



BDEC: Past, Present and Future

Mark Asch

Technical Meeting EXDCI-2

December 2nd-3rd, 2019



What is BDEC?

- An international “think-tank” started by **HPC people** (circa 2009) to address **exascale** convergence of compute and (big) data.
- We meet 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, IHPCA, 25(1), 2011.
 - Pathways to Convergence, IHPCA, 32(4), 2018.
- Follow us: www.exascale.org/bdec



What is BDEC-2?



1. BDEC2 (Next Generation) Planning, Chicago, IL, March 26-28, 2018
2. BDEC2 Bloomington Indiana, Applications Focus, November 28-30, 2018
3. BDEC2 Kobe, Japan, Platforms, February 19-21, 2019
4. BDEC2 Poznan, Poland, Demonstrator Projects, May 14-16, 2019
5. BDEC2 San Diego, AI/CI Convergence, October 15-17, 2019
6. BDEC2 Porto, Portugal, Application defined workflows, March 24-26, 2020

with BoF's at SC and ISC meetings



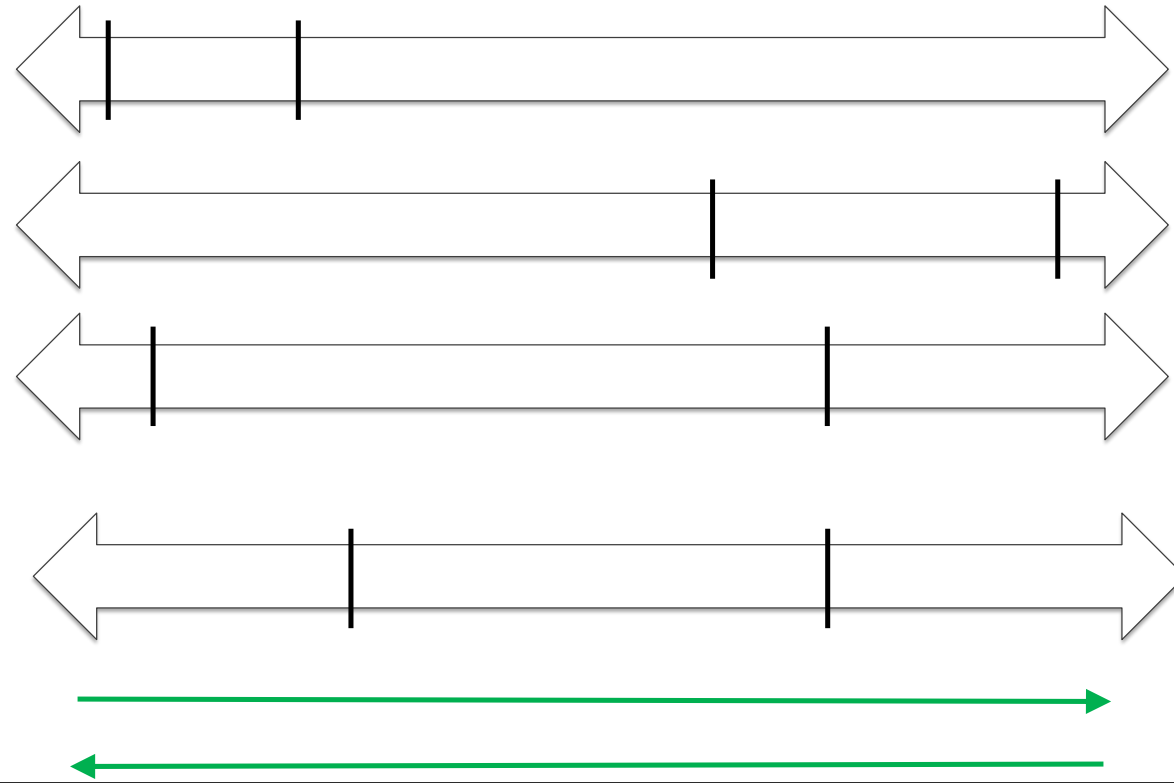
In the beginning...

- 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!



So, what is the Digital Continuum?

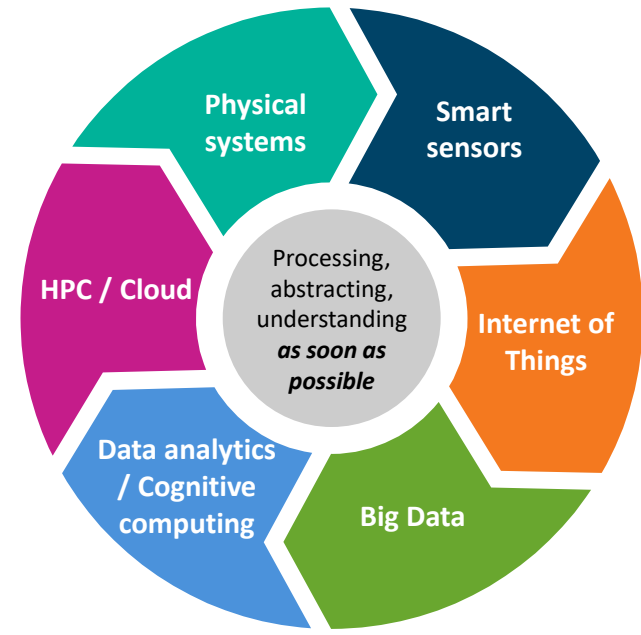
Building fluid cyberinfrastructure



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

Challenge: Programming The Computing Continuum



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^{11}

Count = 10^1
Size = 10^9

How will we write programs?

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**.



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...**



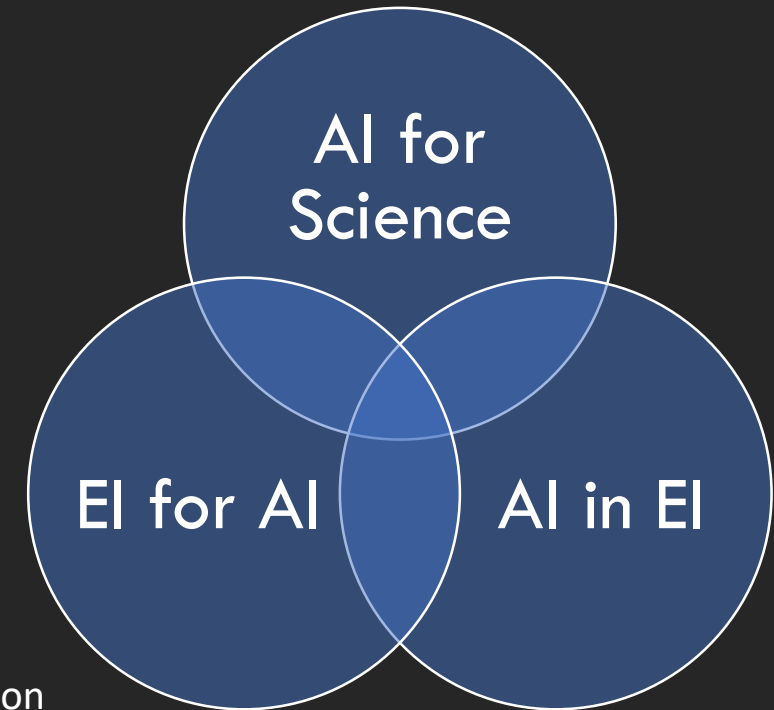
What are the next steps?

- 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

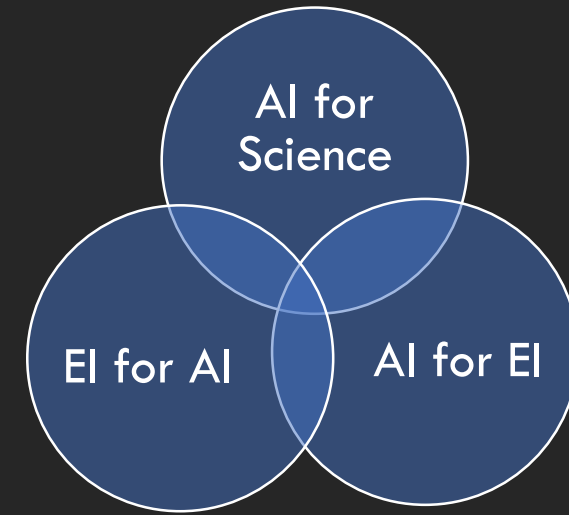


AI in the Continuum

- 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



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



Next Steps for BDEC



- Porto Workshop:
 - March 24th-26th 2020.
 - Organised by EXDCI-2, WP5.
 - Theme: Application Defined Workflows.
 - Opening session with EuroHPC
- Community Report:
 - Messages to be passed to stakeholders in ALL regions.
 - Distilled from Bloomington, Kobe, Poznan, San Diego, Porto and Asia workshops.
 - Will be presented at SC'20.
- And after that????



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.

