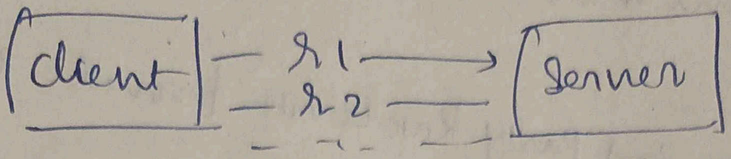
Rate limiting



problems:-

(i) Client can send too many request within a timeframe.

DDOS / DOS

(ii) Server cost expensive
"Each req is expensive".

Hum rate limiter ko implement kaha kare?

option:

(i) Client side → Isme dikkat ye h ki, client side ka code access client ke paas hota to wo use change kar sakta

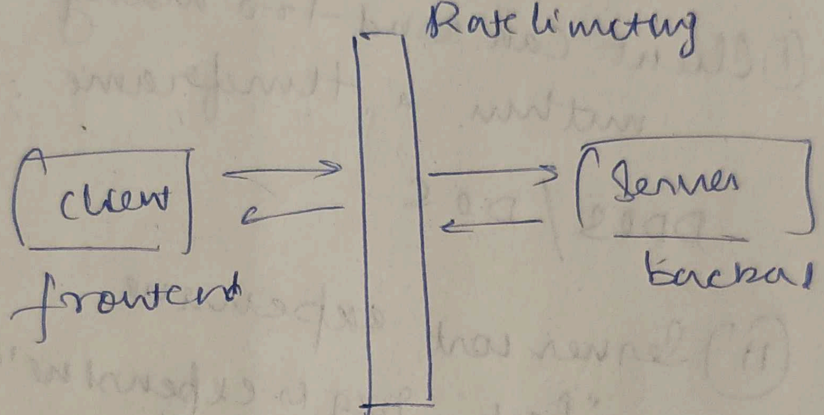
(ii) Server side

(iii) middle ware.

Ye dono aache tareeke h

Case ①:- Agar aapka application fast language
Java -, C, C++, Java
mai bana h to Server Side bhi chalega.

Case ②:- Generally 3rd Party Rate limiter provide like Amazon
aws (API Gateway)



Aacha to Rate limiting karni h, Samjh gaya par
Kispe karni h?

Option:-

① IP based

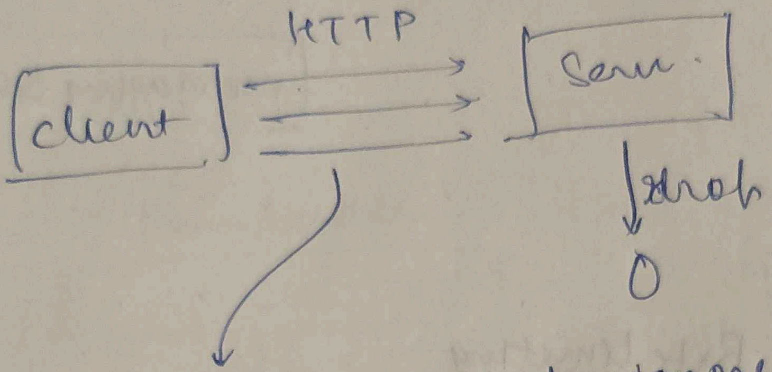
(ii) UserId based.

(iii) other way.

Aap teen tareeko se Rate limit ko handle kar Sakte ho!

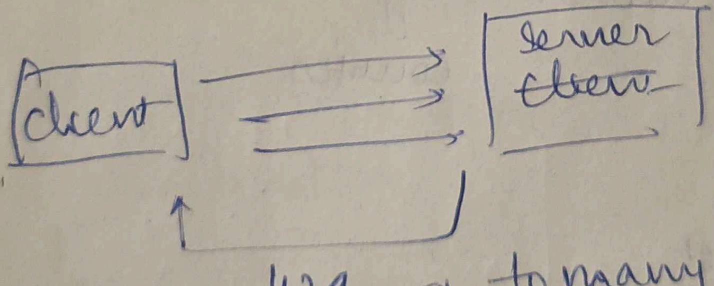
let say mai 2 req/sec serve karunga bas. (from same user)

① Drop



Is request ~~ko~~ ko ignore kar diya, aur client ko koi bhi response ni diya.

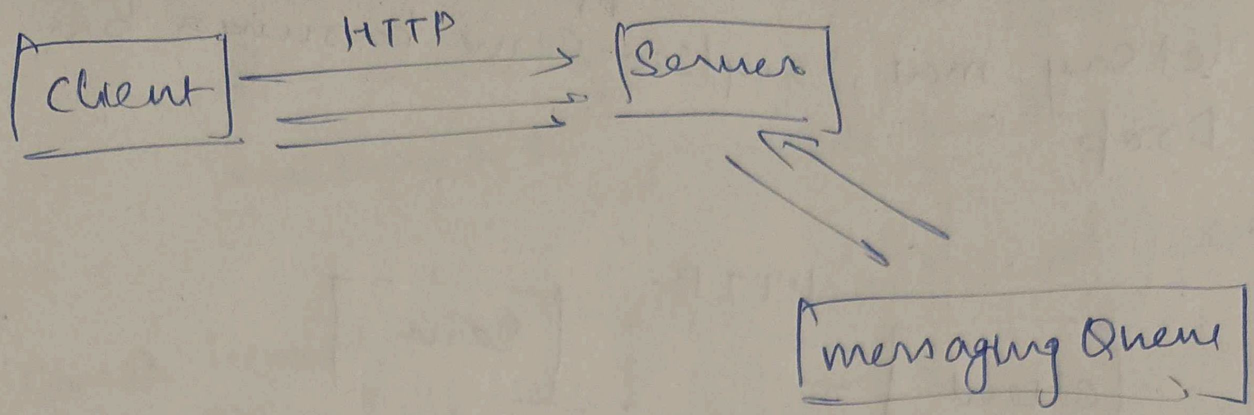
(ii) Responsed with drop



429 → to many requests

(iii) messaging Queue

→ Request ko ek messaging queue ~~to~~ mai daal de, taki baad mai jab server free ho to req ko serve karde.



Algorithms of Rate limiting

(i) Token Bucket

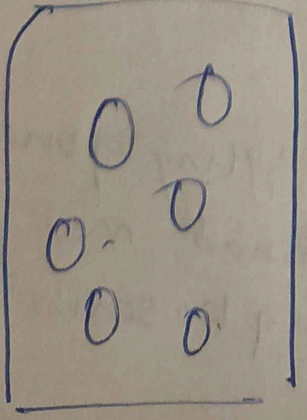
(ii) leaky Bucket

(iii) Fixed window Counter

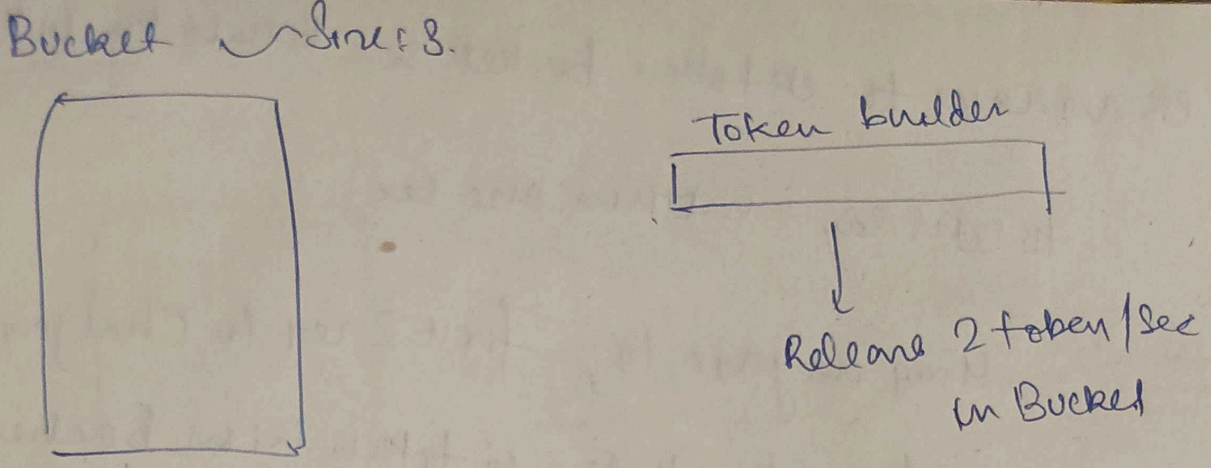
(iv) Sliding window log

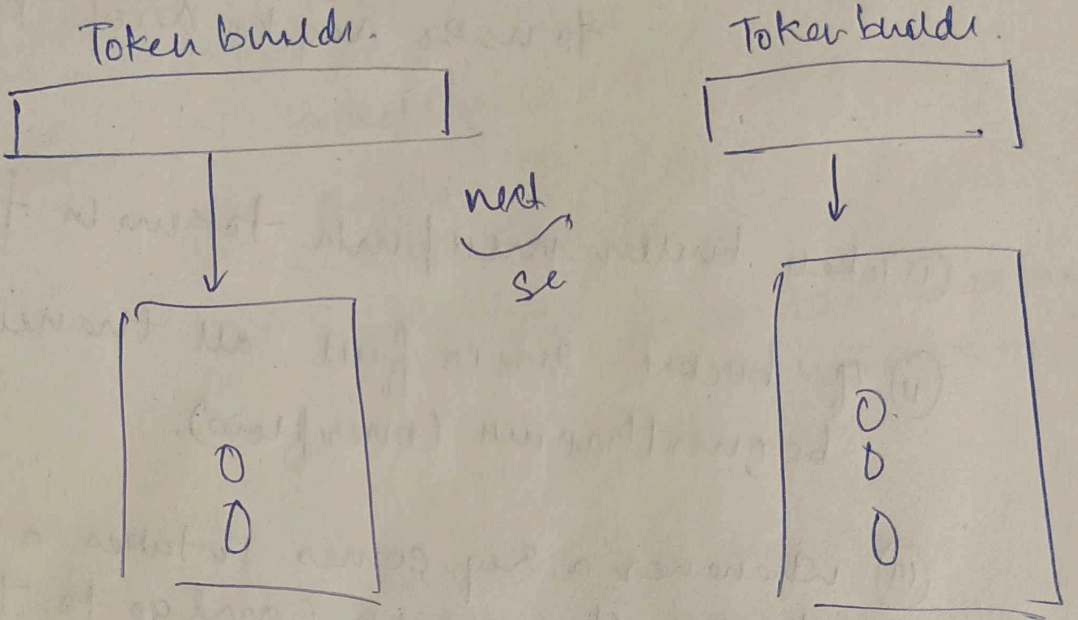
(v) " " counter.

(i) Token Bucket

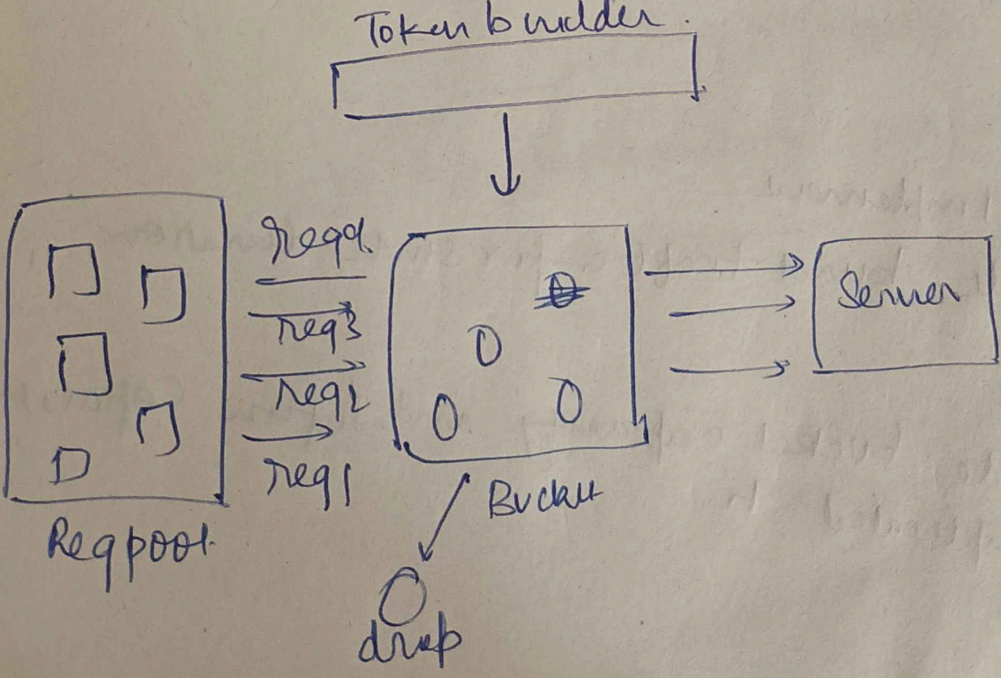


Request Pool





4th one overflow ho jayega as size 3 hi h.



ek req aaya to ek token bucket se khali karta h,

to let's say within one sec,

4 req aa gaye to, first 3 req to chal jayenge

par 4th ke liye to token hi ni bachenge

to usko req ko drop kar denge.

(i) Token builder will push tokens in fixed intervals

(ii) If bucket size is full all the new tokens will be overthrown (overflow).

(iii) whenever a req comes it takes a token from bucket if available and go to the server,

(iv) If bucket is empty, req is dropped.

PRO's

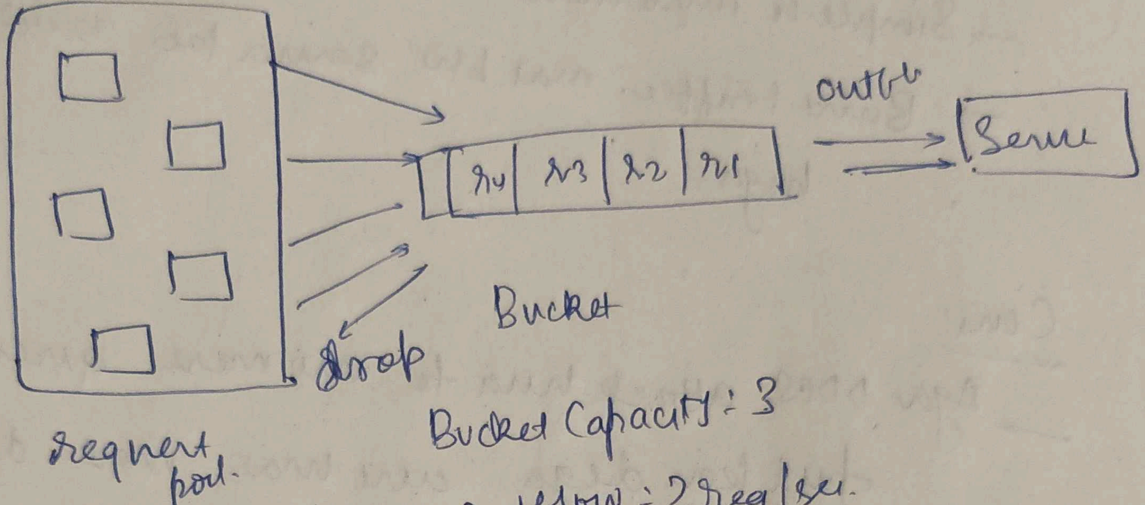
→ Simple to implement

→ Can handle burst traffic for small duration

Cons:

(i) To decide bucket capacity and inflow capacity complicated h.

Leaky Bucket



Bucket Capacity: 3

Outflow: 2 req/sec.

Server ko ek baar mai 2 req uthata h ba

to ye burst traffic handle ni kar sakta kyuki chahe kitne bhi req aa jaye ye bass 2 req uthayega ek sec mai

Ye tab lena chahiye jab aap Server ko Safe tareeke se chalna chahte ho

Pros:

→ Simple to implement

→ Burst traffic mai bhi server ko kuch nhi hoga.

Cons:

→ Agar DDOS attack hua to wo mere queue ko full kar dega aur mai meri dusre kaam ke req ko drop kar dunga ~~to~~ jo ki mera hi loss h.

58:00