ARE CLOUD ENABLED VIRTUAL LABS ECONOMICAL?

A Case Study Analyzing Cloud based Virtual Labs for Educational Purposes

Nane Kratzke

Lübeck University of Applied Sciences, Mönkhofer Weg 239, 23562 Lübeck, Germany

Keywords: Cloud, Computing, Virtual Labs, Higher Education, IaaS, Decision, Making, Cost, Estimation.

Abstract: Cost efficiency is an often mentioned strength of cloud computing (Talukader et al., 2010). In times of decreasing educational budgets virtual labs provided by cloud computing might be an interesting alternative for higher education organizations or IT training facilities. This contribution analyzes the cost advantage of virtual educational labs provided via cloud computing means and compare these costs to costs of classical educational labs provided in a dedicated manner. This contribution develops a four step decision making model which might be interesting for colleges, universities or other IT training facilities planning to implement cloud based training facilities. Furthermore this contribution provides interesting findings when cloud computing has economical advantages in education and when not. The developed four step decision making model of general IaaS* applicability can be used to find out whether an IaaS cloud based virtual IT lab approach is more cost efficient than a dedicated approach.

1 INTRODUCTION

Cloud computing is one of the latest developments within the business information systems domain and describes a delivery model for IT services based on the Internet, and it typically involves the provision of dynamically scalable and often virtualized resources. Nowadays cloud computing is used in e-learning scenarios as well because it fits very well to e-learning requirements. When e-learning has a distinctive remote aspect why delivering necessary educational resources like labs has to be delivered still in a classical and dedicated manner?

Accompanying the increasing relevance of cloud computing in research literature and media there arise manifold publications covering the application of cloud computing to e-learning. E.g. (Cayirci et al., 2009) present a training and education cloud. But only some publications covering aspects how to use cloud computing to provide virtual labs. (Oberg et al., 2011) presents a "virtual lab" system architecture for academic high-performance computing but without educational purposes. And a lot of other authors reflect about to use cloud computing as technical infrastructure for providing e-learning systems (Dong et al., 2009), (Ko and Young, 2011), (Masud and Huang, 2011) digital campus systems (Li and Li, 2011) or personalized learning environment systems (Liang and Yang, 2011) but this has nothing to do with virtual laboratories in the understanding of this contribution. Only (Thiébaut et al., 2011) show a very interesting example of processing wikipedia dumps by applying cloud computing technologies and compare them to dedicated cluster solutions in practical college courses. This could be named a "virtual lab" but it is not done by the authors due to a different focus of their paper. So this contribution is about a case study how to use cloud computing in computer science study programs providing virtual IT labs for practical courses in the higher education domain or for IT training facilities. According to (Barr, 2010) we define:

> *A **Virtual Educational IT Lab** is a collection of compute, storage and net-working resources provided by an educational organization for educational or research purposes. A virtual lab can be provided to a single or a small group of students to support these student(s) in solving practical problems by providing a necessary IT infrastructure. Provided resources are available for short-term use and are billed by actual resource consumption generated by educational or research ac-*

*Infrastructure as a Service. This contribution follows the IaaS, PaaS and SaaS definition of NIST (Mell and Grance, 2011).

Kratzke N..
ARE CLOUD ENABLED VIRTUAL LABS ECONOMICAL? - A Case Study Analyzing Cloud based Virtual Labs for Educational Purposes.
DOI: 10.5220/0003888503530362
In *Proceedings of the 4th International Conference on Computer Supported Education* (CSEDU-2012), pages 353-362
ISBN: 978-989-8565-07-5

tivities. Typically all provided resources are rented by the educational organization from a cloud service provider.

Beside use cases like hosting websites, support software development cycles, short-term system demonstrations, data storage, disaster recovery and business continuity, media processing and rendering, overflow processing or large-scale scientific data processing (Barr, 2010) mentions training use cases as very cloud compatible and economical use cases. This paper does not denial this postulation in general but advocates a more critical view like (Weinmann, 2011) or (Mazhelis et al., 2011) do. Ongoing research (Kratzke, 2011a), (Kratzke, 2011b), (Kratzke, 2012a) shows that cost advantages of cloud computing are deeply use case specific and you should be aware of comparing non comparable use cases. And nevertheless we have to consider some shortcomings of cloud computing which are mentioned in literature and should be considered in planning phases of virtual labs.

Shortcomings of Cloud Computing. (Kratzke, 2012a) showed by analysing popular IT management frameworks like COBIT, TOGAF or ITIL that – from an IT management point of view – cloud computing provides more benefits than disprofits. Nevertheless Kratzke showed also the presence of so called substantial show stoppers for cloud based approaches. Especially security and compliance management aspects may come along with substantial "showstoppers". But these aspects are of minor relevance regarding cloud based virtual labs.

According to (Truong and Dustdar, 2010) or (Kratzke, 2012a) cloud computing is also characterized by an ex ante cost intransparency. This very important weakness (from an IT management and decision making point of view) is even little reflected in literature so far. To answer the question whether a cloud-based approach is more cost efficient than a dedicated approach it has to be answered the question what costs will be generated per month **before** a cloud based approach enters operation (Walker et al., 2010). This is very difficult to answer ex ante because it is influenced by a bunch of interdependent parameters. But for profound decision making for or against cloud based virtual labs exactly this question has to be answered before a virtual lab is established in a dedicated or cloud based manner.

Outline. Therefore this contribution has the following outline. The research methodology is described in section 2. An economical decision making model for or against cloud based virtual labs is described briefly in section 3. A corresponding case study of a computer science study program is analyzed in section 4. Section 5 shows derived findings from the analyzed case study which are useful to plan virtual lab supported practical courses in higher education. Section 5 shows furthermore how to transfer the presented use case to other computer science related lectures and related practical courses and operationalizes simple rules for controlling costs pragmatically. The contribution ends with a summary and outlook in section 6.

2 RESEARCH METHODOLOGY

Especially the work of (Weinmann, 2011) is very interesting from this decision making point of view because it shows how to decide for or against a cloud based approach very pragmatically. This contribution is about using the work of Weinman to build up a simple decision making model for or against cloud based approaches. This decision making model has been applied in a presented case study covering the higher education domain.

For practical courses in higher university or college education we want to find out whether it is more economical to provide classical dedicated educational labs or to use IaaS[2] providing virtual labs for student practical courses.

The analyzed case study was a web technology lecture for computer science students being held at the Lübeck University of Applied Sciences in 2011. During the practical courses of this lecture students formed groups of 5 or 6 persons in order to build up a website for a scientific conference on robotic sailing[3] (project 1) or establish a google map based automatic sailbot tracking service (project 2) for the same conference. All groups were assigned cloud service accounts provided by Amazon Web Services. The groups were asked to use these accounts in order to fulfill their projects in a complete cloud based manner. The resource consumption of all groups were measured by analyzing billing as well as usage data provided by the cloud service provider Amazon Web Services (AWS).

Both projects had three main phases (see table 1). The training phase (calendar week 13 to 15) was about to get all students into touch with the cloud service provider tool suite. In the project phase (calendar week 16 to 23) all groups had to develop a cloud based solution for project 1 or project 2. All groups had to

[2]This contribution follows the IaaS, PaaS and SaaS definition of NIST (Mell and Grance, 2011).

[3]World Robotic Sailing Conference 2011, see http://www.wrsc2011.org.

proof that their solutions were adequate for a 24x7 operation within a 24x7 phase (calendar week 21 to 24). So project and live phase overlapped within a three week period. Calendar week 24 was used for the system presentations. The practical course of the lecture finished with a migration phase. The best solutions were used as website (project 1) or google maps based automatic sailbot tracking service (project 2) for the conference and were migrated to the destination environment (which also was a cloud based environment) in calendar week 25.

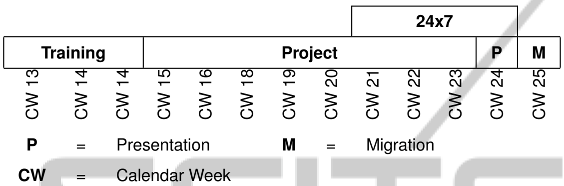

Figure 1: Project phases.

3 DECISION MAKING MODEL

(Weinmann, 2011) is stressing the following interesting fact which is a crucial input for pragmatic decision making for or against cloud based system implementations especially on a IaaS level of cloud computing:

> *A pay-per-use solution obviously makes sense if the unit cost of cloud services is lower than dedicated, owned capacity. [...]*
>
> *[...]* ***a pure cloud solution also makes sense even if its unit cost is higher****, as long as the peak-to-average ratio of the demand curve is higher than the cost differential between on-demand and dedicated capacity. In other words, even if cloud services cost, say, twice as much, a pure cloud solution makes sense for those demand curves where the peak-to-average ratio is two-to-one or higher.* (Weinmann, 2011)

According to Weinman the peak-to-average ratio is the essential indicator whether a cloud based approach is economical reasonable or not. So it is not necessary to estimate costs per month of a cloud based solution exactly. It is sufficient to proof that cloud based costs are smaller then a dedicated system implementation. And this can be figured out by analysing the peak usage as well as the average usage of a system.

Due to page limitations we will use in this contribution the average to peak ratio atp in a simplified form and refer to (Kratzke, 2012b) where the decision making model is presented more detailly.

$$atp := \frac{averageusage}{maximumusage} = \frac{1}{pta} \tag{1}$$

According to (Weinmann, 2011) we have to compare the costs of a classical dedicated approach with the costs of a cloud based approach. On the IaaS level it is common to be billed per service usage with a granularity of the atomic timeframe t level. Which would be in case of Amazon Web Service that you are billed for a server instance per complete (or partial) hour usage. Let us name our dedicated costs per atomic timeframe d and our cloud costs per atomic timeframe c. Cloud costs c can be easily figured out because they are provided as pricing by their cloud service providers[4]. Dedicated costs per atomic timeframe d are a little more complex to calculate. Nevertheless for estimations we can assume, that they can be defined via their amortizations. If a dedicated instance can be procured for p value units[5] their dedicated costs per atomic timeframe (ATF) can be calculated as follows[6]:

$$\begin{aligned} d_{ATF}(p) &= \frac{p}{ATF} \\ d_{3year}(p) &= \frac{p}{3 \cdot 365 \cdot 24h} \\ d_{5year}(p) &= \frac{p}{5 \cdot 365 \cdot 24h} \end{aligned} \tag{2}$$

Typical amortization timeframes are a 3 year or a 5 year hardware regeneration interval (see equation 2). So a 500 \$ server would have the following dedicated costs per atomic timeframe of 1h over a amortization interval of 3 years.

$$d_{3year}(500\$) = \frac{500\$}{3 \cdot 365 \cdot 24h} \approx 0.019\frac{\$}{h} \tag{3}$$

According to (Weinmann, 2011) the peak-to-average ratio pta should be greater than the relation between the variable costs per atomic timeframe c and the dedicated costs per atomic timeframe d which can be expressed in the following form:

$$pta > \frac{c}{d} \Leftrightarrow pta \cdot d > c \Leftrightarrow c < \frac{d}{atp} \tag{4}$$

[4]E.g. Amazon Web Services publishes these cloud costs per atomic timeframe here: http://aws.amazon.com/de/ec2/#pricing.

[5]E.g. US Dollars \$ or Euro €.

[6]Be aware – this assumption do not account typical additional operational efforts like administration, cooling or electricity. Nevertheless we do not want to calculate exact costs we only want to know whether a cloud based approach is more economical than a dedicated one. In this case it is OK to give the dedicated side an advantage by not accounting aspects like administration, cooling, electricity, etc. although these costs are included in the variable costs on the cloud service provider side.

In other words equation 4 provides a clear decision criteria to decide for or against a cloud based approach. By knowing your average to peak ratio atp, your hardware procurement costs per instance p as well as your hardware amortization timeframes (which is typically 3 or 5 years) it is possible to calculate a maximum of cloud costs per atomic timeframe c_{MAX} until a cloud based approach is economical (see equation 5). Whenever a cloud service provider can realize instance pricings below c_{MAX} we decide for a cloud based approach[7] – in all other cases we should realize the system in a dedicated approach[8].

$$c_{MAX} := \frac{d}{atp} \tag{5}$$

4 ANALYSED CASE STUDY

All data used in this section has been collected by the supporting cloud service provider AWS. The billing data as well as usage data are standard informations provided by AWS for each client. They are provided by AWS in CSV or XML format and can be therefore easily downloaded and processed by anyone. We used billing as well as EC2 usage data provided by AWS. No special preprocessing of the data were performed by AWS for this special use case analysis. So anyone can run the same analysis by using a standard toolset[9].

4.1 Cost Analysis

Table 1 shows all costs per group within the analysed timeframe. In total the Lübeck University of Applied Sciences had to spend 847.01$ in providing (virtually) unlimited amount of server instances to 49 students organized in 9 groups within a timeframe of 13 calendar weeks. This sounds impressive but says in fact nothing about how cost efficient this performed cloud based approach was. Could we had reached the same results with a classical dedicated approach?

Let us have a look at the cost analysis provided in figure 2. Figure 2(A) shows the costs per month. AWS is only providing billing data on a monthly basis so we had to align months to the corresponding analysed calendar weeks. By interpreting these graphs we must have in mind the general project phases performed by all student groups (see table 1). So please keep the following phases in mind:

- Week 13 - 15: initial training phase.

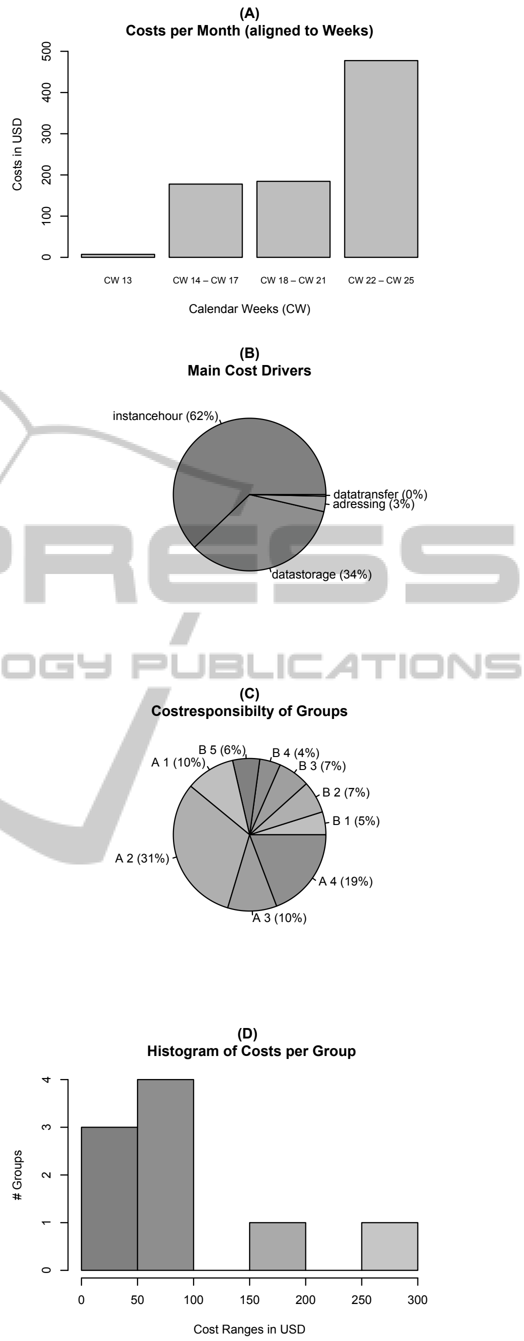

Figure 2: Cost analysis.

- Week 16 - 23: project/development phase.
- Week 21 - 24: (overlapping) 24x7 phase.
- Week 24: presentation.
- Week 25: migration.

So figure 2(A) shows that most of the costs were generated in 24x7 phase (calendar week 21 - 24). **So 24x7**

[7] Only from an economical point of view.

[8] Also only from an economical point of view.

[9] We used only open source software. MySQL as a database and R as a analytical and visualization tool.

Table 1: Group overview.

Group	Size	Project	Costs in $
A 1	5	WRSC Website	88.39$
A 2	6	WRSC Website	265.37$
A 3	4	WRSC Website	88.14$
A 4	6	WRSC Website	162.88$
B 1	6	Sailbot Tracking	41.17$
B 2	6	Sailbot Tracking	57.58$
B 3	6	Sailbot Tracking	57.46$
B 4	5	Sailbot Tracking	37.42$
B 5	5	Sailbot Tracking	48.58$

seems to be expensive.

Figure 2(B) shows the main cost drivers. Almost $2/3$ of the costs were generated by instancehours – that means running servers and being billed for per hour. Almost $1/3$ of the costs were generated by data storage – that means all costs which have to do with the provision of server hard drives, storing backups or other data storage services. Other costs like addressing (requesting IP addresses, DNS names, etc.) or even data transfer had no relevant impact to the total costs. **So the main cost driver were instance hour costs – or box usage as it is sometimes called.**

Figure 2(C) and (D) shows that analyzed groups had quite different cost responsibilities. The expected value of cost responsibility would be $100\%/9 \approx 11.11\%$.

Nevertheless the most cost efficient group was only responsible for about 4% and the most cost "lavish" group was responsible for about 31% of the total costs. It turned out that the A groups – groups responsible for setting up the website – consumed more resources than the sailbot tracking B groups (check table 1). **So cloud generated costs are use case specific.** Different problems result in different architectural solutions which results in different cost behaviour.

4.2 Usage Analysis

As we have found out in our analyzed use case – main cost driver was server/box usage (or instance hours – see figure 2(B)). Thats why a detailed box usage analysis has been performed and is shown in figure 3.

Figure 3(A) shows the maximum and average box usage per calendar week measured within the analyzed timeframe (calendar week 13 - 25). **Figure 3(B)** shows the total sum of all processing hours generated by all used server boxes/instances per calendar week. **Figure 3(C)** shows the average to peak load ratio (according to equation 1) per calendar week.

Within the **initial training phase (calendar week 13 - 21)** the usage characteristic shows an extremely

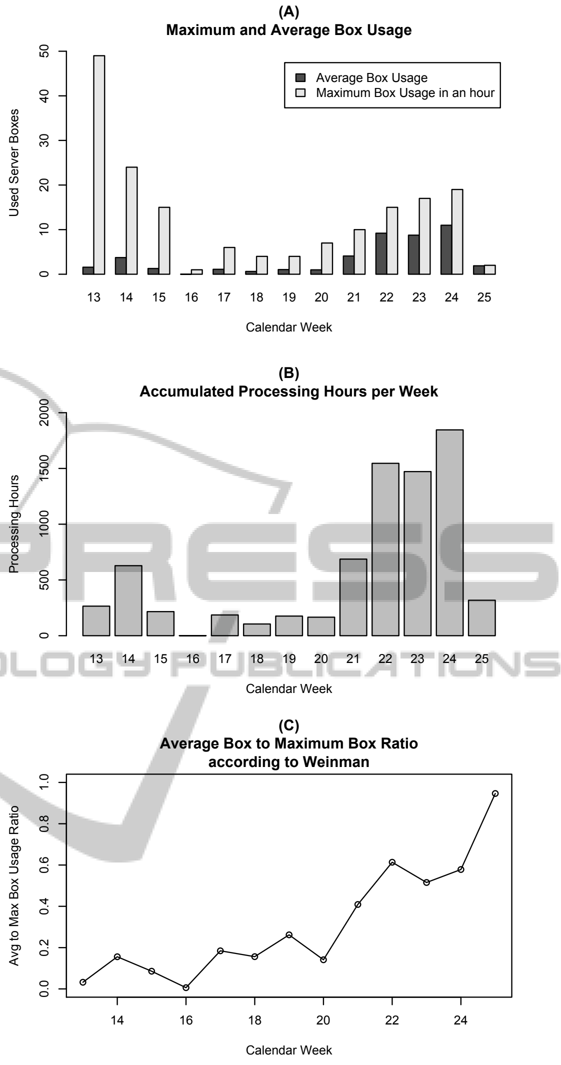

Figure 3: Analysed box usage.

high maximum box usage but an astonishing low average usage. This characterizes an extreme peak load situation and results in extreme low atp ratios (see equation 1 and figure 3(C)). According to equation 4 or the definition of c_{MAX} (equation 5) this shows a very ideal cloud computing (peaky) situation. **So training phases seems to be very economical interesting cloud computing use cases** which is also stated by (Barr, 2010)[10].

Within the **project phase (calendar week 13 - 15)** the usage characteristic shows dramatically reduced maximum as well as average box usages. Nevertheless the atp ratios (see equation 1 and figure 3(C))

[10](Barr, 2010) is chief evangelist of Amazon Web Services – so this source can not be rated objective. Nevertheless our case study hardens his point.

stay in a very comfortable situation for cloud computing approaches. So we still have a peaky usage situation but on a dramatical lower level. **So also development phases seems to be very economical interesting cloud computing use cases** which is also stated by (Barr, 2010)[11].

The **24x7 phase** (calendar week 21 - 24) shows raised maximum as well as average box usages (see figure 3(A)). Also the atp ratios are raising within this phase – nevertheless the peaky usage characteristic remains but less distinctive. The 24x7 phase can be clearly seen in the accumulated processing hours (see figure 3(B)) which shows a clear peak in calendar weeks 22 - 24. We have stated already in section 4.1 that **24x7 seems to be expensive** and we can harden this conclusion by our usage analysis. This might surprise some readers[12].

The **migration phase** (calendar week 25) was characterized by transferring the best solutions for the website and sailbot tracking service into the operational environment. Within this phase the systems run in a steady load scenario as can be seen in figure 3(A) (almost the same average box as well as maximum usage) and in figure 3(C) (an atp ratio near 1.0). According to equation 5 this shows an extreme uncomfortable situation for cloud computing situations – **so steady loads seem to be no economical interesting cloud computing use cases.**

4.3 Economical Decision Analysis

As we have seen in sections 4.1 and 4.2 we can identify different phases which are more cloud compatible than others from an ecomomical point of view. Training and development phases show very low atp ratios (see figure 3(C)) and therefore indicate peaky usage characteristics of resources which advantages cloud computing realizations[13] (check equation 5). Other phases with less peaky usage characteristics (like our 24x7 or migration phase) disadvantage cloud computing realizations.

So we have identified pro and cons for a cloud based realization of our educational labs. But how to decide? Now we are going to apply our decision model presented in section 3.

Step 1: Determine the atp Ratio.

First of all we have to calculate our overall average to peak load ratio atp. Our analysis timeframe covered the calendar weeks 13 - 25. So an intuitive timeframe for average building would be 13 weeks – but this implicates a continual usage of an educational lab over a complete year which is very uncommon. In the university or college business educational labs are typically used one time per semester. In most cases an educational lab can be used only one time a year (per semester – that means average building over 26 weeks) or even only one time per year (every second semester – that means average building over 52 weeks which is a year). In our analyzed timeframe 7612 hours of instance usage were generated (the sum of all bars of figure 3(B)). So equation 7 shows the average amount of servers which would be necessary to provide 7612 processing hours within a 26 or 52 week timeframe.

$$avg_{26w} = \frac{7612h}{26 \cdot 7 \cdot 24h} \approx 1.74$$
$$avg_{52w} = \frac{7612h}{52 \cdot 7 \cdot 24h} \approx 0.87 \tag{6}$$

Now we can build up our average to peak ratio. Our maximum server usage within an atomic timeframe of 1 hour was 49 servers (please check figure 3(A)). So we get the following atp ratios for a 26 or 52 week timeframe.

$$atp_{26w} = \frac{1.74}{49} \approx 0.035$$
$$atp_{52w} = \frac{0.87}{49} \approx 0.018 \tag{7}$$

Step 2: Determine your Dedicated Costs.

First of all we have to find out how much would cost us a dedicated server. At the Lübeck University of Applied Sciences our procurement office could purchase the smallest possible server version[14] for about 3055\$ (2167€[15]). Equation 2 told us to calculate our dedicated costs per atomic timeframe ($1h$) in the following way for a 5 year amortization interval:

$$d_{5year}(3055\$) = \frac{3055\$}{5 \cdot 365 \cdot 24h} \approx 0.0697 \frac{\$}{h} \tag{8}$$

[11] Remember (Barr, 2010) is chief evangelist of Amazon Web Services. He is not objective. Nevertheless our case study hardens this point too.

[12] It is often stated that cloud computing is good for hosting websites (a 24x7 job). This is true but only if not provided on IaaS level or when provided on IaaS level with a clear peaky usage characteristic. Steady loads are seldom economical IaaS use cases. So you should be very clear about your use case, the related (un)peaky usage characteristic and the IaaS, PaaS or SaaS provision level. No one states that economical decision making for or against cloud computing is obvious.

[13] From an economical point of view.

[14] Dell PowerEdge Server R610, 2.13 GHz Intel Xeon processor, 8GB memory, 140 GB hard drive (valid on 28th October 2011) with approximately 2 ECU – so the AWS instance type pendant would be something between a Standard Small (1 ECU) or Large (4 ECU) instance type. Check out detailed instance type informations of AWS here: http://aws.amazon.com/de/ec2/instance-types/.

[15] Exchange rate $1.41\$ = 1.00e$ from 28th October 2011.

Step 3: Determine your Maximal Economical Cloud Costs.
Equation 5 told us to calculate our c_{MAX} costs in the following way:

$$\begin{aligned} c^{26w}_{MAX} &= \frac{d_{5year}(3055\$)}{atp_{26w}} = \frac{0.0697}{0.035}\frac{\$}{h} \approx 1.99\frac{\$}{h} \\ c^{52w}_{MAX} &= \frac{d_{5year}(3055\$)}{atp_{52w}} = \frac{0.0697}{0.018}\frac{\$}{h} \approx 3.87\frac{\$}{h} \end{aligned} \quad (9)$$

In other words. In our analyzed case study a cloud service provider (regarding a 5 year amortization time frame for servers) could be

- **28.57 times more expensive** in case of a one time per semester usable educational lab or even
- **55.56 times more expensive** in case of a only one time per year usable educational lab

then own dedicated costs – a cloud based educational lab would be still more economical then a dedicated one. Ok these figures are really impressive but do we find appropriate resources within our maximal costs? We have to figure this out in our last step 4 to make a profound decision for or against a cloud based approach for our virtual lab.
Step 4: Determine Appropriate Cloud Resources.
Now we know our maximal cloud costs and have to look if our cloud service provider can deliver appropriate and comparable resources. In our case this is Amazon Web Services, but it could be any other IaaS cloud service provider as well. We do this exemplarily for a 26 week timeframe[16]. But it works absolutely the same for all other timeframes or IaaS cloud service providers as well.
Table 2 shows all instance types of AWS and their allocated costs. Remember section 4.3 told us, that all instance types cheaper than 1.99 $\$/h$ result into cloud based solutions which are more economical than dedicated approaches.
As you can see in table 2 all provided instance types (except one[17]) of AWS in the EU Region are economical in the sense of section 3 and equations 4 and 5. The most *similar* instance types listed in table 2 are marked as **'o'**. (**'-'** stands for *worser*, **'+'** for *better* and **'++'** for *much better* than a dedicated reference system[18]).
So in our analysed use case the *Medium (High CPU)* or may be even the *Small (Standard)* AWS instance types (see table 2) are the most comparable systems to our dedicated reference system (Dell PowerEdge Server R610). Both provide variable cloud costs clearly below our maximum costs of 1.99$\$/h$ (see equation 9). **So in our analysed use case a virtual IT lab for education is much more economical than a dedicated approach. Thats why we decided to implement cloud based IT labs for our practical courses instead of classical dedicated educational labs.**

[16] Because in our special case we can use our educational lab every semester – so two times a year.

[17] 4x XL (High Memory) instance type is not economical reasonable but this instance type is not comparable to our reference system because it is much more powerful.

[18] In our case the Dell PowerEdge Server R610

Table 2: AWS Instance types and pricings, according to AWS pricing informations on 28th Oct. 2011, EU (Ireland) region, on-demand instances, Linux/UNIX operating system.

AWS Instance Type	ECU	Price/h	Economical	Comparable
Micro	< 1	0.025$	yes	-
Small (Standard)	1	0.095$	yes	o
Large (Standard)	4	0.38$	yes	o
XL (Standard)	8	0.76$	yes	+
XL (High Memory)	6.5	0.57$	yes	+
2x XL (High Memory)	13	1.14$	yes	++
4x XL (High Memory)	26	2.28$	no	++
Medium (High CPU)	5	0.19$	yes	o
XL (High CPU)	20	0.76$	yes	++

5 CONCLUSIONS

In section 5.1 we derive some general findings from our use case analysis which might be useful to find interesting practical courses in higher education of computer science which are likely to show similar cost characteristics like the analyzed use case. Nevertheless our analysis shows also in section 5.2 that the here presented approach is transferable to other lectures and related practical courses as well. We operationalize some general advices in section 5.3 for setting up cloud based virtual labs and develop some rules for virtual lab users to control costs pragmatically.

5.1 General Findings

Cloud Computing Economics Love Peak Load Scenarios. As equation 4 showed cloud computing becomes more and more economical as the average to peak ratio nears 0 - this indicates that peaky as well as seldom usage of educational labs argue for cloud based virtual lab approaches, extreme continual usage of labs argue against cloud based approaches from an economical point of view. Nevertheless we always

have to analyze the individual characteristic of a practical course by measuring its average to peak ratio in real world.
Cloud Generated Costs are use Case Specific. Different problems result in different realization architectures generating different costs as well as usage characteristics. So be aware of comparing non comparable use cases! You should run for each practical course the here mentioned cost analysis and decide for or against a cloud computing based virtual lab approach after you have figured out your practical course specific average to peak ratio.
24x7 Seems to be Expensive and of no Economical Interest. for cloud computing if not associated with a peaky usage characteristic. So try to avoid 24x7 tasks in practical courses whereever possible. This will save a lot of money.
One of the Main Cost Driver is Box Usage per Complete (or Partial) Hour. Second Relevant Cost Driver is Data Storage. Data transfer seems to have a minor impact to costs. This might be only valid for our analyzed use case and should be handled with care. Please check out other publications like (Mazhelis et al., 2011) for communication intensive use cases if you plan to use communication/data transfer intensive use cases in your practical courses.
Constant Loads Over Time Seem to be no Economical Interesting use Cases for Cloud Computing. If you have to deal with this kind of use cases in practical courses (which is typically very unlikely) think about classical dedicated approaches. Cloud computing approaches might be only feasible in these cases due to their convenience but not necessary economical reasonable.

5.2 Transferability of the Approach

Because of the identified cost advantages in the analyzed use case there was one more question arising. Is it possible to transfer made experiences to other lectures and practical courses as well? Therefore all module descriptions of computer science (near) study programs of the Lübeck University of Applied Sciences were analyzed. The following lectures and related practical courses showed potential being supported by virtual labs[19].

- Industrial Networks and Databases.
- Informationtechnology.
- Webtechnologies.
- Database Management (and Engineering).
- Integrated Informationsystems.
- Distributed Systems.
- Operating Systems (if Linux/UNIX based).

It turned out that in each analyzed computer science related study program between 8.33% up to 23.10% of all practical courses could be supported by virtual labs according to the module descriptions of the study programs. Whether these practical courses will show similar cost advantages is up for ongoing research. Nevertheless it seems very likely from made experiences. So it should by obvious to the reader that the presented approach is easily transferable to other lectures and related practical courses.
To identify interesting lectures and practical courses it is good to know what is making a practical course virtual lab compatible. Our identified courses have the following virtual lab relevant educational requirements in common:

- a practical course requires databases,
- a practical course requires distributed processing instances,
- a practical course requires webtechnologies on server instances,
- a practical course requires linux/unix based instances,
- a practical course requires parallel processing,
- a practical course requires large-scale data processing.

Whenever there exist a practical course with one of the above mentioned educational requirements a virtual lab might be a reasonable option.

5.3 Advices for Transferring

We learned a lot about providing virtual labs to students via cloud computing means. Some of our made experiences are operationalized as advices for everyone who is planning similar approaches.
Advice Nr. 1: It is likely that Cloud Computing will be new for students. So plan an **initial training phase** with the cloud tooling of the service provider which is necessary to solve the problems of the practical course. We did this with by providing very detailed step by step manual. This seems to be an effective way. And force your students to be present in the first initial training phase.
Advice Nr. 2: If you are planning to use virtual clouds in more than one practical course you might think about a **virtual lab training course** in one of the first semesters of a study program. This might avoid double and triple trainings.

[19]Most of the mentioned lectures and related practical courses are in german. So lecture titles are translated.

Advice Nr. 3: Let your students play within their virtual lab in presence phases, at home, at university, wherever they want. **Cloud computing gives flexibility. Use it!** You are not bounded to presence phases any more if training phase is passed.

Advice Nr. 4: Watch your costs! We identified two main cost drivers. First cost driver is instance operating hours. Whenever an instance is running money flows. Make this clear to students by setting up rules. *Rule Nr. 1: Shut down all instances after finishing your experiments.* Some cloud providers have APIs to control resources. You may think about developing garbage collector scripts shutting down every running instance at midnight? This will discipline even lazy students. Most cloud service providers provide different instance types which can be formalised in another rule. *Rule Nr. 2: Use the smallest instance types for experiments. It is always possible to upscale.*

Second most relevant cost driver is data storage generated by not running instances. *Rule Nr. 3: Use as few instances as possible. Every instance produces storage costs.* Some cloud providers provide detailed access control settings. Use them to let no group (account) instantiate more than a limited amount of instances in parallel.

Advice Nr. 5: And to **increase a general cost awareness** it is good to establish another rule. *Rule Nr. 4: If two groups show the same performance. The one with less costs will get the better grade.* So resource consumption will not become a key criteria for grading. But our experiences show that this rule disciplines a lot.

Table 3 summarizes all relevant rules to be considered by virtual lab users.

Table 3: Cost control rules for Virtual Lab users.

Rule	Description
1	Shut down all instances after finishing experiments.
2	Use the smallest possible instance types for experiments.
3	Delete all unnecessary stopped instances.
4	Generated costs are considered for grading.

6 SUMMARY AND OUTLOOK

Not mentioned so far but very important. **Cloud based virtual labs are scalable.** The here presented approach is working with 10 students. It is also working with 100 or even 1000 students. We only have to provide more virtual labs but do not have to invest into dedicated hardware with a typical three or five year or even longer financial commitment. That provides flexibility and provides the possibility to manage periods of time which a significant step-up or step-down of students.

Nevertheless not all practical courses are virtual lab compatible. This might be of technical, functional or of economical reasons. This contribution presented a **pragmatical four step decision making model** for or against IaaS based virtual IT labs for educational purposes primarily from an economical point of view. The decision making is pragmatical and inspired by (Weinmann, 2011). We applied this decision making model (see section 3) in a concrete use case of practical educational labs in the higher education domain (colleges, universities, etc., see section 4) and showed that it can be very economical to use cloud based educational labs. It turned out that cloud based educational labs (in the analysed use case) have a more than **25 to 50 times cost advantage** (see section 4.3 [step 3]) to classical dedicated approaches. So cloud computing seems to be a very promising and **economical variant of providing educational labs** for university or college practical courses which is mainly due to an inherent peaky usage characteristics of practical university or college courses (see figure 3). This is a very essential finding for the Lübeck University of Applied Sciences and might be of **interest for other colleges, universities or IT training facilities as well.**

Furthermore this contribution refined some general findings to identify cloud compatible lectures and practical courses (see section 5.1) and showed that the approach of virtual labs is transferable to a significant amount of practical courses or lectures in computer science study programs (see section 5.2). Other study programs were not analyzed but the here presented virtual lab approach is not necessarily bound to computer science study programs. Our transferability analysis showed that all courses with requirements of providing databases, distributed processing capabilities, parallel processing or large-scale data processing are interesting candidates. These requirements might be also typical for engineering study programs.

Finally this contribution refined some advices for applying virtual labs in practical courses and proposes some easy mind-keeping rules for efficient cost control (see section 5.3 and table 3).

Outlook. In our ongoing research we plan to transfer the here mentioned cloud based virtual lab approach to more lectures and their related practical courses in order to double check our conclusions and to harden our postulations. By collecting more usage as well as billing data from different courses we hope to get an even better understanding of relevant cost driving parameters. By collecting this data we hope to improve a more precise cost estimation. It is furthermore planned to develop a kind of management software for setting up and managing virtual labs. A key feature would be an automatic average to peak ratio calculation based on provided usage data by a cloud service provider in order to answer the question whether a cloud based approach is more economical than a dedicated one.

ACKNOWLEDGEMENTS

Thanks to Amazon Web Services for supporting our ongoing research with several research as well as educational grants. Thanks to my students and Michael Breuker for using cloud computing in practical education. This contribution would not exist without their engagement. Let me thank Alexander Schlaefer and Uwe Krohn for organizing the World Robotic Sailing Championship 2011 in Lübeck and their confidence in our students.

REFERENCES

Barr, J. (2010). *Host your Web Site in the Cloud*. sitepoint.

Cayirci, E., Rong, C., Huiskamp, W., and Verkoelen, C. (2009). Snow leopard cloud: A multi-national education training and experimentation cloud and its security challenges. In *LNCS CloudCom 2009*, pages 57 – 68. Springer.

Dong, B., Zheng, Q., Qiao, M., Shu, J., and Yang, J. (2009). Bluesky cloud framework: An e-learning framework embracing cloud computing. In *LNCS CloudCom 2009*, pages 577–582. Springer.

Ko, C.-C. and Young, S. S.-C. (2011). Explore the next generation of cloud-based e-learning environment. In *LNCS Edutainment 2011*, pages 107–114. Springer.

Kratzke, N. (2011a). Cloud-based it management impacts. In *Proceedings of CLOSER 2011*, pages 145–151.

Kratzke, N. (2011b). Overcoming ex ante cost intransparency of clouds. In *Proceedings CLOSER2011 (special session on Business Systems and Aligned IT Services - BITS 2011)*, pages 707–716.

Kratzke, N. (2012a). Cloud computing costs and benefits (to be published). In Ivanov, van Sinderen, and Shishkov, editors, *Cloud Computing and Service Science*, Lecture Notes in Business Information Processing. Springer.

Kratzke, N. (2012b). What cost us cloud computing? how to decide for or against iaas using the average to peak ratio atp (to be published). In *Proceedings of CLOSER2012*.

Li, N. and Li, G. (2011). Research on digital campus based on cloud computing. In *Proceedings of CESM2011*, pages 213–218. Springer.

Liang, P.-H. and Yang, J.-M. (2011). Virtual personalized learning environment (vple) on the cloud. In *Proceedings of WISM 2011 (LNCS)*, pages 403–411. Springer.

Masud, A. H. and Huang, X. (2011). Esaas: A new education software model in e-learning systems. In *Proceedings of ICCIC 2011*, pages 468–475. Springer.

Mazhelis, O., Tyrväinen, P., Eeik, T. K., and Hiltunen, J. (2011). Dedicated vs.. on-demand infrastructure costs in communications-intensive applications. In *Proc. of the 1st international Conference on Cloud Computing and Services Science (CLOSER2011)*, pages 362–370.

Mell, P. and Grance, T. (2011). The nist definition of cloud computing (draft). Technical report, National Institute of Standards and Technology (U.S. Department of Commerce).

Oberg, M., Woitaszek, M., Voran, T., and Tufo, H. M. (2011). A system architecture supporting high-performance and cloud computing in an academic consortium environment. *Computer Science - Research and Development*.

Talukader, A. K., Zimmermann, L., and Prahalad, H. (2010). Cloud economics: Principles, costs and benefits. In *Cloud Computing - Computer Communications and Networks*, volume 4, pages 343–360. Springer.

Thiébaut, D., Li, Y., Jaunzeikare, D., Cheng, A., Recto, E. R., Riggs, G., Zhao, X. T., Stolpestad, T., and Nyguyen, C. L. T. (2011). Processing wikipedia dumps - a case-study comparing xgrid and mapreduce approaches. In *Proc. of the 1st international Conference on Cloud Computing and Services Science (CLOSER2011)*.

Truong, H.-L. and Dustdar, S. (2010). Cloud computing for small research groups in computational science and engineering. *Computing*.

Walker, E., Brisken, W., and Romney, J. (2010). To lease or not to lease from storage clouds. *Computer*, pages 44–50.

Weinmann, J. (2011). Mathematical proof of the inevitability of cloud computing. http://www.JoeWeinman.com/Resources/Joe_Weinman_Inevitability_Of_Cloud.pdf.