



Case Study: Demonstrating 80Gb load performance

How to solve the challenge of high-throughput full-stack mobile traffic testing for Tier 1 operators

Introduction:

The need to demonstrate load performance

A leading network equipment vendor needed to demonstrate to a Tier 1 operator customer that a solution it had delivered could handle the promised load and perform as expected with different traffic models. It had tried to do so with an existing test solution, but this was unable to reach the required performance levels and lacked flexibility to model the specific traffic flows demanded. With time running out to prove that they could deliver, the vendor turned to the PureLoad solution.

PureLoad offered the required scalability and flexibility, allowing the vendor to prove that its solution could deliver the performance demanded by the operator.

The project was completed in the nick of time, enabling the vendor to meet its commitments and successfully deliver to a key customer account.

Background

Testing performance and scalability from an early stage is an essential part of today's software development projects. No organisation can afford to lose business by discovering performance-related issues after the launch of a service.

Operators have every expectation that that the service will scale and reach the desired audience. Trouble-shooting and fixing performance-related errors post-launch is not only costly, but can also create bad-will among customers and end users, and significantly damage your reputation.

Networks are becoming increasingly application dependent as network functions account for a larger share of content and awareness grows. Networks are evolving entities as customers demand the addition of more devices and apps, which can lead to new and disruptive traffic patterns that, in turn, create new challenges for operators and service providers.

That's why it's increasingly important to ensure that testing solutions can simulate realistic usage patterns that mimic real-life user actions, behaviours and devices. These simulations need to be scalable and easily adaptable to ensure that they can quickly respond to changes in user behaviour, device types, characteristics and content types. In short, testing at this complex level requires an extreme performance testing environment, backed with flexibility.

"A leading network equipment vendor needed to demonstrate to a Tier 1 operator customer that a solution it had delivered could handle the promised load and perform as expected with different traffic models. It had tried to do so with an existing test solution, but this was unable to reach the required performance levels and lacked flexibility to model the specific traffic flows demanded."

The scope:

Deliver a realistic 80Gb test in three weeks

PureLoad provides the ability to simulate extreme volumes of application traffic (OSI Layer 7) to ensure absolute scalability and performance in any application infrastructure, across your subscriber base.

A leading network equipment vendor needed to demonstrate to a Tier 1 operator customer that a solution it had delivered could handle the promised load and perform as expected with different traffic models. It had tried to do so with an existing test solution, but this was unable to reach the required performance levels and lacked flexibility to model the specific traffic flows demanded.

With time running out to prove that they could deliver, the vendor turned to the PureLoad solution.

In this case, PureLoad was required to simulate web traffic (HTTP and HTTPS) – at a level equating to that of a small country (80 Gbit/s throughput). Our challenge was to design, build and execute a realistic 80Gbit test within three weeks, so that the vendor could meet its obligations.

We completed the project ahead of schedule and to the delight of the customer. This paper outlines below how we did this.

The requirements:

Traffic model and simulation

At the start, we based our traffic model on long-term measurements taken from a live network with genuine end users. Our simulation used a 19kB average payload with two requests per TCP connection and an average session length of 40 seconds. The goal was to generate 80 Gbit/s throughput for the system under test.

- When securing networks from an application perspective, the traffic generated must mirror genuine traffic generated by real users. Sending large chunks of payload just to fill the network is not sufficient. A base for a realistic simulation must include the use of a full-stack client for traffic generation, which provides realistic network behaviour with the client having to react to ever-changing network conditions, such as TCP retransmissions.
- It's also essential to create varying (non-static) user behaviour. As well as creating a large number of users, the simulation needs to ensure that end users all behave differently and their usage has different characteristics, as in the real world. People do different things.
- In this case it was important that users were simulated using different IP addresses – we used around 10 million unique IP addresses. Other client characteristics, such as UAProf and number of connections used, can also be built into any test simulation.
- Different user behaviour is another important consideration. For example, longer and shorter browsing sessions, staying on one app, or moving slowly / quickly between different apps. The content used should also be unique, such as the use of text, images and movies in different formats.

“The solution was up and running and the relevant tests executed in less than three weeks, allowing the vendor to prove that it could meet the performance requirements demanded by the operator, and ensure acceptance of the delivered solution.”

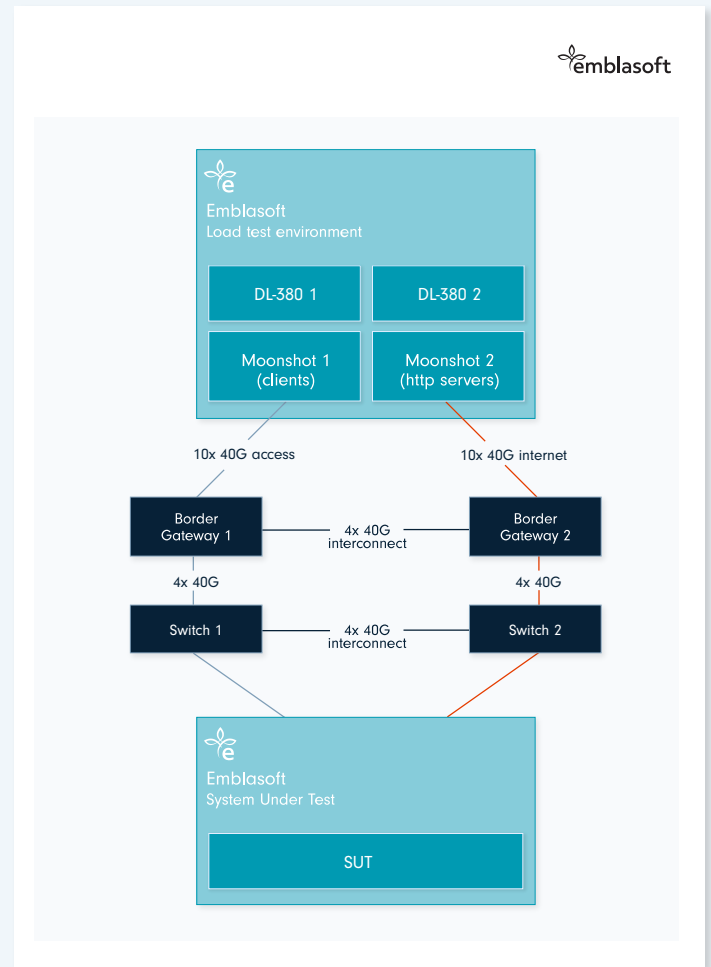
The solution: Designing the platform

Considering the high-load levels involved, hardware requirements for the test environment were moderate. The solution used two standard HP Moonshot chassis, each with 40 HP 710x blades (4-core, 1 CPU). The blades were used for generating client traffic and for running the terminating HTTP servers (nginx-based).

The console was run from 'DL-380 1' and the administration of the environment from 'DL-380 2.' Traffic flow came from PureLoad clients via 'Border Gateway 1' and 'Switch 1' to the SUT.

From the SUT, traffic flowed via 'Switch 2' and 'Border Gateway 2' to the terminating HTTP servers in the Moonshot 2 chassis. The traffic then followed the corresponding path back to the clients. In order to implement such a large system in a short time span, a virtualised environment was used.

This provided good flexibility when it came to scalability of OS resources. The load test environment was implemented on KVM running on Ubuntu 14.04 LTS. The compute hosts were two HP Moonshot chassis with 710x cartridges. All compute and admin hosts also ran Ubuntu 14.04. The total number of VMs was 92 load servers and 70 web servers.



The scope: Solution design

The sole aim of this performance test was to prove that the system could handle 80 Gbit/s throughput of constant traffic.

Test results

The customer achieved their goal of generating 80 Gbit/s throughput with a behaviour that mimicked the desired traffic model.

The model used a 19kB average payload and two requests per TCP connection with an average session length of 40s. Ten million unique IP addresses were used.

This provided an Extreme Performance Test of:

- 500,000 HTTP(S) transactions per second
- 250,000 new TCP connections per second
- 11 million established TCP connections with the rate of connection per second fluctuating with an average of 40s

Timing

The solution was up and running and the relevant tests executed in less than three weeks, allowing the vendor to prove that it could meet the performance requirements demanded by the operator, and ensure acceptance of the delivered solution.

Summary

Thanks to the flexible design of PureLoad and our strong competence in performance testing and modelling, it was possible to set up and execute an extreme performance test generating 80 Gbit/s throughput using 10 million unique IP addresses within the required timescales, enabling the vendor to prove that it could deliver the promised performance to the operator.