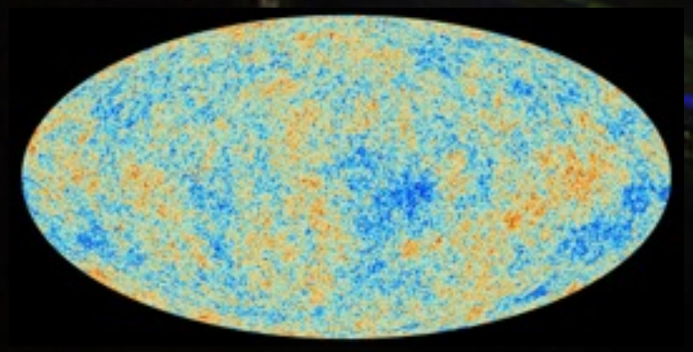
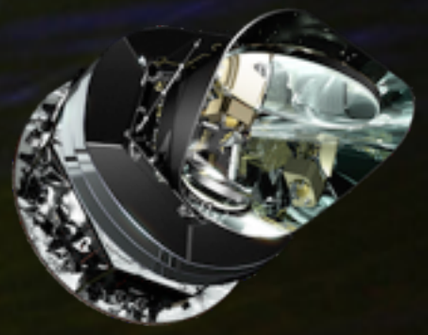
Phi in the Sky
and the Extreme COSMOS





Phi in the Sky and the Extreme COSMOS

(Part of the STFC DiRAC Facility)

Professor Paul Shellard

Stephen Hawking Centre for Theoretical Cosmology
University of Cambridge

Institutions include Cambridge, Sussex, Portsmouth, IC, UCL, UCLan,Oxford, Durham, Manchester, Nottingham, Edinburgh

Mathematical Sciences

Intel Parallel Computing Centre collaboration
with James Briggs, Juha Jaykka (COSMOS)
John Pennycook (Intel) and Cheng Liao (SGI)

COSMOS@DiRAC

DiRAC-2 system configuration: SGI UV2000 (+ UV1000)

- 1856 Intel Xeon processor cores (Sandy Bridge)
- 31 Xeon Phi coprocessors (1860 KNC cores)
- 14.5 TB global shared memory
- At delivery in 2012 (upgrade COSMOS IX):
  - ◇ World's largest single-image SMP system
  - ◇ First SMP system accelerated with Xeon Phi

COSMOS@DiRAC

COSMOS focus on code innovation and development

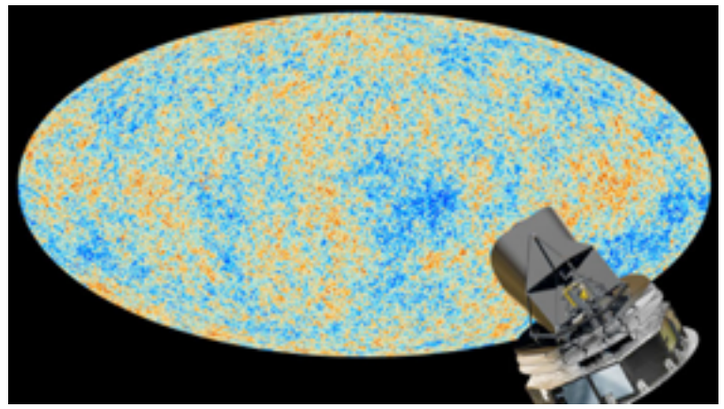
Highly competitive field, large data, new user influx, so SMP!

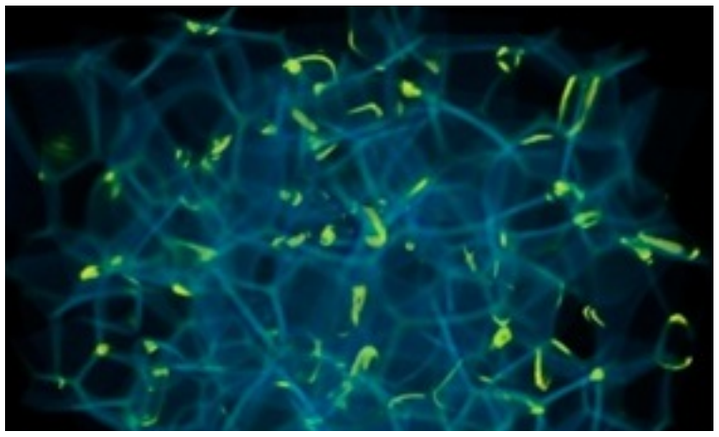
The most flexible HPC platform in the industry (inclusive)

- *Single system image (OpenMP, scalable IO),*
- *MIC coprocessors (new hybrid paradigms),*
- *Cluster (connectionless MPI),*
- *PGAS (UPC, etc)*

COSMOS@DiRAC Extreme Science

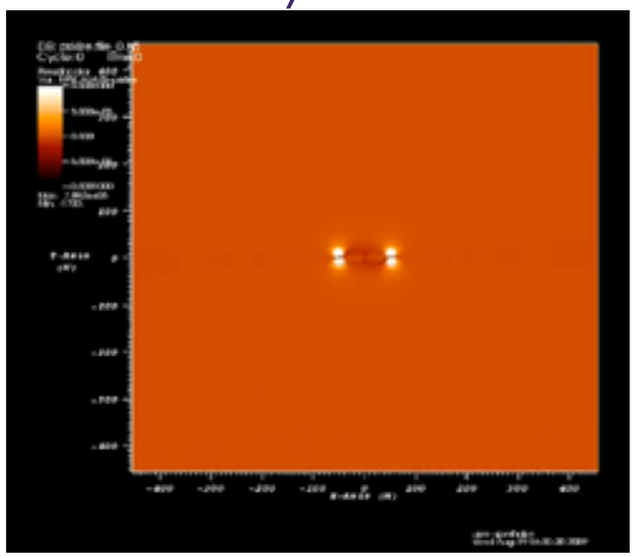
I. SPACE SCIENCE: Analysis of maps of the cosmic microwave sky and surveys of galaxies





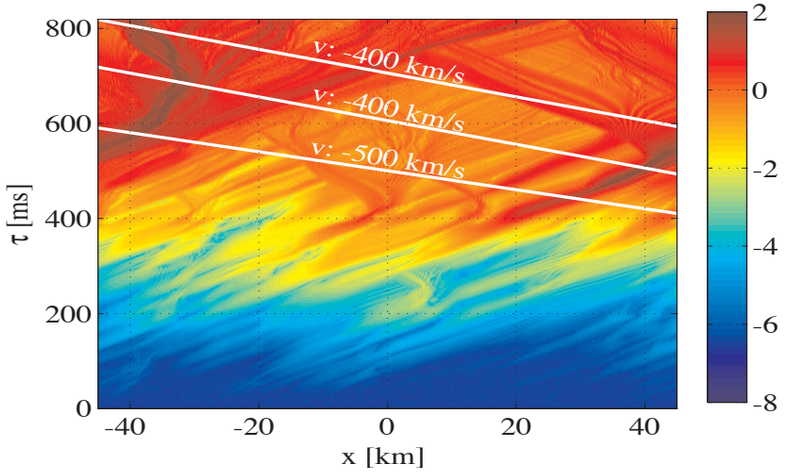
II. MODELLING INFLATION

Observational consequences of inflation and early universe theory.



III. BLACK HOLE COLLISIONS:

BH mergers, the gravitational waves produced and backreaction.



IV. EXOPLANET SEARCHES:

Detection, observation and characterisation of sub-stellar obj.

Discovery of Gravitational Waves

BICEP2 results: World media - 18th March 2014

U.S. INTERNATIONAL 中文网

The New Yo

Monday, March 17, 2014 | Today's Paper

WORLD U.S. NEW YORK BUSINESS OPINION SPORTS

West Steps Up Sanctions on Russia After Crimea Vote

By PETER BAKER 11:14 AM ET

"We're making it clear there are consequences for these actions," President Obama said in announcing sanctions on those deemed to be

theguardian

News | Sport | Comment | Culture | Business | Money | Life & style

News › Science › Physics

Primordial gravitational wave discovery heralds 'whole new era' in physics

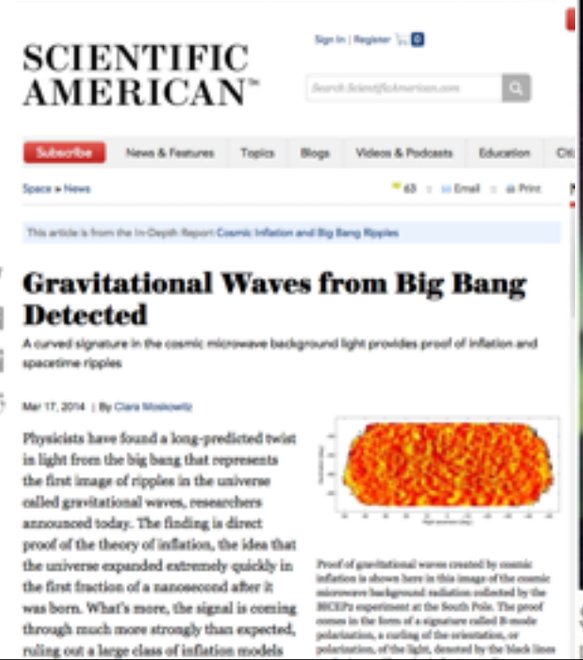
Gravitational waves could help unite general relativity and quantum mechanics to reveal a 'theory of everything'

Stuart Clark

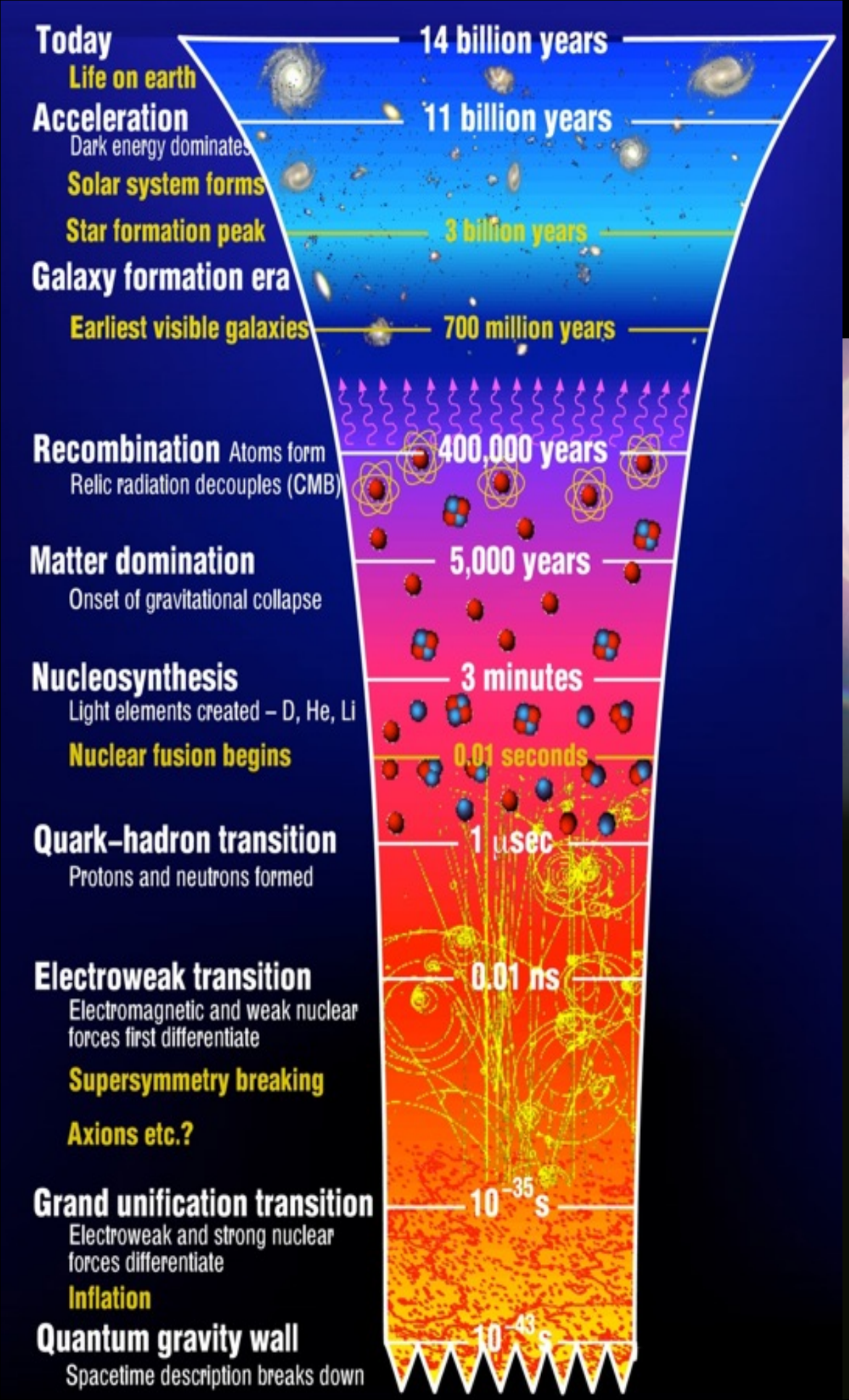
The Guardian, Monday 17 March 2014 18.08 GMT

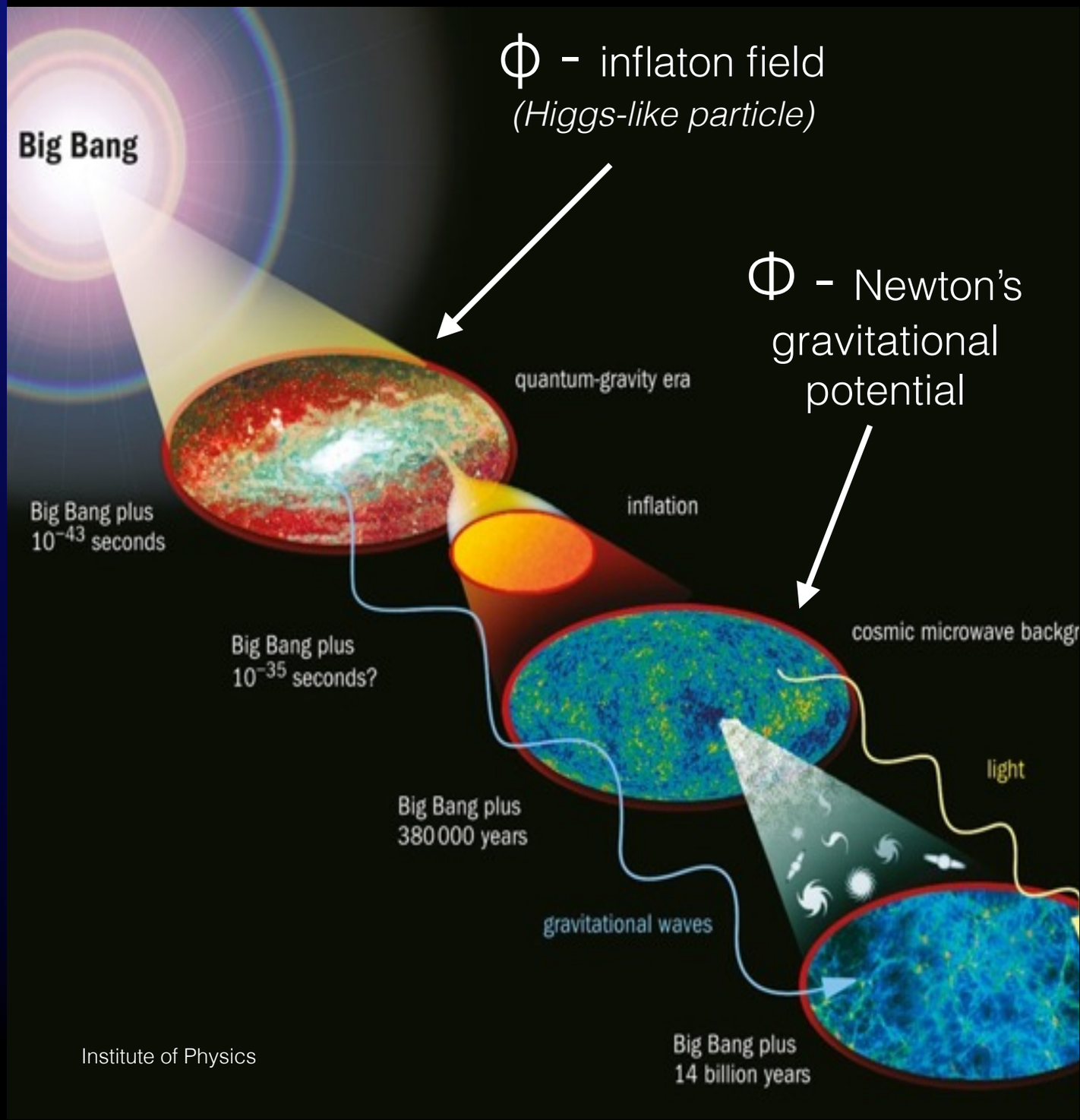
Jump to comments (1675)

Scientists detected telltale signs of gravitational waves using the Bicep2 telescope (far left) at the south pole. Photograph: Keith Vanderlinde/NSF

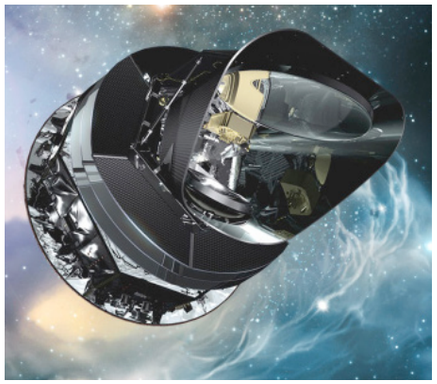


A brief history of the extreme Universe …



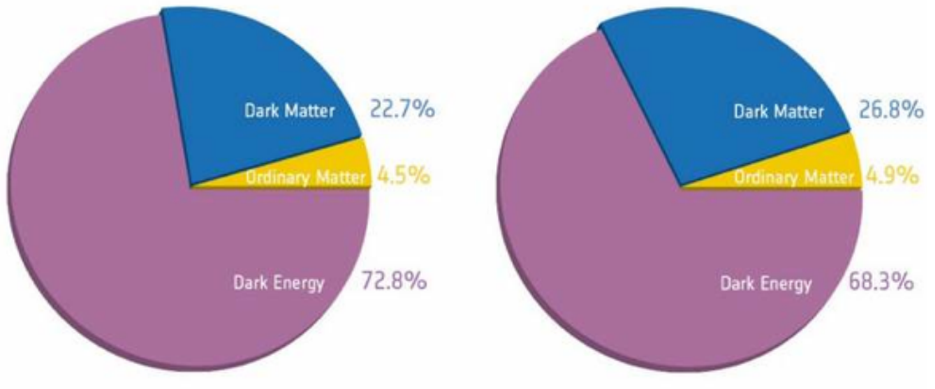


Planck Satellite Science

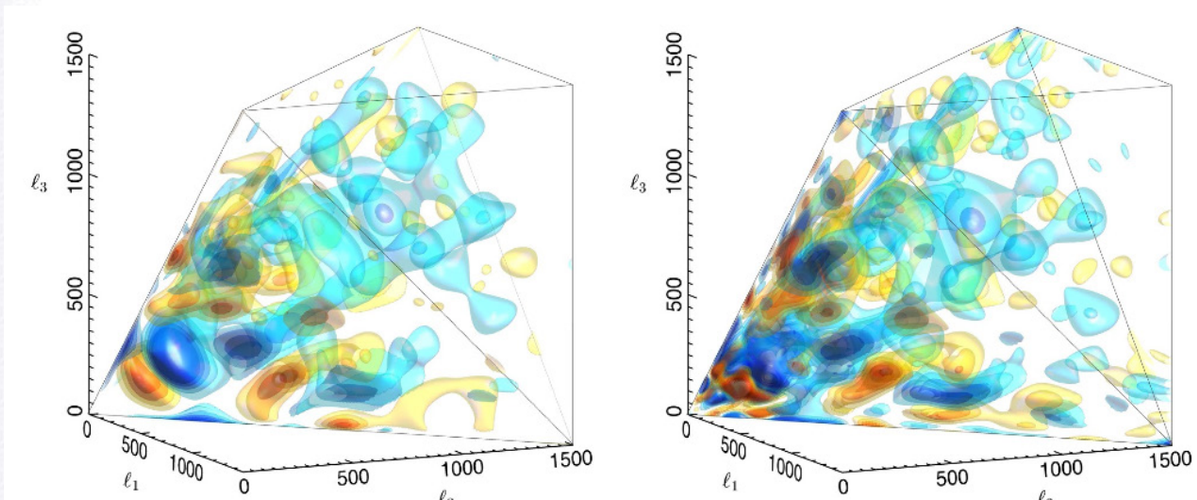


- ESA satellite offers unprecedented view of the relic light left over from the Big Bang 13.8 billion years ago.
- Finest cosmological dataset available - new *high precision estimates of cosmological parameters* with work led on DiRAC

- Publishing on BICEP polarisation results soon!
- New qualitative thresholds crossed using unique DiRAC capabilities.
- *Non-Gaussian statistics used to test inflation*
  Reconstruction of the 3-point correlator *or CMB bispectrum* ("triangles in the sky").

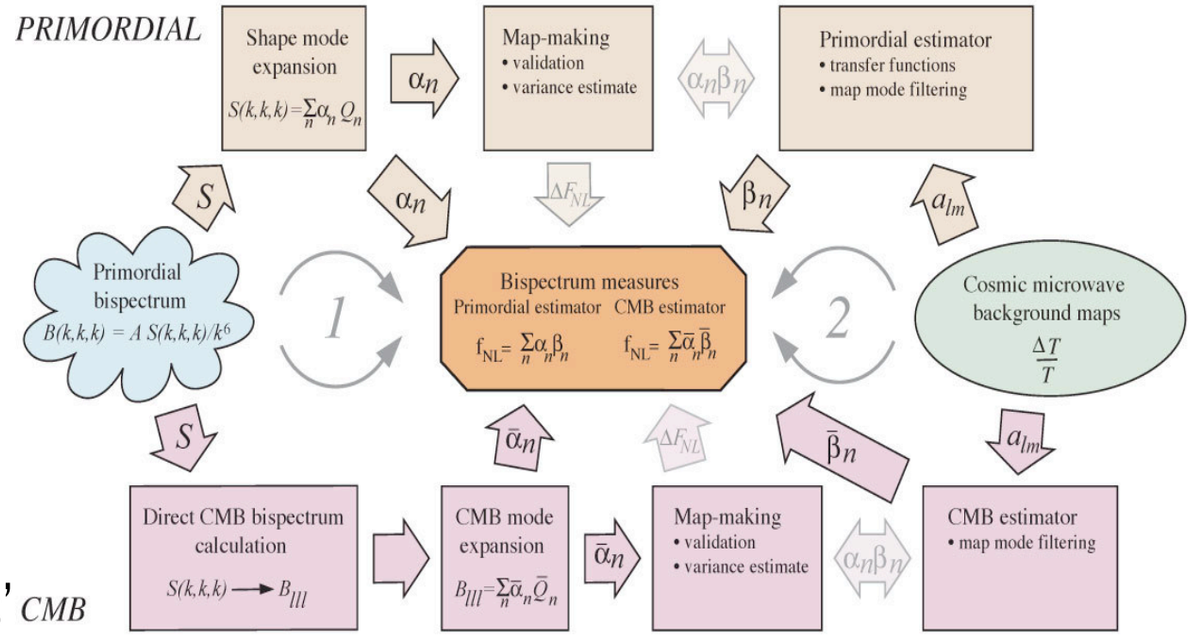


Before Planck (left) and After Planck (right)



3D Planck bispectrum - Fourier (left) & polynomial basis (right)

- Discovery of lensing bispectrum and 'hints' of new physics being investigated further



Complex Planck CMB pipeline for three-point correlator

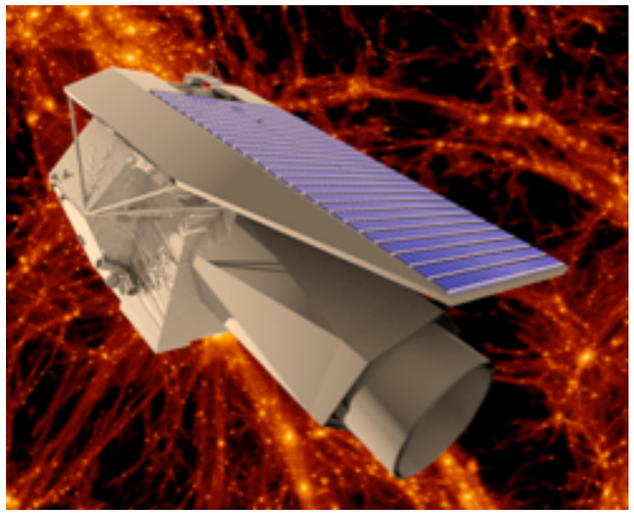
COSMOS Data-driven Discovery

Primary science driver: Advancing the confrontation between new cosmic datasets (e.g. Planck, Euclid) and precision theoretical modelling.

Competitive science exploitation requires powerful and flexible HPC platforms for rapid pipeline development: new algorithms, prototyping and post-processing. *Maximise discovery potential.*

Monte-Carlo simulations of the Universe needed in vast quantities for accurate statistical analysis of experimental and systematic effects, with requirements dwarfing the final science data products! *BigData solutions needed.*

The Dark Energy Survey is currently mapping 300 million galaxies



ESA Euclid satellite (construction began July 2013)

Data storage virtualization: Key constraint as cosmic datasets grow

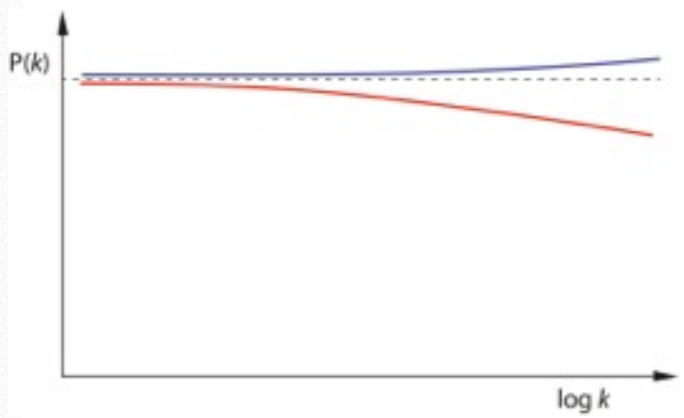
- ◇ 2D Planck temperature and polarization analysis (x3)
- ◇ 3D galaxy surveys, e.g. Dark Energy Survey (x100) and then Euclid
- ◇ Black hole collisions and gravitational waveforms (LIGO, eLISA)

Cool data: Disks full with necessary, but infrequently accessed, data

The "BigData without BigData" paradigm: Developing end-to-end analysis pipelines which hold real and simulated data products in (shared-)memory, using real-time MC generation and accelerating compute 'hotspots' on coprocessors - thus minimising IO/storage

Confrontation of theory & observation

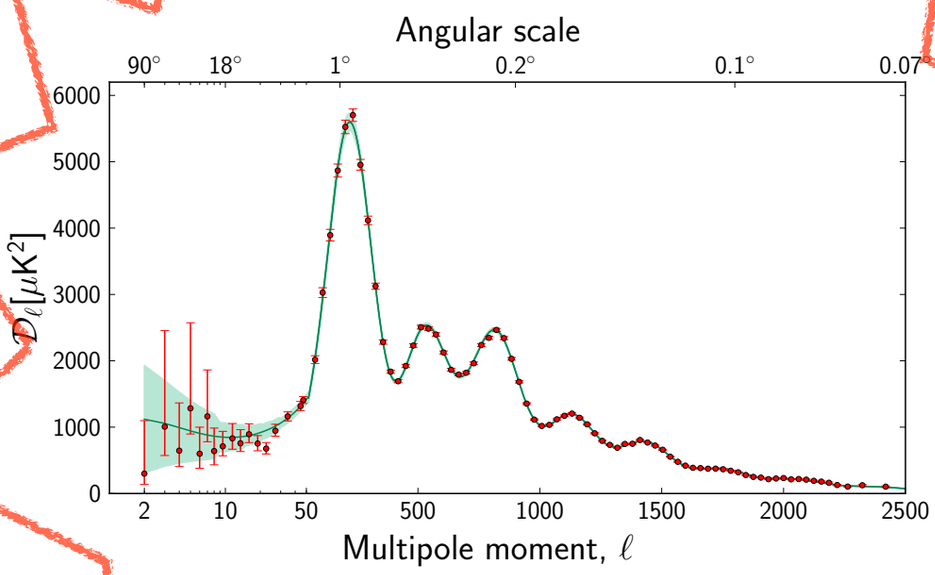
Primordial power spectrum



ENLIGHTEN

Plasma physics of the standard cosmology e.g. CAMB

CMB angular power spectrum



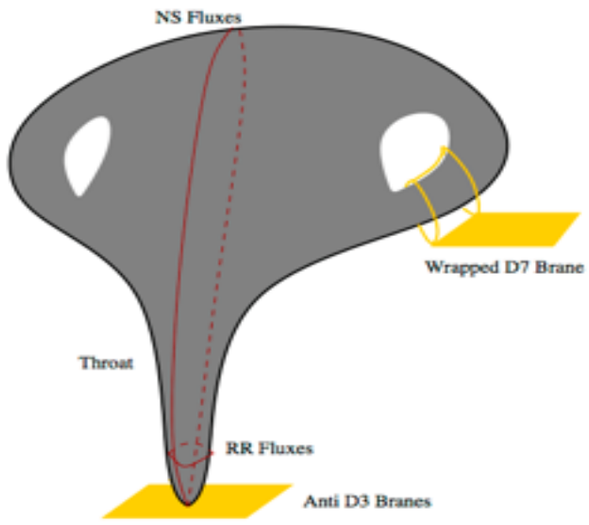
EUREKA

Modelling inflation and early universe dynamics e.g. WALLS

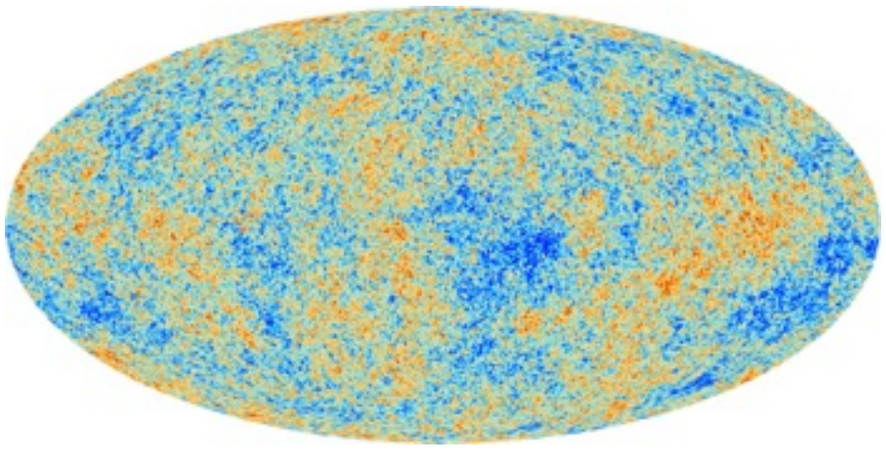
New insights or constraints

EXPOSE

Determining signal power on different angular scales e.g. MODAL



Fundamental Theory (Inflationary Model)



Observational data (Planck CMB Maps)

Intel Parallel Computing Centre

SGI collaboration since 1997

- early access to first SMP
- Origin2000, Altix 3000, UV
- Design/feedback interaction

Intel collaboration since 2003

- early access Itanium, Xeon Phi
- parallel programmer support

Vanguard IPCC status (announcement April 2014)

- COSMOS Parallel Programmer support - James Briggs (approx. 60%)
- Intel applications support for KNC - John Pennycook (50%)
- Direct links to Intel engineering teams (notably offload), priority IPS
- Access to Xeon Phi KNL simulator now
- Early access Xeon Phi KNL system 2015/16 (when available)
- *Partners: Joe Curley, Jim Jeffers (Intel), Karl Feind, Mike Woodacre (SGI)*

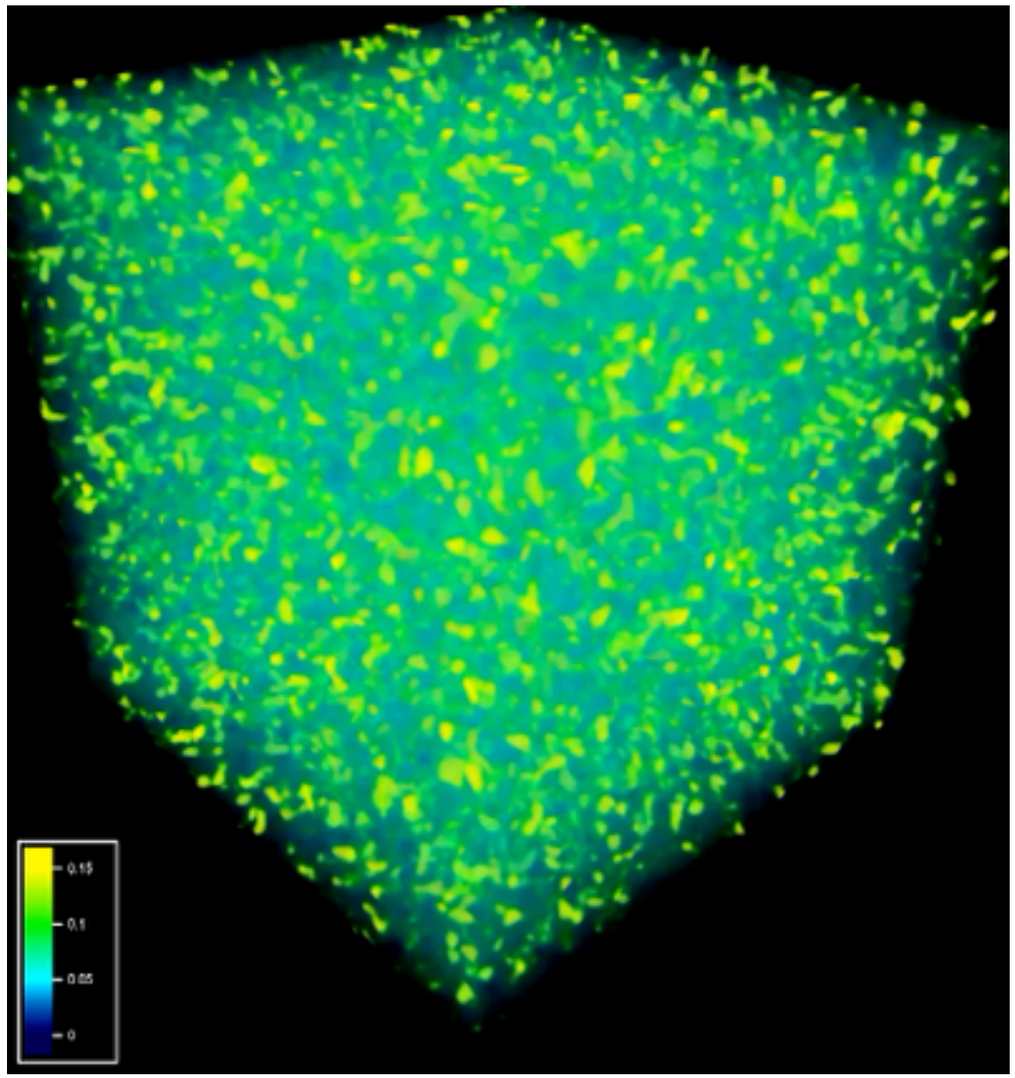
Early Universe - WALLS

Code by Martins: Reporting MIC optimisation work by James Briggs & John Pennycook

- Simulates the evolution of domain wall networks in the early universe.
- To find out more about domain walls - see CTC Public Outreach pages:

  www.ctc.cam.ac.uk/outreach/origins/cosmic_structures_two.php

- Used at COSMOS for 10 years to study hybrid networks
- Possible contribution to BICEP result!?
- Pure SMP code using OpenMP
  - Simple leapfrog time step algorithm
    Calculate area of domain walls
  - Compiled for 3D or 4D simulations
  - Periodic boundary conditions
- Benchmark for acceptance testing on previous machines

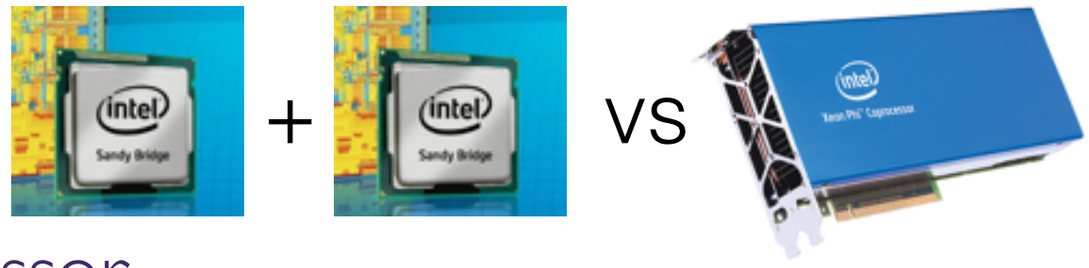


WALLS baseline

Experimental Setup: 480^3 problem

2 x Intel® Xeon® E5-4650L processor

vs. 1 x Intel® Xeon Phi™ 5110P coprocessor

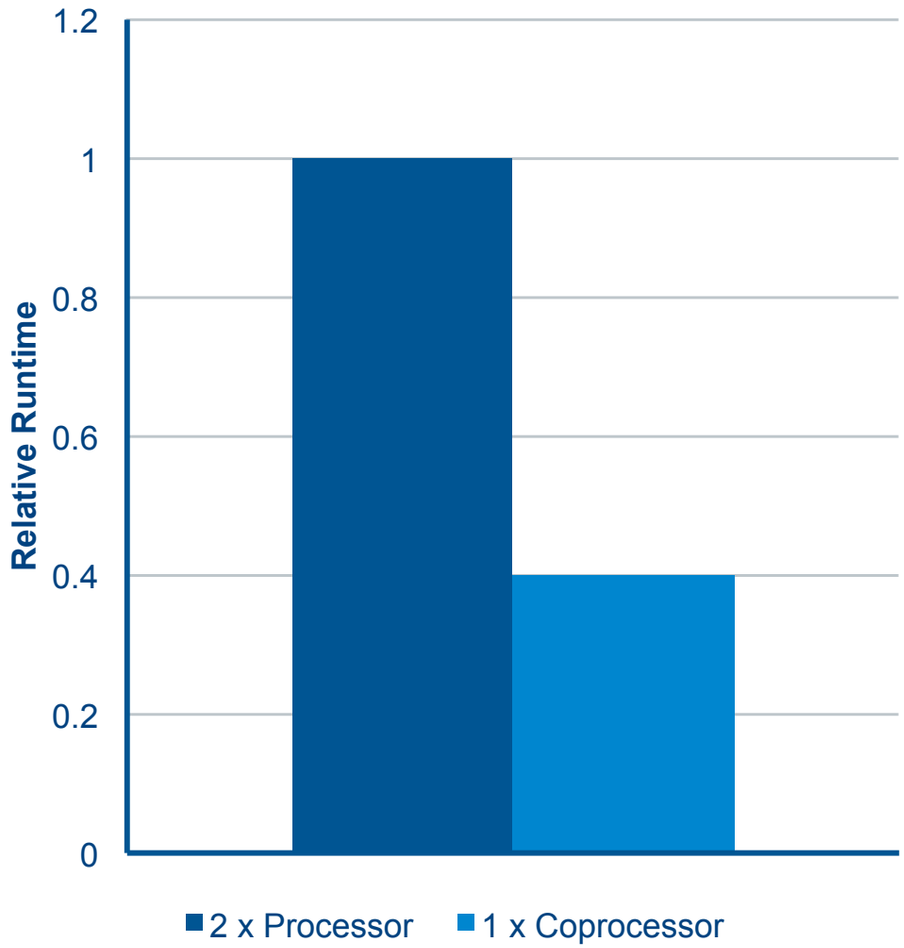


"Out-of-the-Box" Comparison:

Porting = recompile with -mmic

Processor is over 2x faster than coprocessor!

Why? Poor vectorisation & poor memory



The "Challenge":

Optimize and modernize the code

No "ninja" programming

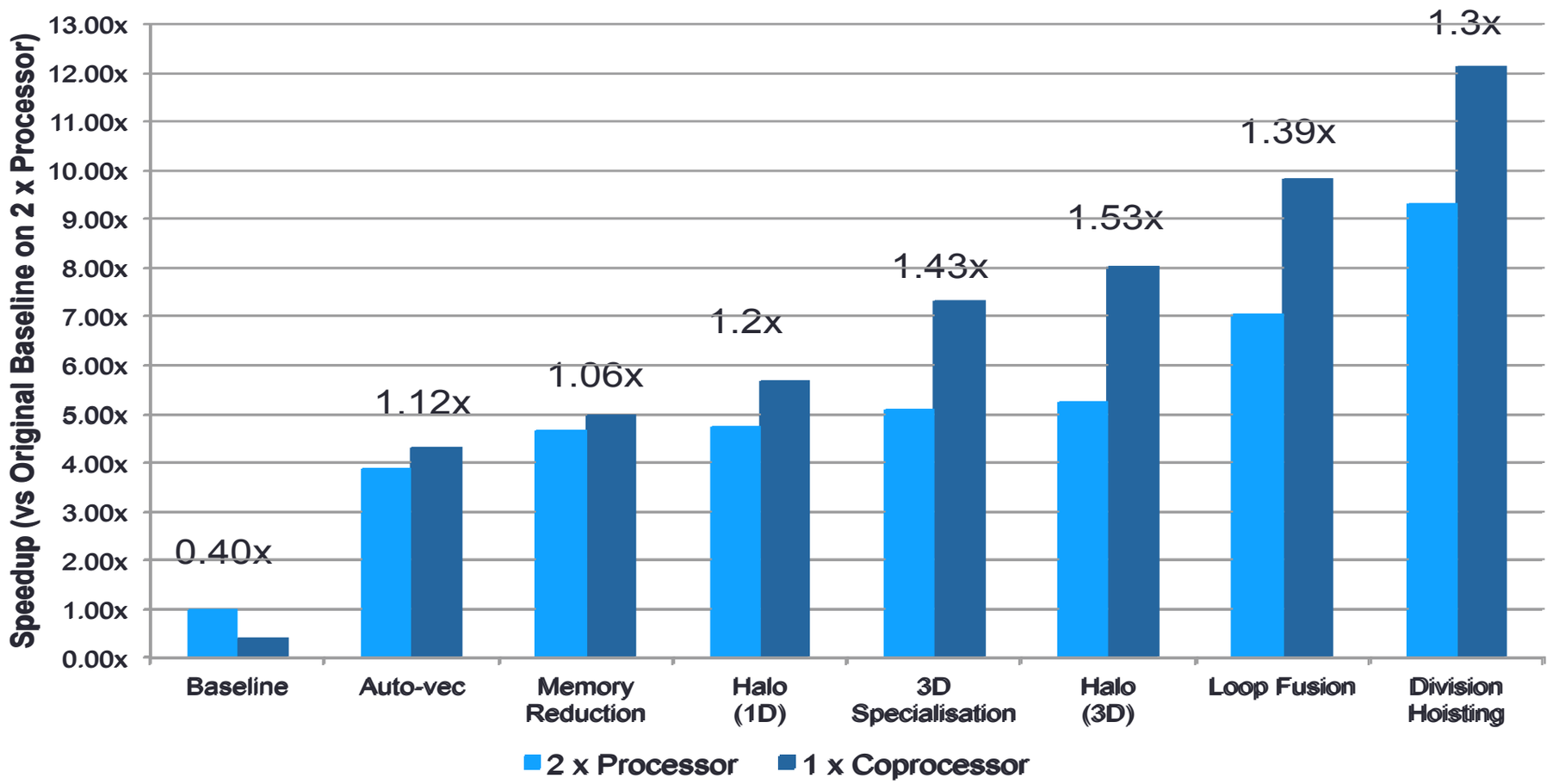
The Result:

Significant performance improvements in ~3-4 weeks

Clear, readable code-base

"Template" stencil code transferable to other simulations

WALLS Xeon Phi results

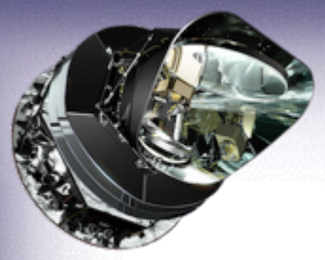


Observations IV: (Intel whitepaper drafted)

- Straightforward code changes can have dramatic impact
- Both Xeon and Phi speed-up from the same changes
- Coprocessors benefit more (30x) than processors (9x)

Future work: • Transfer to other 3D codes

- Using offload runtime to stream large problems through one coprocessor
- Sharing between multiple Xeons & Phis



Planck Non-Gaussianity - MODAL

Code by Fergusson: Reporting MIC optimisation work by James Briggs & John Pennycook

Hybrid OpenMP/MPI generalising CAMB for non-Gaussian theories

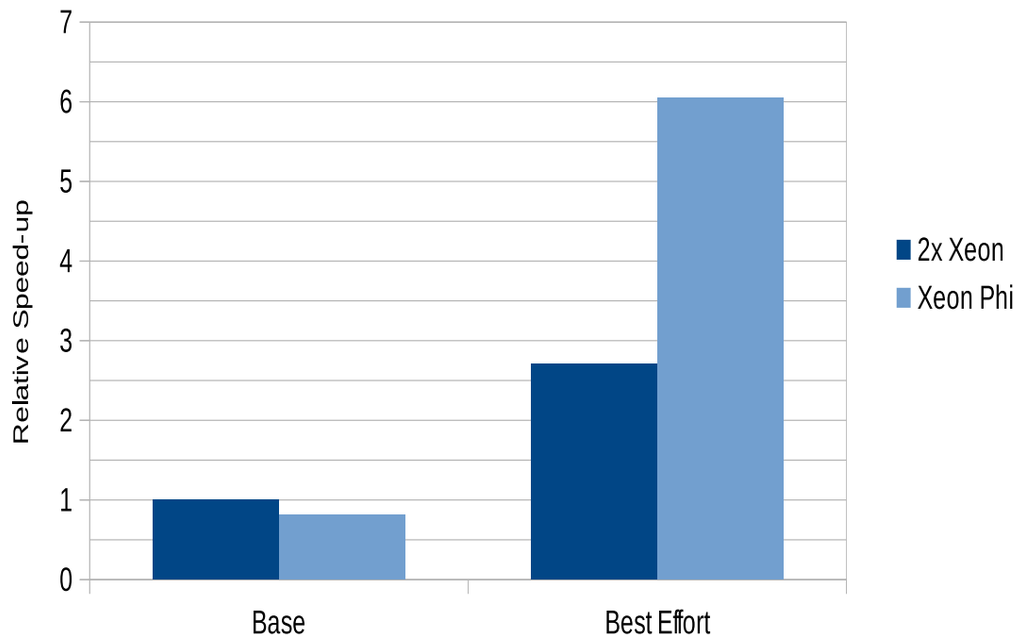
Part of key non-Gaussian pipeline for Planck satellite data analysis

Repeated integrations of early and late-time basis functions (2D Gauss-Legendre)

Improvement - unrolling loops to allow auto-vectorization (7x speed-up on MIC)

Stage	2x Xeon Time (s)	2x Xeon Speed-up	MIC Time (s)	MIC Speed-up	MIC vs 2x Xeon Speed-up
Base	453.5	1	560.7	1	0.8
Unroll/vector	171.3	2.64	79.5	7.05	2.15
Alignment	171.0	2.65	78.0	7.18	2.19
Maths Tweak	167.7	2.70	75.0	7.48	2.24

Speed-up versus Xeon Base Case



Timescales:

3x on Xeon from loop unrolling and data align. = 1 hour

2x on Xeon Phi from OpenMP = 1 week

Final Result:

Single Xeon Phi = 4.5 Xeon Sandybridge

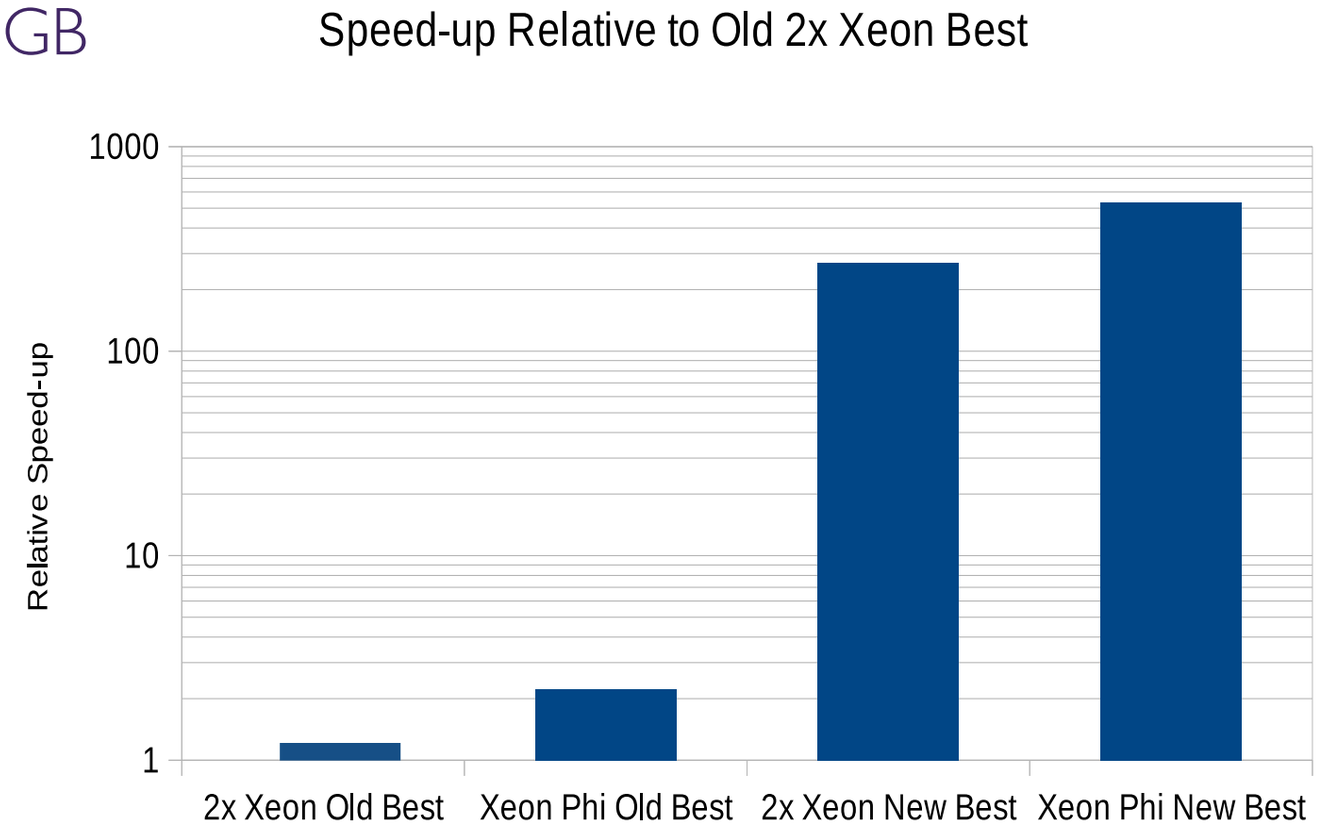
New algorithms: MODAL frontiers

During optimisation, identified hidden repetition
- can precompute look-up array of size 1GB
- enabled through rapid implementation

730x speed-up over optimised 2 Xeons

4600x speed-up over original Xeon Phi

Serendipity …



Future work:

Port CMB 3D integration code to MIC

Apply to large datasets for large-scale structure modal methodology (galaxy surveys)

Step-change:

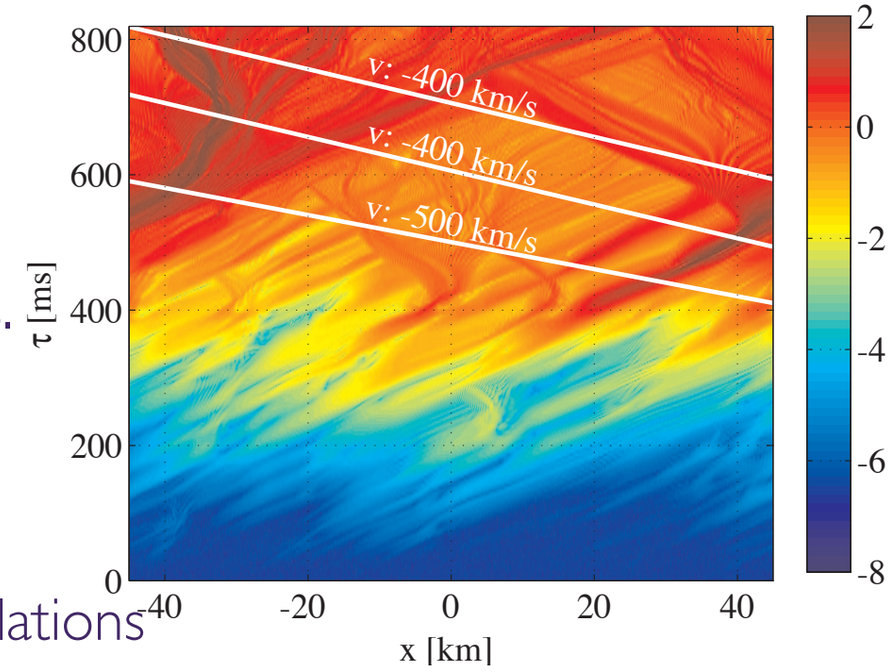
Feasibility joint Markov Chain Monte Carlo analysis with 2-point and 3-point corrs.

The importance of programming support demonstrated - from vendors to the HPC ecosystem
- showing huge benefits to HPC productivity and to advancing core science goals

SGI collaboration: Exoplanets

Code by Yurchenko: Reporting MIC optimisation work by Cheng Liao (SGI) and James Briggs

ExoMol exoplanets project lead by Jonathan Tennyson and Sergey Yurchenko of UCL.



Aim:

- Create a large database of simulated molecular line lists.
- Input to atmospheric models of exoplanets and cool stars.
- Use to search for exoplanets with water & organic molecules.

(see: http://www.exomol.com/)

SMP Pipeline:

- OpenMP code performs complex quantum-mechanical calculations which generates a *large matrix between $50K^2$ and $1M^2$ dense.*
- Tailored algorithm difficult to rewrite to a distributive model.
- Solve matrix for eigenvalues and eigenvectors which represent molecular rotation-vibration states.
- Unify pipeline into a single program to reduce runtime, IO, storage requirement, and general user time.

Required scalable SMP eigensolver:

But poor scaling found for LAPACK DSYEV(+D) with multi-threaded BLAS (for Intel MKL, NAG, & OpenBLAS)

TROVE Benchmark on COSMOS

PLASMA library: DSYTRD solver with tridiagonal solver replaced faster MR3-SMP library.

- 'Tiled' linear algebra – task based, very cache friendly.
- Customized the memory page placement for NUMA.
- Customized thread counts and thread placement for different parts of the calculation.
- Written an Autotuning harness for thread counts and tile sizes.

Performance data: Trove with ScaLAPACK vs PLASMA+MR3SMP.

Matrix Dimension	#Cores	Runtime pdsyev	Runtime SGI PLASMA	Ratio
16k	40	95s	66s	1.43x
32k	40	712s	397s	1.79x
130k	64	~7hr	~5hr	1.4x
250k	256	~30hr	~18hr	1.66x

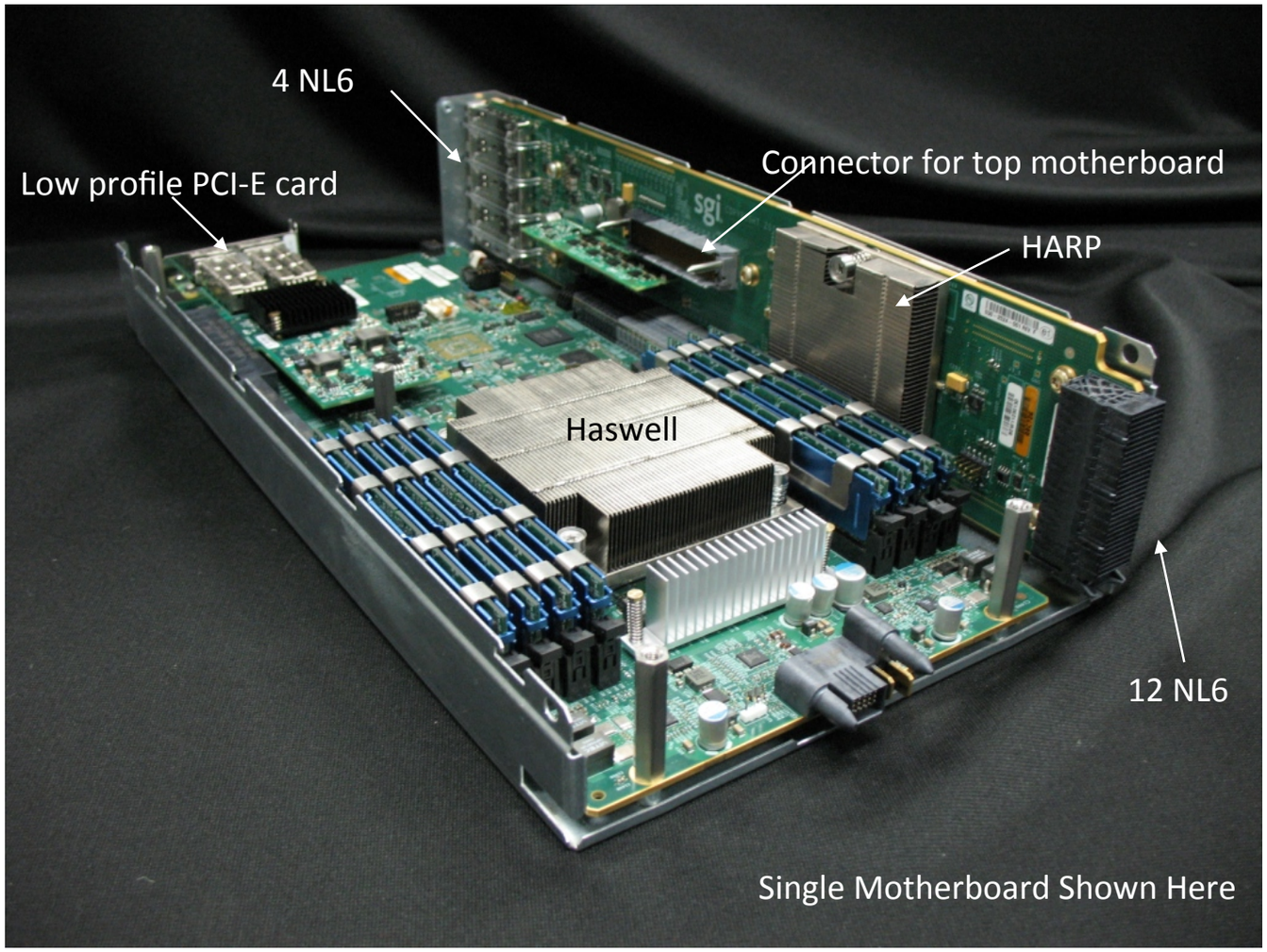
Observations:

- Trove an exemplar of BigData pipelines rapidly created in SMP
- Many applications for LaPACK and other SMP libraries - high efficiency possible
- Shared-memory is very important - Intel take physical address bits beyond 46!

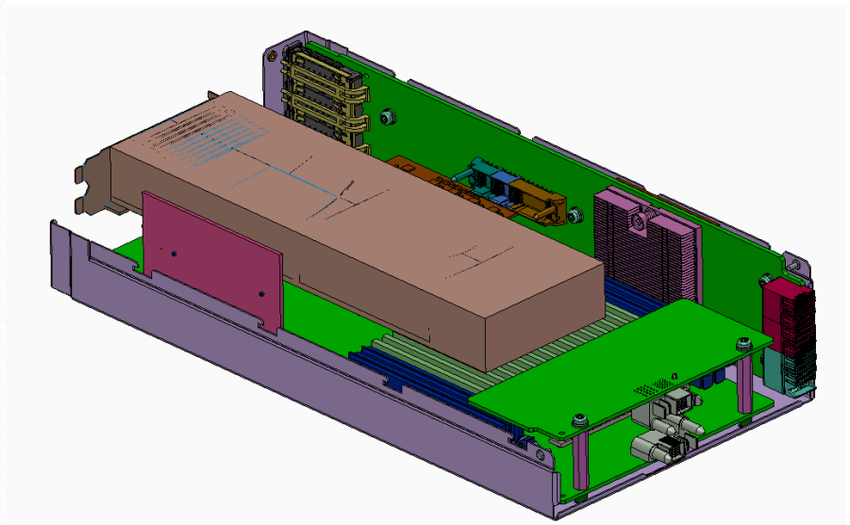
SGI co-design: UV MG Blade

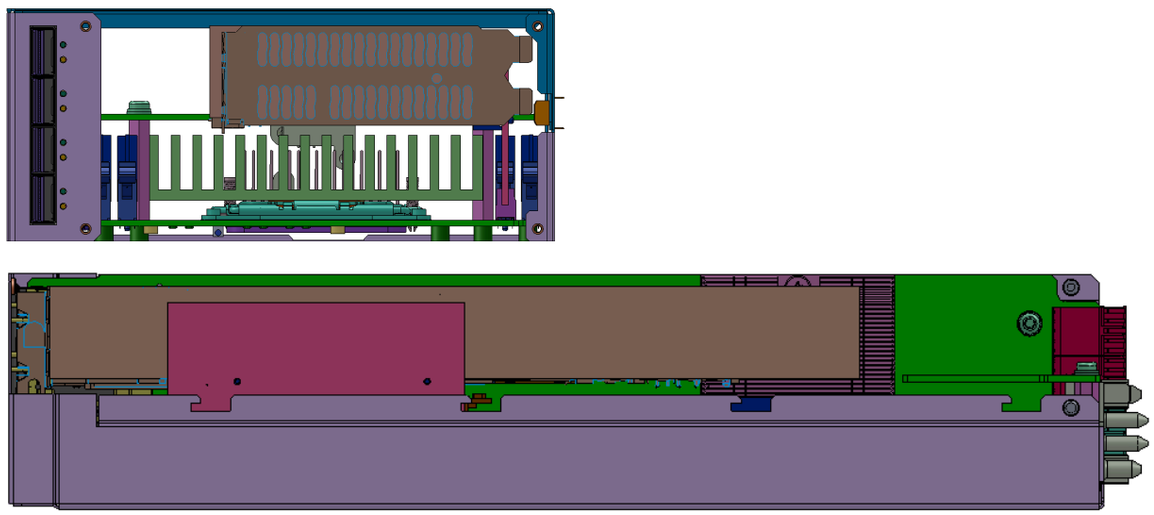
SGI UV architecture

Blade configured in 16 socket/8 blade IRU



SGI UV/MIC architecture defined by COSMOS UV2 needs = MG blade





DiRAC-3 UV3 KNL Integration?

KNL key features:

- 2 PCIe x16 connections
- External DIMMs for additional capability
- Desired ratio of network bandwidth …
- Programming usage profile for applications of multiple-KNLs …

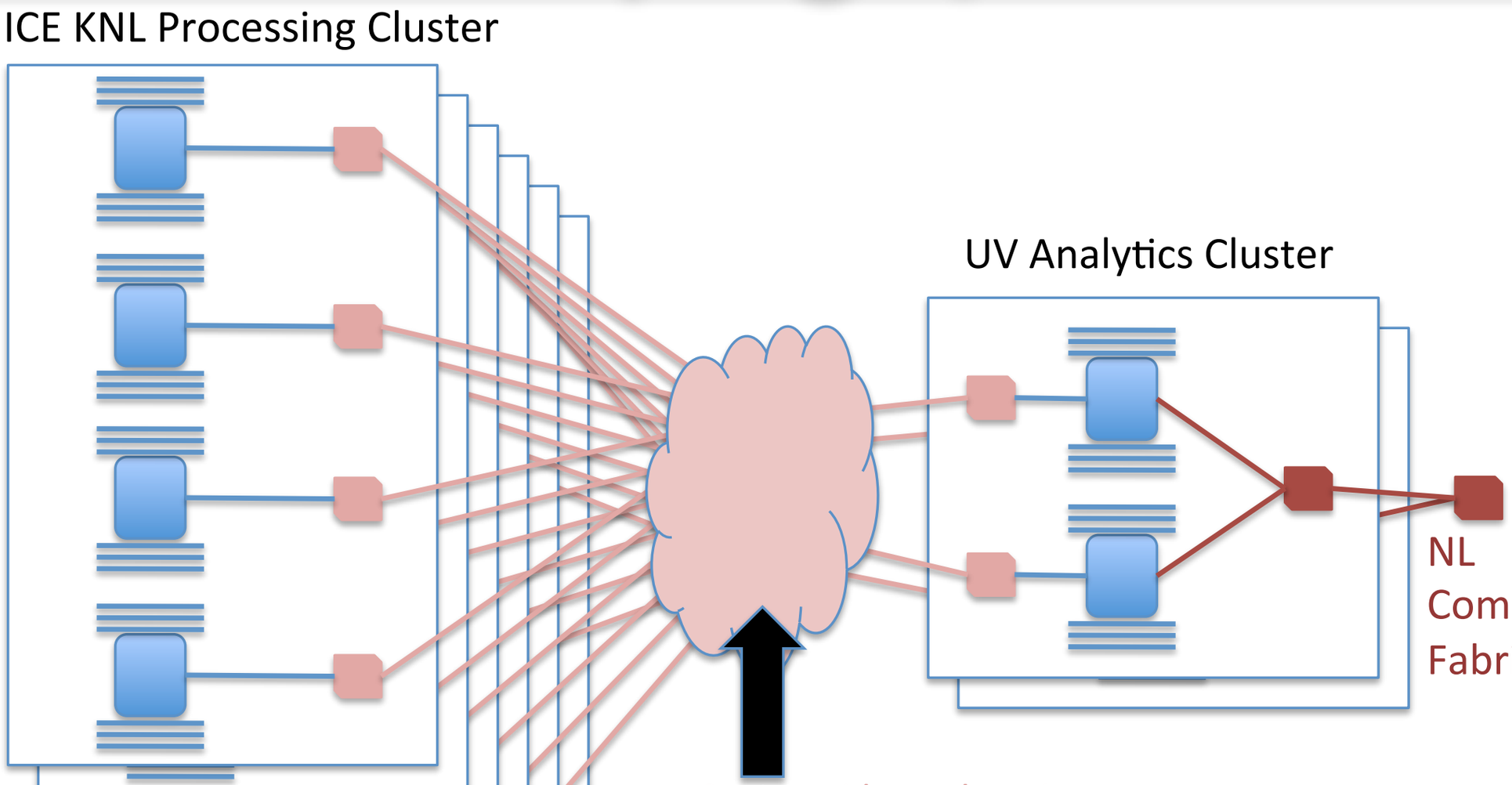
Option 1: KNL integrated similar to KNC on UV2

- Using standard PCIe card on MG-blade

Option 2: Multiple KNL on PCIe/Infiniband/Intel fabric, connect to UV

- Enables direct communications between multiple KNL skts direction on the fabric of choice without going into UV/NL fabric
- Enables offload over fabric with Intel MPSS support – potential to offload from UV Xeon to KNL or from KNL to UV-Xeon

Fabric coupling of KNL and UV



UV Single System Image supports up to 64TB and 4096 cores within a 4-rack, 256-skt NUMAlink connected system

Most flexible HPC platform in industry: 3-in-1 system supporting SMP, offload acceleration, and/or MPI KNL

COSMOS Programming Paradigms

Support for flexible paradigms to maximise competitiveness and inclusivity:

- welcoming and offering traction to new users whatever their HPC background
- rapid prototyping, code development, post-processing with limited programmer resources
- hybrid SMP exemplar pipelines for data analysis and science exploitation (Trove & Planck)
- unified vision for end-to-end SMP analysis pipelines deploying offload to coprocessors

IPCC Code development and porting to MIC architectures with Intel/SGI (co-design):

- codes can be easily ported to MIC, allowing rapid assessment (e.g. CAMB)
- optimisation efforts can yield very substantial speed-ups for both Xeon and Phi
- good MIC performance achieved for WALLS and MODAL (ongoing COSMOS/DiRAC codes)

The importance of programming support demonstrated

EU extreme scales network example: CMB Planck and beyond

E.g. Cambridge - George Efstathiou, Anthony Challinor, EPS; Sussex - Hindmarsh, Lewis
Paris - Julien Lesgourges, Geneva - Durrer, Kunz; Padua - Liguori, Matarrase
Portsmouth/UCL (DES, LSST, Euclid) etc

and/or Black hole dynamics …