Nimbus
or
an Open Source Cloud Platform
or
the Best Open Source EC2 No
Money Can Buy ;-)
What Are Clouds?

Infrastructure-as-a-Service (IaaS)
- e.g., AWS, GoGrid, Flexiscale, Science Clouds

Platform-as-a-Service (PaaS)
- e.g., GoogleApp Engine

Software-as-a-Service (SaaS)
- e.g., Animoto
What is Nimbus?

- An extensible open source Infrastructure-as-a-Service implementation
  - Turns your cluster into a cloud
- Why open source IaaS?
  - **Experiment and use:** make your own cloud or configure a private cloud
  - **Customize:** try new things, make the IaaS paradigm work for your application domain
- Our particular interest in customization: scientific computing
Nimbus Features

- **Cloud computing infrastructure**
  - WSRF and EC2 interfaces
  - Xen implementation (KVM in preparation for release)
  - Launches flexibly defined groups of VMs and configures their networking

- **Can be configured to use familiar schedulers like PBS and SGE to manage VMs**
  - The workspace pilot

- **Launches self-configuring virtual clusters**
  - The context broker

- **Defines an extensible architecture**
  - And has been extended by various projects

**News Flash:** Ian Gable from UVIC becomes a Nimbus committer!
5-click Guide to Nimbus

- storage service
- workspace service
- IaaS gateway
- workspace resource manager
- workspace pilot
- workspace control
- potentially other providers
- EC2
- context broker
- context client
- workspace client
- cloud client
Science Clouds

News Flash: the Masaryk cloud officially available online today!
Science Clouds: Goals

- Make it easy for scientific projects to experiment with cloud computing
  - Can cloud computing be used for science?
- Customize software in response to the needs of scientific projects
  - Start with EC2-like functionality and evolve to serve scientific projects: virtual clusters, diverse resource leases
  - Federating clouds: moving between cloud resources in academic and commercial space
Who Runs on Science Clouds?

Project diversity: Science, CS, education, build&test…
CS research: investigate latency-sensitive apps, e.g. hadoop
Need access to distributed resources, and high level of privilege to run a ViNE router
Virtual workspace: ViNE router + application VMs
Paper: “CloudBLAST: Combining MapReduce and Virtualization on Distributed Resources for Bioinformatics Applications” by Andréa Matsunaga, Maurício Tsugawa and José Fortes, accepted to eScience 2008.
• STAR: a high-energy physics experiment
• Needs resources **with the right configuration**
  ◦ Complex environments: correct versions of operating systems, libraries, tools, etc all have to be installed.
  ◦ Consistent environments: require validation
• Solution: deploy a virtual OSG STAR cluster
  ◦ OSG CE (headnode) plus STAR worker nodes: SL4 + STAR conf
• Requirements
  ◦ One-click virtual cluster deployment
  ◦ Migration: Science Clouds -> EC2
    • Last September: EC2 runs of up to 100 nodes (production scale, non-critical codes)
    • Testing for critical production deployment
• Work by Jerome Lauret, Doug Olson, Leve Hajdu, Lidia Didenko
Infrastructure Testing

- Project: tests of various Globus components on different platforms for correctness and scalability
- Need short-term but flexible access to diverse platforms
- Workspaces: Globus 101 + others
- Work by various members of the Globus Toolkit (Tom Howe, John Bresnahan, Stu Martin, Martin Feller)
- Resulted in provisioning a private Nimbus cloud for Globus
Alice HEP Experiment at CERN

Collaboration with the CERNVM project
Other Projects

- Evaluating a cloud from user’s perspective

- Economics
What Our Users Are Saying

Our Guestbook: see what users are saying!

"Xen was the obvious choice for the virtualization part, but it took me a fair amount of time to find a piece of software to manage the deployment of the VMs. I came across Globus virtual workspaces (Nimbus) and the Globus Toolkit. This was by far the best solution for deploying and managing the developer environments."

Scott Haskell, AdBrite.com, October 29, 2008

"**VERY COOL** .. you have done really nice work."

Nicholas Karonis, Northern Illinois University, October 20, 2008

"I've been around HPC for 10 years and this is without question the coolest thing I've seen in a long time!"

Ron Price, The University of Utah, October 16, 2008
Getting Started

- Available to scientific projects
- Send us mail to get access
- Do the quickstart: you should be running in 15 minutes!
- Move on to create virtual clusters and virtual Grids
The Last Slide

- Nimbus is an extensible, easy-to-use, open source tool for configuring clouds
- **What has our impact been?**
  - Utilization, time used per project, etc.
  - Scientific results, papers written and in preparation, ongoing discussions
- **Another kind of impact: we are doing things we could not do before**
  - Deploying network routers on remote platforms
  - Easily finding the right environment in distributed environment
  - Provisioning resource when we need them
- **We’re learning what’s possible**