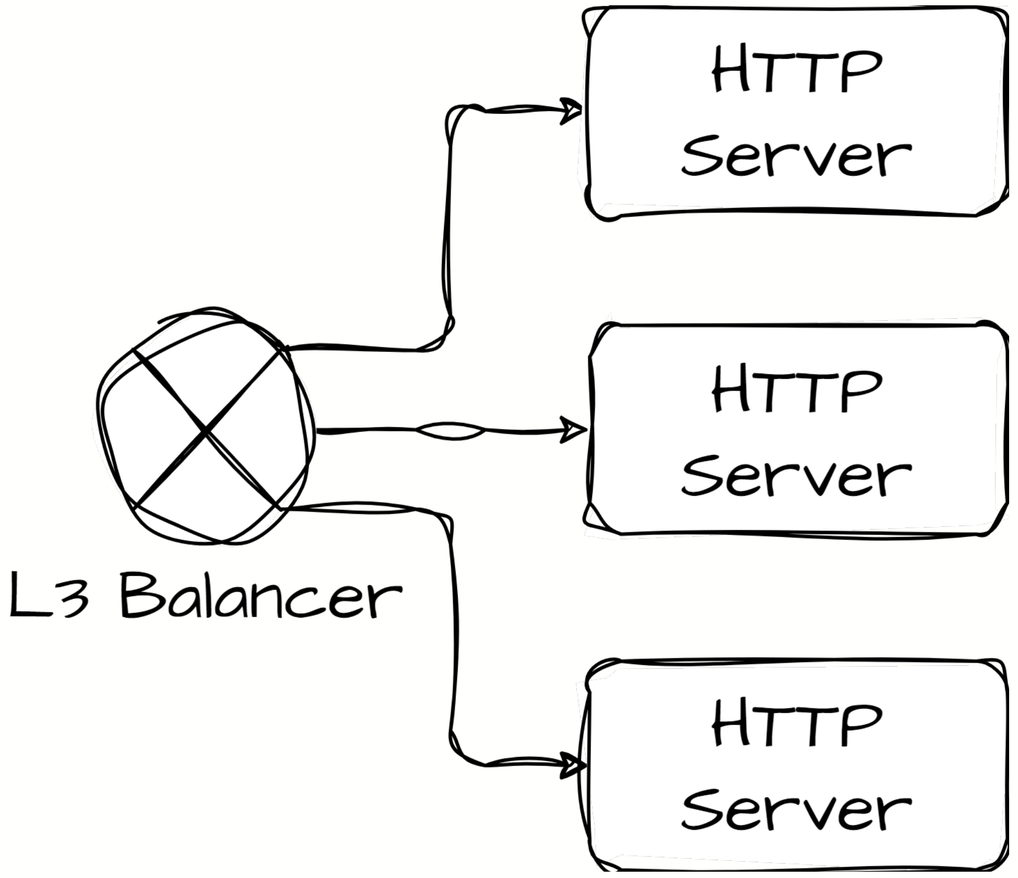
Infrastructure

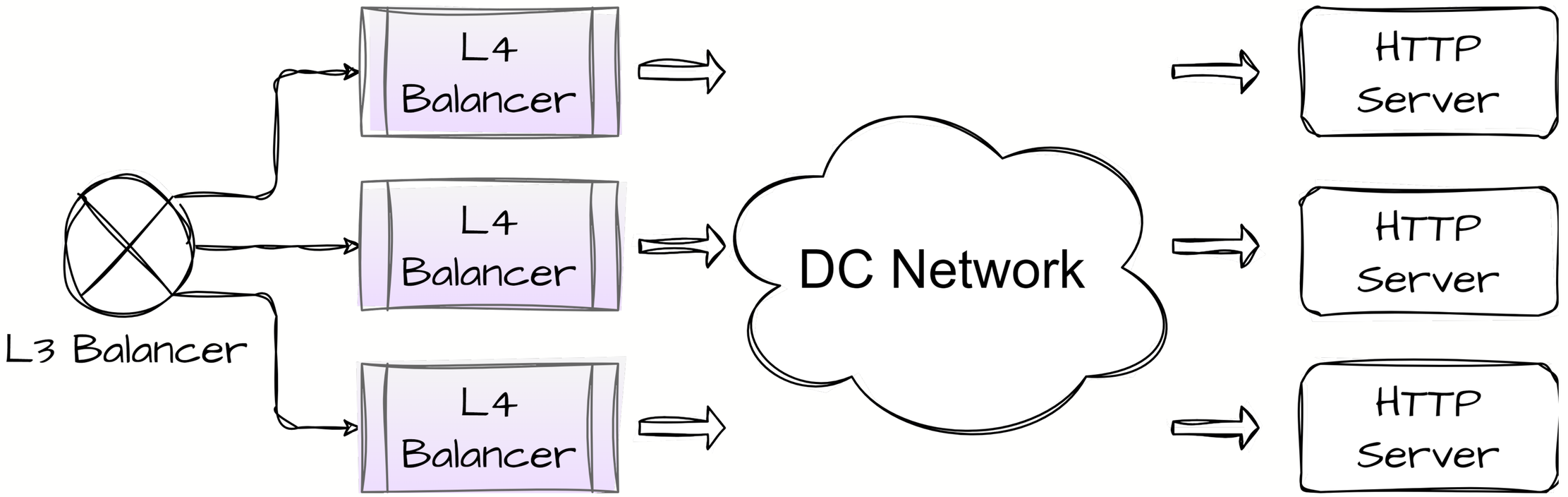
Nearly Stateless L4 Balancer

Alexander Azimov
Lev Pantiukhin

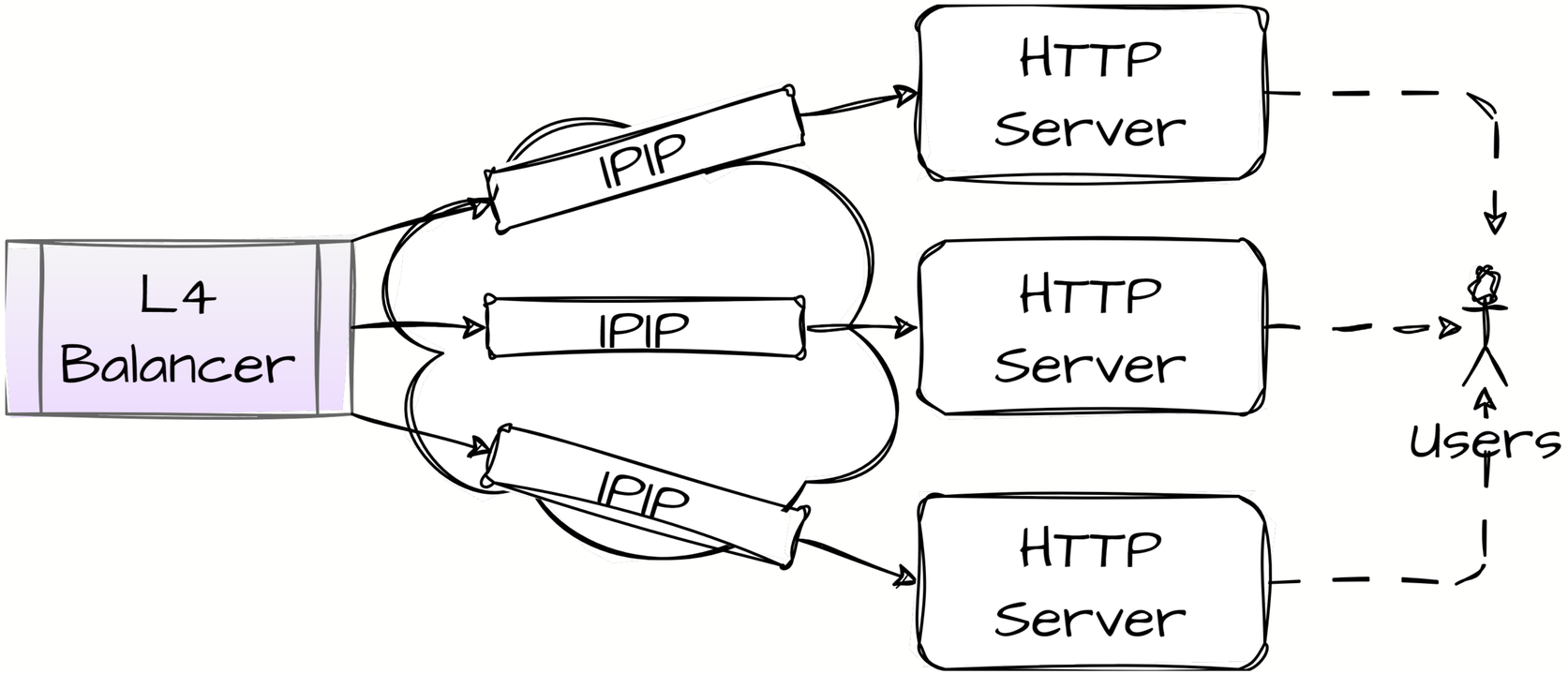
Load Balancing



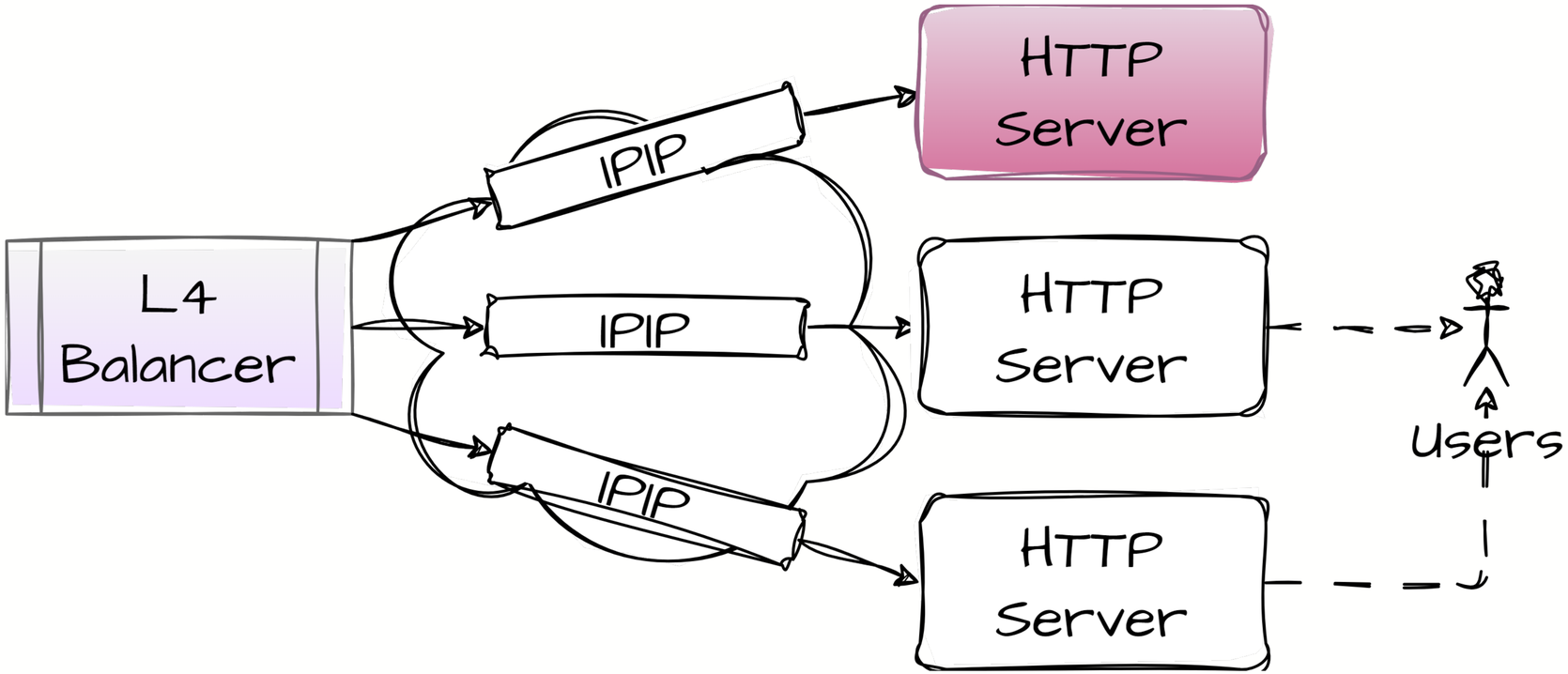
Multilayer Load Balancing



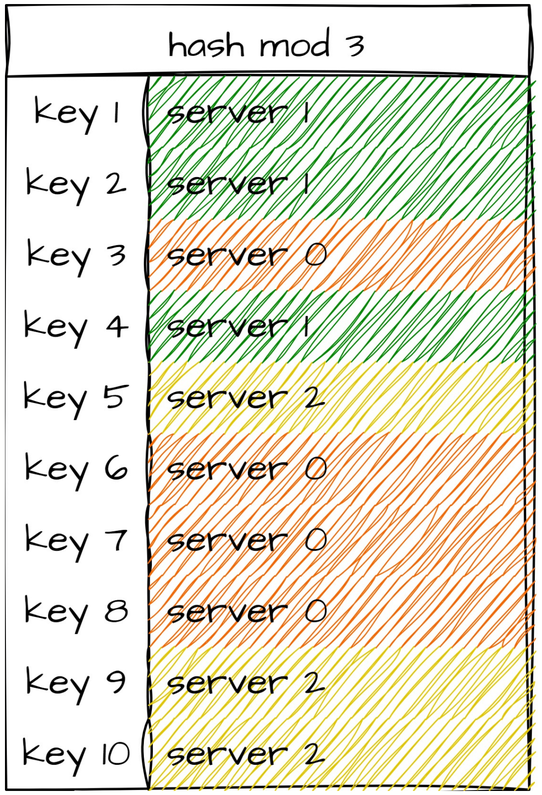
L4 Balancer: DSR

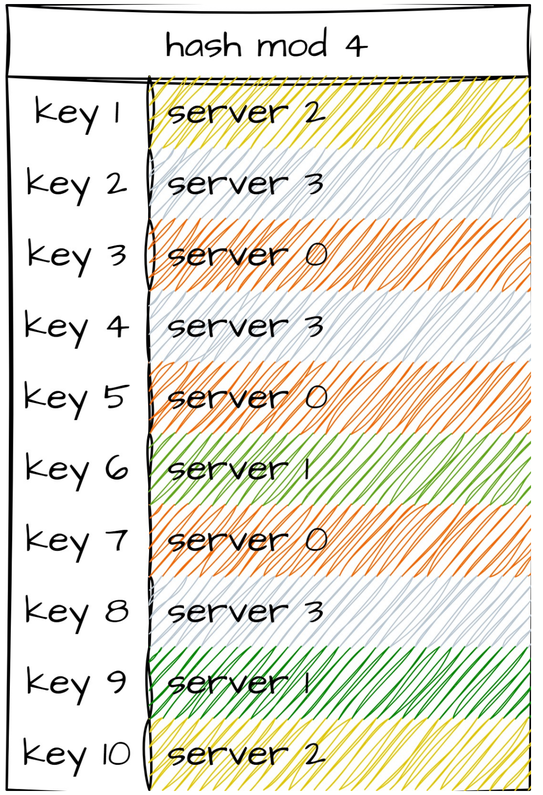


L4 Balancer: Health Checks

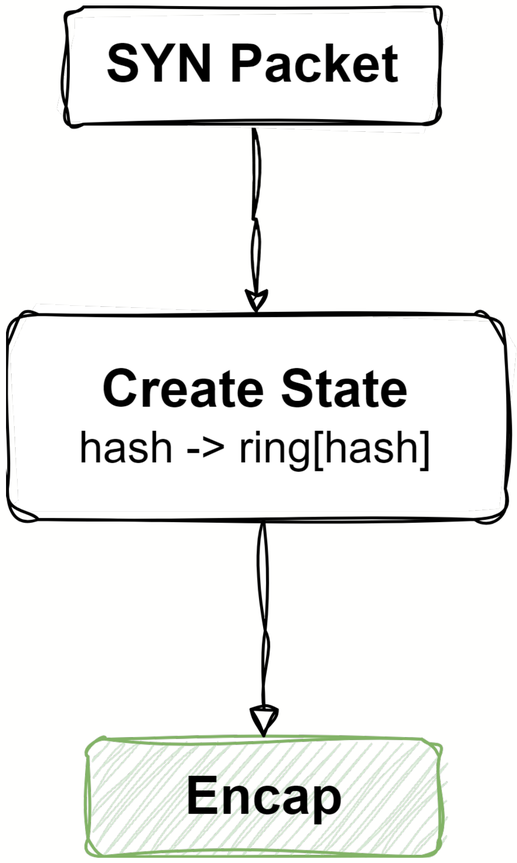


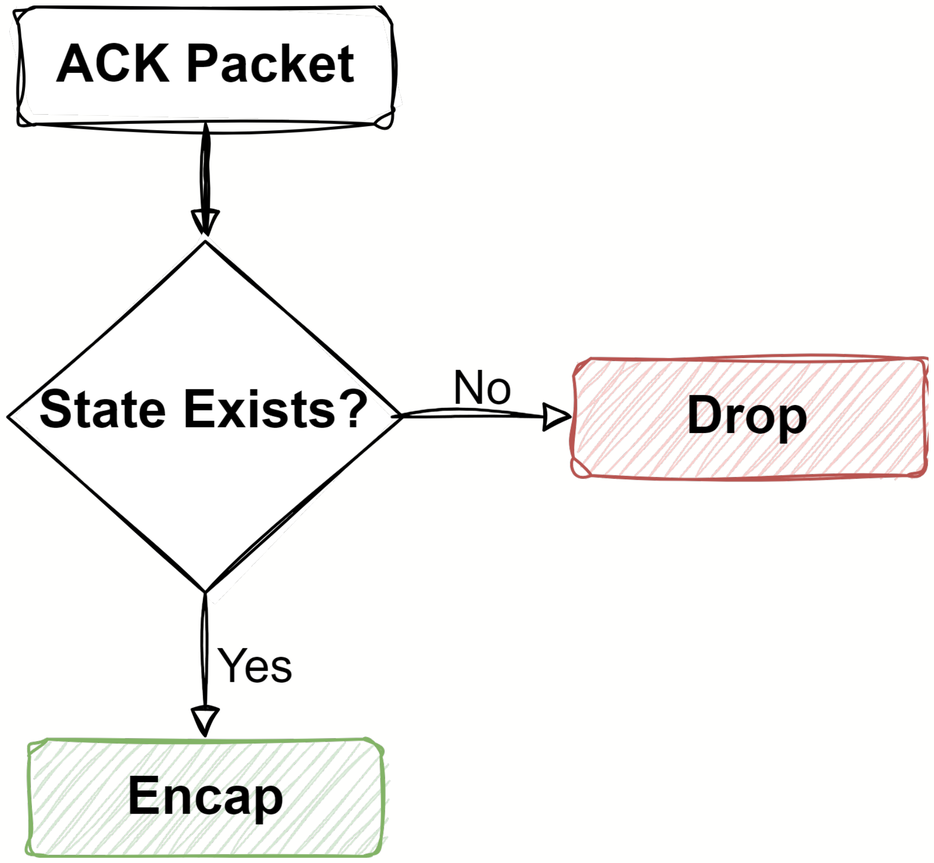
Hash and State



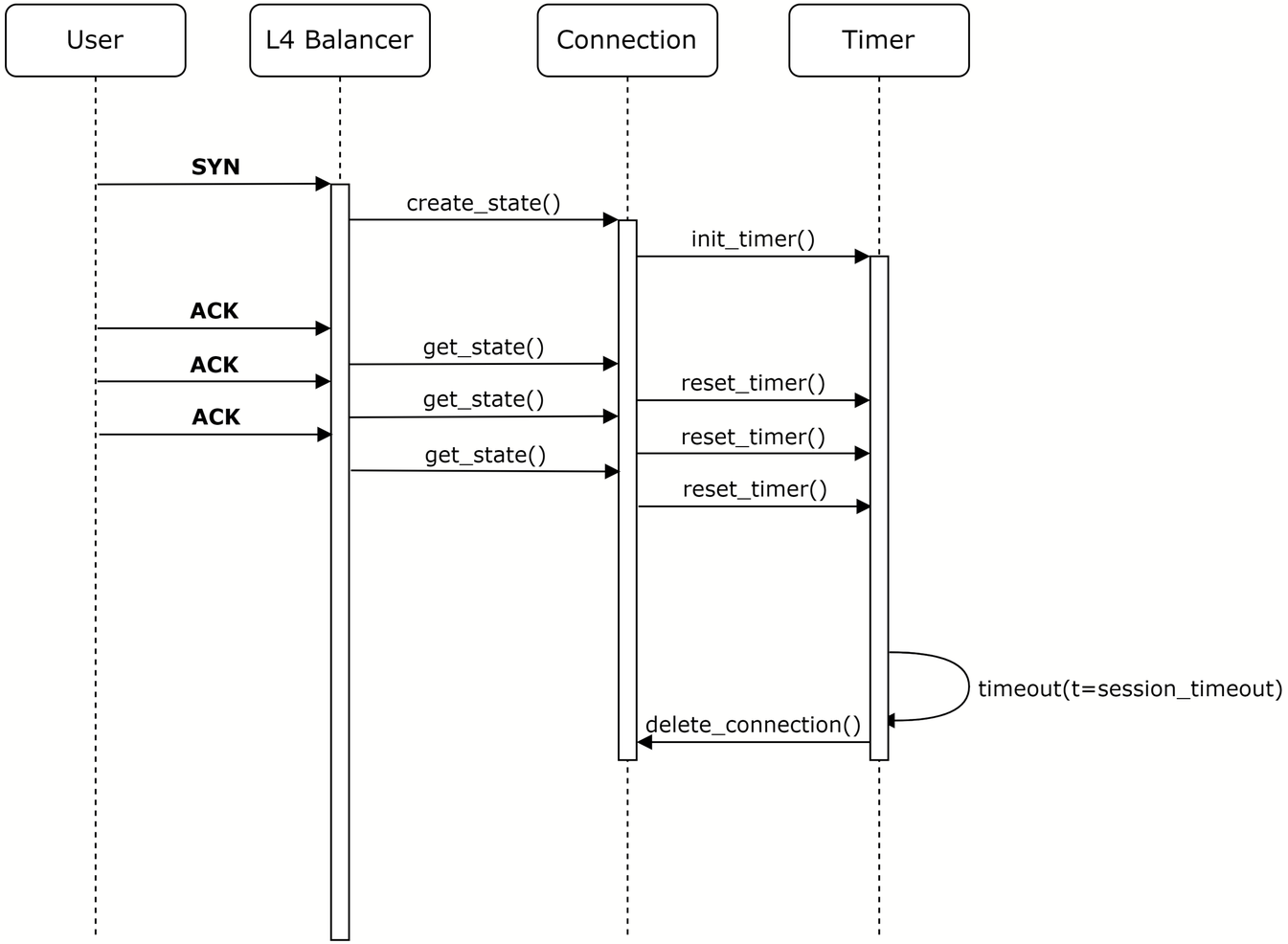


Classic Packet Flow





Connection State Machine

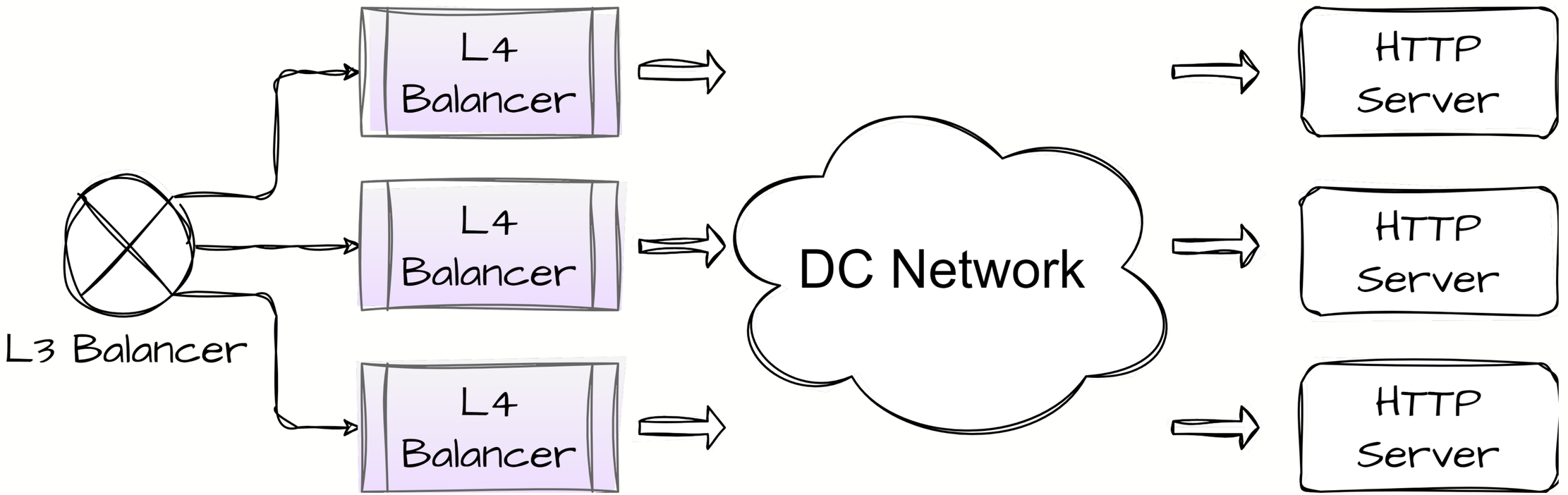


Classic L4 Balancer

1. Processes only ingress traffic
2. Relies on health checks
3. Stateful

Classic L4 Balancer

1. Processes only ingress traffic
2. Relies on health checks
3. Stateful
4. Vulnerable to DDoS

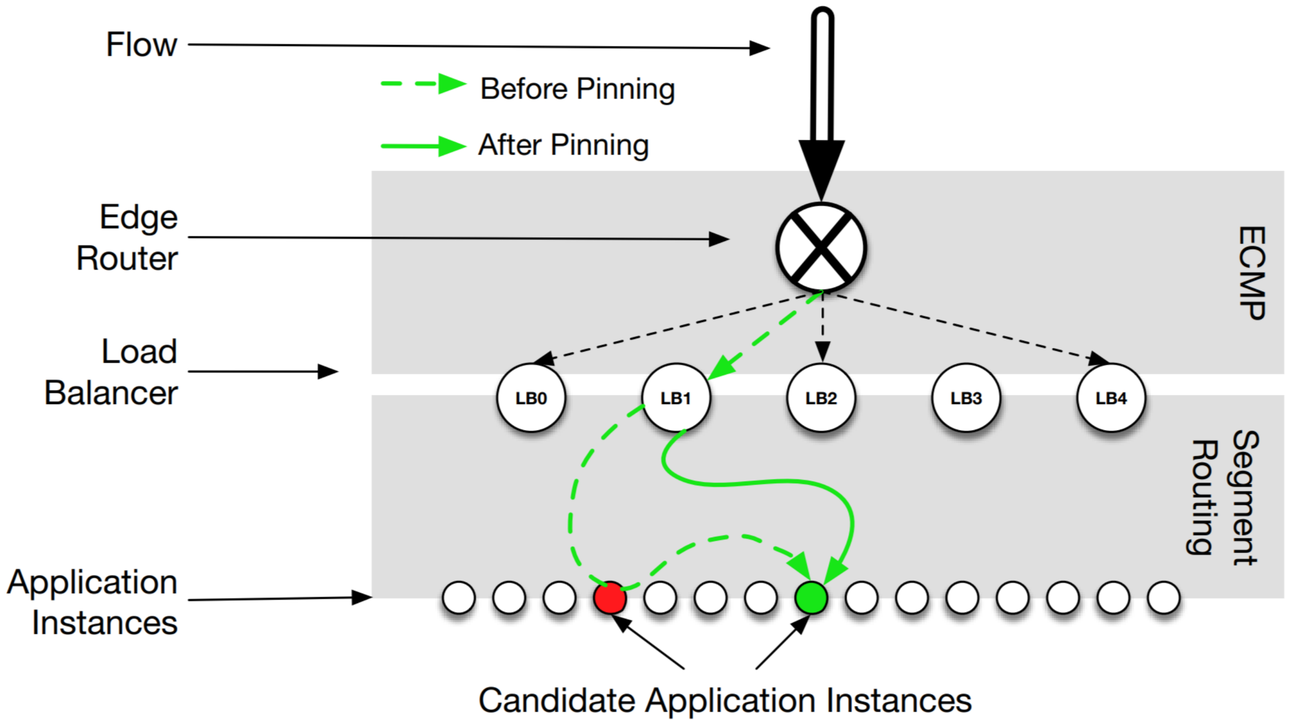


L3 Balancer – invulnerable

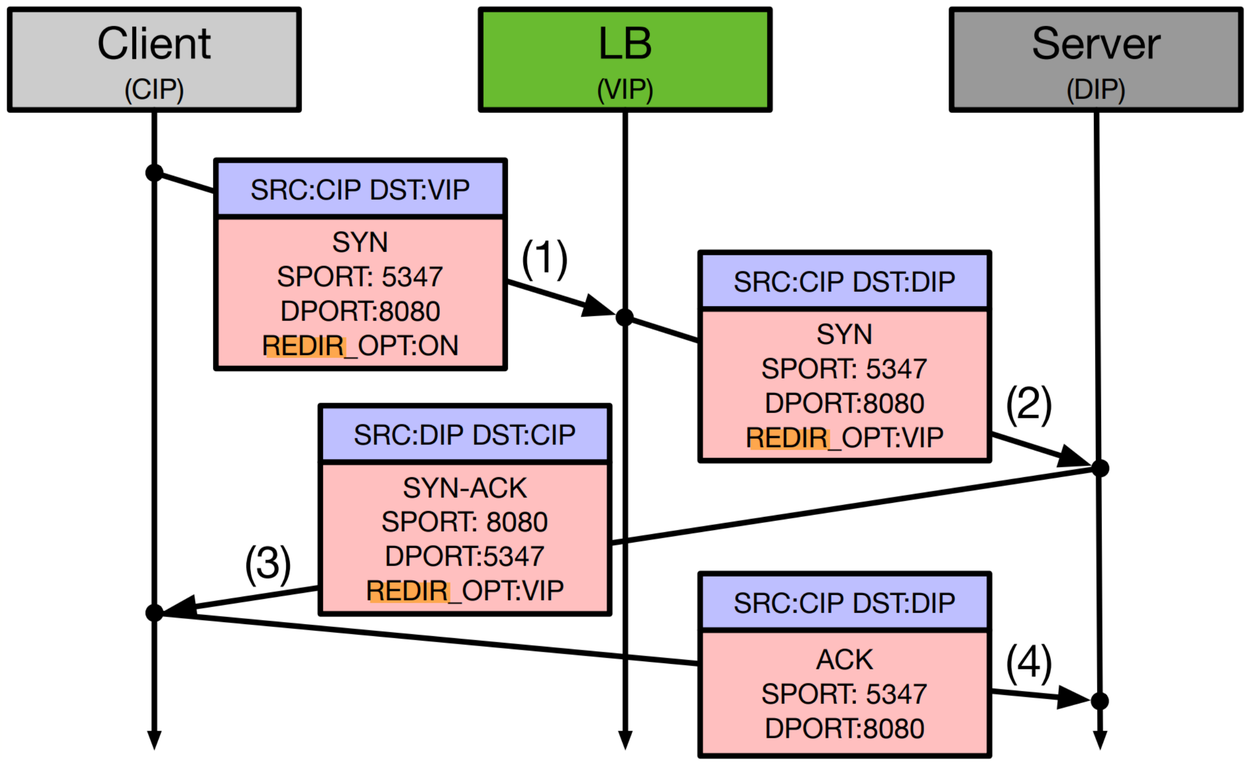
HTTP Server – SYN Cookies, TCP state machine

L4 Balancer – ???

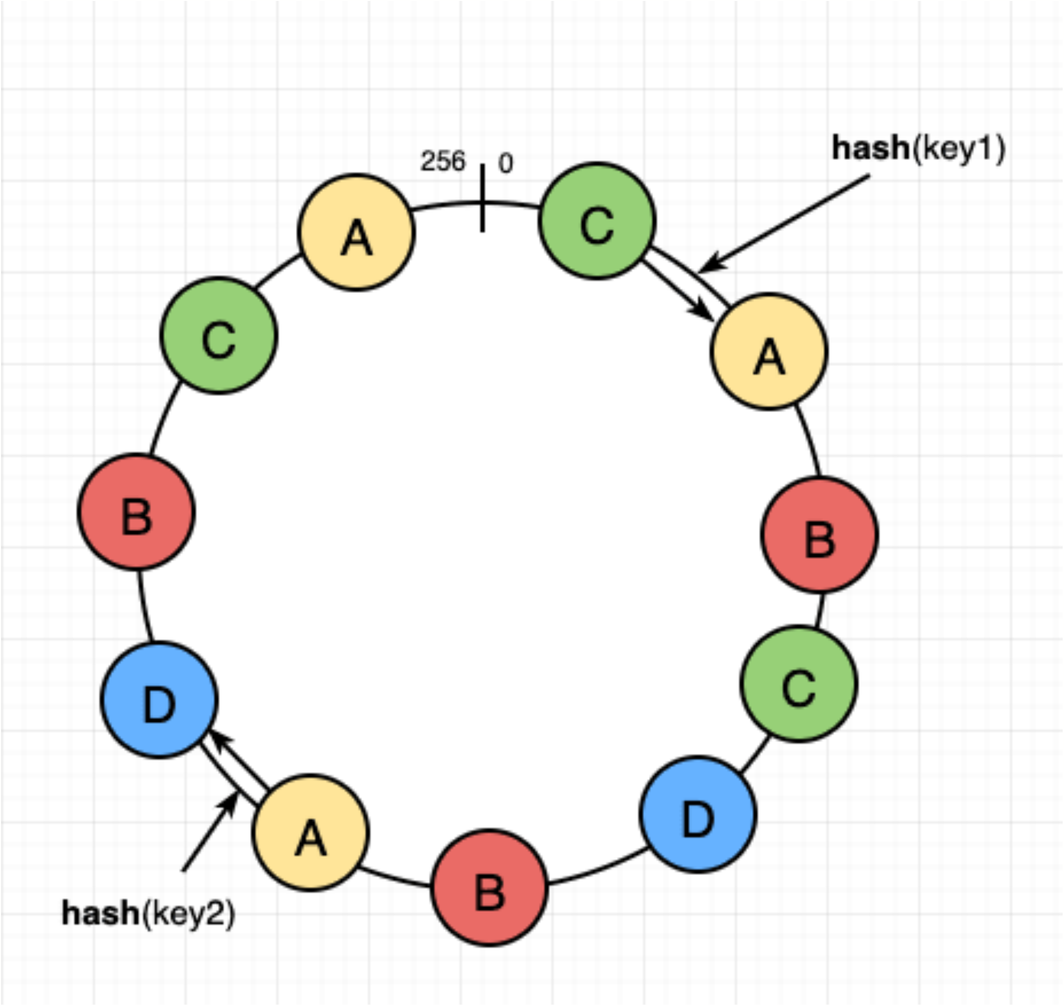
Stateless L4 Balancer: Network Offload



Stateless L4 Balancer: TCP Offload



Stateless L4 Balancer requires integration with network and/or TCP stack

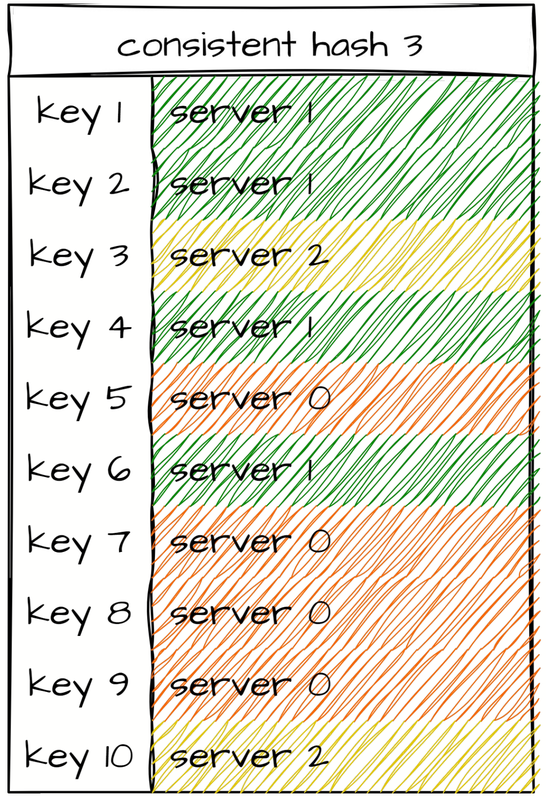


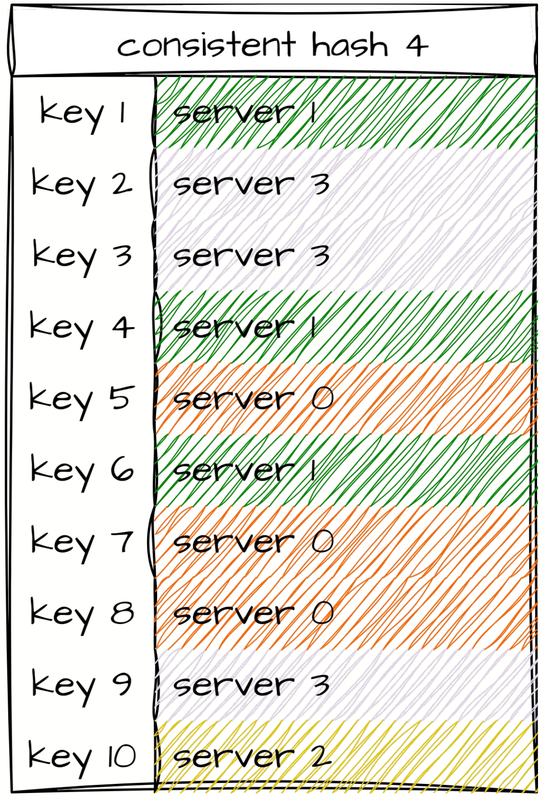
Consistent Hash

Remaps only n/m keys, where

- n – is the number of keys
- m – is the number of slots

Consistent Hash

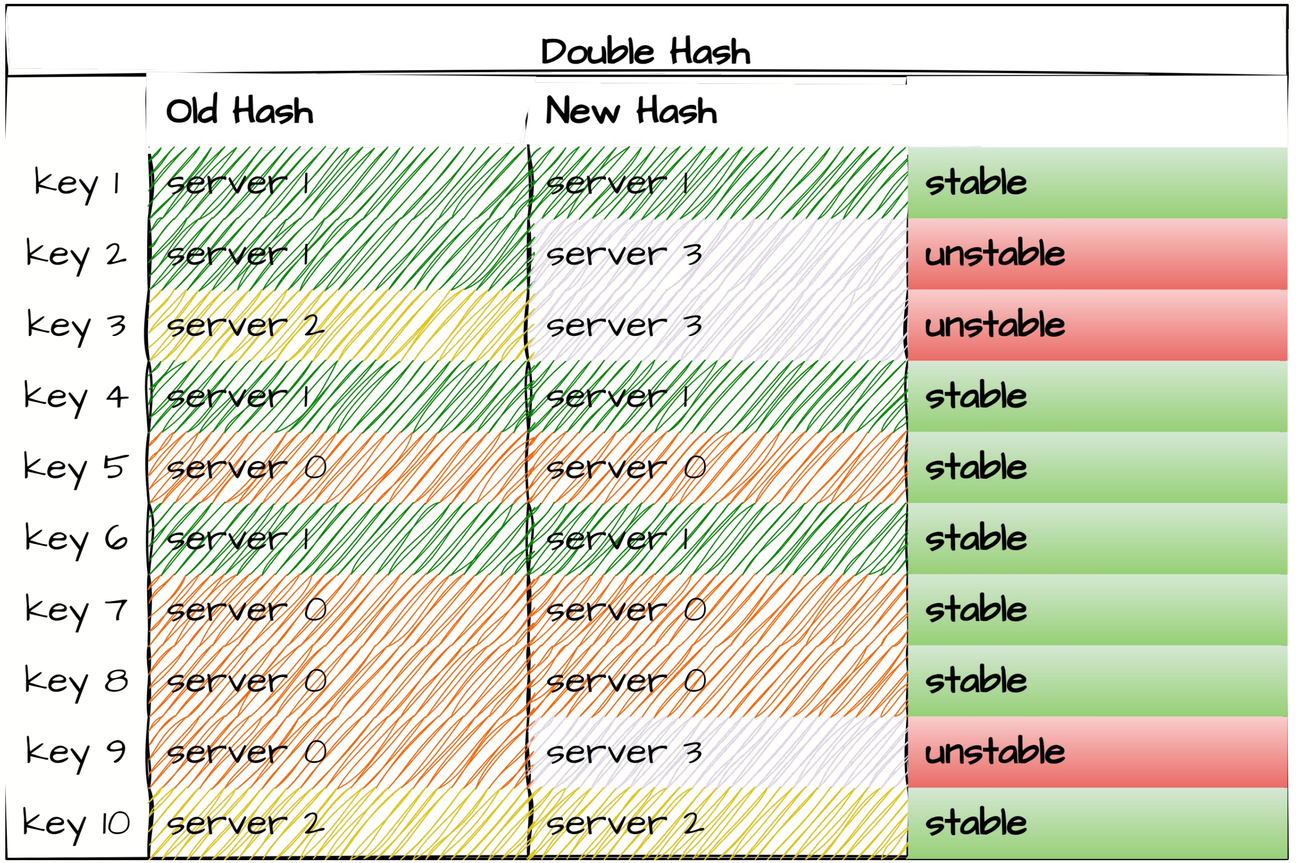


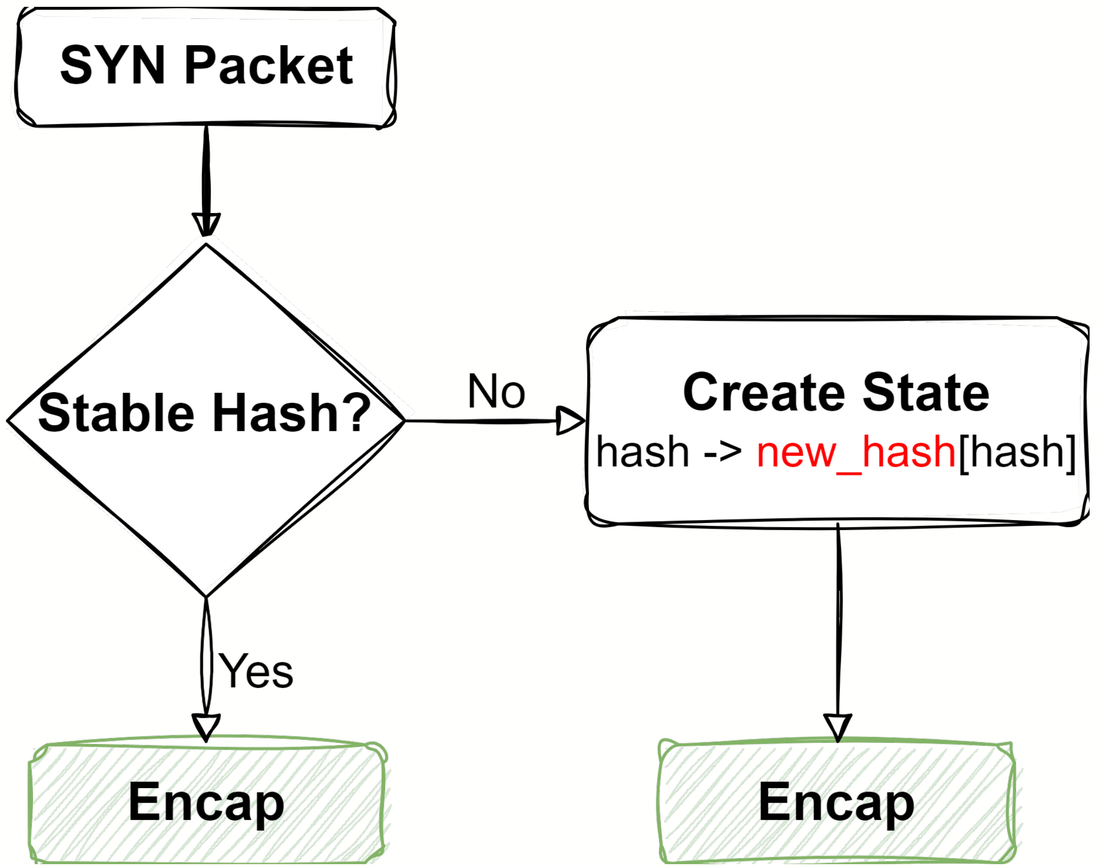


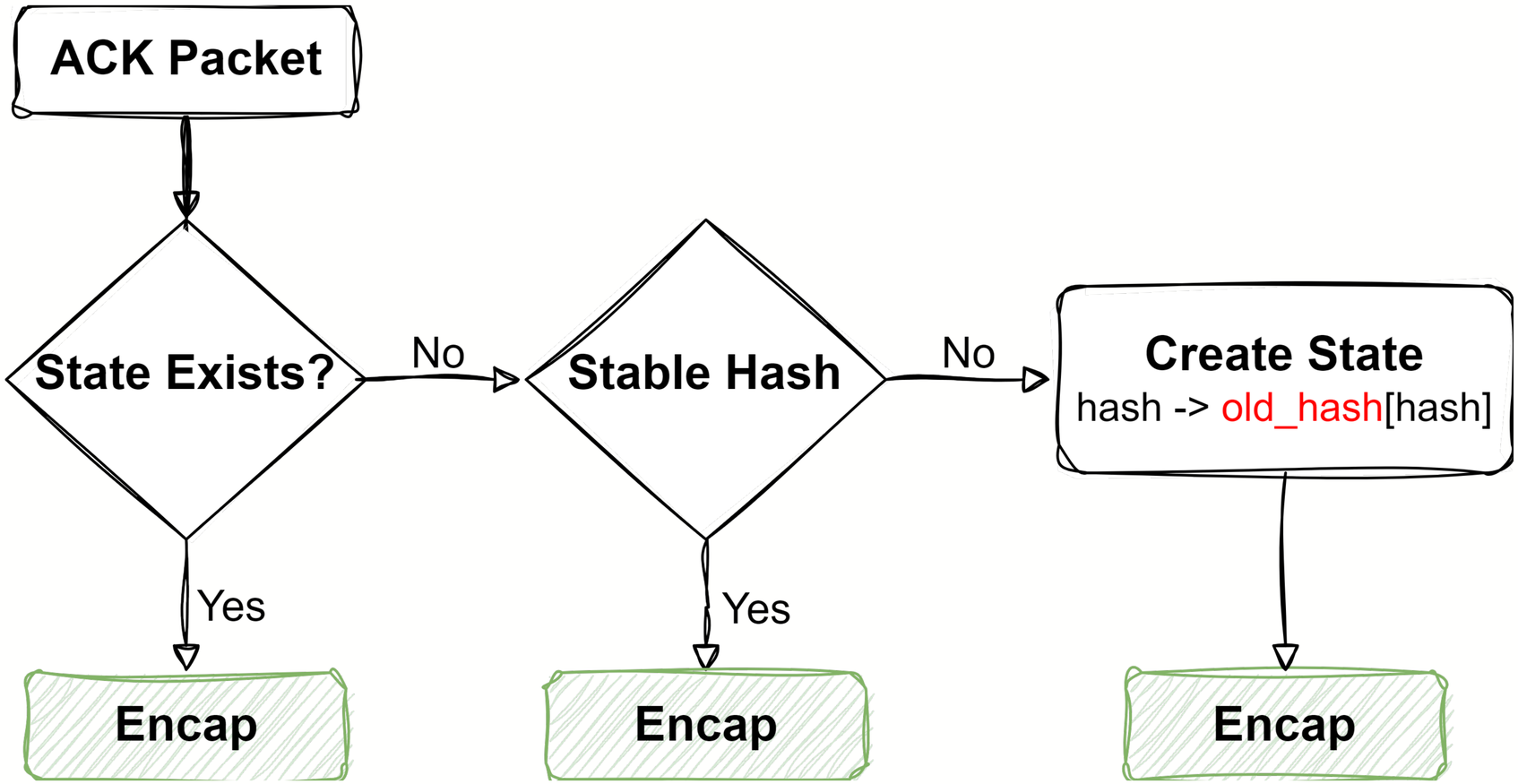
Double Hash

Double Hash		
	Old Hash	New Hash
key 1	server 1	server 1
key 2	server 1	server 3
key 3	server 2	server 3
key 4	server 1	server 1
key 5	server 0	server 0
key 6	server 1	server 1
key 7	server 0	server 0
key 8	server 0	server 0
key 9	server 0	server 3
key 10	server 2	server 2

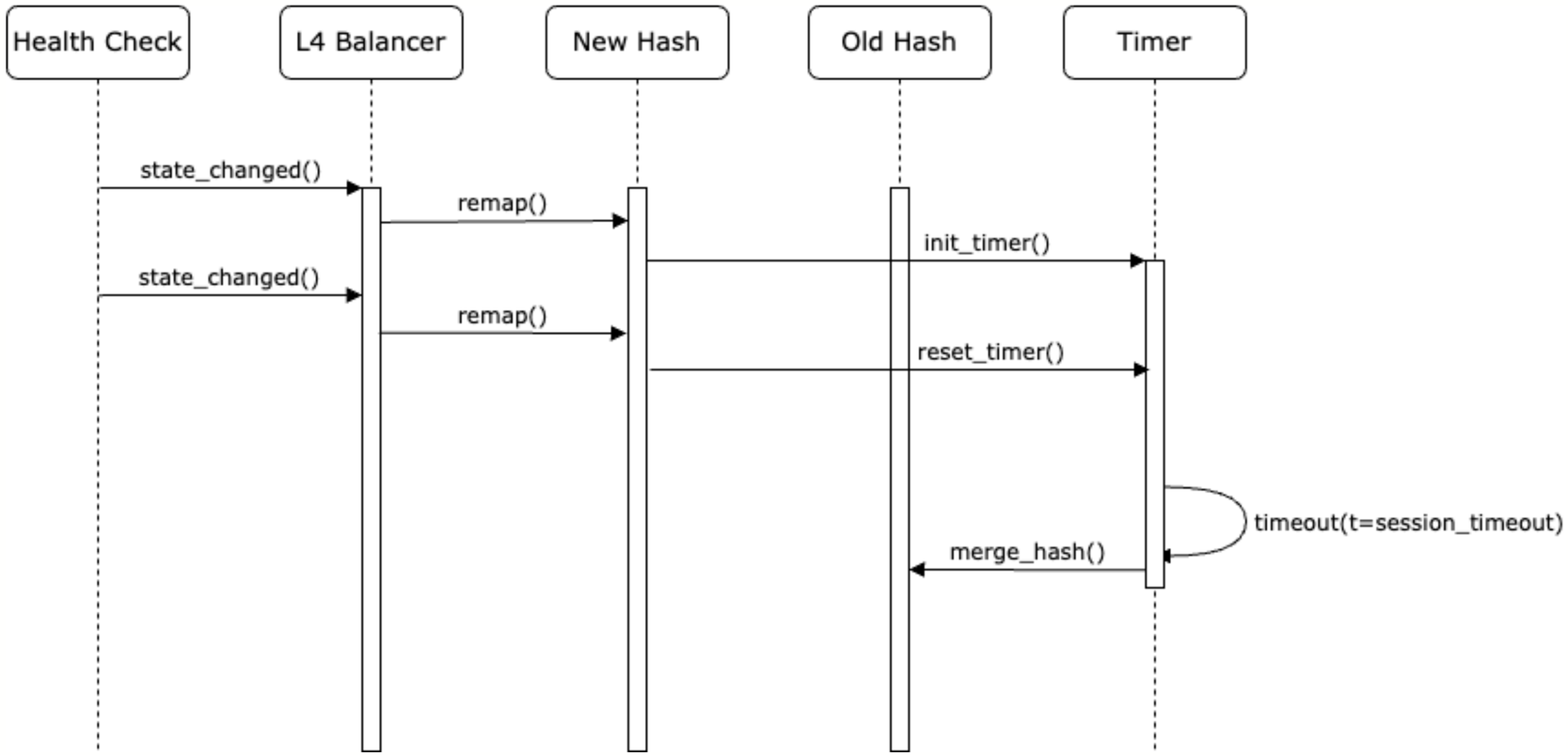
Double Hash: Stability Property




SYN Packet
Stable Hash?
No
Create State
hash -> new_hash[hash]
Yes
Encap
Encap


ACK Packet
State Exists?
No
Stable Hash
No
Create State
hash -> old_hash[hash]
Yes
Yes
Encap
Encap
Encap

Double Hash: State Machine



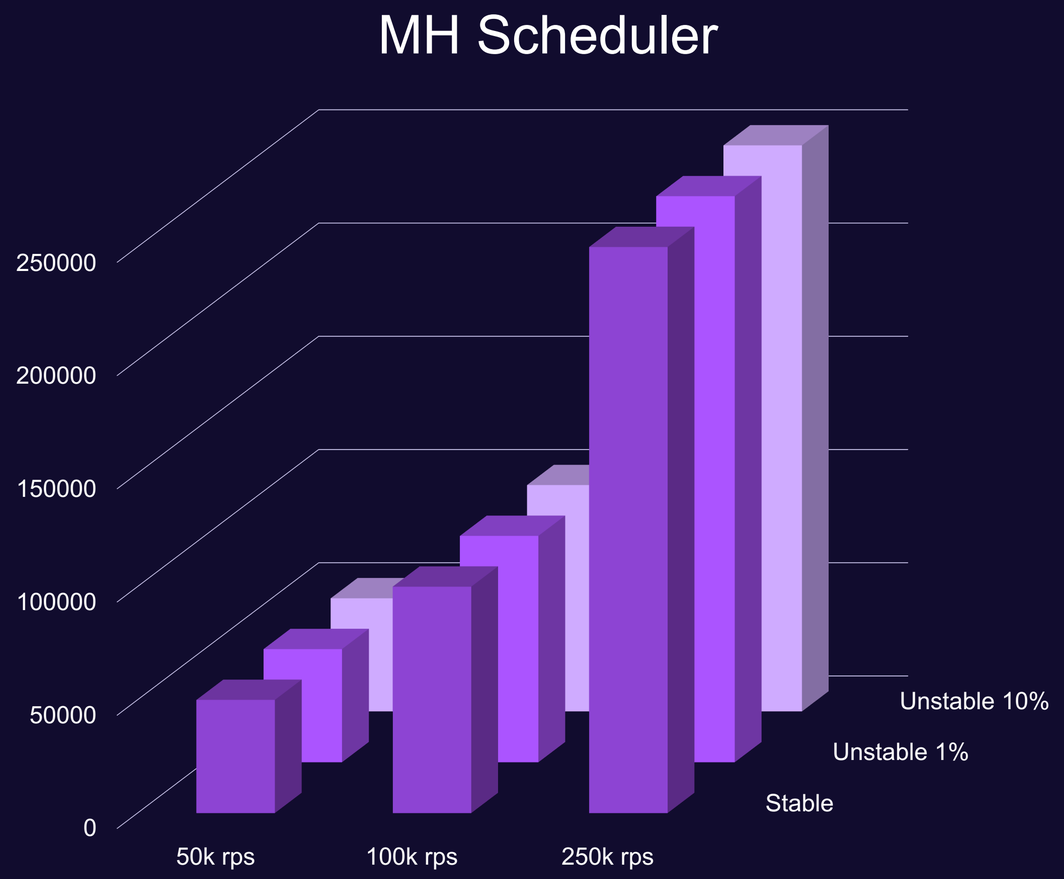
IPVS: MH Scheduler

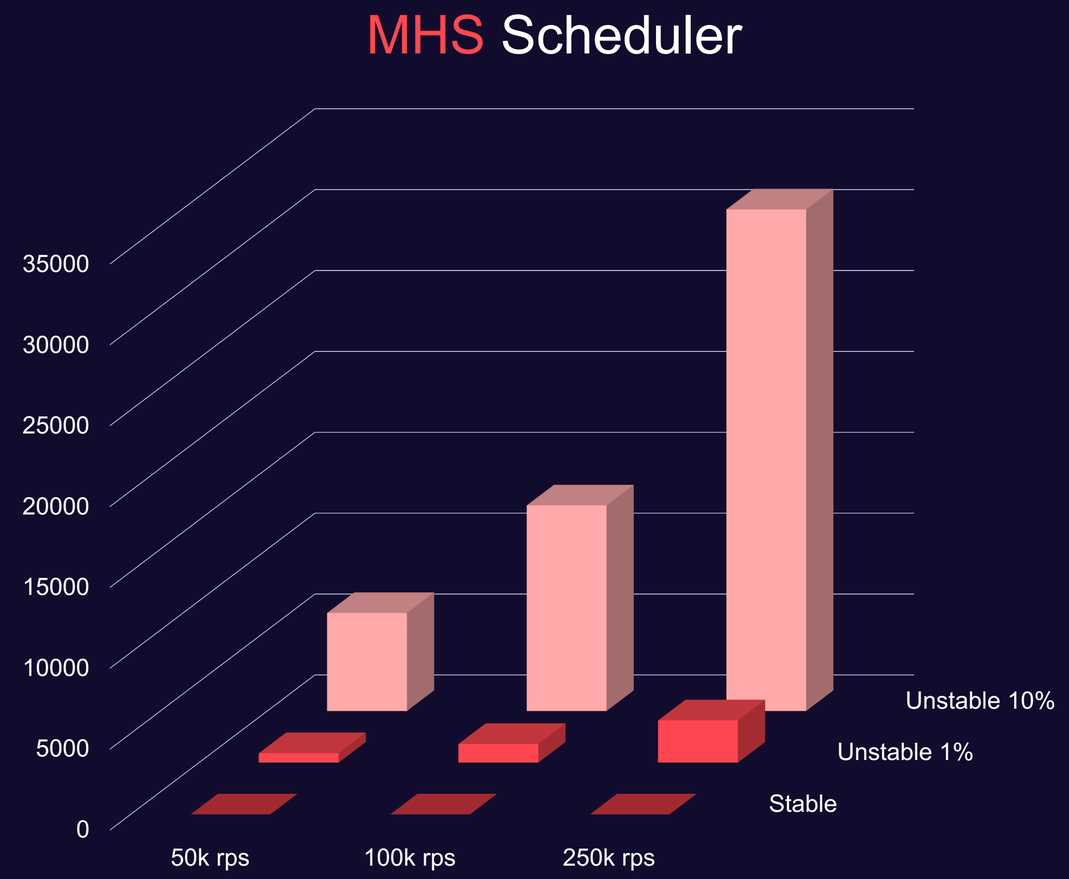
Maglev: A Fast and Reliable Software Network Load Balancer

Danielle E. Eisenbud, Cheng Yi, Carlo Contavalli, Cody Smith,
Roman Kononov, Eric Mann-Hielscher, Ardas Cilingiroglu, Bin Cheyney,
Wentao Shang[†*] and Jinnah Dylan Hosein[‡*]

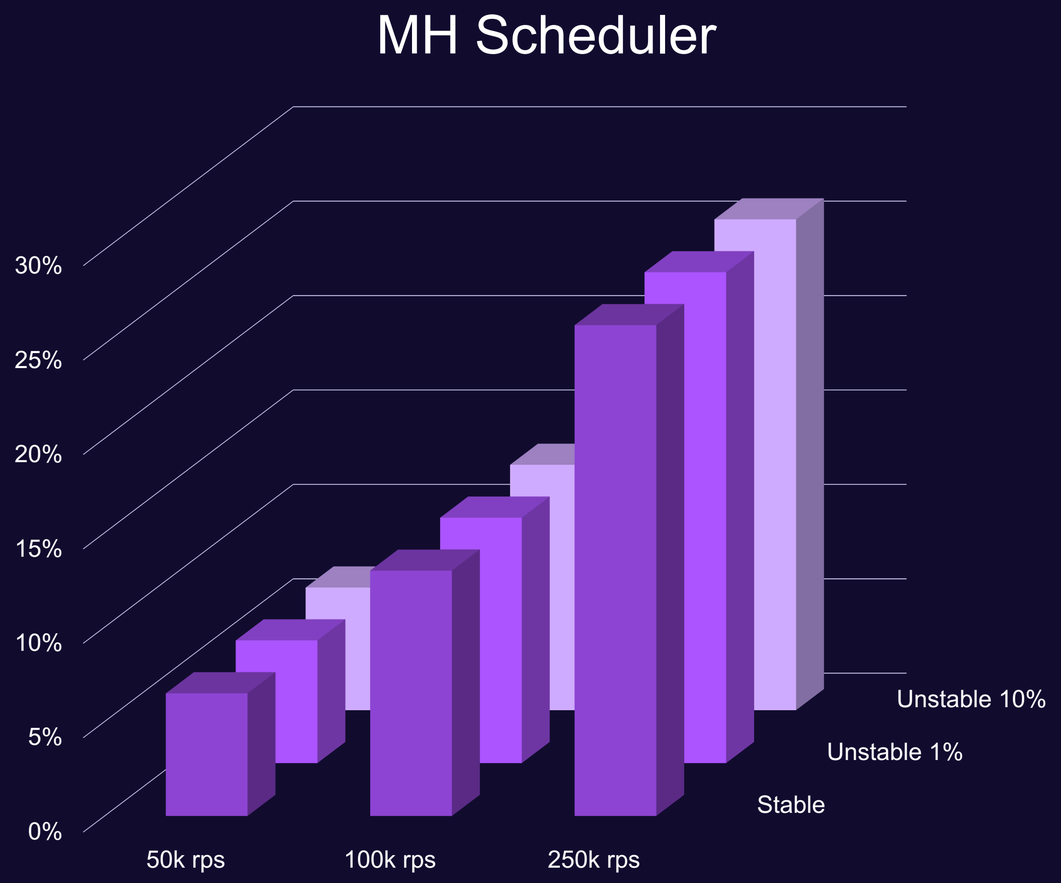
Google Inc. [†]UCLA [‡]SpaceX
maglev-nsdi@google.com

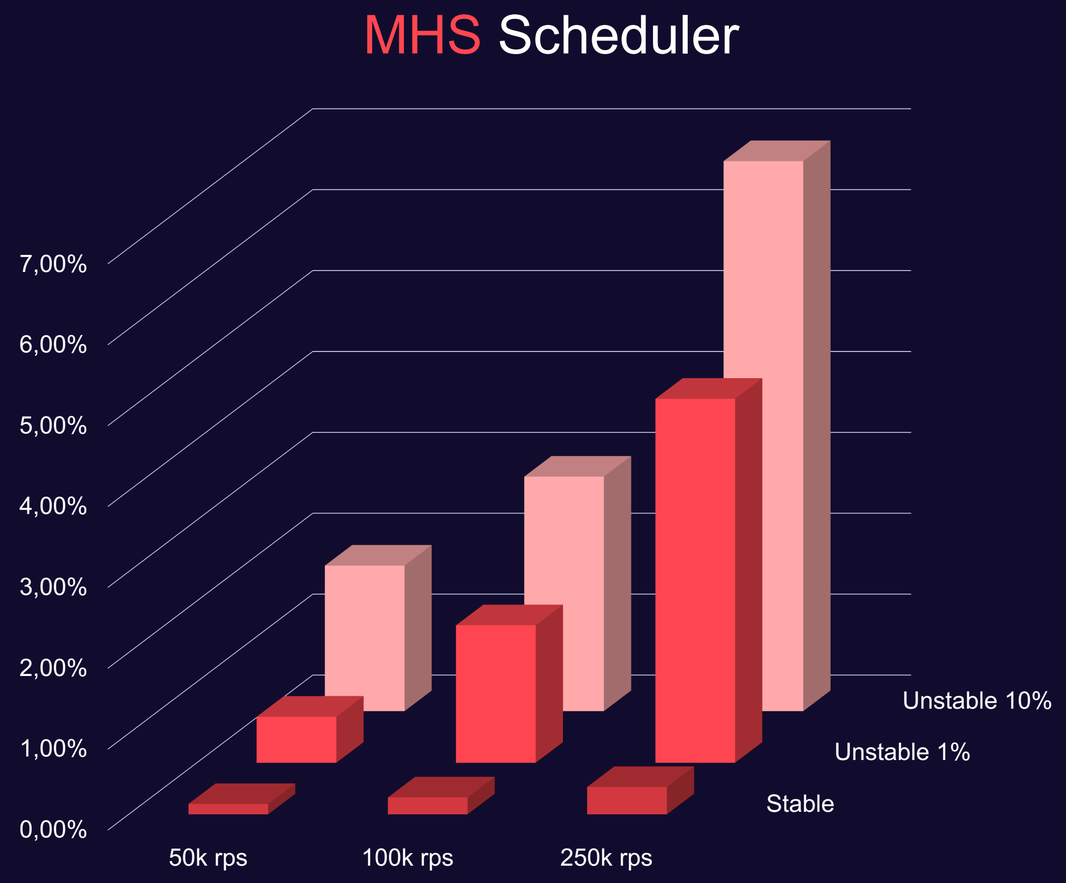
Number of States





CPU Usage


MH Scheduler
30%
25%
20%
15%
10%
5%
0%
50k rps
100k rps
250k rps
Stable
Unstable 1%
Unstable 10%


MHS Scheduler
7,00%
6,00%
5,00%
4,00%
3,00%
2,00%
1,00%
0,00%
50k rps
100k rps
250k rps
Stable
Unstable 1%
Unstable 10%

Nearly Stateless L4 Balancer

1. Reduces the number of states by order of magnitude
2. Significantly reduces the CPU usage
3. Makes L4 balancer less vulnerable to DDoS
4. Doesn’t require changes in network or application

MHS Scheduler at GitHub!

https://github.com/kndrvt/mhs