

Clouds of Grids or Grids of Clouds, identifying the suitability of each for different types of task

2nd December, 2009

David Wallom
University of Oxford
OGF-Europe
VP e-Research, OGF, www.ogf.org

Overview



- OGF
- Grids & Clouds
- Current patterns and standards

The OGF Mission Statement



" OGF is an open community committed to driving the rapid evolution and adoption of **applied distributed computing**. Applied Distributed Computing is critical to developing new, innovative and scalable applications and infrastructures that are essential to productivity in the enterprise and within the science community."

OGF Strategy & Activities



- Strategy
 - Proactively lead the development of new distributed computing technologies relevant to our stakeholders
 - Engage with all relevant organizations and user groups working to the same goals
- Current OGF Strategic Activities
 - **Grids and Clouds**
 - Green IT
 - e-Infrastructure requirements
 - Production grid interoperability
 - Application domains and utilization of DCI

How Do Grids and Clouds Relate?



- Grids came from “big science” and the desire to collaborate in a *federated environment*
 - Manage sharing of resources
 - New technologies developed to cope
- Clouds are coming from industry and the desire to dynamically provision resources *in the cloud*
 - Simple APIs for using abstracted or virtualized resources
 - Already using existing technologies
 - Economies of scale in the data center
 - Aka, utility computing, internet computing, ...
- “Grids are an access model; Clouds are a business model”
 - Chris Smith, Platform Computing, OGF VP Standards
- Distributed applications need and can use capabilities being developed under both *grid* and *cloud*
 - There is no real *grid vs. cloud dichotomy*

What is the Cloud?

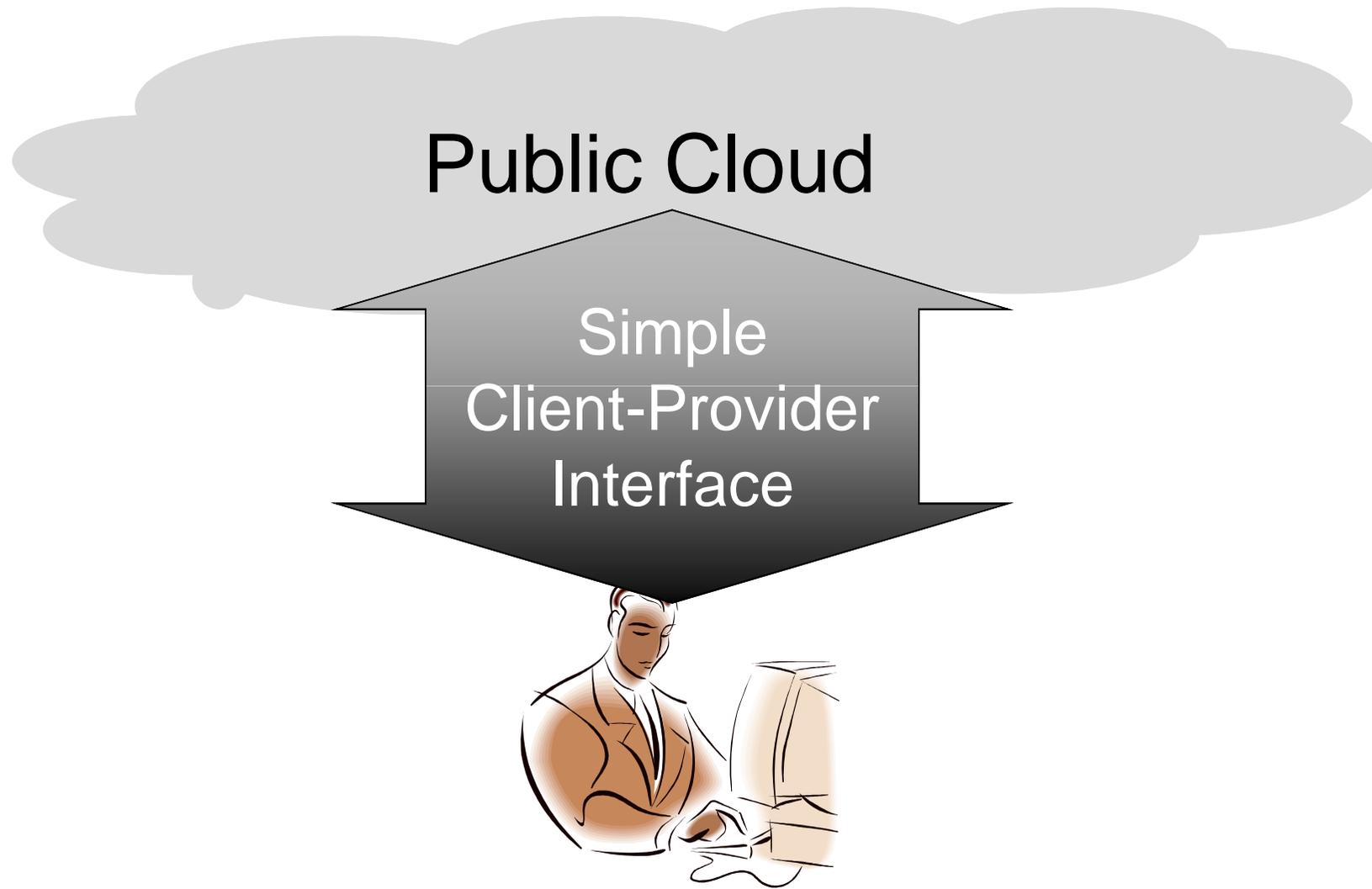
Typically three types are described

- Infrastructure as a Service (IaaS)
 - provide (virtualized) resources on demand
 - Amazon AWS, Rightscale, GoGrid
- Platform as a Service (PaaS)
 - build applications using a provided toolkit so that the application can be run in the Cloud
 - Google App Engine, Microsoft Azure
- Software as a Service (SaaS)
 - offer an entire application “in the cloud”
 - Salesforce.com

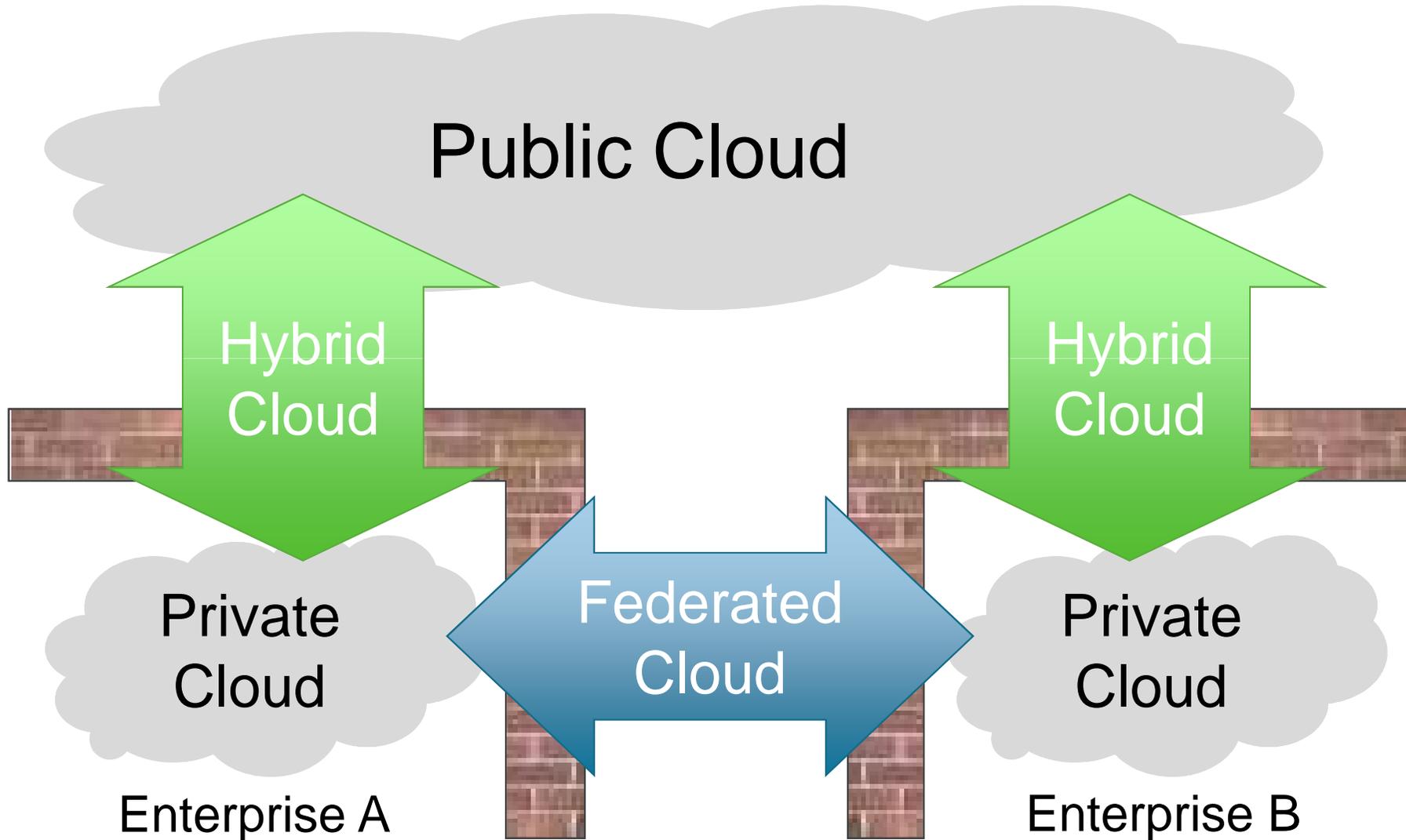
Is Cloud Disruptive?

- Definitely! But not a technology disruption. It's a disruptive business model.
- Cloud is characterized by services made available on-demand, and payment for these services based on actual usage.
 - no up front contracts
- So a disruptive use of technology, but the technologies used to build Clouds existed already, and are well understood.

Deployment Models Blur the Boundaries of Grids and Clouds



Deployment Models Blur the Boundaries of Grids and Clouds



Cloud Issues...



- Cost & Cost Predictability
- Users expect to monitor & manage "their" infrastructure
 - Will a public cloud provider expose enough information for a client to troubleshoot when something goes wrong?
- Security & Privacy
 - You can store encrypted data in a clouds, but can you compute on it?
- Regulation
 - Physical location of data
 - Long-term audit trails (15-20 years)
- Individual vs. Corporate Requirements
 - Corporate use of public clouds may entail legal & contracting overheads
 - Ease of use and quick provisioning may tempt individuals to ignore corporate procedures
 - Trade-off between quick results and longer term risk exposure
- Internal IT departments may want to offer their own "seed cloud"
- *Interoperability & portability* between private and public clouds

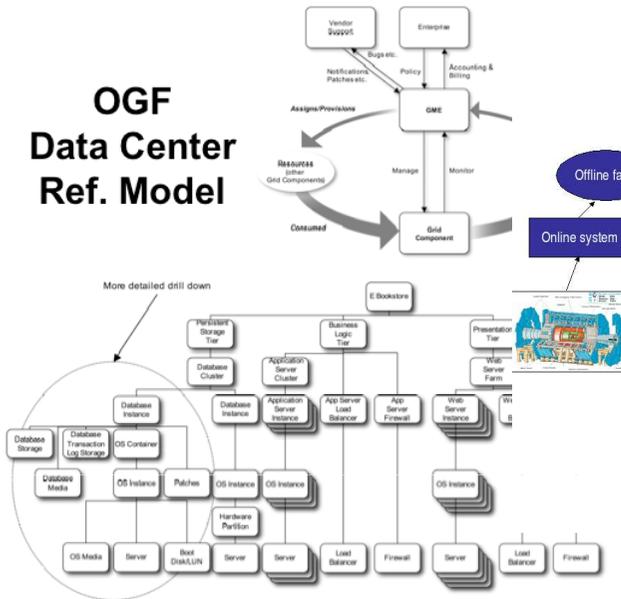
Distributed Computing Patterns



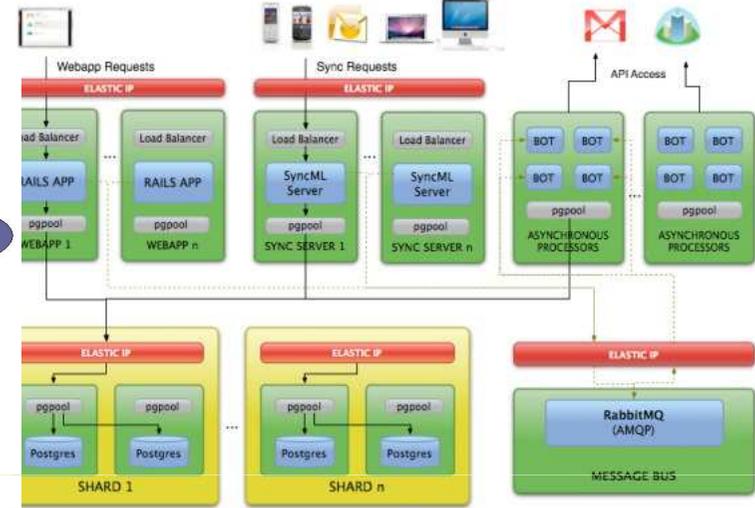
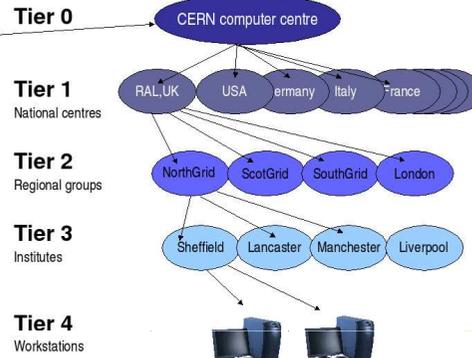
- “It’s déjà vu all over again”
 - The lessons learned from solving previous problems in Grids can be applied to Cloud computing (e.g. identity federation)
- Use the well known concept of “design patterns” to help capture distributed computing best practices

Patterns in Practice

OGF Data Center Ref. Model



Tier Structure



HDFS Architecture

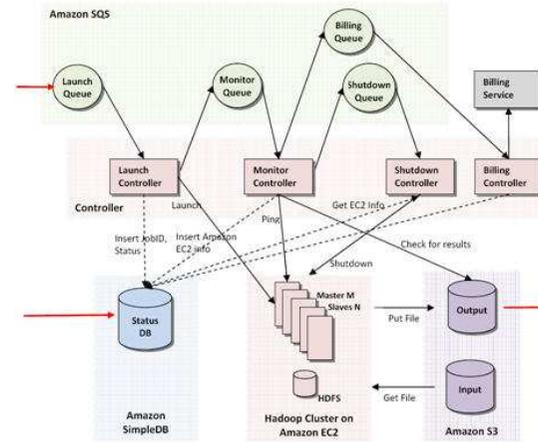
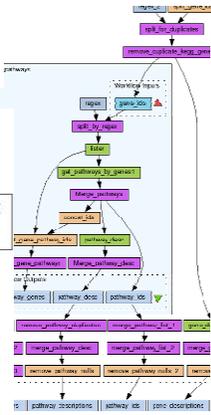
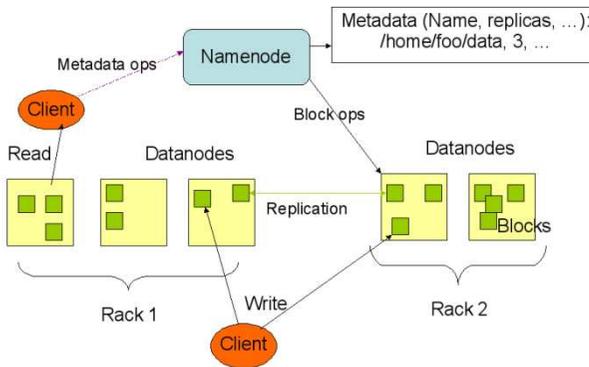
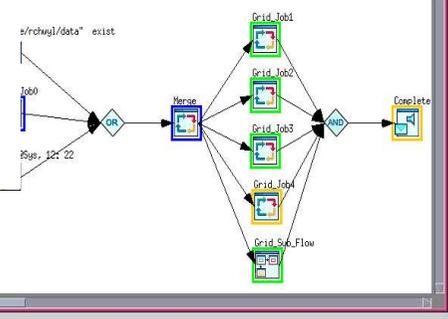
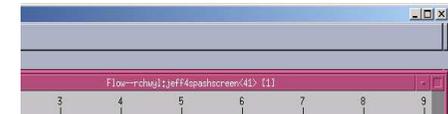


Figure 4: GrepTheWeb Architecture - Zoom Level 3



Is this Grid or Is this Cloud?



- These patterns help us identify whether an application is better suited to Grid or to Cloud
 - Scientific workflows – typically can be run on Grids
 - Service Oriented Application – very suitable for Cloud
 - Single instrument data acquisition – probably doesn't distribute very well

No Shortage of Challenges Remain



- For each domain the following considerations must be made;
 - Data access and interoperability
 - Must be done at the application domain level, by the domain users
 - Security
 - Different models will expose different security threats
 - Reliability
 - Managing redundancy, live migration, etc., across the infrastructure
 - Frameworks
 - How to manage sets of resources, e.g., VMs and VOs?
 - Performance management
 - What job mix needs to be supported, e.g., e-commerce, HPC, transactional, database, data streaming?
 - Costing models
 - How to compare your own infrastructure costs with a cloud providers?

Summary



- The OGF is bringing its deep experience in building and using Grids to help address the challenges of building and using Clouds
- Need to have a set of well documented patterns to assist new communities
- Wide spectrum of applications can use the capabilities being developed under the categories of grid and cloud
- Ongoing standards activities will facilitate interoperability among public, private, hybrid and federated clouds as both IaaS and PaaS
- Complaints about grid have been primarily about its complexity, we must ensure that Clouds are treated seriously as well, every IaaS owner is now a systems administrator, not just an application user
- Having detailed knowledge of the capabilities of cloud/grid/HPC will allow us to ensure that the right applications end up using the right Distributed Computing Infrastructure

Upcoming OGF Events!

- OGF28
 - Munich, Germany
 - March 15th – 19th, 2010

