Open Parallel

Software for Multicore and Parallel Computing

# TOPS: An Open Platform for the SKA?

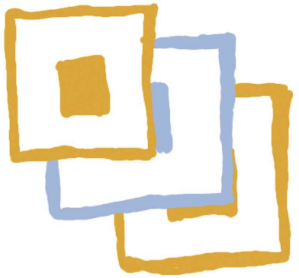
**Nicolás Erdödy**
Founder, CEO – Open Parallel Ltd

**Computing for SKA**
**Colloquium – AUT University**

Auckland, New Zealand
February 12, 2016

# Open Parallel

**Software for Multicore and Parallel Computing**



# MULTICORE WORLD 2016



catalyst

open source technologists

# Outline

- Work in progress...

# Brief

- The Problem: “data deluge”
- An Opportunity: the SKA's SDP compute model as general case
- TOPS (The Open Parallel Stack) - **A Distributed Operating System for Rack Scale Computing**.
- How to start: Open Source & OpenStack
- **Independence** – Think differently
- “This time, we have time”
- Let's work together...

# The Open Parallel Stack (TOPS)

- TOPS is something we need but we don't have yet
- The idea is to assemble a framework from the OS up to enable testing and debugging HPC programs on a small to medium scale before deploying them to systems like the SKA in high demand
- It's not about intensive R&D or significant development from scratch but to collect, preserve and build on Open Source work

# Open Parallel Ltd.

- **NZ Company** – involved with SKA since 2011.
- Formally pre-selected in 2012 by NZ Government as viable prospect for engagement in SDP and CSP.
- **Since 2013 Open Parallel is formally:**

  - Work Package Manager of the **Software Development Environment for the CSP**,

  **- Contributing to SDP Compute Platform**,

  - Member of the **New Zealand SKA Alliance**

# Success takes time


CHAMPIONS


10
10
Coca-Cola


VISA


NEVADA

Could the SKA and other HPC projects generate an ecosystem that triggers the next generation of “world champions” from our countries?

# Part 2 – Where are we going?

As today's HPC
becomes tomorrow's
Cloud computing platform
it will enable a wider application of
**Machine Understanding**
-the near real-time
complex modelling
and analysis of data
that leads to insight
and faster decisions.

# What is the SKA?

- The world's largest radio telescope
- The ultimate big data project
- The largest supercomputer in the world
- A technological management challenge

  and...
- **The general case of future HPC + Cloud...**

# SKA Context

- The SKA needs exascale computing
- There is an architecture for the system
- Processor details are not finalised
- Radio telescopes last for decades
- Processors will be replaced/upgraded
- Programming can't wait for the hardware

# Major requirements

- Longevity
- Adaptability
- Acceptability
- Manageability
- Availability

# Longevity

- Exascale may/will need new computing models
- The old ones aren't going away
- New languages like Chapel and X10 exist (remember Fortress?)
- But C, C++, and Fortran have a proven track record.  Climate models typically use Fortran.
- UNIX is the pre-eminent multiplatform OS and has been around since 1970s

# Programming

- Software must be ready when hardware is
- So it must be developed on **other** hardware
- Impractical to develop on SKA at any time
- Must write, test, and profile on smaller systems
- The Open Parallel Stack is needed on them too

# Acceptability

- Almost all the TOP500 use Linux
- Including Cray, Blue Gene, Tianhe-2
- Compute nodes may use a small kernel
- Compute island managers use a Linux variant
- System management may use a standard Linux

# Adaptability

- Stack must scale from lab machines to the SKA
- Stack should not be bound to one CPU type
- Nor to one storage system
- Nor to one interconnect
- System needs to be maintainable
- Efficient communication is vital
- Linux has drivers for Infiniband, Thunderbolt, ...

# Management

- Power, communication, software.
- Power use must be monitored
- and controlled.
- Communication must be monitored
- and controlled.
- Software must be packaged, deployed, and scheduled.

# Management (II)

- Ways to measure power exist
- Ways to slow machines down or turn off this or that exist
- Power management was especially important for Android (phones, tablets)
- Policies suitable for exascale machines still have to be written
- Ways to measure communication already exist
- Ways to control the use of communication devices exist
- Policies for deciding which computations should get what share of the bandwidth, that scale to exascale, need to be developed
- **Packaging and deployment are where OpenStack and Catalyst come in**

# Communication with humans:

- Understanding the behaviour of massively parallel programs is difficult for people
- Performance visualisation tools can help
- What's your experience?

# Availability

- If the SKA is down, data are lost forever.
- Storage devices and processors *will* fail.
- Software *will* need correction.
- New applications *will* be developed.
- Need to deploy software to many islands.
- Need to restart work from failed devices.

# Standing on others' shoulders

- Use OpenStack
- open source scalable "cloud computing"
- can support TOPS deployment needs
- can support monitoring needs
- shared filesystems
- containers

# Containers

- Can provide fault isolation
- By taking snapshots, can provide restart
- TOPS will need to choose from several
- LXC is particularly interesting

# Standing on others' shoulders (2)

- OpenHPC is important
- TOPS will need to track its abstraction interfaces
- Some scientific data visualisation tools might be included in TOPS
- BTW, it seems that “open” is the fastest and most effective way to commoditisation and COTS equivalence...

# Could SKA's IT be a Black Swan?

- **“Black Swan”** = high-impact events that are rare and unpredictable but in retrospect seem not so improbable
- **One in six** IT projects (…) is a black swan, with a cost overrun of **200%**, on average (*)
- **Developers struggle to combine different software systems**
- **61% of managers report major conflicts between project and line organisations**

- (*) “Why your IT Project may be riskier than you think”. B. Flyvbjerg et al. HBR, Sept. 2011

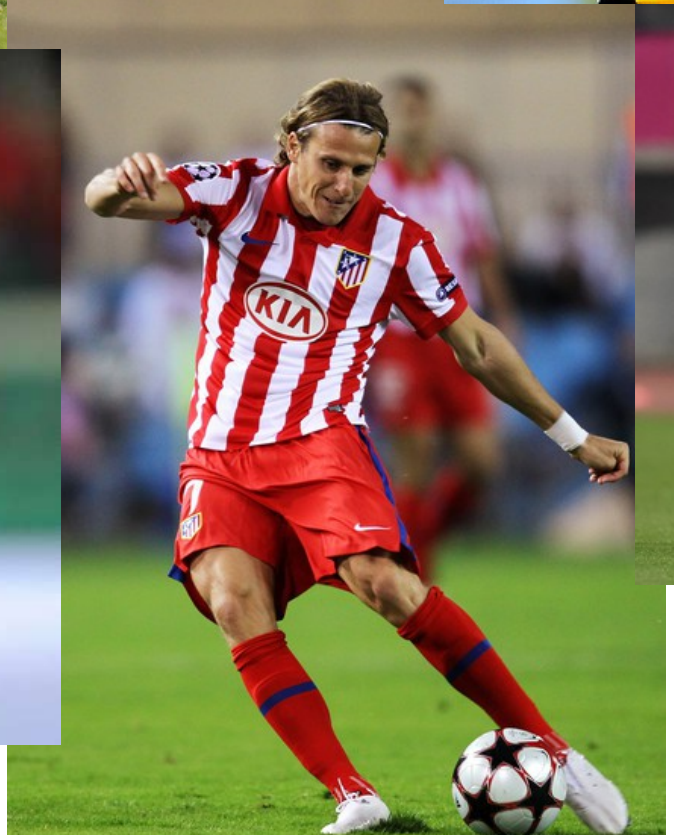
Would software have longevity, adaptability, acceptability, manageability and availability as **Diego Forlán**?


10


vodafone


TRAMONTINA
antel


PIRELLI


KIA


YANMAR
10
28

## 15-16-17 February 2016

# 5th Multicore World - Wellington

- Peter Kogge (Notre Dame, IBM Fellow, DARPA Exascale report)
- Alex Szalay (Johns Hopkins, Sloan)
- Geoffrey C Fox (Indiana)
- John Gustafson (A*STAR, Gustafson's Law, Singapore)
- Happy Sithole (Director CHPC, South Africa)
- Tshiamo Motshegwa (HPC, SKA, Botswana)
- Chun-Yu Lin (NCHC, Taiwan)
- Balazs Gerofi (RIKEN – K Computer, Japan)
- VMware, DELL, Oracle, NVIDIA, INTEL, Altera, Catalyst
- Cassandra, LMAX, SCION, ICRAR
- MacDiarmid-VUW, AUT, Otago, Melbourne

# Multicore World 2017

- **20 – 23 February 2017, Wellington**
- Pete Beckman, Director Exascale Technology Institute. Project – Argo (Argonne Labs)
- Barbara Chapman, Head of Computer Science at DoE Brookhaven Institute -collaboration w/DoD
- Filippo Spiga, Head of Research of Software Engineering at University of Cambridge
- Michelle Simmons, Director Centre for Quantum Computing, UNSW, Australia
- Hermann Hartig, Lead OS – TU Dresden, Germany

# Thank you!

- OpenParallel.com
- MulticoreWorld.com
- Nicolas.Erdody@openparallel.com
- about.me/nicolas.erdody
- Oamaru, South Island, New Zealand