European Translational Information and Knowledge Management Services

eTRIKS Deliverable report

Grant agreement no. 115446

Deliverable D1.3

10 Major system updates

Due date of deliverable: September 2017

Actual submission date: 25th of july 2017 for the first draft version

Dissemination Level		
PU	Public	PU
PP	Restricted to other programme participants (including Commission Services)	
RE	Restricted to a group specified by the consortium (including Commission Services)	
CO	Confidential, only for members of the consortium (including Commission Services)	

DELIVERABLE INFORMATION

Project	
Project acronym:	eTRIKS
Project full title:	European Translational Information and Knowledge Management Services
Grant agreement no.:	115446
Document	
Deliverable number:	D1.3
Deliverable title:	10 Major system updates
Deliverable version:	2.0
Due date of deliverable:	30 September 2017
Actual submission date:	7 September 2017
Leader:	Ghita Rahal
Editors:	
Authors:	Benjamin Guillon, Gino Marchetti, Ghita Rahal
Reviewers:	Jay Bergeron, David Henderson, Peter Rice
Participating beneficiaries:	CNRS (CC), ICL
Work Package no.:	1

Work Package title:	WP1: Platform Service Delivery
Work Package leaders:	Ghita Rahal
Work Package participants:	Ghita Rahal, Pengfei Lui, Gino Marhcetti, Benjamin Guillon, Andreas Tielmann, David Henderson, Denny Verbeeck, Florian Guitton, Wei Gu, Chris Marshall, Peter Rice,
Estimated person-months for deliverable:	12
Nature:	Report
Version:	2.0
Draft/Final:	FINAL
No of pages (including cover):	9
Keywords:	eTRIKS, Security, Openstack cloud

Contents

I. Introduction

This document describes the main updates that have been made in order to provide access to the most advanced features and keep the platform up-to-date.

II. Executive Summary

The goal of this deliverable is to give an overview of the evolution of the platform by listing the major updates that have been taking place during its lifetime. Most of the updates pertain to the technical components of the platform, while others are upgrades to the applications allowing calculations on and analysis and visualisation of biomedical data.

III. Motivation

The eTRIKS platform is a complex assembly of software components provided by various vendors as well as from the open source software community. It provides a diverse set of services to the eTRIKS, and the wider scientific research, communities. The platform is delivered using web technologies, with some services being publicly accessible while others restricted to specific users.

Security of the system is a major concern for the eTRIKS consortium and more specifically for Work Package 1, which is responsible for the hosting and operation of the eTRIKS platform hosting data and applications for the research community. In order to provide the security level expected for the eTRIKS platform, during the five years of the eTRIKS project, various software updates were undertaken by the WP1.

Among all these, the ten most significant updates are detailed in the following document.

IV. Major updates

1 - The Openstack Cloud Framework

The eTRIKS platform is a cloud based IaaS platform, that leverages the Openstack Cloud framework in order to spin up virtual machines for all the hosting needs of the eTRIKS project.

When the eTRIKS platform was first deployed in 2013, it used Openstack Folsom (released in September 2012). To provide new functionalities and ensure better reliability as well as security, several delicate Openstack upgrades were carried out.

In 2014, Openstack was upgraded to the Havana version (released in October 2013), followed shortly after by the Icehouse version (released in April 2014). Then in 2015 the Kilo version (released in April 2015) was installed and at the end of the project in 2017, the Newton version (released in September 2016) is the last upgrade installed.

The Openstack Kilo update in 2015 enabled eTRIKS, a few months later, to migrate the platform from the outdated nova-network to the neutron networking module. This migration to the neutron networking module was mandatory to accept and release future upgrades.

Release name	Release date	Deployment date in eTRIKS
Folsom	September 2012	June 2013
Havana	October 2013	June 2014
Icehouse	April 2014	November 2014
Kilo	April 2015	November 2015
Newton	September 2016	August 2017

2 - Operating Systems

We distinguish two types of machines in the eTRIKS platform: the physical nodes and the virtual nodes. The physical nodes are part of the Openstack cloud platform which then manages the creation of virtual machines.

While the physical nodes were originally running Scientific Linux, they were migrated to CentOS, during a datacenter wide Linux distribution harmonization in 2015. Both Scientific Linux and CentOS are RedHat based distribution.

However, the virtual machines of the eTRIKS platform, that host the eTRIKS services, are running Ubuntu since the beginning of the project.

All these operating systems were updated regularly, mostly for stability and security reasons. These updates encompassed both kernel and package upgrades.

Scope	**Operating System**	**Release date**	**Deployment date in eTRIKS**
Virtual Machines	Ubuntu 12.04 LTS	April 2012	June 2013
	Ubuntu 14.04 LTS	April 2014	From May 2014 to November 2015
	Ubuntu 16.04 LTS	April 2016	From June 2016 to November 2016
Physical Machines	Scientific Linux 6.4	March 2013	June 2013

	Scientific Linux 6.6	November 2014	December 2014
	CentOS 7	July 2014	June 2015

3 - TranSMART

TranSMART [TF] is the main business software of the eTRIKS project, hence it's also the primary service provided by the eTRIKS platform. A specific eTRIKS flavour of the tranSMART application was built up to accommodate requirements critical to eTRIKS' clients. The modification that became the foundation of the eTRIKS platform was the adaptation of tranSMART to use the open source PostgreSQL database engine instead of the commercial Oracle database engine. Oracle had been the only option for tranSMART adopters prior to PostgreSQL integration. PostgreSQL integration was the key achievement that made adoption of the eTRIKS platform viable for the many academic clients that incorporated the eTRIKS platform into their research processes. During the course of the project, several upgrades were successfully conducted, including both the TranSMART core upgrades from the TranSMART Foundation and those specific to the eTRIKS flavour.

eTRIKS flavored TranSMART	**TranSMART version base**	**Deployment/Release date in eTRIKS**
v1	TM v1.1.0	September 2013
v2	TM v1.2.2	November 2014
v3	TM v16.1	August 2016
v4	TM v16.2	May 2017

While other versions were essentially native transmart plus configuration settings, v3 included some eTRIKS-specific plugins that were not available in the TF released 16.1: smartR, HiDome, XNAT viewer.

4 - Java

The Java development kit (JDK) was used to develop TranSMART. As a consequence, the Java Runtime Environment (JRE) is an essential component of any TranSMART deployment. Initially, the OpenJDK distribution (the default in Ubuntu) was used, but for performance reasons the Oracle JRE was introduced in its stead.

Since the JRE is often a source of security vulnerabilities, it was often necessary to update JREs on every TranSMART host to ensure data protection.

JDK/JRE Distribution	Adoption date
OpenJDK 7	June 2013
Oracle JRE 7	January 2015
Oracle JRE 8	October 2015

5 - Apache HTTPd

The Apache HTTPd web server acts as a web frontend/proxy for TranSMART deployments. This is the only piece of software that faces the wider Internet directly, so it must be routinely updated for security-related purposes.

During the entire duration of the eTRIKS project, the major version 2.4 was installed and run for each eTRIKS web service. Of course, during the span of nearly five years, it evolved through several minor updates, from v2.4.4 in June 2013 to v2.4.26 in mid 2017.

6 - Apache Tomcat

Tomcat, a web server application released by the Apache foundation, is key to any TranSMART deployment. To be precise, this is the java container server that runs the TranSMART java bytecode.

Tomcat version 7 was used during most of the duration of the eTRIKS project, with initial eTRIKS deployments on Tomcat 6 in mid 2013 and preliminary tests undertaken on Tomcat 8 starting in May 2017.
As Apache HTTPd, it went through several updates during the span of four years, starting with v6.0.32 in 2013, subsequently migrating through v7.0.30 to v7.0.52 and ending on v8.0.37.

Tomcat Version	**Release date (on Ubuntu)**	**Adoption date**
Tomcat 6	October 2008	June 2013
Tomcat 7	March 2014	September 2014
Tomcat 8	February 2016	May 2017

7 - PostgreSQL

The TranSMART code base on which all eTRIKS versions of the the core TranSMART application is built uses postgresql v9.3 libraries. However, the postgresql database engine provides some level of backward compatibility, allowing the upgrade of the postgres database servers as desired.

This backwards compatibility is supported as long as all components use the same major version (eg. v9.x.y). This feature allowed the upgrade of the database servers in concert in order to benefit from the stability and performance improvements added in the latest versions of PostgreSQL.

PostgreSQL Version	Release date	Adoption date in eTRIKS
v9.2	October 2012	June 2013
v9.3	September 2013	November 2013
v9.4	December 2014	June 2015
v9.5	January 2016	March 2016
v9.6	September 2016	February 2017

8 - R/Rserve

R and its TCP/IP based server counterpart, Rserve, are responsible for running TranSMART analytics in the background. As such they are a performance-critical component of the TranSMART stack.

Original TranSMART deployments used the standard Ubuntu R packages but for performance reasons it was decided to replace Ubuntu R with the Microsoft R Open distribution (MRO), formerly known as R Revolution Open (RRO). MRO leverages some Intel CPU optimization libraries, among other things, and provides increased performance.

R distribution	Release date	Adoption date in eTRIKS
R Base v2.15.3	March 2013	June 2013
RRO v3.1.3	March 2015	May 2015
MRO v3.2.5	April 2016	June 2016

9 - Gitlab

The Gitlab platform is a software development oriented collaboration platform which, among other things, allows centralized code versioning through git. Gitlab was a critical tool for all developers of the eTRIKS project.

Gitlab version 5 was first deployed at the very beginning of the project in 2013. After several years, it became critical to update the Gitlab version to ensure stability. Gitlab Version 8 replaced version 5 and, since then, Gitlab has been updated regularly with new releases to

provide a highly robust platform for eTRIKS' clients (Version 9 being in use during the writing of this document).

Gitlab version	Release date	Deployment date in eTRIKS
v5.2	May 2013	June 2013
v5.4.2	November 2013	December 2013
v8.6	March 2016	March 2016
v9.0	March 2017	March 2017
v9.2	June 2017	June 2017

10 - Operations Portal

The Operation Portal is a java based application that uses a technology stack similar to that of TranSMART. The operations portal enables users from projects hosted by eTRIKS to manage and access the various eTRIKS services. The portal was developed in house at CC2INP3 shortly after the initial eTRIKS platform was released mid-year 2013.

The portal does not utilize a strict version numbering scheme but has undergone various phases of upgrades. One upgrade enabled the portal to include the newly-developed eTRIKS-Labs, a collection of emerging applications and services specializing in the analysis and visualization of biomedical data.

V. Conclusion

During the course of the eTRIKS project, the eTRIKS platform has undergone multiple upgrades to keep it up-to-date and to consolidate its robustness and performances. This constant evolution has made it easier to maintain and to share with other centres to permit its sustainability post-project.