Scientific Computing at SLAC

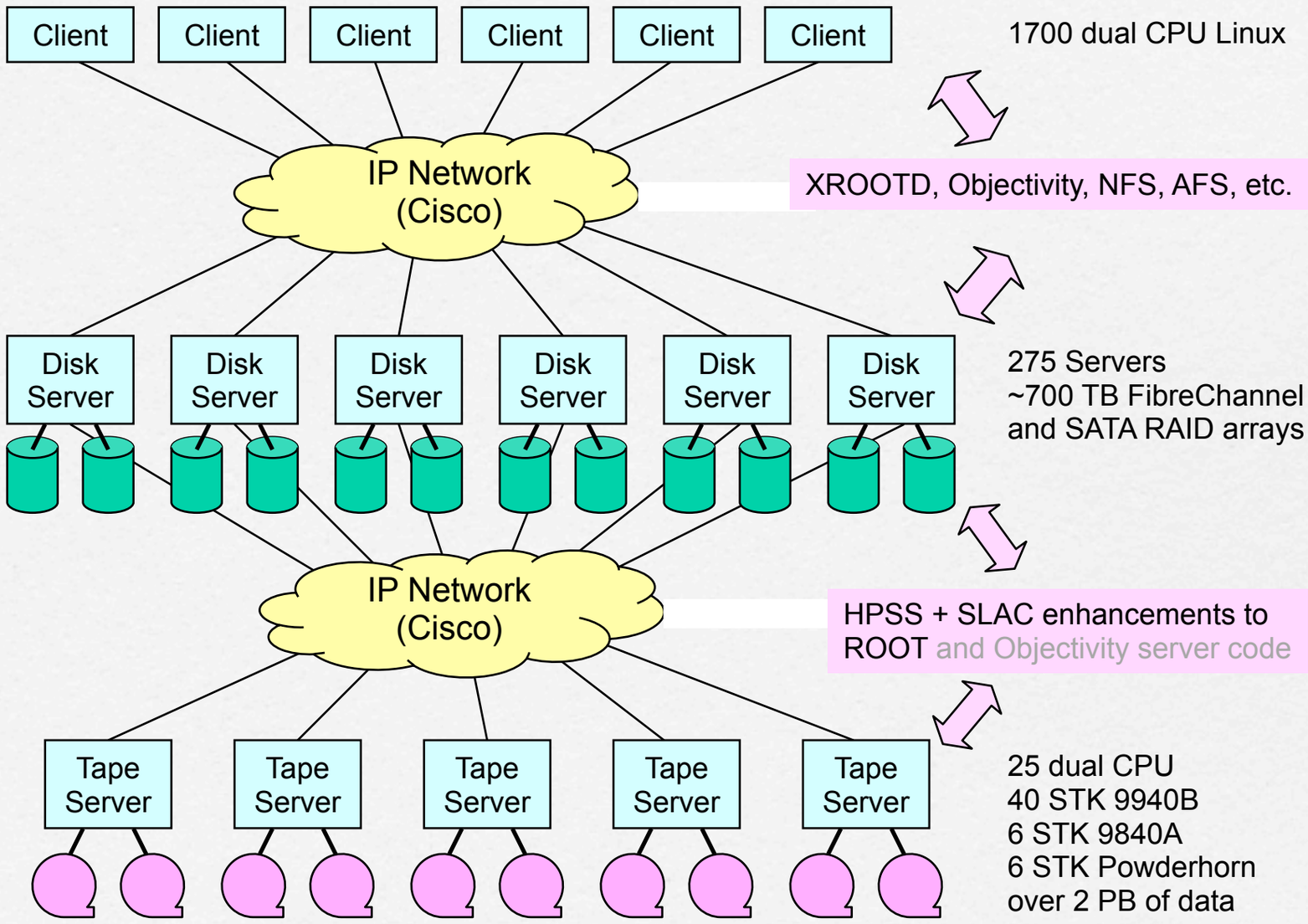
Chuck Boeheim
Asst Director: Scientific Computing and Computing Services

Drivers for SLAC Computing

- Computing to enable today's data-intensive science
  - clusters, interconnects, networks, mass storage, ...
- Computing research to prepare for tomorrow's challenges
  - massive memory, low latency, petascale databases, detector simulation, ...

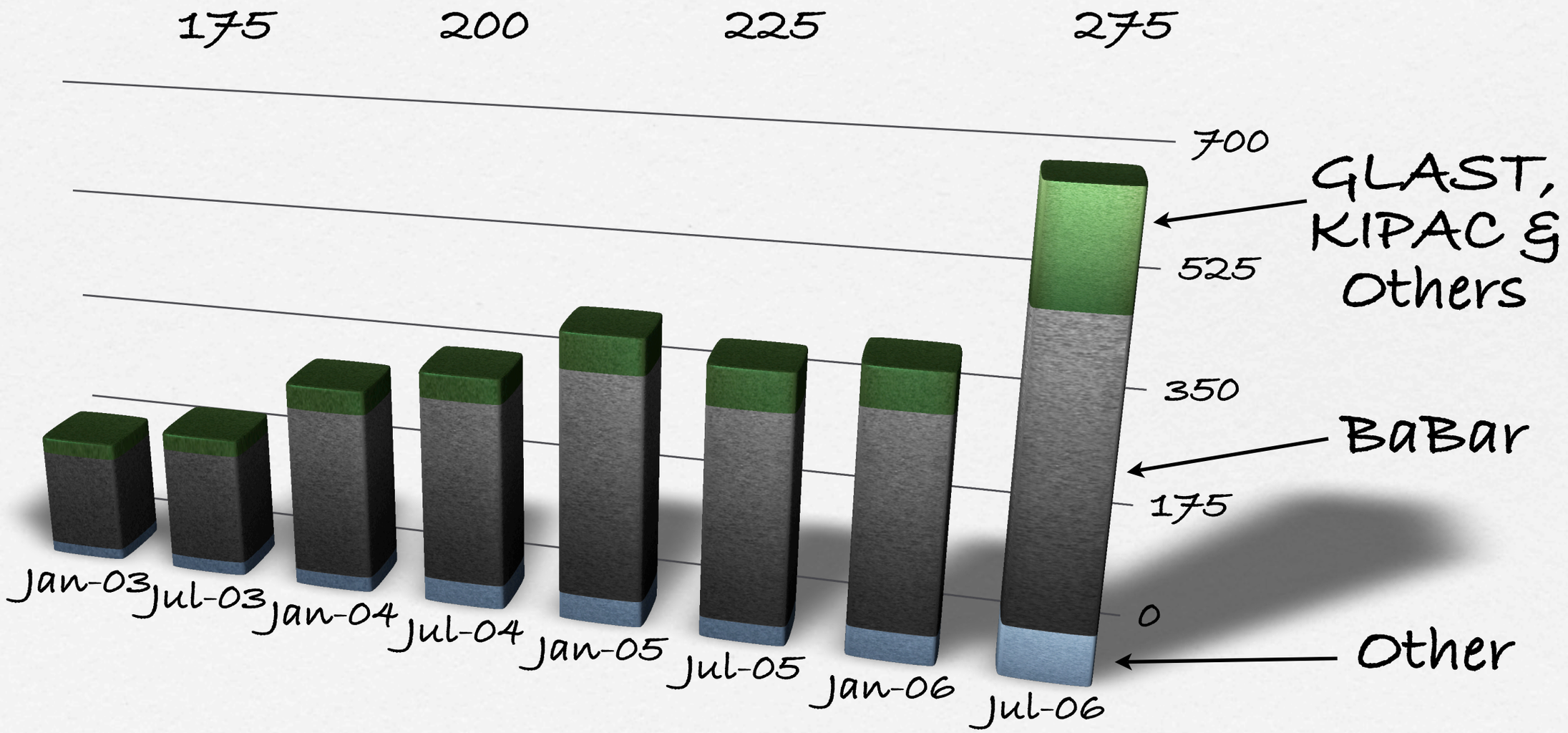
2

SLAC Computing Fabric



Storage Growth

File Servers



Computing Growth (and Diversification)

Number of Nodes

SpecInt2000

Growth and Diversification

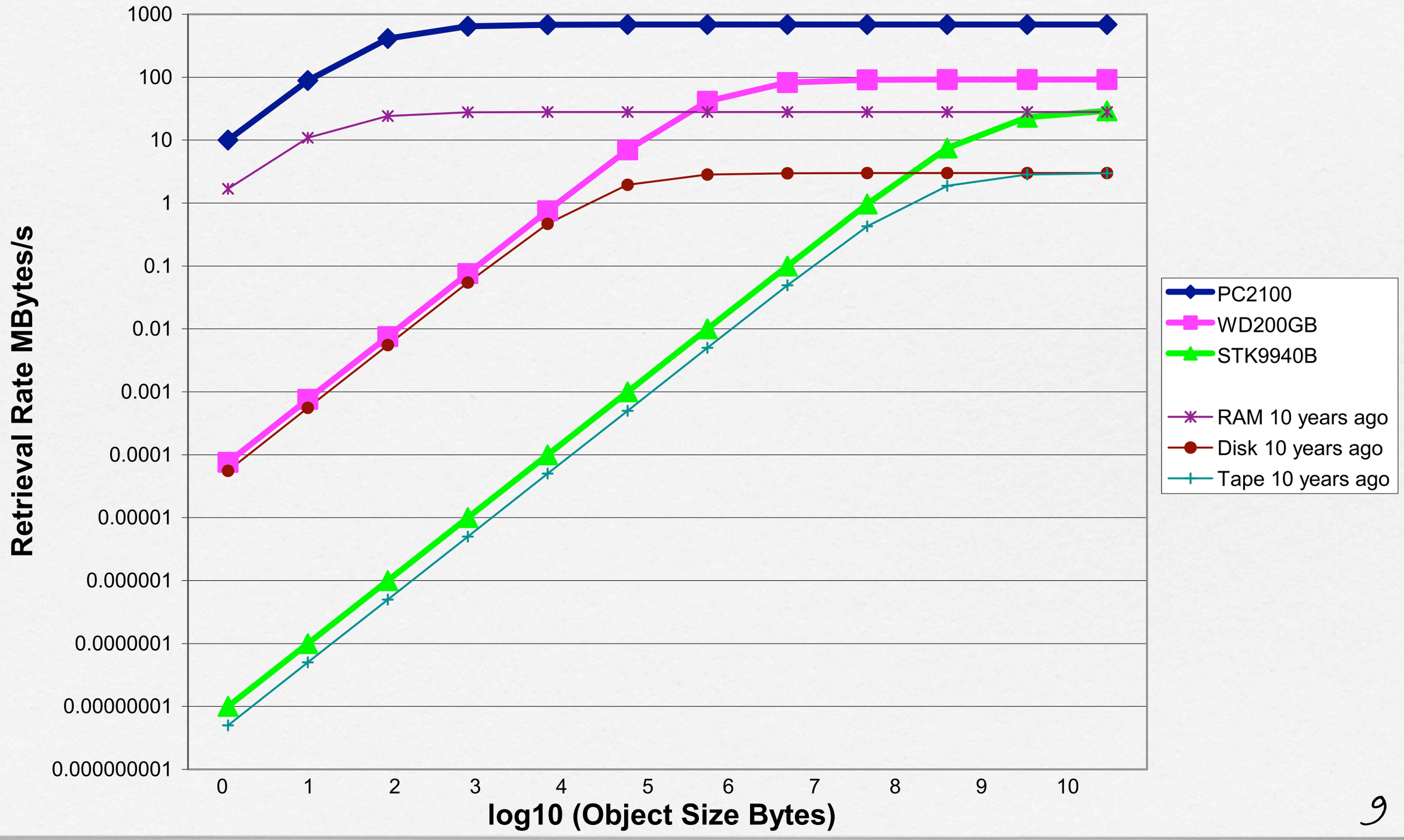
- Continue shared cluster growth as much as practical
- Increasing MPI (Parallel) capacity and support (Astro, Accelerator, and more)
- Grid interfaces and support (Atlas et.al.)
- Large SMPs (Astro)
- Visualization

Research - Petacache

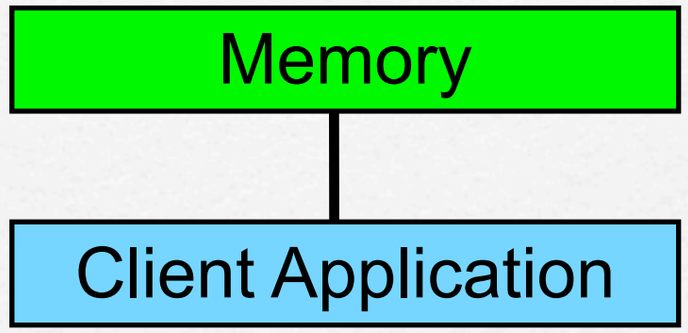
PetaCache Goals

- The PetaCache architecture aims at revolutionizing the query and analysis of scientific databases with complex structure.
  - Generally this applies to feature databases (terabytes-petabytes) rather than bulk data (petabytes-exabytes)
- The original motivation comes from HEP
  - Sparse (~random) access to tens of terabytes today, petabytes tomorrow
  - Access by thousands of processors today, tens of thousands tomorrow

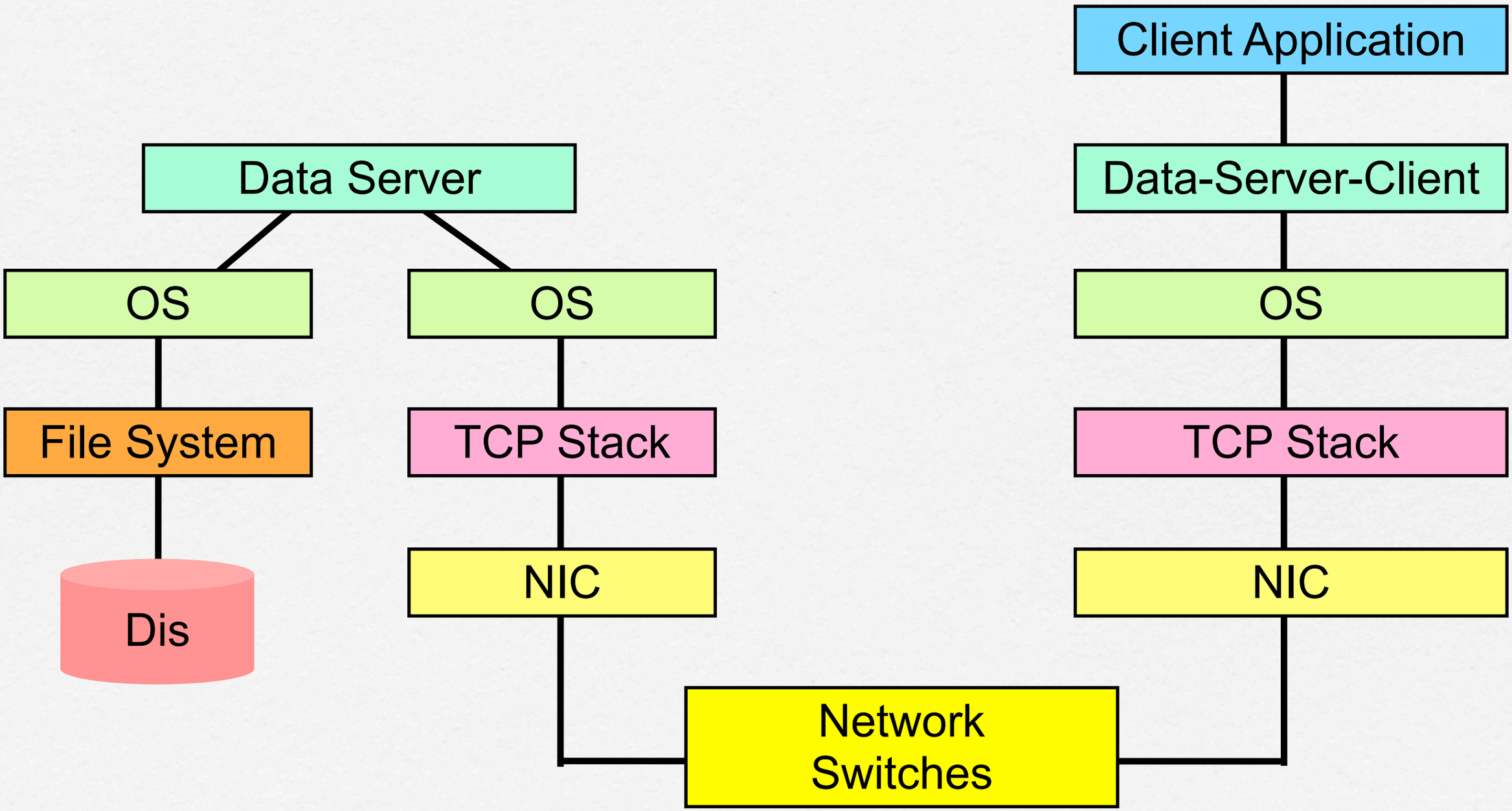
A Decade's Progress in Access Times



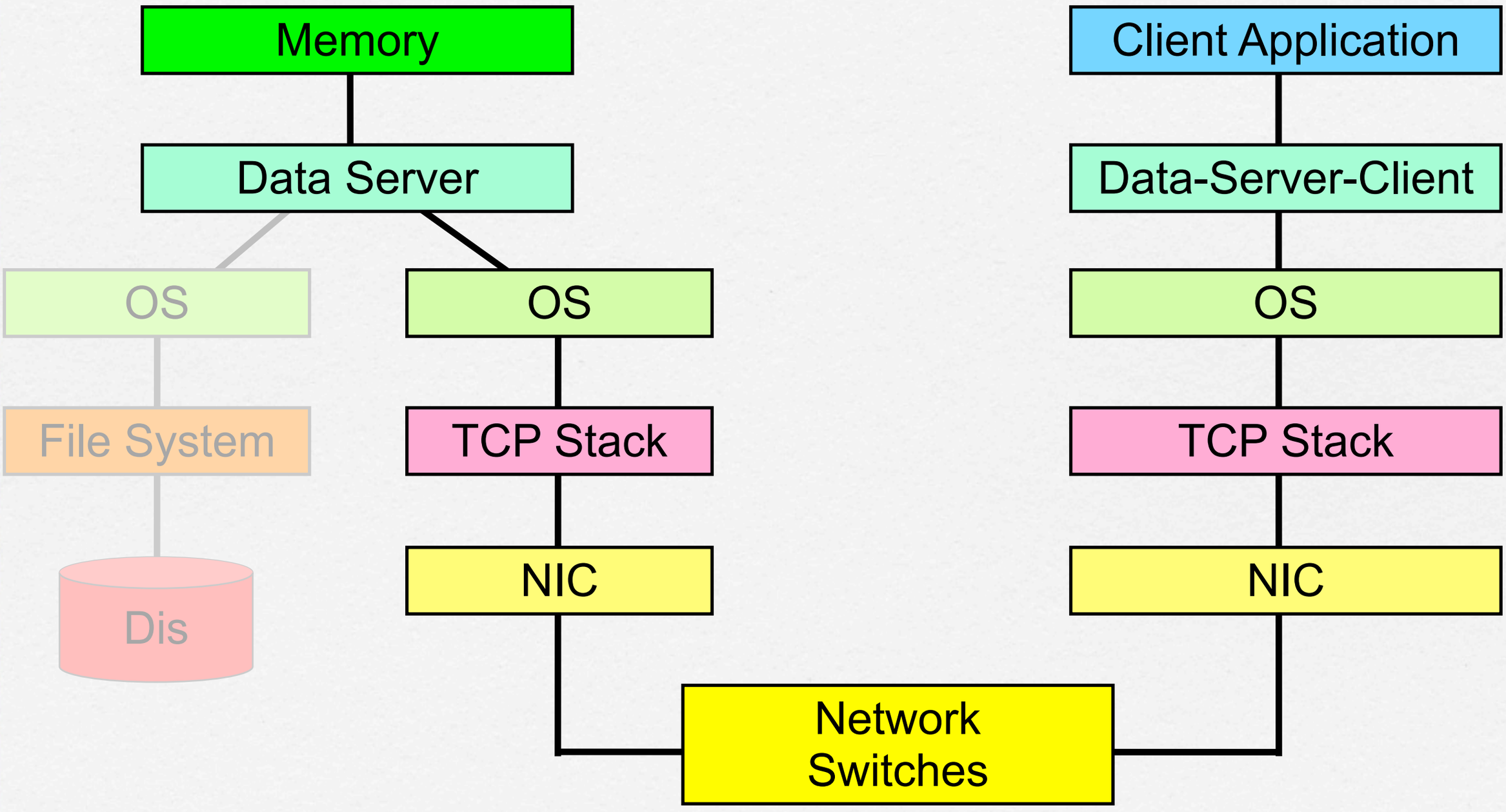
Latency (1) Ideal



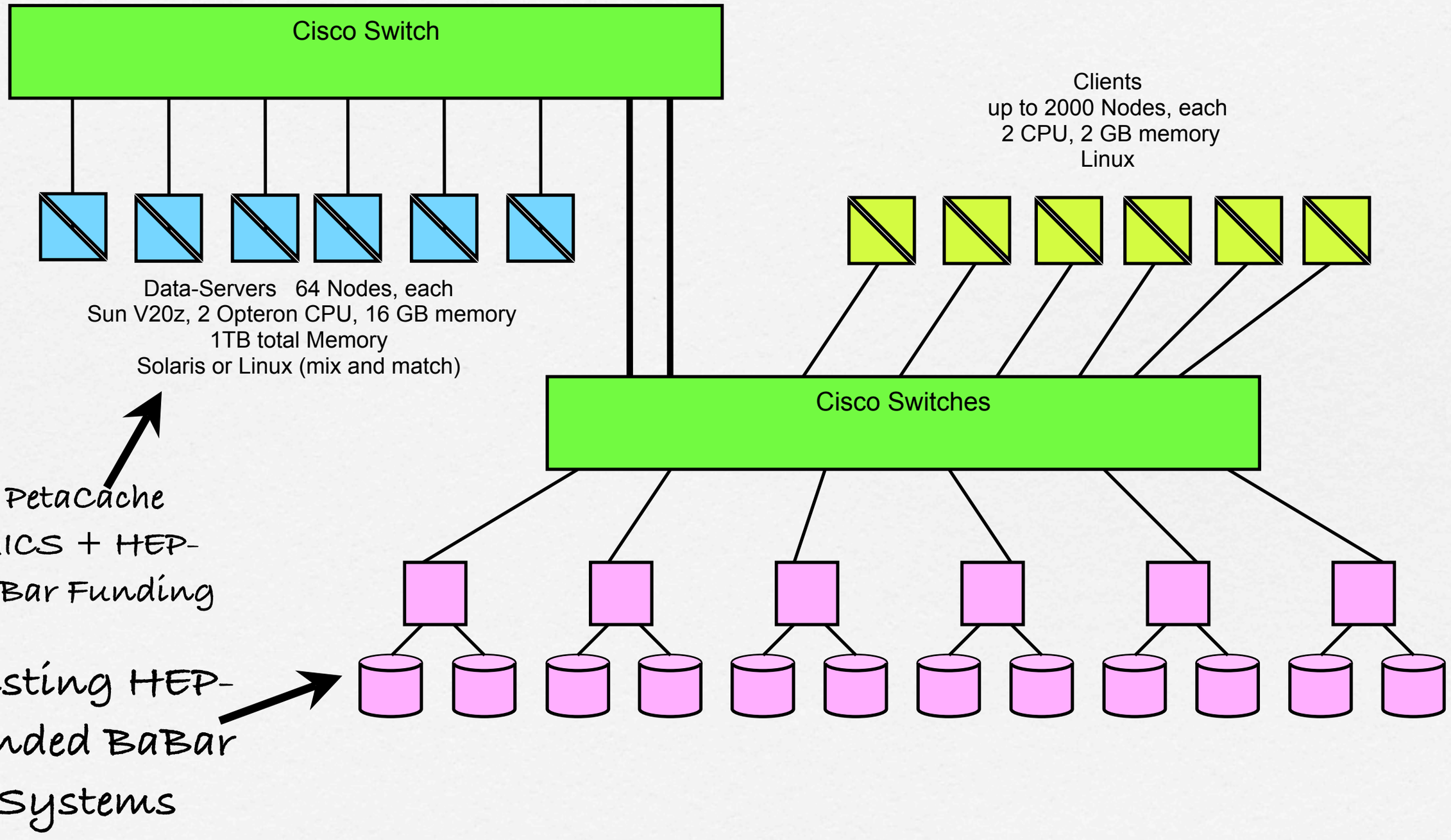
Latency (2) Current reality



Latency (3) Practical Goal



Prototype Machine



Object-Serving Software

- Scala (Xrootd/olbd) (Andy Hanushevsky/SLAC)
  - Optimized for read-only access
  - File-access paradigm (filename, offset, bytecount)
  - Make 1000s of servers transparent to user code
  - Load balancing
  - Self-organizing
  - Automatic staging from tape
  - Failure recovery
- Allows BaBar to start getting benefit from a new data-access architecture within months without changes to user code
- The application can ignore the hundreds of separate address spaces in the data-cache memory

Proposals from SLAC Science

- BaBar
  - Host part of the (~30 TB) microDST data
  - Access data via "pointer skims"
- GLAST
  - Will require a ~2TB intensely accessed database. Have asked to test concepts on a PetaCache Prototype
- LSST Database Prototyping
  - Proposed tests using a PetaCache prototype

Summary

- Data-intensive science increasingly requires low-latency access to terabytes or petabytes
- Memory is one key:
  - Commodity DRAM today (increasing total cost by ~2x)
  - Storage-class memory (whatever that will be) in the future
- Revolutions in scientific data analysis will be another key
  - Current HEP approaches to data analysis assume that random access is prohibitively expensive
  - As a result, permitting random access brings much-less-than-revolutionary immediate benefit
- Use the impressive motive force of a major HEP collaboration with huge data-analysis needs to drive the development of techniques for revolutionary exploitation of an above-threshold machine.

Research - Very Large Databases

SCCS & Very Large Databases

- 10-year, unique experience with VLDB
  - designing, building, deploying & managing peta-scale production data sets / database world's largest database, 1.4 PB
  - innovative approaches to managing BaBar's data proposed, designed and developed by SCCS
  - SCCS leverages VLDB experience to assist LSST in solving data-related challenges
    - effort started ~end of 2004

18

LSST: Data Related Challenges

- Large volumes
  - 7 PB/year (image and catalog data)
  - 500 TB/year (database)
    - Most today's "very large databases": ~10s TB range
- High availability
  - petabytes → 10s of 1000s of disks → daily disk failures
- Real-time requirement
  - Transient alerts generated in < 60 sec
- Spatial & temporal aspects
  - Most surveys focus on a single dimension
- All data made public with minimal delay
  - Wide range of users: professional & amateur astronomers, students, general public

19

SCCS Role in LSST Data Management

- **Current effort**
  - **Architecting data access system**
  - **Database technology choice, design and prototyping**
- **Coordinating 2 LSST teams**
  - **Data Access Working Group**
  - **Database Group**
    - ~12 members from SLAC, NCSA, U. Washington, Johns Hopkins U., U. Arizona, LLNL, SDSC

SCCS leads Data Access and Database related effort

20

More Details on VLDB Work by SCCS

- Prototyping at SCCS
  - to uncover peta-scale limitations, bottlenecks & deficiencies, understand needed improvements
  - current focus on MySQL, discussing / planning DB2 tests with IBM
- Close collaboration with key MySQL developers
  - to develop tools and techniques for supporting petascale databases
  - work include large scale tests, providing input and feature requests
- Working closely with world-class database gurus
  - includes face-to-face discussions, frequent email exchanges and co-writing papers

21

Growing Visibility of SCCS in Astronomical Community

- Actively working with astronomical community on developing standards
  - working with 4 large astronomical surveys (LSST, SDSS, PanSTARRS, 2MASS) on spatial / GIS[1] support by RDBMSes
  - providing input to NVO[2]
- Involved in reviewing computing systems of other astronomical surveys
- Presented at recognized astronomical conferences
  - SPIE Astronomical Telescopes and Instrumentation
  - American Astronomical Society

[1]Geographic Information Systems [2]National Virtual Observatory

22

Research - Geant4

SLAC Geant4 team

- Geant4 is a toolkit simulating elementary particles passing through and interacting with matter, and modeling the detector apparatus measuring the passage of elementary particles and recording the energy and dose deposition.
- Geant4 is developed and maintained by an international collaboration.
  - SLAC is the second largest center next to CERN.

SLAC Geant4 Team

- The SLAC Geant4 team is strongly involved in the development and maintenance of Geant4 toolkit.
  - Makoto Asai
    - Deputy-spokesperson, coordinator of Global Architecture, Run&Event and Detector Response WGs
  - Dennis Wright
    - Coordinators of Hadronics and Documentation WGs
  - Joseph Perl
    - Coordinator of Visualization WG
  - Tatsumi Koi
    - Major developer of neutron physics
  - Jane Tinsley
    - Major developer of event biasing options and visualization

25

User support and technology transfer

- BaBar, the pioneer experiment, adopted Geant4 in 2000. They have simulated more than 10^{10} events so far at 20 sites in US and Europe.
- All experimental programs at SLAC including GLAST, EXO and ILC detector study rely on Geant4. All of four LHC experiments at CERN and most of other experiments in US and in the world have already adopted it.

26

User support and technology transfer

- Use of Geant4 not limited to high energy and nuclear physics. Users in space engineering, medical and industrial applications are rapidly expanding.
  - Geant4 is an excellent showcase of technology transfer from particle physics to other fields.
- The SLAC Geant4 team supports Geant4 for all experiments at SLAC. It also acts as the support center for all other DOE Labs and other users in US.

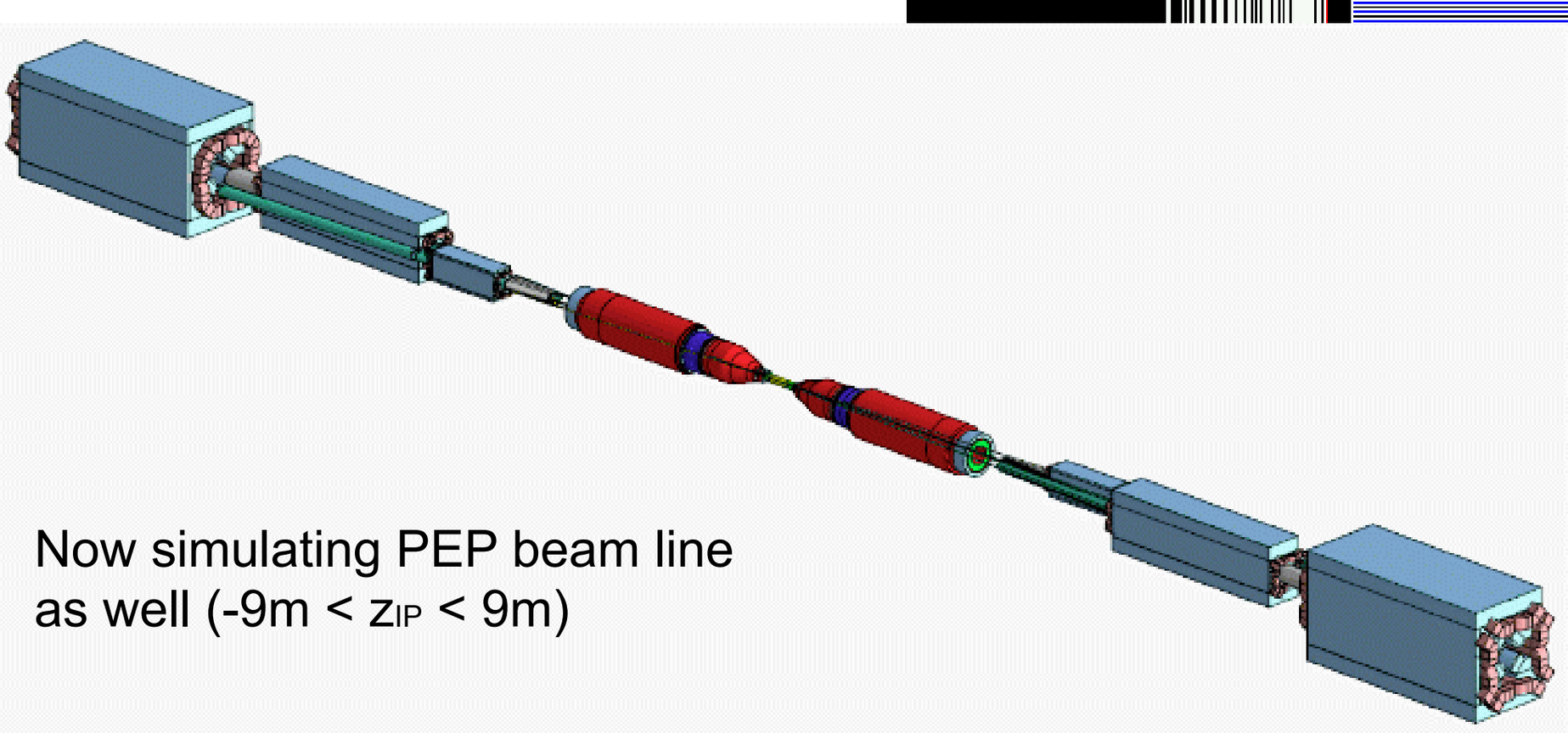
Current Uses of Geant4

- BaBar
- PEP Beamline
- ATLAS
- CMS
- LHCb
- GLAST LAT
- EXO
- Beam Transport
- ILC LCD
- Astro
- Medical

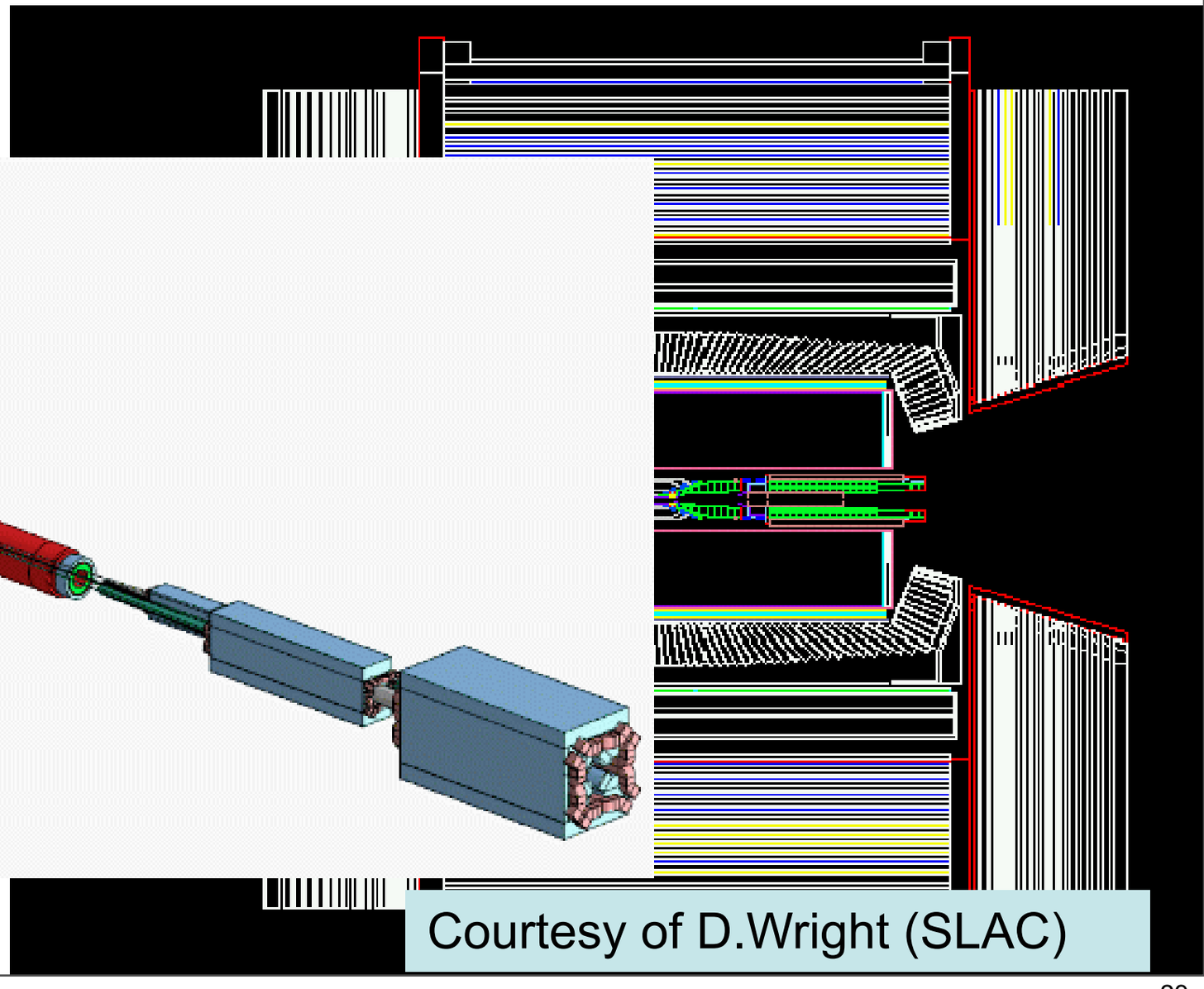
28

BaBar

- BaBar at SLAC is the pioneer experiment in HEP in use of Geant4
  - Started in 2000
  - Simulated 10^{10} events so far
  - Produced at 20 sites in North America and Europe
  - Current average production rate 6.1 x 10^{7} events/week



Now simulating PEP beam line as well (-9m < z_{IP} < 9m)



Courtesy of D.Wright (SLAC)

Geant 4

Geant4 at the LHC Today

Now Geant4 has become the standard simulation for ATLAS, LHCB, and CMS

	ATLAS	CMS	LHCb
Transition to Geant4 (G3 stopped)	DC02 '04	Nov '03	May '04
Produced # of events in DC	12 M	40 M	80 M
CPU time (sec)/ event (2.8 Ghz)	600 ($pp \rightarrow Z \rightarrow ee$) 700 (SUSY)	200 (QCD jets) 60 (min bias)	22-65
Memory used	400 Mb	220 Mb	220 Mb
# of placed volumes	5 M	1.2 M	18 M

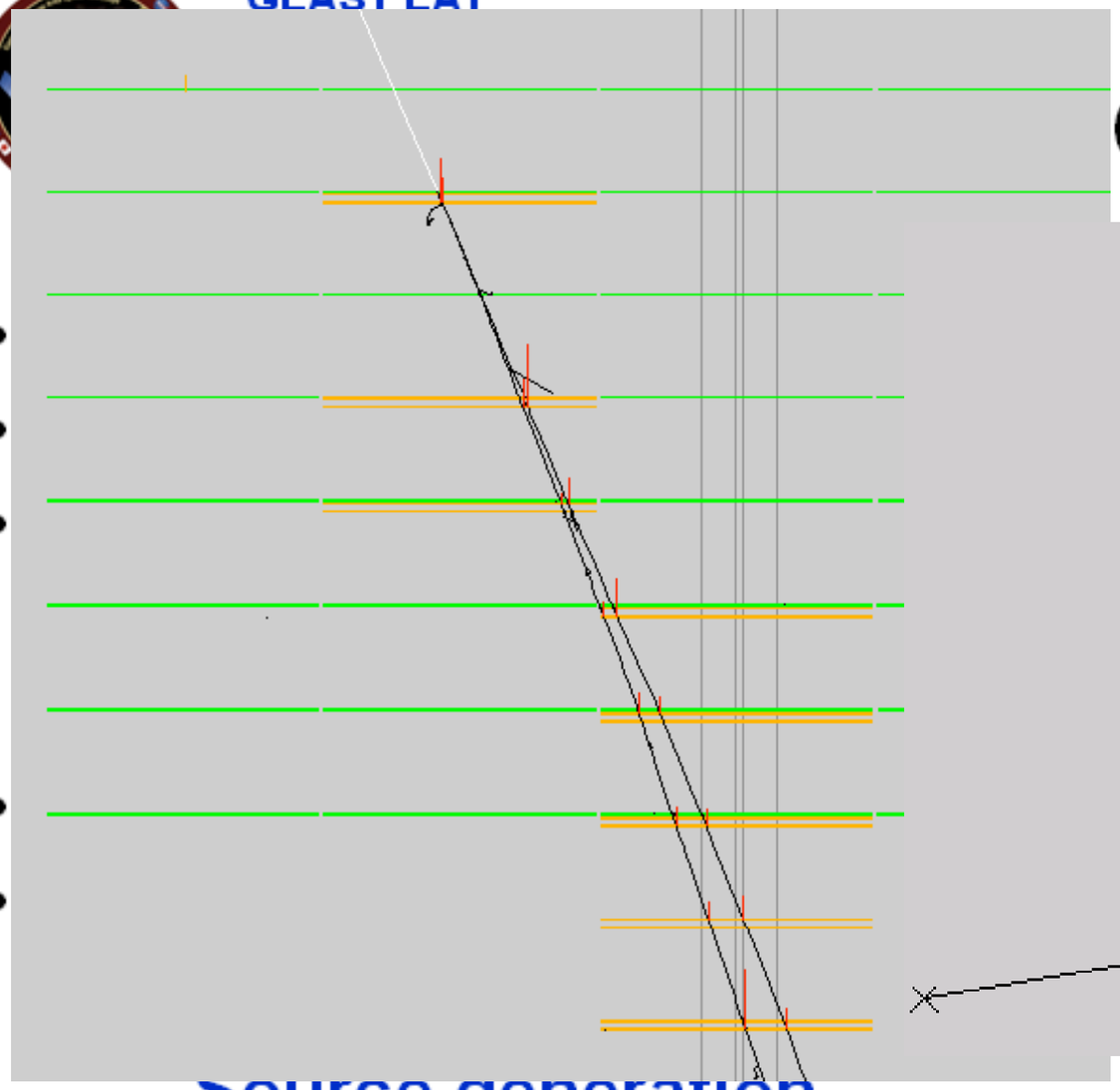
No memory leaks!!

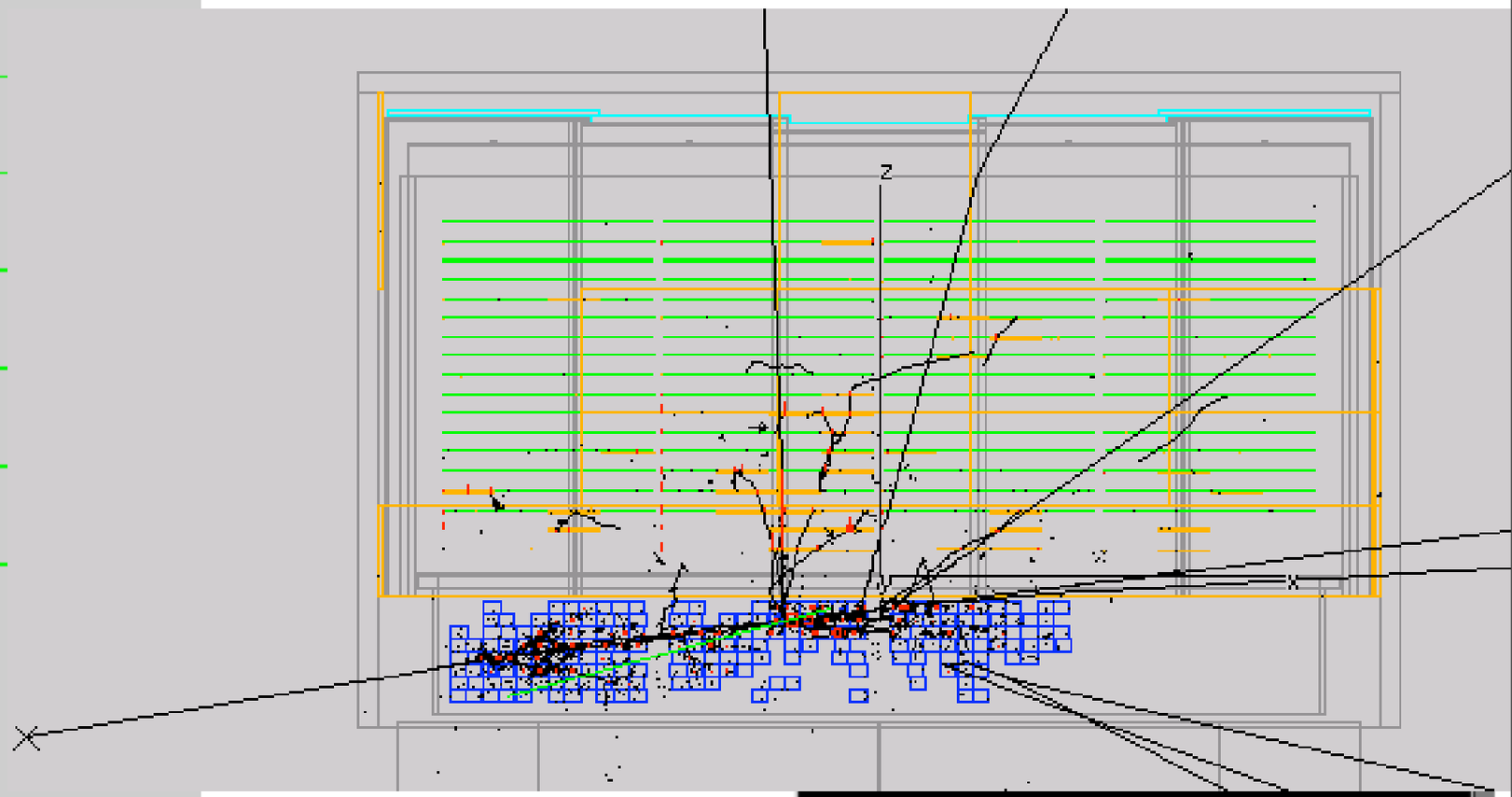
⇒ Observations:

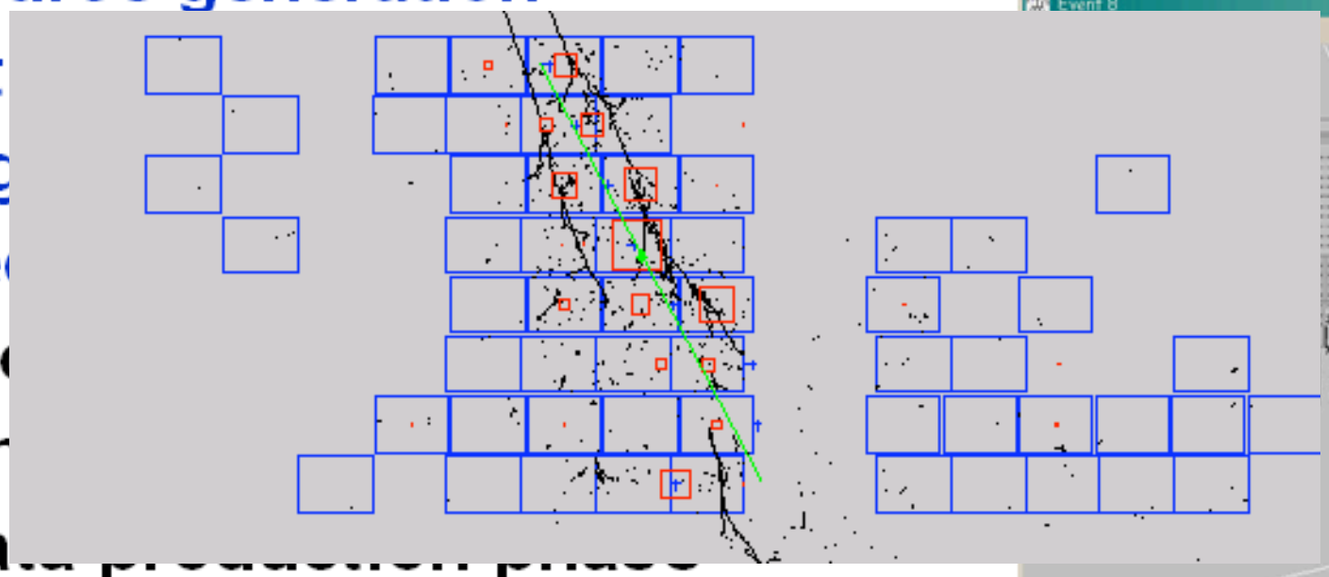
- Geant4 in production is running now very stable/very few problems (~ 10^{-5})
- Transition to Geant4 has been a very smooth process for all experiments

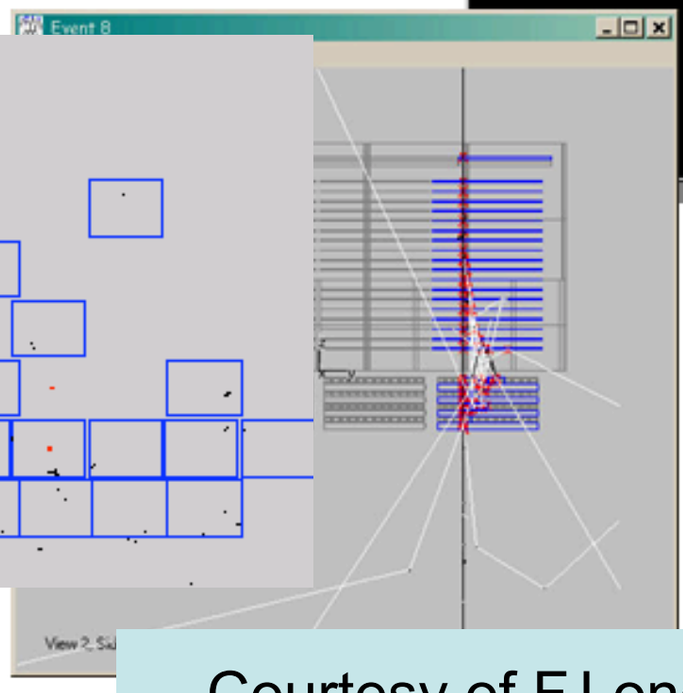
C Simulation

-
-
-
-
-
  - Source generation
  - Hit
  - Dig
  - Re
- G4 ad
- Contir
- MC da







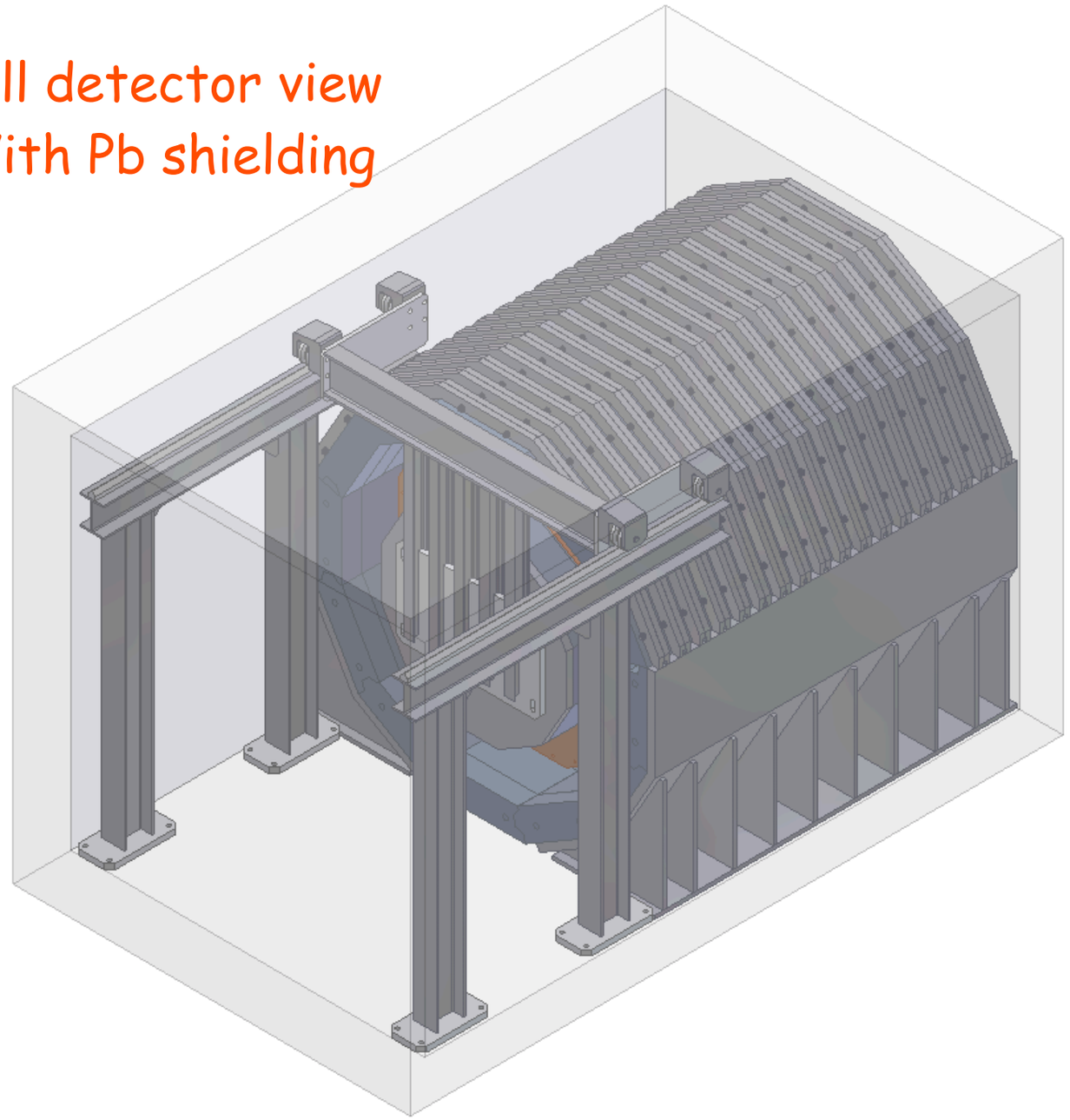


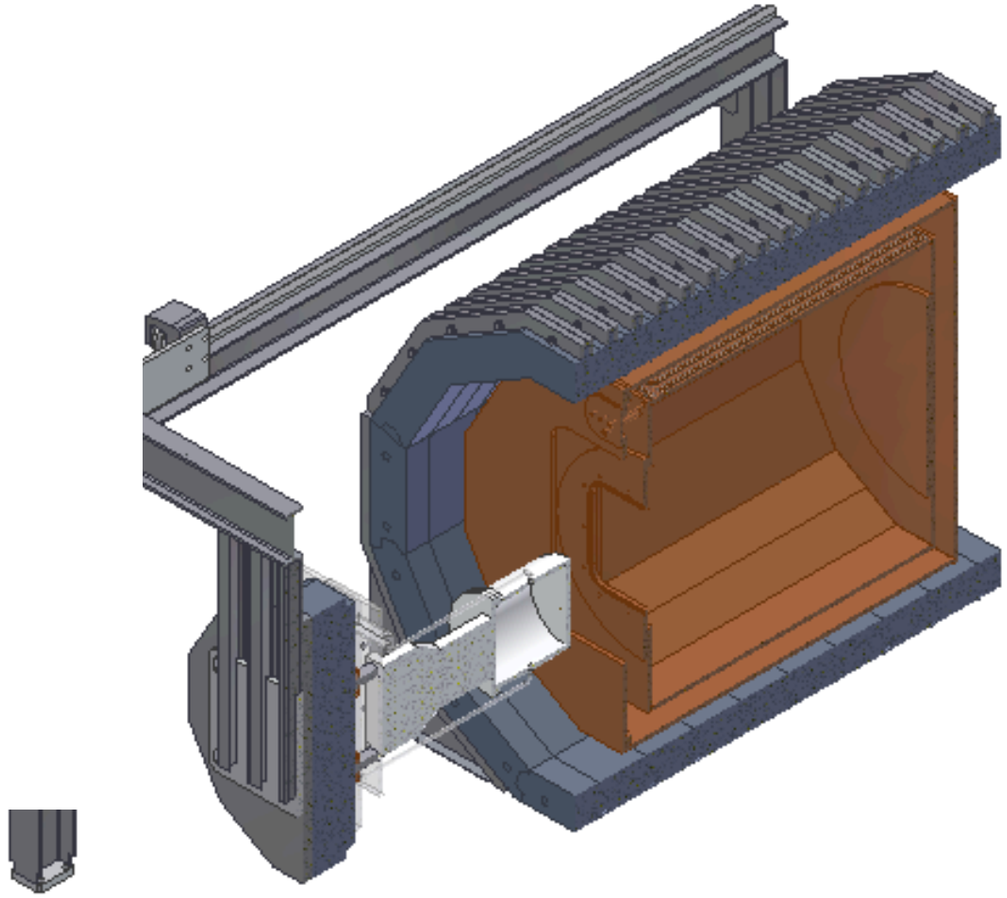
Courtesy of F.Longo (INFN/Trieste)

Geant4 simulation of EXO

- Enriched Xenon Observatory for double beta decay

Full detector view
With Pb shielding





Courtesy of J-L.Vuilleumier (Neuchatel)
and M.Breidenbach (SLAC)

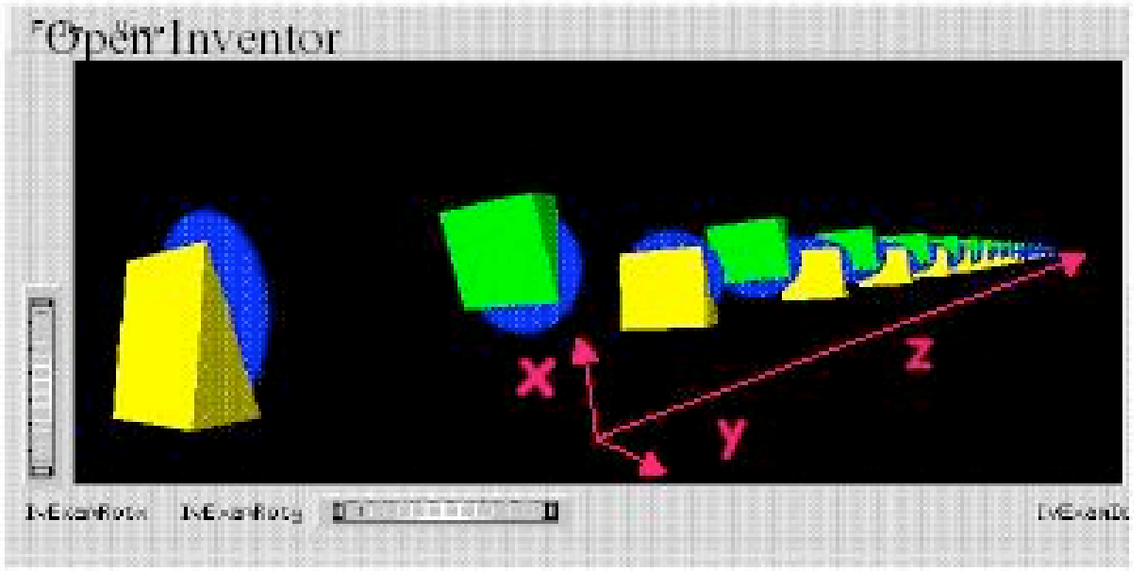
Geant4 for beam transportation

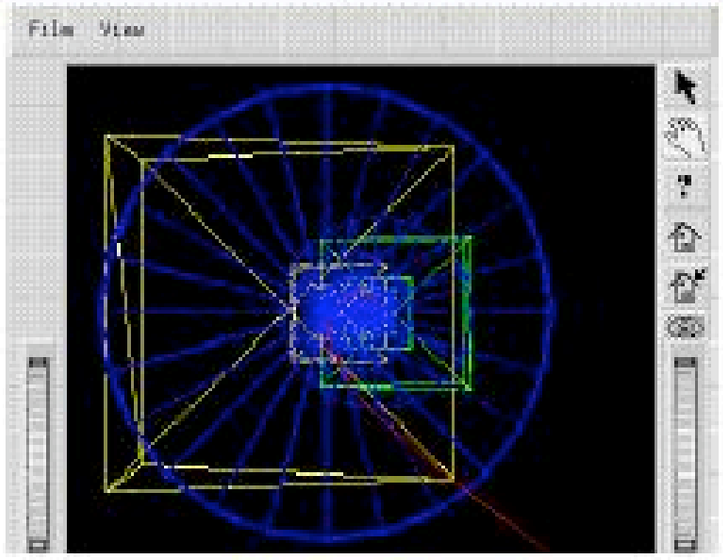
Example: Helical Channel

Published in proc. of PAC 2001
(Fermilab-Conf-01-182-T)

72 m long solenoidal + dipole field with wedge absorbers and thin cavities

$$B_{xy} = B_T \cos,\sin\left(\frac{2p}{L} z\right) \qquad B_z = B_0$$





Other simulations:

- Alternate Solenoid Channel (sFoFo), published in proceedings of PAC2001 and Feasibility Study II for a Neutrino Factory at BNL (2001)
- Bent Solenoid Channel, presented at Emittance Exchange Workshop, BNL 2000
- Low Frequency r.f. Cooling Channel, presented at International Cooling Experiment Workship, CERN 2001
- Cooling Experiment (MICE) Simulation (in progress)

G4 Users Meeting, February 21st, 2002 *V. Daniel Elvira, Fermilab*

Courtesy of V.D.Elvira (FNAL)

Full Simulations

For International Linear Collider experiment

LCDROOT/LCDG4 MOKKA JUPITER

Common GEANT4 executable

Runtime geometry

Generic Hit output

Full Support of SLAC Geant4 Team

Courtesy of N.Graf (SLAC)

Smart-

AC

INTEGRAL

Cassin

LIS

GLAS

Herschel

Bepi Colombo

SWIF

Astro-E2

XMM-Newton

GAI

JWS

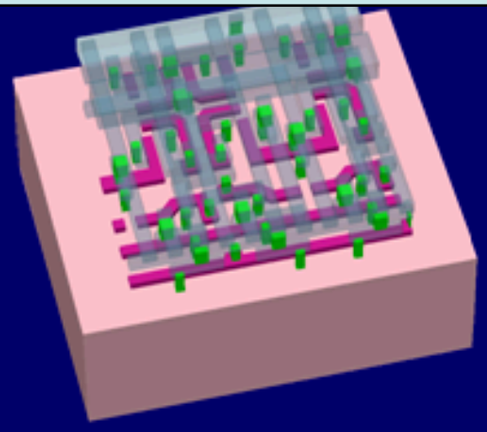
EUSO

AMS

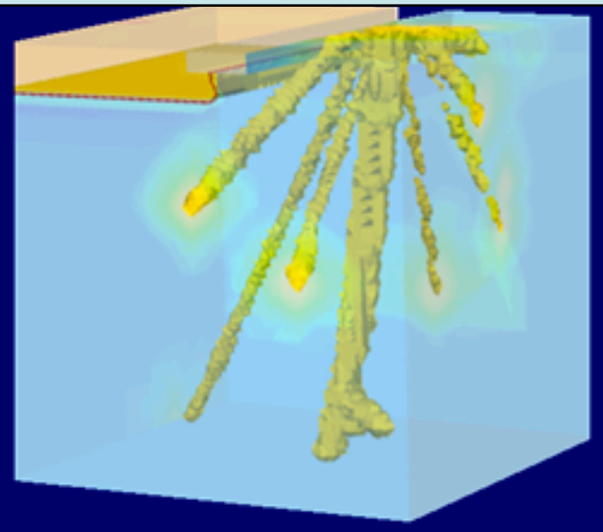
MAX

ISS Columbus

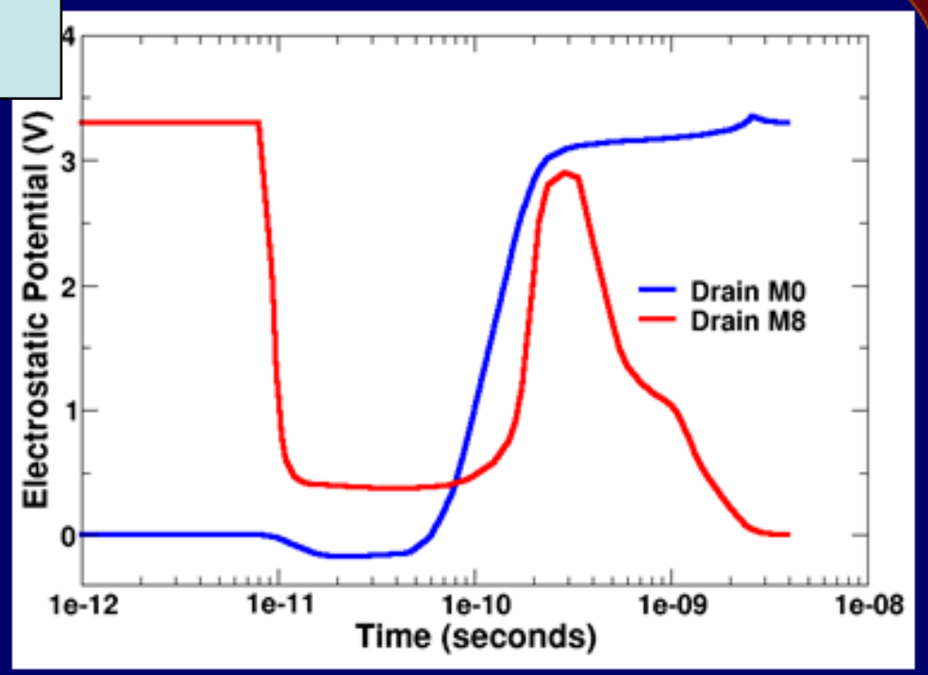
Simulation of Single Event Effects on SRAM in space radiation environments



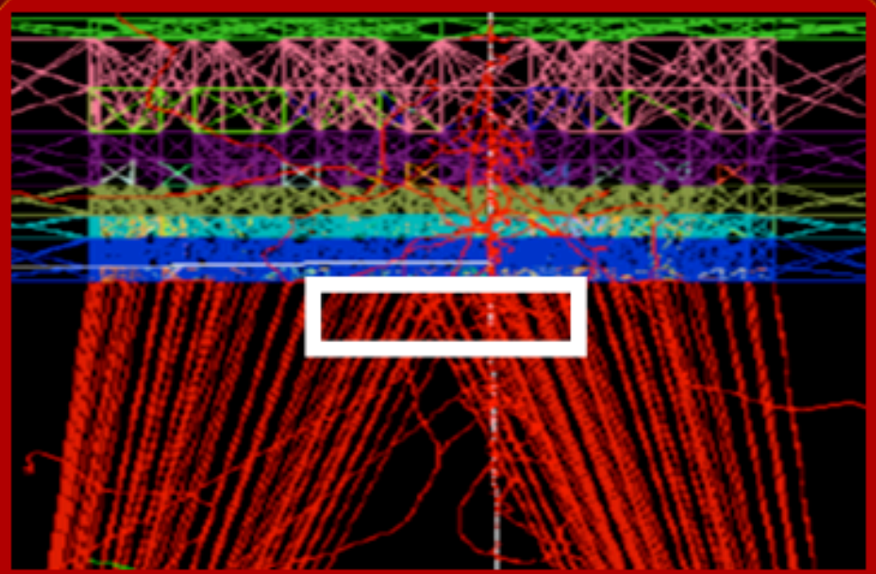
TCAD Cell Structure: SRAM Cell



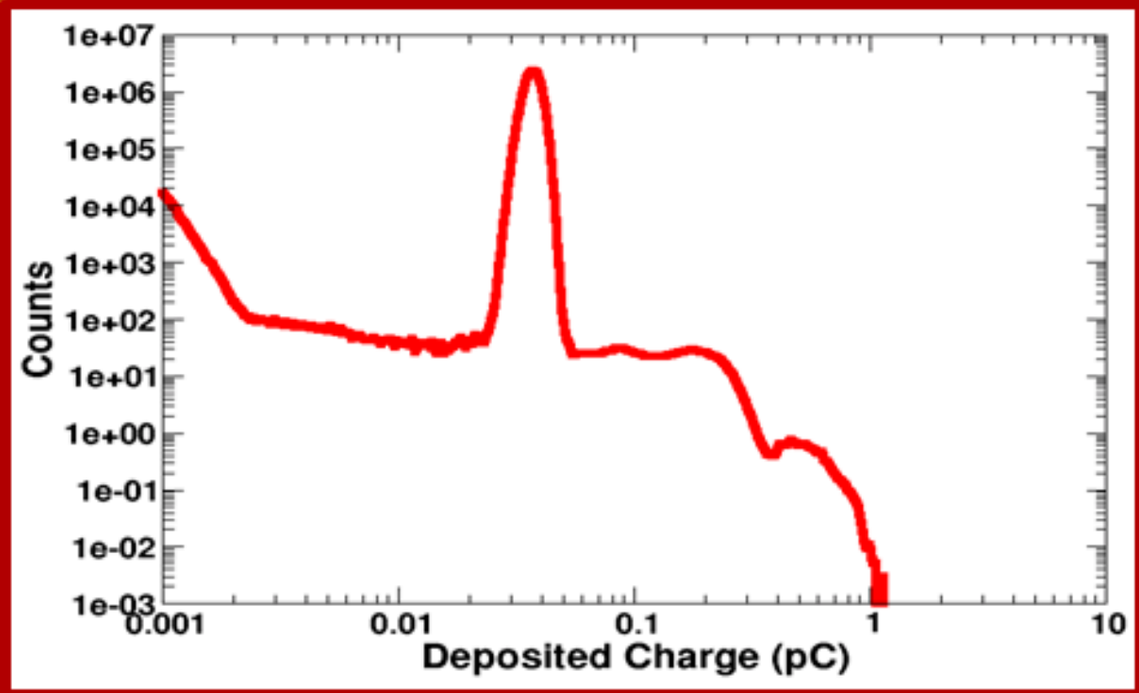
Single Charge Deposition in TCAD: Ne+W Event



SRAM Cell Upset

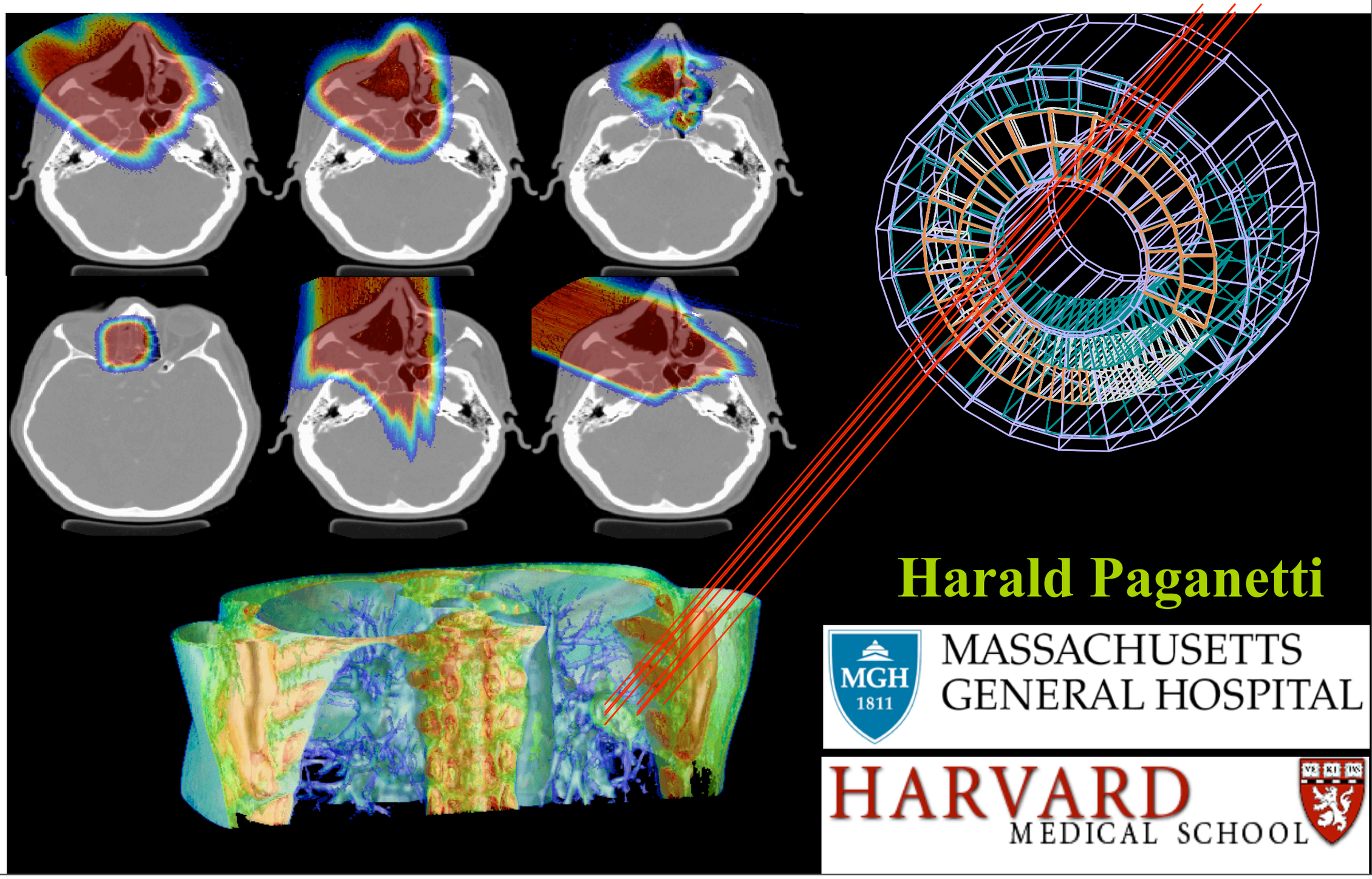


Geant4 Geometry and 523 MeV Neon Event



MRED Energy Deposition for 10^8 Events

GEANT4 based proton dose calculation in a clinical environment: technical aspects, strategies and challenges



Harald Paganetti

MASSACHUSETTS GENERAL HOSPITAL

HARVARD MEDICAL SCHOOL

G4NAMU - the Geant4 North American Medical Users Organization

What...?

G4NAMU was launched in May of 2005 to provide a meeting place for the rapidly growing Geant4 medical user community of North America. The purpose of G4NAMU is to bring this community together to share issues and advice, to develop regional collaboration and to communicate as a group to the Geant4 developers.

Who...?

G4NAMU's current membership includes 69 members from 31 institutions throughout Canada and the United States.

How...?

G4NAMU communicates through a mailing list, geant4-namu@slac.stanford.edu. The list is lightly moderated just to maintain focus and prevent spam.

- To join the list, send mail to perl@slac.stanford.edu.
- For details on list usage, see the Mailing List User Guide.

While many G4NAMU resources are provided by the Stanford Linear Accelerator Center's Geant4 team, G4NAMU is intended to be a consensus organization that evolves as needed by its members.

Another useful resource for medical users is the Geant4 Medical Applications Forum, a resource for medical users worldwide.

Tutorial courses

Geant4 tutorial conducted by SLAC Geant4 team

Date	Place	Capacity	Comments
Feb.18.2002 - Feb.20.2002	SLAC, CA	80	-
Oct.27.2003 - Oct.29.2003	FermiLab, IL	50	-
Jan.12.2004 - Jan.14.2004	Vanderbilt University, TN	50	-
Mar.08.2004 - Mar.10.2004	SLAC, CA	80	-
Mar.03.2005	ESA/ESTEC, Netherlands	-	Co-organized with ESA/ESTEC
Jun.06.2005 - Jun.07.2005	Helsinki Institute of Physics, Finland	-	Co-organized with HIP
Mar.07.2006 - Mar.10.2006	SLAC, CA	50	-
May.22.2006 - May.25.2006	Jefferson Lab, VA	100	Registration closed
Sep.25.2006 - Sep.28.2006	McGill University, Quebec, Canada	60	Registration starts soon
Nov.10.2006	NASA JPL, CA	40	Short cource associated with Geant4 space users workshop

http://top25.sciencedirect.com/index.php?subject_area_id=21

INTRODUCTION

The **ScienceDirect TOP25 Hottest Articles** is a **free** quarterly service from ScienceDirect. When you subscribe to the **ScienceDirect TOP25**, you'll receive an e-mail every three months listing the ScienceDirect users' 25 most frequently downloaded journal articles, from any selected journal among more than 2,000 titles in the ScienceDirect database, or from any of 24 subject areas.

Now you can keep track of the latest trends in your speciality and find out what your colleagues are reading. A simple click on any of the listed articles will take you to the journal abstract, and of course, you have the option to download any article straight to your desktop, depending on your access rights.

Read more about how the TOP25 is generated and what it reflects.

HOTTEST ARTICLES ON ScienceDirect

Check out the ScienceDirect TOP25 Hottest Articles within your area of interest.

Subject Area: Physics and Astronomy

Journal: [all journals]

With these drop-down menus, the ScienceDirect TOP25 Hottest Articles are selected. Please refine your selection if necessary. To see the overall TOP25 within a certain subject area or journal, please select 'all subjects' or 'all journals' from the drop-down menus.

TOP25 articles within the subject area: Physics and Astronomy

1. **Nanoscience and engineering in mechanics and materials** • **Article**
   Journal of Physics and Chemistry of Solids, Volume 65, Issue 8-9, 1 August 2004, Pages 1501-1506
   Chong, K.P.

2. **Geant4-a simulation toolkit** • **Article**
   Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, Volume 506, Issue 3, 1 July 2003, Pages 250-303
   Agostinelli, S.; Allison, J.; Amako, K.; Apostolakis, J.; Araujo, H.; Arce, P.; Asai, M.; Axen, D.; Banerjee, S.; Barrand, G.; Behner, F.; Bellagamba, L.; Boudreau, J.; Broglia, L.; Brunengo, A.; Burk

3. **Radiation pneumonitis and pulmonary fibrosis in non-small-cell lung cancer: Pulmonary function,**

http://www.in-cites.com/hotpapers/2004/november04-eng.html

http://www.in-cites.com/hotpapers/2005/jan05-eng.html

http://www.in-cites.com/hotpapers/2005/mar05-eng.html

http://www.in-cites.com/hotpapers/2005/may05-eng.html

http://www.in-cites.com/hotpapers/2005/july05-eng.html

The Top 3 Hot Papers

Published In The Last 2-Years For:

Engineering

Previous | Field Menu 2006 | Field Menu 2005 | Field Menu 2004
"Super Hot" Papers in Science Published Since 2003

Engineering

(Sorted by citations. 3 of 128)

1 Citations: 133
Title: GEANT4-A SIMULATION TOOLKIT
Authors: AGOSTINELLI S; ALLISON J; AMAKO K; APOSTOLAKIS J; ARAUJO H; ARCE P; ASAI M; AXEN D; BANERJEE S; BARRAND G; BEHNER F; BELLAGAMBA L; BOUDREAU J; BROGLIA L; BRUNENGO A; BURKHARDT H; CHAUVIE S; CHUMA J; CHYTRACEK R; COOPERMAN G; COSMO G; DEGTYARENKO P; DELL'ACQUA A; DEPAOLA G; DIETRICH D; ENAMI R; FELICIELLO A; FERGUSON C; FESEFELDT H; FOLGER G; FOPPIANO F; FORTI A; GARELLI S; GIANI S; GIANNITRAPANI R; GIBIN D; CADENAS JJG; GONZALEZ I; ABRIL GG; GREENIAUS G; GREINER W; GRICHINE V; GROSSHEIM A; GUATELLI S; GUMPLINGER P; HAMATSU R; HASHIMOTO K; HASUI H; HEIKKINEN A; HOWARD A; IVANCHENKO V; JOHNSON A; JONES FW; KALLENBACH J; KANAYA N; KAWABATA M; KAWABATA Y; KAWAGUTI M; KELNER S; KENT P; KIMURA A; KODAMA T; KOKOULIN R; KOSSOV M; KURASHIGE H; LAMANNA E; LAMPEN T; LARA V; LEFEBURE V; LEI F; LIENDL M; LOCKMAN W; LONGO F; MAGNI S; MAIRE M; MEDERNACH E; MINAMIMOTO K; DE FREITAS PM; MORITA Y; MURAKAMI K; NAGAMATU M; NARTALLO R; NIEMINEN P; NISHIMURA T; OHTSUBO K; OKAMURA M; O'NEALE S; OOHATA Y; PAECH K; PERL J; PFEIFFER A; PIA MG; RANJARD F; RYBIN A; SADILOV S; DI SALVO E; SANTIN G; SASAKI T; SAVVAS N; SAWADA Y; SCHERER S; SEIL S; SIROTENKO V; SMITH D; STARKOV N; STOECKER H; SULKIMO J; TAKAHATA M; TANAKA S; TCHERNIAEV E; TEHRANI ES; TROPEANO M; TRUSCOTT P; UNO H; URBAN L; URBAN P; VERDERI M; WALKDEN A; WANDER W; WEBER H; WELLISCH JP;