

70 YEARS OF CREATING TOMORROW



Los Alamos
NATIONAL LABORATORY

Digital Archiving and Storage at LANL

Designing Storage Architectures for Digital Collections

Library of Congress

September 23-24, 2013

Christopher Mitchell, Los Alamos National Laboratory, CCS-7

UNCLASSIFIED



Operated by Los Alamos National Security, LLC for the U.S. Department of Energy's NNSA

LA-UR-13-27317



Context of LANL's Digital Storage Needs

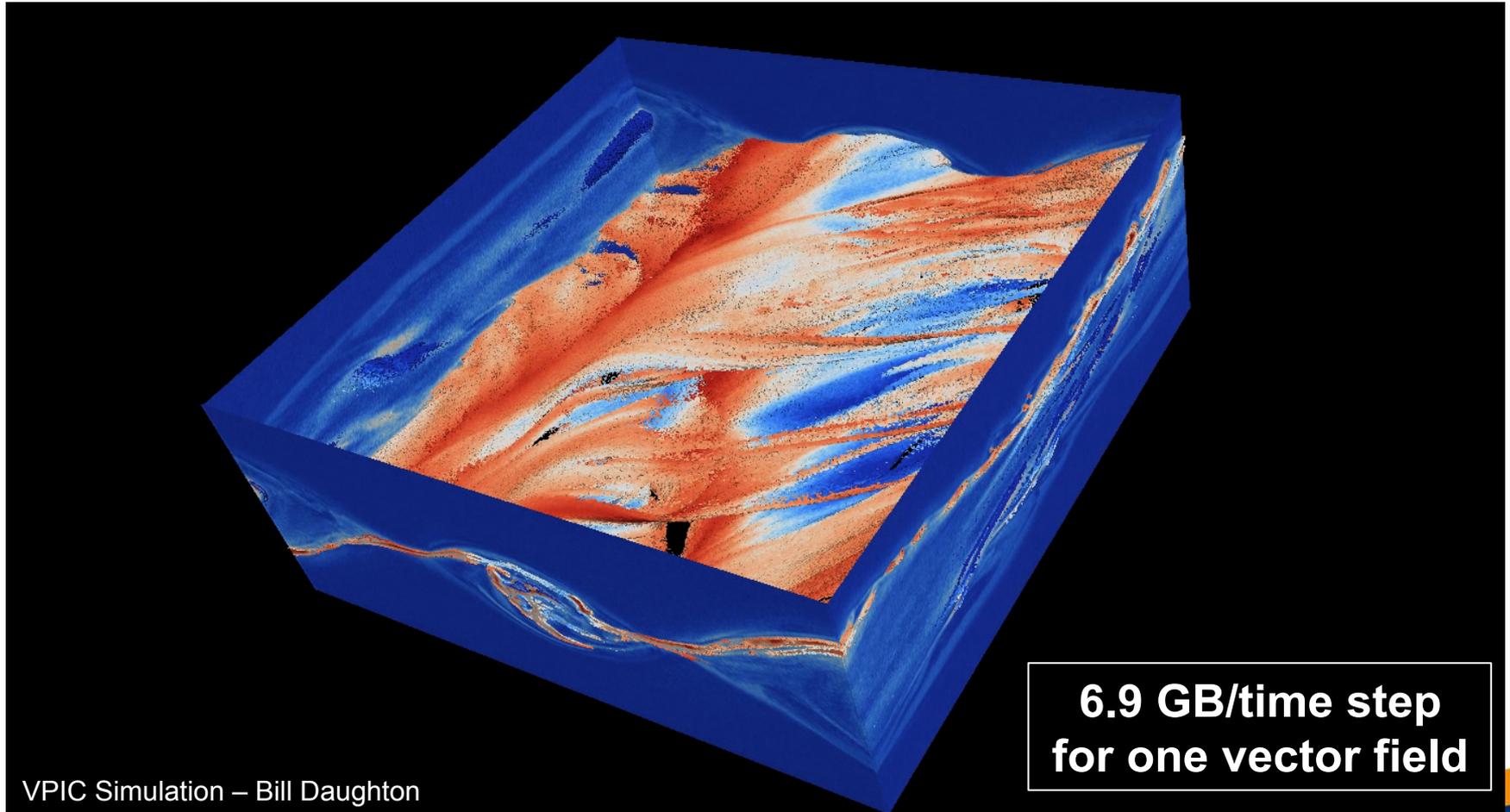
- LANL's mission is to develop and apply science and technology to ensure the safety, security, and reliability of the U.S. nuclear deterrent; reduce global threats; and solve other emerging national security and energy challenges.
- We run both experiments and large HPC simulations across a wide variety of disciplines.
- Our largest supercomputers routinely place in the Top 10 of the TOP500 list (Cielo, Roadrunner, back to the CM-5) and run simulations at scale.

UNCLASSIFIED





Example: Plasma Simulation Data



UNCLASSIFIED





Current HPC Storage Infrastructure

- Looking at our largest archive system, HPSS, with all instantiations summed:
 - 5 SL8500 tape libraries with 138 tape drives
 - 38.6 PB of used space for 410 million archived files.
 - 30,074 tape cartridges in use.
- Other archive systems in use have differing balances of tape to disk for permanent storage and use other technologies (ex. GPFS & TSM combined)
- **Takeaway:** We have large amounts of scientific data from our HPC systems across many storage mechanisms.

UNCLASSIFIED





Sampling of Relevant R&D Efforts

- In-Situ Analysis
 - Move from a “save all raw data” to “save data analysis products” mode of thinking.
 - Decreases storage resource requirements.
 - Forces scientists to treat simulation the same as an in-the-field experiment. Leverages “FLOPS are free” mindset with new HPC systems.
- VisIO & Burst Buffers
 - Make data analysis faster to enable easier data product creation and exploration.
- Archive Metadata Search
 - Working with LANL Library to develop a search tool & curation mechanism for our scientific data archives to allow for search on tags, data features, etc.

UNCLASSIFIED

