

Positive Energy Worker

Autonomous Positive Energy Machine Running on The Blockchain



Zied Guesmi

Supervisors: Prof. A.B. Supervisor
Prof. C.D. Supervisor

Advisors: Dr. A. Advisor
Dr. B. Advisor

Ministry of Higher Education And Scientific Research.
University of Tunis ELMANAR, Tunisia
Faculty of Mathematical, Physical and Natural Sciences of Tunis
Department of Computer Science Engineering

This dissertation is submitted for the degree of
Computer Engineer

Department of Computer Science
Engineering

September 2018

I can not dedicate something that does not belong to me, so dear parents . . .

Acknowledgements

And I would like to acknowledge ...

Abstract

This is where you write your abstract ...

Table of contents

List of figures	xiii
General Introduction	1
1 Project Context	3
1.1 Introduction	3
1.2 Project background	3
1.3 Host company	3
1.4 Problematic	5
1.5 Suggested solution	5
1.6 Conclusion	5
2 State of The Art	7
2.1 Introduction	7
2.2 Critical overview of the existing	7
2.2.1 Solar-Powered Data Centers	7
2.2.2 Wind-Powered Data Centers	8
2.2.3 Geothermal Data Centers	8
2.2.4 Limitations of previous solutions	9
2.3 State of the art	9
2.4 Conclusion	11
3 Requirements Analysis	13
3.1 Introduction	13
3.2 Global System Analysis	13
3.2.1 iExec ecosystem	13
3.2.2 General Sequence Diagram	15
3.3 Requirements	18
3.3.1 Functional Requirements	18

3.3.2	Non-Functional Requirements	18
3.4	Conclusion	19
4	Design & technical specifications	21
4.1	Introduction	21
4.2	Desing and architecture	21
4.2.1	xtremweb	22
4.2.2	blockchain	22
4.2.3	energy	22
4.2.4	IoT	22
4.3	Energy	23
4.4	Geth node	23
4.5	Use cases	23
4.5.1	Surveillance Camera	23
4.5.2	Weather Station	23
4.6	Technical Specifications	23
4.6.1	specs of the board	23
4.6.2	automate configuration and deployment	23
4.6.3	develop dapps (use cases)	23
4.6.4	deployment of all the environment	23
4.6.5	cross-compiling	23
4.6.6	cluster	23
4.7	Conclusion	23
5	Implementation	25
5.1	Introduction	25
5.2	Work environment	25
5.2.1	Hardware environment	25
5.2.2	Software environment	25
5.2.3	Development Technologies	25
5.2.4	Documentation	25
5.2.5	cluster + automation	25
5.3	Illustration	25
5.4	Conclusion	25
	General Conclusion And Perspectives	27

Table of contents	xi
-------------------	-----------

References	29
-------------------	-----------

List of figures

1.1	iExec logo	3
2.1	Emerson Network Power solar array	7
2.2	QTS Princeton datacenter photovoltaic solar panels	8
2.3	The 131-foot tall wind turbine at Other World Computing in Woodstock, Illinois	8
2.4	Ethereum blockchain	9
2.5	RLC token	10
3.1	iExec decentralized cloud architecture	14
3.2	iExec marketplace	15
3.3	iExec core sequence diagram - part 1	16
3.4	iExec core sequence diagram - part 2	17
4.1	xxxxxxxxxxx	21

General Introduction

Whether we are online shopping, streaming videos or reading the feed on social media, every internet activity involves huge amounts of data that needs to be stored and processed somewhere. Datacenters are there all over the world to accomplish this mission. They have already spread from virtually nothing 10 years ago to consuming about 3 per cent of the global electricity supply and accounting for about 2 per cent of total greenhouse gas emissions. That gives it the same carbon footprint as the airline industry[1].

“ If we carry on going the way we have been it would become unsustainable – this level of data centre growth is not sustainable beyond the next 10 to 15 years. The question is, what are we going to do about it? ”

*Professor Ian Bitterlin*¹

We have already improved everything around compute architecture, the only thing left to improve is the compute side of the equation. It's either a breakthrough in our perspective about it, or we need to get deadly serious about doubling the number of power plants on the planet. One way to curb their carbon footprint is to increase the amount of renewable energy they use. But even if the industry was able to shift to 100 per cent renewable electricity, the volume of energy they would need would put intolerable pressure on the world's power systems. So which way the industry decides to go could have a huge bearing on whether renewable energy receives the huge investment that drives innovation – bringing down the cost of green electricity to everybody's benefit. It could also play a large role in determining whether the world can avoid the worst ravages of global warming.

¹ Professor Ian Bitterlin is a Chartered Engineer with more than 25 years' experience in data-center power and cooling, CTO of Emerson Network Power Systems in EMEA and a Visiting Professor at University of Leeds in the School of Mechanical Engineering.

In this context, this project, originally suggested by iExec, is performed as part of the preparation to obtain the computer science engineering degree from the Faculty of Mathematical, Physical and Natural Sciences of Tunis, University of Tunis ELMANAR, Tunisia. It aims to suggest a solution to this problem in an innovative approach by embracing Edge computing concept combined with solar energy.

This report illustrates the work that has been done, from design to requirements and implementation. Those topics are covered in five chapters, we start by contextualizing the project and discussing the state of the art with a comparison between our perspective and some existing ones. Then, we present the different requirements and technical specifications which leads us to the implementation details. Finally we conclude and suggest some improvements to the current limitations.

Chapter 1

Project Context

1.1 Introduction

The aim of this chapter is to contextualize the project by giving its background, to present the host company and to define in a later section the main problem as well as the way we address it.

1.2 Project background

This project is about designing and implementing a positive energy worker. A Raspberry Pi based system that powers iExec's infrastructure and serves - at the same time - as an IoT device to embrace the Edge computing concept. It is achieved in the context of the preparation of the end of studies project submitted to obtain computer science engineering degree.

1.3 Host company



Fig. 1.1 iExec logo

Started in 2016, iExec[2] was co-founded by Dr. Gilles Fedak and Pr. Haiwu He. Ph.D, CEO and Co-Founder, Dr. Gilles Fedak has been a permanent INRIA research scientist since 2004 at the ENS in Lyon, France. His research interests lie in Parallel and Distributed Computing, with a particular emphasis on the problematic of using large and loosely-coupled distributed computing infrastructures to support highly demanding computational and data-intensive science. He co-authored about 80 peer-reviewed scientific papers and won two Best Paper awards. Pr. Haiwu He, Ph.D, Co-Founder and Head of Asian-Pacific Region was a research engineer expert at INRIA Rhone-Alpes in Lyon, France from 2008 to 2014. He has published about 30 refereed journal and conference papers. His research interest covers peer-to-peer distributed systems, cloud computing, and big data.

iExec aims at providing decentralized applications running on the blockchain a scalable, secure and easy access to the services, data-sets and computing resources they need. This technology relies on Ethereum smart contracts and allows the building of a virtual cloud infrastructure that provides high-performance computing services on demand.

iExec leverages a set of research technologies that have been developed at the INRIA and CNRS research institutes in the field of Desktop Grid computing. The idea of Desktop Grid (aka. Volunteer Computing) is to collect the computer resources that are underutilized on the Internet to execute very large parallel applications at the fraction of the cost of a traditional supercomputer. iExec relies on XtremWeb-HEP[3], a mature, solid, and open-source Desktop Grid software which implements all the needed features: fault-tolerance, multi-applications, multi-users, hybrid public/private infrastructure, deployment of virtual images, data management, security and accountability, and many more.

iExec is developing a new Proof-of-Contribution[4] (PoCo) protocol, that will allow off-chain consensus. Thanks to the Proof-of-Contribution, external resource providers will have the usage of their resources certified directly in the blockchain.

iExec aims to deploy a scalable, high-performance, secure and manageable infrastructure sidechain that will promote a new form of distributed governance, involving key HPC, big data and cloud industry leaders.

iExec is using its ERC20-compliant token to provide standard and secure payments. RLC[5] which stands for "Run on Lots of Computers", can be securely and easily stored, transferred, traded, divided and used to make payments. This widely adopted cryptocurrency (87 million RLC are currently in circulation) is used to access all iExec's services.

From a research project, iExec is now a company, whose headquarters are in Lyon, France, with a subsidiary in Hong Kong.

1.4 Problematic

With the huge increase of human dependency on Information Technology for even daily activities, comes the massive consumption of electricity. In 2016, global data centers used roughly 416 terawatts (more than 90 billion kilowatt-hours for just U.S. data centers) or about 3% of the total electricity in terms of percentage, which is nearly 40% more than the entire United Kingdom. Predictions have shown that this consumption will double every four years[6].

As processing-power demanding technologies like artificial intelligence and blockchain have been appearing, the network of data centers that have sprung up in the past decade will spread. Moreover, internet-connected devices is changing the entire landscape because IoT is projected to exceed 20 billion devices by 2020. Given there are currently 10 billion devices, doubling that will require huge increases to our data center infrastructure, which will massively grow our electricity consumption, and that is just adding fuel to the fire.

There has been some trials to remedy the situation such as using other alternatives to silicon in data storage and counting on virtualisation to reduce the use of physical machines, but all of that still can not keep up against the extreme consumption demand.

1.5 Suggested solution

Although the existing trials are aiming to make datacenters greener, in this project we are taking another perspective. The idea is to push computation (or part of it) to the edge of the network instead of sending it to the cloud (fog/edge computing principles) and use solar energy to power devices that execute this computation. iExec plays a key role in this approach because the execution process will be handled by its software, in other words we are creating an iExec's worker that is completely autonomous and energy positive. The prototype of this system is a Raspberry Pi based device powered by a solar panel but the concept can be applied to a wider range of use cases.

1.6 Conclusion

After talking about the project's context and background, we presented the host company and discussed the problematic we are dealing with as well as the solution to resolve this issue, which is the subject of this project.

Chapter 2

State of The Art

2.1 Introduction

Before elaborating our solution, we discuss in this section some existing approaches and their limitations. We define, after, multiple technologies and concepts we use in this project.

2.2 Critical overview of the existing

There has not been other trials that use the same approach as this project, but there are some other attempts to remedy the problem of energy consumption in the IT field. Those solution suggest that datacenters should be powered using renewable energy sources.

2.2.1 Solar-Powered Data Centers

Emerson Network Power, i/o Data Centers and AISO are among the companies that have implemented on-site solar solutions. Emerson Network Power has installed a 7,800 square foot solar array on the roof of its new St. Louis data center[7].



Fig. 2.1 Emerson Network Power solar array

Also, QTS Princeton data center campus installed more than 57,000 solar panels in New Jersey countryside. The installation generates up to 14.1 megawatts of power. That's more than enough to supply the daytime energy needs of the data center[8].



Fig. 2.2 QTS Princeton datacenter photovoltaic solar panels

2.2.2 Wind-Powered Data Centers

Only a handful of companies have implemented wind turbines in working data centers. A small ISP and hosting company in Woodstock, Illinois, Other World Computing (OWC) may be the first data center operator in the U.S. to power its facility entirely with wind power from an on-site turbine. In 2009, OWC began using a 131-foot-tall wind turbine to provide all the electric power for its building, which includes the company's headquarters and a data center supporting its web hosting and ISP services[9].



Fig. 2.3 The 131-foot tall wind turbine at Other World Computing in Woodstock, Illinois

2.2.3 Geothermal Data Centers

A number of big companies like Google and Facebook are using geothermal cooling in their data centers. They have built them in icy areas like Sweden and Finland taking advantage of the cold air in those regions.

2.2.4 Limitations of previous solutions

Even though most of these approaches exist in real world, they still have problems reaching the scale required to successfully support the power requirements of big data centers. Solar power hasn't been widely used in data centers because it takes a very large installation of photovoltaic (PV) solar panels to produce even a fraction of the energy required by most data centers. In addition to that, the cost of deploying such solutions is still exorbitant.

Furthermore, working on the datacenter itself does not take advantage of the vast network of connected devices all over the world and keeps these important resources idle and not effectively used.

2.3 State of the art

In this section, we explain some key concepts and technologies we used during this project. Some of them are distributed systems related and some others describe some blockchain notion.

- **Ethereum blockchain:**



Fig. 2.4 Ethereum blockchain

Ethereum[10] was initially described by Vitalik Buterin[11] in late 2013 as a result of his research and work in the Bitcoin community. Vitalik published the Ethereum white paper, where he describes in detail the technical design and rationale for the Ethereum protocol and smart contracts architecture. In January 2014, Ethereum was formally announced by Vitalik at the The North American Bitcoin Conference in Miami, Florida, USA.

It is an open Blockchain platform that lets anyone build and use decentralized applications. Like Bitcoin, no one controls or owns Ethereum – it is an open-source project built by the community around the world. But unlike the Bitcoin protocol, Ethereum was designed to be adaptable and flexible. It is easy to create new applications on the Ethereum platform.

Ethereum is a programmable Blockchain. Rather than giving users a set of pre-defined operations (e.g. Bitcoin transactions), Ethereum allows users to create their own operations of any complexity they wish. In this way, it serves as a platform for many different types of decentralized applications, including but not limited to cryptocurrencies.

So basically, ethereum allows exchanging ether which makes it behave like Bitcoin does. However, ethereum allows anybody to write any piece of code (Smart Contract) and upload it to the Blockchain so anyone can interact with it which brings us to cite that there are two types of accounts that live within the ethereum Blockchain: externally owned accounts which are controlled by private keys and contract accounts which are controlled by a piece of code.

It is the Ethereum Virtual Machine (EVM) that executes the code of smart contracts and it is generally expensive to run applications on top of ethereum because of gas price, so iExec's platform takes this execution offchain in order to allow on-demand, secure and low-cost access to competitive computing infrastructures.

- **XtremWeb[3]:** a data driven volunteer cloud middleware used by iExec to manage the offchain computations. The middleware implements many features needed such as authentication, tasks scheduling, fault-tolerance, security (...).

It was mainly developed by Oleg Lodygensky[12] at the LAL laboratory in IN2P3 based on XtremWeb 1.8.0 by INRIA.

- **RLC[5]:**

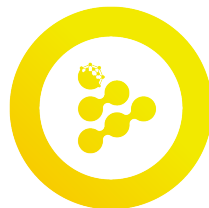


Fig. 2.5 RLC token

It is the token used to access the resources provided by the market network. It is the unique way of payment for application providers, server providers and data providers.

- **iExec Software Development Kit[13] (SDK):** this tool gives the ability to deploy any legacy applications in the iExec infrastructure, and execute them through calls to Ethereum smart contracts.
- **PoCo:** the Proof-of-Contribution protocol[4] ensures that the results computed by the workers are valid and can be trusted by the user who required them. It controls the way multiple workers achieve consensus on a computation result by leveraging several mechanisms[14]. This protocol was developed by Hadrien Croubois[15] a Ph.D. student at ENS Lyon, France.
- **IoT:** The internet of things, or IoT, is a system of interrelated computing devices, mechanical and digital machines, objects, animals or people that are provided with unique identifiers (UIDs) and the ability to transfer data over a network without requiring human-to-human or human-to-computer interaction[16].
- **Fog computing:** Fog computing, also known as fog networking or fogging, is a decentralized computing infrastructure in which data, compute, storage and applications are distributed in the most logical, efficient place between the data source and the cloud. Fog computing essentially extends cloud computing and services to the edge of the network, bringing the advantages and power of the cloud closer to where data is created and acted upon[17].
- **Edge computing:** Edge computing in IT is defined as the deployment of data-handling activities or other network operations away from centralized and always-connected network segments, and toward individual sources of data capture, such as endpoints like laptops, tablets, smartphones or IoT devices[18].

2.4 Conclusion

iExec ecosystem provides an extremely important platform and set of tools that serves perfectly the purpose of this project. Implementing the Energy Positive Worker would make use of the architecture and the business model to find the incentive to push this project toward a bright future.

Chapter 3

Requirements Analysis

3.1 Introduction

This chapter is about the overall view of the project analysis and requirements. We start with analysing the global system and explaining its basic architecture then we illustrate that with a general sequence diagrams. In a second section we specify the project requirements and give a brief explanation of each one of them.

3.2 Global System Analysis

3.2.1 iExec ecosystem

iExec creates a global and open market where computing power is traded like a commodity. It uses the blockchain as the laying underground so everything is certified on top of it. This guarantees multiple services such as the payments and the PoCo (...). This platform has different agents[19]:

- **The user** who issues computing resources to run decentralized applications. He is identified by his wallet and uses RLC to pay for the rented resources.
- **The worker** or resource provider who lends his machine power and monetize it. The worker builds his reputation according to executions he does.
- **The scheduler** manages the match-making of the demand and supply in the marketplace, does the necessary controls during the different phases of the process and finalizes the payments as well as sending results back to users.
- **The application developer** designs applications running on the Ethereum[10] blockchain.

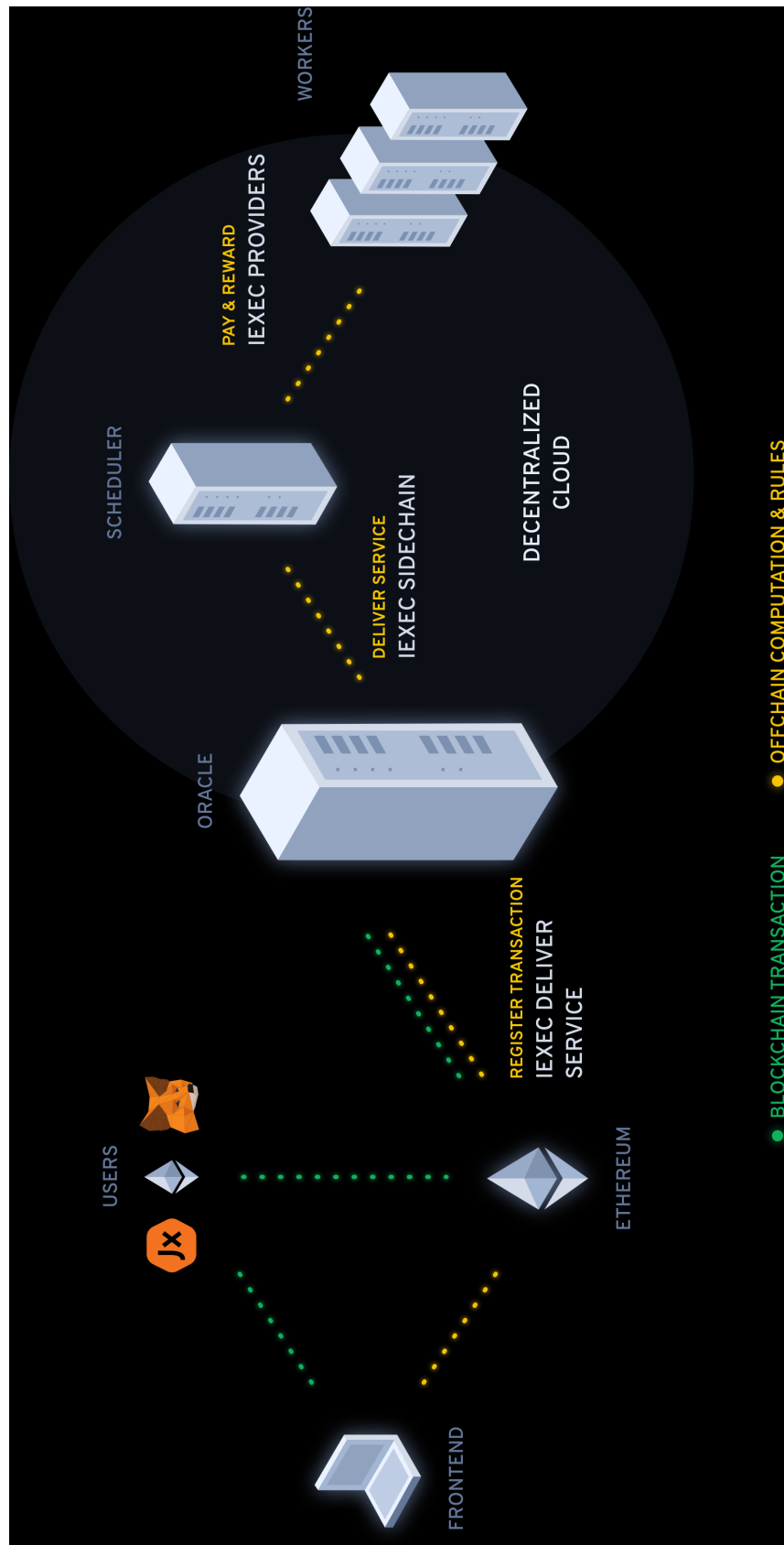


Fig. 3.1 iExec decentralized cloud architecture

The following figure[20] presents the iExec ecosystem and highlights the different parts of it including the marketplace and the worker pools.

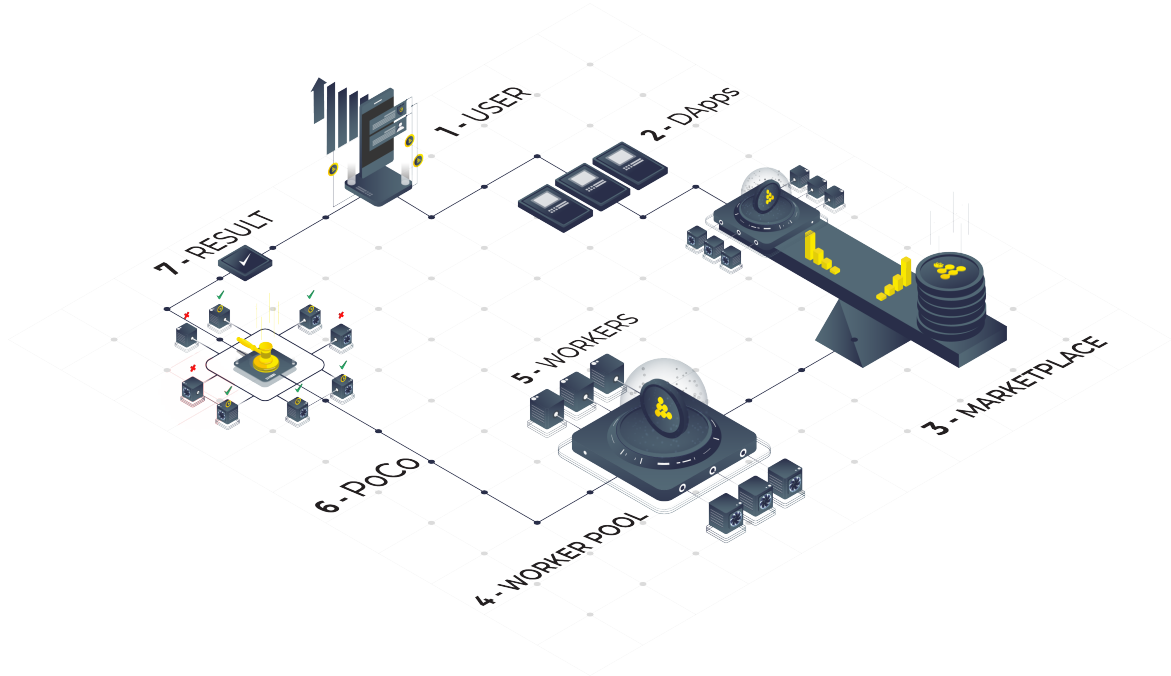


Fig. 3.2 iExec marketplace

3.2.2 General Sequence Diagram

The worker is one of three components of the global architecture, in addition to the scheduler and the user. The user starts the process by issuing an order (sending a job to be executed). The scheduler does the verification of the funds for all parts. It checks the accounts of the user and the worker and insures that they have enough RLC, delegates the job to the worker, gets the result, runs the PoCo to verify the execution, if no problem occurs it finalizes the payment and sends the result back to the user. The worker's mission starts once the job is delegated, it downloads the docker image of the application, downloads also the data needed by the execution, executes the task and uploads the result back to the scheduler. The worker has always to communicate with the scheduler to update his status (alive, working, waiting for work...) and give the scheduler the ability to track the execution process.

The sequence diagram illustrates the details of all the process:

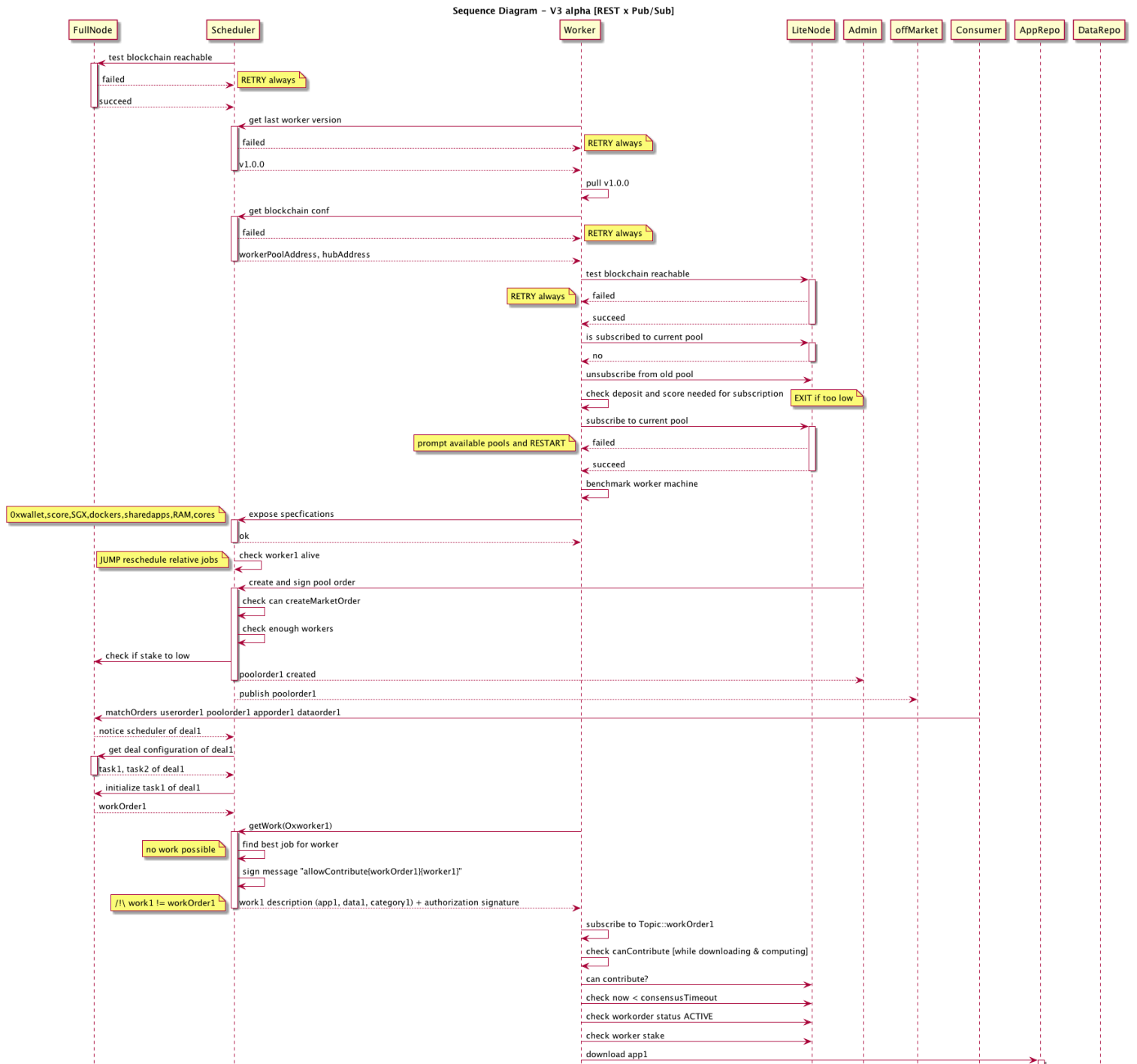


Fig. 3.3 iExec core sequence diagram - part 1

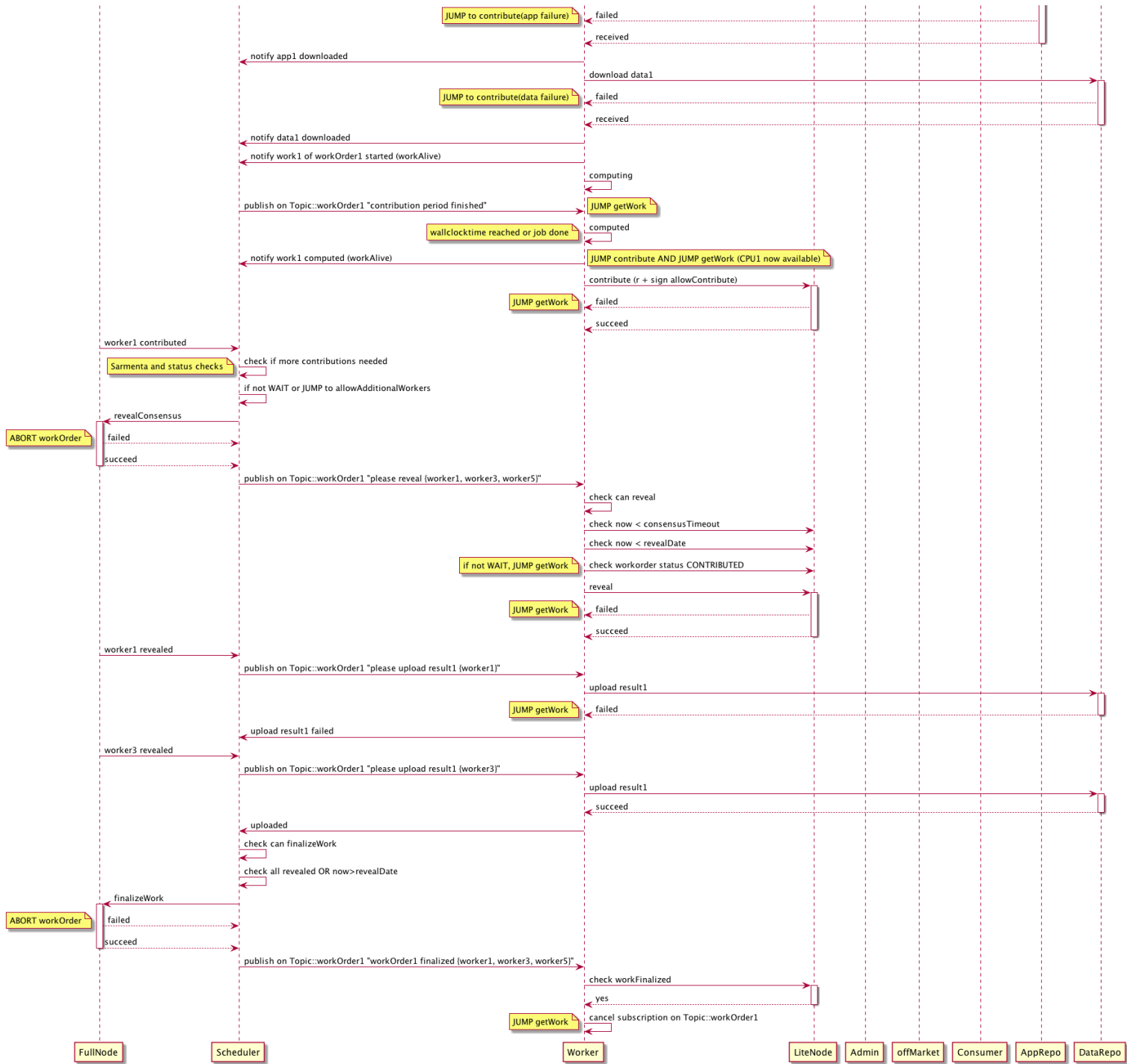


Fig. 3.4 iExec core sequence diagram - part 2

3.3 Requirements

3.3.1 Functional Requirements

The project aims to provide a certain number of functionalities, some of them are absolutely mandatory and some others depend on the use case. The main features are listed below.

- **Execute jobs sent by the middleware:** The worker is a part of iExec's infrastructure, so it should execute any type of job sent by XtremWeb middleware, the source of these tasks is blockchain dapps that need computation resources. The only constraint that those tasks should respect is the specs limitation of the device, Raspberry Pi cards for instance. It's the responsibility of the dapps developer and the task sender to deal with this constraint.
- **Ability to prove execution:** The proof-of-contribution[4] is the protocol that verifies the execution of tasks, so the worker should be able to contribute to the process of the verification. In order to achieve that, the worker should have an identity represented by its blockchain wallet and its account in the iExec ecosystem.
- **Accomplish the mission of an IoT device:** The worker is always available to execute tasks sent by users, but it can also be considered as an IoT device that processes/sends data collected by its sensors. For instance, the worker can be a surveillance camera and analyse the video stream to detect motions or a weather station that collects informations about the temperature, the humidity, the wind speed ...etc.
- **Ability to manage a cluster of workers:** The project is meant to be applicable on a large number of workers, so it is hard to manage them separately, that is why it is necessary to have a way to manage them easily as a cluster.

3.3.2 Non-Functional Requirements

The project emphasizes some extremely important non-functional requirements. Those key specifications should be respected in order to maintain the spirit of the project.

- **Positive energy:** One of the main features of this project is to eliminate the energy cost and design a system that is totally green and nature friendly. The worker is powered by solar energy and would produce more energy than it consumes. The electricity is saved in a battery to avoid climate's change effects. The objective is to charge the battery with enough electricity for two days.

- **Support ARM architecture:** Who says IoT says ARM[21] because this architecture is the omnipresent architecture in the IoT ecosystem, so the project has absolutely to support it. The challenge would be to build binaries for ARM using non-ARM hardware.

3.4 Conclusion

After understanding the global architecture of the iExec environment and defining the roles of the different components, we identified the functional and non-functional requirements of the project. Those requirements would lead us to the next part where we design our system and detail its technical specifications.

Chapter 4

Design & technical specifications

4.1 Introduction

Once the system requirements are identified, the next step would be to prepare the conceptual representation of the project. This part focuses just on the worker, going all the way from its mission and details of communication with the scheduler to the energy consumption part.

4.2 Desing and architecture

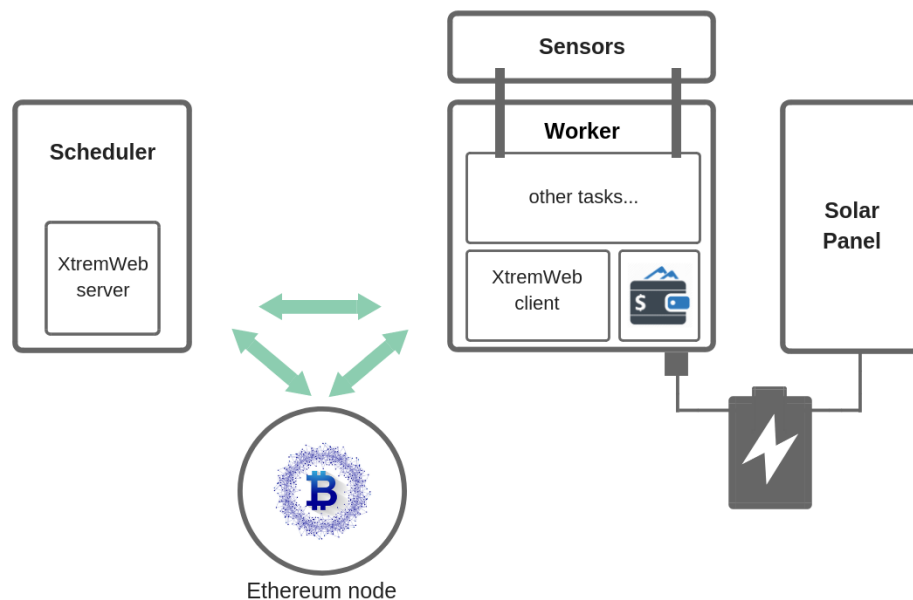


Fig. 4.1 xxxxxxxxxxxx

4.2.1 xtremweb

4.2.2 blockchain

4.2.3 energy

4.2.4 IoT

talk about the diagram

blockchain: wallet using sdk → ansible to verify

xtremweb: add arm support

cluster → swarm

Energy: solar panel, battery, voltage, watts, hours board consumption → tests

challenges: x-compiling → Qemu, optimize docker image size → cache best practices,

specs → heap size

develop dapps: names

docker(optimisation) java shell

4.3 Energy

4.4 Geth node

4.5 Use cases

4.5.1 Surveillance Camera

4.5.2 Weather Station

4.6 Technical Specifications

4.6.1 specs of the board

4.6.2 automate configuration and deployment

4.6.3 develop dapps (use cases)

4.6.4 deployment of all the environment

4.6.5 cross-compiling

4.6.6 cluster

4.7 Conclusion

Chapter 5

Implementation

5.1 Introduction

- intro

5.2 Work environment

add arm support to the scheduler

5.2.1 Hardware environment

5.2.2 Software environment

5.2.3 Development Technologies

5.2.4 Documentation

5.2.5 cluster + automation

5.3 Illustration

5.4 Conclusion

General Conclusion And Perspectives

References

- [1] Tom Bawden. Global warming: Data centres to consume three times as much energy in next decade, experts warn. *Independent*, January 2016.
<https://www.independent.co.uk/environment/global-warming-data-centres-to-consume-three-times-as-much-energy-in-next-decade-experts-warn-a6830086.html>.
Last access date: 24/08/2018 11:00.
- [2] iExec <https://www.iex.ec>.
- [3] XtremWeb <https://www.xtremweb-hep.lal.in2p3.fr/>.
- [4] PoCo https://www.iex.ec/news_categories/poco-series/.
- [5] RLC token <https://www.iex.ec/rlc-token/>.
- [6] Radoslav Danilak. Why energy is a big and rapidly growing problem for data centers. *Forbes*, December 2017.
<https://www.forbes.com/sites/forbestechcouncil/2017/12/15/why-energy-is-a-big-and-rapidly-growing-problem-for-data-centers/#682a483c5a30>.
Last access date: 24/08/2018 11:00.
- [7] Solar-Powered Data Centers <https://www.datacenterknowledge.com/solar-powered-data-centers/>.
- [8] Solar-Powered Data Centers <https://www.datacenterknowledge.com/archives/2014/10/14/super-sizing-solar-power-for-data-centers>.
- [9] Wind-Powered Data Centers <https://www.datacenterknowledge.com/wind-powered-data-centers/>.
- [10] Ethereum blockchain <https://www.ethereum.org/>.
- [11] Vitalik Buterin, Ethereum creator https://about.me/vitalik_buterin.
- [12] Oleg Lodygensky, XtremWeb main developer <https://users.lal.in2p3.fr/lodygens/>.
- [13] iExec SDK <https://github.com/iExecBlockchainComputing/iexec-sdk>.
- [14] PoCo definition <https://www.medium.com/iex-ec/poco-series-2-on-the-use-of-staking-to-prevent-attacks-2a5c700558bd>.
- [15] Hadrien Croubois, PoCo designer <https://www.hadriencroubois.com/>.

- [16] What is IoT <https://www.internetofthingsagenda.techtarget.com/definition/Internet-of-Things-IoT>.
- [17] What is Fog computing <https://www.techopedia.com/definition/32472/edge-computing>.
- [18] What is Edge computing <https://www.techopedia.com/definition/32472/edge-computing>.
- [19] iExec decentralized cloud <https://www.iex.ec/overview/>.
- [20] iExec marketplace <https://www.iex.ec/marketplace/>.
- [21] ARM architecture <https://www.arm.com>.