Real-time Pulsar Pipeline

Alessio Sclocco

2nd eAstronomy Meeting, Netherlands eScience Center

April 29, 2013

Introduction

The goal of this project is to design and build a pipeline for real-time pulsar searching

We want a parallel pipeline that uses many-core as accelerators

Layout of the presentation:

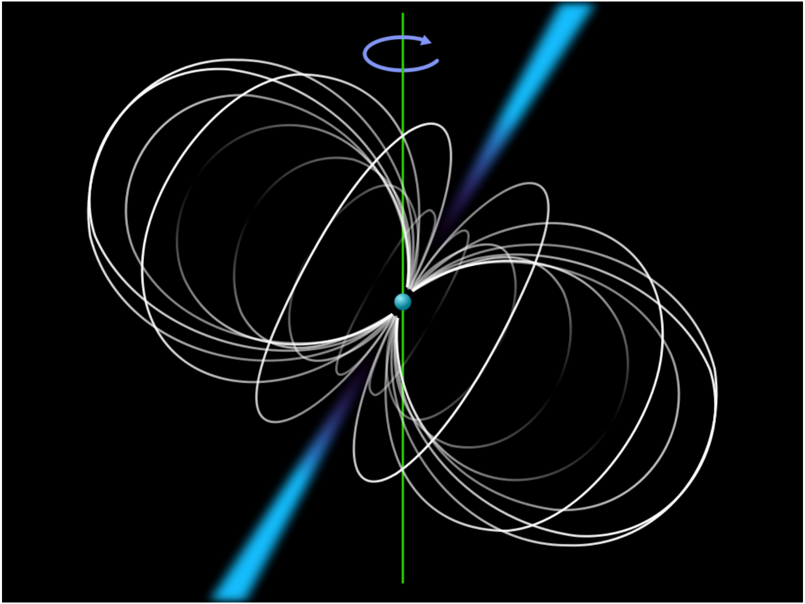
1. The Project
2. Current Status
3. Future Plans

The Project

Pulsars (1/3)

- A pulsar is a highly magnetized, rapidly rotating neutron star
- If its magnetic axis lies on Earth's line of sight we perceive a pulse
- The effect is similar to a lighthouse

Pulsars (2/3)



Pulsars (3/3)

Why are pulsars so important for scientists ?

- They are high resolution clocks in space
- A precise clock and gravity: test for general relativity
- Understand the equation of state of neutron stars
- Measure the effect of gravitational waves
- Measure distances in space
- Probe the interstellar medium

Pulsars are labs with conditions impossible to recreate on Earth

Being so important, scientists would like to know many of them

And use them for their experiments

Challenges (1/2)

Being so important, scientists would like to know many of them

And use them for their experiments

But, 46 years after the discovery of the first pulsar, we only know two thousand

Challenges (1/2)

Being so important, scientists would like to know many of them

And use them for their experiments

But, 46 years after the discovery of the first pulsar, we only know two thousand

Why ?

Challenges (2/2)

Because there is plenty of challenges in the search for pulsars

- Radio frequency interference
- Dispersion, scintillation, scattering
- Binary systems
- Fast periods
- Faint signals

Challenges (2/2)

Because there is plenty of challenges in the search for pulsars

- Radio frequency interference
- Dispersion, scintillation, scattering
- Binary systems
- Fast periods
- Faint signals

To increase the rate of pulsars discover we propose to use
many-core accelerators

Current Status

Algorithms

Three are the algorithms in our pipeline:

1. Dedispersion
2. Folding
3. Signal-to-noise Computation

The algorithms are implemented with OpenCL

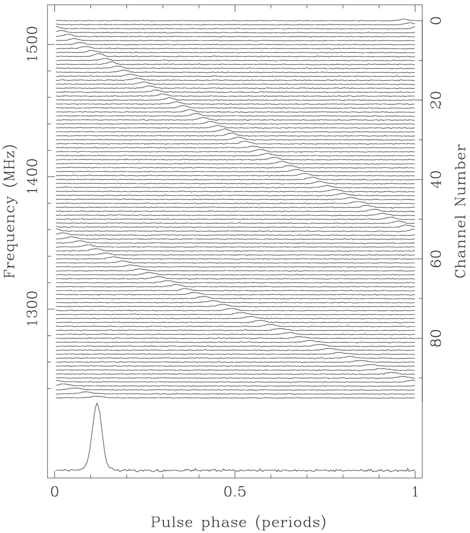
Performance portability is achieved with auto-tuning

Memory access is the main bottleneck

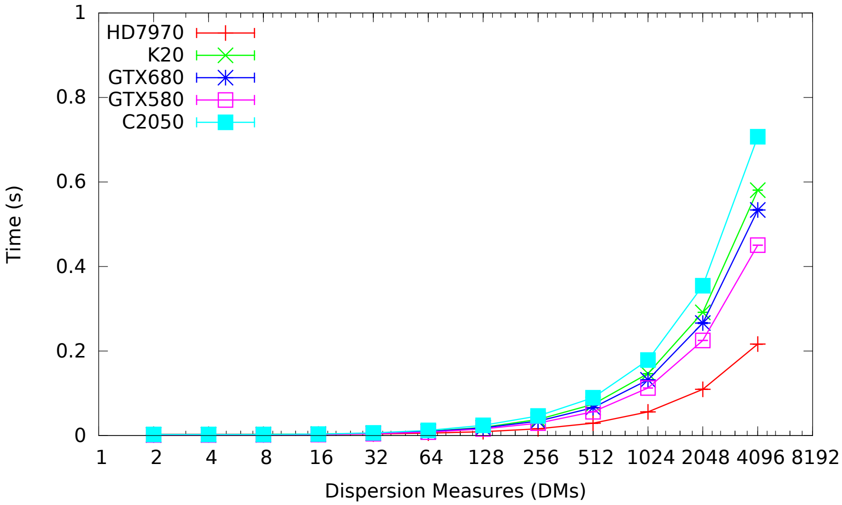
Observation Setup

- Sampling rate: 20,000 samples per second
- Frequency interval: 1,425 – 1,724 MHz
- Frequency channels: 1,024
- DM interval: 2 – 4,096
- DM step: 0.25
- Period interval: 32 – 1,024
- Period step: 32
- Period bins: 32

Dedispersion (1/3)

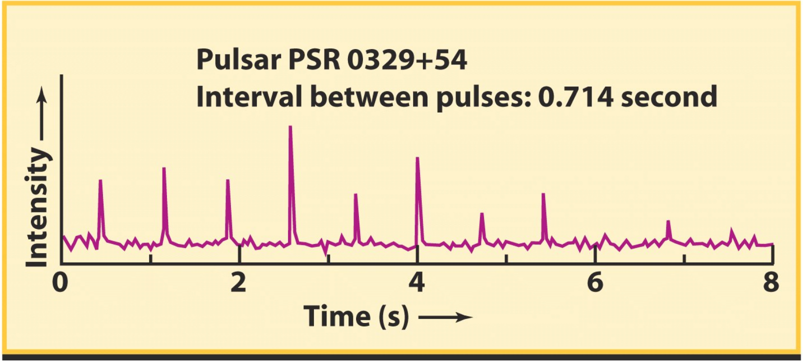


Dedispersion (2/3)

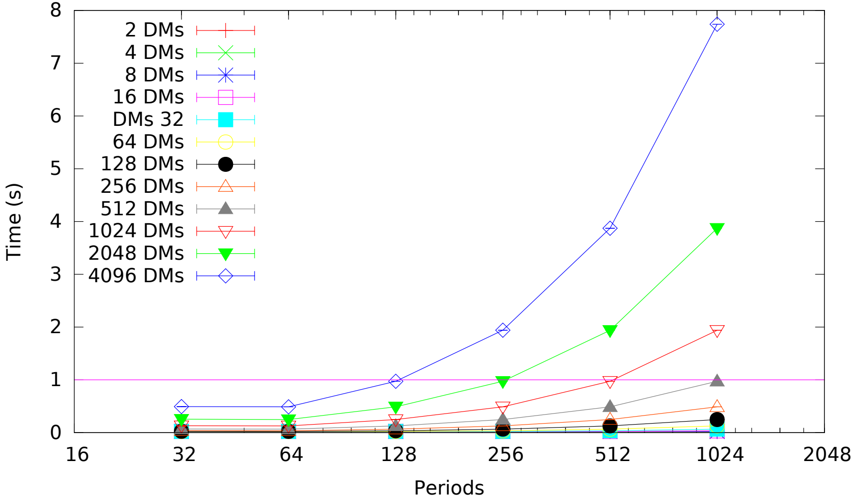


Dedispersion (3/3)

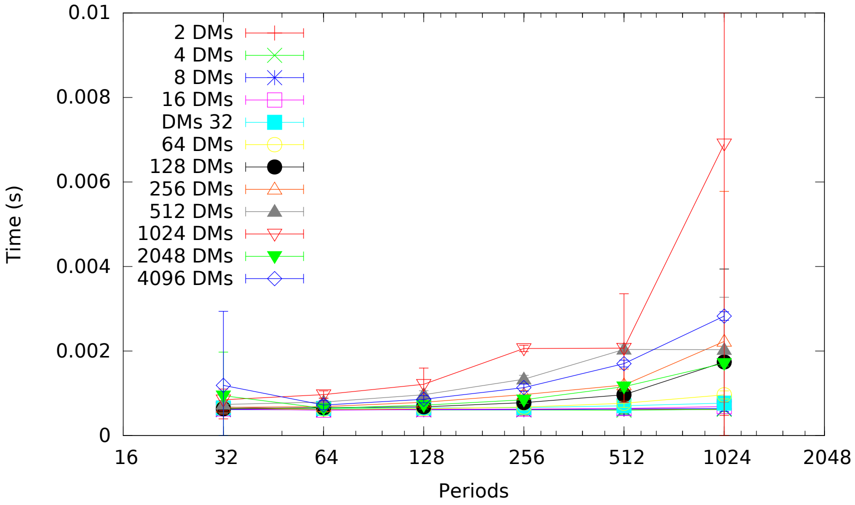
Platform	DM/s
AMD HD7970	18,922
NVIDIA GTX580	9,107
NVIDIA GTX680	7,775
NVIDIA C2050	5,792
NVIDIA K20	7,052


Pulsar PSR 0329+54
Interval between pulses: 0.714 second
Intensity
0
2
4
6
8
Time (s)

Folding (2/2)


8
7
6
5
4
3
2
1
0
Time (s)
2 DMs
4 DMs
8 DMs
16 DMs
DMs 32
64 DMs
128 DMs
256 DMs
512 DMs
1024 DMs
2048 DMs
4096 DMs
16
32
64
128
256
512
1024
2048
Periods

SNR Computation



Conclusions

In the first year of the eAstronomy project we

1. Analyzed the problem
2. Designed the fundamental many-core algorithms
3. Implemented all the components with OpenCL
4. Measured their performance

Future Plans

Pipeline Integration

This year we are going to build the pipeline out of the components we have

Challenge: how to auto-tune a pipeline ?

How will it compare with current state of the art software ?

Distributed Application

A single node will never be enough for real science

This year we are going to scale the pipeline out of a single node

Where: Cluster ? Grid ? Cloud ?