PUBLIC CLOUD FOR AI RESEARCH IN HEP

Sofia Vallecorsa – CERN IT Innovation

March 10th, 2023

WHY DEPLOYING ON PUBLIC CLOUD?

- Possibility to **burst out** to **large distributed systems**
- **Pay on demand**
- Access to **state-of-the-art (specialized) hardware** (GPUs, TPU, IPU, Quantum, ...)
- Access to **state-of-the-art services**: MLaaS/DLaaS (and QCaaS)
- **Flexibility** in choosing the level of deployment automation

DEEP LEARNING ON PUBLIC CLOUD

Enables DNN **R&D,** extends the **range of applications,** allows **“explainability” and systematic studies**

DL workflows **run optimally on Public Cloud**

- Typical **burst-out** pattern → O(100) GPUs for a short time
- Specialized **hardware improves efficiency and sustainability**
- Profit from **MLaaS/DLaaS solutions: how “custom”** are our models?

Study optimal deployment strategy:

- Take into account all communication **cost**
- Keep **physics accuracy** under control
- **Reduce monetary cost**

WHAT ARE WE TESTING?

2020-2021: A set of benchmarks through the **CloudBank EU project**. Examples:

- Generative Adversarial Networks for detector simulation (*IT*)
- Quantum Reinforcement Learning on the DWave annealer (*with ATS)*

2022-2023: ***CERN openlab Oracle project***

Cardoso, R., Golubovic, D., Lozada, I. P., Rocha, R., Fernandes, J., & Vallecorsa, S. (2021). Accelerating GAN training using highly parallel hardware on public cloud. In *EPJ Web of Conferences* (Vol. 251, p. 02073).

CERN: BURSTING DLAAS TO PUBLIC CLOUDS **@ISC2021**

*"**Very relevant for the research community** (...) Will be interesting to see the subsequent publications"*

Reviewer 1

*"**..interesting to ISC HPC audiences** working on scientific applications."*

Reviewer 2

*"**... very interesting and relevant for ISC (...)** a complex setup that is generally the **main burden of industrial applications.... "***

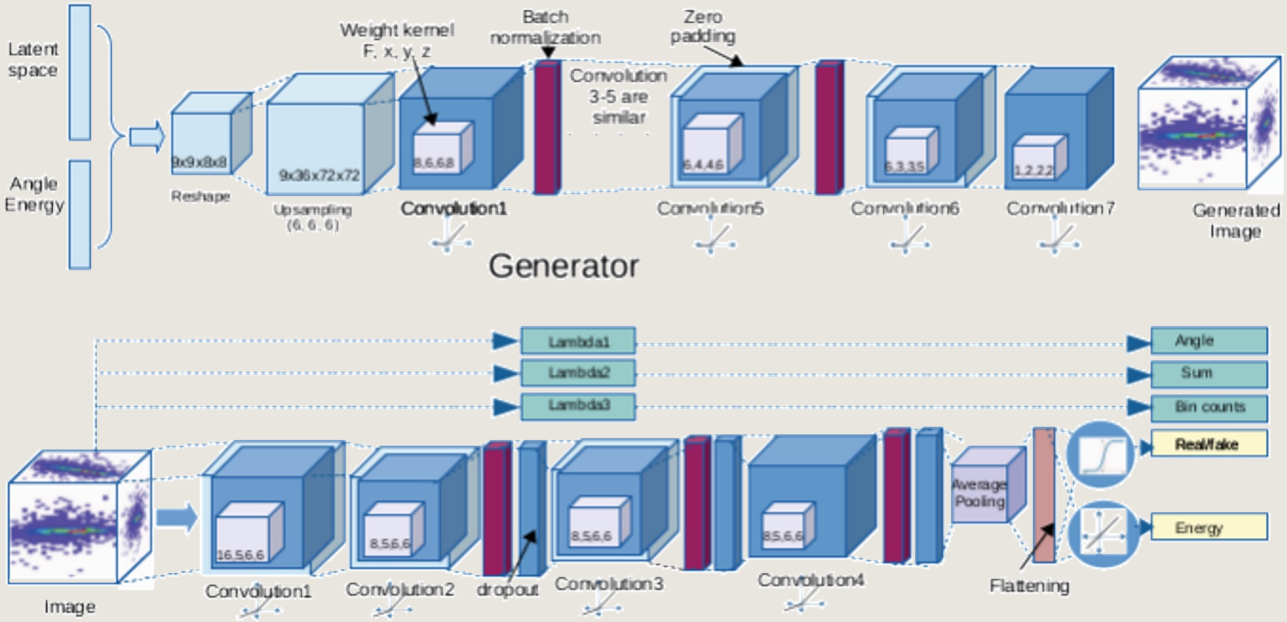
Reviewer 3

Also presented at vCHEP2021

THE 3DGAN BENCHMARK

- **3D convolutional GAN**[1] for simulating electromagnetic calorimeters
  - Relatively **small dataset** (few 10s GB)
  - Relatively **small model** (few millions parameters)
  - Heavily **compute bound**
- Data parallel training

3DGAN trains in 1 week on a NVIDIA P100



[1]G. Khattak. EPJ Web of Conferences. Vol. 214. EDP Sciences, 2019.

DEPLOYMENT VIA PUBLIC CLOUD SERVICES

- Direct low-level management of virtual machines
  - significant fraction of the infrastructure management burden on the user
- Open platform (Kubernetes)
  - Abstract infrastructure through APIs and delegate most of the operations to cloud service itself
  - **Test on Google Cloud**
- Vendor MLaaS frameworks
  - Infrastructure and workload optimized
  - **Test on Microsoft Azure**

KUBEFLOW BASED DEPLOYMENT ON GCP

Existing **on-premises deployment** and configuration are **reused and directed** to Google Kubernetes Engine.

From 1 to 128 GPUs

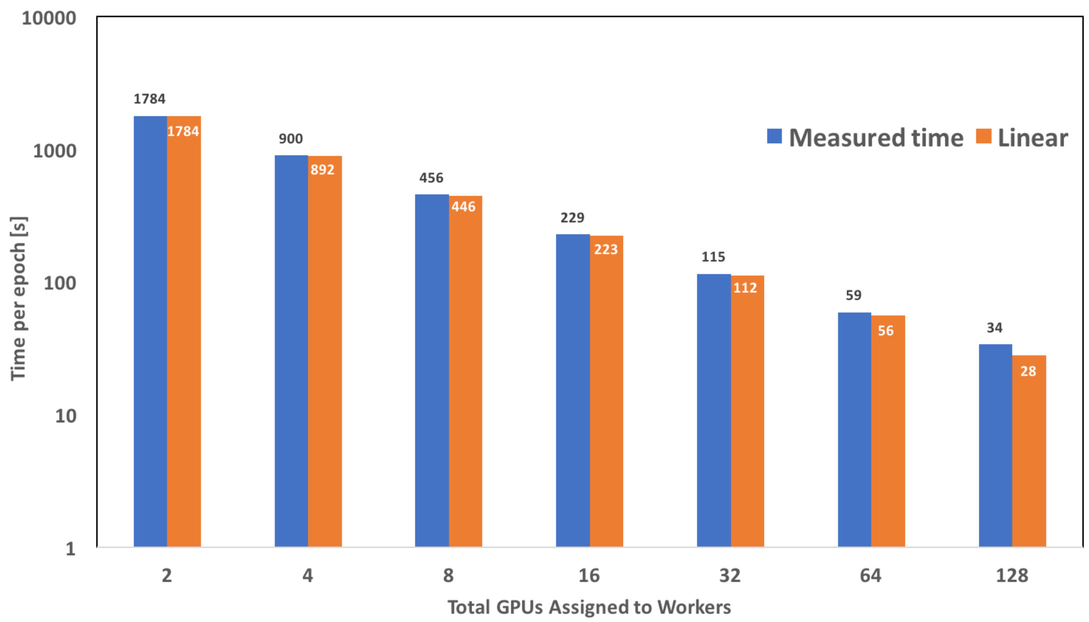
Multiple node groups, with a number of V100 GPUs per node varying from 1 to 8

Optimise configuration

GPUs per node, per workers, batch size, etc..

x100 near linear speed-up
3DGAN training down to 1 hour

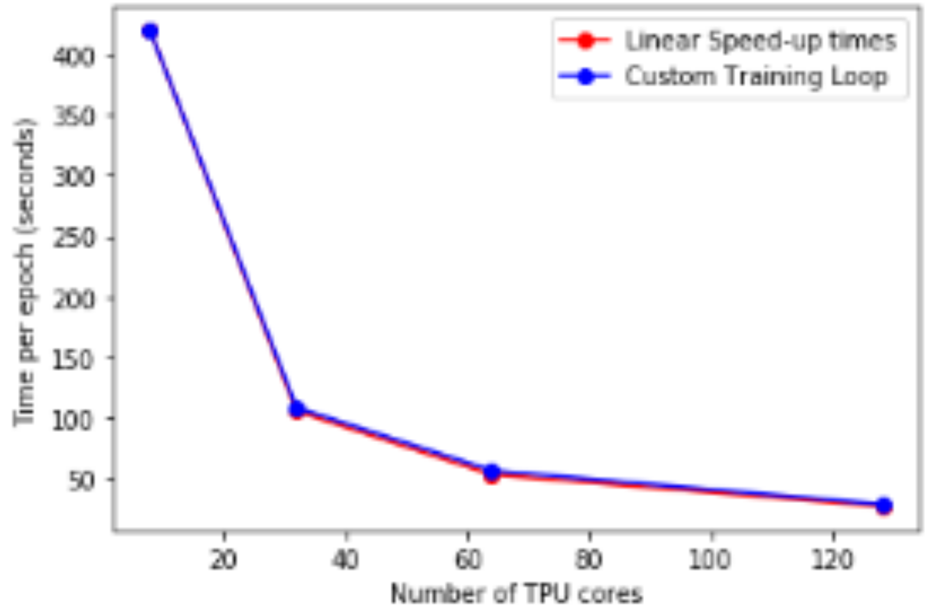
We could train GAN-based simulation at the scale of the ATLAS calorimeter in a matter of days instead of months!

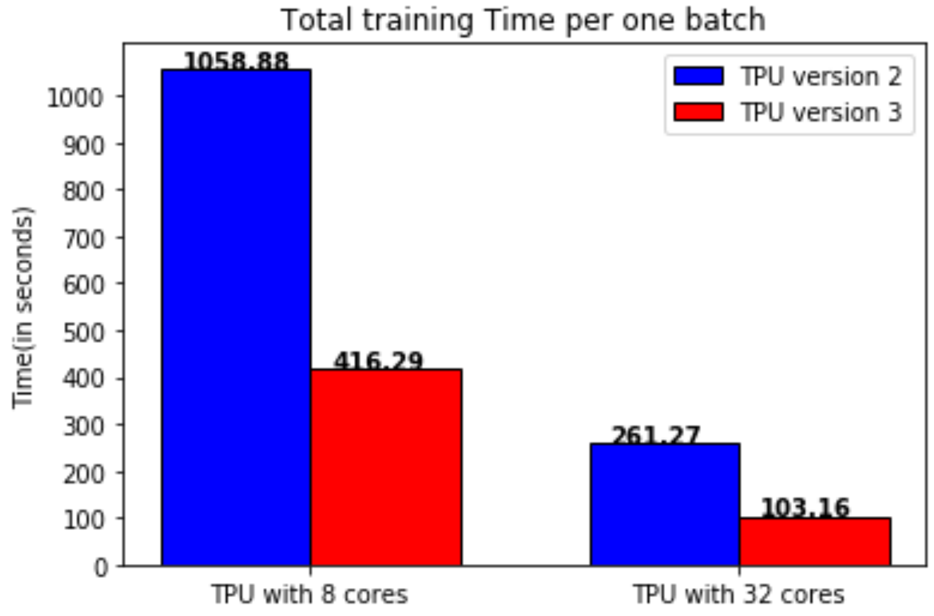


Access through CloudBank EU

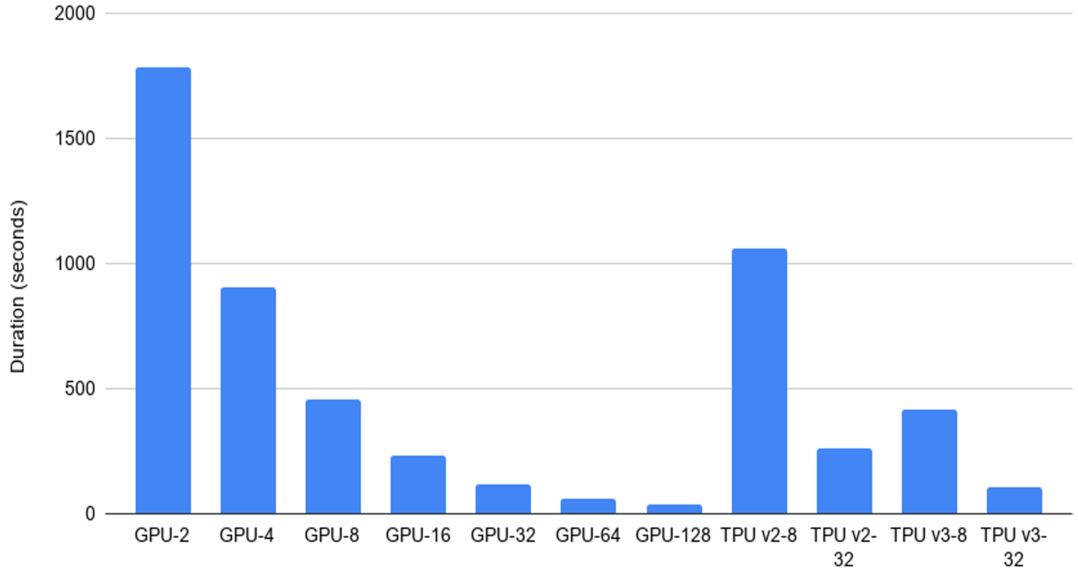
TRAINING ON TPUS

Access to different hardware beyond GPUs extends the range of optimization
Example: TPUs





TPUs versions comparison



TPU vs multi-GPU comparison

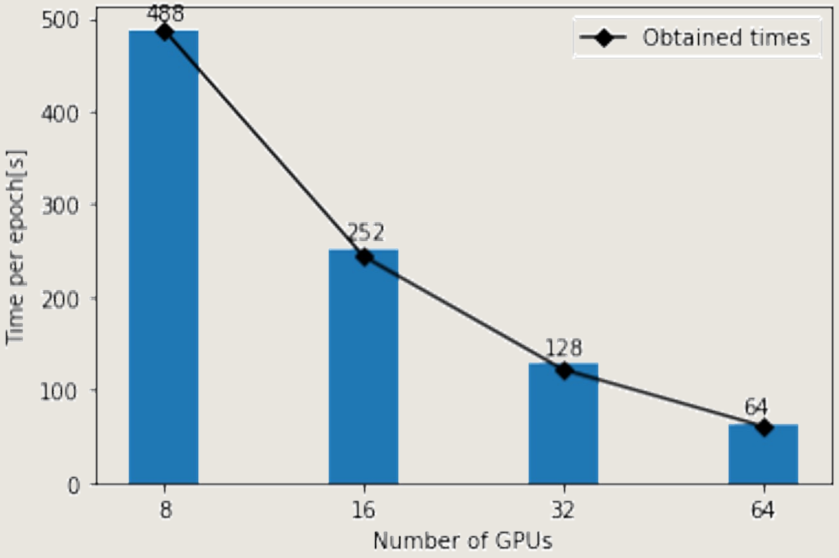
MICROSOFT AZURE MLAAS

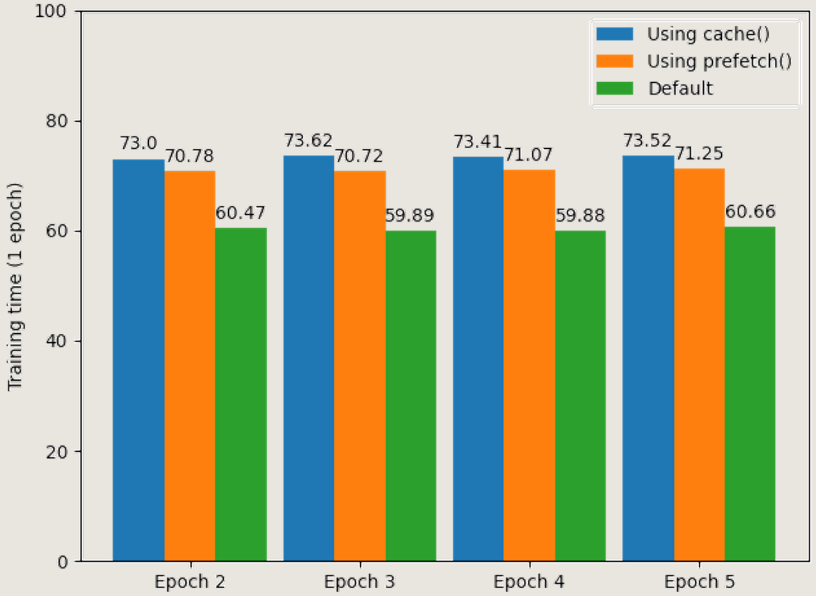
Managed service:

- Compute cluster **provisioning** is entirely **operationalized by the service**
- Automatic optimization of data set management (**caching, pre-fetching and parallel data loading**).

Our test:

- **24 cores VMs** with 448 GiB memory and **4 V100 GPUs each**.
- **Comparison to manual tuning** of caching, pre-fetching and auto-tuning





Azure Machine Learning Service used through CERN openlab's Azure cloud research enrollment

A DEEPER LOOK AT MLAAS: 2022-2023 CERN OPENLAB ORACLE PROJECT

09.03.23

AI MODEL CATALOG

Increasingly popular components in AI ecosystems

- Help **introduce** ML/DL in areas where expertise is limited or still building
- **Accelerate** R&D, improve **reproducibility** and **quality monitoring**
- **Simplify** deployment

Model registry in Oracle is available through Oracle Accelerated Data Science SDK (ADS)

Investigate deployment **on CERN cloud and OCI**, following a "hybrid cloud" model

Study ADS **functionalities wrt to HEP use cases**

TOWARD ”SMART” CATALOGS?

How much we can **realistically re-use a pretrained model** on different data?

Possibility is limited by the nature of the **input data** and the **model generalization capabilities**.

In HEP directly applying a model trained for analising the output of a specific detector to a different use case is usually **unfeasible**, **inefficient** or **too costly** in terms of the data pre-processing step.

Can we leverage current studies on DL generalisation to build a catalog of adaptable models?

WHAT ABOUT COSTS?

Cardoso, R., Golubovic, D., Lozada, I. P., Rocha, R., Fernandes, J., & Vallecorsa, S. (2021). Accelerating GAN training using highly parallel hardware on public cloud. In *EPJ Web of Conferences* (Vol. 251, p. 02073).

A FEW WORDS ON COSTS ON GCP

It is critical to understand costs **upfront**

- Reach **optimal service settings** and **sustainable run** configurations

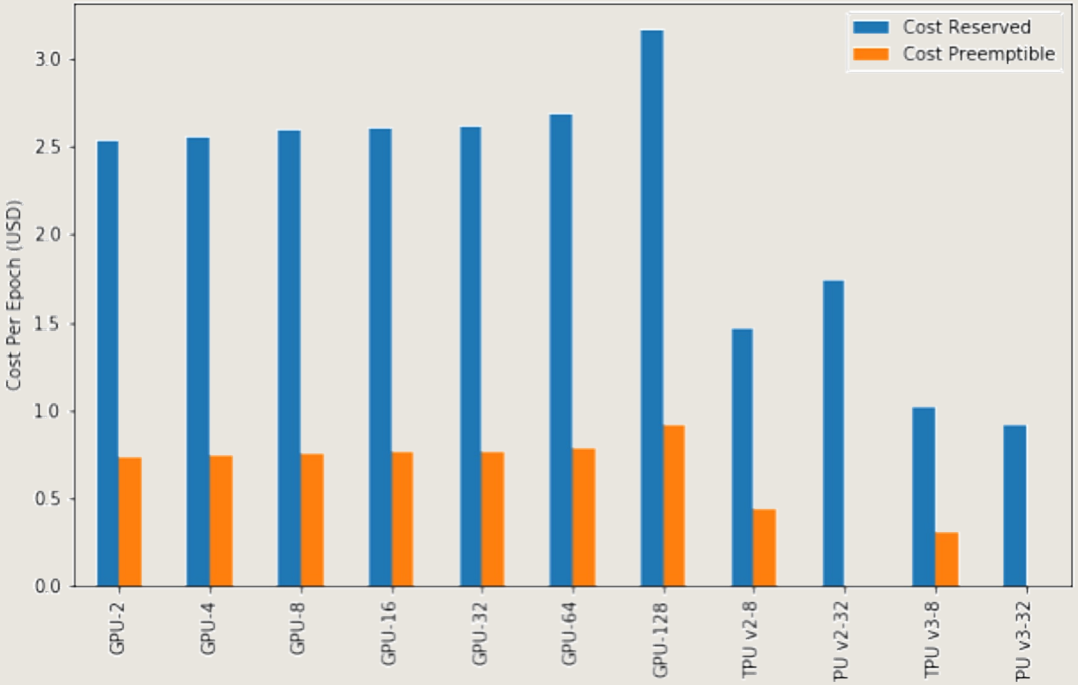
Infrastructure costs are driven by the GPUs

- Similar for all configuration, while **reducing the training time**

Best results use **pre-emptible TPU v3-8**

- 2.4 times cheaper than their GPU equivalent

3DGAN training down to 1 hour for as little as ~25 USD
"the cost of a pizza … in Geneva" !



KUBECONN HIGGS ANALYSIS

CERN Analysis Run

Kubernetes 1.12

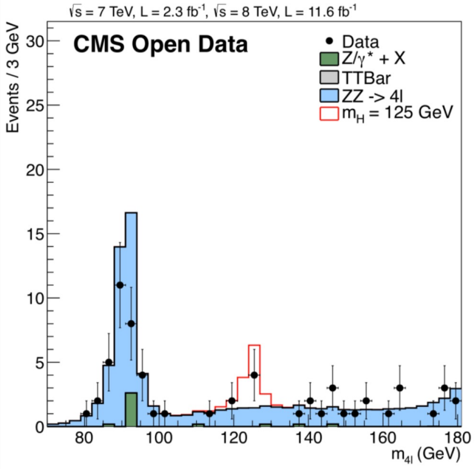
61 Nodes (VMs)

- 4 Cores / 8 GB nodes
- 40GB disks (SSDs)

Running on 36TB (half the dataset)

Total time: 19h with 244 cores (~1GB/s)

Goal for GCP: 250x speedup to run it in <10min



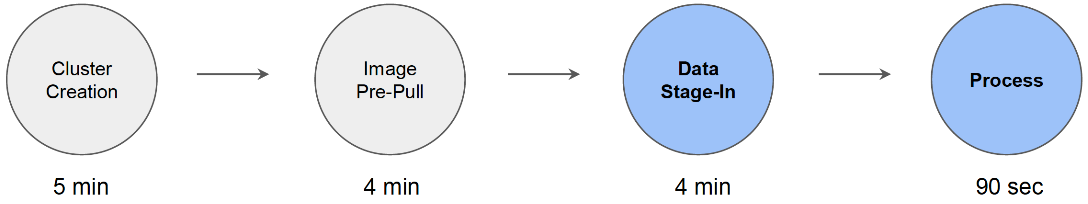
Estimated cost < 300 USD

GCP Analysis Run

Kubernetes clusters on GKE (Managed Kubernetes service on GCP)

Today's run included (real demo run was ~2x that)

- 660 nodes: n1-highmem-16, 104 GB RAM
  - 10560 cores, 69 TB RAM



https://www.youtube.com/watch?v=CTfp2woVEkA

PUBLIC CLOUD AND SUSTAINABLE AI ?

ML/DL footprint is becoming more and more relevant

AI community are starting to define **best practices**[1]

- **Efficient ML architectures**
- **Efficient training strategies (**self-supervision, few-short learning, meta-learning, ...)
- **Processors and systems optimized for ML training** versus general-purpose
- **Centralised computing on large Public Cloud data center**
- Possibility to choose locations using clean energy
- **New hardware** (neuromorphic, quantum)

1 Patterson, David, et al. "The Carbon Footprint of Machine Learning Training Will Plateau, Then Shrink." (2022).

CONCLUSIONS

Missing easy-to-use, reliable tools and a mechanism to access them is detrimental to the research process

Public cloud can provide such tools

- Essential ingredient to **Deep Learning (**and now **Quantum Computing)**

Need to validate services (offered as commodity) for scientific use cases

- Must **establish global strategy** combining technical performance with cost optimization

In 2021 we demonstrated the use of public cloud services at scale for DL in scientific application thanks to CloudBank EU project

- **Hybrid provisioning and orchestration**, profiting from the variety of technology that each cloud provider offers.

A DL researcher wish-list:

Well-defined unified cloud access mechanism ....
...dynamic, adaptive...
... **possibly lightweight** ;-)

Clear way of **understanding costs**

THANK YOU