**Need**

Hour glass figure

Tables of NIST use cases and their properties: Done

Tables of components in OGRE facets: Done

HPC-ABDS figure

**Sources of Information**

In discussing the structure of Big Data Applications, let us first discuss the inevitably incomplete input that we used to do our analysis. We have gained of course quite a bit of experience from our research over many years but 3 explicit sources that we used were a recent use case survey by NIST from Fall 2013; a survey of data intensive research applications (Distributed Computing MetaPatterns) by Jha et al.; and study of members of data analytics libraries including R, Mahout and MLLib. We follow with a summary of first two sources.

The NIST Big Data Public Working Group (NBD-PWG) was launched in June 2013 with a set of working groups covering Big Data Definitions, Taxonomies, Requirements, Security and Privacy Requirements, Reference Architectures White Paper Survey, Reference Architectures, Security and Privacy Reference Architectures and Big Data Technology Roadmap. The Requirements working group gathered 51 use cases from a public call and then analyzed in terms of requirements of a reference architecture. Here we will look at them differently to identify common patterns and characteristics which can be used to guide and evaluate Big Data hardware and software.

The use cases are organized into nine broad areas as follows, with the number of associated use cases in parentheses:

* Government Operation (4)
* Commercial (8)
* Defense (3)
* Healthcare and Life Sciences (10)
* Deep Learning and Social Media (6)
* The Ecosystem for Research (4)
* Astronomy and Physics (5)
* Earth, Environmental and Polar Science (10)
* Energy (1)

Note that the majority of use cases come from research applications but commercial, defense and government operations have some coverage. A template was prepared by the requirements working group, which allowed experts to categorize each use case by 26 features that included

Use case Actors/Stakeholders and their roles and responsibilities; use case goals and description. Specification of current analysis covering compute system, storage, networking and software. Characteristics of use case Big Data with Data Source (distributed/centralized), Volume (size), Velocity (e.g. real time), Variety (multiple datasets, mashup), Variability (rate of change). The so-called Big Data Science (collection, curation, analysis) with Veracity (Robustness Issues, semantics), Visualization, Data Quality (syntax), Data Types and Data Analytics. These detailed specifications were complemented by broad comments including Big Data Specific Challenges (Gaps), Mobility issues, Security & Privacy Requirements and identification of issues for generalizing this use case.

The complete set of 51 responses with in addition a summary from the working group of applications, current status and futures as well as extracted requirements can be found in []

Need a discussion of Jha et al.

**Lessons from Parallel Computing**

Before discussing features and patterns of Big Data applications, it is instructive to consider the better understood parallel computing situation.

**Commercial Big data is big**

**HPC-ABDS Overview**

Layers

Target 5 Architectures

Principles including Hourglass

**NIST Review**

Template

Table of 51 use cases

Examples of V's properties

20 additional use case

**Building Ogres**

Summary of mini-apps

NPB

Berkeley dwarfs

Distributed Computing MetaPatterns (see CloudDB talk)

**Properties of properties of 51 use cases**

What is parallelism Over

Low-Level (Run-time) Computational Types

Higher-Level Computational Types or Features

Global Machine Learning aka EGO – Exascale Global Optimization

**A set of Image data related use cases (Better summarized in Cetraro)**

Facebook etc.

17:Pathology Imaging/ Digital Pathology

18: Computational Bioimaging

26: Large-scale Deep Learning

27: Organizing large-scale, unstructured collections of consumer photos

36: Catalina Real-Time Transient Survey (CRTS)

43: Radar Data Analysis for CReSIS Remote Sensing of Ice Sheets

44: UAVSAR Data Processing, Data Product Delivery, and Data Services

**Facets of the Ogres**

Application Class Facet of Ogres

Problem Architecture Facet of Ogres (Meta or MacroPattern)

4 Forms of MapReduce

Computational Features

Data Source and Style Facet

**Analytics Facet (kernels) of the Ogres**

Several examples

**Parallel Global Machine Learning Examples**

MDS and Clustering pictures

Harp WDA-SMACOF efficiency results

Summary of Harp

Spherical Phylogram from MDS

More results from Azure and original hpc-abds paper (In Spare Slides and CloudDB talk)

**Comparing Data Intensive and Simulation Problems**

Communication structure

Graphs v Particle Dynamics

**Comparing Commercial and Science uses**

Cetraro: Note key difference between commercial and science. Former has huge number of users which map into classic mapreduce. Latter needs much of this capability for management/access functions but different scale. However backend use issues comparable in 2 cases with of course different algorithms

**Sensors and the Internet of Things - Streaming**

Sensors/Streaming as a Service

Relevance of cloud

Projection of IoT size

IOTCloud and performance

39: Particle Physics: Analysis of LHC Large Hadron Collider Data: Discovery of Higgs particle

13: Large Scale Geospatial Analysis and Visualization

50: DOE-BER AmeriFlux and FLUXNET Networks

51: Consumption forecasting in Smart Grids

**Java Grande**

Kernel mpiJava performance

DAVS Java clustering performance C# v Java

Results of DAVS

**Conclusions**

Integrate (don’t compete) HPC with “Commodity Big data” (Google to Amazon to Enterprise Data Analytics)

i.e. improve Mahout; don’t compete with it

Use Hadoop plug-ins rather than replacing Hadoop

Enhanced Apache Big Data Stack HPC-ABDS has ~120 members

Opportunities at Resource management, Data/File, Streaming, Programming, monitoring, workflow layers for HPC and ABDS integration

Data intensive algorithms do not have the well developed high performance libraries familiar from HPC

Strong case for high performance Java (Grande) run time supporting all forms of parallelism

|  |  |  |  |
| --- | --- | --- | --- |
| 9 Image-based NIST Use Cases | | | |
| Use Case | **Title** | **Application** | **Features** |
| 17 | Pathology Imaging/ Digital Pathology | Moving to terabyte size 3D images, Global Classification | PP, LML, MR for search |
| 18 | Light sources | Biology and Materials | PP, LML |
| 26 | Large-scale Deep Learning | Stanford ran 10 million images and 11 billion parameters on a 64 GPU HPC; vision (drive car), speech, and Natural Language Processing | GML |
| 27 | Organizing large-scale, unstructured collections of photos | Fit position and camera direction to assemble 3D photo ensemble | GML |
| 36 | Catalina Real-Time Transient Synoptic Sky Survey (CRTS) | Processing of individual images for events based on classification of image structure (GML) | PP, LML |
| 43 | Radar Data Analysis for CReSIS Remote Sensing of Ice Sheets | Identify glacier beds and snow layers | PP, LML moving to GML for full ice-sheet |
| 44 | UAVSAR Data Processing, Data Product Delivery, and Data Services | Find and display slippage from radar images | PP |
| 45, 46 | Analysis of Simulation visualizations | Find paths, classify orbits, classify patterns that signal earthquakes, instabilities, climate, turbulence | PP LML ?GML |

|  |  |  |  |
| --- | --- | --- | --- |
| Internet of Things (8) and Streaming Apps: Properties and Examples | | | |
| There will be 24 (Mobile Industry Group) to 50 (Cisco) billion devices on the Internet by 2020. | | | |
| The cloud is natural controller of and resource provider for the Internet of Things. | | | |
| Broad categories are PC’s/tablets/Smart phones/watches, Wearable devices (Smart People), “Intelligent River” “Smart Homes and Grid” “Ubiquitous Cities” and Robotics. | | | |
| Majority of NIST use cases are streaming – experimental science gathers data in a stream – sometimes batched as in a field trip. | | | |
| Use Case | **Title** | **Application** | **Features** |
| 10: | Cargo Shipping | Tracking as in UPS, Fedex | PP GIS LML |
| 11-13 | Military Sensor Networks | Large Scale Geospatial Analysis and Visualization | PP GIS LML |
| 28 | Truthy | Information diffusion research from Twitter Data | PP MR for Search, GML for community determination |
| 39 | Experimental Particle Physics | Analysis of LHC Large Hadron Collider Data: Discovery of Higgs particle | PP Local Processing Global statistics (MRStat) |
| 50 | Environmental Networks | DOE-BER AmeriFlux and FLUXNET | PP GIS LML |
| 51 | Smart Grids | Energy consumption forecasting | PP GIS LML |

|  |  |
| --- | --- |
| What is Parallelism Over for NIST Use Cases? | |
| General Class | **Examples** | |
| People | Users (but see below) or Subjects of application and often both |
| Decision makers | Researchers or doctors (users of application) |
| Items | Experimental observations |
| Contents of online store |
| Images or “Electronic Information nuggets” |
| EMR: Electronic Medical Records (often similar to people parallelism) |
| Protein or Gene Sequences |
| Material properties, Manufactured Object specifications, etc., in custom dataset |
| Modelled entities | Vehicles and people |
| Sensors | Internet of Things |
| Events | Detected anomalies in telescope, credit card or atmospheric data |
| Graph Nodes | RDF databases |
| Regular Nodes | Simple nodes as in a learning network |
| Information Units | Tweets, Blogs, Documents, Web Pages, etc. and characters/words in them |
| Files or data | To be backed up, moved or assigned metadata |
| Particles/cells/ mesh points | Used in parallel simulations |

|  |  |  |
| --- | --- | --- |
| Low-Level (Run-time) Computational Types of NIST Use Cases | | |
| Label | **#** | **Description** |
| PP | 26 | Pleasingly Parallel or Map Only |
| MR | 18 | Classic MapReduce MR (add MRStat below for full count) |
| MRStat | 7 | Simple version of MR where key computations are simple reduction as found in statistical averages such as histograms and averages |
| MRIter | 23 | Iterative MapReduce or MPI |
| Graph | 9 | Complex graph data structure needed in analysis |
| Fusion | 11 | Integrate diverse data to aid discovery/decision making; could involve sophisticated algorithms or could just be a portal |
| Streaming | 41 | Some data comes in incrementally and is processed this way |

# is count in 51 use cases

|  |  |  |
| --- | --- | --- |
| Higher-Level Computational Types or Features of NIST Use Cases | | |
| Label | **#** | **Description** |
| Class | 30 | **Classification:** divide data into categories |
| S/Q | 12 | Index, Search and Query |
| CF | 4 | Collaborative Filtering for recommender engines |
| LML | 36 | Local Machine Learning (Independent for each parallel entity) |
| GML | 23 | **Global Machine Learning:** Deep Learning, Clustering, LDA, PLSI, MDS,  Large Scale Optimizations as in Variational Bayes, Lifted Belief Propagation, Stochastic Gradient Descent, L-BFGS, Levenberg-Marquardt  Sometimes call EGO or Exascale Global Optimization with scalable parallel algorithm |
|  | 51 | **Workflow**: Universal so no label |
| GIS | 16 | Geotagged data and often displayed in ESRI, Microsoft Virtual Earth, Google Earth, GeoServer etc. |
| HPC | 5 | Classic large-scale simulation of cosmos, materials, etc. generating (visualization) data |
| Agent | 2 | Simulations of models of data-defined macroscopic entities represented as agents |

|  |  |
| --- | --- |
| # | Applications Facet of Big Data Ogres |
| 30 | Classification divide data into categories including collaborative filtering |
| 12 | Search Index and query |
|  | Maximum Likelihood or χ2 minimizations |
|  | Expectation Maximization (often Steepest descent) |
| 36 | Local (pleasingly parallel) Machine Learning |
| 23 | GML or (Exascale) Global Optimization such as Learning Networks, Variational Bayes and Gibbs Sampling |
| 2 | Use Agents as in epidemiology (swarm approaches)? |

|  |
| --- |
| Problem Architecture Facet of Ogres (Meta or Macro Pattern) |
| Pleasingly Parallel as in BLAST, Protein docking, some (bio-)imagery including Local Analytics or Local Machine Learning with pleasingly parallel filtering, as in light source data, radar images |
| Classic MapReduce for Search and Query |
| Global Analytics or Machine Learning requiring iterative programming models |
| Problem set up as a graph as opposed to vector, grid |
| SPMD (Single Program Multiple Data) |
| Bulk Synchronous Processing: well-defined compute-communication phases |
| Fusion: Knowledge discovery often involves fusion of multiple methods. |
| Workflow (often used in fusion) |

|  |
| --- |
| Computational Features Facet of Ogres |
| Flops per byte: important for performance |
| Communication Interconnect requirements; |
| Is application (graph) constant or dynamic? |
| Most applications consist of a set of interconnected entities; is this regular as a set of pixels or is it a complicated irregular graph? |
| Is communication BSP or Asynchronous? In latter case shared memory may be attractive; |
| Are algorithms Iterative or not? |
| Data Abstraction: key-value, pixel, graph, vector |
| Are data points in metric or non-metric spaces? |
| Core libraries needed: matrix-matrix/vector algebra, conjugate gradient, reduction, broadcast …. |

|  |
| --- |
| Data Source and Style Facet of Ogres |
| SQL |
| NOSQL based |
| Other Enterprise data systems (10 examples from NIST []) |
| Set of Files (as managed in iRODS) |
| Internet of Things |
| Streaming |
| HPC simulations |
| Involve GIS (Geographical Information Systems) |
| Before data gets to compute system, there is often an initial data gathering phase which is characterized by a block size and timing. Block size varies from month (Remote Sensing, Seismic) to day (genomic) to seconds or lower (Real time control, streaming) |
| There are storage/compute system styles: Shared, Dedicated, Permanent, Transient |
| Other characteristics are needed for permanent auxiliary/comparison datasets and these could be interdisciplinary, implying nontrivial data movement/replication |

|  |  |
| --- | --- |
| Distinctive Architectures for Data Analytics | |
| Pleasingly Parallel | Includes local machine learning (LML) as in parallel decomposition over items and apply data processing to each item. Hadoop could be used but also many other High Throughput Computing and Many task tools |
| Search | Includes applications using collaborative filtering and motif finding implemented using classic MapReduce (Hadoop) |
| Map-Collective | Iterative MapReduce using Collective Communication as in clustering – Hadoop with Harp, Spark etc. |
| Map-Communication | Iterative MapReduce such as Giraph with point-to-point communication and includes most graph algorithms such as maximum clique, connected component, finding diameter, community detection). Vary in difficulty of finding partitioning (classic parallel load balancing) |
| Shared Memory | Thread-based (event driven) graph algorithms such as shortest path and Betweenness centrality |

|  |
| --- |
| Core Analytics Facet of Ogres (microPattern) |
|  |
| Pleasingly Parallel (Map Only) or Local Machine Learning: ~any algorithm |
| Map-Reduce |
| Search, Query, Index: Dominant commercial use and important in Science with less users |
| Recommender Systems including Collaborative filtering: Dominant commercial use, Little Science |
| Summarizing statistics (MRStat) as in LHC Data analysis (histograms) |
| Linear Classifiers: Bayes, Random Forests |
| Global Analytics – Nonlinear Solvers (Structure depends on Objective Function) |
| Stochastic Gradient Descent SGD |
| (L-)BFGS approximation to Newton’s Method |
| Levenberg-Marquardt solver |
| Global Analytics – Map-Collective (See Mahout, MLlib) |
| Outlier Detection |
| Clustering (many methods) |
| Mixture Models, LDA (Latent Dirichlet Allocation), PLSI (Probabilistic Latent Semantic Indexing) |
| SVM and Logistic Regression |
| PageRank (find leading eigenvector of sparse matrix) |
| SVD (Singular Value Decomposition) |
| MDS (Multidimensional Scaling) |
| Learning Neural Networks (Deep Learning) |
| Hidden Markov Models |
| Global Analytics – Map-Communication (targets for Giraph) |
| Graph Structure (Communities, subgraphs/motifs, diameter, maximal cliques, connected components) |
| Network Dynamics - Graph simulation Algorithms (epidemiology) |
| Global Analytics – Asynchronous Shared Memory |
| Graph Structure (Betweenness centrality, shortest path) |



|  |  |
| --- | --- |
| Table: Kaleidoscope of Apache Big Data Stack (ABDS) and HPC Technologies | |
| |  | | --- | | Cross-Cutting Functionalities | | Message Protocols: Thrift, Protobuf | | Distributed Coordination: Zookeeper, JGroups | | Security & Privacy: InCommon, OpenStack Keystone, LDAP | | Monitoring: Ambari, Ganglia, Nagios, Inca | | |  | | --- | | **Workflow-Orchestration:** Oozie, ODE, Airavata, OODT (Tools), Pegasus, Kepler, Swift, Taverna, Trident, ActiveBPEL, BioKepler, Galaxy, IPython | | **Application and Analytics:** Mahout , MLlib , MLbase, CompLearn, R, Bioconductor, ImageJ, Scalapack, PetSc | | **High level Programming:** Hive, HCatalog, Pig, Shark, MRQL, Impala, Sawzall | | **Basic Programming model and runtime**, **SPMD, Streaming, MapReduce, MPI:** Hadoop, Spark, Twister, Stratosphere, Tez, Hama, Storm, S4, Samza, Giraph, Pregel, Pegasus | | **Inter process communication Collectives, point-to-point, publish-subscribe:** Hadoop, Spark, Harp, MPI, Netty, ZeroMQ, ActiveMQ, QPid, Kafka, Kestrel | | **In-memory databases/caches:** GORA (general object from NoSQL), Memcached, Redis (key value), Hazelcast, Ehcache | | **Object-relational mapping:** Hibernate, OpenJPA and JDBC Standard | | **Extraction Tools:** UIMA, Tika | | **SQL:** Oracle, MySQL, Phoenix, SciDB | | **NoSQL:** HBase, Accumulo, Cassandra, Solandra, MongoDB, CouchDB, Lucene, Solr, Berkeley DB, Azure Table, Dynamo, Riak, Voldemort. Neo4J, Yarcdata, Jena, Sesame, AllegroGraph, RYA | | **File management:** iRODS | | **Data Transport:** BitTorrent, HTTP, FTP, SSH, Globus Online (GridFTP) | | **Cluster Resource Management**: Mesos, Yarn, Helix, Llama, Condor, SGE, OpenPBS, Moab, Slurm, Torque | | **File systems:** Swift, Cinder, Ceph, FUSE, Gluster, Lustre, GPFS, GFFS | | **Interoperability:** Whirr, JClouds, OCCI, CDMI | | **DevOps:** Docker,Puppet, Chef, Ansible, Boto, Libcloud, Cobbler, CloudMesh | | **IaaS Management from HPC to hypervisors:** OpenStack, OpenNebula, Eucalyptus, CloudStack, vCloud, Amazon, Azure, Google | |

|  |
| --- |
| The 71 NIST Use Cases with number in each broad area |
| Government Operation(4): National Archives and Records Administration, Census Bureau |
| Commercial(8): Finance in Cloud, Cloud Backup, Mendeley (Citations), Netflix, Web Search, Digital Materials, Cargo shipping (as in UPS) |
| Defense(3): Sensors, Image surveillance, Situation Assessment |
| Healthcare and Life Sciences(10): Medical records, Graph and Probabilistic analysis, Pathology, Bioimaging, Genomics, Epidemiology, People Activity models, Biodiversity |
| Deep Learning and Social Media(6): Driving Car, Geolocate images/cameras, Twitter, Crowd Sourcing, Network Science, NIST benchmark datasets |
| The Ecosystem for Research(4): Metadata, Collaboration, Language Translation, Light source experiments |
| Astronomy and Physics(5): Sky Surveys including comparison to simulation, Large Hadron Collider at CERN, Belle Accelerator II in Japan |
| Earth, Environmental and Polar Science(10): Radar Scattering in Atmosphere, Earthquake, Ocean, Earth Observation, Ice sheet Radar scattering, Earth radar mapping, Climate simulation datasets, Atmospheric turbulence identification, Subsurface Biogeochemistry (microbes to watersheds), AmeriFlux and FLUXNET gas sensors |
| Energy(1): Smart grid |
| Enterprise Data Systems(10): Multiple users performing interactive queries and updates on a database with basic availability and eventual consistency (BASE); Perform real time analytics on data source streams and notify users when specified events occur; Move data from external data sources into a highly horizontally scalable data store, transform it using highly horizontally scalable processing (e.g. Map-Reduce), and return it to the horizontally scalable data store (ELT); Perform batch analytics on the data in a highly horizontally scalable data store using highly horizontally scalable processing (e.g MapReduce) with a user-friendly interface (e.g. SQL like); Perform interactive analytics on data in analytics-optimized database; Visualize data extracted from horizontally scalable Big Data store; Move data from a highly horizontally scalable data store into a traditional Enterprise Data Warehouse; Extract, process, and move data from data stores to archives; Combine data from Cloud databases and on premise data stores for analytics, data mining, and/or machine learning; Orchestrate multiple sequential and parallel data transformations and/or analytic processing using a workflow manager |
| Security & Privacy(10): Consumer Digital Media Usage; Nielsen Homescan; Web Traffic Analytics; Health Information Exchange; Personal Genetic Privacy; Pharma Clinic Trial Data Sharing; Cyber-security; Aviation Industry; Military - Unmanned Vehicle sensor data; Education - “Common Core” Student Performance Reporting |