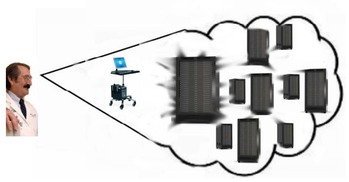
On-Demand Scientific Computing

ILLINOIS
UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
NM STATE UNIVERSITY
CALIFORNIA STATE UNIVERSITY SAN BERNARDINO

I. Fitting Virtualization

By carefully selecting how virtualization is applied, HPC systems can run more efficiently. Scientists perform supporting tasks, such as compiling and post-processing, which may not require a high performance computing system and may even interfere with running the simulations. Therefore, the hardware can be better used by moving the supporting tasks to a more appropriate, virtualized platform.



II. Benefits of Virtualization

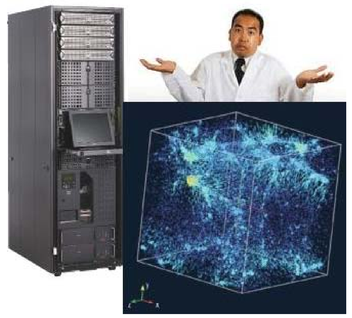
Virtualization allows consolidation of various supporting tasks and flexible allocation of resources. Virtualized systems also have reliability and maintenance advantages over physical systems.

Computations can run uninterrupted without functional interference by other computations and can even be paused or migrated. Furthermore, an on-demand virtualization service can provide personalized software stacks that fit a scientist's specific needs.

III. Project Objectives

- To provide a template of the current Tri-lab Operating System Stack (TOSS) and use it to automatically generate virtual machines for scientists.
- To provide the ability for small jobs such as pre-processing, post-processing and legacy serial applications to run on a virtualized system.
- To make available extra cycles that may enable a scientist to get their work done.
- To enable a program that does not support check pointing to run on an extended period of time.
- To have the ability to run scientific computations for one week instead of one day with full reliability.

IV. Science on Traditional HPC Systems



The traditional way to run scientific computations is to run on a physical cluster on one of the HPC systems. Many times, the HPC has an overload of jobs waiting.

For the scientist that wants to run a large job, they have to wait for other jobs to finish. Many of those jobs are small and could run somewhere else.

V. Science on Supplementary Virtualized Systems

Scientist running small jobs and users who don't have time constraints on their jobs can run on virtual systems.

Running suitable jobs on a separate virtual system reduces the load for the HPC, makes better utilized and produces a more reliable and enjoyable HPC experience for both the user and the administrator.

Los Alamos NATIONAL LABORATORY EST. 1943 | ISTI INFORMATION SCIENCE & TECHNOLOGY INSTITUTE | National Security Education Center A CONSORTIUM OF LANL INSTITUTES | New Mexico Consortium's IAS Institute for Advanced Studies at Los Alamos National Laboratory

VI. Scientific Computing on Demand

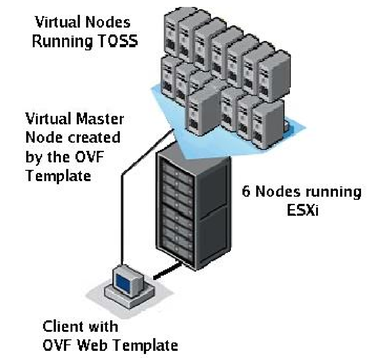
Here are the steps:

1) User signs into a web portal in the web browser
2) Chooses the cluster size (CPU, RAM) of the new system
3) System creates virtual cluster fast
4) User can now run jobs on their dedicated virtual cluster

VII. How Scientific Computing on Demand Work

Hardware For Virtualizing
- Powerful Clusters
- 10GigE Interconnect
- Terminal (Any PC)

Software For Virtualizing
- Tri-Lab Operating System Stack
- VM Ware Hypervisor
- Infrastructure On-Demand
- OVF Web Template



VIII. What is OVF?

OVF stands for Open Virtualization Format. It contains the metadata information of a virtual machine in XML format.

Appealing features of OVF for our project:

- **Supports multi virtual configurations**
- **Enables portable VM packing**
- **Offers future extensibility**

OVA (Open Virtual Appliance). It is a package that contains 3 files that describe a virtual machine:

Putting It Together and Packing It Up

- **.OVF file, descriptor file**
- **.MF file, optional manifest**
- **.VMDK file, virtual disk image**

Students: Travis Johnson | Gregorio Hinojos | Benjamin Cheng
Mentors: Andrew Shewmaker | Dane Gardner | Graham Van Heule | Andree Jacobson

LA-UR 10-05200