NEXT GENERATION INTERNET

Open Call 1

CacheCash

Deliverable 3: Experiment Results and Final Report

Authors	Timur Friedman, timur.friedman@sorbonne-universite.fr, Sorbonne Université Olivier Fourmaux, olivier.fourmaux@sorbonne-universite.fr, Sorbonne Université Berat Senel, berat.senel@lip6.fr, Sorbonne Université Maxime Mouchet, maxime.mouchet@lip6.fr, Sorbonne Université Justin Cappos, jcappos@nyu.edu, NYU Tandon School of Engineering Rick McGeer, rick.mcgeer@us-ignite.org, US Ignite
Due Date	31 March 2021
Submission Date	2 June 2022
Assigned Reviewers	
Keywords	

Deliverable 3: Part I

Analysis, results, and wider impact

The information contained in sections 1-11 will be used in part to update the NGI Atlantic's public deliverables (including the Experiment Catalogue on the website).

1 Abstract

The experiment was based on two cutting-edge technologies.

CacheCash, the CDN technology that we were testing, and which is developed by our US partner at NYU, has the potential to change the nature of CDNs by involving the end-users themselves directly in serving content through machines that are under their control. It provides a service where interested users run caches and are incentivized to participate by receiving a cryptocurrency in exchange for serving content.

EdgeNet, which we have been developing in Europe, brings service deployment to the edge cloud, which is the key enabling technology, with customized extensions for Kubernetes. It has been successfully piloted in Fed4FIRE+ open call experiments and is now ready for larger-scale experiments.

2 Project Vision

Much has been made for the potential for blockchain technologies to improve the security, scalability, and economics of distributed data delivery. We propose to understand the real-world feasibility of such technologies by deploying a cutting edge blockchain-based content distribution network (CDN) to distribute open source software to real world users. In this way, we will be able to evaluate the strengths of such approaches, the weaknesses, and to identify research challenges that must be tackled to allow these technologies to become widespread.

Delivery of large chunks of content, such as video, accounts for a substantial percentage of all internet traffic. Such content, when distributed by large corporations, benefits from CDNs, with nodes hosted by internet service providers (ISPs) providing local service to end users in those ISPs. CacheCash (https://ssl.engineering.nyu.edu/projects#cachecash), the CDN technology that we test, has the potential to change the nature of CDNs by involving the end users themselves directly in serving content through machines that are under their control. Putting users in charge can lead to a wider range of content benefitting from CDNs. CacheCash provides a service in which interested users run caches, and they are incentivised to participate by receiving a crypto-currency (Cachecoin) in exchange for serving content to other users. Both cryptographic and economic analyses demonstrate that the incentives lead CacheCash users to honestly serve content. Analysis has also shown that CacheCash can scale to meet the workload of even the most popular services used today.

Our experiment put these ideas to the test by using EdgeNet, an edge cloud with about 50 nodes in distributed locations. This also allowed us to strain and evaluate EdgeNet's capabilities, at the same time, to derivate decryption strategies and reveal implementation issues of CacheCash thanks to the data harvested.

Co-funded by the Horizon 2020
Framework Programme of the European Union

3 Details on participants (both EU and US)

EU/PI Timur Friedman is a faculty member at Sorbonne Université, and a member of the LIP6 and LINCS computer science laboratories. His research expertise is in the area of internet measurements. Starting in 2008, he was co-responsible for setting up and running the PlanetLab Europe testbed, part of the EU's Fed4FIRE+ testbed federation. He has extensive experience in EU projects (FP6, FP7, H2020), most recently as leader of the EMBERS project (https://embers.city/). EU/PI Friedman brought the PlanetLab Europe testbed to the CacheCash experiment, and joined US/PI in bringing the EdgeNet technology. EU/PI Timur Friedman is joined by his colleague Olivier Fourmaux, also a faculty member at Sorbonne Université, and a member of the LIP6 computer science laboratory. Dr. Fourmaux co-directs the work on EdgeNet with EU/PI Friedman.

The principal technical work on the project was carried out by Berat Senel, currently a research engineer at Sorbonne Université, where he is principal developer of EdgeNet. Mr. Senel has an engineering degree and he was the founder of an information technology start-up in the transportation sector, which he sold his shares after got started at Sorbonne Université.

US/PI Justin Cappos is an associate professor in the Computer Science and Engineering department at New York University's Tandon School of Engineering. His research focuses on practical systems security research advances that can be used in production. His research advances are adopted into production use by Docker, git, Google, Python, VMware, automobiles, Cloudflare, Microsoft, Amazon, Digital Ocean, and most Linux distributions. He is the security assessment lead for the Linux Foundation's Cloud Native Computing Foundation. Due to the practical impact of his work, Justin was named to Popular Science's Brilliant 10 list in 2013. US/PI Cappos brought the CacheCash technology, which was tested.

US/PI Rick McGeer holds a Ph.D. from the University of California (UC) Berkeley. He was instrumental in the development of PlanetLab and GENI. He is the chief scientist of US Ignite and an adjunct professor at the University of Victoria. He has held research positions at SAP, HP, Cadence, and UC Berkeley, founded Softface, and was an assistant professor at the University of British Columbia. He has written over 100 refereed papers and two books in the fields of computer-aided design, distributed systems, networking, and cloud computing. US/PI McGeer brings the EdgeNet technology, which he pioneered, and which is being adopted by PlanetLab Europe, and which was used for the CacheCash experiment.

Co-funded by the Horizon 2020
Framework Programme of the European Union

4 Results

We created six setups at different times to analyze the effect of cache location on content retrieval for clients distributed in diverse geolocations. Each setup contains caches on an EdgeNet node located at a different location. Throughout the measurement, for each setup, we made repeated CacheCash-specific curl requests from EdgeNet nodes. We performed these measurements to collect data regarding overall transaction duration for content retrieval, time-to-first-byte (TTFB), and remote procedure call (RPC) service latencies using the gRPC protocol.

TTFB in the 300-500 ms range is the standard expectation. A shorter TTFB than this range means good performance. Metrics confirm that TTFB in CacheCash is within the standard range despite encryption and decryption operations that are required as part of content delivery. Moreover, if a cache is close to a client, e.g., both are in the same city, the TTFB outcome indicates good performance.

On the other hand, operations regarding encryption, decryption, and puzzle-solving (the computation required for solving puzzles is a disincentive to cheat in CacheCash) add considerable overhead, causing the overall transaction duration to be nearly twice what it would otherwise be. However, this is not an unsolvable problem. The cryptography design can be altered to achieve a shorter transaction duration while still meeting security requirements.

In addition, we prepared a setup dedicated to throughput analysis. In this setup, we used the nodes in the LIP6 Lab, where we are based, to host essential services with caches, and we used a compute cluster at LIP6 to emulate CacheCash clients. The goal was to make as many concurrent requests as possible. Cache throughput reached 45 MiB/s with 150 concurrent clients. CacheCash's publisher daemon is memory intensive, which limited us from increasing the number of clients that make requests in order to reveal the maximum throughput capability.

Measurement data is publicly available at the following URL:

https://drive.google.com/drive/folders/1AIkyOmovtFZn3gZHQWPz_Uf7SX8xeWST?usp=sharing.

Deploying CacheCash on EdgeNet uncovered implementation issues and required features for both the CacheCash system and the EdgeNet testbed. At first, caches in CacheCash were not reachable over the Internet, revealing modifications that were required to allow them to be reached via a node's external IP address. We also found that secure communication between the CacheCash services was not possible to start with, and we implemented this feature. Regarding EdgeNet, there wasn't a component to automate secure communication between containers within the cluster. We therefore integrated a third-party solution in EdgeNet to enable secure communication. We also developed ready-to-use configuration files to allow CacheCash to be deployed on any Kubernetes cluster, and scripts to automate the collection of data from measurements.

Co-funded by the Horizon 2020
Framework Programme of the European Union

5 Discussion and Analysis on Results

In Deliverable 2, we presented initial findings that indicated CacheCash might be able to compete with conventional CDNs. This hypothesis was borne out by the results presented in Section 4 of the current deliverable. These show that TTFB in CacheCash aligns with the expectations for a CDN. On the basis of this evidence, we believe that if a widely distributed CacheCash network across the world were to be established, it is likely to have shorter TTFB than conventional CDNs.

However, with respect to content retrieval time, which includes cryptography and blockchain processes in CacheCash, we saw a different story. Despite good TTFB results, the current CacheCash implementation still requires optimization for content retrieval. The reason is that data chunks are delivered to a client in sets of four in an encrypted state, then the client exchanges information with the cache that distributed the content in order to decrypt the data. Because of this design, cryptography-related operations take up almost 50% of overall transaction duration in most cases. We believe that modifications to the CacheCash architecture will allow these issues to be overcome.

In Deliverable 2, we described a novel design for the puzzle solution aspect of CacheCash: requiring a client to find puzzle solutions for three chunks out of each four in order to decrypt data chunks. Our measurement data, described in the present deliverable, indicates that this design may improve TTFB by approximately 40 ms. The CacheCash team is developing an interface for changing these parameters so that they can study the optimal trade-off between performance and security.

Additionally, spikes occurred at gRPC handling times in gRPC services, resulting in delays. Implementation errors in the Golang code cause this. It is currently under investigation by the CacheCash team.

One of the goals of this experiment was to work with outside projects to distribute their content. In this way, CacheCash would officially start the public distribution of content from these projects. For example, we aimed to have CacheCash deliver Python packages as part of our experiment. We were not able to achieve this milestone due to two causes.

1. Our work in the CacheCash experiment revealed implementation issues that affect the CDN's performance. The CacheCash team is currently working to resolve these issues.
2. We discovered that certain software components, such as the 'pip' tool, needed to be extended before CacheCash could natively distribute Python packages. Other projects, besides Python, for which we hoped to distribute content, also required extensions to their specific tools. It is entirely possible to create these extensions, but not within the time and effort limits of the CacheCash experiment.

The CacheCash team aims to achieve this milestone at a later moment, once the implementation issues have been handled.

Co-funded by the Horizon 2020
Framework Programme of the European Union

6 Present and Foreseen TRL

As for the technology readiness levels of EdgeNet and CacheCash, we evaluated EdgeNet at about TRL 3 Experimental Proof of Concept, as there is a functioning instance of EdgeNet that is supporting experiments. CacheCash is closer to TRL 2, a technology concept that has been formulated. We aimed to move CacheCash to TRL 3 and also EdgeNet closer to TRL 4.

As an outcome of the experiment, CacheCash is indeed closer to TRL 3, as most of the implementation issues have been resolved in order to run a functioning instance. And EdgeNet is closer to TRL 4 as it becomes lab-validated and now offers a service to establish secure communication between deployed containers.

7 Exploitation, Dissemination and Communication Status

The purpose of the experiment was to trial CacheCash in the real world with real internet users consuming content that is of interest to them. In addition to the Python distributions mentioned in Section 5, we aimed to provide the content from The Update Framework, or TUF (https://theupdateframework.io/). We planned to slowly start by first evaluating the properties of CacheCash with a small subset of TUF users. However, as explained at in Section 5, we were not able to achieve these hoped-for exploitation results.

Our work did contribute to future exploitation for two start-up companies that are collaborating with US/PI Cappos in the content delivery field and that plan to adopt CacheCash.

Sorbonne Université plans to use the EdgeNet software to run its PlanetLab Europe testbed. With this successful trial of CacheCash on EdgeNet, one exploitation result is that we are starting to migrate PlanetLab Europe's users to EdgeNet. The migration of PlanetLab Europe's existing infrastructure into the EdgeNet cluster will follow that.

As for the dissemination and communication strategy, the partners use a variety of methods to achieve maximum outreach for the activities and results of the experiment. The principal actions are undertaken as follows:

Project websites

Experiment achievements and results will be published on the websites. The CacheCash team is in preparation to publish a website shortly.

Blog posts

We prepared eight blog posts, which are now ready to be published. Sorbonne Université and the NYU Tandon School collaborated to polish them. We will start publishing these blog posts in the following weeks.

Spreadsheets

Co-funded by the Horizon 2020
Framework Programme of the European Union

The measurement data accompanied by the data description is ready to be publicly available via Google Sheets.

Conferences and Workshops

We had a paper is accepted at the EdgeSys 2021 workshop, which will take place in April in conjunction with ACM EuroSys 2021. The paper describes EdgeNet and describes its role in supporting the CacheCash experiment. The evaluation work that we performed on CacheCash informed our evaluation of EdgeNet for this paper. Partners plan to prepare another paper dedicated to this experiment in the coming months.

Networking and community building actions

US/PI Cappos is already well anchored in the Linux Foundation open source community, and is advancing CacheCash in that ecosystem. Sorbonne Université's work with him has given us perspective on how EdgeNet can better integrate itself into the open source efforts of the Linux Foundation and other efforts surrounding the Kubernetes open source code on which EdgeNet is based.

Social Media

We made use of the official accounts of the partners on social media platforms to increase the interaction with end-users and to publicize the results. Once the measurement data is published, the social media accounts will be used actively to bring attention to this.

8 Impacts

Impact 1: Enhanced EU – US cooperation in Next Generation Internet, including policy cooperation.

Sorbonne Université's collaboration with US Ignite and the NYU Tandon School is an outcome of the Global Experimentation for Future Internet (GEFI) workshop, 26-27 October 2017, in Rio de Janeiro, Brazil. It became clear that there was a shared interest in working towards a common successor platform to the existing LXC-based PlanetLab technology, to move to the contemporary de facto industry standards of Docker and Kubernetes. This collaboration had, until this experiment, not been formalized through joint funding, but it is a very real and enduring collaboration, featuring weekly meetings to advance our joint goals.

Impact 2: Reinforced collaboration and increased synergies between the Next Generation Internet and the Tomorrow's Internet programmes.

The US side has in the past benefitted from an NSF EAGER grant to support some of this work, and the EU side has received partial support from French national funding. At times, the development initiative has been taken in the United States and Canada, and at times in France. Our ability to continue together had, however, constantly been in question due to the uncertainties on both sides of the Atlantic regarding ongoing financing. This

Co-funded by the Horizon 2020 Framework Programme of the European Union

NGIatlantic.eu funding is the first official recognition of our collaboration, enabling the US partners to seek similar funding, and it sets the entire collaboration on a firmer footing for the future. This experiment also increased the level of collaboration between EU and US partners, allowing them to exchange information about the NGI Generation Internet and the Tomorrow's Internet programmes to seek further potential collaborations in the future.

Impact 3: Developing interoperable solutions and joint demonstrators, contributions to standards.

This successful CacheCash trial on EdgeNet means that CacheCash is capable of involving any Kubernetes-based cluster such as EdgeNet to run caches on its nodes, and EdgeNet can host any CDN solution. This mutual interest allows both software-based systems to be demonstrated to the trans-Atlantic community. Because both are open source solutions, they can either be adopted by high-tech startups or empower spin-off startups. Moreover, EdgeNet, as a testbed, can attract the EU-US ecosystem of top researchers to run their experiments.

Impact 4: An EU - US ecosystem of top researchers, hi-tech start-ups / SMEs and Internet-related communities collaborating on the evolution of the Internet

EdgeNet has now received funding from the US-based company VMware, which has now based much of its business strategy on Kubernetes, the same underlying technology that EdgeNet uses. Through our VMware connections, we have been introduced into the Kubernetes working groups that are relevant to our EdgeNet work: the Multitenancy working group and the IoT-Edge working group. We have made presentations to these working groups, and we are preparing code that we plan to present to the Multitenancy working group and that we hope will be taken up by the Kubernetes community.

9 Conclusion and Future Work

CacheCash is a blockchain-based content delivery network technology that aims to establish a broader network at the edge than conventional CDNs. The measurement data gathered throughout this experiment demonstrated that CacheCash is a promising framework for next generation CDN infrastructures. Despite the performance issues, which are understandable at this TRL for such a complex system, CacheCash meets general expectations for a conventional CDN. Furthermore, it has the potential to improve CDN performance in terms of TTFB and content retrieval time by incentivizing end-user with CacheCoin in exchange for serving content as a part of the network. To do so requires further optimizing the current design of the cryptography processes.

Co-funded by the Horizon 2020
Framework Programme of the European Union

From EdgeNet's point of view, this experiment showed the testbed has the ability to welcome advanced experiments like CacheCash. A feature added to the testbed is a certificate manager to enable secure communication between containers. EdgeNet could improve its fair sharing of storage on nodes as it starts to welcome more experimenters.

As future work, the CacheCash team focuses on solving the implementation issues revealed by this experiment. Following that, they plan to enable an interface to change the cryptographic design easily, so that they can study performance versus security trade-offs. At this point, EdgeNet again will be the place to put CacheCash into the test. Once done, the blog posts will be published on relevant communication platforms, including the CacheCash website that is under construction. In addition to these, the preparation of a paper that presents the findings is underway. We also look forward to further collaboration to strengthen ties between the institutions, including researcher visits when the current pandemic is past.

10 Glossary

CDN	Content Delivery Network
Fed4FIRE+	Federation For Fire Plus
ISP	Internet Service Provider
PI	Principal Investigator
TTFB	Time To First Byte
TRL	Technology Readiness Level
TUF	The Update Framework
NSF	National Science Foundation
LXC	Linux Containers
NGI	Next Generation Internet

Co-funded by the Horizon 2020
Framework Programme of the European Union

Deliverable 3: Part II

Financial and cost information

This part is to be treated as a consortium confidential deliverable, and access is restricted to consortium partners and EU commission operatives.

Co-funded by the Horizon 2020
Framework Programme of the European Union