
HIGH PERFORMANCE COMPUTING
Research Cyberinfrastructure
UNIVERSITY OF SOUTH CAROLINA

RESEARCH CYBERINFRASTRUCTURE

Dr. Phil Moore Director

Paul Sagona **HPC Systems Architect**

Ben Torkian **Senior Applications Scientist**

Nathan Elger HPC Systems Analyst

Tower Building, Suite 750
1301 Gervais St.

UNIVERSITY OF SOUTH CAROLINA

RESEARCH CYBERINFRASTRUCTURE

Reporting to the Office of Information Technology and working in close partnership with the Office of the Provost, Office of VP for Research and the RCI Faculty Advisory Committee, the RCI program provides faculty and students specialized computational resources and expertise in support of the research mission at the University of South Carolina.

RESEARCH CYBERINFRASTRUCTURE

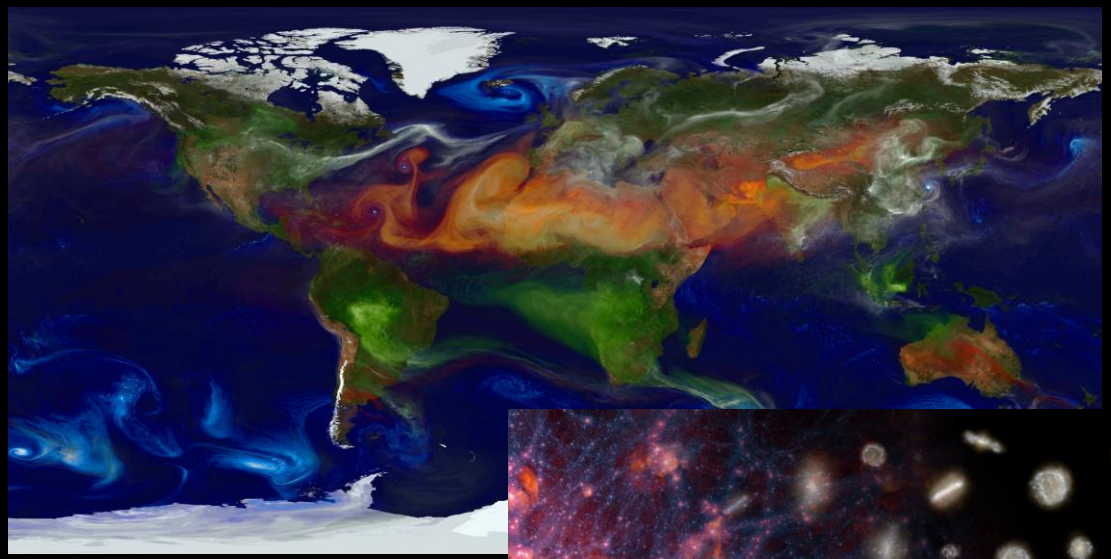
- High performance computing clusters
- Access to national supercomputing facilities
- Scientific and research applications
- High-speed networks
- Software engineering expertise
- Workshops and seminars
- Cyberinfrastructure for grant proposals

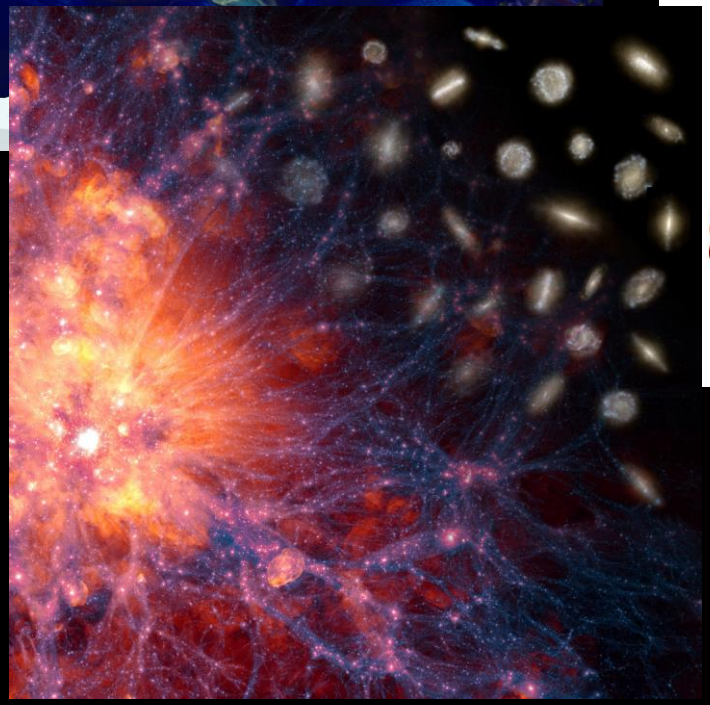
WHAT IS HIGH PERFORMANCE COMPUTING?

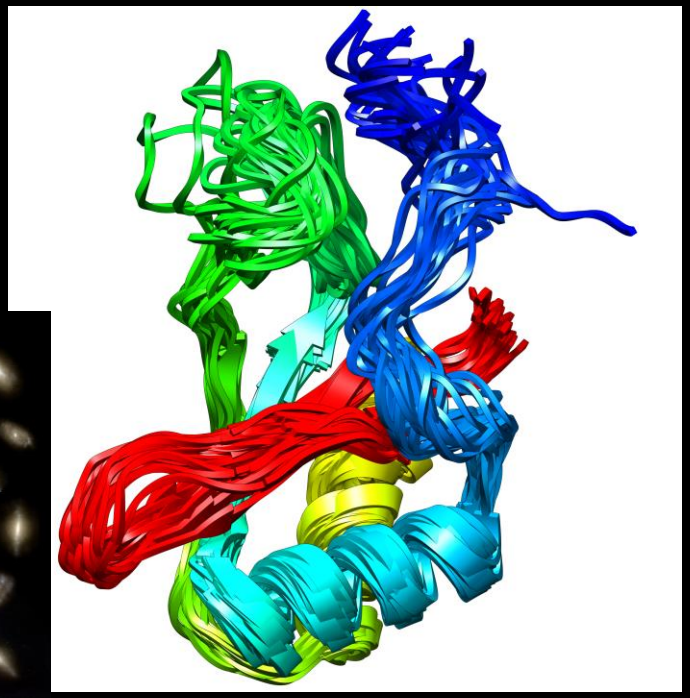
... The application of "supercomputers" to scientific computational problems that are either too large for standard computers or would take them too long.

Richard Gerber, NERSC, https://www.nersc.gov/assets/Uploads/Gerber-HPC-2.pdf

HIGH PERFORMANCE COMPUTING







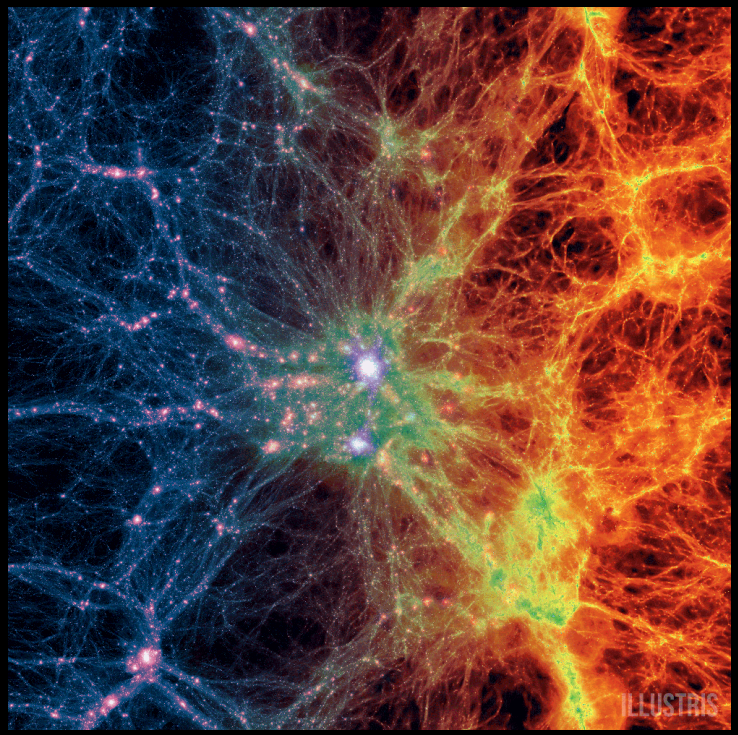
Sources:
http://www.nas.nasa.gov/SC13/assets/images/content/21_Duffy_D_Climate_Nature_Run_Chem_2006-09-13_06-00_SC13_big.jpg
http://www.gauss-centre.eu/gauss-centre/EN/Projects/Astrophysics/2014/springel_illustris.html;jsessionid=C79D2B1BE3B2161A053E3C9FBEB4BEAA?nn=1345700
https://www.nersc.gov/assets/ScienceAtNERSC/DaggettHiResWhitebg.png

UNIVERSITY OF SOUTH CAROLINA

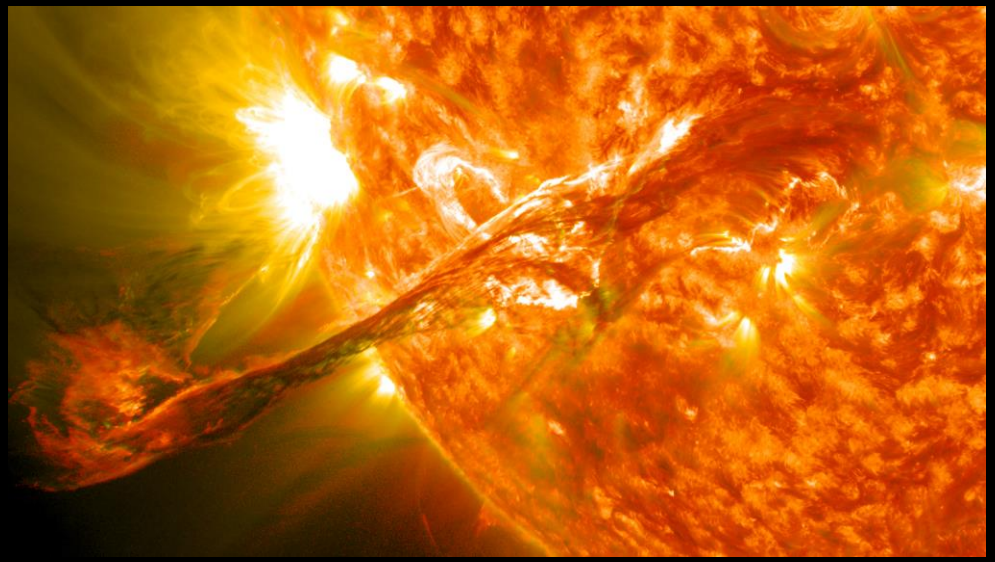
HIGH PERFORMANCE COMPUTING

Why use HPC?

- Access to the Universe. Past, Present, Future… Way past the realm of human experience.
- Extreme time-scales and sizes
- Explore dangerous or inaccessible domains



http://www.gauss-centre.eu/gauss-centre/EN/Projects/Astrophysics/2014/springel_illustris.html;jsessionid=C79D2B1BE3B2161A053E3C9FBEB4BEAA?nn=1345700



http://en.wikipedia.org/wiki/Solar_flare#mediaviewer/File:Magnificent_CME_Erupts_on_the_Sun_-_August_31.jpg

UNIVERSITY OF SOUTH CAROLINA

HIGH PERFORMANCE COMPUTING

What is a supercomputer?

A supercomputer is not so different from a super high-end desktop computer.

Or rather, a lot of super high-end desktop computers.

http://2.bp.blogspot.com/-TNIuXCaY2iA/T6WIRQDuR3I/AAAAAAAABIM/aksHvb7MLk0/s1600/CyberpowerPC6.jpg

http://upload.wikimedia.org/wikipedia/commons/4/44/Titan_render.png

UNIVERSITY OF SOUTH CAROLINA

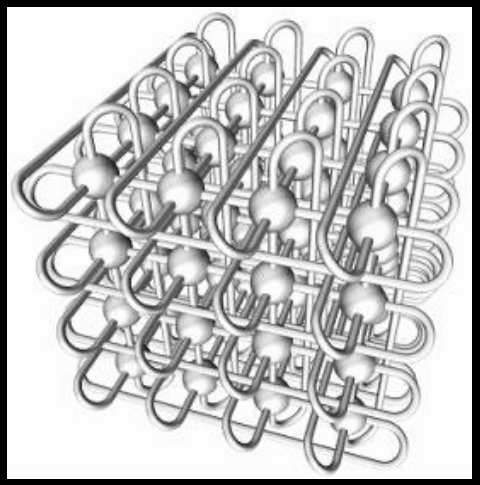
HIGH PERFORMANCE COMPUTING

More than just powerful servers…

The nodes are all connected to each other with an ultra high bandwidth, low latency network.

Typical point-to-point bandwidth
Supercomputer: 10 GBytes/sec
Your home: 0.02 GBytes/sec

Latency
Supercomputer: 1 µs
Your home computer: 20,000 µs



http://antikhali.com/cap/torus-network

Titan uses a three-dimensional torus network.

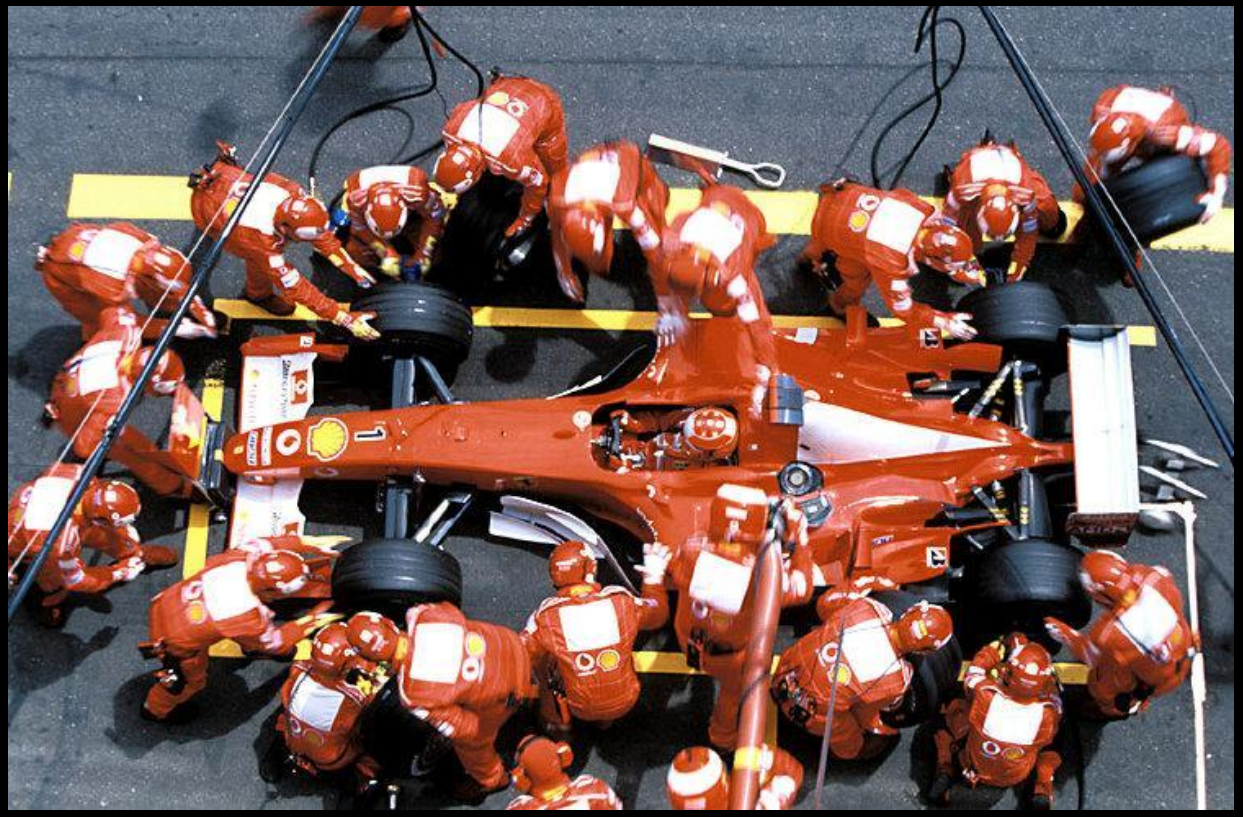
THE TITAN SUPERCOMPUTER

Active	Became operational October 29, 2012
Location	Oak Ridge National Laboratory
Architecture	18,688 AMD Opteron 6274 16-core CPUs
	18,688 Nvidia Tesla K20X GPUs
Power	8.2 MW
Storage	40 PB, 1.4 TB/s IO Lustre filesystem
Speed	17.59 petaFLOPS (LINPACK)
	27 petaFLOPS theoretical peak
Cost	$97 million
Ranking	TOP500: #2, June 2014
Purpose	Scientific research
Legacy	Ranked 1 on TOP500 when built.

UNIVERSITY OF SOUTH CAROLINA

PARALLEL COMPUTING

Programs for HPC systems must be split up into many smaller "programs" called threads, corresponding to each core. To piece the larger program together, the cores must be able to communicate with each other efficiently, and the system as a whole must be organized well. (NICS)



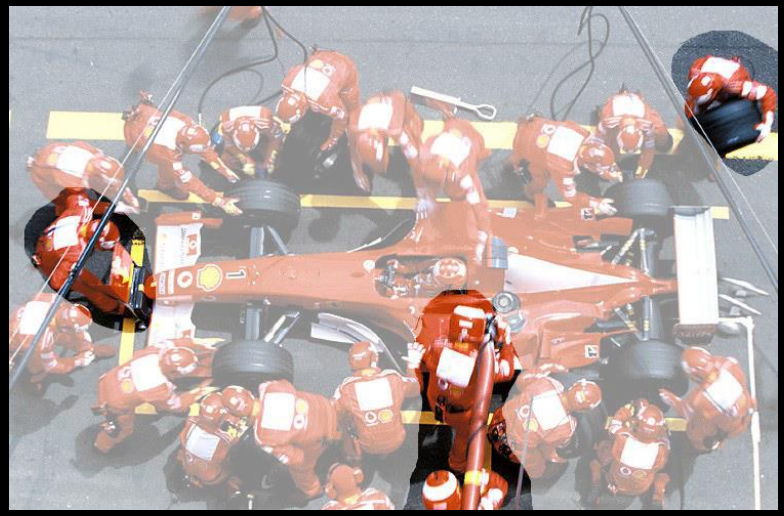
Source: http://www.galeforcef1.com/?p=30

"Divide and Conquer"

TWO TYPES OF PARALLELISM

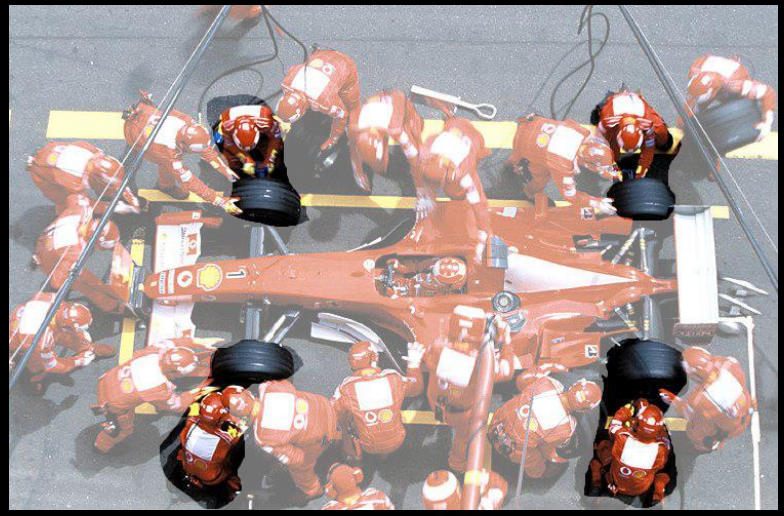
Functional parallelism (task parallelism):

- different people are performing different tasks at the same time



Data parallelism:

- different people are performing the same task, but on different equivalent and independent objects



CHALLENGES OF HPC

- Difficult to program
- Complex networks to overcome bottlenecks
- Massive amounts of data
- Speeds to read and write data
- Heat
- Power

NEW TECHNOLOGY

GPUs and PHI Coprocessors

NEW TECHNOLOGY

GPUs and CUDA

Fire sim:

https://www.youtube.com/watch?v=QY-5ijpkkB8

Galaxy smash

https://www.youtube.com/watch?v=aPgzo9Mvk6o&list=UUV1Uc6fQgEU4Q8tecW6975w

Lots-o-particles

https://www.youtube.com/watch?v=-QJ4bAtS2rk

HIGH PERFORMANCE COMPUTING @USC

Bolden Shared Cluster

- 14 total HP SL Series nodes, 20 cores, Intel Xeon 2.8 GHz
- 280 Total CPU Cores
- 1 node with 256 GB RAM
- 6 nodes with 128 GB RAM
- 6 nodes with 64 GB RAM
- 1 SL250 node with NVIDIA K20X
- 1 Dell R920: 60 Cores, Intel Xeon 2.8 GHz, 256 GB RAM
- FDR Infiniband
- 100TB Lustre File system
- Located in the UTS Annex Datacenter

UNIVERSITY OF SOUTH CAROLINA

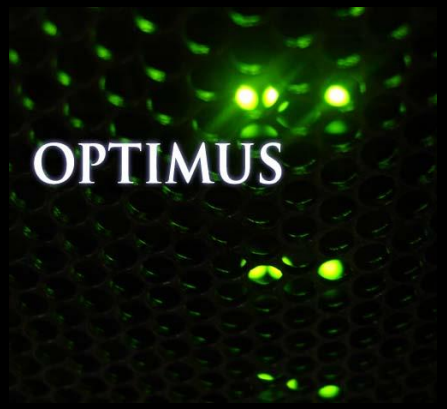
HIGH PERFORMANCE COMPUTING

- Maxwell/Planck Cluster
  - Heterogenous Intel based cluster with shared storage
  - 65 Total Nodes, 804 total CPU Cores
  - 40 HP SL390: 12 cores per node, Intel Xeon 2.4 GHz, 24 GB RAM
  - 19 HP SL250: 12 cores per node, Intel Xeon 2.8 GHz, 24 GB RAM
  - 6 HP SL250: 16 cores per node, Intel Xeon 2.6 GHz, 32 GB RAM
  - 45 NVIDIA M1060
  - 9 NVIDIA M2070
  - NVIDIA m2090
  - Intel Xeon Phi 5110p
  - QDR Infiniband
  - 24TB attached storage

UNIVERSITY OF SOUTH CAROLINA

HIGH PERFORMANCE COMPUTING

- Optimus Cluster
  - Retiring in 2015
  - AMD, 256 cores
- Zia Shared-memory
  - Retiring in 2015
  - SGI Itanium SMP, 128 cores
- Nick Cluster
  - Teaching Cluster
  - 154 CPU cores, 1 TB disk



HPC INCLUDES

- High Performance Computing Clusters
- Scientific Applications
- New accelerator technologies
- Local and national resources
- High-speed networks
- High-throughput data storage
- Visualization

XSEDE

- Extreme Science and Engineering Discovery Environment
- The most powerful integrated advanced digital resources and services in the world. Funded by NSF.
- Scientists and Engineers around the world use XSEDE resources and services: supercomputers, collections of data, help services.
- XSEDE lowers technological barriers to the access and use of computing resources. Using XSEDE, researchers can establish private, secure environments that have all the resources, services, and collaboration support they need to be productive.
- Consists of supercomputers, high-end visualization, data analysis and storage around the country.

BIOINFORMATICS APPLICATIONS

- A5pipeline
- Ampliconnoise
- BEAGLE
- BEAST 1.6.1
- Bioconductor
- BIOM 0.9.3
- Biopython
- BLAST
- BLAT
- Bowtie
- BWA
- cdbtools
- cdhit 3.1
- Chimeraslayer
- clearcut
- CLUSTALW
- Cogent
- DRISEE
- exonerate
- FastTree
- gg_otus
- IDBA-UD
- IMa2
- Infernal
- LIBSVM
- MaSuRCA
- Mothur
- mpiBLAST
- MUSCLE
- ParsInsert
- pplacer
- pprospector
- MUSCLE
- ParsInsert
- pplacer
- Pprospector
- Prottest
- Qiime
- QUAST
- RAxML
- rdpclassifier
- RTAX
- SOAPdenovo
- SPAdes
- TopHat
- trilinos
- Trinity
- uclust
- vienna

HPC APPLICATIONS

Chemistry/Material Engineering

- ADINA
- Amber
- autostem
- delft3d
- gaussian
- gromacs
- LAMMPS
- molden
- mpcq
- NAMD
- NWchem
- openMM
- q-chem
- qhull
- espresso
- TeraChem
- TURBOMOLE
- Zephyr

Mathematics

- ACML
- AUTO
- COMSOL
- CULA
- fftw
- GPUmat4
- Matlab
- Metis
- Octave
- suitesparse
- mathematica
- R
- qrupdate

UNIVERSITY OF SOUTH CAROLINA

DEVELOPMENT HPC APPLICATIONS

- A++/P++
- boost
- CUDA
- GCC
- glpk
- GMAC
- gmp
- hdf5
- Intel
- Java JDK
- Jinja2
- libjacket
- Magma
- mpi4py
- mpiCH2
- mvapich2
- netcdf
- Numpy
- open64
- opencurrent
- openFOAM
- openmpi
- PGI
- Pygments
- Pynast
- Python
- Pytz
- Pyzmq
- sphinx
- SQLAlchemy
- sunstudio
- tornado
- visit
- zmq

HPC IN INDUSTRY

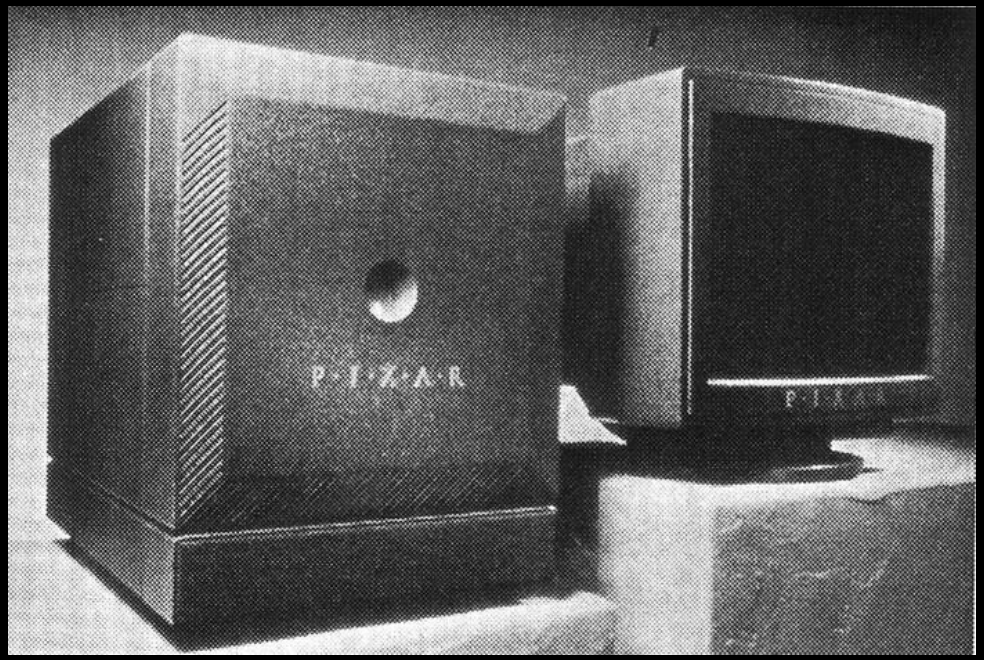
Major industry needs large-scale computing

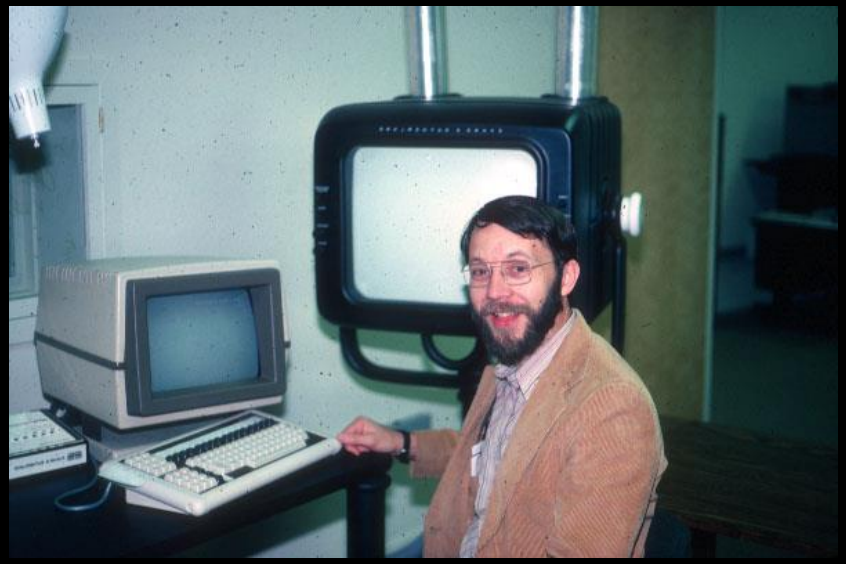
- Oil/Gas/Energy (Exxon, BP,…)
- Pharmaceuticals
- Household goods (Johnson & Johnson)
- Financial Sector
- Aerospace
- Automotive
- Gene Sequencing
- Film/Animation

FILM/ANIMATION STUDIOS

Pixar Image Computer

- 2D and 3D
- PIXAR made GPU
- PIXAR = GPU

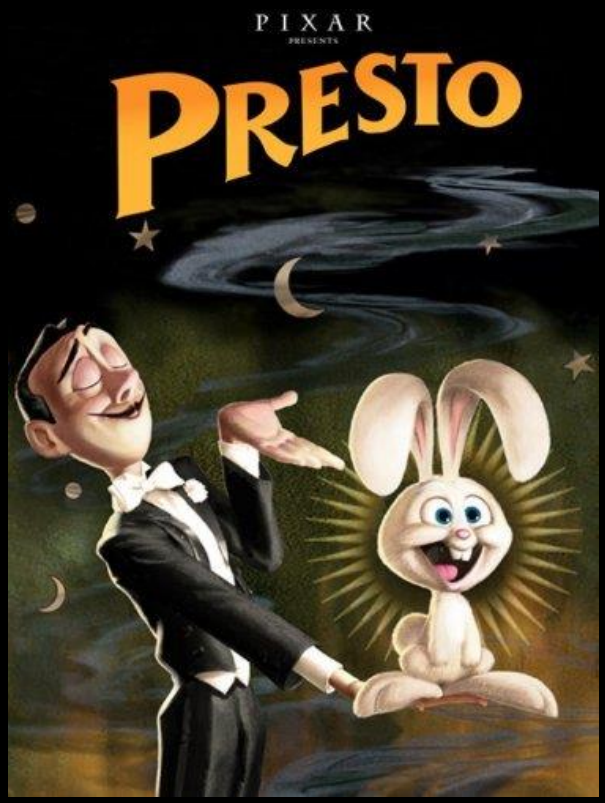




Ed Catmull

PIXAR

- 80s Vector Graphics Machines
- 90s Silicon Graphics four films
  - Toy Story
  - A Bug's Life
  - Monster, Inc.
  - Toy Story 2
- 2001 to now GPU more than 10 films
  - Finding Nemo first port
  - Today Presto K20
  - Next K40



WORKSHOPS

- Python Users Group
- OpenFOAM
- Bioinformatics
- XSEDE
- Software Development
- Version Control
- Linux Operating System
- Vendor training

RCI PROJECTS

- BLAST (distributed speedup)
- MOTHUR (multicore scaling)
- Brain imaging calculation (GPU speedup)
- Stratigraphic simulation(multicore speedup)
- Heat flow simulation (multicore speedup)
- Marsh simulation (.NET to Python conversion)
- Grading application (python development)
- Units calculator (python development)

HPC USERS @ USC

- Pharmacy
- Mechanical Engineering
- Civil Engineering
- Nuclear Engineering
- Computer Science and Engineering
- Chemical Engineering
- Public Health
- Business
- Biology
- Chemistry
- Biochemistry
- Philosophy
- English
- Psychology
- Geology
- Marine Science
- Mathematic
- Statistics
- Physics

CAREERS

Careers in HPC (OR CSCE)

Fields

- Research
- Finance
- Logistic Services
- WWW
- Defense
- Weather and Climate
- Energy
- Telecomm
- Geophysics
- …

Typical hiring under

- Military
- Industry
- State Government
- Federal Government

CAREERS

Military

- Pay
- Benefits
- Travel
- Excellent Technical Training
- Hazardous Duty
- Administration Overhead
- Stability
- Retire Early
- Tuition Assistance

CAREERS

Industry

- Pay
- Benefits
- Travel
- Workshops and Training
- Stability
- For Profit
- Reorganizations

CAREERS

State Government
Agency or University?

- Pay
- Benefits
- Travel
- Technical Training
- Stability
- Jump start for gaining experience

CAREERS

Federal Government
Federal Employee or Contractor?

- Pay
- Benefits
- Travel
- Technical Training
- Stability
- Politics
- Administrative Overhead
- Security Clearance (DOE,DOD,NSA)

CAREERS

Interview Hints

- Don't be too passive or too aggressive
- GPA is important!
- Show what you've done vs. say what you know
- Exploit free software to give yourself experience

CAREERS

People Networking

- Local User Groups
- Google Groups
- Open Source contributor
- Conferences and Presentations
- Volunteer Work
- Web site or Blog

CAREERS

Technologies to KNOW

- Open source experience
- Hands-on class projects are good
- Technical blog or web site
- Integrate several technologies that work together

QUESTIONS?

www.sc.edu/rci

rci@sc.edu

UNIVERSITY OF SOUTH CAROLINA