



BDEC

Mark Asch - mark.asch@u-picardie.fr

ETP4HPC - SRA-3 kickoff meeting

IBM IOT, Munich, March 20th 2017

March 20, 2017



Outline

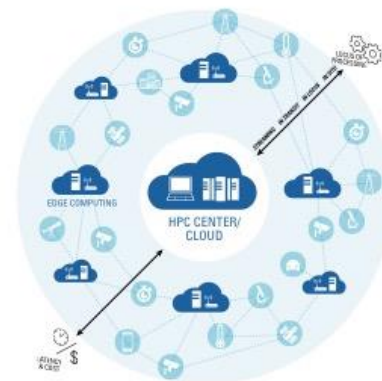
- What is BDEC?
- Why convergence?
- What are the pathways forward?
- Where does Europe stand today and what are the implications for the new SRA?

- Successor to the **IESP** (International Exascale Software Project) that played a central role in exascale roadmaps.
- **BDEC** holds 1-2 meetings per year since 2013 (US, Japan, EU)
- Latest meeting (#5) was hosted (for the 1st time!) by **China**...
- **Objective**: elaborate international roadmap and guidelines for achieving **convergence** between HPC and HDA



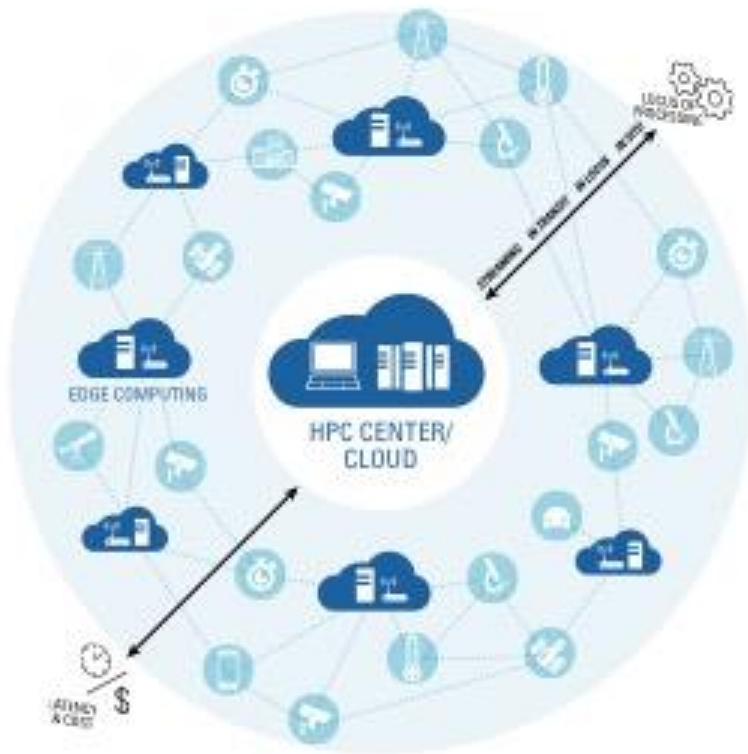
Convergence - why and how?

- Why?
 - we are in the era of Big Data Analytics / ML / DL
 - but there are 2 parallel worlds – both for software and infrastructure
 - public stakeholders will not finance both!
 - scientifically, we can advance by sharing competences and resources between the 2 worlds...



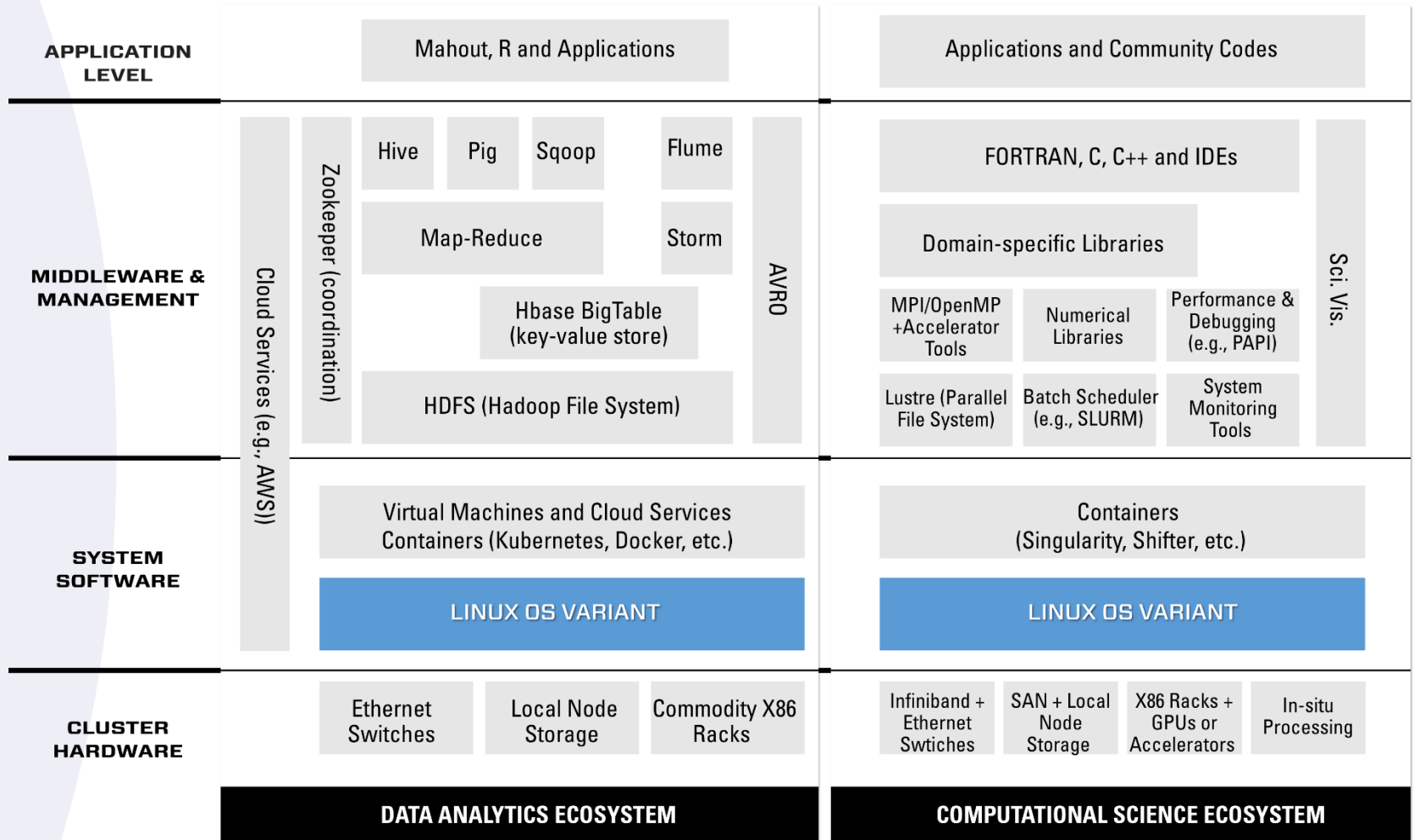
APPLICATION LEVEL	Mahout, R and Applications	Applications and Community Codes
MIDDLEWARE & MANAGEMENT	Hive Pig Sqoop Flume Map-Reduce Storm Hbase BigTable (key-value store) HDFS (Hadoop File System)	FORTRAN, C, C++ and IDEs Domain-specific Libraries MPI/OpenMP + Accelerator Tools Numerical Libraries Performance & Debugging (e.g., PAPI) Lustru (Parallel File System) Batch Scheduler (e.g., SLURM) System Monitoring Tools
	Virtual Machines and Cloud Services Containers (Kubernetes, Docker, etc.) LINUX OS VARIANT	Containers (Singularity, Shifter, etc.) LINUX OS VARIANT
	Ethernet Switches Local Node Storage Commodity X86 Racks	Infiniband + Ethernet Switches SAN – Local Node Storage X86 Racks + GPUs or Accelerators In-situ Processing
CLUSTER HARDWARE	DATA ANALYTICS ECOSYSTEM	COMPUTATIONAL SCIENCE ECOSYSTEM

The 2 worlds: hardware, data and architecture



- instruments and sensors are measuring more and more data (IoT)
- but these are out on the “edge”
- HPC and Cloud infrastructures are in the centre
- how and where are we going to do the processing?
- how can we limit communication costs?

The 2 worlds: software stacks



Pathways to Convergence: new realities

- rising importance of AI and ML;
- edge (IoT) is the next big thing...
- CDN and other in-transit processing must be considered;
- what does this mean for future hardware, software, architecture and applications?

Pathways to Convergence – how to advance?

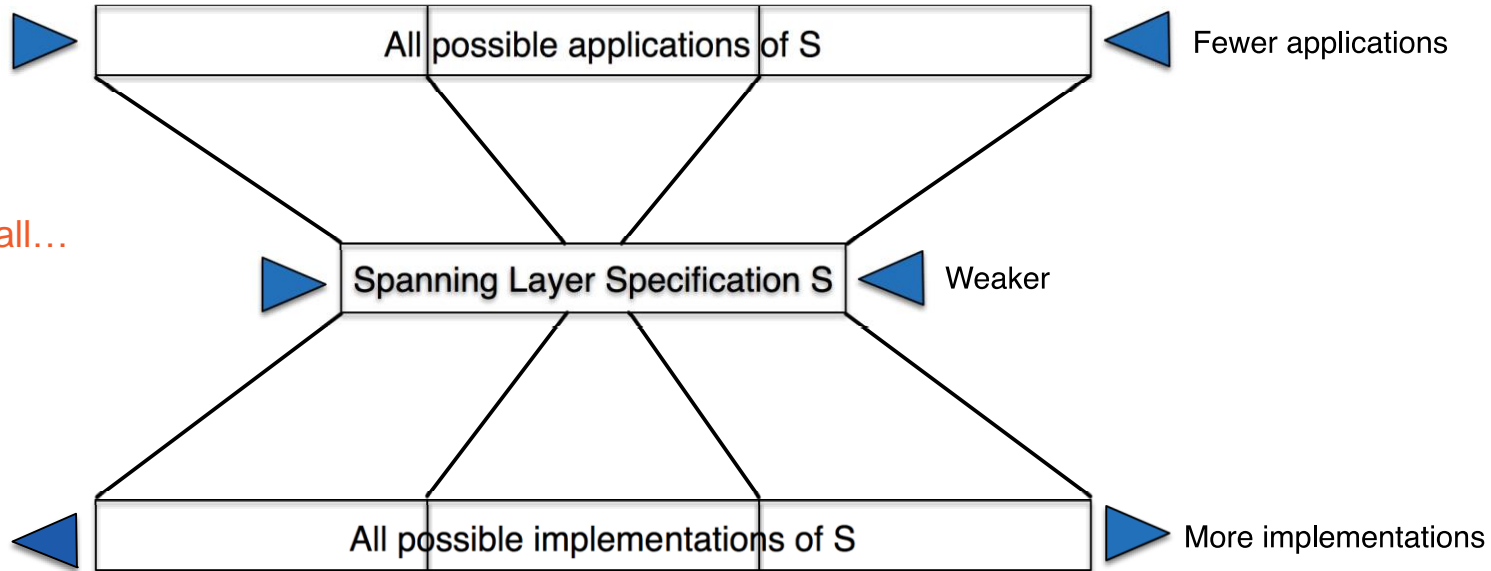
- HPC/Cloud Infrastructures (Physically or logically centralized resources)
 - hardware
 - software
 - applications
 - data
- Edge/Cloud Infrastructures
 - hardware
 - software
 - applications
 - data
- Applications and Workflows
- What does this mean for future hardware, software, architecture and applications?

Pathways to Convergence – some answers

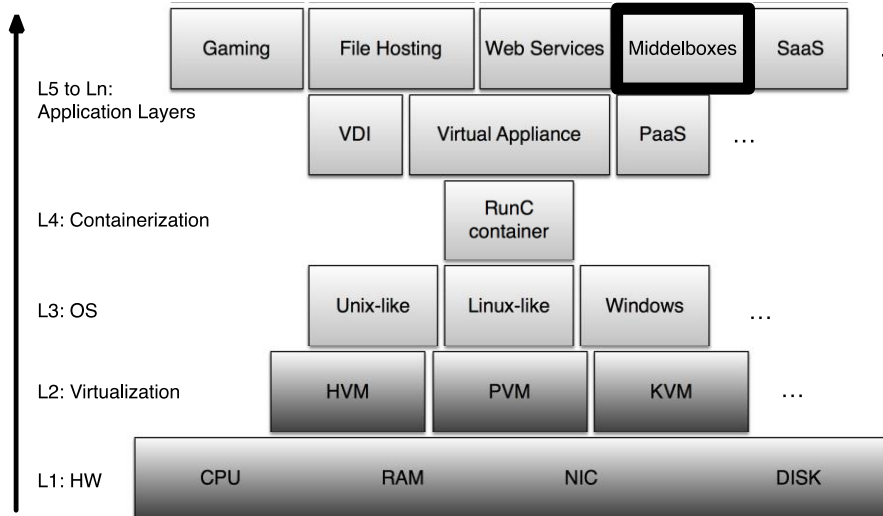
- **Answer #1:** China and (especially) Japan are building “converged” machines with both Big Data and HPC capacities... so just copy them?
- Answer #2: look for a new **narrow waist**? (see below)
- Answer #3: develop and concentrate on new **workflows**.
- Answer #4: **borrow** ideas from BD world and implement them in HPC (see also #1)
 - New file systems.
 - New data and communication procedures, resources and protocols.
- Answer #5: **new software** for data reduction, data analytics, dataflows, Function as a Service (FaaS)

Pathways to Convergence – narrow waist

You can't win 'em all...



Containers as narrow waist?



Where do we stand in Europe?

- In the (pure) exascale race, we're **3-4 years** behind...
- We need to **concentrate** on domains where we have (consensual) strengths and move forward on these:
 - federated platforms
 - algorithm and software development
 - analytical methods and tools
 - applications
- The **ETP4HPC-SRA** is the best vehicle for this! Use it...
- **EXDCI** is playing a central role in the BDEC initiative. Use it...

Thank you!

- Please consult www.exascale.org for reports and presentations from all BDEC meetings, where you can download the document “Pathways to Convergence”
- You can contact me at: mark.asch@u-picardie.fr