

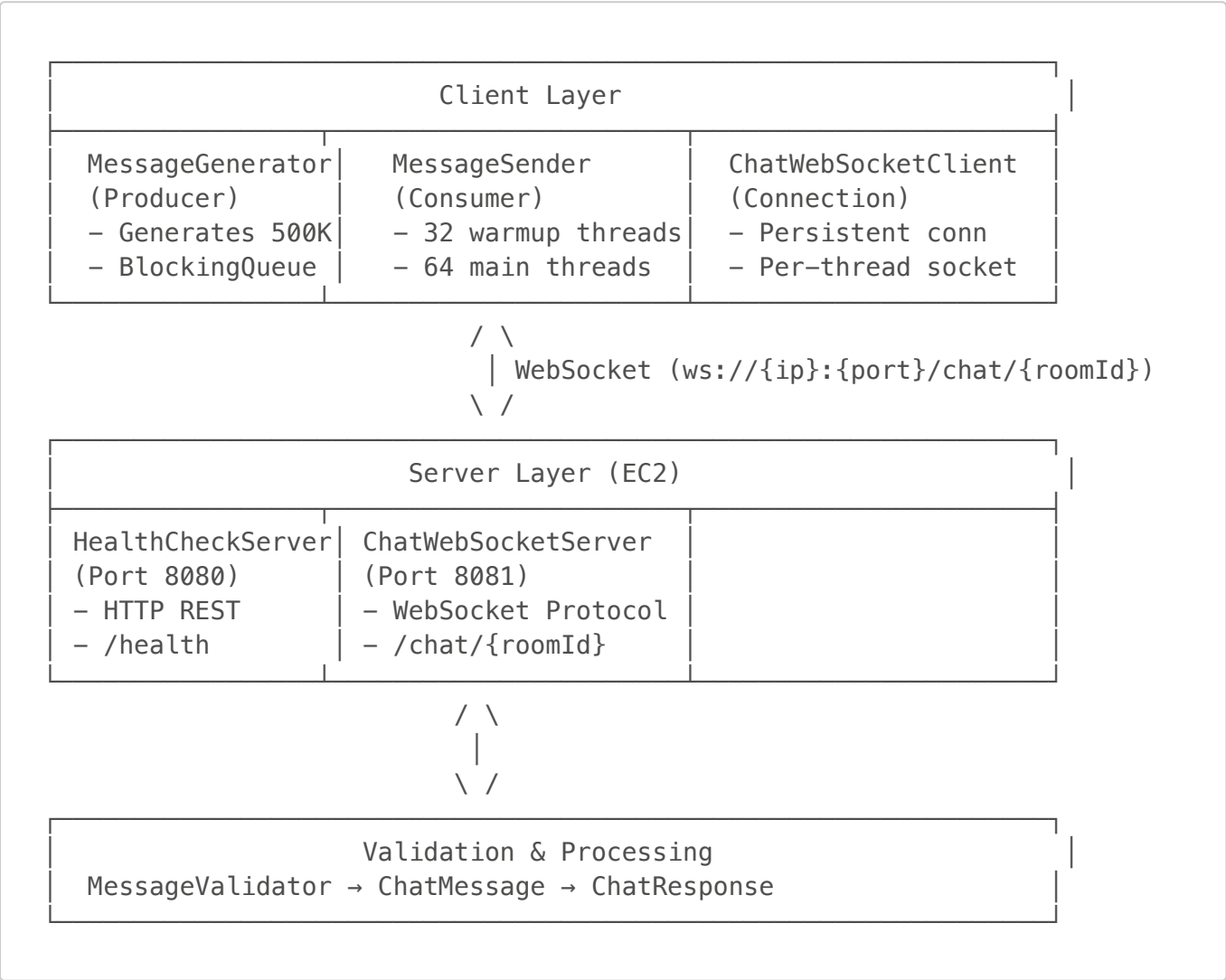
# ChatFlow WebSocket Server

## 1. github

<https://github.com/yimdx/ChatFlow>

## 2. System Architecture

Architecture Diagram



### Major Classes and Relationships

#### Server-Side Classes

##### ChatWebSocketServer (Main Server)

- Extends `WebSocketServer` from Java-WebSocket library
- Manages WebSocket lifecycle: `onOpen`, `onMessage`, `onClose`, `onError`
  - Build connection with client, receive `ChatMessage` and return `ChatResponse`
- Validates room IDs (1-20) using regex pattern `^/chat/(\\d+)$`

##### HealthCheckServer (HTTP Health Endpoint)

- Uses `HttpServer` from JDK
- Serves `/health` endpoint on port 8080
- Returns JSON: `{"status":"healthy"}`

## MessageValidator

- Validates `ChatMessage` fields (username, message length, etc.)
- Returns list of validation errors

## Models:

- `ChatMessage`: Request model (userId, username, message, timestamp, messageType, roomId)
- `ChatResponse`: Response model (includes serverTimestamp, status)
- `ErrorResponse`: Error handling model

## Client-Side Classes

### MessageGenerator (Producer Thread)

- Implements `Runnable`
- Generates 500,000 random pre-defined messages
- Populates `BlockingQueue<ChatMessage>`
- Random message types: 90% TEXT, 5% JOIN, 5% LEAVE

### MessageSender (Consumer Thread)

- Implements `Runnable`
- Creates ONE persistent `WebSocket` connection per thread
- Consumes messages from `BlockingQueue`
- Sends messages sequentially with response waiting
- Tracks success/failure counts atomically

### ChatWebSocketClient (WebSocket Connection)

- Extends `WebSocketClient` from Java-WebSocket library
- Manages connection lifecycle and message serialization
- Tracks metrics: success count, failure count, latency

## Part 2 Additional Classes:

- `MetricRecord`: Records per-message metrics (timestamp, latency, room)
- `CsvWriter`: Writes metrics to CSV for analysis
- `PerformanceAnalyzer`: Calculates statistics (mean, median, p95, p99, throughput per room)
- `ThroughputVisualizer`: Generates throughput-over-time analysis

## Threading Model

### 1. Generator Thread (1 thread)

- Runs independently
- Produces all 500K messages upfront
- Non-blocking producer to `BlockingQueue`

## 2. Warmup Phase (32 threads)

- Fixed thread pool: `Executors.newFixedThreadPool(32)`
- Each thread: 1,000 messages (32K total)
- Each thread maintains ONE persistent WebSocket connection
- Duration: ~20 seconds (with 20ms latency)

## 3. Main Phase (64 threads)

- Thread count: 64
- Each thread: ~14,625 messages (468K total)
- Duration: ~150 seconds (with 20ms latency)

### Thread Lifecycle:

```
Thread Start → Create WebSocket → Connect (blocking)
    ↓
For each message:
    → Take from queue (blocking)
    → Send message
    → Wait for response (1s timeout)
    → Increment success/fail counter
    ↓
Close WebSocket → Thread Exit
```

### Synchronization Mechanisms:

- `BlockingQueue<ChatMessage>`: Thread-safe message queue
- `AtomicInteger`: Lock-free counters (success, failure, reconnections)
- `ExecutorService`: Managed thread lifecycle
- `ConcurrentHashMap` for thread-safe room management

## 4. WebSocket Connection Management Strategy

### Client-Side Strategy: Thread-Persistent Connections

**Design Decision:** Each thread maintains ONE persistent connection for its entire lifecycle.

### Key Properties:

- **Connection Count:** 96 total (32 warmup + 64 main, sequential phases)
- **Connection Reuse:** No pooling or sharing between threads
- **Room Distribution:** Uniform random (each thread picks room 1-20)
- **Failure Handling:** Exponential backoff retry (up to 5 attempts)

## 5. Little's Law Calculations and Predictions

### Little's Law Formula

$$L = \lambda \times W$$

Where:

$\lambda$  (lambda) = Throughput (requests/second)

L = Concurrency (number of simultaneous requests)

W = Average response time (seconds)

## System Parameters

### Measured Values from Warmup:

- Total messages: 32k
- Total runtime: ~34 seconds (warmup)
- average latency 34ms
- Threads: 32 concurrent

## Performance Predictions

### Ideal Throughput

```
# for warmup phase
λ = 32 concurrent threads/ 34 ms
λ ≈ 1k messages/second

# for main phase

λ = 64 concurrent threads/ 34 ms
λ ≈ 1.9k messages/second
```

## Test Results

### Part1

```
ection
13:33:15.837 [main] INFO cs6650.assignment1.Main - Main Phase completed in 263507 ms
13:33:15.837 [main] INFO cs6650.assignment1.Main -
13:33:15.837 [main] INFO cs6650.assignment1.Main - =====
13:33:15.837 [main] INFO cs6650.assignment1.Main - PERFORMANCE RESULTS
13:33:15.837 [main] INFO cs6650.assignment1.Main - =====
13:33:15.837 [main] INFO cs6650.assignment1.Main - 1. Successful messages sent: 500001
13:33:15.837 [main] INFO cs6650.assignment1.Main - 2. Failed messages: 6
13:33:15.837 [main] INFO cs6650.assignment1.Main - 3. Total runtime: 307915 ms (307.915 seconds)
13:33:15.837 [main] INFO cs6650.assignment1.Main - - Warmup phase: 34399 ms
13:33:15.837 [main] INFO cs6650.assignment1.Main - - Main phase: 263507 ms
13:33:15.837 [main] INFO cs6650.assignment1.Main - 4. Overall throughput: 1623.8247568322427 messages/second
13:33:15.837 [main] INFO cs6650.assignment1.Main - - Warmup throughput: 930.259600569784 messages/second
13:33:15.838 [main] INFO cs6650.assignment1.Main - - Main phase throughput: 1776.043900162045 messages/second
13:33:15.838 [main] INFO cs6650.assignment1.Main - 5. Connection statistics:
13:33:15.838 [main] INFO cs6650.assignment1.Main - - Total persistent connections: 96
13:33:15.838 [main] INFO cs6650.assignment1.Main - - Reconnections: 0
13:33:15.838 [main] INFO cs6650.assignment1.Main - =====
```

### Part2

```
14:05:15.308 [main] INFO cs6650.assignment1.Main - =====
14:05:15.308 [main] INFO cs6650.assignment1.Main - BASIC PERFORMANCE RESULTS
14:05:15.308 [main] INFO cs6650.assignment1.Main - =====
14:05:15.308 [main] INFO cs6650.assignment1.Main - 1. Successful messages sent: 500000
14:05:15.308 [main] INFO cs6650.assignment1.Main - 2. Failed messages: 0
14:05:15.308 [main] INFO cs6650.assignment1.Main - 3. Total runtime: 303600 ms (303.6 seconds)
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Warmup phase: 34511 ms
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Main phase: 264991 ms
14:05:15.308 [main] INFO cs6650.assignment1.Main - 4. Overall throughput: 1646.9038208168643 messages/sec
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Warmup throughput: 927.2405899568254 messages/sec
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Main phase throughput: 1766.0977165262216 message
14:05:15.308 [main] INFO cs6650.assignment1.Main - 5. Connection statistics:
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Total persistent connections: 96
14:05:15.308 [main] INFO cs6650.assignment1.Main - - Reconnections: 0
14:05:15.308 [main] INFO cs6650.assignment1.Main - =====
14:05:15.308 [main] INFO cs6650.assignment1.Main -
Performing statistical analysis...
14:05:15.311 [main] INFO c.a.util.PerformanceAnalyzer - Analyzing metrics from: results/metrics_20260213_
14:05:15.572 [main] INFO c.a.util.PerformanceAnalyzer - Analysis completed
```

=====

## STATISTICAL ANALYSIS

=====

Total Messages: 500000  
Mean Response Time: 34.02 ms  
Median Response Time: 32.00 ms  
95th Percentile: 50.00 ms  
99th Percentile: 68.00 ms  
Min Response Time: 1 ms  
Max Response Time: 281 ms

Message Type Distribution:  
LEAVE: 24923 (5.0%)  
JOIN: 24857 (5.0%)  
TEXT: 450220 (90.0%)

Message Count Per Room:

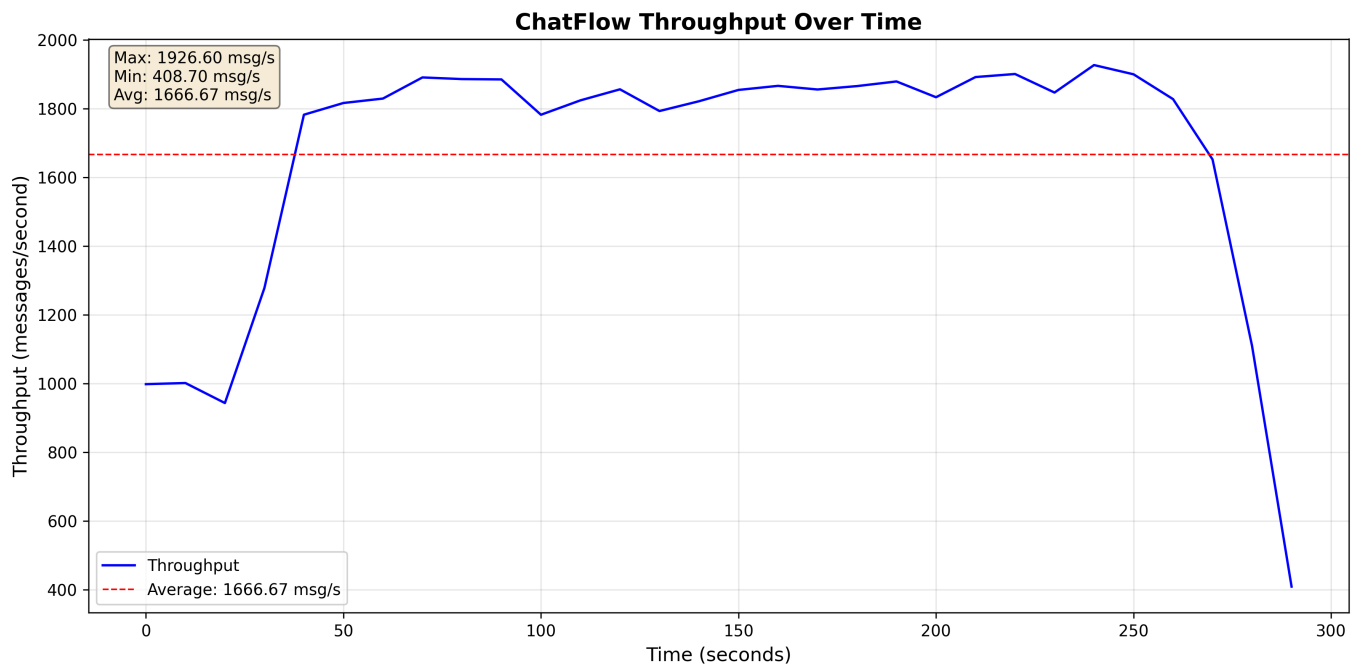
- Room 1: 40560 messages
- Room 2: 36560 messages
- Room 3: 7312 messages
- Room 4: 9312 messages
- Room 5: 17624 messages
- Room 6: 30248 messages
- Room 7: 7312 messages
- Room 8: 8312 messages
- Room 9: 7312 messages
- Room 10: 22936 messages
- Room 11: 38560 messages
- Room 12: 28936 messages
- Room 13: 3000 messages
- Room 14: 22936 messages
- Room 15: 73152 messages
- Room 16: 22936 messages
- Room 17: 66808 messages
- Room 18: 23936 messages
- Room 19: 22936 messages

Room 20: 9312 messages

#### Throughput Per Room:

Room 1: 136.28 messages/second  
Room 2: 138.56 messages/second  
Room 3: 29.09 messages/second  
Room 4: 33.87 messages/second  
Room 5: 59.95 messages/second  
Room 6: 104.90 messages/second  
Room 7: 29.36 messages/second  
Room 8: 30.62 messages/second  
Room 9: 30.43 messages/second  
Room 10: 79.19 messages/second  
Room 11: 128.97 messages/second  
Room 12: 100.21 messages/second  
Room 13: 87.79 messages/second  
Room 14: 76.65 messages/second  
Room 15: 278.75 messages/second  
Room 16: 77.35 messages/second  
Room 17: 226.77 messages/second  
Room 18: 80.30 messages/second  
Room 19: 77.59 messages/second  
Room 20: 32.94 messages/second

=====



EC2

