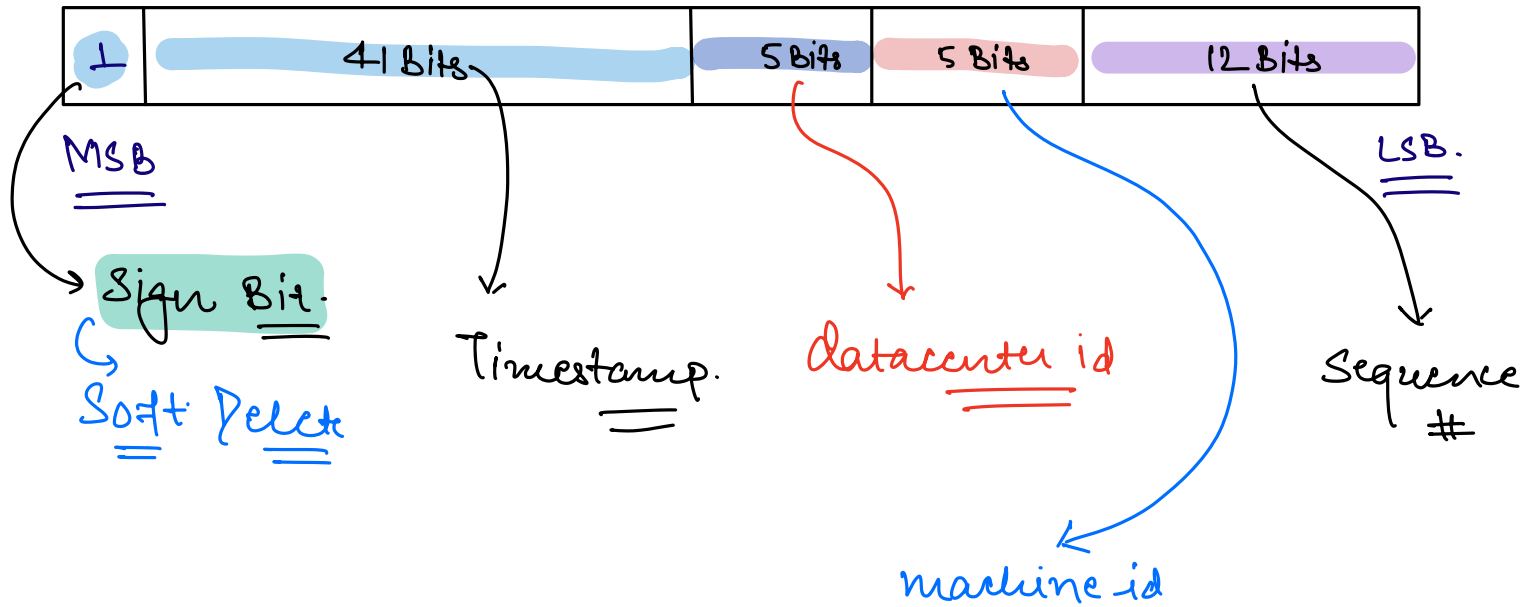
timestamp ⇒ epoch time

↳ How many milli seconds have passed after the cut off date.

01-01-1970.

Twitter ⇒ 4th Nov, 2010.

41 bits $\Rightarrow$ 2^{41}

$\Rightarrow$ 2199023255552 ms.

$\Rightarrow$ $\dfrac{2199023255552}{1000}$ sec

$\Rightarrow$ $\dfrac{2199023255552}{1000 \times 60 \times 60 \times 24 \times 365}$ years.
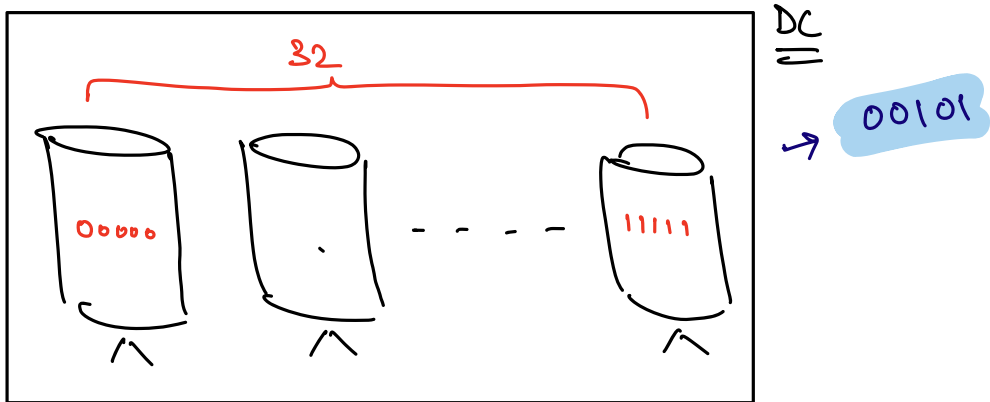
$\simeq$ 70 years.

5 Bits $\Rightarrow$ Datacenter id.

$\hookrightarrow$ 0-31 $\Rightarrow$ 32 no's.
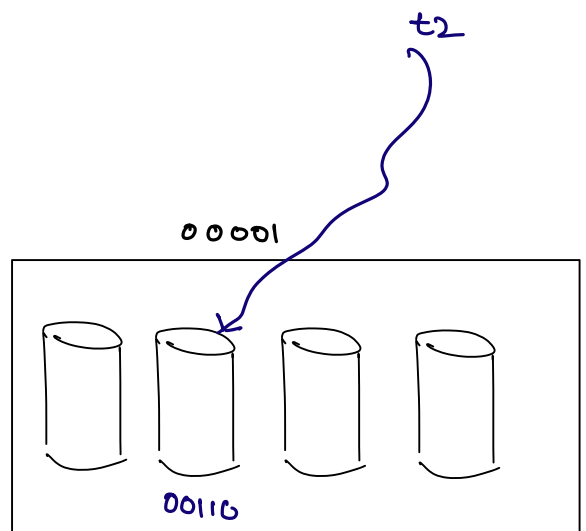
$\rightarrow$ 32 Datacentres.

5 Bits $\Rightarrow$ m/c id.
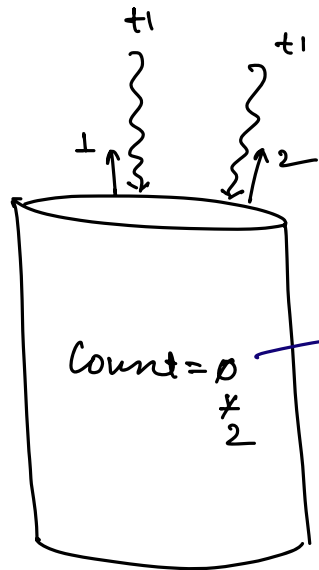
$\hookrightarrow$ 0-31 $\Rightarrow$ 32 no's.

$\rightarrow$ 32 m/c per datacenter.

DC

00101

$$\text{Max no. of m/c} = 32 \times 32 = 2^{10}$$

$$= \underline{\underline{1024.}}$$

DC

10 Bits → M/c

t_1 t_1

00000

00101   00110

t_2

00001

00110
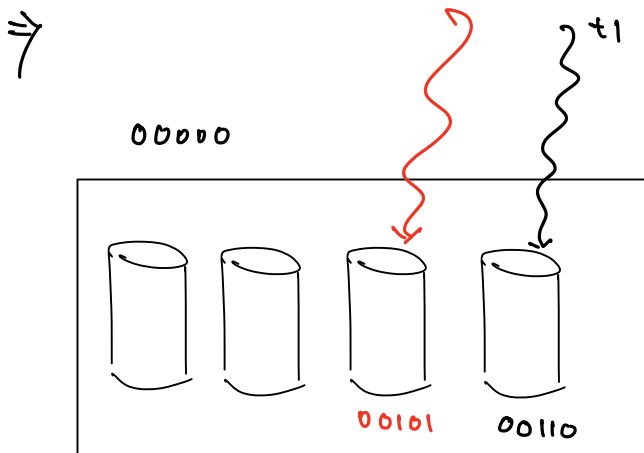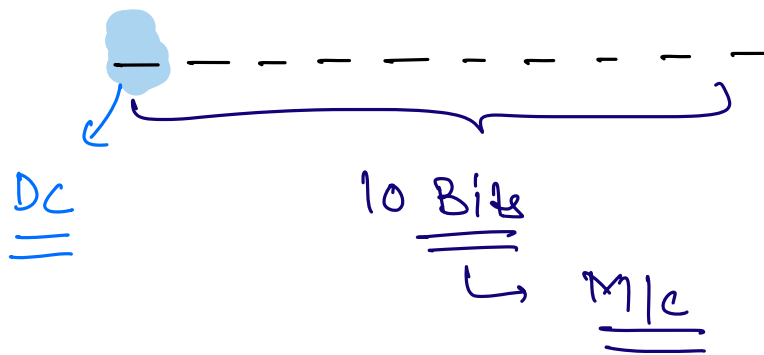
12 Bit : Sequence No.

Race Condition

Count = 0
*
2

12 Bits.
$\hookrightarrow 2^{12} = $ (4096.)

get reinitialized
after every ms.

$1ms \Rightarrow 4000$
$1sec \Rightarrow 4 \times 10^6$ (4M 9PS)