Architetture SoC e FPGA per Calcolo Scientifico e Riduzione Dati

Sara Bertocco

David Goz, Luca Tornatore, Giuliano Taffoni

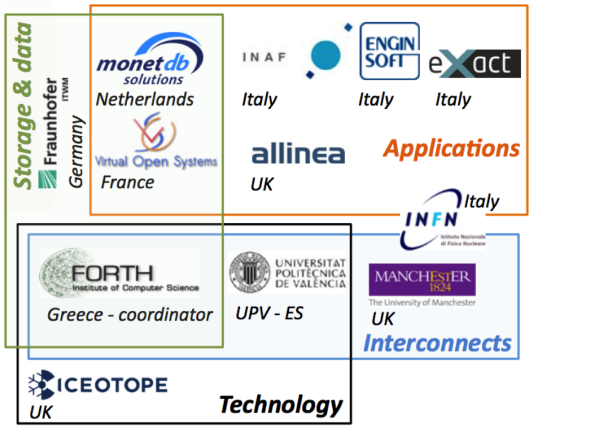
ICT Workshop INAF 2018 – Catania, 10-14 Settembre

Come nasce l'idea

- Collaborazione
  - Software HPC
    - Porting di software
  - Infrastrutture di calcolo

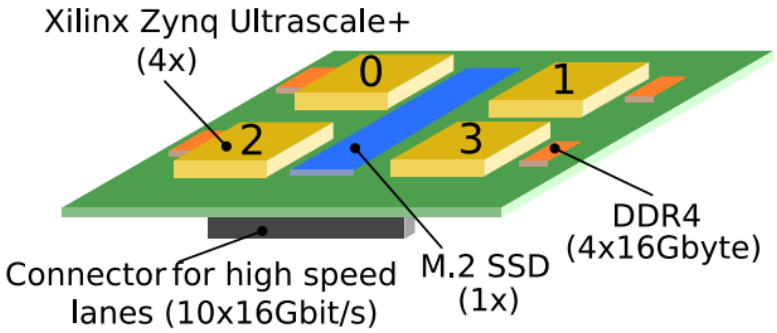
Progetto ExaNest

- Progetto Horizon2020 che vuole dimostrare la fattibilità di sistemi HPC ExaScale basati su tecnologia europea



- Consorzio tra realtà accademiche e industriali

- Riprogettare applicazioni software HPC per sfruttare nuove architetture con potenzialità exascale

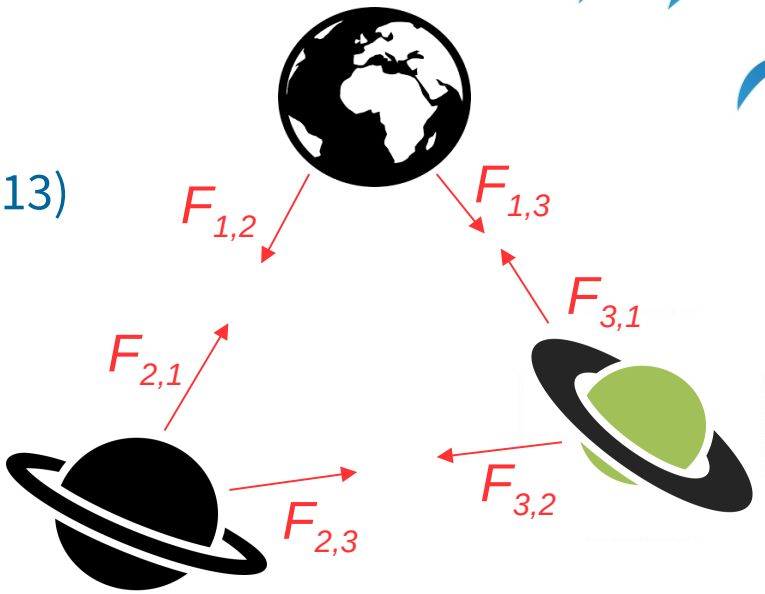


ExaNeSt compute unit:
- 4 Xilinx Zynq Ultrascale+ FPGAs;
- 4 ARMv8 cores @1.5GHz per FPGA;
- 16 GB of DDR4 memory per FPGA;
- one NVM SSD storage device.

INAF
ISTITUTO NAZIONALE DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI TRIESTE

Codice N-body diretto

Codice N-body diretto ricavato da HiGPUs (R.Capuzzo-Dolcetta, M.Spera, D.Punzo, 2013) un software usato per lo studio dell'evoluzione dinamica di sistemi stellari (complessità $O(N^2)$)



Device utilizzabili:

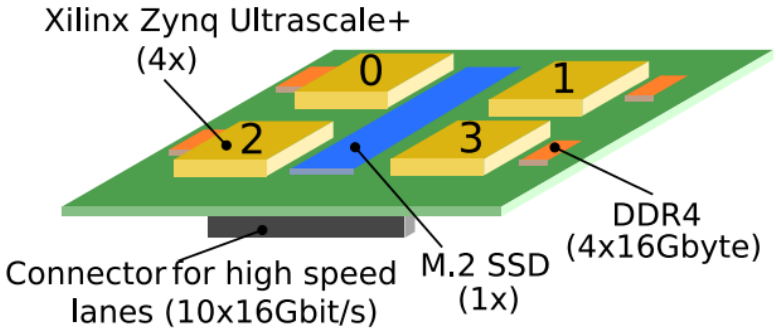
- OpenCL compatibili (es. CPU, GPU, FPGA, SoC)

Schema di parallellizzazione:

- MPI+OpenMP nell'host
- OpenCL nel device

Riassumendo:

- Codice N-body diretto reingegnerizzato e ottimizzato per componenti SoC OpenCL compatibili
- Hardware target (progetto ExaNest):

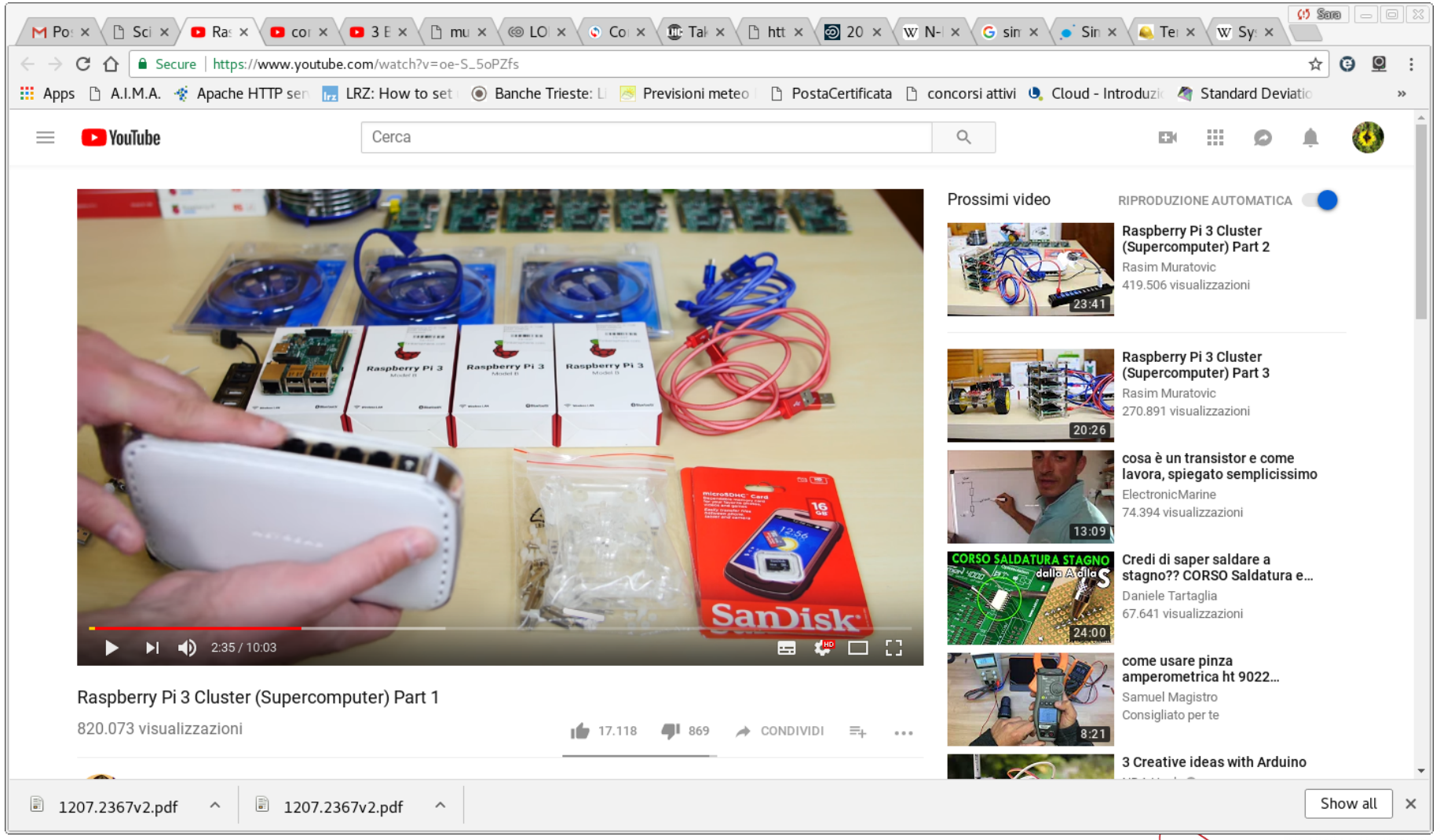


ExaNeSt compute unit:
- 4 Xilinx Zynq Ultrascale+ FPGAs;
- 4 ARMv8 cores @1.5GHz per FPGA;
- 16 GB of DDR4 memory per FPGA;
- one NVM SSD storage device.

Primo step

- Porting software su OpenCL eseguito
- Test su CPU intel eseguiti
- Test su singola GPU dedicata eseguiti
- Hardware ibrido?

Technology watch



Raspberry Pi

Raspberry Pi è una famiglia di single board computer, computer della dimensione di una carta di credito, a basso costo e basso consumo, che possono essere collegati a monitor, tastiera e mouse.
Usati per fare formazione, internet delle cose (IoT), macchine per calcolo HPC

Technology watch

https://www.lanl.gov/discover/news-release-archive/2017/November/1113-raspberry-pi.php

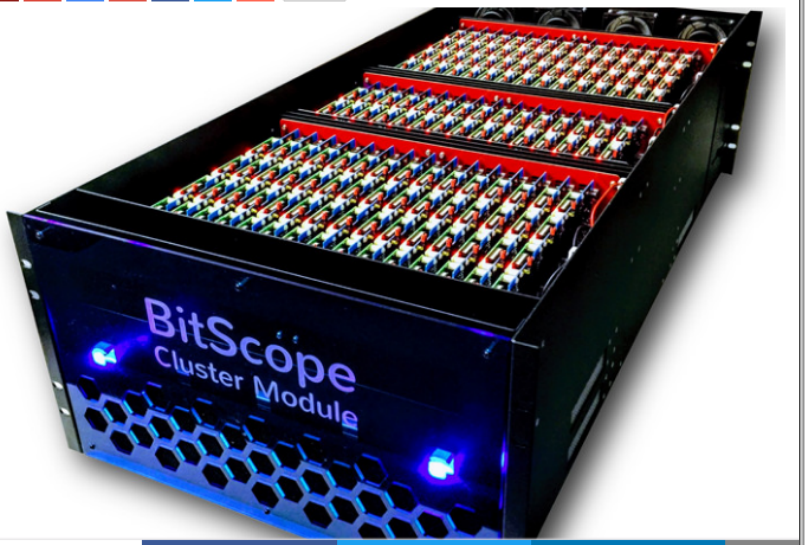
Los Alamos National Laboratory
Delivering science and technology to protect our nation and promote world stability

SCIENCE & INNOVATION | COLLABORATION | CAREERS | COMMUNITY

Our Stories » News Releases » News Releases - 2017 » November » Raspberry Pi

Scalable clusters make HPC R&D easy as Raspberry Pi

It brings a powerful high-performance-computing testbed to system-software developers, researchers and others who lack machine time on the world's fastest supercomputers.



https://www.punto-informatico.it/il-supercomputer-da-750-raspberry-pi/?refresh_ce

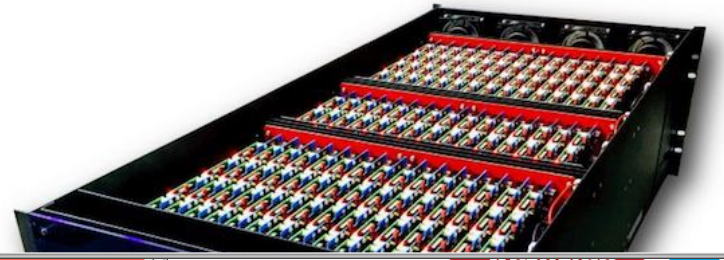
Pi

Il supercomputer da 750 Raspberry Pi

Il supercomputer basato su ARM acquistato dal laboratorio di Los Alamos garantisce costi e consumi nettamente inferiori rispetto ad altri sistemi di High Performance Computing. Il suo scopo principale? Servire da banco di prova per gli sviluppatori, lasciando così liberi gli altri HPC per test e simulazioni

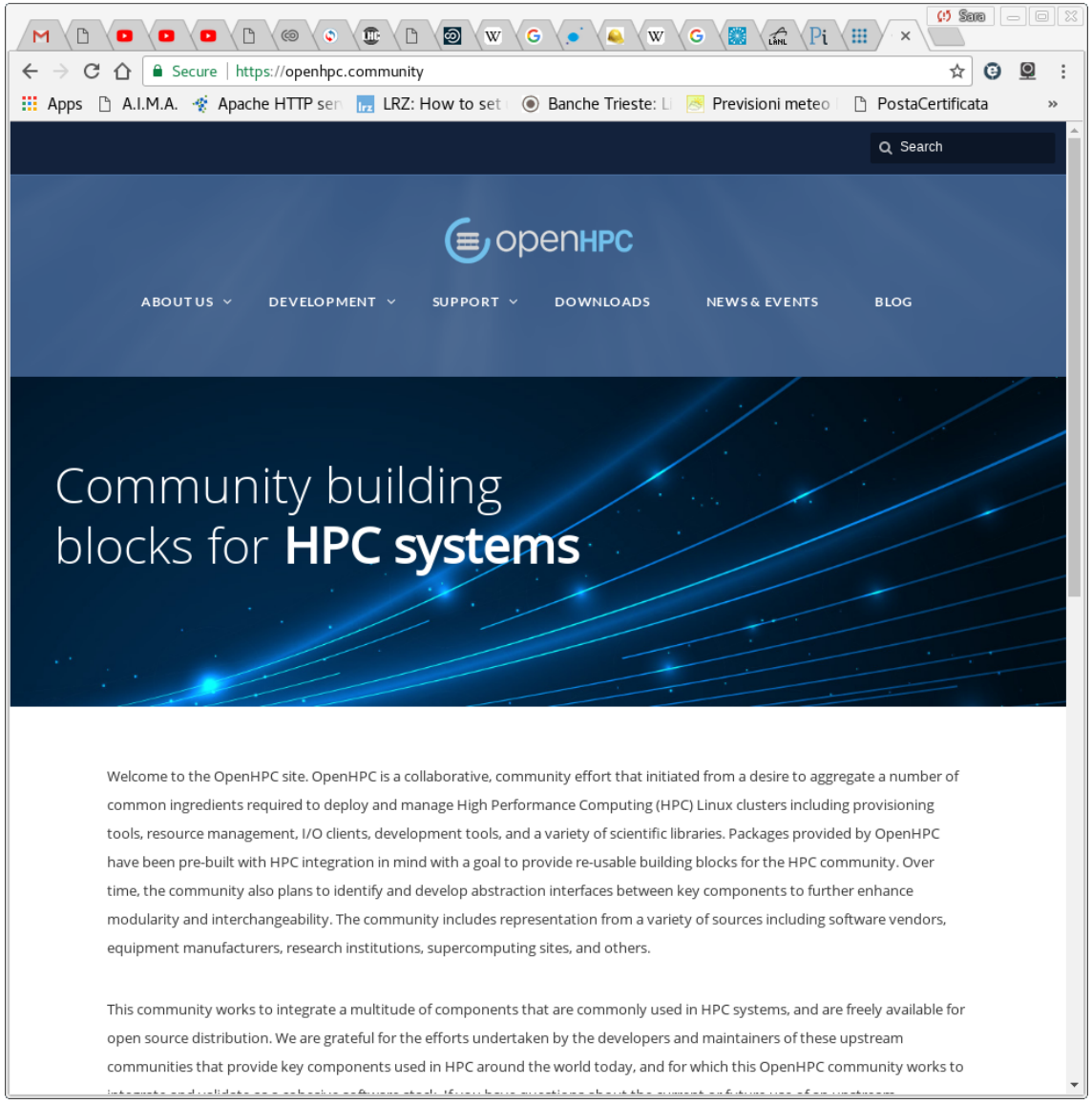
Roma – Il laboratorio nazionale di Los Alamos ha presentato alla conferenza SuperComputing 17 il suo nuovo supercomputer **costruito utilizzando 750 Raspberry Pi** . L'HPC in questione è basato su un rack composto da cinque moduli prodotti da BitScope Designs , a loro volta costituiti da 150 Raspberry Pi ciascuno: **il risultato finale è una piattaforma di ben 3.000 core** dai consumi relativamente esigui se paragonati agli altri supercomputer basati su architetture differenti già in possesso del laboratorio, come Trinity e Crossroads.

Elia Tufarolo
05 Dicembre 2017



DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI TRIESTE

Technology watch



Technology watch

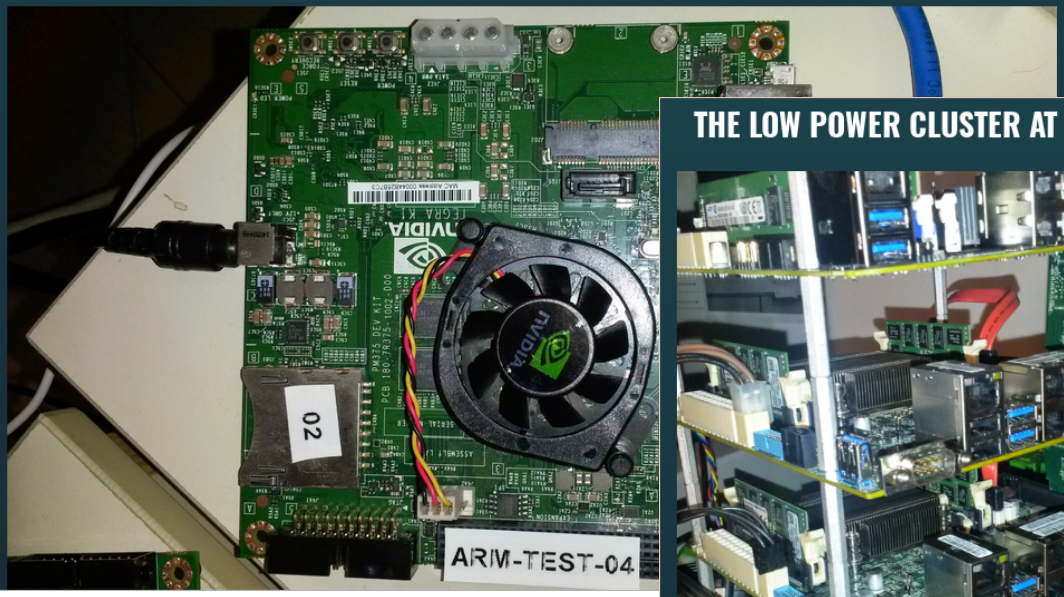
Home - Mozilla Firefox

www.cosa-project.it/home.html

HOME | THE COSA CLUSTERS | APPLICATIONS | DOCUMENTS | CONTACTS

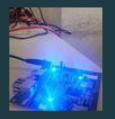
INFN COSA Project

COMPUTING ON SOC ARCHITECTURE



THE LOW POWER CLUSTER AT CNAF

THE CLUSTER IS MADE OF TWO SUB-CLUSTERS:
- A 32BIT ARM BASED CLUSTER WITH 1 GB/S INTERCONNECTION
- A 64BITCLUSTER BASED ON INTEL LOW POWER SOCS WITH 10GB/S INTERCONNECTION

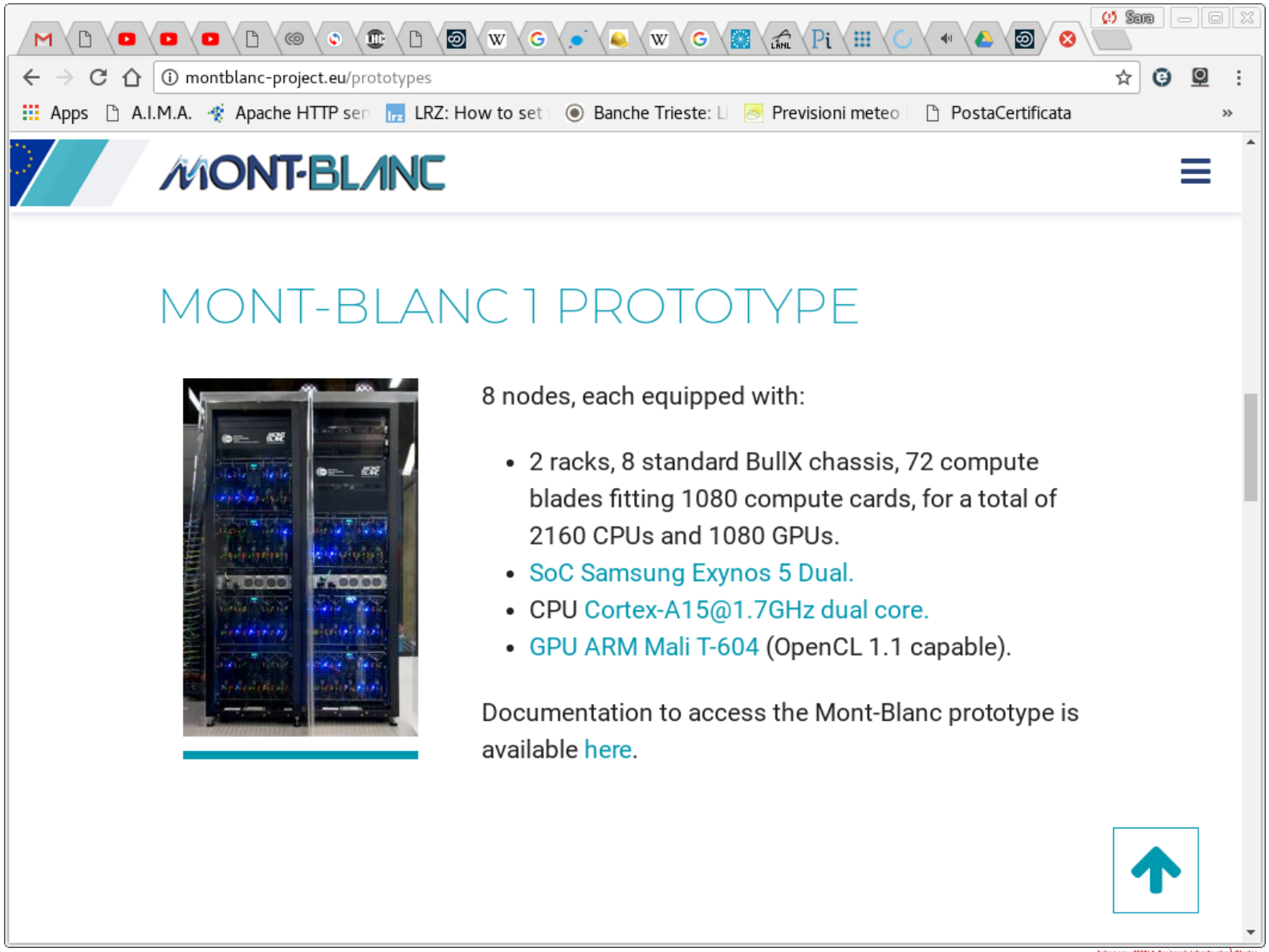


FEATURES

The ARM cluster is composed by 8 NVIDIA JETSON TK1, 4 NVIDIA JETSON TX1, two ODROID-XU3. 1 CUBIEBORD, 1 SABREBOARD, 1 ARNDALE OCTA board. Boards are interconnected with standard 1Gb/s ethernet.

The 64bit INTEL cluster is composed by 4 mini-ITX boards with the C2750 AVOTON SoC plus 4 mini-ITX motherboard based on the new XEOND-1540 CPU plus 4 N3700-ITX boards, plus 2 J4205 PENTIUM boards. Interconnection are based on 1Gb/s and 10Gb/s ethernet.

Technology watch



Raspberry Pi: caratteristiche

- Raspberry Pi 3 Model B+
- SoC Broadcom BCM2837B0
- CPU: 1.4 GHz 64/32-bit quad-core ARM Cortex-A53
- Grafica: Broadcom VideoCore IV 300 MHz/400 MHz (GPU)
- Memoria: 1 GB LPDDR2 RAM at 900 MHz
- Potenza: 1.5 W (media in idle), 6.7 W (max sotto stress)
- Storage: MicroSDHC slot
- Wireless LAN Dual-band 802.11ac e Bluetooth 4.2
- Gigabit Ethernet (over USB)
- Costo: $35

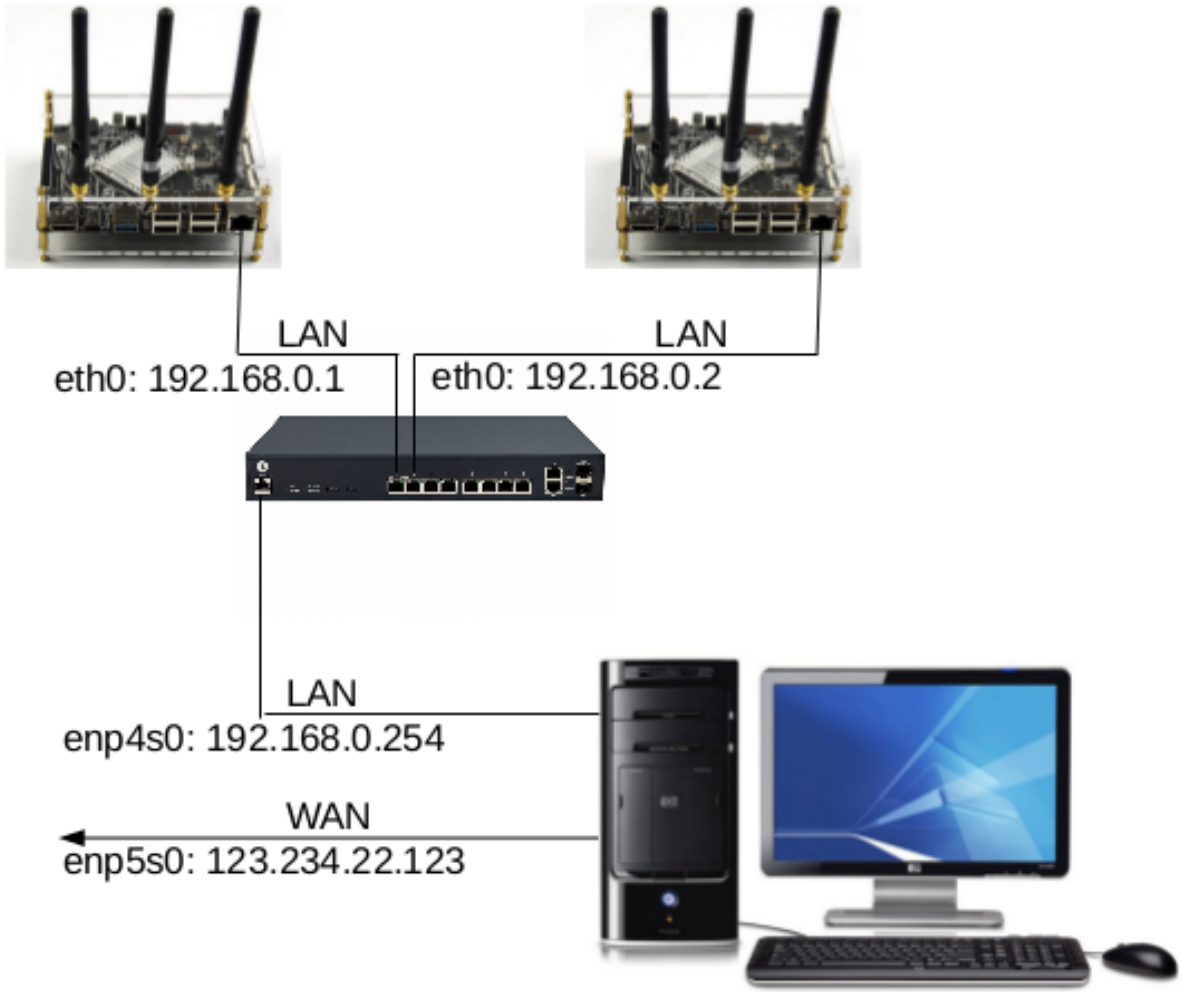
Punti deboli

- Ethernet è limitata a ~ 300 Mbit/s dal bus USB 2.0 tra l'adattatore Gigabit Ethernet e il SoC
- La memoria arriva a 1G di RAM massima per scheda
- Nessun device utilizzabile con OpenCL

Firefly-RK3399

- CPU: due Cortex-A72 2.0GHz, e 4 Cortex-A53 1.42GHz
- GPU: Mali-T864 quad-core con supporto OpenGL ES1.1/2.0/3.0/3.1 and OpenCL
- Memoria: 4GB Dual-Channel DDR3
- Potenza: 4.3 W idle mode (24W max)
- Storage: 16GB eMMC
- Wireless LAN 802.11ac e Bluetooth 4.1
- Gigabit Ethernet: 1 porta
- Costo: 300 euro su Amazon, 300 X 1.5 per noi :(

Cluster layout e networking



Cluster Manager SLURM

- SLURM: Simple Linux Utility for Resource Management è un job scheduler open source e free
- Usato in molti supercomputer nel mondo
- Fornisce:
  - Allocazione esclusiva o meno delle risorse di calcolo (nodi)
  - E` un framework dove eseguire e monitorare dei job (tipicamente paralleli – MPI) di calcolo su un insieme di risorse allocate (nodi)
  - Ha un sistema di accodamento in caso di carico

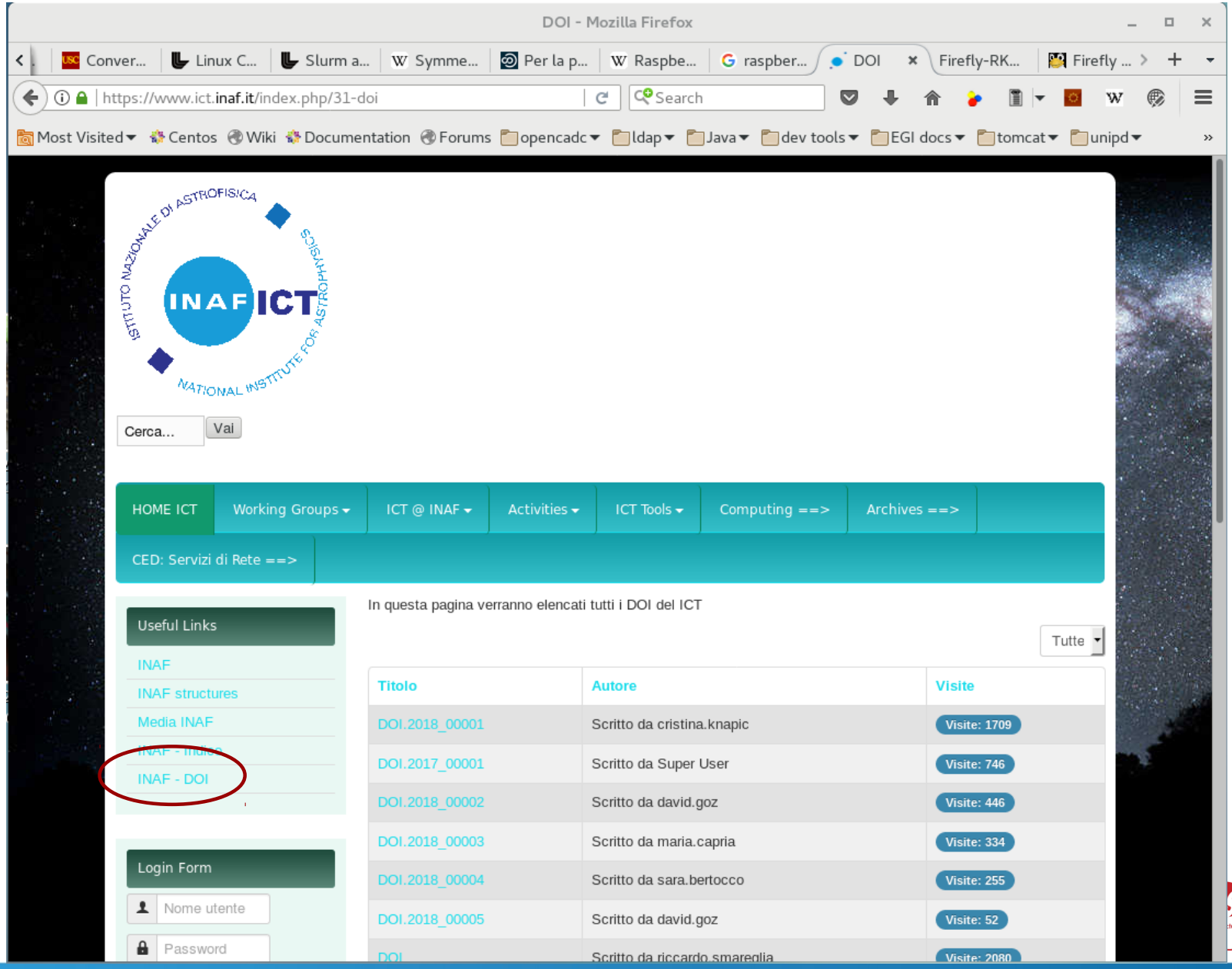
Pubblicità

"INAF-OATs Technical Report 222 - INCAS: Intensive Clustered ARM SoC - Cluster Deployment"

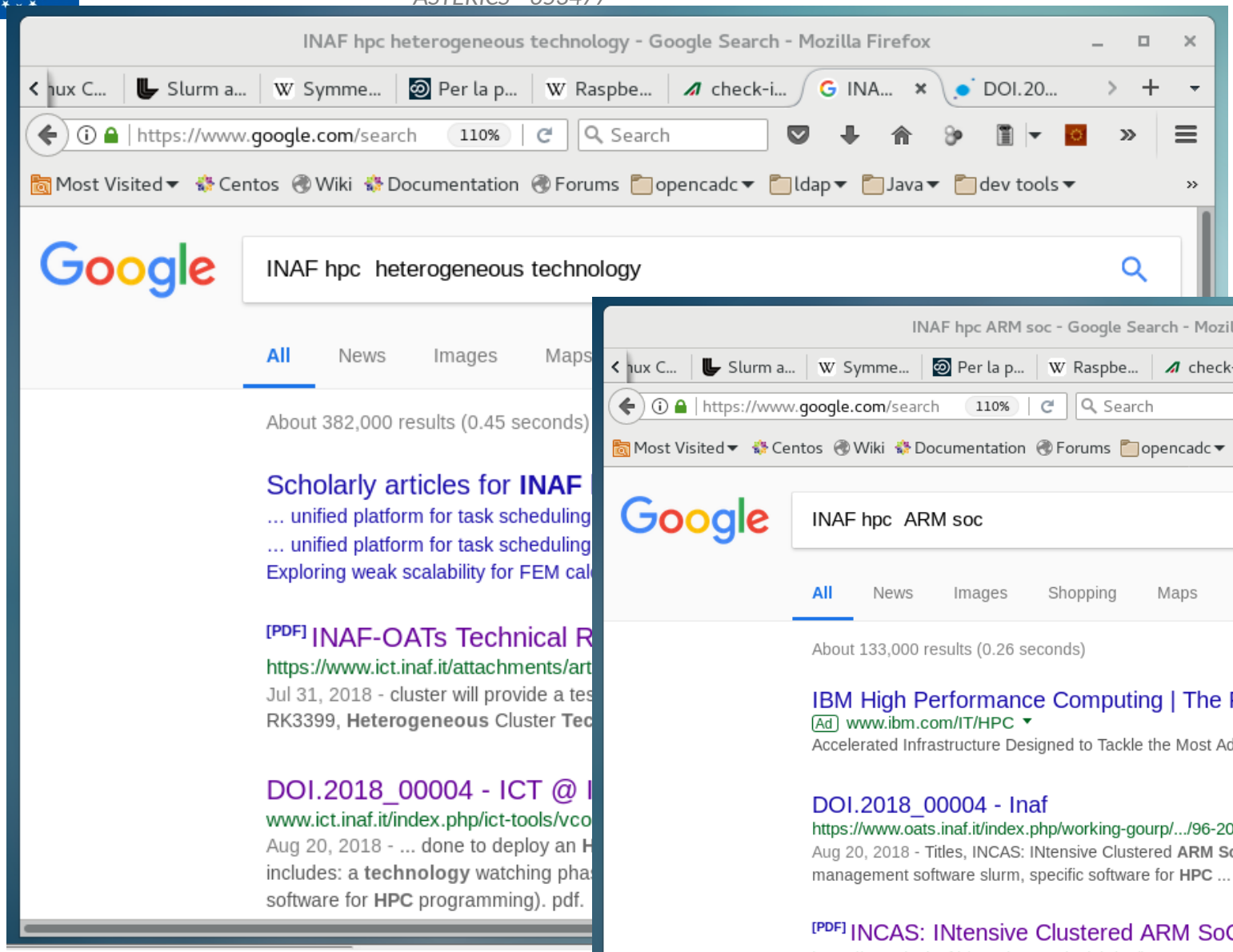
https://doi.org/10.20371/INAF/PUB/2018_00004

https://www.ict.inaf.it/attachments/article/96/bertocco_TechRepOATs_222.pdf

Pubblicità



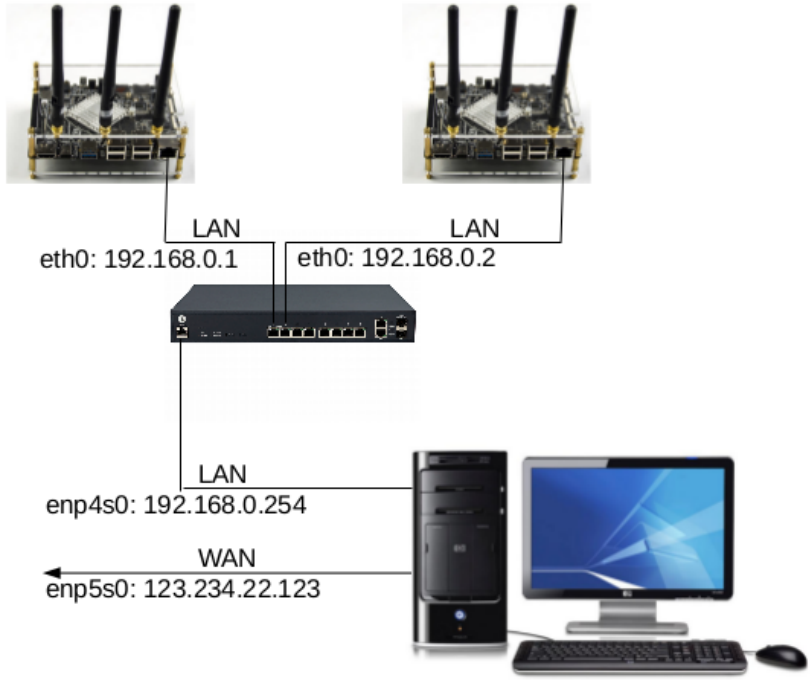
INAF
ISTITUTO NAZIONALE DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI TRIESTE


INAF hpc heterogeneous technology - Google Search - Mozilla Firefox
https://www.google.com/search
110%
Search
Most Visited
Centos
Wiki
Documentation
Forums
opencadc
ldap
Java
dev tools
Google
INAF hpc heterogeneous technology
All
News
Images
Maps
About 382,000 results (0.45 seconds)
Scholarly articles for INAF
... unified platform for task scheduling
... unified platform for task scheduling
Exploring weak scalability for FEM cal
[PDF] INAF-OATs Technical R
https://www.ict.inaf.it/attachments/art
Jul 31, 2018 - cluster will provide a tes
RK3399, Heterogeneous Cluster Tec
DOI.2018_00004 - ICT @ I
www.ict.inaf.it/index.php/ict-tools/vco
Aug 20, 2018 - ... done to deploy an H
includes: a technology watching pha
software for HPC programming). pdf.

INAF hpc ARM soc - Google Search - Mozilla Firefox
https://www.google.com/search
110%
Search
Most Visited
Centos
Wiki
Documentation
Forums
opencadc
ldap
Java
dev tools
Google
INAF hpc ARM soc
All
News
Images
Shopping
Maps
More
Settings
Tools
About 133,000 results (0.26 seconds)
IBM High Performance Computing | The Future of Data Analytics
Ad www.ibm.com/IT/HPC
Accelerated Infrastructure Designed to Tackle the Most Advanced Workloads. Simplified Software.
DOI.2018_00004 - Inaf
https://www.oats.inaf.it/index.php/working-gourp/.../96-2018-4 - Translate this page
Aug 20, 2018 - Titles, INCAS: INtensive Clustered ARM SoC - Cluster Deployment ... (cluster management software slurm, specific software for HPC ...
[PDF] INCAS: INtensive Clustered ARM SoC - Cluster ... - ICT @ INAF
https://www.ict.inaf.it/attachments/article/96/bertocco_TechRepOATs_222.pdf
Jul 31, 2018 - Clustered ARM SoC - Cluster Deployment. S. Bertoccoa ... cluster will provide a testbed to validate and optimize several HPC codes developed ...
ARM in Datacenter/HPC - Ars Technica OpenForum

Passi per l'installazione

- Installazione OS (ubuntu 16.04 LTS)
- Configurazione rete
  - Nomi host
  - NAT (iptables)
- SLURM master e nodi di calcolo
  - Sincronizzazione clock
  - NFS per aree condivise (home utenti e storage)
  - Demoni SLURM e configurazione (slurm.conf)
- Strumenti per HPC



Strumenti (ottimizzati)

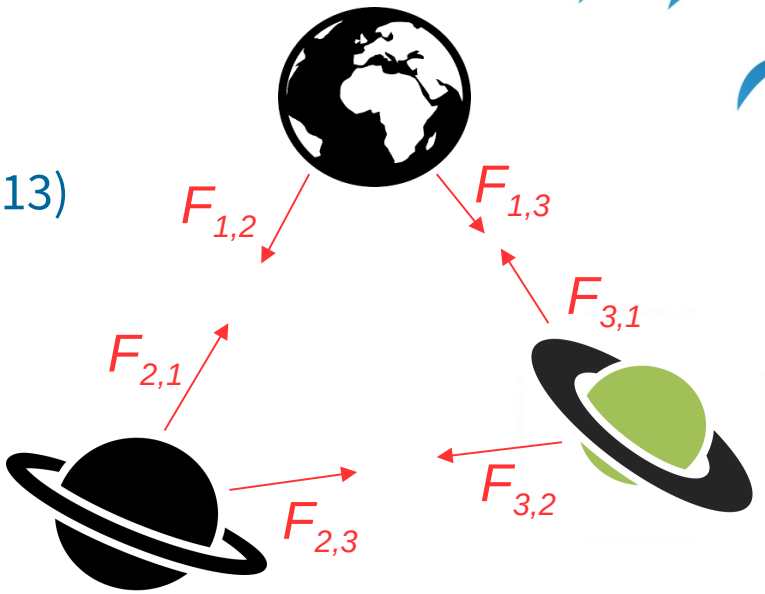
- Compilatore C/C++
- Debugger e strumenti di profiling
- OpenMPI
- Configurazione di SLURM per supporto MPI

Software di test: “Hello world”

```c
#include <mpi.h>
#include <stdio.h>

int main(int argc, char** argv) {
    // Initialize the MPI environment
    MPI_Init(NULL, NULL);

    // Get the number of processes
    int world_size;
    MPI_Comm_size(MPI_COMM_WORLD, &world_size);

    // Get the rank of the process
    int world_rank;
    MPI_Comm_rank(MPI_COMM_WORLD, &world_rank);

    // Get the name of the processor
    char processor_name[MPI_MAX_PROCESSOR_NAME];
    int name_len;
    MPI_Get_processor_name(processor_name, &name_len);

    // Print off a hello world message
    printf("Hello world from processor %s, rank %d out of %d processors\n",
           processor_name, world_rank, world_size);

    // Finalize the MPI environment.
    MPI_Finalize();
}
```

Codice N-body diretto

Codice N-body diretto ricavato da HiGPUs (R.Capuzzo-Dolcetta, M.Spera, D.Punzo, 2013) un software usato per lo studio dell'evoluzione dinamica di sistemi stellari (complessità $O(N^2)$)



Device utilizzabili:

- OpenCL compatibili (es. CPU, GPU, FPGA, SoC)

Schema di parallellizzazione:

- MPI+OpenMP nell'host
- OpenCL nel device

Software di test (altra pubblicità)

- "INAF-OATs Technical Report 223: Direct N-body code designed for heterogeneous platforms"
  DOI: https://doi.org/10.20371/INAF/PUB/2018_00002

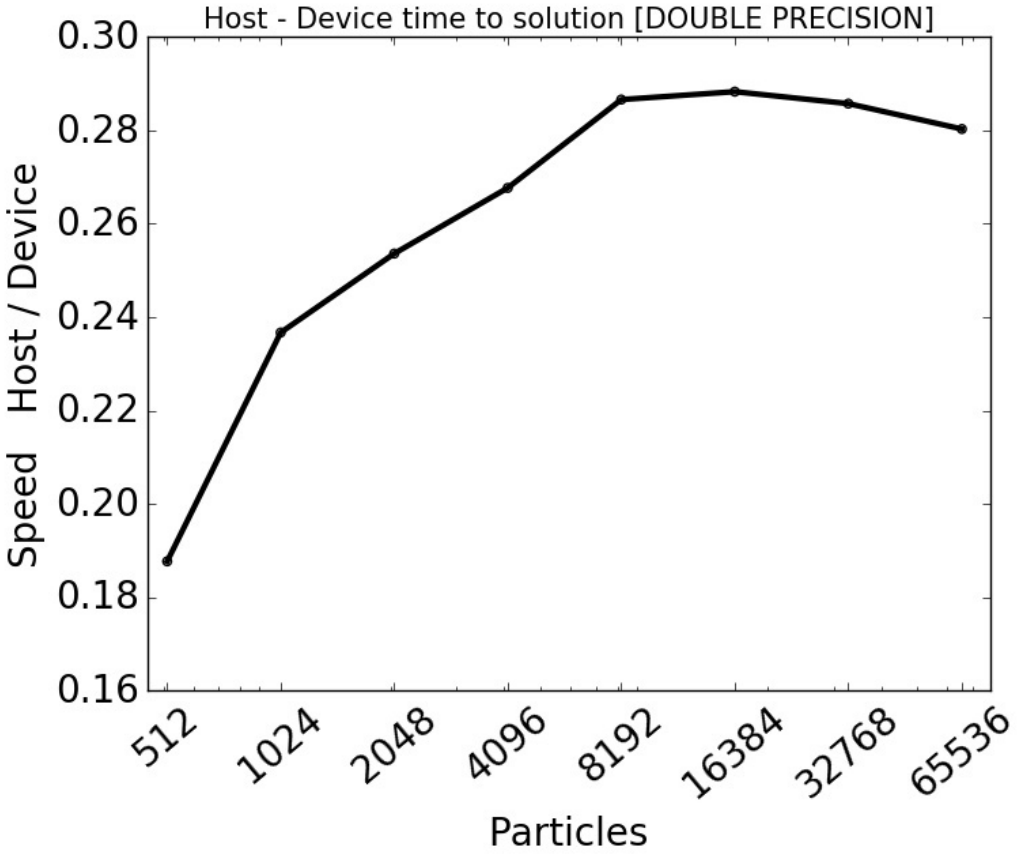
- "INAF-OATs Technical Report 224: Direct N-body code designed for a cluster based on heterogeneous computational nodes"
  DOI: https://doi.org/10.20371/INAF/PUB/2018_00005

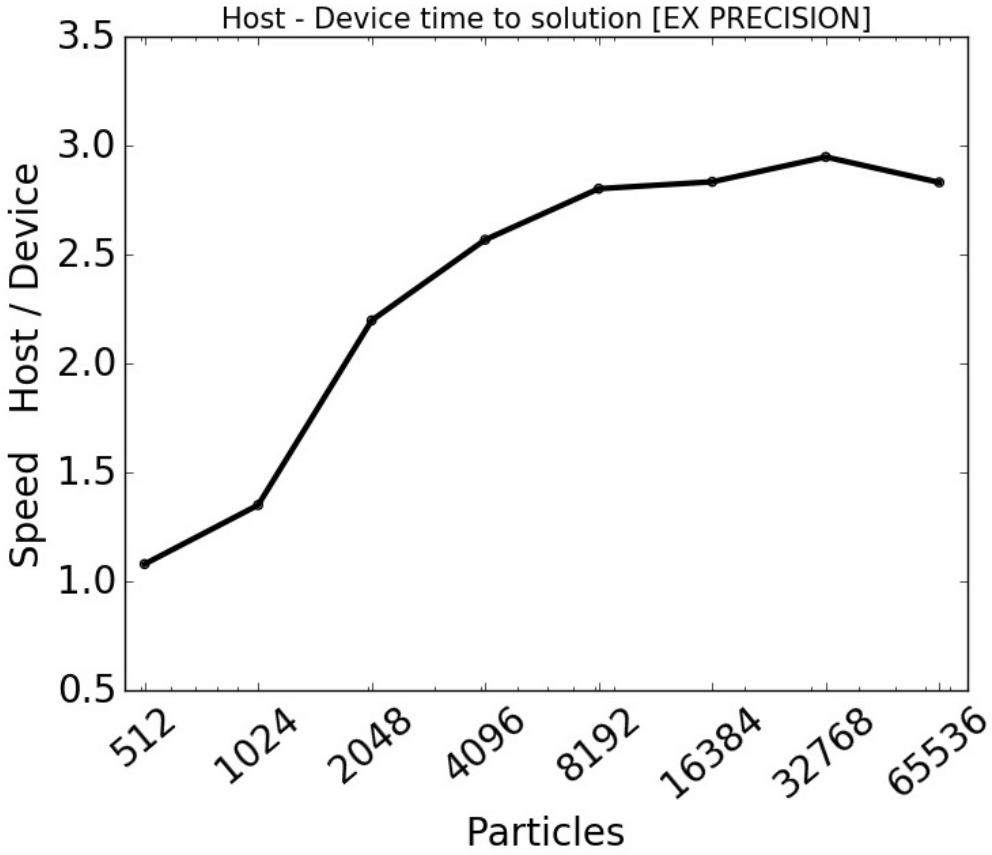
- Continua …

Scopo

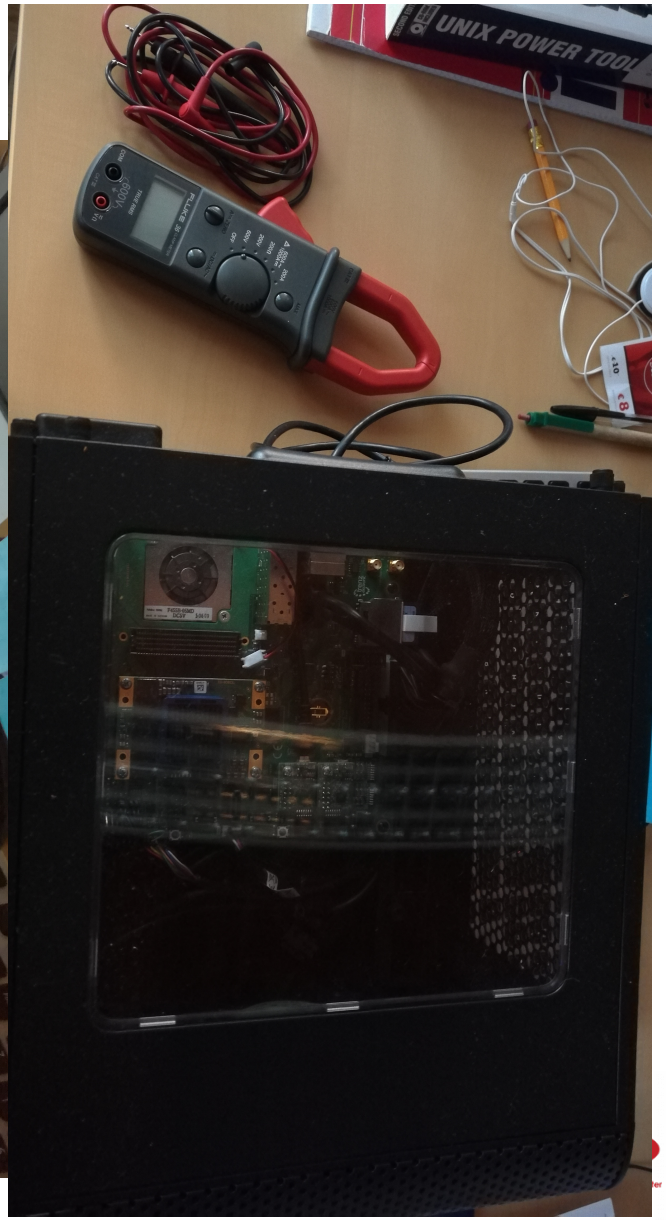
- Usare un tipico codice astrofisico (non un benchmark standard o una libreria) per:
  - capire come si dovranno scrivere i codici (ovvero pensare algoritmi) per architetture ibride e per tipologie diverse all'interno della stessa famiglia
    es. ho un codice che gira su CPU. Come lo ottimizzo per ARM piuttosto che per Intel?
  - ha senso farlo? ci sono delle strategie di massima?
  - quale incremento di prestazioni ottengo se lo faccio?
  - per sfruttare devices diversi, architetture diverse ha senso sviluppare algoritmi oblivious in cui si cerca di mitigare architetture diverse facendo in modo di avere prestazioni decenti ovunque?
  - Che prestazoni si ottengono su SoC?
  - Come far incontrare power efficiency e performance pura?
  - studiare linguaggi o tool diversi, vedendone pro e contro.
    Es. se voglio utilizzare una GPU posso usare OpenCL o OpenACC.... quali sono i pro/contro?

Performance of N-body code on ARM64 SoC





Prossimo step: FPGA e misure



INAF
ISTITUTO NAZIONALE DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI TRIESTE

Acknowledgements

This work has been supported by the European Commission through the projects ASTERICS (grant no. 653477) and ExaNeSt (FET-HPC) (grant no. 671553), funded in the framework of the European Union's Horizon 2020 research and innovation program.

Astronomy ESFRI & Research Infrastructure Cluster
ASTERICS - 653477

Grazie per l'attenzione!

Domande?

sara.bertocco@inaf.it
david.goz@inaf.it
luca.tornatore@inaf.it
giuliano.taffoni@inaf.it

Backup Slides

Nomenclatura

- SoC: System on a Chip

  è un circuito integrato che integra tutte le componenti di un computer (CPU, memoria, porte di I/O, grafica)

- OpenCL: Open Computing Language

  è un framework per scrivere programmi eseguibili su piattaforme eterogenee (CPU, GPU, DSP, FPGA)

- FPGA: Field Programmable Gate Array

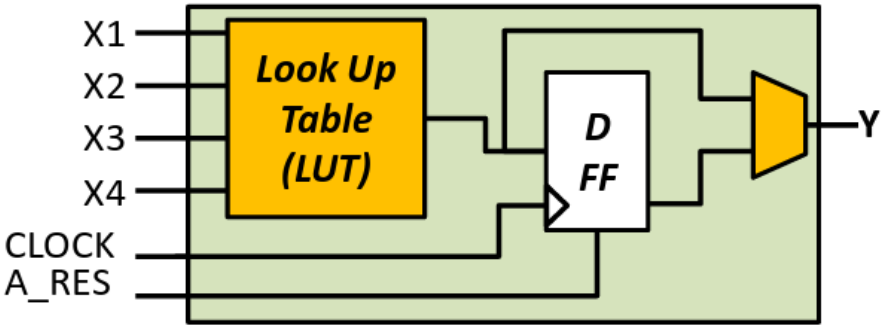
  è un circuito integrato progettato per essere programmabile

FPGA - Architettura

ALMA MATER STUDIORUM – UNIVERSITY OF BOLOGNA
DEIS - DEPARTMENT OF ELECTRONICS, COMPUTER ENGINEERING AND SYSTEMS

Architettura di un FPGA

Logic cell



Modello generale di riferimento

I dettagli delle celle logiche in realtà variano in base al produttore ed al modello dell'FPGA.

Look-Up Table

X1	X2	X3	X4	Y
0	0	0	1	?
0	0	1	0	?
...	...	...	...	?
1	1	1	1	?

Tipicamente ogni cella comprende:

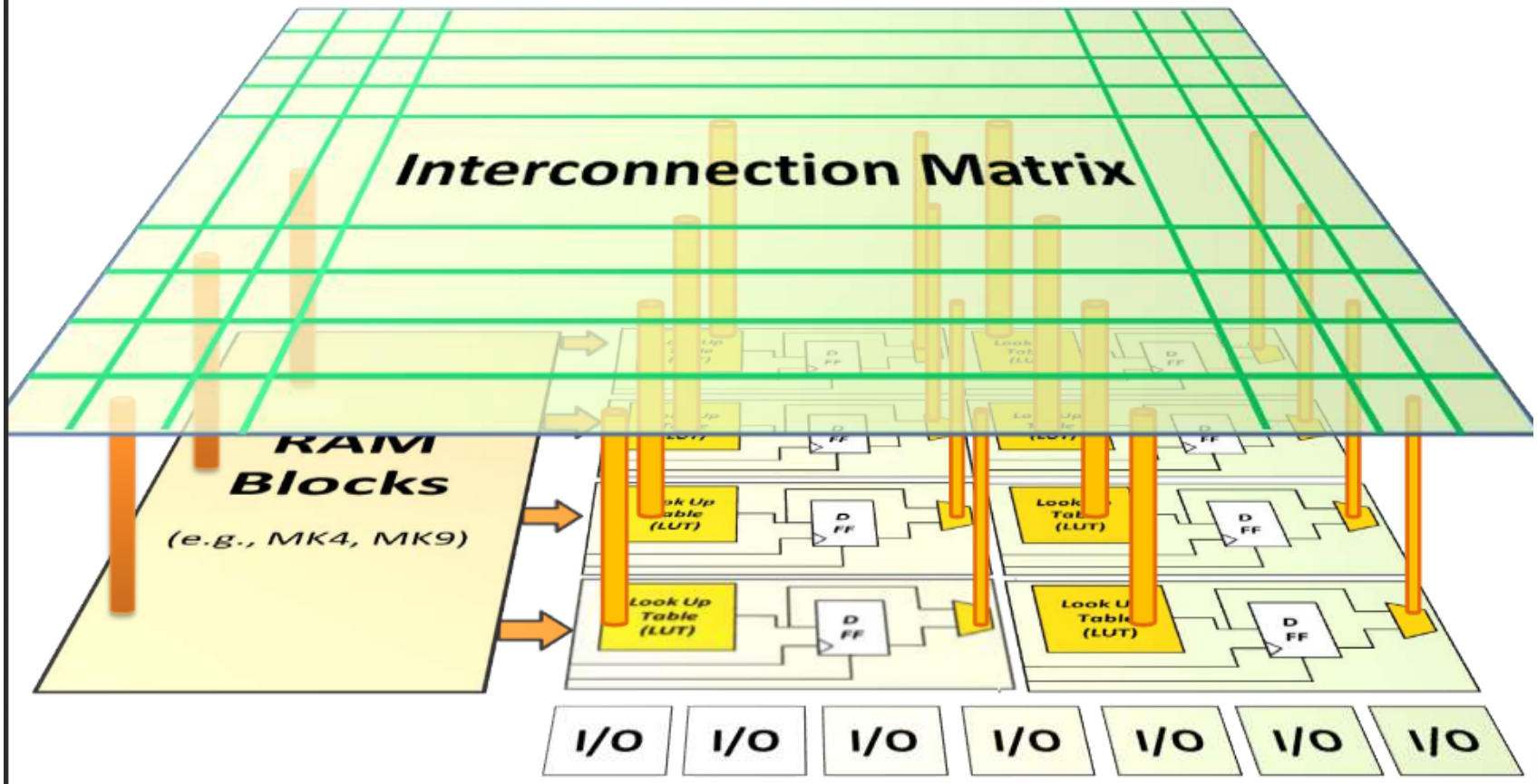
- **Una look-up table:** che consente di mappare una qualsiasi funzione combinatoria 4 ingressi 1 uscita
- **Un FF di tipo D** (con set e clr asincroni)
- **Un mux 2 -> 1:** per bypassare il FF in caso di celle puramente combinatorie

Elemento programmabile

FPGA: programmazione

ALMA MATER STUDIORUM – UNIVERSITY OF BOLOGNA
DEIS - DEPARTMENT OF ELECTRONICS, COMPUTER ENGINEERING AND SYSTEMS

Architettura di un FPGA



INAF
ISTITUTO NAZIONALE DI ASTROFISICA
OSSERVATORIO ASTRONOMICO DI TRIESTE