Physics plans and ILDG usage in Italy

Francesco Di Renzo
University of Parma & INFN Parma

The MAIN ILDG USERS in Italy are still the ROME groups @RM123

- A (by now) well long track of *ILDG-based* projects …
- … within ETMC Collaboration
- Lots of configs for $n_f=2$ (TLS) and (Iwa) $n_f=2+1+1$
- Current work on CLOVER-improved $n_f=2+1+1$. It's not yet the time to store configs
- A snapshot on CLOVER-improved $n_f=2$ was presented at LAT2013

PROCEEDINGS OF SCIENCE

A first look at maximally twisted mass lattice QCD calculations at the physical point

A. Abdel-Rehim[a], **Ph. Boucaud**[b], **N. Carrasco**[c], **A. Deuzeman**[d], **P. Dimopoulos**[e,f], **R. Frezzotti**[e], **G. Herdoiza**[g], **K. Jansen**[h], **B. Kostrzewa***[i,h], **M. Mangin-Brinet**[j], **I. Montvay**[k], **D. Palao**[l], **G.C. Rossi**[e], **F. Sanfilippo**[b], **L. Scorzato**[m], **A. Shindler**[n], **C. Urbach**[o], **U. Wenger**[d]

[a] *CaSToRC, The Cyprus Institute, 2121 Aglantzia, Nicosia, Cyprus*
[b] *Laboratoire de Physique Théorique, Université de Paris XI, 91405 Orsay-Cedex, France*
[c] *Departamento de Física Teòrica and IFIC, Univ. de València-CSIC, E-46100 València, Spain*
[d] *Albert Einstein Center for Fund. Physics, University of Bern, CH-3012 Bern, Switzerland*
[e] *Dip. di Fisica, Università di Roma Tor Vergata and INFN, I-00133 Roma, Italy*
[f] *Centro Fermi, Piazza del Viminale 1 I-00184 Rome, Italy*
[g] *Institut für Kernphysik, Johannes Gutenberg-Universität, D-55099 Mainz, Germany*
[h] *NIC, DESY, Zeuthen, Platanenallee 6, D-15738 Zeuthen, Germany*
[i] *Humboldt Universität zu Berlin, Institut für Physik, Newstonstr. 15, 12489 Berlin, Germany*
[j] *Theory Group, Lab. de Physique Subatomique et de Cosmologie, 38026 Grenoble, France*
[k] *Deutsches Elektronen-Synchrotron DESY, Notkestr. 85, D-22607 Hamburg, Germany*
[l] *Goethe-Universität, Institut für Theoretische Physik, D-60438 Frankfurt am Main, Germany*
[m] *INFN -TIFPA, via Sommarive 14 - 38123 Trento, Italy*
[n] *IAS, IKP and JCHP, Forschungszentrum Jülich, 52428 Jülich, Germany*
[o] *HISKP (Theory), Rheinische Friedrich-Wilhelms Universität Bonn, Germany*

In this contribution, a first look at simulations using maximally twisted mass Wilson fermions at the physical point is presented. A lattice action including clover and twisted mass terms is presented and the Monte Carlo histories of one run with two mass-degenerate flavours at a single lattice spacing are shown. Measurements from the light and heavy-light pseudoscalar sectors are compared to previous $N_f=2$ results and their phenomenological values. Finally, the strategy for extending simulations to $N_f=2+1+1$ is outlined.

HU-EP-13/61, DESY 13-218, SFB/CPP-13-92

L/a	48
T/a	96
β	2.10
κ	0.13729
$a\mu_l$	0.0009
$a\mu_s^{(\mathrm{val})}$	0.0245, 0.0252
$a\mu_c^{(\mathrm{val})}$	0.2940, 0.3058
c_{sw}	1.57551
N_{traj}	> 2000
$P(acc)$	~ 0.75
$\langle P \rangle$	0.603531(6)
$\tau_{\mathrm{int}}(\langle P \rangle)$	10.0(3.5)
am_{PCAC}	0.00004(2)
$m_\pi L$	3.00(2)
a	0.091(5) fm
r_0/a	~ 5.3

Table 1: Run parameters and the values of the valence strange and charm quark masses. In addition, preliminary measurements of the auto-correlation time of the plaquette, the PCAC quark mass, the pion mass (in lattice units), the lattice spacing and the Sommer scale.

This does not exhaus LQCD research in Italy …

- ✓ There is an italian branch of CLS based in Milano Bicocca, Roma1 and Roma2. Refer to Stefan's talk (same comments on ILDG apply)

- ✓ … there is other research *outside ILDG core-business* (I mean *e.g.* finite temperature, confinement, topology …)

- … in Parma there is some usage of (old, "licensed") ILDG configurations to test mechanisms to detect finite size effects and to perform (relative) scale setting.

Computing facilities: Fermi still there …

Top500 List - November 2013

R_{max} and R_{peak} values are in TFlops. For more details about other fields, check the TOP500 description.

R_{peak} values are for the normal CPU clock rate. For the effeciency of the systems you should take the Turbu CPU clock rate into account.

previous 1 2 3 4 5 next

Rank	Site	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	National Super Computer Center in Guangzhou China	**Tianhe-2 (MilkyWay-2)** - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P NUDT	3120000	33862.7	54902.4	17808
2	DOE/SC/Oak Ridge National Laboratory United States	**Titan** - Cray XK7 , Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA K20x Cray Inc.	560640	17590.0	27112.5	8209
3	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM	1572864	17173.2	20132.7	7890
4	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu	705024	10510.0	11280.4	12660
5	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	786432	8586.6	10066.3	3945
6	Swiss National Supercomputing Centre (CSCS) Switzerland	**Piz Daint** - Cray XC30, Xeon E5-2670 8C 2.600GHz, Aries interconnect , NVIDIA K20x Cray Inc.	115984	6271.0	7788.9	2325
7	Texas Advanced Computing Center/Univ. of Texas United States	**Stampede** - PowerEdge C8220, Xeon E5-2680 8C 2.700GHz, Infiniband FDR, Intel Xeon Phi SE10P Dell	462462	5168.1	8520.1	4510
8	Forschungszentrum Juelich (FZJ) Germany	**JUQUEEN** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	458752	5008.9	5872.0	2301
9	DOE/NNSA/LLNL United States	**Vulcan** - BlueGene/Q, Power BQC 16C 1.600GHz, Custom Interconnect IBM	393216	4293.3	5033.2	1972
10	Leibniz Rechenzentrum Germany	**SuperMUC** - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM	147456	2897.0	3185.1	3423
11	GSIC Center, Tokyo Institute of Technology Japan	**TSUBAME 2.5** - Cluster Platform SL390s G7, Xeon X5670 6C 2.930GHz, Infiniband QDR, NVIDIA K20x NEC/HP	74358	2843.0	5609.4	1399
12	National Supercomputing Center in Tianjin China	**Tianhe-1A** - NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050 NUDT	186368	2566.0	4701.0	4040
13	DOE/SC/Pacific Northwest National Laboratory United States	**cascade** - Atipa Visione IF442 Blade Server, Xeon E5-2670 8C 2.600GHz, Infiniband FDR, Intel Xeon Phi 5110P Atipa Technology	194616	2345.8	3388.0	1384
14	Total Exploration Production France	**Pangea** - SGI ICE X, Xeon E5-2670 8C 2.600GHz, Infiniband FDR SGI	110400	2098.1	2296.3	2118
15	CINECA Italy	**Fermi** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM	163840	1788.9	2097.2	822
16	NASA/Ames Research Center/NAS United States	**Pleiades** - SGI ICE X, Intel Xeon E5-2670/E5-2680v2 2.6/2.8GHz, Infiniband FDR SGI	96192	1541.3	2107.0	2015

CINECA INFN

- A BlueGene/Q system was installed at CINECA (Bologna) in the first half of 2012. CINECA is the major computing consortium in Italy (a TIER-0 site within PRACE!)
- There is an INFN-CINECA agreement which gives us access to some BG/Q computing time.
- Most of Fermi's computing power goes into PRACE allocations (so, our community well involved!)

Computing facilities: SUMA activities going on

SUMA

INFN

INFN got money from the Research Ministry (MIUR) for the SUMA project

SUMA plans to support computational physics goals, and at the same time aims to explore all suitable ways in which the technological developments made at INFN can be put to good use for the present and future needs of computational physics.

THEOCLUSTER ZEFIRO

CINECA Eurora

APE QUONG

COKA

SUMA activities

- SUMA groups all the INFN groups active in Lattice QCD
- see POSTER at the SC13 conference

The SUMA Project: HPC Support for the Theoretical Physics Community

SUMA Collaboration

Introduction

The INFN community in theoretical physics is active in several scientific areas that require significant computational support; these areas stretch over a wide spectrum, requiring in some cases fairly limited computing resources – for instance in nuclear physics, high-energy physics phenomenology, spin-system simulations -- while, at the other end of the spectrum huge computing power, that can only be provided at the transnational level, is needed; examples in this class are Lattice Quantum Chromodynamics (LQCD), dynamical systems and classical and ab initio simulations of bio-systems; some research groups work on areas that are acknowledged grand-challenges of High Performance Computing.

At the same time, for most groups active in these areas, it is becoming more and more difficult to independently develop their computational strategies and algorithms in a way that allows to adapt to the increasingly fast changes happening in high performance computing architectures.

Last but not least, several existing INFN projects have produced significant progress on technological developments that may be crucial building blocks for new generation HPC systems, if it can be shown that they are efficient solutions to (at least some) large scale computational problems.

SUMA plans to support all these physics goals, by providing computing resources - both in-house and through access programmes, helping develop and share the know-how needed to efficiently use new computing systems and validating and assessing performances, and at the same time aims to explore all suitable ways in which the technological developments made at INFN can be put to good use for the present and future needs of computational physics.

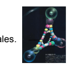
The SUMA project works in close collaboration with academia, research centers and computer centers in Italy, such as the Universities of Ferrara, Parma, Pisa and Rome, SISSA (Trieste) and CINECA (Bologna).

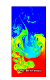
Computational Theoretical Physics at INFN

Theoretical Physics at INFN is heavily supported by numerical simulations. In several cases INFN computational applications are recognized grand-challenges of high-performance computing.

INFN groups are active in the following projects:

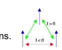
Lattice Quantum Chromodinamics (LQCD): studying the structure and the dynamics of matter at its tiniest scales.

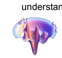




Computational fluid-dynamics (CFD): discovering the statistical properties of fluids in the turbulent regime.

Complex systems dynamics: understanding the behaviour of systems governed by conflicting interactions.



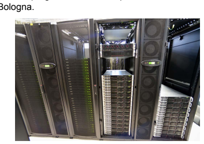


Quantitative biology (QBIO): understanding how fundamental physics interactions shape the building blocks of life.

Work Horses

INFN researchers use a variety of HPC computer systems supporting their investigations.

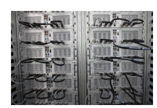
INFN operates a number of Tier1 HPC clusters, and uses Tier0 facilities made available by the PRACE access program of the European Union, as well as the Blue Gene/Q system installed at CINECA in Bologna.



INFN Tier1 cluster installed in Pisa.

The Fermi BG/Q system installed at CINECA.

INFN carry on several experimental projects to optimize codes, test and assess performances on new architectures, including GPUs, MICs and FPGAs.



The FPGA-based Janus II system.

The Eurora system installed at CINECA.

QCD on GPUs

The goal of LQCD is to compute numerically, by Monte-Carlo simulations, the theory of Quarks and Gluons, Quantum Chromodynamics, on a discretized space-time Lattice. The computational task is extraordinary, typically a N x N sparse matrix must be inverted $O(10^6)$ times, with N going up to 10^8-10^9. For that reason, since the '80s LQCD has fostered the development of parallel HPC (e.g., APE machines, BlueGene machines).

The marriage between LQCD and GPUs was unavoidable. Pioneers entered the (video)game in 2006, at the time of OpenGL [1]. More friendly programming frameworks have then made GPUs widespread in our community.

OUR ACTIVITY: in 2009 we started the development of a production code entirely running on GPUs [2]

- Single GPU version: to cut CPU-GPU memory transfer, whole Molecular Dynamics (MD) runs on GPU.
- Mixed precision strategy: MD in single precision, Metropolis test and measurements in double.
- Performance is limited by internal memory bandwidth. O(100) sustained Gflops attained on single GPU.
- Code already in production on GPU farms (Pisa, Rome) for studying QCD matter in extreme conditions (Universe right after the Big Bang, heavy ion collision experiments) [3]
- Extension to many-GPU parallel architectures essential for studies with larger RAM requirement. Main challenge is internode communication during MD. Sinergy between communication-masking strategies at the programming level and custom interconnection architectures will be essential.

New Programming Approaches

Code refactoring to get extreme performances for any new architecture is not always the only choice.

We are testing the effectiveness of simpler solutions provided by modern compilers, *e.g.*

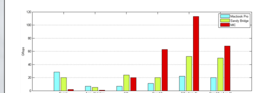
- Intel ICC makes array notation available, to help express vector parallelism within codes:

```
A[0:N] = B[0:N] + C[0:N]
```

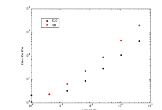
- **OpenMP** is well established for multithreading, **cilk** is also available in icc.
  OpenACC provides support for GPU mutithreading in much the same style.

All this open the way to easy, portable codes which can be as simple as in the following example:

```
# ifdef MYopenMP
# pragma omp parallel for
# endif
# ifdef MYopenACC
# pragma acc parallel  loop present ( latti )
# endif
# ifdeg MYcilk
cilk_for ( int i=0; i < setup.sz2 ; i++) {
# else
for ( int i=0; i < setup.sz2 ; i++) f
# endif
latti[y].theta1 = latti[i].theta + dt*latti[i].pi ;
}
```



(Big) Matrix multiplications on different architectures, with different combination of array notation and OpenMP or cilk

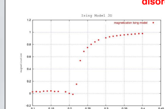


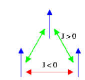
Hybrid MonteCarlo simulations of spin systems: the code is much the same for the nVidia K20 GPU (OpenACC pragmas) and for the Intel Xeon SB (OpenMP pragmas)

Spin-Glass on Many-Core and GPUs

The Spin-glass is a statistic model to study behaviours of complex macroscopic systems like

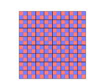
disordered magnetic materials

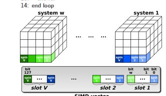




Require: set of $\{S\}$ and $\{J\}$

1: **loop** {loop on Monte Carlo steps}
2: **for all** $s_i \in \{S\}$ **do**
3: $s'_i = (s_i == 1)\ ?\ -1\ :\ 1$ {flip tentatively value of s_i}
4: $\Delta E = \sum_{\langle ij \rangle}(J_{ij} \cdot s'_i \cdot s_j) - (J_{ij} \cdot s_i \cdot s_j)$ {compute energy change}
5: **if** $\Delta E \leq 0$ **then**
6: $s_i = s'_i$ {accept new value of s_i}
7: **else**
8: $\rho = \mathrm{rnd}()$ {compute a random number $0 \leq \rho \leq 1,\ \rho \in \mathbb{Q}$}
9: **if** $\rho < e^{-\beta\Delta E}$ **then** {$\beta = 1/T,\ T = $ Temperature}
10: $s_i = s_i'$ {accept new value of s_i}
11: **end if**
12: **end if**
13: **end for**
14: **end loop**





System	Core 2 Duo	CBE (16 cores)	Janus	C1060	NH (8 cores)	C2050	SB (16 cores)	K20X	Xeon-Phi	Janus 2
Year	2007	2007	2008	2009	2009	2010	2012	2012	2013	2013
Power (W)	150	220	35	200	220	300	300	300	300	25
SUT (ps/flip)	1000	150	16	720	200	430	60	230	52	2
Energy/flip (nJ/flip)	150	33	0.56	144	244	129	18	69	15.6	0.05

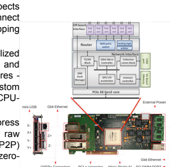
Spin-update-time (SUT) of EA simulation codes on a 64^3 lattice on several architectures.

Networks for HPC

In the race towards exaFLOPS systems one of the most critical aspects is the design of smart, efficient and robust network able to interconnect the huge number of many-core high performance processors equipping modern HPC platforms.

Leveraging on past APE experiences - 3D Torus network for specialized systems, proven effective for large scale scientific computing - and taking into account the emerging of powerful many-core architectures - mainly GPU processor - we designed APEnet+. APENet+ is a custom 3D torus interconnection architecture optimized for hybrid clusters CPU-GPU [4].

The APENet+ interconnect fabric is built on a FPGA-based PCI-express board with 6 bi-directional off-board links showing 34 Gbps of raw bandwidth per direction, and leverages upon peer-to-peer (P2P) capabilities of Fermi and Kepler-class NVIDIA GPUs to obtain real zero-copy, GPU-to-GPU low latency transfers.



The minimization of APENet+ transfer latency is achieved through the adoption of a simple RDMA protocol implemented in FPGA with specialized hardware blocks tightly coupled with embedded microprocessor.

In the framework of SUMA project we will refine the architecture and improve the performances of APEnet+. First, the adoption of the last generation 28nm FPGA will enable us to switch to Gen3 PCIe protocol on host side and to integrate faster Torus channels, doubling the I/O bandwidth on both sides of the NIC. Furthermore, the huge amount of HW resources available in the 28nm FPGA will be used to develop specific computational task accelerators in the form of an ASIP (Application Specific IP) or as a custom hardware blocks. Lastly, we will perform customization of APEnet+ IP in order to evaluate the use of torus-based unconventional interconnect topologies for computing systems dedicated to large-scale simulation in INFN non-traditional research areas (Bio-computing and Brain simulation).

References

[1] G. I. Egri, Z. Fodor, C. Hoelbling, S. D. Katz, D. Nogradi, K. K. Szabo, Lattice QCD as a video game,", Comput. Phys. Commun. 177, 631 (2007) arXiv:hep-lat/0611022.

[2] C. Bonati, G. Cossu, M. D'Elia and P. Incardona, QCD simulations with staggered fermions on GPUs," Comput. Phys. Commun. 183, 853 (2012) [arXiv:1106.5673 [hep-lat]].

[3] C. Bonati, G. Cossu, M. D'Elia and F. Sanflippo, Phys. Rev. D 83, 054505 (2011) [arXiv:1011.4515 [hep-lat]]; M. D'Elia, M. Mariti and F. Negro, Phys. Rev. Lett. 110, 082002 (2013) [arXiv:1209.0722 [hep-lat]]; C. Bonati, M. D'Elia, M. Mariti, F. Negro and F. Sanflippo, arXiv:1307.8063 [hep-lat].

[4] R. Ammendola et al, "APEnet+: a 3D Torus network optimized for GPU-based HPC Systems", (2012) J. Phys.: Conf. Ser. 396 042059; R. Ammendola et al, "GPU peer-to-peer techniques applied to a cluster interconnect", Proceedings of CASS2013 workshop, accepted for publication (arXiv:1307.8276).