DOI: 10.15439/2014F465
ACSIS, Vol. 3

Identity Providers-as-a-Service built as Cloud-of-Clouds: challenges and opportunities

Diego Kreutz
LaSIGE/FCUL, Lisbon, Portugal
Email: kreutz@ieee.org

Eduardo Feitosa
IComp/UFAM, Manaus, Brazil
Email: efeitosa@icomp.ufam.edu.br

***Abstract*—In our previous work we designed and evaluated the feasibility of highly secure and dependable identity providers (IdPs) for the increasing requirements of future IT infrastructures. In this position paper we extend our previous work by analyzing and discussing the benefits of deploying highly secure and dependable identity providers-as-a-service (IdP-as-a-Service), without compromising the confidentiality of sensitive data and operations. In order to achieve this goal, we discuss some of the forefront challenges of deploying IdP-as-a-Service as a cloud-of-clouds model to ensure important properties such as the resistance against different types of threats and attacks, arbitrary faults, and make it more realistic to improve the system availability up to the three-nines mark. Notwithstanding, the main opportunities towards IdP-as-a-Service are also analyzed. We finish the paper proposing a sustainable business model based on our previous deployments and results, showing that it can be a win-win opportunity, i.e., both IdP-as-a-Service providers and customers can benefit from it.**

Keywords*—identity providers, IdP-as-a-Service, business model and opportunities, security, dependability, high availability, cloud providers, multi-cloud, telco cloud, hybrid cloud.

I. Introduction

FUTURE IT infrastructures will further combine and foster the interoperability of several computing, storage and networking technologies, such as those driven by new concepts and architectures like software-defined networks (SDN), software-defined storage (SDS), software-defined computing (SDC) and software-defined management (SDM), orchestrated by the control elements of software-defined environments (SDE) [1], [2]. In other words, on-demand provisioning will become the rule rather than the exception in all layers of current and future IT infrastructures. In particular, SDN was one of the key missing pieces to complete the SDE puzzle, enabling on-demand provisioning of network resources [1], [3].

In this new software-defined world (SDW), comprised by a myriad of advanced technologies and world-wide interoperability/integrations (e.g., cross-domain authentication, federations of authentication and authorization infrastructures, and so forth), enterprises (from small to large) can benefit from the on-demand service and IaaS provisioning offered by flexible and dynamic software-driven IT architectures and infrastructures. Different critical sectors, traditionally more resistant to changes, have been investing on cloud-based (outsourced) IT infrastructures and services, such as oil and gas industry and banks [4], [5].

Yet, some of the major challenges are related to the need of providing and ensuring higher degrees of security and dependability on different types of systems [6], [7], [8], [9]. Strictly speaking, it is essential to have in mind that the dynamic provisioning of IT services for future environments will have to consider different crucial aspects, such as performance, high availability, confidentiality, privacy, fault tolerance, and automated security, also from a conceptual and design perspective and not only as optional bolted-on features appended to the systems in an ad-hoc mode when it is already too late, i.e., data leakage or intrusions have already happened [6], [10], [11]. In fact, security and dependability of essential services is becoming a first class concern for enterprises that depend on systems connected to the Internet. One of the reasons is that the number and criticality of security threats have been rising at the same time that attacks are getting more sophisticated and challenging to deal with. Advanced persistent threats, large scale distributed denial of service and data leakage are becoming more frequent and dangerous to the enterprise business, governments and nearly all sorts of institutions [12], [13], [14], [15], [16].

Identity providers (e.g., OpenID providers) are not an exception. Recent research indicates that there is a significant gap on the provisioning of highly secure and dependable IdPs, representing one of the top concerns for future IT infrastructures [17], [18], [19]. Therefore, to best of our knowledge, we are proposing the first IdP-as-a-Service based on a cloud-of-clouds model for deploying highly secure and dependable IdPs. This can be achieved by combining different advanced techniques from distributed systems, dependability and security. Furthermore, we have already shown how it is possible to take advantage of multi-infrastructures (e.g., data centers) to increase the robustness of the system for tolerating different types of faults and attacks [18], [19], [20].

An OpenID-as-a-service has been proposed before [21]. However, it fails at addressing several security and dependability issues of current identity providers, as we have further depicted in our previous work [18], [19], [20]. Moreover, it uses only a single cloud, based on OpenStack, to scale the OpenID service, which is susceptible to several threats and performance issues [7], [8], [22].

Our main contribution is a cloud-of-clouds IdP model for achieving the technical and financial requirements of future IT infrastructures. In other words, the model has to be capable of taking advantage of the benefits provided by diverse infrastructures and still being cost effective, i.e., represent an interesting business opportunity for both providers and customers. Therefore, in this position paper we discuss the five essential challenges for deploying IdP-as-a-Service on a cloud-of-clouds model. First, deployment and operation challenges, advantages and disadvantages on different scenarios, such as collocation, private cloud, public cloud, and telco cloud are discussed. Second, we analyze the main trade-offs of different

© 2014, PTI

deployment alternatives. Third, we introduce the challenges and potential solutions for ensuring the confidentiality and privacy of sensitive data in a cloud-of-clouds environment. Forth, we dissect the main challenges of small and medium enterprise on building highly secure and dependable systems for future IT infrastructures. Finally, we provide a step-by-step cost analysis, based on our previous deployments and a real IT infrastructure use case, of the IdP-as-a-Service model and the new opportunities for both providers and customers.

II. Problem Statement

IdPs are arguably important services of current and future IT infrastructures. However, existing solutions have several vulnerabilities and weaknesses, being incapable of ensuring high levels of security and dependability required by the new and dynamic world of enterprises that day-after-day depend more heavily on IT systems. For instance, currently available and deployed identity providers can be threatened by different advanced persistent threats, large scale DDoS attacks, security breaches caused by software or infrastructure vulnerabilities, and so forth [12], [13], [15], [16], [17].

RADIUS and OpenID are examples of services with different weakness regarding security and dependability [17], [18], [19], [20], as summarized in Table I. Current implementations and deployments are highly susceptible to: (i) common vulnerabilities in different parts of the IT stack; (ii) sensitive data leakage due to the fact that keys and certificates are commonly stored in the operating system's file system; and (iii) resource depletion attacks when deployed on the same physical server with current virtualization technologies (e.g., Xen hypervisor).

Table I. Vulnerabilities and properties.

Vulnerability/Support	RADIUS	OpenID
Tolerates crash faults (using back-end clusters)	Yes	Yes
Tolerates arbitrary faults	No	No
Tolerates infrastructure outages	No	No
Tolerates DDoS attacks	No	No
Risk of common vulnerabilities	High	High
Risk of sensitive data leakage	High	High
Diverse security-related vulnerabilities	Yes	Yes
Susceptible to resource depletion attacks	Yes	Yes

Another issue worth mentioning is the fact that those services are commonly deployed in a single physical infrastructure (e.g., single physical machine or multiple servers in a single data center). This is still a wide spread practice in most enterprise environments, with some forefront exceptions such as cloud providers, which have several data centers in different locations. Notwithstanding, there is also a very high inherent complexity in deploying and managing current networks and services. This fact has been one of the main propeller of initiatives towards outsourcing of middleboxes [23], routing control logic [24], among other network functionality [25], [26].

In an interconnected and interdependent world, an energy outage in one IdP could impact many uses in distinct locations. This is the case of *eduroam* [27], where an energy outage in one university can simply deny the access of thousands of users to resources geographically dispersed. Therefore, it is of paramount importance to design IdPs that can be easily deployed across multiple physical infrastructures, such as different data centers or clouds.

This scenario is becoming even more critical once identity providers are being deployed for ensuring the security (e.g., access control) of large-scale distributed virtual networks, controlling not only users but also virtual routers and virtual machines [28]. Therefore, their security and dependability is a crucial issue for ensuring different security, availability and safety properties of future IT infrastructures [6], [29].

To tackle with some of those issues regarding security and dependability of critical services, we have proposed and evaluated the feasibility of resilient and trustworthy IdP services [18], [19], [20], [30]. While these system designs and implementations solve different of the afore mentioned issues of OpenID and RADIUS-like services, they pose also additional challenges for small and medium enterprise. Highly secure and dependable services are not simple and easy to develop, deploy and operate because they require different high skilled professionals (e.g., security specialists, distributed systems experts, system operation specialists, and highly skilled network operators) that are costly to afford and maintain, in particular for enterprise that do not have IT as their main business goal. Notwithstanding, both the IT infrastructure and human resources are much more expensive for small and medium enterprise when compared to larger companies [31].

At a first glance, one could think of outsourcing identity providers to one cloud provider. But, there are also several limitations of cloud providers that can have a significant impact on the system security and reliability, such as privacy, confidentiality, product/cloud provider lock-in, possibility of data loss, application interfaces and interoperability, geographical location of data, higher levels of failure when compared to traditional in-house machines, and so forth [7], [8], [22], [32].

Therefore, in this position paper we argue for secure and dependable IdP-as-a-Service using a cloud-of-clouds model. For different applications, such as storage, database systems, and experimentation testbeds, it has already been shown that multi-cloud systems can help to free the consumer from technical and business failures that any single cloud provider might be experiencing, prevent vendor lock-in, increase security and dependability, and reduce overall service cost by choosing the most cost-effective cloud providers [33], [34], [35], [36].

III. Towards Secure and Dependable IdP-as-a-Service

A. A Resilient and Secure IdP Architecture

Previously, we depicted a functional model, system design artifacts and essential techniques required for developing and deploying more secure and dependable identity providers for future IT infrastructures [18], [19], [20]. Furthermore, we analyzed some of the main trade offs of deploying robust and resilient services on a single physical machine or on multiple physical infrastructures. Whereas the performance can be boosted when deploying the system on a single physical infrastructure (e.g., data center), a deployment on multiple physical infrastructures (e.g., three different data centers) can significantly augment the system's resistance against physical and logical failures, first class DDoS attacks and resource depletion attacks, i.e., improve in orders of magnitude the system robustness and overall availability.

Figure 1 illustrates our proposed architecture with replicated components for higher availability. The IdP service replicas are capable of tolerating arbitrary faults, i.e., any accidental or malicious faults such as those caused by different types of attacks. Compared with a tradicional identify provider architecture (e.g., OpenID), two new components, IdP gateways and secure elements, are employed to safeguard the authentication systems without compromising backward compatibility. Further details regarding the system elements, protocols and technologies can be found in our previous work [18], [19], [20].

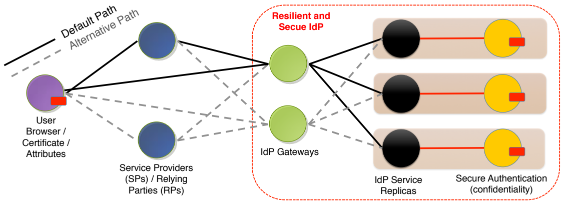
Default Path
Alternative Path
User Browser / Certificate / Attributes
Service Providers (SPs) / Relying Parties (RPs)
Resilient and Secue IdP
IdP Gateways
IdP Service Replicas
Secure Authentication (confidentiality)

Figure 1. Resilient and secure IdPs (based on the OpenID 2.0 model).

Figure 2 illustrates a deployment of our prototype in a multi-cloud enviroment. In this case, there are five clouds, being three public clouds (IdP replicas: IdP-R1, IdP-R2, IdP-R3), one private cloud (gateway GW1) and one colocation (gateway GW2). Additionally, there are also two service providers (SP1/RP1 and SP2/RP2), each of them running on a different public clouds, and end users using the IdP service to access the resources provided by SP1 and SP2, respectively.

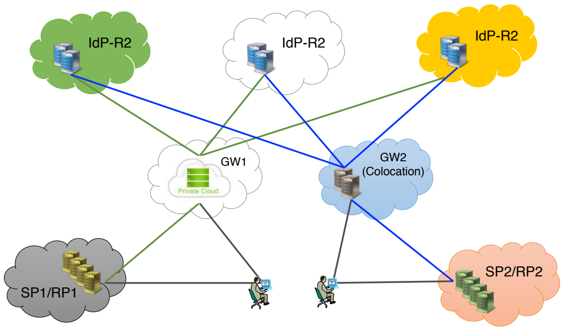

Figure 2. IdP-as-a-Service build as a cloud-of-clouds (or hybrid infra).

This deployment example is only to give an idea of the possibilites realized through our proposed architecture. In fact, we have already experienced its deployment across multiple data centers geographically distributed.

B. Multi-environment Deployment Trade Offs

Our resilient IdP architecture is designed to be deployable on diverse environments. A deployment in a single data center can achieve a high throughput (e.g., 2,200 authentications/s), while a multi-data center deployment can impose a higher communication latency, reducing the number of authentications/s [19], [20]. Yet, a multi-data center deployment can ensure higher levels of availability and provide protection mechanisms for resisting against different kinds of resource exhaustion attacks, large scale DDoS, and so forth [19], [20]. Therefore, decide between one or another deployment depends on the specific requirements of the enterprise, i.e., the IdP customer.

Figure 3 advances a step further our previous research and analysis [19] by extending the trade offs to multi-data center and multi-cloud environments. Here we only consider deployments targeting high availability and resistance against multiple types of threats and attacks. As we previously analyzed in details and numbers, the performance tends to degrade with the geographic distribution of machines due to (mainly) the increase in the network latency [20]. Different data centers of a single cloud provider can perform better than a multi-cloud setup if the provider has dedicated and efficient network links between the data centers. Otherwise, the performance of a multi-data center deployment will be similar to a multi-cloud environment. In this position paper, we assume only public networks in a multi-cloud configuration, which is still the most common case. Nevertheless, we are also aware that this scenario is likely going to change in a near future due to the advancements being fostered by SDN, which have already reached the backbone of telco companies such as NTT.

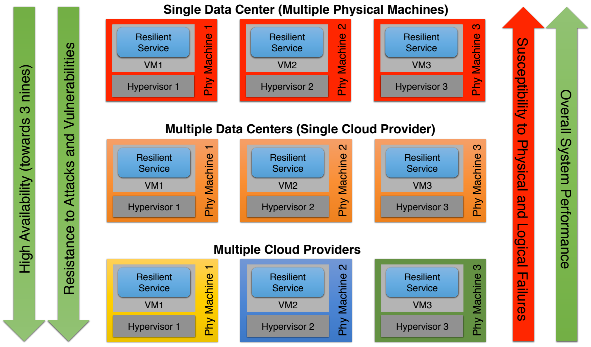

Figure 3. Diverse physical infrastructures: deployment trade-offs.

The susceptibility to physical and logical failures reduces with the increasing diversity of physical infrastructures. While one single data center can suffer an energy or communication blackout, it is much less likely to affect multiple data centers or cloud providers at the same time. Additionally, multi-environment configurations also increase the system availability. Moreover, multi-infrastructure deployments can give a considerable push towards achieving the "three-nines" mark of system availability [37].

Multi-infrastructure deployments also augment the system resistance against different types of attacks and vulnerabilities. Large scale DDoS attacks and common vulnerabilities (e.g., in security mechanisms, hypervisors, operating systems, physical network, physical servers, etc.) have less chance of affecting the system in a multi-cloud deployment. Each cloud provider has its own systems and protection mechanisms, making the task of any attacker much harder. In fact, some cloud providers have already shown how it is possible to protect cloud services against large scale DDoS by taking advantage of the huge amount of resources offered by different data centers and the right defense mechanisms, such as anycast methods [14], [38].

C. Where to deploy the system elements?

Should we use colocation, private cloud, public cloud, hybrid cloud or telco cloud? What are the most common *advantages and disadvantages of each choice?* There is no simple answer to these questions. Each environment is capable of providing different properties and benefits. Therefore, choose which one is the most adequate to deploy your IdP-as-a-Service is highly dependent on the specific requirements of the enterprise. For instance, while it is hard to ensure data confidentiality and privacy against a malicious sysadmin of a public cloud provider, to build a private cloud or rent a space in a data center for your own physical machines (colocation) will enable higher levels of protection regarding the confidentiality and privacy of sensitive data. But, at the same time, private clouds and colocation can also increase CAPEX and OPEX [39]. Differently, public clouds are less risky and more elastic in terms of scaling and costs.

There are IT companies specialized in offering colocation-as-a-service (CaaS), with data centers geographically dispersed, i.e., customers can rent spaces in different locations to increase their system availability. However, the main drawbacks of this approach is that you still need to manage your own infrastructure, as well as have a plan for scaling, which can be tricky and susceptible to the market reactions to the enterprise's products and services.

Cloud providers (e.g., Figure 4) are a much more flexible (and impose less risks) alternative to most enterprises. One can simple start with a single computing unit and dynamically scale up (on-demand) to thousands of nodes. Consequently, this reduces risks, CAPEX and OPEX. In the worst case, if the business fails, it is as simple as to scale down to zero resources. On the other hand, it would be harder get rid of the infrastructure if using colocation or building your own private cloud. Independently of the business success or not, you would still have the IT infrastructure. Eventually, if the business fails, sell it to a cloud provider at a relatively low price compared to the CAPEX and OPEX spent in the IT infrastructure.

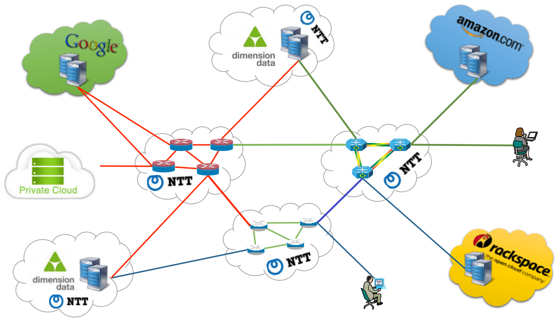

Figure 4. IaaS: private cloud, public cloud, hybrid cloud or telco cloud?

Typical cloud providers, such as Google, Amazon, and RackSpace, allow customers to choose between public and private clouds. As an example, Amazon have special contracts for companies willing to have their own private cloud, with higher security guarantees (e.g., confidentiality, privacy, and so forth). In spite of the fact that you are not the owner of the private cloud, you have contractual guarantees, i.e., the contract is the critical piece of the puzzle. Alternatively, it is also possible to build your own private cloud(s). Going a bit further, it can be interesting to consider a hybrid model, i.e., use your (potentially small) private cloud for critical and sensitive data and operations, while the less sensitive data and services of the business can be deployed on a public cloud. The main advantages of private clouds are with respect to security properties such as confidentiality and privacy, which are a bit harder to achieve or guarantee in public clouds.

Interestingly, another type of cloud model is rising, the telco cloud [40]. For a long time, telecom companies have not realized the opportunity to become big IT players (e.g., cloud providers). A typical telecom company owns large network infrastructures and has data centers in different locations. Consequently, these companies have already the main building blocks required to become cloud providers. More than that, Telcos have the network in their hands. By interconnecting their data centers through dedicated high speed links and offering on-demand network provisioning for their customers, these companies have the unique opportunity to become big cloud provider players in this competitive market. No other cloud provider (with eventual few exceptions) is capable of providing flexible, scalable and dynamic network services (inter-data centers) as telecom companies are. In fact, some global Telcos like NTT have already realized this opportunity and jumped into the cloud provider market by offering both computing and network resources provisioning as you go. For instance, NTT has more than 140 data centers spread across the world and interconnected by their own network infrastructure, providing global virtualization services [41].

Telco clouds have a clear advantage on the provisioning of network-as-a-service (NaaS), network-level security and QoS at the network level. This is particularly true between geographically dispersed data centers or cloud infrastructures. Most of the other cloud providers cannot afford to have their own network infrastructure between data centers spread across the globe. Therefore, they are tied to the contracts, services, prices and transport technologies of carrier network providers. Yet, inside a single data center, all cloud providers have the same capabilities of providing an infrastructure capable of offering NaaS and ensuring security and QoS properties.

D. Confidentiality and Privacy: more challenges ahead

Ensure the confidentiality and privacy of sensitive data and operations in public clouds is one of the major challenges for deploying IdP-as-a-Service. Therefore, multi-cloud deployments require new methods and techniques for ensuring the confidentiality of the most critical parts of the system.

In traditional deployments, server keys, session keys, authentication assertions, and so forth, can be easily compromised with access to the authentication servers. Therefore, we proposed in our previous work to split these services in two parts. The first is comprised by the traditional service/protocol itself, while the second is a special purpose component for isolating the essential operations and data required to authenticate and authorize clients. This isolation can be achieved through secure elements of different types, depending mainly on the specific requirements of the target environment [18], [19], [20]. The secure elements, proposed in our previous work, are specialized components that significantly enhance the security (in particular confidentiality) of the system without sacrificing system properties such as high availability and performance. To this purpose, we have designed a simple and common interface which can be easily implemented in different hardware or software platforms, such as grids of smart cards, tamper resistant FPGAs, virtual TPMs and highly secure hypervisors [30]. However, secure components are not quite suitable for all kinds of deployments in public or telco clouds. For instance, if you do not completely trust the cloud provider (even with a well-defined and strict contract), then you need alternative methods for protecting the confidentiality and privacy of your sensitive data and operations.

Some potential alternatives are tough available. However, each of them has certain requirements that can lead to low system performance, changes on the traditional protocol specifications or careful deployment and modification of some essential elements of our previously proposed architecture, such as the gateway. Let's briefly introduce and discuss a few possibilities and their implications, such as cryptographic cloud storage [42], shared cloud-backed file systems [35], encrypted database systems [43], [44], and secret sharing schemes [45], [46].

Cryptographic cloud storage [42] and shared cloud-backed file systems [35] can be used to provide integrity and confidentiality properties between users and the backend authentication servers. Those services offer secure and reliable storage functionality by ensuring properties such as privacy and confidentiality through multi-cloud environments and cryptographic schemes. However, to use them as an anchor of trust in an IdP-as-a-Service infrastructure, we would face at least three problems. First, shared and cryptographic storage protect only the data, but not the operations (e.g., sensitive cryptographic operations). Second, while on the client side they are simple to use (because they are designed to be Dropbox-like solutions), it is more tricky to adapt those solutions on the authentication backend components. In particular, the authentication protocols (e.g., RADIUS, OpenID) would have to be adapted. Third, these solutions require special purpose component (e.g., data processor) on the client side. This component is responsible for encrypting data (and ensuring a fair data distribution) before sending it to the multi-cloud storage system.

Encrypted databases [43], [44] can be considered as a second alternative. However, in order to effectively use an encrypted database on the backend authentication server for ensuring data and operations confidentiality and privacy, it would be necessary to change both clients and relying parties of OpenID-based architectures. Furthermore, the authentication protocol would have to be changed as well to accommodate the new authentication operations based on encrypted data.

Alternatively, secret sharing algorithms [45], [46] can be used to ensure confidentiality of sensitive data. Differently from the other two approaches, secret sharing does not impose backward compatibility problems. It poses only two challenges. First, the gateway element has to be modified in order to "join the shares" of the sensitive data (e.g., message signatures) from the different IdP service replicas. As secret sharing guarantees that with only t (threshold) shares one is able to build a valid piece of data, this would restrict the activity of any malicious cloud provider or attacker in possession of less than t shares of the secret. The second issue is that the gateway becomes now a critical element of the architecture because is assembles valid messages based on t shares received from different authentication replicas. Additionally, the gateway may also be the "dealer" of the secret sharing algorithm, i.e., the element responsible for distributing the shares to each replica or secure element of the system. Therefore, the gateway has to be protected, i.e., cannot be deployed on public or telco clouds. It should be deployed on private clouds or in collocation mode to ensure that the client (e.g., enterprise company A) is the only one with access and control over the gateway elements. This characterizes a hybrid cloud scenario.

E. Main challenges for small and medium enterprises

Secure and resilient IdPs are capable of supporting different types of threats, such as advanced persistent threats, large scale DDoS, physical and logical disruption, and ensuring critical security properties (e.g., confidentiality) [18], [19], [20]. However, to develop and deploy such kind of systems is not a simple or easy task. It requires different specialized skills in security, distributed systems, networking, systems operation, and so forth. Yet, most of small and medium enterprises are not able (or willing) to afford the cost of highly specialized IT teams to support those kind of advanced systems. In general, only large IT companies such as Google, Amazon, Rackspace, Microsoft, among others, can afford to spend considerable sums of money on highly skilled teams. Moreover, as it has already been shown, the overall infrastructure (e.g., CPU, storage, etc.) and human resources costs reduces significantly with the size of the enterprise [31]. For instance, in large scale IT providers the server admin ratio is 800 to 1k, while in small enterprises this ratio is approximately eight times smaller. This means that small enterprises have a nearly 8x higher OPEX cost considering only human resources.

Furthermore, we also need to add other variables, which add extra complexity to the provisioning of highly secure and dependable services across multiple infrastructures. Examples include the variation in cost models from provider to provider, unclear contracts (i.e., do not specify all the details and tools available to the customer), significant performance variation and quality of service guarantees between different cloud providers, diversity of technical tools and resources for deploying systems across multiple infrastructures, different levels of failures between distinct providers, and so forth [31], [47], [48]. In other words, there are too many risks and costs in building and owning their own secure and dependable identity providers for small and medium enterprises.

On the other hand, large IT providers already have large and globally spread physical infrastructures and highly skilled IT teams in various areas, such as operating systems, distributed systems, security, database systems, and so forth. Therefore, for them it is not costly or risky to provide new kinds of systems-as-a-service. On the contrary, this can add more revenue opportunities to their portfolio of products and services.

IV. Secure and Dependable IdP-as-a-Service: a win-win opportunity ahead

In this section we start by summarizing the first experimental results regarding to the performance of our prototype implementation, a multi-environment deployable secure and dependable identity provider. After that, we discuss the scaling capacity of the system based on data analysis of a real enterprise environment. Following, we analyze, discuss and propose secure and dependable IdP-as-a-Service as a viable and interesting win-win opportunity for cloud providers, new IT startups and customer, i.e., normal enterprises.

A. First experiments and results

Table II summarizes the first results of deploying our resilient and secure IdP prototype on three distinct computing environments, one single physical machine (UFAM-VMs), one single data center with multiple computing instances (Amazon-EC2, using `m3.xlarge` nodes [49]), and multiple data centers (Amazon-ECs, data centers of N. Virginia, N. California and Oregon). A complete description of the environments and results regarding fault tolerance and attacks can be found in our previous work [30], [20].

Table II. AUTHS/S WITH 20, 40, 80 AND 100 CLIENTS

Environment	20 clients	40 clients	80 clients	100 clients
UFAM-VMs	867.73	984.59	995.12	960.11
Amazon-EC2	1969.17	2166.58	2244.30	2244.04
Amazon-DCs	26.66	50.72	92.42	114.05

The best throughput is achieved by the Amazon-EC2, reaching over 2,200 OpenID 2.0 authentications/s. The main difference between UFAM-VMs and Amazon-EC2 is the computing power of the nodes, which are twice 2x more powerful in the Amazon-EC2 environment. In the UFAM-VMs we achieved a throughput of approximately 1k authentications/s. Lastly, the inter-data center deployment (Amazon-DCs), despite using the same computing nodes of the Amazon-EC2 environment, achieved the lowest performance. The main cause of this substantial drop in performance was the network latency between data centers, which was around 94x higher than the latency in the other two environments. Nevertheless, the performance (up to 114 authentications/s with 100 clients) is still enough to support the demand of enterprises with thousands of users, as shown by our statistics analysis of a real environment (see next section). Another interesting thing to mention is the fact that we have already identified several optimizations and improvements that can be done to significantly increase the system performance in all three setups [30].

B. Discussing the capacity and scaling of the system

We used the authentication statistics of a real IT infrastructure to statistically estimate the capacity of our prototype, i.e., to estimate the number of users that it can support. The reference institution has two authentication systems, multiple OpenLDAP and Active Directory (AD) servers.

Both authentication services are used by almost all the services and protocolos of the institution, such as SMTP and HTTP servers, Windows system services and components, IEEE 802.1X in wireless infrastructures, and Web content management systems. The authentication of dozens of online systems, provided by the institution to approximately 11.5k users, is also integrated through OpenLDAP. Furthermore, all logons on Windows and Linux labs, as well as other PCs, are also controlled by these two systems.

By analyzing seven days of logs from the OpenLDAP and AD servers, we identified: (a) 143,907 authentications during the worst peak hour (OpenLDAP + AD authentications per hour), which means an average of nearly 40 authentications per second; (b) 118 authentications in the worst possible case (worst OpenLDAP peak second + worst AD peak second), which in practice did not happened because the hour/second was not the same for both systems; (c) less than 102 authentications/s throughout all the seconds of the analyzed period.

Table III summarizes the capabilities of our prototype considering all three environments. Furthermore, we also estimate the number of users supported by increasing the number of instances (gateways and replicas) of the system. For instance, two gateways and eight replicas, four replicas working with each gateway, can potentially increase the system's performance by 2x (nearly linear performance). In this case, it is necessary to split the users among the instances (e.g., 50%) and/or apply load balancing techniques [50] on the gateways and database sharding techniques [51] on the replicas.

Table III. SCALING UP TO 1M USERS.

Environment	10k users	100k users	500k users	1M users
UFAM-VMs	4.16%	41.66%	208.30%	416.61%
Amazon-EC2	1.78%	17.82%	89.11%	178.22%
Amazon-DCs	35.07%	350.72%	1753.61%	3507.23%

A total of 10k users requires an average of around 40 authentications/s, considering an environment similar to the one described. Consequently, all three environments support this demand, requiring only 4.16% of the computing power of UFAM-VMs, 1.78% of Amazon-EC2, and 35.07% of Amazon-DCs. The Amazon-EC2 environment is capable of supporting an IT infrastructure with more than 500k users. This takes into consideration only the current implementation of the prototype. Though, there are different optimizations that could be done on the system and computing environment, such as (a) use the most recent version of BFT-SMaRT [52], which has several performance and durability optimization; (b) use optimized pools of thread on the gateway; (c) use multiple gateways, since the replicas are capable of processing more than 70k raw messages per second [52]; (d) use more powerful computing nodes such as `m3.2xlarge` [49], which nearly double the computing power of the nodes used in Amazon-EC2 environment; and (e) send requests in batches between the gateway and replicas. Therefore, we can arguably say that our system can be extended for supporting environments with more than 1M users and/or networked devices and systems that require more resilient and trustworthy authentication services.

C. Towards IdP-as-a-service: a win-win opportunity

Next, we provide cost estimations and discuss the possibilities of building reliable third party IdPs using a cloud-of-clouds model. Based on our previous evidence and experiments, we believe that there is a promising opportunity for the provisioning of secure and dependable IdP-as-a-Service. Our discussion and cost estimations corroborate in demonstrating the feasibility and benefits of IdPs-as-a-service.

With our Amazon-EC2 environment, we are capable of handling authentication requests of IT infrastructures with more than 500k users without any further optimizations on the system. As one Amazon EC2 `m3.xlarge` node costs 0.45 dollars an hour, and we used five of those nodes, we have a total cost of 54 dollars per day (at full operation/capacity). In one year we would have to spend $17,541.90, which is barely enough to buy the machines with the required computing power. If we look at an average salary of a *security specialist* [53] ($46,000/year), we can easily conclude that the cloud infrastructure costs is much less than the average salary of a single *security specialist*. Yet, we are not considering the infrastructure and maintenance costs (CAPEX, OPEX, TCO). Consequently, the outsourcing of critical services to specialized companies (e.g., Amazon) would be an interesting option to consider.

The infrastructure (virtual servers) at Amazon costs 17,541.90 per year. If we add to that 50% to provide the service (e.g., an OpenID provider), we come up with a total spending of 26,312.85 dollars, which is still cheaper than to have your own infrastructure and human resources. Yet, to Amazon, it could be an attractive business because the most complicated and costly part is the infrastructure, which is already provided. Therefore, as the company already has the required expertise (specialized infrastructure operation and security teams, and so forth), it is reasonable to add just another 50% to provide a IdP-as-a-Service. Consequently, both Amazon and customers would benefit from the technical and business model.

If we think about large scale demands (e.g., millions of users), things get even more interesting. For instance, with two instances of our Amazon-EC2 environment, we are capable of supporting an IT infrastructure with 1M users. In this case, the infrastructure provisioning would cost 35,083.80 dollars per year, which is not enough to pay one single *security specialist*.

Table IV gives a roughly idea of the spending costs of using secure, dependable and optimized identity providers from third parties, which could be provided as a service by companies such as Amazon, Rack Space, Google, or by a startup building a cloud-of-clouds service. We calculated the costs following the estimations presented in Table III. The infrastructure costs represents the real market practice (values) of Amazon. As an estimative, we consider that a company such as Amazon, which already provides the elastic infrastructure, is providing as well the IdP service. Thus, we added a service cost of $0.055037 per user, which we think is a reasonable value based on some OPEX costs estimation, with a good margin of net revenues. In the end, as can be observed, we have an average cost (IaaS + service) of $0.090077 per user/year. This value can arguably be considered as inexpensive for the customer and profitably for the service provider. Therefore, we could easily increase the service costs to make the business (from the service provider perspective) even more attractive. However, it is worth emphasizing that the key issue is to optimize the resources of the IdP-as-a-Service, i.e., do not over allocate resources before the demand. An IdP-as-a-Service has to be highly elastic and dynamically allocate the resources based on the business growth and customer needs.

Based on our numbers, one single customer, considering the average number of authentications/s and a demand of around 500k users, would generate an income of $90,120.80 per year (89.11% of an Amazon-EC2 like setup + service cost) for the service provider. If we extrapolate the costs and assume that a single user costs half a dollar per year, which is still extremely cheap for the customer, we would reach an astonishing 250k dollars for one single customer with 500k users, making this business highly profitable. For instance, in practical terms, a company like Facebook, which has over one billion active users [54], can reach a profit of 1.86 billion dollars only from user generated content [55]. This means that such companies would most probably be able to pay $0.090077 for having outsourced reliable and secure authentication and authorization infrastructures. Of course that a company like Facebook, due to the fact that it already owns a huge computing infrastructure and specialized human resources, could opt to have its own IdP service. However, other companies, whose main business is not IT, such as eBay, PayPal, among many others, could more likely opt for outsourced high quality authentication and authorization services. Notwithstanding, most of small and medium scale businesses would easily benefit from outsources IdP services.

Table IV. COSTS ESTIMATION FOR IDP-AS-A-SERVICE.

Cost/Users	**10k**	**100k**	**500k**	**1M**	**10M**
IaaS cost	$350.40	$3,507.65	$17,541.90	$35,083.80	$350,838.00
Service cost	$550.37	$5,503.70	$27,518.50	$55,037.00	$550,370.00
Total cost/y	$900.77	$9,011.35	$45,060.40	$90,120.80	$902,169.00

It is worth emphasizing that the service cost estimated in Table IV represents the revenue of a potential startup company specialized in building and provisioning of secure and dependable IdP-as-a-Service. In other words, a single enterprise customer (e.g., online social network system/business) with 10M users would generate an income of $902,169.00, which is reasonable enough to keep a team of specialized IT engineers.

V. FINAL REMARKS

In our previous work, we have dissected how it is possible to build more secure and dependable identity providers, which are one of the key components for ensuring the security of most IT infrastructures and systems. By developing different prototypes (e.g., resilient RADIUS, resilient OpenID) we have shown that it is possible to tolerate arbitrary faults (e.g., energy disruptions, connectivity failures, common software vulnerabilities, and so forth) and different types of attacks. However, small and medium enterprises cannot afford highly specialized IT teams to deploy and operate sophisticated systems, capable of ensuring critical properties such as confidentiality (despite eventual intrusions or malicious sysadmins), privacy and high availability (e.g., three-nines). To overcome the complexity and costs of deploying secure and dependable IdPs, we proposed IdP-as-a-Service as a viable and interesting win-win opportunity. IdP-as-a-Service can be built as a cloud-of-clouds model to achieve high levels of availability, scalability, elasticity, cost-effectiveness and robustness against a large diversity of threats and accidental or intentional problems.

Based on our first analysis, taking into account data and statistics from real environments and deployments, we discussed how IdP-as-a-Service can represent a new opportunity for cloud providers (or startups) willing to invest in this market niche. Some of the main challenges to overcome are related to identify and deploy the most appropriate system components for achieving elastic environments, high performance, and high levels of security, in particular regarding confidentiality and privacy of sensitive data and operations.

Acknowledgment. We would like to thank Vinicius Cogo and Douglas Macedo for their feedback. This work is supported by EU's 7th Framework Program (FP7), through the project SecFuNet (FP7-ICT-STREP-288349), and by CNPq/Brazil, through grants 590047/2011-6 and 202104/2012-5.

REFERENCES

[1] S. Racherla, D. Cain, S. Irwin, P. Ljungstrom, P. Patil, and A. M. Tarenzio, *Implementing IBM Software Defined Network for Virtual Environments*. IBM RedBooks, May 2014.

[2] C. Dixon, D. Olshefski, V. Jain, C. DeCusatis, W. Felter, J. Carter, M. Banikazemi, V. Mann, J. Tracey, and R. Recio, "Software defined networking to support the software defined environment," *IBM Journal of R&D*, vol. 58, no. 2, 2014.

[3] D. Kreutz, F. M. V. Ramos, P. Verissimo, C. Esteve Rothenberg, S. Azodolmolky, and S. Uhlig, "Software-Defined Networking: A Comprehensive Survey," *ArXiv e-prints*, Jun. 2014.

[4] A. Khajeh-Hosseini, D. Greenwood, and I. Sommerville, "Cloud migration: A case study of migrating an enterprise it system to IaaS," in *CLOUD, 2010 IEEE 3rd International Conference on*. IEEE, 2010.

[5] G. Sattiraju, S. Mohan, and S. Mishra, "Idrbt community cloud for indian banks," in *ICACCI, 2013 International Conference on*, Aug 2013.

[6] D. Kreutz, F. M. Ramos, and P. Verissimo, "Towards secure and dependable software-defined networks," in *SIGCOMM HotSDN*, 2013.

[7] C. Morgan, "Cloud security concerns relating to the difference in infrastructure and data operational control," itSMF International, Tech. Rep., May 2011.

[8] L. M. Vaquero, L. Rodero-Merino, and D. Moran, "Locking the sky: A survey on iaas cloud security," *Computing*, vol. 91, no. 1, Jan. 2011.

[9] A. J. Kornecki, N. Subramanian, and J. Zalewski, "Studying interrelationships of safety and security for software assurance in cyber-physical systems: Approach based on bayesian belief networks," in *Proceedings of the FedCSIS*. IEEE, 2013.

[10] J. Torres, M. Nogueira, and G. Pujolle, "A survey on identity management for the future network," *IEEE Comm. Surveys Tut.*, 2013.

[11] P. Verissimo, N. Neves, C. Cachin, J. Poritz, D. Powell, Y. Deswarte, R. Stroud, and I. Welch, "Intrusion-tolerant middleware: the road to automatic security," *IEEE Security & Privacy*, vol. 4, no. 4, 2006.

[12] C. Tankard, "Advanced Persistent threats and how to monitor and deter them," *Network Security*, no. 8, 2011.

[13] "All about Stuxnet," 2013, http://stuxnet.net.

[14] M. Prince, "The DDoS that almost broke the internet," 2013, http://goo.gl/oeDrMY.

[15] Verizon, "Data breach investigations report," Tech. Rep., 2013.

[16] IBM, "IBM X-Force 2012 trend and risk report," Tech. Rep., 2013.

[17] S.-T. Sun, K. Hawkey, and K. Beznosov, "Systematically breaking and fixing openid security," *Computers & Security*, vol. 31, no. 4, 2012.

[18] D. Kreutz, H. Niedermayer, E. Feitosa, J. da Silva Fraga, and O. Malichevskyy, "Architecture components for resilient networks," SecFuNet.eu, Tech. Rep., 2013.

[19] D. Kreutz, O. Malichevskyy, E. Feitosa, K. R. S. Barbosa, and H. Cunha, "System design artifacts for resilient identification and authentication infrastructures," in *ICNS*. IARIA, 2014.

[20] D. Kreutz, E. Feitosa, H. Cunha, H. Niedermayer, and H. Kinkelin, "Increasing the resilience and trustworthiness of openid identity providers for future networks and services," in *ARES/ECTCM*. IEEE, 2014.

[21] R. Khan, J. Ylitalo, and A. Ahmed, "Openid authentication as a service in openstack," in *IAS*, Dec 2011.

[22] S. Wattal and A. Kumar, "Cloud computing - an emerging trend in information technology," in *ICICT*, Feb 2014.

[23] J. Sherry, S. Hasan, C. Scott, A. Krishnamurthy, S. Ratnasamy, and V. Sekar, "Making middleboxes someone else's problem: Network processing as a cloud service," *SIGCOMM Comput. Commun. Rev.*, vol. 42, no. 4, pp. 13–24, Aug. 2012.

[24] V. Kotronis, X. Dimitropoulos, and B. Ager, "Outsourcing the routing control logic: Better internet routing based on sdn principles," in *HotNets-XI*. ACM, 2012, pp. 55–60.

[25] S. K. Fayazbakhsh, M. K. Reiter, and V. Sekar, "Verifiable network function outsourcing: Requirements, challenges, and roadmap," in *Hot-Middlebox '13*. New York, NY, USA: ACM, 2013, pp. 25–30.

[26] G. Gibb, H. Zeng, and N. McKeown, "Outsourcing network functionality," in *SIGCOMM HotSDN*. ACM, 2012, pp. 73–78.

[27] GEANT & TERENA, "eduroam," 2012, https://www.eduroam.org/.

[28] D. M. F. Mattos and O. C. M. B. Duarte, "Authentication and access control architecture for software defined networks," in *WNetVirt*, 2013.

[29] D. Kreutz, A. Casimiro, and M. Pasin, "A trustworthy and resilient event broker for monitoring cloud infrastructures," in *IFIP DAIS*, 2012.

[30] H. Niedermayer, D. Kreutz, E. Feitosa, O. Malichevskyy, A. Bessani, J. Fraga, H. A. Cunha, and H. Kinkelin, "Trustworthy and resilient authentication service architectures," SecFuNet.eu, Tech. Rep., 2014.

[31] Y. Chen and R. Sion, "To cloud or not to cloud?: Musings on costs and viability," in *2nd ACM SOCC*. ACM, 2011.

[32] A. Rot and M. Sobinska, "It security threats in cloud computing sourcing model," in *Proceedings of the FedCSIS*. IEEE, 2013.

[33] M. Correia, "Clouds-of-clouds for dependability and security: Geo-replication meets the cloud," in *Euro-Par 2013: Parallel Processing Workshops*, ser. Lecture Notes in Computer Science, 2014, vol. 8374.

[34] M. AlZain, E. Pardede, B. Soh, and J. Thom, "Cloud computing security: From single to multi-clouds," in *HICSS*, Jan 2012.

[35] A. Bessani, R. Mendes, T. Oliveira, N. Neves, M. Correia, M. Pasin, and P. Verissimo, "Scfs: A shared cloud-backed file system," in *USENIX ATC*. USENIX Association, Jun. 2014.

[36] A. Hume, Y. Al-Hazmi, B. Belter, K. Campowsky, L. Carril, G. Carrozzo, V. Engen, D. Garcia-Perez, J. Jofre Ponsato, R. Kabert, Y. Liang, C. Rohr, and G. Seghbroeck, "Bonfire: A multi-cloud test facility for internet of services experimentation." Springer, 2012, vol. 44.

[37] D. Durkee, "Why cloud computing will never be free," *Commun. ACM*, vol. 53, no. 5, May 2010.

[38] M. Prince, "Ceasefires don't end cyberwars," 2012, http://goo.gl/Vkljbi.

[39] B. Golden, "Capex vs. Opex: Most People Miss the Point About Cloud Economics," 2009, http://goo.gl/peS91k.

[40] X. Zhiqun, C. Duan, H. Zhiyuan, and S. Qunying, "Emerging of telco cloud," *Communications, China*, vol. 10, no. 6, June 2013.

[41] NTT, "Data center - nexcenter," 2014, http://goo.gl/t7uwX3.

[42] S. Kamara and K. Lauter, "Cryptographic cloud storage," in *Financial Cryptography and Data Security*, 2010, vol. 6054.

[43] R. A. Popa, C. M. S. Redfield, N. Zeldovich, and H. Balakrishnan, "CryptDB: Protecting confidentiality with encrypted query processing," in *ACM SOSP*. ACM, 2011.

[44] Z. Dayioglu, "Secure database in cloud computing - cryptdb revisited," *International Journal of Info. Security Science*, vol. 3, no. 1, 2014.

[45] A. Shamir, "How to share a secret," *Commun. ACM*, vol. 22, 1979.

[46] L. Goubin and A. Martinelli, "Protecting aes with shamir's secret sharing scheme." in *CHES*, vol. 6917. Springer, 2011.

[47] A. Iosup, S. Ostermann, N. Yigitbasi, R. Prodan, T. Fahringer, and D. Epema, "Performance analysis of cloud computing services for many-tasks scientific computing," *IEEE Trans. Parallel Distrib. Syst.*, vol. 22, no. 6, Jun. 2011.

[48] A. Khajeh-Hosseini, I. Sommerville, J. Bogaerts, and P. Teregowda, "Decision support tools for cloud migration in the enterprise," in *IEEE CLOUD*. IEEE, 2011.

[49] Amazon, "Amazon EC2 Pricing," 2014, http://goo.gl/WNEVvS.

[50] M. Das, S. Yadav, A. Kandhare, S. Malpani, R. Rathinam, and J. Thiagarajan, "Load balancing by endpoints," Sep. 14 2011, uS Patent App. 13/232,894.

[51] R. Cattell, "Scalable sql and nosql data stores," *ACM SIGMOD Record*, vol. 39, no. 4, 2011.

[52] A. Bessani, J. Sousa, and E. Alchieri, "State Machine Replication for the Masses with BFT-SMaRt," FCUL, Tech. Rep., Dec. 2013.

[53] Indeed, "Security specialist salary," 2014, http://goo.gl/RN9nR3.

[54] D. Tam, "Facebook by the numbers: 1.06 billion monthly active users," 2013, http://cnet.co/1jWlKJM.

[55] L. Fisher, "How much do social networks make from user-generated content?" 2011, http://tnw.to/1CUUS.