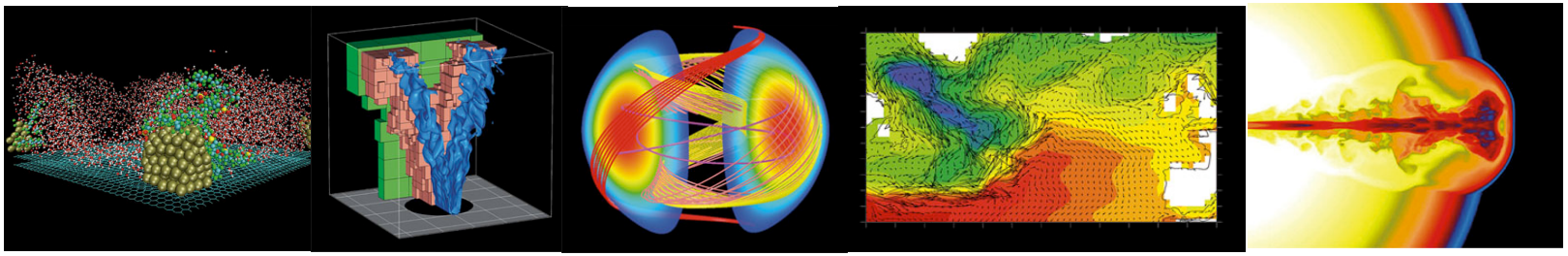


NERSC Workload Analysis

John Shalf
SDSA Team Leader
jshalf@lbl.gov

NERSC User Group Meeting

September 17, 2007

BERKELEY LAB

NERSC

Office of Science
U.S. DEPARTMENT OF ENERGY

Workload Analysis

- Purpose
  - Understand NERSC User Requirements
  - Prioritize software library and tool optimization and deployment
  - Inform NERSC Sustained System Performance (SSP) Benchmark Selection
    - "Benchmarks are only useful insofar as they model the intended computational workload." Ingrid Bucher & Joanne Martin, LANL, 1982
    - Effective performance on SSP to reflect effective performance on NERSC workload

Balancing Requirements

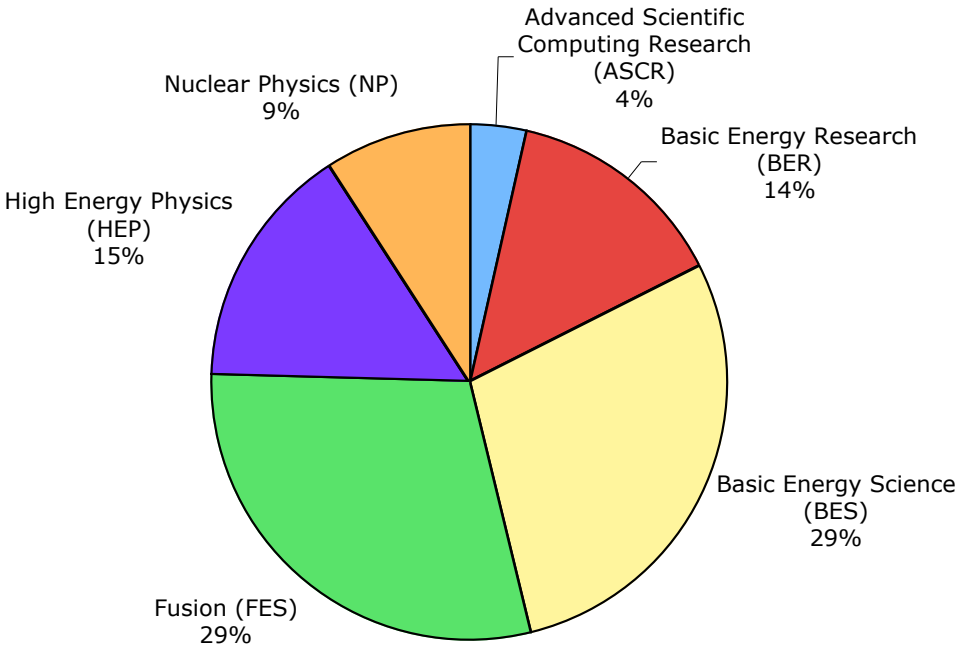
- NERSC Workload overview
  - ~3000 users
  - ~300 projects respresenting a broad range of science
  - ~700 codes (>2 codes per project on average!)
  - 15 science areas for 6 Office of Science divisions
- Select a subset (<10) codes to represent the requirements of the workload
  - Weight based on code's contribution workload? (workload coverage)
  - Weight equally for each area of science? (algorithm/science-area coverage)
- Attempt to cover both dimensions
  - Still daunting
  - Search for islands of coherence in the codes or algorithm selection by different scientific disciplines

Workload Overview

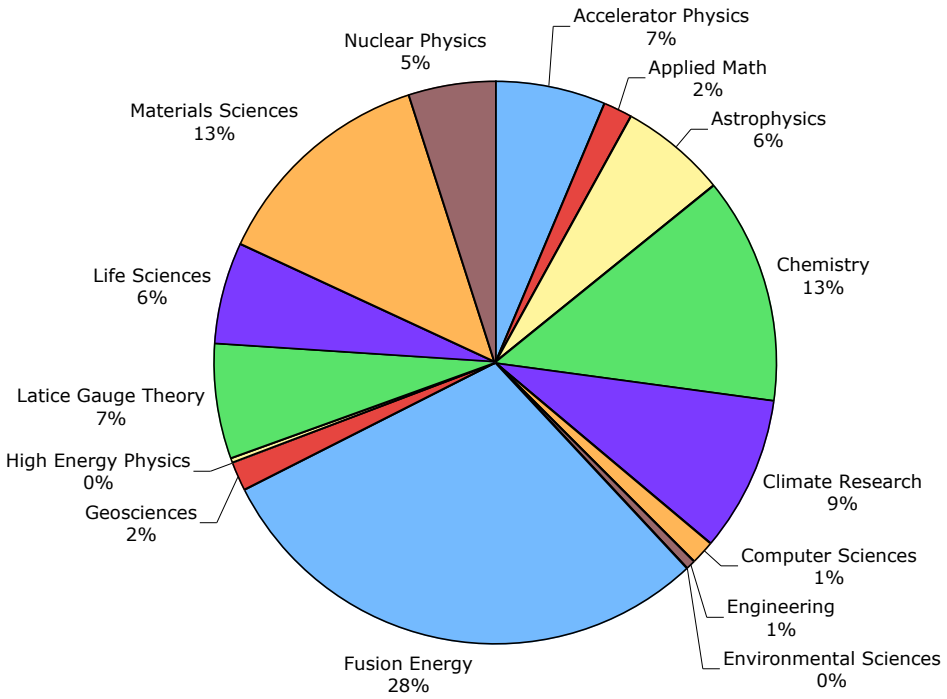
		DOE Ofice						
Science Area	Data	ASCR	BER	BES	FES	HEP	NP	Grand Total
Accelerator Physics	Normalized Sum of MPP Request			10200000		9768000	7580009	11546000
	MPP Request Count			3		32		44
Applied Math	Normalized Sum of MPP Request	10107000		10000				10117000
	MPP Request Count	23		1				24
Astrophysics	Normalized Sum of MPP Request					9888000	2602000	12490000
	MPP Request Count					42	6	48
Chemistry	Normalized Sum of MPP Request	20000		22534000				22554000
	MPP Request Count	1		106				107
Climate Research	Normalized Sum of MPP Request	400000	15066000					15466000
	MPP Request Count	1	44					45
Computer Sciences	Normalized Sum of MPP Request	1400000						1400000
	MPP Request Count	41						41
Engineering	Normalized Sum of MPP Request	20000		1910000				1930000
	MPP Request Count	2		7				9
Environmental Sciences	Normalized Sum of MPP Request		170000					170000
	MPP Request Count		3					3
Fusion Energy	Normalized Sum of MPP Request	350000			40200990			40550990
	MPP Request Count	2			135			137
Geosciences	Normalized Sum of MPP Request			2704000				2704000
	MPP Request Count			15				15
High Energy Physics	Normalized Sum of MPP Request					500000		500000
	MPP Request Count					1		1
Lattice Gauge Theory	Normalized Sum of MPP Request					16910000	17463000	34373000
	MPP Request Count					9	5	14
Life Sciences	Normalized Sum of MPP Request	50000	11019000	40000				11109000
	MPP Request Count	1	27	1				29
Materials Sciences	Normalized Sum of MPP Request	570000		22640750				23210750
	MPP Request Count	3		131				134
Nuclear Physics	Normalized Sum of MPP Request	50000					12696000	12746000
	MPP Request Count	1					31	32
Total Normalized Sum of MPP Request		12967000	26255000	50858750	40200990	37066000	33519000	200866740
Total MPP Request Count		75	74	264	135	84	51	683

ERCAP Allocations 2007 By Office and Science Area

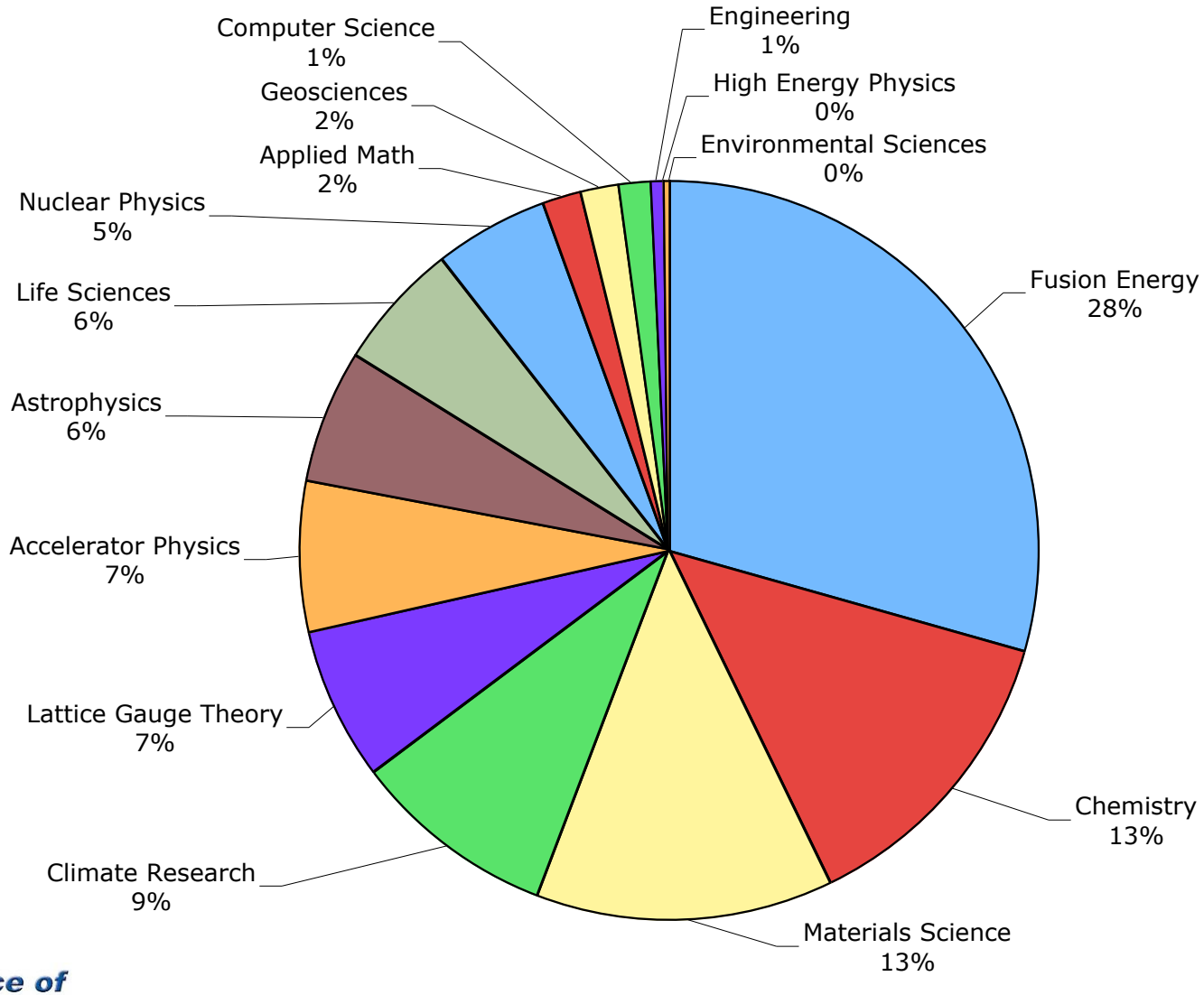
Awards by DOE Office



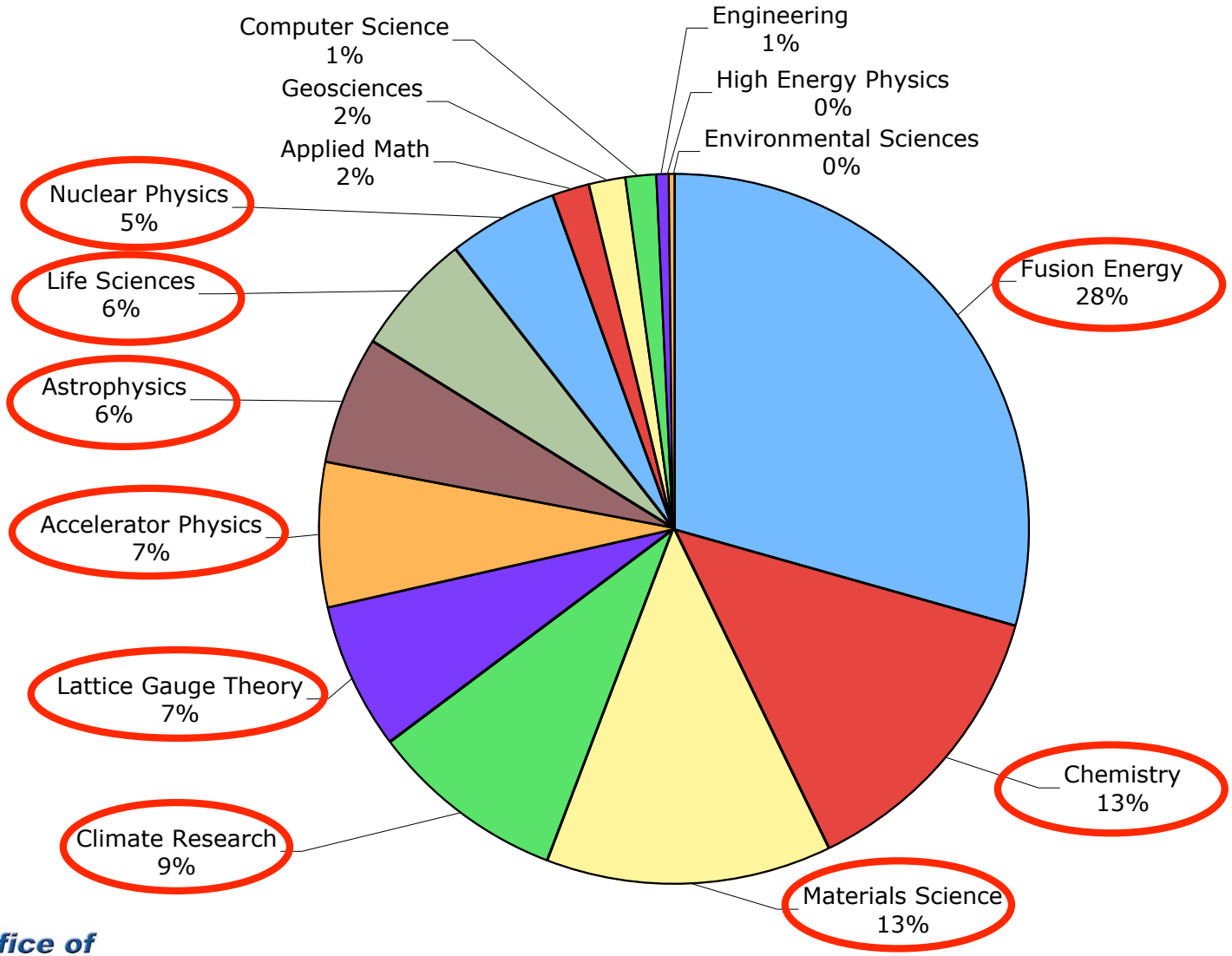
Awards By Science Area



Focus on Science Areas



Focus on Science Areas



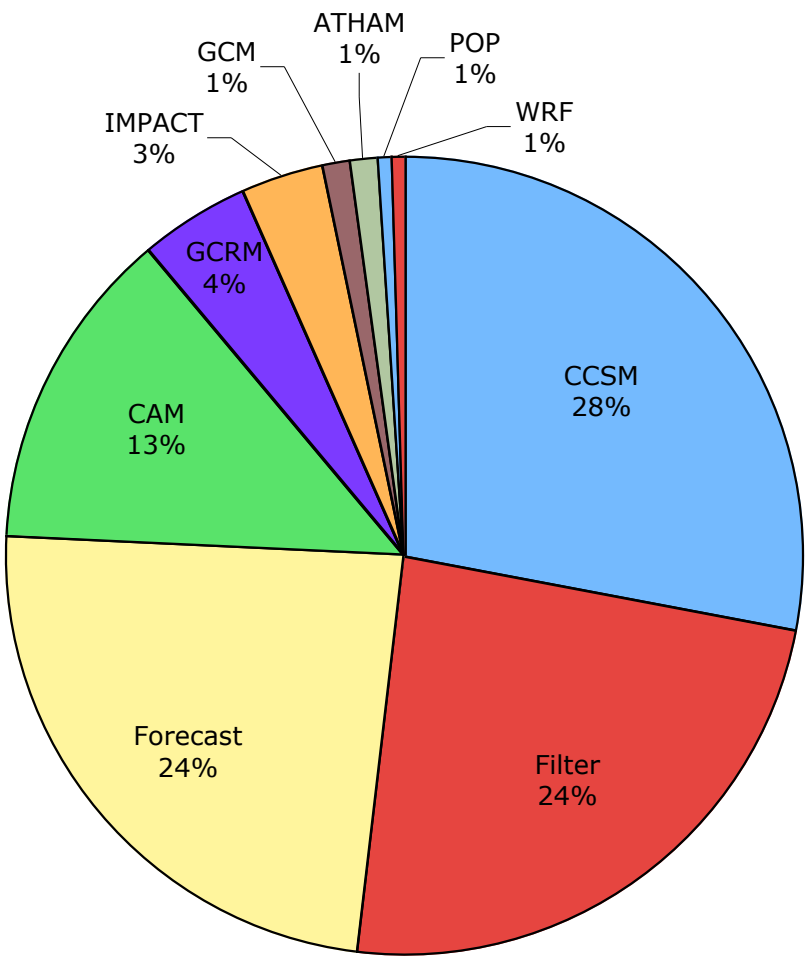
Modeling the NERSC Workload

Science Category	Award MPP Hours	Number of Projects
Fusion Energy	16,509,000	51
Chemistry	7,405,000	41
Materials Science	7,385,000	58
Climate Research	4,911,000	29
Lattice Gauge Theory	3,812,500	6
Accelerator Physics	3,650,000	16
Astrophysics	3,319,000	16
Life Sciences	3,149,000	20
Nuclear Physics	2,790,000	16
Applied Math	930,000	10
Geosciences	920,000	6
Computer Science	731,500	13
Engineering	390,000	4
High Energy Physics	70,000	1
Environmental Sciences	24,000	1

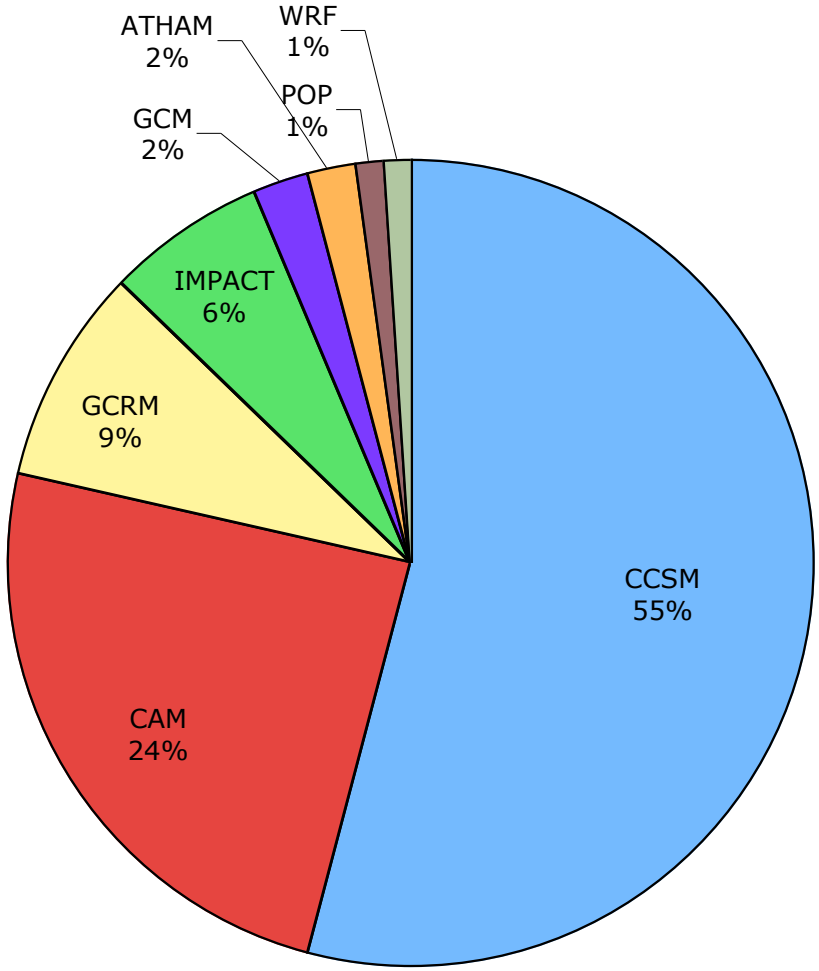
- **Target: choose 6 representative applications from 288 projects, 684 code descriptions**

Climate Modeling (BER)

Climate with INCITE



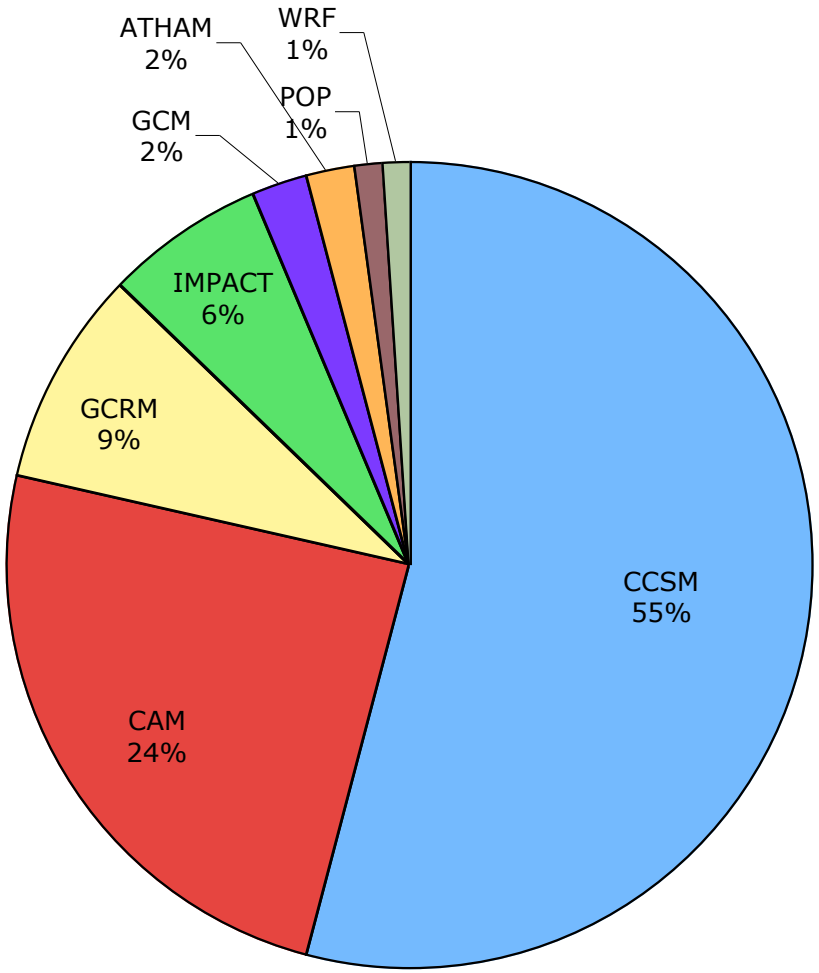
Climate without INCITE



Climate Modeling (BER)

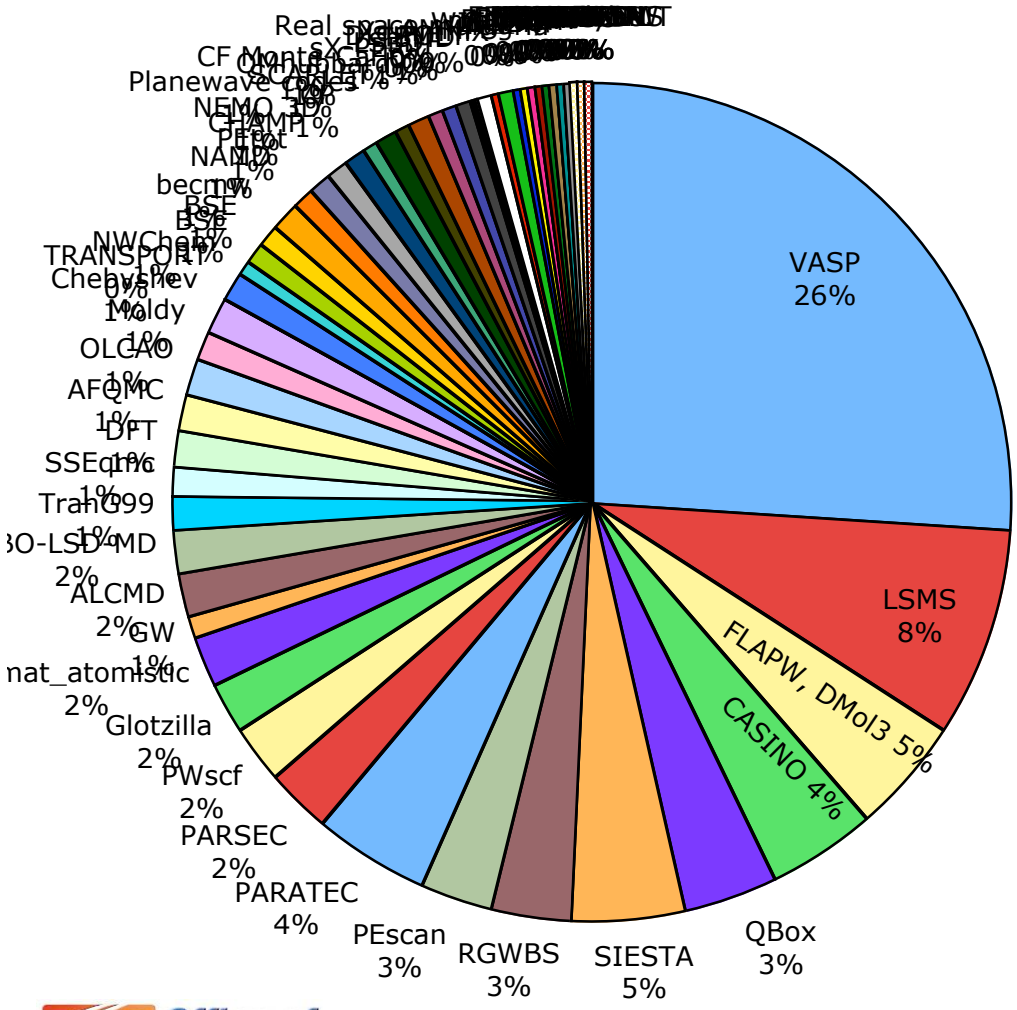
	Code	MPP Award	Percent	Cumulative%
1	CCSM	2,342,000	51%	51%
2	CAM	2,000,000	23%	74%
3	GCRM	2,000,000	8%	82%
4	IMPACT	1,085,000	6%	88%
5	GCM	375,000	2%	90%
6	ATHAM	280,000	2%	92%
7	POP	100,000	1%	93%
8	WRF	80,000	1%	94%

Climate without INCITE



- **CAM and POP dominate CCSM computational requirements**
- **FV-CAM increasingly replacing Spectral-CAM in future CCSM calculations**
- **FV-CAM with D-Mesh selected (coordinate w/NCAR procurement)**

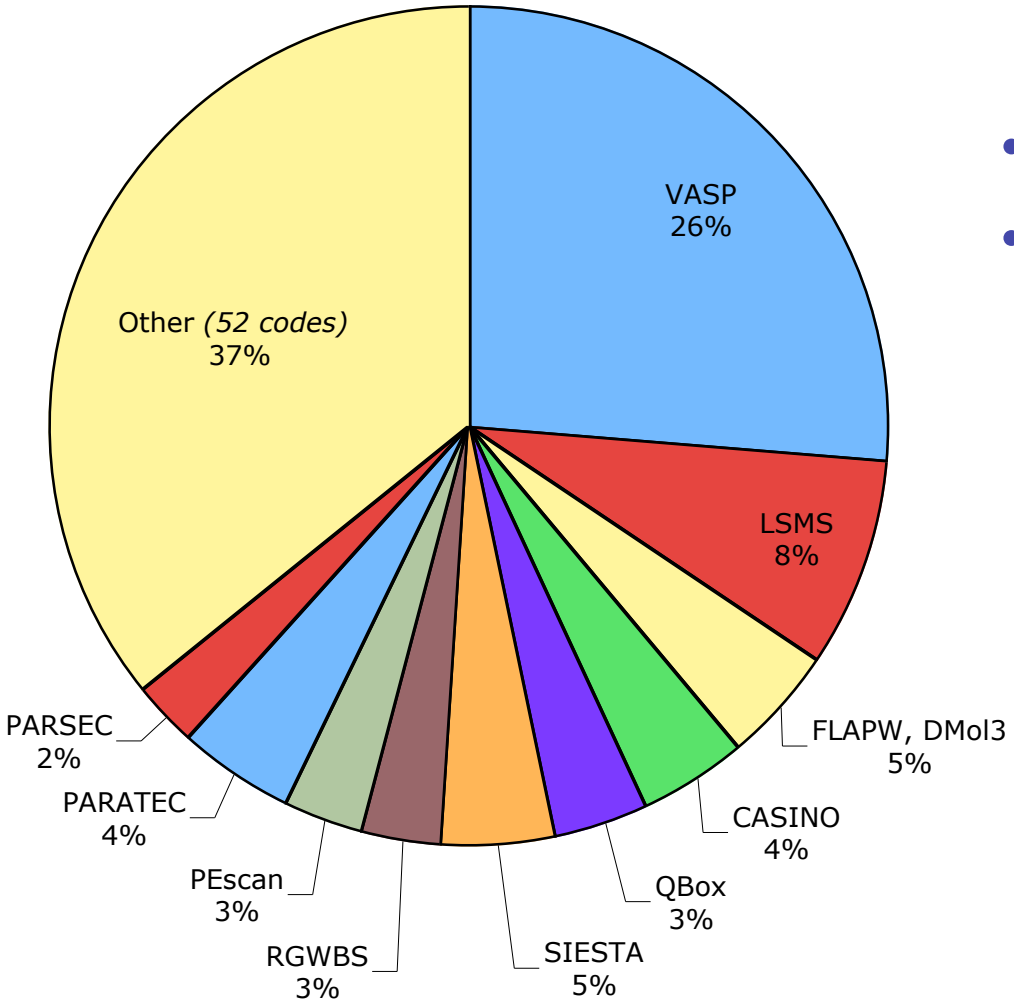
Material Science



- **7,385,000 MPP hours awarded**
- **62 codes, 65 users**
- **Typical code used in 2.15 allocation requests**

	Code	MPP Hours	Percent	Cumulative%
1	VASP	1,992,110	26%	26%
2	LSMS	600,000	8%	34%
3	FLAPW, DMol3	350,000	5%	39%
4	CASINO	312,500	4%	43%
5	QBox	262,500	3%	46%
6	SIESTA	346,500	5%	51%
7	RGWBS	232,500	3%	54%
8	PEscan	220,000	3%	57%
9	PARATEC	337,500	4%	61%
10	PARSEC	182,500	2%	64%
	Other	167,300	34%	66%

Material Science

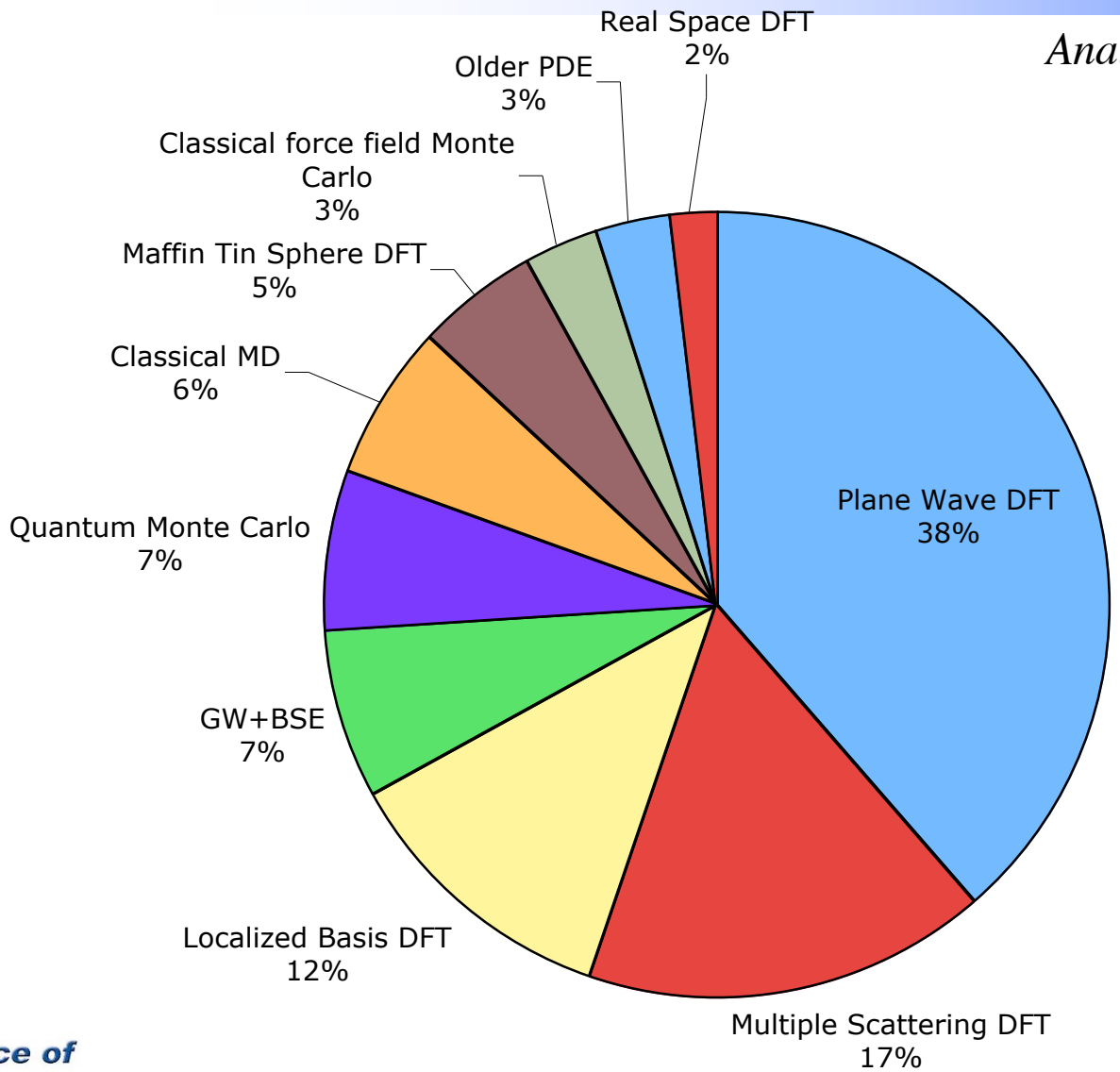


- 7,385,000 MPP hours awarded
- 62 codes, 65 users
- Typical code used in 2.15 allocation requests

	Code	MPP Hours	Percent	Cumulative%
1	VASP	1,992,110	26%	26%
2	LSMS	600,000	8%	34%
3	FLAPW, DMol3	350,000	5%	39%
4	CASINO	312,500	4%	43%
5	QBox	262,500	3%	46%
6	SIESTA	346,500	5%	51%
7	RGWBS	232,500	3%	54%
8	PEscan	220,000	3%	57%
9	PARATEC	337,500	4%	61%
10	PARSEC	182,500	2%	64%
	Other	167,300	34%	66%

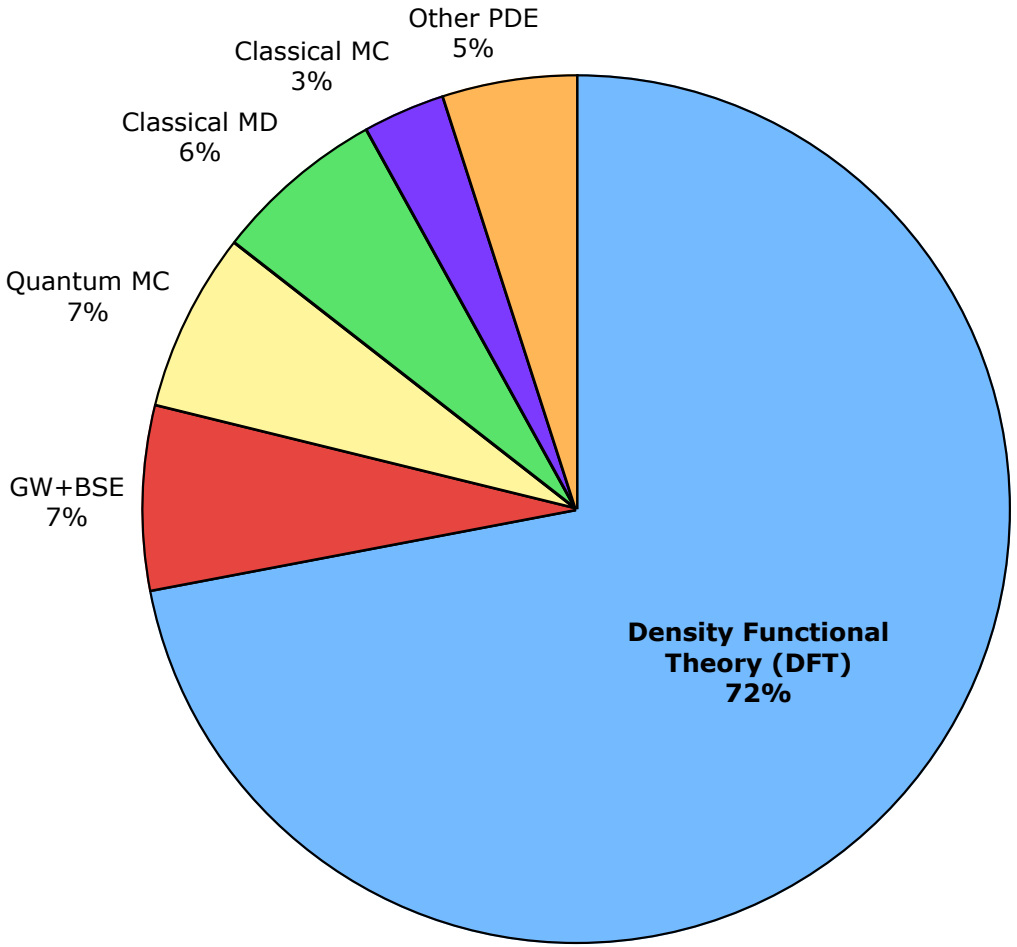
Materials Science *(by algorithm)*

Analysis by Lin-Wang Wang

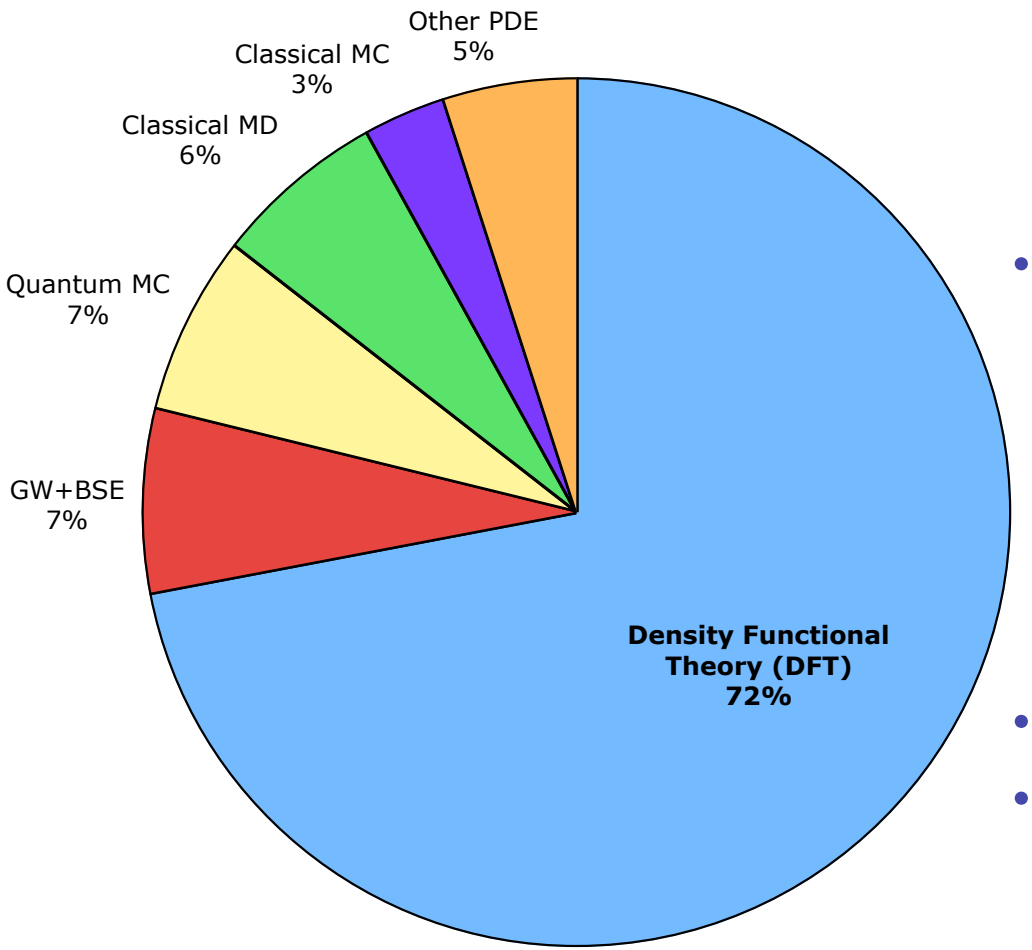


Materials Science (by algorithm category)

Analysis by Lin-Wang Wang

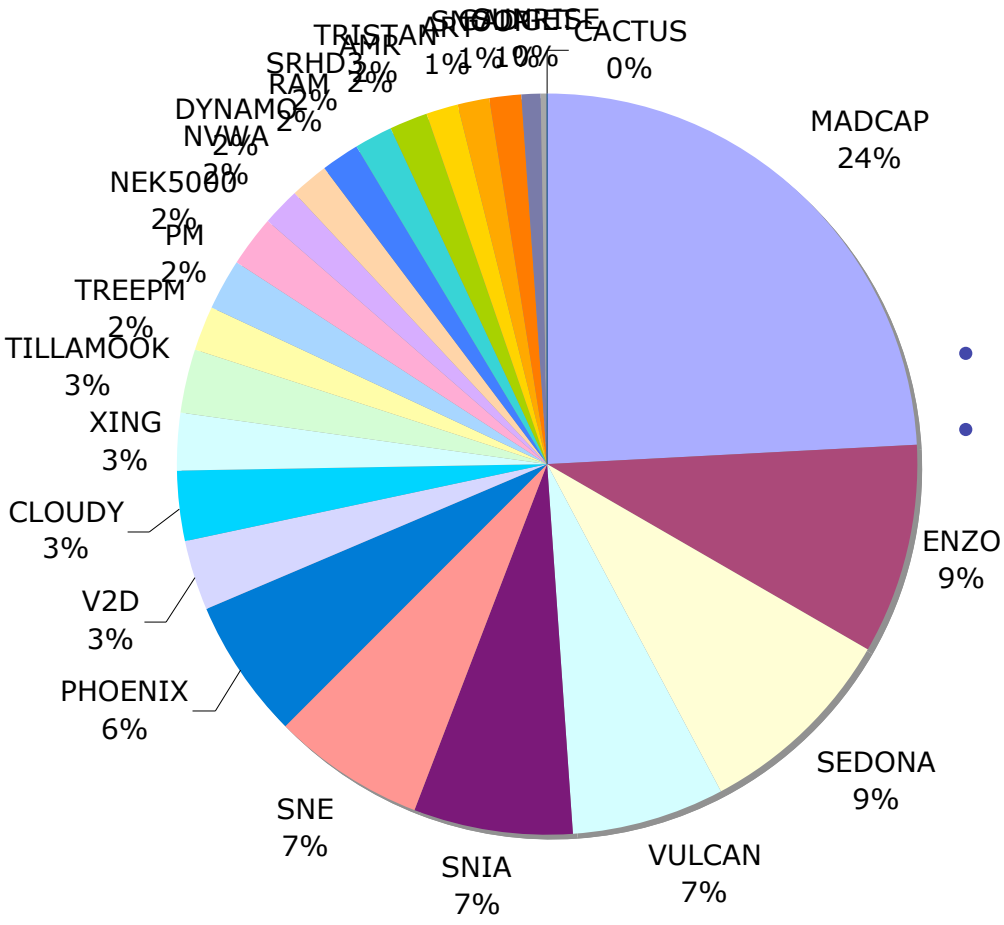


Materials Science *(by algorithm category)*



- Density Functional Theory codes
  - >70% of the workload!
  - Majority are planewave DFT!
- Common requirements for DFT
  - 3D global FFT
  - Dense Linear Algebra for orthogonalization of wave basis functions
  - Dense Linear Algebra calculating pseudopotential
- Dominant Code: VASP
- Similar Codes (planewave DFT)
  - QBox
  - PARATEC
  - PETOT/PESCAN

Astrophysics



- **MADCAP: CMB Analysis Suite**
  - **Dominates allocations even though it is not INCITE**
  - **I/O dominated: Now covered in separate I/O benchmarking tests**
- **ENZO: INCITE AMR code**
- **SciDAC Astrophysics Codes dominant**
  - **Coverage of MHD + combustion**
  - **Suggested to look at codes that use implicit methods rather than explicit timestepping (better representative of future codes)**
  - **Might help with Fusion coverage**

Other Application Areas

- **Fusion: 76 codes**
  - 5 codes account for >50% of workload: OSIRIS, GEM, NIMROD, M3D, GTC
  - Further subdivide to PIC (OSIRIS, GEM, GTC) and MHD (NIMROD, M3D) code categories
- **Chemistry: 56 codes for 48 allocations**
  - INCITE award (S3D) eclipses other chemistry codes: put in separate category
    - Planewave DFT: VASP, CPMD, DACAPO
    - Quantum Monte Carlo: ZORI
    - Ab-initio Quantum Chemistry: Molpro, Gaussian, GAMESS
  - Planewave DFT dominates (but already covered in MatSci workload)
  - Small allocations Q-Chem category add up to dominant workload component
- **Accelerator Modeling**
  - 50% of workload consumed by 3 codes VORPAL, OSIRIS, QuickPIC
  - Dominated by PIC codes

code	MPP Award	Percent	Cumulative%
OSIRIS	2,112,500	11%	11%
GEM	2,058,333	11%	22%
NIMROD	2,229,167	12%	34%
M3D	1,921,667	10%	45%
GTC	1,783,333	10%	54%

Code	Award	Percent	Cumulative%
ZORI	695,000	12%	12%
MOLPRO	519,024	9%	21%
DACAPO	500,000	9%	29%
GAUSSIAN	408,701	7%	36%
CPMD	396,607	7%	43%
VASP	371,667	6%	49%
GAMESS	364,048	6%	56%

Code	MPP Award	Percent	Cumulative%
VORPAL	**1,529,786**	33%	33%
OSIRIS	**784,286**	16%	49%
QuickPIC	**610,000**	13%	62%
Omega3p	**210,536**	4%	66%
Track3p	**210,536**	4%	70%

Selecting Benchmarks

- Coverage
  - Cover science areas
  - Cover algorithm space
- Portable
  - Robust 'build' systems
  - Not architecture specific implementation
- Scalable
  - Do not want to emphasize applications that do not justify scalable HPC resources
- Distributable
  - No proprietary or export-controlled code
- Availability of Developer for Assistance/Support

Narrowing Selection

First Cut (Primary Secondary Kernel):

Fusion PIC	**GEM (4.9) / GTC (4.4) / XGC (3.7) / SUMMIT (1.5)**
Accelerator modeling	**OSIRIS (2.9) / VORPAL (2.6) / IMPACTZT / BeamB3D / QuickPIC**
Fusion MHD	**M3D / NIMROD / GS2 / GYRO / BOUT / Dynamo**
DFT	**VASP / LSMS / PWscf / CPMD / PARATEC /**
Quantum Chem	**NWChem / SIESTA / CASINO / GAMESS / Gaussian**
Climate	**CAM / Forecast Model / IMPACT**
LG Physics	**MILC / RHMC**
Other astro	**MADCAP / SN1A / VULCAN / SRH3D / FLASH/MAESTRO**
Combustion	**S3D**
~~**LifeSci**~~	~~**BLAST / Forge / RepeatMasker**~~

Narrowing Selection

First Cut (Primary Secondary Kernel):

Fusion PIC	**GEM (4.9) / GTC (4.4) / XGC (3.7) / SUMMIT (1.5)**
Accelerator modeling	**OSIRIS (2.9) / VORPAL (2.6) / IMPACTZT / BeamB3D / QuickPIC**
Fusion MHD	**M3D / NIMROD / GS2 / GYRO / BOUT / Dynamo**
DFT	**VASP / LSMS / PWscf / CPMD / PARATEC /**
Quantum Chem	**NWChem / SIESTA / CASINO / GAMESS / Gaussian**
Climate	**CAM / Forecast Model / IMPACT**
LG Physics	**MILC / RHMC**
Other astro	**MADCAP / SN1A / VULCAN / SRH3D / FLASH / MAESTRO**
~~**LifeSci**~~	~~**BLAST / Forge / RepeatMasker**~~

Benchmark Summary

Benchmark	Science Space	Algorithm Space	Proposed Concurrency	Parameters	Lang	Libraries
CAM	Climate (BER)	Navier Stokes CFD, FFT	1024	Nx, ny	F90	netCDF
GAMESS	Quantum Chem (BES)	Dense linear algebra, DFT	512	# of electron/ atoms	F90	DDI, BLAS
GTC	Fusion (FES)	PIC, finite difference	512, 2048	# of particles and grid points	F90	FFT (opt)
MILC	LG Physics (NP)	Conjugate gradiant, sparse matrix	2048	# of grid points, steps per trajectory	C	None
PARATEC	Mat Science (BES)	DFT	512, 2048, 4096	# of atoms	F90	Scalapack
IMPACTZ ? VORPAL ?	Accelerator Physics (HEP)	Largely PIC, fft component Finite difference, PIC		# of particles, grid points	F90/ C++ C/C++	H5Part (HDF5) HDF5, aztec, lapack, petsc
FLASH ? MAESTRO ?	Astrophysics (HEP)	PPM, AMR Low Mach Hydro	1024, 2048	# of grid points	F90/C	HDF5, Para-mesh