Convergence entre Big Data & HPC : Enjeux et Défis

Mark Asch
Université de Picardie, Amiens

Journées Scientifiques Equip@Meso Janvier 30, 2017

The EXDCI project has received funding from the European Union's Horizon 2020 research and innovation programmed under grant agreement No 671558.

What is EXDCI?

- EXDCI is a 30-month project starting from September 2015 with
  - a budget of € 2.5 million
  - with 173,5 PMs
  - funded in call FET-HPC 2014.
- EXDCI's objective is to support the coordination of the development and implementation of a common strategy for the European HPC Ecosystem.
  - https://exdci.eu

exdci European Extreme Data & Computing Initiative

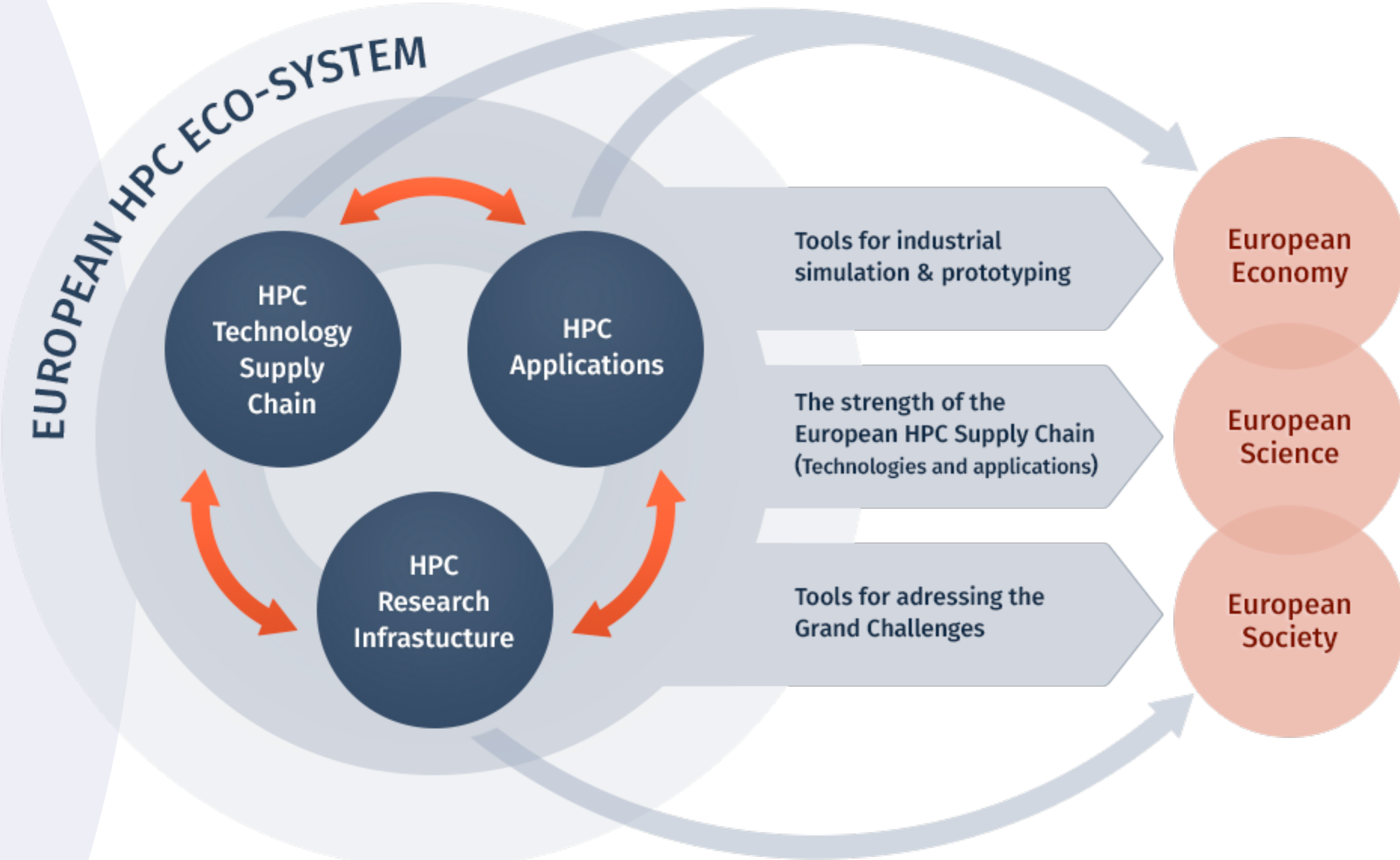
Strategic Goals of EXDCI

Development of a common European HPC Strategy - EXDCI will support the implementation of a common European HPC strategy through the coordination of activities of stakeholders such as the European Technology Platform for HPC (ETP4HPC), PRACE, application owners and users (including emerging HPC applications), the European exascale computing research community, the open source HPC community, and other related activities in other parts of H2020.

EXDCI aims to support the road-mapping, strategy-making and performance-monitoring activities of the ecosystem, i.e.:

- ✓ **Producing and aligning** roadmaps for HPC Technology and HPC Applications;
- ✓ **Measuring** the implementation of the European HPC strategy;
- ✓ **Building and maintaining** relations with other international HPC activities and regions;
- ✓ **Supporting** the generation of young talent as a crucial element of the development of European HPC.

European HPC ECO-SYSTEM



IESP: 2009-2012

International Journal of High Performance Computing Applications, **February 2011,** vol. 25 no. 1 **3-60**

The International Exascale Software Project Roadmap[1]

Jack Dongarra, Pete Beckman, Terry Moore, Patrick Aerts, Giovanni Aloisio, Jean-Claude Andre, David Barkai, Jean-Yves Berthou, Taisuke Boku, Bertrand Braunschweig, Franck Cappello, Barbara Chapman, Xuebin Chi, Alok Choudhary, Sudip Dosanjh, Thom Dunning, Sandro Fiore, Al Geist, Bill Gropp, Robert Harrison, Mark Hereld, Michael Heroux, Adolfy Hoisie, Koh Hotta, Yutaka Ishikawa, Zhong Jin, Fred Johnson, Sanjay Kale, Richard Kenway, David Keyes, Bill Kramer, Jesus Labarta, Alain Lichnewsky, Thomas Lippert, Bob Lucas, Barney Maccabe, Satoshi Matsuoka, Paul Messina, Peter Michielse, Bernd Mohr, Matthias Mueller, Wolfgang Nagel, Hiroshi Nakashima, Michael E. Papka, Dan Reed, Mitsuhisa Sato, Ed Seidel, John Shalf, David Skinner, Marc Snir, Thomas Sterling, Rick Stevens, Fred Streitz, Bob Sugar, Shinji Sumimoto, William Tang, John Taylor, Rajeev Thakur, Anne Trefethen, Mateo Valero, Aad van der Steen, Jeffrey Vetter, Peg Williams, Robert Wisniewski, and Kathy Yelick

Abstract

Over the last twenty years, the open source community has provided more and more software on which the world's High Performance Computing (HPC) systems depend for performance and productivity. The community has invested millions of dollars and years of effort to build key components. But although the investments in these separate software elements have been tremendously valuable, a great deal of productivity has also been lost because of the lack of planning, coordination, and key integration of technologies necessary to make them work together smoothly and efficiently, both within individual PetaScale systems and between different systems. It seems clear that this completely uncoordinated development model will not provide the software needed to support the unprecedented parallelism required for peta/exascale computation on millions of cores, or the flexibility required to exploit new hardware models and features, such as transactional memory, speculative execution, and GPUs. This report describes the work of the community to prepare for the challenges of exascale computing, ultimately combing their efforts in a coordinated International Exascale Software Project.

Keywords

High Performance Computing, Software Stack, Exascale computing

[1] The International Exascale Software Project was organized by and has received ongoing support from a variety of national agencies: In the United states, the Department of Energy Office of Advance Scientific Computing Research (DOE-ASCR) and the National Science Foundation Office of CyberInfrastructure (NSF-OCI); In France, the Commissariat à l'énergie atomique et aux énergies alternatives (CEA), Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS), Agence nationale de la recherche (ANR), INRIA and Teratec; In the United Kingdom, Engineering and Physical Sciences Research Council (EPSRC); In Japan, The University of Tsukuba, RIKEN, Kyoto University, Tokyo University and the Tokyo Institute of Technology. Corporations contributing to the staging of different IESP meetings have included Cray, EDF/EESI, IBM, Intel, Fujitsu Ltd., and NVIDIA.

What is BDEC? – exascale.org

- IESP (2009-2012)
  - IESP Roadmap
- BDEC (2013-…) = { EU, USA, JAPAN, CHINA }
  - EESI – European Exacscale Software Initiative
  - EESI2 – European Exacscale Software Initiative
  - EXDCI…

1.Create an international collaborative process focused on the co-design of software infrastructure for extreme scale science, addressing the challenges of both extreme scale computing and big data, and supporting a broad spectrum of major research domains, 2.Describe funding structures and strategies of public bodies with Exascale R&D goals Worldwide 3.Establishing and maintaining a global network of expertise and funding bodies in the area of Exascale computing

exdci European Extreme Data & Computing Initiative

EXDCI WP6

Task 6.3: EXDCI in Europe and Worldwide (aka BDEC)

Journées Scientifiques Equip@Meso Janvier 30, 2017

The EXDCI project has received funding from the European Union's Horizon 2020 research and innovation programmed under grant agreement No 671558.

Achievements to date

- BoF in SC'15.
- 2-day BDEC workshop in Frankfurt, June 2016.
  - 70 invited participants from 10 countries.
  - 6 keynotes, 4 national roadmaps (EU, USA, JP, CN), 21 presentations.
  - Dynamic breakout sessions.
- ½-day BDEC workshop in ISC'16, June 2016.
  - Reporting on BDEC workshop.
  - Perspectives and roadmaps.
  - Discussion on new roadmap.
- Intensive drafting of a new international roadmap.
- BoF in SC'16

BDEC @ Europe

European Extreme Data & Computing Initiative

Conclusions to date

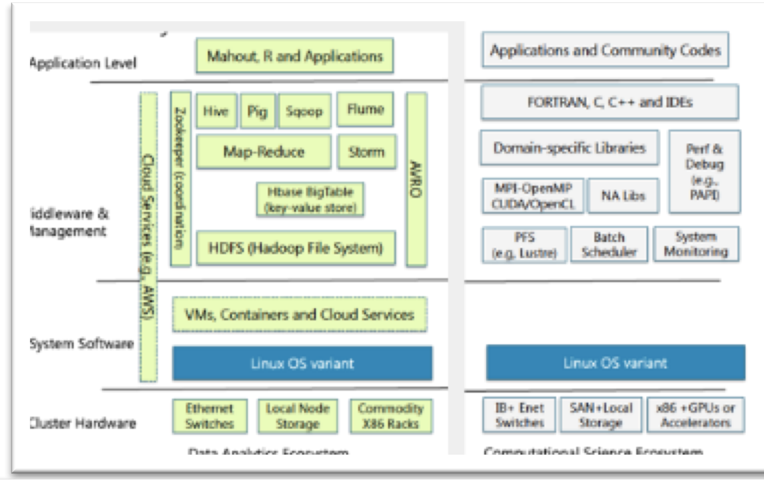
- BDEC is alive and well: Europe, through EXDCI, is an central player in international road-mapping.
- EXDCI has been an important engine.
- Europe is the only continent that has this type of CSA.
- Sharing of international perspectives is vital for future research and funding strategies in e-infrastructures.
- Convergence of Big science and Big data is going to play an increasingly important role on the global scene.

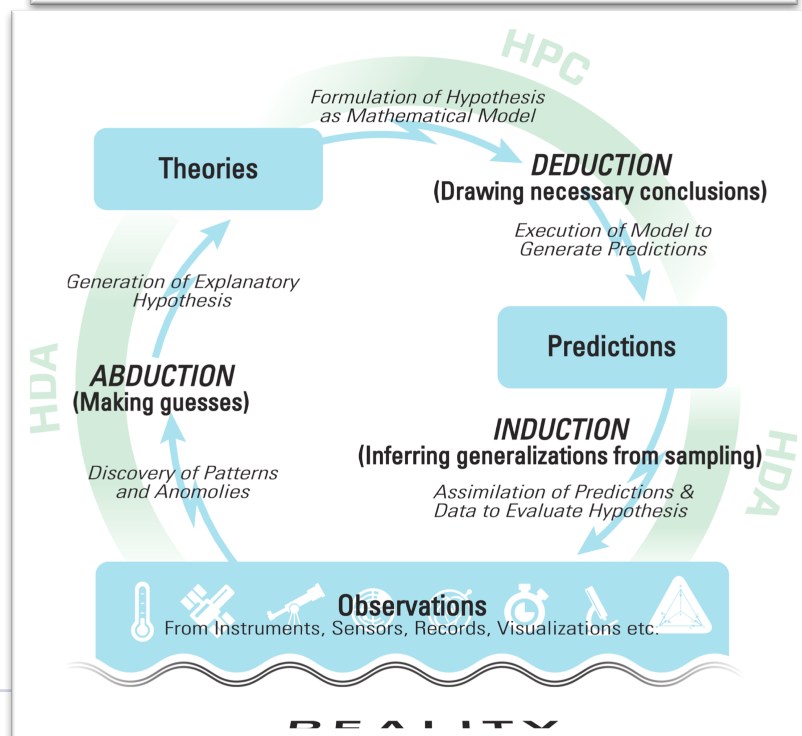
Going forward

- BoF in SC'16 where an initial draft of the new report was presented.
- Coordination of contributions to the final draft, entitled "Pathways to Convergence" – see next slides.
- Publication of the Pathways document in the most highly visible journals.
- Publication of white papers targeted at funding organisations and decision makers.
- Close collaboration with the EOSC initiative.
- Close collaboration with the BDVA cPPP.
- Preparation of 2017-18 BDEC meetings (USA, Japan).

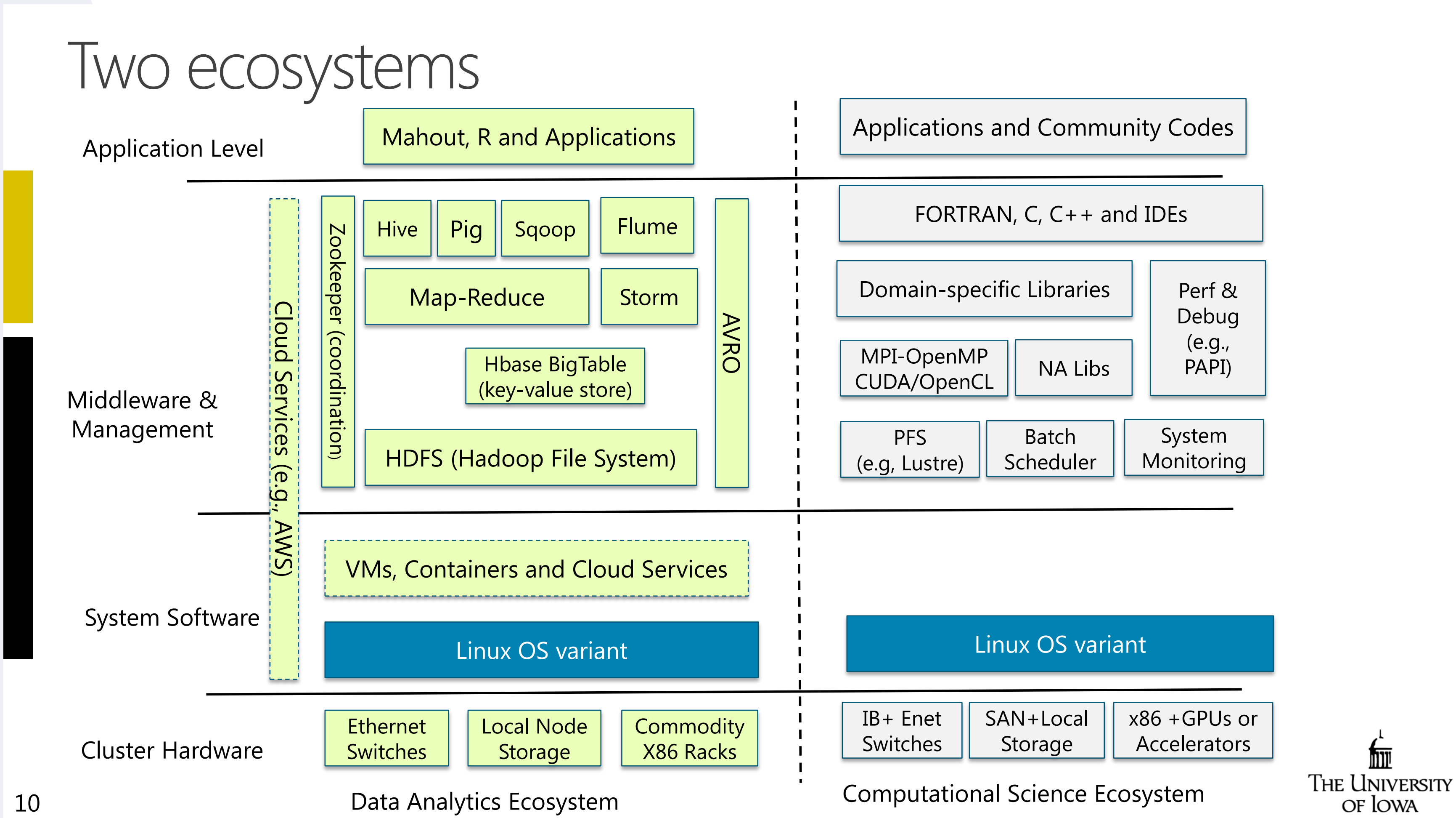
Pathways to Convergence

- What is the problem?
  - 2 separate software stacks – 1 for HPC, 1 for Big Data.
  - Impossibility to continue investing in both.
  - Outdated paradigm for scientific process.
- What is the solution?
  - Convergence of the software stacks.
  - Coordinated investment strategies.
  - Updated scientific process paradigm…



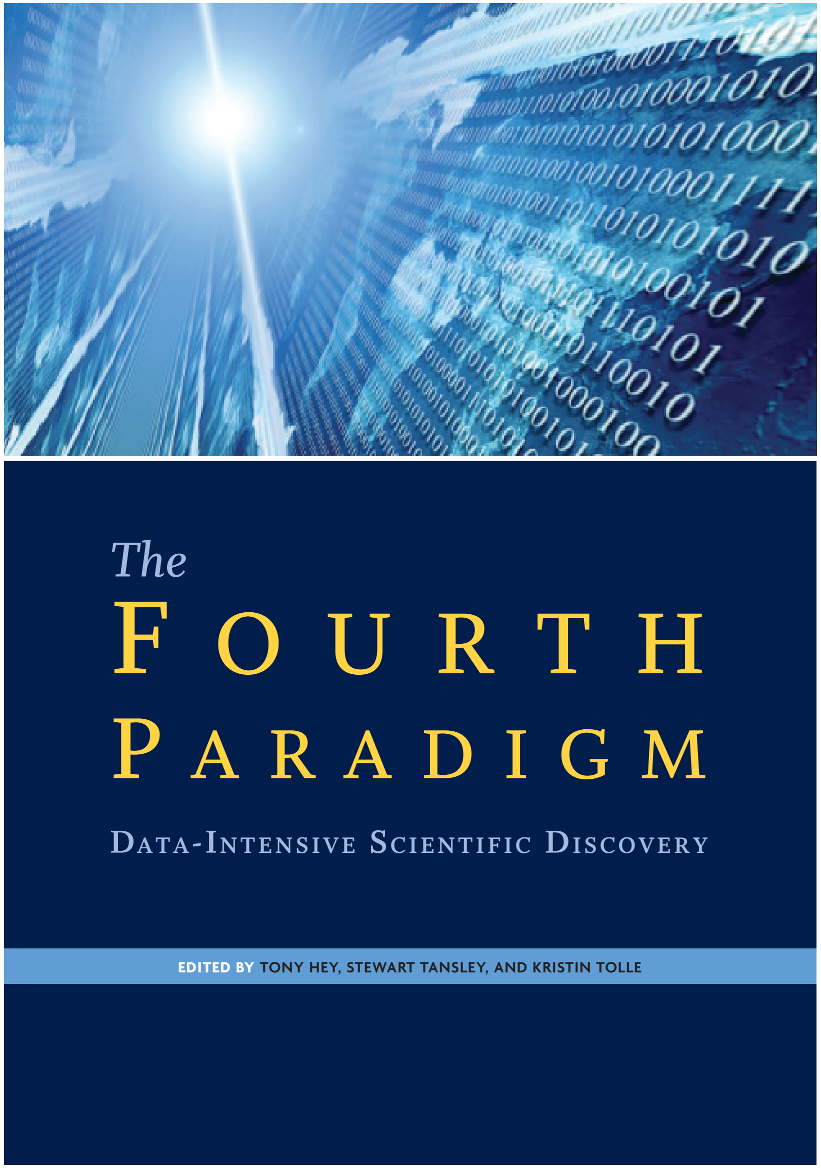


1. Bifurcation ahead:



Credit: J. Dongarra, D. Reed

2. Paradigm 2009:



4th silo...

Experimental

Thousand years ago

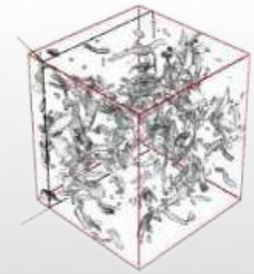
Description of natural phenomena

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G\rho}{3} - K\frac{c^2}{a^2}$$

Theoretical

Last few hundred years

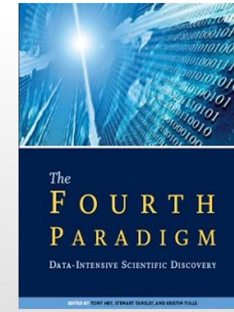
Newton's laws, Maxwell's equations…



Computational

Last few decades

Simulation of complex phenomena



The Fourth Paradigm

Today and the Future

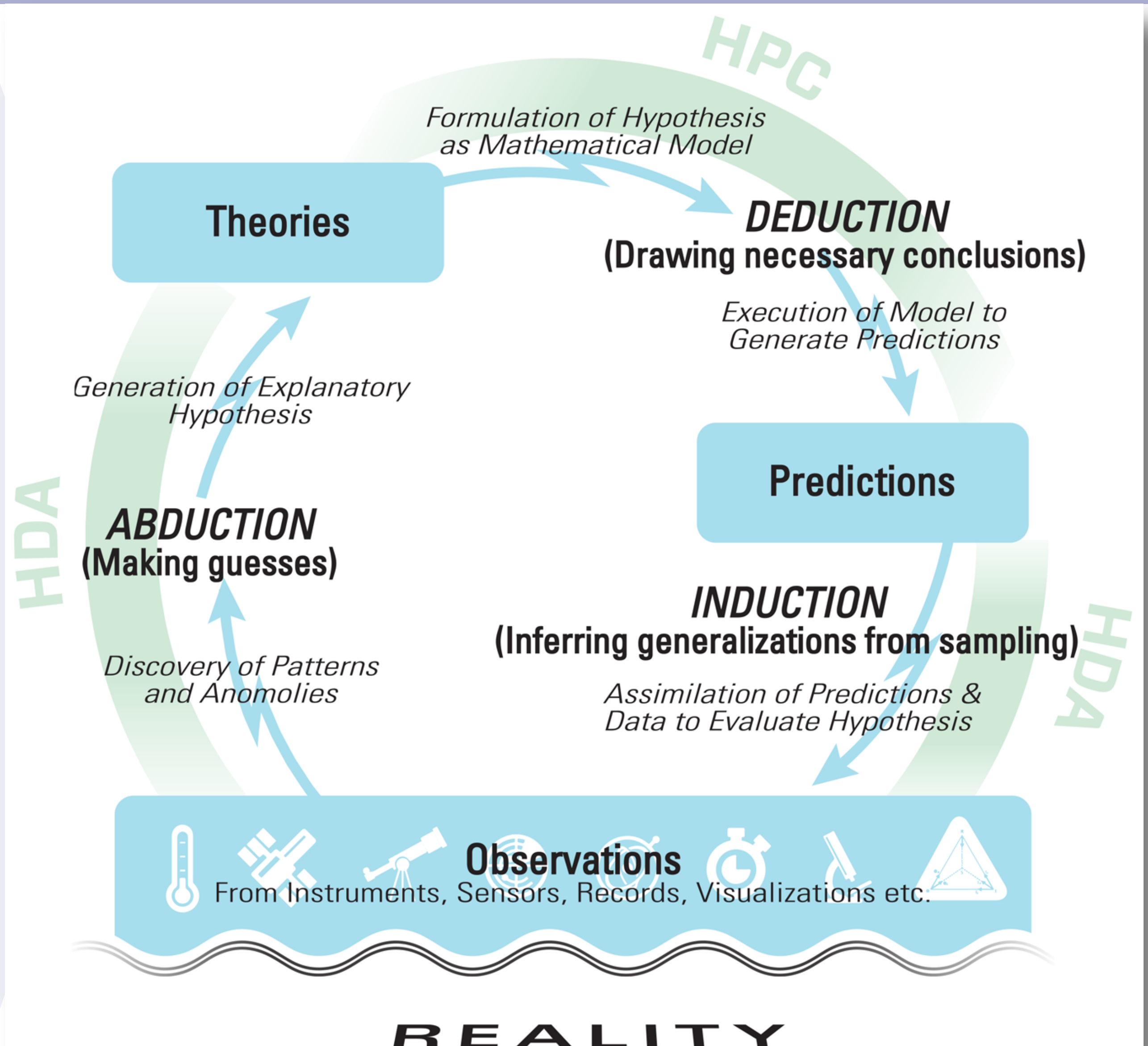
Unify theory, experiment and simulation with large multidisciplinary Data

Using data exploration and data mining (from instruments, sensors, humans…)

Distributed Communities

Crédits: Dennis Gannon

Paradigm 2017:



exdci European Extreme Data & Computing Initiative

Pathways to Convergence: outline of the report*

- 1. Introduction
  - 1.1 Convergence: Challenges and Opportunities
- 2. Common Context for Planning Cyberinfrastructure Convergence
  - 2.1 A Shared Model of Scientific Inquiry for Cyberinfrastructure Planning
  - 2.2 Confronting the General Problem of Data Logistics
  - 2.3 Software Infrastructure for Science as a Social Investment
  - 2.4 The overarching challenge of sustainability
- 3. Architectures and Strategies
  - 3.1 One traditional meaning of convergence: Building a new hourglass
  - 3.2 Containerization: The new "narrow waist"?
  - 3.3 Containers with adequate HPC and security capabilities . .
  - 3.4 Alternative Strategies for Data Logistics

* Available at exascale.org

Pathways to Convergence: outline of the report

- 3.4 Alternative Strategies for Data
  - 3.4.1 Alternative 1: Streaming
  - 3.4.2 Alternative 2: Content Delivery Networks (Processing In-transit)
  - 3.4.3 Alternative 3: Computing at the Edge (at the Source)
  - 3.4.4 Alternative 4: Logically Centralized Data (in the Cloud)
  - 3.4.5 Research Computing moves to Big Data stack
  - 3.4.6 Continue in the Bifurcated world, but do containerization to achieve some commonality
- 3.5 Trends in the Hardware Platform Substrate

- 4 Scientific Software Stack: layers and components
  - 4.1 Software Libraries for Common Intermediate Processing Tasks
  - 4.2 Math Libraries, Software Ecosystems for application development

Pathways to Convergence: outline of the report

- 4.2 Math Libraries, Software Ecosystems for application development
  - 4.2.1 Leveraging HPC Math Libraries in HDA
  - 4.2.2 New Efforts for Dense Linear Algebra Standards
  - 4.2.3 Challenges in the HPC Software Ecosystem
- 4.3 Interoperability between programming models and data formats
- 5 Applications
  - 5.1 Taxonomy of Application/workflow patterns and templates
  - 5.2 Science at the boundary of observation and simulation
  - 5.3 Numerical laboratories
  - 5.4 Data as a source of unexpected theories
  - 5.5 Common Reference Object Multi-modal Sensing(CROMS)

Pathways to Convergence: outline of the report

- 6 Operations and System Management
  - 6.1 Radically improved resource management
  - 6.2 QoS and Performance
- 7 Conclusions and Recommendations

Pathways to Convergence: Data Logistics & Strategies (1)

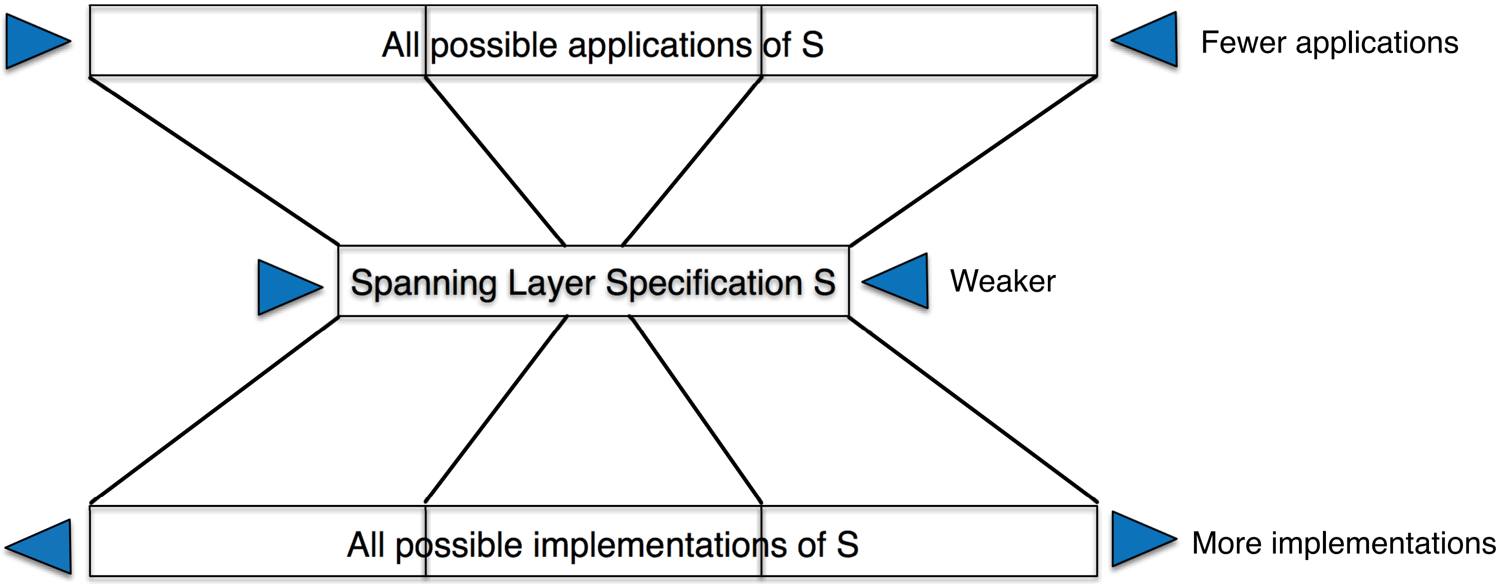
- Data logistics =
  - the management of the time-sensitive positioning and encoding/layout of data relative to its intended users and the resources they can use;
  - satisfy all the international communities that want to analyze data that is generated at one (or a few) locations, but is worked on somewhere else (eg. SKA, LHC, ESGF, etc.)
  - a continuum, with I/0 issues inside the IDC or supercomputing facility falling at one end, and BD workflows that begin at locations (possibly distributed) that are remote from the HPC system(s) targeted for the workload, falling at the other.
- A significant challenge to the classical HPC model!
  - data exists or is generated locally, it is loaded when a job is getting ready to run, and then it is written back after job completion.

Pathways to Convergence: Data Logistics & Strategies (2)

- 3 hypotheses:
  - (H1) data of interest may not be (and often are not) in the right form and/or condition
  - (H2) data of interest are physically distributed
  - (H3) data of interest may come from multiple sources
- 4 pathways:
  - (P1) data streaming,
  - (P2) in-transit processing,
  - (P3) processing at the edge of the distributed system, i.e. as close as possible to the data sources, and
  - (P4) logically centered cloud-like processing

Pathways to Convergence: Hourglass Model (1)

- History: let's look back to the pre-divergence era…
  - IP
  - Unix-like OS
- Hourglass model:



Pathways to Convergence: Hourglass Model (2)

- Spanning Layer is critical…
  - can be implemented on a wide variety of technology platforms;
  - can support an equally wide variety of applications.
- The challenge: provide a service specification (i.e. an API) for the spanning layer that can satisfy both these requirements!
- Optimization… covering communication, computation, and storage.

Pathways to Convergence: Containers (1)

- Question: Can containerization form the new "narrow waist"???
  - Building-from-source is well known in HPC, but unknown in HDA communities – therefore not an option!
  - Containers are more efficient than VM's (they avoid runtime overheads) and can attain native-installation execution times.
  - Packaging reusable HPC software components in containers dramatically reduces the usage barrier for the HDA community.
- But, adequate HPC and security capabilities still have to be developed.

Pathways to Convergence: Containers (2)

- Observations:
  - Low-level (starting from HW) convergence is a reachable goal.
  - How high can we make progress with the definition of common layers (e.g. resource management, scheduling tools, storage abstractions, etc.) that would benefit a world where Big Data analytics and Scientific computing get closer and closer?
  - All this goes beyond the containerization topic, of course…

Pathways to Convergence: Scientific Software Stack (1)

- Software Libraries for Common Intermediate Processing Tasks
  - cleaning, sub-setting, filtering, mapping, object segmentation, feature extraction, registration, etc. common to many domains;
  - can a common set of software tools, appropriately layered and modularized, be developed to serve the diverse purposes of a number of different communities at the same time?
  - methods that integrate or synthesize data of multiple types and/or from multiple sources or sensor modalities – eg. Google car

Pathways to Convergence: Scientific Software Stack (2)

- Math Libraries, Software Ecosystems for application development
  - Leveraging HPC Math Libraries in HDA.
  - Need containers to encapsulate complexity of input parameters and usage.
  - Efforts for dense linear algebra standards for ML and DL.
  - Challenges in the HPC Software Ecosystem – OpenMP is 30 years old…
  - Interoperability between HDA/HPC programming models and data formats:
    - OpenMP/MPI vs. Hadoop/MapReduce/Spark.
    - Programming languages differ.
    - Data interoperability between these different approaches needs to be addressed.

CONCLUSIONS

- The issue of convergence is a major challenge for us (HPC community) – the others "don't give a damn"…
- We need to ensure our future infrastructure and research funding – this is going to pass only through convergence.
- We welcome your contributions:
  - download the draft from www.exascale.org
  - mail me your comments/ideas/remarks/criticisms mark.asch@u-picardie.fr
- We plan to:
  - finalize the report in Wuxi (March 2017)
  - organize a major international meeting on Convergence (Q4-2017)
  - prepare a proposal for EXDCI-2… (Q3-2017)

European Extreme Data & Computing Initiative

Thank you!

mark.asch@u-picardie.fr

https://exdci.eu

The EXDCI project has received funding from the European Union's Horizon 2020 research and innovation programmed under grant agreement No 671558.