The Leadership Computing Alliance

Addressing the HPC Challenges of M-OSRP Algorithms

Michael Perrone, PhD
Data-Centric Systems, IBM Research
mpp@us.ibm.com

Executive Summary

- **Processing challenges** driving new, compute and data intensive algorithms
- **Technical challenges** put next generation algorithms at risk
- **IBM** is committed to HPC and to creating the next generation of imaging systems

- **Need a coordinated effort to design and deliver the next generation of upstream petroleum computing systems**

- **Next Steps: Set up a meeting to discuss the best path forward.**

GOAL: Assist in making M-OSRP algorithms **realizable**

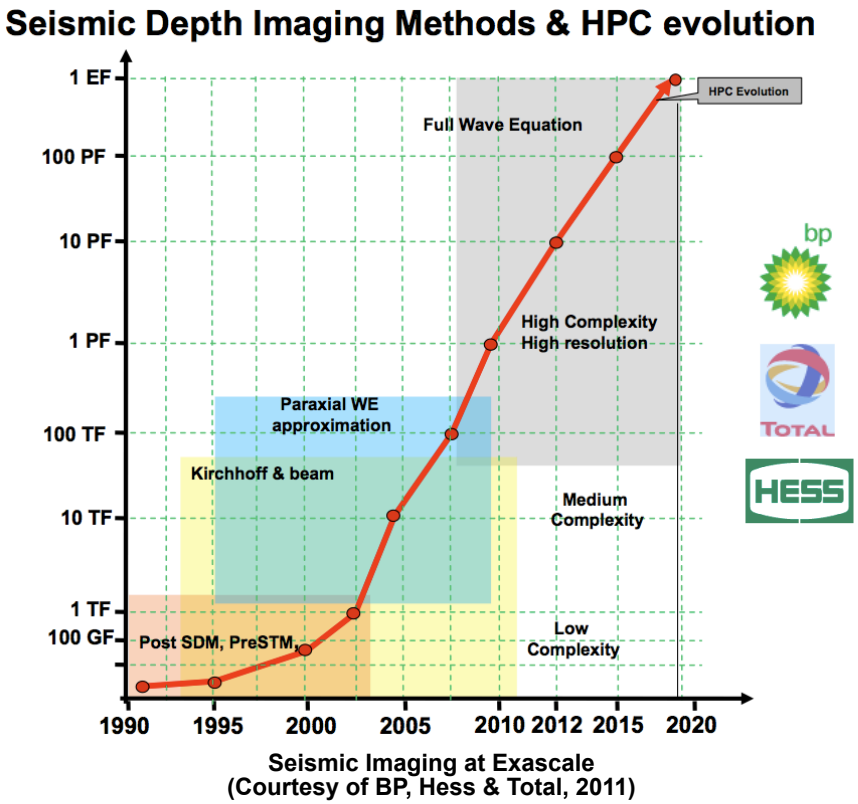
- **Long history of IBM participation in the M-OSRP**
  - Technical papers published
  - Analysis of M-OSRP algorithms
  - Code optimization
  - Participation in meetings

- **Continued IBM support of the M-OSRP mission**
  - Hardware
    - IBM HW in U of Houston data center: Cell blades
    - Remote access to IBM HW in IBM data centers: BlueGene, P8/P9, etc.
  - Software
    - ISS IMA on Blue Gene and Cell, including comparison of various architecture
    - ISS IME in process
  - Training
    - Host visiting scientists and lecturers
    - Host students for training on IBM HW

Need for New System Design

Algorithmic Complexity

- RTM Iso = 5x Kirchhoff
- RTM TTI = 3x RTM Iso
- FWI Iso = 10x RTM Iso
- FWI TTI = 10x RTM TTI
- Elastic = 6x RTM TTI



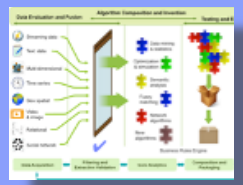
Historical View

- 1960s-80s: NMO/DMO (mainframe)
- 1970-80s: 2D post-stack processing (mainframe vector)
- 1980s: 3D post-stack (VAX, UNIX)
- 1990's: 3D pre-stack (IBM SP, Sun, SGI SMPs)
- Late 1990s: 3D PSTM (Linux Clusters)
- 2000's: 3D PSDM, WEM (x86)
- Today: 3D RTM, ISS IMA (x86+GPU)
- Future
  - Higher frequency RTM
  - Elastic RTM
  - ISS Internal Multiple Removal
  - Other ISS & Green's function methods

Data-Centric Systems: Application Domains

IBM

Business Analytics



System G
Big Insights

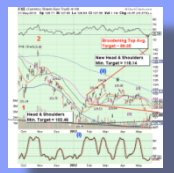
Business Intelligence

Watson

Social Analytics

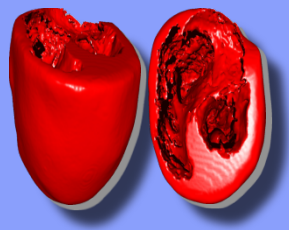
DOD All-Source Intelligence

Financial Analytics



Integrated Trading and VaR

Life Sciences



Health Care Analytics

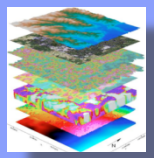
Complex analytics

HPA

Technical Computing

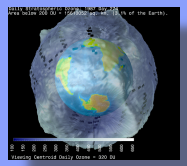
Engineering Design, Prototyping, Analysis, Optimization

Oil and Gas



Integrated Wide Azimuth Imaging and Interpretation

Climate & Environment



Production Weather

Science

DOE NNSA and Office of Science

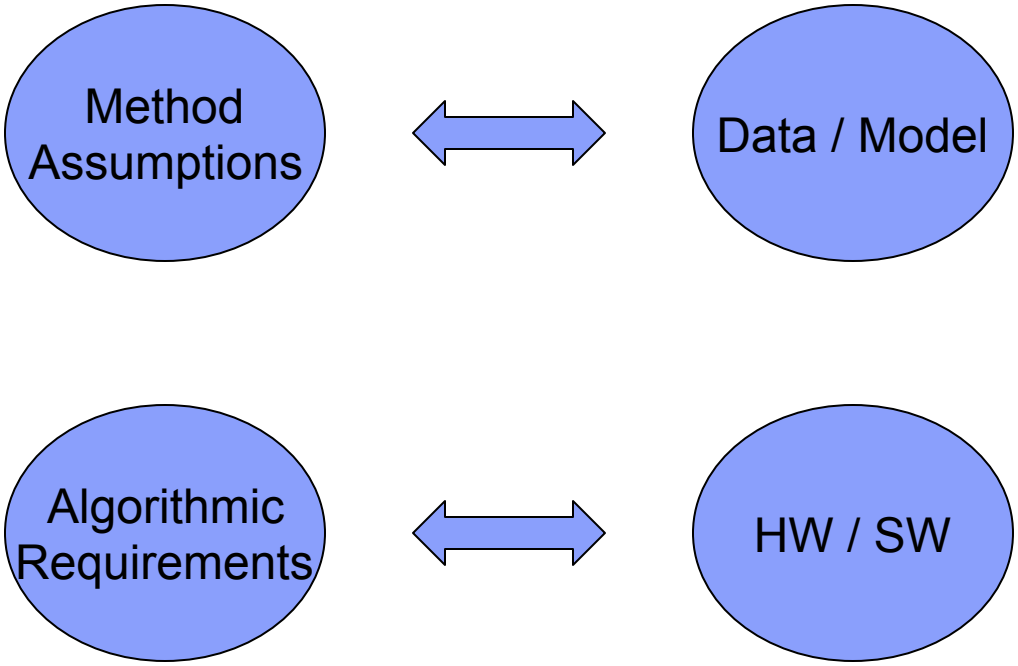
Modeling, Simulation

HPC

DCS

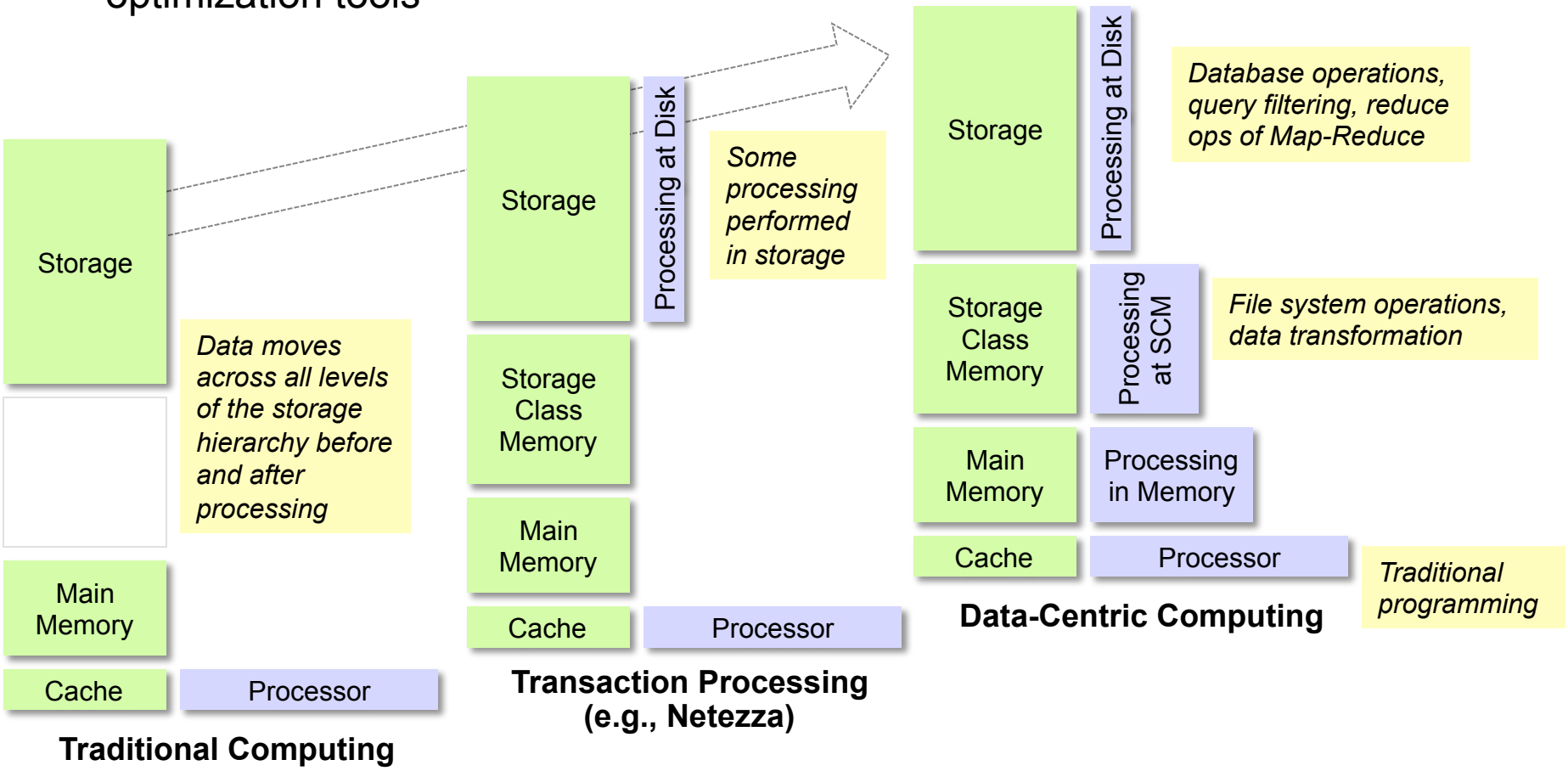
Key Domain Characteristics: Big Data, Complex Analytics, Scale and Time to Solution Requirements
Overlapping Requirements in HPC and HPA enable an converged solution

The Weglein Midrash

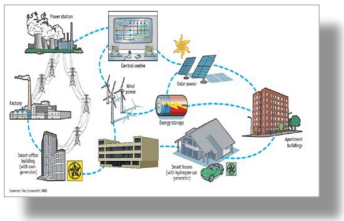


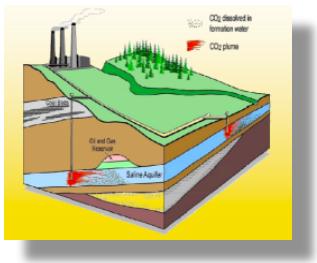
Data-Centric System Vision

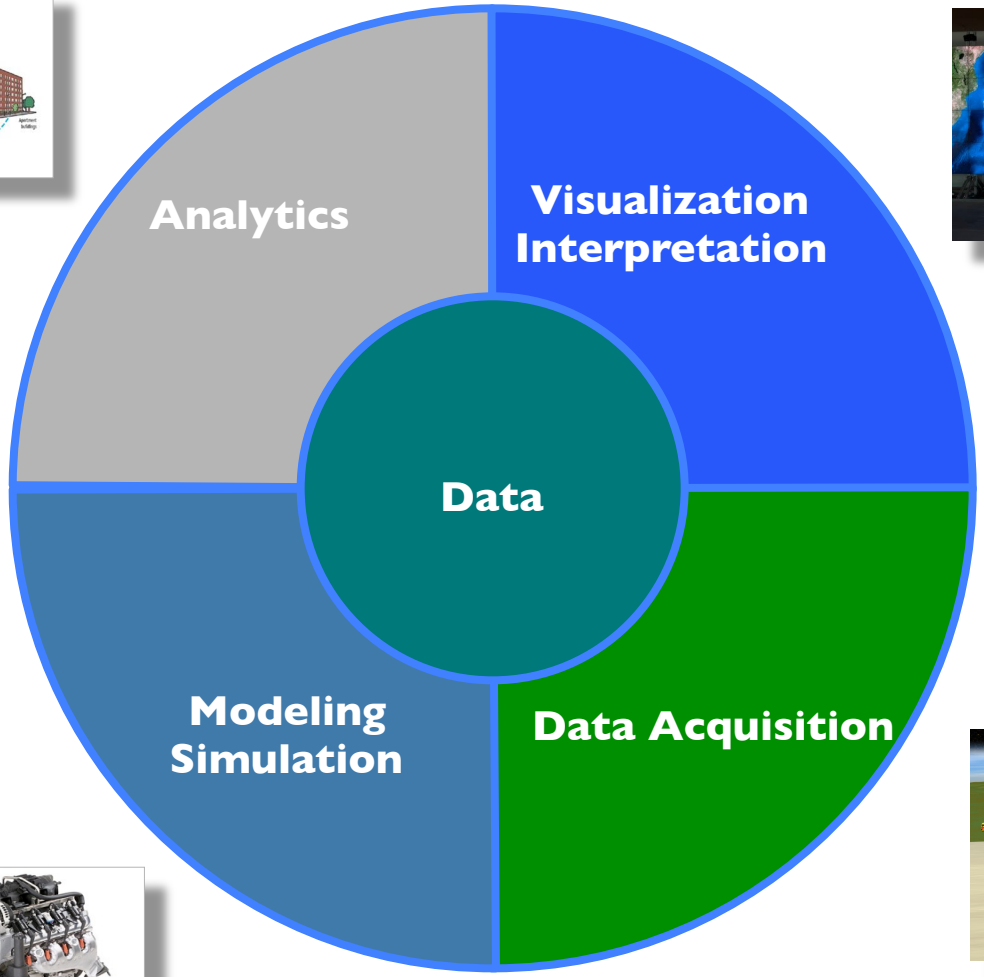
- Integration of massive data management and compute with complex analytics
- Optimized workflow components (compute and dataflow) across the system
- Data centric systems move computation to the data
  - Software provides middleware, programming models, APIs and workflow optimization tools

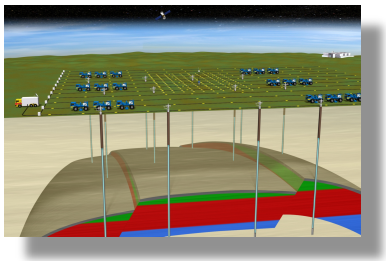


Solutions must be complete, robust, reliable, easy to use









Evolving to Exascale Class Data, Exascale Class Compute

OpenPOWER Foundation

MISSION: The OpenPOWER Foundations's mission is to **create an open ecosystem**, using the POWER Architecture to share expertise, investment and validated and compliant server-class IP **to serve the evolving needs of customers**.

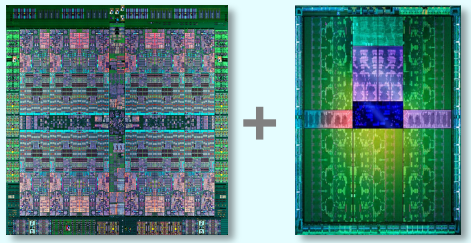
OpenMP

OpenACC
Directives for Accelerators

- Opening the architecture to give the industry the ability to innovate across the full Hardware and Software stack
  - Includes SOC design, Bus Specifications, Reference Designs, FW OS and Hypervisor Open Source
- Driving an expansion of enterprise class Hardware and Software stack for the data center
- Building a vibrant and mutually beneficial ecosystem for POWER

Example:

POWER CPU + Tesla GPU



Platinum Members

ALTERA, Google, IBM, Mellanox TECHNOLOGIES, Micron, NVIDIA, PowerCore, SAMSUNG, TYAN

Gold Members

TEAMSUN 华胜天成, VeriSilicon, ZTE中兴

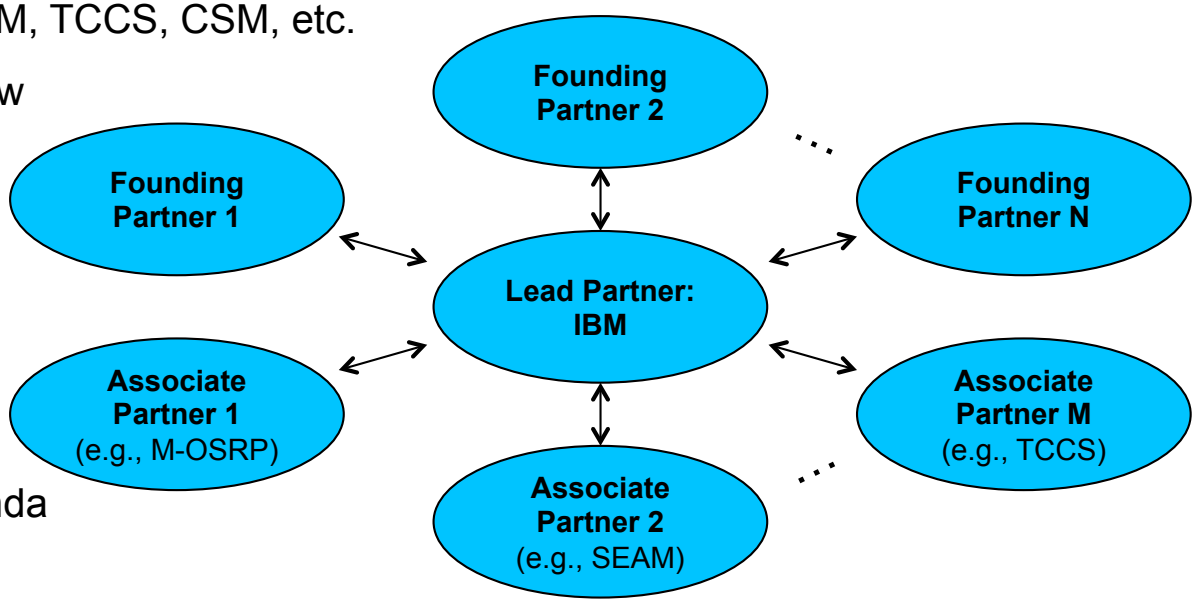
PROPOSAL: Leadership Computing Alliance (LCA)

An association of like-minded partners focused on defining and enabling the requirements for the next generation of seismic imaging data centers.

▪**Mission**: Through a multi-year, multi-phase, research co-design and collaboration process between IBM and other industry-leading LCA members, the LCA will (1) identify, quantify, test and evaluate the systems requirements of LCA member workflows and use these requirements to identify, explore and evaluate key system design points to guide IBM's DCS design for specific market segments, and (2) leverage IBM's broad ecosystem to address a new generation of challenges and opportunities driven by Big Data, real-time analytics, and Cloud computing.

▪**Industry Consortia:** M-OSRP, SEAM, TCCS, CSM, etc.

▪**Duration:** Open-ended, yearly review



▪**Partners**

- Lead Partner: IBM
- Founding Partners
  - Mainly from industry
  - Help define co-design agenda
- Associate Partners
  - Not able define agenda
  - University, professional society, industry/academic consortia, etc.
  - Partner with M-OSRP, SEAM, IPARS and others of value to Founding Partners

DCS Summary

Proposed meeting: EAGE'14 June 16-19, Amsterdam

- **DCS development motivated by market shift**
  1) Recent and continuing explosion of data
  2) Competitive advantage lies within the data
  3) All problems demand efficient systems to manage/analyze/compute/visualize data intense workloads

- **Complexity of addressing #'s 1-3 above at scale requires a new system design**
  - Bring compute and data together
  - Optimize bandwidth, throughput, latency, reliability and efficiency (/watt)

- **IBM identified this trend early on with our years of research on exascale systems**
  - This has evolved into a focused R&D effort led jointly by IBM Research & Systems Technology Group (STG) called Data-Centric Systems

- **IBM is seeking strategic industry, government, education, and OEM partners to jointly develop and optimize their application code to leverage this new generation of systems.**
  - Partner investment of people, direct IP funding and sharing of actual/proxy code
  - Partner commitment to multi-year relationship
  - Scale of investment will drive:
    - Scope of IBM technology included and/or partner applications addressed
    - Level of geographic or industry exclusivity
    - Level of unique IBM development: HW, SW, tools, compilers or applications

Thank you!

Questions?