Instagram

# Performance Testing Strategy

## Introduction

The purpose of this document is to specifically describe how the performance requirements for Instagram application will be tested and verified in test environment (System Under Test, SUT). The document will outline the scenarios, tests, parameters and data used in evaluating the capacity of the included features. The scope of tests described in current document is to verify how SUT behaves under load stress. All deliverables should be used for internal (in-house) usage only and should not be presented to public.

This document describes the strategy of performance testing for the Instagram project. It consists of outlines for the following items:

* Scope of testing, test descriptions
* Non-functional requirements (NFR) related to performance
* Pass/fail criteria
* Approach
* Main scenarios
* Test data
* Requirements for test environment

## Items to be tested

Mainly, Instagram application will be tested via UI with server side.

|  |  |
| --- | --- |
| **#** | **Modules name** |
|  | Home Page (Feed) |
|  | Search and Recommended |
|  | Activity |
|  | Account Page |
|  | Created Posts |
|  | Created Stories |
|  | Posts and Story creation |
|  | Direct messages |

## User flow

That diagram describes general flow for logged in user.

Diagram

Description automatically generated

User flow for anonymous user will not be presented, because almost all functionality is available only for logged in users.

## Approach

### Test types assumed for conducting

The following activities

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **Test type** | **Demand** | **Period** |
| 1 | Smoke test | Mandatory | Regular |
| 2 | Capacity | Mandatory | After significant changes |
| 3 | Load test | Mandatory | Regular |
| 4 | Stress | Optional | Rare |
| 5 | Scalability | Optional | Rare |
| 6 | Volume | Optional | Rare |
| 7 | Durable | Optional | Rare |

1. Smoke testing

Should be performed every time when functionality of the application and the script need to be checked. Also, if needed can be used as warming up test before main testing step.

1. Capacity testing

Should be performed to find the number of virtual users which the application support in stable state. The test must be performed as one of first main tests to find optimal load model for other types of performance tests and should be performed each time after significant changes in the application or its configuration.

1. Load test

Load test is a kind of the most regular tests to check benchmark of the application and its components. Can be performed after each insignificant change in code base or configuration.

1. Stress testing

Stress testing supposed to run occasionally to check application’s stability under high load. Can be performed close to after code complete or by special request.

1. Scalability testing

Can be performed once or more in order to get multiplier(s) for different number of front-end/other servers.

1. Volume testing

Is to run with small, planned and huge amount of data with regular load to get indicators on application’s responsiveness/metrics change. Should be performed at least once or by special request.

1. Longevity testing

Supposed long time running the test with the load lower than average. Should be performed occasionally after significant code changes or by special request to make sure the application’s responsiveness and key performance indicators do not change significantly after long time running, and to check on memory leak as well.

### Level of testing

The following levels of performance testing are supposed:

|  |  |  |
| --- | --- | --- |
| **#** | **Level** | **Description** |
| 1 | Purely mocked version | Used to test only application’s tier with no engaged services: basic client-side testing, smoke testing, checking test scripts etc. |
| 2 | Partially mocked version | Used to test particular set(s) of application’s components with test data to get their benchmark with no additional load for mocked parts. |
| 3 | Full set of test data | Used for end-to-end testing of the system |

## Non-Functional requirements

* Highly scalable
* Highly consistency
* High availability
* High reliability
* User data should be durable (Any uploaded photos should never be lost)
* The maximum latency for generating the News Feed is 150 ms.

## Environmental needs (facilities, hardware, software, network, supplies, level of security, special tools)

### Test environment(s)

For performance testing a dedicated environment is needed. The configurations of the servers should be as much as close to production’s ones.

Expected numbers of the servers:

* Front-End servers
* Back-End servers
* Databases (PostgreSQL, Hive, Pgbouncer, Redis, Memcache)

To perform partial and full end-to-end performance testing it’s recommended to have a separated database and other services, i.e. they should not be located on the same server where related services are being tested by another team.

**Test environment specification(s):**

Close to Production environment as much as possible.

### Testing tools

| **Module** | **Software** | **Distribution** | **Where to install** | **Supported OS** | **Description** |
| --- | --- | --- | --- | --- | --- |
| **Module** | **Software** | **Distribution** | **Where to install** | **Supported OS** | **Description** |
| CI/CD | Jenkins | Manual installation | Load Generator host | Linux/ Windows | Module that starts and control execution of your performance tests |
| Monitoring agent | Telegraf | Custom installation package | All involved in testing hosts | Linux/ Windows | Module that gets all performance and health metrics from host where it installed (including Load Generator) to get more accurate picture of your performance tests impact |
| Load Generator (LG) | JMeter | Manual installation | Dedicated host | Linux/ Windows | Module that generates load on your application |
| Collector | InfluxDB | Docker image | Dedicated/Load Generator host | Linux/ Windows | Module that collects performance metrics from monitoring agents and send them to our cloud |
| Metrics visualization | Grafana | Manual installation | Dedicated host | Linux/ Windows | Module that visualizes all collected performance metrics |

## Test data

To have whole cycle of performance testing test data for Instagram should be:

1. Reusable
2. Generated in necessary amount for different stage of testing at any time:
   1. User accounts
   2. Posts with photos and videos
   3. Daily and saved stories.
   4. Some history and log records for volume testing
   5. Other
3. Cleanable (for example history, log records)

## Performance Entry, Exit, and Suspension Criteria

### Entry Criteria

* Test plan is complete and approved by the client.
* Correct version is installed in performance testing environment, i.e. the version previously functionally tested and fixed if needed
* Test data is complete and in the performance testing environment in sufficient time to allow test scripts to be completed.
* Test accounts have been created in the performance testing environment in sufficient time to allow test scripts to be completed.
* Test scripts complete.
* All assigned resources are available to monitor the test.
* All parameter sets used in the script are generated based on the Database values.

### Exit Criteria

* All test scripts completed successfully
* No critical problems encountered
* All non-critical problems are logged
* All test logs are captured
* All post-test notifications sent

### Suspension Criteria

* Not all test scripts will complete
* Critical problems are encountered and logged
* Hardware errors prevent the completion of the test

## Responsibilities

Oleksandr Maksymenko – Performance Engineer

Oleksandr Maksymenko – Automation Quality Assurance Engineer

## Risks and contingencies

* A significant difference in configuration from the production environment
* Performance testing results can be essentially different even in case of minor difference in think times, arrival rate and test duration
* During the execution of the tests, some major performance or functional problems that may require code changes, creation of a new build may be discovered and in that case it may be necessary to repeat the load test from the beginning
* Load test should be performed against a build that is solid enough, and that has been functionally tested, after code is complete. Failure to follow this rule may result on rework to update test scripts for every new build, plus the load test may need to be repeated from the beginning. This will affect the schedule
* Performance testing tool is not capable of identically reproducing real life scenarios - so results could only be trusted as having limited reliability level
* Network/systems latency issues
* Environment’s unavailability