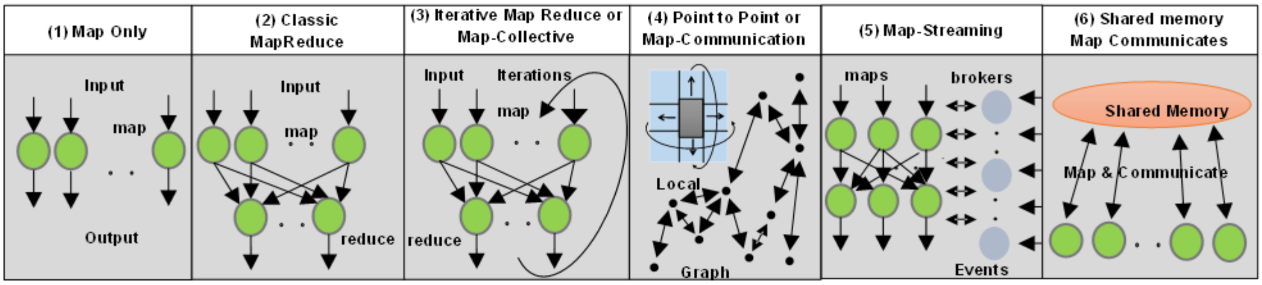
XPS Meeting Questions: Answers from RaPyDLI Collaborative Research: Rapid Prototyping HPC Environment for Deep Learning NSF 1439007

1. *Applications: What are the applications that motivate future systems? What advances in algorithms and programming systems will support these applications?*
   We assume this corresponds to NSF supported applications although commercial applications will be relevant to indicate synergies where research applications can take advantage of the large commercial investments in software (easiest as often open source) and algorithms (usually proprietary). Future systems will need of course to address a wide range of applications from high end exascale simulations to large scale data intensive problems and include both "big research" and the "long tail". Useful sources of data intensive applications are the National Academy study [1] and the NIST report [2]. Whereas there is a reasonable consensus on the architectures needed to support leading edge simulations, the same is not true for data. Is it commodity clouds with HDFS; distributed grids as used to analyze LHC data; HPC clusters with Lustre; high GPU but small node deep learning clusters or large shared memory graph engines? These different choices partly reflect different applications and more research is needed to understand application requirements and their mapping to different system classes. Here benchmarks would be valuable and these are well studied in database arena [3] with some extensions but without systematic study of research data analysis [4]. The National Academy Study has a good discussion of new algorithm research. One important highlight is the study of $O(N^2)$ algorithms and deriving related O(N) or O(NlogN) fast methods. This is known in some examples but not broadly understood – especially for practical problems. Sampling methods are good here and in other algorithms although some say that research data is so valuable that sampling is less effective. A second highlight of the National Academy Study, is streaming applications which vary from large things (such as light sources and telescopes) through the much-anticipated Internet of Things where up to 70 billion devices are forecast by 2020. The exascale simulation studies have shown that the data visualizations form intense streaming problems with potentially 100's of petabyte/sec bandwidth. Streaming applications have new hardware, software and algorithm considerations. The National Academy study [1] highlights new online algorithms needed to analyze such data. More broadly we can note that some of today's libraries like R or Mahout do not have good parallel performance and better parallel machine learning is needed across the board. Issues such as use of parameter servers – common in machine learning but not HPC -- need to studied.

2. *Systems: What hardware architectures and distributed systems will support future applications? What challenges do we face in design and management?*
   We mentioned in answer to 1) that there is little consensus as to system architecture for data intensive problems although we earlier introduced 6 classes in the figure below [5] which appears

to describe a large number of commercial and research use cases.



These 6 classes can of course be further refined with high throughput computing (class 1) having centralized or grid (distributed) implementations. Further as discussed above, a major RaPyDLI deep learning target is a subset of HPC cluster (class 4) with a few nodes and many GPUs per node. We see a major challenge in getting consensus as to the critical architectures; and for example, recent NSF data infrastructure awards don't clearly cover this spectrum well. In data management area, the research community has made limited use of commercially dominate solutions such as object stores, NoSQL and NewSQL. Research makes much heavier direct use of files than the commercial systems; however we expect a movement to commercial data architectures and away from files. The importance of hypervisors (clouds), and DevOps technologies like Docker, needs to be understood better. We expect growing use of sophisticated scripting languages like Python (or Julia perhaps) and we need to design and implement such languages so they perform well for all parallel computing; i.e. for all 6 classes of the figure and not just pleasingly parallel (Map only) case 1).

3. *Technologies: What emerging technologies will change fundamental assumptions in hardware and software? What constraints disappear? What challenges arise?*
There are many interesting driving forces such as the march to Exascale (with task oriented run time), the Internet of Things, Software defined Networks (growing to software defined systems), Clouds and commodity big data, growing number of cores per node and use of nonvolatile memory on nodes. These allow better performance but many challenges as these disruptive technologies are not well coordinated with themselves and with the research Big Data arena.

4. *Methodologies: How should we perform interdisciplinary research that spans applications, systems, and technologies? How should abstraction layers evolve?*
We don't fully understand this question. Both abstractions and interdisciplinary research are important but are not directly correlated. Nearly always a good system is built around abstractions although at the moment, it is not clear that there are many abstractions that are heavily used in data intensive case. Well respected application data standards like HDF, OPeNDAP, and FITS probably need to be reworked for modern technology such as NoSQL and object stores. MPI in parallel computing is likely to end up disappearing and perhaps wrapped in Hadoop [6] or Spark. Probably the hardest part of interdisciplinary work is collaboration where tools are still fragmented and glitchy. For example, Google docs is popular as a collaborative tool but its simple formatting capability limits its value. Even telecons or videocons are pretty unreliable especially

with growing use of mobile phones. We haven't advanced that much over the Access Grid.

5. *Risks: What are the risks that threaten the success of XPS research directions? How do we guard and hedge against these threats?*
Some risks come from the vibrant commercial big data world. Their application mix (with search and recommender systems and giant numbers of users) is quite different from research although there are some commonalities. It is hard to find the best balance between re-using commercial technologies and enhancing research value. Collaboration with industry and identifying those cases where commercial approaches are inadequate [7] seem good steps; XPS needs to focus here. These dangers/difficulties can be illustrated by the workflow area where commercial big data system appear to ignore the excellent grid/HPC work in this area, and are developing new solutions such as NiFi and Crunch. It is important that both simulation and data sides of parallel research reach out and work with the commodity approaches illustrated by Apache Foundation projects.

References

1. Committee on the Analysis of Massive Data; Committee on Applied and Theoretical Statistics; Board on Mathematical Sciences and Their Applications; Division on Engineering and Physical Sciences; National Research Council, *Frontiers in Massive Data Analysis*. 2013: National Academies Press. http://www.nap.edu/catalog.php?record_id=18374
2. NIST, *NIST Big Data Public Working Group (NBD-PWG) Use Cases and Requirements.* 2013. http://bigdatawg.nist.gov/usecases.php
3. Chaitan Baru and Tilmann Rabl. *Tutorial 4 " Big Data Benchmarking" at 2014 IEEE International Conference on Big Data*. 2014 [accessed 2015 January 2]; Available from: http://cci.drexel.edu/bigdata/bigdata2014/tutorial.htm.
4. Geoffrey C. FOX , Shantenu JHA, Judy QIU, Saliya EKANAYAKE, and Andre LUCKOW, *Towards a Comprehensive Set of Big Data Benchmarks*. February 15, 2015. http://grids.ucs.indiana.edu/ptliupages/publications/OgreFacetsv9.pdf.
5. Geoffrey C. Fox, Shantenu Jha, Judy Qiu, and Andre Luckow, *Towards an Understanding of Facets and Exemplars of Big Data Applications*, in *20 Years of Beowulf: Workshop to Honor Thomas Sterling's 65th Birthday* April 13, 2015. Annapolis http://dsc.soic.indiana.edu/publications/ogre.v1.2.pdf.
6. Bingjing Zhang, Yang Ruan, and Judy Qiu, *Harp: Collective Communication on Hadoop*, in *IEEE International Conference on Cloud Engineering (IC2E)*. March 9-12, 2015. Tempe AZ. http://grids.ucs.indiana.edu/ptliupages/publications/HarpQiuZhang.pdf.
7. Geoffrey Fox, Judy Qiu, Shantenu Jha, Supun Kamburugamuve, and Andre Luckow, *HPC-ABDS High Performance Computing Enhanced Apache Big Data Stack*, in *Invited talk at 2nd International Workshop on Scalable Computing For Real-Time Big Data Applications (SCRAMBL'15) at.CCGrid2015, the 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*. 2015, IEEE. Shenzhen, Guangdong, China. http://dsc.soic.indiana.edu/publications/HPC-ABDSDescribedv2.pdf.