The Form of High-Performance Computing: A Survey

Priati Assiroj1, H L H S Warnars2, R Kosala3, B Ranti4, S Supangat5, A I Kistijantoro6, E Abdurrahman7

1Binus Graduate Program, Doctor of Computer Science, Bina Nusantara University Jakarta, Information System Dept. Buana Perjuangan University, Karawang
2Computer Science Department, BINUS Graduate Program, Doctor of Computer Science, Bina Nusantara University Jakarta
3Computer Science Department, BINUS Graduate Program, Doctor of Computer Science, Bina Nusantara University Jakarta
4Information System Department, Bina Nusantara University Jakarta
5Information System Technology Department, Institut Teknologi Bandung
6School of Electronics Engineering and Informatics, Institut Teknologi Bandung
7Computer Science Department, BINUS Graduate Program, Doctor of Computer Science, Bina Nusantara University Jakarta

Email: priati@binus.ac.id

Abstract. Computational technology is an important thing that needed in the whole of human life. Along with the development of computer technology, there are a significant data created and wait to be processed; we will get more advantages if it well-processed, if not we will see the data explosion, the data is just a useless stack. Research mitigates the other research paper and makes a conclusion and a report by PRISMA method. The development of computer technology will support the development of an organization. Considering any factors in choosing the right and suitable high-performance computer resource due to our organization is a must. The existing high-performance computing forms are Supercomputer, Grid Computing, Cluster Computing, and Cloud Computing. Based on this work, the last form of high-performance computing, Cloud Computing is one of the most used forms thus we mitigate its advantages and disadvantages. Recently, to process this vast data, we must have a high-performance computer and to provide it, this is not a cheap resource.

Keywords. Technology, Computing, Network, Data, Services.

1. Introduction

The development of computer technology is an important thing that can support organizational growth. Computations are needed at the whole of human life as a sophisticated decision support system that helps people when they decide something. The fact is, as it develops of computer technology, an explosion of the data will occur caused by the growth of the data that very fast. This significant number of data will be an advantage if well-processed. To process large data, we need considerable computational resources, and we need a huge budget to develop this resource. Therefore, consideration is required when we decide to choose computational technology. The existing computational techniques are Cluster Computer, Supercomputer, Grid Computing, Cloud Computing. Then we call it a high-performance computing (HPC) technology. The research domain on this filed is increasing rapidly in Indonesia. As a developing country, Indonesia has many works that involve many stack holders or large volume projects thus high-performance computing is beneficial [1].

Clusters in the computer & network science as shown at figure 1, are a set of independent computer servers working and visible to the network clients as a unit computer. The process to connect multiple computers to work together called Clustering. Computer clusters are the most economical alternative technique for developing high-performance computer systems. Cluster Computer technology is one of the distributed or parallel systems, consisting of computers that are interrelated, work together and merge into a single computer resource [2]. The existing Internet services are provided with large-scale cluster computers thus the biggest supercomputers are made with computer cluster technology [3].
Supercomputers, based on figure 2, are the latest technology based on the large computing systems. Supercomputers make computing performance work well [4] to complete duties that require high computational environment [5]. Supercomputing runs huge projects or applications that cannot be run on common systems [6]. Some computers are distributed in cluster form, and some processors are used [7]. Supercomputers technology started in the 1940s when High Performance Computer was introduced at first. There were radical changes in the 1960s when parallel compilation of computers was first built [8]. Generally, the use of supercomputers extends to many sectors such as Life Science, Medicine, Mechanics, Economics, Electronics, Telecommunication, Computer Science and Engineering and many more. UNESCO classifies the field of knowledge that important in investment that related to the use of supercomputers technology, Engineering [9] [10], Economics and Linguistics [11] and Science Knowledge [12] [13].

In the mid-90s grid computing was introduced [14]. Grid computing is a technology that collects computer resources all over the world. Grid computing coordinates the distributed computing resources to create virtual organizations that utilize idle resources [15]. Research conducted by Chris and Giorgos [16] estimated the economic value obtained from grid computing and evaluated its benefits through web service distribution. The economic value obtained can be significant for industries that use grid computing. The industrial players will be more competitive if they use this technology then the prices can be a critical factor that determines the resources [17]. According to The 451 Group, grid computing resource institutes, one of the crucial success factors for establishing commercial grids is the allocation models and application of resource trading [18]. The research conducted by Shin Yeo and Buyya [19]
focus on how to set reasonable prices that are by resource use. Service providers must have the right way to set competitive prices and provide competitive services. Grid computing as described in figure 3 is to collect computer resources in the world.

Figure 3. Grid Computing

Cloud Computing is combining computer technology with internet-based development. As shown in figure 4, Cloud is a metaphor of web that also describes the hidden complex infrastructure. With cloud computing, we can increase efficiency regarding cost and energy consumption. The High Performance Cloud Computing Model (HPC2) provides a solution to overcome the limitations of existing resources. HPC2 uses the concept of cloud computing and can produce an Infrastructure model as a Service (IaaS) [20,21]. Cloud Computing [22,23,24,25] uses the Internet and virtual technology that is available on demand quickly can be reconfigured and can be accessed anywhere [26]. Paper [27] provides information that various pricing techniques can be done to know the advantages and disadvantages as shown on table 1.

We must consider any factor in choosing the right and suitable high performance computer resource due to our organization. Cloud Computing is one of the most used forms thus we mitigate its advantages and disadvantages. Along with the development of computer technology, there are a significant data created and wait to be processed; we will get more advantages if it well-processed, if not we will see the data explosion, the data is just a useless stack.

Figure 4. Cloud Computing
2. Methodology
This work uses PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-analysis) [28] to create a systematic review protocol. The stages are: Define Eligibility Criteria, Define Information Resource, Literature Selection, Data Collection, and Data Item Selection.

2.1. Define Eligibility Criteria
The eligibility criteria can be determined by Inclusion Criteria (IC) which are:
IC1: the work selects the original research articles in English.
IC2: articles are published between 2016 and 2018.
IC3: articles propose research about high-performance computing cloud computing, cluster computing, grid computing, and supercomputer.

2.2. Define Information Resource
Articles are found in online academic studies database such as IEEE Explore, Google Scholar, and Science Direct. This work chooses the appropriate papers that suitable for IC.

2.3. Literature Selection
The first thing is determining the keywords. The keywords are High-performance Computing, Clustering, Cloud, Grid and Supercomputer. Then choose titles and explore the paper by understanding the abstracts, read the paper in full or partial and determine the feasibility of the report for the next review. We still reassessing and re-evaluating the remaining articles.

2.4. Data Collection
This work collects data manually by creating a data extraction and assesses 443 research articles due to keywords “high performance computing, supercomputer, cluster, grid, and cloud computing”.

| Table 1. Data Collection |
|--------------------------|
| Resource | Found | Candidate | Selected |
| Science Direct | 245 | 22 | 12 |
| IEEEXplore | 14 | 14 | 5 |
| Springer Link | 180 | 8 | 5 |
| Total | 443 | 44 | 22 |

44 articles are eligible to be a reference candidate according to the title and abstract, and 13 articles are selected to do this research.

2.5. Data Item Selection
Data are obtained from articles that have researched about high performance computer, supercomputer, grid, cluster and cloud computing. By detail, this is shown in Table 1. From 22 selected papers then we classify due to the topics of each. The classification must be done to mitigate the form of high-performance computing that explained in every article. The research found two papers that discuss high-performance computing form in collaboratively, Grid computing, Cloud Computing, Cluster computing in [28] and Grid and Cloud computing in [29]. The classification result is shown as figure 5.

| Table 2. High Performance Computing Form |
|------------------------------------------|
| Title | Type | High Performance Computing Form |
| Toward Weather… [30] | Proceeding | Cloud Computing |
| Evaluation of… [31] | Proceeding | Cloud Computing |
| EDISON… [32] | Proceeding | Cloud Computing |
| Cost Optimization… [33] | Journal | Cloud and Grid Computing |
| Toward Data… [34] | Journal | Cloud Computing |
| Brokering in… [35] | Journal | Cloud Computing |
| Managing high… [36] | Journal | Cloud Computing |
Table 2 shows the list of the form of high-performance computer from selected papers. From the 4 proposed form, Grid, Cloud, Cluster Computing and Supercomputers, we found that the most used is high-performance cloud computing.

| Title                                    | Type         | Benefits based on pricing                                                                 | Problems based on pricing                      |
|------------------------------------------|--------------|------------------------------------------------------------------------------------------|------------------------------------------------|
| Broker as a service                      | Conference   | predicts from user’s historical data                                                      | Considering few parameters                     |
| Cloud customer                           | Journal      | by considering user’s historical data and                                                | Unmentioned calculation, underutilization resource|
| Historical…[51]                         |              |                                                                                          |                                                 |

3. Result and Discussion

Figure 5 explains papers that discuss high-performance computer form in Cloud Computing exclusively is 16 papers or 64%. The documents that address Supercomputer is about 5 papers or 20%, articles that discuss Grid Computing is 2 papers or 8% and the last is 2 or 8% papers discuss Cluster Computing.

From the literature study that has been done, the high-performance computing form that most used and has the highest percentage is cloud computing thus this work wants to know the detail about the advantages and disadvantages of this form by reviewing other research. From table 2 we know that cloud computing is the most used form. Then we propose table 3 in which the survey of cloud computing based on its pricing, also the advantages and disadvantages. This result answers why many cloud computing are developed and used.

Table 3. Advantage and disadvantage on cloud computing

| Title                                    | Type       | Benefits based on pricing                                                                 | Problems based on pricing                      |
|------------------------------------------|------------|------------------------------------------------------------------------------------------|------------------------------------------------|
| Broker as a service                      | Conference | predicts from user’s historical data                                                      | Considering few parameters                     |
| Cloud customer                           | Journal    | by considering user’s historical data and                                                | Unmentioned calculation, underutilization resource|
| Historical…[51]                         |            |                                                                                          |                                                 |
A genetic algorithm… [52] Conference provider operations, the
cost can be reduced
High scalability on a
proposed method
Homogenous
environment
The value of cooperation…
[53] Journal Cooperative is better than
competitive brokering
maximize profit by
allocating dynamic
resources
Considering single
parameter user cost
Considering
homogenous resources
Towards an autonomic… [54] Journal
Economic and energy… [55] Journal
A mechanism design… [56] Journal
An economical and… [57] Journal
Double-Sided bidding…[58] Journal
Dynamic cloud instance…
[59] Journal
Online resource scheduling...
[60] Journal
4. Conclusion
The High-Performance Computing existing forms are Cluster Computing, Supercomputer, Grid
Computing, and Cloud Computing. This paper summarizes these forms based on the literature collected
from online academic databases. The result shows among these 4 forms; cloud computing is one of the
most used caused by its price. Cloud Computing is secure to be developed and maintained. Those papers
can be a reference to an organization which want to implement high-performance computer form to
choose the right and suitable way.

References
[1] Hadiana A. (2014). High Performance Computing https://lipi.go.id/berita/single/Komputasi-Berkinerja-Tinggi/7904 (Accessed December 2, 2018)
[2] Buyya R (1999). High Performance Cluster Computing: Architectures and Systems Vol. 1, Prentice Hall
[3] Top500 Supercomputer Lists, http://www.top500.org/lists (Accessed 29-11-2018)
[4] Baun C, Kunze M, Kurze T and Mauch V (2011). High performance computing as a service High Performance Computing: From Grids and Clouds to Exascale IOS Press
[5] Baun C, Kunze M, Nimis J and Tai S (2011). Cloud Computing—Web-Based Dynamic IT Services Springer, Germany
[6] Fox, A., Griffith, R., Joseph, A., Katz, R., Konwinski, A., Lee, G., ... & Stoica, I. (2009). Above the clouds: A berkeley view of cloud computing. Dept. Electrical Eng. and Comput. Sciences, University of California, Berkeley, Rep. UCB/EECS, 28(13), 2009. [7] Buyya R, Yeo C.S, Venugopal S 2008 Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities, in: 10th IEEE International Conference on High Performance Computing and Communications, pp. 5–13
[8] Nurmi D, Wolski R, Grzegorczyk C, Obertelli G, Soman S, Youseff L and Zagorodnov D. (2009). The eucalyptus open-source cloud-computing system in: 9th IEEE/ACM International Symposium on Cluster Computing and the Grid, pp. 124–131.
[9] Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. (2008). A break in the clouds: towards a cloud definition. ACM SIGCOMM Computer Communication Review, 39(1), 50-55.
[10] Mell P and Gumble T. (2011). The NIST definition of cloud computing 2011, http://www.nist.org
[11] Verschelde, J. (2012). Introduction to supercomputing. Computing Memory Coalescing Techniques.

[12] Resch, M. M. (2005). Remarks on Supercomputing in Germany. Praxis der Informationsverarbeitung und Kommunikation, 28(4), 233-237.

[13] Jones, J. P., & Nitzeberg, B. (1999, April). Scheduling for parallel supercomputing: A historical perspective of achievable utilization. In Workshop on Job Scheduling Strategies for Parallel Processing (pp. 1-16). Springer, Berlin, Heidelberg.

[14] Segall, R. S. (2013). Computational Dimensionalities of Global Supercomputing. Journal of Systems, Cybernetics and Informatics, 11(9), 75-86.

[15] Gonzalez A F, Rosilro R, Davila J A M and Olivera V M 2015 *Historical Review and future challenges in supercomputing and Networks of scientific communication* Revista The Journal of Supercomputing.

[16] Smith J 2014 *High Performance Computing and Mathematics* On-line: https://www.math.iastate.edu/pdfs/MM2014.pfd math.iastate.edu

[17] Fabricius U, Freundl C Kostler H and Rude U. (2005). *High Performance Computing education for students in computational Engineering*. In Lecture Notes in Computer Science, vol. 3515. No. II, pp. 27-35

[18] Moses G A and Mariasingam M A. (2006). *The Development of a high quality internet-based degree program in computational science* 5th Global congress on engineering education, 146-150, Brooklyn, New York, USA, 17-21.

[19] Farian H, Anne K M and Haas M. (2008). *Teaching high performance computing in the undergraduate college CS curriculum* Journal of Computing Science in Colleges, 23(3), 135-142.

[20] Gaziza B, Salima N, Guldina K and Elena K. (2012). *Parallel Computing in training of informatics teachers* Procedia-Social and Behavioral Science, 51, 883-887.

[21] Roure D D, Baker M A and Jennings N R. (2003). *The evolution of the grid* in: F. Berman, G. Fox, T. Hey (Eds.), Grid Computing: Making the Global Infrastructure a Reality, John Wiley & Sons Ltd., pp. 65–100.

[22] Neumann D, Stößer J, Weinhardt C and Nimis J. (2008). *A framework for commercial grids—economic and technical challenges* Journal of Grid Computing 6 325–347.

[23] Chris K and Giorgos C (2003). Grid resource commercialization: economic engineering and delivery scenarios IBM Zurich Research Laboratory.

[24] Fellows W and Wallage S (2007). *Grid computing—the state of the market the 451 Group*

[25] Shin Yeo C, Buyya R (2007). *Pricing for utility-driven resource management and allocation in clusters* International Journal of High Performance Computing Applications 21 405–418.

[26] Chauhan S S, Pilli E S, Joshi R C, Singh G and Govil M C 2018 *Brokering in interconnected cloud computing environments: A survey J*. Parallel Distrib. Comput., https://doi.org/10.1016/j.jpdc.2018.08.001.

[27] Aazam M and Huh E N 2014 *Broker as a service (baas) pricing and resource estimation model in: IEEE 6th International Conference on Cloud Computing Technology and Science, CloudCom*, pp. pp. 63–468

[28] Aazam M, Huh E N, St-Hilaire M, Lung C H and Lambadaris I 2016 *Cloud customer historical record based resource pricing* IEEE Trans. Parallel Distrib. Syst. 27 (7) 1929–1940.

[29] Chamorro L, López-Pires F and Barán B 2016 *A genetic algorithm for dynamic cloud application brokerage in: IEEE International Conference on Cloud Engineering, IC2E*, pp. 131–134.

[30] Guan Z and Melodia T 2017 *The value of cooperation: Minimizing user costs in multi-broker mobile cloud computing networks* IEEE Trans. Cloud Comput. PP (99) 1–1.

[31] Mehrotra R, Srivastava S, Banicescu I and Abdelwahed S 2016 *Towards an autonomous performance management approach for a cloud broker environment using a decomposition-coordination based methodology* Future Gener. Comput. Syst. 54 (C) 195–205.

[32] Nir M, Matrawy A, St-Hilaire M 2018 *Economic and energy considerations for resource augmentation in mobile cloud computing* IEEE Trans. Cloud Comput. 6 (1) 99–113.

[33] Prasad A S and Rao S 2014 *A mechanism design approach to resource procurement in cloud computing* IEEE Trans. Comput. 63 (1) 17–30.
[34] Shen H, Liu G and Wang H 2017 An economical and SLO-guaranteed cloud storage service across multiple cloud service providers IEEE Trans. Parallel Distrib. Syst. PP (99) 1–1.
[35] Tang L, He S and Li Q 2017 Double-Sided bidding mechanism for resource sharing in mobile cloud IEEE Trans. Veh. Technol. 66 (2) 1798–1809.
[36] Wang W, Niu D, Liang B and Li B 2015 Dynamic cloud instance acquisition via iaas cloud brokerage IEEE Trans. Parallel Distrib. Syst. 26 (6) 1580–1593.
[37] Zhang R, Wu K, Li M and Wang J 2016 Online resource scheduling under concave pricing for cloud computing IEEE Trans. Parallel Distrib. Syst. 27 (4) 1131–1145.
[38] Zakarya M and Gillam L 2017 Energy Efficient Computing, Clusters, Grids and Clouds: A Taxonomy and Survey, <[CDATA][Sustainable Computing: Informatics and Systems]]>, http://dx.doi.org/10.1016/j.suscom.2017.03.002
[39] Alkhanak E N, Lee S P, Rezaei R and Parizi R M 2015 Cost optimization approaches for scientific workflow scheduling in cloud and grid computing: A review, classifications, and open issues The Journal of Systems & Software, doi:10.1016/j.jss.2015.11.023
[40] Moher D, Liberati A, Tetzlaff J and Altman D G 2009 “The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement,” PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097
[41] Carreno E D, Roloff E and Philippe O A 2016 Towards Weather Forecasting in the Cloud, 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing, doi:10.1109/PDP.2016.80
[42] Salaria S 2017 Evaluation of HPC-Big Data Applications Using Cloud Platforms 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, doi:10.1109/CCGRID.2017.143
[43] Suh Y K 2016 EDISON: A Web-based HPC Simulation Execution Framework for Large-scale Scientific Computing Software 16th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, doi:10.1109/CCGrid.2016.31
[44] Chang V 2017 Toward data analysis for weather cloud computing Knowledge-Based System (2017), http://dx.doi.org/10.1016/j.knosys.2017.03.003
[45] Hou Z, Wang Y, Sui Y, Gu J, Zhao T and Zhou X 2018 Managing high-performance computing applications as an on-demand service on federated clouds Computers and Electrical Engineering, https://doi.org/10.1016/j.compeleceng.2018.02.036
[46] Varghese B and Buyya R 2018 Next generation cloud computing: New trends and research directions Future Generation Computer Systems 79 849–861 http://dx.doi.org/10.1016/j.future.2017.09.020
[47] Wang L, Ma Y, Yan J, Chang V, Zomaya A Y 2016 pipsCloud: High performance cloud computing for remote sensing big data management and processing Future Generation Computer Systems, http://dx.doi.org/10.1016/j.future.2016.06.009
[48] Bittencourt L F, Goldman A, Madeira E R M, da Fonseca N L S and Sakellariou R 2018 Scheduling in distributed systems: A cloud computing perspective Computer Science Review 30 31–54 https://doi.org/10.1016/j.cosrev.2018.08.002
[49] Hassan H A, Mohamed S A and Sheta W M 2016 Scalability and communication performance of HPC on Azure Cloud Egyptian Informatics Journal 17,175–182, http://dx.doi.org/10.1016/j.eij.2015.11.001
[50] Sharkh M A 2016 Building a Cloud on Earth: A Study of Cloud Computing Data Center Simulators, Computer Networks, doi:10.1016/j.comnet.2016.06.037
[51] Atif M, Kobayashi R, Menadue B J, Lin C Y, Sanderson M and Williams A 2016 Breaking HPC Barriers with the 56GbE Cloud 6th International Conference On Advances In Computing & Communications, ICACC, 6-8 September 2016, Cochin, India doi: 10.1016/j.procs.2016.07.174
[52] Shakil K A 2018 Cloud Computing in Bioinformatics and Big Data Analytics: Current Status and Future Research Big Data Analytics, Advances in Intelligent Systems and Computing 654, https://doi.org/10.1007/978-981-10-6620-7_60
[53] Nair P R and Anbuudayasankar S P 2018 Tackling Supply Chain Management Through High-Performance Computing: Opportunities and Challenges, Silicon Photonics & High Performance
Computing, Advances in Intelligent Systems and Computing 718, https://doi.org/10.1007/978-981-10-7656-5_1

[54] Huansong F 2017 High-Performance Key-Value Store On OpenSHMEM, 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing 2017 IEEE DOI 10.1109/CCGRID.2017.49

[55] Markidis S, Peng I B, Iakymchuk R, Laure E, Kestor G and Gioiosa R 2016 A Performance Characterization of Streaming Computing on Supercomputers, The International Conference on Computational Science Volume 80, Pages 98–107

[56] Feng J, Liu G, Zhang J, Zhang Z, Yu J and Zhang Z 2018 Workload Characterization and Evolutionary Analyses of Tianhe-1A Supercomputer, ICCS 2018, LNCS 10860, pp. 578–585, https://doi.org/10.1007/978-3-319-93698-7_44

[57] Kondratyuk N, Snirnov G, Dlinnova E, Biryukov S and Stagailov V 2018 Hybrid Supercomputer Desmos with Torus Angara Interconnect: Efficiency Analysis and Optimization CCIS 910, pp. 77–91, Springer Nature Switzerland AG 2018 https://doi.org/10.1007/978-3-319-99673-8_6

[58] Cheng P, Lu Y, Du Y and Chen Z 2018 Experiences of Converging Big Data Analytics Frameworks with High Performance Computing Systems, LNCS 10776, pp. 90–106, https://doi.org/10.1007/978-3-319-69953-0_6

[59] Heinrich F C, Cornebize T, Degomme A, Legrand A, Carpen-Amarie A, Orgerie A 2017 Predicting the Energy Consumption of MPI Applications at Scale Using a Single Node, IEEE International Conference on Cluster Computing, DOI 10.1109/CLUSTER.2017.66

[60] Bajuadji A A 2012 High Performance Computing. https://slideplayer.info/slide/3056239/ (Accessed December 20, 2018)