



WhatsApp System Design (Basic) Cheat Sheet

V2021.03.17
(Dr Yan Xu)

Overview

- One of the most Challenging Services
- Point-to-Point Chat
- Group Chat

Core Services

- User Service
 - user information CUID
- Online Service
 - user online / offline
- **Push Service**
 - fast server to user data flow
- **Message Service**
 - store and query message
- **Thread Service**
 - user & thread relationship
- Contact Service
 - friendship CUID

System Requirement

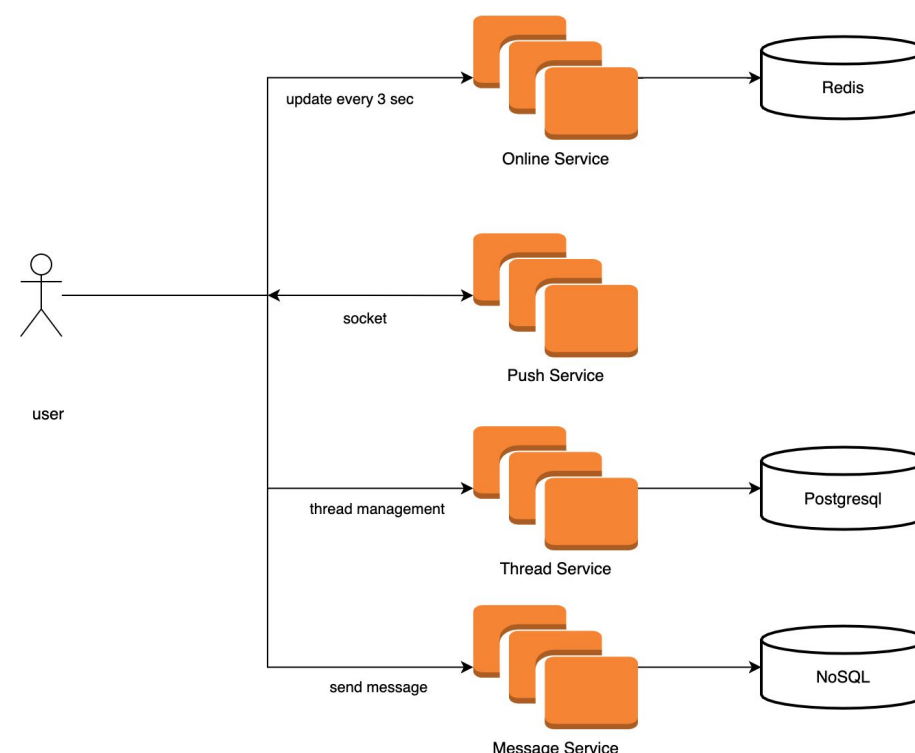
- Fast (<1 sec), Secure, Scalable and High SLA

System Non-Functional Requirement

- Users: **1B**
- Daily Message: **200B** (counting group message times)
- Peak Message: **~8M QPS [F**K]**
- Message Storage:
 - $50\text{ B} * 2\text{M} * 3600 * 24 \sim 8\text{ TB per day}$
- Groups: **500K**
- Group Size: **<500**
- Average Group Size: **10**
- Group Message Percentage: **50%**

System High-level Design

- Combination of Push & Pull
 - Push for "update notification"
 - Pull for "get the message"

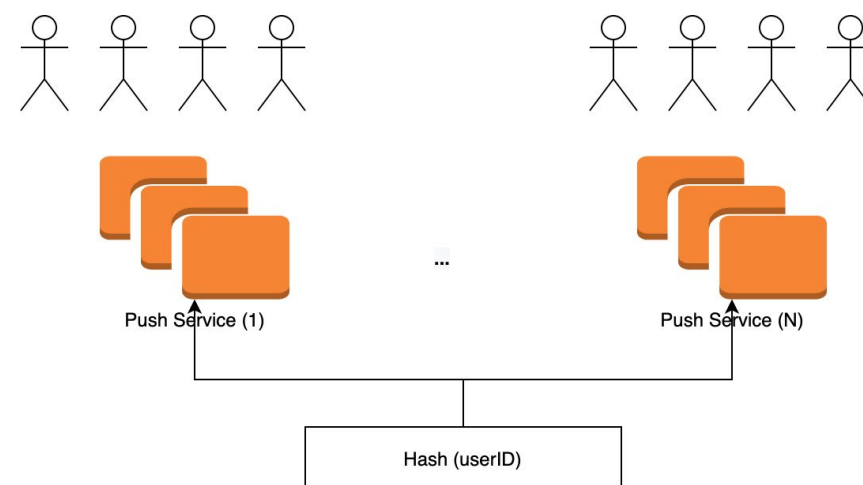


Online Service

- Pull
- Once every 3 seconds (heartbeat)
- Storage: In-Memory + Sharding (userID)

Push Service

- Socket
 - by default, a single server can handle 65,536 socket connections, because it's the max number of TCP ports available.
- Sharding (userID)



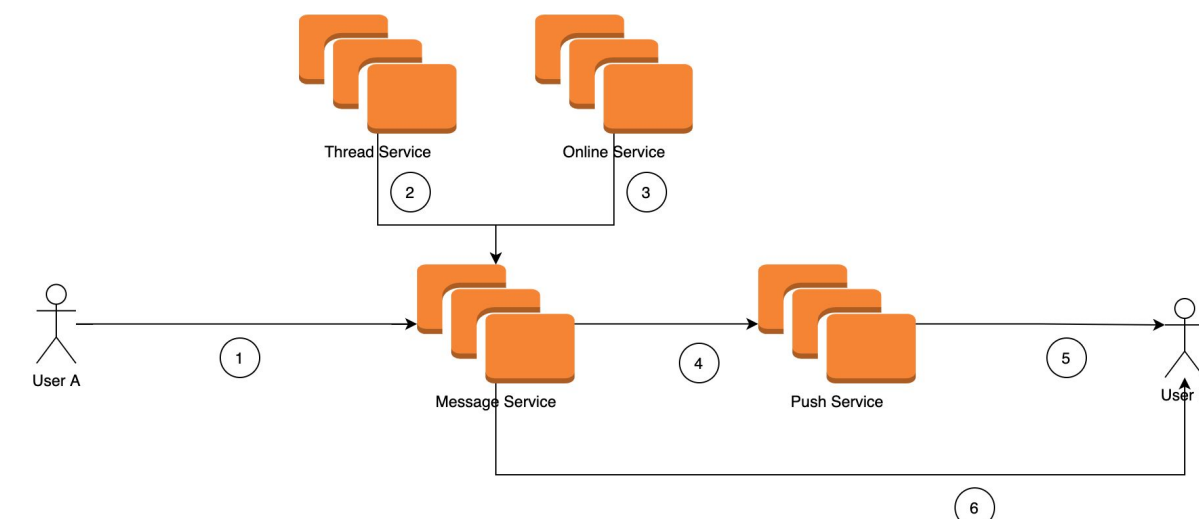
Thread Service

- REST API
- Functions: mute, last active time, etc
- Sharding (userID)

Message Service

User A -> message -> User B

- Step 1: User A calls Message Service API
- Step 2: Message Service checks End Users
- Step 3: Only send to Active Users
- Step 4: Update End User (B) to Pull
- Step 5: User B pulls latest Thread message
(the process is the same for groups)



Scalability

- Load Balancer
- NoSQL
 - key: thread-hash-code
- Shared Database
 - key: user id

Role of Machine Learning

- Voice2Text and Text2Voice
- Language Translation
- Message Prediction and Auto-Completion
- User Profile Extraction for Ad. (scary!!)
- Chat Assistant (e.g. to figure out who the other side is)

Real-world Considerations

- messages should be compressed and encrypted
- browser-version chat (no socket connections)
- images and files sending in message
- video conference
- > 500 users group and group management