BOOK OF ABSTRACTS

of the

1st Italian Grid Training Workshop

of the

Italian Grid Infrastructure (IGI)

GARR, Rome, Italy, January 20-21, 2014

Edited by Alessandro Costantini, Emidio Giorgio, Vania Boccia and Daniele Cesini

Virt&I-Comm Special Issue (2014). Published by: MASTER-UP S.R.L. ISSN: 2279-8773

Virt&l-Comm is a new international online electronic scientific magazine offering a forum for presenting the work carried out by Virtual Innovation, Research, Teaching & Learning Communities. This magazine is the result of a joint endeavour of the COMPCHEM Virtual Organization (VO), of the European Chemistry Thematic Network (ECTN) Association and of the University of Perugia spinoff Master-Up.

Virt&l-Comm is a scholarly open access online Magazine requiring no payment neither from the authors nor from the readers.

Virt&l-Comm publishes at present two issues per year to promote Molecular and Materials Science, Teaching and Learning, Computer Science research. The magazine is also specialized in education and innovation and focuses on Service Oriented approaches. Therefore, in addition to articles, news, projects, reports of (successful or failed) attempts to build services relevant to the field of interest, lists of best practices, products used, the advantages and disadvantages of the solutions adopted will be considered for publication. In particular, all the information useful to build the puzzle of innovative complex applications in education (EDU), information and communication tecnology (ICT) and research and development (R&D) in Molecular and Materials science are welcome.

BOOK OF ABSTRACTS

of the

1st Italian Grid Training Workshop

of the

Italian Grid Infrastructure (IGI)

GARR, Rome, Italy, January 20-21, 2014

Organizing Committee

Antonio Laganà, Chair (Università di Perugia)
Daniele Cesini (INFN-CNAF)
Alessandro Costantini (INFN-PERUGIA)
Emidio Giorgio (INFN-CATANIA)
Vania Boccia (INFN-NAPOLI)

INSTM Committee

Vincenzo Barone (Scuola Normale Superiore, Pisa)
Orlando Crescenzi (Università degli studi di Napoli Federico II)
Antonio Laganà (Università di Perugia)

Administrative Assistance

Laura Pirelli (GARR)

Technical Assistance

Edoardo Angelucci (GARR)

Welcome to Rome!

Rome is the capital of Italy and also of the Province of Rome and of the region of Lazio. With 2.7 million residents in 1,285.3 km^2 (496.3 sq mi), it is also the country's largest and most populated *comune* and fourth-most populous city in the European Union by population within city limits. The urban area of Rome extends beyond the administrative city limits with a population of around 3.8 million. Between 3.2 and 4.2 million people live in Rome metropolitan area. The city is located in the central-western portion of the Italian Peninsula, on the Tiber within Lazio (Latium). The Vatican City is an independent country within the city boundaries of Rome, the only example of a country within a city existing.

Rome's history spans more than two and a half thousand years, since its legendary founding in 753 BC. Rome is one of the oldest continuously occupied cities in Europe. It is referred to as "The Eternal City", a notion expressed by ancient Roman poets and writers. In the ancient world it was successively the capital city of the Roman Kingdom, the Roman Republic and the Roman Empire, and is regarded as one of the birthplaces of Western civilization. Since the 1st century AD, Rome has been considered the seat of the Papacy and in the 8th century it became the capital of the Papal States, which lasted until 1870. In 1871 Rome became the capital of the Kingdom of Italy, and in 1946 that of the Italian Republic.

After the Middle Ages, Rome was ruled by popes such as Alexander VI and Leo X, who transformed the city into one of the major centers of the Italian Renaissance along with Florence. The current version of St Peter's Basilica was built and the Sistine Chapel was painted by Michelangelo. Famous artists and architects, such as Bramante, Bernini and Raphael, resided for some time in Rome, contributing to its Renaissance and Baroque architecture.

Rome has a status of the global city. In 2007, Rome was the 11th-most-visited city in the world, 3rd most visited in the European Union, and the most popular tourist attraction in Italy. Its historic centre is listed by UNESCO as a World Heritage Site. Monuments and museums such as the Vatican Museums and the Colosseum are among the world's most visited tourist destinations with both locations receiving millions of tourists a year. Rome hosted the 1960 Summer Olympics.

INTRODUCTION TO THE 1ST ITALIAN GRID TRAINING WORKSHOP

The COMPCHEM Virtual Organisation (https://www3.compchem.unipg.it/compchem/) in collaboration with the Italian NGI User Support and Training unit (IGI-FUS) held the *First Training Workshop on Grid porting of computational chemistry applications*. The event, held from 20th to 21st of January, was hosted in Rome by the GARR consortium (http://www.garr.it/b/eng), the Italian National Research and Education Networks (NREN), and saw the active participation of 22 young researchers from 12 different departments of the various national institutes, including a remote participant from the Spanish National Research Council (CSIC) in Madrid.
The National Institute of Materials Science and Technology (INSTM - http://www.instm.it/en/instm.aspx) also supported the event by funding three scholarships for the attendance of the courses.
The event was preceded by two on-line meetings, where the User Support team demonstrated to the participants how to cope with the most common issues when accessing distributed resources for the first time.
During the 2 days event, after a general introduction on the EGI distributed systems, the workshop focused on three specific use cases that the students were asked to prepare in advance. The use cases were ported onto the grid, mainly by exploiting the NGI_IT services and resources. The porting was implemented starting with the command line interface, made available by the gLite middleware, to end with the creation of basic graphic interfaces relying on the Italian Grid Portal (https://portal.italiangrid.it).
The three real life use cases were based on three different applications popular among the Computational Chemistry community members: VENUS, a classical trajectory direct dynamics computer program developed at the Texas Tech University (https://cdssim.chem.ttu.edu/nav/htmlpages/licensemenu.jsp), CRYSTAL, a quantum chemistry program for solid state physics and chemistry developped at the University of Torino, Italy (http://www.crystal.unito.it/index.php) and Quantum Espresso, an open-source software for electronic-structure calculations and materials modeling at the nanoscale (http://www.quantum-espresso.org/).
Professor A. Laganà, from the University of Perugia and leader of COMPCHEM VO commented that "*this training workshop is part of a dissemination model that abandons the general purpose approach to focus on the porting of specific use cases done by a synergy of grid supporters and experts of the scientific application that for a limited period of time work side by side*". "*A follow up to improve the module and a special issue of the VIRT&L-COMM electronic journal (https://www3.compchem.unipg.it/ojs/index.php/ojs) containing the event material will be released*" continues Professor Laganà.

The participants expressed a general satisfaction for the event, evaluating it with an overall average rate of 3.5 out of 4. One of the received comments we liked most was "*Experience to be repeated. Very useful to exchange information and share experiences with users belonging to different fields of the chemistry realm.*"

ITALIAN GRID INFRASTRUCTURE

IGI (http://www.italiangrid.it) is a Joint research Unit (JRU) made of more than 20 Italian academic and research institutions that is based on a Memorandum of Understanding (MoU) signed for the purpose of participating of some national and European Grid projects. At present IGI is actively participating to the EGI-InSPIRE project (https://www.egi.eu/about/egi-inspire/) and is one the largest National Grid Initiatives (NGI) of the European Grid

Infrastructure (EGI, https://www.egi.eu) with a recognized leadership both in the Grid technology development and in the management of the distributed computing infrastructure operations supporting research communities.
IGI currently relies on more than 50 geographically distributed sites, provides about 33000 computing cores
plus 30PB of storage capacity and supports more then 50 Virtual Organizations (VOs) with thousands of active users. The infrastructure implements a customized version of the gLite \cite{glite} middleware distributed by the European Middleware Initiative (EMI) project (http://www.eu-emi.eu/). One of the roles of IGI is to satisfy the compute and storage demand of various user communities such as high energy physics, computational chemistry, bioinformatics, astronomy and astrophysics, earth science.
IGI provides user support services for already consolidated Grid users but provides also consultancy, application porting and training support for new users and new communities that would like to exploit the Grid infrastructure. The IGI "User Support and Training" unit takes care of new user communities. User Support activities are carried on in close collaboration with the support teams of other national Grid initiatives and with the EGI.eu user support activities.

CONSORTIUM GARR

GARR is the Italian Academic and Research telecommunication network. Its main goal is to provide high-bandwidth connectivity and advanced services to the national scientific and academic community. The GARR network is fully integrated into the worldwide Internet. Consortium GARR is a non-profit organization constituted under the aegis of the MIUR (Ministry of Education and Scientific Research).The founding Members are: CNR (National Research Council), ENEA (Italian National Agency for New Technologies, Energy and Sustainable Economic Development), Fondazione CRUI (Conference of Italian University Chancellors), INFN (National Institute of Nuclear Physics). All Academic and major Scientific organizations in Italy connect to the GARR network. GARR plans and operates the national high-speed telecommunication network for University and Scientific Research. All Academic and major Scientific organizations in Italy connect to the GARR network.

GARR institutional mandate includes the following:

- to implement and operate the national high-speed telecommunication network ;
- for University and Scientific Research, and to interconnect to other NRENs in Europe and worldwide, as well as with the Global Internet;
- to provide the user community with network operation and application services;
- to facilitate cooperation in the field of research through the exploitation of leading-edge e-Infrastructures, both at a national and international level;
- to disseminate advanced knowledge about network infrastructure, and to stimulate the exchange of technical know how within the user community.

OBJECTIVE

The main objectives of the workshop are:

- give an overview, from an user perspective, of the Grid infrastructure in the European context.

- give the know-how needed to analyse and solve the issues and related to the application porting on distributed computing infrastructures. The following computational chemistry applications will be used to solve real-life GRID use cases:
    - VENUS (molecular dynamics)
    - CRYSTAL (electronic structures of solid materials)
    - QUANTUM ESPRESSO (electronic structures and modelling)

Although such use cases are taken from the computational chemistry realm, the workshop participation has been open to interested users belonging to other communities and disciplines. In order to ensure a common background for all participants, the course has been preceded by web based sessions filling possible knowledge gaps.

PROGRAM OF THE 1st ITALIAN GRID TRAINIG WORKSHOP

20-21 GENUARY 2014

CONSORTIUM GARR, ROME, ITALY

- Monday, January 20, 2014

  - 09:00 - 13:00 Introduction and course overview
    Welcome address, introduction to IGI, grid middleware overview

    - 09:00 Welcome and Registration *1h0'*
    - 10:00 The Italian and European Grid Infrastructures *30'*

      Speaker: Antonio Laganà (University of Perugia)

    - 10:30 Coffee break *30'*
    - 11:00 Grid Middleware Services and User Interfaces *1h30'*

      Speakers: Mr. Daniele Cesini (INFN-CNAF), Dr. Alessandro Costantini (IGI), Dr. Emidio Giorgio (INFN-CATANIA), Vania Boccia (INFN-Napoli)

    - 12:30 Hands on Grid Services *30'*

  - 13:00 - 14:00 Lunch Break

  - 14:00 - 18:00 Use Case 1 : CRYSTAL

    - 14:00 Introduction to CRYSTAL *15'*
      Speaker: Silvia Casassa

    - 14:15 CRYSTAL porting to Grid *30'*
      Speaker: Dr. Alessandro Costantini (IGI)

    - 14:45 Hands on CRYSTAL (Grid Command Line) *1h30'*
    - 16:15 Coffee Break *15'*
    - 16:30 Hands on CRYSTAL (IGI Portal) *1h30'*

- Tuesday, January 21, 2014

  - 09:00 - 13:00 Use Case 2 : VENUS

    - 09:00 Introduction to VENUS *20'*

      Speaker: Andrea Lombardi

    - 09:20 VENUS porting to Grid *30'*

      Speaker: Dr. Alessandro Costantini (IGI)

- 09:50 Hands on VENUS (Grid Command Line) *1h30'*
- 11:20 Coffee Break *20'*
- 11:40 Hands on VENUS (IGI Portal) *1h20'*

- 13:00 - 14:00 Lunch Break
- 14:00 - 17:45 Use Case 3 : QUANTUM ESPRESSO
  - 14:00 Introduction to QUANTUM ESPRESSO *20'*

    Speakers: Stefano Cozzini, Michele Ceotto
  - 14:20 QUANTUM ESPRESSO porting to Grid *30'*

    Speaker: Mr. Daniele Cesini (INFN-CNAF)
  - 14:50 QUANTUM ESPRESSO Hands on (Grid Command Line) *1h20'*
  - 16:10 Coffee break *15'*
  - 16:25 QUANTUM ESPRESSO Hands on (IGI Portal) *1h20'*

- 17:45 - 18:00 Feedback and farewell

LIST OF PARTICIPANTS

Name	Institute
Bartolomei Massimiliano	CSIC, Madrid (Spain)
Carlotto Silvia	Department of Chemistry, University of Padova (Italy)
Casassa Silvia	Department of Chemistry, University of Torino (Italy)
Ceotto Michele	Department of Chemistry, University of Milano (Italy)
Ciriaco Fulvio	Department of Chemistry, University of Bari
Degli Esposti Boschi Cristian	CNR-IMM, Bologna (Italy)
Filomia Federico	Department of Chemistry, University of Perugia (Italy)
Giusepponi Simone	ENEA, Frascati (Italy)
La Manna Pietro	ICTP, Napoli (Italy)
Lombardi Andrea	Department of Chemistry, University of Perugia (Italy)
Longobardi Anna	Department of Physics, University of Napoli Federico II (Italy)
Lunghi Alessandro	Department of Chemistry, University of Firenze (Italy)
Luni Camilla	INSTM, Padova (Italy)
Mavelli Fabio	Department of Chemistry, University of Bari (Italy)
Rampino Sergio	CNR-ISTM, Perugia (Italy)
Spinello Angelo	Department of Biology, University of Palermo (Italy)

LIST OF CONTRIBUTIONS

Acknowledgements

Organizers acknowledge past and present grid staff, expert users, participants and related groups and Institutions for the success of the Italian Grid Training Workshop event.
The material presented in Section 1 has been collected and assembled from the outcomes carried out by the EGEEI/II/III and EGI project.

1. Introduction and course overview

1.1 The italian and european grid infrastructure

Antonio Laganà, University of Perugia (Italy)

Italy has been one of the most active European countries in promoting the development of Grid computing in physics. It has been also particularly active in gathering on the Grid infrastructure computational scientists from other disciplines and in particular those of the Molecular and Materials Sciences and Technologies community. The resulting cluster of grid-active Italian communities has established a Joint Research Unit (JRU), called Italian Grid Infrastructure (IGI). IGI is, therefore, the National Grid Infrastructure (NGI) that represents the Italian Grid community before EGI, delivers the instances of such community to the Italian government, operates the national compute platform.

A goal of IGI is also that of training Italian researchers to use the Grid by means of its FUS (Formazione e User Support) unit. In this spirit FUS and the Italian section of the COMPCHEM VO, with the support of GARR and INSTM Consortia, have organized a Training Workshop for the Italian Chemistry community. The training workshop has been preceded by two webinars aimed at teaching the attenders how to get Personal Certificates and follow the procedures for authorization and authentication on Grid.

The peculiarity of this Training Workshop has been that of familiarizing the young members of the Chemistry community with the Grid by exploiting collaboration in the porting and execution of three popular chemistry computational applications:
- VENUS (molecular dynamics);

- CRYSTAL (electronic structure of solids);

- QUANTUM ESPRESSO (electronic structure and materials modeling).

These three applications, in fact, not only offer a clear perspective of the various types of chemistry applications that can be implemented on the grid but highlight also the importance that collaboration plays in the process of building a Virtual Research Community (VRC) for Chemistry, Molecular and Materials Sciences and Technologies (CMMST). As a matter of fact, the CMMST VRC adopts a synergistic model for carrying out computational research that builds on a Service Oriented Architecture approach in which the work performed is evaluated and rewarded in terms of the contribution offered to the progress of the community.

Comunità virtuali:
dal polo di calcolo scientifico locale alla grid computazionale mondiale in scienze e tecnologie molecolari

Antonio Laganà
Università di Perugia (UNIPG), IT

DIPARTIMENTO DI CHIMICA, BIOLOGIA e BIOTECNOLOGIE: S. Crocchianti, N. Faginas Lago, L. Pacifici, F. Filomia, C. Manuali
DIPARTIMENTO MATEMATICA E INFORMATICA: O. Gervasi, S. Tasso, S. Pallottelli
INFN: A. Costantini

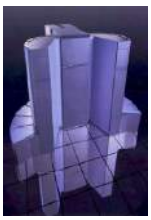
SOMMARIO

1) IL SOGNO DEL RICERCATORE COMPUTAZIONALE

a. CALCOLARE AB INITIO DEGLI OSSERVABILI DI SISTEMI REALISTICI (ESPERIMENTO VIRTUALE)
b. SCEGLIERE (senza fare domanda) LE RISORSE PIU' IDONEE (ovunque siano) PER AUTOGESTIRE LO SVILUPPO E I CALCOLI DI PRODUZIONE
c. COMPORRE I PROPRI CODICI CON PROGRAMMI E ROUTINE DISPONIBILI ALL'INTERNO DELLA COMUNITA' SCIENTIFICA
d. RICONOSCERE L'ALTRUI (e vedere riconosciuto il proprio) LAVORO FATTO A FAVORE DELLA COMUNITA' (Sistema crediti)

Anni '90: Supercalcolatori

Alto parallelismo, accesso mediante grant

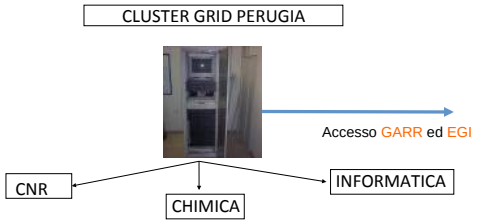


Cray ymp

Connection machine

IBM SP

2) GRID - ANNI 2000: CALCOLO IN RETE



1) Blades con CPU multicore, GPU, …
2) Accesso opportunistico

LA GRID: DEFINIZIONE

Ian Foster

I.Foster, C.Kesselman: The Grid: Blueprint for a New Computing Infrastructure", 1998

Carl Kesselman

" La grid computazionale è una infrastruttura hardware e software in grado di offrire accesso a capacità di calcolo di alto livello affidabile, riproducibile, pervasiva ed economica"

CARATTERISTICHE

La Grid

1) **Coordina risorse eterogenee senza controllo centralizzato**
2) **Usa protocolli e interfacce standard, non proprietarie e general purpose**
3) **Offre un servizio di alta qualità**

(Ian Foster, 2002)

[1] Foster, I. and Kesselman, C. eds. The Grid: Blueprint for a New Computing Infrastructure, Morgan Kaufmann, 1999, 259-278

[2] Ian Foster, Carl Kesselman, and Steven Tuecke. 2001. The Anatomy of the Grid: Enabling Scalable Virtual Organizations. Int. J. High Perform. Comput. Appl. 15, 3 (August 2001), 200-222. DOI=10.1177/109434200101500302

[3] What is the Grid? A Three Point Checklist. I. Foster, GRIDToday, July 20, 2002.

SISTEMA DISTRIBUITO

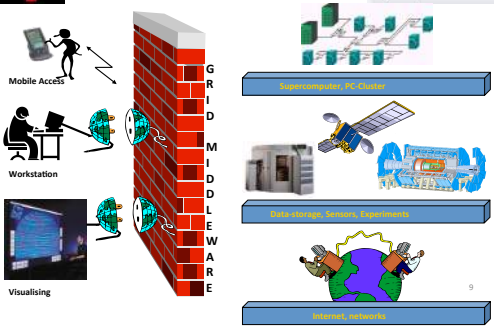
L'UTENTE HA PIENO POSSESSO DEL SUO AMBIENTE DI LAVORO

LE GRANDI RISORSE DI CALCOLO SONO GESTITE DAI RELATIVI AMMINISTRATORI MA SONO CONDIVISE PER L'USO CONDIVISO ON LINE

L'UTENTE SOTTOMETTE I JOB CHE VENGONO ESEGUITI SULLA GRID: somehow, somewhere, (after) some time.

LA GRID NON HA PADRONI

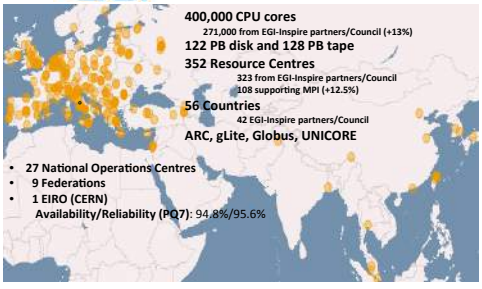
SISTEMA INCLUSIVO



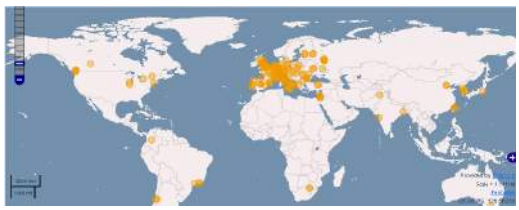
La Grid Italiana (IGI)

IGI is the Italian NGI (National Grid Infrastructure) and is part of **EGI** the European Grid Infrastructure.
IGI is currently a Joint Research Unit (JRU) established in December 2007, and made by Research institutions (4 Universities among which Perugia) and is represented in the Coordination body.
The leading Institution is INFN

INFRASTRUTTURA Grid Europea



THE WORLD GRID



USA – Technical University of Texas in Lubbock, University of Oklahoma in Norman
BRASILE – Universidad Estadual Paulista, Bauru, San Paulo
TAIWAN – Academia Sinica, Taipei

3) LE ORGANIZZAZIONI VIRTUALI

- **Virtual Organization (VO): un insieme di individui e/o istituzioni con interessi scientifici simili che condividono le risorse in forma regolamentata indipendentemente dalla loro collocazione geografica**
  - *La condivisione è strettamente controllata*, I fornitori e i consumatori di risorse definiscono chiaramente cosa viene condiviso, chi può condividere e le condizioni per la condivisione

LA VO COMPCHEM & LA PIATTAFORMA

contatto Alessandro Costantini (alex.costantini@gmail.com)

website http://compchem.unipg.it

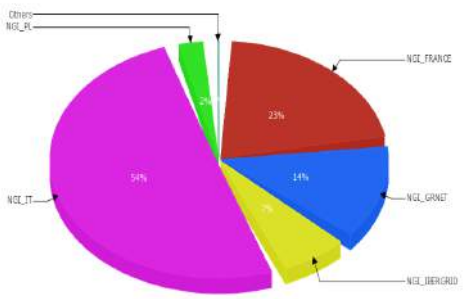
fornitori risorse circa 10000 cpu da IGI (Italia), CESGA (Spagna), IN2P3 (Francia), CYFRONET (Polonia), POZNAN Supercomputing Center (Polonia), GRNET (Grecia), Università Cipro

utenti registrati 58 (soprattutto Italia, Spagna, Grecia) + 20 via GriF

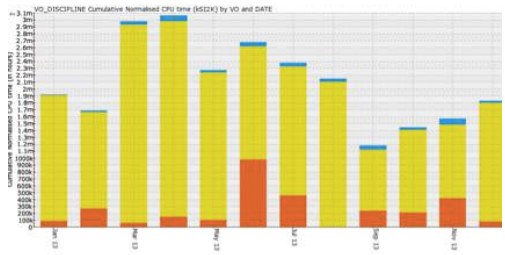
link a NGI (soprattutto Italia, Spagna, Grecia)

VO connections Gaussian, weNMR

A. Lagana', A. Riganelli, O. Gervasi, On the structuring of the computational chemistry virtual organization COMPCHEM,
Lecture Notes in Computer Science 3980, 665-674 (2006).



TEMPO CPU USATO



4) SERVICE: UTENTE FORNITORE

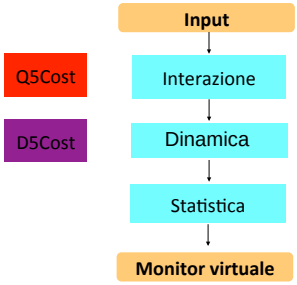
1. **UTENTE SEMPLICE**
   - **PASSIVO: Utilizza programmi altrui (possibilmente mediante un portale)**
   - **ATTIVO: Implementa almeno un programma per uso personale**
2. **FORNITORE SW (si possono guadagnare crediti)**
   - **PASSIVO: Implementa almeno un programma per uso altrui**
   - **ATTIVO: Gestisce almeno un programma per uso cooperativo**
3. **FORNITORE HW**
   - **PASSIVO: Conferisce all'infrastruttura almeno un cluster di processori**
   - **ATTIVO: Contribuisce a istallare e gestire la struttura**
4. **GESTORE (AZIONISTA): Partecipa allo sviluppo e alla gestione della organizzazione virtuale**

- Ulteriori informazioni in http://compchem.unipg.it

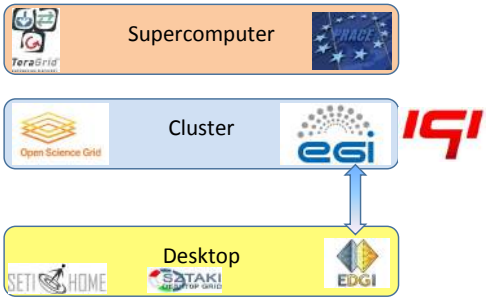
I SERVIZI IMPLEMENTATI

- **(quantum chemistry) COLUMBUS, GAMESS-US, ..**
- **(quantum dynamics) ABC, RWAVEPR, FLUSS, MCTDH, DIFF REAL WAVE, …**
- **(molecular dynamics) VENUS, DL_POLY, GROMACS, …**
- **(multiscale chemistry simulations) CHIMERE, CLEAN-COMBUSTION**
- **(education) LEARNING OBJECTS REPOSITORY**
- **ISTALLAZIONE DI ULTERIORE SOFTWARE (PREFERIBILMENTE OPEN SOURCE O ACCADEMICO)**

GEMS (IL SIMULATORE MOLECULARE SU GRID
DIAGRAMMA DI FLUSSO E FORMATO DATI)



INTEROPERABILITA' SELETTIVA SU GRID



Mutual Aid: A Factor of Evolution (Dover Value Editions)
Kropotkin, Peter
Editore: Dover Publications, 2006
ISBN 10: 0486449130 / ISBN 13: 9780486449135

Usato / Paperback / Quantità: 3
Da Net Textstore, LLC. (Lincoln, NE, U.S.A.)
Valutazione libreria:

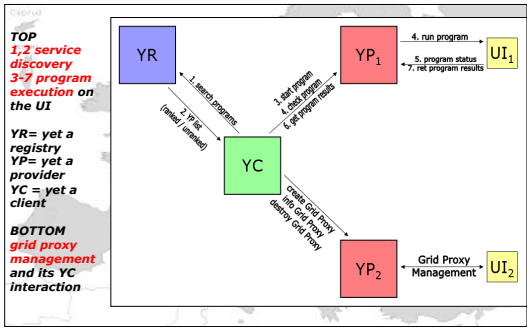
Foto dell'editore

The Generous Man: How Helping Others is the Sexiest Thing You Can Do
Norretranders, Tor
Libreria: Better World Books
(Mishawaka, IN, U.S.A.)
Valutazione libreria:
Quantità: 1
ISBN: 9781560257288

GriF: UN FRAMEWORK PER VALUTARE LA QUALITA' DEL SERVIZIO (QoS)

- Rende le applicazioni del tipo **scatola chiusa**
- Rende la distribuzione su grid dei job **automatica**
- Adotta un approccio **Service Oriented (SOA)** articolato come insieme di **servizi grid**
- Offre agli utenti modalità operazionali **facili e autogestite**
- Consente la costruzione **a la carte** delle applicazioni
- Indirizza in modo opportuno i job al **computer appropriato**

Due JAVA server, un JAVA client



5) SVILUPPARE UNA ECONOMIA DI CREDITI

- Adottare un approccio "service oriented"
- Introdurre una metrica nella VO
- calcolare QoS e QoU
- Implementare un sistema di crediti e di costi dei servizi

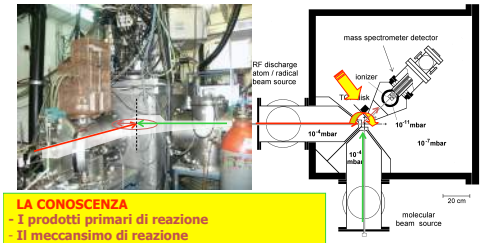
C. Manuali, A. Lagana'
GRIF: A New Collaborative Framework for a Web Service Approach to Grid Empowered Calculations
Future Generation of Computer Systems, 27(3), 315-318 (2011) DOI 10.1016/j.future.2010.08.006

6) COSA FARE

- POSSO ADERIRE A COMPCHEM?
- POSSO PARTECIPARE AL POLO DI CALCOLO SCIENTIFICO?
- ESEMPI DI RICERCA PER L'INNOVAZIONE E LA DIDATTICA

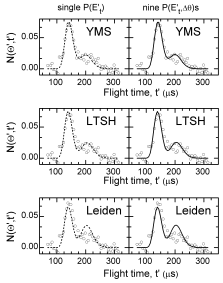
L'ESPERIMENTO REALE A FASCI MOLECOLARI

LE MISURE
- Il tempo di volo e l'angolo di dispersiones



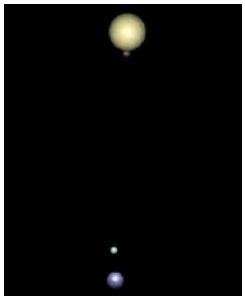
LA CONOSCENZA
- I prodotti primari di reazione
- Il meccansimo di reazione
- La struttura e il tempo di vita degli intermedi
- La distribuzione dell'energia interna
- Il potenziale di interazione

L'ESPERIMENTO VIRTUALE VS LA MISURA



A. Lagana , E. Garcia , A. Paladini, P. Casavecchia, N. Balucani,
The last mile of molecular reaction dynamics virtual experiments: the case of the OH (N=1-10) + CO (j=0-3) → H + CO_2 reaction
Faraday Discussion of Chem. Soc. 157, 415-436, 2012

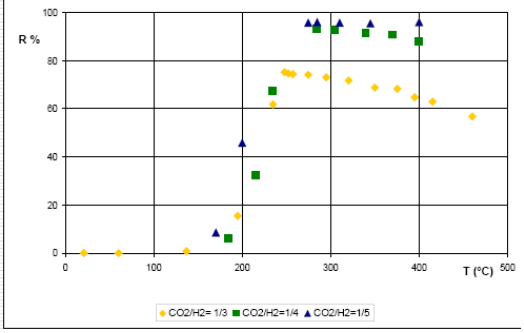
LA VISUALIZZAZIONE DEL MICROSCOPICO



STATO DI TRANSIZIONE
Esiste una superficie di separazione oltre la quale il sistema non torna indietro?

OH + HCl → H_2O + Cl

LA RIDUZIONE DI CO_2 con H_2 (PAUL SABATIER)



IL PROGETTO PROGEO

Utilizzazione delle energie rinnovabili: un tema fondamentale per la sostenibilità della nostra società

Unica fonte energetica e "motore" di cui disponiamo: il **Sole**

I PARTNER

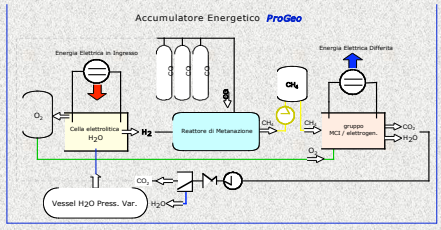
UNIVERSITA'
Una ricerca della conoscenza in sé, sia pure applicata a sistemi reali

ENEA
Una ricerca orientata alla produzione di applicazioni produttive

INDUSTRIA
Ingegnerizzazione, produzione e marketing di prodotti innovativi in campo energetico.

PROGEO

Usare l'energia prodotta da fonti rinnovabili per produrre H_2 con cui ridurre la CO_2 e formare CH_4 (metano verde)

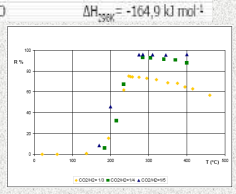


LE RESE DI PROGEO

I rendimenti dei primi due processi: elettrolisi dell'acqua e reazione di metanizzazione sono molto alte.
Gli attuali elettrolizzatori raggiungono rendimenti del 70-75%, considerando tale valore come rapporto tra il potere calorifico dell'idrogeno prodotto (11.9 MJ/Nm3) e l'energia elettrica necessaria per produrlo

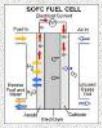
Per quanto attiene la reazione di metanizzazione, si segue la reazione di Sabatier: $CO_2 + 4H_2 \Rightarrow CH_4 + 2H_2O \quad \Delta H_{298K} = -164{,}9\ kJ\ mol^{-1}$

Test sperimentali sono stati condotti in ENEA Casaccia confermando che per rapporti $CO_2/H_2 = ¼$ in poi, la resa di trasformazione è > al 95%.
La reazione si auto-sostiene essendo leggermente esotermica.



Per quanto attiene la fase di conversione del metano in energia elettrica è necessario mettere in evidenza due punti fondamentali:

1) gli attuali economici (~50$/kW) motori a combustione interna (**MCI**) a metano hanno raggiunto rendimenti ottimali, ma comunque intorno al 30% a giri costanti;



2) le più costose Fuel Cells tipo **MCFC/SOFC** possono raggiungere rendimenti intorno al 60% e configurazione totalmente statica. Una terza via sembra molto attrattiva ed è quella delle microturbine a gas, il cui costo è, al momento, molto elevato.

Da ciò deriva che il rendimento globale di ***ProGeo*** come rapporto tra energia *on demand* in uscita ed energia "*non producibile*" in ingresso, può variare dal **21** al **43%,** in funzione dell'investimento.

IL BILANCIO ENERGETICO e DI MASSA

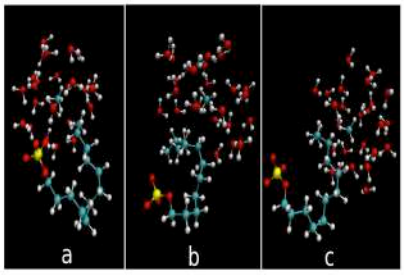
IL BILANCIO ECONOMICO DI PROGEO

Attrezzature		materiali 10%	personale	TOTALI
versione base prototipale da 50kW (valori in €)	versione base di serie da 100 kW (0.1 MW)		20 giorni/uomo	
137.100 €	88.140 €	8.814 €	6.480 €	103.354 €
	versione base di serie da 500 kW (0.5 MW)		25 giorni/uomo	
	295.340 €	29.534 €	8.000 €	233.874 €
	versione base di serie da 1MW		30 giorni/uomo	
	343.398 €	34.338 €	9.600 €	387.318 €
	versione base di serie da 2.5MW		35 giorni/uomo	
	680.916 €	69.092 €	11.200 €	672.297 €
	versione base di serie da 5MW		40 giorni/uomo	
	841.281 €	84.128 €	12.800 €	938.209 €
	versione base di serie da 10 MW		50 giorni/uomo	
	1.177.793 €	117.779 €	19.200 €	1.314.773 €

Tab.1 Previsione di COSTO per varie potenze nella versione base (storage in bombole ed elettrolizzatori commerciali)

Come esempio possiamo prendere una wind farm con 10 MW installati. Ipotizzando che l'energia notturna (es. 8 ore) possa venire accumulata, perché non richiesta, abbiamo 80.000 kWh disponibili da accumulare in metano. Ipotizzando un rendimento minimo globale di ProGeo (da energia elettrica in ingresso ad energia elettrica "differita") ad pari al 21%, potremo fornire 17.000 kWh aggiuntivi "*on demand*". Se il prezzo diurno del kWh è fissato in 15c€ circa, otterremo 2.700 €/giorno ed un tempo di ammortamento inferiore ai 3 anni.

IMMAGAZZINAMENTO DEL METANO IN GHIACCIO: EFFETTO AGITATORE DI SDS

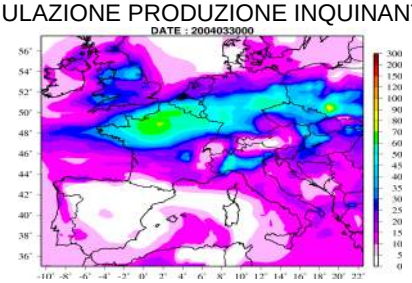


FORMAZIONE DEL GHIACCIO DI METANO COME CLATRATO IDRATO

- Cl_2
- H_2S
- CO_2
- CH_4
- H_2
- *etc.*

UNA PROVA SPERIMENTALE

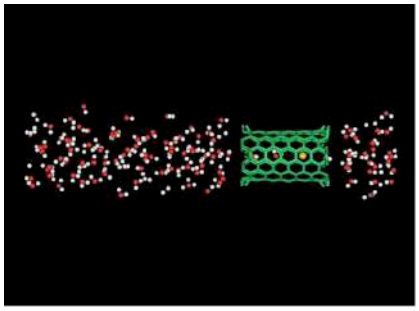
SIMULAZIONE PRODUZIONE INQUINANTI



CPM10 Concentration from CHIMERE-aerosols

SERVIZIO ARPA MARTEDI MATTINA ALLE 7:30 RAI TRE

FLUSSO IONICO IN MICROPORI

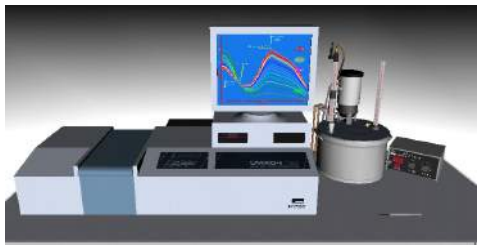


L. Arteconi, A. Laganà, A molecular dynamics study of ion permeability through molecular pores, Lecture Notes in Computer Science 3482, 1093-1100 (2005).

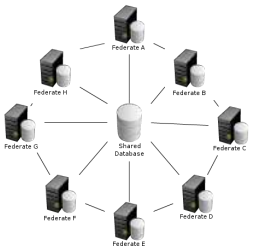
VMSLab-G Prototype: http://www.vmslab.org

VMSLab: HVR - Abstract model for a real Lab

VMSLab-G: UV-VIS Spectroscopy



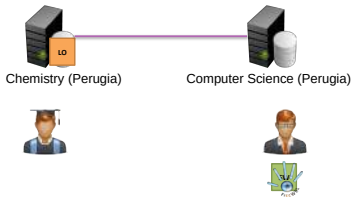
Go to the VRML animation



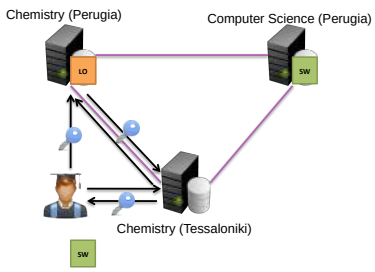
Our results

- Federation of autonomous repositories
- Automatic content sharing
- Downloadable content (for registered users)
- Simplified content import from Moodle
- LO Dependency management

A use case



A use case



CONCLUSIONI

- ESTENDERE IL POLO DI CALCOLO SCIENTIFICO DI PERUGIA AD ALTRI GRUPPI (HARDWARE, APPLICAZIONI E STRUMENTI)
- APPLICARE IL CALCOLO COLLABORATIVO A SISTEMI DI SCIENZE MOLECOLARI SEMPRE PIU' COMPLESSI
- INCENTIVARE I SERVIZI E LE UTENZE MEDIANTE CREDITI LEGATI A PARAMETRI DI QUALITA'
- SVILUPPARE L'INNOVAZIONE E L'INSEGNAMENTO BASATI SULLA RICERCA

1.2 Grid middleware services and user interfaces

Emidio Giorgio[1], Vania Boccia[2]
[1]INFN-Catania and IGI (Italy)
[2]INFN-Napoli and IGI (Italy)

The difficulty of porting an application to the Grid environment is due to several factors, among which: the codes must be able to run on heterogeneous (in terms of architecture, OS, compilers, libraries, etc.) resources, as well as handle different input and output les in the distributed environment; finally, the complex command line interface of some applications and services have to be masked by more user friendly graphical interfaces (e.g. GUI) and workflows.
This difficulty is even greater as the community of users is away from ICT world. In such cases, even if the user is generally conscious of the advantages offered by large computing infrastructures, it is very unlikely that he comes close to them for some reason: the user is not interested in becoming an expert in the use of these infrastructures, but rather wants to learn only what he needs to progress with his research.
Within the IGI framework, the User Support Unit plays a main role in the recruitment of new users belonging to several scientific fields and in the support of those activities aimed at executing efficiently their codes on the Grid.
To this end, over the years, the support unit has promoted various training activities to teach communities how to interact with the Grid: among these initiatives are certainly to be considered the training workshop on the use of Grid and meetings with individual community for porting applications. Here is presented an activity promoted for the first time from the support unit, with the aim of organizing a training workshop on computational chemistry theme that inheriting all the positive aspects of previous initiatives but also at the same time would provide the opportunity for different communities to share the learning experience in their own scientific field.

1. Preparatory meetings

Two preliminary meetings have been scheduled within the workshop organizers and participants, with two main goals : sort out the authorization process for the access to the Grid, and provide to participants a basic understanding of the Grid framework. In order to minimize the effort requested for participation, the meeting have been done online, using a web conferencing tool, that still allowed interactions among students and tutors.
Obtaining an X509 certificate from an accredited CA can be difficult, depending from the candidate home institute. For this reason, school organizers evaluated in advance and send out suggestion to each student for the most convenient way for him/her to obtain the X509 certificate, mandatory for the access to Grid; therefore, the first meeting, it was assessed which was the status of this process and blocking problems, if any. Those who succeeded in obtaining the certificate were invited to fulfil the other mandatory step for the grid access, i.e. the subscription of the Virtual Organizations.
In the second meeting, one week later, after a quick reckoning of the certificate and vo subscriptions status, some basic concepts about Grid were exposed, followed by a short practice both on the command line and web interface.
The outcome of this preparatory meetings was satisfactory, as it allowed to start the live meeting focusing directly on porting application concepts, assuming that preliminar topics had been addressed.

2. Grid Overview, Security, Information System, Data Managements

Although the course was focused on application porting, and preliminary sessions had introduced the grid middleware architecture, a quick introduction to grid was delivered during the live course, for those who couldn't join the online meetings. Within the very range of grid middleware topics, they were selected only those functional to the application porting sessions: therefore concepts like VO, certificate management and information system were just outlined, in order to provide a general understanding of the grid architecture, while more focus (still not too technical) was given to Job and Data management, being these instruments widely used for application porting and usage.

3. Job Management

An exhaustive treatment of the job management is often not necessary in contexts like the one discussed in this document.
The risk you run is in fact to deviate from the key targets which are:

1. to put the new community in conditions of use as soon as possible the Grid infrastructure and
2. to avoid giving too many details on operating procedures that then users will never use.

With these objectives, the tutorial on the Job Management has been designed by treating only general information about grid services involved and the specific terminologies and reporting, through simple examples, the procedure to be used preparing and executing a job in the cases of interest of the communities involved in the workshop .
In particular, the examples focused on job sequential, parallel and parametric, have been reported with the procedures to be followed for the proper submission of the job on the Grid.
Following are shown the slide used during the workshop.

4. Italian Grid Portal

An easy to use web interface is exposed by the IGI portal. This one guides the user from the login phase, until the end of his work session during which a simple web interface allows the user to configure and to run a simple or a parallel job and, eventually, manage even a complex workflow, select the run name (a unique identifier for the entire chain), manage the input files (that should be pre-loaded on a defined Grid location).Since the submission starts, and for the entire workflow lifecycle, the user is notified via mail of the events occurring, including when the workflow completes and the final output data are made available on the predefined Grid storage path. The web interface hides the inner complexity of a parallel jobs workflow in a Grid environment, although it limits workflow customization.

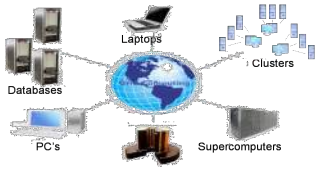
Middleware Overview

Vania Boccia
INFN Napoli

Emidio Giorgio
IGI-INFN Catania

What is a Grid?

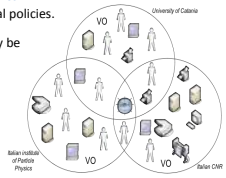
- Collection of geographically distributed heterogeneous **resources**.
- Flexible, secure, **coordinated resource sharing** among dynamic collections of individuals, institutions, and resources



Grid User

Overview

- Members of the grid can dynamically be organized into multiple virtual organizations.
- Each of these VOs may have different policies and administrative control.
- All the VOs are part of a large grid and can share resources accordingly to several policies.
- The resources shared among VOs may be data, special hardware, processing capabilities, Software and licenses.
- Members of a grid can be part of multiple VOs at the same time.



Grid Architecture Requirements

- Heterogeneity (OSes, Devs, Apps.)
- VO Resource Sharing (Management, Security and Accounting)
- Resource Utilisation (Reservation, Metering, Monitoring and Logging, Priorities)
- Job Execution (VO access, SLA)
- Data Services (Integration, Provisioning, Cataloguing, Metadata)
- **Security** (Authentication, Authorization)
- Administrative Costs (Provisioning, Deployment and Configuration)
- Scalability
- Availability (Disaster Recovery, Fault Management)
- Application Specific Requirements….

Terminology

- Computational Resource
  - Physical machines on which users wants to run their programs and store (or access) data files, without worrying about their physical location
- Job
  - A **computational task** (a binary application or script) that a user wants to run on the Grid, to retrieve results later
- Job Submission
  - It is the action of delegating the application to the Grid middleware for its execution.

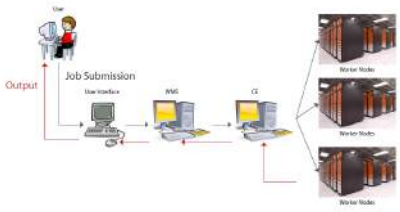
Main components services

- The User Interface (UI) is the user entry point.
- The Workload Management System (WMS) is a set of services which are in charge of finding the suitable resources for the submitted jobs and schedule the jobs.
- The Computing element (CE) is the computational resource, the entry point to a cluster or PCs handled by a job queue management system; in particular: TORQUE/PBS, LSF, CONDOR, SGE

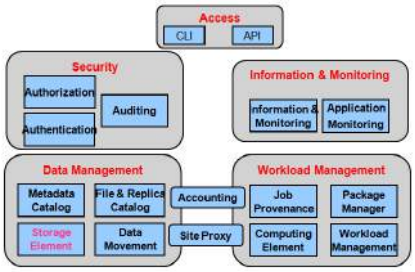
Main components services

- The Worker Nodes (WNs) are the machines where jobs are really executed.
- The Information System and Monitoring maintain data related to available grid resources and their health statuses, in order to optimize resource exploitation
- The Virtual Organization Management service (VOMS) is the way gLite improves the management of authentication and authorization to the Grid resources. The VOMS allows to their own members to define different access rights to VO's resources.
- The Storage element (SE) and the File catalogue (LFC) allow to manage Grid files and offer a mechanism to locate them easily for users and jobs.

Job life cycle

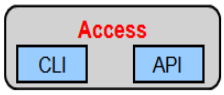


Middleware Services



gLite – Grid access

- Two possibilities: **APIs or CLI**
- Built on top of **GridPortals and GUIs**



gLite Security

- **Authentication based on X.509 PKI infrastructure**
  - Certificate Authorities (CA) issue (long lived) certificates identifying individuals (much like a passport).
  - Certification Authorities (CA) can generate service certificates which identify Grid services or hosts in the whole Grid.
  - In order to reduce vulnerability, on the Grid user identification is done by using (short lived-12 hours) proxies of their certificates.
- **The use of Proxy certificates allows the following:**
  - Delegation:
    - in which a remote service acts on behalf of the user as long as the delegation remains valid.
    - It is used to securely transfer user's credentials to a service.
    - This happens by making signed copies of the original proxy. (Single Sign On)
  - Add additional info (Add VO specific information provided by VOMS).
  - Possible to store a long term proxy on a secure server (MyProxy).

gLite Security: AutH/AhtZ

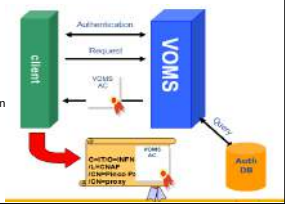
- **Authentication**
  - The user receive a certificate from a CA.
  - He/She connects to the UI (via SSH, portal).
  - He/She **Creates the proxy (single sign on)** – All grid services will use this proxy to identify the user.
- **Authorization**
  - The user has to subscribe to a VO (VOMS).
  - The VO establishes the user rights.
  - In any Grid service it will be verified if the user belongs to the VO and assigns the proper access rights to the user.

VOMS

- Virtual Organization Membership Service is a service that keeps track of the members of a VO.
- Support MyProxy (stored proxies).
- VOMS Grants users authorization to access the resource at VO level.

VOMS

- It extends the proxy certificates with info on VO membership, groups, roles(e.g. administrator, software engineer, student) & capabilities.
  - Each VO has a database containing group membership, roles and capabilities information for each user.
  - User contacts VOMS server requesting his authorization information.
  - Server sends signed authorization info to the client.
  - Client includes them in a proxy certificate.
  - Proxy is mapped on resources to local accounts



Clients

- voms-proxy-init –voms <VONAME> [options]
- voms-proxy-info
- voms-proxy-destroy

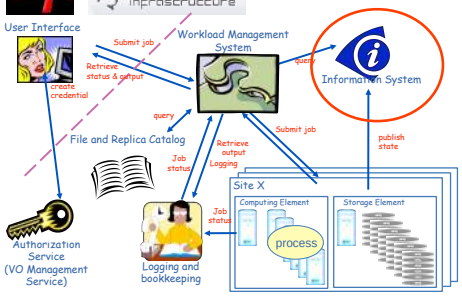
The gLite Information System

Emidio Giorgio

Information System and Monitoring

- **What?**
  - System to collect information on the state of resources.
- **Why?**
  - To discover resources of the grid and their nature.
  - To check for health status of resources.
  - To provide data in order to manage the workload more efficiently.
- **How?**
  - Monitoring and publishing fresh data on the state of resources.
- **Who?**
  - User searching specific resources for their activity.
  - Workload Management System.
  - Other monitoring system

Scope



2

Uses of the Information System

If you are a user

Retrieve information about
- Grid resources and status
- Resources that can run your job
- Status of your jobs

If you are a middleware developer

Workload Management System:
Matching job requirements and Grid resources

Monitoring Services:
Retrieving information about Grid Resources status and availability

If you are site manager or service

You "generate" the information for example relative to your site or to a given service

3

Evolution

- The data published in the Information System (IS) conforms to the GLUE (*Grid Laboratory for a Uniform Environment*) Schema. The GLUE Schema aims to define a common conceptual data model to be used for Grid resources.
- The BDII (*Berkeley DB Information Index*), based on an updated version of the Monitoring and Discovery Service (MDS), from Globus, is adopted as main provider of the Information Service.

4

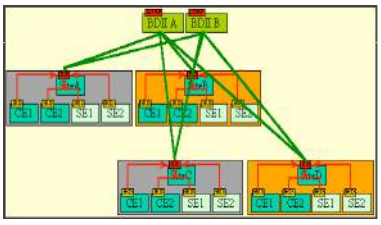
Interfaces

- **BDII Information System**
  - *main Information System for the current production grid*
  - **Sets of commands:**
    - **lcg-infosites:** simple, meets most of the user needs
    - **lcg-info**: supports more complex queries
    - **ldap clients** (ldapsearch, ldap browser) : mainly used by (advanced) sysadmins and developers, to check in depth details

5

LCG Information Service

- a user or a service can query
  - the BDII (usual mode)
  - LDAP servers on each site



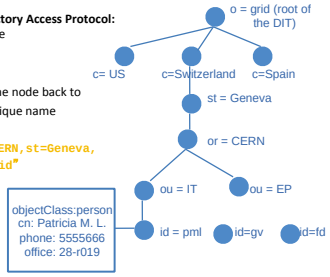
7

The LDAP Protocol

- ► **Lightweight Directory Access Protocol:** structures data as a tree
- ►

Following a path from the node back to the root of the DIT, a unique name is built (the DN):

`"id=pml,ou=IT,or=CERN,st=Geneva, c=Switzerland,o=grid"`



8

lcg-infosites

- The lcg-infosites command can be used as an easy way to retrieve information on Grid resources for most use cases.

USAGE:

lcg-infosites --vo <vo name> options -v <verbose level>
--is <BDII to query>

9

lcg-infosites options

ce	The information related to number of CPUs, running jobs, waiting jobs and names of the CEs are provided. All these data group all VOs together. With "-v 1" only the names of the queues will be printed while with "-v 2" The RAM Memory together with the operating system and its version and the processor included in each CE are printed.
se	The names of the SEs supported by the user's VO together with the kind of Storage System,the used and available space will be printed. With "-v 1" only the names of the SEs will be printed.
closeSE	The names of the CEs where the user's VO is allowed to run together with their corresponding closest SEs are provided.
lfc	Name of the lfc Catalog for the user's VO.
tag	The names of the tags relative to the software installed in site is printed together with the corresponding CE.
all	It groups together the information provided by ce, se, lrc and rmc.
is	If not specified the BDII defined in default by the variable LCG GFAL INFOSYS will be queries. However the user may want to query any other BDII without redefining this environment variable. This is possible specifying this argument followed by the name of the BDII which the user wants to query. All options admits this argument.

10

PRACTICAL

- Will introduce use of the commands to access the information system
- You will be exploring the VO gridit, checking which sites (computing and storage elements) support this VO.

11

lcg-info options

--list-attrs	Prints a list of the attributes that can be queried.
--list-ce	Lists the CEs which satisfy a query, or all the CEs if no query is given.
--list-se	Lists the SEs which satisfy a query, or all the SEs if no query is given.
--query	Restricts the output to the CEs (SEs) which satisfy the given query.
--bdii	Allows to specify a BDII in the form :. If not given, the value of the environmental variable LCG_GFAL_INFOSYS is used. If that is not defined, the command returns an error.
--sed	Print the output in a "sed-friendly" format.
--attrs	Specifies the attributes whose values should be printed.
--vo	Restricts the output to CEs or SEs where the given VO is authorized. Mandatory when VO-dependent attributes are queried upon.

12

Get the list of supported attributes

$ lcg-info --list-attrs

Attribute name	Glue object class	Glue attribute name
MaxTime	GlueCE	GlueCEPolicyMaxWallClockTime
CEStatus	GlueCE	GlueCEStateStatus
TotalJobs	GlueCE	GlueCEStateTotalJobs
CEVOs	GlueCE	GlueCEAccessControlBaseRule
TotalCPUs	GlueCE	GlueCEInfoTotalCPUs
FreeCPUs	GlueCE	GlueCEStateFreeCPUs
CE	GlueCE	GlueCEUniqueID
WaitingJobs	GlueCE	GlueCEStateWaitingJobs
RunningJobs	GlueCE	GlueCEStateRunningJobs
CloseCE	GlueCESEBindGroup	GlueCESEBindGroupCEUniqueID
CloseSE	GlueCESEBindGroup	GlueCESEBindGroupSEUniqueID
SEVOs	GlueSA	GlueSAAccessControlBaseRule
UsedSpace	GlueSA	GlueSAStateUsedSpace
AvailableSpace	GlueSA	GlueSAStateAvailableSpace
Type	GlueSE	GlueSEType
SE	GlueSE	GlueSEUniqueID
Protocol	GlueSEAccessProtocol	GlueSEAccessProtocolType
ArchType	GlueSL	GlueSLArchitectureType
Processor	GlueSubCluster	GlueHostProcessorModel
OS	GlueSubCluster	GlueHostOperatingSystemName
Cluster	GlueSubCluster	GlueSubClusterUniqueID
Tag	GlueSubCluster	GlueHostApplicationSoftwareRunTimeEnvironment
Memory	GlueSubCluster	GlueHostMainMemoryRAMSize

13

lcg-info examples

List all the CE(s) that can run MPICH, giving the number of free CPUs and the tags of installed software

$ lcg-info --vo gridit--list-ce --query 'Tag=MPICH' --attrs 'FreeCPUs,Tag'

Careful here!

No space allowed here!

```
-.....
CE: grid-ce.bio.dist.unige.it:2119/jobmanager-lcgpbs-long
 - FreeCPUs        6
 - Tag             LCG-2
                   LCG-2_1_0
                   LCG-2_1_1
....
```

14

Grid data management

Emidio Giorgio

Outline

- Motivation
- The Grid data management challenge
- Data management components
- LFC File Catalog
  - Hands on LFC
- LCG Utils commands
  - Hands on lcg-utils

Motivation

Data Intensive Sciences **depend** on Grid Infrastructures

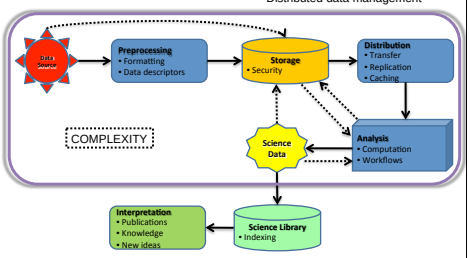
Characteristics:

- Data is inherently **distributed**
- Data is produced in **large quantities**
- Data is produced at a **very high rate**
- Data has **complex** interrelations
- Data is **needed** by many people

single (person / computers) cannot do all the work
Several groups collaborates in data analysis

High-level data processing scenario

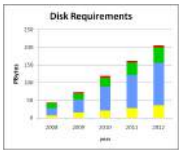
Distributed data management

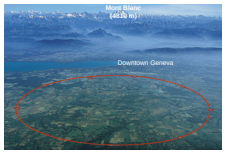


High Energy Physics

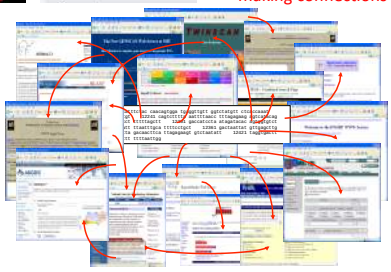
Large Hadron Collider (LHC) at CERN

- One of the most powerful instruments ever built to investigate matter
- 4 Experiments: ALICE, ATLAS, CMS, LHCb
  - 4 Virtual Organizations
- 27 km circumference tunnel
- Generating 10PB/year





Biomedical data – making connections

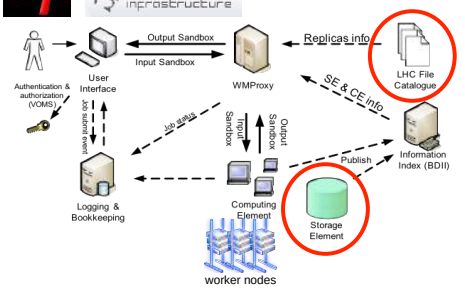


The Grid data management challenge

- Heterogeneity
  - Data are stored on different storage systems using different access technologies
- Distribution
  - Data are stored in different locations – in most cases there is no shared file system or common namespace
  - Data need to be moved between different locations

- Need common interface to storage resources
  - Storage Resource Manager (SRM)
- Need to keep track where data are stored
  - File and Replica Catalogs
- Need scheduled, reliable file transfer
  - File Transfer Service

gLite/EMI components

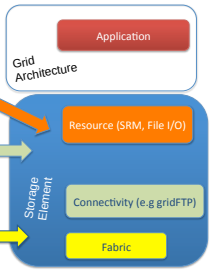


Data management components

- The **Storage Element (SE)** is the service which allows users and applications to store/retrieve data (files)
- Data Management Services provide services for location, access and transfer of files
  - Users do not need to know file location, just its logical name
  - Files can be replicated or transferred to other SE, as needed
  - Files are shared within VO
- Files are write-once, read-many
  - Files cannot be changed unless removed or replaced

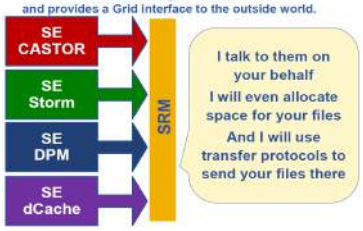
Storage Element

- Functions:
  - File I/O service: needed by the end-user application in order to access its data
  - SRM service implementation on the given storage
  - Transfer service for various transfer protocol(s)
  - Auxiliary Security and Logging services
  - Storage back-end with all the associated hardware and drivers



Storage Resource Manager (SRM)

The SRM is a single interface that takes care of *local storage interaction* and provides a Grid interface to the outside world.



Data management components : Files

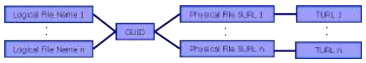
- Files located in SE:
  - Write once, read many (cannot be changed unless replaced)
  - Accessible by users and application from anywhere
  - Several replicas of one file can be replicated at different sites
- SE:
  - Provide storage space
  - Support protocols for remote IO-transfer-access
  - Provide interface for management of disk and tape storage resources (SRM)

File Naming

- Files (and in general, data) may be replicated at many Grid sites if they are heavily used.
  - Upon creation, each file also acquires an immutable **Grid Unique ID (GUID)**. The GUID is *unique* by construction.
  - The user application does not need to know which locations these are as long as it is able to run and read the data as if it were local.
  - Therefore the name by which the user refers to the file has to be a location-independent logical file name (LFN). The users will only operate on this LFN namespace, as if it was a single virtual file system.
- The applications may use either the LFN or GUID to identify their files:
  - GUID, always mandated by the system
  - LFN, optionally assigned by the users.
    - Some files may not have LFNs at all.

Name conventions

- **Logical File Name (LFN)**
  - An alias created by a user to refer to some item of data, e.g. lfn:/grid/gilda/budapest23/run2/track1
- **Globally Unique Identifier (GUID)**
  - A non-human-readable unique identifier for an item of data, e.g. guid:f81d4fae-7dec-11d0-a765-00a0c91e6bf6
- **Site URL (SURL) (or Physical File Name (PFN) or Site FN)**
  - The location of an actual piece of data on a storage system, e.g. srm://pcrd24.cern.ch/flatfiles/cms/output10_1 (SRM)
- **Transport URL (TURL)**
  - Temporary locator of a replica + access protocol: understood by a SE, e.g. rfio://lxshare0209.cern.ch//data/alice/ntuples.dat



Name conventions

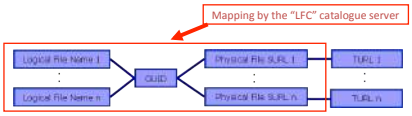
Users primarily access and manage files through "logical filenames" - LFN

LFN has a directory tree structure

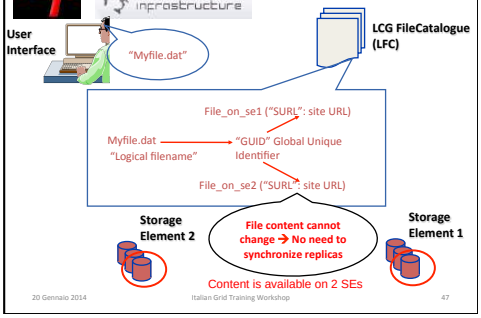
lfn:/grid/<VO_name>/ <you create it>

LFN Namespace | **Defined by the user**

Mapping by the "LFC" catalogue server



Resolving LFN



File Naming /2

- The LFN is the key by which the users usually locate the actual locations of their files.
  - The replicas are identified by **Site URLs (SURLs)**. Each replica has its own SURL, specifying implicitly which Storage Element needs to be contacted to extract the data. The SURL is a valid URL that can be used as an argument in an SRM interface. An example SURL is:
    - srm://castorgrid.cern.ch:8443/srm/managerv1?SFN=/castor/cern.ch/file1
- SRM end-point : Site URL (SURL)
  - Usually, users are not directly exposed to SURLs, but only to the logical namespace defined by LFNs.
    - The Grid Catalogs provide mappings needed for the services to actually locate the files.
    - The user perceives a single file system.

File Naming /3

- **LFN (Logical File Name):** *logical*(*human* readable) *identifier* for a file. LFNs are unique, but mutable, i.e. they can be changed by the user. The namespace of the LFNs is a global hierarchical namespace.
  - Each Virtual Organisation can have its own namespace (everybody in the VO needs to agree with this convention).
  - One LFN may have many symbolic links pointing to it. Symbolic links have to be specified using absolute paths.
- **GUID (Global Unique Identifier):** A logical identifier, which guarantees its uniqueness by construction.
  - Each LFN also has a GUID (1:1relationship).
  - GUIDs are immutable, i.e. they cannot be changed by the user, otherwise consistency cannot be assured.

File Naming/4

- **SURL (Site URL)** specifies a physical instance of a file replica. Formerly known as the Physical File Name (PFN).
  - SURLs are accepted by the Storage Element's SRM interface. Their schema is therefore srm.
  - A file may have many replicas, hence the mapping between GUIDs <-> SURLs is 1:N.
- **TURL (Transport URL)** It is an URL that can be used to actually transfer a file using any standard transport protocol.
  - The TURL is a fully qualified URL starting with the protocol to be used for transfer.

LFC Catalog commands

lfc-chmod	Change access mode of the LFC file/directory
lfc-chown	Change owner and group of the LFC file-directory
lfc-delcomment	Delete the comment associated with the file/directory
lfc-getacl	Get file/directory access control lists
lfc-ln	Make a symbolic link to a file/directory
lfc-ls	List file/directory entries in a directory
lfc-mkdir	Create a directory
lfc-rename	Rename a file/directory
lfc-rm	Remove a file/directory
lfc-setacl	Set file/directory access control lists
lfc-setcomment	Add/replace a comment

LFC hands on

Several environment variables need to be set before you start to ensure that the correct catalog service is used. The default settings for these variables in your account should be correct, however this needs to be checked.
The variables which need to be checked having these exact values are:

```
$LCG_CATALOG_TYPE
$LFC_HOST
$LCG_GFAL_INFOSYS
```

If one or more of them has different or empty value, please set it (them) in this way:

```
export LCG_CATALOG_TYPE=lfc
export LCG_GFAL_INFOSYS=egee-bdii.cnaf.infn.it:2170
```

Now, you are ready to start (replace dteam with your_vo and LFC_HOST with the LFC server of your VO).

lcg-infosites –vo <your_vo> lfc

```
export LFC_HOST=lfcserver.cnaf.infn.it #VO dependent
```


LFC hands on

```
$ lfc-mkdir /grid/gridit/emidio

$ lfc-setcomment /grid/gridit/emidio "Emidio Work Dir"

$ lfc-ls -l --comment /grid/gridit | grep emidio
drwxrwxr-x   0 536      107                          0 May 04 08:53
   emidio Emidio Work Dir
$ lfc-getacl /grid/gridit/emidio
# file: /grid/gridit/emidio
# owner: /C=IT/O=INFN/OU=Personal Certificate/L=Catania/CN=Emidio Giorgio
# group: gridit
user::rwx
group::rwx              #effective:rwx
other::r-x
default:user::rwx
default:group::rwx
default:other::r-x
$ export LFC_HOME=/grid/gridit/emidio
```


LCG Utils hands on

```
$ dd if=/dev/urandom of=bigfile.dat bs=100k count=100

$ lcg-infosites –vo gridit se |grep infn
lcg-infosites --vo gridit se | grep infn
      297481969         96356864  SRM   aliserv6.ct.infn.it
     6339900447       2169063236  SRM   atlasse.lnf.infn.it
     3969623523              n.a  SRM   boalice5.bo.infn.it
       98318539          1681460  SRM   emi-se.pr.infn.it
       66752607        471977687  SRM   grid-se.lns.infn.it
    10001905090              n.a  SRM   grid-se2.pr.infn.it
```

1. Upload to Grid (lcg-cr);
2. Test SURL (lcg-lr)
3. Replicate a file (lcg-rep);
4. Download from Grid (lcg-cp);

LCG Utils hands on upload file to Grid

```
$ lcg-cr -v --vo gridit -d prod-se-02.pd.infn.it -l lfn:my_first_gridfile.dat \
    file:/home/giorgio/bigfile.dat
Using grid catalog type: lfc
Using grid catalog : lfcserver.cnaf.infn.it
Checksum type: None
SE type: SRMv2
Destination SURL : srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/file204ce41b-8aa3-457b-80a2-
    a265cc33a595
Source SRM Request Token: 2a1fa104-56e0-4b08-b3f1-259e9ce498e6
Source URL: file:/home/giorgio/bigfile.dat
File size: 10240000
VO name: gridit
Destination specified: prod-se-02.pd.infn.it
Destination URL for copy: gsiftp://prod-se-02.pd.infn.it/prod-se-02.pd.infn.it:/flatfiles/SE00/gridit/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595.1519405.0
# streams: 1
      9437184 bytes   2645.70 KB/sec avg   4266.67 KB/sec inst
Transfer took 4050 ms
Using grid catalog type: lfc
Using grid catalog : lfcserver.cnaf.infn.it
Site URL to be registered: srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595
File size: 10240000
Using LFN: lfn:/grid/gridit/emidio/my_first_gridfile.dat
Using GUID: guid:3b2b10fb-1285-43a6-8149-3b1b1b7002d4
Registering LFN: /grid/gridit/emidio/my_first_gridfile.dat (3b2b10fb-1285-43a6-8149-3b1b1b7002d4)
Registering SURL: srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/file204ce41b-8aa3-457b-80a2-
    a265cc33a595 (3b2b10fb-1285-43a6-8149-3b1b1b7002d4)
$ lcg-lr lfn:my_first_gridfile.dat
srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595
```

LCG Utils hands on replicate a file

```
$ lcg-rep -v -vo gridit  -d se.scope.unina.it  lfn:my_first_gridfile.dat
Using grid catalog type: LFC
Using grid catalog : lfcserver.cnaf.infn.it
VO name: gridit
Checksum type: None
Trying SURL srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/file204ce41b-8aa3-457b-80a2-
    a265cc33a595 ...
Source SE type: SRMv2
Source SRM Request Token: f2aa2c26-188b-4eac-ad51-eed84ea5b8e1
Destination SE type: SRMv2
Destination specified: se.scope.unina.it
Source URL for copy: gsiftp://prod-se-02.pd.infn.it/prod-se-02.pd.infn.it:/flatfiles/SE00/gridit/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595.1519405.0
Destination URL: gsiftp://se.scope.unina.it/se.scope.unina.it:/egee02/gridit/2012-05-04/filed5a00900-d7c2-42b5-
    a1b5-7595a9599b52.2767446.0
# streams: 1
      10240000 bytes   7142.86 KB/sec avg   7142.86 KB/sec inst
Transfer took 3630 ms
Using LFN: lfn:/grid/gridit/emidio/my_first_gridfile.dat
Using GUID: guid:3b2b10fb-1285-43a6-8149-3b1b1b7002d4
Registering SURL: srm://se.scope.unina.it/dpm/scope.unina.it/home/gridit/generated/2012-05-04/filed5a00900-d7c2-42b5-
    a1b5-7595a9599b52 (3b2b10fb-1285-43a6-8149-3b1b1b7002d4)
Destination URL registered in file catalog: srm://se.scope.unina.it/dpm/scope.unina.it/home/gridit/generated/2012-05-04/
    filed5a00900-d7c2-42b5-a1b5-7595a9599b52
$ lcg-lr lfn:my_first_gridfile.dat
srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595
srm://se.scope.unina.it/dpm/scope.unina.it/home/gridit/generated/2012-05-04/filed5a00900-d7c2-42b5-
    a1b5-7595a9599b52
```


LCG Utils hands on download from Grid

```
$ lfc-ls -l /grid/gridit/emidio
-rw-rw-r--   1 536      107             10240000 May 04 09:27 my_first_gridfile.dat
$ lcg-cp --vo gridit -v lfn:my_first_gridfile.dat file:/home/giorgio/
    bigfile2.dat
Using grid catalog type: LFC
Using grid catalog : lfcserver.cnaf.infn.it
VO name: gridit
Checksum type: None
Trying SURL srm://prod-se-02.pd.infn.it/dpm/pd.infn.it/home/gridit/generated/2012-05-04/
    file204ce41b-8aa3-457b-80a2-a265cc33a595 ...
Source SE type: SRMv2
Source SRM Request Token: c9af51b7-6329-45fa-af01-95da17e5673a
Source URL: /grid/gridit/emidio/my_first_gridfile.dat
File size: 10240000
Source URL for copy: gsiftp://prod-se-02.pd.infn.it/prod-se-02.pd.infn.it:/flatfiles/SE00/
    gridit/2012-05-04/file204ce41b-8aa3-457b-80a2-a265cc33a595.1519405.0
Destination URL: file:/home/giorgio/bigfile2.dat
# streams: 1
      5242880 bytes   2558.81 KB/sec avg   2558.81 KB/sec inst
$ ls -la bigfile2.dat
-rw-r--r-- 1 giorgio users 10240000  4 mag  2012 bigfile2.dat
```

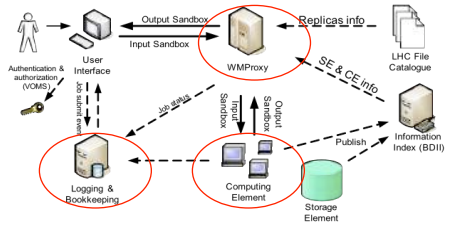

Thank you

Questions ?

Grid Job management

Vania Boccia
INFN-Napoli

gLite/EMI components

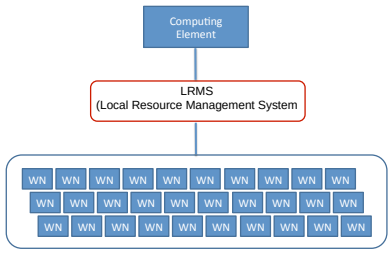


Workload Management System (WMS)

Workload Management System (WMS) comprises a set of Grid middleware components responsible for distribution and management of tasks across Grid resources:

- The Workload Manager(WM) aims to accept and satisfy requests for job management coming from its clients.
- WM will pass the job to an appropriate CE for execution taking into account requirements and the preferences expressed in the job description.
- The decision of which resource should be used is the outcome of a matchmaking process.
- The Logging and Bookkeeping service tracks jobs managed by the WMS. It collects events from many WMS components and records the status and history of the job.

What is a Computing Element in GRID



Ways to submit a job on GRID

There are two way to submit a job on the GRID:

- Directly to a Computing Element (if the user knows the right endpoint, given from the information system
- By means of Workload Management System that is devoted to transparently manage the whole job life:
  - Resource allocation
  - Job submission
  - Proxy renewal (together with MyProxy service)
  - Job monitoring (together with Logging&Bookkeeping service)
  - Job output retrival

How to describe a JOB?

- By means of the Job Description Language (JDL) a special syntax able to describe all the job needs:
  - physical (disk space, RAM, CPUs, …)
  - logical (operative system, libraries, software, compilers, …)
  - job data (input, output)
- By using the JDL syntax the user has to write a text file with extension *.jdl*

JDL

- The Job Description Language is a high-level language used to describe jobs and aggregates of jobs with arbitrary dependency relations.
- The JDL is used to specify the desired job characteristics and constraints, which are taken into account by the WMS to select the most suitable resource to execute the job.
- The JDL file consists of a certain number of lines having the format:

  attribute = expression;

The JDL attributes

Two categories:

- job attributes:
  - define job characteristics
- resource attributes:
  - computing resource attributes: specify job requirements in terms of computing resources needs
  - storage resource attributes: specify job requirements related to storage resources (i.e. access protocol, logical file names, ...)

JDL attributes overview

- Executable:
  - specifies the command to be run by the job. If the command is already present on the WN, it must be expressed as a absolute path; if it has to be copied from the UI, only the file name must be specified, and the path of the command on the UI should be given in the InputSandbox attribute.
- Arguments:
  - can contain a string value, which is taken as argument list for the executable
- The attributes StdOutput and StdError define the name of the files containing the standard output and standard error of the executable, once the job output is retrieved.

JDL attributes overview

- If files have to be copied from the UI to the execution node, they must be listed in the InputSandbox attribute:

  InputSandbox = {"test.sh", .. ,"fileN"};

- The files to be transferred back to the UI after the job is finished can be specified using the OutputSandbox attribute:

  OutputSandbox = {"std.out","std.err"};

- The Requirements attribute can be used to express constraints on the resources where the job should run. Its value is a Boolean expression that must evaluate to true for a job to run on that specific CE.

Simple JOB: Hello.jdl

```
# the name of the file which will be actually executed
Executable = "hello.sh" ;
# the name of the file where standard output stream will be redirected
StdOutput = "stdout.txt";
# the name of the file where standard error stream will be redirected
StdError = "stderr.txt";
# list of local files which will be transferred to the  remote resource before the execution
InputSandbox = {"hello.sh"};
# list of remote files which will be transferred from the resource # after the execution
OutputSandbox = {"stdout.txt", "stderr.txt"}
```


Hello.sh

```
#!/bin/sh
echo "Hello World - from $HOSTNAME"
echo -n "It's " date
```

Hello.sh is a Bash script to be executed.

The script is transferred to the remote resource, executed, and the standard output/error are streamed on the files *stdout.txt* and *stderr.txt*, which are included in the job Output Sandbox for the successive retrieval.

Parametric JOB: param.jdl

```
JobType = "Parametric";
Executable = "Stampa-param.sh";
Arguments = "_PARAM_";
InputSandbox =  "Stampa-param.sh";
Parameters = {0,2,4,6};
StdOutput = "std.out";
StdError = "std.err";
OutputSandbox  = {"std.out", "std.err"};
```

Assume for example that we want to execute an executable named Stampa-param.sh with different command line parameters every time (i.e. similar to executing ./Stampa-param.sh 0, . ./Stampa-param.sh 2, ./Stampa-param.sh 4 and ./ Stampa-param.sh 6 one after the other).

The executable Stampa-param.sh is a simple bash script that prints the parameter value.

Parallel JOB: mpi.jdl

```
Executable = "mpi-start-wrapper.sh";
Arguments = " cpi";
CpuNumber = 4;
StdOutput = "mpi.out";
StdError = "mpi.err";
InputSandbox = {"mpi-start-wrapper.sh","mpi-hooks.sh", "cpi.c"};
OutputSandbox = {"mpi.err","mpi.out"};
Requirements = Member("MPI-START",
other.GlueHostApplicationSoftwareRunTimeEnvironment);
```

The executable mpi-start-wrapper.sh is a bash script invoking an other bash script (mpi-hooks.sh to compile code cpi.c on Worker Node before execution.

Note the use of the resource attribute Requirements used to require for a Computing Element that supports MPI.

Job Monitoring

JOB states:

- Submitted (arrived on WMS)
- Waiting (authenticated on WMS)
- Ready (processed by WMS)
- Scheduled (on CE's LRMS queue)
- Running (in execution on WN)
- Done (Job executed)
- Aborted (job aborted by the WMS, i. e. due to a proxy expiration)
- Cancelled (job canceled by the user)
- Cleared (when the job output has been removed from the user or from the WMS)

CLI to submit a JOB

Obtains the CEs list that matches job requirements
- **glite-wms-job-list-match myfile.jdl**

Submits a job and prompt the unique JOB ID
- **glite-wms-job-submit –a myfile.jdl**

Checks the JOB status
- **glite-wms-job-status <job_id>**

When a JOB is in “Done” state, allows you to retrieve JOB output
- **glite-wms-job-output <job_id>**

Allows to cancel submitted Jobs
- **glite-wms-job-cancel <job_id>**

hands on download from Grid

```
[egiorgio@ui2 ~]$ glite-wms-job-submit -a hello.jdl
Connecting to the service https://gridrb.fe.infn.it:7443/glite_wms_wmproxy_server
====================== glite-wms-job-submit Success ======================
The job has been successfully submitted to the WMProxy Your job identifier is: https://gridrb.fe.infn.it:9000/nW8jSAEeZYKtO0j91ZU8YQ
==========================================================================

[egiorgio@ui2 ~]$ glite-wms-job-submit -a -o jobid.txt hello.jdl
Connecting to the service https://wms005.cnaf.infn.it:7443/glite_wms_wmproxy_server
====================== glite-wms-job-submit Success ======================
The job has been successfully submitted to the WMProxy Your job identifier is:
https://wms005.cnaf.infn.it:9000/a-UJrAT2A3zSkScTngAc6Q
The job identifier has been saved in the following file: /home/egiorgio/jobid.txt
==========================================================================
```


hands on download from Grid

```
[egiorgio@ui2 ~]$ glite-wms-job-status https://gridrb.fe.infn.it:9000/nW8jSAEeZYKtO0j91ZU8YQ
====================== glite-wms-job-status Success ======================
BOOKKEEPING INFORMATION:
Status info for the Job : https://gridrb.fe.infn.it:9000/nW8jSAEeZYKtO0j91ZU8YQ
Current Status: Scheduled
Status Reason: unavailable
Destination: gridce4.pi.infn.it:8443/cream-lsf-grid
Submitted: Tue Jan 14 10:00:19 2014 CET
==========================================================================

[egiorgio@ui2 ~]$ glite-wms-job-status -i jobid.txt
====================== glite-wms-job-status Success ======================
BOOKKEEPING INFORMATION:
Status info for the Job : https://wms005.cnaf.infn.it:9000/a-UJrAT2A3zSkScTngAc6Q
Current Status: Running Status
Reason: unavailable
Destination: ce-1.le.infn.it:8443/cream-lsf-gridit
Submitted: Tue Jan 14 11:34:42 2014 CET
==========================================================================
```


hands on download from Grid

```
[egiorgio@ui2 ~]$ glite-wms-job-status -i jobid.txt
====================== glite-wms-job-status Success ======================
BOOKKEEPING INFORMATION:
Status info for the Job : https://wms005.cnaf.infn.it:9000/a-UJrAT2A3zSkScTngAc6Q
Current Status: Done(Success)
Logged Reason(s): - job completed - Job Terminated Successfully Exit code: 0
Status Reason: Job Terminated Successfully
Destination: ce-1.le.infn.it:8443/cream-lsf-gridit
Submitted: Tue Jan 14 11:34:42 2014 CET
==========================================================================

[egiorgio@ui2 ~]$ glite-wms-job-output -i jobid.txt --dir .
Connecting to the service https://wms005.cnaf.infn.it:7443/glite_wms_wmproxy_server
==========================================================================
JOB GET OUTPUT OUTCOME Output sandbox files for the job: https://wms005.cnaf.infn.it:9000/a-UJrAT2A3zSkScTngAc6Q have been successfully retrieved and stored in the directory: /home/egiorgio/egiorgio_a-UJrAT2A3zSkScTngAc6Q
==========================================================================

[egiorgio@ui2 ~]$ cat /home/egiorgio/egiorgio_a-UJrAT2A3zSkScTngAc6Q/std stderr.txt stdout.txt

[egiorgio@ui2 ~]$ cat /home/egiorgio/egiorgio_a-UJrAT2A3zSkScTngAc6Q/stdout.txt
Hello World - from lx25.le.infn.it it's Tue Jan 14 11:35:01 CET 2014
```


Thank you

Questions ?

2. Use Case 1: CRYSTAL

2.1 Introduction to CRYSTAL

CRYSTAL14. Program overview
Silvia Casassa, University of Tourin, Italy

The capabilities of the Crystal14 program have been presented in the light of its possible use a standard Grid tool.
Crystal14 is an ab initio code that uses a Gaussian-type basis set: both pseudopotential and all-electron strategies are permitted. A variety of density functionals is available, including as an extreme case Hartree-Fock; hybrids of various nature (global, range-separated, double) can be used. The program can treat on the same grounds 0D (molecules), 1D (polymers), 2D (slabs), as well as 3D (crystals) systems. Symmetry is fully exploited at all steps of the calculation: this permits, for example, to investigate nanotubes of increasing radius at a nearly constant cost or to perform Self-Consistent-Field calculations on fullerenes with 6000 atoms, 84000 atomic orbitals and 20 SCF cycles, on a single core in one day. Three versions of the code exist, serial, parallel and massive-parallel. Many tensorial properties can be evaluated in a fully automated way by using a single input keyword: elastic, piezoelectric, photoelastic, dielectric, as well first and second hyperpolarizabilies, electric field gradients, Born tensors, etc. Many tools permit a complete analysis of the vibrational properties of crystalline compounds. The infrared and Raman intensities are now computed analytically and related spectra can be generated. Isotopic shifts are easily evaluated, frequencies of only a fragment of a large system computed and nuclear contribution to the dielectric tensor determined.
New algorithms have been devised for the investigation of solid solutions and disordered systems. The topological analysis of the electron charge density, according to the Quantum Theory of Atoms in Molecules, is now incorporated in the code via the integrated merge of the Topond package. Electron correlation can be evaluated at the M oller-Plesset second order level (namely MP2) and a set of double-hybrids are presently available via the integrated merge with the Cryscor program.

In order to show limits and capabilities of CRYSTAL14, as concern in particular its use on the Grid, ice XI has been chosen as a test case. Ice XI is a hydrogen-bonded solid, namely a proton-ordered phase of Ice stable at low pressure and temperature. The hybrid B3LYP functional has been chosen as default and different input files parameters in any Gaussian-type basis functions code, on the fundamental properties (geometry and energy) as well as on the vibrational spectrum of ice XI. The following procedure has been envisaged:

- starting from the experimental geometry, a convergence with respect to the selected computational parameters, i.e.: integral tolerances and shrinking factors (please refer to the CRYSTAL Manual on http : www:crysta:unito:it, for further details);
- then, a full geometry optimization is performed in order to achieve the minimum in the surface potential;
- finally, a frequency calculation on the optimized geometry will be submitted to get the complete IR and Raman spectrum in gamma and the related intensities.

The users are requested to the computational strategy is also reported:

1. at the end of each CRYSTAL14 calculation it is worth to save the Fortran unit 9 (fort.9);
2. fort.9 can be processed by CRYSTAL14 to get a lot of properties (see the Manual) or as a guess for a next run;

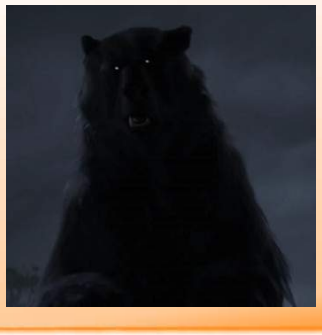
3. moreover, at the end of a geometry optimization, the final configuration is saved on the Fortran unit 34 (fort.34) which can be used as a starting point for a frequency run;
4. frequencies and intensities are reported on CRYSTAL14 output on the unit FREQINFO which can be easily interfaced to many visualization software (see the web site for references).

Ice XI	Energy	CPU time (8 proc)	OPTG	FREQ
631dp		save 9,34	res from fort.34; save FREQINFO..	
6311dp		save 9,34	res from fort.34; save FREQINFO..	
Tzp (pol)		save 9,34	res from fort.34; save FREQINFO..	

Table 1: Basis set and computational strategies.

Aknowledgments.
S. Casassa and the Organizers aknowledge the CRYSTAL group for the support given

→in order to face a problem ....



... in the forest of Quantum-Mechanical code

we need to follow a Wisp!

CRYSTAL14

a Program for the *ab initio* investigation of Crystalline Solids

S. Casassa

Università di Torino, Theoretical Chemistry Group

Summary

- Background →original tools
  - ⋆ Basis set
  - ⋆ Symmetry
  - ⋆ Hamiltonians
- Old and new Features →what can be calculated
  - ⋆ geometry and energy
  - ⋆ one-electron
  - ⋆ IR and Raman spectra, optical, elastic, topological properties, ..
- Performances →at which price
  - in time and
  - accuracy
- Technicalities →and how

Basis set

- Gaussian, *s,p,d and f*
- all electron + pseudopotential
- `http://www.crystal.unito.it/basis-sets.php` - M. Peintinger
- chemical insight
- basis set superposition error (BSSE)

$$\Psi_i^{\boldsymbol{k}}(\boldsymbol{r}) = \sum_\mu \sum_{\boldsymbol{g}} c(\boldsymbol{k})_{\mu i} \exp(i\boldsymbol{k}\boldsymbol{g})\, \chi_\mu^{\boldsymbol{g}}$$

Geometry

- Periodicity and periodic boundary condition
- 0D → 1D → 2D → 2D
- Full exploitation of Symmetry
  - ⋆ in the reciprocal space **k**
  - ⋆ in the direct space **g**
- geometry manipulation, i.e insertion, displacement, substitution, deletion of atoms

Hamiltonians

♡ DFT

↓

♡ Hybrid ↣ Grimme

↑

♡ HF ↣ MP2

now that we know our weaponry



the adventure can begin

General Features (1)

✓ Single-point energy calculation
- ⋆ one-electron properties (Band, DOS, potential)

✓ Automated geometry optimization
- ⋆ transition state search
- ⋆ constant volume and pressure OPTG

✓ Harmonic vibrational frequencies
- ⋆ at G point
- ⋆ phonon dispersion with a supercell approach
- ⋆ IR and Raman intensity

General Features (2)

✓ Equation of State

✓ Elastic Tensor

✓ Dielectric Constants calculation via CPHF/KS

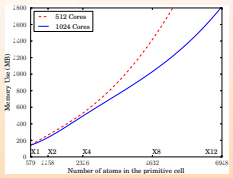
✓ Automatic treatment of Solid Solution

✓ Compton profiles

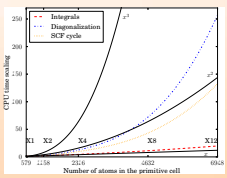
✓ Topological analysis of $\rho(\mathbf{r})$

✓ Perturbative approach to electron correlation (MP2)

Performances

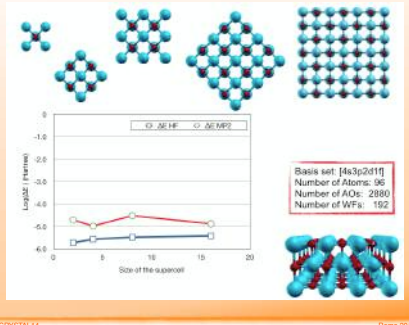


Memory Use

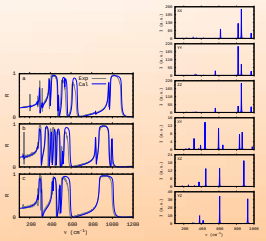


CPU time

Accuracy



Results



IR Spectra Raman Spectra

http://www.crystal.unito.it/animations-of-vibrational-modes.php

Technicalities

- Input/Output
  - title/geometry/BS/SCF-keywords
- Executables
  - > *crystal* < nomefile.d12 > nomefile.out
  - > *properties* < nomefile.d3 > nomefile.outp
- Fortran Units
  - fort.9
  - fort.25
  - FREQ.DAT, OPTG.DAT, INTENS.DAT, ..

http://www.crystal.unito.it/

Acknowledges

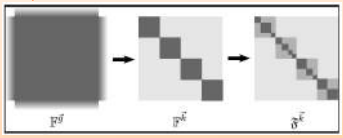
☺ the CRYSTAL group

- R. Dovesi, C. Pisani, C. Roetti, V.R. Saunders,
- R. Orlando, C.M. Zicovich-Wilson,
- B. Civalleri, A. Erba, S. Casassa, L. Maschio,
- M. Rerat, B. Kirtman, N. Harrison, P. D'Arco,
- M. Llunell

Acknowledges

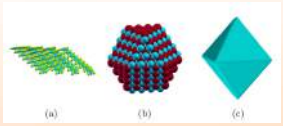
... ☺☺☺ and You all for Your kind attention!

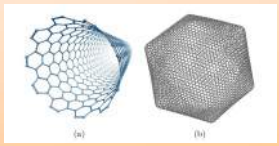
⇝ in reciprocal space



↤

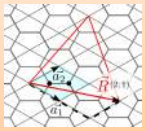
⇝ in direct space

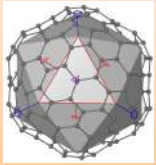




⇝ the construction of fullerene from graphene

```
SLAB
77
2.4612
1
6  -0.33333333333 0.33333333333 0.
FULLE
2 2
IH
ICOSA
...
```





↤

Input

GEOMETRY

```
GIACCIO ORTOROMBICO 36 / BULK / BASE 6-31 Pople
CRYSTAL
0 0 0
36
    4.6019      7.9978      7.528
  5
  8     0.5    -0.1691   -0.0691
  8     0.0    -0.3362    0.0648
  1     0.5     0.1586    0.1885
  1     0.5     0.0576    0.0230
  1    -0.333  -0.2167   -0.0191
ATOMINSE
1
17 0.4 0.4 0.0
OPTG
FULLOPTG
FREQCALC
INTENS
ENDG
```


Input

BASIS SET

```
8 3
0 0 6 2.0 1.00
  0.8588500000D+04  0.1895150000D-02
  0.1297230000D+04  0.1438590000D-01
  0.2992960000D+03  0.7073200000D-01
  0.8737110000D+02  0.2400610000D+00
  0.2567890000D+02  0.5947970000D+00
  0.3740040000D+01  0.2808020000D+00
0 1 3 6.0 1.00
  0.4211750000D+02  0.1138890000D+00  0.3651140000D-01
  0.9628370000D+01  0.9208110000D+00  0.2371530000D+00
  0.2853320000D+01 -0.3274470000D-02  0.8197020000D+00
0 3 1 0.0 1.00
  0.1292000000D+01  0.1000000000D+01
1 2
0 0 3 1.0 1.00
  0.3386500000D+02  0.2549380000D-01
  0.5094790000D+01  0.1903730000D+00
  0.1158790000D+01  0.8521610000D+00
0 0 1 0.0 1.00
  0.3258400000D+00  0.1000000000D+01
99 0
ENDBS
```


Input

SCF PARAMETERS

```
DFT
B3LYP
XLGRID
END
TOLINTEG
7 7 7 10 18
SHRINK
4 4
FMIXING
45
TOLDEE
10
SCFDIR
ENDSCF

↤
```

2.2 CRYSTAL porting to Grid

CRYSTAL. Porting to grid
Alessandro Costantini, INFN-Perugia and IGI (Italy)

In the adopted use case we considered the ice XI structure which is a hydrogen-bonded solid, namely a proton-ordered phase of ice, stable at low pressure and temperature.

The approach followed to build for CRYSTAL14 a graphical interface suitable for the Grid environment, able to execute both the sequential and parallel versions of the program (and related CRYSTAL tools) on Grid resources, has been based on the use of the WS-PGRADE framework already implemented into the IGI Portal used for submit and monitor the job.

The framework, in fact, allows a user friendly management of the execution of the various tasks, collection of Grid resources, transfer to the computing elements of the code considered with the corresponding input files, start of the jobs, observation and supervision of their execution and finally even staging out of the result files after successful completion. In this environment the IGI Portal acts as a central component to instantiate workflows, manages their execution and performs the file staging involving input and output processes.

In the above described use case the execution process may take several hours, in contrast with Grid policies that fix the CPU time to a limited granted amount. For such reason, special care was taken in handling the checkpointing of the calculation where a set of specialized bash functions have been developed with the purpose of copying selected files from the Worker Node (WN) where the job is physically running to Grid Storage Elements (SE)s, where the files are stored in a temporary or permanent way.
The same solution enables the user to access directly via the Web portal the outcomes, making them available at runtime.

In the present version of the developed GUI interface, special fields enable the user to define the type of calculation, the kind of resources to be used and the number of required cores.
At the end of each job the user had to retrieve the related output files directly from the IGI Portal, analyze them and submit a new job to continue the current work.

Italian Grid Training Workshop

CRYSTAL

porting to grid

Alessandro COSTANTINI

INFN/IGI, Perugia (IT)

alessandro.costantini@pg.infn.it

Summary

- CRYSTAL application porting
- Needs and solutions
- Hands-on
  - CLI
    - JDL
    - Submission and job monitoring
  - IGP
    - Graphical interface
    - Submission and job monitoring

Application porting

CRYSTAL can be compiled for sequential and parallel calculations

- Serial
  - Ifort, MKL
- Parallel
  - SL5 (ifort, MKL, MPICH2 and OPENMPI)
  - SL6 (ifort. MKL, MPICH2)
- Requirements:
  - A set of resources (MPI support)
  - Storage area to save the results coming from the computations (few GB/calculation)
  - Checkpointing and restart
  - Check outputs at runtime

Needs and solutions

- 6 versions of the executable handled by the user
  - Simple mechanism to choose the right one
    - Check on the host (SLC/SL6)
    - Check on the MPI support (OPENMPI/MPICH2)
    - FLAG: -t <type>
- Long simulation running
  - Limited runtime in grid
  - Checkpointing at fixed time intervals
  - Check the output at runtime
- Different modules used in CRYSTAL
  - Check on the input files to define the right module to use

JDL

```
-ui2.grid.unipg.it

# cp /tmp/GTW-CRYSTAL.tar

# Job
Type = "Job";
JobType = "normal";

#Myproxy
MyProxyServer="myproxy.cnaf.infn.it";

#stdout
StdOutput = "std.out";
#stderr
StdError = "std.err";
```


JDL

```
############
# CRYSTAL #
############

#Executable
Executable = "crystal-prod.sh";

#Arguments
Arguments = " -t serial -u acostantini -i ice_631dp.d12";

#Files to be sent on grid
InputSandbox = {"ice_631dp.d12","crystal-prod.sh","crystal-hook.sh"};

#Files to be retrieved from grid
OutputSandbox = {"std.out","std.err"};
```

JDL

```
##########################################
# Requirements for parallel execution #
##########################################

# Number of CPU
CPUNumber = 8;

#MPI-START Req
Requirements = Member("MPI-START", other.GlueHostApplication
SoftwareRunTimeEnvironment)

#&& Member("MPICH2", other.GlueHostApplicationSoftwareRunTimeEnvironment)
#&& Member("OPENMPI", other.GlueHostApplicationSoftwareRunTimeEnvironment)
```


Hands-On

CRYSTAL can be compiled for sequential and parallel calculations

- 1 Serial run of CRYSTAL
  - ice_631dp.d12
  - FLAG: -t serial
- 2 Parallel run CRYSTAL
  - ice_631dp_optg.d12
  - FLAG: –t parallel
  - Requirements: CPUNUMBER, MPI-START
- 3 Parallel run with 2 input files
  - ice_631dp_freq.d12
  - ice_631dp_optg.f34
  - FLAG: –t parallel -k ice_631dp_optg.f34
  - Requirements: CPUNUMBER, MPI-START

3. Usecase 2: VENUS

3.1 Introduction to VENUS

VENUS. Program overview
Andrea Lombardi, University of Perugia, Italy

Grid computing is particularly suited for calculations that can be divided into different independent tasks that do not need to exchange information between them. As such, classical trajectory calculations, widely used in theoretical chemistry to simulate the molecular dynamics, naturally benefit of the Grid. In a typical Grid implementation, the computing load can be distributed according to different values of some integral of the motion (e.g. total energy, angular momentum, or a set of vibrational quantum numbers) that labels a bunch of trajectories or also by different values of some mathematical parameter, for example the seed of the random number generator that is used to produce the set of initial conditions. Each set of initial conditions will correspond to a particular batch of trajectories. After the calculations end, output data can be collected and processed to obtain the desired results.

One of the case studies proposed in the workshop consisted in handling precisely a classical trajectory model tool for the CO_2 –CO_2 molecular collision simulations, a real Grid-empowered molecular simulator.
Such processes are relevant in aircraft and spacecraft design studies (like the reentry modeling into planetary atmosphere). The input entries are the collision energy, the rotational temperature and the initial vibrational quantum numbers of the colliding molecules (as just seen internal energy partitioning is quite more complex than for atom-diatom systems).
Different batches of trajectories to be assigned to different Grid nodes and having common energy, rotational temperature and vibrational quantum numbers, were identified by a different seed of the random number generator used to create sets of initial conditions.

Introduction to VENUS

Andrea Lombardi

Dipartimento di Chimica, Università di Perugia
Via Elce di Sotto 8, 06123 Perugia, Italy

Italian Grid Training Workshop - GARR, Rome, 20-21 January 2014

Outline

- Venus package for molecular dynamics
- Initial conditions
- Interactions
- Possible applications
- Use case

"VENUS" © B. Hase research group, http://monte.chem.ttu.edu/venus.html

is a general classical trajectory computer program and a popular chemical dynamics simulation program package

Performs classical simulations under different initial conditions with the use of the available set of analytic potential functions

Atomistic simulations

Solving Hamilton's equations of motion

$$\frac{\partial p_i}{\partial t} = -\frac{\partial H}{\partial q_i}$$

$$\frac{\partial q_i}{\partial t} = \frac{\partial H}{\partial p_i}$$

$$E = H(q,p) = T(p) + V(q)$$

1. Initial conditions, coords and momenta, random selection
3. Trajectory numerical integration
5. Results: final coords and momenta used to calculate product energies and properties

Meaning of the trajectory method

Trajectories generate a mapping relationship between initial and final classical states

Set of initial states, E,J, internal energy → *Set of final states*

If the sample is large enough one can get probabilities, cross sections and rates. Some quantisation method is needed to "project" classical states onto quantum states (simple binning, projection onto normal modes, Gaussian binning ……)

Initial conditions

Polyatomic Rotational Energy

Rotational energy and angular momentum for a polyatomic molecule is selected by assuming separability of vibrational and rotational motion.

The amount of energy defined by a rotational temperature T_{rot}

Two options

Rotational energy about each axis is RT/2.

$$E_{rot} = \frac{3}{2} k_b T$$

Rotational energy from a thermal distribution by assuming a symmetric top

$$E_{rot} = \sum_{i=x,y,z} \frac{J_i^2}{2 I_i}$$

Polyatomic Vibrational Energy

Normal Modes

- In mass-scaled Cartesian coordinates the kinetic T and potential energy V are:
  $2T = \dot{\mathbf{c}}_i^{\mathrm{T}} \dot{\mathbf{c}}_i, \quad 2V = \mathbf{c}_i^{\mathrm{T}} \mathbf{V} \mathbf{c}_i$
  where $\mathbf{c}_i$ is a 3N-dimensional column vector of the Coordinates and $\mathbf{V}$ is the mass-weighted Hessian matrix (second derivatives, numerically calculated)

- find a matrix $\mathbf{A}$ that diagonalizes the $\mathbf{V}$ matrix, obtain normal mode frequencies and eigenvectors

- $\mathbf{A}$ can be used for the transformation of Cartesian coordinates and normal modes

Microcanonical Normal Mode Sampling.

n normal modes, with fixed E_v total energy, a set of n individual n mode energies is sampled (n R_i random numbers)

$$E_i = \left[E_v - \sum_j^{i-1} E_j\right]\left[1 - R_i^{\frac{1}{n-i}}\right] \qquad E_v = \sum_{i=1}^{n} E_i$$

using normal mode frequencies, the normal mode amplitudes can be obtained. A random phase is then assigned to each normal mode

$$A_i = \frac{\sqrt{2E_i}}{\omega_i}$$

$$Q_i = A_i \cos(2\pi R_i)$$

$$\dot{Q}_i = -\omega_i A_i \cos(2\pi R_i)$$

The corresponding Cartesian coordinates and momenta are then generated using the eigenvector matrix **L**

Fixed Normal Mode Energy Sampling

For this option the normal mode quantum numbers, n_i are read in and these values are used to set the energy of each normal mode

Thermal Sampling

According to a given vibrational temperature T_{vib}, normal mode energies are sampled from a Boltzmann distribution

Diatomic molecules

Diatomic Hamilton function of a rotating oscillator (no couplings)

$$H(q,p) = \frac{p^2}{2\mu} + V(r) + \frac{J^2}{2\mu r^2}$$

the vibrational n and rotational J quantum numbers are supplied by the user. The vibrational energy is calculated semiclassically, and the linear momentum is then obtained as follows (E_{nJ} rotovibrational energy)

$$p = \pm(2\mu)^{1/2}\left[E_{nJ} - \frac{J(J+1)}{2\mu r^2} - V(r)\right]^{1/2}$$

A rejection method is used to choose the value of the interatomic distance *r*

Initial conditions for collisions

impact parameter b, supply a fixed value of b or randomly select b in the interval [0,b_{max}], b_{max} to be determined as a cutoff radius

orientation of the reactants: mutual orientation of the reactants randomly generated

different reaction paths: the properties of the reactive channel products have to be specified

Relative Translational Energy

The value of the relative translational E_{rel} can fixed or chosen from the Boltzmann distribution at the given temperature T_{rel}

Interaction

different analytical potential energy functions are implemented for intramolecular and intermolecular interactions are implemented

Stretching

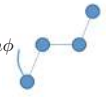
$$V = f(r - r_0)^2$$

Bending

$$V = f_\theta(\theta - \theta_0)^2$$

Torsion

$$V(\phi) = V_0 \sin n\phi$$



Intermolecular

Lennard-Jones ...

HARTREE-FOCK DISPERSION

Interaction types

Type 1: Harmonic stretch, $V = 1/2\ f_r (r - r_e)^2$

Type 2: Morse stretch, $V = D\{1 - \exp[-\beta(r)(r)]\}^2$, $\beta = c_1 + c_2\,\Delta r + c_3\Delta r^2 + c_4\Delta r^3$, $\Delta r = r - r_e$

Type 3: Harmonic bend, $V = 1/2\ f_\theta(\theta-\theta_e)^2$, $f_\theta = f_r^{ij} S(r_{ij})\, S(r_{ik})$ where r_{ij} and r_{ik} define the bond. The $S(r)$ attenuation terms are unity if $\Delta r < 0$. Otherwise, $S(r) = \exp[-C\,\Delta r^2]$.

Type 4: Harmonic alpha bend (wag), $V = 1/2\ f_\alpha\,(\alpha - \alpha)^2$

Type 5: Generalized Lennard-Jones, $V = a/r^n + b/r^m + c/r$

Type 6: 2-fold torsion (double bond), $V = V_\phi\,(1-\cos 2\gamma)/2 = V_\phi \sin^2\gamma$
3-fold torsion (methyl group), $V = V_\phi \cos 3(\gamma/2)$

Type 7: Generalized exponential repulsion and attraction, $V = a\exp(-br) + c/r^6$

Type 8: Ghost pair interaction, $V = q_i/r_{ij} + q_j/(1/r_{ij} + 1/r_{jk} + 1/r_{jk})$
G. C. Lie and E. Clementi, *J. Chem. Phys.* **62**, 2195 (1975)

Type 9: Tetrahedral center
R. J. Duchovic, W. L. Hase, and H. B. Schlegel, *J. Phys. Chem.* **88**, 1339 (1984).
Revised according to W. L. Hase, S. L. Mondro, R. J. Duchovic, D. M. Hirst, *J. Am. Chem. Soc.* **109**, 2916 (1987)] and X. Hu and W. L. Hase, *J. Chem. Phys.* **95**, 8073 (1991).

Type 10: Non-diagonal stretch-stretch interaction, $V = f_{ij,ij}\,(r_{ij}-r_{ij}^e)\,(r_{ij}-r_{ij}^e)$, $f_{ij} = f_{ij}^r\, S(r_{ij})\, S(r_{ij})$. The $S(r)$ terms have the same functional form as in Type 5 above.

Interaction types

Type 12: Non-diagonal bend-bend interaction,
$V = f_{\theta\theta}\,(\theta_{ijk} - \theta_{ijk}^e)\,(\theta_{mno} - \theta_{mno}^e)$, $f_{\theta\theta} = f_{\theta\theta}^r\, S(r_{ij})\, S(r_{ik})\, S(r_{mo})\, S(r_{no})$. The $S(r)$ terms are defined in Type 3 above and have the same exponential constant C as in Type 3.

Type 13: Torsion for dihedral angle, $V = \sum_n \frac{k_n^\phi}{2}\,(1+\cos(n\phi-\gamma_n))$

Type 14: Axilrod-Teller three-body potential, B.M. Axilrod and E. Teller, *J. Chem. Phys.* **11**, 299 (1943).
$$V = Z\,\frac{1+3\cos\theta_i\cos\theta_j\cos\theta_k}{(r_{ij}\,r_{ik}\,r_{jk})^3}$$

Type 15: S_N1 potential energy functions.
S. K. Vitale Linde and W. L. Hase, *J. Phys. Chem.* **94**, 2778 (1990) and H. Wang, L. Zhu and W. L. Hase, *J. Phys. Chem.* **98**, 1608 (1994).

Type 16: Rydberg potential, $V = -D\,(1+a\rho)\,e^{-a\rho}$ where $\rho = (r-r_e)/r_e$.

Type 17: Hartree-Fock Dispersion potential, R. Ahlrichs, R. Penco and G. Scoles, *Chem. Phys.* **19**, 119 (1977).
$$V = \Delta E_{HF} - \left[\frac{C_6}{r^6} + \frac{C_8}{r^8} + \frac{C_{10}}{r^{10}}\right] F(r)$$
where $F(r) = \begin{cases} \exp[-(1.28 r_m/r - 1)^2] & \text{for } r \le 1.28 r_m \\ 0 & \text{for } r \ge 1.28 r_m \end{cases}$
and the Hartree-Fock repulsion is given by $\Delta E_{HF} = A\,e^{-ar}$.

Most often

For dynamics not in the neighbourhood of the equilibrium configurations (i.e. large amplitude motions, reactive processes ..) potential energy surfaces obtained by force fields, although well parametrized, are in general not satisfactory

Use of analytic potential energy surfaces fitted to experimental and computational data.

Particularly, for collision dynamics and reactive systems specific potential energy surfaces must be obtained for each system

This is costly

since molecular energy and its derivatives need to be computed frequently during the integration of the equations of motion, cumbersome expressions can slow down calculations

--A possible solution is to use directly ab initio or semi-empirical molecular orbital calculations to obtain the energies and derivatives as they are needed (no fitting)

USE CASE

Carbon dioxide collisions, energy transfer

In gas phase, molecular collisions control energy exchange (disposal) and state population

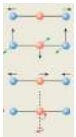
Relevant in the **design of aircraft and spacecraft**, reentry studies (TPS etc..)

Hypersonic flows
High temperatures

Relevant in **modeling of earth and planetary atmospheres** (composition, appearance of life, climate ....)
plasma kinetics,
CTM applications

CO_2 model for collision simulations

•**To a first approximation**, CO_2 **vibrations behave like harmonic oscillators** and can be described by normal coordinates



Four fundamental modes (bending degenerate)

Standard notation for a vibrational states

667.5 cm^{-1} 1340 cm^{-1} 2349 cm^{-1}

quantum n. v_2 v_1 v_3 $\left(v_1\ v_2^{\,l}\ v_3\right)$

•**Rotation:** similar to the motion of a **diatomic rigid rotor**
Superimposed on each vibrational energy level is the rotational energy structure

$$E_{rot} = Bj(j+1), \quad B \simeq 0.39\,cm^{-1}$$

PES

a dedicated PES is used

•**Potential energy surface (PES) for CO_2 + CO_2**

$$V = V_{intra} + V_{inter}$$

Intramolecular part: use analytical potential energy surface (S. Carter, J. N. Murrel, CCACCA 57, 355-365 1984)

The intermolecular PES assumes flexible monomers:
Intermolecular PES, including the dependence on the internal CO_2 states*

*F. Pirani, M. Bartolomei, A. Lombardi, A. Laganà, J. Comp. Chem., 33, 1806, 2012

Collision process

$$CO_2\,(v_{a1}, v_{a2}, v_{a3}) + CO_2\,(v_{b1}, v_{b2}, v_{b3}) \rightarrow CO_2\,(v'_{a1}, v'_{a2}, v'_{a3}) + CO_2\,(v'_{b1}, v'_{b2}, v'_{b3})$$

Computation scheme

- generate many trajectories
- for each traj. get the triples of final quantum numbers v_i, w_i
- use data to calculate energy distributions, collision probabilities and cross sections

Trajectory initial conditions:

- set initial collision energy -impact parameter -rotational energy -vibrational excitation
- Reactants are given random orientation
- Project final coords and momenta on Normal Modes directions
- Calculate final Normal Modes excitation

Venus options:

Fixed Normal Mode Energy Sampling

Rotational energy from a thermal distribution by assuming a symmetric top

Input file

Output

3.2 VENUS porting to Grid

VENUS. Porting to grid
Alessandro Costantini, INFN-Perugia and IGI (Italy)

In the adopted use case the molecular collision in gas phase of the CO_2-CO_2 triatom-triatom system has been considered. Internal relevant motions are symmetric stretching, rotation around the molecular axis and an axis perpendicular to it, bending and asymmetric stretching.

To a first approximation, CO_2 vibrations are treated as harmonic oscillators and normal coordinates are used. In this particular use case two VENUS options have been used: the Fixed Normal Mode Energy Sampling and the Rotational energy thermalized distribution (by assuming a symmetric top geometry).

In the application porting process, VENUS was compiled as a static binary using the gfortran compiler and other open source libraries. The static compilation of the package ensures that the program is binary compatible with the Grid Worker Nodes (WNs), preventing the incompatibility errors associated with the usage of dynamically loaded libraries. Computationally speaking, the sequential VENUS program can be executed on multiple Grid resources with different input files simultaneously. This pattern is a very popular way of utilizing the Grid and it is called "parameter study" (or "parameter sweep") approach.

For the above mentioned reasons a massive job submission has been performed by making use of the Distributed Infrastructure with Remote Agent Control (DIRAC) framework, able to assist the user in the management of a multitude of sequential jobs via a simple and intuitive interface. DIRAC provides a Grid middleware stack that integrates heterogeneous computing resources and provides a solution for both job submission and data management tasks.

Italian Grid Training Workshop

VENUS

porting to grid

Alessandro COSTANTINI

INFN/IGI, Perugia (IT)

alessandro.costantini@pg.infn.it

Summary

- VENUS application porting
- Needs and solutions
- Hands-on
  - CLI
    - JDL
    - Submission and job monitoring
  - IGP
    - Graphical interface
    - Submission and job monitoring

Application porting

VENUS can be compiled for sequential calculations

- Serial
  - Gfortran, system libraries
- Embarassing parallel
  - Each run is independent from the other
  - Parameter study approach
- Requirements:
  - A set of resources
  - Storage area to save the results coming from the computations (few MB/calculation)

Needs and solutions

- Versions of the executable produced by the user
  - SW can be compiled locally (UI) or in the grid (WN)
- Long simulation running
  - Calculations calibrated by the user
- Different outputs for different files
  - Input files stored in grid
  - Outputs stored separately

JDL

```
-ui2.grid.unipg.it

# cp /tmp/GTW-VENUS.tar

# Job
Type = "Job";
JobType = "Parametric";

#Myproxy
MyProxyServer="myproxy.cnaf.infn.it";

#stdout
StdOutput = "venus__PARAM_.out";
#stderr
StdError = " venus__PARAM_.err";
```


JDL

```
############
# VENUS#
############

#Executable
Executable = "venus1lfc.sh";

#Arguments
Arguments = " venus  _PARAM_  test  /alex/venus ";

#Files to be sent on grid
InputSandbox = {""venus1lfc.sh","venus""};

#Files to be retrieved from grid
OutputSandbox =
{"venus_grid__PARAM_.err","venus_grid__PARAM_.out"};
```

JDL

```
############
# VENUS#
############

#-----
#example: total of 4 jobs
Parameters = 5;
ParameterStart = 1;
ParameterStep = 1;
#-----
```


JDL

```
########################################
# Requirements for execution #
########################################

#Requirements = other.GlueCEUniqueID == "ng-ce.grid.unipg.it:8443/cream-pbs-grid"
#&& other.GlueCEUniqueID == "cream-01.cnaf.infn.it:8443/cream-pbs-prod-sl5"
```


Hands-On

VENUS can be compiled for embarrassing parallel calculations

- 1 define your LFC path
  - /grid/<VO>/<user>
- 2 copy inputs in the defined path
  - As a collaction in a tarfile
- 3 play with the JDL
  - Parameters
  - Requirements

4. Usecase 3: QUANTUM ESPRESSO

4.1 Introduction to QUANTUM ESPRESSO

QUANTUM ESPRESSO. Program overview
Michele Ceotto, University of Milano, Italy

My current research focuses mainly on two topics: ab initio semiclassical molecular dynamics and density functional theory (DFT) calculations of doped and molecule-adsorbed titania nanoparticles.

In the first case, a semiclassical approximation of the quantum propagator is calculated via phase space integration of the coherent state representation of the Feynman Path Integral exact formulation of the propagator. This boils down to a classical Monte Carlo integration performed with thousands of classical trajectories, including their positions, velocities, Hessians and actions.

The second part of my research includes the usage of commercial and academic quantum chemistry codes for solid state and surface state. These include Quantum Espresso (QE) and Crystal. The goal of this research is both to explain and to predict photocatalytic experiments of titania doping. Experimental physical observables are calculated via DFT calculations and compared with the experiments in order to elaborate a common electron-hole separation photocatalytic mechanism. Simulations are relevant for this type of research since experimental characterization are far from providing a definite and clear nanoscopic picture. Instead, ab initio calculations can provided all geometrical and electronic details about the solid states simulations. However, these are limited by the DFT intrinsic approximations and several different level of functional approximation are needed in order to have a clear physical picture.
During this Workshop, I briefly introduced the QE inputs file and provided the participants with three QE usecases. The main goal was to give a flavor of QE potentialities if implemented on the GRID architecture, without getting into the code details.

Figure 1 shows the atoms arrangements for a typical titania semiconductor. The specific usecase reported is for the Anatase polymorph and it is a supercell, i.e. the compositions of several primitive cell of solid state titania. The computational time required depends on the size of the supercell and we generated three size of supercells for testing the effect of the GRID parallelization over the computational time. Clearly, by increasing the number of atoms per cell, the computational effort augments. Two usecase geometries are reported in Figure 2. The smaller one (left side) is a primitive titania cell, which is composed of 2 Ti atoms and 4 O atoms. This calculations was employed during the Workshop to test and possible debug both the command lines and the web user interfaces provided by the Italian Grid Infrastructure (IGI) websites. In this case, the IGI personnel (Daniele Cesini) designed a user-friendly portal for launching QE calculations. Clearly the parallelization speed up is not significant and necessary in this case, since the computational time is mainly due to latency GRID time. However, this usecase helped each user to get familiar quickly with the IGI web interface.

Figure 1: The solid state Ti and O atoms arrangment for the Anatase titania polymorph. The periodic supercell boundaries are outlined.

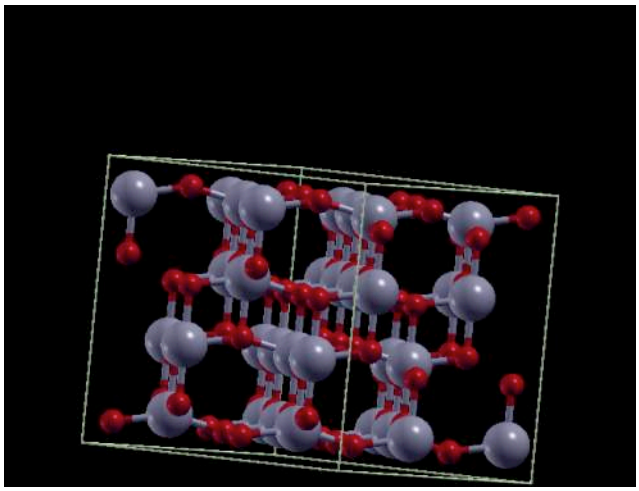
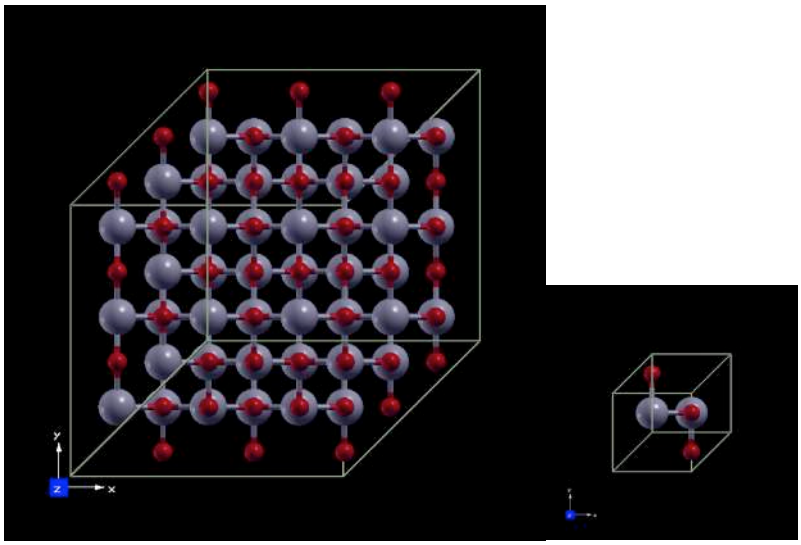


Figure 2: Usecase supercell arrangments for titania Anatase polymorph. Left: Elementary primitive cell of 2 Ti atoms and 4 O atoms. Right: 54 Ti atoms and 108 O atoms supercell.



Instead, the right side primitive cell can take as long as several hours to be processed. This usecase was left to the users for testing the performance of different GRID architecture. One significant advantage of the GRID infrastructure is that many different typo of machines are available. It is left to the user to select a specific machine and test not only the average waiting time but also the code performance.

The Training workshop on application porting has surely being useful to my current research for the following reasons:
1) it avoids direct computational grant application, where a future initial, final and total time of computation is fixed a priori, without taking into account the researcher's availability;

2) the training was focusing on three software (VENUS, Crystal and Quantum Espresso), which are currently and fully part of my research tools and which embrace important theoretical chemistry fields;
3) the opportunity to learn about the Italian and the European grid infrastructure, understand how it works and how better use it;

In my opinion the training has been delivered in a profitable way and it has been very useful to be able to run your own job even during the workshop.

4.2 QUANTUM ESPRESSO

Invited speaker: Stefano Cozzini

QUANTUM-ESPRESSO and the GRID
Stefano Cozzini, IOM/CNR Democritos & eXact lab S.R.L., Trieste (Italy)

Italian GRID Training workshop
Roma 20/21 gennaio 2014

Quantum-Espresso and the GRID

Stefano Cozzini
IOM/CNR Democritos & eXact lab srl

Outline

- Quantum/Espresso suite
- Some Ideas how to use the GRID for Q/E
- Q/E on the GRID: phonon calculations

Quantum Espresso

Quantum opEn-Source Package for Research in Electronic Structure, Simulation, and Optimization

Quantum ESPRESSO is an open-source suite of computer codes that implements state-of-the-art electronic-structure methods for materials modeling at the atomic/nano-scale.

www.quantum-espresso.org

Q/E in numbers*

*(slide from S.Baroni)

www.quantum-espresso.org

qe-forge.org

What I can do with Q/E (version 5.0.3) ?

Structure optimisation
Phonons
Elastic constants
Quantum transport
Ab-initio molecular dynamics
Dielectric and Raman tensors
Infrared spectra
NMR spectra

Q/E computational features

Q/E is CPU/RAM intensive.

Parallelization is a must

Q/E developed with HPC infrastructure as primary target…

Different level of parallelism !

mpirun -np 4096 ./neb.x -nimage 8 -npool 2 -ntg 4 -ndiag 144 -input my.input

-nimage: number of images (loosely coupled: each one uses 512 cpus)
-npool :k-points are distributed across 2 pools of 256 processors each (for image)
-ntg: 3D FFT is performed using 4 task groups (64 processors each)
-ndiag: diagonalization of the subspace Hamiltonian is distributed to a square grid of 144 processors

From Q/E developer meeting minutes:

- GPU-enabled 5.0.x version of PWscf released
- new parallelization levels added or extended
- BGQ porting
- Serious progress achieved in massive-parallelization PRACE project

Q/E and the GRID

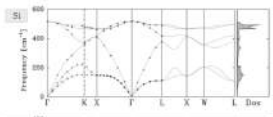
Large-scale computations with Quantum ESPRESSO require HPC resources (tightly coupled clusters):

BUT:

often many smaller-size, loosely-coupled or independent computations are required. A few examples:

- the search for transition pathways (Nudged Elastic Band method);
- calculations under different conditions (pressure, temperature) or for different compositions, or for different values of some parameters;
- full phonon dispersions in crystals

Phonon in crystals



Force Costant Matrixes

Phonon frequencies $\omega(\mathbf{q})$ are determined by the secular equation:

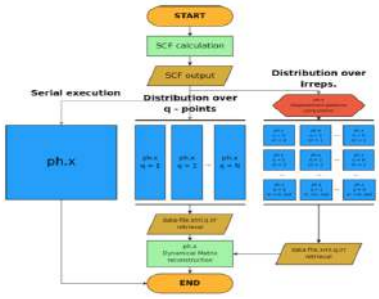
$$\| \tilde{C}^{\alpha\beta}_{st}(\mathbf{q}) - M_s\omega^2(\mathbf{q})\delta_{st}\delta_{\alpha\beta} \| = 0$$

Calculations of phonons

- Force constants are computed by ph.x code on a grid of n **q** vectors.
- For each **q** one has to perform a set of linear response calculations one of each irrep (irreducible representation) generally proportional to the number of atoms
- BUT:
  - Force constant calculation are independent for each q
  - Irreps are almost independent for q: only some data should collected at the end

FIT NICELY ON GRID INFRASTRUCTURE

Phonon workflow



Practical implementation (I)

On Q/E Package: minor changes needed in the phonon code, namely

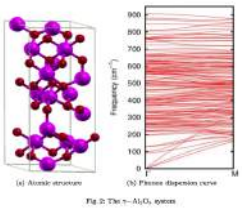
- possibility to run one q-vector at the time (already there)
- possibility to run one irrep (or one group of irreps) at the time and to save partial results.

On GRID infrastructure:

Implemented a Python server-client application which takes care of dispatching ph.x jobs and of collecting results

NOTE: server-client is independent from jobs submission mechanisms: it could run everywhere if plugins for job submission are provided.

Case study: gamma-Al_2O_3



- 11 q points
- 120 irreps for q

A few weeks on modern workstation

Comparing HPC vs GRID

- HPC client /server approach
  - For each q an independent parallel job.
  - Parallel jobs can use 4/8/16 CPUS
- GRID client/server approach
  - each client computes serially one or more (1,4,6) irreps

HPC Results

- Time to results: best cases
  - 16 CPUs: ~ 40 hours
  - 08 CPUs: ~ 50 hours
  - 04 CPUs: ~110 hours

NOTES: HPC scheduling policies taken into accounts:

-maximum 128 cpus at time
-maximum 12 hours for jobs
-maximum 10 jobs running for users

Grid experiment

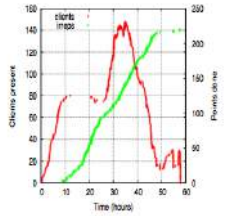
3000 jobs submitted in chunks of 500: clients contact back the server, receive input data and starting data files (hundreds of Mb).

Jobs lost in cyberspace (~ 60% of all contacted servers! of which 30- 40% due to failure in downloading starting data files) are resubmitted.

VOs involved: COMPCHEM and EU-IndiaGRID

GRID Results (6 irreps per clients)

- Up to 145 independent jobs simultaneously running
- Less 60 hours to complete the experiments :
- Comparable with 8 CPUs parallel runs!



Conclusions

- A realistic application of Quantum ESPRESSO to first-principle calculations at the nanoscale was demonstrated on the GRID
- Results produced in a relatively short time in spite of a rather high job failure rate: GRID can be competitive with conventional High-Performance Computers !
- Full exploitation of GRID infrastructure requires however the possibility to select HPC (with MPI), or large multicore machines (with OpenMP), In order to enable High throughput calculation on medium size HPC problem

4.3 QUANTUM ESPRESSO porting to Grid

Porting QUANTUM ESPRESSO to the Grid
Daniele Cesini, INFN-CNAF and IGI (Italy)

In the proposed use case the prediction of photocatalytic experiments of titania doping is performed. Experimental physical observables are calculated via DFT calculations and compared with the experiments in order to elaborate a common electron-hole separation photocatalytic mechanism. Computer simulations are relevant for this type of research since experimental characterization are far from providing a definite and clear nanoscopic picture. Instead, ab initio calculations can provided all geometrical and electronic details about the solid states simulations. However, these are limited by the DFT intrinsic approximations and several different level of functional approximation are needed in order to have a clear physical picture.

From a computation point of view, increasing the size of the titania supercell described in the previous paragraph, translates into long, parallel Quantum Espresso runs; the size of the supercell is also directly correlated to the size of the data that need to be handled (quickly reaching the order of 10 GB).

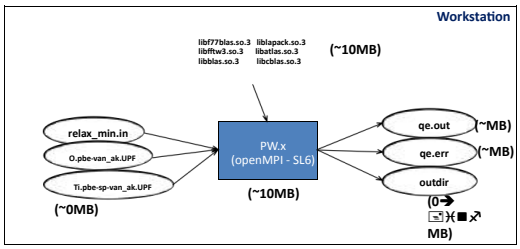
The presented Quantum Espresso use case is based only on the "plain wave" module (pw.x) of the software suite. It was ported to the Grid rebuilding the executable on Scientific Linux 6 (which is the most common operating system available on IGI) and making use of OpenMPI to implement the parallelism. To complete the simulation bypassing the mentioned walclock time limitations per single job, an automatically submitted linear workflow of parallel runs was implemented using the output (i.e the checkpoint file) of the i_{th} job as input of the i_{th+1} job for a defined number of steps. The workflow submission to the Grid has been implemented through the IGI portal.

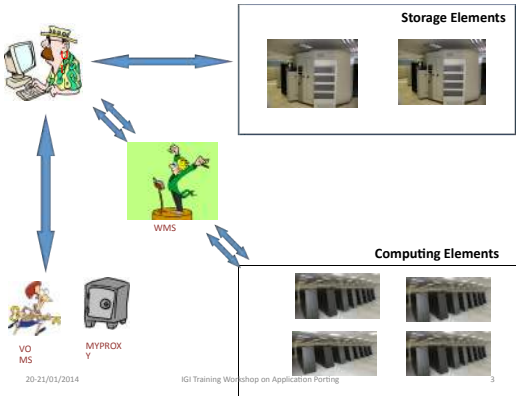
An easy to use web interface (see slide 14) is exposed by the IGI portal allowing the user to select the run name (a unique id for the entire chain), the input files (that should be pre-loaded on a defined Grid location) and the maximum number of checkpointing steps before stopping the workflows. After the submission the user is notified via mail of the events occurring to the workflow during its entire lifecycle and when it completes the final output data are made available on a predefined Grid storage path. The web interface hides the inner complexity of running a parallel jobs workflow in the Grid environment at the cost of limiting the possibility of customizing the submissions, i.e. changing the number of MPI nodes, changing the resource ranking algorithm, selecting the resubmission policies in case of failures, etc. However a more advanced interface, as an extension of the current one is currently under study by the IGI portal development team.

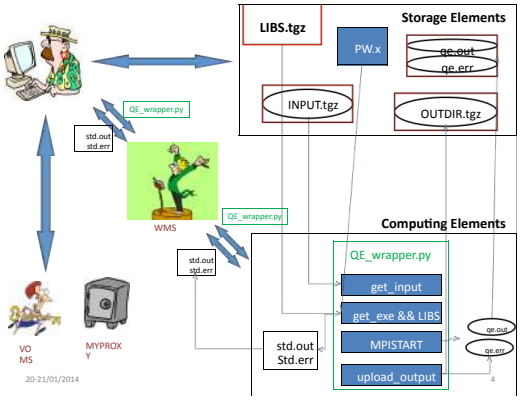
Porting Quantum Espresso to the Grid

D.Cesini – IGI-FUS

./mpirun –np <NCPU> pw.x < relax_min.in > qe.out 2> qe.err







INPUT.tgz

```
[cesini@igi-ui test_corso_dir]$ tar -tvzf input/Ti.pbe-sp-van_ak.tgz

-rw-rw-r-- cesini/cesini 277529 2013-12-09 15:42 O.pbe-van_ak.UPF
-rw-rw-r-- cesini/cesini   1460 2013-12-09 16:49 relax_min.in
-rw-rw-r-- cesini/cesini   1482 2013-12-10 14:13 relax_min.in_0
-rw-rw-r-- cesini/cesini   1478 2013-12-10 14:52 relax_min.in_1
-rw-rw-r-- cesini/cesini   1386 2013-12-09 15:42 scf_sec.in
-rw-rw-r-- cesini/cesini 571123 2013-12-09 15:42 Ti.pbe-sp-van_ak.UPF
-rw-rw-r-- cesini/cesini   5918 2013-12-09 17:34 relax_ore.in
-rw-rw-r-- cesini/cesini   5944 2014-01-14 14:34 relax_ore.in_0
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_1
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_2
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_3
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_4
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_5
-rw-rw-r-- cesini/cesini   5939 2014-01-14 14:34 relax_ore.in_6
```

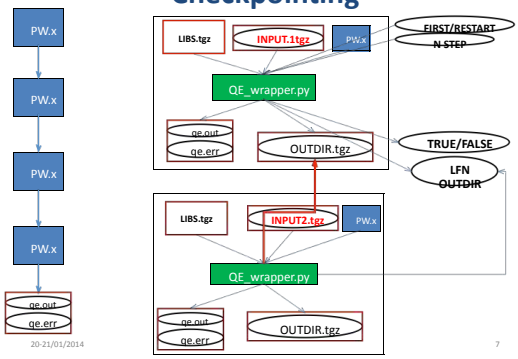

Quale Sito?

- Con WN SL6
- Che support
- Che support
- Che abbia u
- Che abbia >
- Che permet
- Che

```
RHEL6 = ( (  other.GlueHostOperatingSystemName  == "CentOS" ||
        other.GlueHostOperatingSystemName  == "RedHatEnterpriseAS"  ||
        other.GlueHostOperatingSystemName  == "ScientificSL" ||
        other.GlueHostOperatingSystemName  == "ScientificCERNSLC"
        )
&&
      ( other.GlueHostOperatingSystemRelease >= 6.0
&&
      other.GlueHostOperatingSystemRelease < 7.0 )
      ) ;

Requirements = (
Member("MPI-START",

(other.GlueCEInfoTotalCPUs > 16 )
&&
(other.GlueHostMainMemoryRAMSize >= 2000)
&&
RHEL6
);
```

- cream-02.cnaf.infn.it:8443/cream-pbs-prod-sl6
- emi-ce01.scope.unina.it:8443/cream-pbs-hpc (IB)

Checkpointing



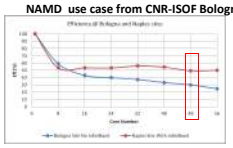
Convenzione Input Filename

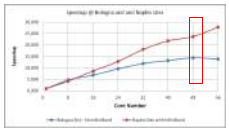
```
[cesini@igi-ui t]$ cat input/relax_min.in_0
&CONTROL
 calculation =  "vc-relax"  ,
 restart_mode = 'from_scratch' ,
 outdir='./out/' ,
 pseudo_dir = './' ,
 nstep       = 3,
 etot_conv_thr = 1.0E-5 ,
 forc_conv_thr = 1.0D-4 ,

[....]
```

```
[cesini@igi-ui t]$ cat input/relax_min.in_1
&CONTROL
 calculation =  "vc-relax"  ,
 restart_mode = 'restart' ,
 outdir='./out/' ,
 pseudo_dir = './' ,
 nstep       = 55,
 etot_conv_thr = 1.0E-5 ,
 forc_conv_thr = 1.0D-4 ,

[....]
```

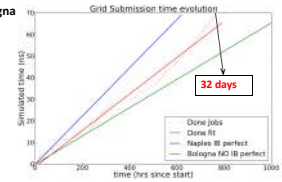

UNINA & IGI-Bologna Sites

NAMD use case from CNR-ISOF Bologna

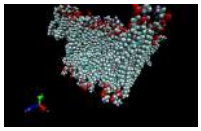




Higher the core number higher the waiting times on the sites queues



Completed in 32 days using 48cores and two sites



QE_wrapper.py e QE.jdl

Sottomissione del primo job

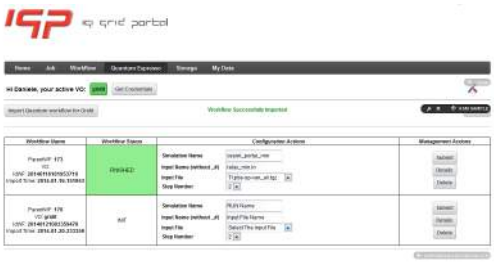
- Preparare il tar con gli input di tutti i restart richiesti (TAR_WITH_INPUTS)
- Caricare il tar in INPUT_GRID_PATH (Sia da UI che da portale)
- Lanciare il primo job con pw_portal_FIRST.jdl e
  Arguments = "<ID_NAME> <Base_input_name> <TAR_with_inputs> SL6 FIRST <user_dir> 2";
- Recuperare gli output (Sia da UI che da portale)

Sottomissione del restart

- Recuperare dall'output del job precedente i file di controllo del job di restart e rinominarli secondo la convenzione
- Lanciare il job di restart con pw_portal_RESTART.jdl e
  Arguments = "<ID_NAME> <Base_input_name> <TAR_with_inputs> SL6 FIRST <user_dir> 2";
  Use the same ID_NAME of the previous job
- Recuperare gli output e verificare se il job è terminato con successo dallo standard output di QE

Sottomissione di un intero workflow da portale

- Preparare e caricare gli input
- Settare i parametri richiesti
- Lanciare il RUN
- Recuperare gli output



5. Abstract from participants

5.1 Research activities by Silvia Carlotto

Silvia Carlotto
Dipartimento di Scienze Chimiche, Università degli Studi di Padova
e-mail:silvia.carlotto@unipd.it

My principal research activity, which is carried out within the FP7 Project "NextGEnCat", is the theoretical modeling of metal oxide materials, with a special attention to surfaces and to their catalytic properties. Currently, I'm investigating the effects on defects and impurities on the properties of lanthanide-free perovskites (in particular $SrTiO_3$). To this end, the interaction of small molecules of interest for catalytic converters (CO, NO and O_2) with the pure, doped and/or reduced surfaces are considered. These studies are performed with density functional calculations, using supercell models and a pseudopotential plane-wave approach. The Quantum-ESPRESSO (QE) open-source suite is used for all calculations.

The Italian Grid Training Workshop has given an exhaustive overview on the Grid computing infrastructure. There are several positive aspects to point out: i) a good general introduction, ii) the real time interactions with expert people (not a trivial aspect for researchers without any experience on Grid infrastructure), iii) the possibility to test directly the job submission by Grid Command Line and by IGI Portal. The clarity in explanation has permitted to switch between two options. In my opinion, the workshop was most useful for people with low or without experience on grid. More expert attendants would need one full day for each software so that specific/technical issues could be addressed.

Concerning QE, which I'm using in my research activities, the seminar was very effective in comparing the performance in HPC systems and in the Grid infrastructure, showing how the Grid could be a viable alternative for time consuming calculations.

5.2 Studies of many-body quantum systems on the italian GRID By Cristan Degli Esposti Boschi

Cristian Degli Esposti Boschi (CNR-IMM, Sezione di Bologna) and
Daniele Cesini (CNAF-INFN, Bologna)

In the study of quantum properties of many-body interacting systems to ideal extremes may be represented by Density Functional Theory (DFT) on the one side and by Hubbard- or Heisenberg-like models on the other side, the former being a widely used framework for quantitative calculations in realistic materials with moderate correlations and the latter being a route to investigate the role of strong correlations at the price of focusing mainly on universal or qualitative features. A celebrated computational method in this second case is the so-called Density Matrix Renormalization Group (DMRG).

During the training worskhop on application porting to the GRID the use case of a popular DFT parallel code – Quantum Espresso – was presented. There is evidence that for systems with tens of atoms and periodic boundary conditions or for problems in which the computational load can be splitted in an almost parallel way, such as the calculation of the entries of a vibrational dynamic matrix, the GRID approach proves to be successful. However, when the number of atoms reaches O(100) a massively parallel treatment is required and the overall performance of the GRID approach may be significantly affected by intercommunication latencies and priority policies.

On the DMRG side, instead, the execution is typically serial but an advantageous use of the GRID may be the istance of many independent runs for different system's parameters. For example, using the DIRAC submission tool within portal.italiangrid.it we could collect thousands of runs to compute the ground-state entanglement entropy of a class of spin-1 quantum Hamiltonians with spatial anisotropy, whose phase diagram is still controversial in some parts.

5.3 The interaction of small molecules with nucleic acids by Angelo Spinello

Angelo Spinello, PhD student
Dipartimento di Scienze e Tecnologie Biologiche, Chimiche e Farmaceutiche
Università degli Studi di Palermo

My research project concerns the computational study of the interaction of small molecules, in particular metal complexes, with oligonucleotide structures. The models under study are RNA, double stranded B-DNA and more recently the G-quadruplex (G4) DNA. G4 are a non canonical DNA conformations recently observed in human cells, considered potential biotarget of novel anticancer drugs. The interaction of small molecules with DNA is associated to interesting chemical and biological properties of the resulting complexes. For this reason, we have been recently involved in the synthesis of novel metal complexes and organic compounds as potential DNA binders. My research project is part of a wider research line that involves a combined experimental-computational approach to study the interaction of polynucleotides with metal complexes. Fine-structural details of the molecule-nucleotide binding mode in solution have been obtained through MD simulations and QM/MM calculations. For example, we have recently simulated the interaction between metallo-intercalators and DNA models, obtaining atomic level structural details of the interaction and of rationalizing the results of the experimental measurements obtained in aqueous solutions [1,2].
The Italian Grid Training Workshop was a precious opportunity for me to learn how the grid infrastructures work. I think it could be a precious tool for my research as a valid alternative to HPC computing.

References

1. A. Spinello, A. Terenzi, G. Barone: Metal Complex–DNA Binding: Insights from Molecular Dynamics and DFT/MM Calculations, J. Inorg. Biochem. 124 (2013) 63-69.
2. A. Lauria, R. Bonsignore, A. Terenzi, A. Spinello, F. Giannici, A. Longo, A.M. Almerico, G. Barone: Nickel(II), copper(II) and zinc(II) metallo-intercalators: structural details of the DNA-binding by a combined experimental and computational investigation, Dalton Trans. DOI:10.1039/C3DT53066C

5.4 Time-resolved FTIR spectra and gravimetric analysis by Pietro La Manna

Pietro La Manna [1,2]
1 Department of Chemical, Materials and Industrial Production Engineering, University of Naples Federico II, p.le Tecchio 80, 80125 Naples, I
2 Institute of Polymers Chemistry and Technology, National Research Council of Italy, via Campi Flegrei, 80134 Pozzuoli, I

Time-resolved FTIR spectra collected in-situ and gravimetric analysis constitute the experimental basis for a molecular level characterization of the diffusion process of low molecular-weight compounds in polymeric substrates (polyimides, poly(ε-caprolactone, PCL). The spectroscopic data are analysed by different and complementary techniques [difference spectroscopy (DS), least-squared curve-fitting (LSCF) and 2D-FTIR correlation spectroscopy (2D-COS)] which provide information on the overall diffusivity, the nature of the molecular interactions among the system components and the dynamics of the various molecular species. Spectroscopic measurements on thin films (≤ 1.0 μm) allow us to identify the interaction site(s) on the polymer backbone and to propose likely structures for the H-bonding molecular aggregates. The complex band profile of absorbed water in the O–H stretching region is interpreted in terms of normal modes contributions with the aid of ab-initio calculations performed on representative model systems. These are able to capture the main features of the experimental spectra and the perturbations brought about by the interactions of the polymer substrate with the penetrant. In the light of the ab-initio/Normal-coordinate analysis it is possible to exclude the direct involvement of sp^3 oxygen to H-bonding interaction with water.
The Italian Grid Training Workshop has been very useful and informative and served to improve my computational skills and my knowledge on ab initio methods for the study of the systems I am interested in.

5.5 Research activities by Fulvio Ciriaco

Fulvio Ciriaco
Department of Chemistry, University of Bari (Italy)

Research interests: my research mainly involves electronic structure calculations, the main projects. I currently deal with:
- arene-dithyol adsorption above Au surfaces
- carbon diffusion processes on metal surfaces of the X and XI group
- characterization of transition pathways for the oxydrilation of piperazines by means of oxiranes
- deprotonation processes in MALDI spectroscopy
- SQDG fragmentation processes in ion trap mass spectrometers
- interaction between cyclodextrins and 4-thiothymidine

The computational software I use most frequently:
- NWChem, GAMESS-US for molecular studies,
- CRYSTAL, Quantum-Espresso for periodic system calculations.

I am also interested in recent developments of QMC electronic structure computations, currently using CASINO-QMC for benchmarking how well DFT reproduces electronic properties on a selection of the above mentioned systems.
I perform all my computations on a ~100 cores cluster and on CINECA Fermi through submission of ISCRA-C projects.

When I applied for the Grid Training Workshop, I had attended a conference on grid resources in computational biology and shortly thereafter obtained an account on the cluster of INFN-Ba. The workshop had just the right format for me: a hands on tutorial, with the possibility to pose the several questions on the alien platform to the helpful user support team.
I also liked the vision of a community that pervaded the workshop. Except for the first two projects above, involving large node resources, the systems above could partially or completely be adapted for running on the Grid.

My first plan is to attack the problem of SQDG fragmentation by means of a program exploring part of the configurational space, residing on the u.i. submitting (several) DFT calculations to Grid nodes to obtain the energies.

5.6 Research activities by Simone Giusepponi

Simone Giusepponi
ENEA Frascati (Italy)

My research activities involve mainly the materials for energy applications:

Hydrogen storage:
Nanostructured magnesium hydride MgH2 is considered an attractive hydrogen storage material. However, this material displays too high temperatures of decomposition, mainly owing to high thermodynamic stability and slow decomposition kinetics. Since the desorption mechanism is strongly influenced by the chemical and mechanical properties at the interface between MgH2 and Mg, a detailed study of this interface is needed. Clear indications of both the equilibrium properties and the behavior of the Mg and H atoms in terms of total energy calculations have been given using extensive first-principle molecular dynamics simulations of the MgH2-Mg interface. The interfaces and the hydrogen desorption are studied as functions of the temperature. The atomic environment of the Mg atoms at the interface and hydrogen paths for desorption are characterized and studied [1,2].

Materials for fusion applications
There is a widespread consensus in considering tungsten and its alloys as good candidates for structural and shield material in the future nuclear fusion reactors. Indeed W alloys are characterized by thermal properties, such as good thermal conductivity, high melting temperature and low sputtering erosion, of great interest in fusion applications. Thus a deep understanding of the mechanical properties of W in presence of impurities is unavoidable step for the use of W under extreme conditions. Moreover, in view of their application as plasma-facing materials, need to be characterized when radiation defects, such as vacancies or interstitials, are formed. For these reasons first-principles total energy methods based on density functional theory are used to investigate mechanical and structural properties of bcc crystal tungsten in presence of point defects [3,4].

The Training Workshop on Application Porting was a beneficial tool for sharing information about the application fields of computational chemistry codes. Moreover, it mainly showed that the EGI and IGI Grid-like distributed infrastructures provide a way to the normal user to access to large computing resources through a facilitated procedure, without being an expert in computer science.

1. S. Giusepponi, M. Celino. Hydrogen Desorption from Mg Hydride: An Ab Initio Study. Crystals 2 (2012) pp.845-860.
2. S. Giusepponi, M. Celino. DFT model of hydrogen desorption from MgH2: The role of iron catalyst. Int. J. Hydrogen Energy 38 (2013) pp.15254-15263.
3. S. Giusepponi, M. Celino. The ideal tensile strength of tungsten and tungsten alloys by first-principles calculations. J. Nucl. Mater. 435 (2013) pp.52-55.
4. S. Giusepponi, M. Celino. The effects of vacancies in the mechanical properties of tungsten: a first-principles study. Submitted to J. Nucl. Mater.

5.7 Research activities by Sergio Rampino

Sergio Rampino
CNR-ISTM Perugia (Italy)

My research interests are in the field of relativistic quantum chemistry. In particular, I am working at the development of the Dirac-Kohn-Sham (DKS) module of the all-electron four-component program BERTHA [Belpassi et al., Phys. Chem. Chem. Phys. 12, 12638 (2011)]. The DKS approach, combining the Dirac equation with the Kohn-Sham model of Density Functional Theory (DFT), allows the inclusion of both relativistic and electron corrletation effects in a mean field theory and represents therefore the most accurate and efficient way to treat systems were heavy elements, and as a consequence, many electrons are involved. DKS calculations, however, have an intrinsically larger computational cost compared to their non-relativistic analogue. My work has been so far concerned with the overcoming of both time and memory barriers to DKS calculations.

In this respect, on one hand we recently achieved an efficient parallel implementation [Storchi et al., J. Chem. Th. Comp. 9, 5356 (2013)] of the method allowing to perform DKS calculations with up to 39k basis functions on a reasonable time scale. On the other hand, more recently we addressed our work to the memory bottleneck towards an open-ended version of the program based on a complete distributed matrices approach and allowing the applicability of the DKS model to arbitrarily large systems.

The Italian Grid Training Workshop was a great chance for me to consolidate my experience with the European grid computing environment.