



IICT

INSTITUTE OF INFORMATION AND
COMMUNICATION TECHNOLOGIES
BULGARIAN ACADEMY OF SCIENCE



1869

Научна инфраструктура на ИИКТ-БАН – достъп, възможности, обучение

проф. д-р Емануил Атанасов
р-л секция ВСМА

email: emanouil@parallel.bas.bg



Academic HPC resources in Bulgaria

AVITOHOL at IICT-BAS

150x HP ProLiant SL250s Gen8 each with
 2x Intel Xeon E5-2650 v2 (8C/16T),
 64 GB DDR3-1866 RAM and
 2x Intel Xeon Phi 7120P
 6x HP ProLiant DL380p Gen8 nodes with
 2x Intel Xeon E5-2650v2 (8C/16T),
 64 GB DDR3-1866 RAM
 Infiniband 56 Gb/s FDR
 Storage system with 96 TB



Total Performance:
 RPeak: 412.3 TFlop/s
 RMax: 264.2 TFlop/s
 Top 500 position: 389

HPCG cluster at IICT-BAS

36 blades BL 280c(2x Intel X5560(4C/8T); 24GB DDR3);
 8 management nodes HP DL 380 G6(2x Intel
 X5560(4C/8T); 32GB DDR3);
 2 HP ProLiant SL390s G7(2x Intel E5649(6C/12T);96GB
 DDR3)
 8x nVidia TESLA M2090 per server;
 2 HP SL270s Gen8 (2x Intel Xeon E5-2650 v2(8C/16T);
 128GB DDR3)
 Total number of Xeon Phi 5110P coprocessors: 9
 Total 132TBs of system storage



TOTAL PERFORMANCE:
 RPEAK: 22.94 TFlop/s

NCSA IBM Blue Gene/P

8192 PowerPC 450 processors
 4TBs of system memory
 12TBs of system storage
 IBM proprietary interconnect with
 2.5 µs latency and 10GBps bandwidth

TOTAL PERFORMANCE:
 RPEAK: 27.85 TFlop/s
 RMAX: 23.45 TFlop/s



PHYSON at Sofia University

53 Intel Xeon x86_64 processors
 524Gibs of system memory
 6.5TBs of system storage
 2x nVidia Tesla M2090 graphics processors



TOTAL PERFORMANCE:
 RPEAK: 3.57 TFlop/s
 RMAX: 3.22 TFlop/s

MADARA at IIOCCP-BAS

54 Primergy RX200 S5 servers with
 2 Intel Xeon E5520(4C/8T) each
 and a total of 800GB DDR3 1066MHz
 20Gb/s DDR Infiniband
 108TB System Storage by Fujitsu FibreCat SX100



The Avitohol supercomputer



Hardware features

Servers	150 x Dual CPU HP ProLiant SL250s Gen8
CPUs	Intel Xeon E5-2650v2 2.6GHz – 8 cores /16 HT
RAM	64 GB per node
Coprocessors	300 x Intel Xeon Phi 7120P(x86) - 61 cores / 244 HT
Total CPUs	2400 cores/4800 HT + 18300 cores/ 73200 HT
Total RAM	14400 GB (9600 + 4800)
Disk Storage	100 TB
Interconnect	Non-blocking FDR Infiniband
Latency	1.1 μ s
Bandwidth	56 Gbps

Environment and performance

Peak Performance CPU	50 Tflop/s
Pear Performance Accelerators	362 Tflop/s
Total Peak Performance	412 Tflop/s
Real Measured Performance	264 Tflop/s
Max Power	250 kW



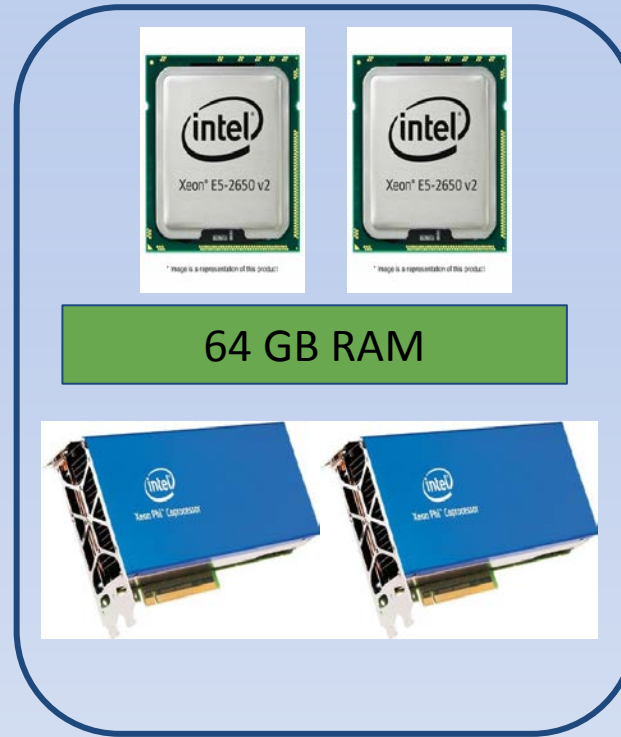
The system consists of 4 dual racks of type HP MCS 200. Each one provides power and cooling for up to 50 kW of equipment, cooled by water. Last on Top500 List on 389 place (Nov 2015) <http://www.top500.org/system/178609>

Xeon Phi accelerators

- x86 Architecture
- 61 cores
- 244 threads
- 512 bit SIMD
- 16 GB RAM
- 352 GB/s bandwidth
- Works synergistically with Intel Xeon processors
- Give about 90% of the theoretical performance



Architecture of one node



↕ ↕

Low-Latency, High-Bandwidth Network -
Infiniband FDR 56 Gbps

- RedHat Enterprise for HPC
- Intel Cluster Studio (compilers, MPI, MKL)
- GNU Compiler Collection
- Torque/moab for resource management with web and command line interface
- Monitoring and accounting tools
- Services running on virtual machines
- Rich set of applications and libraries like:
 - Boost, mumps, octave, OpenFOAM, WRF, WRF-Chem, NetCDF, Gromacs,...



- 36 blade servers 2 Xeon X5560 @ 2.80GHz, 24 GB RAM – 576 cores total
- 132 TB - three SAN systems with Fibre Channel interconnects
- 2 servers with 16 NVIDIA Tesla M2090 graphic cards
- 2 server with up to 8 Intel Xeon Phi 5110P Coprocessors
- Interconnected with non-blocking InfiniBand
- Peak performance 22.93 Tflop/s



Една карта NVIDIA Tesla V100 PCIe 32 GB има 5120 графични ядра, което осигурява 14,131 Gigafllops single precision, 7.0 Tflops double precision, 32 GB HBM2 RAM, 900.1 GBytes/sec bandwidth, 250 W Max power consumption

- В момента имаме 12 карти, т.е. над 168 Teraflops single precision.
- Процесори Intel(R) Xeon(R) Gold 5118 CPU @ 2.30GHz – 12 ядрени, по 2 на сървър
- Памет 128 GB
- 2x800GB SSD
- 3*12 TB HDD
- Софтуер NVIDIA GRID
- Операционна система RED HAT Enterprise Linux 7



- GPGPU означава използване на графични карти (GPU) за изчисления от общ характер.
- При това се използва големия брой транзистори и съответно изчислителни елементи, налични в графичните карти, за извършване на изчисления, които обикновено се правят на централния процесор
- Съвременните графични карти имат по-голям брой транзистори и много по-високо ниво на паралелизъм на от CPU, което води до по-добра ефективност от гледна точка на цена, енергопотребление, заемано място.
- За GPGPU изчисления могат да се използват масови карти, използвани предимно за компютърни игри, но основните производители NVIDIA и AMD (ATI) предоставят специализиран хардуер за инсталацията на HPC
- Най-популярни методи за програмиране на GPU
- OpenCL – стандарт за паралелно програмиране, който може да се използва също за паралелна обработка върху многоядрени процесори
- NVIDIA CUDA - достъпна само за графични процесори на NVIDIA



- За да използваме GPGPU трябва да разполагаме с GPU хардуер и съответна софтуерна поддръжка
- Няколко начина за използване на предимствата на GPGPU
- Използване на цяло приложение – GROMACS, NAMD, Amber
- Използване на библиотеки – FFT, BLAS -> cuFFT, NVBLAS
- Пренаписване на собствен код
- Различни нива на паралелизъм
- Една карта
- Няколко карти на един сървър (OpenMP)
- Няколко сървъра с по няколко карти (MPI)

- Достъпът е безплатен за учени и докторанти
- Попълват се апликационни форми, в зависимост от желаниа хардуер
 - http://ict.acad.bg/?page_id=41
 - <http://www.hpc.acad.bg/bg/access-policies/>
- Могат да се изтеглят онлайн, но се предават подписани, заедно със съответната политика

- Курсове към ЦОД към БАН
- http://edu.bas.bg/doctorant_school/spec_courses/1_inf_com_sci_tech/information.html
 - Високопроизводителни изчисления – проф. Т. Гюров
 - Разпределени пресмятания – проф. Е. Атанасов
- Редовни обучения по различни проекти – може да се следи на страницата на секция ВСМА или
 - <http://www.hpc.acad.bg/bg/events/>
- Индивидуални консултации