



From Project clusters to cloud standards profiles

# Picking up the data

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

# Making sense of the data

## CloudWatch's methodology

- 1. Data quality**  
Avoiding the “rubbish in, rubbish out” problem
- 2. NIST cloud characteristics**  
Functional vs. non-functional
- 3. Cloud service models**  
Do projects use the same service model?
- 4. Standards vs. service models**  
Which service models do standards imply?
- 5. Review standards for profiling**  
In which way can the selected standards be profiled?

# 1. Data quality

**Does the data provide information of statistical relevance?**

Three indicators are examined per cluster:

**1. Agreement Coefficient (AC)**

Average value per characteristic over all projects

**2. Cluster Cohesion (CC)**

Sample-based standard deviation per characteristic over all projects

**3. Signal-to-Noise (SNR)**

Calculated as AC over CC

**SNR values  $\geq 2$  indicate good data quality**

## 2. Functional & Non-functional cloud characteristics

Which of the 13 NIST characteristics are *most likely* in scope for standardisation?

### Functional characteristics

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

### Non-functional characteristics

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

# 3. Projects service & deployment models

**Does the data hide further cluster segmentation?**

- ◆ Service models: IaaS, (IaaS+), PaaS, SaaS
- ◆ Deployment models: Public, Private, Hybrid
- ◆ Cloud characteristics express different on different service & deployment models!

# 4. Which service models do standards address?

**This might ask the obvious, but ...**

- ◆ Standards addressing IaaS:
  - ◆ CDMI, OCCl, OVF
  - ◆ TOSCA (complex VM provisioning manifests)
  
- ◆ Standards addressing PaaS
  - ◆ TOSCA, CAMP
  - ◆ OCCl (via future extensions?)
  
- ◆ Standards addressing SaaS:
  - ◆ HTML5, JavaScript, HTTP, ...



# 5. How can candidate standards be profiled?

**Compliance does not guarantee interoperability!**

**Which compliance requirements can be scoped down further?**

- ◆ The standards profiler's toolkit
  - ◆ **Notes** – clarify ambiguous (non-)normative text
  - ◆ **Restrictions** – scope down alternatives and options
    - ◆ MAY, MAY NOT, etc. → MUST, MUST NOT, etc.
  - ◆ **Extensions** – define the undefined
    - ◆ Forbid extension points where required
    - ◆ Define (exhaustively!) allowed extensions



# Cluster 1 – scientific computing

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
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

- ◆ Most important characteristics

On demand self service – Massive Scale – Homogeneity

- ◆ Strawman profile standards

- ◆ OCCI/CIMI & CDMI

# Cluster 2 – Trusted public clouds for governments

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
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

## ◆ Most important characteristics

Measured service – Advanced security – Resource pooling

## ◆ Strawman profile standards

- ◆ Usage Record 2, NIST SP 500-307, CIMI, (AMQP)
- ◆ ISO/IEC 27000, NIST SP 800-53, CCM 3.01

# Cluster 3 – High performance dedicated purpose applications

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
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

## ◆ Most important characteristics

Geographic distribution – Massive Scale – Measured Service – Massive Scale

## ◆ Strawman profile standards

◆ OCCI

◆ Usage Record 2, NIST SP 500-307, CIMI, (AMQP)



## Break-out sessions

Let's get things done!

# Three breakout groups

- ◆ Group 1
  - ◆ Cluster 1 – Scientific computing
  - ◆ Chair: Peter Deussen, Fraunhofer FOKUS
  
- ◆ Group 2
  - ◆ Cluster 2 – Public trusted clouds for governments
  - ◆ Chair: Damir Savanovic, Cloud Security Alliance
  
- ◆ Group 3
  - ◆ Cluster 3 – High performance dedicated purpose applications
  - ◆ Chair: Neil Caithness, University of Oxford

# Breakout objectives

- ◆ Verify & confirm clustering and mapping
- ◆ Confirm the need for standards and profiles
  - ◆ Do requirements overlap or complement?
- ◆ Which standards are missing in the strawman profiles? Which are superfluous/obsolete?
- ◆ Is the strawman profile worth pursuing?
- ◆ Which SDO would be available and willing to support/host the profiling process and work?

# Any questions?

- ◆ Lunch will wait for you after the breakouts!

