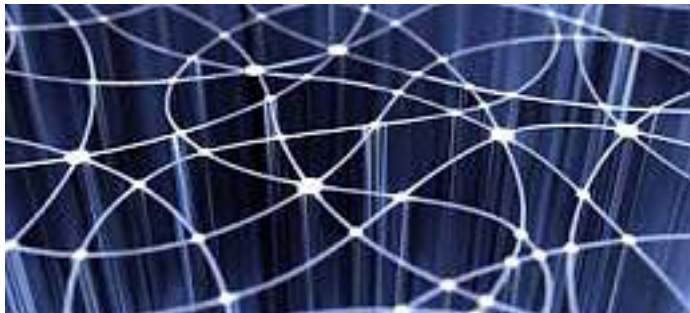# High-performance Signal and Data Processing: Challenges in Astro- and Particle Physics and Radio Astronomy Instrumentation

Monday 27 January 2014 - Friday 31 January 2014

University of the Witwatersrand

Book of Abstracts

Contents

0

An intuitive Parallel Programming Tool-flow for Software Defined Radio Signal Processing on Heterogeneous Architectures

Author: Lerato Mohapi[1]

Co-authors: Amit Mishra [1]; Michael Inggs [1]; Simon Winberg [1]

[1] *University of Cape Town*

We present a domain specific language (DSL) for software defined radio (SDR), referred to as SDR-DSL. The main objective of SDR-DSL is to provide an intuitive parallel programming tool-flow (PPTF) for digital signal processing (DSP) algorithms, which combines SDR and parallel programming domains to implement a PPTF that matches high level signal processing abstraction to generic parallel executable patterns. As languages that contain constructs for specific problem spaces, DSLs generally provide substantial gains in expressiveness and ease of use for even complex heterogeneous architectures. We demonstrate our approach in tackling parallel DSP programming complexity issues using an embedded DSL compiler with code generation kernels for multiple heterogeneous targets. We also discuss how a high-level DSP algorithms implementation using SDR-DSL can be validated prior to hardware implementation using software design cycles and functional verification tools.

Summary:

SDR-DSL is intended for intuitive yet scalable design and assembly of parallel DSP routines for SDR. Through its data types and execution semantics, SDR-DSL is expected to reflect the desired abstraction hierarchy in SDR. The SDR-DSL compiler, based on Delite DSL framework, provides a constraint system which is used to drive creation of new DSP datapaths for SDR easily. SDR-DSL accomplishes these while being part of a larger tool-chain which supports validation, code generations for heterogeneous architectures, and physical device assembly.

1

Noise Sensitivity of a VHF Broadband Interferometers

Author: Chih-Fong (Brady) Wen[1]

Co-author: Andrew Collier [1]

[1] *University of KwaZulu Natal*

Corresponding Author: wenbrado@gmail.com

A VHF interferometer can be used to measure the three-dimensional source of radiation emitted by lightning discharges. This is achieved by analysing the phase delay between the signals recorded at each of the three antennas.

Using a numerical model of an interferometer coded in R we simulated a simple interferometer, including the source, antennas and a data processing unit. First, a monochromatic and isotropic point source was simulated in order to validate the structure of the model. The model was then expanded to include multiple monochromatic signals and finally a truly broadband signal.

Using the model we were able to simulate the effects of noise on the resolving power of the interferometer and determine under what conditions the observations become unreliable.

2

Application of Data Intensive Science to the Electrical Energy Grid.

Authors: Albert Dove[1]; Ivan Hofsajer[1]

[1] *University of Witwatersrand*

Corresponding Author: ivan.hofsajer@wits.ac.za

The fields of radio astronomy and high energy partial physics are characterized by very large amounts of data. This can be seen in the CERN ATLAS experiment and the SKA radio telescope which need to gather and process very large data sets. The nature of these data sets is that they:

A. Are big
B. Stream, in the sense that they continually produce data over long periods of time.
C. The data source is localized
D. Some form of preprocessing is performed at the source to reduce the data in need of processing.
E. The detail data analysis is not required in real time.

This type of experiment with a large throughput has been rigorously analyzed from both the front end high throughout electronics to the data aggregating and processing or storage. It is proposed here that this type of experimental system may be used to inform the design of a very different application, but with some common elements.

As energy is fast becoming a scarce resource, it is important to be able to carefully control its use. This is required both for efficiency as well as for maximizing the utilization of existing infrastructure. All forms of energy are under pressure, but it is electrical energy in particular that will be the focus of this investigation.

Currently the electrical energy generation and distribution grid is characterized by a control system that has relatively few measurement points and a centralized control center. This has work very well in the past, but as the grid is coming under increasing pressure, it is important to be able to achieve better control, which in turn needs better and finer measurement. This is especially important within the South African context, where nationally the electricity supply grid runs with almost no reserve margin.

Many initiatives abound concerning the so-called smart grid. With the term meaning different things to different people. However it is generally considered to have a distributed network of sensors. With the additional data, "smarter" control and decisions can be made.

It is proposed that every single consumer of electrical energy in South Africa be metered on a sub 50Hz cycle time base. This could produce a complete picture of energy consumption in a real time manner, producing a data stream of about 10G-bytes per second. This is a substantial amount of data and possibly the techniques developed in the data intensive science experiments could be successfully used here. The energy metering project could be characterized as:
A. Big
B. Streams continuously
C. Data source is distributed.
D. Preprocessing is required to alleviate the amount of data.
E. Some analysis is required in real time, but not all.

3

Evaluation of High-Level Open-Source Tool-Flows for Rapid Prototyping of SDR Applications on Heterogeneous Platforms

Author: Khobatha Setetemela[1]

Co-authors: Michael Inggs [2]; Simon Winberg [2]

[1] *University of Cape Town*

[2] *University of Cape Town*

Corresponding Author: sttori001@myuct.ac.za

Traditional radios have a fixed hardware-defined functionality and typically support only one radio standard. Though resulting in simple, easy-to-optimise designs, the fixed architecture is not sustainable given the rapidly changing landscape of radio protocols and the increasing demand for seamless access. Software-defined radios (SDRs) were coined two decades ago to meet these and other demands.A SDR is a flexible radio architecture where some or all of the physical layer functions that are traditionally realised using fixed hardware processors are implemented through user-modifiable software that runs on programmable chips. These chips include GPPs, DSPs and FPGAs. Although SDRs are generally applauded for their attractive potential, they remain a niche technology. It is virtually agreed that since SDR is an inherently complex domain, the high skills set required to build efficient SDRs is a prime reason why many communication systems designers are reluctant to adopt the technology. Currently, conventional SDR design methodologies such as the popular HDL-based approaches for FPGA firmware development do not provide mechanisms to effectively abstract the software and hardware complexity of designs for the average designer. Rather, these tools tend to force developers to work at a low-level of coding, often building up solutions from first principles, which can be a tedious and error prone process that may yield inefficient designs and poor productivity levels. There is therefore a need for easy-to-use and efficient high-level tools to simplify and accelerate SDR applications design. While there is high and varied activity in literature and industry to address this need, very little work has been done to survey the various existing approaches and evaluate their effectiveness. Where this has been done, focus has been limited to homogeneous target architectures and few open-source tools were considered. In this paper, we comprehensively evaluate potential open-source high-level tool-flows (Delite DSL framework, Migen, MyHDL and Ptolemy) for rapid prototyping of SDR applications on heterogeneous target platforms. The tools are evaluated against an ideal high-level SDR flow that we designed; particular strengths and lacks of each tool are discussed.

4

ARM processors as a low cost alternative tool for computation of FFTs for radio astronomy

Author: Mitchell Cox[1]

[1] *University of the Witwatersrand*

Corresponding Author: mitch@enox.co.za

The Fast Fourier Transform (FFT) has many uses in science and in particular, radio astronomy, by the F-Engine of the correlator. This operation must be done for all polarisations of all antennas and is highly parallel. A possible alternative to the use of expensive GPUs and FPGAs is a cluster of ARM processors that can perform FFTs in parallel, cost effectively and with low power consumption. ARM Cortex-A7 and Cortex-A9 CPUs are benchmarked using FFTW, a high-performance, open-source FFT library. Single- and multi-thread as well as multi-processor tests are done. The results are used to characterise the theoretical processor throughputs in Bytes/s for one-dimensional complex FFTs of various sizes. It is found that a single 1 GHz quad-core Cortex-A9 processor is able to process a 32768 point one-dimensional complex FFT at up to 250 MB/s with a power consumption of approximately 5 W.

5

The Data Pipeline of the AGILE Space Telescope

Author: Andrew Chen[1]

[1] *University of the Witwatersrand*

Corresponding Author: andrew.chen@wits.ac.za

AGILE is an Italian Space Agency mission dedicated to observing the gamma-ray Universe, combining a gamma-ray imager (sensitive in the energy range 30 MeV–50 GeV), a hard X-ray imager (sensitive in the range 18–60 keV), a calorimeter (sensitive in the range 350 keV–100 MeV), and an anticoincidence system. AGILE was successfully launched on 2007 April 23 and continues to observe the gamma-ray sky today. Here we present the data pipeline of AGILE, from the silicon strip detectors, electronics, and on-board triggers of the instrument, through the telemetry, transmission to the Data Center, processing, analysis, and archiving, to the provision of data, software, and catalogs to the instrument team, guest observers, and the general public.

6

Data Processing for ALICE at the LHC

Author: Tom Dietel[1]

[1] *University of Cape Town*

Corresponding Author: thomas.dietel@uct.ac.za

The ALICE collaboration studies the Quark-Gluon Plasma, a deconfined state of strongly interacting matter at extreme temperatures and densities, which is created in high-energy collisions of heavy nuclei at the Large Hadron Collider (LHC) at CERN.

The ALICE apparatus was designed to inspect Pb-Pb collisions at an interaction rate of 8000 Hz, to read out up to 1000 events per second, and to record data at a rate of more than 1 GB/s. ALICE also analyzes p-p and p-Pb collisions at interaction rates up to 200 kHz, with similar limits on read-out and data rates as for Pb-Pb collisions. ALICE employs a multi-level hardware trigger system to select interactions for read-out and a High-Level Trigger to perform online reconstruction, data compression and further event selection.

In parallel to the luminosity upgrade of the LHC during the second long shutdown in 2018/19, a major uppgrade to ALICE is planned, including replacements for the inner silicon tracker and the read-out chambers of the time-projection chamber. These detectors will feature continuous read-out, pushing the data rate above 1 TB/s. A new computing system will reduce the data rate to tape while keeping the
full event sample of up to 50 kHz Pb-Pb collisions, using online reconstruction, storage of partially reconstructed data and compression.

I will present the main physics goals and data taking strategy for the previous and upcoming runs until 2017. I will then present the upgrade of the ALICE experiment, pointing out how the changing physics goals influence the technology selection for the upgrade. I will focus on the design of the new computing system to process a data stream of more than 1 TB/s.

7

Prometeo, the next generation test-bench for the front-end electronics of the ATLAS Tile Calorimeter upgrade phase-II

Author: XIFENG RUAN[1]

[1] *WITS*

Prometeo is a portable test-bench for the full certification of the front-end electronics of the ATLAS Tile Calorimeter upgrade phase-II. It is a high throughput electronics system designed to simultaneously read-out all the samples from 12 channels at the LHC bunch crossing frequency. The core of

the system is a Xilinx VC707 evaluation board extended with a dual QSFP FMC module to read-out and control the front-end boards. The rest of the functionalities of the system are provided by a HV mezzanine board that to turn on the gain of the photo-multipliers, an LED board that sends light to illuminate the them, and a 12 channel ADC board that samples the analog output of the front-end. The system is connected by ethernet to a GUI client from which QA tests are performed on the electronics such as noise measurements and linearity response to an injected charge.

8

The RHINO Digital Processing Skills Development Initiative: An integrated review of the platform, resources and training structures

Author: Simon Winberg[1]

[1] *University of Cape Town*

Corresponding Author: simon.winberg@uct.ac.za

This paper presents an integrated perspective, and progress review, of the Reconfigurable Hardware Interface for computiNg and radiO (RHINO) project[1] and the skills development programme and related development resources for this initiative. The RHINO platform is designed around providing a comparatively low cost FPGA-based reconfigurable computing platform suited for a variety of Software Defined Radio (SDR) and Radio Astronomy (RA) back-end processing applications. RHINO is planned to provide a level of compatibility with the more powerful Reconfigurable Open Architecture Computing Hardware (ROACH) platform, and is intended to accommodate a trajectory for novice developers that decide to delve more deeply into RA processing to transition to ROACH and other high-end FPGA-based platforms.

The work carried out on the RHINO HDL skills development initiative is separated into four main interdependent aspects: 1) development of firmware and software, and hardware revisions, of the RHINO platform and its system software; 2) establishing an effective tool flow (integration and configuration of existing tools) and programming solutions that are appropriate for both training and development of applications; and 3) providing compatibility with existing frameworks such as the CASPER MSSGE Simulink toolflow[2] - tools and GNU Radio[3], and 4) example applications, performance testing, and preparation of training materials including support for migration to, and integration with, other platforms. A holistic systems view of these components is explained, and briefly mentions the various projects that fit into this higher-level initiative.

The paper proceeds to report on the progress achieved, challenges encountered and proposed solutions to these. We hope that this presentation will inspire comment and feedback to us in terms of our design choices, and we invite requests or suggestions for training support resources, either at the undergraduate or postgraduate level, which would be useful to the broader field of reconfigurable computing data processing.

REFERENCES

[1] S. Winberg, A. Langman, and S. Scott, "The RHINO platform: charging towards innovation and skills development in software defined radio," in South African Institute for Computer Scientists and Information Technologists, 2011, pp. 334–337.
[2] C. Chang, "Design and applications of a reconfigurable computing system for high performance digital signal processing," University of California, Berkeley, 2005.
[3] E. Blossom, "GNU radio: tools for exploring the radio frequency spectrum," Linux journal, vol. 2004, no. 122, p. 4, 2004.

9

The national SA-CERN programme

Author: Jean Cleymans[1]

[1] *University of Cape Town*

Corresponding Author: jean.cleymans@uct.ac.za

The national SA-CERN programme was launched in December 2008. The talk will
review the focus areas of the programme and its achievements thus far
and the future developments.

10

Performance Analysis of Virtualization for High Performance Computing

Author: Matthew Cawood[1]

[1] *UCT*

Corresponding Author: matthewcawood@gmail.com

The field of High Performance Computing (HPC) is growing rapidly to meet the need of solving Big
Data problems. As HPC systems grow more powerful and complex, administrative and utilization
challenges are introduced. The Cloud computing paradigm offers promising Virtualization tech-
nologies for managing large HPC systems and optimizing their usage. However there are still major
obstacles which need to be addressed in order to efficiently run HPC applications within a Cloud
system.

One of the main obstacles of implementing Virtualized computer environments in HPC clusters is
the impact they have on performance. HPC clusters rely heavily on computational throughput and
memory bandwidth, as well as on high performance networks, such as Infiniband, to provide com-
munications between interconnected hardware. Virtualization has a negative performance impact
on all of these factors, with network performance being particularly affected by such technologies.
With the recent release of Single Root I/O Virtualization (SR-IOV) for Infiniband;
virtual machines (VMs) can directly and more efficiently access the network.

This paper presents results of an in-depth performance evaluation of the KVM hypervisor deployed
within an HPC cluster environment. The HPC Challenge benchmark was used to assess Virtualiza-
tion impact on various aspects of cluster performance. Focus was placed on establishing a good
baseline performance, then on comparing virtual machine performance in a number of tests. An
evaluation was also done on other relevant topics such VM to CPU mapping policies and Gigabit
Ethernet versus Infiniband performance, as well as the impacts of hyper threading and software
optimization.

Key Words: Virtualisation, HPC, Cloud Computing, Infiniband, SR-IOV

11

The ATLAS readout electronics and timing

Author: Alberto Valero[1]

[1] *Instituto de Física Corpuscular (Universidad de Valencia-CSIC)*

Corresponding Author: jvalero@cern.ch

The ATLAS detector has been designed to study the proton-proton collisions produced by the Large Hadron Collider (LHC) at CERN. The trigger based data acquisition system of ATLAS is composed by a first level with detector specific hardware followed by a common software trigger. The events selected by the first level of trigger are transmitted to the off-detector electronics where the data is reconstructed and transferred to the common High Level Trigger system. A series of upgrades are scheduled for the next ten years to increase the LHC instantaneous luminosity.The overall readout architecture and the trigger structure is being revised to cope with the new LHC parameters. This presentation summarizes the readout electronics of the different ATLAS sub-detectors in the present system and the evolution for the ATLAS upgrade. It will include a detailed description of the custom front-end and back-end electronics, the signal reconstruction algorithms and the different timing procedures used to synchronize the acquired data.

12

LED Board for the Tile Calorimeter of the ATLAS Detector at the LHC

Authors: Reto Suter[1]; Titus Masike[1]

[1] *University of the Witwatersrand*

Corresponding Author: tmasike27@yahoo.com

Reviewing the architecture of the MobiDick4 system, this is used in the development of a test bench which will be used for the readout and control electronics of the Tile Calorimeter. Emphasis on the LED Board

13

A Scalable Heterogeneous Architecture for Software Defined Radio on the RHINO platform

Author: Matthew Bridges[1]

[1] *University of Cape Town*

Corresponding Author: matthewbridges88@gmail.com

Since the adoption of Software Defined Radio (SDR) into the fields of Wireless Communications, RADAR and Radio Astronomy there has been an ever-growing need for high performance computer systems optimised for large bandwidth Digital Signal Processing. Some of the key factors which affect the performance throughput of these systems are:
1. The number of processing elements
2. Processing frequency
3. Memory access latency
4. External Interfacing (Ease of and data throughput)
5. Ease of programming

FPGAs perform very well as a SDR platform due to their reconfigurable and highly parallel nature as well as simple, high-throughput external interfacing. However, FPGAs are renowned for being very difficult platforms to program and the designs created for them are often fixed after compile time and are not interactive. At the other end of the scale microprocessor are integrated circuits designed to perform general purpose computation and are therefore very versatile and interactive. However, their fixed data path and limited parallelism makes them ineffective for many applications including SDR.

The Reconfigurable Hardware Interface for computiNg and radiO (RHINO) Board is an open-source platform developed at the University of Cape Town for the purpose of SDR and computation. It's main processing element is a large Xilinx Spartan 6 FPGA which is coupled to an ARM microcontroller from Texas Instruments. The board also has high speed networking and Input/Output interfaces as well as a large amount of memory.

The aim of this research was to build a framework for simplifying the development of interactive and parameterized logic cores. On the RHINO, these logic cores are then memory mapped to the processor creating a heterogeneous system. An architecture capable of performing the tasks of an SDR Transceiver was then developed from logic cores created using this framework and serves as the starting point for SDR on the RHINO allowing more complicated systems to be built on top.

Although this project was focused on Software Defined Radio, the framework was designed for general computation and is therefore applicable to other fields which require complicated computation.

14

Optimising LInux Operating System on Arm Architecture

Author: Jonathan Padavatan[1]

[1] *iThemba Labs*

Corresponding Author: padavatanj@gmail.com

Optimising the linux environment for higher performance usually takes place once a working system is running to specification and unforeseen bottlenecks or bugs occur. However, once a system is up and running, it is frequently difficult to differentiate between application specific problems and operating system problems. Ideally, tuning any given system should start early in the development stage. This investigation into tuning the Linux Linaro 13.05 release based on Linaro Stable Kernel (LSK) preview 3.9.4-2013.0 running on ARM architecture is focused on the I/O intensive environment of high throughput computing. Four areas under consideration are CPU usage, memory, disk I/O and GPU usage. The choice of parameters best suited for each area will differ as the default system tools under the /proc directory offer wide variety of process specific tools for profiling and tweaking the system, and choosing the correct tool will be critical. Benchmark tools for an I/O process for use are Oprofile, Nice, VMStat and IOSTAT. Using Oprofile to analyze and evaluate the basic configuration system, a simple disk copying test will identify relevant processes and parameters to provide a solid baseline. Since optimizing the operating system for an I/O intensive application, it is noted that poor performance will be observed for different workload characteristics; hence a change management process will be required

15

Observing to the very edge of a black hole using wideband signal processing

Author: Jonathan Weintroub[1]

[1] *Harvard-Smithsonian Center for Astrophysics*

Corresponding Author: jweintroub@cfa.harvard.edu

A broad international collaboration is building the Event Horizon Telescope (EHT). The aim is to test Einstein's theory of General Relativity in one of the very few places it could break down: the strong gravity regime right at the edge of a black hole. The EHT is an earth-size VLBI array operating at the shortest radio wavelengths, that has achieved unprecedented angular resolution of a few tens

of micro-arcseconds. For nearby super massive black holes (SMBH) this size scale is comparable to the Schwarzschild Radius, and emission in the immediate neighborhood of the event horizon can be directly observed.

In 2007 the EHT detected SgrA*, the SMBH at the Milky Way's center, on event-horizon scales. Three VLBI stations were located on Mauna Kea, Hawaii, Cedar Flat, California, and Mount Graham, Arizona. Subsequent observations have used larger collecting areas, greater bandwidth, and additional stations forming more and longer baselines. These developments have improved sensitivity, resolution and imaging power. The EHT has detected the base of the jet powered by the SMBH in the galaxy Messier 87 (Virgo A), and, most recently, dual polarization VLBI has probed magnetic fields on event-horizon scales.

Digital signal processing (DSP) technology from the Collaboration for Astronomy Signal Processing and Electronics Research (CASPER) have been key to the success of the EHT throughout. Single dish digital back ends, and multi-dish phased arrays have used successive generations of CASPER platforms, including iBOB, BEE2, ROACH1 and ROACH2. I will summarize the ground breaking science results obtained with the CASPER-enabled EHT, and outline future technical developments, with emphasis on the secret sauce of CASPER DSP.

for the Event Horizon Telescope Collaboration
http://eventhorizontelescope.org/

16

Major Unresolved Matters about Blazars and Active Galactic Nuclei

Author: PROSPERY C. SIMPEMBA[1]

Co-authors: Chinyama Kaumba [2]; Sergio Colafrancesco [3]

[1] School of Physics, University of Witwatersrand, Private Bag 3, Wits, 2050, South Africa

[2] The Copperbelt University, School of Mathematics and Natural Sciences, Department of Physics, P.O. Box 21692, 10101 Kitwe, Zambia

[3] School of Physics, University of Witwatersrand, Private Bag 3, Wits, 2050, South Africa.

Corresponding Author: pcs200800@gmail.com

Studies about active galactic nuclei (AGN) and radio jets of galaxies and blazars focus on seeking for the more acceptable physics, explaining energy jet formation, speeding up and collimation of these outflows. We review several articles on this subject to establish the unresolved matters about blazars and AGN in the already conducted studies. The challenge about the theories explaining the origin of ultra high energy cosmic rays (UHECR) is still there for the current day astronomers. The high energy generation and emission in the core of AGN depicts an accelerator that nature presents to us which we all seek to understand and cherish.

Key words: Galaxies, blazars, jets, cosmic rays

17

ATLAS GPU Trigger Studies

Author: Timothy Bristow[1]

[1] University of Edinburgh

Corresponding Author: timothy.michael.bristow@cern.ch

The ATLAS trigger system is required to filter collisions in the detector and reduce the millions of collisions per second down to a few hundred events
which are stored and analysed further. The trigger system is split into a number of subsystems which run at different levels of abstraction. The High Level
Trigger is required to decode the bytestream from the detector into space points and run clustering algorithms on these results. This must be done
efficiently and quickly. This problem lends itself well to a parallel processing solution as all of the data points are independent. NVIDIA Tesla GPUs with thousands of cores, multi-co\
re processors
and co-processors, such as the Intel Xeon Phi, have been investigated as possible tools to enhance the ATLAS High Level Trigger. This allows for the main CPU to distribute
large, computationally intensive problems to the GPUs, or co-processors. ARM processors are being investigated for high throughput computing with low energy consumption, however, they
have limited processing power. Many of the latest ARM/ mobile processors have built in GPUs which offer substantial computing power. This would allow for
a similar construct that is used in the ATLAS High Level Trigger GPU study. The OpenCL programming langauge allows for multi-core code to be ported between
devices fairly easily, which would allow for benchmarks to be run to compare the performance of the ARM processors with the built in GPUs with the larger, more powerful NVIDIA GPUs.

18

An ATCA framework for the ATLAS Front to Back End Electronics for the Phase II Upgrade at the LHC

Author: Robert Reed[1]

[1] *University of the Witwatersrand*

The Large Hadron Collider at CERN is scheduled to undergo another major
upgrade in what is called phase II in the year 2022. During this upgrade the ATLAS team will do major modifications to the detector to account for the increased luminosity factor of ten. Almost the entire read out electronics, situated on the front end, will be relocated to the back end as well as upgraded. A radically new system will be required to house, manage and connect this new hardware. The proposed solution will be an Advanced Telecommunication Computing Architecture or ATCA which will not only house but also allow advanced management features and control at a hardware level through the Intelligent Platform Management Interface. The details and current setup of the ATCA and how it will be part of the TileCal upgrade demonstrator program will be presented in full.

Summary:

Keywords: LHC, CERN, ATLAS, ATCA, TileCal, LHC Phase II Upgrade

19

CPU Benchmarking performance for ARM Processors

Authors: Gerhard Harmsen[1]; Robert Reed[1]; Thomas Wrigley[1]

Co-authors: Jonathan Padavatan [1]; Mitchell Cox [1]

[1] *University of the Witwatersrand*

Corresponding Author: thomas.wrigley@cern.ch

The Large Hadron Collider (LHC) at CERN is currently undergoing a major
upgrade to handle higher energies. This will be the first of two upgrades and the expected amount
of data produced by this upgraded system will far exceed current data throughput capabilities. It is
expected that the same will be so for the Square Kilometre Array (SKA) Radio Telescope. A poten-
tial alternative to current high performance computing systems involves using low-cost, low-power
ARM processors in large arrays to provide massive parallelisation and hence large data throughput.
The central advantage in using ARM processors is found in the central processing unit (CPU). As
such, a thorough evaluation and benchmarking of the CPUs of three different models of ARM pro-
cessor, namely the Cortex-A7, Cortex-A9,and Cortex-A15, has been prepared. Results have been
obtained for single and multiple (cluster-configuration)
processors and an attempt has been made to compare benchmark performance in standardised tests
such as High Performance Linpack (HPL) to "real-world"
performance applications.

Summary:

Keywords: ARM, benchmark, CPU, high-throughput computing, high-performance computing, LHC,
ATLAS, SKA, CERN.

20

RAM Benchmarking performance for ARM Processors

Authors: Gerhard Harmsen[1]; Robert Reed[1]; Thomas Wrigley[1]

[1] *University of the Witwatersrand*

Corresponding Author: thomas.wrigley@cern.ch

The Large Hadron Collider (LHC) at CERN is currently undergoing a major
upgrade to handle higher energies. This will be the first of two upgrades and the expected amount
of data produced by this upgraded system will far exceed current data throughput capabilities. It is
expected that the same will be so for the Square Kilometre Array (SKA) Radio Telescope. A potential
alternative to current high performance computing systems involves using low-cost, low-power
ARM processors in large arrays to provide massive parallelisation and hence large data throughput.
As memory performance will play an important role in high-throughput computing, several tests and
applications such as the STREAM benchmark have been used to thoroughly evaluate and benchmark
the memory (RAM) performance of three different models of ARM processor, namely the Cortex-A7,
Cortex-A9, and Cortex-A15. Various aspects of memory performance have been evaluated, including
the effects of using multiple processors in a cluster-configuration.

Summary:

Keywords: ARM, memory, RAM, benchmark, high-throughput computing, high-performance
computing, LHC, ATLAS, SKA, CERN.

21

Imaging the radio sky

Author: Andreas Faltenbaher[1]

[1] *University of the Witwatersrand*

Our knowledge of the Universe is almost entirely based on electromagnetic
waves arriving from distant sources, such as stars, galaxies, quasars, etc.
Most of the electromagnetic radiation is shielded by the atmosphere of the
earth. Only visible light ($\lambda \approx 10^{-7}$m) and radio waves
($\lambda \approx 1$m) arrive at sea level. This is why only optical and
radio telescopes are used for ground-based observations, other wavebands
require satellite missions. In my presentation I will review basic differences
between optical and radio imaging and discuss the resulting data processing

22

The long journey to the Higgs boson and beyond at the LHC

Author: Peter Jenni[1]

[1] *CERN*

Since three years the experiments at the Large Hadron Collider (LHC) investigate particle physics at
the highest collision energies ever achieved in a laboratory. Following a rich harvest of results for
Standard Model (SM) physics came in 2012 the first spectacular discovery, by the ATLAS and CMS
experiments observing a new, heavy particle which is most likely the long-awaited Higgs boson.
The latest results with the full data sets accumulated over the first three-year running period of
the LHC will be presented, including recent refined measurements on Higgs properties. Other, far-
reaching results can be reported for exploratory new physics searches like Supersymmetry (SUSY),
Extra Dimensions, and the production of new heavy particles.

However, with this recent discovery of a heavy scalar boson the exciting journey into unexplored
physics territory, within and beyond the SM, has only just begun at the LHC. Besides the first results
and the future prospects, the talk will also touch on the motivation, history and the challenges of the
whole LHC project, as well as on the fruitful collaboration with the South African teams.

23

PERFORMANCE ANALYSIS OF VIRTUALIZATION FOR HIGH PERFORMANCE COMPUTING

Author: Matthew Cawood[1]

[1] *University of Cape Town*

The field of High Performance Computing (HPC) is growing rapidly to meet the need of solving Big
Data problems. As HPC systems grow more powerful and complex, administrative and utilization
challenges are introduced. The Cloud computing paradigm offers promising Virtualization tech-
nologies for managing large HPC systems and optimizing their usage. However there are still major
obstacles which need to be addressed in order to efficiently run HPC applications within a Cloud
system.

One of the main obstacles of implementing Virtualized computer environments in HPC clusters is
the impact they have on performance. HPC clusters rely heavily on computational throughput and
memory bandwidth, as well as on high performance networks, such as Infiniband, to provide com-
munications between interconnected hardware. Virtualization has a negative performance impact
on all of these factors, with network performance being particularly affected by such technologies.
With the recent release of Single Root I/O Virtualization (SR-IOV) for Infiniband; virtual machines
(VMs) can directly and more efficiently access the network.

This paper presents results of an in-depth performance evaluation of the KVM hypervisor deployed within an HPC cluster environment. The HPC Challenge benchmark was used to assess Virtualization impact on various aspects of cluster performance. Focus was placed on establishing a good baseline performance, then on comparing virtual machine performance in a number of tests. An evaluation was also done on other relevant topics such VM to CPU mapping policies and Gigabit Ethernet versus Infiniband performance, as well as the impacts of hyper threading and software optimization.

24

Mathematizing the problem of high-throughput computing

Author: Bruce Mellado[1]

Co-author: Vishnu Jejjala [1]

[1] *University of the Witwatersrand*

The introduction of ARM processors to high-throughput computing requires quantifying the output rate. Data-flow in high-throughput computing can be evaluated analytically provided a number of assumptions. Provided that data distribution to the RAM, or input rate, can be sustained at a rate higher than the output rate, a number of expressions can be derived. These formulae can be expressed in terms of a dimensionless quantity related to the RAM frequency and the CPU clock frequency. Features of these formulae will be discussed. The problem of data-flow in a more general application can be approximated to a problem of fluid dynamics. The prospects of developing the corresponding differential equations will also be discussed.

Plenary Overview Session I / 25

Welcome

Corresponding Author: john.carter@wits.ac.za

Plenary Overview Session I / 26

Workshop Overview

Corresponding Authors: bruce.mellado@wits.ac.za, mikings@gmail.com

Plenary Overview Session I / 27

Cyber-infrastructure in South Africa

Corresponding Author: daniel.adams@dst.gov.za

Plenary Overview Session I / 28

The SKA & MeerKAT

Plenary Overview Session I / 30

Welcome from the DST

Corresponding Author: thomas.aufderheyde@dst.gov.za

Plenary Overview Session II / 31

The MeerKat Correlator & Beamformer

Plenary Overview Session II / 32

The Long Journey to the Higgs Boson and Beyond at the LHC

Corresponding Author: peter.jenni@cern.ch

Plenary Overview Session II / 33

The national SA-CERN programme

Corresponding Author: jean.cleymans@uct.ac.za

Plenary Session I / 34

Observing to the Very Edge of a Black Hole Using Wideband Signal Processing

Corresponding Author: jweintroub@cfa.harvard.edu

Plenary session II / 38

Readout Electronics of the Alice Detector

Corresponding Author: thomas.dietel@uct.ac.za

Plenary Session III / 39

The upgrade of the ATLAS readout system

Corresponding Author: jvalero@cern.ch

Plenary Session III / 40

Upgrade of the ATLAS TileCal Electronics

Corresponding Author: carlos.solans@cern.ch

41

GRID in South Africa

Corresponding Author: brucellino@gmail.com

Plenary session II / 42

The Data Pipeline of the AGILE Space Telescope

Corresponding Author: andrew.chen@wits.ac.za

43

Likelihood Analysis of Higgs anomalous couplings

Author: Gilad Amar[1]

Co-author: Stefan von Buddenbrock [1]

[1] *University of Witwatersrand*

Corresponding Author: giladamar@gmail.com

An investigation into Higgs production, this study explores beyond Standard Model anomalous couplings in weak vector boson fusion. It is found that in the HWW vertex, effective strengths for anomalous couplings can be described by two constants lambda and lambda prime respectively. In an effort to discover how much data need be accrued by electron/positron colliders, to discriminate between the SM and such BSM physics the concept of likelihood from statistics is included. The study involves the Monte Carlo generation of millions of events, and hundreds of thousands of CPU intensive test statistics required to determine the sensitivity of model discrimination. In addition, this processor-heavy analysis makes for a useful benchmarking tool of Intel and ARM processors being considered in the WITS Massive Affordable Computing project.

Plenary Session I / 44

Imaging the Radio Sky

Corresponding Author: andreas.faltenbacher@wits.ac.za

45

Correlator/Beamformer

Corresponding Author: andrew@ska.ac.za

46

RFI Issues

Corresponding Author: jason.manley@ska.ac.za

48

The Computing Model of ATLAS

Corresponding Author: yacoob@ukzn.ac.za

Plenary Session IV / 49

Opti-NUM Solutions (Matlab): Distributed Computing

Plenary Session III / 50

FPGA Toolflows: CASPER and Beyond

Corresponding Author: wesley@ska.ac.za

51

Signal Processing Challenges

Corresponding Author: paul.prozesky@ska.ac.za

52

Massive Parallelism for Combinatorial Problems

Authors: Edward Steere[1]; Scott Hazelhurst[1]

[1] *University of the Witwatersrand*

Corresponding Author: edward.steere@gmail.com

Massive parallelism is a design paradigm in computer architecture which trades the complexity of sequential processing units for many simpler units operating in parallel [1]. Massively parallel architectures have a higher theoretical throughput than a similar sequential architecture, because many of the transistors which would otherwise be committed to inefficient optimisations are instead available for additional cores; thus, increasing the overall throughput of the device [1].

Many contemporary research projects have investigated the use of massively parallel computer devices to accelerate computation [2, 3, 4]. One such device is the Graphics Processing Unit (GPU.) The GPU has become ubiquitous with compute driven science due to its wide support base for HPC and its low cost. GPU driven HPC applications are found in a number of data driven research fields such as bioinformatics [4].

Our work leverages massively parallel platforms to accelerate combinatorial problems. A central idea of the research is a model. The model is used to describe the structure of algorithms which solve combinatorial problems on massively parallel platforms.

The model emphasises two central ideas:
- Data parallelism
- Collaboration between solvers
These two mechanisms ought to lead to a faster execution of each step of the algorithm and an increase in the rate of convergence of the system as a whole.

We present an overview of the model as well as a review of two contemporary massively parallel platforms – the GPU and the Convey FPGA Hybrid Computer – investigating their architectures and providing a brief review of contemporary research being aided by these platforms.

[1] Asanovic, K. et al. "A View of the Parallel Computing Landscape," 2009, Communications of the ACM 52, pp 56–67.
[2] Ujaldon, M. "High performance computing and simulations on the GPU using CUDA," 2012, HPCS, pp 1–7.
[3] Yang, B. et al. "GPU accelerated Monte Carlo simulation of deep penetration neutron transport," 2012, PDGC, pp 899–904.
[4] Z. Ying et al. "GPU-Accelerated DNA Distance Matrix Computation," 2011, Chinagrid Conference, pp 47–47.

Parallel Session II - High-throughput supercomputing & performance testing (1) / 53

Application of Data Intensive Science to the Electrical Energy Grid

Corresponding Authors: albertdove@yahoo.com, ivan.hofsajer@wits.ac.za

Plenary Session IV / 54

The RHINO Digital Processing Skills Development Initiative: An integrated review of the platform, resources and training structures

Corresponding Author: simon.winberg@uct.ac.za

Parallel Session I - Digital Backend Processing Hardware & Tools / 55

A Scalable Heterogeneous Architecture for Software Defined Radio on the RHINO platform

Corresponding Author: matthewbridges88@gmail.com

Parallel Session I - Digital Backend Processing Hardware & Tools / 57

An intuitive Parallel Programming Tool-flow for Software Defined Radio Signal Processing on Heterogeneous Architectures

Corresponding Author: lerato.mohapi@uct.ac.za

Parallel Session I - Digital Backend Processing Hardware & Tools / 58

Evaluation of Open-Source High-Level Tool-Flows for Rapid Prototyping of SDR Applications

Corresponding Author: khobatha.setetemela@gmail.com

Parallel Session II - High-throughput supercomputing & performance testing (1) / 59

ARM processors as a low cost alternative tool for computation of FFTs for radio astronomy

Author: Mitchell Cox[1]

[1] *University of the Witwatersrand*

Parallel Session II - High-throughput supercomputing & performance testing (1) / 60

ATLAS GPU Trigger Studies

Author: Timothy Bristow[1]

[1] *University of Edinburgh*

Parallel Session II - High-throughput supercomputing & performance testing (1) / 61

CPU Benchmarking performance for ARM Processors

Author: Robert Reed[1]

[1] *University of the Witwatersrand*

63

Mathematizing the problem of high-throughput computing

Author: Bruce Mellado[1]

[1] *University of the Witwatersrand*

Parallel Session III - High-throughput supercomputing & performance testing (2) / 64

Likelihood Analysis of Higgs anomalous couplings

Author: Gilad Amar[1]

[1] *University of Witwatersrand*

Parallel Session III - High-throughput supercomputing & performance testing (2) / 65

Massive Parallelism for Combinatorial Problems

Author: Edward Steere[1]

[1] *University of the Witwatersrand*

66

Data Processing for ALICE at the LHC

Author: Tom Dietel[1]

[1] *University of Cape Town*

67

Major Unresolved Matters about Blazars and Active Galactic Nuclei

Author: PROSPERY SIMPEMBA[1]

[1] *WITS/CBU*

Plenary session II / 68

Interferometry

Corresponding Author: davidson@sun.ac.za

69

TBD

Corresponding Author: francois.kapp@ska.ac.za

Parallel Session IV - High-speed Signal Processing Platforms and Software / 70

Noise Sensitivity of VHF Broadband Interferometers

Author: Chih-Fong (Brady) Wen[1]

[1] *University of KwaZulu Natal*

71

The ATLAS readout electronics and timing

Author: Alberto Valero[1]

[1] *Instituto de Física Corpuscular (Universidad de Valencia-CSIC)*

72

GRID in South Africa

Corresponding Author: bbecker@csir.co.za

73

The ATLAS Computing Model

Corresponding Author: yacoob@ukzn.ac.za

Parallel Session IV - High-speed Signal Processing Platforms and Software / 74

Prometeo, the next generation test-bench for the front-end electronics of the ATLAS Tile Calorimeter upgrade phase-II

Author: Xifeng Ruan[1]

[1] *University of the Witwatersrand*

Parallel Session IV - High-speed Signal Processing Platforms and Software / 75

An ATCA framework for the ATLAS Front to Back End Electronics for the Phase II Upgrade at the LHC

Author: Robert Reed[1]

[1] *University of the Witwatersrand*

Parallel Session IV - High-speed Signal Processing Platforms and Software / 76

LED Board for the Tile Calorimeter of the ATLAS Detector at the LHC

Author: Titus Masike[1]

[1] *University of the Witwatersrand*

Parallel Session III - High-throughput supercomputing & performance testing (2) / 77

Optimising Linux Operating System on ARM Architecture

Author: Jonathan Padavatan[1]

[1] *Ithemba Labs*

78

The MeerKAT radio telescope

Author: Justin Jonas[1]

[1] *SKA*

Corresponding Author: justin@ska.ac.za

The MeerKAT radio telescope is currently under construction, with completion of the 64 dish array scheduled for 2016. This will be the largest centimetre wavelength telescope in the southern hemisphere, and one of the largest in the world. Construction of the first phase of the SKA is due to commence in 2016 and the MeerKAT will be incorporated into SKA-mid, contributing about 25% of the total sensitivity for Phase 1. The science case, design and implementation of the MeerKAT will be introduced, with particular emphasis being placed on the digital signal path.

Because the cosmic radio signals are noise-like and have very wide bandwidths, the aggregate digital data rate coming from the receptors is very large. Practical issues such as radio frequency interference (RFI) multiply this data rate because high resolution ADCs are required to ensure sufficient headroom to maintain linearity. The various observing modes require the receptor data to be processed in real time by the central signal processor (CSP), and this DSP processing load has a polynomial scaling law that depends on the array and receptor parameters (e.g. number of antennas, ratio of dish diameter to dish spacing, maximum dish spacing). The various CSP modes and implementations are discussed.

Parallel Session III - High-throughput supercomputing & performance testing (2) / 79

Performance Analysis of virtualization for high performance computing

Corresponding Author: matthewcawood@gmail.com

80

RAM Benchmarking of ARM-based SoCs

Corresponding Author: gerhard.harmsen5@gmail.com

Parallel Session II - High-throughput supercomputing & performance testing (1) / 81

RAM Benchmarking of ARM-based SoCs

Corresponding Authors: thomas.wrigley@cern.ch, gerhard.harmsen5@gmail.com

82

The sROD Module for the ATLAS Tile Calorimeter Phase-2 Upgrade Demonstrator

Author: Pablo Moreno[1]

[1] *University of the Witwatersrand*

TileCal is the central hadronic calorimeter of the ATLAS experiment at the Large Hadron Collider at CERN. The main upgrade of the LHC to increase the instantaneous luminosity is scheduled for 2022. The High Luminosity LHC, also called upgrade phase-2, will imply a complete redesign of the read-out electronics in TileCal. In the new read-out architecture, the front-end electronics aims to transmit full digitized information to the back-end system in the counting room. Thus, the back-end system will provide digital calibrated information with enhanced precision and granularity to the first level trigger to improve the trigger efficiencies. The demonstrator project has been envisaged to qualify this new proposed architecture. A reduced part of the detector, 1/256 of the total, will be upgraded with the new electronics during 2014 to evaluate the proposed architecture in real conditions.The sROD module is designed on a double mid-size AMC format and will operate under an AdvancedTCA framework. The module includes one Xilinx Kintex 7 and one Xilinx Virtex 7 for data receiving and processing, as well as the implementation of embedded systems. Related to

optics, the sROD uses 4 Avago MiniPODs to receive data from the front-end electronics and 2 Avago MiniPODs to send control commands to the front-end and for communication with the first level trigger. A QSFP optical module is also included for expansion functionalities and a SFP module to maintain compatibility with the existing hardware.A complete description of the sROD module for the demonstrator including the main functionalities, circuit design and the control software and firmware will be presented.

83

The Wits Astro Data Center: a hands-on, interactive data center for multi-frequency astronomy

Author: Sergio Colafrancesco[1]

[1] *Wits University*

Corresponding Author: sergio.colafrancesco@wits.ac.za

Wits Astro Data Center for multi-frequency astronomy.

Summary:

The Wits Astro Data Center (WADC) has been recently created to respond to the challenges posed by the development of the Southern Africa platform for muulti-frequency astronomy, including MeerKAT, SKA, HESS, CTA, SALT and their links to the multi-frequency archival data sparse all over the world. The WADC also is a key response to the research strategy set up by the Wits astrophysics group that is involved in the exploitation of the data coming from these experiments as well as from several other round-based and space-borne experiments in astrophysics, astro-particle physics and cosmology.
We provide a hands-on tour to the potential, the technical characteristics and the applications of the interactive data analysis viable at the WADC.

84

Correlator and Beamformer

Corresponding Author: andrew@ska.ac.za

85

RFI Issues

Corresponding Author: jason.manley@ska.ac.za

86

The GRID in South Africa

Corresponding Author: bbecker@csir.co.za

Plenary Session I / 87

The Wits Astronomy Data Center: A hands-on data analysis center for multi-frequency astronomy

Corresponding Author: sergio.colafrancesco@wits.ac.za

88

RFI Issues

Corresponding Author: jason.manley@ska.ac.za

89

The Cherenkov Telescope Array

90

TBA

Corresponding Author: francois@ska.ac.za

Plenary Session III / 91

MeerKAT RFI Issues and signal processing challenges

Corresponding Author: jason.manley@ska.ac.za

92

Upgrade of the ATLAS TileCal Electronics

Author: Carlos Solans[1]

[1] *CERN*

Corresponding Author: carlos.solans@cern.ch

The Tile Calorimeter (TileCal) is the central hadronic calorimeter for the ATLAS experiment at the LHC.
TileCal is a key detector for the measurement of hadrons, jets, hadronic tau decays and the determination of the missing transverse energy. Its performance in Run 1 has been excellent. The absolute energy scale for all of the 5182 read-out cells has been preserved through the different calibration systems. The main upgrade of TileCal will occur for the High Luminosity LHC phase (phase 2) which

is scheduled around 2022. The upgrade aims at replacing the majority of the on- and off- detector electronics so that all calorimeter signals are digitized and sent to the off-detector electronics in the counting room. An ambitious upgrade development program is pursued where three different options are being evaluated for the front-end electronics. The option choice will be decided after extensive test beam studies. High speed optical links are used to read out all digitized data to the counting room. For the off-detector electronics a new back-end architecture is being developed. A demonstrator prototype read-out for a slice of the calorimeter with most of the new electronics, but also compatible with the present system, is planned to be inserted in ATLAS already in mid 2014 (at the end of the phase 0 upgrade).

93

White Rabbit

94

White Rabbit

95

The White Rabbit project

Author: Grzegorz Daniluk[1]

Co-authors: Erik Van Der Bij [1]; Javier Serrano [1]; Maciej Lipiński [2]; Tomasz Włostowski [1]

[1] *CERN*
[2] *CERN / Warsaw University of Technology*

Corresponding Author: grzegorz.daniluk@cern.ch

White Rabbit (WR) is a multi-laboratory, multi-company collaboration for the development of a new Ethernet-based technology which ensures sub-nanosecond synchronization and deterministic data transfer. It was initiated at CERN to provide a successor of the currently used General Machine Timing system for the accelerator complex. The project uses an open source paradigm for the development of hardware, gateware and software components. The presentation will give a general overview of the project, its origin, architecture and applications. It will describe how the three main technologies used in WR (IEEE1588, layer-1 syntonization and precise phase measurements) are combined to achieve sub-nanosecond accuracy of synchronization in the entire network. Methods to ensure high reliability and deterministic data delivery will be also outlined. Two of the main components of White Rabbit will be introduced: the WR Switch and the WR PTP Core. The presentation will then give an overview of the first White Rabbit installation for the CERN Neutrinos to Gran Sasso project and discuss future applications of White Rabbit in different places around the World. The last part of the presentation will explain the ongoing effort in IEEE1588 to include WR solutions into the standard.

Plenary Session IV / 96

White Rabbit

Corresponding Author: grzegorz.daniluk@cern.ch

Plenary Session I / 97

User Perspective of Data Analysis Flow with the ATLAS Detector

Corresponding Author: montoya@cern.ch

Plenary session II / 98

The Computing Model of ATLAS

Corresponding Author: yacoob@ukzn.ac.za

Plenary session II / 99

An Overview of SANSA Earth Observation Data Processing and Storage: Challenges and Opportunities

Parallel Session III - High-throughput supercomputing & performance testing (2) / 100

Massively Parallel High-Performance Ray-Tracing in Astrophysics Simulations

Corresponding Author: wcarlson@cern.ch

101

The MeerKAT Digital Back End

Author: Francois Kapp[1]

Co-authors: Jason Manley [2]; Sias Malan [2]

[1] *SKA*

[2] *SKA SA*

Corresponding Author: francois@ska.ac.za

The context, history and current status of the MeerKAT Digital Back End development