Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Performance evaluation for the design of a hybrid cloud based distance synchronous and asynchronous learning architecture

Enrico Barbierato a, Lelio Campanile b, Marco Gribaudo c, Mauro Iacono b,*
Michele Mastroianni b, Stefania Nacchia b

a Università Cattolica del Sacro Cuore, Facoltà di Scienze Matematiche, Fisiche e Naturali, Brescia, Italy
b Dipartimento di Matematica e Fisica, Università degli Studi della Campania “Luigi Vanvitelli”, Caserta, Italy
c Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, Milano, Italy

A R T I C L E I N F O

Keywords:
Hybrid cloud
Performance evaluation
Distance learning
VARFIMA
Moving average
Simulation

A B S T R A C T

The COVID-19 emergency suddenly obliged schools and universities around the world to deliver on-line lectures and services. While the urgency of response resulted in a fast and massive adoption of standard, public on-line platforms, generally owned by big players in the digital services market, this does not sufficiently take into account privacy-related and security-related issues and potential legal problems about the legitimate exploitation of the intellectual rights about contents. However, the experience brought to attention a vast set of issues, which have been addressed by implementing these services by means of private platforms.

This work presents a modeling and evaluation framework, defined on a set of high-level, management-oriented parameters and based on a Vectorial Auto Regressive Fractional (Integrated) Moving Average based approach, to support the design of distance learning architectures. The purpose of this framework is to help decision makers to evaluate the requirements and the costs of hybrid cloud technology solutions. Furthermore, it aims at providing a coarse grain reference organization integrating low-cost, long-term storage management services to implement a viable and accessible history feature for all materials. The proposed solution has been designed bearing in mind the ecosystem of Italian universities. A realistic case study has been shaped on the needs of an important, generalist, polycentric Italian university, where some of the authors of this paper work.

1. Introduction

On-line universities and training facilities are widely spread and offer distance learning services on a regular basis on the market. In many countries, the COVID-19 emergency suddenly forced traditional universities and schools to offer on-line courses, creating an unexpected workload on proper platforms, which has been matched by many software and hardware vendors. Resorting to external infrastructures for on-line lectures poses, at least in the European Union, a relevant issue related to privacy, and the exposure of an enormous volume of valuable multimedia education-related data, which undoubtedly can be processed by a set of platforms to automatically extract information about the state and practice of education in whole countries. The last point also raises an intellectual property issue.

* Corresponding author.

E-mail addresses: enrico.barbierato@unicatt.it (E. Barbierato), lelio.campanile@unicampania.it (L. Campanile), mario.gribaudo@polimi.it (M. Gribaudo), mauro.iacono@unicampania.it (M. Iacono), michele.mastroianni@unicampania.it (M. Mastroianni), stefania.nacchia@unicampania.it (S. Nacchia).

https://doi.org/10.1016/j.simpat.2021.102303

Available online 1 March 2021
1569-190X/© 2021 Elsevier B.V. All rights reserved.
While on-line universities natively use dedicated means, and ad-hoc solutions, verified and maintained with the needed planning and evaluations, this approach has not been feasible for universities and schools that had to implement on-line services in a short time and with few resources. Physical interaction is the ordinary way by which lectures are delivered according to a given schedule, while on-line universities already offer pre-recorded lectures that can be accessed individually and autonomously according to the needs of their students. As the current situation with COVID-19 in some countries suggested policy makers the need to maintain active on-line services, the acquired experience of the first months of the crisis may be leveraged for the design of solutions that avoid the critical points raised in those months. 

Firstly, attention is given to the fact that needed on-line interactions involve also underage students, so that their image is potentially exposed. Possible misconduct is not excluded due to security threats. Cyberbullying episodes have also been reported in Italy on the public platforms that allow link sharing based access privileges. Besides obvious methodological and organizational issues, issues about privacy and security have been raised and got attention of media and public opinion. Similar points had been raised about universities, which are in Italy structurally very different from schools for organization, autonomy, dimensions, devices, budget and legal powers of the management.

Secondly, as license agreements of such platforms include the right to automatically process data streams to deliver them and to carry on the service, content-related issues have been raised. The main reason is the possibility of extracting valuable market data from an information-intensive mass of contents by means of Machine Learning (ML) or Artificial Intelligence (AI) techniques included in the automatic processing activities. In this way, they represent an enormous potential business enabler empowering the owners of the platforms in an unprecedented way. Basically, they handle in a digital form, information about all the knowledge transmission procedures of many countries, including metadata about delivery modes, time, usage, reuse, and a map of users, institutions and their relationships. These data can be coupled, in some cases, with other information about users to produce marketing services, but can also be used to produce contents and education offer without any involvement, nor any explicit authorization, nor compensation of those that actually hold the intellectual property of the information, benefitting of the investments on education of national governments, teachers or private subjects.

Thirdly, on-line examinations open another critical line of issues. University exams are regarded as a public act in Italy, so that, although publicity has been ensured, the opportunity of performing them on private platforms may be questioned. Specifically, the identification of involved parties should be taken seriously, as well as the conditions in which students are in their environment, which must ensure no interference from other parties nor the possibility of plagiarism. Oral exams are far more an interesting issue, because they can be recorded and analyzed as well as lectures. The availability of massive amount of information can be used to analyze and profile questions asked by the professors, to design added value services helping students in predicting questions and answering without any study, which is a borderline behavior that may produce relevant revenues. Moreover, malicious use of public platforms, which are not controlled nor controllable by the university, can occur, to illegally help students during exams. As a result, the legal validity of exams cannot be granted.

This amount of information should be managed carefully, and should probably not transit on public platforms.

The design of suitable private platforms, such as those used by an on-line university, is a viable solution. However, the live and interactive nature of lectures poses a different problem, which is worth to be analyzed in terms of cost estimation and reference architectures, and needs to be evaluated case by case, because the needed effort to own and manage such a system is not trivial.

This paper proposes a reference model for a private cloud-based platform to support schools and universities in delivering on-line teaching and a methodology to estimate costs according to a set of teaching-oriented operational parameters. The reference model is based on hybrid cloud to minimize costs of peak usage while ensuring elasticity and the delivery of services to all users. A second advantage consists of the possibility to use external low-cost massive storage systems to provide multimedia contents from courses with a high latency access strategy. The methodology is based on a simulation approach that evaluates workload dynamics on the long, medium and short term, covering off-line contents lifetime, yearly organization of courses, semesters and exams periods and daily accesses. The simulation model exploits a Vectorial Auto Regressive Fractional (Integrated) Moving Average (a non-scalar version of ARFIMA, see [1]) approach to encompass all elements and ensure a smooth behavior of the model. The originality of this paper is twofold. On one side, it brings to attention the problems related to distance learning platforms; on the other, it presents a novel application of VARFIMA methods to this class of problems.

This article is organized as follows: following this Introduction, Section 2 presents related work, Section 3 describes the proposed organization for a hybrid cloud based solution, Section 4 describes the simulation approach adopted for the study and its results, followed by conclusions and future work.

2. Background and related work
Cloud technology has become a standard of IT solutions in industry, businesses, government as well as academia. It has also significantly transformed the digital world through its effective mechanism of hosting services and performing computations. In the recent past, these technologies have produced a significant impact on our ways of living, working, communicating, and even educating, as they have replaced traditional libraries with digital ones, traditional computer laboratories with virtual environments. The latter are more attractive to academic institutions in terms of cost, scalability, readiness and availability. Due to the COVID-19 emergency, the e-learning facilities provided through cloud computing are swiftly replacing the traditional teaching and learning methods that can be further revolutionized by mitigating mobility, collaboration, and user perception [2].

Usually, e-learning systems are based on client–server architecture and web-based technology thus the majority of schools and universities tend to choose the private cloud as they can easily monitor their company IT security and reliability. These solutions
Table 1
E-learning solutions comparison.

| Research work | Main focus |
|---------------|------------|
| [4]           | Utilization of pre-scheduled resources for the hot contents and applications |
| [5]           | Introduction of a semantic knowledge base to facilitate learners in finding educational services |
| [6]           | Analysis of the effectiveness of traditional versus cloud-based E-learning model |
| [8]           | Explanation of the challenges within the context of integrating E-learning and Cloud computing |
| [7]           | Analysis of the benefits and limitations of using cloud computing in the education |
| [9]           | Presentation of a novel model for accessing the E-materials in the educational community cloud |

are feasible because they mostly rely on a relatively large-scale IT infrastructure, avoiding the use of the public cloud and issues related to security and data privacy issues.

However, there are some disadvantages that must be addressed prior to the full integration of e-Learning architecture into the academic framework:

- e-Learning systems are still weak on scalability at the infrastructure level;
- The number of computational resources is quite limited compared to more sophisticated IT companies;
- The cost of dynamically add new resources and their management might be too expansive.

These are few of the main reasons that have forced to consider, first IT companies, later most schools and universities to heavily rely on hybrid cloud model to optimize their resources and to increase their essential abilities by merging outside enterprise functions into the cloud and still controlling essential activities in the enterprise, maintaining service levels in the face of rapid workload variations, through private cloud [3].

2.1. Cloud based solutions for E-learning

Scientific literature proposes many e-learning architectures solutions involving optimal cloud resources utilization. However, to the best of our knowledge, the solutions involving hybrid cloud configurations are less numerous, being studied only recently.

In terms of resource use, scalability, availability and efficiency, the authors in [4] proposed an e-Learning platform, BlueSky embracing the cloud to enhance conventional learning systems. The core of BlueSky is focused on tracking the e-Learning systems’ run-time status, dynamically claiming resources based on clear thresholds for resource use, and managing loads based on institutional priorities.

In [5] the authors have built a semantic framework for cloud learning environments, since the lack of semantic explanations of learning services makes the task of allocating sufficient resources more difficult. This paper proposes a four-layered semantic knowledge base modeling learners, learning tools and contexts, and lexical concerns to facilitate this challenge.

It is worth noticing that most research activities dealing with e-learning based system and cloud computing, mainly focus on exploring the benefits of e-learning systems from a pedagogical point of view rather than exploring new architectural solutions for improving performance or utilization [6–9]. In Table 1 we summarize the above mentioned research work, highlighting the main aspects and features pointed out by the examined literature about cloud based e-learning systems.

2.2. Hybrid cloud

Since hybrid cloud-based solutions involve a composition of two or more clouds (private, community, or public), they are more challenging than other deployment models.

However more and more enterprises and companies are, nowadays, relying heavily on such solutions for processing highly dynamic workloads of most applications developed and deployed on a cloud infrastructure [10]. Cloud bursting is a particular example of an effective hybrid cloud-based approach [11], where the application system uses fixed resources in a private data center for the most part of its computing and bursts into a public data center. Specifically, it temporarily integrates on-demand resources when private resources are inadequate. Moreover, based on whether workload demand is known ahead of time, research for automating cloud bursting is loosely divided into two groups. In the first group, the number of tasks is planned in advance and there is a trade-off between the completion time of the tasks and the amount of resources available. Researchers have proposed ways to plan resources in an agile manner to reach time limits; this category involves high-performance computing for scientific applications as in [12,13]. The potential workload in the second group is uncertain. It is therefore important to predict future demand and to optimize the trade-off between application requirements, such as response time and throughput, and resource economics, such as overhead costs and configuration, e.g. in the case of video streaming services (such as e-learning based systems) [14]. In [15],
the authors present a cloud bursting approach focused on long-term and short-term projections of requests for a business-critical web system to classify the optimum resources of the system deployed in private and public data centers. A dedicated pool of virtual machines (VMs) is allocated to the Web system in a private data center depending on one-week forecasts. In addition, VMs are enabled in both private and public data centers based on one-hour predictions. The approach proposed enables all physical servers and VMs to be re-allocated using a software-defined networking (SDN) system.

In [16], the authors present the CloudWard Bound to partially outsource enterprise services to a hybrid cloud environment. Based on the application performance criteria as well as the privacy constraints, a selection of candidate components is selected for migration to the public cloud. Authors use a series of metrics, including the amount of workload needed, storage space, and transaction delays, to provide efficient hybrid cloud deployment. Another interesting work is discussed in [14], where a hybrid cloud computing model is proposed featuring a two-zone system architecture. The Internet-based application that must be deployed on the architecture, is divided into two naturally different components, namely the base load and flash crowd load.

The base load refers to the smaller and simpler workload encountered by the application most of its time, while flash crowd load refers to the much larger, yet transient load spikes encountered at a rare time (e.g., the 5%-percentile heavy load time).

It can be set up and operated in the (small) enterprise data center with the intention of efficient planning and high usage, while the flash crowd load platform can be supplied on demand via a cloud provider, taking full advantage of the flexible quality of the cloud infrastructure. The workload factoring service is based on a simple feature that splits the workload into two parts on (unpredictable) load spikes and ensures that the base load component stays within the volume plan and that the flash crowd load component incurs limited cache/replication requirements for the application data needed.

In [17], the authors discuss the issue of making IaaS more flexible by proposing a structure for the allocation of cloud resources to enable the use of external clouds. Under this system, the IaaS cloud has its own private cloud and can outsource its tasks to other cloud providers called external clouds (ECs) where its local resources are not adequate. Each task has a strict deadline to be met, so that the resource allocation issue can be considered a time-limited task scheduling (DCTS) one. A detailed programming formulation of the DCTS problem is developed with the goal of optimizing the income of the private cloud on the premise of guaranteeing QoS and a self-adapting PSO (SLPSO)-based programming approach is proposed to solve the problem and improve the efficiency of the solution.

The correct option of the best strategy for optimizing the usage of the internal data center and minimizing the cost of performing overloaded tasks in the cloud is discussed in [18]. The authors present an optimization issue in a multi-provider hybrid cloud environment with time-limited and preventable but non-provider-movable workloads characterized by memory, CPU and data transmission requirements. The authors suggest the formulation of the scheduling problem using a binary integer program and analyze the computational costs of this technique with regard to the key parameters of the problem.

2.3. Performance modeling & evaluation

Performance modeling is paramount in building trust and confidence in predicting how a system or network will actually perform. Especially within cloud architectures, all performance indices are in advance defined by service level agreements (SLAs). Any breach can lead to penalties and business loss to the cloud providers. Researchers have often modeled performance either via measurement, analytical or simulation methods focusing on particular Quality of Service (QoS) parameters as an indication of how service requests are being processed in the cloud environment [19].

The challenges related to the performance modeling and prediction of cloud-based applications, which are fine-grained and highly distributed, both from a geographic and logic perspective, are extensively discussed in [20], where a performance model suitable for analyzing the service quality of large sized IaaS clouds, is done using interacting stochastic models and suitable simulation and modeling techniques.

Analytical modeling is often used in the performance modeling scenarios, as it usually overcomes the limitations posed by rules-based policies, as discussed in [21]. In the domain of dynamic provisioning, analytical performance models are often built on queuing theory, where Queueing Networks (QN) become essential for modeling multi-tier applications to describe application architecture [22].

In [23] authors combine the power of both analytical formulation and simulation for modeling the performance of cloud-based system. Specifically, the authors present a novel modeling method oriented to the prediction of cloud architectures performance, suitable for joining the advantages of high-level modeling abstractions and of the detail of a specialized simulator. In [24], Generalized Stochastic Petri Net, are used to describe the workload and the behavior of users and applications, while Cloudsim [25] is adopted for the cloud simulation. In [26] a multi-formalism approach is used within a standard simulator to evaluate performances of security monitoring in multi-tenant cloud architectures. To cope with the numerous challenges rising from the management of highly dynamic workloads scheduled on the cloud infrastructure, accurate evaluations of workloads and of their resource demands are compelling.

Workload evaluations have been addressed under different perspectives. Roy et al. [27] evaluate the incoming workload of a system for future time periods by means of an Autoregressive Moving Average model considering the workload patterns up to the current time period. Similarly, the authors in [28] address the evaluation of workloads characterized by a seasonal behavior using an Autoregressive Integrated Moving Average model. The model, which is based on historical workload data, is updated on the run by applying feedback from latest observed loads.
3. System organization and modeling parameters

3.1. Problem statement

A university delivers real time on-line lectures on a regular schedule for all courses of all offered degrees to students. Lectures and other contents should be available in real time. They are recorded in order to be available off-line.

Courses are organized in lectures, which occur during the morning and the afternoon of weekdays, thanks to a scheduling that prevents overlapping.

This paper proposes a technique to evaluate the effects of the behavior of the users of a hybrid cloud system according to the needs of the university, in order to understand what are the dynamics of the relevant parameters over time and support the planning, design, development and management operations.

3.2. Users

The system consists of two categories of users that are relevant to understand the workloads: professors and students. Both categories only can access the system after authentication, which is managed locally in the private cloud.

Professors are granted the possibility to access the system any time to schedule or reschedule lectures, to upload pre-recorded contents or documents, to start, stop and record real-time lectures, to schedule, define, start and stop exam procedures. Students are granted the possibility to access the system at any time.

They can access recent recorded lectures that are already available for immediate fruition, and they can book long term stored lectures, which will be available as soon as possible and notified; finally, they can take exams.

For the sake of simplicity, the modeling process considers, for both users, the operations that are more computationally expensive for the system in terms of frequency and bursts of execution (e.g. login operations, which, at the beginning of the daily lecture schedule, involve all users simultaneously) or heavy workload (e.g. multimedia operations). All other operations will be assimilated to the most similar critical ones, with no loss of generality.

3.3. System organization

As the system is owned by the university, standard workloads must be estimated to define the specifications. Since very large volumes of streaming traffic and stored multimedia contents are involved, and the workload of the system is very different during daytime and night hours, during weekdays and weekends, during semesters, exam periods, pauses and holidays, elasticity is an important feature of the desired system, so a cloud-based approach is chosen. As all users must be correctly served in every situation, and the workload must be always correctly handled, a hybrid cloud logic is adopted. The idea is to use a public cloud to balance the peaks and the overloads, so that when there are not enough resources to ensure a correct response of the system, part of the workload is overflowed towards a chosen public cloud provider on a pay-per-use logic. Moreover, as only the most recent lectures of each course are likely to be frequently accessed, to save storage and to avoid the expense for an excess of low-accessed installed storage devices, older materials are encrypted and stored on a public, high-latency, low-cost cloud storage service, so that they can be accessed via a booking system that collects students' requests and brings back on-line requested contents when possible, notifying the students when ready to deliver. Old contents are kept for a given number of years on the external cloud storage service, to limit costs: some courses can be stored for more or for less years according to the needs of the university, so that an average time limit is considered.

The general organization of the system is represented in Fig. 1. It shows the relevant user types on the left, the private cloud in the center and the third-party services on the right. Both public cloud services and long-term storage services may be delivered by means of cloud technologies. Both of them are billed on a commercial service basis and reachable by a connection with the needed bandwidth.1 The private cloud provides local authentication services (managed in compliance of EU General Data Protection Regulation or GDPR), local cloud scheduling support (to manage provisioning of local cloud computing resources and elasticity, and the need for overflowing towards the public cloud services), local storage (to store and provide recent contents), a long term storage management module (that automatically applies policies for transfers of materials towards the long term storage when time comes, and manages requests for old materials retrieval) and a offline materials booking management module (that collects and organizes retrieval requests for the long term storage manager).

---

1 It is assumed that the reference is the RIMIC network (http://www.rimic.it) as link between the private cloud and external services.
3.4. Modeling parameters

Due to the goals of the modeling and evaluation process, a set of high-level parameters has been designed to allow a sufficiently effective description of the needs and the management policies for the system. Parameters are organized into 3 categories, namely Organizational parameters, Behavioral parameters and Service parameters.

Organizational parameters include all parameters describing how the academic year is organized. In this perspective, the number, start date and length of teaching periods, exam and holiday periods per year are fixed on the macroscopic time scale. The organization of each period in terms of weekly organization, number of teaching days per week are fixed on the medium time scale. Number and distribution of lecture hours per day are fixed on the microscopic scale. Finally, the number of real-time and pre-recorded courses (or additional contents) are fixed. Also, the number of professors, students and related mapping to courses and curricula are given.

Behavioral parameters denote users in terms of average number and variance of (i) students attending a lecture per hour, (ii) professors delivering a lecture per hour, (iii) students that use real-time lecture and (iv) recent recorded or pre-recorded materials and old lectures from the long term storage. Further parameters include the distributions of users within the system during the day, the week and weekends, the holidays and the year.

Service parameters include all parameters describing technical aspects of the system: (i) average lecture length and variance, (ii) average lecture size (in number of megabytes) and variance, (iii) the average and variance of computing needs in MIPS requested to stream or deliver 60 min of a lecture, and (iv) the average local storage time and variance of lectures. Moreover, as these values are actually not independent (some course categories have different characteristics than others, depending of the type of contents delivered and the actual tools and teaching style used), some correlation parameters between technical parameters are discusses in Section 4.

4. Simulation

In order to account for the time taken by the university to properly plan the requested resources and investment, the system is studied over a time span of 10 years. The parameters have been set as per the description provided in Section 3 to approximately match the regular teaching activity of one of the authors of this paper working at Università degli Studi della Campania “Luigi Vanvitelli”, Caserta, Italy. This Italian University, with about 25,000 students, offers 68 undergraduate and postgraduate courses, which are organized in two semesters. For each semester, 30 study credits (assumed of 8 lecture hours each) are provided for each course. Parameters values are shown in Table 2.

To perform a cost analysis on the system proposed, we associate a cost for each generated value. The cost values, grossed up from market value of public cloud Services and server (for internal private cloud operations), are summarized in Table 3. Experiments have been conducted with a custom simulator written in Matlab. Simulations required few minutes on a high-end 2016 MacBook Pro laptop.

4.1. Simulation scenario and modeling approach

Simulations consists in generating $N$ traces accounting for the hourly resource consumption during the considered time frame. In particular, the temporal horizon is divided into time slots, each one covering for one hour. For each time slot $t_i$, a state characterized by the following 7 components is generated:
Table 2
Model parameters.

| Variable description | Variable name   | Value       |
|----------------------|----------------|-------------|
| Avg. num. of students enrolled to courses in parallel | AveStudNum | 2200        |
| Average num. of parallel courses per hour | AveProfNum | 21          |
| Average lesson duration in hours | AveLessLen | 1.5         |

Table 3
Costs * GB associated to cloud operations (USD).

| Operation              | Public cloud | Private cloud |
|------------------------|--------------|---------------|
| MIPS required          | 0.0069 * month | 0.00016 * month |
| Storage read           | 0            | 0             |
| Storage write          | 0            | 0             |
| Remote read            | 0.012        |               |
| Remote write           | 0            |               |
| Remote delete          | 0.135        |               |
| Remote access          | 0.005        |               |
| Data storage           | 0.045 * month | 0.0046 * month |

Table 2
Model parameters.

| Variable description | Variable name   | Value       |
|----------------------|----------------|-------------|
| Avg. num. of students enrolled to courses in parallel | AveStudNum | 2200        |
| Average num. of parallel courses per hour | AveProfNum | 21          |
| Average lesson duration in hours | AveLessLen | 1.5         |

Table 3
Costs * GB associated to cloud operations (USD).

| Operation              | Public cloud | Private cloud |
|------------------------|--------------|---------------|
| MIPS required          | 0.0069 * month | 0.00016 * month |
| Storage read           | 0            | 0             |
| Storage write          | 0            | 0             |
| Remote read            | 0.012        |               |
| Remote write           | 0            |               |
| Remote delete          | 0.135        |               |
| Remote access          | 0.005        |               |
| Data storage           | 0.045 * month | 0.0046 * month |

- $x_1$: **MIPS required** - CPU usage required to support the compression and on-line diffusion of the incoming stream;
- $x_2$: **Storage read** - Read from private cloud due to students either following on-line lessons or accessing locally archived content;
- $x_3$: **Storage write** - Write to the private cloud by lecturer to provide either on-line or off-line content;
- $x_4$: **Requests from remote storage** - Requests of file transfers from high latency remote storage by students requesting older lessons;
- $x_5$: **Remote storage read** - Read from remote storage (as for $x_4$);
- $x_6$: **Local storage occupancy** - Occupancy of the local storage;
- $x_7$: **Remote storage occupancy** - Occupancy of the high latency remote storage.

The first 5 components $x_1 = [x_{11}, \ldots, x_{15}]$ are generated by using a VARFIMA model to capture the temporal correlation between the various variables. The following procedure is used: first, for each time slot $t_i$, a sample $y_i = [y_{i1}, \ldots, y_{i5}]$ from a 5-dimensional Gaussian distribution characterized by mean vector $m_i$ and co-variance matrix $S_i$ is generated. The values of $m_i$ and $S_i$ are dependent on time, and reflect the particular moment of the day/year. In particular, seven possible scenarios are considered: days, nights and week-ends during lesson time; days, nights and week-ends during the exam periods; holiday season. The values of each component is derived from the parameters presented in Table 2, with additional considerations, such as that write to the storage is done only during the lesson periods, and preferably in day time; access to remote stage occurs mainly during the exams, and the workload during holidays is limited. The seven scenarios are chosen as shown in Fig. 2: in particular we consider a year composed of 2 periods, each one composed of 14 weeks of lessons and 10 weeks of exams, plus one single period of 4 weeks of holidays. Each week has two days of week-end load, and we consider 11 h of daily traffic, and 13 h of nightly workload. The VARFIMA model is then obtained applying a digital filter to the set of Gaussian Samples:

$$x_t = \sum_{j=1}^{n_a} a_j x_{t-j} + \sum_{k=0}^{n_b-1} b_k y_{t-k}$$

(1)

Coefficients $[a_1, \ldots, a_{n_a}]$ and $[b_0, \ldots, b_{n_b-1}]$ have been set according to [29] to reflect the correlation structure of a typical video streaming service.
Values $x_{i6}$ and $x_{i7}$ are then computed in a post processing phase. In particular, they take into account two parameters: the average number of hours $\alpha$ a content is hosted locally, and the average number of hours $\beta$ a video remains on the remote storage. They are then computed as:

$$x_{i6} = x_{i-1} + x_{i3} - \frac{x_{i-1}}{\alpha}$$

(2)

$$x_{i7} = x_{i-1} + x_{i6} - \frac{x_{i-1}}{\beta}$$

(3)

Component $x_{i1}$ is used to plan the computation capacity of the local environment, and the external requirements to support peaks of requests. The total network load is determined by components $x_{i2}$, $x_{i3}$ and $x_{i5}$. Since as shown in Table 3 high latency storage charges also the number of requests, these are accounted in $x_{i4}$. The costs related to read, write and deletion on both private and public cloud, can instead be computed from the evolution of variables $x_{i6}$ and $x_{i7}$.

4.2. Results

Results were computed generating $N = 1000$ simulation runs. The Figs. 3 and 4 respectively show the trend of the values of local write speed and remote file requests (Fig. 3), and MIPS requests, local read and remote read (Fig. 4). The simulation spans over two weeks around an exam session.

In Fig. 3, file requests drop periodically in correspondence with the absence of lectures during the weekend. Starting from the beginning of the exams session, as the lectures have ended, the main write workload has decreased, while it can be observed a slight increase of file requests. This activity can be explained by considering that students are preparing the exam by reviewing the material related to the current and previous study programs (according to Italian university regulations, late students can take exams referring to classes offered during preceding academical years).

This interpretation is confirmed in Fig. 4: in this figure, the blue line shows workloads in MIPS, which are clearly mirroring write requests and confirming the pattern emerged during the weekends. It is interesting to note the difference in scale of the local and remote reads, which denote peaks in the same periods but obviously describe a lower request for reads from remote storage. This is expected firstly, as only a minority of students needs older study materials and secondly, by the fact that the level of remote reads line has a basically uniform behavior before and after the end of lectures. Finally, the local read line, which is dominant on the remote read ones, shows an increase or workload after the end of lectures and also during weekends during the finalized of the study preparation.

With regard to the mid weeks within the same time period, the following figures plot MIPS (Fig. 5) and network usage (Fig. 6) divided in percentiles.
Fig. 3. Local/remote write requests.

Fig. 4. MIPS requests and Local/remote read speed.

Fig. 5 tracks down live lectures, as the significant part of the workload of the system in terms of CPU activity is due to video compression and writing in the model. The end of the courses decreases the workload on the system, which becomes negligible. The figure shows the percentiles obtained by the simulation campaign of level of MIPS that is required on the system, in terms of 5 levels of percentiles (10th, 25th, 50th, 75th and 90th) obtained by the campaign. The workloads are obviously negligible in the hours when there is no lecture, overnight, and in the weekend. Fig. 6 also reveals traffic related to students activities: besides the same oscillations due to the hours without lectures, the percentile distribution around 50th percentile (that is related to a bandwidth of 64 Gbps) shows a symmetric distribution of results.

The probability distribution, estimated in peak hours, is shown for MIPS usage in Fig. 7 and for network usage in Fig. 8.
Fig. 5. Percentile diagrams — MIPS distribution.

Fig. 6. Percentile diagrams — network usage distribution.

Fig. 7 confirms the total dominance of the lecture-related workload over the others. The figure shows the probability distribution of being in need of a given amount of MIPS, considering the aggregate computing power available at the local site. A different relationship between the four cases (respectively, the peak of the requests during a day and a night in the lecture period and in the
exam period) emerges in Fig. 8 with respect to network workload distribution. It shows a quite low activity overnight during the exam period (that anyway amounts to 10 Gbps and is thus non-negligible) with respect to night activity during the exam period, and a minor activity during lectures with respect to exam periods during the daytime. Note that the distributions are similar and quite symmetric.

Fig. 7. Probability distribution of MIPS usage needed in peak hour.

Fig. 8. Network usage distribution in peak hour.
Fig. 9 denotes the correlation between couples of metrics from the computation results. The first five plots show the presence of different periods occurring during the usage of the system, denoted by multiple groups of dots. In fact, in the top left plot, the left, vertical blue aggregate is related to the period in which there is no lecture, so that the requested MIPS are negligible with respect to the lecture period, which is instead depicted by the central cloud. Analogously, in the top center, top right and bottom left plot similar phenomena are observable. In the bottom center plot, more significant correlations emerge between local read/write (in 3 periods) and remote requests/remote read, as expected.

Fig. 10 plots the shape of local and remote storage amount during the 10 years of lifetime of the system that have been simulated. The remote storage, as expected, grows with time, while the local storage plot shows a “sawtooth” shape, due to the offloading of lecture videos after exam sessions. After a saturation period of 2 years, due to the limit designed for local storage, the usage of its resources has a constant average. As storage lifetime is a critical issue and its replacement in intensively used systems is part of the resource planning, the limit that has been assumed as a design parameter should be intended as temporary, as the usual dynamics of storage units costs and the need for replacement naturally allow an expansion of the total amount of storage volume available in short periods.

The last results of the simulation of the system are meant to identify, in an hybrid cloud environment, which part of the system will be implemented in public cloud or on premise private cloud infrastructure. To do this, a set of thresholds is introduced. Since the system uses by default internal resources (until threshold), and, in the case of lack of those resources, uses costly external resources, an evaluation of the costs is of great importance as the threshold varies. It is important to underline only the bare metal costs are taken in account, while a real Total Cost of Ownership (TCO) approach will be considered in future work.

Firstly, the probability of the system to stay below the threshold level is evaluated. Fig. 11 denotes the probability of staying below threshold versus the threshold value (in MIPS). The third axis represents the simulation runs. It is immediate to see that, if the internal systems can provide more than $5 \times 10^5$ MIPS, the probability to be below threshold is 100%, i.e. the system may be implemented only using internal resources.

Fig. 12 represents the average yearly costs related to the imposed threshold, which represents the internal maximum system performance. The blue line shapes the costs (actual cost) as the threshold varies. The other lines represent a possible analysis in case of provider’s cost rising (cost × 1.5) or decreasing (half, one third, one fourth).

5. Conclusions and future work

This paper provided an example of simulation approach supporting the decision-making process in exploring the advantages of hybrid cloud based solutions for Universities willing to adopt in-house synchronous and asynchronous distance learning services in order to avoid the risks and the exposure due to external providers. The presented scenario includes realistic parameters, using as reference values those originated from a real case, and the results of its evaluation, including technical workload distributions, requirements and a cost-related analysis. The main use of external cloud resources is the remote storage of encrypted materials.
that are not needed to be available online, but with a user reservation-based distribution model, and the use of external computing resources to cope with load peaks that may arise during the life of the system.

Future work includes the evaluation of cost trade-offs by applying a more sophisticated approach, based on the total cost of ownership and other managerial metrics considering all the factors that drive the decision process. Furthermore, another area of investigations concerns the evaluation of network performances in terms of quality of service for a better understanding of
operational conditions and a more detailed model of multimedia service provisioning. As a result, we expect to account for the features that enable the video distribution servers to lower quality to cope with higher demand on the short period.

References

[1] C.W.J. Granger, R. Joyeux, An introduction to long-memory time series models and fractional differencing, J. Time Series Anal. 1 (1) (1980) 15–29.
[2] H.H. Yang, L. Feng, J. MacLeod, Understanding college students’ acceptance of cloud classrooms in flipped instruction: integrating UTAUT and connected classroom climate, J. Educ. Comput. Res. 56 (8) (2019) 1258–1276.
[3] T. Dillon, C. Wu, E. Chang, Cloud computing: issues and challenges, in: 2010 24th IEEE International Conference on Advanced Information Networking and Applications, Ieee, 2010, pp. 27–33.
[4] B. Dong, Q. Zheng, M. Qiao, J. Shu, J. Yang, Bluesky cloud framework: an e-learning framework embracing cloud computing, in: IEEE International Conference on Cloud Computing, Springer, 2009, pp. 577–582.
[5] A. Mikroyannidis, A semantic framework for cloud learning environments, in: Cloud Computing for Teaching and Learning: Strategies for Design and Implementation, IGI Global, 2012, pp. 17–31.
[6] Q. Alajmi, A. Sadqi, A. Kamaludin, M.A. Al-Sharafi, E-learning models: The effectiveness of the cloud-based E-learning model over the traditional E-learning model, in: 2017 8th International Conference on Information Technology (ICIT), IEEE, 2017, pp. 12–16.
[7] A. Ghazizadeh, Cloud computing benefits and architecture in e-learning, in: 2012 IEEE Seventh International Conference on Wireless, Mobile and Ubiquitous Technology in Education, IEEE, 2012, pp. 199–201.
[8] I. Ewuzie, A. Usoro, Exploration of cloud computing adoption for e-learning in higher education, in: 2012 Second Symposium on Network Cloud Computing and Applications, IEEE, 2012, pp. 151–154.
[9] I. Carol, G.G.R. Roy, A.J.P. Prassanna, A cloud model for effective e-learning, in: 2014 World Congress on Computing and Communication Technologies, IEEE, 2014, pp. 167–169.
[10] M.C. Calzarossa, M.L. Della Vedova, L. Massari, D. Petcu, M.I. Tabash, D. Tesserai, Workloads in the Clouds, in: Principles of Performance and Reliability Modeling and Evaluation, Springer, 2016, pp. 525–550.
[11] S. Engdahl, Cloudbursting - Hybrid Application Hosting, 2008, https://aws.amazon.com/blogs/aws/cloudbursting/.
[12] H.Y. Chu, Y. Simmhan, Cost-optimal and resilient job life-cycle management on hybrid clouds, in: 2014 IEEE 28th International Parallel and Distributed Processing Symposium, IEEE, 2014, pp. 327–336.
[13] M.R. Hoseinyfarahabady, H.R. Samani, L.M. Leslie, Y.C. Lee, A.Y. Zomaya, Handling uncertainty: Pareto-efficient bot scheduling on hybrid clouds, in: 2013 42nd International Conference on Parallel Processing, IEEE, 2013, pp. 419–428.
[14] H. Zhang, G. Jiang, K. Yoshihira, H. Chen, Proactive workload management in hybrid cloud computing, IEEE Trans. Netw. Serv. Manag. 11 (1) (2014) 90–100.
[15] Y. Ogawa, G. Hasegawa, M. Murata, Cloud bursting approach based on predicting requests for business-critical web systems, in: 2017 International Conference on Computing, Networking and Communications (ICNC), IEEE, 2017, pp. 437–441.
[16] M. Hajiof, X. Sun, Y.W.E. Sung, D. Malts, S. Rao, K. Sripandikulchai, M. Towasmalani, Cloudward bound: planning for beneficial migration of enterprise applications to the cloud, ACM SIGCOMM Comput. Commun. Rev. 40 (4) (2010) 243–254.
[17] X. Zuo, G. Zhang, W. Tan, Self-adaptive learning PSO-based deadline constrained task scheduling for hybrid IaaS cloud, IEEE Trans. Autom. Sci. Eng. 11 (2) (2013) 564–573.
[18] R. Van den Bossche, K. Vanmechelen, J. Broeckhove, Cost-optimal scheduling in hybrid IaaS clouds for deadline constrained workloads, in: 2010 IEEE 3rd International Conference on Cloud Computing, IEEE, 2010, pp. 228–235.
[19] P. Singleton, Performance modelling—what, why, when and how, BT Technol. J. 20 (3) (2002) 133–143.
[20] H. Khazaee, J. Misic, V.B. Misic, A fine-grained performance model of cloud computing centers, IEEE Trans. Parallel Distrib. Syst. 24 (11) (2012) 2138–2147.
[21] O. Ibidunmoye, M.H. Moghadam, E.B. Lakew, E. Elmo, Adaptive service performance control using cooperative fuzzy reinforcement learning in virtualized environments, in: Proceedings of the 10th International Conference on Utility and Cloud Computing, 2017, pp. 19–28.
[22] C. Qu, R.N. Calheiros, R. Buyya, Auto-scaling web applications in clouds: A taxonomy and survey, ACM Comput. Surv. 51 (4) (2018) 1–33.
[23] E. Barbierato, M. Gribaudo, M. Iacono, A. Jakóbik, Exploiting CloudSim in a multiformalism modeling approach for cloud based systems, Simul. Model. Pract. Theory 93 (2019) 133–147.
[24] M.A. Marsan, G. Balbo, G. Conte, S. Donatelli, G. Franceschinis, Modelling with Generalized Stochastic Petri Nets, Vol. 292, Wiley, New York, 1995.
[25] R.N. Calheiros, R. Ranjan, A. Beloglazov, C.A. De Rose, R. Buyya, CloudSim: a toolkit for modeling and simulation of cloud computing environments and evaluation of resource provisioning algorithms, Softw. Pract. Exp. 41 (1) (2011) 23–50.
[26] L. Campanile, M. Iacono, S. Marrone, M. Mastroianni, On performance evaluation of security monitoring in multitenant cloud applications, Electron. Notes Theor. Comput. Sci. 353 (2020).
[27] N. Roy, A. Dubey, A. Gokhale, Efficient autoscaling in the cloud using predictive models for workload forecasting, in: 2011 IEEE 4th International Conference on Cloud Computing, IEEE, 2011, pp. 500–507.
[28] R.N. Calheiros, E. Masoumi, R. Ranjan, R. Buyya, Workload prediction using ARIMA model and its impact on cloud applications’ QoS, IEEE Trans. Cloud Comput. 3 (4) (2014) 449–458.
[29] A. Biernacki, Analysis and modelling of traffic produced by adaptive http-based video, Multimedia Tools Appl. 76 (2017).