



Nuno Morais

Master of Science

Monitoring and deploying services on edge devices

Dissertação para obtenção do Grau de Mestre em
Engenharia Informática

Orientador: João Leitão, Assistant Professor,
NOVA University of Lisbon

Júri

Presidente: Name of the committee chairperson
Arguente: Name of a rapporteur
Vogal: Yet another member of the committee



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

December, 2019

Monitoring and deploying services on edge devices

Copyright © Nuno Morais, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa.

A Faculdade de Ciências e Tecnologia e a Universidade NOVA de Lisboa têm o direito, perpétuo e sem limites geográficos, de arquivar e publicar esta dissertação através de exemplares impressos reproduzidos em papel ou de forma digital, ou por qualquer outro meio conhecido ou que venha a ser inventado, e de a divulgar através de repositórios científicos e de admitir a sua cópia e distribuição com objetivos educacionais ou de investigação, não comerciais, desde que seja dado crédito ao autor e editor.

RESUMO

Lorem ipsum em Português.

Palavras-chave: Palavras-chave (em Português) ...

ABSTRACT

Lorem ipsum in english.

Keywords: Keywords (in English) ...

ÍNDICE

Lista de Figuras	xi
Lista de Tabelas	xiii
Listagens	xv
1 Introduction	1
1.1 Context	1
1.2 Motivation	2
1.3 Expected Contribution	2
1.4 Document Structure	3
2 Motivation	5
2.1 What is the problem	5
3 Related Work	7
3.1 Topology Management	7
3.1.1 Random overlays	7
3.1.2 Structured overlays	7
3.1.3 Self-adapting overlays	7
3.2 Aggregation	7
3.2.1 Types of aggregation	7
3.2.2 Relevant aggregation protocols	7
3.3 Resource Discovery	7
3.4 Offloading computation to the edge	7
4 Planning	9
4.0.1 Proposed solution	9
4.0.2 Scheduling	9
Bibliografia	11
A Appendix 2 Lorem Ipsum	13
I Annex 1 Lorem Ipsum	15

LISTA DE FIGURAS

LISTA DE TABELAS

LISTAGENS

INTRODUCTION

1.1 Context

Nowadays, the Cloud Computing paradigm is the standard for development, deployment and management of services, it has proven to have massive economic benefits that make it very likely to remain permanent in future of the computing landscape. It provides the illusion of unlimited resources available to services, and has changed the way developers, users and businesses rationalize about applications [1]. Currently, most software present in our everyday life such as Google Apps, Amazon, Twitter, among many others is deployed on some form of cloud service.

However, currently, the rise in popularity of mobile applications and IoT applications differs from the centralized model proposed by the Cloud Computing paradigm. With recent advances in the IoT industry, it is safe to assume that in the future almost all consumer electronics will play a role in producing as well as consuming data. The number of devices at the edge and the data they produce will increase rapidly, and transporting the data to be processed in the Cloud will become unfeasable in the future.

When all computations reside in the data center (DC), far from the source of the data, problems arise: from the physical space needed to contain all the infrastructure, the increasing ammount of bandwidth needed to support the information exchange from the DC to the client, the latency in communication from the client to the DC as well as the security aspects that arise from offloading data storage and computation, have directed us into a post-cloud era where a new computing paradigm emerged, Edge Computing.

Edge computing takes into consideration all the computing and network resources that act as an "edge" along the path between the data source and the DC and addresses the increasing need for supporting interaction between cloud computing systems and mobile or IoT applications [5]. However, when accounting for all the devices that are external to

the DC, we are met by a huge increase in heterogeneity of devices: from Data Centers to private servers, desktops and mobile devices to 5G towers and ISP servers, among others.

1.2 Motivation

The aforementioned heterogeneity implies that there is a broad spectrum of computational, storage and networking capabilities along the edge of the network that can be leveraged upon to perform different types of computation that rely on the individual characteristics of the devices performing the tasks, which can vary from generic computations to aggregation, summarization, and filtering of data. [3]

Systems that require real-time processing of data may not even be feasible with Cloud Computing. When the volume of data increases, transporting the data in real time to a Data Center is impossible, for example, a Boeing 787 will create around 5 gigabytes of data per second [2], and Google's self-driving car generates 1 Gygabyte every second [4], which is infeasible to transport to the DC for processing and responding in real-time.

There have been efforts to move computation towards the Edge of the network, Fog Computing [6], which is an extension of cloud computing from the core of the network to the edge of the network, has shown to benefit web application performance [7], additionally, Content Distribution Networks [] and Cloudlets [] are an extension of this paradigm and are extensively used nowadays.

To fully materialize the Edge Computing paradigm, applications need to be split into small services that cooperate to fulfill applicational needs. Additionally, tools must be developed that allow service and resource discovery in addition to service and device monitoring. All of this performed over the inherent scale considered by the infrastructure of the Edge Computing paradigm.

paragrafo com sobre para quem e que isto e (e.g. Google, Amazon, Smart City, Smart Country ?? also porque e que isto e diferente do que ja existe

These tools must be based on protocols that are able to federate all the devices and aggregate massive amounts of device data in order to perform efficient service deployment and management. Additionally, they must handle the churn and network instability that arises when relying on devices that do not have the same infrastructure as those in Data Centers.

1.3 Expected Contribution

To achieve this, we propose to create a new novel algorithm which employs a hierarchical topology that resembles the device distribution of the Edge Infrastructure. This topology is created by assigning a level to each device and leveraging on gossip mechanisms to build a structure resembling a FAT-tree [].

The levels of the tree will be determined by ...undecided... and will the tree be used to employ efficient aggregation and search algorithms. Each level of the tree will be

composed by many devices that form groups among themselves, the topology of the groups ...undecided...

The purpose of this algorithm is to allow:

1. Efficient resource monitoring to deploy services on.
2. Offloading computation from the cloud to the Edge and vice-versa through elastic management of deployed services.
3. Service discovery enabled by efficiently searching over large ammount of devices
4. Federate large ammount of heterogeneous devices and use hetereogeneity as an advantage for building the topology.

We plan to research existing protocols (both for topology management and aggregation) and enumerate their trade-offs along with how they behave across different environments. Then, employ a combination of different techniques inspired in the performed research in a unique way that is tailored for this topology.

1.4 Document Structure

The document is structured in the following manner:

Chapter 2 focuses on the related work, first section covers the different types of topology management protocols, with an emphasis on random and self-adapting overlays, second section studies the different types of aggregation and popular implementations for each aggregation type. Third section adresses resource discovery and how to perorm efficient searches over networks composed by a large number of devices. Finally, fourth section discusses recent approaches towards enabling Edge Computing, along with discussion about Fog, Mist and Osmotic Computing.

Chapter 3 further explains the proposed contribution along with the work plan for the remainder of the thesis.

MOTIVATION

2.1 What is the problem

RELATED WORK

3.1 Topology Management

3.1.1 Random overlays

3.1.2 Structured overlays

3.1.3 Self-adapting overlays

3.2 Aggregation

3.2.1 Types of aggregation

3.2.2 Relevant aggregation protocols

3.3 Resource Discovery

3.4 Offloading computation to the edge

PLANNING

4.0.1 Proposed solution

4.0.2 Scheduling

BIBLIOGRAFIA

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica e et al. “A View of Cloud Computing”. Em: *Commun. ACM* 53.4 (abr. de 2010), 50–58. ISSN: 0001-0782. DOI: [10.1145/1721654.1721672](https://doi.org/10.1145/1721654.1721672). URL: <https://doi.org/10.1145/1721654.1721672>.
- [2] M. Finnegan. *Boeing 787s to create half a terabyte of data per flight, says Virgin Atlantic*. 2013. URL: <https://www.computerworld.com/article/3417915/boeing-787s-to-create-half-a-terabyte-of-data-per-flight--says-virgin-atlantic.html>.
- [3] J. Leitão, P. Á. Costa, M. C. Gomes e N. M. Preguiça. “Towards Enabling Novel Edge-Enabled Applications”. Em: *CoRR* abs/1805.06989 (2018). arXiv: [1805.06989](https://arxiv.org/abs/1805.06989). URL: <http://arxiv.org/abs/1805.06989>.
- [4] *Self-driving Cars Will Create 2 Petabytes Of Data, What Are The Big Data Opportunities For The Car Industry?* URL: <https://datafloq.com/read/self-driving-cars-create-2-petabytes-data-annually/172>.
- [5] W. Shi, J. Cao, Q. Zhang, Y. Li e L. Xu. “Edge Computing: Vision and Challenges”. Em: *IEEE Internet of Things Journal* 3 (out. de 2016), pp. 1–1. DOI: [10.1109/JIOT.2016.2579198](https://doi.org/10.1109/JIOT.2016.2579198).
- [6] S. Yi, Z. Hao, Z. Qin e Q. Li. “Fog computing: Platform and applications”. Em: *2015 Third IEEE Workshop on Hot Topics in Web Systems and Technologies (HotWeb)*. IEEE. 2015, pp. 73–78.
- [7] J. Zhu, D. Chan, M. Prabhu, P. Natarajan e H. Hu. “Improving Web Sites Performance Using Edge Servers in Fog Computing Architecture”. Em: mar. de 2013, pp. 320–323. ISBN: 978-1-4673-5659-6. DOI: [10.1109/SOSE.2013.73](https://doi.org/10.1109/SOSE.2013.73).

A P Ê N D I C E



APPENDIX 2 LOREM IPSUM



ANNEX 1 LOREM IPSUM