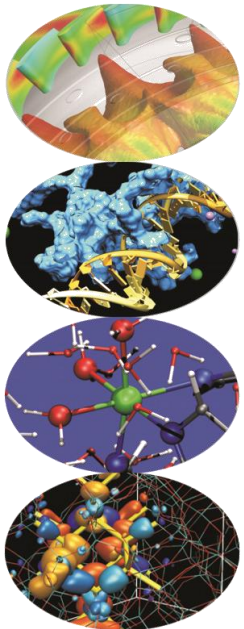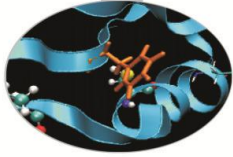# How to obtain HPC resources

A. Emerson, HPC, Cineca.

How do I get access to a supercomputer?

- With the exception of commercial agreements, virtually all access to HPC systems is via peer-reviewed calls to national or international resource providers.

- Depending on the call and provider, usually necessary to write a project proposal detailing the scientific case, how the CPU hours will be used and the application codes which will be run.

- Projects are then evaluated scientifically (for high quality research) and technically (for feasibility).

- In Europe, the principal provider of computer time is PRACE.

PRACE
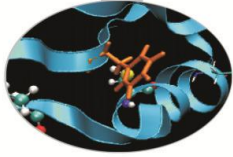
- Partnership for Advanced Computing in Europe
- http://www.prace-ri.eu/
- The mission of PRACE is to enable high impact scientific discovery and engineering research and development across all disciplines to enhance European competitiveness for the benefit of society.
- PRACE seeks to realize this mission through world class computing and data management resources and services through a peer review process.
- PRACE is established as an international non-profit association with its seat in Brussels. It has 25 member countries .
- Four Hosting Members (France, Germany, Italy and Spain) provide Multi-PFlop/s Tier-0 Systems.

PRACE resources
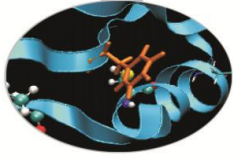
PRACE offers two types of hardware resources:

1. Tier-0 on Petascale supercomputers currently in Germany, France, Italy and Spain.

2. Tier-1 ("DECI")on Terascale or Petascale clusters available in most PRACE centres.

Tier-0 calls are managed by the PRACE AISBL association but Tier-1 calls are usually funded by PRACE projects.

PRACE Tier-0 Systems (10th Call, 2014)
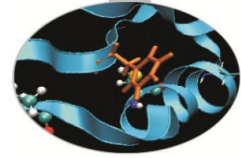
- Curie (CEA, France, Bull x86, 2 Pflop/s)

- Hornet (HLRS, Germany, Cray XC40, 4 Pflop/s)

- Super MUC (LRZ, Germany, IBM DataPlex, 3 Pflop/s)

- Fermi (Cineca, Italy, IBM BG/Q, 2.1 Pflop/s)

- MareNostrum (BSC, IBM DataPlex, 1 Pflop/s)
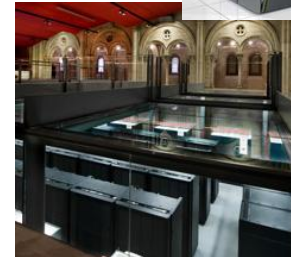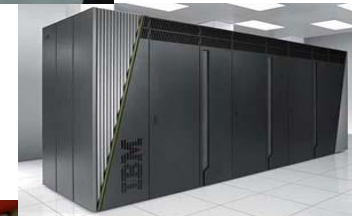
PRACE Tier-0 calls

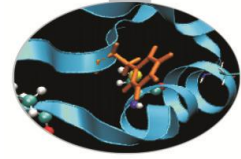PRACE offers 3 different forms of access to Tier-0 resources:

- **Project Regular Access calls**

  Calls for Proposals are issued twice a year and are evaluated by leading scientists and engineers in a peer-review process. Tier-0 proposals typically request many millions of core hours and must demonstrate high parallel scalability.

- **Multi-Year Project Access**

  It is available to major projects or infrastructures that can benefit from PRACE resources and for which more than a single year of access is needed.

- **Preparatory Access**

  It is a simplified form of access for limited resources for the preparation of resource requests in response to Project Access Calls for Proposals.

  Type **A** (scalability tests)

 Type **B** (Enabling + Scalability tests)

  Type **C** (Enabling + Scalability tests with PRACE involvement)

PATC HPC Molecular Dynamics Course 22-24 Sept. 2015

How to apply for Tier-0 project access - procedure
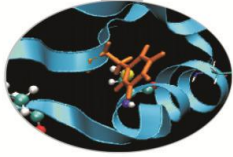
- Consult the application guide from the PRACE website. (http://www.prace-ri.eu/application-guide/)

- Register to obtain username+password for the application portal.

- Fill in the on-line application including abstract and technical details.

- Prepare and attach separate project document according to template.

- Submit before deadline (possible to save preliminary versions and even un-submit before final deadline).

- *For Italian applicants (i.e. based in Italy) we strongly recommend you contact us (i.e. Cineca) first before preparing the application.*

- Researchers based in other countries may try contacting their national representative. PRACE staff can also be contacted.

How to apply for Tier-0 project access –some advice
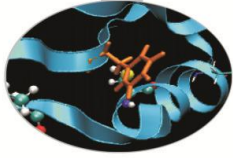
You will have to provide a **Workplan** where you justify the budget you are asking for and how the simulations will be performed in the timescale of the project. Thus in the project document you should include:

– A GANTT chart detailing the activities during the project duration.

– A table demonstrating how you arrive at the requested budget
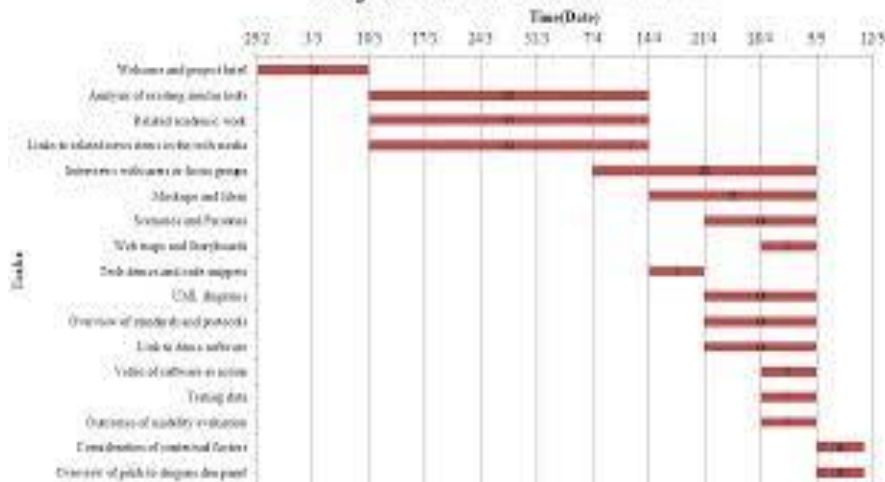
Simulation type	#cores/run	#runs	Walltime/ run (h)	Core hours
Simple MD APoa1	2048	10	1000	20M
MD complex	4096	5	1000	20M
MD complex 2	2048	1	100	2M
			TOTAL	42M

How to apply for Tier-0 project access – some advice

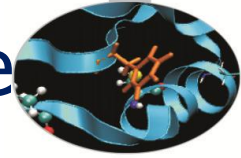You **must** demonstrate adequate parallel scaling on the chosen computer system.

- – Ideally, you will have benchmark results of the proposed input systems (or similar) from a preparatory access or other project on the chosen computer.

- – If you don't have such data find scaling data which matches as closely as possible what you wish to do considering hardware, software and input. For classical molecular dynamics, always cite clearly the number of atoms in the input – this gives the reviewers a clue as to the scalability (typically max performance at 100-150 atoms/core)

- – On some systems (e.g. Fermi) possible to run separate, multiple simulations in the same job ("sub-blocking"). Not usually considered acceptable for a PRACE Tier-0 project.

- – If possible, consider the use of MPI/OpenMP versions of the applications to save on memory/core and make use of multihreaded hardware.

- – Be careful with Replica-Exchange and biased MD-algorithms
  - • For NAMD2.9 and lower, REMD is via TCL scripts (not possible on BG/Q).
  - • In NAMD 2.9 (2.10?), Targeted and Steered MD scale poorly because of rank 0 communication.

PATC HPC Molecular Dynamics Course 22-24 Sept. 2015

How to apply for Tier-0 project access – some advice
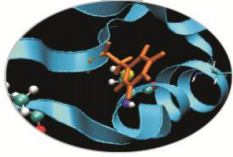
For classical MD projects other sections in the form are not critical:

- Memory requirements are usually low.

- For most MD simulations I/O is almost negligible so no need to mention MPI/IO, HDF5, etc. Also number of files needed is low.

- Archival of trajectories should be within guidelines (mistake anyway to generate large trajectories)

- All MD codes allow checkpoints (restarts) and so job walltimes < 24h.

- Typical allocations are 30-40M core hrs for BG/Q, ~20M core hrs for other architectures. For <5M hrs you must justify Tier-0 resources.

Features of PRACE 10th Call
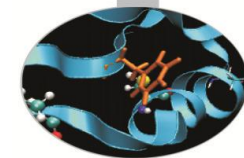
Technical requirements for call 10

https://prace-peer-review.cines.fr/proposal/PRACE_Technical_Guidelines_for_Applicant_CALL10.pdf

Computer System	Minimum Parallel Scaling	Max memory/ core (Gb)
Curie	Fat Nodes 128 Thin Nodes 512 Hybrid 32	4 4 3
Fermi	2048 (but typically >=4096)	1
SuperMUC	512 (typically >=2048)	*
Hornet	2048	*
Mare Nostrum	1024	2Gb

Other requirements include:
- max no. of files
- storage and archive space
- checkpoint frequency
- simultaneous jobs

* should use a substantial fraction of available memory

PATC HPC Molecular Dynamics Course 22-24 Sept. 2015

Typical PRACE Tier-0 call life cycle (project access)

Example CALL 10

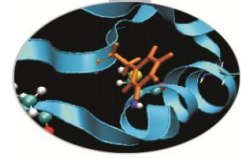| Flow Step	Date
Call opens	10th Sept 2014
Call closes (after 6 weeks)	22nd Oct 2014
Admin check by PRACE staff (1 week)	
Technical Review by PRACE centres (2 weeks)	
Scientific Review (2 months)	
Applicant Response (1 week)	15-21 Jan 2015
PRACE prioritisation Panel (1 week)	
Allocation Starts	10th March 2015
Allocation Ends (1 year after start)	9th March 2016
Final Report (2 months after production end)	

PATC HPC Molecular Dynamics Course 22-24 Sept. 2015

New features of Call 10

- ## Multi-year project access
  - 2-3 years instead of 1 year.
  - Same eligibility criteria as 1 yr calls.
  - Has to demonstrate the need for more than one year. Resources allocated 1 year at a time.
  - Annual review procedure based on report and F2F meeting.
- ## Programmatic Access
  - For a period of up to 3 years.
  - Open to research groups or research projects (e.g. EU flagships, FET projects, etc) or similar.
  - Will consist of various computational experiments (which do not need to be defined at time of application).

NEXT CALL (12) EXPECTED Late September 2015 – PLEASE CONTACT US!

Preparatory Access Calls
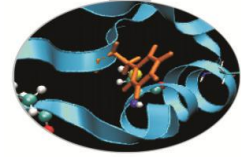
- Designed for code optimisation and benchmarking, possibly with PRACE staff.

- Three types:
  - Type A (scalability tests)
  - Type B (Enabling + Scalability tests)
  - Type C (Enabling + Scalability tests with PRACE involvement)

- Calls every quarter (March; July; September; December). Start date 2 months after submission (if successful).

- Allocation periods normally 2 months (Type A) and 6 months (Types B and C).

- Budget allocations depend on type and computer and partition (GPU, MIC, etc). For example, 100K (Type A) and 250K core hours (Type B,C) for BG/Q, between 50K-200K core hours for other computers.

- Proposals evaluated using a "lightweight" evaluation procedure. Applications should include description of issues preventing scalability.

PRACE Tier-1 (DECI) Calls
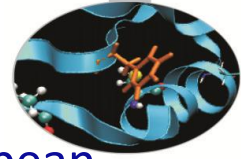
- Inherited into PRACE from the DEISA project, DECI=(Distributed European Computing Initiative)

- Like Tier-0, two calls/year and 1 year in duration. Subject to scientific and technical review.

- Unlike Tier-0:
  - Projects are smaller (e.g. 1M core hours) and have more flexible parallel scaling requirements.
  - Applicants apply for a particular architecture (e.g. GPU, BG/Q, SGI, etc) rather than a computer site.

- Most European countries contribute Tier-1 resources. Cineca will provide access to the Galileo cluster.

- Future of Tier-1 under discussion but a call likely to appear in March-May. (PLEASE CONTACT US IF INTERESTED)

DECI-13 Call closed September 21st. (allocation start data 18-Jan-2015)

Call for Proposals for DECI-12 (Tier-1) (DECI (Tier-1) Access)
Monday 23 December 2013

Opening date: 18th December 2013, 12:00 CET
Closing date: 20th January 2014, 17:00 CET
Allocation start date: 1st May, 2014
Allocation period: 1 year for Project Access
Type of access: DECI (Tier-1)

Tier-1 machines available: Cray XE/XC, and a range of large clusters including GPU resources made available from Cyprus, Finland, Hungary, Ireland, Italy, Poland, Serbia, Sweden, Switzerland, the Czech Republic, the Netherlands, the United Kingdom and Turkey.

CONTINUE

PATC HPC Molecular Dynamics Course 22-24 Sept. 2015
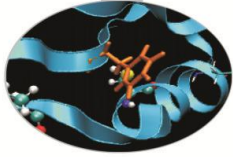
EUDAT resources

- EUDAT is a European funded project with the aim of providing research data services.

- EUDAT provides data resources via two types of call:
  1. Directly via EUDAT.
  2. Asking for EUDAT support via Tier-0 or Tier-1 (DECI) calls (there are separate boxes in the PRACE application forms).

- Resources include persistent disk space, e.g 150Tb for 24 months, + tools for managing and sharing data between researchers. (*but no CPU time*)

- Expected to be used mainly in e.g. astrophysics, bioinformatics, CFD communities, etc. For MD may be useful for very large system trajectories.

National Resources - Italy
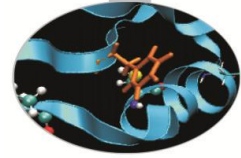
- For Italy-based researchers, Cineca provides computer time via the ISCRA calls. *http://www.hpc.cineca.it/services/iscra*

- Two types of call (B and C) available for accessing:
  - Fermi (for type B: 1-10M hrs, type C: 1M hrs)
  - Galileo(type C, 200k hrs)
  - PICO (type C, 50K hours for bioinformatics, data analytics and visualisation projects)

- For type B two calls/year, type C continuous submission and reviewed once/month.

- Applications must be submitted in English and are evaluated both scientifically and technically.

PICO - Features

.
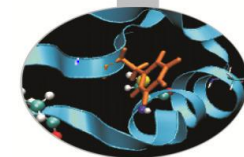
	Total Nodes	CPU	Cores per Nodes	Memory (RAM)	Notes
Compute/login node	66	Intel Xeon E5 2670 v2 @2.5Ghz	20	128 GB	
Visualization node	2	Intel Xeon E5 2670 v2 @ 2.5Ghz	20	128 GB	2 GPU Nvidia K40
Big Mem node	2	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16	512 GB	1 GPU Nvidia K20
BigInsight node	4	Intel Xeon E5 2650 v2 @ 2.6 Ghz	16	64 GB	32TB of local disk

- ❑ 4Pb storage area based on GSS technology with tape library of 12 Pb
- ❑ Storage area organized by project areas
- ❑ "Multi-level memory" , i.e. data can be automatically migrated to tape

PATC HPC Molecular Dynamics Course 22- 24 Sept. 2015

Galileo Features
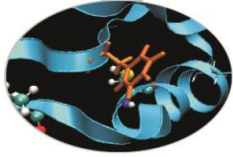
- **Model**: IBM NeXtScale
  Architecture: Linux Infiniband Cluster

- **Nodes**: 516
  Processors: 2 8-cores Intel Haswell 2.40 GHz per node
  Cores: 16 cores/node, 8256 cores in total
  GPU: 2 Intel Phi 7120p per node on 384 nodes (768 in total)
  RAM: 128 GB/node, 8 GB/core
  Internal Network: Infiniband with 4x QDR switches
  Disk Space: xxx TB of local scratch
  Peak Performance: xxx TFlop/s (to be defined)

Final comments
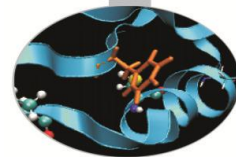
Do not neglect the technical description of the project!

- Like experimental work it is important before submitting any application to understand what resources are needed, particularly CPU time but also memory, disk, accelerators, etc.

- For any call, important to have good estimates of the performance at the parallelism you need (i.e. number of cores) so you can plan the simulations and know how long they are likely to take. If you don't know, you can try PRACE preparatory access or ISCRA-C. _At the very least_ include the number of atoms in the project description (but not sufficient for PRACE T0).

- Include also MD-specific optimisations or algorithms – time step, implicit/explicit solvent, SHAKE, REMD, metadynamics, etc.

- In case of doubt for any of these calls- PRACE Tier0, DECI, Eudat, ISCRA- feel free to contact us.