The Evolution of e-Infrastructures in Research, Industry, and Education

Wolfgang Gentzsch
The DEISA Project & Board of Directors of OGF
gentzsch at rzg.mpg.de

“It’s hard to make predictions, especially about the future”

Yogi Berra
Service Infrastructures: nothing new
Building Blocks: Computers, Grids, Clouds
Example: DEISA Ecosystem for Science Applications
Next-Generation e-Infrastructures
- e-Learning on e-Infrastructures = e-Learning 2.0
Service Infrastructures

Ancient Rome: ten aqueducts, some 150,000 m³ of water each day

Electrical Power Grid Infrastructure

EGEE – Enabling Grids for E-Science
HPC Centers

- HPC Centers are **service providers**, for past 35 years
- Computing, storage, applications, data, etc IT services
- Serve (local) research, education, and industry
- Very professional: to end-users, they appear almost as Cloud services (AWS Cloud definition: easy, secure, flexible, on demand, pay per use, self serve)
- But: no virtualization, semi-automatic, static
- They could become a Cloud customer for dynamic scaling and adopting to changing business and user demands
RoadRunner, Today’s Fasted SC

- **1986**, Cray-2 breaking the *Giga*flop/s barrier
- **1997**, Intel ASCI Red, breaking *Tera*flop/s barrier
- **2008**, IBM RoadRunner, breaking *Peta*flop/s
  - At DOE’s Los Alamos National Laboratory
  - 1.026 Linpack Petaflop/s solving 2 Mio equations
  - 6912 dual-core Opteron & 12960 IBM Cell
  - #1 on the Top500 list of June 2008
Grids

1998: The Grid: Blueprint for a New Computing Infrastructure:
“... hardware and software infrastructure ... dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.”

2002: The Anatomy of the Grid:
“. . . coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations.”

2002: Grid Checklist:
1) coordinates resources not subject to centralized control …
2) … using standard, open protocols and interfaces
3) … to deliver nontrivial qualities of service.

Quotes: Ian Foster, Carl Kesselman, Steve Tuecke
Research is using Grids
Industry is using Grids
Clouds

- IT resources provisioned outside corporate data center
- Resources accessed over the internet
- Variable cost of services
- Service oriented: SaaS, PaaS, IaaS, HaaS
- A virtual computing environment (VMware, Xen,...)
- Build and deliver, always-on, pay-per-use IT services
- Scaling: computing, storage, database, services, and users, up and down
- Abstraction of the hardware from the service
The Cloud of Cloud Companies

- Amazon
- Google
- Salesforce
- Microsoft
- Sun
- IBM
- Oracle
- EMC
- Cloudera
- Cloudsoft
- Akamai
- Areti Internet
- Enki
- Fortress ITX
- Joyent
- Layered Technologies
- Rackspace
- Terremark
- Xcalibre
Example of an e-Infrastructure:

The DEISA Ecosystem for HPC Grand-Challenge Applications

Distributed European Infrastructure for Supercomputing Applications
DEISA Partners

DEISA1: May 1st, 2004 – April 30th, 2008
DEISA2: May 1st, 2008 – April 30th, 2011
DEISA: Vision - Mission - Strategy

Vision:
Establishing persistent European **HPC ecosystem**
integrating national Tier-1 (Tflop/s) centres and the new European Tier-0 (Pflop/s) centres

Mission:
Enhance Europe’s capability in computing and science by integrating most powerful **supercomputers** into a European HPC e-infrastructure
Build European Supercomputing **Service** on top of existing national services, based on the deployment and operation of a persistent, **production** quality, distributed supercomputing environment with continental scope
One Example of Virtual Communities: Joint Research Activity “Life Sciences”

The DEISA Life Science Portal

Joint Research Activity (JRA)

Promoting parallel apps in the life science community

Running big simulations on DEISA infrastructure that couldn’t be done locally

Providing ease of access to resources

Application support for life science portal
Cluster/Grid/Cloud Portal
Example: NICE EnginFrame

Provides remote, interactive, transparent, and secure access to applications and data on your corporate Intranet or Internet, or in the Cloud.

Users and administrators can access and control computing resources via an intuitive and standard Web interface virtually anywhere using a standard Web browser.
Aerodynamic Shape Optimization

- 4 parameters to be optimized
- cubic face centered DOE
- 25 cases + 16 extra cases for error estim.
- polynomial response function
- 70 hours wall clock time on 64 cpus

Each of these steps need to be fully automated and controlled by the optimizer
– Study the impact of radiative heat transfer (RHT) on the combustion process (2D)
– Couple combustion (AVBP), the RHT (Rayon) codes and the pollutant formation (AVBP)
– Parallelization of the Rayon code and improvement of the coupling part
– Load balancing issue
– 3D extension proposed to DECI and accepted
Climate Research
Statistics of Climate Variability

Project to study climate trends, each 50 TB output data

H. Dijkstra, U Utrecht, and W. Hazeleger, KNMI
Environmental Application

- Study the impact of water cycles of the hydrological and vegetation models on climate models
- Coupling area in West Africa
- Best performances with a vector and scalar platform
- Improve extensibility of the architecture and the coupling part
- AMMA project, PhD thesis, 2 publ. and 2 comms.
Materials Science
First-principles statistical mechanics for molecular switches at surfaces (MolSwitch)

Azobenzene on copper, silver and gold surfaces

Controlled reversible switching should be possible on Ag surfaces

Courtesy: K. Reuter, FHI
Curvy membranes make proteins attractive

For almost two decades, physicists have been on the track of membrane mediated interactions. Simulations in DEISA have now revealed that curvy membranes make proteins attractive.


a) proteins (red) adhere on a membrane (blue/yellow) and locally bend it;
b) this triggers a growing invagination.
c) cross-section through an almost complete vesicle

Cosmology Project

- Study galaxy formation in cosmology
- Physics / modules: Gravitation, Hydrodynamics, Chemistry
- Best performance on heterogeneous platforms
- Load balancing issue and improvement of the coupling part
- Proposed to DECI

Image:

- Hydrodynamics (Baryonic gas)
- Abundance for each species
- Gravity (Dark Matter)
- Chemistry (Baryonic gas)
- Mass phi
- Gas mass phi
- a)
Next-Generation e-Infrastructures for

ACHI: Advances in Computer-Human Interactions

eKNOW: Information, Process, and Knowledge Management

eL & mL: Mobile, Hybrid, and On-line Learning

eTELEMED: Health, Telemedicine, and Social Medicine

GEOWS: Advanced Geographic Information Systems & Web Services

ICDS: Digital Society

ICQNM: Quantum, Nano, and Micro Technologies
Connected: anyone, anywhere, anytime, any device

- Integration of new devices, data and information sources
- Cell phones, PDAs, smart sensors, sensor arrays, health monitors
- Devices embedded in cars, engines, roads, bridges, clothes,...
- Huge amount of data for real-time analysis
- Policies, grid economy, to maintain stability and efficiency
- Support organizational and societal structures, to bridge political and social boundaries . . .
Example: e-Learning on e-Infrastructures

The Challenges:

World-wide data and knowledge explosion
We need more scientists and engineers, but not enough students are interested in science
Schools and teachers are not prepared

The Solution:

New ways of teaching and learning for our digital natives
e-Infrastructures for enriched learning
Working with didactic and pedagogic experts

The Prototype:

e-School, interactive science laboratory, the digital sand-box for life-long learning in the sciences
Today’s Education Challenges

- **Information** & knowledge growing exponentially
- **Teaching** methods + materials do not keep pace
- **Learning** is too passive and static, life is highly active and dynamic
  
  Prof. Srivathsan, India: “...education today is not learning centric, it is exams centric.”
- **Students** become de-motivated and lack creativity
- **e-Learning** environments just scratching the surface

We need **100Ks** of new jobs in science and engineering

We have to focus on K-12 students and their teachers
e-School Prototype

A **Virtual Laboratory** based on an e-Infrastructure and a distributed digital repository for science and engineering applications for students and educators.

Bridging the Chasm between

Education  
Science
Vision: e-School Science Collaboratory for a better learning experience

- **Inter-active** learning tools for creative students (edutainment)...
- ...**same tools** engineers & scientists are using in the 21\textsuperscript{st} century
- Edu portal provides seamless access to **virtual laboratory**
- 100s of real-world computer simulations available for all ages
- On dynamic, shared, remote resources, **at your finger tip**

e-School: empowering education

- **Learning by doing** -
Welcome to the e-School Digital Laboratory!

This is a prototype of e-School which will demonstrate how you can easily and interactively build, modify and run physics applications on remote computers in the Internet.

Therefore, after some initial reading, the most important action for YOU is to go to the Application section and run at least one application.

PLEASE TRY IT, IT'S EASY!

In the future, here, you will find a wealth of digital experiments in areas such as math, physics, chemistry, biology, weather, climate, environment, bioinformatics, biophysics, medicine, aeronautics and humanities, especially prepared for your specific interest, school grade and curriculum.

You are now in your personal and secure e-School website. You can personalize this site further by adding your favorite links on the right side of this page, or by sharing information with others, in the Wiki, or by including in the application library your own digital examples which you may have designed and developed yourself.

Next, you want to configure your system for the e-School prototype in an optimal way. For that, we recommend to follow the 'Configuration' steps on the 'Downloads' section.

Next, you want to learn how to use this e-School prototype. For this, we have included a so-called "e-School Primer" which you can get from the "Downloads" page. From the same place you can get e-School's newsletter.

Now, you are ready to go! Click on the red "Applications" tab, and let you take through the world of digital virtual laboratory experiments!

Thank you very much for visiting us. We appreciate any comment! Please send to wgentzsch@d-grid.de.
e-School: Your personal workspace
Example: interactive real-time fluid flow
Distance Learning on e-Infrastructures

**Distance Learning...**
- Independent of time and space
- Self-paced learning
- Teacher-independent learning
- Deductive science education
- Mostly single-learner environment.
- Linear inter-reactivity, at best
- Mostly static and repetitive

=> Improved (but similar to) class-room learning

**...on e-Infrastructures**
- Independent of time and space
- Self-paced learning
- Teacher-independent learning
- Inquiry-based science education
- Allows for collaborating groups
- Fully nonlinear interactivity
- Highly dynamic and nonlinear
- Allows for complex simulations, data processing, and visualization
- Students and teachers become more creative, motivated and committed

=> Paradigm Shift
Read more in:

Thank You!

GRACIAS POR SU ATENCIÓN

Gentzsch@rzg.mpg.de