Square Kilometre Array ...its Software Defined Supercomputer ... and a very fast parallel file system

John Garbutt @johnthetubaguy

StackHPC

What is the **SKA**?



SKA1 MID

SKA1 LOW



How to do Science with all that data?

Part 1: Science Data Processor Architecture

SKA Science Data Processor







SDP Operational System



Data Islands



Key Challenges

- Massive data rates
- Sized for average load
- Science Workflows evolving over time
- 24x7 Telescope Operation

Part 2: Software Defined Supercomputer

ALaSKA Performance Prototype Platform



Scientific OpenStack



k8s cloud-provider-openstack



Data Islands



Part 3: Cambridge Data Accelerator



Top500 Nov18		35	244	AIGIS - Sugon Cluster W780I, Xeon E5-2 NVIDIA Tesla P100, Infiniband FDR , Sug Government China	630v4 10C 2.20 on	GHz, 24,4	00	1,307.0	300	4.357	
		36 87		Cumulus - PowerEdge C6420/C6320, Intel Xeon Phi 7210/Xeon Gold 6142F 16C 2.6GHz, Intel Omni-Path , Dell EMC University of Cambridge United Kingdom			76	2,271.4	530	4.286	
12	443	Hua V10 Tele Chin	37	28	Cray XC50, Xeon Platinum 8160 24C 2.10 interconnect , Cray Inc./Hitachi Japan Meteorological Agency Japan	GHz, Aries	135,	792	5,730.5	1,354	4.232
13	284	Wilk Infini Univ Unite	es-2 - iband E ersity o ed King	Dell C4 EDR, NV of Camb gdom	130, Xeon E5-2650v4 12C 2.2GHz, IDIA Tesla P100 , Dell EMC ridge	21,240	<mark>1</mark> ,193.0	114	10.428		
14	15	HPC 2.1G Eni S Italy	4 - Pro Hz, Me 5.p.A.	liant DL Ilanox I	.380 Gen10, Xeon Platinum 8160 24C nfiniBand EDR, NVIDIA Tesla P100 , HPE	253,600	12,210.0	1,320	9.804		

IO-500

This is the official IO-500 ranked list¹⁾ from November 2018 (from SC 2018).

Please see also the 10 node challenge ranked list from November 2018.

The list shows the best result for a given combination of system/institution/filesystem.

#		io500								
	institution	system	storage vendor	filesystem type	client nodes	client total procs	<mark>data</mark>	score	bw	md
									GiB/s	klOP/s
1	Oak Ridge National Laboratory	Summit	IBM	Spectrum Scale	504	1008	zip	366.47	88.20	1522.69
2	Korea Institute of Science and Technology Information (KISTI)	NURION	DDN	IME	2048	4096	zip	160.67	554.23	46.58
3	University of Cambridge	Data Accelerator	Dell EMC	Lustre	528	4224	zip	158.71	71.40	352.75
4	JCAHPC	Oakforest- PACS	DDN	IME	2048	16384	zip	137.78	560.10	33.89
5	WekalO	WekalO	WekalO		17	935	zip	92.95	37.39	231.05
6	KAUST	Sha <mark>heenII</mark>	Cray	DataWarp	1024	8192	zip	77.37	496.81	12.05
7	University of Cambridge	Data Accelerator	Dell EMC	BeeGFS	184	5888	zip	74.58	58.81	94.57
8	Google	Exascaler on GCP	Google	Lustre	120	960	zip	56.77	23.06	139.74
9	JCAHPC	Oakforest- PACS	DDN	Lustre	256	8192	zip	42.18	20.04	88.78

IO⁵⁰⁰

Data Accelerator Platform

- 24 x Dell EMC PowerEdge R740xd
- 12 Intel SSD P4600 1.4 TB
- 2 Intel Omni-Path Adaptors
- User requests capacity
- Partitioned into Storage & Metadata
- BeeGFS: Pick NUMA local IB card
- Lustre: multi-rail



B

B

• Max IOR Easy: ~4Tb/s Read, ~2Tb/s Write

Slurm Integration



Slurm Burst Buffer Workflow



Non-Persistent Burst Buffer Mounts



How to do Science with all that data?

SKA SDP's Buffer

- Flexible Architecture
- Control Flow Prototypes
- Data Rate Prototypes

How to get involved?





THE CROSSROADS OF CLOUD AND HPC

Exploring OpenStack Cloud Computing for Scientific Workloads

Scientific OpenStack SIG

https://www.openstack.org/science/

Thank You!

StackHPC @johnthetubaguy john.garbutt@stackhpc.com