Big Data Use Cases and Requirements

GEOFFREY FOX, Indiana University, School of Informatics and Computing
Co-Chair, Use Cases & Requirements Subgroup, NIST Big Data Public Working Group

WO CHANG, National Institute of Standards and Technologies
Co-Chair, NIST Big Data Public Working Group

Abstract

We formed a community of interest from industry, academia, and government, with the goal of developing a consensus set of Big Data requirements across all stakeholders. The major activities were gathering various use cases from diversified application domains and extracting requirements. Initially we developed a use case template with 26 fields, and these were completed for 51 areas. They were spread over application sectors as follows: Government Operations (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). These are, of course, only representative, and miss many important cases; but they form an interesting set for initial study. After gathering the use cases, a multi-step process was used to extract requirements. First, specific requirements were extracted from each use case. Those specific requirements were then mapped to broad characteristics, which were motivated by the structure of the NIST PWG reference architecture. These characteristics were Data sources, Data transformation, Capability infrastructure, Data Consumer, Security & Privacy, Lifecycle management and "Other" where the latter catch-all largely included mobile access. Then we aggregated all the 439 specific requirements into 35 high-level generalized requirements, which are vendor-neutral and technology agnostic.

1 Introduction

On June 19, 2013, the NIST Big Data Public Working Group (NBD-PWG) was launched with participation from industry, academia, and government from across the nation. The scope of NBD-PWG involves forming a community of interests from all sectors—including industry, academia, and government—with the goal of developing a consensus on definitions, taxonomies, secure reference architectures, and a technology roadmap. Such a consensus would create a vendor-neutral, technology- and infrastructure-independent framework that would enable Big Data stakeholders to identify and use the best analytics tools for their processing and visualization requirements on the most suitable computing platform and cluster, while also allowing value-added from Big Data service providers.

NBD-PWG currently has created five subgroups: Definitions and Taxonomies, Security and Privacy, Reference Architecture, Technology Roadmap plus the Use Case and Requirements subgroup whose work is discussed here.

The initial focus of the NBD-PWG Use Case and Requirements Subgroup was to form a community of interest from industry, academia, and government, with the goal of developing a consensus list of big data requirements across all stakeholders. This included gathering and understanding various use cases from diversified application domains.

Tasks assigned to the subgroup include the following:

- Gather input from all stakeholders regarding big data requirements. A goal that turned into gathering use cases
- Analyze/prioritize a list of challenging general requirements derived from use cases that may delay or prevent adoption of big data deployment.
- Develop a comprehensive list of big data requirements.

This report was produced by an open collaborative process involving weekly telephone conversations and information exchange using the NIST document system. The 51 use cases came from participants in the calls (subgroup members), and from others informed of the opportunity to contribute. The activity culminated in a public meeting September 30 2013 at NIST.

The use cases are organized into the nine broad sectors/areas (application domains) listed below with examples and number of use cases in parentheses:

- **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)** Self-driving cars, Geolocate Images, Twitter, Crowd Sourcing, Network Science, NIST Benchmark Datasets
- **Ecosystem for Research (4):** Metadata, Collaboration, Language Translation, Light Source Experiments
- **Astronomy and Physics (5):** Sky Surveys (and comparisons to simulation), LHC 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

Many presentations have been given [1-6] on the use case activity and follow-on analysis has been presented identifying common patterns [7, 8] and mapping into the well-known Apache [9] software environment [10, 11].

2 Use Cases

The working group began by defining a use case template, shown in Table 1. The template was valuable for gathering consistent information, thus supporting analysis and comparison of the use cases. However, the use cases are described in varying levels of detail. In addition, while some use cases are described in (mostly) qualitative terms, others include detailed quantitative data/metrics. The individual use cases may be downloaded from the NIST document library [12]. In addition, all 51 use cases are compiled in a single document [13] and are published by NIST as part of their Big Data document collection [14]. We stress that all use cases have been submitted openly, and no significant editing has been performed. There are differences in scope and interpretation, but the benefits of free and open submission outweigh those of greater uniformity.

Table 1: NBD (NIST Big Data) Requirements WG Use Case Template Aug 11 2013

Use Case Title		
Vertical (area)		
Author/Company/Email		
Actors/Stakeholders and their roles and responsibilities		
Goals		
Use Case Description		
Current Solutions	Compute(System)	
	Storage	
	Networking	
	Software	
Big Data Characteristics	Data Source (distributed/centralized)	

	Volume (size)	
	Velocity (e.g. real time)	
	Variety (multiple datasets, mashup)	
	Variability (rate of change)	
Big Data Science (collection, curation, analysis, action)	Veracity (Robustness Issues, semantics)	
	Visualization	
	Data Quality (syntax)	
	Data Types	
	Data Analytics	
Big Data Specific Challenges (Gaps)		
Big Data Specific Challenges in Mobility		
Security & Privacy Requirements		
Highlight issues for generalizing this use case (e.g. for ref. architecture)		
More Information (URLs)		
Note: <additional comments>		
Notes to preparer: No proprietary or confidential information should be included. Please ADD picture of operation or data architecture of application below table		

The 51 use cases are given in Table 2 divided into 9 categories and listing counts of requirements derived by the process described in Section 3. Table 2 also gives the institutional source of each use case. Note these use cases have a higher fraction of scientific research than a sample that reflected actual sizes of data. For example probably the largest science data sizes are <~100 petabytes (use case 39 – LHC) which is <0.00002 of the size of digital universe of shared electronic data [15, 16]. The latter are dominated by data from commercial sector in Table 2 with only 8 of 51 use cases. The 51 use cases are organized into the nine application domains introduced in section 1. For some domains, multiple similar big data applications are presented, providing a more complete view of big data requirements in that domain. The full report of the working group [14] has substantial detail on each use case including:

- The completed use case template of Table 1
- A relatively uniform summary of each use case consisting of 3 components: A**pplication Overview**, **Current** processing approach and **Futures** of application and analysis.
- Diagrams and figures illustrating some of use cases
- Summary of key properties for each use case: Data Volume, Velocity, Variety and Software Analytics

Table 2: The 51 use cases and Counts of Specific and Generic Requirements

#	Use Case	G is # Generic Requirements		# Specific Requirements						
		Source	**G**	**S**	**T**	**C**	**U**	**P**	**L.**	**O**
	Government Operation									
1	Census 2010 and 2000 - Title 13 Big Data	NARA	8	1		1		1	5	
2	National Archives and Records Administration Accession NARA, Search, Retrieve, Preservation	NARA	10	5	5	2	3	1	4	1
3	Statistical Survey Response Improvement (Adaptive Design)	U.S. Census Bureau	9	1	1	1	1	2	1	1
4	Non-Traditional Data in Statistical Survey Response Improvement (Adaptive Design)	U.S. Census Bureau	4		1	1	1	1	1	
	Commercial									

5	Cloud Eco-System, for Financial Industries (Banking, Securities & Investments, Insurance) transacting business within the United States	Compliance Partners, LLC	1	1	1			1		1
6	Mendeley - An International Network of Research	Mendeley	13	2	3	6	2	1	4	1
7	Netflix Movie Service	Indiana University	21	1	5	6	1	1	1	1
8	Web Search	Indiana University	15	3	2	1	3	2	2	1
9	IaaS (Infrastructure as a Service) Big Data Business Continuity & Disaster Recovery (BC/DR) Within A Cloud Eco-System	Compliance Partners, LLC			2	2		1		
10	Cargo Shipping	MaCT USA	7	1	2	1		1		
11	Materials Data for Manufacturing	R&R Data Services	8	3	1		2	2	1	
12	Simulation driven Materials Genomics	DoE LBNL	14	2	4	8	1	2	2	1
	Defense									
13	Large Scale Geospatial Analysis and Visualization	Data Tactics	7	1	2	1	1	1		
14	Object identification and tracking from Wide Area Large Format Imagery (WALF) Imagery or Full Motion Video (FMV) - Persistent Surveillance	Data Tactics	10	1	1	3	2	1	1	
15	Intelligence Data Processing and Analysis	Data Tactics	11	3	1	3	1	1	1	
	Healthcare and Life Sciences									
16	Electronic Medical Record (EMR) Data	Indiana University	15	5	2	4	1	5	3	1
17	Pathology Imaging/digital pathology	Emory University	15	4	3	4	1	1	1	1
18	Computational Bioimaging	DoE LBNL	11	2	4	4	1	1	1	
19	Genomic Measurements	NIST	15	3	2	3	1	1		1
20	Comparative analysis for metagenomes and genomes	DoE LBNL	12	5	6	1	5	3	3	
21	Individualized Diabetes Management	Indiana University	12	6	6	5	1	2	3	1
22	Statistical Relational Artificial Intelligence for Health Care	Indiana University	11	6	5	5	1	1	1	
23	World Population Scale Epidemiological Study	Virginia Tech	12	3	3	5	1	2	1	
24	Social Contagion Modeling for Planning, Public Health and Disaster Management	Virginia Tech	13	3	3	5	2	2	3	1
25	Biodiversity and LifeWatch	University of Amsterdam	10	6	7	2	3	2	5	
	Deep Learning and Social Media									
26	Large-scale Deep Learning	Stanford University	3			3				
27	Organizing large-scale, unstructured collections of consumer photos	Indiana University	8	1	2	1	1	1		
28	Truthy: Information diffusion research from Twitter Data	Indiana University	13	5	1	4	3	1	1	1
29	Crowd Sourcing in the Humanities as Source for Big and Dynamic Data	Max-Planck-Institute Nijmegen	3		2			1		
30	CINET: Cyberinfrastructure for Network (Graph) Science and Analytics	Virginia Tech	9	2	4	5	1			
31	NIST Information Access Division analytic technology performance measurement, evaluations, and standards	NIST	5	2	1	1	1	1		
	The Ecosystem for Research									
32	DataNet Federation Consortium DFC	UNC Chapel Hill		2	1	2	1	2		
33	The 'Discinnet process', metadata <-> big data global experiment	Discinnet Labs	1	1		1		1	1	
34	Semantic Graph-search on Scientific Chemical and Text-based Data	NIST		1	2	1	1			
35	Light source beamlines	DoE LBNL		2	1	1		1		
	Astronomy and Physics									
36	Catalina Real-Time Transient Survey (CRTS): a digital, panoramic, synoptic sky survey	Caltech	3	1	2		1			
37	DOE Extreme Data from Cosmological Sky Survey and Simulations	DoE Argonne, University of Washington	6	1	1	2	1			
38	Large Survey Data for Cosmology	DoE LBNL	7	1	2	3			1	
39	Particle Physics: Analysis of LHC Large Hadron Collider Data: Discovery of Higgs particle	DoE BNL, FNAL, LBNL, Indiana University	18	3	3	3	1	1	1	
40	Belle II High Energy Physics Experiment	DoE PNNL	7	1		4		1		
	Earth, Environmental and Polar Science									
41	EISCAT 3D incoherent scatter radar system	Cardiff University, EISCAT Science	11	3	3	1	1		1	1

42	ENVRI, Common Operations of Environmental Research Infrastructure	Cardiff University	11	2	1	2	5	1	4	1
43	Radar Data Analysis for CReSIS Remote Sensing of Ice Sheets	Indiana University	19	3	2	3	2	2	1	1
44	UAVSAR Data Processing, Data Product Delivery, and Data Services	NASA JPL	9	2	4	4	1		2	1
45	NASA LARC/GSFC iRODS Federation Testbed	NASA Langley	4	1	1	3	1			
46	MERRA Analytic Services MERRA/AS	NASA GSFC	5	4	1	3	1			2
47	Atmospheric Turbulence - Event Discovery and Predictive Analytics	NASA HQ	8	2	4	2	1		1	
48	Climate Studies using the Community Earth System Model at DOE's NERSC center	NCAR	7	3	1	1	2			1
49	DOE-BER Subsurface Biogeochemistry Scientific Focus Area	DoE LBNL	5	3		1	1			1
50	DOE-BER AmeriFlux and FLUXNET Networks	DoE LBNL	6	5	1	2	1			1
Energy										
51	Consumption forecasting in Smart Grids	University of Southern California	7	2	1	1		1		1

The specific requirements are divided into 7 categories motivated by features in the NIST Big Data PWG reference architecture. The abbreviations used in Table 2 are spelt out in Table 3. This also gives total counts for each requirement type and average counts per use case. Note that summing bottom 7 rows of column 3 implies that the average total number of specific requirements per use case is 10.92

Table 3: Specific and Generic Requirements and their Number

Specific Requirement Category	Abbreviation in Table 1	Count	Average per Use Case
Generic	G	439	8.61
Data Sources	S	121	2.37
Data Transformation	T	113	2.22
Capability Infrastructure	C	128	2.51
Data Consumer	U	62	1.22
Security & Privacy	P	53	1.04
Lifecycle Management	L	57	1.12
Other	O	23	0.45

3 Use Case Requirements

Our process for derivation of requirements extraction involved two steps. The first step is to extract specific requirements based on each application's characteristics, including detailed information on:

- **Data sources** (data size, file formats, rate of growth, at rest or in motion, etc.)
- **Data transformation (**data fusion, analytics)
- **Capability** infrastructure (software tools, platform tools, hardware resources such as storage and networking)
- **Data Consumer** (usage, processed results in text, table, visual, and other formats)
- **Security & Privacy**
- **Lifecycle Management** (curation, conversion, quality check, pre-analytic processing, etc.)
- with a final "**Other**" category which largely corresponds to mobile access requirements.

The second step is to aggregate each application's specific requirements into high-level generalized requirements that are vendor-neutral and technology-agnostic. Note that these use cases and requirements are not exhaustive.

The use case and requirements information is presented online at the following links:

- Index to all use cases [12]
- List of specific requirements versus use case [17]

- List of general requirements versus architecture component [18]
- List of general requirements versus architecture component with record of use cases giving requirements [19]
- List of architecture components and specific requirements plus use case constraining the components [20]

There were 35 generic requirements [18] summarizing 439 specific requirements from the 51 use cases. This implies that on average 12.54 specific requirements were summarized in each generic requirement and this corresponded to an average of 8.61 specific requirements per use case. This was a subset of the 557 total specific requirements or 10.92 per use case. Column 2 of Table 4 gives the number of specific requirements driving this generic requirement.

Table 4: Generic Requirements with Count of Number of Motivating Specific Requirements

#	Count	Generic Requirement
Data Source Requirements (DSR)		
1	28	Needs to support reliable real time, asynchronous, streaming, and batch processing to collect data from centralized, distributed, and cloud data sources, sensors, or instruments.
2	22	Needs to support slow, bursty, and high-throughput data transmission between data sources and computing clusters.
3	28	Needs to support diversified data content ranging from structured and unstructured text, document, graph, web, geospatial, compressed, timed, spatial, multimedia, simulation, and instrumental data.
Transformation Provider Requirements (TPR)		
1	38	Needs to support diversified compute-intensive, analytic processing, and machine learning techniques.
2	7	Needs to support batch and real-time analytic processing.
3	15	Needs to support processing large diversified data content and modeling.
4	6	Needs to support processing data in motion (streaming, fetching new content, tracking, etc.).
Capability Provider Requirements (CPR)		
1	29	Needs to support legacy and advanced software packages (software).
2	17	Needs to support legacy and advanced computing platforms (platform).
3	23	Needs to support legacy and advanced distributed computing clusters, co-processors, input output (I/O) processing (infrastructure).
4	14	Needs to support elastic data transmission (networking).
5	35	Needs to support legacy, large, and advanced distributed data storage (storage).
6	13	Needs to support legacy and advanced executable programming: applications, tools, utilities, and libraries (software).
Data Consumer Requirements (DCR)		
1	4	Needs to support fast searches (~0.1 seconds) from processed data with high relevancy, accuracy, and high recall.
2	16	Needs to support diversified output file formats for visualization, rendering, and reporting.
3	2	Needs to support visual layout for results presentation.
4	11	Needs to support rich user interface for access using browser, visualization tools.
5	20	Needs to support high-resolution multi-dimension layer of data visualization.
6	1	Needs to support streaming results to clients.
Security and Privacy Requirements (SPR)		
1	32	Needs to protect and preserve security and privacy on sensitive data.
2	12	Needs to support multi-level policy-driven, sandbox, access control, authentication on protected data.
Lifecycle Management Requirements (LMR)		
1	20	Needs to support data quality curation including pre-processing, data clustering, classification, reduction, format transformation.
2	2	Needs to support dynamic updates on data, user profiles, and links.
3	6	Needs to support data lifecycle and long-term preservation policy, including data provenance.
4	4	Needs to support data validation.
5	4	Needs to support human annotation for data validation.
6	3	Needs to support prevention of data loss or corruption.
7	1	Needs to support multi-site archival.
8	2	Needs to support persistent identifier and data traceability.
9	1	Needs to support standardizing, aggregating, and normalizing data from disparate sources.
Other Requirements (OR)		
1	6	Needs to support rich user interface from mobile platforms to access processed results.

2	2	Needs to support performance monitoring on analytic processing from mobile platforms.
3	13	Needs to support rich visual content search and rendering from mobile platforms.
4	1	Needs to support mobile device data acquisition.
5	1	Needs to support security across mobile devices.

4 Conclusions

The use cases are exemplars, and there are several areas where additional coverage would be important. The NBD-PWG has produced the current V1.0 collection summarized in this paper to present a coherent description and send information to the other working groups. The current collection of use cases includes the topics of Table 2. The NBD-PWG plans to add categories and use cases to this collection. The recommendations in Section 3 were abstracted from the use cases. These recommendations need more study both within this working group, and with other working groups.

The use cases have already spawned other activities including backdrop for Big Data standards, patterns [7, 8] based on earlier work in parallel computing [21-23], benchmarks and software architectures [10, 11].

References

1. Geoffrey Fox, *Multi-faceted Classification of Big Data Uses and Proposed Architecture Integrating High Performance Computing and the Apache Stack*, in *Sixth International Workshop on Cloud Data Management*. March 31, 2014. Chicago, ILL. http://grids.ucs.indiana.edu/ptliupages/presentations/OgresABDS-CloudDB-Mar31-2014.pptx.
2. Geoffrey Fox. *Big Data Applications and Analytics Online course: Section 5 - Big Data Use Cases Survey: Unit 12 - Overview of NIST Big Data Public Working Group (NBD-PWG) Process and Results*. 2014 [accessed 2014 August 8]; Available from: https://bigdatacoursespring2014.appspot.com/unit?unit=12.
3. Geoffrey Fox. *Big Data Applications and Analytics Online course: Section 5 - Big Data Use Cases Survey: Unit 13 - 51 Big Data Use Cases*. 2014 [accessed 2014 August 8]; Available from: https://bigdatacoursespring2014.appspot.com/unit?unit=13.
4. Geoffrey Fox. *Big Data Applications and Analytics Online course: Section 5 - Big Data Use Cases Survey: Unit 14 - Features of 51 Big Data Use Cases*. 2014 [accessed 2014 August 8]; Available from: https://bigdatacoursespring2014.appspot.com/unit?unit=14.
5. Geoffrey Fox, *Big Data at the Intersection of Clouds and HPC*, in *The 2014 International Conference on High Performance Computing & Simulation (HPCS 2014)*. July 23, 2014. The Savoia Hotel Regency Bologna (Italy). http://grids.ucs.indiana.edu/ptliupages/presentations/HPCS-July23-2014.pptx.
6. Geoffrey C Fox, Judy Qiu, and Shantenu Jha, *Integrating the Apache Big Data Stack with HPC for Big Data*, in *AGU Fall Meeting*, IN025: Leveraging Enabling Technologies and Architectures to enable Data Intensive Science - Abstract of Invited Talk. August 7, 2014. San Francisco, 15-19 December 2014. http://grids.ucs.indiana.edu/ptliupages/publications/AbstractAGUDec2014.pdf.
7. Shantenu Jha, Judy Qiu, Andre Luckow, Pradeep Mantha, and Geoffrey C. Fox, *A Tale of Two Data-Intensive Approaches: Applications, Architectures and Infrastructure*, in *3rd International IEEE Congress on Big Data Application and Experience Track*. June 27- July 2, 2014. Anchorage, Alaska. http://arxiv.org/abs/1403.1528.
8. Geoffrey C.Fox, Shantenu Jha, Judy Qiu, and Andre Luckow, *Towards an Understanding of Facets and Exemplars of Big Data Applications*. July 20, 2014. http://grids.ucs.indiana.edu/ptliupages/publications/OgrePaperv9.pdf.
9. *HPC-ABDS Kaleidoscope of 120 Apache Big Data Stack and HPC Tecnologies*. [accessed 2014 April 8]; Available from: http://hpc-abds.org/kaleidoscope/.
10. Geoffrey Fox, Judy Qiu, and Shantenu Jha, *High Performance High Functionality Big Data Software Stack*, in *Big Data and Extreme-scale Computing (BDEC)*. 2014. Fukuoka, Japan. http://www.exascale.org/bdec/sites/www.exascale.org.bdec/files/whitepapers/fox.pdf.
11. Judy Qiu, Shantenu Jha, Andre Luckow, and Geoffrey C.Fox, *Towards HPC-ABDS: An Initial High-Performance Big Data Stack*. August 8, 2014. http://grids.ucs.indiana.edu/ptliupages/publications/nist-hpc-abds.pdf.
12. NIST, *NIST Big Data Public Working Group (NBD-PWG) Use Cases and Requirements*. 2013. http://bigdatawg.nist.gov/usecases.php
13. NIST. *NBD Requirements WG Use Case V1.0 Full Listing*. 2013 Sept. 25 [accessed 2013 Sept. 30]; Available from: http://bigdatawg.nist.gov/_uploadfiles/M0180_v15_1305888391.docx.
14. NIST. *Big Data Working Group Reports from V1*. 2013 [accessed 2014 March 26]; Report at http://bigdatawg.nist.gov/V1_output_docs.php Available from: http://bigdatawg.nist.gov/home.php.
15. Mary Meeker and Liang Wu (Kleiner Perkins Caufield Byers), *Internet Trends*, in *D11 Conference (All Things Digital)*. May 29, 2013. Rancho Palos Verdes, California. http://www.slideshare.net/kleinerperkins/kpcb-internet-trends-2013.

16. Mary Meeker, *INTERNET TRENDS 2014*, in *CODE CONFERENCE* May 28, 2014. http://www.kpcb.com/internet-trends.
17. NIST. *NIST Big Data Public Working Group (NBD-PWG) List of specific requirements versus use case*. 2013 [accessed 2014 August 8]; Available from: http://bigdatawg.nist.gov/uc_reqs_summary.php.
18. NIST. *NIST Big Data Public Working Group (NBD-PWG) List of general requirements versus architecture component*. 2013 [accessed 2014 August 8]; Available from: http://bigdatawg.nist.gov/uc_reqs_gen.php.
19. NIST. *NIST Big Data Public Working Group (NBD-PWG) List of general requirements versus architecture component with record of use cases giving requirements*. 2013 [accessed 2014 August 8]; Available from: http://bigdatawg.nist.gov/uc_reqs_gen_ref.php
20. NIST. *NIST Big Data Public Working Group (NBD-PWG) List of architecture components and specific requirements plus use case constraining the components*. 2013 [accessed 2014 August 8]; Available from: http://bigdatawg.nist.gov/uc_reqs_gen_detail.php
21. Timothy G. Mattson, Beverly A. Sanders, and Berna L. Massingill, *Patterns for Parallel Programming*. 2013: Addison-Wesley Professional. ISBN:0321940784
22. R. Barrett, M. Berry, T. F. Chan, J. Demmel, J. Donato, J. Dongarra, V. Eijkhout, R. Pozo, C. Romine, and H. Van der Vorst, *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods, 2nd Edition*. 1994, Philadelphia, PA: SIAM. http://www.netlib.org/linalg/html_templates/Templates.html
23. Armbrust, M., A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Konwinski, G. Lee, D.A. Patterson, A. Rabkin, and M. Zaharia, *Above the Clouds: A Berkeley View of Cloud Computing.* 2009. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.149.7163