Grille et Cloud...
Ou en sommes nous?

Marc-Elian Bégin

Six² Sàrl

JI08

Obernai, France, Septembre 2008

Six2
Software
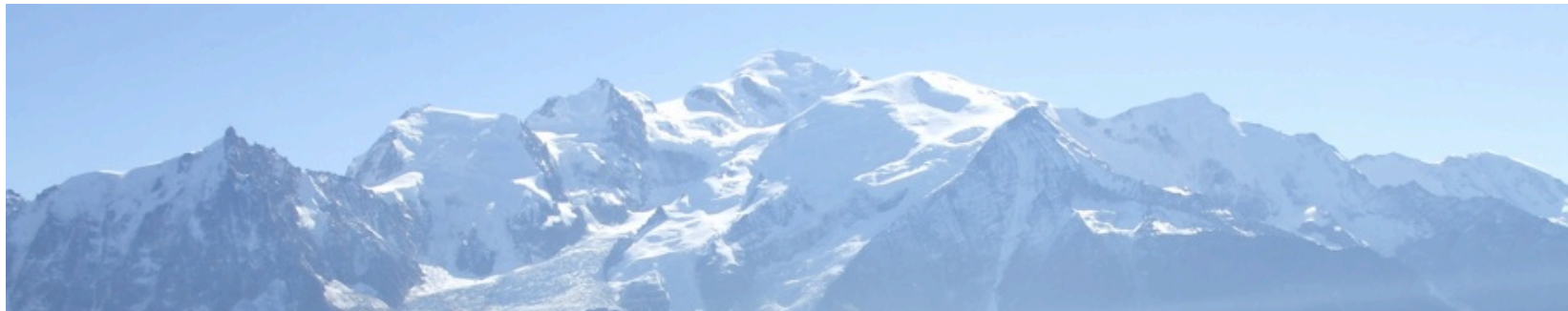Integration
eXperts

Content

- **Background and Motivation**
- **Grid: Unique Features**
- **Cloud: Unique Features**
- **Our Experience with the Cloud**
- **Build the Grid on the Cloud !**
- **Conclusions**

Who we are

- **Six² Sàrl (limited)**
- **Based in Geneva, Switzerland**

- **Agile tools and application for automated build, integration, deployment and test of distributed software development**
- **Agile software consultancy and services**

Photo by: Taraskas

Six2 Software Integration eXperts

Background and Motivation

- **Grid and cloud comparative study**
  - "An EGEE Comparative study: Grids and Clouds-evolution or revolution?", by Marc-Elian Bégin

- **Six² builds agile tools for software development**
  - Deployed as Software as a Service (SaaS)
  - We need to transparently deploy on
    - Public or "outer" clouds (e.g. Amazon AWS)
    - Private or "inner" clouds
  - Need open source cloud distribution, following the Amazon cloud model (EC2 and S3)

Six2 Software Integration eXperts

Grid: Unique Features

- **What makes the Grid unique and valuable?**
- **Collaboration**
  - Platform for dynamic and distributed collaborations: Virtual Organization
- **Federation**
  - Platform for federating existing resources into a whole
  - Together offers more than the sum of its parts

Six2 *Software Integration eXperts*

eGee
Enabling Grids
for E-sciencE

Scheduled = 21539
Running = 25374

Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
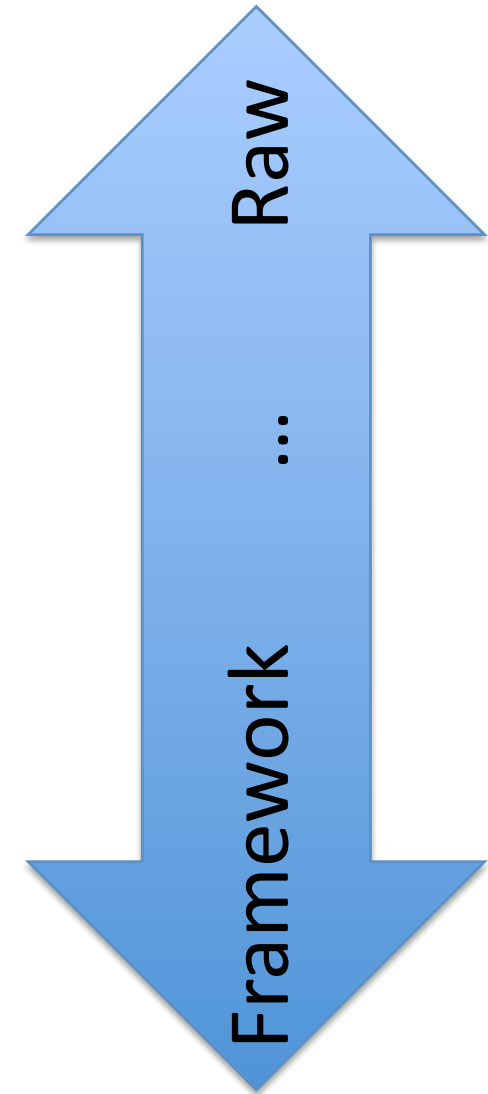Material Sciences
…

21:13:50 UTC
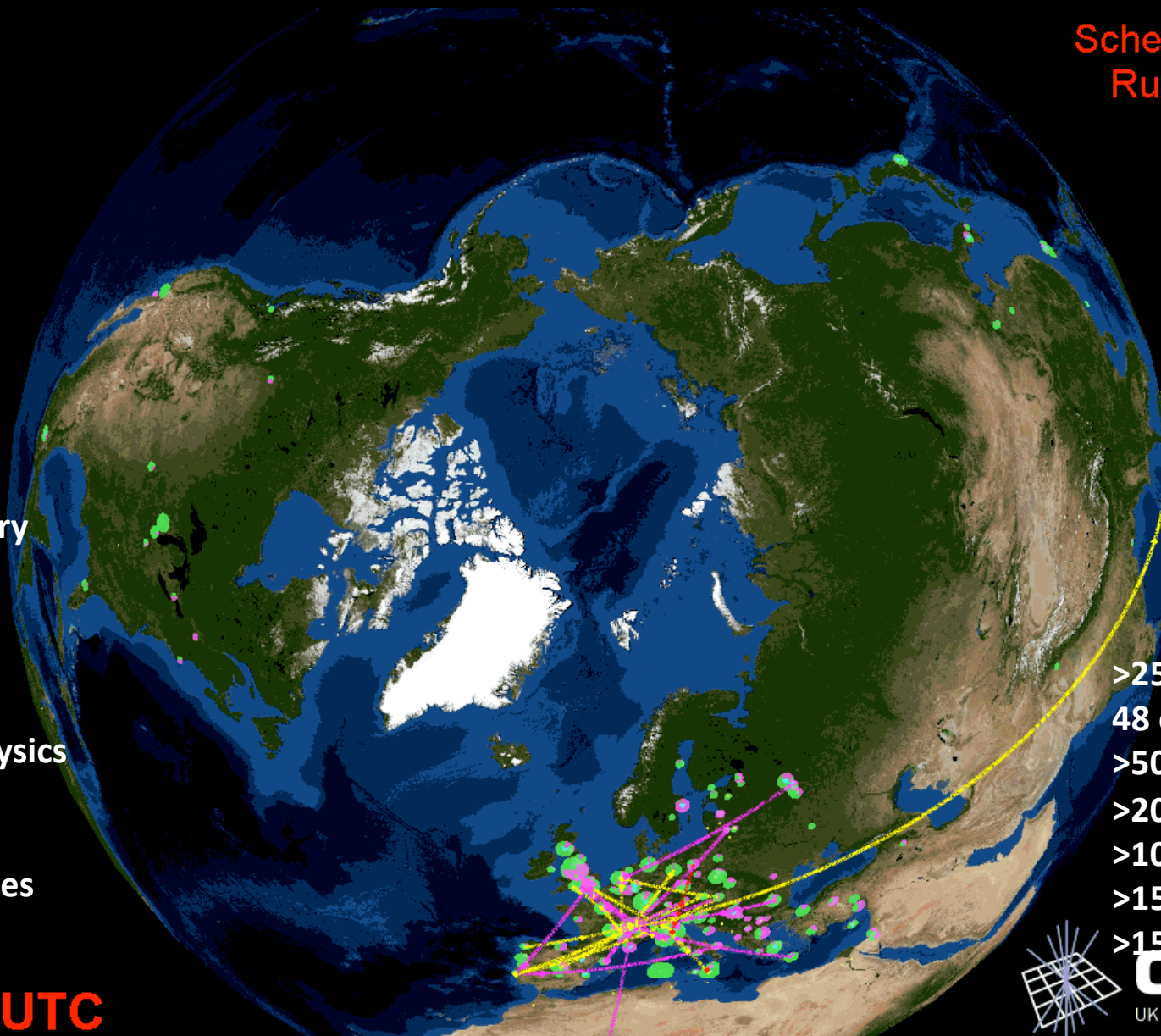
>250 sites
48 countries
>50,000 CPUs
>20 PetaBytes
>10,000 users
>150 VOs
>150,000 jobs/day

GridPP
UK Computing for Particle Physics

Six2 Software Integration eXperts

Different Clouds

- **Amazon Web Services**
  - Elastic Computing Cloud (EC2)
  - Simple Storage Service (S3)

- **AppNexus**

- **GoGrid**

- **Google App Engine**

Raw

...

Framework

Cloud: Unique Features

- **Ease of use**
  - REST and HTTP(S)

- **Runtime environment**
  - Hardware virtualisation
  - Gives users full control

- **Elasticity**
  - Pay-as-you-go
  - Cloud providers can buy hardware faster than you!

Cloud: Amazon Web Services

- **EC2 (Elastic Computing Cloud) is the computing service of Amazon**
  - Based on hardware virtualisation (Xen)
  - Users request virtual machine instances, pointing to an image (public or private) stored in S3
  - Users have full control over each instance (e.g. access as root, if required)
  - Requests can be issued via SOAP and REST

Cloud: Amazon Web Services

- **S3 (Simple Storage Service) is a service for storing and accessing data on the Amazon cloud**
  - From a user's point-of-view, S3 is independent from the other Amazon services
  - Data is built in a hierarchical fashion, grouped in buckets (i.e. containers) and objects
  - Data is accessible via SOAP, REST and BitTorrent
- **Elastic Block Storage**
  - Locally mounted storage
  - Highly available
  - Possible to breakpoint to S3

Six2 Software Integration eXperts

Cloud: Amazon Web Services

- **Other AWS services:**
  - SQS (Simple Queue Service)
  - SimpleDB
  - Billing services: DevPay
  - Elastic IP (Static IPs for Dynamic Cloud Computing)
  - Multiple Locations
  - SLAs

Performance

- ## **EC2, S3 bandwidth performance summary**

Test type	Transfer (MB/sec)	Remarks
EC2 -> EC2	75.0	Using curl on 1-2 GB files, without SSL
S3 -> EC2	49.8	Using 8 x curl on 1 GB files, with SSL
	51.5	Using 8 x curl on 1 GB files, without SSL
EC2 -> S3	53.8	Using 12 x curl on 1 GB files, with SSL

- ## **The conclusions from [6] regarding the EC2 -> EC2 transfers are that "basically we're getting a full gigabit between the instances".**

Six2 Software Integration eXperts

Why Amazon Started This

- **They started this for themselves**
- **Virtualise their own infrastructure**
  - Separate the infrastructure from the "user runtime environment"
- **Optimise resource utilisation**
- **Provide better support to their own developers**
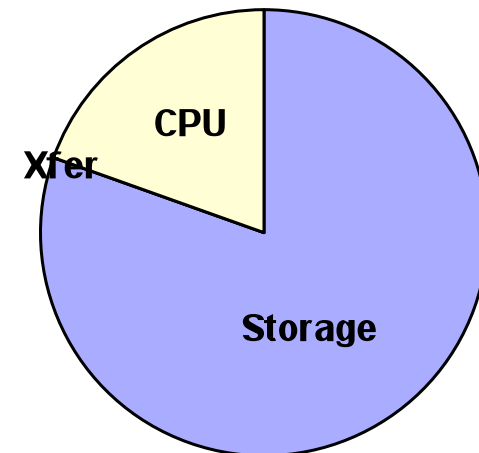- **Be more agile in their development**

Do This At Home

- **Time for a quick demo…**

Six2 Software Integration eXperts

Costs: EGEE workload in 2007

Data:

25PB stored

11PB transferred
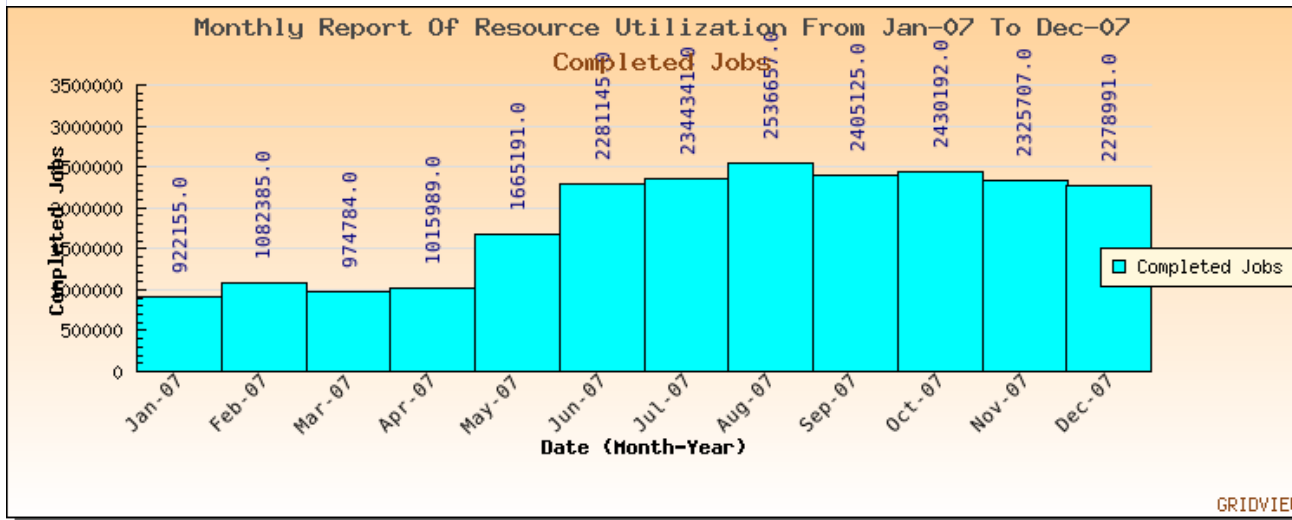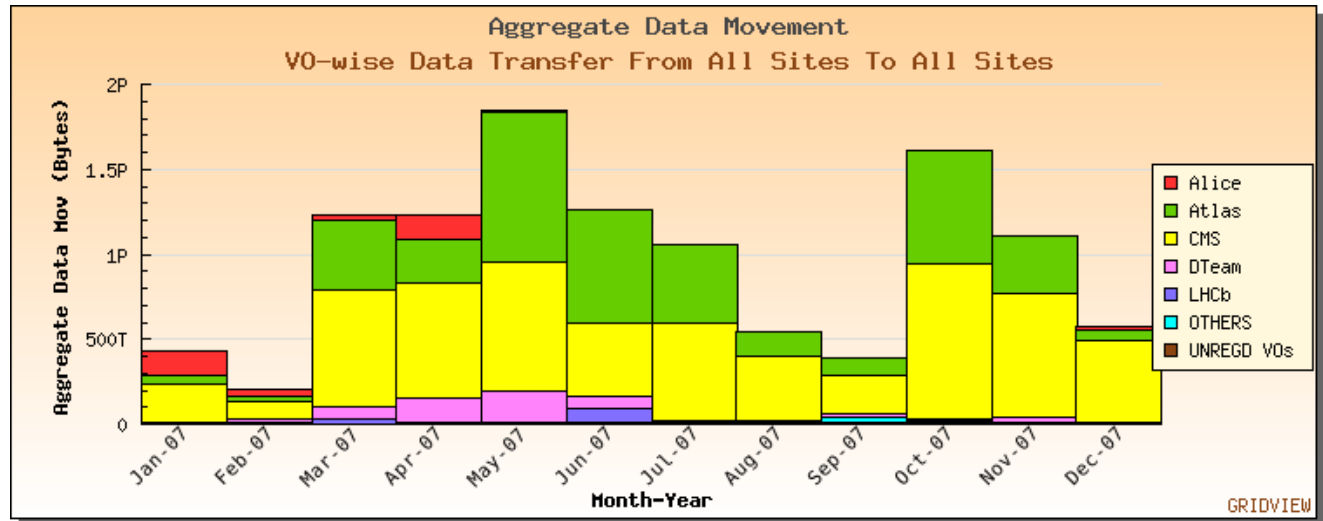
CPU: 114 Million hours

Estimated cost if performed with Amazon's EC2 and S3: ~38 M€

http://gridview.cern.ch/GRIDVIEW/same_index.php   http://calculator.s3.amazonaws.com/calc5.html? 17/05/08 $58688679.08

Six2 Software Integration eXperts

Ease of Use

- **Key to the success of AWS is the choice of technologies**
- **Hardware virtualisation (Xen based)**
- **HTTP(S)/REST**
  - RESTful Web Services
  - Resource Oriented Architecture (ROA)
  - Everybody has an HTTP client on their machine (if not many)
  - Language of the Web

Ease of Use

- **This backs-up the claim from Amazon that AWS requires "no middleware" (for the user!)**
- **However, the level of service provided by AWS is lower than EGEE**
- **For EGEE/gLite, several MB are required to use the grid**

Six²'s Experience with the Cloud

- **²One Build and Test Application build on the Cloud**
- **Goal: "*SaaS for software stakeholders to automatically, and continuously, build, integrate, deploy and test distributed software systems*"**
- **Benefit:**
  - Maintain at all time the ability to release
  - Reduce risk: minimises integration and deployment risk at the end of a development cycle
  - Reduce time to market: can go from idea to product faster since production pipeline automated
- **This means**
  - Provide hardware on-demand
  - Let users define their own runtime environment
  - Give users full control on their build and test environment
  - … in other words provide them a **Cloud application**!

²One Process Overview

²One process
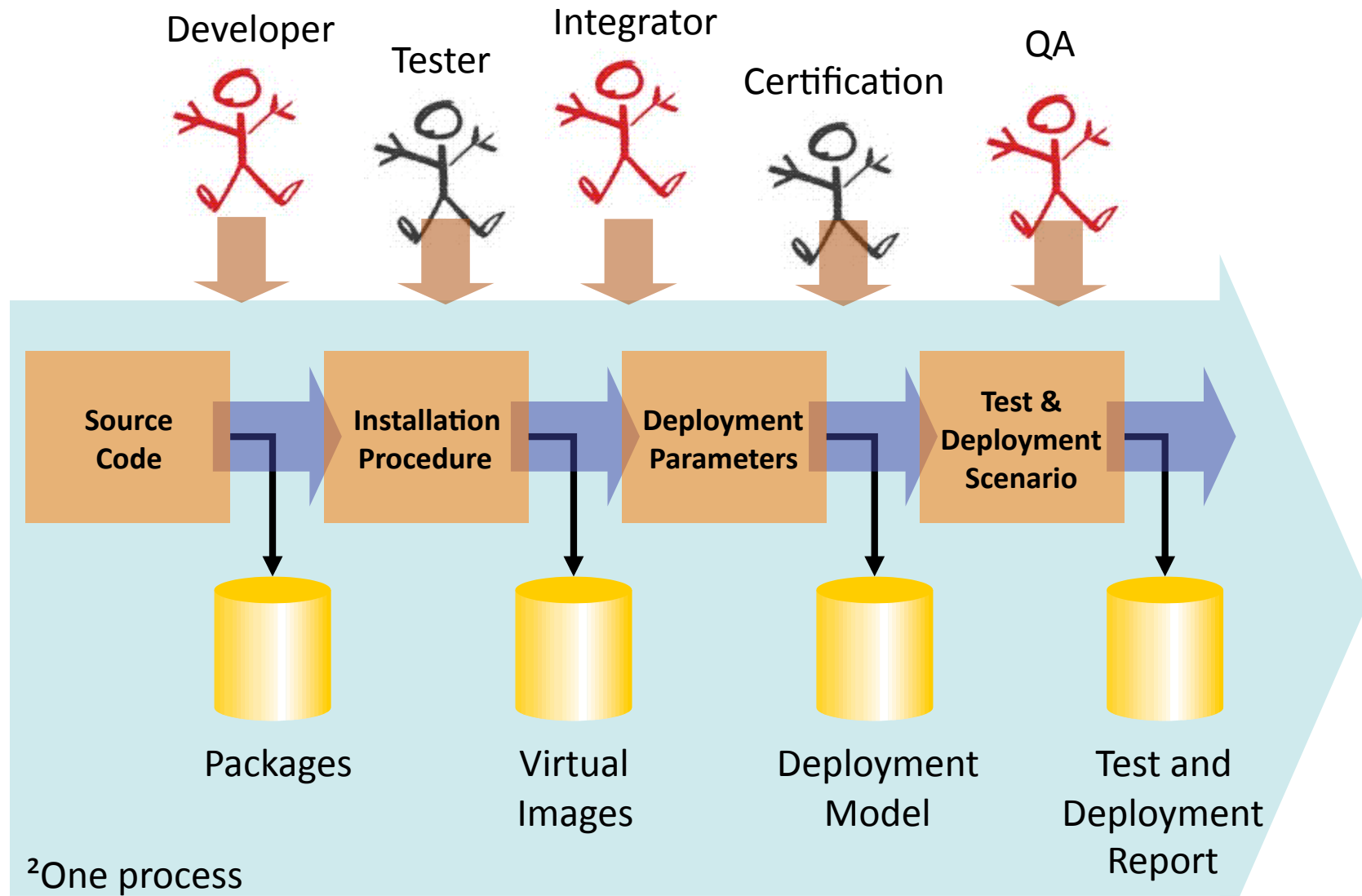
Deployment in the Cloud

Deployments are isolated

Resources only used when needed

Six2 Software Integration eXperts

How is the Cloud Enabling ²One?

- **Virtualisation provides separation between infrastructure and user runtime environment**
- **Users specify virtual images as their deployment building blocks (generated by ²One or not)**
- **Pay-as-you-go allows users to use the service when they want and only pay for what they use**
- **Elasticity of the cloud allows users to start simple and explore more complex deployment over time**
- **Simple interface allows easy integration with existing systems (no need to take over the world)**

Combining Technologies

- **Cloud is not only a "commercial offering"**
- **Cloud is especially a set of technological choices**
- **Cloud is not grid**
- **Cloud is a subset of grid**
- **But it does a better job at it!**

Six2 *Software Integration eXperts*

Combining Technologies

- **Grids should be built on clouds !!**
  - Keep the Grid promise
  - Enable more and stronger collaboration
  - Simplify site management
  - Improve user experience
- **Advantages**
  - Provide a clean separation between infrastructure and user runtime environment
  - Provide a uniform user runtime environment
  - Enable labs to offload peak demand to commercial sites ("inter" cloud)
- **Disadvantage**
  - We're back on the learning curve!

Cloud Standards

- **Need for standard?**
- **Yes but 'defacto' standards already exist, e.g.:**
  - EC2 and S3 for service interface and semantic (REST)
  - HTTP(S)
  - Virtualization formats
    - Xen or KVM for image virtualisation format
- **Since the cloud is simple and based on several existing standards we're not stuck waiting for standardisation bodies**

Build the Grid on the Cloud !

- **Obstacle**
  - Middleware must be
    - Simple to operate
    - Robust and resilient
    - Adapt to existing infrastructures
  - Open source
- **What's missing?**
  - Opportunity to demonstrate these concepts in a grid context
  - Open source cloud distribution
  - Partners willing to get it done!

Six2 *Software Integration eXperts*

Conclusions

- **Grid / Cloud fusion makes sense**
  - Cloud computing is getting traction, especially with Amazon Web Services (AWS) commercial offering
  - Grid (e.g. EGEE) has a larger scope than cloud, however, technological choices and simple interfaces of clouds like AWS is relevant to the grid world
- **What makes cloud services like AWS successful**
  - Hardware virtualisation
  - REST and HTTP(S)
  - These technologies should displace existing grid middleware

  **Grid made the cloud possible…
  it's part of the same "next big thing"!!**

References

- **Cloud Webcasts: http://tech.slashdot.org/article.pl?sid=08/07/23/1853218**

- **Cloud / Grid Position Paper: https://edms.cern.ch/file/925013/3/EGEE-Grid-Cloud.pdf**

- **Book: RESTful Web Services, Richardson, Ruby**

- **Java RESTful Web Service framework: restlet (restlet.org)**

Merci

☺

Six² Sàrl, Geneva, Switzerland

www.sixsq.com

Questions, comments, feedback:

meb@sixsq.com