Cray Solutions for Earth System Modeling

Presented to the International Workshop on Computing in the Atmospheric Sciences 2013

Per Nyberg
nyberg@cray.com
Director, Business Development

CRAY®
THE SUPERCOMPUTER COMPANY

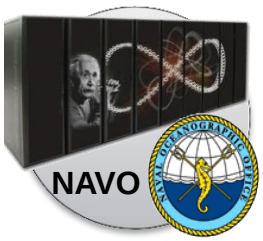
Copyright 2012 Cray Inc. All rights reserved. Cray and the Cray logo are trademarks of Cray Inc.

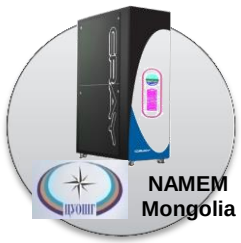
Topics

- Cray's Presence in the Earth System Modeling Community
- Earth System Modeling on Cray systems
- Evolution of Science and Knowledge Discovery
- Cray Vision

Cray's Presence in the Weather and Climate Community

- **Earth System Modeling is a key area on Cray Systems worldwide:**
  - Dedicated operational NWP and research centers
  - Multi-disciplinary research centers
  - From Teraflops to Petaflops
- **Recent Cray XC30 wins at DWD, ECMWF and NEA**





Deutscher Wetterdienst (DWD)

- **The petascale facility will be composed of two Cray XC30s and a Cray Sonexion storage system.**
  - Configured for full redundancy with a shared global file system.
- **To be located at DWD's facility in Offenbach, Germany**
- **Will enable DWD to produce higher resolution and more accurate global and regional weather forecasts.**

European Centre for Medium-Range Weather Forecasts (ECMWF)

- **Cray will provide a multi-petaflops supercomputing infrastructure at ECMWF designed for operational resiliency:**
  - Composed of two Cray XC30 systems and a multi-petabyte Cray Sonexion storage system.
  - Operational in 2014.
- **The new system will help improve ECMWF's forecasting capabilities by providing high performance computing to support:**
  - Increased resolution and model enhancements.
  - Development of ensemble-based data assimilation methods in conjunction with the Ensemble Prediction System.
  - Better representation of physical processes and further increased use of satellite observation data.

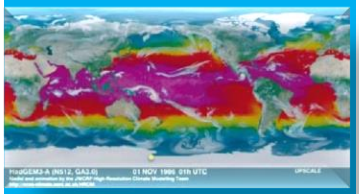
Singapore National Environment Agency (NEA)

- NEA recently selected Cray for the provision of their next supercomputing facility.
- NEA is the leading public organization responsible for improving and sustaining a clean and green environment in Singapore.
- Provides timely weather information to support public safety and socio-economic activities, including haze alerts.
- 55+ Tflop/s Cray XC30.

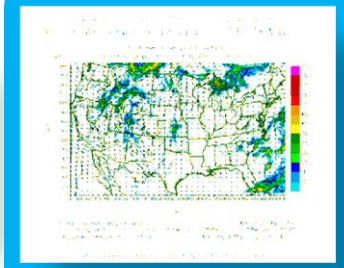
Some Additional Recent Installations and Wins

- **Leadership facilities worldwide that support significant earth system modeling communities, including**
  - NCSA Blue Waters Cray XK7: >13 PF
  - ORNL Titan Cray XK7: >27 PF
  - NERSC Edison Cray XC30: >2 PF
  - EPSRC/EPCC Archer Cray XC30

Leading Earth System Modeling Cray Systems

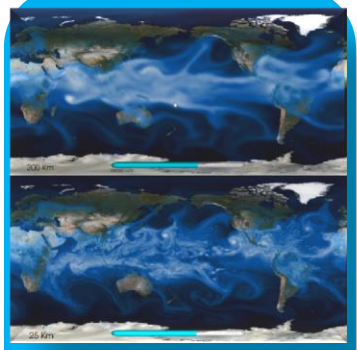


- UPSCALE project on HLRS Cray XE6 “Hermit”: UK on PRACE: weather-resolving Simulations of Climate for globAL Environmental risk
- UK Met Office and National Centre for Atmospheric Science (NCAS)
- Improve understanding of climate processes through 25 year simulations at “weather resolution”

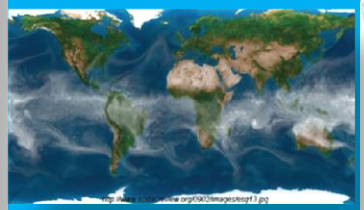


- NOAA/CAPS Hazardous Weather Testbed (HWT) 2013 Spring Experiment
- Used NICS' Cray XC30 “Darter”
- Real-time baseline storm-scale ensemble forecasts.
- Aimed at improving hazardous-weather prediction lead to increased severe-storm-warning times for the public

- Understanding Tornadoes and Their Parent Supercells Through Ultra-High Resolution Simulation & Analysis
- NCSA Blue Waters Petascale Computing Resource Allocation (PRAC)



- CAM 25km Simulations on NERSC Cray XE6 “Franklin”
- Principal motivations was to quantify hurricane statistics in a changing climate.
- Also interested in other extreme weather statistics.



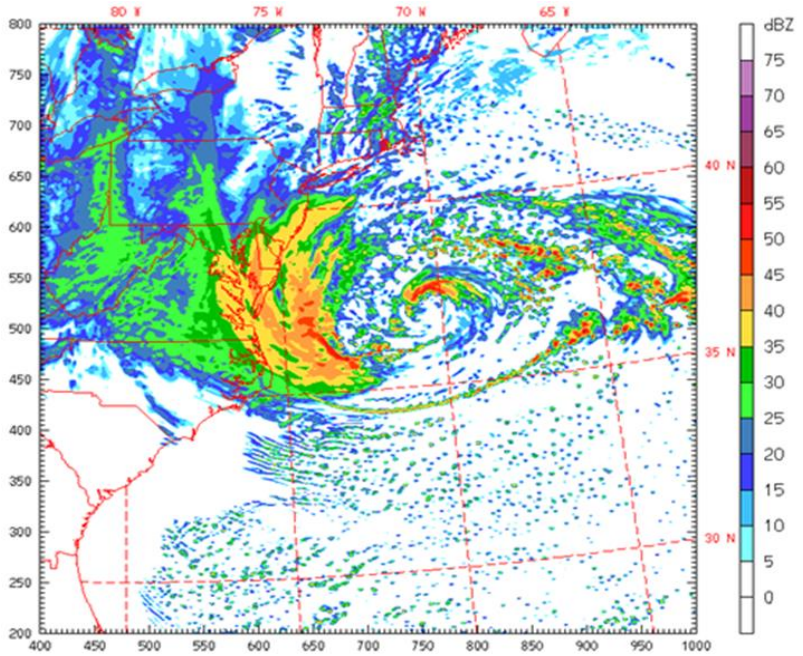
- CAM/HOMME cubed sphere spectral element community atmospheric model
- ORNL Titan early science application
- Answer questions about specific climate change adaptation and mitigation scenarios

WRF Hurricane Sandy Simulation on Blue Waters

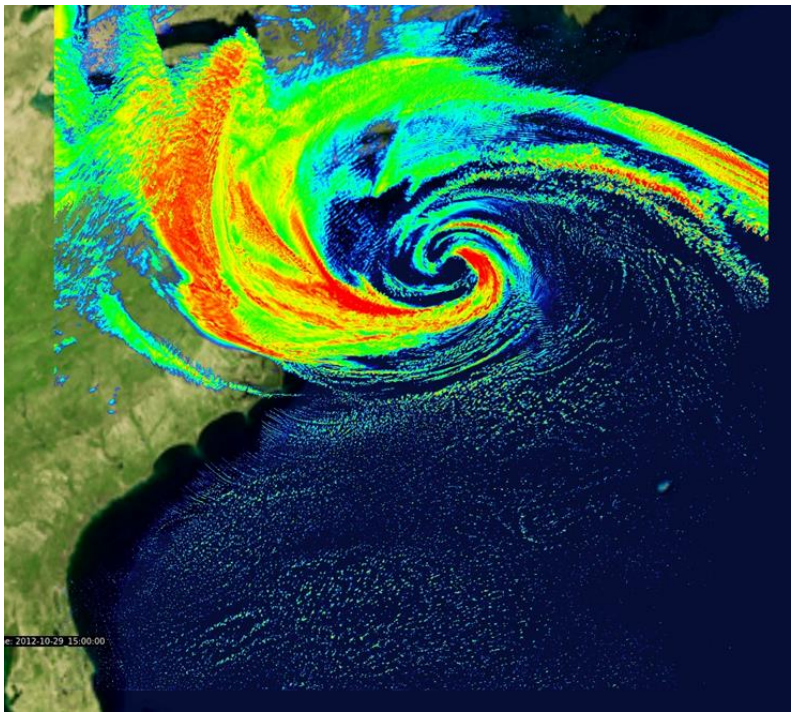
- **Largest ever storm prediction model using real data of over 4 billion points used to simulate the landfall of Hurricane Sandy with WRF application.**
  - Grid size of 9120x9216x48 @ 500m resolution
- **NCSA "Blue Waters" Cray XK7 sustained 285 Tflops using 437,760 cores simulating an 18-hr forecast.**
- **SC13 paper has been submitted: "Petascale WRF Simulation of Hurricane Sandy", M. Shapiro et al**
- **Performance will enable research into potential for greater accuracy when predicting exact landfall time and place, and wind and water damage.**

WRF Hurricane Sandy Simulation on Blue Waters

- Initial analysis of WRF output is showing some very striking features of Hurricane Sandy. Level of detail between a 3km WRF simulation and BW 500meter run is apparent in these radar reflectivity results



3km WRF results



Blue Waters 500meter WRF results

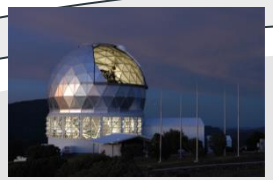
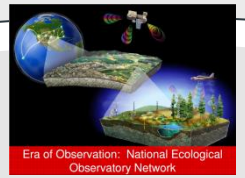
Applications R&D for the Extreme Scale

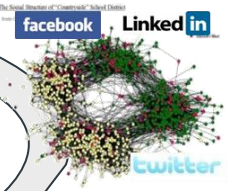
- **Cray has been engaging in a number applications focused petascale and exascale activities involving the earth system modeling community, including:**
  - Swiss Platform for High-Performance and High-Productivity Computing - HP2C
    - Implementations of key applications on hybrid architectures
  - UK Met Office / NCAS Project UPSCALE
    - Science on current Petascale systems through improving scalability of UM
  - ICOsahedral-grid Models for EXascale Earth system simulations (ICOMEX)
    - Exascale implementations of next generation models
  - Titan Early Science Applications
    - Pre-exascale implementations of key applications using GPUs
  - Blue Waters Petascale Computing Resource Allocations (PRAC) Teams
    - Sustained Petaflop applications
  - CRESTA Project – EU funded exascale co-design
    - Co-design approach to exascale application implementation

Evolution of Science and Knowledge Discovery

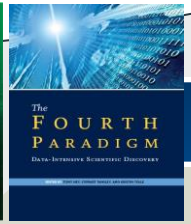
Sensors, devices, simulations, social,…



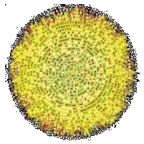



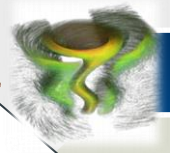


Multi-disciplinary, multi-institutional



Data Intensive / Data Driven Science





Computational Simulation of Complex Phenomena

$$\mathbf{x}_a - \mathbf{x}_b = \mathbf{B}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}(\mathbf{y} - \mathbf{H}\mathbf{x}_b)$$

Theoretical Research



Experimental Research

U.S. DEPARTMENT OF ENERGY
Office of Science

Heterogeneous Data: Environmental Science Example

- Mobile stations
- High-resolution weather stations
- Full-size snow/weather stations
- External weather stations
- Satellite imagery
- Weather radar
- Mobile weather radar
- Stream observations
- Citizen-supplied observations
- Ground LIDAR
- Aerial LIDAR
- Nitrogen/methane measures
- Snow hydrology & avalanche probes
- Seismic probes
- Distributed optical fiber temperature sensing
- Water quality sampling
- Stream gauging stations
- Rapid mass movements research
- Run-off stations
- Soil research

Source: Lehning, Michael et al, "Instrumenting the Earth: Next-Generation Sensor Networks and Environmental Science" in The Fourth Paradigm: Data-Intensive Science, ed. Tony Hey, http://research.microsoft.com/en-us/collaboration/fourthparadigm/4th_paradigm_book_complete_lr.pdf

14

Data and Computational Drivers

- **Today's science is:**
  - Data-intensive
  - Data-driven
  - Compute-intensive
  - Multi-scale, multi-physics
- **Resource requirements are driven by multiple dimensions.**
- **Data Tsunami is defying standard approaches to interpretation**
  - Volume and complexity of data are too much for either humans or current technologies for effective analysis

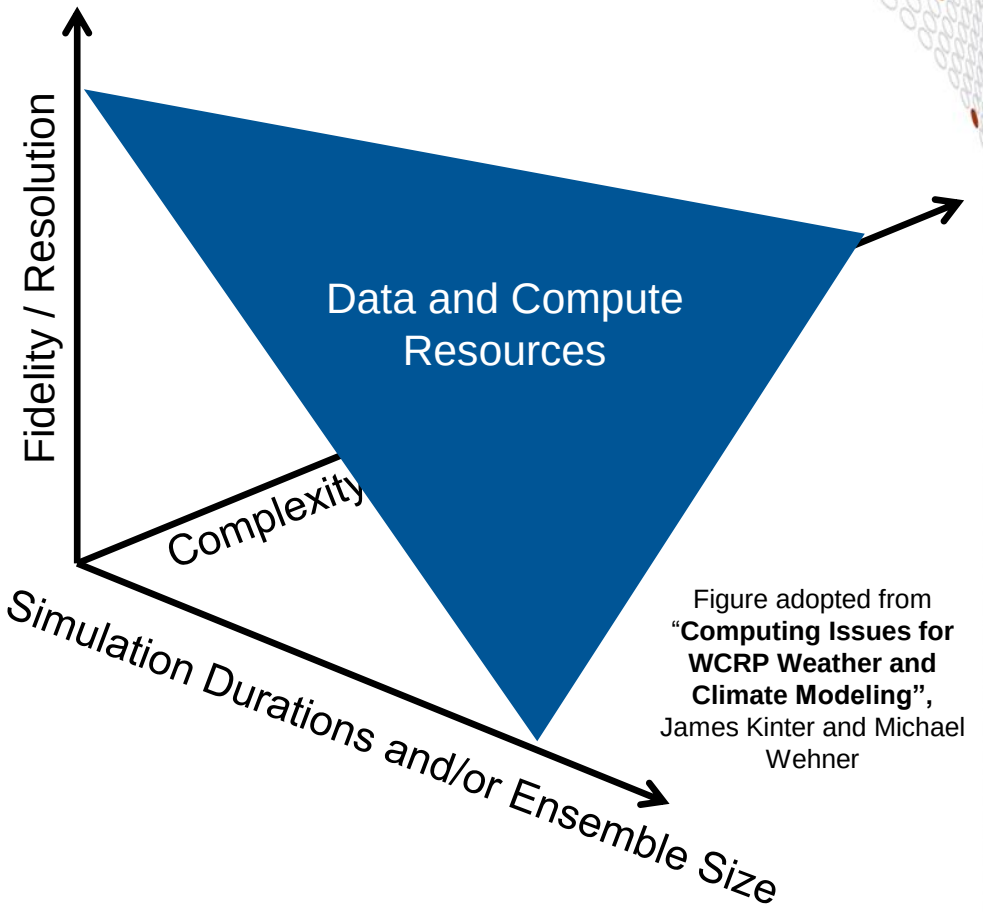
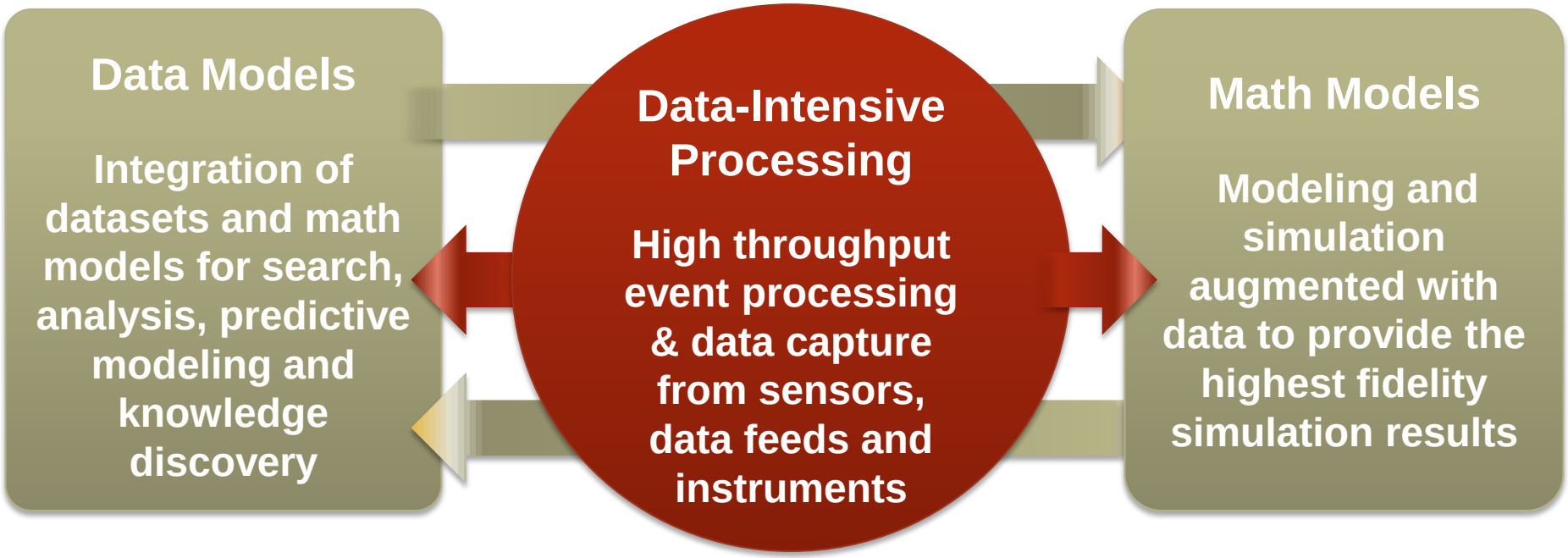


Figure adopted from "**Computing Issues for WCRP Weather and Climate Modeling**", James Kinter and Michael Wehner

Data driven discovery and advanced analytics are rapidly becoming a competitive differentiator providing insight and predictive capabilities.

Cray's Vision:
The Fusion of Supercomputing and Big & Fast Data

Modeling The World
Cray Supercomputers solving "grand challenges" in science, engineering and analytics



Advanced Analytic Appliances

Storage & Data Management

Supercomputers

Modeling The World

Data Models

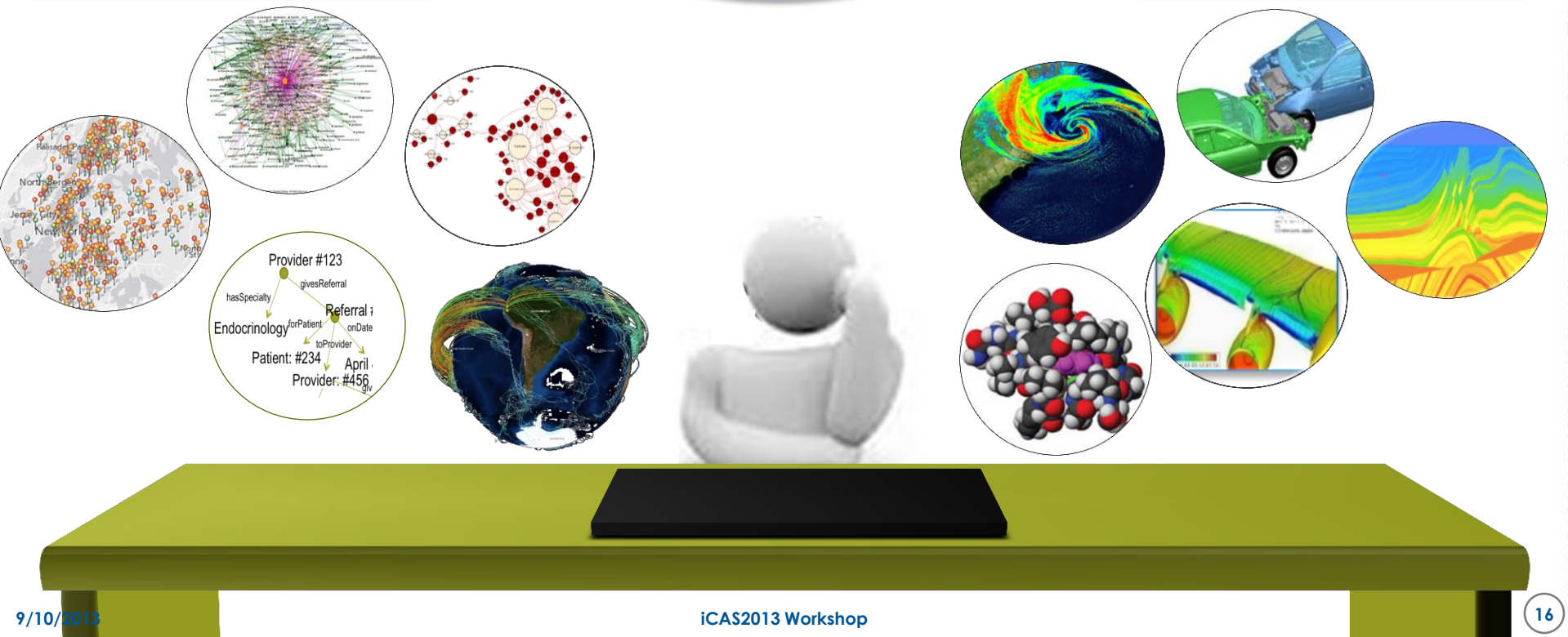
Integration of datasets and math models for search, analysis, predictive modeling and knowledge discovery

Data-Intensive Processing

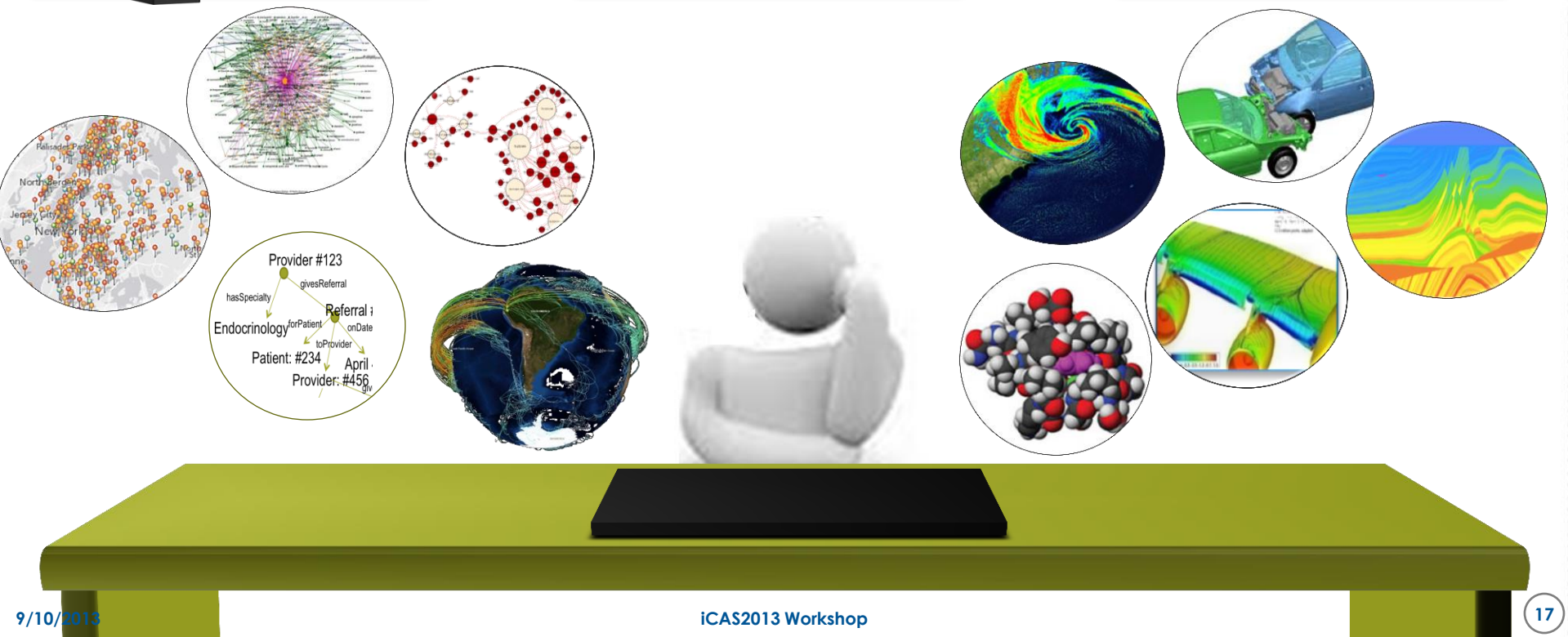
High throughput event processing & data capture from sensors, data feeds and instruments

Math Models

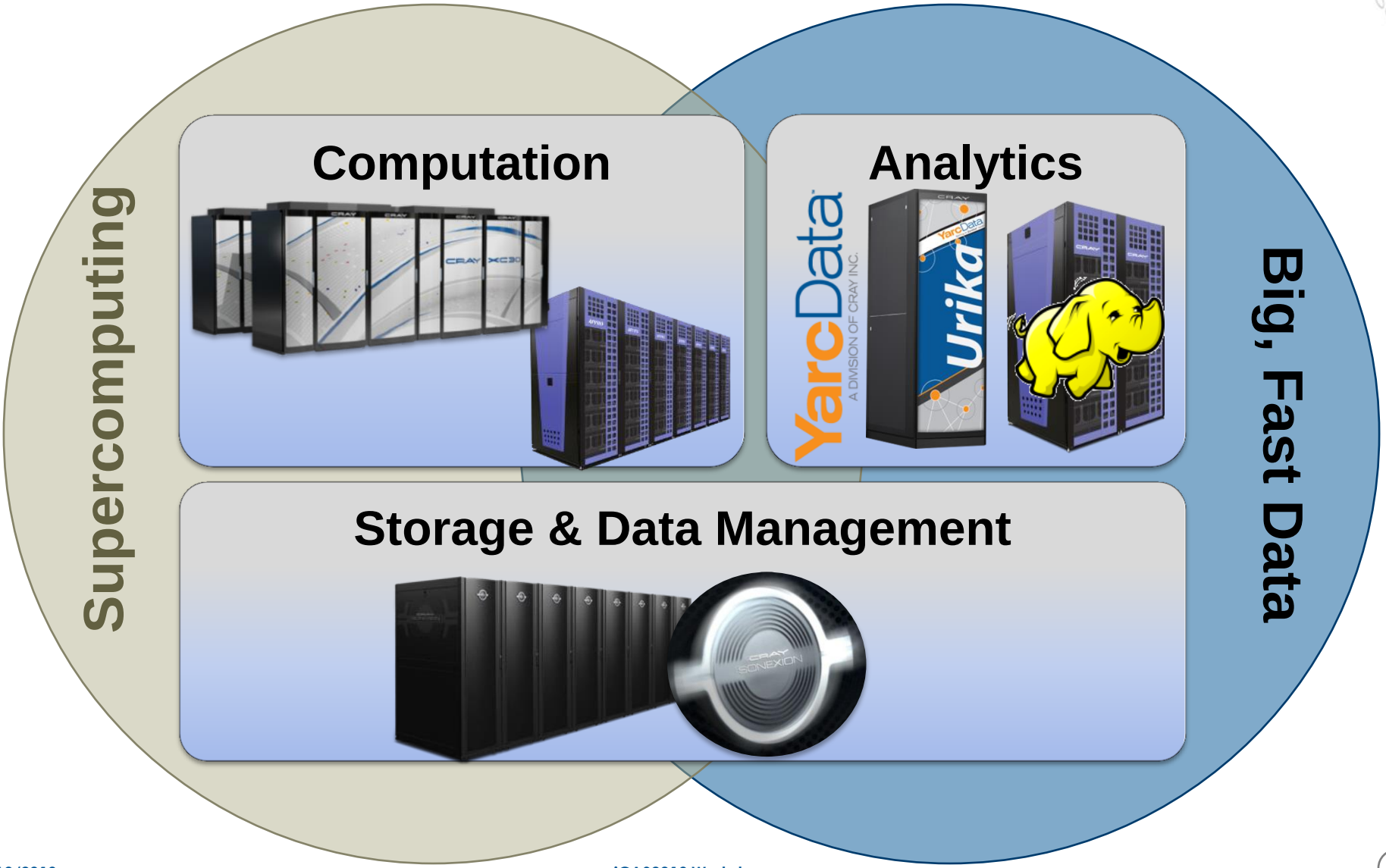
Modeling and simulation augmented with data to provide the highest fidelity simulation results



Integrated HPC Environments are the capability that will turn data in to insight and discovery.



Compute, Store, Analyze



Cray HPC Computing Solutions

Cray CS300 Cluster Supercomputers

Configurable Industry standard scale out computing

Flexible configurations to provide application oriented characteristics

TCO optimized energy-efficient design

Complete turn-key system with integrated HPC software stack powered by Advanced Cluster Engine (ACE)

Cray experience in large cluster deployment

Cray XC30 Supercomputers

Highly differentiated architecture for the most challenging HPC environments

Highly integrated architecture designed for robustness and extreme scalability

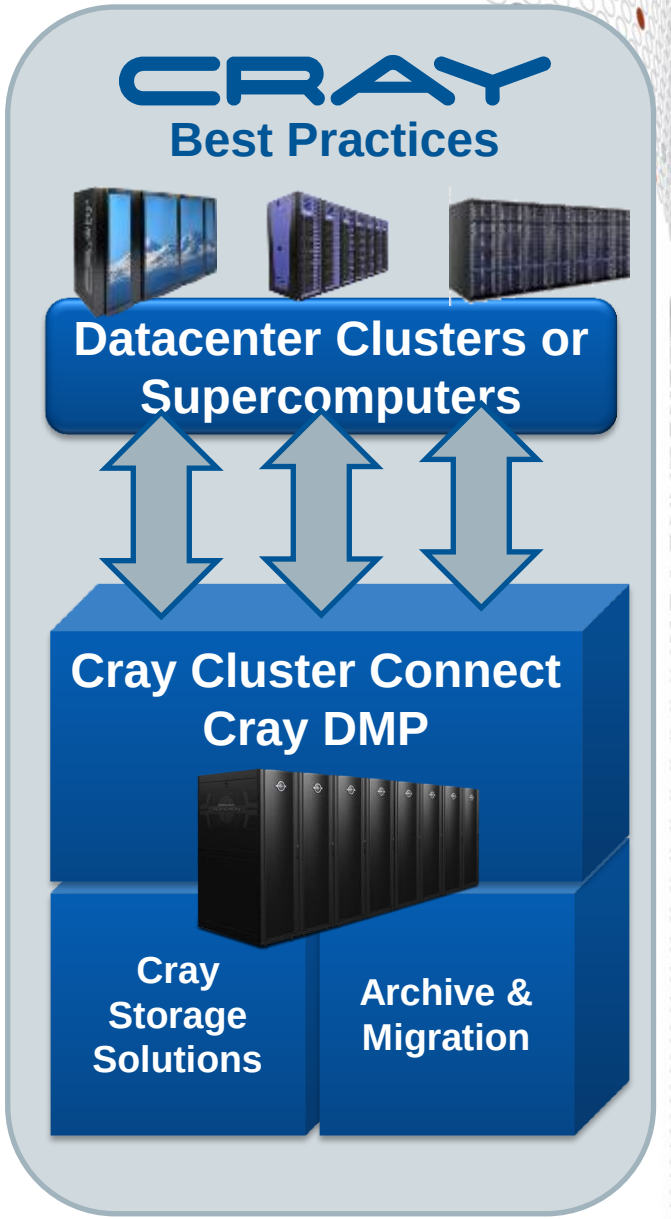
Field proven Cray Linux Environment

TCO optimized energy-efficient design

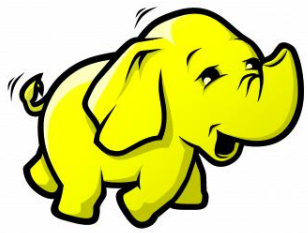
Designed to span multiple generations of technologies and performance increases.

Scalable Storage & Data Management Solutions

- **Data management and scalable parallel I/O performance for the most demanding HPC environments**
- **Cray Cluster Connect**
  - Complete storage and data management solutions for cluster environments
- **Cray Data Management Platform**
  - Lustre File Systems by Cray
  - Cray Development and Login
  - Cray Management Services
  - Cray Migration and Archive
- **Cray Storage Systems**
  - Cray Sonexion™
- **Cray Monitoring and Management Tools**



Cray Cluster Supercomputers for Hadoop: Purpose-Built, Turnkey, Hadoop Solutions



Best Hadoop Distribution

- **Security**– Comprehensive, and fast, encryption
- **Performance** – Faster Hive, Cache acceleration, etc.
- **Management** – Intel Manager for Hadoop Software

+

Performance of a Cray

- **Proven HPC** – Cray HPC technology and expertise
- **Vast Scale** – Grow to meet any mission requirements
- **Holistic Design** – Balanced: Compute, networking, & Storage

+

Turnkey Solution

- **Reliable** – Rapid ROI… runs as-advertised
- **Support** – One throat to choke, for the whole stack
- **Maintenance** – Update & evolve, without concerns

=

High Value Hadoop

- **Performance** – Power to accommodate current & future goals
- **Reliability** – Will meet any challenge, without surprises
- **Maintenance** – Easy to maintain & accommodate change

Discovery Through Graph Analytics…

Discover Unknown and Hidden Relationships in Big Data

- Relationship Warehouse supporting Inferencing/Deduction, Pattern-based queries and Intuitive Visualization

Perform Real-time Analytics on Big Data Graph Problems

- High-performance, Graph Appliance with large shared-memory, massive multi-threading and scalable I/O

Realize Rapid Time to Value on Big Data Solutions

- Ease of Enterprise adoption with industry-standards, open-source software stack enabling reuse of existing skillsets and no lock-in

"We are also impressed with… its product strategy, which combines the advantages of a pre-integrated hardware appliance with the flexibility of a subscription model."

451 Research

Summary

- **Cray is a technology provider with a proven record of deploying production solutions and delivering customer oriented innovation.**
- **Our commitment to long-term partnerships provides significant value to our customers.**
- **Cray offers a complete, single-vendor solution with systems to meet any size need in computing, storage and data management, and large-scale data analytics.**
- **In the context of weather and climate centers, Cray systems support:**
  - A broader and more detailed range of modeling capabilities
  - Reliable on-time delivery of products and services
  - Shortened deployment time and R&D cycles, improving time to operational introduction of new science


CRAY
THE SUPERCOMPUTER COMPANY