KAUST Confidential (when completed)

KAUST Supercomputing Laboratory (KSL)

SAMPLE Project Proposal (PP)

Project Title	Porting of ANSIS Code to BG/P
Principal Investigator (PI)	Dr RG Watson
PI Signature	
Date of Proposal	27 Feb 11

KSL Use Only

Proj Type	☐ **Development / ☐ Production**	**Approval Date:**	

System	Volume	Duration
BG/P		
X86 Linux		
Disk		
Tape		

Document Owner(s)
Richard Orme – Business Manager, KAUST Supercomputing Laboratory (KSL)

Version Control:

Version	Date	Author	Change Description
1.0	2 Oct 09	Richard Orme	Final Draft (under review by KAUST Legal)
1.1	24 Oct 09	Richard Orme	Revision 1
1.2	28 Oct 09	Richard Orme	Revision 2
1.3	15 Nov 09	Richard Orme	Allow scanned submission
1.4	22 Feb 10	Richard Orme	General revision
1.5	1 Apr 10	Richard Orme	Add submission details & PI Signature
1.6	20 Apr 10	Richard Orme	Typo correction
1.7	9 Aug 10	Richard Orme	Updated helpdesk email address
1.8	27 Feb 11	Richard Orme	Format update
1.9	4 Apr 11	Richard Orme	Added SCPM core-hrs required
1.10	3 May 11	Richard Orme	Added SCPM Jobs description
1.11	9 Jul 11	Richard Orme	Amended SCPM descriptions to CCS

Available Systems (as at 9 Jul 11):

1. A 16-rack IBM Blue Gene/P system, comprising 16,384 nodes, each with four cores and 4GB of memory - for a total of 65,536 cores and 64TB of memory..
2. A 96-node Linux x86 cluster, each 8-way node with 32 GB of memory.
3. 1.5 PB of IBM DCS9900 disk storage with an aggregate data transfer bandwidth of 16 GBps.

Submission

Please send a scanned copy of the completed Project Proposal to:

help@hpc.kaust.edu.sa

Definitions:

- **Development Project** - A development project provides access for system familiarization, code porting, performance assessment, and other pre-production work. Development Projects will not be allocated significant computing resources.
- **Production Project** - A production project requires that applications have been ported and tuned, and that performance assessments have been completed. Production Projects will be allocated significant computing resources.

Principal Investigator (PI):

Name:	RG Watson
Email:	rob.watson@kaust.edu.sa
Tel:	555-8080
Organisation:	KAUST
Department:	Clean Combustion Research Center (CCRC)
Organisation Address:	Building UN1550, Office 0111 4700 King Abdullah University of Science and Technology Thuwal 23955-6900 Kingdom of Saudi Arabia
KAUST Cost Center :	99999 ***Note:*** *KAUST departments only*

PI Signature:		**Date:**	27 Feb 11

Additional Investigators

Please provide the same information for any further Additional Investigators in 'Additional Information'.

1	**Name:**	Dr S Holmes
	Email:	steve.holmes@kaust.edu.sa
	Tel:	555-8081
	Organization:	Clean Combustion Research Center (CCRC)
	Org Address:	N/A
2	**Name:**	Dr B Moriarty
	Email:	bruce.moriarty@kaust.edu.sa
	Tel:	555-8082
	Organization:	Clean Combustion Research Center (CCRC)
	Org Address:	N/A

Collaborators (External):

Please provide the same information for any other Collaborators in 'Additional Information'.

1	**Name:**	Dr P Jeckyll
	Email:	paul.jeckyll@imperial.ac.uk
	Tel:	+44-207-555-1234
	Organization:	Imperial College London
	Org Address:	The Strand, London
2	**Name:**	
	Email:	
	Tel:	
	Organization:	
	Org Address:	

Project Description:

Please describe the activities proposed, including current state of art, research work proposed, expected milestones, and deliverables, as well as a summary description in the box below, and include the scientific field of the investigation as part of the description.

Note: Citations of the scientific literature are encouraged in order to show where the proposed simulations stand with respect to the 'state of the practice' in terms of such factors as model generality, resolution, and advantages of simulation versus experiment and theory.

The ANSIS code was developed to provide accurate and highly resolved subsurface velocity structures of the earth by inverting the seismic data acquired from the earth surface. Accurate velocity imaging tools are strongly needed for seismic imaging and interpretation in order to find oil and gas deposits. Therefore, ANSIS is of interest to the oil and gas industry. However, it is time-consuming and requires a

huge amount of computing resources. This leads to the goal of this project which is to develop a waveform inversion tool (ANSIS) for large-scale seismic exploration.

The most time-consuming part of waveform inversion is numerical forward modeling which will be performed several times during the inversion process to find a velocity model that minimizes the misfit between the observed and calculated data. One way to efficiently run the forward modeling is to apply a domain decomposition method to our time-domain finite-difference algorithm to make use of the large number of computing cores in the Shaheen supercomputer.

Currently, the ANSIS code performs very well on up to 256 cores. Now we need to port it to BG/P and test the performance at different scales.

Once ported, the ANSIS code will be tested on three-dimensional synthetic and field data sets prior to submission of a Production Project Proposal.

Scientific Impact:

Please detail the expected scientific impact of the proposed research.

We expect the ANSIS code to become a really useful, computationally efficient tool for finding oil and gas in areas with complex subsurface structures.

Codes & Libraries:

- *Please provide the following information for each code or library which will be used.*
- *If needed, please include the same information for any other codes or libraries to be used in 'Additional Information' at the end of this proposal, or attached on a separate sheet.*

1	**Name of Code/Library:**	ANSIS
	Ownership / Licensing:	KAUST
	URL (*for Open Source codes*)	N/A
	Function:	3D waveform inversion using domain decomposition
2	**Name of Code/Library:**	SciPy
	Ownership / Licensing:	Open source
	URL (*for Open Source codes*)	http://www.scipy.org/
	Function:	N-dimensional array manipulation

Code Readiness:

- *Please provide details of code performance and scalability achieved, and note any known issues which might impact production execution.*
- *If possible, please provide a simple table/graph showing the required 'wall time' versus the number of cores used.*
- *If needed, please include the same information for any other codes or libraries to be used in 'Additional Information' at the end of this proposal, or attached on a separate sheet.*

1	**Name of Code/Library:**	ANSIS
	Scalability:	Code ran on our local Linux x86 cluster and got the following walltimes for a typical production job
	Known Issues:	# cores wallclock time (hrs)
2	**Name of Code/Library:**	
	Scalability:	
	Known Issues:	

Resource Requirements:

Compute Resource Requirements:		**Standard Jobs Requirement**	**CSS Jobs Requirement**	**Duration (in Days)**
MINIMUM Requirements:	**BG/P (core-hours)**	100,000	5,000,000	6 months
	Linux x86 (core-hours)	50,000	1,000,000	6 months

MAXIMUM Requirements:	**BG/P (core-hours)**	500,000	10,000,000	12 months
	Linux x86 (core-hours)	100,000	2,000,000	12 months
Memory Requirements:	**BG/P (GB/core)**	1		
	Linux x86 (GB/core)	4		
Time-to-Solution at different core counts:				

Notes:

1. Standard Jobs are jobs submitted to the systems via the LoadLeveler Job Scheduler that are scheduled and run as soon as possible.
2. 'CCS' stands for 'Compute Cycle Scavenging'. CCS Jobs are pre-emptible back-fill jobs that can be pre-empted by 'Standard Jobs'. Therefore, when CCS Jobs will be run cannot be guaranteed. We recommend that CCS Jobs are check-pointed.
3. The 'Shaheen' BG/P and Linux x86 systems are allocated in 'core-hours', i.e. one core for one hour.
4. The 'Shaheen' BG/P system has 1GB of memory per core.
5. The Linux x86 cluster has 4GB of memory per core.

Storage Resource Requirements:		**Requirement (in TB)**	**Duration (in Days)**
MINIMUM Requirements:	**Disk**	1	6 months
	Tape	5	6 months
MAXIMUM Requirements:	**Disk**	5	12 months
	Tape	10	12 months

Notes:

1. KSL policy states that all project data stored on disk will be removed 6 months after the completion of the project.
2. KSL policy states that all project data stored on the tape archive will be removed upon the completion of the project *unless* special arrangements have been requested and granted..

Other Resource Requirements:

e.g. prep time required to characterize scalability; any human/machine interaction required during the computation, etc.**:**

1	May require intervention to stop slow performing jobs
2	

Consultancy Support Required from KSL

Please indicate the number of man-days and type of any support required from KSL staff, which can include:

- *Code development*
- *Code porting*
- *Code performance tuning*
- *Algorithm development*
- *Pre- and Post-Processing code development*
- *Data analysis and visualization support.*
- *Research program development.*
- *Project management support*

Please note that KSL may be entitled to a share in the Intellectual Property Rights to any research results produced as a result of support provided by KSL.

Estimate 10 man-days of code performance tuning assistance required.

Confidentiality and Legal Issues:

Please provide details of any potential confidentiality or legal issues, e.g.*:*

- *Is the project proposal confidential? If so, how?*
- *Is the data confidential? If so, how?*
- *Are any other aspects of the project confidential? If so, how?*
- *If the project is successful, could it be the subject of publicity?*
- *Do any third parties have ownership of any codes or data being used?*

The code developed under this project will be only used within our research group, but this may change in the future.

Project could produce a useful new tool for finding oil and gas in areas with complex subsurface structures, resulting in significant publicity for KAUST.

Other Information in Support of the Proposal:

Please include any other considerations you feel would support of this proposal, e.g.:

- *Would the proposed project have any social impact?*
- *Would the proposed project have potential for generating good publicity for KAUST and/or partner organizations?*
- *Would the proposed project develop any useful tools that might be shared with others inside or outside of KAUST?*

This project has potential for generating good publicity for KAUST, and could develop a tool that would be used throughout the oil & gas industry.

Additional Information

Below is a graph of the wallclock times required for a typical production run of ANSIS on our local Linux cluster.

For KAUST SL Office Use Only:

Date of Review:		
Project Review Board Members:		
Result:	***Approved***	***Not Approved***

Resources Allocated:

System	Volume	Duration
BG/P (core-hours)		
Linux x86 (core-hours)		
Disk (TB)		
Tape (TB)		

Services Allocated:

Service Description	Volume (Man-Days)	Duration

Comments/Notes:
