EARTH SCIENCES: CORE DRIVERS

Performance: Drive continued improvements in fidelity of weather & climate simulations

Reliability: Maintain performance, reliability & serviceability as systems grow in size & complexity

Analysis: Derive greater value from environmental data, both observations and simulation

Vision: Manage 10-year transition/rewrite of applications to exploit 2025 machine architectures
EARTH SCIENCES: WHY CRAY?

Over 80% of the world’s operational weather forecast centers use Cray systems.

- Reliability
  - Operationally proven, unrivalled experience
- Performance
  - Balance performance & throughput across workflow
  - Software development environment, performance tools & application support experts
- Long-term customer partnerships
- New analysis approaches through converged systems
CRAY GROWTH IN WEATHER, CLIMATE & OCEANOGRAPHY


Global NWP Centre  Met Office  Ministry of Earth Sciences  Government of India  Japan Meteorological Agency  IBS Center for Climate Physics  Taihoru Nukurangi  ZAMG  ECMWF

© 2018 Cray Inc.
Supercomputers are Critical to Simulation

- **Largest ever storm prediction model**
  - Over 4 billion points used to simulate the landfall of Hurricane Sandy
  - Urban scale grid resolution of 500m (compared to standard 3km)
  - Enables the research to understanding fine grained properties of hurricanes

- **Key to the development of new antiretroviral drugs**
  - Determined the precise chemical structure of the HIV capsid – the protein shell that protects the virus’s genetic material and is a key to its virulence.
  - Requires the assembly of more than 1,300 identical proteins – in atomic-level detail.

- **Studying crop devastation by whiteflies to address a major cause of hunger in East Africa**
  - Understanding the DNA of the species by generating phylogenetic trees
  - With only 500 whiteflies in a genetic dataset, the possible relationships between these flies run into the octillions \(10^{25}\)
Crop data is key to decision makers
- Applying Deep Learning on satellite data, the two major crops can be distinguished with 95% accuracy just a few months after planting and well before harvest.
- More timely estimates could be used for a variety of applications, including supply-chain logistics, commodity market future projections, and more.

Quantitatively assess how extreme weather will change in the future
- A single climate simulation can produce over 100TB of data with archives reaching over 5PB.
- Deep learning techniques are ideal for pattern recognition over large data sets

Development of systems for connected cars and autonomous technologies
- These advances could not have been realized without the application of deep learning to object detection in image and full motion video.
MUCH More Data to Serve, Store and Manage

**MORE INPUT DATA**
- More powerful instruments
- More powerful sensors
- Advanced Analytics in the workflow
- AI (ML/DL) in the workflow

**MORE OUTPUT DATA**
- Bigger problem sizes
- Higher fidelity models
- Higher resolution models
- New data-intensive algorithms

Compute

© 2018 Cray Inc.
Not Just More Data But Also Different I/O Patterns

- Large, streaming I/O (HDDs shine)
- Small, random I/O (SSDs shine)

Modeling & Simulation
Advanced Analytics
Artificial Intelligence
AS DATA VOLUME INCREASES

• More focus on ability to handle and use data efficiently
  • Analyze ensemble output
  • Apply AI techniques to both observations and model output to identify severe weather features
• Improve accuracy through better data assimilation methods
• Interconnect focus evolves to meet these needs
  • Less on how it enables single job MPI scalability
  • More focus on how the interconnect provides excellent data access
  • More focus on congestion management to enable MPI- and I/O-dominated jobs to get reproducible performance
Cray Vision
Convergence of Supercomputing and Artificial Intelligence

**Supercomputing**
- Large and Latency Sensitive Data Movement
- Massively Scalable Computing
- Dense Accelerated Computing
- Parallel Storage and I/O

**Artificial Intelligence**
- AI Workflow
- Machine Learning and Deep Learning
- Graph Analytics
- Apache Spark™ and Python-based data science

**Data-Intensive Processing**

- Simulation
- AI & Machine Learning

**Big Data Analytics**
CRAY CONVERGED ARCHITECTURE VISION
Slingshot will be a great interconnect for our earth sciences customers.

This is Cray’s 8th supercomputing interconnect.

- MPI Performance
- Traffic Classes
- Congestion Management
- Flexibility
- Reliability
- Ethernet Compatibility

Very low average and TAIL latencies improve runtime reproducibility.

Better data ingress/egress.
New era – end of Dennard scaling – can’t get faster performance just from shrinking silicon
→ specialization for different kinds of tasks
Need for performance portability
STORAGE TRENDS – FLASH COSTS DECLINE

Today

- **DataWarp** (stage/de-stage via WLM, different namespace)
- **ClusterStor L300N and NXD** (cache - transparent to users)
- **ClusterStor L300F** (all flash, small, random I/Os, 500,000 IOPs)
- Local SSDs

Future

- Lustre tiering and additional Lustre features

---

**ECONOMIES OF STORAGE MEDIA**

~60% of the cost of any storage system is in the media

**SSD Media =**
Most of performance + some initial capacity

**HDD Media =**
Some performance + rest of capacity

**SOFTWARE =**
Workflow-accelerating data placement on the right media at the right time

$0.400
$0.350
$0.300
$0.250
$0.200
$0.150
$0.100
$0.050
$0.000

2018 2019 2020 2021 2022 2023 2024

<table>
<thead>
<tr>
<th>Year</th>
<th>SSD $/GB</th>
<th>HDD $/GB</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>$0.380</td>
<td>$0.250</td>
</tr>
<tr>
<td>2019</td>
<td>$0.250</td>
<td>$0.180</td>
</tr>
<tr>
<td>2020</td>
<td>$0.180</td>
<td>$0.140</td>
</tr>
<tr>
<td>2021</td>
<td>$0.140</td>
<td>$0.110</td>
</tr>
<tr>
<td>2022</td>
<td>$0.110</td>
<td>$0.081</td>
</tr>
<tr>
<td>2023</td>
<td>$0.081</td>
<td>$0.060</td>
</tr>
<tr>
<td>2024</td>
<td>$0.060</td>
<td>$0.000</td>
</tr>
</tbody>
</table>

May 2018, IDC

© 2018 Cray Inc.
Hyper-scale web companies have developed many interesting technologies:

- Driven by data velocity, volume & variety, combined with resiliency requirements at scale

Several are of particular interest in weather & climate science:

- Containerization
- Flexible & scalable data analysis platforms
- New analysis techniques and machine learning
 CONTAINERS FOR FLEXIBILITY

The ability to create your own world supports a diverse workload on a shared supercomputer

Build with Certified Software Stacks

Bundle libraries and dependencies

Build a consistent Environment from desktop to supercomputer
URIKA-XC MAKE ANALYTICS AND GRAPH “FIRST CLASS” CITIZENS ON XC SERIES SYSTEMS

Expanding to Analytics and Open Data Science

**Python Open Data Science**
- Production Supercomputing
- Weather Forecasting
- Seismic Imaging
- Manufacturing CAE
- Scientific Supercomputing
- Climate Science
- Chemistry & Materials Science

**Spark Big Data Analytics**
- Data Preparation
- Analysis
- Visualization
- Machine Learning
- Deep Learning

**Large-Scale Graph Discovery**
- Cancer Cell Morphology
- Fraud and Insider Threat Detection

Deployed using containers for portability and ease of use
CRAY® URIKA-CS AI SUITE

Pre-integrated and supported AI stack with popular open source AI frameworks and libraries delivered as container images for ease of development and deployment

<table>
<thead>
<tr>
<th>UIs: Jupyter Notebooks, TensorBoard</th>
</tr>
</thead>
<tbody>
<tr>
<td>Java, Scala, R, Python</td>
</tr>
<tr>
<td>MLlib, GraphX, Spark SQL, Spark</td>
</tr>
<tr>
<td>Streaming</td>
</tr>
<tr>
<td>BigDL</td>
</tr>
<tr>
<td>Anaconda Python</td>
</tr>
<tr>
<td>TensorFlow™</td>
</tr>
<tr>
<td>Apache Spark™</td>
</tr>
<tr>
<td>Dask</td>
</tr>
<tr>
<td>Cray Distributed Training Framework</td>
</tr>
<tr>
<td>(CrayPE ML Plugin)</td>
</tr>
<tr>
<td>Intel® MKL, Intel MKL-DNN, cuDNN,</td>
</tr>
<tr>
<td>NVDIA SDK, OpenMPI</td>
</tr>
</tbody>
</table>

**Urika-CS:**
- Pre-integrated software suite for AI workflow
- Distributed deep learning training for heterogeneous systems (CS500 & CS-Storm)
- Supported stack – for open source components and for distributed training framework

© 2018 Cray Inc.
FLEXIBLE & SCALABLE DATA ANALYSIS PLATFORMS

• Equipped with rich set of analysis libraries & easily extensible
• Enables higher-productivity languages (Scala, R, Python) and Jupyter interactive notebooks
• Leverage Python scalable data analysis infrastructure
  • Met Office IRIS library
  • DASK Parallel Python engine
  • Pangeo software stack
• Performance with SPARK, DASK, R, Cray Graph Engine
NEW ANALYSIS APPROACHES: WEATHER/CLIMATE INFORMATICS

• Applying more complex analysis treatments to find new insights or increase accuracy of predictions
  • Machine/Deep Learning
  • Regression/Clustering/Optimization methods

• Challenges:
  • Meteorological Data very different from classic “Big Data” sets
    • Multivariate, with Spatial & Temporal locality
    • Can be very sparse/limited in some dimensions
    • Limited/missing observations over unpopulated areas
  • Data-driven predictions must be linked physics/first-principles based theory
PANGEO DATA: TOWARD A BIG DATA ANALYSIS PLATFORM

Pangeo Goal: create an open-source toolkit for the analysis of climate datasets, built on the Python language ecosystem, Xarray multi-dimensional array tools, and Dask parallel analytics system.

Parallelism is key: single device performance is falling behind!

Xarray

Xarray is an open source project and Python package that provides a toolkit for working with labeled multi-dimensional arrays of data. Xarray adopts the Common Data Model for self-describing scientific data in widespread use in the Earth sciences. Xarray datasets are an in-memory representation of a netCDF file. Xarray provides the basic data structures used by many other Pangeo packages, as well as powerful tools for computation and visualization.

Iris

Iris seeks to provide a powerful, easy to use, and community-driven Python library for analysing and visualising meteorological and oceanographic data sets.

With iris you can:

- Use a single API to work on your data, irrespective of its original format.
- Read and write CF-netCDF, GRIB, and PP files.
- Easily produce graphics and maps via integration with matplotlib and cartopy.

Iris is an alternative to Xarray. Iris is developed primarily by the UK Met Office Informatics Lab.

Dask
MACHINE/DEEP LEARNING IN WEATHER/CLIMATE

• Arduous to train, but comparatively quick to run
• Data producer vs data consumer
• Complementary use cases
  • Rapid classifiers or predictors for radar/observations
  • Advanced MOS systems – DICast system
  • Pattern recognition in model outputs or observations
  • Emulation: Replacement of expensive parameterizations
    • train NN using expensive radiation model
    • train NN using cloud-resolving model, replace convection parameterization
## AI, MACHINE LEARNING & DEEP LEARNING

### ARTIFICIAL INTELLIGENCE

<table>
<thead>
<tr>
<th>Sense</th>
<th>Comprehend</th>
<th>Predict</th>
<th>Act and Adapt</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ANALYTICS</strong></td>
<td><strong>MACHINE LEARNING</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search datasets for insights</td>
<td>Learn patterns from the past to predict future</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Descriptive</td>
<td><strong>What happened?</strong></td>
<td><strong>Unsupervised</strong>&lt;br&gt;Group, cluster and organize content with domain-specific heuristic models</td>
<td><strong>Supervised</strong>&lt;br&gt;Train mathematical predictive models with labelled data</td>
</tr>
<tr>
<td>Diagnostic</td>
<td><strong>Why did it happen?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Predictive</td>
<td><strong>What will happen?</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prescriptive</td>
<td><strong>How to make it happen?</strong></td>
<td><strong>Vision</strong>&lt;br&gt;Vision</td>
<td><strong>Speech</strong>&lt;br&gt;Speech</td>
</tr>
</tbody>
</table>

© 2018 Cray Inc.
Challenge:

• Climate simulations run at 10000x faster than real-time and high resolution required to reproduce extreme weather events – generate 100’s TB of data

Supercomputer + AI Solution:

• Semi-supervised convolutional architectures can identify extreme weather events such as Tropical Cyclones, Atmospheric Rivers, Weather Fronts with 90% accuracy

• 15-PetaFLOP Deep Learning system used to scale training of a single model to ~9600 Xeon-Phi nodes; obtaining peak performance of 11.73-15.07 PFLOP/s

Can neural networks enhance the prediction of damaging hail? If so, why?

NN identifies physically relevant features: temperature, dewpoint, elevation, confluent warm moist air, mid-level rotation.

*D.J. Gagne, NCAR*
IMPROVING SATELLITE DATA UTILIZATION Through Deep Learning at NOAA

- Satellites provide more data than can be assimilated, ~3% of available data is used today
- Use DL object detection to identify areas of atmospheric instability from satellite observation data, focus extraction of observations on these regions of interest

Run on Theia – Cray CS-Storm system

100 nodes, each with 8 NVIDIA Tesla P100 GPUs

Jebb Stewart, 2018 ECMWF workshop on HPC in Meteorology
SOIL MOISTURE
Through Deep Learning at NOAA

Run on Theia – Cray CS-Storm system

100 nodes, each with 8 NVIDIA Tesla P100 GPUs

Jebb Stewart, 2018 ECMWF workshop on HPC in Meteorology
“As in many industries, we are challenged with increasing data volumes and are turning to large-scale analytics, machine learning and deep learning applications to drive new insights and innovation,” said Charles Ewen, director of technology and CIO at the Met Office. “The Met Office already has one of the world’s largest Cray XC supercomputing systems. Now with our implementation of Cray’s Urika-XC software, we are applying AI and analytics to deliver ever-more accurate and detailed weather forecasts and climate change analyses, while also developing new commercial products.”
MIXED SIMULATION/ANALYTICS WORKLOADS

Computational Modeling

Math Models
- Simulation and modelling of the natural world via mathematical equations

Data-Intensive Processing
- Hybrid workflows with a mix of simulation and analytics

Data Models
- Analysis of large datasets for knowledge discovery, insight, and prediction
Run Any Workflows

Converged Network

Containers & Virtualization

Extensibility