



www.csiro.au

Report on International Exascale Software Project meeting, Maui, Oct 18 - 19, 2010

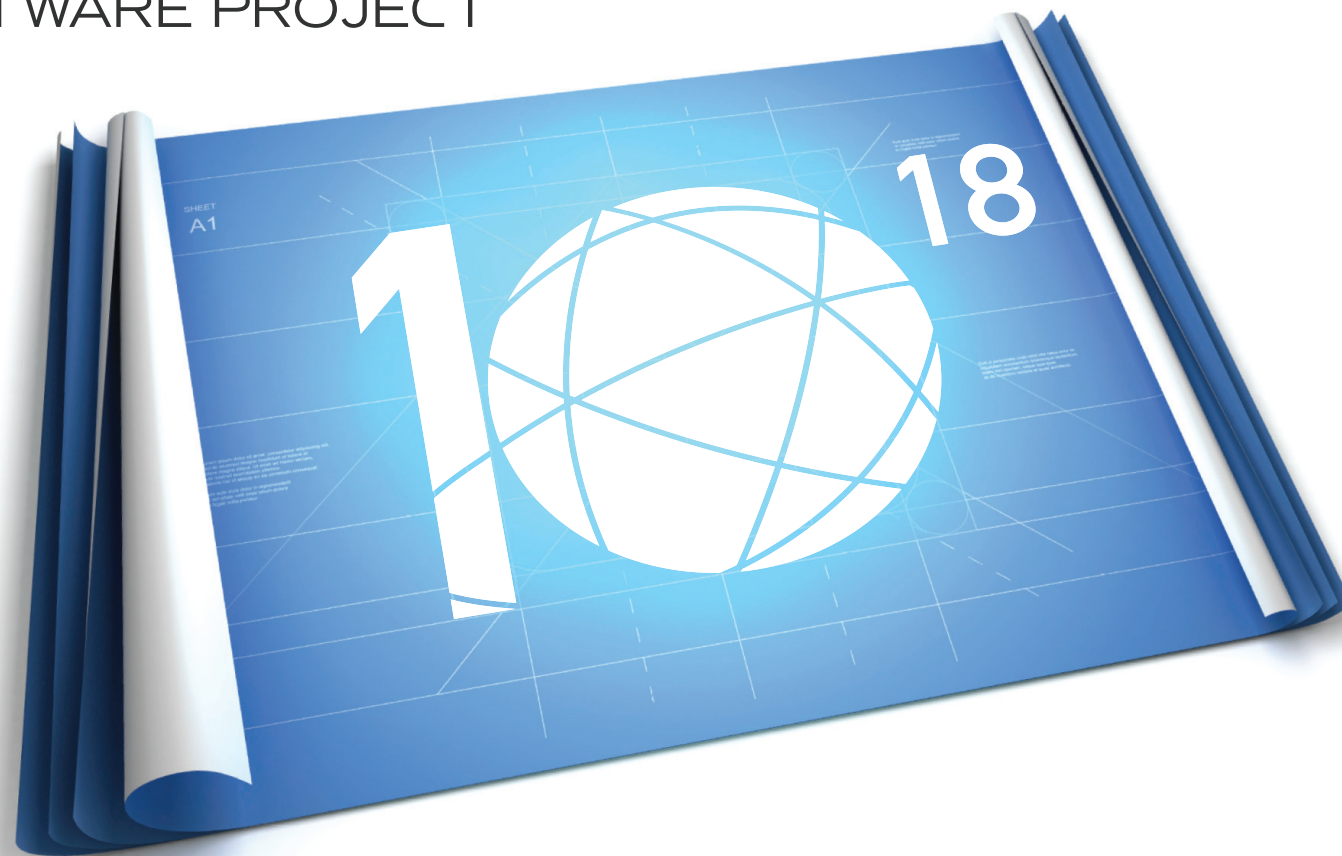
Tim Cornwell, ASKAP Computing Project Lead
Australian Square Kilometre Array Pathfinder



International Exascale Software Project

- <http://www.exascale.org>
- Organised by Jack Dongarra and Pete Beckman
- Also European equivalent

INTERNATIONAL **EXASCALE** ROADMAP 1.0 SOFTWARE PROJECT



Co-design

- Systems development paradigm

- Wikipedia:

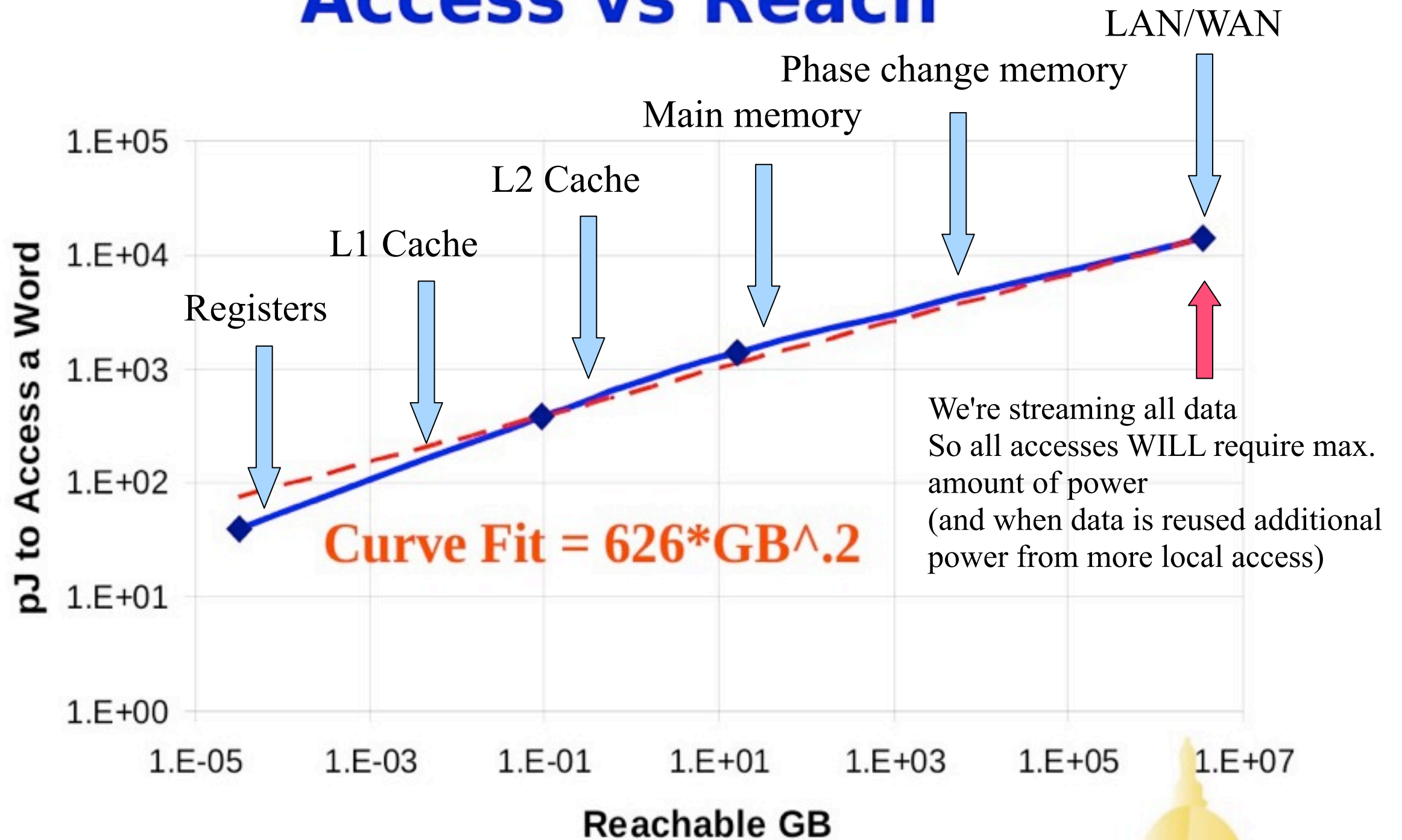
In co-design there is an understanding that all human artifacts are designed with a purpose. In co-design one tries to include those perspectives that are related to the design in the process. It is generally recognized that the quality of design increases if the stakeholders interests are considered in the design process. Co-design "begins when first you view the world through the eyes of another."

- All relevant areas of expertise involved in system design
- Computing hardware/software/applications
- IESP to select limited number of Co-Design Vehicles
- Presented the case that SKA should be a CDV
- See SKA memo 128

Does SKA meet the criteria for CDV?

<i>Petascale or near-petascale application today with a demonstrated need for exascale performance</i>	<ul style="list-style-type: none"> • ~ 100TF by end 2010 • ~ 1PF by end 2011 • ~ Operations: 100TF, 100PB/year by 2013
<i>Significant scientific goals in an area that is expected to be a scientific or societal driver for exascale computing</i>	<ul style="list-style-type: none"> • Astronomy is widely recognised as a societal good and technology driver • Demonstrated by successful funding of SKA
<i>Realistic and a definable set of steps to exascale that can be mapped out over 10 years or less</i>	<ul style="list-style-type: none"> • Pathfinders (e.g. ASKAP, LOFAR) • SKA1 • SKA2
<i>Community experienced in algorithm, software and/or hardware developments and is willing to engage in the exascale co-design process.</i>	<ul style="list-style-type: none"> • Community is experienced • Exascale is necessary for science goals • Strongly motivated to participate in co-design
<i>Modular and open enough to stimulate the development of additional modules addressing related questions in the area</i>	<ul style="list-style-type: none"> • Astronomy is very open, nearly all OSS • Design, software shared across community
<i>Fills a slot in the portfolio of extreme scale application needed to test all these dimensions</i>	<ul style="list-style-type: none"> • Algorithms (near EP and non EP) • Data graphs, flow, and management

Access vs Reach



Source: Energy at ExaFlops, Peter M. Kogge, SC09 Exa Panel

Exascale=GigaHz KiloCore MegaNode

Systems	2009	2018	Difference Today & 2018
System peak	2 Pflop/s	1 Eflop/s	O(1000)
Power	6 MW	~20 MW	
System memory	0.3 PB	32 - 64 PB [.03 Bytes/Flop]	O(100)
Node performance	125 GF	1,2 or 15TF	O(10) – O(100)
Node memory BW	25 GB/s	2 - 4TB/s [.002 Bytes/Flop]	O(100)
Node concurrency	12	O(1k) or 10k	O(100) – O(1000)
Total Node Interconnect BW	3.5 GB/s	200-400GB/s (1:4 or 1:8 from memory BW)	O(100)
System size (nodes)	18,700	O(100,000) or O(1M)	O(10) – O(100)
Total concurrency	225,000	O(billion) [O(10) to O(100) for latency hiding]	O(10,000)
Storage	15 PB	500-1000 PB (>10x system memory is min)	O(10) – O(100)
IO	0.2 TB	60 TB/s (how long to drain the machine)	O(100)
MTTI	days	O(1 day)	- O(10)

Software costs for exascale

- Software effort does not scale as computing flops!
- Scaling upwards to Exascale dependent on relatively few individuals
 - Contract out some work
- ~ All calibration and imaging processing software must be rewritten
 - Was true for ASKAP
 - ASKAP expects to spend O(20) FTE-years, halfway through now
 - ASKAP scaling work dependent on 3 individuals
 - Aided by end-to-end control of operational model and software stack
- We must partner with large industry/research lab
 - Building relationships now e.g. ORNL, LLNL
 - Important to lead with our accomplishments - LOFAR and ASKAP 100TF
- Data management must be integral part of Exascale effort

My conclusions

- Exascale will happen by ~ 2018...
- ... but we may not be able to use it unless we act coherently
- Technical points
 - Flops are essentially free
 - Memory (50%), memory bandwidth/networking (50%) are limiting
- ~ Evolving consensus/strawperson Exascale architecture: GigaKiloMega
- SKA is a unique data-driven application
 - Substantial interest in SKA as co-design vehicle
 - Data intensive applications under-represented in HPC community
- Co-design is essential
- SKA Exascale Co-design Centre?
 - Astronomers, computer scientists, application scientists, calibration and imaging specialists, computing hardware engineers, data management engineers, operations engineers, vendors

ATNF/ASKAP

Tim Cornwell
ASKAP Computing Project Lead

Phone: +61 2 9372 4261

Email: tim.cornwell@csiro.au

Web: www.atnf.csiro.au

www.csiro.au

Thank you

Contact Us

Phone: 1300 363 400 or +61 3 9545 2176

Email: enquiries@csiro.au Web: www.csiro.au