IOStack

(H2020-644182)

Software-Defined Storage for Big Data on top of the OpenStack platform

D6.3 Final dissemination, exploitation, and standardization report

Due date of deliverable: 30-11-2017
Actual submission date: 06-12-2017

Start date of project: 01-01-2015   Duration: 36 months

Summary of the document

Document Type	Deliverable
Dissemination level	Public
State	v1.0 FINAL
Number of pages	27
WP/Task related to this document	WP6 / T6.2 & T6.3
WP/Task responsible	Arctur
Leader	Nejc Bat (ARC)
Technical Manager	Luka Živulović (ARC)
Quality Manager	Pedro García (URV)
Author(s)	Nejc Bat, Luka Živulovič, Jasna Vuksanovič
Partner(s) Contributing	URV, IBM, BSC, EUR, IDI, ARC, GRP, MPS
Document ID	IOStack_D6.3_Public.pdf
Abstract	This deliverable is the final report on dissemination activities done during the IOStack project.
Keywords	Dissemination, Conference, Video, Workshop

History of changes

Version	Date	Author	Summary of changes
0.1	15-11-2017	Nejc Bat	First version
0.9	05-12-2017	Nejc Bat	Added additional exploitation descriptions and revised dissemination activities
1.0	05-12-2017	Nejc Bat	Last adjustments and finalizing

Table of Contents

1   Executive summary

This deliverable is the final report on dissemination activities done during the IOStack project. Being a three-year project we first present the work that has been done in the first two years of the project. This deliverable overlaps also with the previous WP6 deliverables, namely D6.1 and D6.2.

The deliverable then proceeds to present the dissemination activities done in the final year of the project as well as the result of the Exploitation workshop that has been done internally amongst the partners to define an exploitation strategy.

2   Previous work

IOStack is a three year project. To fully illustrate the dissemination efforts of the project we will first recap the key decisions and activities that were done during the project.

2.1   Process and strategy for dissemination activities

At the very beginning of the project the following dissemination plan was established:

Our first guideline for dissemination is to specialize dissemination events to reach well-defined user and practitioner communities. Therefore, in order to communicate properly and not to overlap themes between partners, partners should focus on their own Work Packages (WPs), which are related to a specific topic in the project. By doing this, each partner can be specialized on particular aspects of Software-Defined Storage (SDS). Such a degree of expertise and specialization will provide high value to the dissemination actions and communication activities of partners. Furthermore, partners should benefit from their specialized background to infer the best alternatives to disseminate their own WP results, maximizing the visibility of IOStack and its impact to society.

We distinguish four specific topics that will drive the communication plan:

- SDS for object storage

- SDS for block storage

- Compute or virtualisation orchestration

- Cooperative storage and data analytics strategies

As a second guideline, since the IOStack project relies on an association of industrial and research partners, it is very important to take advantage of this plurality for broadcasting each novel development in both communities. To do so, each partner should disseminate its results using suitable channels for each action. It means that research partners should publish research papers and participate in events or workshops and for the industrial partners they should send press releases to mass media and spread the word of early results in local or international events related to Software-Defined Storage (SDS) technologies.

In third place, apart from the single communication plan of each partner and its dissemination actions, joint collaborations among partners and other research projects or industries are also important, because they could result into a new outcome. Furthermore it is an opportunity to propagate the achievements to other stakeholders.

Apart from these high-level guidelines, in IOStack we define a battery of specific actions to materialize the project dissemination. To this end, Fig. 1 describes the different types of dissemination events considered in this project. In the following, we describe their objective and how they can promote the visibility and adoption of IOStack.

Conferences and Publications: Research results that have been achieved in the different work packages will be presented to the scientific community in the form of publications in journals, presentations and demonstrations in conferences, workshops and exhibitions. This is key to promote the ideas behind IOStack, and to have the scientific backing for introduction in the relevant standards later on. The research quality of the academic partners and the good synergy with industrial ones should ensure high impact publications in top conferences and journals.

Some of the conference where we are planning to attend are: IEEE Cloud, IEEE P2P, ACM Middleware, ACM CCS, ACM Eurosys, USENIX FAST, among others. Regarding journals, we expect to publish in some of the following ones: Elsevier Future Generation Computer Systems, Elsevier Computer Networks, Elsevier Computer Communications, IEEE Internet Computing, IEEE Network Magazine, IEEE/ACM Transactions on Networking, VLDBJ, and ACM/IEEE Transactions on Database Systems.

As described in deliverable D2.1, these published results will follow a green open access strategy where articles and related data will be self-archived by the authors in an online repository (project's webpage and/or personal homepage) after the corresponding scientific publisher grants open access
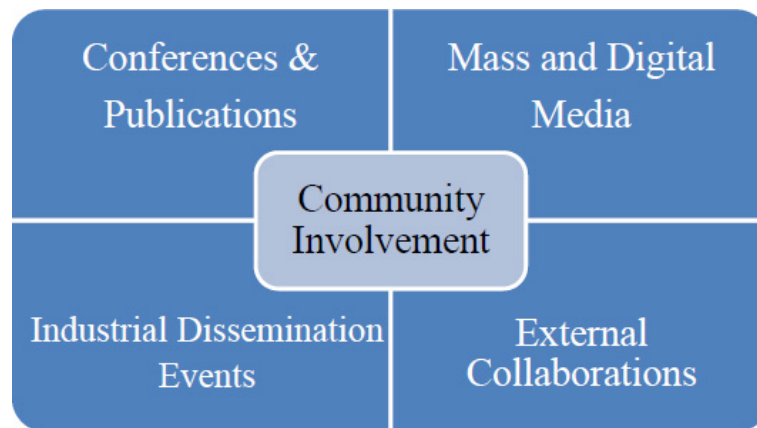
Figure 1: Areas of dissemination plan for each topic/aspect of IOStack

to authors. Most conferences and journals allow publishing the results after the peer-review period as long as the final copy-editing of the article is kept to subscribers or attendees of the journal and conferences respectively.

Mass and Digital Media: The IOStack partners will engage in the usual forms of dissemination to large audiences. Today, this is clearly related to mass and digital media, which includes channels such as the Internet, TV and radio, among others. We should take into account that this form of dissemination is general and broad, as it is not specially tailored to a technical audience. This yields that we should carefully craft the contents of these mass dissemination events in order to be easily accessible for non-technical audiences.

Therefore, IOStack partners will develop the required collateral to promote the adoption of the IOStack project including social media videos, user manuals and marketing collateral. We plan to expose some of these materials to the commissioners in the next project reviews in order to acquire high-quality feedback.

Industrial Dissemination Events: This category of dissemination contain those events that are neither research nor mass/digitalmedia events. For instance, industrial meetings or technical/developer conferences fall into this category. In fact, category of dissemination activities is of great importance in IOStack; it includes, for instance, OpenStack Summit or DockerCon events, which are premier technical meetings for the open-source community that supports the technologies behind IOStack.

During the project, assisting to such events (OpenStack Summit, DockerCon) will be a dissemination priority. This will make the community interested in IOStack, which may lead in contributions from the open-source community to the project. Ideally, we expect to up-stream the code of IOStack as an official part of the OpenStack ecosystem.

External Collaborations: In our experience with other EU projects, to weave alliances with other research project, industries and initiatives will be a key element to propagate the impacts of IOStack rapidly. Moreover, the influence of partners within a collaboration may enable a virtuous circle in which all collaborators are benefited. The mutual benefits of collaborations among projects includes the discovery of new use cases, the inter-operation of various EU software toolkits, and the increment of user/customer communities, to name a few.

For these reasons, we will also focus on other European Projects in the second year of IOStack. As we will mention later on, we started several contacts with NESUS initiative and various companies from the energy utility sector during the first year, so we are planning the activities to develop with all of them. We expect astonishing results to come in the next months.

Community Involvement: In IOStack, we target 3 different communities: datacenters, data analytics companies and open-source communities. To reach these communities will boost the adoption of IOStack by real-world companies. Moreover, as well as it will help further developments of the project toolkit beyond the project life time.

First, we target the involvement of datacenters once the IOStack prototype is ready. Concretely, during the second year of the project, we will resort to the partners that operate datacenters (MPStor, Arctur) in order to make the IOStack prototype a usable tool for administrators. In this way, we aim to disseminate our achievements and to gather the feedback and interest from the participants.

Second, we also aim at involving our data analytics use cases to do pilot tests with IOStack. In particular, we are currently developing SDS services for improving, both in terms of flexibility and performance, the data analytics tasks of Idiada and GridPocket. Once achieved our objective, we need to publicize these success stories to other other companies, so that they may also adopt IOStack in their own business processes. Clearly, appropriate industrial dissemination events and venues are key to spread these success stories to increase our potential customer community.

In addition, open-source community involvement will be very important for the project too. There is another interesting community that we will involve in the IOStack project. We are pointing to the OpenStack community, because OpenStack is becoming the "de facto" world standard as a cloud storage platform, with partners such as: NASA, RackSpace, Intel, Paypal, CERN and many more. As IOStack is supported on this platform, we plan to involve the OpenStack community in order to get their feedback and seek valuable collaborations around the globe. Similarly, the Docker community will also be very important for the compute framework of IOStack. We will also execute specific actions to attract the Docker community to the technical development of the project.

Finally, for the third year we plan to consolidate the activities started in the second year, seeking a major impact in industrial field, enforcing the academic research results and finally getting a bigger user communities involvement. Depending on the outcomes achieved in the second year, we will adapt the dissemination activities in order to take the best advantage of these results. It means that getting a positive feedback of the user community could help us define the best way to market our achievements and collaborations with other EU projects could derive to new goals into the IOStack project. In summary, we can define some guidelines for the third year goals and activities, but we will be flexible on that point and set up a fine plan by the end of the second year.

2.2   Year 1 dissemination activities

In this section we will summarize the first year dissemination activities, specifying the partners participating in each activity. We divided them into the following subsections:

2.2.1   Academic conferences and publications

- **Conference paper** (February 16-19, 2015, Santa Clara, CA, USA): "SDGen: Mimicking Datasets for Content Generation in Storage Benchmarks". Authors: Raúl Gracia-Tinedo, Danny Harnik, Dalit Naor, Dmitry Sotnikov, Sivan Toledo and Aviad Zuck. *13th USENIX Conference on File and Storage Technologies (FAST'15)*[1].

- **Magazine paper**: "IOStack: Software-Defined Object Storage". Authors: Raúl Gracia-Tinedo, Pedro García-López, Marc Sánchez-Artigas, Josep Sampé, Yosef Moatti, Eran Rom, Dalit Naor, Ramon Nou, Toni Cortés, William Oppermann, Pietro Michiardi. *IEEE Internet Computing* (to appear)[2]. This paper will be published in a high visibility magazine. This publications has a twofold benefit: First, a magazine enables us to spread the ideas of IOStack to a broad audience; and second, the article offers a general and coherent view of the project, as all partners have participated in its elaboration.

- **USENIX FAST'15 Conference** (February 16-19, 2015, Santa Clara, CA, USA): URV and IBM presented a full research paper related to realistic data generation for storage benchmarking, specifically in terms of compression. Apart of being a joint work between two partners of the consortium, it was a good opportunity to promote and augment the visibility of the IOStack project though face-to-face meetings with several attendees.

- **2nd URV Doctoral Workshop in Computer Science and Mathematics (Tarragona, November 2015)**: Pietro Michiardi from Eurecom was invited as keynote speaker at the 2nd URV Doctoral

Workshop in Computer Science and Mathematics (November 13th 2015). Apart of being an interesting joint dissemination activity between Eurecom and URV, the talk of Pietro Michiardi exposed the scheduling and allocation problems currently found in many industrial and real world compute environments. This is directly related to the challenges to be solved in IOStack. More details can be found at the workshop's webpage[1].

2.2.2 Mass Media and Digital Media

To be present in digital media is key to disseminate the results of IOStack. In this sense, our actions will be focused both on publishing the news related to IOStack and to participate in mass media for a broader impact on less technical audience (tv, radio, etc.).

Regarding digital media actions, in IOStack we already executed the following actions:

- **IOStack official webpage**[2]: We set up an official project webpage. The IOStack site offers general information about the project objectives, the consortium and the software and dataset outcomes. Moreover, the IOStack webpage offers means to specialized and general public of auditing the progress of the project via public deliverables (open data). The webpage contains a news section to link the most important events related to the project, as well as integration with IOStack Twitter account.

- **IOStack Twitter account**[3]: IOStack is present in Twitter as a mean of rapidly disseminating news related to the project. The account is being updated frequently and it is being linked with other accounts of previous EU projects, in order to maximize the impact of IOStack tweets.

- **IOStack Github account**[4]: Another way to disseminate the project is via promoting our source code outcomes. For this reason, we are actively working in the IOStack Github account that is publicly available.

- **IOStack links in official partners' sites**: As a mean of expanding the dissemination of IOStack, several partners are publicizing the project in their respective websites. For instance, BSC has posted a permanent notice about IOStack[5]. Similarly, MPStor has also integrated publicity of IOStack in its own webpage[6]. Other partners such as URV[7], Eurecom[8] and GridPocket have included information on its participation to IOStack project on their corporate website.

- **MPStor Newsletter**: MPStor is also making use of newsletters and mailing lists to spread the word of IOStack across customers and other companies. For instance, the newsletter of February 2015 contains an article about IOStack and it has been pushed to over $4,000$ emails[9].

- **IBM Research blog spot**: IBM has also contributed with a technology blog in which the Storlet framework, as a par of IOStack project, has been discussed. The blog spot can be found at: "Storlets: From research prototype to open source technology"[10] (published April 30th 2015).

2.2.3 Industrial Dissemination events

- **DockerCon'15 (Barcelona, November 2015)**: Daniele Venzano from Eurecom presented Zoe in a 45 minutes talk at the DockerCon'15 EU Conference held in Barcelona in 2015. The project

---

[1]http://deim.urv.cat/~dcsm/program.html

[2]http://iostack.eu

[3]https://twitter.com/iostackproject

[4]https://github.com/iostackproject

[5]http://www.bsc.es/about-bsc/press/bsc-in-the-media/bsc-brings-expertise-data-management-iostack-project

[6]http://www.mpstor.com/news-top/news/175-horizon-2020-iostack-project

[7]http://ast-deim.urv.cat/web/projects/active-projects/iostack

[8]https://distsysgroup.wordpress.com/collaborations/

[9]http://www.mpstor.com/media-center-mainmenu/newsletter

[10]http://ibmresearchnews.blogspot.com.es/2015/04/storlets-from-research-prototype-to.html

IOStack was mentioned in the presentation as the environment in which the architectural improvements lead by Zoe will take place. Both the video and the slides of the presentation are available on Youtube and on the Docker Web Site[11].

- **OpenStack Israel (June 2015)**: The OpenStack Israel event provides diverse tracks on everything OpenStack —from compute, networking, and storage, through hands-on workshops and even a conference that discusses real world stories. In this event, Eran Rom presented "Bringing Compute to Openstack Swift"[12][3].

- **OpenStack Summit (Vancouver, May 2015)**: The OpenStack Summit is a five-day conference for developers, users, and administrators of OpenStack Cloud Software. In this international event, Eran Rom from IBM presented Storlets and their applications to Software-Defined Storage[13].

- **OpenStack Summit (Tokyo, November'15)**: Eran Rom presented "Storlets: Making Swift More Software Defined than Ever" at OpenStack Summit. The talk was about the applications of the Storlet framework to SDS Services[14][4].

- **IBM InterConnect 2015**: IBM InterConnect is a premier cloud & mobile conference, a rendevouz point for users, developers and companies targeted to the cloud business. In this event, Dalit Naor from IBM presented "The perfect match: Apache Spark meets OpenStack Swift"[15]. The presentation described the synergies of combining Spark and Swift. In fact, in IOStack IBM is committed to make Swift to cooperate with Data Analytics platforms (e.g., Spark) via active storage strategies for reducing data processing times and boost performance.

- **OneM2M workshop at ETSI (Sophia Antipolis, December 2015)**: GridPocket did a talk and demonstration of smart grid data processing solutions at the European Telecommunication Standards Institute[16]. The talk was about cloud storage an data privacy with international visitors, which are central points if the GridPocket use case in IOStack.

- **Universal Exposition (Milan, October 2015)**: Participation of GridPocket (as a winner H2020 SME Innovation Instrument support) to the side event of Expo Milan, involving talks and demonstration of smart grid data processing, storage and privacy to international companies[17].

- **Cop21 (Lyon, July 2015)**. Presentation on energy data privacy and processing by GridPocket at the Lyon EDF pre-event of the Global Climate Summit organized by United Nations.

- **Innovative City (Nice , June 2015)**. Demonstration of GridPocket's smart grid data processing solution and presentation of IOStack project plans at the international conference and trade show dedicated to smart cities in Nice[18].

- **MedCop21 (Marseille, June 2015)**. GridPocket presentation on smart grid data processing technologies to professionals, politics and institutions of the Mediterranean region countries.

- **SmartGrid Paris (Paris, May 2015)**. GridPocket technology demonstration, including data processing platform and presentation of IOStack project at SmartGrid Paris professional tradeshow[19].

---

[11] http://dockerconeu2015.sched.org/event/88d8ac8cc2193aec4cf73caa7ff76e2b

[12] http://www.slideshare.net/openstackil/ibm-swift

[13] https://openstacksummitmay2015vancouver.sched.org/eran_rom.1sq7eu5t?iframe=no#.VkHrBL9OdSA

[14] https://www.openstack.org/summit/tokyo-2015/videos/presentation/storlets-making-swift-more-software-defined-than-ever

[15] https://www.youtube.com/watch?v=4iez3otZH4o

[16] http://www.etsi.org/news-events/events/966-2015-12-etsi-m2m-workshop-2015-featuring-onem2m

[17] http://www.expo2015.org/

[18] http://www.innovative-city.com/

[19] http://www.smart-energies-expo.com/

- **Meeting with RTCC (North Carolina, April 2015)**. Presentation on European clean tech industry, and research projects, including IOStack at the workshop with Research Triangle CleanTech Cluster.

- **Meeting with SKEMA (Raleigh, April 2015)**. Presentation on Smart Grid technologies,and related research projects, including IOStack to students and staff of SKEMA business school in Raleigh.

- **Meeting with ASPROM (Paris, March 2015)**. Presentation on BigData, with special focus on data privacy including IOStack project at the (professional conference on Smart networks for electricity, gas and water organized by ASPROM association in Paris.

2.2.4   Year 1 Community involvement and Exploitation Activities

At the beginning of the project IOStack was aimed to become a framework to automate storage provisioning and provide SDS services at large scale. In terms of community involvement, this meant that datacenter administrators and companies with data analytics business processes would have been our main customer communities. Therefore, the first activity regarding community involvement was focused on testing and deploying IOStack within our use case partners.

On the one hand, we focused on datacenter administrators and operators. In this sense, Arctur worked on making the IOStack dashboard suitable and usable for a datacenter administrator. This was a required step towards making the IOStack toolkit an exploitation-ready prototype that could be adopted by other companies. Of course, the role of Arctur was crucial; Arctur would be a pilot customer of IOStack in order to facilitate storage management in the datacenter. The customer experience of Arctur was expected to lead us to the provision of a real-world toolkit with actual commercial possibilities. Moreover, Arctur is in contact with other datacenters that may also adopt IOStack, which would be a solid basis for implantation in the market.

Similarly, MPStor also planned to exploit IOStack in their own business strategy. MPStor was also developing an automated volume management and a filter framework to deploy advanced SDS services in block volumes. These cutting-edge services represent a value added to their operation that might have lead to advantages w.r.t. competitors. On the other hand, we studied in depth the Big Data necessities of our use-case companies. In particular, Idiada, an automotive company that stores and processes large amounts of information related to vehicle tests. Thus, there are two potential capabilities of IOStack that may enable Idiada of exploiting it as a commercial product. First, the management of large amounts information in Idiada is becoming a problem. To solve such an scalability problem, IOStack promises advanced and software-defined data management techniques, which can reduce the amount of data stored in Idiada's servers. Second, Idiada is also executing parallel simulations of several vehicle crush tests, which requires advanced scheduling and allocation strategies in the compute cluster. IOStack also focused on providing flexible and intelligent compute cluster strategies via VMs and containers (i.e., Zoe[20]). Executing real workloads of Idiada in our experimental IOStack deployments to study the best strategies to solve their problems. In fact, by solving the problems of Idiada in both compute and storage sides, IOStack demonstrated to be a feasible solution to be adopted by many other SMEs around Europe. As another data analytics use-case, worked with GridPocket. Essentially, GridPocket works with very large amounts of temporal, comma-separated value (CSV) data that represents the electricity usage of users. In this sense, GridPocket presents two main problems that IOStack aimed at solving:

1. To efficiently execute complex machine learning algorithms in CSV datasets, and

2. To be able of keeping the privacy and confidentiality of such a sensitive data from external access.

The GridPocket use-case opens the door in IOStack to exploit SDS services in cooperation with the compute cluster. This novel perspective of SDS may attract other companies, perhaps outside the energy utility sector, to explore IOStack as a practical tool in their data analytics business processes.

---

[20]http://zoe-analytics.eu/

2.3 Year 2 Dissmeination activities

In year 2 of the IOStack project the dissemination activities have been done as described below.

2.3.1 Academic conferences and publications

- R. Gracia-Tinedo, P. Garcia-Lopez, M. Sanchez-Artigas, J. Sampé, Y. Moatti, E. Rom, D. Naor, R. Nou, T. Cortes, W. Oppermann and P. Michiardi (2016). IOStack: Software-Defined Object Storage. IEEE Internet Computing, 20(3), pp.10-18.

- M. Sánchez-Artigas. "Oblivious RAM as a Substrate for Cloud Storage - The Leakage Challenge Ahead". ACM CCSW '16. 2016. pp. 49-53.

- R. Chaabouni, M. Sánchez-Artigas, A. Chaabouni and P. García-López. "Improving the QoE in Personal Clouds with Cross-Swarm Bundling". IEEE LCN'16. 2016. To appear.

- M. Ruiz, P. García-López and M. Sánchez-Artigas. "TallyNetworks: Protecting Your Private Opinions with Edge-centric Computing". LSDVE@EuroPar'16. 2016. To appear.

- J. Sampé, M. Sánchez-Artigas and P. García-López. "Vertigo: Programmable Micro-controllers for Software-Defined Object Storage". IEEE CLOUD'16. 2016. To appear.

- R. Gracia-Tinedo, P. García-López, A. Gómez and A. Illana. "Understanding Data Sharing in Private Personal Clouds". IEEE CLOUD'16. 2016. To appear.

- R. Chaabouni, M. Sánchez-Artigas and P. García-López. "The Power of Swarming in Personal Clouds Under Bandwidth Budget", Journal of Network and Computer Applications, Vol. 65. 2016, pp. 48-71.

- P. García-López, R. Gracia-Tinedo, Alberto Montresor. "Towards Data-driven Software-defined Infrastructures", 2nd International Conference on Cloud Forward: From Distributed to Complete Computing, CF2016, 18-20 October 2016, Madrid, Spain.

- R. Gracia-Tinedo, J. Sampé, E. Zamora-Gómez, M. Sánchez-Artigas, P. García-López, Y. Moatti, E. Rom, "Crystal: Software-Defined Storage for Multi-tenant Object Stores". In Proceedings of USENIX FAST'17. Joint work of URV and IBM.

- Pace, Francesco; Milanesio, Marco; Venzano, Daniele; Carra, Damiano; Michiardi, Pietro, "Experimental Performance Evaluation of Cloud-Based Analytics-as-a-Service", in Proc. of IEEE Cloud, San Francisco, USA, June 2016.

- Lulli, Alessandro; Dell'Amico, Matteo; Michiardi, Pietro; Ricci, Laura, "NG-DBSCAN: Scalable Density-Based Clustering for Arbitrary Data", Proceedings of the VLDB Endowment, Volume 10, 2016-2017.

2.3.2 Mass Media and Digital Media

- Diari de Tarragona; `http://www.diaridetarragona.com/tarragona/66242/la-urv-disena-un-software-que-facilita-la-gestion-de-datos`.

- La Vanguardia; `http://www.lavanguardia.com/tecnologia/20160726/403488642292/la-urv-coordina-el-primer-proyecto-europeo-para-gestionar-datos-masivos.html`.

- Agencia Sinc; `http://www.agenciasinc.es/Noticias/Europa-desarrolla-un-software-que-abarata-la-gestion-masiva-de-datos-en-empresas`.

- Innovaticias; `http://www.innovaticias.com/innovacion/37501/innovador-software-abarata-gestion-masiva-datos-empresas`.

- Madri+D; `http://www.madrimasd.org/informacionidi/noticias/noticia.asp?id=67057`.

- DatacenterDynamics; `http://www.datacenterdynamics.es/focus/archive/2016/07/la-urv-de-tarragona-coordina-un-proyecto-europeo-para-la-gesti%C3%B3n-masiva-de-dat`.

- Tarragona 21; `http://diaridigital.tarragona21.com/la-urv-desenvolupa-un-programari-que-facilita-la-gestio-massiva-de-dades/`.

- Diari Més; `http://www.diarimes.com/es/noticias/tarragona/2016/07/27/la_urv_coordina_primer_proyecto_europeo_para_gestionar_big_data_6515_1091.html`.

- URVactiv@; `http://diaridigital.urv.cat/es/urv-desarrolla-software-facilita-gestion-masiva-datos/`.

- Crystal - My first storage policy. This tutorial explains how to write a storage policy with Crystal and install a storage filter. The video shows how this enables dynamic reconfiguration of OpenStack Swift, which can be exploited to optimize storage workloads; `https://www.youtube.com/watch?v=vbNxCbQbKWM`.

- Crystal - Playing with Dynamic Storage Automation Policies. This demo video shows how to use dynamic storage automation policies that are triggered by workload monitoring metrics; `https://www.youtube.com/watch?v=7DPhB9zN9zo`.

- Crystal - Multi-tenant Bandwidth Differentiation. This demo video shows how Crystal can provide bandwidth differentiation in a multi-tenant OpenStack Swift deployment; `https://www.youtube.com/watch?v=6JixYX3yXwY`.

- IOStack: Software-Defined Storage for Big Data. This video presents the project objectives, use cases and proposed solutions; `https://www.youtube.com/watch?v=b6gjjn7Kz7A`.

- IOStack web page: `http://iostack.eu/` The web page was updated with the new deliverables, scientific publications and datasets produced during the last year. The web also describes the main software results of the project to date.

- Crystal web page: `http://crystal-sds.org/`; This web page presents Crystal (an open and extensible Software-Defined-Storage for OpenStack Swift), describing its motivation, architecture and implementation. It also includes tutorials and demo videos.

- Crystal documentation at GitHub: `https://github.com/Crystal-SDS/controller`; Crystal source code is available at GitHub along with the complete API specification, installation instructions and code samples.

- IOStack Twitter account: `https://twitter.com/iostackproject`; URV manages the IOStack Twitter account, publicizing events and activities attended by the consortium members.

- The GitHub for "stocator" at the IBM Spark Technology Center mentions IOStack (at bottom): `https://github.com/SparkTC/stocator`.

- Youtube video for Pushdown and IOSTACK; `https://www.youtube.com/watch?v=47_fDBzMZQQ`.

- Idiada IOStack usa-case video: `https://www.youtube.com/watch?v=n5VvG76sUos`.

- IBM OSIC team published a blog on the IOStack pushdown technology; `https://osic.org/blogs/ibm-research-osic-speed-big-data-queries`.

2.3.3   Industrial Dissemination events and external collaborations

- CloudScape 2016, Brussels, March 8-9, 2016. Dr. Raúl Gracia (Universitat Rovira i Virgili) presented a poster of IOStack in this venue. He explained an overall vision of IOStack objectives and presented the first results of the project.

- ClusterTICSUD, Tarragona, November 10, 2016. Dr Pedro Garcia Lopez was invited by Tarragona's Enterprise Cluster to give a talk about Big Data Technology Challenges. During this talk, Dr. Pedro Garcia Lopez explained the overall vision of IOStack project, and presented the major results at this date. The event represented a good technology transfer opportunity and it was a remarkable dissemination event with more than 50 relevant IT companies in the area.

- A "stocator" related talk was submitted to be given by Gil Vernik of IBM at the Hadoop Strata 2017 San Jose conference.

- BSC: External collaborations with other research projects and industries e.g. BSC proposed IOStack-based object storage for NextGenIO H2020 EU research project.

- GridPocket contribution at conference "Innovation City Convention" in Nice (June 2016).

- Cloud Forward Conference 2016, Madrid, October 18-20, 2016. Dr. Pedro Garcia Lopez represented IOSTACK in this conference. In particular, he presented the results of the IOStack project to the NATRES cluster `https://eucloudclusters.wordpress.com/new-approaches-for-infrastructure-services/`. We initiated conversations with H2020 Mikelangelo `https://www.mikelangelo-project.eu/` representatives from Intel. Dr. Pedro Garcia Lopez also outlined the IOSTACK vision and results in a position paper about Software Defined Data Driven Infrastructures.

- OpenStack Summit, Barcelona, October 25-28, 2016. Dr. Raúl Gracia gave a talk at vBrown-Bag TechTalks presenting Crystal, one of the main software outcomes of the IOStack project. "Crystal: Open and Extensible Software-Defined Storage for OpenStack Swift" `https://www.openstack.org/summit/barcelona-2016/summit-schedule/events/16751` Video of the talk: `https://www.youtube.com/watch?v=TlYIJgAHfpc`.

- OpenStack Summit, Barcelona, October 25-28, 2016. "Plethora of Use Cases with Openstack Storlets"; details and video at: `https://www.youtube.com/watch?v=JkPp1FZEjn4`.

- OpenStack Summit, Barcelona, October 25-28, 2016. "Crystal Open and Extensible Software Defined Storage"; details and video at: `https://www.youtube.com/watch?v=TlYIJgAHfpc`.

- GridPocket contribution at conference at "EUW Barcelona" (November 2016).

- Ramon Nou talk at BOS session of HPC, Cloud, Big data convergence; Talk is about QoS on Storage and includes material about Iostack (filters and SDS and how they could be used with HPC workloads). JLESC 2016, 1 December, session 6.

- URV is in preliminary talks with Storage Made Easy (SME, web: `https://www.storagemadeeasy.com/`) about a potential collaboration. URV is exploring ways IOStack toolkit could fit SME use cases.

- University of Verona: collaboration on scheduling theory.

- Arctur: Design and printed copies of IOStack poster and IOStack leaflet for distribution in various academic and industrial events as well as other face-to-face meetings.

- INRIA Sophia Antipolis: collaboration on convex optimization.

- University of Pisa: collaboration on distributed algorithms, running on Zoe.

Figure 2: The IOStack leaflet

- Principality of Monaco: real-life experimental deployment of Zoe, collaboration on additional back-ends for Zoe.

- presenting Zoe open-source project to various interested parties, both academic and industrial.

- GridPocket Demo at "Innovative City Convention" in Nice (June 2016).

- GridPocket Demo at "Smart Energy Paris tradeshow" (June 2016).

- GridPocket Demo at "European Utility Week" in Barcelona (November 2016).

- Arctur announces support for IOStack platform on the Arctur-2 hyper-converged supercomputer at Supercomputing'16 conference, USA, Utah, Salt Lake City, 2016.

- Arctur promotes IOStack future usage to various industrial partners at face-to-face meetings.

- Idida promoted IOStack internally by distribution of video presentations and relevant documents to Idiada worldwide branch offices.

2.3.4 Community Involvement and Exploitation Activities

In IOStack we are aiming at involving various communities. These are from the field of Big Data users, such ash data analytic use cases, but also from the field of data center operators. Our main targets are open source communities since our project is fully open source based. Of course, through various already mentioned activities we are also addressing commercial companies and enterprises.

- IBM is actively contributing to the OpenStack community with several already fully integrated solutions such as Storlets. We had strong presence at OpenStack Summit 2016 in Barcelona with several talks and face-to-face meetings.

- IBM Stocator has become the default connector to Object Store service in IBM BlueMix. That is, Stocator is in IBM production use, which sets a solid ground for future production ready releases of IOStack through widely known and used IBM BlueMix platform.

- Since Zoe is one of the key components in IOStack, we are also active in the Docker open source community. EUR attended a Docker conference to present Zoe. Also, Zoe is already being used in production by Air France and KPMG. KPMG has two engineers developing new features that will eventually be merged in the open source version. EUR is holding regular meetings with Air France to receive feedback and ideas for improvement.

- It is also important to note that we are actively updating our public GitHub repositories with a wide range of open source components, developed and used within IOStack project and available to any interested open source or commercial community: `https://github.com/iostackproject`.

2.4   Contribution to Open Data

IOStack project is committed to open data and we are following the path, set in the data management plan. Datasets are publicly available for researchers as green open data. Currently available datasets that we have are available here:

- Arctur Web Workload. This dataset holds webservers log entries which are used to work on the Arctur's use case: `http://iostack.eu/datasets-menu/download/6-datasets-menu/17-arcturtrace`.

- Idiada Document Database holds a collection of documents which will be used to work on the Idiada's use case: `http://iostack.eu/datasets-menu/download/6-datasets-menu/16-idiadatrace`.

- GridPocket provided a data generator which is used to generate simulated data for the GridPocket use case. The simulator is based on real-life data and follows the same patterns and attributes. We have to use data generator in this case since we are not allowed to publicly expose the sensitive end-users data. Even though only generator is publicly we will do internal experiments on real data; GridPocket data generator: `https://github.com/gridpocket/project-iostack`.

3 Year 3 Dissemination Activities

3.1 Academic conferences and publications

- Gracia-Tinedo R, Sampé J, Zamora E, Sánchez-Artigas M, García-López P, Moatti Y, Rom E. Crystal: Software-Defined Storage for Multi-Tenant Object Stores. In15th USENIX Conference on File and Storage Technologies (FAST 17) 2017 Feb. USENIX Association.

- Moatti Y, Rom E, Gracia-Tinedo R, Naor D, Chen D, Sampe J, Sanchez-Artigas M, Garcıa-Lopez P, Gluszak F, Deschdt E, Pace F. Too Big to Eat: Boosting Analytics Data Ingestion from Object Stores with Scoop. InData Engineering (ICDE), 2017 IEEE 33rd International Conference on 2017 Apr 19 (pp. 309-320). IEEE. [21]

- Vernik G, Factor M, Kolodner EK, Ofer E, Michiardi P, Pace F. Stocator: a high performance object store connector for spark. InProceedings of the 10th ACM International Systems and Storage Conference 2017 May 22 (p. 27). ACM. [22]

- Pace F, Venzano D, Carra D, Michiardi P. Flexible scheduling of distributed analytic applications. InProceedings of the 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing 2017 May 14 (pp. 100-109). IEEE Press. [23]

- Sutra P, Rivière E, Cotes C, Artigas MS, Lopez PG, Bernard E, Burns W, Zamarreño G. CRESON: Callable and Replicated Shared Objects over NoSQL. InDistributed Computing Systems (ICDCS), 2017 IEEE 37th International Conference on 2017 Jun 5 (pp. 115-128). IEEE.

- In July 2017 BSC had a Summer School lecture at Heraklion (GR) where they explained JIT pre-fetching inside as ML session.

- In September 2017 IBM and Eurecom prepared a poster titled "Stocator: An Object Store Aware Connector for Spark" that was published at the SOCC 2017 [24] conference in Santa Clara (USA)

- In November 2017 URV held a presentation "IOStack: Software Defined Storage for Big Data" at the Imperial College London (UK)

- In December 2017 URV published a conference paper titled "Data-driven Serverless Functions for Object Storage" for the Middleware 2017[25] conference in Las Vegas (USA)

3.2 Mass and Social media

- In March 2017 URV published an article titled "Giving Wings to Your Data: A First Experience of Personal Cloud Interoperability" in the FGCS Journal

- URV regularly updated the project webpage and Twitter feed.

- Arctur produced new videos that present the project as a whole as well as individual use cases. The videos were done by a professional animator and were in the final stages of production during the writing of this document. Once finalized they will be put on the internet and shared to the public.

---

[21] http://ieeexplore.ieee.org/document/7929987/
[22] https://www.systor.org/2017/posters.html
[23] https://dl.acm.org/citation.cfm?id=3101126
[24] https://acmsocc.github.io/2017/accepted-posters.html
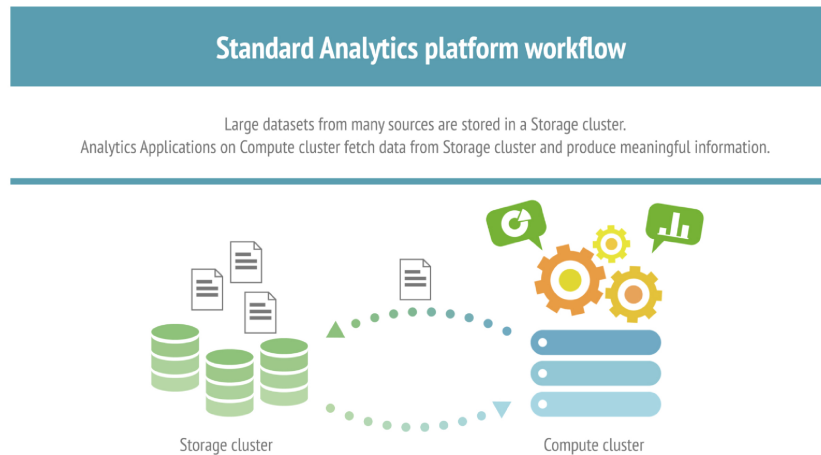[25] http://2017.middleware-conference.org/accepted.html

Figure 3: Screen shot from one of the new animations

3.3   Industrial events and external collaborations

- In January 2017 GridPocket attended the CES 2017 international trade show in Las Vegas (USA).

- In March 2017 URV contacted with the OpenIO project. OpenIO is a Object Storage provider interested in offering Big Data services in the next future. OpenIO and IOStack has started a collaboration to integrate technologies in OpenIO (like Crystal and Stocator), but also to participate together in joint project proposals.

- In March 2017 IBM presented the results of the IOStack project at the Big Things Meet up in Haifa (Israel)

- In June 2017 URV held a remote presentation to the team lead at Pacific Northwest National Laboratory about IOStack achievements. They expect to organize a future remote presentation to the PNNL team about Software Defined Storage for Big Data technologies.

- In June 2017 GridPocket attended the Smart Energies Summit Paris national trade show in France.

- In June 2017 URV attended the ICDCS 2017 at the Aragonne National Laboratory: Argonne Laboratory are developing software defined technologies that can benefit from our approach.

- In June 2017 IBM presented the Stocator technology at the STRATA '17 Conference in London (UK)

- In June 2017 IBM presented a conference paper on Stocator at the Spark Summit in San Francisco (USA)

- In June 2017 MPS and Arctur attended the ISC 17 [26]event in Frankfurt

- In July 2017 BSC made contact with the BigStorage project in Heraklion (GR) and shared the S3QL modification for an ESR.

- In July 2017 GridPocket attended the Innovative City 2017 event in Nice, France.

- In September 2017 IBM presented the IOStack results at the Haifa Cloud Meet up in Haifa (Israel)

---

[26]www.isc-hpc.com

- In September 2017 Arctur attended the OpenStack Days UK event in London (UK) and gave a presentation titled "IOStack: Software-Defined Storage for Big Data"

- In October 2017 URV held a remote presentation to the Pacific Northwest National Laboratory (Advanced Computing, Mathematics, and Data Division): "IOStack: Software Defined Storage for Big Data"

- In October 2017 GridPocket attended the European Utility Week 2017 in Amsterdam (NL)

- In October 2017 IBM gave a talk titled "Using Pluggable Apache Spark SQL Filters to Help GridPocket Users keep up with the Jones and save the planet" at the Spark Summit Europe 2017 [27] in Dublin (IE)

- In November 2017 BSC published a conference paper titled "Improving OpenStack Swift interaction with the I/O Stack to enable Software Defined Storage". The paper contributes to the IEEE SC2 conference held in Kanazawa, Japan.

- In November 2017 URV presented the IOStack project at their stand during the Connect-EU Conference in Barcelona(E)

- BSC also evaluated the option to use Swift as an Object Storage option for NextGENIO EU Project in WP5, discarded because the environment layout was not appropriate for OpenStack.

3.4 IOStack workshop

In September 2017 Arctur attended the OpenStack Days IT event in Milano (IT) where in cooperation with the local organizer held a small workshop as part of the OpenStack Days event. The session was attended by only a small subset of all attendees of the event, but this fact is not unexpected since the main topic of the workshop was intended to deal with a very specific topic which is Software Defined Storage.

Figure 4: Opening Session of the OpenStack IT Days event

The workshop hosted a number of speakers who covered various topics. The silver lining of topics revolved around production environments on OpenStack, how to deploy such environments and of course also about Software Defined Storage and the way froward.

The talks that were held and their titles and the respective organizations were as follows:

---

[27] https://spark-summit.org/eu-2017/events/using-pluggable-apache-spark-sql-filters-to-help-gridpocket-users-keep-up-with-the-jones-and-save-the-planet/

- **Mirantis:** Hybrid Clouds APPS without pain: Using Spinnaker to orchestrate applications

- **Prozeta:** World's fastest production-ready software defined storage for OpenStack

- **Switch:** Deploy Kubernetes on OpenStack using Cloud LBAAS and the Cinder volumes

- **Sardinia Systems:** Full life-cycle lessons from a large scale government private Cloud

- **Huawei:** Kuryr Services at scale or how to reach the clouds with K8S and Neutron

- **Arctur:** IOStack: An Integrated Toolkit on top of OpenStack

After the talks the workshop continued and was concluded with an open discussion where all the attendees were involved. The topics of the discussion revolved about the experiences and problems incurred during deploying OpenStack as a whole or even individual components. The general feeling that was shared amongst all involved was that OpenStack is at a stage where Linux was in the mid 90's. The common opinion was that OpenStack isn't really mature yet but all agreed that it is rapidly gaining momentum and stability needed to gain a stable foothold. The general consensus was that this will for sure come true in the next five to ten years. The discussion also covered the topic of Software Defined Storage. All the involved generally agreed that SDS are promising technologies but there is still some work to be done for it to be applicable to production environments.

The audience that attended the workshop was from Italy, France, Spain, the Untied Kingdom, Ukraine and Slovenia. The attendees came form academic as well as industrial backgrounds.

3.5   Update on community involvement

In June 2017 URV and Arctur joint efforts to contact the OpenStack community and got the Storlet description published in the official OpenStack documentation: `https://wiki.openstack.org/wiki/Storlets#Related_Projects`

The consortium is also in contact with the main development team of OpenStack trying to get Crystal published on the official OpenStack page as an ˝Associated project˝, but at the time of writing of this proposal we still haven't managed to have it published.

4   Exploitation

To better understand and unify the exploitation ideas from different partners an "exploitation work-shop" was organized during the project consortium meeting in Slovenia in June 2017. To gather and analyse the available data the Osterwalder Business Model Canvas method was chosen. This method gathers inside views of the owners of the product or service at offer for which the Business Model is being formulated. This data is then systematically organized and categorized into a pre-defined table. In such way all the influences and actors that act upon our service or product. So we are able to fully understand the processes and dependencies before going to market. We do this my method-ically analysing nine key components of the product or service. Formal descriptions of the business become the building blocks for its activities. Many different business conceptualizations exist; Os-terwalder's work and thesis propose a single reference model based on the similarities of a wide range of business model conceptualizations. With his business model design template, we can easily describe their business model. The key components are as follows:

- **Infrastructure**

  - *Key Activities:* The most important activities in executing a company's value proposition. An example for Bic, the pen manufacturer, would be creating an efficient supply chain to drive down costs.

  - *Key Resources:* The resources that are necessary to create value for the customer. They are con-sidered an asset to a company, which are needed in order to sustain and support the business. These resources could be human, financial, physical and intellectual.

  - *Partner Network:* In order to optimize operations and reduce risks of a business model, orga-nization usually cultivate buyer-supplier relationships so they can focus on their core activity. Complementary business alliances also can be considered through joint ventures, strategic al-liances between competitors or non-competitors.

- **Offering**

  - *Value Propositions:* The collection of products and services a business offers to meet the needs of its customers. According to Osterwalder, (2004), a company's value proposition is what distinguishes itself from its competitors. The value proposition provides value through var-ious elements such as newness, performance, customization, "getting the job done", design, brand/status, price, cost reduction, risk reduction, accessibility, and convenience/usability. The value propositions may be: Quantitative (price and efficiency) or Qualitative (overall cus-tomer experience and outcome).

- **Customers**

  - *Customer Segments:* To build an effective business model, a company must identify which cus-tomers it tries to serve. Various sets of customers can be segmented based on the different needs and attributes to ensure appropriate implementation of corporate strategy meets the character-istics of selected group of clients. The different types of customer segments include:

    Mass Market: There is no specific segmentation for a company that follows the Mass Market element as the organization displays a wide view of potential clients. e.g. Car

    Niche Market: Customer segmentation based on specialized needs and characteristics of its clients. e.g. Rolex

    Segmented: A company applies additional segmentation within existing customer segment. In the segmented situation, the business may further distinguish its clients based on gender, age,

and/or income.

Diversify: A business serves multiple customer segments with different needs and characteristics.

Multi-Sided Platform / Market: For a smooth day-to-day business operation, some companies will serve mutually dependent customer segment. A credit card company will provide services to credit card holders while simultaneously assisting merchants who accept those credit cards.

- *Channels:* A company can deliver its value proposition to its targeted customers through different channels. Effective channels will distribute a company's value proposition in ways that are fast, efficient and cost effective. An organization can reach its clients either through its own channels (store front), partner channels (major distributors), or a combination of both.

- *Customer Relationships:* To ensure the survival and success of any businesses, companies must identify the type of relationship they want to create with their customer segments. Various forms of customer relationships include:

Personal Assistance: Assistance in a form of employee-customer interaction. Such assistance is performed either during sales, after sales, and/or both.

Dedicated Personal Assistance: The most intimate and hands on personal assistance where a sales representative is assigned to handle all the needs and questions of a special set of clients.

Self Service: The type of relationship that translates from the indirect interaction between the company and the clients. Here, an organization provides the tools needed for the customers to serve themselves easily and effectively.

Automated Services: A system similar to self-service but more personalized as it has the ability to identify individual customers and his/her preferences. An example of this would be Amazon.com making book suggestion based on the characteristics of the previous book purchased.

Communities: Creating a community allows for a direct interaction among different clients and the company. The community platform produces a scenario where knowledge can be shared and problems are solved between different clients.

Co-creation: A personal relationship is created through the customer's direct input in the final outcome of the company's products/services.

- **Finances**

- *Cost Structure:* This describes the most important monetary consequences while operating under different business models. A company's DOC.

Classes of Business Structures:

Cost-Driven – This business model focuses on minimizing all costs and having no frills. e.g. Low cost airlines

Value-Driven – Less concerned with cost, this business model focuses on creating value for their products and services. e.g. Louis Vuitton, Rolex

Characteristics of Cost Structures:

Fixed Costs – Costs are unchanged across different applications. e.g. salary, rent

Variable Costs – These costs vary depending on the amount of production of goods or services. e.g. music festivals

Economies of Scale – Costs go down as the amount of good are ordered or produced.

Economies of Scope – Costs go down due to incorporating other businesses which have a direct relation to the original product.

- *Revenue Streams:* The way a company makes income from each customer segment. Several ways to generate a revenue stream:
Asset Sale – (the most common type) Selling ownership rights to a physical good. e.g. retail corporations

Usage Fee – Money generated from the use of a particular service e.g. UPS

Subscription Fees – Revenue generated by selling a continuous service. e.g. Netflix

Lending/Leasing/Renting – Giving exclusive right to an asset for a particular period of time. e.g. Leasing a Car

Licensing – Revenue generated from charging for the use of a protected intellectual property.

Brokerage Fees – Revenue generated from an intermediate service between 2 parties. e.g. Broker selling a house for commission

Advertising – Revenue generated from charging fees for product advertising.

The project partners were asked to analyse and provide feedback and their points of view on each individual component of the Canvas. The results were collected and inserted into the canvas. Based on the results it has become quite apparent that individual exploitation expectations differ from partner to partner, but two general exploitation niches clearly stand out from the final canvas. We can say that we have an academic/research oriented exploitation plan and an end market oriented exploitation plan. This not an unexpected result. The two different exploitation plans logically come from the fact that the project consortium is composed both of academic/research partners and of industrial partners focusing on the actual end users.
Due to this fact the two Business Model Canvases were created and are presented in the following chapters.
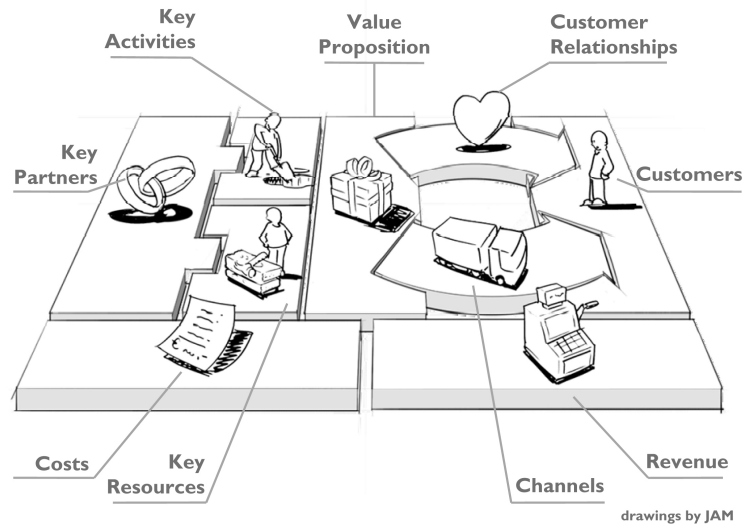
Figure 5: The Osterwalder Business Model Canvas

4.1   Academic Business Model Canvas

It is only logical that academic business partners of the project are more research focused when compared to the industrial partners. The business model therefore reflects the fact that the Academic partners are interested to share the knowledge and to further develop the IOStack technologies. The academic business model canvas is graphically represented in the following figure.

As we can see in the canvas the Academic partners focus more on the community involvement in
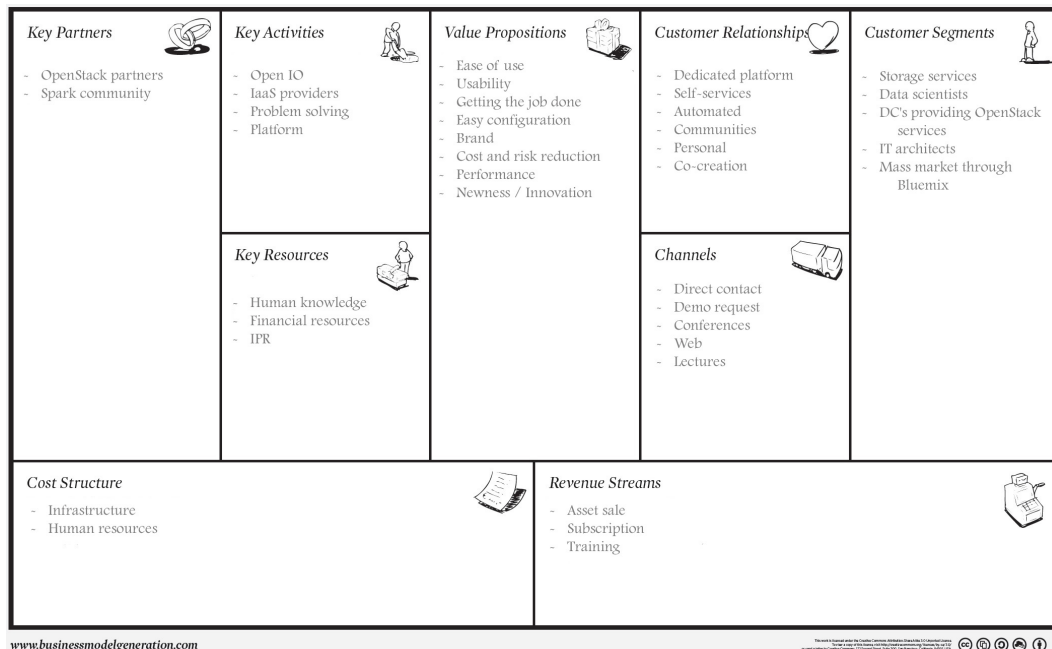

Figure 6: Academic Osterwalder Business Model Canvas

terms of further development of the technology. The resources focus on the the human resources an on the intellectual property rights to leverage further leverage current state of development. The key customers are similar communities whereas the cost structure is focused on the human resources implemented in delivering the services. Also the revenue streams are focused on further disseminating

the know-how gained during the project.

Based on the overall assessment of the Academic Business Model Canvas we can formulate the following exploitation statement:

The academic partners of the IOStack project will leverage the developments and gained knowledge from the project to further develop the platform and expertise. To achieve this they will take advantage of the OpenStack and Spark communities to push forward the developments and knowledge. The results will be offered to similar communities and to Data Centre providers. Revenue will primary be generated from consulting services.

Going beyond the general exploitation statement the partners have started to go beyond the project to exploit the results of the project. BSC plans to use the results of the project into 3 different ways. First, they plan to further develop ideas such as just-in-time pre-fetching, SDS in object stores, and bandwidth differentiation on future projects of the group. Second, they intend to use the experience gained in developing solutions in OpenStack, and especially with Swift and Cinder, in order to better develop the consulting actions BSC undergoes with local industry. An third, BSC will include SDS as part of the MIRI (Master in Innovation and Research in Informatics) where some of the participants of the projects teach. At the same they will include lessons learned of the project in BSC training events whenever suitable.

Similarly URV is especially interested in bringing forward the Crystal component. Crystal is now an innovative Software Defined Storage solution for OpenStack Swift. URV as the major developer and promoter of Crystal is exploring different exploitation paths:

1. Production-ready: Idiada and URV are collaborating together to make Crystal production-ready so that it can be used as the storage solution in Idiada. There is now a pilot in Idiada that will enter in production in January 2018. From that point, Idiada and URV can work together to sustain and commercialize the product.

2. Venture Capital and start-up: URV has applied to the open call called "The Collider" (thecollider.tech) offered by Mobile World Capital Barcelona. It is a program to invest venture capital money to create and boost start-ups from promising research groups. Our group has been selected this year to create a company around Crystal, but due to lack of personal resources, we plan to follow that path in 2018.

3. Collaborations with well-established players in the Storage market: We outline now two major potential collaborations: OpenIO and SwiftStack. URV group has already established an open collaboration with OpenIO to promote IOStack technologies in their company. In particular, OpenIO is interested in entering the HPC and Big Data market, and IOStack results (Crystal, Stoccator) will be considered. We have also applied to EU H2020 projects with OpenIO to reinforce our collaboration. Finally, we have initial talks with SwiftStack to begin collaborations with our research group.

4.2 Industrial Business Model Canvas

Similarly to the Academic, also the Industrial partners have identified the potential of the IOStack developments. Of course due to the fact that the Industrial partners are mare market oriented, also the Business Model Canvas reflects this fact. The industrial Business Model Canvas is presented in the figure below.

As we can see in the Business Model Canvas the industrial partners focus more on generating additional business revenue and leverage the results to try to gin new customers. That is why the Key Partners are both the development community as well as existing customers that can be used as reference cases for new customers. Also the industrial partners admit the the Key Activities are to maintain and further develop the platform while the price/performance aspect of the offer is the
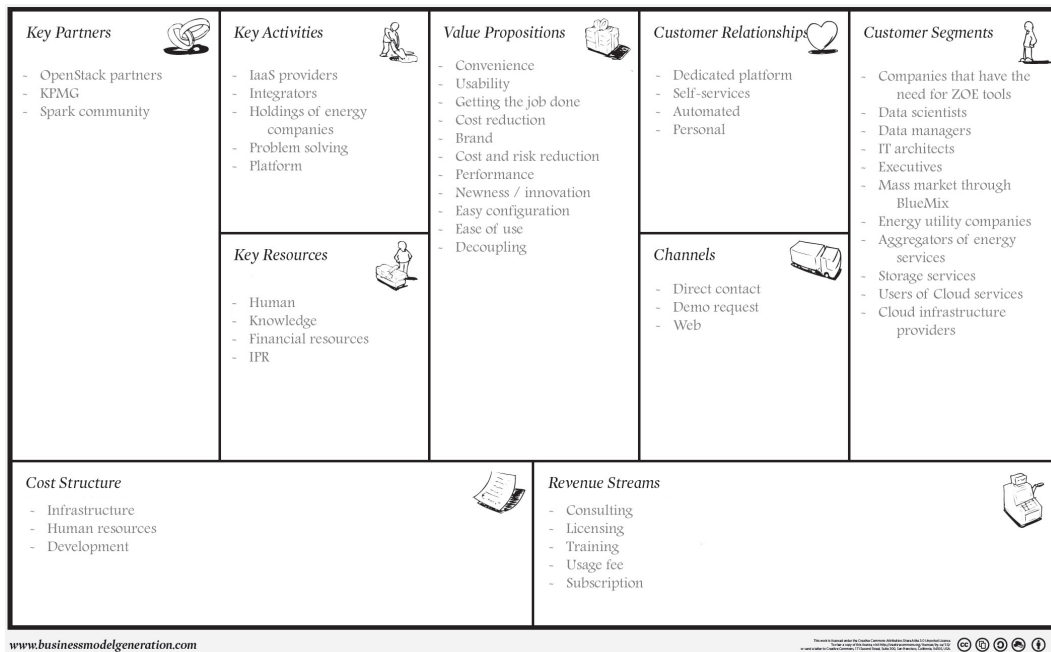
Figure 7: Industrial Osterwalder Business Model Canvas

most important in the Value Proposition. The Customers will most definitely be found amongst companies with large datasets and Data Centre service providers as well as by up-selling existing customers. The most important resources are both the human resources and the infrastructure used to run the services. The most important components of the price are the costs of development and hosting infrastructure. Whereas for selling the services the partners foresee various licensing models.

based on this Business Model we can form the following Exploitation statement for the industrial partners:

The IOStack industrial will leverage the results and the knowledge gained during the project to expand their offer or service portfolio. By implementing various licensing models different end customers will be focused. The price will be adapted based on the resources that were used to maintain and develop the platform. The key aspect of the solution for the end customers of the solution is the maximum price-performance of the solution. The partners will therefore use the results of the project to gain new customers or offer new services.

Individual partners have already started some concrete steps for exploitation plans and some initial exploitation results are already showing.

Arctur at the time of writing of this document had established the IOStack platform on their infrastructure. Being a rather small DC services provider - when comparing to global DC providers - Arctur at this time does not have actual users of the IOStack platform features at this moment. Although Arctur is firmly convinced that the results of the IOStack will come in to play with the development of the infrastructure and business in the future.

Achievements and developments of IOStack have also been implemented by Eurecom. The Zoe project has been conceived as an open-source project, and received attention by several companies. In particular, two companies are already using Zoe in production to manage and deploy analytics applications: KPMG Germany and AirFrance/KLM.

Both companies have different ways of participating to an open source project.

For what concerns KPMG, the company required a tight integration of Zoe in their production environment which called for the definition of a sub-project to establish: a sound development method-

ology (unit test, continuous integration, ...), some custom components (support for K8s, application marketplace, web user interface), and consulting services. As such, we established a research contract to achieve such tasks.

For what concerns AirFrance/KLM, requirements were different: first, the company currently can not publicly commit to participate to open-source projects. As a consequence, the interaction Eurecom had with AF/KLM was mostly taking the form of meetings whereby Eurecom collected feedback on performance issues, and bugs in general. Secondly, AF/KLM datacenter engineers developed their own components to integrate Zoe with their existing infrastructure.

Furthermore GridPocket needs to store and process large amount of data coming from smart meters almost in real time. This processing is required for tasks such as enabling end energy consumers to track their consumptions and compare themselves to their neighbours, or tasks related to the data scientists performing interactive analytic requests. The main performance metrics that will need to be improved are :

- volume and execution time of requests that need to be run to have a specific result (e.g. average consumption in a given area or around a given meter ID),

- duration of reading and writing processes of the real time big data ingestion (ETL).

With the current data sets, the processing time stays reasonable short and acceptable for customer (both utilities and end-users). Nevertheless, in the coming months this processing time is expected to become longer and longer as the number of GridPocket's consumers and meters will keep increasing. GridPocket anticipate a need to review its architecture within the next 12-18 months, and possibly shift into one of the most appropriate solutions : Objects Storage and Spark.

IOStack project has demonstrated the benefit of using Objects Storage and Spark through tests run on our use case. It also gave GridPocket an opportunity to build hand on experience in integration of these technologies with energy use cases. For an efficient adaptation of these technologies into GridPocket's system architecture, GridPocket will consider further collaboration with IBM, in order to leverage on IBM's advance regarding object storage and spark implementation and exploitation.

Idiada is already using the IOStack toolkit in a pilot on their premises. They have a new OpenStack Swift cluster that uses Crystal as the Software-Defined Storage layer. File system access is provided by two external tools (CloudFuse for Linux and OpenStackDrive for Windows). A number of filters and controllers were implemented in Crystal to fulfil Idiada varied requirements like data reduction, file tagging, encryption, access control or removal of files based on rules.
In this pilot, the IOStack toolkit has proven to be flexible enough to adapt the installation to the policies required by an industry with high security requirements.

MPSTOR has learned that storage provisioning in the data center is a basic requirement. The lowest practical level of provisioning is through storage APIs and command lines tools, low level command line (python based) tools are seen as very powerful means to automate datacenter processes. It is not obvious which virtualisation frameworks will be winners so therefore which higher level frameworks to integrate such as OpenStack, Kubernetes etc. is difficult to judge. The best strategy is to have a set of provisioning tools that range from low level API, storage Array command lines tools, SDS command line tools to finally SDS integration with higher level cloud or virtualisation frameworks. A second key lesson is that end to end provisioning is the valuable part of storage provisioning, it is not simply the creation of storage volumes from storage arrays that is important but the attaching through datacenter fabrics of those volumes to the final consumer of the storage. The final lesson we learned was in-line filters is a useful framework as it allows the storage array functionality to be extended. The filter framework competes for compute resources so careful consideration needs to be given to what filters should be developed, we believe for general workloads

data protection, data security are more appropriate filters than say compression or deduplication. For very specific applications filters provide a very good framework to do in-line processing of the data. Since the start of IOStack many changes have occurred in the datacenter, such as new fabric protocols iSER and new fabrics and protocols such as NVMe. NVMe has the potential to radically change how the disaggregation of compute and storage is implemented. NVMe by itself is almost designed to solve the disaggregation requirement for block storage in IOStack. MPSTOR will therefore try to find the best fit between new technology changes (ex NVMe) , datacenter practice (devops CLI tools) and virtualisation framework changes (ex Docker, Kubernetes, OpenStack) and tailor the IOStack block layer development so that it becomes part of the standard automation tools from MPSTOR.

5   Standardization

Within the efforts of the IOStack project no new standards were produced or determined. This was not the intention of the project. However the project tried to contribute as much as possible to the existing OpenStack standards by extending and contributing to the established components. There are two specific fields where the project partners extensively extended and are worth pointing out:

- OpenStack Storlets: is a public OpenStack project with a specific Middleware and Compute Engine for an OpenStack Swift compute framework that runs compute within a Swift cluster: `https://github.com/openstack/storlets`

- Stocator: Stocator is high performing connector to object storage for Apache Spark, achieving performance by leveraging object store semantics: `https://github.com/SparkTC/stocator`

IOStack partners are noted as contributors in both components.

6   Conclusions

As we can see from the document the Project has invested a considerable amount of effort and energy to disseminate the results and developments of the project. A systematic and well planned approach has over the past three years produced a significant number of well noted results. The most significant result is that IOStack has been recognized as a good input by the OpenStack community and has been listed amongst other outside solutions on the official OpenStack webpage.

The dissemination and development efforts will now have to continue with the exploitation of the results of the project. As we can see the initial exploitation steps were already taken. It is now up to the partners to take advantage of the momentum that the IOStack project has given to convert the results into practical applications of the platform.

References

[1] R. Gracia-Tinedo, D. Harnik, D. Naor, D. Sotnikov, S. Toledo, and A. Zuck, "SDGen: mimicking datasets for content generation in storage benchmarks," in USENIX FAST'15, pp. 317–330, 2015.

[2] R. Gracia-Tinedo, P. García-López, M. Sánchez-Artigas, J. Sampé, Y. Moatti, E. Rom, D. Naor, R. Nou, T. Cortés, W. Oppermann, and P. Michiardi, "Iostack: Software-defined object storage," IEEE Internet Computing, vol. 20, no. 3, pp. 10–18, 2016.

[3] "Openstack (israel, june 2015)." http://www.slideshare.net/openstackil/ibm-swift.

[4] "Openstack (tokyo, november 2015)." https://www.openstack.org/summit/tokyo-2015/videos/presentation/storlets-making-swift-more-software-defined-than-ever.