D4.3 Final report on Cloud standards profile development (Update 1)

A European Cloud observatory supporting cloud policies, standards profiles & services

www.cloudwatchhub.eu|info@cloudwatchhub.eu

This report provides a discussion on the standards profile development activities within the CloudWATCH project. The report does not describe a fixed set of standards profiles but a methodology that allowed the development of standards profiles according to specific characteristics that are of interest for a group of participants in a cloud computing environment that aim on using a common profile. This methodology is exercised by three examples.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

CloudWATCH Mission

The CloudWATCH mission is to accelerate the adoption of cloud computing across European private and public organisations. CloudWATCH offers independent, practical tips on why, when and how to move to the cloud, showcasing success stories that demonstrate real world benefits of cloud computing. CloudWATCH fosters interoperable services and solutions to broaden choice for consumers. CloudWATCH provides tips on legal and contractual issues. CloudWATCH offers insights on real issues like security, trust and data protection. CloudWATCH is driving focused work on common standards profiles with practical guidance on relevant standards and certification Schemes for trusted cloud services across the European Union.

The CloudWATCH partnership brings together experts on cloud computing; certification schemes; security; interoperability; standards implementation and roadmapping as well as legal professionals. The partners have a collective network spanning 24 European member states and 4 associate countries. This network includes: 80 corporate members representing 10,000 companies that employ 2 million citizens and generate 1 trillion in revenue; 100s of partnerships with SMEs and 60 global chapters pushing for standardisation, and a scientific user base of over 22,000.

Disclaimer

CloudWATCH (A European Cloud Observatory supporting cloud policies, standard profiles and services) is funded by the European Commission's Unit on Software and Services, Cloud Computing within DG Connect under the 7th Framework Programme.

The information, views and tips set out in this publication are those of the CloudWATCH Consortium and its pool of international experts and cannot be considered to reflect the views of the European Commission.

Document information summary

Document title:	Final report on Cloud standards profile development
Main Author(s):	Peter Deussen, Fraunhofer FOKUS
Contributing author(s):	Michel Drescher, EGI.eu David Wallom, UOXF Neil Caithness, UOXF Jesus Luna, CSA Patrice Chazerand, DE
Reviewer(s):	Silvana Muscella & Nicholas Ferguson, Trust-IT
Target audiences:	Cloud computing projects and initiatives including projects funded by Unit E2 Software, Services and Cloud Computing projects Standards Development Organisations. In particular IEEE P2301, OGF OCCI.
Keywords:	Cloud standards profiles, interoperability, clustering
Deliverable nature:	R
Dissemination level: (Confidentiality)	PU
Contractual delivery date:	M20
Actual delivery date:	Final version M22; Final - Updated version M25
Version:	Final – Updated version
Reference to related publications	

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Executive Summary

This report provides a discussion on the standards profile development activities within the CloudWATCH project. The report does not describe a fixed set of standards profiles as originally planned. Instead, it provides a more flexible methodology that allows a certain group of participants in a cloud computing eco-system (providers, customers, 3rd party service providers, legal organisations, etc.) (a) to identify common interests regarding standard profiling, and (b) to derive a focussed set of use cases as a basis for profiling. These use cases form the basis for the selection of standards that may contribute to the profile under development, and for understanding how to restrict or extend these standards as part of the actual profiling work.

This methodology is exercised by means of three examples. Stakeholder representatives have been chosen as 38 cloud computing projects including a number of European Commission funded projects. These have been grouped into five clusters of which three have been analysed further:

- **Cluster 1 - Scientific computing.** This cluster comprises a number of projects that aim at highly distributed data processing in an academic context.
- **Cluster 2 – Trusted public clouds for government.** This cluster consists of a set of initiatives driven by public sector organisations.
- **Cluster 3 – High performance, dedicated purpose applications.** This cluster is similar to Cluster 1 but comprises projects concentrating on high performance computing that are more focussed regarding their objectives.

Moreover, we have discussed possible combinations of standards for each of these clusters, and provided some guidelines on how to profile them. For Cluster 1, a complete profile based on the CIMI (Cloud Infrastructure Management Interface) standard has been provided as a reference model for future profiling activities.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Table of Contents

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Table of Tables

Table of Figures

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

1 Introduction

This report provides a discussion on the standards profile development activities within the CloudWATCH project. The report does not describe a fixed set of standards profiles as originally planned. This approach has proven unsuitable for the following reasons:

- A standard profile for a given application domain (our original list comprises public sector, private industry, and academia) is likely to be too general to be useful. Generic requirements from such domains are not strong enough to (a) select standards that build the basis of such a generic profile and (b) result in meaningful restrictions and extension.
- Generic standard profiles are a moving target. Any methodology to derive standards profiles must be based on the collection and analysis of use cases. But depending on the chosen granularity, the number of use cases that can be applied can be very large, and provide contradicting requirements. Hence, a more focussed approach is required to identify the scope and target audience of standard profiles.

Therefore, we decided to employ a more flexible approach. We assume that a standard profile will be of use for a certain subset of participants in a cloud computing eco-system (providers, customers, 3rd party service providers, legal organisations, etc.). For this report, we have chosen a number of initiatives including EC funded projects as sample set. Our goal is to provide these stakeholder groups not with a given set of profiles of arguable usefulness but with a general approach to derive standard profiles matching the specific requirements of these groups. This methodology is based on the following steps:

- Obtaining a ranking of stakeholder interests with regard to 13 cloud computing characteristics (an extension of the five characteristics provided in the final NIST definition of cloud computing) through a questionnaire or an automated tool.[1]
- Applying a clustering algorithm to the resulting set of stakeholder/characteristics vectors. These clusters identify groups of stakeholders that share a common interest for some set of characteristics, while at the same time agreeing that certain other characteristics are of minor importance for them.
- In some cases, a further analysis is required to understand the degree of cohesion within these clusters in more detail. This analysis may result in splitting a cluster that is to general into several smaller ones.
- Now a clear understanding of which groups of stakeholders agree on the importance/unimportance of which characteristics. This information is used to identify use cases that will be considered for standard profile definition more precisely. Use case analysis aims now on (a) the identification of standards that are suitable for a standards profile useful for the stakeholder group in question, and (b) to annotate these standards with additions that form the actual profile. Annotations are:
    - **Notes** clarifying normative or non-normative text in a standard.

[1] At the time of writing this report, an online version of such a ranking tool is under development and is planned to be made available at the CloudWATCHHub.eu web site.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

- **Restrictions** altering the allowed interpretations from the original underlying set to a smaller set of interpretations and implementations.
- **Extensions** utilizing intentional extension points.

This report provides three “case studies” to evaluate the suitability of our approach. The set of stakeholders’ chosen data points for the clustering activity comprises of 38 cloud projects, including EC-funded projects. Details of these projects will be included in an updated version of D2.3 Final User Stories published in August 2015. Initiatives from our original target domains public sector, private industry, and academia, as outlined in D2.1 Reference Model Framework Report, are included. From this set of projects, five clusters are derived. Three of them have been selected as illustrative examples to show how our methodology works.

- **Cluster 1 - Scientific computing.** This cluster comprises a number of projects that aim on highly distributed data processing in an academic context.
- **Cluster 2 – Trusted public clouds for government.** This cluster consist of a set of initiatives driven by public sector organisations
- **Cluster 3 – High performance, dedicated purpose applications.** This cluster is similar to Cluster 1 but comprises projects concentrating on high performance computing that are more focussed regarding their objectives.

The derivation of standard profiles for each of these clusters is beyond the scope of this report. We therefore discuss possible combinations of standards for each of these clusters, and provide some guidelines on how to profile them.

This report is organized as follows. Section 1 provides the reader with a brief overview of the objectives and results described in this deliverable. Section 2 summarises the methodology the project used to cluster 38 EU projects in a meaningful manner. Section 3 describes the methodology developed to efficiently generate straw-man cloud standards profiles. Section 4 applies said methodology to three example clusters of those presented earlier in the document. Section 5 looks forward, outlining work that still lays ahead, and which CloudWatch legacy can help along this path. The deliverable concludes with section 6, and is complemented with two appendices providing ancillary information.

To this updated version of D4.3, three new appendices are included. Firstly, appendix 3 provides a comprehensive overview of developing a security cloud standards profile. Appendices 4 and 5 looks at a small business and enterprise (SME/SMB) perspective of the importance of standards based on collaboration with National Trade Associations (NTA).

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

2 Brief review of the project clustering methodology

The methodology described in this report was not planned as part of the original CloudWATCH project, but has been developed as an emergent need. The method is described in detail in the CloudWATCH deliverable D2.4 (Policy and compliance requirements Report) and only a brief description is provided here, along with a presentation of the first-iteration results that form the working basis for the rest of this report.

One approach to the task of deriving standards profiles for cloud computing applications requires an understanding of the landscape of applications and the identification of groups of applications within this landscape that have similar requirements, or even more ambitiously, similar aspirations. With no way to effectively identify such groups from the broad array of applications in the European cloud environment, CloudWATCH sought to develop an empirical approach. Possible empirical metrics are suggested by the NIST definition of cloud computing[2]. In an early draft of the definition, NIST present a list of 13 characteristics comprised of five that were considered essential, and eight that were considered to be common characteristics. In its final publication NIST dropped the eight common characteristics from the definition retaining only the five essential. (See NIST special Publication 800-145 [NIST-800-145].) We retain the full list for our initial development of the method presented here, and we will report on a comparative study using both the long and short versions of the definition in D2.4. We also present the method and the comparative study at the *NIST Cloud Computing Forum & Workshop VIII* being held in Gaithersburg, MD, USA July 7-10, 2015[3].

Five "Essential Characteristics"

1. On-demand self service
2. Broad network access
3. Resource pooling
4. Rapid elasticity
5. Measured service

Eight "Common Characteristics"

6. Massive Scale
7. Homogeneity
8. Virtualization
9. Low Cost Software
10. Resilient Computing
11. Geographic Distribution
12. Service Orientation
13. Advanced Security

[2] http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf
[3] http://www.nist.gov/itl/cloud/cloud_computing_wkshp_viii.cfm

To use the NIST characteristics as an empirical metrics for identifying potential clusters of cloud applications, we need to score the metrics for a representative sample of European and international applications. We refer to these applications from here on as projects. We present here a cluster analysis of these cloud computing projects as a way of gaining further insight into the general landscape of cloud computing. Our intention is two-fold. First, to verify how well the NIST model fits and clarifies the landscape, both in its long and short forms, and second, to provide the much needed insight for our process of developing cloud computing standards profiles. The objectives of this empirical analysis are to discover groups of projects that are consistent in their relationship to a set of well-defined general characteristics, and distinct from other such groups.

We perform our analysis with a dataset representing 38 cloud projects scored against the full set of 13 NIST defining characteristics on an interval scale. Our clustering procedure is based on the outcome of a Principal Components Analysis [PCA] and we interpret the landscape on a simultaneous biplot of the characteristic coefficients and component scores.

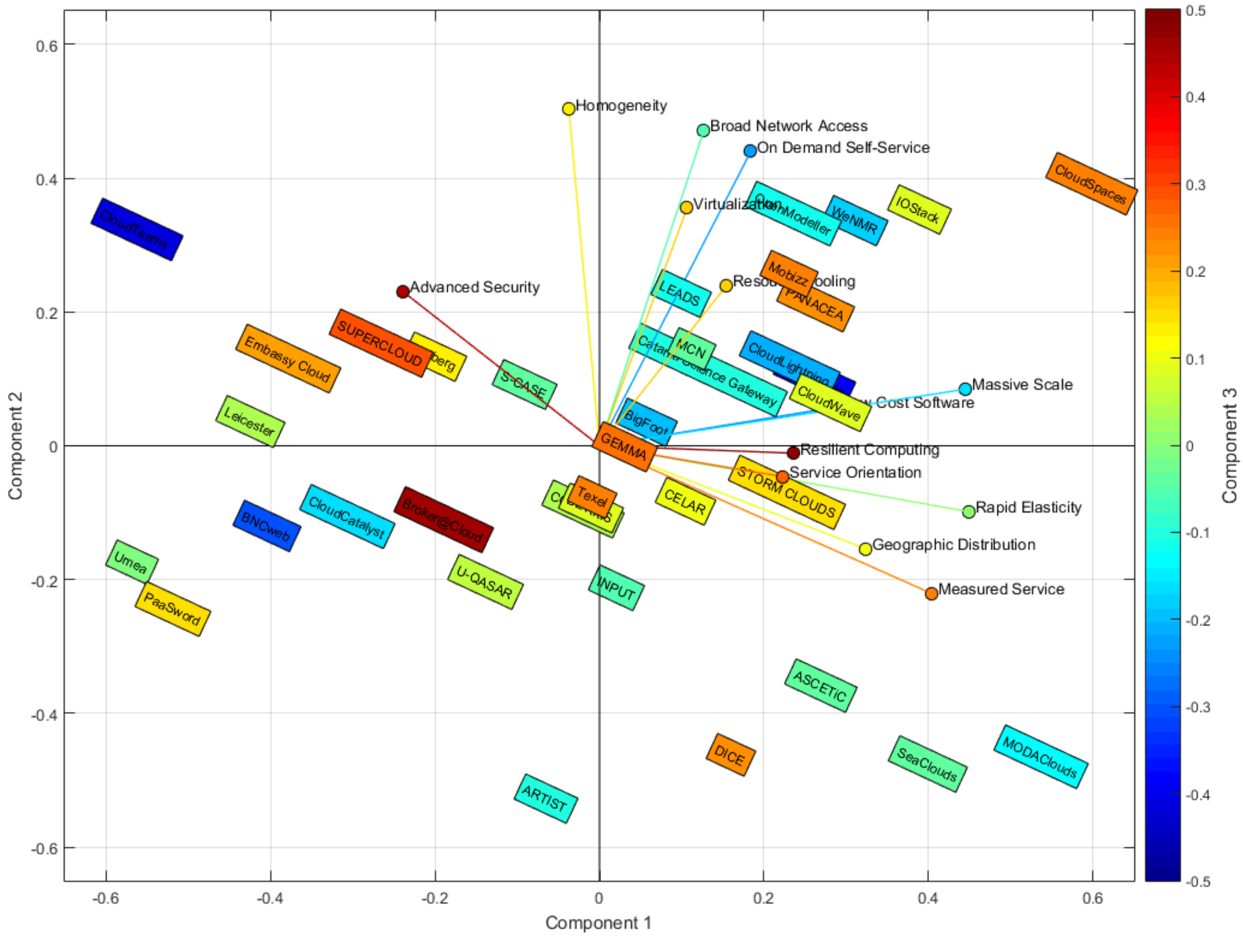


Figure 1. Biplot of 38 European Cloud Projects

The biplot shows simultaneously two features of the analysis. First, the location of each project in the 13-dimensional space of the principal components, here showing components 1 and 2 on the conventional x-, and y-axes, and component 3 colour coded for depth on the z-axis. Second, the orientation of the space is shown with respect to the loadings on each of the original characteristics. The

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

alignment of clusters of projects with the characteristic vectors is the principal interpretive mechanism that we employ.

Next we perform a cluster analysis of projects in the principal components ordination space. There are many clustering algorithms and we choose a simple Euclidian distance algorithm operating on the first five principle components only (i.e. those with eigenvalues greater than one – the well-known Kaiser-Guttman criterion[4].)

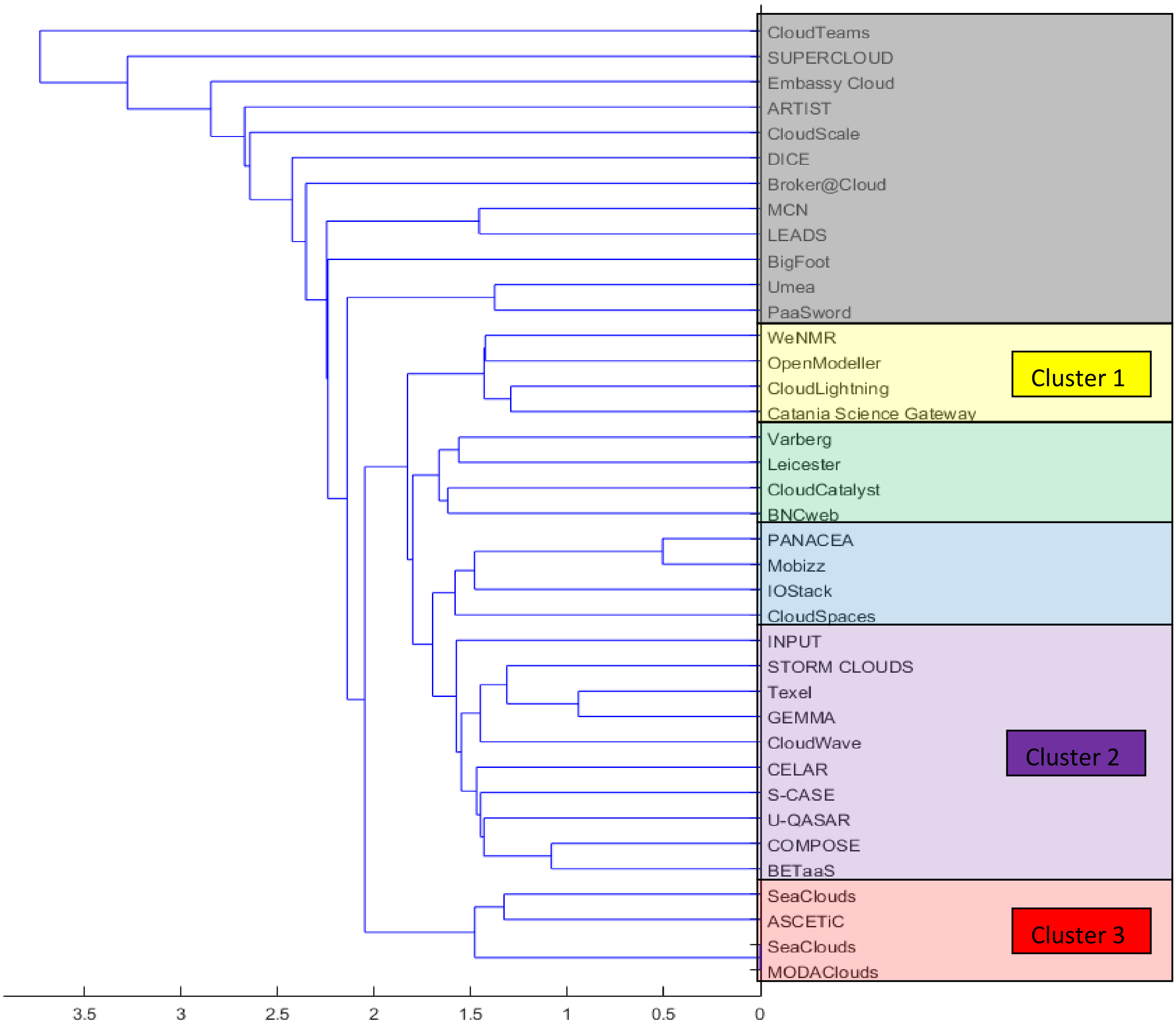
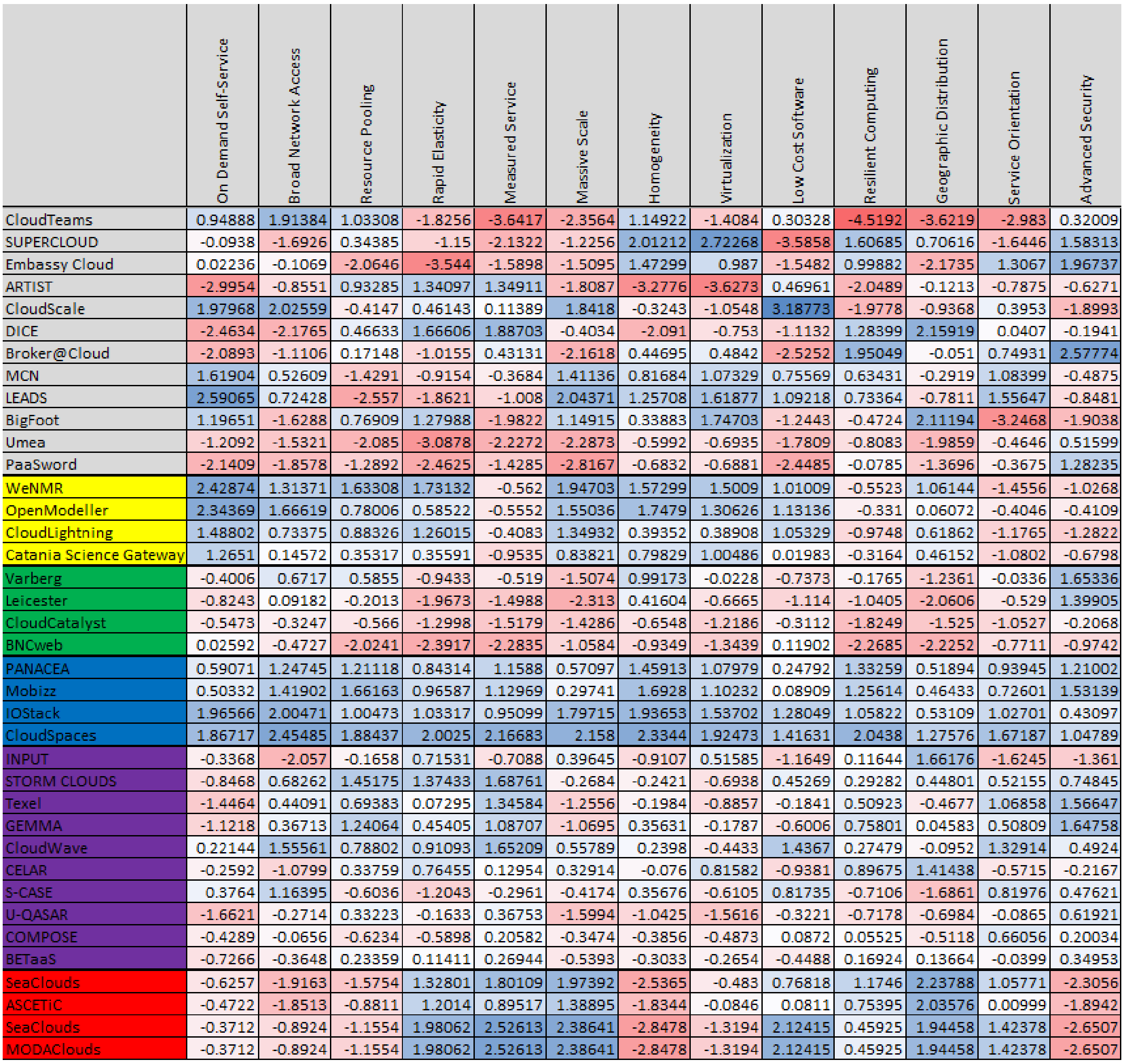


Figure 2. Cluster tree for 38 European Cloud Projects

[4] Yeomans, KA and Golder, PA. 1982. The Guttman-Kaiser Criterion as a predictor of the Number of Common Factors. *The Statistician*, 31(3) 221-229.

A breakdown of clusters is shown in the table below, where, as a natural extension of the biplot we derive a numerical interpretation that provides a ranking of the NIST characteristics for each project. These values are simply the projections of each project onto each characteristic vector in turn. Grouping projects into clusters, we can calculate the cluster mean and standard deviation values which will be used in the interpretations that follows.



	On Demand Self-Service	Broad Network Access	Resource Pooling	Rapid Elasticity	Measured Service	Massive Scale	Homogeneity	Virtualization	Low Cost Software	Resilient Computing	Geographic Distribution	Service Orientation	Advanced Security
CloudTeams	0.94888	1.91384	1.03308	-1.8256	-3.6417	-2.3564	1.14922	-1.4084	0.30328	-4.5192	-3.6219	-2.983	0.32009
SUPERCLOUD	-0.0938	-1.6926	0.34385	-1.15	-2.1322	-1.2256	2.01212	2.72268	-3.5858	1.60685	0.70616	-1.6446	1.58313
Embassy Cloud	0.02236	-0.1069	-2.0646	-3.544	-1.5898	-1.5095	1.47299	0.987	-1.5482	0.99882	-2.1735	1.3067	1.96737
ARTIST	-2.9954	-0.8551	0.93285	1.34097	1.34911	-1.8087	-3.2776	-3.6273	0.46961	-2.0489	-0.1213	-0.7875	-0.6271
CloudScale	1.97968	2.02559	-0.4147	0.46143	0.11389	1.8418	-0.3243	-1.0548	3.18773	-1.9778	-0.9368	0.3953	-1.8993
DICE	-2.4634	-2.1765	0.46633	1.66606	1.88703	-0.4034	-2.091	-0.753	-1.1132	1.28399	2.15919	0.0407	-0.1941
Broker@Cloud	-2.0893	-1.1106	0.17148	-1.0155	0.43131	-2.1618	0.44695	0.4842	-2.5252	1.95049	-0.051	0.74931	2.57774
MCN	1.61904	0.52609	-1.4291	-0.9154	-0.3684	1.41136	0.81684	1.07329	0.75569	0.63431	-0.2919	1.08399	-0.4875
LEADS	2.59065	0.72428	-2.557	-1.8621	-1.008	2.04371	1.25708	1.61877	1.09218	0.73364	-0.7811	1.55647	-0.8481
BigFoot	1.19651	-1.6288	0.76909	1.27988	-1.9822	1.14915	0.33883	1.74703	-1.2443	-0.4724	2.11194	-3.2468	-1.9038
Umea	-1.2092	-1.5321	-2.085	-3.0878	-2.2272	-2.2873	-0.5992	-0.6935	-1.7809	-0.8083	-1.9859	-0.4646	0.51599
PaaSword	-2.1409	-1.8578	-1.2892	-2.4625	-1.4285	-2.8167	-0.6832	-0.6881	-2.4485	-0.0785	-1.3696	-0.3675	1.28235
WeNMR	2.42874	1.31371	1.63308	1.73132	-0.562	1.94703	1.57299	1.5009	1.01009	-0.5523	1.06144	-1.4556	-1.0268
OpenModeller	2.34369	1.66619	0.78006	0.58522	-0.5552	1.55036	1.7479	1.30626	1.13136	-0.331	0.06072	-0.4046	-0.4109
CloudLightning	1.48802	0.73375	0.88326	1.26015	-0.4083	1.34932	0.39352	0.38908	1.05329	-0.9748	0.61862	-1.1765	-1.2822
Catania Science Gateway	1.2651	0.14572	0.35317	0.35591	-0.9535	0.83821	0.79829	1.00486	0.01983	-0.3164	0.46152	-1.0802	-0.6798
Varberg	-0.4006	0.6717	0.5855	-0.9433	-0.519	-1.5074	0.99173	-0.0228	-0.7373	-0.1765	-1.2361	-0.0336	1.65336
Leicester	-0.8243	0.09182	-0.2013	-1.9673	-1.4988	-2.313	0.41604	-0.6665	-1.114	-1.0405	-2.0606	-0.529	1.39905
CloudCatalyst	-0.5473	-0.3247	-0.566	-1.2998	-1.5179	-1.4286	-0.6548	-1.2186	-0.3112	-1.8249	-1.525	-1.0527	-0.2068
BNCweb	0.02592	-0.4727	-2.0241	-2.3917	-2.2835	-1.0584	-0.9349	-1.3439	0.11902	-2.2685	-2.2252	-0.7711	-0.9742
PANACEA	0.59071	1.24745	1.21118	0.84314	1.1588	0.57097	1.45913	1.07979	0.24792	1.33259	0.51894	0.93945	1.21002
Mobizz	0.50332	1.41902	1.66163	0.96587	1.12969	0.29741	1.6928	1.10232	0.08909	1.25614	0.46433	0.72601	1.53139
IOStack	1.96566	2.00471	1.00473	1.03317	0.95099	1.79715	1.93653	1.53702	1.28049	1.05822	0.53109	1.02701	0.43097
CloudSpaces	1.86717	2.45485	1.88437	2.0025	2.16683	2.158	2.3344	1.92473	1.41631	2.0438	1.27576	1.67187	1.04789
INPUT	-0.3368	-2.057	-0.1658	0.71531	-0.7088	0.39645	-0.9107	0.51585	-1.1649	0.11644	1.66176	-1.6245	-1.361
STORM CLOUDS	-0.8468	0.68262	1.45175	1.37433	1.68761	-0.2684	-0.2421	-0.6938	0.45269	0.29282	0.44801	0.52155	0.74845
Texel	-1.4464	0.44091	0.69383	0.07295	1.34584	-1.2556	-0.1984	-0.8857	-0.1841	0.50923	-0.4677	1.06858	1.56647
GEMMA	-1.1218	0.36713	1.24064	0.45405	1.08707	-1.0695	0.35631	-0.1787	-0.6006	0.75801	0.04583	0.50809	1.64758
CloudWave	0.22144	1.55561	0.78802	0.91093	1.65209	0.55789	0.2398	-0.4433	1.4367	0.27479	-0.0952	1.32914	0.4924
CELAR	-0.2592	-1.0799	0.33759	0.76455	0.12954	0.32914	-0.076	0.81582	-0.9381	0.89675	1.41438	-0.5715	-0.2167
S-CASE	0.3764	1.16395	-0.6036	-1.2043	-0.2961	-0.4174	0.35676	-0.6105	0.81735	-0.7106	-1.6861	0.81976	0.47621
U-QASAR	-1.6621	-0.2714	0.33223	-0.1633	0.36753	-1.5994	-1.0425	-1.5616	-0.3221	-0.7178	-0.6984	-0.0865	0.61921
COMPOSE	-0.4289	-0.0656	-0.6234	-0.5898	0.20582	-0.3474	-0.3856	-0.4873	0.0872	0.05525	-0.5118	0.66056	0.20034
BETaaS	-0.7266	-0.3648	0.23359	0.11411	0.26944	-0.5393	-0.3033	-0.2654	-0.4488	0.16924	0.13664	-0.0399	0.34953
SeaClouds	-0.6257	-1.9163	-1.5754	1.32801	1.80109	1.97392	-2.5365	-0.483	0.76818	1.1746	2.23788	1.05771	-2.3056
ASCETiC	-0.4722	-1.8513	-0.8811	1.2014	0.89517	1.38895	-1.8344	-0.0846	0.0811	0.75395	2.03576	0.00999	-1.8942
SeaClouds	-0.3712	-0.8924	-1.1554	1.98062	2.52613	2.38641	-2.8478	-1.3194	2.12415	0.45925	1.94458	1.42378	-2.6507
MODAClouds	-0.3712	-0.8924	-1.1554	1.98062	2.52613	2.38641	-2.8478	-1.3194	2.12415	0.45925	1.94458	1.42378	-2.6507

Table 1. A clustered numerical interpretation of the biplot.

3 From cluster to straw-man standards profile

The project has developed a process to generate repeatable and reliable data to cluster projects and activities into groups of similar interests (see Section 2 for a summary, and D2.4 for a complete description).

To achieve similar quality in arriving at meaningful standards profiles straw-man documents, the project had to develop a methodology that would ensure the quality of the resulting project clusters. In order to determine which standards would be suitable candidates for profiling for a given cluster, we essentially need to answer the following questions:

1. Is there a way to assure, or at least evaluate the quality in our collected data, and cluster assignment? Is our clustering approach going in the right direction? And, if our data sample meets our expectations, what information can we derive from it?
2. Putting the NIST cloud characteristics into the context of standards and their applicability/scope, which of these characteristics provide scope for standardisation? Is there a methodology for determining or estimating the suitability of a NIST cloud characteristic for standardisation which we can focus our effort on?
3. Is there a common service model used within the cluster? Or is the result divergent in that some member projects operate on one service model (e.g. IaaS) while others operate on another (e.g. SaaS)?
4. Which service model is addressed/implied by any given existing Cloud standard? For example, while it is very clear that CIMI and OCCI both address the IaaS[5] model, we conjecture that out of those two, only OCCI would be also suitable to address SaaS[6] properly even with yet to be devised extensions.
5. How suitable is a standard specification for profiling? Are there commonalities in how to analyse a standard specification for profiling? If so, which are these?

The following subsections will describe our approach in more detail.

3.1 Cluster data quality

The methodology for defining clusters as described in D2.4 generates a “heat map” style of visual representation of the data entered by the projects themselves (or CloudWATCH representatives). Next to the visual 3D bi-plot visualisation of the clustering methodology, the 2-dimensional spreadsheet provides a numeric representation of the statistical analysis of the data. Instead of ranging from 0 to 9 as originally entered into the data gathering tool, values now represent the relative importance of the given cloud characteristic for the respective project. Values close to or equal to zero represent a neutral stance, negative values symbolise lesser or the least importance, and positive values denote higher to highest importance of the characteristic.

[5] Infrastructure as a Service
[6] Software as a Service

	On Demand Self-Service	Broad Network Access	Resource Pooling	Rapid Elasticity	Measured Service	Massive Scale	Homogeneity	Virtualization	Low Cost Software	Resilient Computing	Geographic Distribution	Service Orientation	Advanced Security
MODAClouds	-0.371	-0.892	-1.155	1.981	2.526	2.386	-2.848	-1.319	2.124	0.459	1.945	1.424	-2.651

Figure 3. An example of relative importance representation of cloud characteristics

Figure 3 provides an example of how the result might look like. The MODAClouds project has a very developed opinion about the importance of certain cloud characteristics: While Measured Service, Massive Scale, and Low Cost Software are the three most important characteristics for this project, it Homogeneity or Advanced Security is of less importance. The MODACloud project is relatively neutral towards Resilient Computing and, slightly less, towards On Demand Self-Service.

The statistical project clustering methodology produced an ordered list of projects similar to a genealogical tree, which was used to define the clusters. The result of this work was a set of five clusters of projects, with 12 projects that did not fit any of the clusters. To determine if the cluster has the right grouping or expresses any statistical relevance, we examine the following three indicators per cluster:

1. **Agreement coefficient (AC).** The agreement coefficient of the cluster on a given NIST cloud characteristic is defined as the average value over all cluster projects' individual rating for said characteristic. Thus, there will be 13 agreement coefficients per cluster, for all 13 NIST cloud characteristics.
2. **Cluster cohesion (CC).** The cluster cohesion value on a given NIST cloud characteristic is defined as the standard deviation (based on a sample, not an entire population) for the agreement coefficient. The cluster cohesion value indicates the cluster projects' consensus on the agreement coefficient. Large values indicate poor consensus, whereas small values indicate strong consensus, i.e. all projects rating the importance of the pertinent characteristic consistently and with very similar values.
3. **Signal-to-Noise ratio (SNR).** The signal-to-noise ratio, i.e. the agreement coefficient over the cluster cohesion, allows us to assess the overall quality of the ratings provided by the cluster.

The higher the absolute value of the agreement coefficient and the smaller the cluster cohesion the more significance and meaningful value the agreement coefficient bears for further analysis and work: The more significant the agreement coefficient, the higher the chances that the projects in the respective cluster will engage and take on the profile development in the future.

Similar values for agreement coefficient and cluster cohesion, regardless their value, indicate statistical noise of little or no significance.

In the case of low values for both indicators, it may be advisable to reconsider the composition of the cluster in order to obtain significant results. Currently, there is no clear strategy we can formulate for this scenario. We however conjecture that a strongly and significantly scoped cluster correlates with it being represented very close to a balanced binary tree in the genealogy tree representation of the entire sample. Conversely, degenerate binary tree patterns seem to correlate with very poor or no agreement coefficient for a cluster that includes those projects. Cluster 4 (see below) is an example of this observation.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

We expect that a significant cluster should express a ratio between agreement coefficient and cluster cohesion of 2 or better. However, given the small overall sample size we do not impose this threshold as a strict requirement.

3.2 Functional vs. non-functional cloud characteristics

The NIST definition of Cloud computing (Special Publication SP 800-145 describes five essential characteristics, three service models, and four deployment models for cloud computing. It is very important to understand that these describe three orthogonal aspects of cloud computing, i.e. one can freely combine these to describe a specific instance of cloud computing without any of the defined terms losing any validity. For example, a Software-as-a-Service offering allowing the user to self-enrol and configure the service to their needs (hence satisfying the on-demand self-service characteristic) is just as much a cloud computing service as any other offering combining other service models and characteristics.

Before arriving at a published set of five essential characteristics, earlier drafts of the NIST SP 800-145 publication also described 8 further characteristics that were classified as "common characteristics". All thirteen form the basis for the activities described in this deliverable.

The distinction between functional and non-functional requirements in software engineering has a profound impact on their implementation, in that functional requirements are reflected in designing means of direct user interaction and intentional influence, whereas non-functional systems express intrinsic system behaviour. For example, for a system to successfully provide the result of multiplying two numbers, an obvious functional requirement for such a system would be to allow the user to enter both factors. A non-functional requirement might be for the system to provide the result within 1 second once the user presses a "calculate now" button.

Applying the requirements engineering methodology allows us to determine which of the NIST characteristics provide scope in standardisation conjecturing that a well-defined set of standards would need combining to satisfy any given NIST cloud characteristic. Reviewing the definition for the NIST essential and common cloud characteristics, and applying the requirements engineering methodology provides us with the following classification:

Functional cloud characteristics (E denoting essential characteristics):

- [E] On-demand self service
- [E] Broad network access
- [E] Measured service
- Virtualisation
- Resilient computing
- Geographic distribution
- Advanced Security

Non-functional cloud-characteristics:

- [E] Resource Pooling

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

- [E] Rapid elasticity
- Massive scale
- Homogeneity
- Low-cost software
- Service Orientation

This classification has impact on the cloud standards profile effort in that we see scope for standardisation (and consequently, for profiling) for functional cloud characteristics. Therefore, focus should be on developing straw-man contents around those cloud characteristics.

3.3 Reviewing and condensing project use cases

CloudWatch has used the Use Case cards developed in WP2 for a variety of information gathering – section 0 (Appendix 1) gives a template for these[7]. Here, a condensed version of these use case cards focus on key information around the NIST cloud definition, i.e. the characteristics, the deployment models (private, public, hybrid), and service models (IaaS, PaaS, SaaS).

To derive a useful straw-man standards profile document, we need to know which service models the projects are employing (or are going to), and to confirm project grading of importance of cloud characteristics. Finding that information in a consistent manner is the goal of this step. This mainly comprises of desktop research, reviewing project publications, and other means of data collection.

3.4 Cloud standards service models

This step might be an obvious necessity, but it is important to make this explicit, since not all standard specifications are clear about intended and unintended service model applicability.

While, for example, it is clear that OCCI, CIMI, OVF and CDMI are all standards that primarily target the IaaS model. Indeed, often the name of the standard expresses such intention. This however, is less clear for exemplary standards such as TOSCA or, again, OCCI. While TOSCA clearly addresses the PaaS (Platform-as-a-Service) model, it might actually be used in a hybrid scenario where an IaaS provider offers an add-on that consumes TOSCA manifests in order to automate VM provisioning and lifecycle management for their IaaS customers.

OCCI, as the other example, is so versatile that, with the appropriate extension, one can use OCCI to standardise service management functions that apply to any service model in the cloud computing paradigm.

3.5 Reviewing standards for profiling

Rarely, is a standard specification is perfectly scoped around a well-defined set of use cases. However, almost always, even standard documents in the ICT sector document compromises between the developing parties, or allow a certain level of flexibility in anticipated implementations, perhaps to support a larger set of use cases, thus increasing applicability and relevance of the developed standards. This, unfortunately, negatively impacts on other aspects of the standard: The more use cases it

[7] See also D2.1 Reference Model Framework Report

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

supports, the more relevant it may become, but at the same time it increases the probability of implementations that are *not* interoperable among each other.

Profile documents intend to scope down the incorporated standards, reducing the supported use cases and thus removing variance in possible alternative interpretations and implementation, usually to exactly one interpretation, aspiring to guarantee interoperability.

When profiling a standard, three common best practices have prevailed as accepted techniques of assembling and defining profiles:

- **Notes.** Notes are used in profile documents to clarify normative or non-normative text in a standard. Formally, notes do not alter the interpretation of the referenced part of the standard, but supply clarification for better and easier understanding of the underlying original text.
- **Restrictions.** Standard specifications provide alternative interpretations, message configurations and workflows, as indicated earlier. Restrictions formally alter the allowed interpretations from the original underlying set to a smaller set of interpretations and implementations, usually to exactly 1 – the one that is defined in the profile document.
- **Extensions.** Occasionally, standard specifications intentionally omit defining a certain condition or possible rendering of a message, workflow element or any other element. These definition holes are almost never omissions, but intentional extension points, employing a standardisation technique called “defining the base line”.

Notes, though formal elements of a profile specification, do not alter the meaning and interpretation of underlying standards. Restrictions and Extensions, however, do alter the meaning and interpretation of an underlying specification. There is some debate whether a profile is formally allowed to break conformance to an incorporated standard by intentionally disregarding one or more of the unconditional statements made therein, or altering the interpretation of a normative clause in an incompatible way. Most profiling specifications restrict themselves to maintaining compatibility with the underlying specification(s), i.e. an implementation conforming to the profile in question would also implicitly conform to the profiled original specification. There do exist some examples of profile documents that break the conformance chain: Conforming to the profile would force an implementation to lose conformance to the underlying specification.

For the purpose of CloudWATCH, the latter is strongly advised against since such an approach, in our view, would defeat the purpose of standardisation as such.

3.5.1 Inventorying possible restrictions for a profile

Standard specifications often use their own language to express levels of compliance with them. A very popular specification, if not the one and only specification used to define levels of conformance in this space is [RFC2119], defining the following keywords. They are commonly grouped in two categories, i.e. “unconditional” and “conditional” in terms of conformance. Keywords in the “unconditional” category leave no alternative in interpretation and implementation. They are imperative in their language nature. “Conditional” keywords are those that provide scope for restricting conformance claims in a profile document, as they allow for alternatives depending on the context, which makes achieving interoperability in different implementations difficult or outright impossible.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Unconditional keywords: MUST – MUST NOT – REQUIRED – SHALL – SHALL NOT

Conditional keywords[8]**:** SHOULD – SHOULD NOT – RECOMMENDED – NOT RECOMMENDED – MAY – OPTIONAL

Normative text in standard specifications employing conditional keywords need to be assessed for applicability in the context of the cloud computing cluster for which we are drafting the straw-man profile documents, and added to the inventory of potential conformance claims made in the profile document.

3.5.2 Inventory of extension points

Often, standards do not define every possible use case that might apply. If it did, a standard specification would never be finished. To deal with such situations, specifications often explicitly define "extension points", where the authors intentionally leave it to the reader to complement the standard to fit their respective purpose.

Such elements in a specification are very useful since they can be exploited in many ways, including the purpose of the standard profile specifications illustrated in this deliverable.

Extension points frequently appear in formal language definitions such as XML, in the form of an "<xsd:any/>" element, indicating that, as far as the formal specification is concerned, *any* further content conforming to the language definition is allowed to appear at this point.

As there are no best practices for defining and describing extension points in a standard specification it is difficult to identify them. Very frequently though, extension points are found in specifications that standardise payload-bearing infrastructures, digital envelopes and other wrapping techniques. A good example is the OCCI family of specifications, which uses extension points as a primary design principle for the specifications themselves. Another example standard employing extension points describes MIME-multipart messages that are often used to encode Emails with several attachments, mandating the following E-Mail message structure:

1. MIME envelope
    a. Mail message header
    b. Mail message body (7-bit ASCII encoded)
    c. MIME multipart envelope (defining the message part separator string)
        i. Mail message body (8-bit ASCII encoded to accommodate HTML messages)
        ii. Attachment 1
        iii. Attachment 2
        iv. …

[8] RFC2119 makes a further distinction between MAY and OPTIONAL and the other conditional keywords, in that MAY and OPTIONAL are truly indicating equally valid alternatives, whereas the other conditional keywords indicate exceptional, yet still valid circumstances within which such deviation from the specification are still conformant. But this makes no difference to our current situation.

The definition of the amount of multipart parts is potentially unlimited (one extension point a profile might want to constrain), as well as the type of data contained in each attachment (second extension point).

Arriving at this point in the standards profile development process, enough material is available to construct a straw-man document outlining the intended contents, incorporated standard specifications, and number and targets of conformance claims made in the profile itself.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

4 Straw-man cloud standards profiles

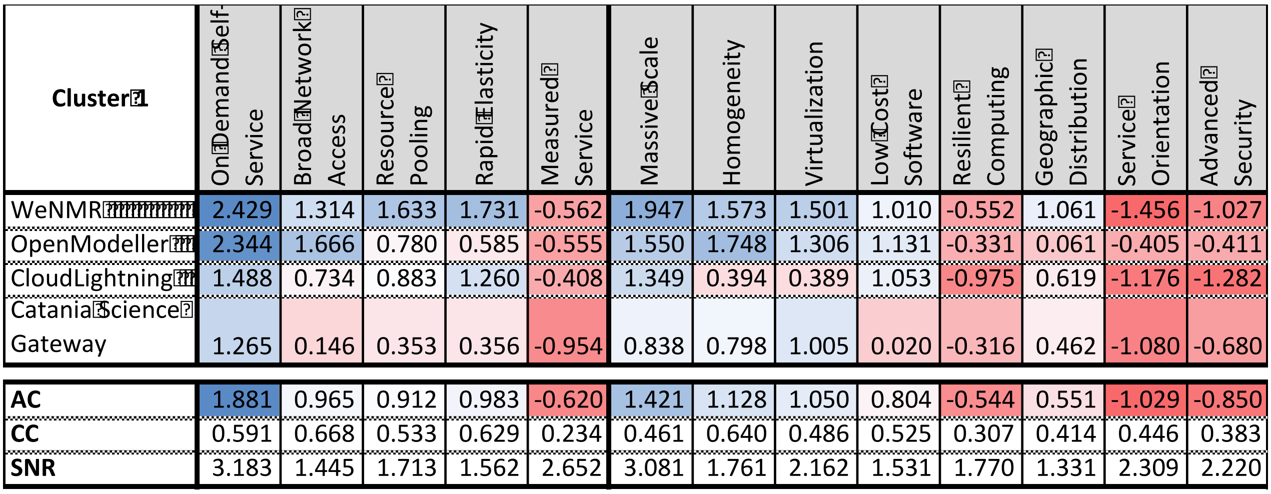
In this section, we provide an in-depth analysis of three of the five clusters identified in Section 2.

4.1 Cluster 1 – Scientific Computing

The "Scientific Computing" cluster is an example of a cluster where the contained projects are at large agreement on the importance and unimportance of characteristics.

4.1.1 Cluster Quality Assessment

Cluster 1 has the following "heat map":



Cluster 1	On Demand Self-Service	Broad Network Access	Resource Pooling	Rapid Elasticity	Measured Service	Massive Scale	Homogeneity	Virtualization	Low Cost Software	Resilient Computing	Geographic Distribution	Service Orientation	Advanced Security
WeNMR	2.429	1.314	1.633	1.731	-0.562	1.947	1.573	1.501	1.010	-0.552	1.061	-1.456	-1.027
OpenModeller	2.344	1.666	0.780	0.585	-0.555	1.550	1.748	1.306	1.131	-0.331	0.061	-0.405	-0.411
CloudLightning	1.488	0.734	0.883	1.260	-0.408	1.349	0.394	0.389	1.053	-0.975	0.619	-1.176	-1.282
Catania Science Gateway	1.265	0.146	0.353	0.356	-0.954	0.838	0.798	1.005	0.020	-0.316	0.462	-1.080	-0.680
AC	1.881	0.965	0.912	0.983	-0.620	1.421	1.128	1.050	0.804	-0.544	0.551	-1.029	-0.850
CC	0.591	0.668	0.533	0.629	0.234	0.461	0.640	0.486	0.525	0.307	0.414	0.446	0.383
SNR	3.183	1.445	1.713	1.562	2.652	3.081	1.761	2.162	1.531	1.770	1.331	2.309	2.220

Figure 4. "Heat map" correlation of EU projects and cloud characteristics importance for Cluster 1.

The following observations can be made:

- **Strong agreement among projects**. The agreement coefficients for the NIST characteristics of high importance, namely "on-demand self-service" (1.881), "massive scale" (1.421), and "homogeneity" (1.128) are comparatively high. Likewise, coefficients for the characteristics of low importance, "service orientation" (-1.029), "advanced security" (-0.850), "measured service" (-0.620), and "resilient computing" (-0.544, with low standard deviation) are comparatively low.
- **Small deviation**. In all cases, the standard deviation is considerably smaller than the absolute value of the agreement coefficients, although not all ratios of agreement coefficient and standard deviation reach the significance threshold of 2 (see Section 3.1), We accept this for the "homogeneity" characteristic since within the cluster there is only a single project that is divergent and then only to a minor degree. Hence, individual results are well grouped around the average.

Therefore, cluster 1 provides a suitable basis for analysis of use cases derived from the four projects, and the subsequent derivation of standard profiles.

4.1.2 Reviewing cluster projects use cases

Cluster 1 comprises four projects:

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

- WeNMR
- OpenModeller
- CloudLighting
- Catania Science Gateway

4.1.2.1 WeNMR

WeNMR - A worldwide e-Infrastructure for NMR and structural biology[9]

Description: WeNMR is a worldwide e-Infrastructure for nuclear magnetic resonance (NMR) spectroscopy and structural biology. It is the largest Virtual Organization in the Life sciences and is supported by EGI.
WeNMR provides services from the following categories:

- Processing
- Assignment
- Analysis
- Structure Calculation
- Molecular Dynamics
- Modelling
- Tools

Goals and aspirations: WeNMR demonstrates how data intensive applications with high potential of parallelization can be deployed on a Grid or Cloud infrastructure.

TECHNICAL ASPECTS	
Criteria for success	Deployment of heterogeneous applications based on a homogeneous resource management model Unified data access for NMR data

4.1.2.2 openModeller

openModeller[10]

Description: openModeller provides a flexible, robust, cross-platform environment to carry out ecological niche modelling experiments. It is comprised of a single framework written in C++ allowing multilpe interfaces on top of it, such as command line programs, Desktop interface, Web interface and Web Service interface. The framework includes facilities for sampling points, creating, testing, evaluating and projecting models into different environmental scenarios, reading species occurrence and environmental data in different formats, among many other features. More than 10 algorithms are available as plugins.

Goals and aspirations: openModeller demonstrates an application allows for the parallel execution of a number of experiments (based on different algorithms/configurations). Data exchange between nodes carrying out individual experiments is not required.

TECHNICAL ASPECTS

[9] https://www.wenmr.eu/
[10] http://openmodeller.sourceforge.net/

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Criteria for success	Deployment of heterogeneous applications based on a homogenious resource management model (algorithm plugins, addressed by a html/soap API). Unified data access for model data

4.1.2.3 CloudLightning

CloudLightning - A self-organising, self-managing heterogeneous cloud[11]
Description: The project proposes to create a new way of provisioning heterogeneous cloud resources to deliver services, specified by the user, using a bespoke service description language. The project addresses energy inefficiencies particularly in the use of resources and consequently to deliver savings to the cloud provider and the cloud consumer in terms of reduced power consumption and improved service delivery, with hyperscale systems particularly in mind. CloudLightning will implement three use cases: Genomics, Oil & Gas Exploration and Ray Tracing. CloudWATCH has selected the Genomics use case for the purposes of this document.
Goals and aspirations: CloudLightning will demonstrate the use of a large-scale simulation for genome processing with the anticipation of greater energy efficiency resulting in lower costs. As the cost of the raw sequencing technology drops, the computing challenge becomes the final significant technology bottleneck preventing the routine use of genomics data in clinical settings. CloudLightning will target this both through the use of heterogeneous computing technologies to offer significantly improved performance/cost and performance/Watt, but also enabling this computation to be hosted at large-scale in the cloud, making it practical for wide-scale use. In addition to realigning the computation cost factors in genome processing with sequencing costs, it can significantly improve the genome processing throughput and speed of genome sequence computation. This will have the effect of reducing the wider cycle time thus increasing the volume and quality of related research.

TECHNICAL ASPECTS	
Criteria for success	• To build a prototype management system and delivery model: CloudLightning will develop a software stack for power efficient cloud infrastructure management, based on the principles of self-organisation and self-management. This will be augmented with a declarative cloud delivery model that promotes access to heterogeneous resources. • Validation of approach with use cases from three application domains – Genomics, Oil & Gas Exploration, and Ray Tracing. The specific use cases will be augmented with an analysis of dense and sparse matrix analysis techniques that have broad application in a wide variety of fields. • Demonstrate scalability: CloudLightning will be designed to manage the ultra- and hyperscale cloud infrastructures of the future. A testbed running the CloudLightning software stack will be used to gather instrumentation data that will form the basis of large-scale simulations of self-organised and self-managed hyperscale heterogeneous clouds.

[11] http://cloudlightning.eu/

4.1.2.4 Catania Science Gateway - DECIDE

The Catania Science Gateway provides a portal to access a number of more specialized science gateways. In this report, we concentrate on the Diagnostic Enhancement of Confidence by an International Distributed Environment (DECIDE).

DECIDE uses a Grid based infrastructure utilizing the SAGA standard. We adapt the DECIDE use case to provide us with an interpretation of the Cluster 1 characteristics from a cloud computing perspective.

<table>
<tr><th colspan="2">Catania Science Gateway – DECIDE - Diagnostic Enhancement of Confidence by an International Distributed Environment[12]</th></tr>
<tr><td colspan="2"><b>Description:</b> The DECIDE Grid-based e-Infrastructure relies on the Pan-European backbone GÉANT and the NRENs (National Research and Education Networks) and offers computing and storage resources and data-intensive processing tools. DECIDE is focused on supporting neurologists and physicians involved in the assessment of neurodegenerative diseases in the diagnosis and prognosis.
The DECIDE platform consists of three different layers: research networks, resources and domain-specific applications.
• Network connectivity, provided by the GÉANT backbone and the National Research and Education Networks, connects , brings different types of computing and storage resources.
• The Grid infrastructure (cloud infrastructure) used as a platform to enable collaboration among all partners, as a technological “glue” to harmonize and unify developments, and as an elastic pool of computing and storage resources where large volumes of data can be hosted and related analyses can be performed.
Four applications are provided by DECIDE: Neurological clinical image analysis; Analysis of Position emission tomography (PET) biomarkers in Neurological and Psychiatric Disorders; Subcortical segmentation of single-subject MRI brain images for hippocampal volume estimation; Detections of early symptoms of AD and distinguishing different forms of degenerative impairment</td></tr>
<tr><td colspan="2"><b>Goals and aspirations:</b> The use case demonstrates how a large number of resources from different institutions can be utilized as a computing platform for various applications.</td></tr>
<tr><td colspan="2"><b>TECHNICAL ASPECTS</b></td></tr>
<tr><td>Criteria for success</td><td>• Deployment of heterogeneous applications based on a homogenious resource management model.
• Unified data access for model data</td></tr>
</table>

4.1.3 Standards Profile Example

From the use cases described in the previous sections, the NIST characteristics

- On-demand self-service
- Massive scale
- Homogeneity

can now be underlined by the following technical requirements:

[12] http://applications.eu-decide.eu/

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

- **Homogeneous deployment and operation of heterogeneous applications**. A cloud infrastructure environment which provides on-demand self-service and massive scale is suitable to fulfil this requirement. Self-service is important for the configuration of infrastructure resources as computing and storage nodes, while massive scale is required to provide sufficient resources.
- **Homogeneous resource management and resource discovery**. Similar to the previous item, on-demand self-service functions allow applications to access and to manage required resources in a unified way. Since the use cases analysed in the previous section are based on incorporating a large variety of different applications, homogeneous resource management and resource discovery mechanisms are mandatory for job assignment and computation node configuration.
- **Data management and efficient data access**. All use cases are related to data intensive computations. Coordination and synchronization between computing steps is of minor importance, most applications perform data processing in batch mode. Hence, Persistent distributed storage and data management and efficient access is important transfer of large amount of data is a secondary concern.

Hence, a suitable combination of standards that is used as starting point for a standards profile development is the following:

- An IaaS based standard that provides advanced resource management functionalities. Candidates are OCCI and CIMI.
  The **Open Cloud Computing Interface** (OCCI) is a set of specifications delivered through the Open Grid Forum for cloud computing service providers. OCCI provides commonly understood semantics, syntax and a means of management in the domain of consumer-to-provider IaaS. It covers management of the entire life-cycle of OCCI-defined model entities and is compatible with existing standards such as the Open Virtualization Format (OVF) and the Cloud Data Management Interface (CDMI). It uses the Representational State Transfer (REST) approach for interacting with services.
  The **Cloud Infrastructure Management Interface** (CIMI) is an open standard API specification for managing cloud infrastructure. CIMI's goal is to enable users to manage cloud infrastructure in a simple way by standardizing interactions between cloud environments to achieve interoperable cloud infrastructure management between service providers and their consumers and developers.
- An IaaS based standard that deals with cloud storage, for instance the **Cloud Data Management Interface** (CDMI). CDMI is a SNIA standard that specifies a protocol for self-provisioning, administering and accessing cloud storage. CDMI defines RESTful HTTP operations for assessing the capabilities of the cloud storage system, allocating and accessing containers and objects, managing users and groups, implementing access control, attaching metadata, making arbitrary queries, using persistent queues, specifying retention intervals and holds for compliance purposes, using a logging facility, billing, moving data between cloud systems, and exporting data via other protocols such as iSCSI (Internet Small Computer System Interface) and NFS (Network File System).

Clearly, other combinations and additions are possible, depending on the specific requirements of the application use case. Appendix 2 (section 1) provides an in-depth by-clause analysis of CIMI as it would have to be exercised for each standard specification included in the profile, as a reference model for the profile development approach described in Section 3.5.

4.2 Cluster 2 – Trusted Public Clouds for Government

The Trusted Public Clouds for Government cluster is the largest cluster coming out of the cluster-forming methodology described earlier in the document, which is analysed by the CloudWATCH project. It has been selected among those available for further analysis as it demonstrates that such a methodology is not fool proof and may generate results that would benefit from human-assessed adjustment afterwards.

4.2.1 Cluster data quality assessment

Analysing the “heat-map” (see Figure 5) type of correlation between projects and importance of cloud characteristics, a number of observations need further explanation:

- **Weak agreement among projects.** On average, the entire cluster demonstrates weak agreement on assessing the importance of cloud characteristics as defined by NIST. Compared to other clusters, the agreement coefficients are low and barely cross the value 0.5.
- **Large deviations.** If the cluster cohesion values were low, weak agreement coefficients may be considered of some significance. However, a standard deviation larger than the coefficient itself indicates a result scattering beyond the average, meaning that it is not valuable beyond white noise of statistical data evaluation.
- **Broad spectrum of (weak) importance.** Across all NIST cloud characteristics, there are few occasions of agreement between sub-clusters, but none across the entire cluster.

Cluster 2	On Demand Self-Service	Broad Network Access	Resource Pooling	Rapid Elasticity	Measured Service	Massive Scale	Homogeneity	Virtualization	Low Cost Software	Resilient Computing	Geographic Distribution	Service Orientation	Advanced Security
INPUT	-0.337	-2.057	-0.166	0.715	-0.709	0.396	-0.911	0.516	-1.165	0.116	1.662	-1.624	-1.361
STORM CLOUDS	-0.847	0.683	1.452	1.374	1.688	-0.268	-0.242	-0.694	0.453	0.293	0.448	0.522	0.748
Texel	-1.446	0.441	0.694	0.073	1.346	-1.256	-0.198	-0.886	-0.184	0.509	-0.468	1.069	1.566
GEMMA	-1.122	0.367	1.241	0.454	1.087	-1.070	0.356	-0.179	-0.601	0.758	0.046	0.508	1.648
CloudWave	0.221	1.556	0.788	0.911	1.652	0.558	0.240	-0.443	1.437	0.275	-0.095	1.329	0.492
CELAR	-0.259	-1.080	0.338	0.765	0.130	0.329	-0.076	0.816	-0.938	0.897	1.414	-0.571	-0.217
S-CASE	0.376	1.164	-0.604	-1.204	-0.296	-0.417	0.357	-0.610	0.817	-0.711	-1.686	0.820	0.476
U-QASAR	-1.662	-0.271	0.332	-0.163	0.368	-1.599	-1.043	-1.562	-0.322	-0.718	-0.698	-0.087	0.619
COMPOSE	-0.429	-0.066	-0.623	-0.590	0.206	-0.347	-0.386	-0.487	0.087	0.055	-0.512	0.661	0.200
BETaaS	-0.727	-0.365	0.234	0.114	0.269	-0.539	-0.303	-0.265	-0.449	0.169	0.137	-0.040	0.350
AC	-0.623	0.037	0.368	0.245	0.574	-0.421	-0.221	-0.379	-0.086	0.164	0.025	0.259	0.452
CC	0.669	1.063	0.703	0.765	0.825	0.726	0.481	0.675	0.803	0.537	0.989	0.873	0.856
SNR	0.931	0.035	0.524	0.320	0.695	0.581	0.458	0.562	0.108	0.306	0.025	0.296	0.528

Figure 5. "Heat map" correlation of EU projects and cloud characteristics importance for Cluster 2.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

A visual inspection of the heat map however shows that if Cluster 2 were constrained to a subset of its original members, we might observe stronger cohesion in characteristics importance with significantly lower deviation from the average value among the chosen projects. Out of many different such arrangements, one stands out in demonstrating high agreement coefficients with low standard deviations as shown in Figure 6.

Cluster⍰2⍰ (revised)	On⍰Demand⍰ Self-Service	Broad⍰Network⍰ Access	Resource⍰ Pooling	Rapid⍰Elasticity	Measured⍰ Service	Massive⍰Scale	Homogeneity	Virtualization	Low⍰Cost⍰ Software	Resilient⍰ Computing	Geographic⍰ Distribution	Service⍰ Orientation	Advanced⍰ Security
STORM⍰CLOUDS	-0.847	0.683	1.452	1.374	1.688	-0.268	-0.242	-0.694	0.453	0.293	0.448	0.522	0.748
Texel	-1.446	0.441	0.694	0.073	1.346	-1.256	-0.198	-0.886	-0.184	0.509	-0.468	1.069	1.566
GEMMA	-1.122	0.367	1.241	0.454	1.087	-1.070	0.356	-0.179	-0.601	0.758	0.046	0.508	1.648
CloudWave	0.221	1.556	0.788	0.911	1.652	0.558	0.240	-0.443	1.437	0.275	-0.095	1.329	0.492
AC	-0.798	0.762	1.044	0.703	1.443	-0.509	0.039	-0.550	0.276	0.459	-0.017	0.857	1.114
CC	0.723	0.546	0.362	0.564	0.283	0.830	0.304	0.307	0.887	0.226	0.378	0.409	0.580
SNR	1.105	1.394	2.883	1.248	5.106	0.613	0.128	1.793	0.311	2.028	0.046	2.095	1.920

Figure 6. A reduced Cluster 2 showing improved agreement coefficient and cluster cohesion.

Despite not ideal, this cluster reconfiguration shows significant benefits over the original selection:

- **Stronger agreement coefficient.** The remaining projects in this new cluster arrangement show stronger agreement and overlap in importance of cloud characteristics, in both extremes, i.e. from being very important to not important at all. The cluster shows more agreement about which characteristics are more important that about those that are not important – a factor we consider a success criterion for further supporting clustered projects in the endeavour to arrive at a common standards profile supported across the cluster.
- **Less spread in characteristic valuation.** The less important a cloud characteristic is considered, the larger the spread (i.e. standard deviation) in assigning a value to it across the participating projects tends to get. As an illustration, compare the cohesion coefficient and the standard deviation for "Measured Service" (1.443 / 0.283) as the most important cloud characteristic for this cluster with the least important characteristic, "On Demand Self-Service" (-0.798, 0.723): The most important characteristic shows the smallest standard deviation, while the least important characteristic shows the second-highest standard deviation from the average. Given the small sample size, one cannot speak of solid statistical analysis. However, we attribute this correlation to the human tendency to pay more attention to what is important rather than to what is considered not or less important, in arriving at a more accurate depiction of importance.
- **Cluster characteristics remain stable.** Determining the top three and bottom three characteristics for the original Cluster 2, the rankings would be as provided in Table 2. Reducing the cluster to the suggested projects as indicated in Figure 6 merely changes the ranking of the least important characteristics, while significantly increasing the agreement coefficient among the top 3 cloud characteristics.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Top 3			
Characteristic	AC	CC	SNR
Measured Service	0.574	0.825	0.695
Advanced Security	0.452	0.856	0.528
Resource Pooling	0.368	0.703	0.524

Bottom 3			
Characteristic	AC	CC	SNR
On-demand Self Service	-0.623	0.669	0.931
Massive Scale	-0.421	0.727	0.695
Virtualisation	-0.379	0.675	0.562

Table 2. Top 3 and bottom 3 cloud characteristics in Cluster 2.

Reducing the cluster to the suggested projects as indicated in Figure 6 merely changes the ranking of the least important characteristics, while significantly increasing the agreement coefficient among the top 3 cloud characteristics:

Top 3			
Characteristic	AC	CC	SNR
Measured Service	1.443	0.283	5.106
Advanced Security	1.114	0.580	1.920
Resource Pooling	1.044	0.362	2.883

Bottom 3			
Characteristic	AC	CC	SNR
On-demand Self Service	-0.798	0.723	1.105
Massive Scale	-0.550	0.307	0.613
Virtualisation	-0.509	0.830	1.793

Table 3: Top 3 & bottom three cloud characteristics in a revised Cluster 2.

Comparing the original and the new allocation in the discussed cluster (“Trusted Public Clouds for Government”) with the other clusters discussed in this deliverable, all clusters maintain their unique combination of most important, and respectively least important, cloud characteristics. This makes it difficult if not impossible to merge clusters to increase traction and setup for success in developing profiles on Cloud standards. Curiously, while three projects in the Trusted Public Clouds for Government cluster (STORM Clouds, Texel, Gemma) are specifically targeting public administration, other projects addressing the same sector (e.g. Umea, Leicester (City Council) are allocated to different clusters, focussing on different Cloud characteristics, which are influenced by the actual challenges being tackled in the respective projects.

Nonetheless, we determine that the reduced cluster provides better focus on the task at hand than the original assignment.

4.2.2 Reviewing cluster projects use cases

The pruned Cluster 2 comprises of the following projects:

- STORM CLOUDS
- Texel
- Gemma
- CloudWave.

In short, while the CloudWave project roots in the agile application engineering domain, the remaining three projects concern themselves with governmental services in the cloud.

The following is an abbreviated analysis of the projects and activities (not all are EC-funded projects) using the project use cases cards developed earlier in the project.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

4.2.2.1 Storm Clouds

STORM CLOUDS – Surfing Towards The Opportunity Of Real Migration To Cloud-Based Public Services[13]	
Description: STORM CLOUDS aims at deeply exploring how the needed shift by Public Authorities to a cloud-based paradigm in service provisioning should be addressed, mainly from the point of view of the end-users, and taking full advantage of edge ICT.	
Goals and aspirations: The purpose of STORM CLOUDS is to define useful guidelines on how to address the process in order to accelerate it, for Public Authorities and policy makers. These guidelines will be prepared based on direct experimentation in 4 European cities, Águeda (PT), Manchester (UK), Valladolid (ES) and Thessaloniki (GR) creating a set of relevant use cases and best practices.	
LEGAL ASPECTS	
Legal domain	Public administration
Legal frameworks & laws	Data protection requirements (Deliverable D4.1.1)
Compliance criteria	EU Data protection directive (95/46/EC)
ORGANISATIONAL ASPECTS	
Organisation domain	Best practices for Cloud adoption for public administration across IaaS & PaaS
Organisational procedures	Provide baseline technical architecture and guidelines for public administration to migrate to using and/or providing Cloud services
Compliance criteria	Local, national and European law binds public administration.
TECHNICAL ASPECTS	
Preconditions	Technical architecture is chosen in the project, while legal requirements are strict preconditions to the project.
Criteria for success	Pilot deployments per participating municipality are accepted for sustenance beyond project

4.2.2.2 Texel

TEXEL – Smart energy services for the future
Description: Texel is a Dutch island in the North Sea with a population of about 14,000 and total area of over 460 square kilometres. The Municipality of Texel, Gemeente Texel, has decided to achieve 'energy neutrality' by 2020. Simply put, the island is aiming to cut the energy cable connected to the mainland. To reach this ambitious goal, it has joined forces with Capgemini and TexelEnergie, the local energy company, in a Smart Energy Program. The program's objectives are to reduce energy consumption, match demand and supply more effectively, and increase the use of renewable energy. The smart energy program aims to create projects that help individuals save energy within their own homes and also reduce public energy consumption.
Goals and aspirations: Texel wanted to become energy neutral by 2020, by eliminating the need to get energy from the Netherlands mainland. CapGemini created a cloud-based Home Energy Management System, used by Cloud Power communities to collaborate and coordinate the use of energy by

[13] http://stormclouds.eu/

households. In addition, deployed smart lighting grid to remotely view and control public street lights. By 2015, the city will reduce 37% of energy use by utilizing intelligent switching and dimming, the energy consumption by matching demand and supply, and implementing tactics like replacing old light bulbs.	
LEGAL ASPECTS	
Legal domain	Public administration & services
Legal frameworks & laws	n/a
Compliance criteria	n/a
TECHNICAL ASPECTS	
Services architecture and model	CloudPower is a cloudified energy creation, storage and load control system, using smart meters at households, a Home Energy Management System optimising use of appliances, and a Central Energy Management System that ensures shortfalls and oversupply of locally produced energy are alleviated. Smart Public Lighting: Sensor networks-based system for controlling and optimising use of public street lighting through the Internet Both systems are designed as SaaS towards the consumer through web portals. Internally, at least some of the services use an IaaS provider (MS Azure).

4.2.2.3 Gemma

GEMMA – Global Emergency Management[14]	
Description: Atos Global Emergency Management (GEMMA), the solution for emergency management, helps first responders deliver on their commitment to protect citizens and keep society safe. GEMMA is the end-to-end solution for emergency management that optimizes resources, reduces response times and, most importantly, saves lives.	
LEGAL ASPECTS	
Legal domain	Public administrative services
Compliance criteria	EU Data protection directive (95/46/EC)
ORGANISATIONAL ASPECTS	
Organisation domain	Emergency response services
Organisational procedures	Tight collaboration between call operator (triage), response unit, and resource (e.g. helicopter). Handles emergency incident patient data, as well as patient history records.
Compliance criteria	EU Data protection directive (95/46/EC)
TECHNICAL ASPECTS	

[14] http://atos.net/content/dam/global/documents/your-business/atos-emergency-management-whitepaper.pdf

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Preconditions	Live and constant link of remote mobile response units via tablet to the emergency coordination centre (SUMMA 112) and target hospital's A/E unit.
Criteria for success	Very tight integration of many different comunication equipment and services, such as data, telephony, mobile broadband, wireless LAN, voice recording, GPS and more.
Service architecture & model	Primary goals are addressing command & control, decision support and communications. Its foundation is very similar to typical call-centre environments, but it is extended with real-time communications with remote mobile units, exact geo-location and prediction (accuracy in announcing time to hospital for A&E staff to prepare). Tight integration requirements almost mandate an end-to-end solution with as few external interfaces as possible. Closest service model is that of SaaS.

4.2.2.4 CloudWave

CLOUDWAVE – Agile Service Engineering for the Future Internet[15]
Description: CloudWave is an EU-funded research project that is enabling a next generation of cloud infrastructure operations and agile development for their hosted applications. Our approach dynamically adapts cloud services to their environment, resulting in improved service quality and optimized resource use. This is supported with an enhanced cloud monitoring that provides holistic analytics of IaaS and SaaS layer services running on the cloud, leading to CloudWave's innovative, automated adaptation of the infrastructure and application, as well as enabling DevOps-like data and interfaces for the developer.
Goals and aspirations: CloudWave empowers cloud infrastructure providers (IaaS) and their hosted applications developers (SaaS) to transparently collaborate to obtain high levels of service at lower costs. Execution Analytics: CloudWave improves existing cloud monitoring solutions with a more holistic and efficient approach towards IaaS and SaaS services. Unified monitoring consolidates infrastructure vs. application data, as well as virtual infrastructure vs. physical hardware. Programmable monitoring enhances filtering and delivery of data for analysis, allowing for better management of cloud-based resources. Coordinated Adaption: CloudWave enables reconfiguration of the infrastructure and application in real-time to compensate for a variety of performance factors, resulting in an increasingly resilient, automated and optimized cloud deployment. Feedback-driven Development: CloudWave advances current DevOps solutions with developer-oriented data based on its innovative monitoring. It mixes automation (coordinated adaptation) with customizable feedback for improved agile development, resulting in quicker time-to-market, shortened maintenance cycles and more reliable cloud applications for their end-user customers.

[15] http://cloudwave-fp7.eu/

4.2.3 A straw-man cloud standards profile

Reviewing the outcome of section 4.2.1 the three most important cloud characteristics are (again, "E" denoting an essential characteristic):

- [E] Measured Service
- Advanced Security
- [E] Resource Pooling

And the least important are:

- [E] On demand self-service
- Virtualisation
- Massive scale

Reflecting these on the project goals and ambitions, this makes sense: At least two (Texel, STORM CLOUDS) if not three (incl. CloudWave) of the four projects operate under the assumption to deploy services on shared, public clouds, hence resource pooling is within their scope.

All three public administration/government projects deal with data that is sensitive or at least personal, with patient medical records (GEMMA) being the most personal and most sensitive type of information of all. Advanced security measures are without doubt a much-needed capability of the deployed solutions.

"Measured Service" ranks highest in importance for this cluster. Different aspects of "Measured Service" are ranked differently by the projects. The Texel project is looking for reliable, real-time gathering and processing of energy consumption, production and storage. However, this type of service measurement is out of scope of the NIST definition of that characteristic which scopes measured service around the consumption of the offered service. Were the service offered in Texel, then it would surely fit the NIST definition. but the Texel project description does not offer any indication into that direction, nor does it offer information on service management and consumption models.

Likewise, the GEMMA project is most likely implementing a "managed service" business and revenue model, where a custom service is developed and subsequently managed by a third party (e.g. CapGemini or Atos, respectively).

Having given "Measured Service" the individually highest ranking (1.688 and 1.652, respectively) within this cluster, STORM CLOUDS and CloudWave projects are operating under a more classic cloud computing model, in that the measuring of the outsourced service as precisely as necessary, plays a role in the overall business model. STORM CLOUDS aims at deploying public administration services into the Cloud, where the PA entity operates the service for its customers, the citizens, on the infrastructure of an underpinning supplier. Similarly, CLoudWave aims to optimise Cloud resource consumption and usage through explicitly integrating cloud monitoring into their portfolio.

This document can merely provide an indication of which direction in potential standards profiling it recommends. The actual work would have to be done by the cluster projects themselves.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Measured service. Closed systems such as GEMMA or Texel enjoy the benefit of controlling all interfaces internal to the solution, giving them free hand in their choices. For these projects, the measured service is most likely related to either internal components (Texel) or relates more Service Level Management within general IT Service Management (ITSM), for which sufficient solutions already exist.

Other systems such as those developed by STORM CLOUDS and CloudWave, scalable and large-scale solutions exist and are close to CloudWave's solution candidate Nagios. Decoupled or loosely integrated systems (unlike e.g. GEMMA) may look into standardised information models for resource usage, particularly on the IaaS layer, and message-payload oriented specifications standardising on aspects of delivering service metrication data, even for real-time applications.

Existing standards supporting **Measured Service**:

- **Usage Record 2**[16]. The Usage Record specification from the Open Grid Forum defines a comprehensive list of resources and their metrication means. It is extensively used in large worldwide scientific collaborations such as the European Grid Infrastructure (EGI), the Worldwide LHC Computing Grid (WLCG) which also uses resources of EGI, the Open Science Grid (OSG), and XSEDE in the US.
- **NIST Special Publication 500-307**[17]**.** SP 500-307 defines a *model* for the development and definition of Cloud service metrics for a number of well-defined use cases. SP 500-307 classifies metrics following three typical service lifecycle phases: Service Selection, Service Agreement, and Service Measurement. Many more measurement scenarios exist, but are out of scope of NIST SP 500-703, or do not follow its metric modelling framework.
- **DMTF Cloud infrastructure Management Interface (CIMI) Model and RESTful HTTP-based Protocol**[18]**.** The CIMI specification defines a number of metrics for cloud services implementing the IaaS model using the CIMI management interface. It is bound to the underpinning DMTF Common Information Model (CIM)[19] specification.
- **AMQP (Advanced Message Queuing Protocol)**[20]**.** AMQP 1.0 is an OASIS standard since 2012, and approved by the International Standards Organisation (ISO) as ISO/IEC 19464. It provides reliable messaging (from fire-and-forget, to exactly once delivery), cross-platform portable data representation, flexible deployments (peer-to-peer, client-broker, broker-broker networks) and is entirely broker-independent (i.e. allowing heterogeneous and inter-provider deployments). It has a strong industry backing including two major Cloud service providers (Microsoft, VMWare). Even though AMQP does not define any metrics by itself and therefore can be argued as not applicable in this section, it is described nonetheless in this context, since at least two of the three fundamental metric scopes defined in NIST SP 500-307 require a metric measurement delivery model (i.e. Service Agreement and Service Measurement) – we consider measurement

[16] https://www.ogf.org/documents/GFD.98.pdf
[17] http://www.nist.gov/itl/cloud/upload/RATAX-CloudServiceMetricsDescription-DRAFT-20141111.pdf
[18] http://dmtf.org/sites/default/files/standards/documents/DSP0263_1.0.1.pdf
[19] http://www.dmtf.org/standards/cim
[20] https://www.amqp.org/

infrastructures as enabling technology to deliver cloud services regardless the service delivery model (IaaS, PaaS, SaaS).

Advanced Security. Advanced security is a horizontal characteristic, most effective when present on all components of the examined solution, and hence manifests in all other cloud characteristics. At the date of writing, there are no security measures or technology available that were designed especially for cloud computing. General-purpose specifications in the area of information security management exist, alongside a number of cloud-specific guidelines and security control collections:

- **ISO/IEC 27000[21] family.** Often called ISO 27k, this family of specifications defines general-purpose security related vocabulary and controls for information security management systems (ISMS). Of particular interest is ISO/IEC 27018 "Code of practice for data protection controls for public cloud computing services" (e.g. for STORM CLOUD in particular).
- **NIST Special Publication 800-53[22].** SP 800-53 is a collection of "Security and Privacy Controls for Federal Information Systems and Organizations". While the title clearly indicates its scope towards US federal government, it also applies to other public administration and governmental IT systems since it provides a comprehensive list of controls and procedures of which a subset may be selected for implementation by Texel, GEMMA, and STORM CLOUD in particular.
- **CSA CCM 3.01[23].** Similarly, the Cloud Security Alliance (CSA) Cloud Control Matrix (CCM) provides a comprehensive list of more than 130 controls for advanced security measures. The CCM maps controls to cloud architecture subsystems, cloud service models, and most importantly, already existing other international security controls such as the two referenced above.

Resource Pooling. As discussed earlier (section 3.2) resource pooling is considered a non-functional cloud characteristic, and as such an internal function of a cloud service not requiring interoperability across providers or in a consumer/provider relationship.

4.3 Cluster 3 – High-performance, dedicated purpose applications

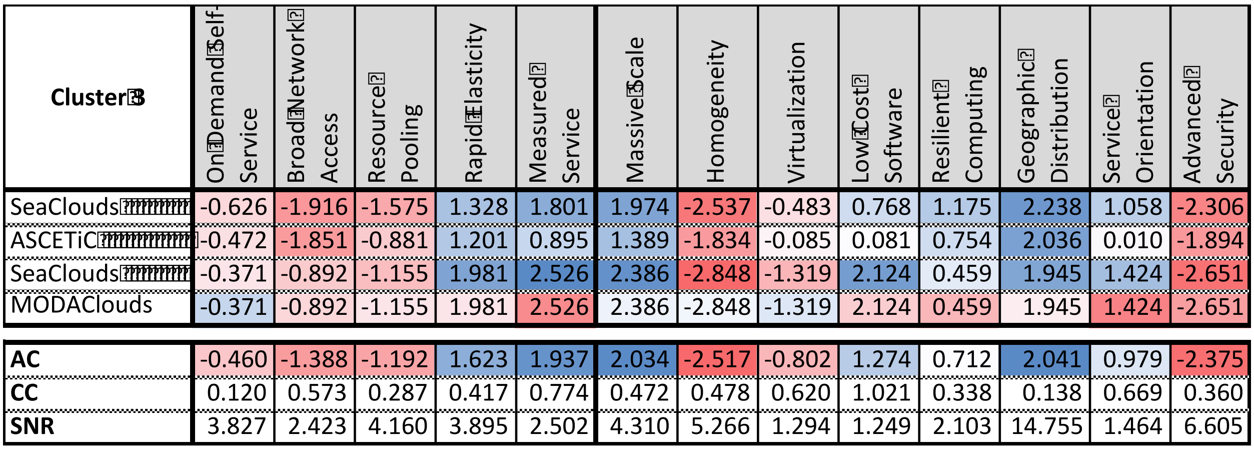
The high-performance cluster is relatively small with just three projects, and shows a superficial similarity to Cluster 1 in that both clusters rate Massive Scale highly, yet are direct opposites in their rating of Homogeneity. It is interesting to note here that the SeaClouds project was scored independently by two different members of the project and that the resulting positions in the cluster tree are remarkably consistent.

4.3.1 Cluster data quality assessment

Inspection of the "heat-map" for Cluster 3 is considerably simpler than for Cluster 2 presented above.

[21] http://www.iso.org/iso/catalogue_detail?csnumber=63411
[22] http://csrc.nist.gov/publications/drafts/800-53-rev4/sp800-53-rev4-ipd.pdf
[23] https://cloudsecurityalliance.org/research/ccm/

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994



Cluster 3	On Demand Self-Service	Broad Network Access	Resource Pooling	Rapid Elasticity	Measured Service	Massive Scale	Homogeneity	Virtualization	Low Cost Software	Resilient Computing	Geographic Distribution	Service Orientation	Advanced Security
SeaClouds	-0.626	-1.916	-1.575	1.328	1.801	1.974	-2.537	-0.483	0.768	1.175	2.238	1.058	-2.306
ASCETiC	-0.472	-1.851	-0.881	1.201	0.895	1.389	-1.834	-0.085	0.081	0.754	2.036	0.010	-1.894
SeaClouds	-0.371	-0.892	-1.155	1.981	2.526	2.386	-2.848	-1.319	2.124	0.459	1.945	1.424	-2.651
MODAClouds	-0.371	-0.892	-1.155	1.981	2.526	2.386	-2.848	-1.319	2.124	0.459	1.945	1.424	-2.651
AC	-0.460	-1.388	-1.192	1.623	1.937	2.034	-2.517	-0.802	1.274	0.712	2.041	0.979	-2.375
CC	0.120	0.573	0.287	0.417	0.774	0.472	0.478	0.620	1.021	0.338	0.138	0.669	0.360
SNR	3.827	2.423	4.160	3.895	2.502	4.310	5.266	1.294	1.249	2.103	14.755	1.464	6.605

Figure 7. "Heat map" correlation of projects to cloud characteristics in Cluster 3.

Here there is broad agreement across characteristics, except perhaps for Low Cost Software with a high standard deviation (Std. Dev.) and the lowest signal to noise ratio (SNR).

4.3.2 Reviewing cluster projects use cases

The use case cards presented below have been compiled with a focus on gathering just that information related to service model and cloud characteristics addressed by the project.

ASCETIC - Adapting Service Lifecycle Towards Efficient Clouds[24]		
Description	ASCETiC address the issue of energy efficient computing, specifically in the design, construction, deployment and operation of Cloud services.	
Goals and aspirations	The project has already identified a number of toolkit components that will lead to direct exploitation, e.g. SaaS modelling tool, energy modeller, Virtual Machine Contextualizer. The exploitation of the entire ASCETiC cloud architecture and its reference implementation (SaaS, PaaS, IaaS) is envisaged through the creation of a spin-off company by the end of the project.	
Legal aspects	Legal domain	N/A
	Legal frameworks, laws, etc., to be taken into account	N/A
	Compliance criteria	**N/A**

[24] http://www.ascetic.eu/

MODAClouds - Model-Driven Approach for design and execution of applications on multiple Clouds[25]	
Description	MODAClouds provides methods, a decision support system, an open source Integration Development Environment and a run-time environment for the high-level design, early proptotyping, semi-automatic code generation, and automatic deployment of applications on Multi-Clouds, with guaranteed quality of service.
Goals and aspirations	The main goal of MODAClouds is to provide methods, a decision support system, an open source IDE and run-time environment for the high-level design, early prototyping, semi-automatic code generation, and automatic deployment of applications on multi-Clouds with guaranteed QoS. Model-driven development combined with novel model-driven risk analysis and quality prediction will enable developers to specify Cloud-provider independent models enriched with quality parameters, implement these, perform quality prediction, monitor applications at run-time and optimize them based on the feedback, thus filling the gap between design and run-time. Additionally, MODAClouds provides techniques for data mapping and synchronization among multiple Clouds.

SeaClouds - Seamless adaptive multi-cloud management of service-based applications[26]	
Description	SeaClouds directly impacts on the way developers are going to build cloud apps without worrying about underlying execution of different PaaS or IaaS providers, relying on its service orchestration capabilities based on informed election among providers.
Goals and aspirations	SeaClopuds main outcome is the implementation of a novel platform which performs a seamless adaptive multi-cloud management of service-based applications.
Existing specifications to rely on	CAMP, TOSCA

These use-case cards are instructive in actually being quite non-committal with regards to specific service models or cloud characteristics. These projects intend to put in place facilitating systems and mechanisms for other applications. These systems aspire to be non-restrictive, whereas the consuming applications may well be very much more specific.

In this case it turns out that the CloudWATCH clustering analysis that identified these projects as belonging to a tightly formed cluster, with similarities and differences clearly identifiable in relation to Cluster 1, proves more useful than the use-case capturing method.

[25] http://www.modaclouds.eu/
[26] http://www.seaclouds-project.eu/

4.3.3 An indicative cloud standards profile model

Reviewing the outcome of 4.3.1 above, we list the four most, and least important of the NIST characteristics. As before, essential characteristics are designated by [E].

Most important characteristics:

1. Geographic Distribution
2. Massive Scale
3. [E] Measured Service
4. [E] Rapid Elasticity

Least important characteristics:

1. Homogeneity
2. Advanced Security
3. [E] Broad Network Access
4. [E] Resource Pooling

"Geographic Distribution" being the highest ranked important characteristic, also has the highest SNR, whereas "Homogeneity" is indicated as the least important characteristic, with moderate to high SNR. This dipole sets Cluster 3 apart from Cluster 1 and gives the key opportunity for fine-tuning a standards profile for these application that aspire to provide management services for high-performance, dedicated purpose applications.

4.3.4 A straw-man cloud standards profile

Even though projects grouped into this cluster are remarkably consistent in their assessment of cloud characteristics, the needs and hence the application and realisation within the projects is very diverse. Moreover, this cluster requires cloud services to provide massive scale and rapid elasticity, both are essential Cloud characteristics, but are non-functional in that the indicated behaviour is expressed in service deployment, management and automation. These are all service operation functions that are all beyond the reach and influence of the service customer. The remaining characteristics are thus Geographic Distribution and Measured Service.

Geographic Distribution. Although common, Geographic Distribution is not an essential characteristic. It is typically associated and even conflated with large-scale cloud services (c.f. Massive Scale characteristics). Geographic Distribution is part of a service deployment architecture to satisfy a number of diverse requirements, such as to enable disaster recovery, provide the same service for different (and usually incompatible) jurisdictions and legal domains, to save energy by leveraging temperature differences between day and night for data centre cooling purposes, and many other. Some of these requirements are service-operation-related and entirely out of scope of this document, while others (such as disaster recovery, and legal jurisdiction) are in scope for service consumers. Typically, a cloud service offering geographic distribution allows the consumer to control the geographic location in a coarse-grained manner. Often these are called cells, regions, or zones. There is no common terminology between cloud providers.

There is no standard specification known to us that would formalise the language, information model and use of such features. It is, however, implemented in non-interoperable and proprietary interfaces (de-facto or industry standard) such as AWS EC2, Google Cloud and Azure, and many others. However, since OCCI is designed as an extremely versatile and extensible specification, the authors anticipate that a formalisation of geographic distribution configuration may be accomplishable with reasonable effort.

- **Open Cloud Computing Interface 1.2.** Although currently in public comment, OCCI 1.2 is considered stable with negligible changes to the specification itself, once all comments have been considered in a published set of documents. OCCI defines a very powerful mix-in concept that can be used to define almost any type of additional features, characteristics, and components of any cloud computing service. Mix-ins can be associated to service instances, and offer the capability of mix-in specific mutable and immutable attributes. Providing a formalised language for a geographic distribution, mix-in allows cloud service providers to offer a wide variety of geographic distribution mechanisms. By associating/attaching one or more of these to an existing cloud service instance, the service consumer instructs the service provider to make the necessary provisions. For example, attaching a “Europe” region to a compute instance would cause an underlying VM (presuming that virtualisation is used) to be provisioned within the European data centre(s). Overriding it with a “North-American” mix-in reference would then cause the same VM instance to migrate to a North American data centre of the same cloud service provider.

Measured Service. This cluster shares the Measured Service characteristic with cluster 2, which looks at trusted cloud services for governments. The same analysis and suggestions for standards apply here as well. Therefore, we repeat in brief the suggested standards to look at, and refer the reader to section **Errore. L'origine riferimento non è stata trovata.** for further information.

- **Usage Record 2**.
- **NIST Special Publication 500-307**.
- **DMTF CIMI.**
- **AMQP (Advanced Message Queuing Protocol)**.

4.4 Discussion: Clusters as a Basis for Standards Development

The results presented in the previous subsections demonstrate value in the chosen methodology for clustering projects and cloud activities into groups with similar or matching importance of cloud characteristics. Nonetheless, a second, much less expected outcome of this work is to suggest that clusters may work together on standardising shared characteristics described as follows:

Typically, profiles on standard specifications mandate the use of a defined set of standards together, and define at the same time the changes and interpretations of incorporated standards within the profile. While this approach often makes sense in an uncoordinated landscape of collaboration, it may not be the best approach for the CloudWATCH undertaking.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

Since the clustering effort discussed in this document makes all participants aware of commonalities among participant projects in a cluster, an alternative approach of developing and defining standards may make much more sense. Instead of providing the profiling of a specific standard together with others into a single profile document targeted at one given cluster discussed in this document, it may be more successful to encourage clusters (and individual projects) working together on one single profile aligned with one single cloud characteristic in one single document. Once all individual profiles are finished, an identified cluster would simply have to write a very brief cluster profile document incorporating by reference any fitting individual profile documents.

However, such a synergetic approach is feasible and achievable only if the use cases of each project and cluster for the characteristic in question are sufficiently overlapping to arrive at a common solution. Otherwise, clusters would have to work on their own cluster-specific standards profile for a given cloud characteristic.

For example, consider the common cloud characteristic “Advanced Security”. Across individual projects, the majority considers it relatively important. However, projects in Cluster 2 consider it in the top 3 of their most important characteristics. Interestingly, Cluster 2 hosts governmental cloud activities: STORM CLOUDS, Texel, Gemma. None of these projects agree on the importance on any of the other cloud characteristics. The same is true of another cluster which was identified but not analysed further as part of this document. Here, two projects were governmental cloud activities: Varberg (Municipality Social Services Administration)[27], Leicester (City Council)[28].

As a result, those five projects might decide to work together in analysing how they achieve “advanced security” in their implementation. In case the implementations are very similar, such collaboration proves a real opportunity to arrive at a profile documentation specifying how interoperable advanced security would be achieved among projects in the cluster – but also across different clusters through the by-reference-incorporation mechanism described above.

Even other projects that are not part of any cluster considering advanced security as very important (such as Broker@Cloud, SUPERCLOUD, Embassy Cloud) may benefit from contributing to a common advanced security profile.

However, if the use cases are *not* compatible enough, the one large collaboration would have to split into smaller collaborations to narrow the scope and to increase the likelihood of common use cases and implementations – very similar to the narrowing exercise conducted for cluster 4 as such.

[27] https://customers.microsoft.com/Pages/Download.aspx?id=15146

[28] https://customers.microsoft.com/Pages/Download.aspx?id=3893

5 Future work and CloudWATCH legacy

Standardization is a time consuming process that exceeds the two-year time frame of collaboration and support actions such as CloudWATCH. Therefore, the original objective to provide a number of standards profile as input for relevant SDOs (Standards Development Organisation) could not be met. Instead, we have developed a methodology for the derivation of standards profiles based on self-assessment of cloud related projects (and by extension, cloud service providers and cloud service customers) based on 13 characteristics of cloud computing identified by NIST.

Therefore, standardization groups which are concerned with the identification of the usage of standards and the relationship between standards is more appropriate for legacy work than an SDO concerned with the development of a singular standard.

A group that meets this description is the IEEE P2301 Cloud Profiles Working Group[29]. This group aims to develop a **Guide for Cloud Portability and Interoperability Profiles (CPIP)**. The guide advises cloud computing ecosystem participants (cloud vendors, service providers, and users) of standards-based choices in areas such as application interfaces, portability interfaces, management interfaces, interoperability interfaces, file formats, and operation conventions. The guide groups these choices into multiple logical profiles, which are organized to address different cloud personalities.

In several conference calls between members of this group and the CloudWATCH consortium, the cluster approach described in this report, as well as the results obtained so far, have been identified as of major interest for IEEE P2301. Current discussion aims on a closer collaboration between both groups and a continuation of the activities that have led to this report under the auspices of IEEE P2301.

The second group undertaking a similar activity is the **ETSI Cloud Standards Coordination (CSC)**[30] group coordinated by the European Commission through unit DG CONNECT E2. Anticipating the group's draft outputs published for public comment in September, CloudWATCH sees great opportunity for synchronising and synergising the work between these two activities addressing standards conformance, interoperability and, eventually, profiling.

The CloudWATCH project has already secured support from the EC for a continuation activity in the CloudWATCH2[31] project. Within this activity one of the three key objectives is the clustering of relevant projects from within the European Cloud ecosystem. As part of this, specific activities will target bringing together recently funded by H2020 calls of relevance through concertation activities.

- **Clustering activities**: facilitating and furthering collaboration on key identified themes.
- **Cross-unit collaboration on re-use and wider uptake**: identifying common areas for the convergence of technologies, re-use and further development of existing software, services and experimental platforms ("develop and test").
- **Ecosystem analysis**: roles in the complete value chain, status of developments, sustainability strategies and business models.

[29] https://standards.ieee.org/develop/project/2301.html
[30] http://csc.etsi.org/
[31] CloudWATCH2 Think Cloud Services for Government, Business & Research – September 2015 - August 2017

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

- **Gap analysis**: technology advancements for future funding, identifying re-use and/or further developments of existing solutions.

Utilising the methodology that has developed through CloudWATCH will simplify the clustering of these new projects and in many cases support the identification of parties that should be contributing in collaboration for the Ecosystem and Gap analyses.

Furthermore, D4.4 Assessment of Cloud Profile interoperability testing will summarise the deployment and testing activities around technology implementing Cloud standards profiles. It will look at the testing methodologies used by the main demand streams and compare them with the formal interoperability testing exercise using the test cases coming from task 4.1

Within CloudWATCH2 there is also the continued long-term commitment to the need to support open standards as a method by which interoperability between different actors within the cloud ecosystem can easily interact. We have seen already that a number of the standards we have discussed within this document are already of interest to newly supported and identified projects and therefore will engage to further support both the profiling work through the methodology that we outlined in section 4.4. CloudWATCH2 will map the usage of standards by FP7 and H2020 projects and also identify necessary extensions. The cloud standards profiles provided as part of CloudWATCH will also be evolved and new clusters and profiles identified. CloudWATCH2 will also support and organise Cloud Plugfest events to promote the practical testing of profiles so that clusters develop.

6 Conclusions

The topics for deliverable D4.3 listed in the CloudWATCH Description of Work includes a review of the finished profile development activities against the initiated profile activities, and an assessment of the effectiveness of the best practices used. Lessons learned and recommendations for future profile activities beyond the project. During the work activities that have led to the results presented in this report, it became clear that the usefulness of a standard profiles (or a set of those profiles) is a moving target. Standards profiles are useful mainly in the context of a given set of use cases. A derivation of "absolute" profiles valid for a specific application domain (e.g., industry, public sector, and academia) proves to be a difficult task.

The work on property clusters however that has been started in the CloudWATCH work package 2 and has led to the definition of the three clusters described in this report has provided us with an alternative approach. Instead of insisting on a fixed set of standard profiles (which are likely to be too general to be useful), we have provided a methodology for the derivation of standards profiles that is now available for collaborating groups of projects (or contributors to a cloud service eco-system) to derive their own standard profiles. This methodology is based on the identification of the importance of 13 characteristics related to cloud computing by means of a ranking for each of these characteristics, and the application of a clustering procedure to the set of considered projects. Project clusters can now be underlined by illustrative use cases demonstrating the identified important characteristics, and thus lead to a better understanding of the requirements on the standards profile to be developed. These requirements lead (a) to the selection of standards as input for the profile, and (b) to the restrictions and extension of these standards that finally comprises the profile.

Although the methodology is neither fool proof, nor fully automatable, the construction of bi-plots and tentative clustering can be performed with a high degree of tool support. A web based prototype of such a tool is currently under development and is intended be published at the CloudWATCHHub.eu web site. Hence, scoring/ranking of characteristics can be performed as a community driven process, and will lead to a better understanding of the current "heat map" of the projects currently funded by the European Commission.

Further analysis of the data within each cluster is achieved using standard statistical methodology, but this is where automation and tool support most likely will end. Further work will have to involve close and intense human interaction, as well as heuristic information until one arrives at a strawman profile. Even then, a stawman profile is often not much more than a discussion paper that can turn out either entirely changed altogether when published as a profile specification, or not further pursued at all by stakeholders. This process from agreeing on joining forces together with the outspoken goal to arrive at a fully specified profile on standards requires the most human effort and interaction of all, and as experience shows [D4.4], typically beyond the scope of a project such as CloudWATCH.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

7 References

CAMP Organization for the Advancement of Structured Information Standards (OASIS), Cloud Application Management for Platforms (CAMP), v1.1, http://docs.oasis-open.org/camp/camp-spec/v1.1/camp-spec-v1.1.html

CDMI Storage Networking Industry Association (SNIA), Cloud Data Management Interface (CDMI), SNIA Technical Position, v.1.1.1, 2015, http://www.snia.org/sites/default/files/CDMI_Spec_v1.1.1.pdf

CIMI Distributed Management Task Force (DMTF), Cloud Infrastructure Management Interface (CIMI) Model and RESTful HTTP-based Protocol, DSP0263, 2012, http://dmtf.org/sites/default/files/standards/documents/DSP0263_1.0.0.pdf

D2.4 CloudWATCH deliverable D2.4 - Policy and compliance requirements Report (PM24) August 2015, To be published

NIST-800-145 The NIST Definition of Cloud Computing, SP 800-145, P. Mell, T. Grance, National Institute of Standards (NIST), 2011, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf

OCCI Open Grid Forum (OGF), Open Cloud Computing Interface (OCCI), three part specification, v1.1, http://occi-wg.org/about/specification/

OVF Distributed Management Task Force (DMTF), Open Virtualization Format (OVF), DSP0243, v 2.1.0, 2013, http://www.dmtf.org/sites/default/files/standards/documents/DSP0243_2.1.0.pdf

PCA Principal Component Analysis, second edition, 2002. I. T. Jolliffe, Springer, 2002.

RFC2119 Key words for use in RFCs to Indicate Requirement Levels, S. Bradner, Internet Engineering Task Force, 1997. https://www.ietf.org/rfc/rfc2119.txt

TOSCA Organization for the Advancement of Structured Information Standards (OASIS), Topology and Orchestration Specification for Cloud Applications (TOSCA), v1.0, 2013, http://docs.oasis-open.org/tosca/TOSCA/v1.0/os/TOSCA-v1.0-os.html

D4.4 CloudWATCH deliverable D4.4 - Assessment of Cloud Profile interoperability testing September 2015, To be published

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

8 Appendix 1 – Use case card

To collect relevant data in a manageable way, the CloudWATCH project uses use case cards to capture information about projects and their use case(s):

Use Case		
Title		
Description		Short summary of the use case
Goals and aspirations for the use case		Background and main message of the use case; context of the use case
Legal aspects	Legal domain	Data privacy regulations, licensing, contracting, etc.
	Legal frameworks, laws, etc., to be taken into account	Laws, policies, etc. which are of relevance
	Compliance criteria	Explanation why the use case is an illustration on how legal requirement can be implemented
Organizational aspects	Organization domain	E.g., security procedures, data privacy procedures, etc.
	Regulations and policies to be taken into account	Policies, standards, best practices to be taken into account
	Description of organization procedures	The "workflow" (or procedures) on organizational level used to achieve the goal of the use case
	Compliance criteria	Explanation why the use case is an illustration on how organisational requirement can be implemented
Technical aspects	Preconditions	Assumptions made prior to the execution of the use case
	Criteria for success	Expected process, outcome, side effects. Described by sequence charts,

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

		etc.
	Failure conditions and responses	Description of what can go wrong, and what to do about it.
Existing specifications to rely on		Specifications and standards already dealing with aspects related to the use case
New specifications required		Specifications and standards needed to establish the goals of the use case
Additional comments		**Add comments, remarks, suggestions, as you see fit**

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

9 Appendix 2 – "By-clause" analysis of the CIMI specification

Clause	Title	Text	Profile change or extension	Type	Supported characteristics	Rationale
4.3	OVF Support	Full Clause	None	NOTE	On-demand self-service, homogeneity, massive scale	Self-service related, important to support homogeneity (parallel deployment of various instances of the same VM)
5.2	Extendibility	The first [extendibility mechanism] allows for a CIMI Consumer to add additional data to a resource. Each resource in the CIMI model has an attribute called "properties." Consumers, when creating or updating a resource, may store any name/value pair in the "properties" attribute. CIMI Providers shall store and return these values to the Consumer. There is no obligation for the Provider to understand or take any action based on these values; they are there for the Consumer's convenience. Providers shall not add elements to this "properties" attribute.	None	NOTE		Mechanism can be used to distinguish between VM configurations representing different types of data processing nodes
5.11	Resource Metadata	Full Clause	None	NOTE	On-demand self-service, homogeneity	Enables unified resource management

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

5.11.2	Capabilities	Full Clause	None	NOTE	On-demand self-service, homogeneity, massive scale	Clause contains a list of capability URIs that indicate whether a specific resource supports an attribute or not. Profile changes with regard to mandatory attribute support are discussed in Clause 5.12 (see entries below)
5.11.2	Capabilities	CloudEntryPoint/ExpandParameter	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	CloudEntryPoint/FilterParameter	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	CloudEntryPoint/firstParameter	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	CloudEntryPoint/SelectParameter	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	System/SystemComponentTemplateByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/DefaultInitialState	Shall not be null	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/InitialStates	Shall not be empty	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/MachineConfigByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/MachineCredentialByValue	May be false	NOTE	Advanced security	Security related feature
5.11.2	Capabilities	Machine/MachineImageByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/MachineVolumeTemplatesByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/MachineStopForce	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management

5.11.2	Capabilities	Machine/MachineStopForceDefault	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/RestoreFromImage	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Machine/UserData	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Credential/CredentialTemplateByValue	May be false	NOTE	Advanced security	Security related feature
5.11.2	Capabilities	Volume/SharedVolumeSupport	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Volume/VolumeConfigByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Volume/VolumeImageByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Volume/VolumeSnapshot	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Volume/VolumeTemplateByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Network/NetworkConfigByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Network/NetworkTemplateByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	NetworkPort/NetworkPortConfigByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	NetworkPort/NetworkPortTemplateByValue	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	ForwardingGroup/MixedNetwork	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

5.11.2	Capabilities	Job/JobRetention	Shall be true	RESTRICTION	On-demand self-service, homogeneity, massive scale	Supports unified resource management
5.11.2	Capabilities	Meter/MeterConfigByValue	May be false	NOTE	Advanced security	Security related feature
5.11.2	Capabilities	Meter/MeterTemplateByValue	May be false	NOTE	Advanced security	Security related feature
5.11.2	Capabilities	EventLog/Linked	May be false	NOTE	Resilient Computing	Resilient related feature
5.12	Cloud Entry Point	Attribute: resourceMetadata	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports customer querying of resource metadata
5.12	Cloud Entry Point	Attribute: systems	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on systems (collections of machines, networks, etc.)
5.12	Cloud Entry Point	Attribute: systemTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on systems (collections of machines, networks, etc.)
5.12	Cloud Entry Point	Attribute: machines	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on machines
5.12	Cloud Entry Point	Attribute: machineTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on machines
5.12	Cloud Entry Point	Attribute: machineConfigs	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on machines
5.12	Cloud Entry Point	Attribute: machineImages	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on machines
5.12	Cloud Entry Point	Attribute: credentials	None	NOTE	Advanced security	Security related feature
5.12	Cloud Entry Point	Attribute: credentialTemplates	None	NOTE	Advanced security	Security related feature
5.12	Cloud Entry Point	Attribute: volumes	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on volumes
5.12	Cloud Entry Point	Attribute: volumeTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on volumes

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

5.12	Cloud Entry Point	Attribute: volumeConfigs	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on volumes
5.12	Cloud Entry Point	Attribute: volumeImages	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on volumes
5.12	Cloud Entry Point	Attribute: networks	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Atrribute: networkTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: networkConfigs	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: networkPorts	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: networkPortTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: networkPortConfigs	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: addresses	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: addressTemplates	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.12	Cloud Entry Point	Attribute: forwardingGroups	None	NOTE	Advanced security	Security related feature
5.12	Cloud Entry Point	Attribute: forwardingGroupTemplates	None	NOTE	Advanced security	Security related feature
5.12	Cloud Entry Point	Attribute: jobs	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on jobs
5.12	Cloud Entry Point	Attribute: meters	None	NOTE	Measured service	Measurement related feature

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

5.12	Cloud Entry Point	Attribute: meterTemplates	None	NOTE	Measured service	Measurement related feature
5.12	Cloud Entry Point	Attribute: meterConfigs	None	NOTE	Measured service	Measurement related feature
5.12	Cloud Entry Point	Attribute: eventLogs	None	NOTE	Resilient Computing	Resilient related feature
5.12	Cloud Entry Point	Attribute: eventLogTemplates	None	NOTE	Resilient Computing	Resilient related feature
5.13	System resources and relationships	Attribute: systems	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on subsystems
5.13	System resources and relationships	Attribute: machines	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on machines
5.13	System resources and relationships	Attribute: credentials	None	NOTE	Advanced security	Security related feature
5.13	System resources and relationships	Attribute: volumes	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on volumes
5.13	System resources and relationships	Attribute: networks	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on networks
5.13	System resources and relationships	Attribute: networkPorts	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on network ports
5.13	System resources and relationships	Attribute: addresses	Mandatory support for provider	RESTRICTION	On-demand self-service, homogeneity	Supports querying of information on network addresses
5.13	System resources and relationships	Attribute: forwardingGroups	None	NOTE	Advanced security	Security related feature
5.13	System resources and relationships	Attribute: meters	None	NOTE	Measured service	Measurement related feature
5.13	System resources and relationships	Attribute: eventLog	None	NOTE	Resilient Computing	Resilient related feature
5.13.1.1.1	SystemSystem Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on system information

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

5.13.1.1.2	SystemMachine Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on machine information
5.13.1.1.3	SystemCredential Collection	Full Clause	None	NOTE	Advanced security	Security related feature
5.13.1.1.4	SystemVolume Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on volume information
5.13.1.1.5	SystemNetwork Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on network information
5.13.1.1.6	SystemNetworkPort Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on network port information
5.13.1.1.7	SystemAddress Collection	Full Clause	None	NOTE	On-demand self-service, homogeneity	Supports querying on network address information
5.13.1.1.8	SystemForwardingGroup Collection	Full Clause	None	NOTE	Advanced security	Security related feature
5.13.1.1.9	SystemMeter Collection	Full Clause	None	NOTE	Resilient Computing	Resilient related feature

10 Appendix 3 – Developing standards profile for security

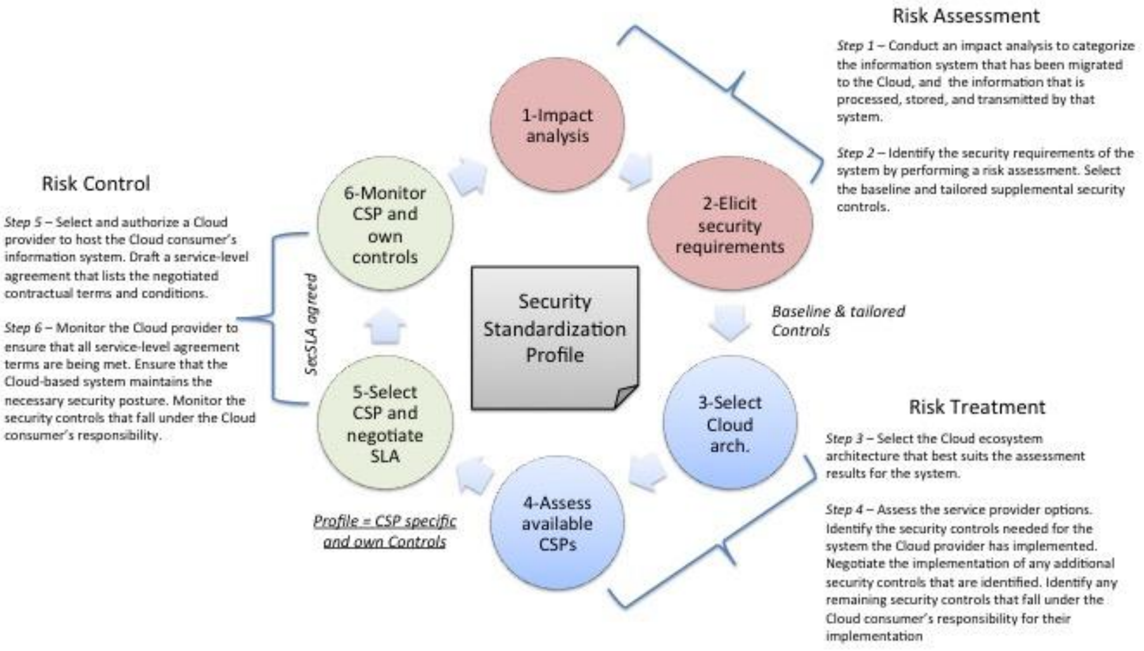
The methodological approach presented in Section 3 to develop standards profile has been validated in some real-world scenarios in Section 4. The steps needed to obtain a standards profile in the security field[32] need to be slightly adapted to fulfil the "good enough security" notion that is widely accepted in this community. The rest of this appendix will elaborate about the approach proposed by CSA to develop standards profile for security, which relies on the usage of risk management techniques capable of allowing users to get awareness about their specific security requirements thanks to an introspective analysis of their systems.

As state of practice, a commonly utilized approach by Cloud Service Providers (CSPs) has relied on the adoption of security certifications based on standardized "controls frameworks" (e.g., ISO/IEC 27002 or the upcoming 27017) to provide customers a reasonable degree of security assurance and transparency. Many CSPs are increasingly adopting Cloud-specific security controls frameworks such as the Cloud Security Alliance's Cloud Control Matrix (CSA CCM, www.cloudsecurityalliance.org/cm.html) and National Institute of Standards and Technology (NIST) Special Publication 800-53 Revision 4. Based on well-known standards, most of these security control frameworks allow for some degree of interoperability between CSPs. However, in order to provide Cloud assurance and transparency in the degree needed by customers, the actual use of security control frameworks has proven rather limited in practice. Over the implementation of their security controls framework, the CSP can only assume the type of data a customer will generate and use; the CSP is not aware of the additional security requirements or the tailored security controls deemed necessary to protect the customer's data. Therefore, in many cases the leveraged certification will either under-/over-provision the security level that is required by the customer which requires mechanisms (and standard profiles) that can enable them to understand and assess what "good-enough security" **Errore. L'origine riferimento non è stata trovata.** means, and especially the new challenges in risk assessment/management that the Cloud entails. This notion of "good enough security" is an empirical indicator related to the need of standard profiles in the security field.

In order to scope the discussion, this Appendix departs from the classical notion of risk management frameworks (RMF) advocated by relevant working groups at ISO/IEC, the European Commission, NIST, and the Cloud Security Alliance. Furthermore, we argue that a standard profile for security can be developed based on controls frameworks (like CSA CCM or NIST 800-53v4), which will in consequence aid customers to request/negotiate their security requirements with CSPs in the form of for example Service Level Agreements (SLAs).

The classical PDCA approaches (Plan-Do-Check-Act) are increasingly being considered by SMEs for assessing and managing their IT risk and security exposure following adoption of Cloud services. Consequently we explore, the synergies across risk management frameworks and standardization profiles as a means to achieve "good enough security" in the Cloud.

[32] The "advanced security" characteristic discussed in Sections 3 and 4.


1-Impact analysis
2-Elicit security requirements
Baseline & tailored Controls
3-Select Cloud arch.
4-Assess available CSPs
5-Select CSP and negotiate SLA
Profile = CSP specific and own Controls
SecSLA agreed
6-Monitor CSP and own controls
Security Standardization Profile

Risk Assessment

Step 1 – Conduct an impact analysis to categorize the information system that has been migrated to the Cloud, and the information that is processed, stored, and transmitted by that system.

Step 2 – Identify the security requirements of the system by performing a risk assessment. Select the baseline and tailored supplemental security controls.

Risk Treatment

Step 3 – Select the Cloud ecosystem architecture that best suits the assessment results for the system.

Step 4 – Assess the service provider options. Identify the security controls needed for the system the Cloud provider has implemented. Negotiate the implementation of any additional security controls that are identified. Identify any remaining security controls that fall under the Cloud consumer's responsibility for their implementation

Risk Control

Step 5 – Select and authorize a Cloud provider to host the Cloud consumer's information system. Draft a service-level agreement that lists the negotiated contractual terms and conditions.

Step 6 – Monitor the Cloud provider to ensure that all service-level agreement terms are being met. Ensure that the Cloud-based system maintains the necessary security posture. Monitor the security controls that fall under the Cloud consumer's responsibility.

Figure 8. Developing a security standardization profile from a risk management perspective.

Organizations targeting the development of a security standard profile as a means to implement good-enough security typically start with an introspective view that identifies both the assets to protect, and the (probabilistic) risks to consider when migrating to the Cloud (cf., NIST SP 800-30 and ENISA's report[33]). The NIST *Guide for Applying the Risk Management Framework to Federal Information Systems* (RMF) provides a structured process that integrates information security and risk management activities into the system development life cycle. The selected Cloud delivery model (public, private, hybrid, community) and the service type (SaaS, PaaS, IaaS), in association with security controls selected for the ecosystem, need to be chosen such that the system preserves its security requirements. Therefore, a systematic risk management cycle helps ensure that the residual risk is minimal, and that the deployed Cloud system achieves a security level that is at least equivalent to the one offered by an on-premise (non-Cloud) technology architecture or solution.

The key elements for the successful development of a security standard profile are the Cloud consumer understanding of the (a) Cloud-specific characteristics, (b) the architectural components for each Cloud service type and deployment model, (c) along with each Cloud actor's precise role in orchestrating a secure ecosystem. The Cloud customer's confidence in accepting the risk from using Cloud services depends on how much trust they place in the entities orchestrating the Cloud ecosystem. The risk management process ensures that issues are identified and mitigated early in the investment cycle and followed by periodic reviews. As Cloud customers and the other Cloud actors involved in securely

[33] Please refer to ENISA's report "Cloud Computing Benefits, risks and recommendations for information security."

orchestrating a Cloud ecosystem have varying degrees of control over Cloud-based IT resources, they need to share the responsibility of implementing and monitoring the security requirements.

Furthermore, it is essential for the Cloud consumers' business -critical processes to identify Cloud-specific risk-adjusted security controls. Cloud consumers need to leverage their contractual agreements to hold the Cloud providers (and Cloud brokers, when applicable) accountable for the implementation of the security controls. They also need to assess the correct implementation and continuously monitor all identified security controls. Draft NIST SP 800-173, *Cloud-Adapted Risk Management Framework (CRMF)*, is a key approach addressing the elements of a successful Cloud risk management strategy to enable the usage of standardization profiles for security. CRMF was first highlighted in NIST SP 500-299 as a cyclically executed process composed of a set of coordinated activities for overseeing and controlling risks. This set of activities consists of the following tasks:

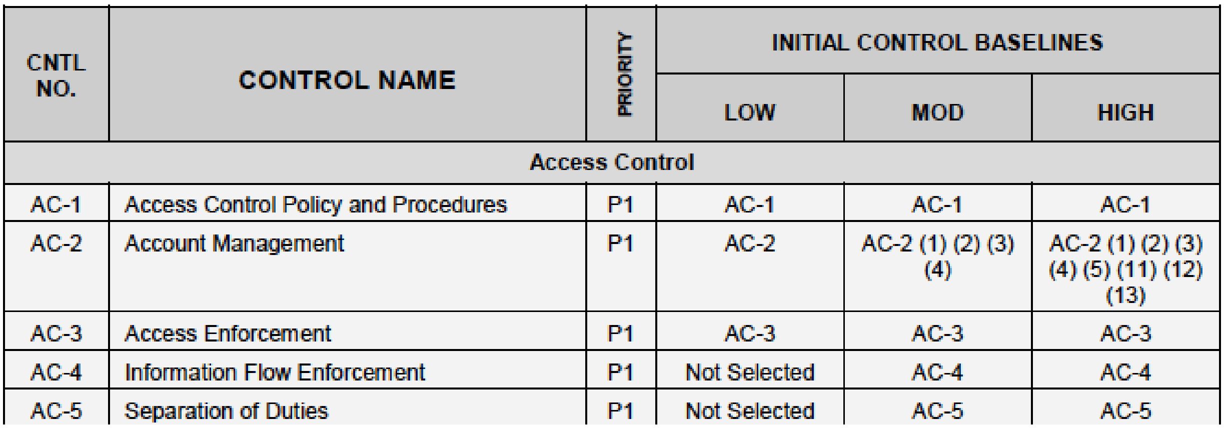
- Risk Assessment
- Risk Treatment
- Risk Control

These tasks collectively target the enhancement/customization of security through standardization profiles (mostly related to SLAs), which goes beyond the capabilities offered by widely used security control frameworks. CRMF provides a consumer-centric approach following the original RMF, identifying the six steps shown in Figure 8.

A risk-based approach to managing information systems is an holistic activity that should be integrated into every aspect of the organization, from planning and system development life cycle processes (Steps 1 – 2 in Figure 8) to security controls allocation (Steps 3 – 4). The resulting set of security controls (baseline, tailored controls, controls inherited from providers and under customer's direct implementation and management) lead gradually to the creation of the standard profile in the CRMF's Step 4. The resulting profile can be then instantiated either as a set of customized security controls, or as a SLA to be negotiated and monitored (Steps 5 – 6).

For example, Figure 9 shows part of the recommended NIST 800-53v4 security controls for three different "profiles" (impact levels equal to low, medium and high) obtained through the presented CRMF process. In this particular case the "advance security" profile would related to the "High" control baseline, which would become the input for Step 4 in Figure 8.

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994



CNTL NO.	CONTROL NAME	PRIORITY	INITIAL CONTROL BASELINES		
			LOW	MOD	HIGH
		Access Control			
AC-1	Access Control Policy and Procedures	P1	AC-1	AC-1	AC-1
AC-2	Account Management	P1	AC-2	AC-2 (1) (2) (3) (4)	AC-2 (1) (2) (3) (4) (5) (11) (12) (13)
AC-3	Access Enforcement	P1	AC-3	AC-3	AC-3
AC-4	Information Flow Enforcement	P1	Not Selected	AC-4	AC-4
AC-5	Separation of Duties	P1	Not Selected	AC-5	AC-5

Figure 9. Excerpt from NIST 800-53 v4 showing a profile

The published ENISA report on security frameworks for Governmental Clouds[34] (GovClouds) highlights the real-world applicability of the process described in this Appendix. The GovClouds analyzed by this report have adopted a similar risk-based approach to elicit the security controls (standard profile) that offer the security level that is adequate for their operation. As mentioned before, the resulting profile (for control frameworks) would be the basis for creating a “good enough” SLA.

11 Appendix 4 – SMEs and the question of standards

Standardization has become a best practice and a reference to the EU Cloud Computing Strategy as part of the drive towards trusted, secure and reliable cloud services. With regard to research and innovation actions, the standards are one of the most important means to bring new technologies to the market[35]. For the purpose of D4.3 Final report on Cloud standards profile development, CloudWATCH collected use cases and clustered them in groups. These use cases served as a basis for the selection of standards that may guide further profile development as well as to help understand how to restrict or extend these standards in the context of the actual profiling work. Why does it matter to create standard profiles? The documentation attached to a typical standard would more often than not lend itself to various interpretations, at least as regards some parts of it. The profile approach allows locking down in more detail the particular part of a standard that suggests possible clustering. This approach has proved to be a booster for interoperability.

In 2014 19 % of EU enterprises used cloud computing, mostly for hosting their e-mail systems and storing files in electronic form. Among those, 46% used advanced cloud services relating to financial and accounting software applications, customer relationship management or to the use of computing power to run business applications[36]. The most widely reported concern of European enterprises that are using the cloud is the risk of a security breach (four out of ten enterprises, 39 %). From the ones that do not

[34] Please refer to http://www.enisa.europa.eu/activities/Resilience-and-CIIP/cloud-computing/governmental-cloud-security/security-framework-for-govenmental-clouds/security-framework-for-governmental-clouds

[35] *Standards and Standardization Handbook*, European Commission. 2013, http://ec.europa.eu/research/industrial_technologies/pdf/handbook-standardisation_en.pdf.

[36] http://ec.europa.eu/eurostat/statistics-explained/index.php/Cloud_computing_-_statistics_on_the_use_by_enterprises

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

use cloud services 42% reported insufficient knowledge of cloud computing as the main factor that prevented them from using it.

Among EU enterprises, in particular Small and Medium Enterprises (SMEs)/Businesses (SMBs) are a very challenging target when it comes to adopting cloud solutions. This is not unrelated to restricted budgets and other limitations, particularly as regards expertise in information security. SMEs are increasingly less likely to adopt cloud because of security concerns, complex terminology, fear of vendor lock-in and lack of transparency in this highly mutable environment. Against this background of widespread uncertainty, it seems that standard-based solutions should have a special attractiveness for the SMEs/SMBs community. For instance a standard-based approach allows designing software according to common standards; as a consequence it enables this software to work across a wide-range of providers. In principle SMEs/SMBs should be interested in cloud standards as at the end they facilitate their business life. With the objective of testing this hypothesis and assessing the understanding and the needs of the SMEs/SMBs community with respect to cloud standards, the CloudWATCH consortium partners ran a short survey[37]. The “Cloud standards and business’s needs” survey was launched on 9 July 2015 through to the end of August. Taking into account the limited time SMEs would predictably have for the survey, the enquiry was limited to these 6 questions shown below. The full survey is shown in appendix 6.

1. Indicate what is the core activity of your business?
2. Please rank the following cloud features on a scale of 1 to 9 depending on how important they are to your business (1 = not important; 9 = very important)
3. Would you support standards developed with EU funds to make sure that these features are optimal at all times to support your operations?
4. How important is the fact that standards are interoperable globally at all times?
5. Which are the areas where standards are needed most?
6. Are you aware of the standards supporting your cloud services? If yes: Can you name at least one?

Despite active engagement in advertising the survey of CloudWATCH consortium partners, inviting DIGITALEUROPE’s National Trade Associations (NTA) and their member SMEs to widen respondents poll, the response rate is very low - 4. Therefore, this cannot serve as a sample of business community.

Therefore, the following recommendations collected in form of SWOT analysis are also based on feedback received from NTAs involved in flagging the survey to their member SMEs.

11.1 SWOT analysis of attempts to raise awareness of SMEs regarding cloud standards

11.1.1 Strengths

Standards in the cloud computing services for SMEs are definitely helpful. In particular in IaaS, standards allow businesses to contract out of their cloud infrastructure providers with the same ease (factoring in a risk management and continuity) as contracting other providers (such as power or Internet connectivity).

[37] https://www.surveymonkey.com/r/StandardsCloud

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994

11.1.2 Weaknesses

The generally low consideration given to our questionnaire betrays an overall lack of interest in whatever happens behind the scenes or in what makes their tools meet their performance requirements. As one member NTA drew a comparison with the car industry: car owners tend not to spend much time under the hood. They display even less patience for this exercise if they lease or rent the car. Likewise, enterprises keen to implement solutions that are expected to optimise their time and to cut costs would rarely demand to be made aware of the details regarding how the solutions work for them.

11.1.3 Opportunities

This being said, this and other surveys seem to make it clear that SMEs fear vendor lock-in and have a strong interest in solutions that can cut costs, which is indeed one of the key motivations to use cloud services. These concerns only vindicate our recommended standard-based approach in creating software, infrastructures or platforms and standard implementations. Indeed, further design of standard profiles will harmonise the current standard landscape and make it more transparent. This will enhance the benefits drawn from existing standards. Users faced with the healthy multiplication of standards can only benefit from their clustering and subsequent profiling which is conducive to enhanced and smoother interoperability between standards from the various cloud fields concerned: IaaS, PaaS, and SaaS. Improved interoperability has long been a leading demand heard from SMEs communities.

11.1.4 Threats

Standardisation is all too often a dry and time-consuming process, which turns the goal to raise awareness of cloud computing services and even more, existing standards among SMEs/SMBs community into a daunting challenge for the long-term that definitely exceeds the two-year time frame of the CloudWATCH project. A legitimate concern on assessing ways to raise the awareness of small European businesses is their deliberately low involvement as users in giving feedback on standards. An issue such as standardization is mostly as a distraction from their main business, especially if they have entrusted a third party with addressing these issues. This being said, the participation of users' communities in standards development organisations is crucial for developing better future versions of standards, which depend on strong and detailed feedback. A more constructive attitude would also encourage standards and standard bodies to open up more widely to the input from the global SMEs/SMBs community. This considered approach would foster a smoother process to develop standards with a global reach.

11.2 The way forward

Surveys show that 41% of European SMEs/SMBs have not adopted any of four advanced technologies (big data, cloud computing, mobile, social media). As it is unrealistic to expect a constituency that is struggling to appropriate new tools that will put them on a par the world's best performers to take the time to monitor how tools are developed, we concluded that the best way to raise the awareness of SMEs/SMBs on cloud standards is through their trade associations at national and EU level. It would not harm if the recommendations of ETSI regarding the EU Rolling Plan for ICT Standardisation would find their way into a coordinated action plan targeting non-ICT focused trade associations of SMEs/SMBs such as Eurochambres, EuroCommerce, UEAPME.

Should this happen the CloudWATCH goal to inform stakeholders, in this case SMEs/SMBs, about what standards are out and to have them assist standardisation activities in other EU funded projects would turn into a very exciting journey.

12 Appendix 5: “Cloud standards and businesses needs” survey

This appendix provides the survey questions and answer options verbatim for the reader’s convenience. Formatting may deviate from the original SurveyMonkey presentation on the Web.

The CloudWATCH project is an EC-funded initiative aimed to accelerate the adoption of cloud services across Europe.

In addition to the tools developed by CloudWATCH, standards seek to identify common interests and to collect a focused set of use cases for profiling. 38 cloud projects, including EC-funded projects, have provided the main ground for this research. Now it is time to check if and how standards address the real needs of businesses.

12.1.1 Indicate what is the core activity of your business?

- □ ICT-related
- □ Non-ICT

2. Please rank the following cloud features on a scale of 1 to 9 depending on how important they are to your business:

Cloud features	**Rank of 1 to 9 (1 = not important; 9 = very important)**
On-demand self service	•
Broad network access	•
Resource pooling	•
Rapid elasticity	•
Measured service	•
Massive Scale	•
Homogeneity	•
Virtualization	•
Low Cost Software	•
Resilient Computing	•

Geographic Distribution	•
Service Orientation	•
Advanced Security	•

[ranks can repeat]

3. Would you support standards developed with EU funds to make sure that these features are optimal at all times to support your operations?

□ Yes

□ No

□ No opinion

4. How important is it that standards are interoperable globally at all times?

(1 = not important; 5 = very important)

5. In which areas are more standards needed (choose as many as you like)?

□ Interoperability

□ Performance

□ Portability

□ Security

□ Accessibility

6. Are you aware of the standards supporting your cloud services?

□ Yes → if yes new open question -> Can you name at least one? _____________

□ No

If you would like to know more about our work please provide us with your contact email.

13 Document Log

DOCUMENT ITERATIONS		
V1.1	First draft	Peter Deussen, Fraunhofer FOKUS
V1.2	Second draft	Peter Deussen, Fraunhofer FOKUS; Michel Drescher, EGI.eu; Neil Caithness UOXF; Nicholas Ferguson, Trust-IT
V1.3	Third draft	Peter Deussen, Fraunhofer FOKUS; Michel Drescher, EGI.eu; David Wallom & Neil Caithness UOXF
V1.4	Fourth draft	Peter Deussen, Fraunhofer FOKUS; Michel Drescher, EGI.eu; David Wallom & Neil Caithness UOXF
V1.5	Fifth draft	Michel Drescher, EGI.eu; David Wallom & Neil Caithness UOXF
V1.6	Internal review	Silvana Muscella & Nicholas Ferguson, Trust-IT
VFinal	Final edits	Peter Deussen, Fraunhofer FOKUS; Michel Drescher, EGI.eu; Nicholas Ferguson, Trust-IT
Update 1	Added Appendix 3 Added Appendices 4 and 5	Jesus Luna, CSA; Nick Ferguson, Trust-IT; Michel Drescher, EGI.eu; Katarzyna Koziol, Patrice Chazerand (both DE).

CloudWATCH is funded by the EC FP7 - DG Connect Software & Services, Cloud. Contract No. 610994