SeaClouds Project

D6.1 Case study extended description

Project Acronym	SeaClouds
Project Title	Seamless adaptive multi-cloud management of service-based applications
Call identifier	FP7-ICT-2012-10
Grant agreement no.	610531
Start Date	1st October 2013
Ending Date	31st March 2016

Work Package	WP6, Case Studies implementation and Validation
Deliverable code	D6.1
Deliverable Title	Case study extended description
Nature	Report
Dissemination Level	Public
Due Date:	M3
Submission Date:	31st December 2013
Lead Partner	Atos
Authors	Francesco D'Andria (ATOS), Christian Tismer (NURO)
Reviewers	Javier Cubo (UMA), Jacopo Soldani (UPI)
Version:	1.0
Status	Final

Table of Contents

List of Figures

List of Tables

Executive Summary

This document, the Deliverable 6.1 (in the following referred to as D6.1), reports on the status and progress of both tasks 6.1 and 6.2, the in-depth as-is analysis of the two SeaClouds industrial use cases and the definition of the scope of the corresponding SeaClouds case studies in terms of cloud infrastructures interoperability and data portability.

For each one of the individual case studies (SoftCare application from ATOS and the Cloud Gaming application from NURO) that are going to be piloted, this document defines the organization of realistic scenarios and identifies the specific SeaClouds technologies that each case study will demonstrate, validate and/or exercise.

D6.1 describes the capabilities to test both applications and makes a first attempt to quantify the expected benefits and to define their assessment measures and metrics.

This report will be continued and extended from the technical point of view in the document D6.2 (month 12).

The structure of D6.1 reflects the organization of the two individual case studies:

- Section 1 and 2 provide a short introduction of the main objectives of the deliverable, positioning the two case studies
- Section 3 focuses on the SoftCare scenario from ATOS
- Section 4 introduces the Cloud Gaming application from NURO
- Section 5 concludes the document.

1. Introduction

The results of SeaClouds aim to benefit both Cloud-based application developers and Cloud PaaS providers.

Cloud-based application developers will be able to deploy, and manage their applications in Cloud offerings, both within the same as well as across different Cloud providers without setting data at risk, alleviating the primary barrier for platform adoption, viz. “the vendor lock-in”.

Moreover, they would be able to compare Cloud offerings with different characteristics and formats (data models) and choose the one that best matches their computational needs taking into account the requirements of the business application.

SeaClouds also opens up the Cloud market to small-medium European Cloud providers, and strengthen their market position by leveraging accessibility (initial matching and deployment) and sustainability (application orchestration/portability) to its Cloud-based application developer customers as a competitive advantage.

The two pilots that will be described in the current deliverable will take advantage of the SeaClouds functionality from the developer’s perspective.

The endmost goal is the deployment, the orchestration including the management and migration of complex applications modules above different PaaS providers.

1.2 Glossary of Acronyms

All deliverables will include a glossary of Acronyms of terms used within the document.

Table 1: Acronyms of Project

Acronym	Definition
GDS	Geriatric Depression Scale
MMSE	Mini Mental State Examination
PaaS	Platform as a Service
PoC	Prove of concept
QoS	Quality of Service
SLA	Service Level Agreement

1. Pilots' Positioning

1.1. SeaClouds at a glance

In order to perform the proposed vision, SeaClouds introduce a user-centric architecture (figure 1) that takes care of different aspects of the cloud development life-cycle, such an open, generic and interoperable foundation to orchestrate parts of cloud-based applications.

It provides services to monitor and manage the underlying providers (both public and private clouds) and thus leverages service level agreement policies in order to guarantee the required performance and QoS on multi-cloud environments.

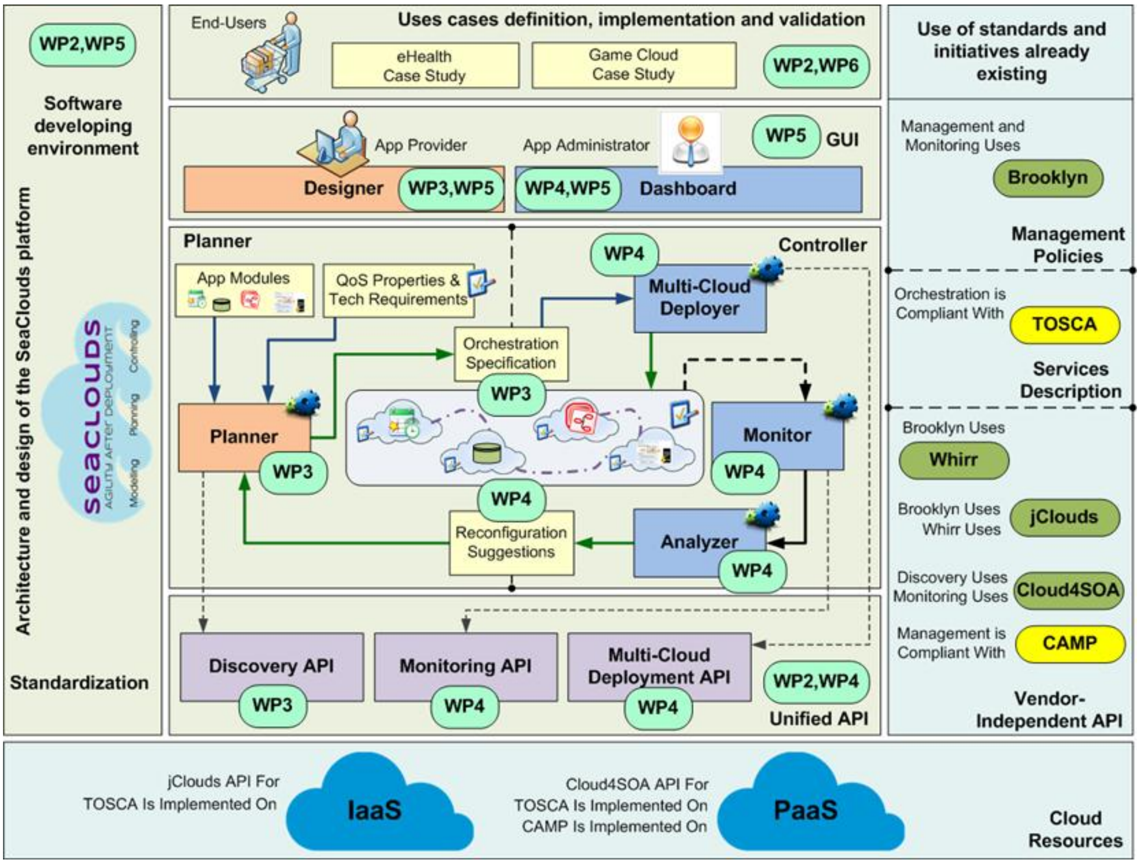

Figure 1 SeaClouds at a glance

2. SoftCare application case study

ATOS aims at providing an e-health and social care case study by developing a full-featured business intelligence solution for assessing disease affecting elderly people.

The objective is to implement a Cloud-based social support network application that provides the following features:

- Supporting maintaining health and functional capability, through the risk assessment and the early detection of deterioration symptoms of the patients and distress signs of their carers;
- Providing the means for the self-care and the self-management of chronic conditions, through the development social networking as well as educational tools;
- Enhancing the home-as-care environment through the provision of user-friendly ICT tools for frequent, unobtrusive monitoring;
- Facilities for high-quality interaction between doctor and patients;
- Added-value features to create and maintain an easy-to-use web-based social network for individual elderly persons, to stimulate elderly person and their careers.

2.1. Current Organization of the Application

2.1.1. System Architecture

The Figure 2 depicts the current organization of the ATOS Case Study, highlighting its architecture and its main entities, which aims to enhance patient and carer safety through the risk assessment and prediction mechanisms supported by the smart devices. The overall system comprises of three main subsystems:

- The web-based environment, which incorporates the social networking utilities for the communication between patients with similar conditions, the communication with carers for sharing similar experiences, and the communication between the aforementioned and specialized medical personnel. It also incorporates educational tools (e.g. electronic library) to facilitate personal motivation and enhance the notion of self-management. An interactive platform for music therapy to improve behavioural and psychological symptoms will also be included.
- The monitoring tools, which involve the development of smart devices for the conduction of remote psychometric tests (Mini Mental State Examination - MMSE or Geriatric Depression Scale – GDS), along with video-conferencing utilities for the provision of a more thorough clinical image of the patient to the medical expert. Patient monitoring will include blood pressure measurements in order to avoid hypertension that is a relevant risk factor for cardiovascular events that could significantly worsen the progression of dementia, as well as body weight measurements to avoid malnutrition, a common problem with dementia progression due to poor appetite or food refusal, in order to allow an early intervention with appropriate alternative strategies (liquid integrators, multivitamins, etc). An electronic recorder of physical activity will also be considered, with the double purpose to evaluate the patient's independence

and to trace the patient's position in case of wandering in public or private areas.

- The risk assessment and analysis tools, incorporating data mining capabilities, retrieving information from diverse resources, including psychometric tests, electronic health records and personal evaluations by medical experts.

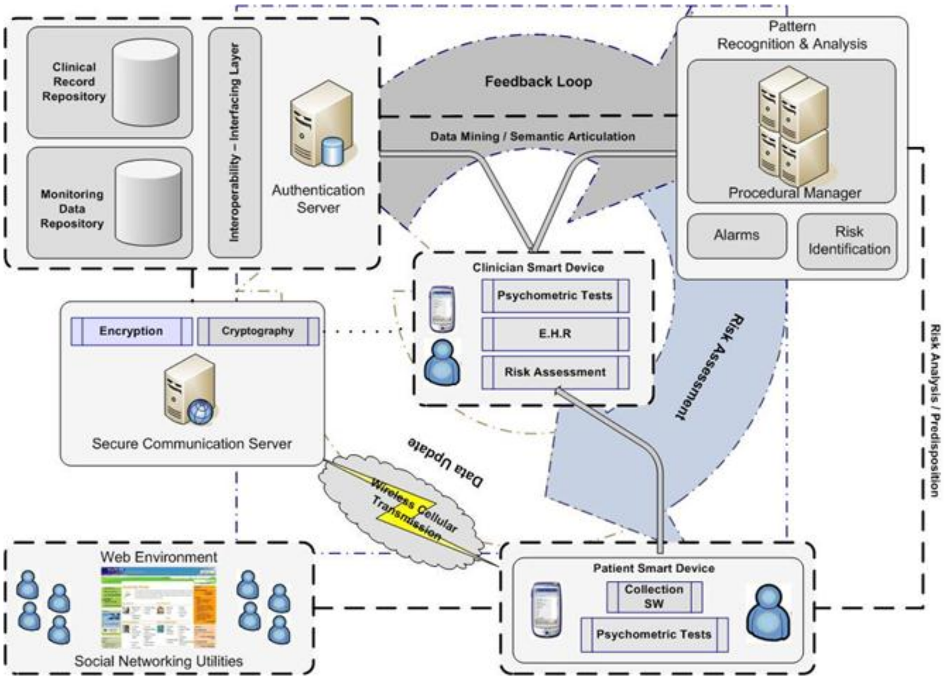

Figure 2 Current Organization of the ATOS Case Study

2.1.2. Current Infrastructure and Resources

Currently, the ATOS case study is still a Prove-of-Concept (PoC).

In the scope of SeaClouds project, ATOS will enhance the application by providing new features and removing bugs.

In terms of infrastructures and resource used, the case study, currently is a monolithic application (code written in one large program, and not modular) that has been installed/tested in the ATOS data centre.

The resources used by the case study are naturally dependent on the growth of the number of features provided, as well as on the number of stakeholders using the service.

2.1.3. Drawbacks, Bottlenecks and Scalability Problems

It is not trivial to accurately estimate the resources required. It is naturally necessary to increase the resources with the growing number of features provided and requests; furthermore, the features themselves may be of varied nature, each with its own resource usage.

To solve this problem, ATOS foreseen that the case study platform will need to scale infrastructures dynamically when more resources are required. A cloud approach is ideal in this situation.

2.2. SeaClouds Atos Case Study: "SoftCare Application"

2.2.1. SoftCare application architecture

The ATOS aims at implementing the SoftCare application as a Cloud-Enabled platform (the monolithic application will be decomposed into a set of web-based services ready to be *cloudified*) built around the needs of elderly people proving the following tools:

- Peer to Peer tools to maintain a close link between public/private health system and elderly people affected by a degenerative disease. This link will be useful to make a daily monitoring of the status of the patient
- Educational tools (e.g. electronic library) to facilitate personal motivation and enhance the notion of self-management
- An interactive platform for music therapy to improve behavioural and psychological symptoms
- Video-conferencing utilities for the provision of a more thorough clinical image of the patient to the medical expert
- Platform that integrate Social Networks. Elderly people are keen to maintain contact with the different generations of their family. They also want to contribute to mainstream debate and discussion and not be boxed into communities with people of a similar age to themselves. Many older people have already invested significant time in building contact lists and relationships within the major social networks.

2.2.2. Expected Benefits from the ATOS SoftCare Application

From re-engineering our health application ready to be ported on the cloud we expect a smooth increase (or decrease) of computational and storage resource usage with increasing (or decreasing) system load and storage demand respectively.

In depth these are the main benefits:

- The vision of SeaClouds is to alleviate the vendor lock-in problem, and manage complex applications on multi-clouds environment which gives the opportunity to choose and change competitive providers according to the changing resource requirements during the lifetime of a service.
- SeaClouds allows also running business applications on robust and reliable infrastructure with a flexible capacity. Through SeaClouds the Atos SoftCare application will run on fully managed cloud infrastructure running in an enterprise-class data center with secure, reliable cloud services to support and deliver hosted software solutions.
- The cloud also offers the ability to deliver flexible resources to accommodate fluctuations in customers' production workloads as well as on-demand capacity for training, demonstration, proof-of-concept or test/development.

3. Cloud Gaming Case Study

Nurogames has many games in the market and in deployment state, some of them based on Game Server Engines developed by Nurogames. The Game Server is responsible for data consistency and cheating protection.

NURO's cloud game case study is based on a Nurogames Engine (e.g. WebRpg or IsoGame) in the following called SeaCloudsGame. These engines are used for games that are online or will be launched soon.

NURO will modify its monolithic server approach to a cloud ready version. Also a testing system for different scenarios will be developed.

3.1. Current Organization of the Application

3.1.1. System Architecture

The SeaCloudsGame consists of a Game Client and a Game Server.

With this study the Game Client is not the focus. Normally Players are using the client to play the game. It is an app for android, iOS or Windows.

Client side for this case study will be based on a test system using scripts to simulate client requests for boom and burst scenarios.

The case study is centred on the Game Server, which is based on webserver technology. All Game Client actions are performed by HTTPS requests to the Game Server. As answer the server sends the result of the action, additional data and events to the Client.

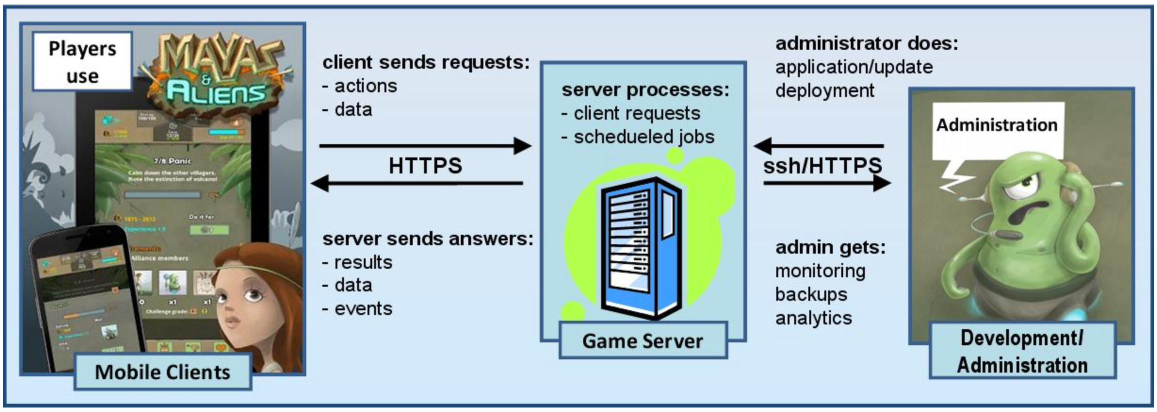

Figure 3 Client, server, administration interaction

The Game Server itself consists on a webserver with PHP, a MySQL database and a scheduler. The Game Server's software is deployed and updated by a SVN repository, containing software, database creation and modification scripts.

There is also a webspace for static data, a separate database for log and analytics data and a backup space.

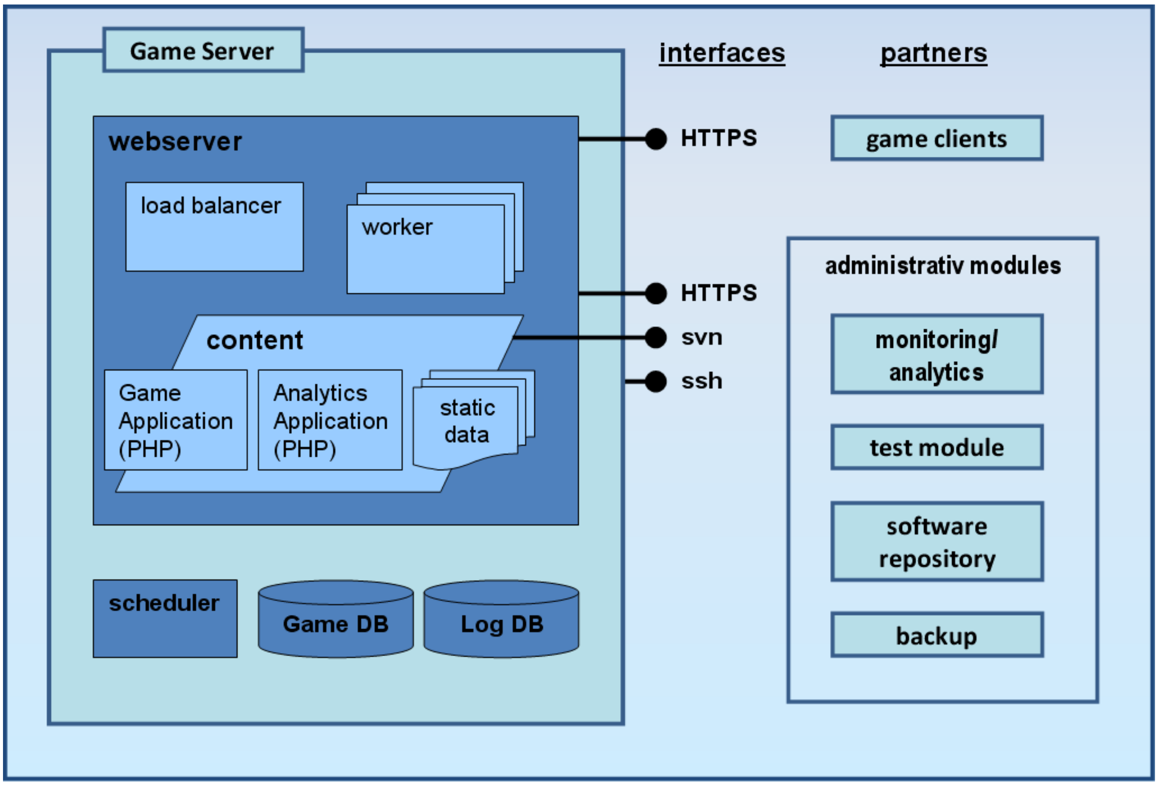

Figure 4 Game Server inside

3.1.2. Current Infrastructure and Resoruces

Deployment and Backend:

The Game Clients of real live games are deployed by Google Play Store for Android and iTunes for Apple iOS. Windows versions are only used internally for testing and demonstration purpose. Actual Game Server software and database updates are deployed manual by a ssh session using svn and scripts. Backup and log export are invoked manually, automatization is under development.

Analytics are done by a web based user interface and data exports.

Components/Software

NURO's actual game installations are XEN based virtual machines (VM) on a dedicated server. One VM can also host different Game Server Installations.

Software:

Operating system: Linux Debian (5 - 7 tested)

Webserver: apache (2.2.x) with PHP (5.2.6, 5.3.3, 5.4.4 tested)

Database: MySQL (5.0.51, 5.1.66 and 5.5.31 tested)

Access: OpenSSH, ssh and rsync

Scheduler: crontab and wget as https client

Deployment: svn console client and shell scripts

Security: several security tools for DoS protection, intrusion prevention and detection.

Hardware:

CPU: Intel Xeon E3-1245 @ 3.30GHz, 4 Cores, 8 Threads

RAM: 16GB

HDD: 2x 3TB Enterprise (Software RAID1, LVM)

VM:

Resource	**Minimum**	**Actual Installation**	**Comment**
Cores	2	up to 8	all VM's share all cores
RAM	600MB	2GB	unused memory is used for buffers and caches
Drivespace	20GB (growing)	100GB	This is the space for operating system and game data without log and analytics data.
Logspace	20GB (growing)	included above	The used log space is depending on log and analytics level.

One apache worker process consumes 20MB.

3.1.3. Drawbacks, Bottlenecks and Scalability Problems

3.1.3.1. Bottlenecks

The database is our primary bottleneck.

With 6 thousand players the throughput is 30 requests per second.

With 130 thousand players the throughput is 2.5 requests per second.

During phases of 2.5 requests per second there is a CPU usage of 50%-80%. Database optimization is in progress.

Secondary bottleneck is speed of hard drive. Depending on using XEN there is a limit of round about 80 MB/s per VM. Writing database dumps is up to 4 times quicker with using pigz (a multi-threaded gzip) before writing. XEN optimization is in progress.

3.1.3.2. Drawbacks

The administrative effort is relative high for managing dedicated servers and VM's.

With respect to avoid downtimes, operating system maintenance is delayed.

With a small team 24x7x365 administration is a huge effort binding human resources.

3.1.3.3. Scalability Problem

The usage intensity of a game server is very volatile. It is affected by many factors like planned campaigns, updates, external promotions and unsteady influences as trends, acceptance, holidays, lunch time, events and also weather.

There are peaks and boom phases that last several days and in an intra-day focus, rhythmical daily boom and burst phases can be discovered.

To explain this behaviour, here are some analytics from two installations of real life game servers. Both games are hosted on VM's sharing the resources of the same dedicated machine as described above.

Multiple day booms

Figure 5 shows multiple day boom phases by sum of client installations and requests per day of Game A and B. Days are normalized to days since game launch.

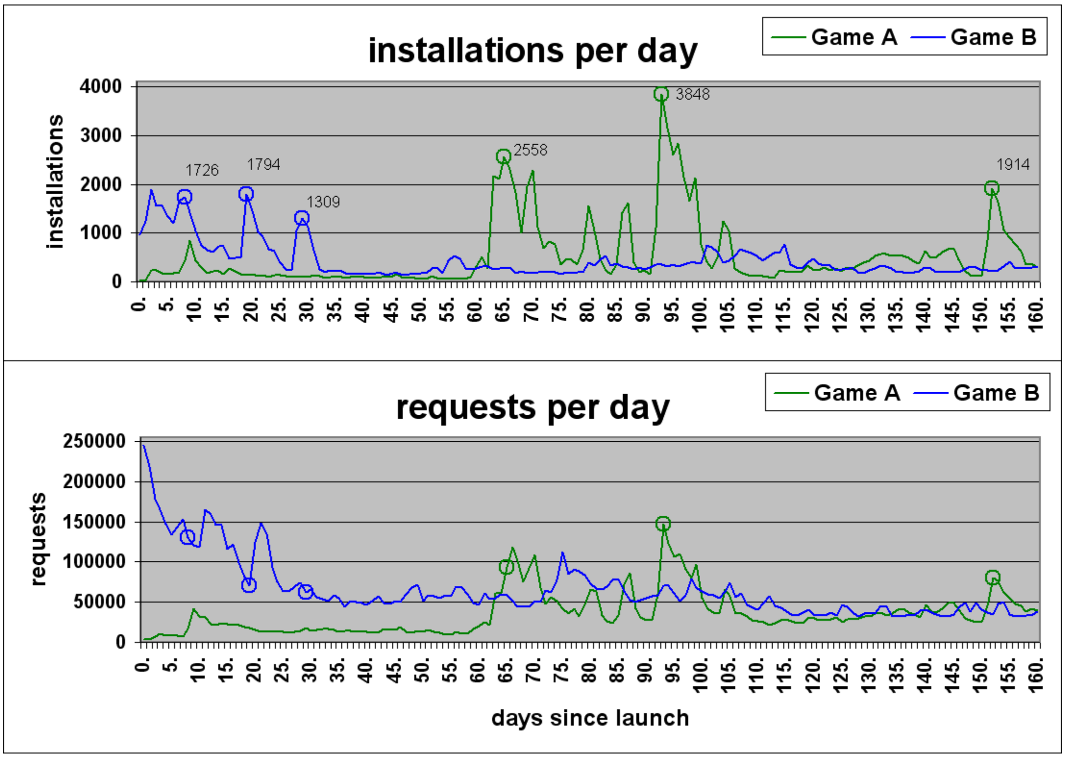

Figure 5: Boom phases

Peaks and boom phases depend on planned marketing campaigns (Game A day 60-70 and 92-100, Game B day 19 and 28-31) and unexpected features by the app stores or reviews in press (Game B day 0-10).

These promotions trigger installations (upper chart) with an effect on number of requests the next days (chart below). Game B illustrates this significantly by an installation peak and request low on day 19 (second mark) followed by a request boom, bursting some days later.

Client updates also cause an increase in requests without additional installations as seen on day 75 of Game B (blue line).

This chart demonstrates a higher need of resources during promoted periods compared to an average period.

Moreover, there is a growing acceptance trend thus there is an expected growth of average requests in the future.

Intra-day usage

Figure 6 shows the request of both games by hour and minute on an average day. The grey line is the sum of requests in both games.

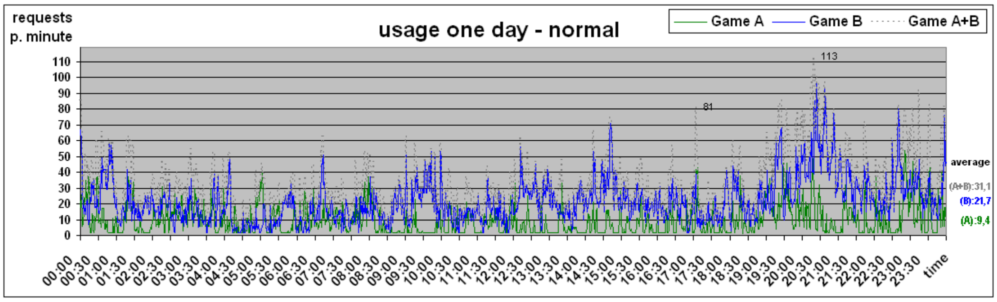

Figure 6: Average day, two games sharing one server

The actual limit of the dedicated server is 2.5 requests per second or 150 requests per minute. At 20:20 there was a peak of 113 requests in the sum of Game A+B. The average of (A+B) is 31.1 requests per minute. On a normal day the server performs well but many resources are unused.

The next Figure displays a day on a boom phase in Game B.

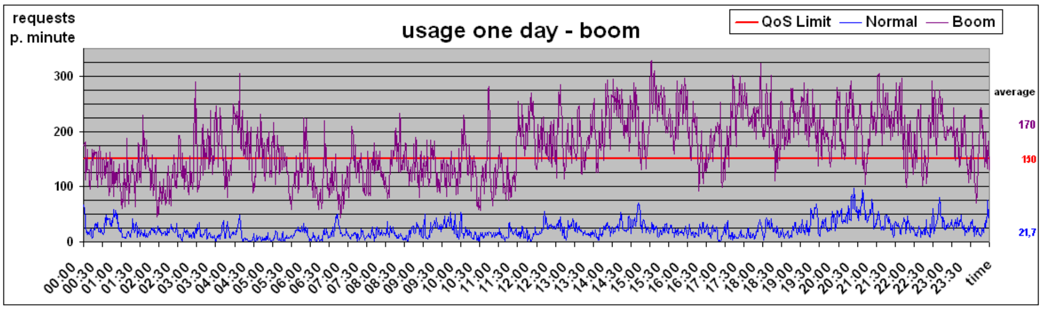

Figure 7: Average vs. boom day of Game B

The purple line shows the requests per minute in a boom situation. The blue line is the average day from Figure 6. The red line is the QoS limit of 150. On that day Game B was almost over the QoS limit of the server and influenced the responsiveness off all games on that machine.

Daily habits

Figure 8 is visualizing the players' habits of Game A and B. It illustrates the periodical resources need throughout a day, depending on requests per minute as indicator. The total itself is just a sum over lifetime and should not be interpreted.

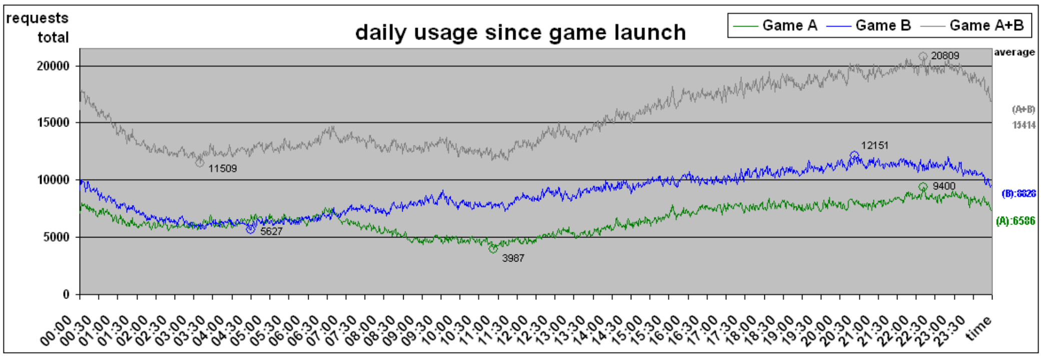

Figure 8: Player habits - daily booms and burst

These significant curves depend on the time zones and regions where the games are favoured. Game A players mainly live in Mexico and USA, some in Europe. Most Game B players are from Europe and Asia. Regular troughs and crests are caused by regional bed, break or lunch time, and also when school or work is over.

Waste of Resources

The next table estimates the waste of resources in the average and minimum usage, if the system is equipped for the maximum boom situation. The data is based on Figure 7.

Table 2: Estimated waste of resources

	Boom vs Average
Max	329
Min	2
Avg	22
Min Usage	1%
Waste on Min	**99%**
Avg Usage	7%
Waste on Avg	**93%**

Interpretation

A Servers prepared for boom phases will waste 93% of resources on average usage and 99% in minimal periods.

Sharing resources with other games, acquiring gamers in special regions and cleverly timed marketing can diminish this problem but that does not solve it.

The actual usage of our live game server is most of the time lower than 0.5 requests per second, thus the server is almost idle.

On peak and boom phases there are significantly more than 2.5 requests per second and the responsiveness of the server will get worse. If a game is accepted by the players, the number of requests per second grow with the number of simultaneously active players.

In this desired situation a dedicated server is a dead end. If more resources are needed to keep the QoS than the dedicated server offers, the game server installation must be moved to a stronger machine or it must be modularized and distributed on several machines.

Nurogames expects that modularization and using cloud based components will prove to be the contemporary, flexible and future save solution to the scalability problems.

3.2. SeaClouds NURO Case Study: "Cloud Gaming Application"

3.2.1. Cloud Enabled case study architecture

The Nurogames SeaCloudsGame case study will implement the Game Server for a cloud based environment.

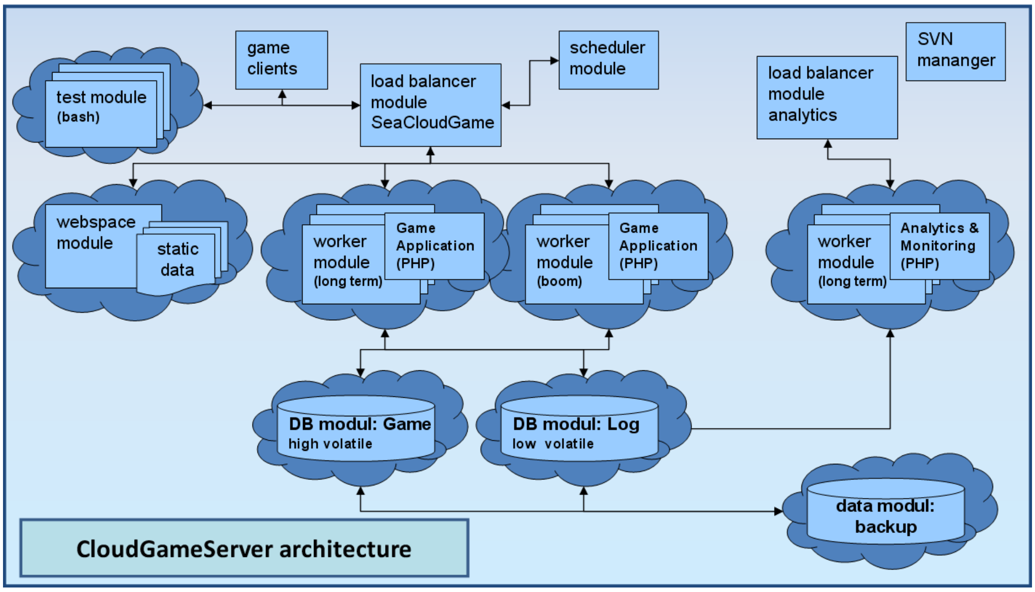

Figure 9: Cloud Game Server Architecture

The architecture of Nuro enabled-cloud gaming platform will have the following modules and controllers:

Game Client (test scripts and mobile installations)

The Client App is having a direct communication with the final user. Translating every user action into requests to the Game Server, which will return the corresponding response, having this way a complete data synchronization for all users.

For this case study client will be simulated. Some Game Client installations will be used to evaluate the look and feel of game responsiveness.

Requirements of Nuros SeaCloudsGame modules

Assumption on security

Security administration (hardening) of all cloud components is expected to be done by the cloud provider. The component should communicate by encrypted connections. This could be done by HTTPS, ssh, VPN or other common security methods.

Fails are expected to be performed automatically by the system or the cloud providers.

3.2.2. Expected Benefits from the NURO Cloud Gaming Application

The goal of the Nurogames SeaCloudsGame case study is to keep the function of the current version, taking advantage of cloud flexibility and helpful tools to optimize and to save administrative time and effort. Reliability, stability and security of the game server application should be improved by the SeaClouds solution. Saving resources in burst scenarios and a flexible pricing of real used resources will be possible.

Stable Quality of Service (QoS)

The flexible use of resources should guarantee the QoS on boom phases. Maintenance time should be optimized on rollout of new versions.

Administration and maintenance

Designer and administrator can focus on the application. They just have to configure the modules according to their needs and invoke deployment, installation and updates of the application.

Administration of hardware and operating system is to be done by the SeaCluds system or the cloud provider.

Scalability over time

If the game is accepted and running for a long time, increasing number of players will cause the server applications to consume more resources. A cloud solution should make it easy to scale the system to the needs. More over on boom phases the QoS limits will be met and on burst phases wasting of resources will be avoided.

Flexibility and control on payment

SeaClouds could considerably reduce costs, only paying for resources which are actually needed. Consuming and releasing resources depends on the needs to satisfy the QoS rules. We expect to set also cost limits to avoid exceeding budgets.

Risks

There is a risk connected with automatic resource consumption. For example if it is invoked by a DoS attack or a system failure. Other possible risks have to be analysed.

4. Conclusion

The document D6.1 shows a first analysis of the two SeaClouds case studies which have quite different requirements for deployment and management on cloud platforms.

The e-health Atos case study is still in an early stage of development specially targeted for deployment in the cloud. However, at this stage, not all requirements are clear yet.

The Cloud Gaming case study leverage on a more mature application, that Nuro is looking forward to port to the Clouds.

The forthcoming report D6.2 “Case Study test-beds and key features mapping” will provide a further and detailed description and key features mapping of the two case studies detailing as well the test-beds used to run the two scenarios.