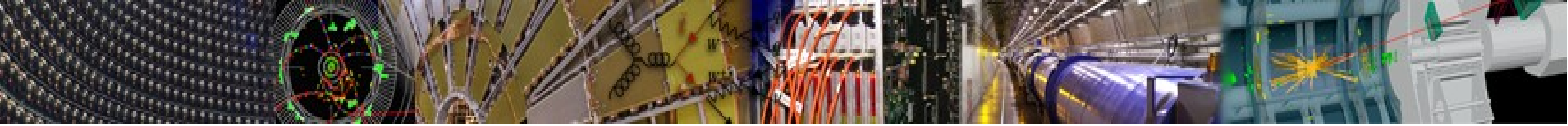


Computing Strategy for Hyper-Kamiokande

Francesca Di Lodovico

14-15 January 2013

Second Hyper-Kamiokande Open Meeting

Kavli IPMU, Kashiwa

Queen Mary University of London

erc

Introduction

- Computing is an essential part of Particle Physics:
  - Experiments take data for decades and computing must have continuing support and planning with constant technological renewal.

➔ Need strategic plan: “Computing Model”.

- A lot can be learnt from current Particle Physics experiments, particularly the LHC. Of course, we can also count on our own experience in T2K. But not everything is the same:
  - New technologies
  - New needs
  - New rules

Introduction

The current strategy at T2K is:

- Super-Kamiokande:
  - Running data and MC in situ at Kenkyuto (Higashi-Mozumi)
  - Copy data directly from Kenkyuto
  - For T2K members, beam data&MC uploaded to the Grid
- ND280:
  - Using a tiered Grid model
- Future:
  - Adapt previous experience to current experiment and technologies as much as possible.
  - Possibly Hyper-K will still run data and MC in situ → but computing strategy needs to be flexible to distributed running.
  - HK storage to be distributed to be easily accessible world-wide.
  - The near detector will likely continue running on a tiered distributed model.

H-K Computing Model Requirements

- Minimize over-heads and maximally exploits avalaible resources
  - Use virtualization
  - Easily expandable to different computing and storage sites
- Flexible to adapt to new technologies:
  - exploit current developments and use the best for our computing model
  - experiments will take data for decades and need to be as insensitive as possible to technology changes
- Comply with new rules in EU/US about long-term data use
  - data needs to be available for at least 10y and keep up with technology changes
- Easy to access data/MC:
  - minimize the burden of e.g. requesting certificates

All requirements can be met through the incorporation of Cloud Computing, Virtualization and Digital Preservation concepts.

The Cloud Computing

- Cloud computing indicates computing resources (hardware and software) that are delivered over a network.
  - It can be described as the umbrella term for the provisioning of IT platforms, infrastructures and applications as services.
- The Cloud provides services according (mainly) to 3 models: SaaS, PaaS, IaaS
- In particle physics:
  - The IaaS is used in most cases (production and storage), which means installing virtual machines.
  - Distributed storage solutions are SaaS and provide uniform access to distributed storage
  - Some clouds at labs as CERN, RAL, TRIUMF, KEK would provide the OS then one could build applications natively (PaaS).
  - The Cloud can be a hybrid of several services.

Cloud Client

↕

Software as a service (SaaS)
Cloud providers install and operate application software.
E.g. Email, virtual desktop, communication, games,...

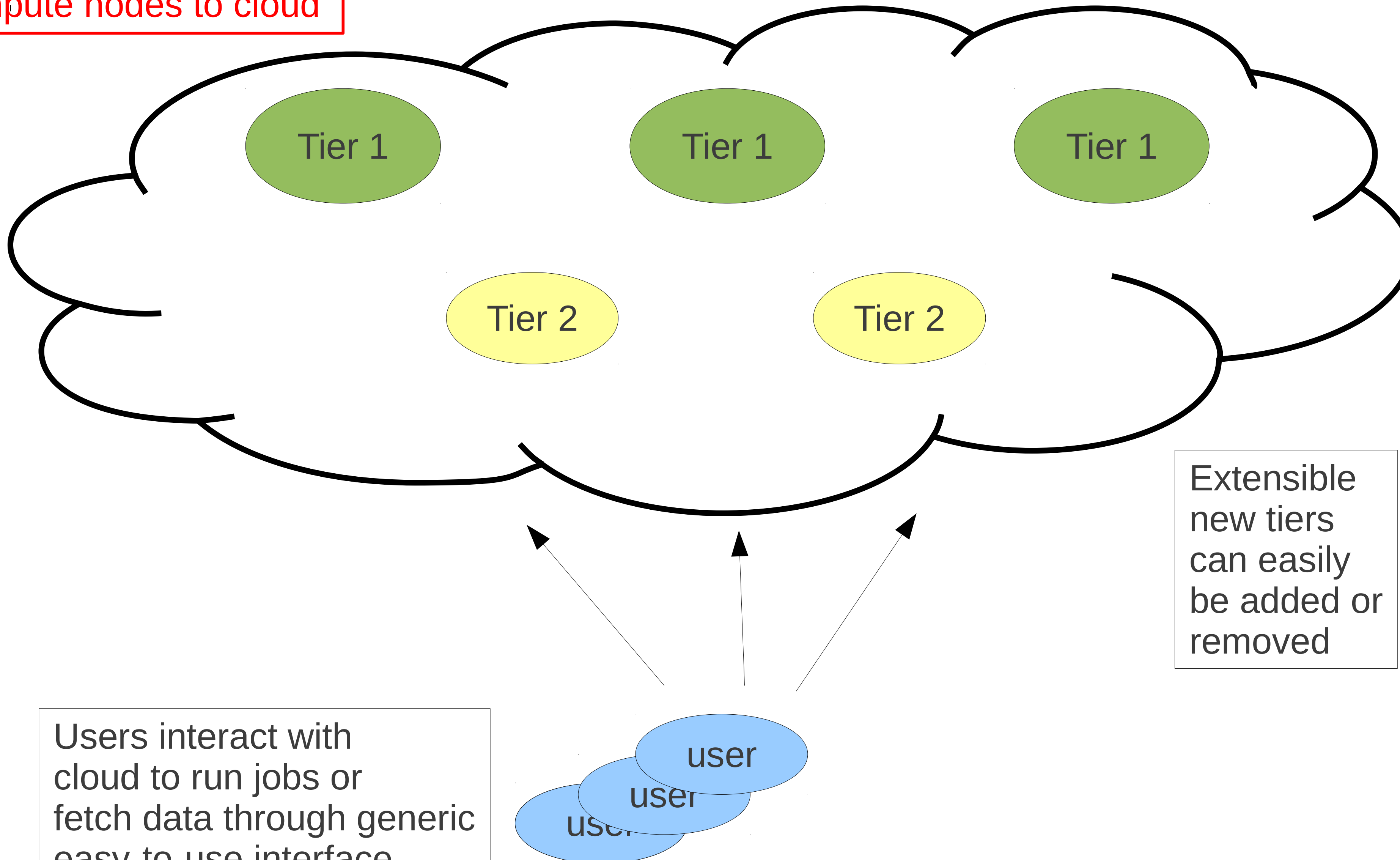
Platform as a service (PaaS)
Cloud providers deliver a computing platform.
E.g. Database, Web server, development tools,...

Infrastructure as a service (IaaS)
Cloud providers offer computers and other resources.
E.g. Virtual machines, servers, storage,..

Cloud Infrastructure

Tier 1 & 2
Provide storage &/or
compute nodes to cloud

It could possibly be a distributed computing Cloud: multiple compute center configured as Clouds



Extensible
new tiers
can easily
be added or
removed

Users interact with
cloud to run jobs or
fetch data through generic
easy-to-use interface

Cloud Storage

- Storage needs to scale with experiment.
- Anticipate cache close to experiment that holds 'current' data (tbd).
- Data transferred as collected to Tier 0 and then replicated to Tier 1 sites.
- Tier 1 sites can hold overlapping datasets.
- Use of cloud storage system provides uniform interface to storage in different locations.
  - To user is just a big pool. They can get their data from most convenient location

Towards the Cloud Production

The Cloud computing is fundamentally an evolution from the Grid. Current Computing Grid has taught us:

- Data Processing and HEP simulation tasks can be efficiently shared between many sites.
- Data can be shipped between centers quickly. Research networks are a very reliable component.
- A 1-3y to stabilize according to past experiences.
- More access to raw data at the beginning up to when detector and reconstruction are completely understood → experimental needs vary with time.
- The compound of a software service and a hardware machine is inflexible and prevents the usage of idle hardware by other sofware services → Wider distribution requires a more homogeneous system environment or a simpler one (or both) to keep things manageable.
- Difficult to run jobs on sites where your experiment does not have local support and without involvement of the experiment

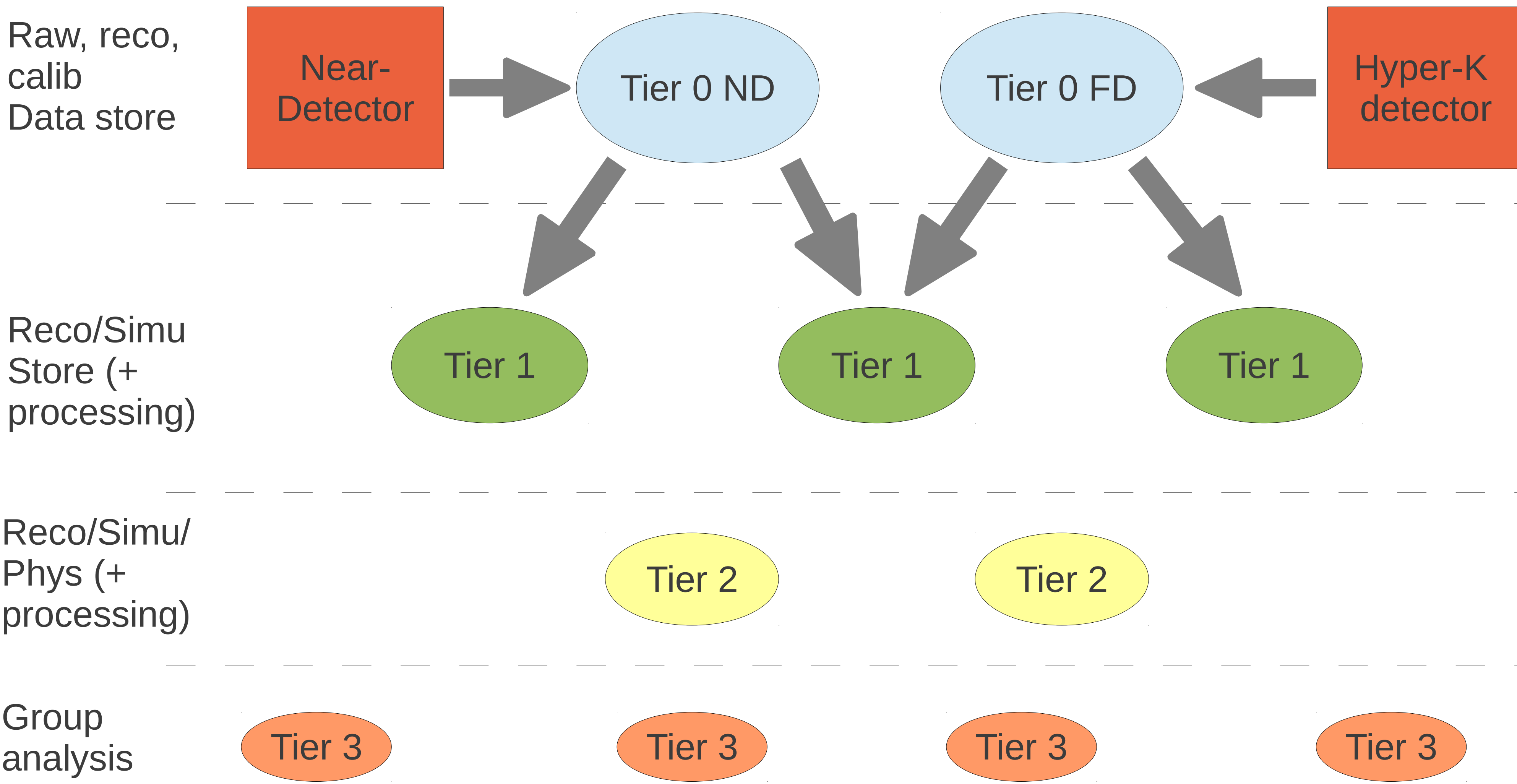
Comparisons Cloud vd Grid

- Cloud want to maximize the use of the resources and decrease the involvement of the experiments in the usage of the resource themselves (e.g. installing code to run)

You need	**Who provides it (Grid)**	**Who provides it (Cloud)**
A computer connected to network, with conditioning and power	Local site staff	Local site staff
An operating system “compatible” with the application	Local site staff, after negotiation with experiments	Comes as a virtual image from the experiment central infrastructure
A local installation of the experiment software (and a local area where to store it)	Local site staff provides area, Experiment support installs software	Downloaded on demand from the experiment central infrastructure
Machines for local experiment facilities (voboxes etc)	Local site staff provide them.	They are also virtual images / not needed locally
A local storage containing the input data	Local site staff needs to have bought storage for the experiment	They are also virtual images / not needed locally
A configuration to be executed	User!	User!

Hierarchical Tiered Computing Model

Model easily expandable to a potential near detector.
Cloud can provide both production and data storage.

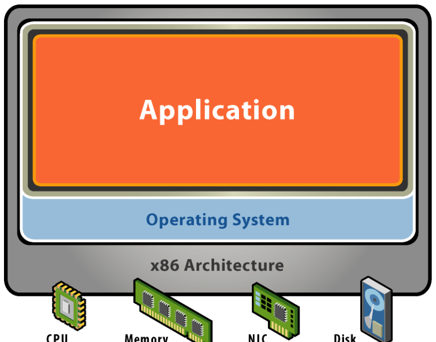


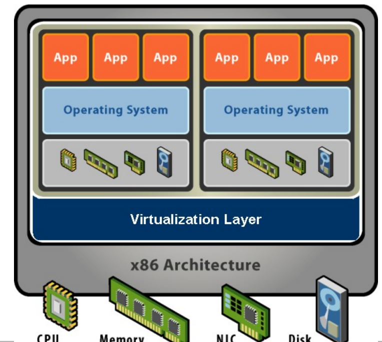
Run virtual machines (<u>efficient, isolated duplicate of the real machine</u>) in the Cloud to optimize resource usage & data preservation

Virtualization

Key issue: applications need to run on different platforms and software environments.

A solution to this problem is virtualization: simulation of the software and/or hardware upon which other software runs.

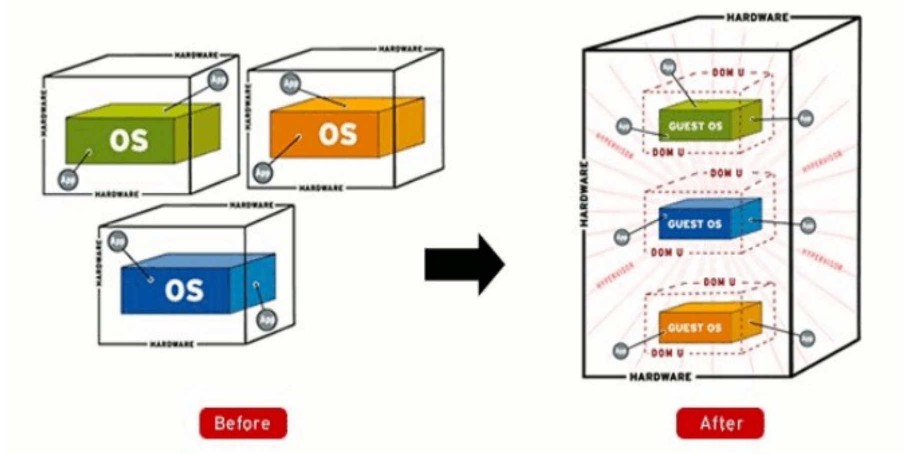




Starting point: a physical machines	Virtual Machine (VM)
Physical hardware:Processors, memory chipset, I/O devices, etc. •Resources often grossly under-utilized	Software Abstraction: •Behaves like hardware •Encapsulate all OS and application state
Software:Tightly coupled to physical hardware •Single active OS instance •OS controls hardware	Virtualization Layer: •Extract level of indirection, enforces isolation •Decouples hardware, OS •Multiplexed physical hardware across VMs

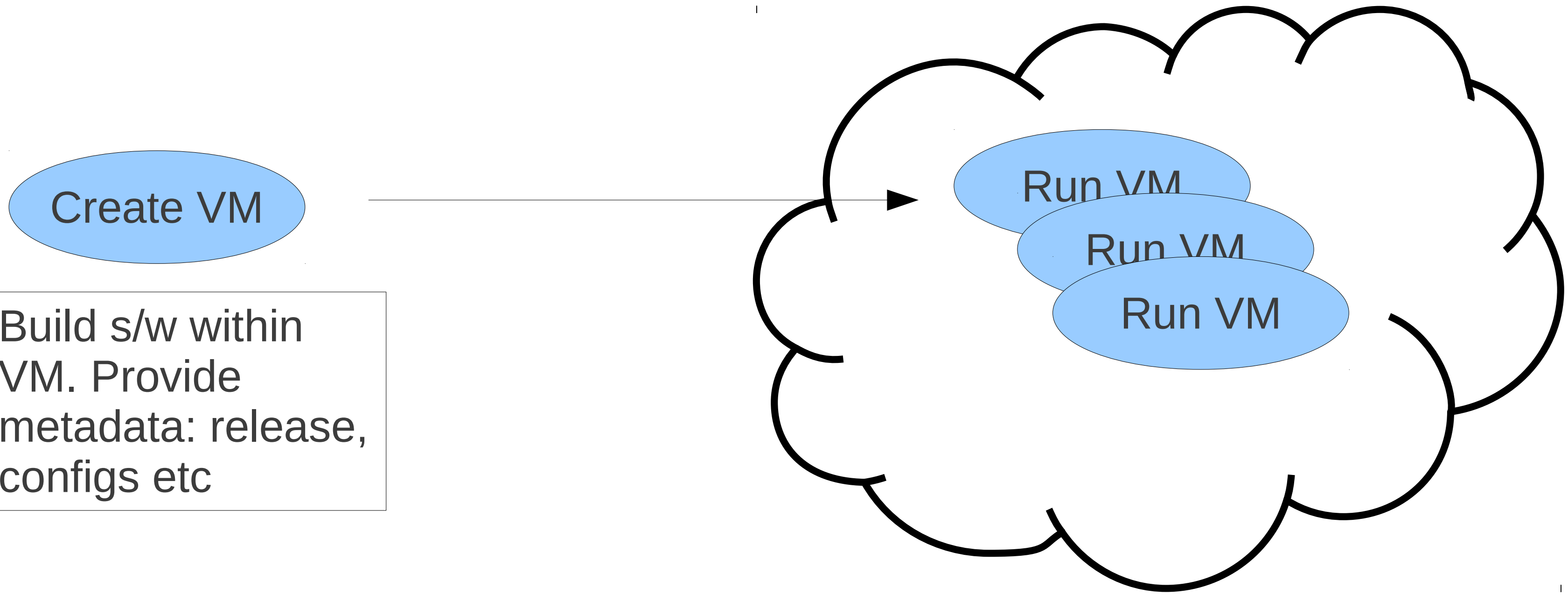
Virtualization Properties

- **Compatibility**:
  - Provide applications with the needed operating system in virtual machines without altering the host system → increase the experimental software life-cycle.



- **Load-Balancing**:
  - Better overall utilization of the hardware resources by shifting virtual machines between available resources.
- **High-Availability**:
  - By migrating or re-starting virtual machines on different hardware resources downtimes due to maintenance or hardware failures can be minimized.
- **OS-Evaluation and Testing**:
  - Possibility to test new systems and setups without touching the production system or providing additional testing hardware.
- **Encapsulation**:
  - Limit risks for host system to be infected by malware and non-disclosure issues by encapsulating users and applications in virtual environments.

Processing Workflow



Build s/w within VM. Provide metadata: release, configs etc

Cloud manager:
- Fills metadata store
- Pre-fetch data,
- Submit job to cloud nodes
- Manage jobs, data

Metadata captured from job creation through to completion helps to maintain provenance of data.

Digital Preservation

- Digital Preservation: ability to retain the possibility to look into data for a long long time
  - It is becoming crucial in particle physics (our data is valuable, both in its physics content and in funding that went into that)
  - Better plan in advance (before hundreds of PBs are stored in fragile ways)
- H-K: provide and maintain the capability to perform the H-K data analysis for at least 10 years
  - Code & repositories, data, databases, documentation, storage, CPU
- Minimize the effort needed to support and maintain the system
  - Think of: hardware,operating system upgrades, tool upgrades, code validation, maintaining documentation, interaction with data center infrastructure, etc.
- This comes automatically with the usage of Virtual Machines, but needs to plan data storage format

Computing Tasks

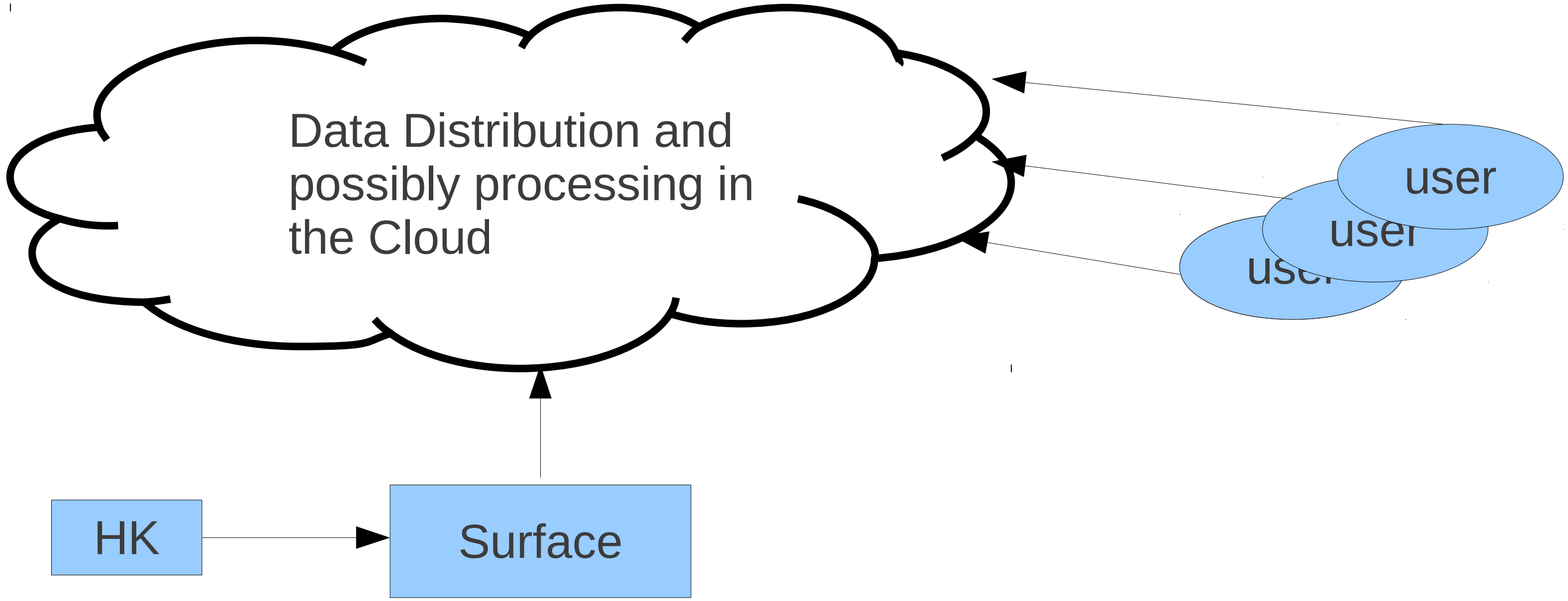
- To define rough initial estimates for HK computing, we need to look at few tasks. They are divided into:
  - Cavern/surface computing issues
  - External data distribution/processing issues

Computing Tasks	Inputs
Temporary data storage and production capability in the cavern	•Event rate/categories. •Storage needed in the cavern (limit dependence on surface connection). •Computing power underground.
Internet connection cavern-surface	•Rate of events transferred to the surface.
Data storage on surface	•Assess whether surface is Tier-0. •Event rate.
Production capability on the surface	•Processes to be run on the surface (low level detector checks, calibrations..). •Event rate.

Computing Tasks

Computing Tasks	Inputs
Clouds	•Several clouds are being created. Choose clouds for testing. •In the future probably there will be a particle physics cloud (not existing now yet).
Virtualization	•Create virtual machines with the current Hyper-Kamiokande simulation and reconstruction code. •Test virtualization also with other exisiting codes.
Digital Preservation	•Look in particular at the long term data storage and current alternatives.
Production capabilities	•Processes to be run. •Simulation production processes. •Tiers availaible. •Production capacity.

Inter-group interactions



- Computing defines the infrastructure: it covers data distribution and data and simulation production.
- To set-up the infrastructure, there is need of inputs from several groups:
  - Cavern: space available for machines
  - DAQ: event rate
  - Network: connection to surface
  - Detector: process to run in situ (monitoring, data quality..)
  - Physics/calibration/software: data and MC to run
  - Collaboration: tier sites

Planning

- Main goals: robustness, stability, low overhead during normal running.

- Collaboration with the H-K groups will allow to define the specifics of the model and needed computing power in the cavern/surface.

- Next years, tests can be performed with existing code (e.g. T2K's) on the clouds, virtualization and digital preservation.

- At strategic points close to data-taking (eg 5, 2 and 1 year before) Data challenges will be performed.

- Data challenges and the commisioning runs will be used to exercise the Computing Model and to test the infrastructure in a simil-production mode.

Expertise

- Have worked on BaBar and ND280 computing.

- Have already utilised Virtual Machine technology within T2K
  - Further work to build VM using approved open Cloud approaches.

- Connections with existing EU preservation projects (EUDat, Pericles) and BaBar Long Term Data Analysis (LTDA).
  - Makes it easier to leverage current approaches and avoid pitfalls.

- Connections with cloud community (cloudscape, open cloud consortium, etc):
  - Leverage existing experience and tools and understand pitfalls

Conclusion

- Computing model that minimizes over-heads and provides data distribution and includes long-term access is essential.
  - This can be achieved with cloud computing, virtualization and data preservation.
- The computing model will need to take advantage of new developments: need new competences and a framework to shield the users. Start now.
- Work starting now to develop the model and will be used first for the MC event distribution and then for the first data (before beam, e.g. commissioning data)
- Solid expertise from T2K and LHC in the UK useful for the developing of the model
- Strong interest from the UK, but other Countries should be involved. If you (or your computing people?) are interested please let me know. Aiming to start regular discussions on computing.

Additional Info

Technology Trends

CPU:

- While transistor count per chip is still increasing, clock speed is not. What you are going to have is many processing units (“cores”) with constant today-like speed
  - Even at transistor count, increase is expected to slow down with smaller CMOS processes;
  - Power consumption can become problematic.

Disk:

- While capacity/$ still seems on the exponential curve, I/O speed is not. A PB system made up with 10 disks you need weeks to be read is not that useful. SSD disks have an initial penalty of 10x in size (or cost). Also, already today, storage budget is an increasing fraction of the total.

Network:

- Indeed seems to keep its “good behavior” for at least some years. 100 Gbit/s geographical networks are “close even today”.

The Cloud Computing

- Characteristics:
  - Availability of large computing infrastructure on need basis
  - Maintenance (virtualization)
  - Device and location independence
  - Reliability
  - Security

- Very similar to the Grid except:
  - Lower barrier to use due to thin-client so less configuration for user.
  - CERN and other computing centres starting to offer Cloud services.
  - Universities also starting to offer Cloud services.

Some Relevant Worl-wide Projects

Just a list of a few projects we use their expertise:

- EUDat (EU FP7 founded) – European Data Infrastructure:
  - Produce the common low-level services that are required to provide the level of interoperation and trust of data that is necessary to support both widespread access to data, and the long-term preservation of data for use and re-use.
- Pericles (EU FP7 founded, strating in Feb 2013):
  - Project aimed at looking at data lifecycle within an archive (including metadata and policy lifecycle).
- BaBar LTDA - Long Term Data Analysis:
  - Data Preservation and Long Term Analysis in High Energy Physics (BaBar stopped data-taking in 2008)
- Cloudscope(http://www.cloudscapeseries.eu/Pages/Home.aspx):
  - A forum for cloud providers and users to discuss current issues
- Open cloud consortium:
  - Supports the development of standards for cloud computing and frameworks for interoperating between clouds

Private Cloud Examples

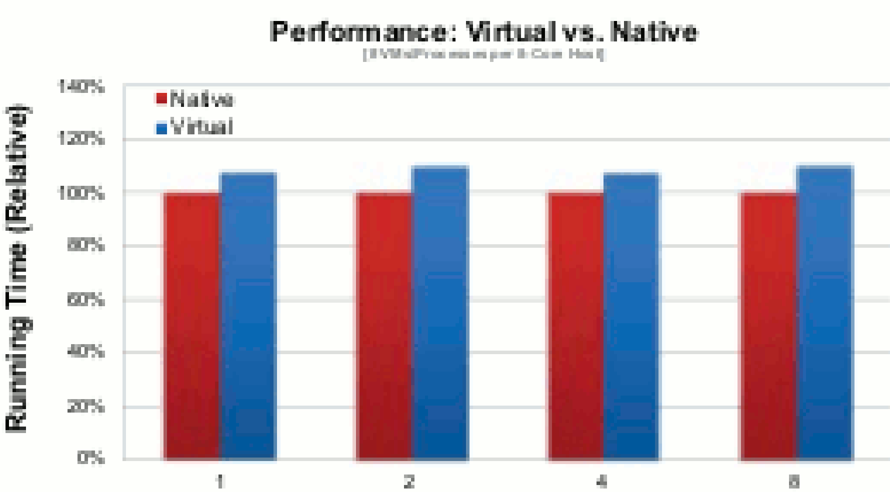
A list of Private Cloud Examples (from Wikipedia):

- Examples of Software as a Service (SaaS) include: Google Apps, Microsoft Office 365, Onlive, GT Nexus, Marketo, and TradeCard.
- Examples of Platform as a Service (PaaS) include: Amazon Elastic Beanstalk, Cloud Foundry, Heroku, Force.com, EngineYard, Mendix, Google App Engine, Windows Azure Compute and OrangeScape.
- Examples of Infrastructure as a Service (IaaS) providers include Amazon CloudFormation, Amazon EC2, Windows Azure Virtual Machines, DynDNS, Google Compute Engine, HP Cloud, iland, Joyent, Rackspace Cloud, ReadySpace Cloud Services, and Terremark.

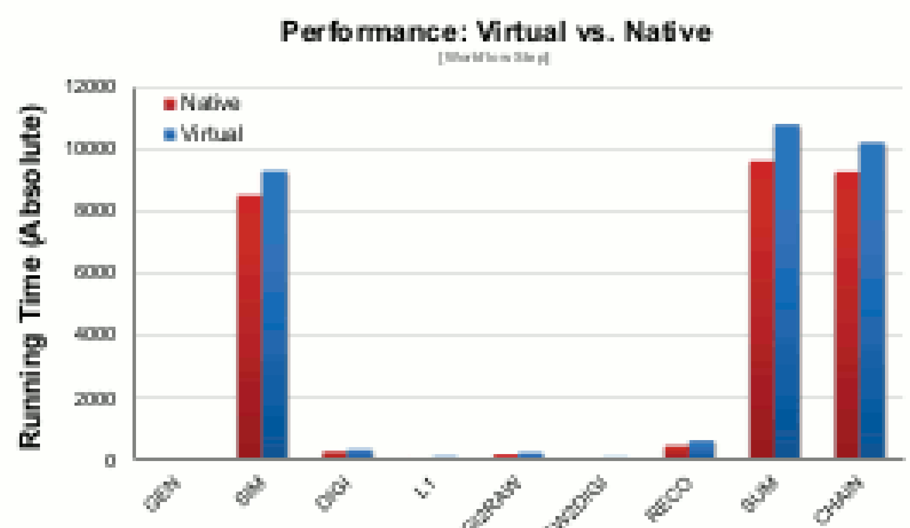
Virtualization Performance

- To estimate the performance impact of the virtualization layer, several standard benchmarking tools are available.
- However, each is specialised on testing only a specific part of the system. To have an overall impression of the performance loss for typical HEP applications, detailed comparisons between native and virtualized systems can be obtained by using experiment specific software and HEP specific analysis workflows.
- Example from CMSSW analysis job. Few percent performance reduction.

Journal of Physics: Conference Series, Volume 331, Issue 6 (2011)



different numbers of parallel running tasks on one host



different analysis steps

Cloud Security

- There are set of policies, technologies, and controls deployed to protect data, applications, and the associated infrastructure of cloud computing.
- There are different types of security issues faced by: cloud providers (organizations providing software-, platform-, or infrastructure-as-a-service via the cloud) and their customers.
- The security management addresses these issues with several security controls:
  - Deterrent Controls
  - Preventive Controlls
  - Corrective Controlls
  - Deterrive Controlls