API Performance Testing

Tool used: Apache JMeter

The test plan simulated a typical user workflow consisting of two critical operations:

1. **User Authentication** (HTTP Request - Login)
   - Endpoint: `/api/v1/auth/login`
   - Method: POST
   - Request body: Username and password credentials

2. **Order Creation** (HTTP Request - Make Order)
   - Endpoint: `/api/v1/orders`
   - Method: POST
   - Request body: Order details including prouduct_id, quantitie, and pricing

Authorization Handling

A critical challenge encountered was implementing proper authorization for authenticated endpoints. This was resolved by:

- Adding an **HTTP Header Manager** component
- Configuring headers:
  - `Content-Type: application/json`
  - `Authorization: Bearer ${jwt_token}`
- This ensured that authenticated requests included the necessary JWT token from the login response

Test Environment

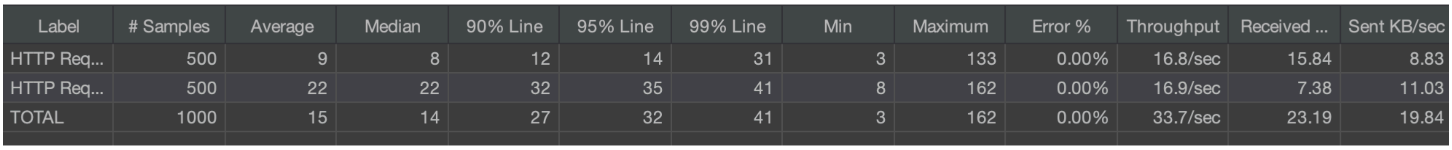
- **Server**: localhost
- **Port**: 8080
- **Protocol**: HTTP

Load Test

The performance test was executed with the following parameters to simulate realistic user load:

Thread Group Configuration:

- **Number of Threads (Virtual Users)**: 50
- **Ramp-up Period**: 30 seconds
- **Loop Count**: 10 iterations per user
- **Total Requests Executed**: 1,000 (500 login + 500 order requests)



Label	# Samples	Average	Median	90% Line	95% Line	99% Line	Min	Maximum	Error %	Throughput	Received ...	Sent KB/sec
HTTP Req...	500	9	8	12	14	31	3	133	0.00%	16.8/sec	15.84	8.83
HTTP Req...	500	22	22	32	35	41	8	162	0.00%	16.9/sec	7.38	11.03
TOTAL	1000	15	14	27	32	41	3	162	0.00%	33.7/sec	23.19	19.84

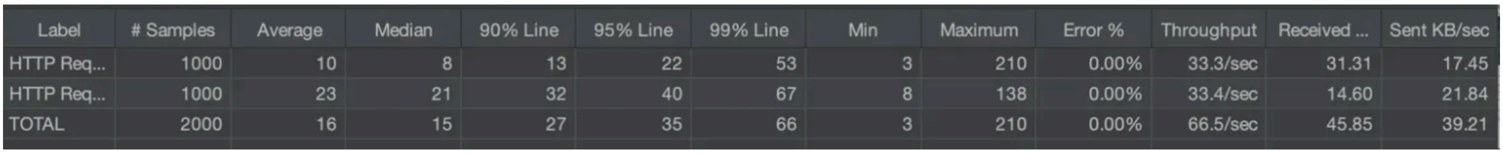
Aggregate Performance Metrics - Load Test

Stress Test

To determine system capacity under heavy load, a stress test was conducted with increased concurrent users:

Thread Group Configuration:

- **Number of Threads (Virtual Users)**: 100
- **Ramp-up Period**: 30 seconds
- **Loop Count**: 10 iterations per user
- **Total Requests Executed**: 2,000 (1,000 login + 1,000 order requests)



Label	# Samples	Average	Median	90% Line	95% Line	99% Line	Min	Maximum	Error %	Throughput	Received ...	Sent KB/sec
HTTP Req...	1000	10	8	13	22	53	3	210	0.00%	33.3/sec	31.31	17.45
HTTP Req...	1000	23	21	32	40	67	8	138	0.00%	33.4/sec	14.60	21.84
TOTAL	2000	16	15	27	35	66	3	210	0.00%	66.5/sec	45.85	39.21

Aggregate Performance Metrics - Stress Test

Analaysis

Performance testing was conducted on the API system using Apache JMeter with two load scenarios: 50 concurrent users and 100 concurrent users, both with a 30-second ramp-up period. Both tests achieved 0% error rates, demonstrating excellent system reliability under load. The results reveal that the system scales effectively with minimal performance degradation as concurrent user load doubles.

At 50 concurrent users, the system performed well with an average response time of 15ms and a median of 14ms, maintaining a throughput of 33.7 requests per second.

The two HTTP requests showed different performance characteristics, with the first request averaging 9ms and the second 22ms, indicating varying complexity or resource requirements between the endpoints.

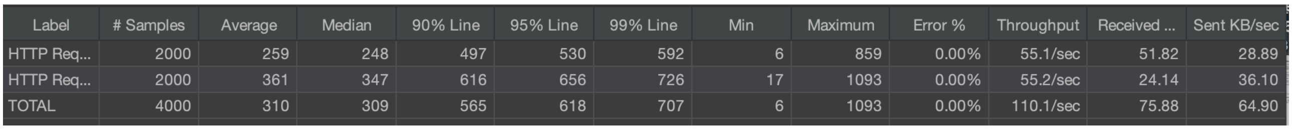
When load increased to **100 concurrent users**, the system maintained strong performance with an average response time of 16ms and median of 15ms—nearly identical to the 50-user baseline. Throughput doubled appropriately to 66.5 requests per second

The test results demonstrate that the system exhibits excellent scalability characteristics, with nearly linear performance scaling from 50 to 100 concurrent use

Spike test

- **Number of Threads (Virtual Users)**: 200

- **Ramp-up Period**: 30 seconds
- **Loop Count**: 10 iterations per user
- **Total Requests Executed**: 4000 (2000 login + 2000 order requests)



Label	# Samples	Average	Median	90% Line	95% Line	99% Line	Min	Maximum	Error %	Throughput	Received ...	Sent KB/sec
HTTP Req...	2000	259	248	497	530	592	6	859	0.00%	55.1/sec	51.82	28.89
HTTP Req...	2000	361	347	616	656	726	17	1093	0.00%	55.2/sec	24.14	36.10
TOTAL	4000	310	309	565	618	707	6	1093	0.00%	110.1/sec	75.88	64.90

Aggregate Performance Metrics - Spike Test

At 200 concurrent users, the system shows clear signs of stress with average response time increasing dramatically to 310ms and median to 309ms—representing a 19x increase from the 100-user baseline.

The test results clearly demonstrate that the system's performance sweet spot lies between 100 and 200 concurrent users. While the system scales linearly and efficiently from 50 to 100 users with minimal degradation, doubling the load again to 200 users causes response times to increase by approximately 20x, indicating a critical bottleneck has been reached.