Big Bang, Big Data, Big Iron

High Performance Computing and
the Cosmic Microwave Background

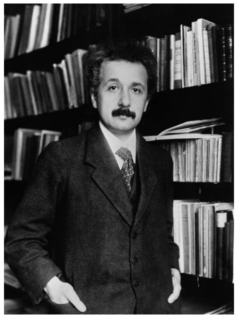
Julian Borrill
Computational Cosmology Center, LBL
Space Sciences Laboratory, UCB

and the
BOOMERanG, MAXIMA, Planck, EBEX & PolarBear collaborations

The Cosmic Microwave Background

Cosmologists are often in error
but *never* in doubt.

1916 – General Relativity



- General Relativity
  - Space tells matter how to move
  - Matter tells space how to bend

$$\underset{\textit{Space}}{G_{\mu\nu}} = 8 \pi G \underset{\textit{Matter}}{T_{\mu\nu}}$$

- But this implies that the Universe is dynamic, and everyone *knows* it's static …
- … so Einstein adds a Cosmological Constant (even though the result is unstable equilibrium)

1929 – Expanding Universe

- Using the Mount Wilson 100-inch telescope Hubble measures nearby galaxies'
  - velocity (via their redshift)
  - distance (via their Cepheids)

  and finds

$$\boxed{v \propto d}$$

- Space is expanding!
- The Universe is dynamic after all.
- Einstein calls the Cosmological Constant "my biggest blunder".

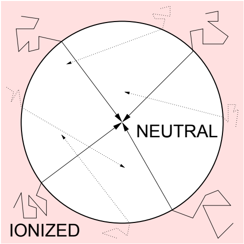
1930-60s – Steady State vs Big Bang

- What does an expanding Universe tells us about its origin and fate?
  - Steady State Theory:
    - new matter is generated to fill the space created by the expansion, and the Universe as a whole is unchanged and eternal (past & future).
  - Big Bang Theory:
    - the Universe (matter and energy; space and time) is created in a single explosive event, resulting in an expanding and hence cooling & rarifying Universe.

1948 – Cosmic Microwave Background

- In a Big Bang Universe the expanding Universe eventually cools through the ionization temperature of hydrogen: $p^+ + e^- \Rightarrow H$.
- Without free electrons to scatter off, the photons free-stream to us today.
- Alpher, Herman & Gamow predict a residual photon field at 5 – 50K

- COSMIC – filling all of space.
- MICROWAVE – redshifted by the expansion of the Universe from 3000K to 3K.
- BACKGROUND – primordial photons coming from "behind" all astrophysical sources.



1964 – CMB First Detection

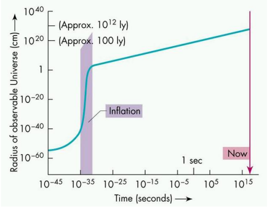
- While trying to zero a Bell Labs radio telescope, Penzias & Wilson found a puzzling residual signal that was constant in time and direction.
- They determined it wasn't terrestrial, instrumental, or due to a "white dielectric substance", but didn't know what it was.
- Meanwhile Dicke, Peebles, Roll & Wilkinson were trying to build just such a telescope in order to detect this signal.
- Penzias & Wilson's accidental measurement killed the Steady State theory and won them the 1978 Nobel Prize in physics.

© 2004 Thomson - Brooks/Cole

1980 – Inflation

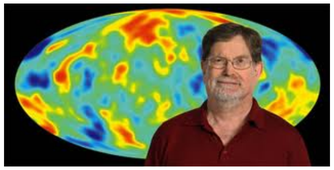
- Increasingly detailed measurements of the CMB temperature showed it to be uniform to better than 1 part in 100,000.
- At the time of last-scattering any points more than 1° apart on the sky today were out of causal contact, so how could they have exactly the same temperature? This is the horizon problem.

- Guth proposed a very early epoch of exponential expansion driven by the energy of the vacuum.
- This also solved the flatness & monopole problems.



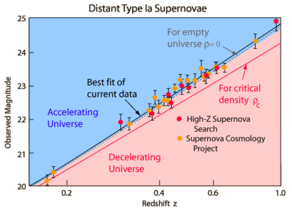
1992 – CMB Fluctuations

- For structure to exist in the Universe today there must have been seed density perturbations in the early Universe.
- Despite its apparent uniformity, the CMB must therefore carry the imprint of these fluctuations.
- After 20 years of searching, fluctuations in the CMB temperature were finally detected by the COBE satellite mission.
- COBE also confirmed that the CMB had a perfect black body spectrum, as a residue of the Big Bang would.
- Mather & Smoot share the 2006 Nobel Prize in physics.



1998 – The Accelerating Universe

- Both the dynamics and the geometry of the Universe were thought to depend solely on its overall density:
  - Critical ($\Omega_{total}=1$): expansion rate asymptotes to zero, flat Universe.
  - Subcritical ($\Omega_{total}<1$): eternal expansion, open Universe.
  - Supercritical ($\Omega_{total}>1$): expansion turns to contraction, closed Universe.
- Measurements of the brightness and distances of supernovae surprisingly showed the Universe is accelerating!
- Acceleration (maybe) driven by a Cosmological Constant!
- Perlmutter and Riess & Schmidt share 2011 Nobel Prize in physics.



2000 – The Concordance Cosmology

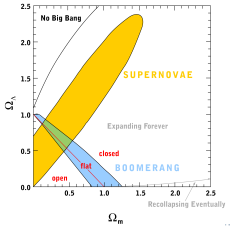
- The BOOMERanG & MAXIMA balloon experiments measure small-scale CMB fluctuations, demonstrating that the Universe is flat.
- The CMB fluctuations encode cosmic geometry ($\Omega_\Lambda + \Omega_m$)
- Type 1a supernovae encode cosmic dynamics ($\Omega_\Lambda - \Omega_m$)
- Their combination breaks the degeneracy in each.

The Concordance Cosmology:

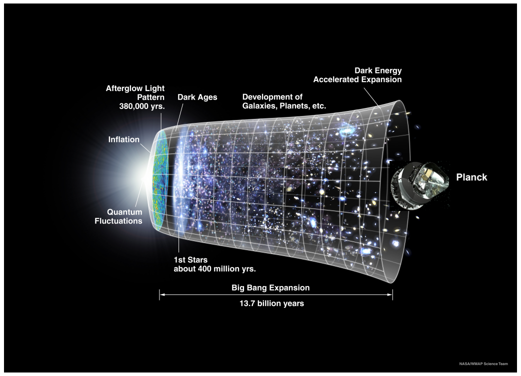
- 70% Dark Energy + 25% Dark Matter + 5% Baryons

  => 95% ignorance!

- What and why is the Dark Universe?



A History Of The Universe

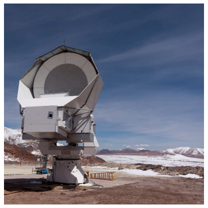


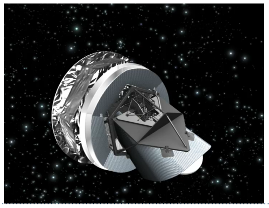
CMB Science

- Primordial photons experience the entire history of the Universe, and everything that happens leaves its trace.
- Primary anisotropies:
  - Generated before last-scattering, encode all physics of the early Universe
    - Fundamental parameters of cosmology
    - Quantum fluctuation generated density perturbations
    - Gravity waves from Inflation
- Secondary anisotropies:
  - Generated after last-scattering, encode all physics of the later Universe
    - Gravitational lensing by dark matter
    - Spectral shifting by hot ionized gas
    - Red/blue shifting by evolving potential wells

CMB Observations

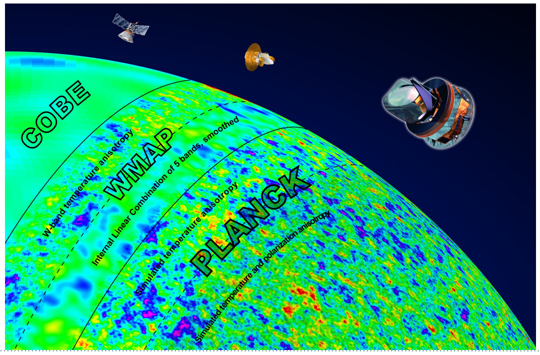
- We are searching for micro- to nano-Kelvin fluctuations on a 3 Kelvin background.
- We need very many, very sensitive, very cold, detectors.
- Scan part of the sky from high dry ground or the stratosphere, or all of the sky from space.





CMB Science Evolution

Evolving science goals require higher resolution & polarization sensitivity.

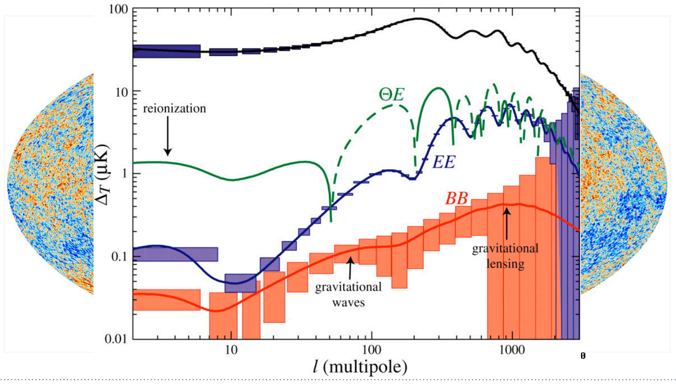


Cosmic Microwave Background Data Analysis

CMB Data Analysis

- A sequence of changes of basis that
    - Reduce the data volume
    - Increase the signal-to-noise
    - Facilitate the removal of systematics
    - Provide a point of comparison with theoretical predictions
- Bases
    - Time-domain: noise-dominated detector samples
    - Frequency maps: foreground-contaminated sky pixels
    - CMB map: single realization of statistical process
    - Angular power spectra: compare with theory predictions

CMB Data Compression



Ideal CMB Analysis – Formalism

- Model data as stationary Gaussian noise and sky-synchronous CMB signal

$$d_t = n_t + P_{tp}\, s_p$$

- Estimate the inverse noise correlations from the (noise-dominated) data

$$N_{tt'}^{-1} = f(|t-t'|) \sim \mathrm{invFFT}(1/\mathrm{FFT}(d))$$

- *Analytically* maximize a Gaussian likelihood for the map given the data

$$m_p = (P^T N^{-1} P)^{-1} P^T N^{-1} d$$

- Construct the pixel domain noise covariance matrix

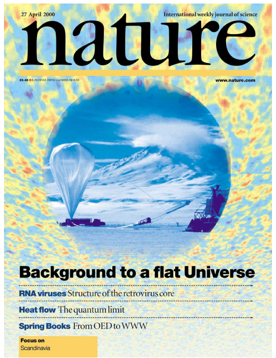
$$N_{pp'} = (P^T N^{-1} P)^{-1}$$

- *Iteratively* maximize a Gaussian likelihood for the CMB power spectrum given the map and its total covariance matrix M = S(c) + N

$$L(c_l \mid m) = -½\,(m^T M^{-1} m + \mathrm{Tr}[\log M])$$

Ideal CMB Analysis – Execution

- Implementation is dominated by dense matrix operations
    - inversion in building $N_{pp'}$
    - multiplication in estimating c_l
- MADCAP software built on ScaLAPACK tools, Level 3 BLAS
    - Scales as $\mathcal{N}_p^3$
- Execution on NERSC's 600-core Cray T3E
    - Achieves ~90% theoretical peak performance
- Spawns MADbench scientific benchmark and procurement software



But …

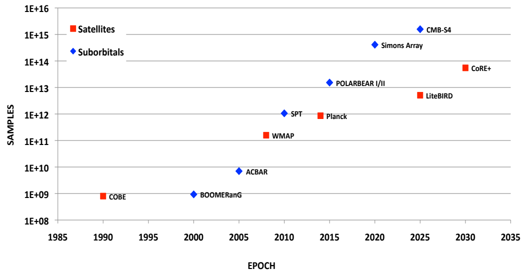
- BOOMERanG:
  - 2,500 sq-degrees at 20 arc-minute resolution in at 1 frequency in temperature only.
- Planck:
  - 40,000 sq-degrees at 5 arc-minute resolution at 9 frequencies in temperature and 2 polarization modes.
- 16x sky coverage, 16x resolution, 9x frequencies, 3x components
  - ➢ O(10^4) increase in $\mathcal{N}_p$
  - ➢ O(10^{12}) increase in operation count
    - Moore's Law provides 1000-fold increase every 15 years
    - We can't wait 60 years for Planck!

Approximate CMB Analysis

- Map-making
  - No explicit noise covariance calculation possible
  - Use PCG instead: $(P^T N^{-1} P)\, m = P^T N^{-1} d$
- Power-spectrum estimation
  - No explicit data covariance matrix available
  - Use pseudo-spectral methods instead:
    - Take spherical harmonic transform of map, simply ignoring inhomogeneous noise, cut-sky!
    - Use Monte Carlo methods to estimate uncertainties and remove bias.
- Dominant cost is now simulating & mapping time-domain data: O($\mathcal{N}_t$)

Time-Domain CMB Data Growth

- The only way to detect fainter signals is to take more samples.
- Exponential data growth for the past and coming 20 years
  - Have to track Moore's Law, however that is achieved.



Simulation & Mapping: Calculations

Given the instrument noise statistics & beams, a scanning strategy, and a sky:

1) SIMULATION: $d_t = n_t + s_t = n_t + P_{tp}\, s_p$
   - A realization of the piecewise stationary noise time-stream:
     - Pseudo-random number generation & FFT
   - A signal time-stream scanned & from the beam-convolved sky:
     - SHT
2) MAPPING: $(P^T N^{-1} P)\, d_p = P^T N^{-1} d_t$ $(A\, x = b)$
   - Build the RHS
     - FFT & sparse matrix-vector multiply
   - Solve for the map
     - PCG over FFT & sparse matrix-vector multiply

Simulation & Mapping: Scaling

- In theory such analyses should scale
  - Linearly with the number of observations.
  - Perfectly to arbitrary numbers of cores.
- In practice this does not happen because of
  - IO (reading pointing; writing time-streams
    reading pointing & time-streams; writing maps)
  - Communication (gathering partial maps from all processes)
- For each new architecture (and often concurrency) the *relative* costs of calculation, communication and I/O change.
- Moore's Law is a constantly moving target!

I/O Details

- Time-ordered data from all the detectors are load-balanced over the processes.
- Each process therefore reads/writes only its samples
  - Detector data are densely sampled per detector
  - Pointing data are
    - Initially sparse-sampled for the instrument boresight
    - Then
      - Interpolated to dense sampling
      - Rotated to each detector's reference frame
- Maps are read/written by a single process.

IO – Before

```
For each MC realization
    For each detector                          ┐
        Read detector pointing                 ├ Sim
        Write detector time-stream             ┘
    For all detectors                          ┐
        Read detector time-stream & pointing   ├ Map
    Write map                                  ┘
```

⇒ Read: Realizations x Detectors x Observations x 2
Write: Realizations x (Detectors x Observations + Pixels)

E.g. for Planck read 500PB & write 70PB.

IO – Optimizations

- Read sparse telescope pointing instead of dense detector pointing
  - Calculate individual detector pointing on the fly.
- Remove redundant write/read of time-streams between simulation & mapping
  - Generate simulations on the fly only when map-maker requests data.
- Put MC loop inside map-maker
  - Amortize common data reads over all realizations.

IO – After

```
Read telescope pointing
For each detector
    Calculate detector pointing
For each MC realization
    For all detectors
        Simulate time-stream
    Write map
```

SimMap

⇒ Read: Sparse Observations
Write: Realizations x Pixels

E.g. for Planck, read 2GB & write 70TB => 10^8 read & 10^3 write compression.

Communication Details

- Time-ordered data from all the detectors are load-balanced over the processes.
- Each process therefore holds
  - *some* of the observations
  - for *some* of the pixels.
- In each PCG iteration, each process reduces its observations.
- At the end of each iteration, each process needs to
  - Send its results to all processes observing the same pixels.
  - Receive the results from all processes observing the same pixels.

Communication – Before

- Initialize a process & MPI task on every core
- Distribute time-stream data & hence pixels
- For each partial- to full-map reduction
  - Each process zero-pads its partial map to a full map
  - Each process calls MPI_Allreduce(map, world)
  - Each process extracts the pixels of interest to it & discards the rest

Communication – Optimizations

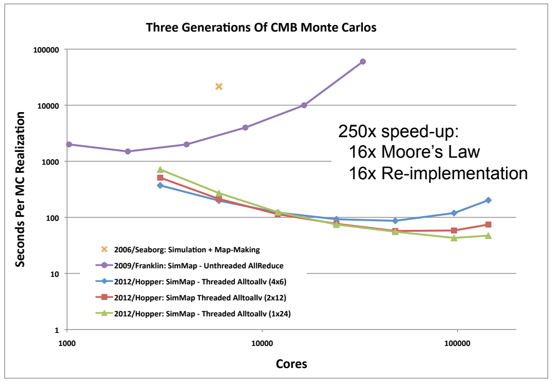
- Reduce the number of MPI tasks
  - Only use MPI for off-node communication
  - Use threads on-node

- Minimize the total volume of the messages
  - Determine all process-pair's pixel overlap
  - If the data volume is smaller, use point-to-point communication of shared pixels instead of global communication of all pixels.

Communication – After

- Initialize a process & MPI task on every node
- Distribute time-stream data & hence pixels
- Calculate common pixels for every pair of processes
- After each PCG iteration
  - If most pixels are common to most processes
    - use MPI_Allreduce(map, world) as before
  - Else
    - Each process prepares its send buffer
    - Call MPI_Alltoallv(sbuffer, rbuffer, world)
    - Each process only receives/accumulates data for its pixels

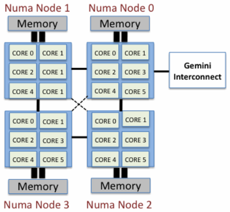
Planck-Sized Simulations Over Time



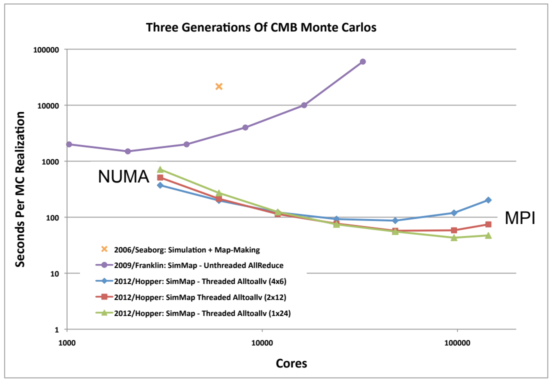
250x speed-up:
16x Moore's Law
16x Re-implementation

Architecture Evolution

- Clock speed is no longer able to maintain Moore's Law.
- Many-core and GPU are two major approaches.
- Both of these will require
  - significant code development
  - performance experiments & auto-tuning
- Eg. NERSC's Cray XE6 system *Hopper*
  - 6384 nodes
  - 2 sockets per node
  - 2 NUMA nodes per socket
  - 6 cores per NUMA node
- What is the best way to run hybrid code on such a system?



Configuration With Concurrency

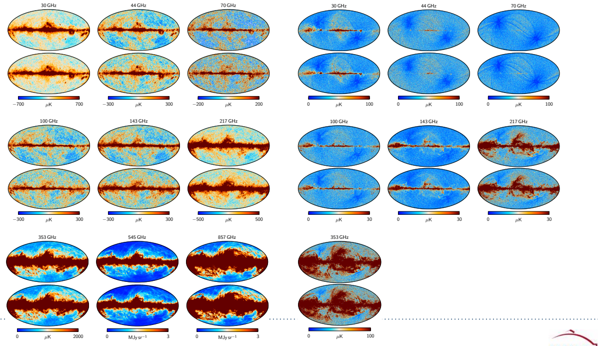


NUMA

MPI

Results: Planck Full Focal Plane 8

- Fiducial mission realization (CMB, foregrounds, noise) to support validation & verification of analysis algorithms & implementations



Results: Planck Full Focal Plane 8

- 10^4-realization CMB and noise Monte Carlo simulation sets reduced to O(10^6) maps to support uncertainty quantification and de-biasing.

Conclusions

- The CMB provides a unique window onto the early Universe
  - investigate fundamental cosmology & physics.
- CMB data analysis is a computationally-challenging problem requiring state of the art HPC capabilities.
- Both the CMB data sets we are gathering and the HPC systems we are using to analyze them are evolving – this is a persistent, dynamic problem.
- The science we can extract from present and future CMB data sets will be determined by the limits on
  a) our computational capability, and
  b) our ability to exploit it.