BDEC2: the digital continuum

Mark Asch

33e Journée CaSciModoT

December 10th, 2020

What is BDEC?

- An international "think-tank" started by HPC people (circa 2009) to address exascale convergence of compute and (big) data.
- We met regularly (2-3 times per year) in intensive 3-day workshops.
- We prepare position papers and roadmaps that are submitted to major national funding agencies and that guide investment strategies.
- Major reports published:
  - IESP Roadmap, IJHPCA, 25(1), 2011.
  - Pathways to Convergence, IJHPCA, 32(4), 2018.
- Follow us: www.exascale.org/bdec

10^{18}

BDEC2

BIG DATA AND EXTREME-SCALE COMPUTING 2

THE INTERNATIONAL JOURNAL of HIGH PERFORMANCE COMPUTING APPLICATIONS

The International Exascale Software Project roadmap

The International Journal of High Performance Computing Applications
25(1) 3–60

Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/1094342010391989
hpc.sagepub.com
SAGE

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, Zhong Jin, Yutaka Ishikawa, 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 S. Mueller, Wolfgang E. 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 20 years, the open-source community has provided more and more software on which the world's high-performance computing systems depend for performance and productivity. The community has invested millions of dollars and years of effort to build key components. However, 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 smoot
within individual petascale systems and between different systems. It seems clear that this con
development model will not provide the software needed to support the unprecedented paralle
exascale computation on millions of cores, or the flexibility required to exploit new hardware mc
as transactional memory, speculative execution, and graphics processing units. This report des
community to prepare for the challenges of exascale computing, ultimately combing their efforts
national Exascale Software Project.

exdci
European Extreme Data & Computing Initiative

Journal Title
XX(X):1–40
©The Author(s) 2018
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI: 10.1177/ToBeAssigned
www.sagepub.com/
SAGE

BIG DATA AND EXTREME-SCALE COMPUTING: PATHWAYS TO CONVERGENCE. Toward a Shaping Strategy for a Future Software and Data Ecosystem for Scientific Inquiry*

M. Asch, T. Moore, M. Asch, R. Badia, M. Beck, P. Beckman, T. Bidot, F. Bodin, F. Cappello, A. Choudhary, B. de Supinski, E. Deelman, J. Dongarra, A. Dubey, G. Fox, H. Fu, S. Girona, W. Gropp, M. Heroux, Y. Ishikawa, K. Keahey, D. Keyes, W. Kramer, J.-F. Lavignon, Y. Lu, S. Matsuoka, B. Mohr, D. Reed, S. Requena, J. Saltz, T. Schulthess, R. Stevens, M. Swany, A. Szalay, W. Tang, G. Varoquaux, J.-P. Vilotte, R. Wisniewski, Z. Xu and I. Zacharov

Asch et al. Pathways to Convergence. Int J. HPC Appl. 32(4), 2018

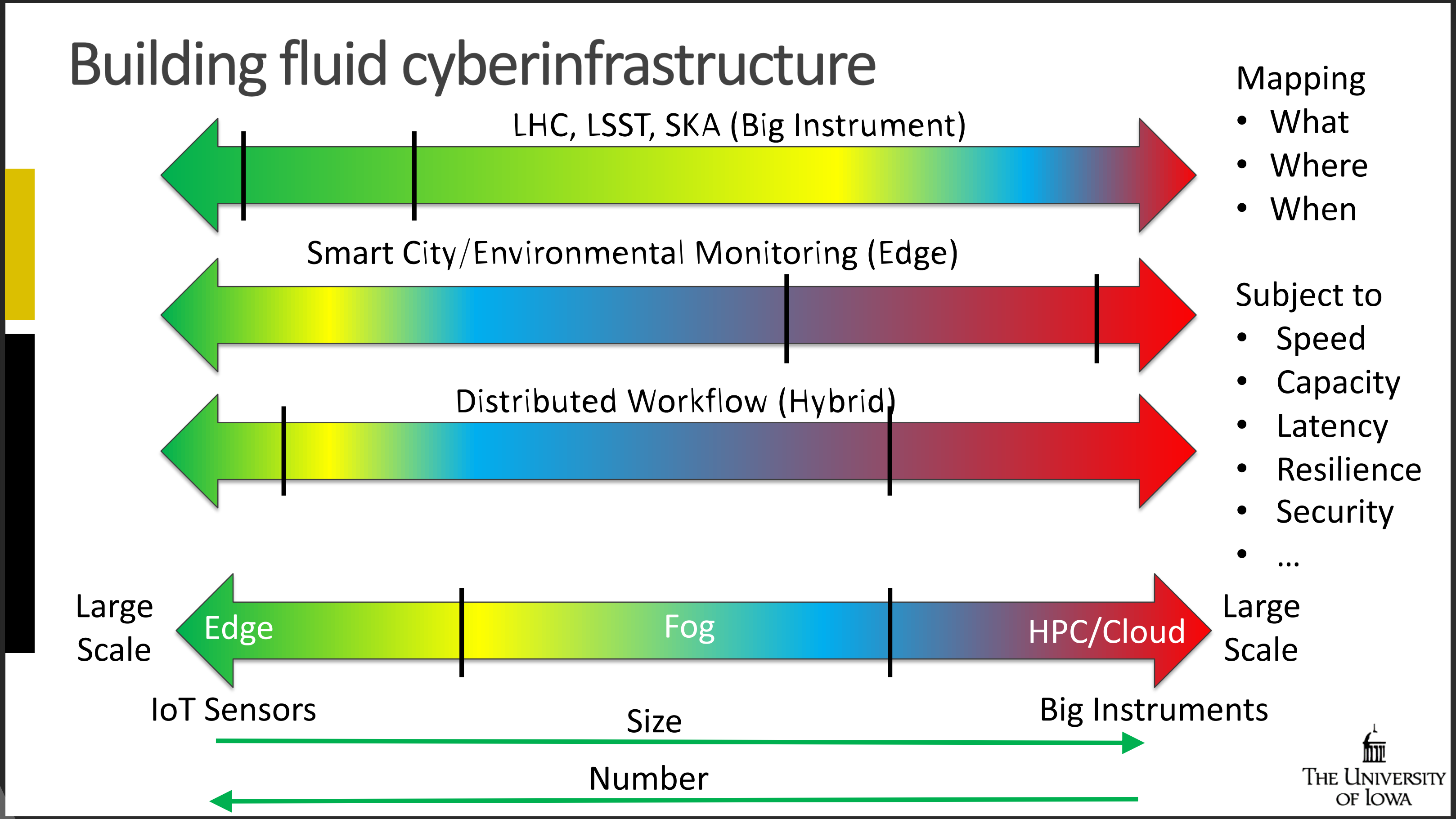
In the beginning...

- Peta- to Exascale transition
- Big Data and IoT (the end of HPC?)
- Convergence of HPC and Big Data, but what about logistics?
- TransContinuum e-infrastructures:
  - Edge-to-edge, from IoT and Big Instruments through to the Centre (cloud, HPC)
  - Data everywhere
  - Compute near the data
  - Workflows
  - AI everywhere... the new (unavoidable) enabler!

BDEC2
BIG DATA AND EXTREME-SCALE COMPUTING 2

exdci
European Extreme Data & Computing Initiative

So, what is the Digital Continuum?



Courtesy: Dan Reed

Challenge: programming the Digital Continuum?

IoT/Edge **HPC/Cloud**

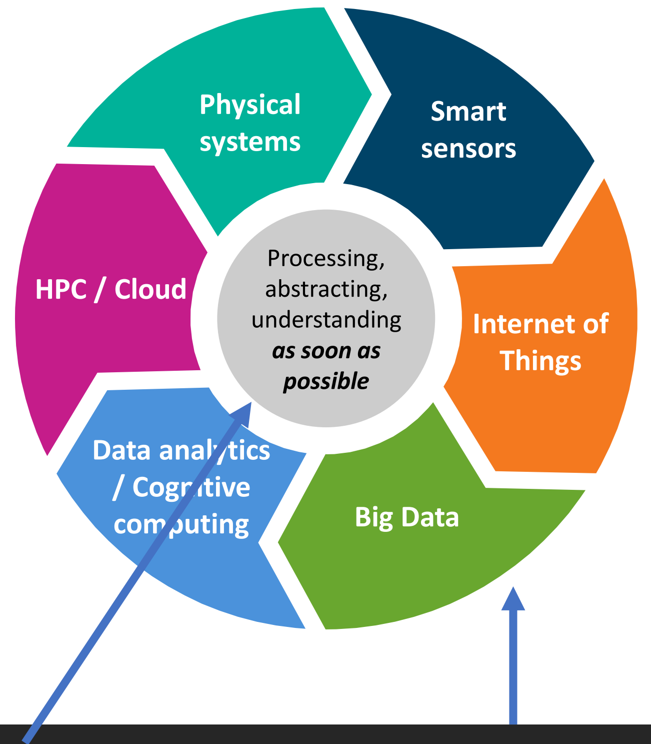
Size	Nano	Micro	Milli	Server	Fog	Campus	Facility
Example	IoT	Smart Device	Sage Node	Linux Box	Co-located Blades	1000-node cluster	Datacenter
Memory	0.5K	256K	8GB	32GB	256G	32TB	16PB
Network	BLE	WiFi/LTE	WiFi/LTE	1 GigE	10GigE	40GigE	N*100GigE
Cost	$5	$30	$600	$3K	$50K	$2M	$1000M

Count = 10^9
Size = 10^1

count = 10^1
Size = 10^9

So, what is the Digital Continuum?

HPC in the loop



Enabling Intelligent data processing at the edge:

- Fog computing
- Edge computing
- Stream analytics

Transforming data into information as soon as possible

Collaboration between edge devices and the HPC/cloud ensuring:

- Data security and Privacy
- Lower bandwidth
- Better use of HPC/Cloud

→ creating a continuous flow

Courtesy: ETP4HPC

BDEC2 Demonstrators

- Definition: a *proof-of-concept* platform designed to demonstrate some **common** capabilities that some of our BDEC2 applications and application communities need.
- Objective: produce a working version of an international, federated, continuum-spanning demonstrator that can be cooperatively operated and managed and that engages stakeholders at all levels.
- Why? Existing cyberinfrastructure was not designed to adequately deal with edge to cloud/HPC workflows, especially not extremely data intensive ones.
- How? A series of international *WORKshops*, bringing together computer scientists, application scientists, big data, IoT, AI and other stakeholders who are focused on achieving this goal.
- Follow us: www.exascale.org/bdec

What are the Challenges?

- There is an end-to-end problem (spanning the continuum) - from AI@Edge to HPC in the Cloud.
- There is a software stack problem (HPC troglodytes).
- There is a resource allocation problem (on demand, shared infrastructure).
- There is a data movement and logistics problem (both directions).
- Robustness, security, sustainability and reliability of large, interlinked, composed infrastructures.
- AI is everywhere, and new infrastructures must support monitoring and control; infrastructure learns (not just the app).

What is a good Demonstrator?

- Could evolve to support multiple application domains.
- Reveals programming model from edge to cloud.
- Shows global workflow (data, resources, users, etc.)
- Architecture is reusable, across multiple scales.
- Could evolve to run across several different composed infrastructures.

10^18 BDEC2 BIG DATA AND EXTREME-SCALE COMPUTING 2

exdci European Extreme Data & Computing Initiative

What are the potential use-cases?

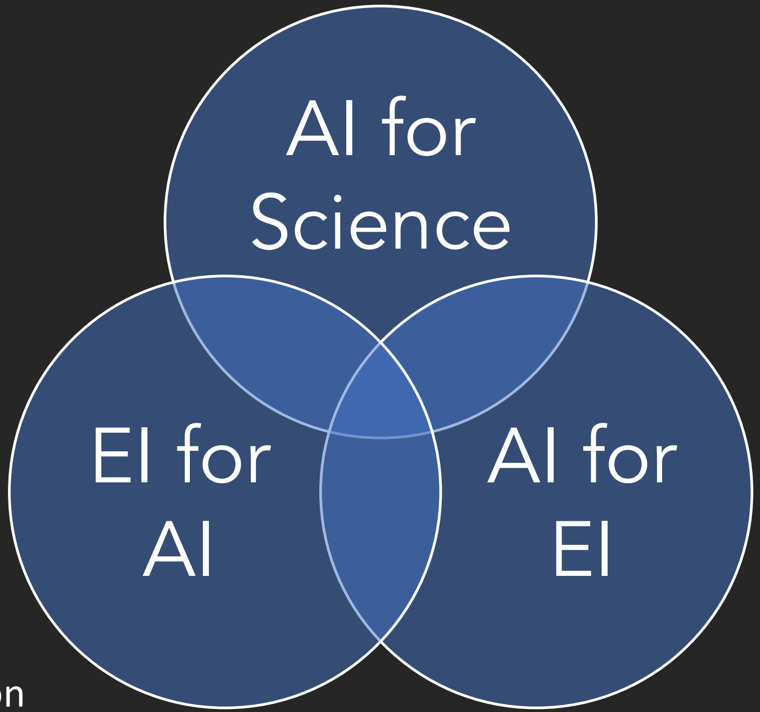
- Big Instruments:
  - Radio telescopes – LOFAR, SKA.
  - High energy physics – LHC .
  - Satellite data – Copernicus, SWOT, HIMAWARI, …
  - Climate, earth sciences, oceanography.
- IoT-like:
  - Personalized medicine.
  - Autonomous vehicles.
  - Predictive maintenance.
  - Precision agriculture.
- Digital Twins...

What are the next steps?

- Two demonstrators to be developed:
  - Global Data Logistics Network.
  - Learning everywhere.
- Multi-lateral funding:
  - Set up an international funding scheme for the above 2 demonstrators.
  - Solicit funding agencies.
  - Encourage private sector involvement (GAFAM, ABC, etc.).
- Strategic Research Agenda for EU (coordinated calls):
  - ETP4HPC
  - BDVA
  - IOTI
  - 5G, ECSO, Robotics

AI in the Continuum

- 3 categories:
  - AI for Science – applications
  - AI for EI – piloting e-infrastructure
  - EI for AI – making AI efficient
- AI for Science:
  - Steering of simulations
  - Embedding ML in simulation methods
  - Customized computational kernels
  - Tuning applications parameters
  - Generative models to compare with simulation
  - Student (AI) Teacher (Sim) models -> learned functions
  - Guided search through parameter spaces
  - Hybrid architectures HPC + Neuromorphic



BIG DATA AND EXTREME-SCALE COMPUTING 2

European Extreme Data & Computing Initiative

BDEC2
BIG DATA AND EXTREME-SCALE COMPUTING 2

In Ten Years…

- **Learned Models Begin to Replace Data**
  – queryable, portable, pluggable, chainable, secure
- **Experimental Discovery Processes Dramatically Refactored**
  – models replace experiments, experiments improve models
- **Many Questions Pursued Semi-Autonomously at Scale**
  – searching for materials, molecules and pathways, new physics
- **Simulation and AI Approaches Merge**
  – deep integration of ML, numerical simulation and UQ
- **Theory Becomes Data for Next Generation AI**
  – AI begins to contribute to advancing theory
- **AI Becomes Common Part of Scientific Laboratory Activities**
  – Infuses scientific, engineering and operations

ARTIFICIAL INTELLIGENCE

DOE readies multibillion-dollar AI push

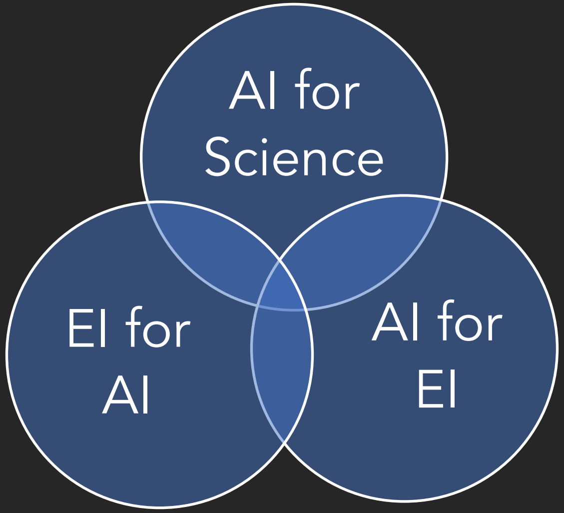
U.S. supercomputing leader is the latest big backer in a globally crowded field

By **Robert F. Service**, *in Washington, D.C.*

The U.S. Department of Energy (DOE) is planning a major initiative to use artificial intelligence (AI) to speed up scientific discoveries. At a meeting here last week, DOE officials said they will likely ask Congress for between $3 billion and $4 billion (10 years, roughly the amount the age is spending to build next-generation “ scale” supercomputers.

exdci
European Extreme Data & Computing Initiative

AI in the Continuum



- AI for EI:
  - Manage AI expectations (users in the loop)
  - Improve system operation: cost, reliability, security
  - Improve app and workflow performance
  - Closed-loop systems
  - Test-beds
- EI for AI
  - Programming: tools to be used across the continuum, performance in the continuum context
  - Distributed service composition: data placement, accomplish complex AI workflows in distributed, unreliable environment
  - Data: retain provenance, enforce access obligations, purpose-driven storage
  - Communications and Protocols: enable communication in this “more Internet than the Internet” environment
  - Authentication and Authorization : establish a chain of trust

BDEC Final Community Report

- Consensus
  - Data volumes are increasing exponentially
  - Data movement is the bottleneck
  - AI will be transformational
  - The Internet is ossified
  - There is no common platform for data
- Recommendations: (made to EC)
  - Design a data ecosystem based on international standards, for the security and transfer of data from devices to computing machines.
  - Conceive and develop transcontinuum workflow-enabling software stacks.
  - Investigate how machine learning can play a role at all stages of the digital continuum, all the way from the edge to the centre.
  - Adapt HPC centres to this new, data-centric science and the machine learning algorithms on which it relies.
  - Build, together, a comprehensive shared, sustainable e-science e-infrastructure to address the major societal challenges in the UN's Sustainable Development Goals in their Agenda 2030. Not forgetting fundamental science challenges.

BDEC2 BIG DATA AND EXTREME-SCALE COMPUTING 2

exdci European Extreme Data & Computing Initiative

Thank YOU

- Contact:
  - mark.asch@u-picardie.fr
- References:
  - www.exascale.org/bdec
  - Asch et al. Pathways to Convergence. Int J. HPC Appl. 32(4), 2018.