HPC Cloud
A tool for research

Floris Sluiter
Project leader
SARA computing & networking services

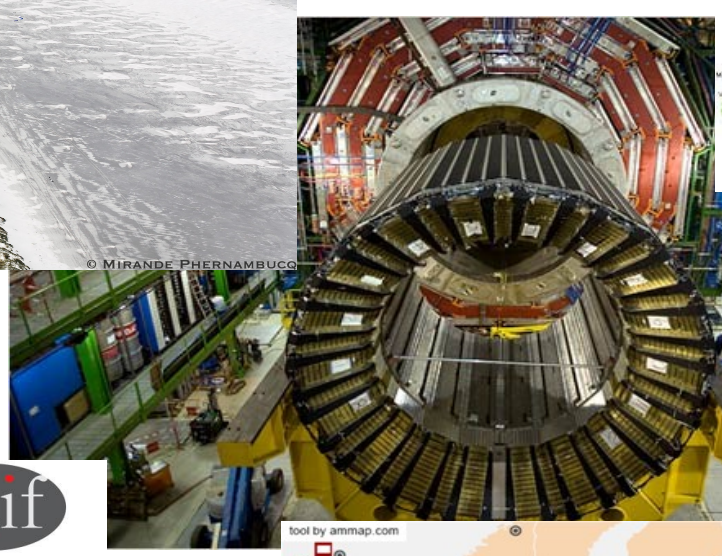
BiG Grid
the dutch e-science grid

SARA Project involvements

DEISA
DISTRIBUTED EUROPEAN INFRASTRUCTURE FOR SUPERCOMPUTING APPLICATIONS

vl·e

PRACE



LCG

egi European Grid Infrastructure
Towards a sustainable grid infrastructure

AMSTERDAM LIGHTHOUSE

CINEGRID AMSTERDAM

glif
Global Lambda Integrated Facility

EPIC

HPDMnet
High Performance Digital Media Network

BBMRI
Biobanking and Biomolecular Resources Research Infrastructure

HPC-Europa2
Pan-European Research Infrastructure on High Performance Computing

GÉANT

OptIPuter

nbic

BiG Grid
the dutch e-science grid

HPC Cloud
Philosophy

HPC Cloud Computing:

Self Service Dynamically Scalable Computing Facilities

Cloud computing is not about new technology, it is about new uses of technology

BiG Grid
the dutch e-science grid

(HPC) Cloud Why?

World

- better utilization for infrastructure
- "Green IT" (power off under-utilization)
- easy management

BiGGrid

- HPC cloud for academic world
- Free choice OS & software environment
- locked software can be used
- easy management

Massive interest and multiple early adopters prove the need for an academic HPC Cloud environment.

- beta-cloud is running “production”
- Popular with “non-HEP” (bio informatics, Psychology, Economics, linguistics, etc)

BiG Grid
the dutch e-science grid

HPC Cloud: Concepts

Clone my laptop!!

HPC cloud

One Environment,
Same image

Broom closet
cluster

Images:
- Software
- Libraries
- Batch systeem
-

Laptop

- HPC Hardware
- No overcommitting
  (reserved resources)
- Secured environment and network
- User is able to fully control their resource
  (VM start, stop, OS, applications,
  resource allocation)
- Develop together with users

BiG Grid
the dutch e-science grid

Our starting point for BiG Grid HPC Cloud

- Easy & standard(familiar) access protocol
  - name&password (or x509 certificates)
  - Support ad hoc collaborations
  - Support Cloud standards (OCCI, OVF, CDMI, WebdDAV)
- Zero client software install
  - Standard browser with java applets & javascript enabled
  - Additional tools optional: VNC viewer, ssh/putty etc
- User has free choice
  - Operating System & applications
  - Root rights in VM and on private network
  - Configuration of private cluster
  - Anything goes: Multi core, multi node, long running (services, databases)
- It doesn't have to be optimal, great is good enough
  - Virtualization overhead acceptible, only thousands of users not millions , only terabytes not petabytes

BiG Grid
the dutch e-science grid

…At AMAZON?

- Cheap?
  - Quadruple Extra Large = 8cores and 64Gb ram: $2.00/h (or $5300/y + $0.68/h)
  - 1024 cores = $2.242.560/y (or $678k + $760k = $1.4M/y)

- Bandwidth = pay extra
- Storage = pay extra
- I/O guarantees?
- Support?
- Secure (no analysis/forensics)?
- **High Performance Computing??**

BiG Grid
the dutch e-science grid

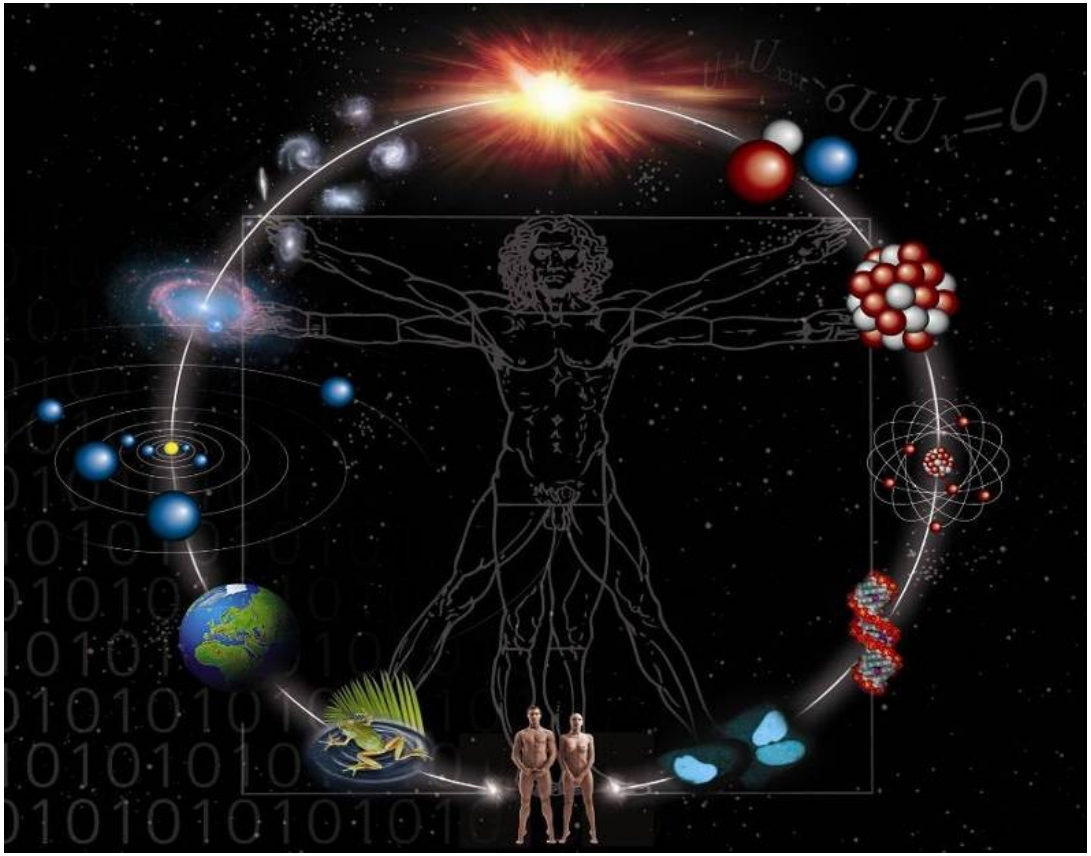
What is needed to create a successful HPC Cloud?

BiG Grid
the dutch e-science grid

Users of Scientific Computing

- High Energy Physics
- Atomic and molecular physics (DNA);
- Life sciences (cell biology);
- Human interaction (all human sciences from linguistics to even phobia studies)

- from the big bang;
- to astronomy;
- science of the solar system;
- earth (climate and geophysics);
- into life and biodiversity.



Slide courtesy of prof. F. Linde, Nikhef

BiG Grid
the dutch e-science grid

Users in pilot and beta phase

- From the start at least 50% in use
- Currently between 70-80%

- 50 user groups
  - 30 % from lifesciences (bio-informatics)
  - Psychology
  - Geography
  - Linguistics
  - Econometrists
- Currently 19 requests on waitinglist (!)
- Festive Launch at 4 th October in Amsterdam (www.sara.nl → Agenda)

BiG Grid
the dutch e-science grid

HPC (Cloud) Application types

Type	Examples	Requirements
Compute Intensive	Monte Carlo simulations and parameter optimizations, etc	CPU Cycles
Data intensive	Signal/Image processing in Astronomy, Remote Sensing, Medical Imaging, DNA matching, Pattern matching, etc	I/O to data (SAN File Servers)
Communication intensive	Particle Physics, MPI, etc	Fast interconnect network
Memory intensive	DNA assembly, etc	Large (Shared) RAM
Continuous services	Databases, webservers, webservices	Dynamically scalable

BiG Grid
the dutch e-science grid

Application models

- Single node (remote desktop on HPC node)
- Pilot jobs
- Master with workers (standard cluster)
- Pipelines/workflows
  - example: MSWindows+Linux
- 24/7 Services that start workers
- User defined

BiG Grid
the dutch e-science grid

HPC Cloud trust (1/2)

Security is of major importance

- cloud user confidence
- infrastructure provider confidence

Protect

- the outside from the cloud users
- the cloud users from the outside
- the cloud users from each other

Not possible to protect the cloud user from himself

- user has full access/control/responsibility
  ex. virus research must be possible

BiG Grid
the dutch e-science grid

HPC Cloud trust (2/2)

- Use virtualization for separation
  - operational from user space
  - users from each other
  - Use Vlans per user to separate network traffic

- Firewall
  - fine-grained access rules (“closed port” policy),
  - Self service and dynamic configuration!
  - non-standard ports open on request only and between limited network ranges

- Monitor (public) network and other access points
  - Scanning of new virtual templates
    - catches initial problems, but once the VM is live...
  - Port scanning
    - catches well-known problems
  - State-full Package Inspection
    - random sample based

BiG Grid
the dutch e-science grid

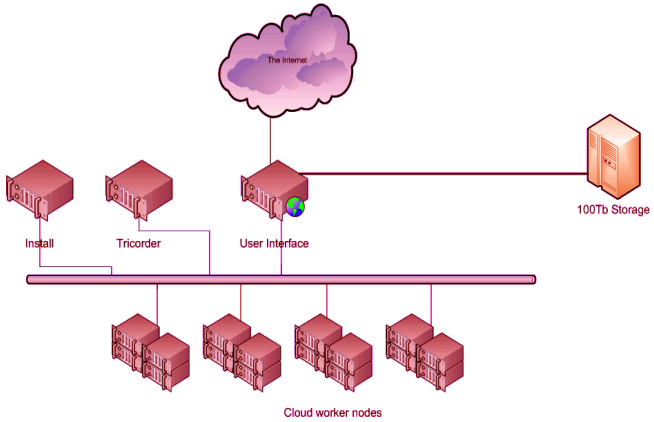
Open Cloud Standards (under construction)
Which ones are needed / Can be used?

Cloud object Type	To describe Configuration	To do Interaction / Change State and Content
Virtual Machine	**OVF** or CIM or Libvirt XML	**OCCI**, VNC, ssh
Storage Volumes, Data management	**CDMI**	WebDAV, NFS, Fuse
Network (VLAN,QOS, ACL&Firewall)	OVF + ??	??internal policy (no dynamic change)?? ??Programmable Network ??
Information on Capabilities (including AAA, quota, billing)	??	??RESTfull??
Information on state of Service and VMs	??CIM??	??RESTfull??

OCCI	http://occi-wg.org/	OCCI is a Protocol and API for all kinds of Management tasks.
CDMI	http://www.snia.org/cdmi	The Cloud Data Management Interface defines the functional interface that applications will use to create, retrieve, update and delete data elements from the Cloud. As part of this interface the client will be able to discover the capabilities of the cloud storage offering and use this interface to manage containers and the data that is placed in them. In addition, metadata can be set on containers and their contained data elements through this interface.
OVF	http://www.dmtf.org/standards/ovf	By packaging virtual appliances in OVF, ISVs can create a single, pre-packaged appliance that can run on customers' virtualization platforms of choice.
CIM	http://dmtf.org/standards/cim	CIM provides a common definition of management information for systems, networks, applications and services, and allows for vendor extensions.
Libvirt XML, WebDAV, NFS, Fuse, VNC, ssh	Industry standards	

The product: Virtual Private HPC Cluster

- We offer:
  - Fully configurable **HPC** Cluster (a cluster from scratch)
  - Fast CPU
  - Large Memory (256GB/32 cores)
  - High Bandwidth (10Gbit/s)
  - Large and fast storage (400Tbyte)
- Users will be **root** inside their own cluster
- **Free** choice of OS, etc
- And/Or **use** existing VMs: Examples, Templates, Clones of Laptop, Downloaded VMs, etc
- **Public** IP possible (subject to security scan)



Platform and tools:

- Redmine collaboration portal
- Custom GUI (Open Source)
- Open Nebula + custom add-ons
- CDMI storage interface

BiG Grid
the dutch e-science grid

HPC Cloud, what is it good for?

- Interactive applications
- High Memory, Large data
- Same data, many different applications
  (Cloud reduces porting efforts!)
- Dynamic, fast changing and complicated applications
- Clusters with Multi Operating Systems

- Collaboration
- Flexible and Versatile

- System architecture is expandable and scalable

BiG Grid
the dutch e-science grid

SNEAK PREVIEW
(What is an ideal system for an HPC Cloud)

BiG Grid
the dutch e-science grid

Calligo

"I make clouds"

19 Nodes:

- CPU Intel 2.13 GHz 32 cores (Xeon-E7 "Westmere-EX")
- RAM 256 Gbyte
- "Local disk" 10 Tbyte
- Ethernet 4*10GE

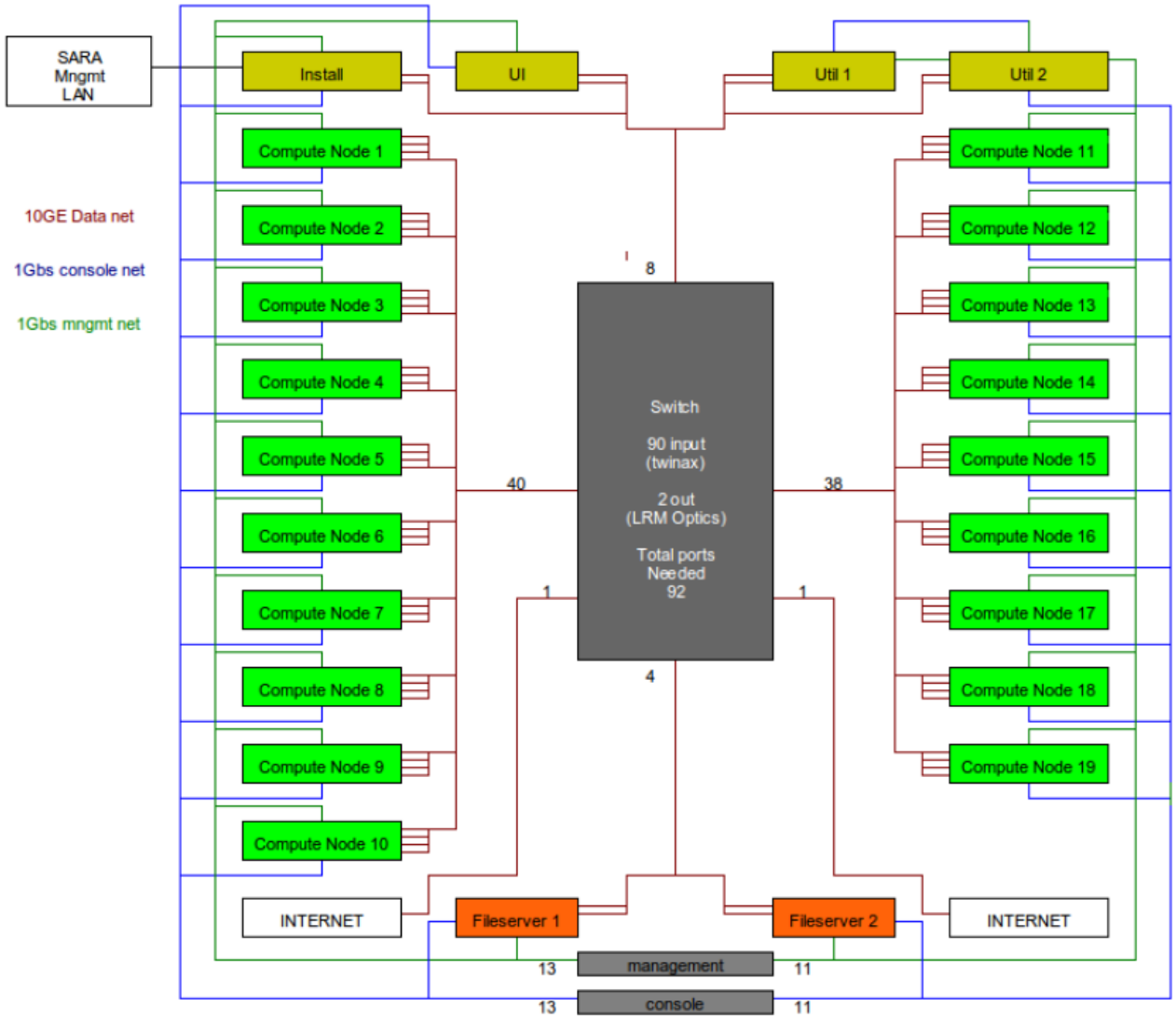
Total System

- 608 cores
- RAM 4,75TB
- 96 ports 10GE, 1-hop, non-blocking interconnect
- 400TB shared storage (ISCSI,NFS,CIFS,CDMI...)
- 11.5K specints / 5TFlops

Platform and tools:
Redmine collaboration portal
Custom GUI (Open Source)
Open Nebula + custom add-ons
CDMI storage interface

BiG Grid
the dutch e-science grid

Calligo, system architecture



BiG Grid
the dutch e-science grid

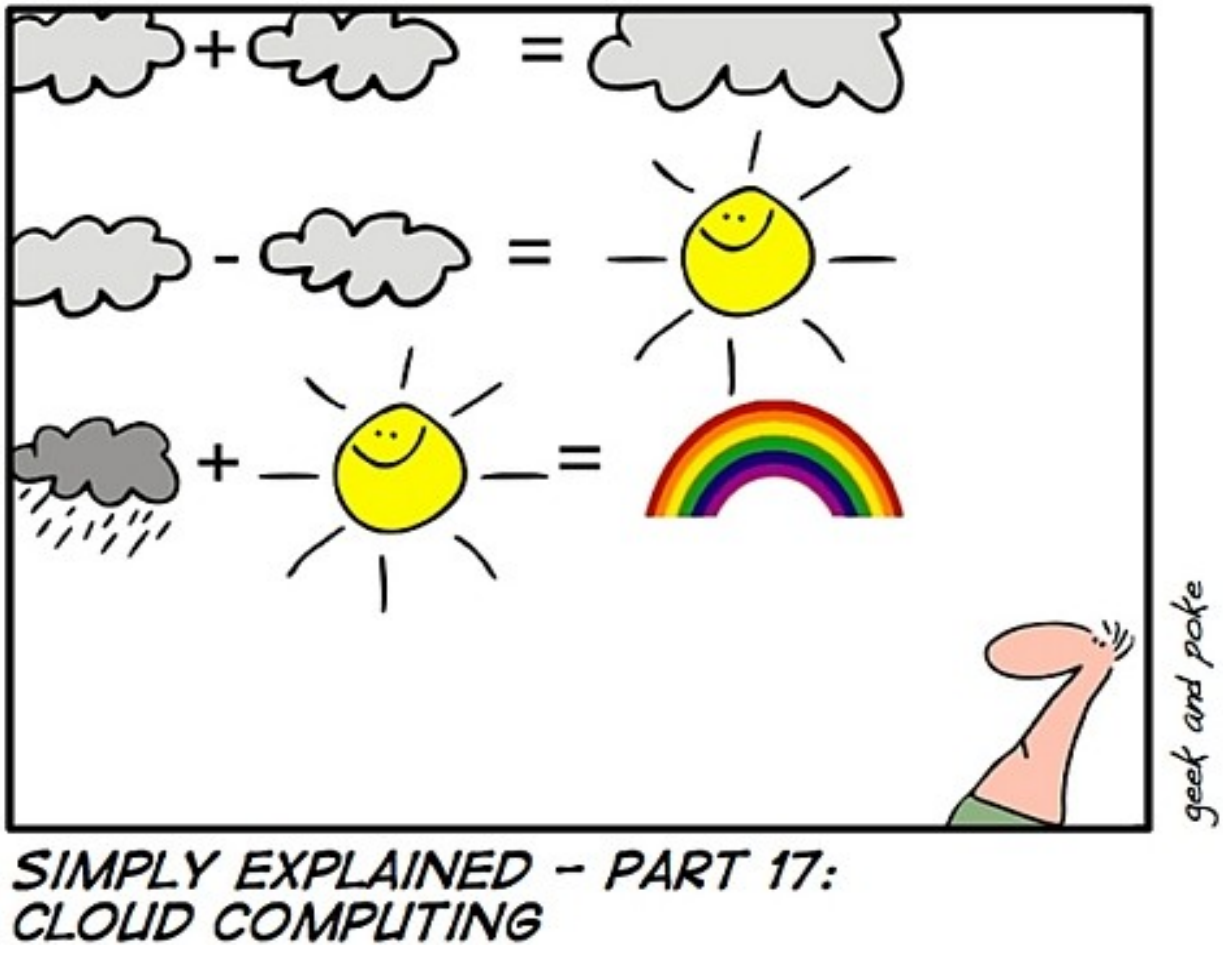
Real world network virtualization tests with qemu/KVM

- 20 gbit/s DDR infiniband (IPoIB) is compared with 1 Gbps Ethernet and 10 Gbps Ethernet
- Virtual network bridged to physical (needed for user separation)
- "real-world" tests performed on non optimized system
- Results
  - 1GE: 0,92 Gbps (1 Gbs)
  - IpoIB: 2,44 Gbps(20Gbs)
  - 10GE: 2,40 Gbps (10Gbs)
- Bottleneck: virtio driver
- Likely Solution: SRIOV

- Full report on www.cloud.sara.nl

BiG Grid
the dutch e-science grid

Thank you!



Questions?

www.cloud.sara.nl

photo: http://cloudappreciationsociety.org/

BiG Grid
the dutch e-science grid