CLOUD COMPUTING

Chat server scalability

NGINX
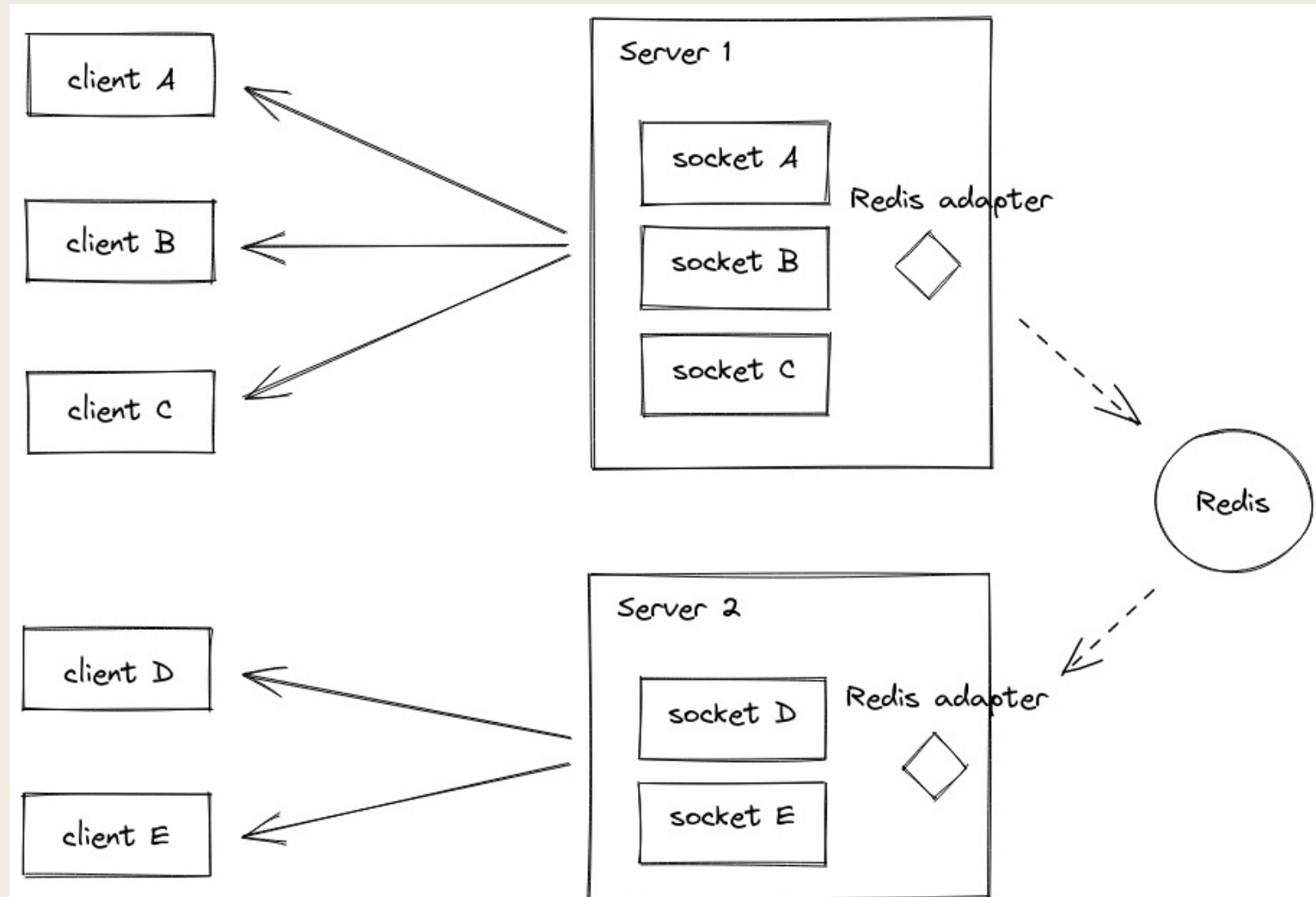
- Load Balancing

- Scaling into multiple Server nodes

- NGINX HTTP Server runs on Port 80 so inbound connection

- IP based
  - *Same IP adresses will always route to the same Server*
  - *Client with different IP will route to a different Server*

NGINX Configuration

- Nginx Configuration File

```nginx
1   worker_processes 4;
2
3   events {
4     worker_connections 1024;
5   }
6
7   http {
8     server {
9       listen 80;
10
11      location / {
12        proxy_set_header X-Forwarded-For $proxy_add_x_forwarded_for;
13        proxy_set_header Host $host;
14
15        proxy_pass http://nodes;
16
17        # enable WebSockets
18        proxy_http_version 1.1;
19        proxy_set_header Upgrade $http_upgrade;
20        proxy_set_header Connection "upgrade";
21      }
22    }
23
24    upstream nodes {
25      # enable sticky session with either "hash" (uses the complete IP address)
26      hash $remote_addr consistent;
27      # or "ip_hash" (uses the first three octets of the client IPv4 address, or the entire IPv
28      # ip_hash;
29      # or "sticky" (needs commercial subscription)
30      # sticky cookie srv_id expires=1h domain=.example.com path=/;
31
32      server server-one:3000;
33      server server-two:3000;
34      server server-three:3000;
35      server server-four:3000;
36    }
```

SOCKET.IO REDIS ADAPTER

Lesson learned

- Load Balancing in general
- NGINX technology