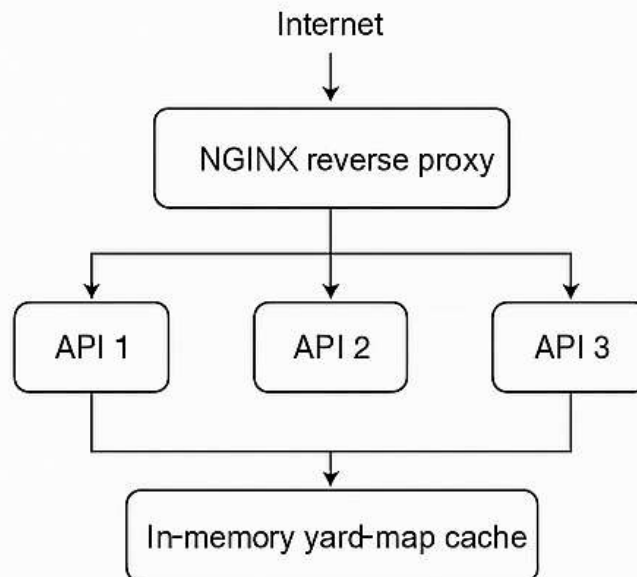


# System Design

## Design a System That Stays Fast Under Heavy Traffic

### Traffic Profile

Requests per second  $\leq 100$  rps  
Peak Hour (monsoon backlogs)  $\leq 500$  ms



### 1. How It Handles Traffic

- Three servers  $\times 400$  rps  $\times 3$  servers/SON; no thing stored between calls
- 55th more if traffic  $>$

### 2. Keeping It Fast

- Stateless — each request carries its own JSON; no thing stored call info.
- Local caching — yard-maps stored in a Hash Map; total 5TB.
- No database in hot path: avoid extra latency

### 3. Simple Monitoring

- Shell script pinging /healthz every 30 seconds; send alert if  $>300$  ms.
- NGINX access logs feed visual reports showing P95 latency, 4xx, 5xx error rates

### 4. Failure Story

- If an API crashes, NGINX stops forwarding traffic to its port, remaining APIs handle traffic.

### 5. Blue-Green (Manual)

- Run new jars on port 9000  $\rightarrow$  hit /pick6poc t locally, add the new jar