为多云数据密集型应用程序定义Intercloud安全框架和体系结构组件

Defining Intercloud Security Framework and Architecture Components for Multi-Cloud Data Intensive Applications

Yuri Demchenko, Fatih Turkmen, Cees de Laat
University of Amsterdam, The Netherlands
{ y.demchenko, F.Turkmen, C.T.A.M.deLaat}@uva.nl

Mathias Slawik
Berlin Institute of Technology, Germany
mathias.slawik@tu-berlin.de

Abstract—**This paper presents results of the ongoing development of the Intercloud Security Framework (ICSF), that is a part of the Intercloud Architecture Framework (ICAF), and provides an architectural basis for building security infrastructure services for multi-cloud applications. The paper refers to general use case of the data intensive applications that indicate need for multi-cloud applications platforms that will require corresponding multi-cloud security services. The paper presents analysis of the general multi-cloud use case that helps eliciting the general requirement to ICSF and identifying the security infrastructure functional components that would allow using distributed cloud based resources and data sets. The paper defines the main ICSF services and functional components, and explains importance of consistent implementation of the Security Services Lifecycle Management in cloud based applications. The paper provides overview of the cloud compliance standards and their role in cloud security. The paper refers to the security infrastructure development in the CYCLONE project that implements federated identify management, secure logging service, and multi-domain Attribute Based Access Control, security services lifecycle management. The paper discusses implementation of the Trust Bootstrapping Protocol as an important mechanism to ensure consistent security in the virtualised inter-cloud environment.**

Keywords-Cloud Security, Intercloud Security Framework (ICSF); Intercloud Federations Framework (ICFF), Dynamic Access Control Infrastructure (DACI), Shared Responsibility Model, Cloud Compliance, Trusted Bootstrapping Protocol

I. INTRODUCTION

Modern cloud based data intensive applications tend to use and integrate services and resources from multiple clouds to allow data collection, transfer and services delivery to distributed or global customers [1, 2]. Current development of the cloud technologies requires the development of hybrid multi-cloud and Intercloud models, architectures and integration tools that could allow integrating heterogeneous cloud based infrastructure services into existing enterprise and campus infrastructures.

Demand for more complex and enterprise or project oriented use of clouds motivates the development of new service provisioning and security models that could allow creating complex project oriented and collaborative infrastructures provisioned on-demand and across multiple providers. As an example, bioinformatics is dealing with the genome sequencing which is compute intensive and often requires using distributed data sets and computing resources from multiple data centers or cloud providers.

Moving company's in-premises datacenter to cloud and using external cloud services requires careful security services and identity management design and deployment as well as well-defined data security infrastructure and protection policy.

Recent Cloud Computing trends analysis [3, 4] identified the growth of hybrid cloud infrastructures, that combine company's cloud infrastructure and involve multiple types of cloud services from different CSPs, as the main factor in changing cloud security paradigm that is becoming more reliant on cloud security services provided by CSPs and trust relations between customer and CSP. This makes the CSP compliance with the cloud security standards and regulations as an important enabling factor in ensuring consistent security in hybrid multi-cloud environment. Complexity of multi-cloud environment will create demand for the 3rd party security services such as cloud access security broker services and managed cloud security services that will interoperate with or be integrated into the enterprise cloud infrastructure [5]. Well defined Intercloud/multi-cloud security architecture will ensure correct consistent services across multi-cloud applications.

This paper presents the results of the ongoing development of the Intercloud Security Framework (ICSF), that is a part of the Intercloud Architecture Framework (ICAF) [6, 7] being developed by the authors. ICSF is intended to provide an architectural basis for building security infrastructure for multi-cloud applications. The paper defines the general multi-cloud use case for data intensive applications (using bioinformatics as such demanding area example) that motivates need for multi-cloud applications platforms that will require corresponding multi-cloud security services. The presented analysis confirms benefits of consistent implementation of the federated multi-cloud security model that can be potentially integrated with the federated access control and federated identify management widely adopted by the major cloud service providers.

The paper provides information about ongoing implementation of the cloud automation platform CYCLONE for multi-cloud applications integration [8] that develops the main multi-cloud infrastructure components proposed in this paper. such as currently implemented federated identify management using eduGAIN, secure shell login using eduGAIN federated identities, and new services being

978-1-5090-6611-7/17 $31.00 © 2017 IEEE
DOI 10.1109/CCGRID.2017.144

developed such as multi-domain Attribute Based Access Control, security services lifecycle management and trust bootstrapping for virtualised cloud environment.

The remainder of the paper is organized as follows. Section II describes the general use cases for multi-cloud data intensive applications that motivate the proposed security infrastructure. Section III summarises requirements to ICSF. Section IV defines the proposed Intercloud Security Framework (ICSF). Section V describes ICSF functional components that provide common security middleware services for multi-cloud applications and services. Section VI provides overview of the cloud compliance standards and how it can be used to ensure consistent security in tightly integrated cloud and enterprise infrastructures. Section VII described the implementation of the ICSF components in the CYCLONE project. The paper concludes with remarks on the future development in section VIII.

II. Use case for Data Intensive Multi-cloud Applications

This section provides background for discussing multi-cloud security requirements and definition of the Intercloud Security Framework (ICSF) in sections III and IV. The presented use case reflects the main infrastructure components for complex bioinformatics applications/workflows that require live remote cloud processing of sequencing data.

Bioinformatics represents one of the most demanding use cases for both high-performance computational infrastructure provisioning and applications deployment automation [2]. Bioinformatics generates huge amount of data produced by multiple research teams from the DNA sequencing. Decreasing prices for DNA sequencing that in a single case produces terabytes of information already cause problems for effective data management.

Bioinformatics deals with the collection and efficient analysis of biological data, particularly genomic information from DNA sequencers, which become increasingly distributed and may be hosted in different private and public or scientific clouds. The terabytes of raw data, produced by the sequencers for each run, require significant computing resources for analysis that may not be available locally. These sequencers are typically located at multiple specialized centers interconnected into bioinformatics Research Infrastructure (RI), while the collaborating researchers are distributed internationally. Some sequencing centers adopt cloud platform for storing data, large public CPS's and RI provides cloud based storage of genome data supporting also federated access control with the industry recognised Identify Providers.

Figure 1 shows the bioinformatics application deployed in Cloud1 in a form of Virtual Private Cloud (VPC) that includes both the actual application that manage the whole scientific workflow and computing cluster.

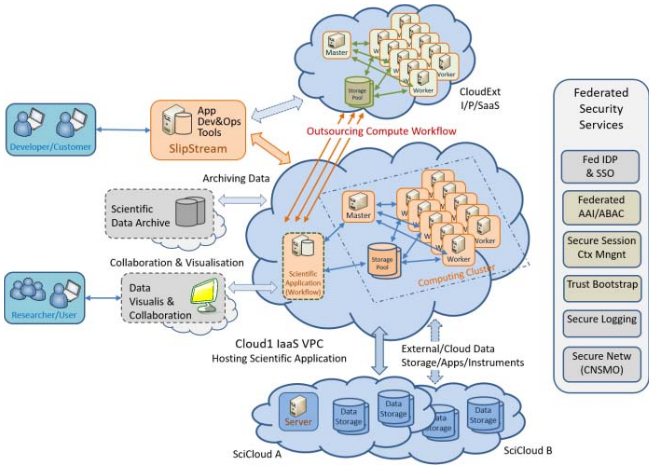

Figure 1. General use case for multi-cloud data intensive application infrastructure (using bioinformatics use case).

The bioinformatics engineer develops and deploys application in Cloud1 using development tools coupled or integrated with the SlipStream cloud automation tools. The application may use external scientific data and applications located in SciCloud A and B. In case of excessive workload, some computational tasks can be outsourced to external cloud CloudExt, in particular in a standard cloudburst scenario. Similarly to original use case definition, Figure 3 includes Scientific Data archive for storing obtained scientific results data. Application user bioinformatician researchers may use data visualisation and collaboration tools that all can be hosted in cloud and provided by specialised SaaS or cloud applications providers.

Suggested security services are combined into the federated security services stack, as depicted on the right side of Figure 1 that include general and specialized security services that are required for multi-cloud applications which are discussed in the next section.

III. General Requirements and Design Principles for Multi-cloud Security Services

The discussed above general use case for multi-cloud applications infrastructure allows us to specify the following general requirements and design principles to multi-cloud security services and Intercloud Security Infrastructure (ICSI) that incorporate and extend current best practices in cloud security [9, 10].

ICSF01. Multi-cloud security infrastructure should provide consistent access control, security credentials and security context management for multi-cloud applications deployment, operation and management, in general covering all application lifecycle, including applications secure session management.

ICSF02. Multi-cloud security services should allow users and applications (internally and on behalf of users) to access all distributed multi-cloud resources using single credentials that should be federated with the individual cloud credentials and access control mechanisms.

ICSF03. ICSI should support federated access control and resource management model, allowing integration with the cloud federation services.

ICSF04. Application based access control must be integrated with the cloud based security services and implement in a consistent way the shared security responsibility model that is defined and implemented by cloud services providers as a standard cloud services security model.

ICSF05. ICSI must ensure data protection during the whole data handling lifecycle, including data transfer between different clouds and security domains as well as data storage in-rest.

ICSF06. ICSI should provide secure trust bootstrapping for the provisioned on-demand cloud based security services that should bind the deployed security services to the applications runtime environment and virtualisation platform, to prevent unauthorised virtual environment cloning.

ICSF07. Security Services Lifecycle Management functionality must support the security context management during the whole security services lifecycle, including binding security context to the provisioning session and virtualisation platform.

ICSF08. Security session synchronization mechanisms should implemented to protect the integrity of the remote run-time environment, including secure session fail-over that should rely on the session synchronization mechanism when restoring the session.

ICSF09. ICSI should support Dynamic Security Associations (DSA) to provide fully verifiable chain of trust from the user client/platform to the virtual resource and the cloud provider platform.

ICSF10. SLA and compliance management, including initial SLA negotiation and further SLA enforcement, must be implemented at the planning/design and operation stages. This functionality can outsourced to and implemented as a part of the user controlled or brokered cloud automation platform.

ICSF11. Brokered and third party security services should ensure cloud compliance with general international and applications domain specific security standards; Cloud Service Broker should include explicit compliance assessment stage when provisioning brokered services.

The presented requirements and design principles use and leverage the best practices in security design of regular Internet and web applications however extend them necessary security mechanisms to ensure bootstrapping of the virtualised environment to the cloud platform. They also reflect changing security paradigm in complex cloud based applications and infrastructures from formal security models to trust based that is in its own turn based on compliance based built trust. Cloud customers must trust cloud services providers and cloud services providers are interested in complying with the industry verified security design principles and standards.

IV. Intercloud Security Framework (ICSF)

ICSF is a part of the general Intercloud Architecture Framework (ICAF) developed in authors' earlier works [6, 7] as a result of cooperative works in a number of European projects. The ICAF defines five complementary components addressing Intercloud integration and interoperability: multi-layer Cloud Services Model (CSM) that combines commonly adopted cloud service models, such as IaaS, PaaS, SaaS, in one multilayer model with corresponding inter-layer interfaces; Intercloud Control and Management Plane (ICCMP) that supports cloud based applications interaction; Intercloud Operation and Management Framework (ICOMF), Intercloud Federation Framework (ICFF), and Intercloud Security Framework (ICSF). All components interact and provide services to each other in an instant multi-cloud infrastructures, in particular ICSF and ICFF are interacting in many scenarios of the multi-cloud and inter-cloud services operation.

A. ICFF and ICSF

The ICFF and its federation models were proposed and described in details in early works of the authors [7]. ICFF provides functionalities for clouds from different administrative domains to create a federation. The federation allows for end-users to access cloud services from multiple domains without the need to obtain a separate identity, while services remain under control of their original operator or home provider. The main components of the federated Intercloud architecture, specifically underlying the Intercloud gateway function (GW), provide translation of the requests, protocols and data formats between cloud domains. At the same time, the federated Intercloud infrastructure requires a number of functionalities, services and mechanisms from ICSF to support its operation. It includes the following components operated jointly by ICFF and ICSF:

- Cloud Service Broker
- Trust Broker and Trusted Introducer
- Service Registry and Discovery
- Federated Identity provider (FedIDP)
- Service and/or inter-domain gateway.

ICFF requires federating customer and cloud provider access control and resource management services. We define two types of federation in cloud: customer side federation dealing with identity federation and access control, and provider side federation that enables using cloud resources from multiple providers [7]. In both cases the security and integrity of the federation is based on the trust establishment between federation members, which in cloud is established as a part or the infrastructure provisioning.

Access control, identity and trust management functions are defined as part of the Intercloud Security Framework that need to support the dynamic resource provisioning in multi-cloud environment.

B. ICSF functionalities and services

The ICSF defines a set of functionalities for identity and trust management, access control and secure communication in the multi-cloud environment. This should be provided in the form of dedicated services that are provisioned over virtual resources. ICSF follows a federated security model and may use ICFF infrastructure services for federating identities and using federated trust services.

The core ICSF functionalities include the followings:

- Policy based access control, security credentials and security context management for multi-cloud applications deployment, operation and management in a federated setting.
- Data protection during the whole data handling lifecycle, including data transfer between different clouds and security domains as well as data storage at rest. When relevant, certain security measures such as encryption should be put in place to remove the incentive for data thefts.
- Secure trust bootstrapping for the provisioned on-demand cloud based security services that should bind the deployed security services to the applications' runtime environment and virtualisation platform, to prevent unauthorised virtual environment cloning and enable secure management of keys/secrets.

The ICSF needs to provide the following functionalities for managing trust in the multi-cloud setting:

- Dynamic trust establishment between indirectly connected cloud entities. Current trust relationships between cloud entities typically established via manual distribution of PKI certificates. ICSF should support establishing dynamic trust relations from the multilateral SLA negotiation process without preliminary existing trust relations.
- ICSF identity management service that interoperate with the cloud deployed security infrastructure and cloud provider identity management services.

C. Security Services Lifecycle Management

In order to ensure the consistency of provisioned security services over virtual multi-cloud infrastructure, we employ the security services lifecycle management (SSLM) approach proposed by the authors in [11] for single cloud deployment. SSLM models the stages of a security service lifecycle from provisioning to decommissioning, and allows for systematic development and management. The lifecycle of each service instance is identified by a session id.

We confirm the importance of the two additional stages introduced in [11]: Reservation stage and Registration&Synchronisation stage, - for complex multi-cloud deployments to handle infrastructure dynamicity. With the wide use of cloud automation tools the binding of the provisioned security services context can be handled by the cloud automation tools. The latter refers to possible scenarios with the provisioned security services migration or failover.

The proposed SSLM incorporates recommendations from existing security lifecycle management frameworks, such as defined in the NIST Special Publication 800-14 "Generally Accepted Principles and Practices in Systems Security" [12] or Microsoft Security Development Lifecycle (SDL) [13], AWS Security design principles. The defined security services lifecycle includes the following typical phases: Initiation, Development and/or Acquisition, Implementation, Operation and Maintenance, and Disposal.

Figure 2 (a) illustrates the proposed Security Services Lifecycle Management (SSLM) model that reflects security services operation in generically distributed multi-domain environment and their binding to the provisioned services, which SLM stages are illustrated in Figure 2 (b). The SSLM includes the following stages:

- **Service request and generation of the Global Reservation ID (GRI)** that will serve as a provisioning session identifier and will bind all other stages and related security context.
- **Reservation stage** that also includes **Reservation session binding** with GRI what provides support for complex reservation processes including required access control and policy enforcement.
- **Deployment & Bootstrapping stage** begins after all component resources have been reserved and includes distribution of the security context and binding the reserved resources or services to the GRI as a common provisioning session ID.
- **Registration & Synchronisation stage** that specifically targets possible scenarios with the provisioned services restoration in case of their failure or migration. In a simple case, the Registration stage binds the local resource or hosting platform run-time processes ID to the GRI as a provisioning session ID.
- During **Operation stage** the security services provide access control to the provisioned services and maintain the service access or usage session.
- **Decommissioning stage** ensures that all sessions are terminated, data are cleaned up and session security context is recycled. Decommissioning stage may also provide information to or initiate services usage accounting.

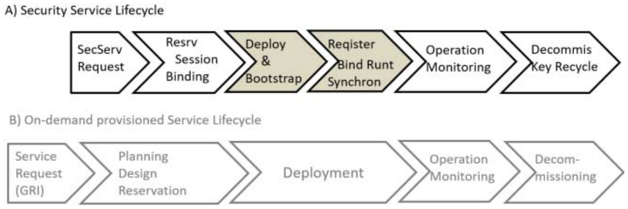

Figure 2. Security Services Lifecycle Management model for on-demand provisioned security services.

The proposed SSLM extensions can be considered as a part of the Deployment stage or they can be executed by the newly deployed application at its initial startup. The Registration & Synchronisation functionality ensures security sessions (re-)synchronization in case of application expansion to external cloud, failure restoration, or application migration (in the framework of the active provisioning session); it can also provide a mechanism for remote data protection by binding them to the session context.

V. ICSF Middleware components

A. Dynamic Access Control infrastructure

In order to address these requirements, the following components are considered in ICSF in the form of services where their implementation is an ongoing process:

- **Authorization Service**: Responsible for managing authorization policies and their enforcement when accessing to sensitive services and resources. The service should follow the existing standards such as eXtensible Access Control Markup Language (XACML) [14] to provide a high-level of interoperability within multiple clouds.
- **Context Management Service**: Manages the contextual information such as trust, delegations and environmental information (e.g. time of day) when enforcing access control policies. Separation of context management from the authorization service enables dynamicity in security enforcement.
- **Encryption and Token/Key Management Service**: Provides cryptographic services to the overall security infrastructure. Encryption of sensitive information at rest or transit, secure bootstrapping and trust establishment all employ this service for their functionalities. The implementation of this service may exploit existing work in this area.

These services serve as a backbone for the implementation of authorization functionality, trust management and information protection at different levels. Besides these basic services, the security infrastructure should support the management, verification and revocation of the security tokens for tenants.

The instantiation and deployment of these services over the cloud resources requires architectural considerations as well as application specific adaptations for efficiency, compliance and reliability.

B. Bootstrapping Trust in Federated Clouds

Trust bootstrapping refers to initialization of cloud nodes with relevant secrets. This functionality and service has been researched in the previous authors work [15]. Other research in this area [16] and [17] use functionality of the Trusted Platform Module (TMP) [18] and python-keylime library [19] for bootstrapping trust within cloud nodes and the services running on them.

In order to bootstrap cloud nodes with security keys and initialize them for integrity monitoring, the tenants rely on a service called Cloud Verifier (CV) [20] that acts as an intermediary between tenants and their nodes. CV is mainly responsible for periodically checking the integrity of resources and it can live in either tenant's or cloud provider's premises. There are three steps involved in keylime to establish trust between a tenant, a cloud node and an external entity called cloud verifier that monitors the nodes/applications for integrity:

Key Generation: The tenant creates a fresh symmetric key K_t for each new node it wants to request.

Node initiation: The cloud provider instantiates a new VM for the tenant with the information $\mathrm{Enc}_{K_t}(d)$. Here d is a form of initialization data such as a cloud-init script sent by the tenant.

Key Derivation Protocol: The final step involves the communications between the tenant, cloud node and cloud verifier for the exchange of keys. This step uses ephemeral keys, specific PCRs (i.e. PCR#16) and TPM_Quote() function to ensure that the keys are securely exchanged through an untrusted network.

CYCLONE project employs keylime [16, 19] for bootstrapping the trust between cloud nodes and the services running on them by trying to resolve several caveats. First of all, keylime has been implemented over Xen hypervisor's TPM features and does not support other hypervisors. The second issue is related to the availability of hardware TPMs on existing chipsets/motherboards. The trust bootstrapping features consume the encryption service described in the previous section to perform the cryptographic operations. Integrity monitoring functionality offered by keylime does not only provide integrity checks for the security services but also applications that run on the cloud resources.

VI. Compliance and Security

Compliance and security are related and in some cases interchangeable. Security is commonly defined as a set of technical, physical, and administrative controls in order to ensure normal operation of a system or application.

Compliance is a certification or confirmation that the system or an organization meets the requirements of the specified standards, established legislation, regulatory guidelines or industry best practices that can be jointly defined as compliance framework.

Why compliance is important for cloud? When moving to cloud, the organization moves from internal security and operational environment (that may not be formally defined) to external operational security that will become a part of SLA (or business requirement) with CSP. Compliance in this case will define the expected level of security and assurance.

When developing cloud based applications, the applications developer must analyse and ensure compliance of the end user applications with the industry related compliance requirements.

It is a common practice in cloud security that Cloud Service Provider (CSP) implements Shared Responsibility Model that splits responsibility for the security of different layers and components between CSP and a customer that can be cloud based application developer, or end user, or both.

As an example, Amazon Web Services (AWS) as an IaaS cloud provider ensures the security of the cloud infrastructure and cloud platform services: facilities, physical security of datacenter, network infrastructure, virtualisation platform and infrastructure. While the customer is responsible for security of the following components: Amazon Machine Instances

(AMI), OS, and applications, data in transit, data at rest, and data stores, credentials, policies and configurations. The customer is specifically responsible to comply with the Acceptable Use Policy (AUP), ensure correct use of the cloud platform, and for security update and patching of the guest OS and installed applications.

Data security and protection is also a shared responsibility that involves:

(1) Cloud provider responsibility to ensure secure data storage, processing and transfer and well as provide necessary security mechanisms to enable application level security;

(2) Application developer responsibility to correctly implement the application security in the cloud multi-tenant virtualised environment (often referred to as Security Development Lifecycle and defined by a number of industry standards and guidelines) to protect user data and personal information, integrate applications security with the provided cloud platform security services and mechanisms, and provide necessary and easy usable security services for end user to correctly use application security;

(3) End user responsibility to ensure security of their application access client (typically browser with hosting OS), access credentials and data.

A. *Compliance standards*

Cloud compliance is generally defined by the Cloud Security Alliance Guidance for Critical Area of Focus in Cloud Computing (CSA3.0) [21] that define 13 domains of the security concerns for Cloud Computing that are divided into two broad categories that define corresponding security controls for cloud governance and operation. The CSA GRC Stack (Governance, Risk Management and Compliance) [22] includes the Cloud Controls Matrix (CCM) and Consensus Assessments Initiative Questionnaire (CAIQ) [23] and other documents. CAIQ provides comprehensive tool that maps general IT and data protection controls and different industry specific requirements to CCM. In particular CAIQ includes mentioned above ISO/IEC 27001:2005, PCI DSS, SOC1-SOC3, FISMA, FedRAMP, EU GDPR, HIPAA/HITECH and in total to 32 different documents.

The cloud providers operating globally need to comply with the different regulations in different countries. In particular, this is important for European Union that has a strict data protection regulation. The new European General Data Protection Regulation (GDPR) adopted in May 2016 will go in action in 2018 and will require many data handling processes to be re-designed [24, 25]. The GDPR will be applied to all businesses and companies operated in the European Union, and would prohibit the transfer of personal data to non-European Union countries that do not meet new EU regulation. In order to bridge these differences in approach and provide a streamlined means for U.S. organizations to comply with the new Regulation, the U.S. Department of Commerce in cooperation with the European Commission developed a so-called “Privacy Shield” that comes in place of the former "Safe Harbor" framework [26].

B. *Compliance analysis and implementation*

The compliance of the cloud platform and applications is an important part of setting up and operating cloud based services. It can be also a part of the automated SLA negotiation and monitoring, in particular this functionality should supported by the Cloud Service Broker. The following sequence can be used for cloud compliance analysis and implementation:

1) Define applications in cloud

2) Identify what data will be moved to the cloud
   - for security and compliance reasons, organisations may decide that some highly confidential data will always remain on an internal network (or private cloud) and will not move to the public cloud

3) For the data moved to cloud, negotiate with the provider about
   - What type of data will reside on the consumer’s own/VPC cloud
   - Back up services
   - Possibility to audit
   - Incident report about data incidents

4) Check what compliance documents or industry best practices are used by CSP (see reference list mentioned above)

5) Check what eDiscovery services and tools are available from the cloud provider and develop incident response plan.

6) Define responsibility of all roles involved into data management and have a corresponding contacts on the cloud provider side

The cloud based application developer must consider all aspects of the security compliance to ensure that the final application provides consistent security including cloud platform security and application security from the point of view of the application end user. In particular, the ones related to access control, user identity management and user data protection.

VII. Federated Security Infrastructure implementation in the CYCLONE Project

CYCLONE project security infrastructure development present an example of application infrastructure evolution from single cloud implementation to multi-cloud operational infrastructure. In developing security services for cloud based applications, we focus on customer controlled security services and rely on the security compliance of the cloud platforms that is ensured by the providers. The CYCLONE security services are motivated by the CYCLONE use cases and are focused on such missing functionality as authentication and authorization for customer developed applications on all cloud layers (e.g., web-based single sign-on as well as SSH login) using federated identities in academic settings.

The CYCLONE security architecture relies on the lower layer and cloud infrastructure security services and provides applications related security services and practical tools that can be adapted and applied for specific applications and implementation platforms – either singularly or in combination. Modular construction and simplicity allows their reusability and composability, including their easy integration with the production-grade tools and established industry-recognized standards, e.g., Keycloak and OpenID Connect.

A. CYCLONE Approaches to Multi-cloud Security

The concrete approaches of CYCLONE to the topic of multi-cloud security have been presented in the project deliverable D4.2. There are two notable implementations that have significant value for the CYCLONE use cases: the CYCLONE Federation Provider [27] and the CYCLONE PAM module [28]. The CYCLONE Federation Provider enables the deployed SaaS solutions to authenticate the bioinformaticians using their federated identities. The second component, the CYCLONE PAM module, enables SSH remote shell access to deployed VMs based on federated identities. The next sections provide more details on the implementation of both components and contrasts them to other related tools.

a) The CYCLONE Federation Provider

While the Bioinformatics end users in the CYCLONE use cases are endowed with a federated identity, using this identity for authentication purposes is quite challenging as there are two main obstacles: First, registering a new application in eduGAIN is not automated in the majority of participating institutions. Depending on the concrete process, this can incur a high delay until applications are ready to be used. Second, each instance of an application needs to be registered with eduGAIN separately. As CYCLONE features a self-service platform where a large number of end users can deploy quite a few applications for themselves, registering each of these applications is not feasible.

Both of these challenges motivate the creation of the CYCLONE Federation provider [27] that can be best described as an “authentication proxy”. It is both a regular SAML 2.0 Service Provider for eduGAIN as well as an OpenID Connect Identity Provider to the relying applications. As new OpenID Connect clients can be created far more easily and rapidly than eduGAIN Service Providers, it accelerates the registration of new instances of federated applications considerably.

The CYCLONE Federation Provider is based on Keycloak that is a comprehensive Web SSO and IDM server. We extended Keycloak with a number of functionality that is required by the CYCLONE use cases, e.g., a periodic data privacy preserving removal process and a self-registration API that allows deployment scripts to automatically register new OpenID Connect Clients.

By using OpenID Connect, CYCLONE relies on a standard for Web SSO that is generally recognized, causing supporting libraries and software to be available widely. Besides the mentioned SAML2.0, a comparable approach would be using a Kerberos together with SPNEGO/GSSAPI HTTP authentication. However, there are a number of drawbacks, for example, configuring Kerberos clients on every user machine, required firewall exceptions for Kerberos communication, the need to setup a complete trust chain, and more. These are the reasons, why this approach would be infeasible in the CYCLONE environment.

b) The CYCLONE PAM module

To access deployed bioinformatics applications via SSH, e.g., to upload research data, every deployed application requires a unique user account and for this, new credentials are established that the end users need to cope with. While this overhead could be reduced by Single Sign On, there is no usable solution for federated SSH login.

The CYCLONE PAM module uses the keyboard-interactive mode of SSH in combination with a custom PAM module to implement such a federated SSH login. The PAM module "pam_openid_connect" [28] starts an embedded web server and displays its URL to the bioinformaticians in the SSH terminal session. When they open the link in their browsers, they are redirected to the CYCLONE Federation Provider where they authenticate with their federated ID using OpenID Connect. After authenticating, the Federation Provider returns the user’s information to the integrated PAM webserver and therefore to the PAM module and the Linux PAM subsystem. The PAM module then compares the user’s account identifier (e.g., email) with a list of user identities allowed to login via the requested system account. This list can be easily modified manually or it can be provided through SlipStream parameters to be used by deployment scripts.

VIII. FUTURE DEVELOPMENT

This paper presents results of the ongoing development of the Intercloud Security Framework (ICSF) that is a part of Intercloud Architecture Framework. ICSF provides a basis for developing security infrastructure services to ensure consistent security of the multi-cloud applications provisioning on demand. In its further theoretical development ICSF address specific requirements for Big Data infrastructures and corresponding paradigm shift to data-centric security that was initially researched in the authors’ paper on new security challenges of the Big Data infrastructures [29]. The paper refers to specific use cases requirements identified in the CYCLONE project for the general bioinformatics use case. Further practical ICSF development will be focused on the development of ICSF components and integration with cloud automation tools such as SlipStream [30].

ACKNOWLEDGEMENT

The research leading to these results has received funding from the Horizon2020 project CYCLONE (funded by the European Commission under grant number 644925).

REFERENCES

[1] NIST Special Publication NIST SP 1500: NIST Big Data Interoperability Framework (NBDIF) [online]

http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-1.pdf

[2] Demchenko, Yuri, Fatih Turkmen, Christophe Blanchet, Charles Loomis, Cees de Laat, Cloud Based Big Data Infrastructure: Architectural Components and Automated Provisioning, The 3rd International Symposium on Big Data Principles, Architectures and Applications (BDAA 2016), as part of The International Conference on High Performance Computing and Simulation (HPCS 2016)

[3] State of the Cloud Report 2016, RigtScale, January 2016 [onine] http://assets.rightscale.com/uploads/pdfs/RightScale-2016-State-of-the-Cloud-Report.pdf

[4] Five trends transforming cloud computing, Oracle + Netsuite [online] http://www.netsuite.com/portal/resource/articles/cloud-computing-trends.shtml

[5] Craig Lawson, Neil MacDonald, Brian Lowans, Brian Reed, Market Guide for Cloud Access Security Brokers, 24 October 2016 Gartner, ID: G00293664 [online] https://www.gartner.com/doc/reprints?id=1-3L0MOC2&ct=161031&st=sb

[6] Demchenko, Y., M. Makkes, R.Strijkers, C.Ngo, C. de Laat, Intercloud Architecture Framework for Heterogeneous Multi-Provider Cloud based Infrastructure Services Provisioning, The International Journal of Next-Generation Computing (IJNGC), Volume 4, Issue 2, July 2013.

[7] Y.Demchenko, C. Lee, C.Ngo, C. de Laat, Federated Access Control in Heterogeneous Intercloud Environment: Basic Models and Architecture Patterns. IEEE Third International Workshop on Cloud Computing Interclouds, Multiclouds, Federations, and Interoperability (Intercloud 2014), In Proc IEEE International Conference on Cloud Engineering (IC2E), March 11, 2014, Boston, USA

[8] Yuri Demchenko, et al, CYCLONE: A Platform for Data Intensive Scientific Applications in Heterogeneous Multi-cloud/Multi-provider Environment, Fifth IEEE International Workshop on Cloud Computing Interclouds, Multiclouds, Federations, and Interoperability (Intercloud 2016), In Proc. IEEE International Conference on Cloud Engineering (IC2E), April 4 - 8, 2016, Berlin, Germany

[9] Todorov, D. & Ozkan, Y. (November 2013) 'AWS security best practices', Amazon Web Services [Online]. Available from: http://media.amazonwebservices.com/AWS_Security_Best_Practices.pdf

[10] Dominic Betts, et al, Developing Multi-tenant Applications for the Cloud on Microsoft Windows Azure, Third Edition, Microsoft. 2012. [online] http://download.microsoft.com/download/D/0/6/D0618696-2F91-4F7F-9477-63FC90D4D29E/Developing%20Multi-tenant%20Applications%20for%20the%20Cloud%203rd%20Edition.pdf

[11] Demchenko, Y., D.R. Lopez, J.A. Garcia Espin, C. de Laat, "Security Services Lifecycle Management in On-Demand Infrastructure Services Provisioning", International Workshop on Cloud Privacy, Security, Risk and Trust (CPSRT 2010), 2nd IEEE International Conference on Cloud Computing Technology and Science (CloudCom2010), 30 November - 3 December 2010, Indianapolis, USA

[12] NIST Special Publication 800-14 - Generally Accepted Principles and Practices for Securing Information Technology Systems. National Institute of Standards and Technology. September 1996. http://csrc.nist.gov/ publications/nistpubs/800-27/sp800-27.pdf

[13] Microsoft Security Development Lifecycle, Version 5.0, March 31, 2010. http://www.microsoft.com/sdl

[14] OASIS, "XACML v3.0: Core specification," OASIS, Tech. Rep., Aug. 2010. [Online]. Available: http://docs.oasisopen.org/xacml/3.0/xacml-3.0-core-spec-cs-01-en.pdf

[15] Membrey, P., K.C.C.Chan, C.Ngo, Y.Demchenko, C. de Laat, Trusted Virtual Infrastructure Bootstrapping for On Demand Services. The 7th International Conference on Availability, Reliability and Security (AReS 2012), 20-24 August 2012, Prague.

[16] Nabil Schear, Patrick T. Cable II, Thomas M. Moyer, Bryan Richard, Robert Rudd, "Bootstrapping and Maintaining Trust in the Cloud", Annual Computer Security Applications Conference (ACSAC), 2016

[17] Bryan Parno, Jonathan M. McCune, Adrian Perrig: Bootstrapping Trust in Commodity Computers. IEEE Symposium on Security and Privacy 2010: 414-429

[18] Trusted Platform Module, Trsuted Computing Group, [online] https://trustedcomputinggroup.org/work-groups/trusted-platform-module/

[19] Python-keylime [online] https://github.com/mit-ll/python-keylime

[20] Joshua Schiffman, Yuqiong Sun, Hayawardh Vijayakumar, Trent Jaeger: Cloud Verifier: Verifiable Auditing Service for IaaS Clouds. SERVICES 2013: 239-246

[21] Cloud Controls Matrix (CCM) [online] https://cloudsecurityalliance.org/research/ccm/

[22] CSA GRC Stack: Governance, Risk Management and Compliance [online] https://cloudsecurityalliance.org/research/grc-stack/

[23] Consensus Assessments Initiative Questionnaire (CAIQ) [online] https://cloudsecurityalliance.org/research/cai/)

[24] Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), Published 4 May 2016 [online] http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&qid=1488147180466&from=en

[25] Overview of the EU General Data Protection Regulation, Hunton&Williams [online] https://www.huntonregulationtracker.com/files/Uploads/Documents/EU%20Data%20Protection%20Reg%20Tracker/HuntonWilliams-GDPR-Management-Guide.pdf

[26] EU-U.S. Privacy Shield: stronger protection for transatlantic data flows Adopted 12 July 2016. Repeals former Safe Harbor Framework [online] http://ec.europa.eu/justice/data-protection/international-transfers/eu-us-privacy-shield/index_en.htm

[27] CYCLONE Federation Provider [online] https://github.com/cyclone-project/cyclone-federation-provider

[28] CYLCONE Python PAM module [online] https://github.com/cyclone-project/cyclone-python-pam

[29] Demchenko, Y., P.Membrey, C.Ngo, C. de Laat, D.Gordijenko, Big Security for Big Data: Addressing Security Challenges for the Big Data Infrastructure, Proc. Secure Data Management (SDM'13) Workshop. Part of VLDB2013 conference, 26-30 August 213, Trento, Italy.

[30] SlipStream Cloud Automation [online] http://sixsq.com/products/slipstream/