Is IBM being Beaten by Cray in Supercomputers or is it Reshaping Supercomputing? -- A Supply Chain view

Andrii Ivanenko, McKinsey & Co, Moscow, andrii.ivanenko@insead.edu

Christian Westlye Larsen, Schlumberger Scandinavia, christian.westlye-larsen@insead.edu

Anne C. Elster, Professor of Computer and Information Science at the Norwegian University of Science and Technology, elster@idi.ntnu.no

Karel Cool, BP Chaired Professor of European Competitiveness, karel.cool@insead.edu

January 15 2015

Resource obsolescence through competition from converging technologies is quickly becoming a more urgent problem for sustaining competitive advantage than resource imitation. Not only can it make a firm's resources obsolete or too costly to be competitive, it may also bring about shifts of power in the supply chain or eco-system that further undermine the competitive position of the targeted firm. This paper analyzes these dynamics and challenges with the situation that IBM is facing in the super computing (or high performance computing) hardware segment and the onslaught of their software / IT business due to cloud providers such as Google and Amazon, and also from much smaller new competitors such as Norwegian Cegal which focuses its cloud services in the oil and gas sector

Keywords: Resource obsolescence, convergence, competitive advantage, sustainability, power shifts, value added, IBM, CRAY, supercomputing, cloud, supply chain, eco-system

Is IBM being Beaten by Cray in Supercomputers or is it Reshaping Supercomputing? -- A Supply Chain view

Andrii Ivanenko, Christian Westlye Larsen, Anne C. Elster and Karel Cool[1]

On April 29th 2014, Cray Inc. won a $70 million bid for supplying the US National Energy Research Scientific Computing Center (NERSC) with their next supercomputer.[2] The computer is likely to become one of the world's fastest computers when it is delivered in 2016. Tenders this large only happen a few times a year worldwide. The NERSC call for bids therefore attracted the entire supercomputing industry, including computing giants such as IBM, HP and Oracle. On June 25th, Cray did it again by winning a $54 million supercomputer contract with the Korean Meteorological Administration (KMA), which will be the computer that prepares future South Korean weather forecasts.[3] July 10 followed with an even more important announcement from Cray: it had been awarded a $174 million contract by the National Nuclear Security Administration to run computations for managing the US nuclear arsenal.[4] On October 29 Cray announced it had won a contract with the British Meteorological Office to build a £97 million supercomputer.

How is it possible that IBM loses such prestigious and important contracts in the supercomputing market to Cray, a company that only has one hundredth of IBM's $6.2 billion annual R&D budget?[5] Is the company that brought computers to the world no longer able to build the fastest machines?

What is a Supercomputer?

Just like a racecar is primarily designed for speed rather than cost, a supercomputer (or 'high performance computer, HPC) is primarily designed for computational performance. The supercomputing market is commonly defined by the machines listed on top500.org of the Top 500 Project. Supercomputers are used in a vast range of science, engineering and business problems that require an extraordinary amount of computations to solve. Example applications include weather forecasting, drug discovery and testing, scientific research in physics and chemistry, vehicle crash collision testing, and data analytics in financial services. In one worldwide study done by IDC, 97% of companies that were using supercomputing said that they could no longer compete or survive without it.[6] Overall supercomputer sales increased from $10 billion in 2007 to an estimated $14 billion in 2016.

In the seventies and eighties, supercomputers were mostly specialized machines, customized and highly distinct from the typical mainframe computers at the time. In 1986, Intel introduced the first supercomputer that was built on commodity processors, but with custom networking and a custom operating system. Starting with the commodity Beowulf cluster[7] at NASA in 1994, clusters of hundreds or even thousands of commodity PCs using commodity networking and system software became a cheap substitute for the low-end custom supercomputers, and eventually overtook the high-end, custom machines. Today, 85% of the 500 fastest supercomputers are commodity clusters, while only 15% are of the traditional custom type, as Figure 1 shows.[8] Most of the custom-designed systems now use individual components (such as the processors) that standard PCs use. Thus, the distinction between the two types has increasingly become blurry.

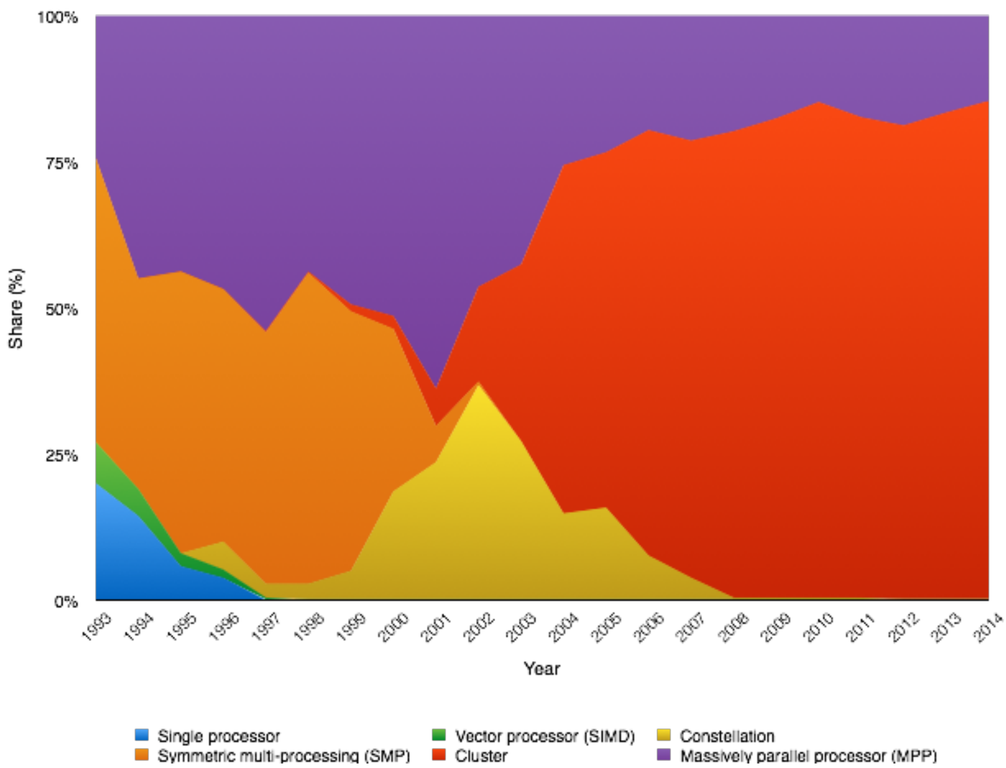

Figure 1. System architecture share of the top 500 supercomputers from 1993 through 2014.

Cray: From David to Goliath – twice

"*I understand that in the laboratory developing the [supercomputer] system, there are only 34 people […], including the janitor. Contrasting this modest effort with our vast development activities, I fail to understand why we have lost our industry leadership position by letting someone else offer the world's most powerful computer*".[9]

Those were the words of IBM President Thomas J. Watson, Jr. in 1965, after he became aware of the massive success of the new, blazingly fast machine from IBM's much smaller competitor, Control Data Corporation (CDC). CDC's machine was ten times faster than anything else on the market and was a huge success, selling in total for over $800 million dollars.[10] Seymour Cray was the chief designer of the new "supercomputer" machine and was a firm believer in the power of small, innovative teams.

By 2006, however, it appeared that Watson's "vast development activities" had paid off. At the end of 2006, 47% of the world's 500 fastest computers were IBM machines.[11] Cray Inc., the company Seymour Cray founded after his success at CDC, had seen its share of the top systems dwindle from over 60% in the 1980s to just 3% in 2006. It had been hurt by several factors, including the founder's focus on technology rather than business leadership and his tragic fatal traffic accident in 1996.

Yet, by 2014 IBM's share of the top 500 list had dropped by fifty percent to only 33% while Cray had made a comeback at IBM's expense, claiming 10% of the top 500 systems. Furthermore, IBM had not yet made an announcement regarding a successor to its popular BlueGene platform, which had powered their lead for the past 15 years.[12] What explained this change of fortune?

The Supercomputer Supply Chain

IBM's surprising defeat -- or retreat -- is partly explained by the fundamental changes that are sweeping through the supercomputing supply chain. While there are thousands of components in a supercomputer, processors (CPUs, accelerators and GPUs) generally account for about 80% of the cost of a system.[13] 95% of the processors used in supercomputers today[14] are of the "x86" family, which is produced by Intel and AMD, with Intel being the dominant player. Smaller processor makers produce the remaining 5%. For example, IBM makes its own POWER series of processors, which are used in some of its highest-end supercomputing systems.

The performance and power usage of a processor is to a large extent constrained by how small its manufacturing process can make the individual transistors that make up the chip. The general trend in recent years has been for processor makers to divest their factories and rely on specialist producers. AMD, for example, spun off its factories in 2009 and is now a "fabless" chipmaker.[15] ARM, whose processor design dominates the smartphone processor market, is a pure IP (Intellectual Property) company that sells its IP to companies such as Freescale, Texas Instruments, Samsung, Toshiba, Atmel, and ST Microelectronics as well as AMD and NVIDIA, which again may outsource their ARM-based designs to external fabrication facilities.[16] Pure-play semiconductor foundries have to a large extent overtaken processor manufacturing, leaving the chipmakers to focus on chip design. The semiconductor makers are mostly Asian players and the most significant is Taiwan Semiconductor Manufacturing Company (TSMC).

The big exception to the rule of owning no factories is Intel, which puts extensive resources into developing a manufacturing process that produces smaller transistors than any other player. This requires them to build new factories with each new process generation, which typically changes every 2-3 years.[17] Also IBM still has its own factories in collaboration with other chipmakers such as Samsung and Toshiba.

Thus, in essence, semiconductor manufacturers (e.g. TSMC) source wafers from wafer manufacturers and produce chips for processor makers (e.g. AMD). These in turn supply to system vendors (e.g. HP, IBM) that do the final assembly of the system. System vendors also design parts such as the cooling and power systems of the machine, and some even design the high-performance storage and connection networks that link one part of the machine to another.

Figure 2 shows the estimated distribution of revenues in the chain among semiconductor manufacturers, processor makers and system vendors in 2003.[18] System vendors captured the largest share (55%) of the $7.2 billion total revenues. The processor makers obtained roughly half that share (27%), while the semiconductor players captured 18% of total chain revenues.

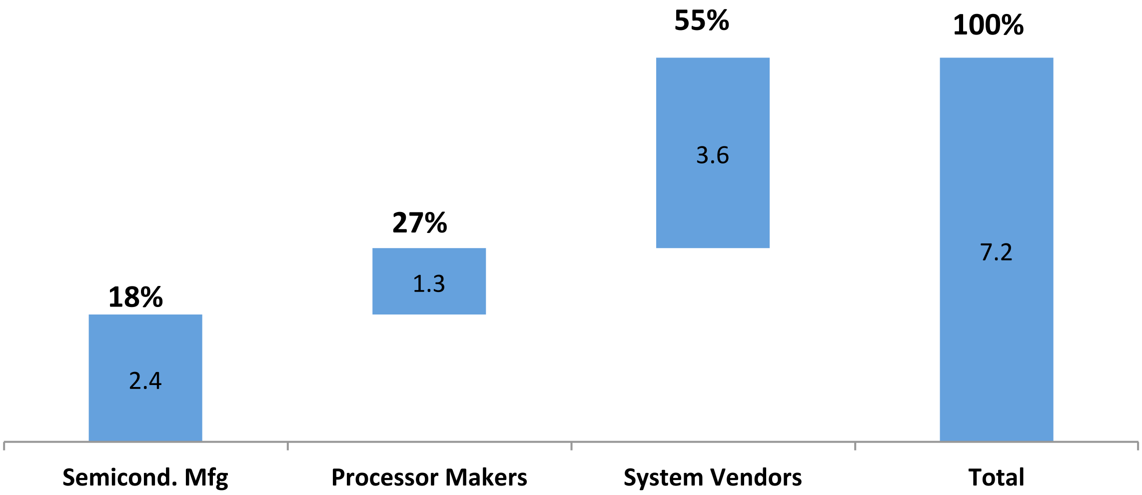

Figure 2. Distribution of revenue among semiconductor manufacturers, processor makers and system vendors in the supercomputer supply chain in 2003, bn. USD

Figure 3 shows the estimated distribution of the $10.3 billion revenues in the chain a decade later. Interestingly, the revenue split had barely changed, showing the system vendors capturing 57%, the processor makers 28%, and the chipmakers 15%.

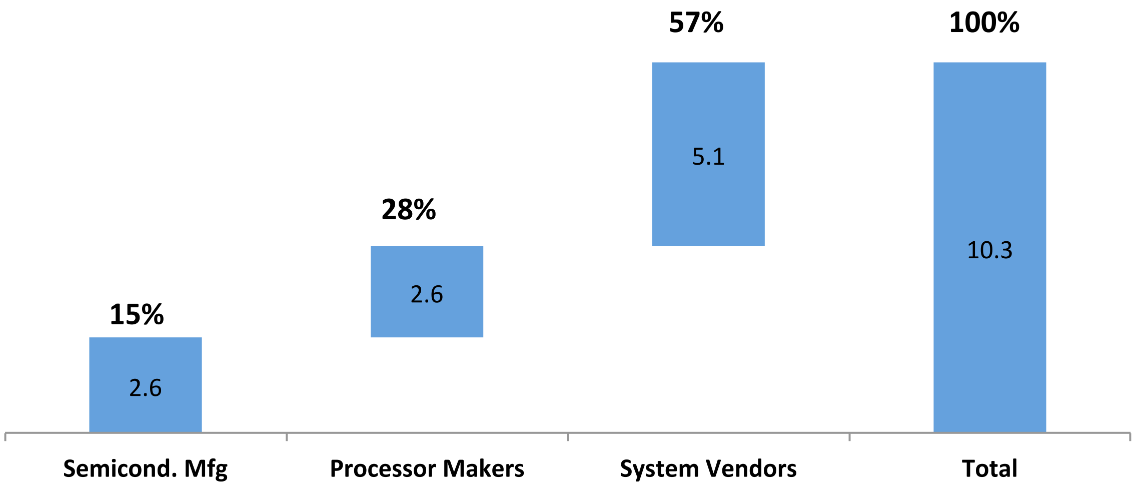

Figure 3. Distribution of revenue among semiconductor manufacturers, processor makers and system vendors in the supercomputer supply chain in 2013,bn. USD

Revenues partly reflect the cost of each activity and the revenue split therefore has to be contrasted to the share of profit in the chain to infer possible shifts in power. Figure 4 provides the estimated EBIT shares for 2003. This shows the system vendors capturing 32% of chain profits, much less than their 55% share of revenues. In contrast, the estimated 2003 share of profits of the processor makers and the semiconductor manufacturers stood at 34% for both, much higher than their respective shares of revenues of 27% and 18%.

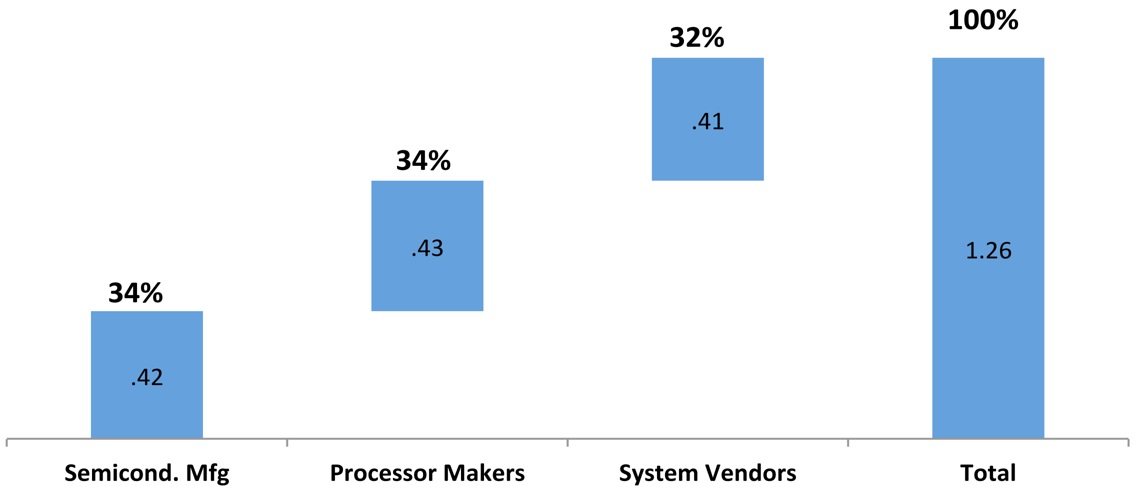

Figure 4. Distribution of EBIT among semiconductor manufacturers, processor makers and system vendors in the supercomputer supply chain, 2003, bn. USD

Figure 5 shows the estimated profit split for 2013. Though the share of revenues for the system vendors had barely changed, their share of profits had more than halved, to 15%. This dramatic drop was much to the benefit of the processor makers that saw their share increase from 34% to 45% of total chain profits, and the semiconductor makers whose share increased from 34% to 40%.

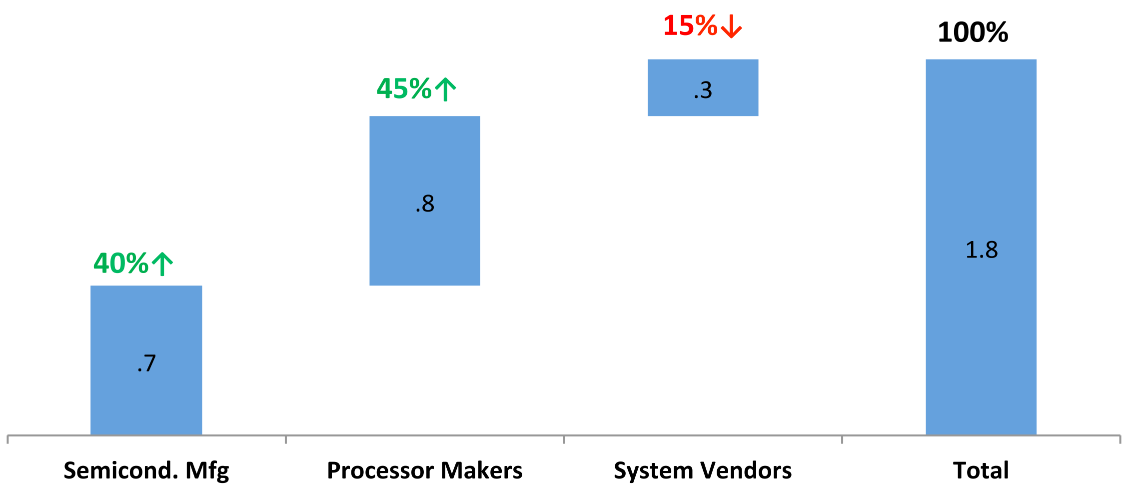

Figure 5. Distribution of EBIT among semiconductor manufacturers, processor makers and system vendors in the supercomputer supply chain, 2013, bn. USD.

The increasing share of total profits for the semiconductor manufacturers, such as TSMC, and processor makers, such as Intel and AMD, shows the rising importance of the processor in the supercomputer – as well as the pricing power of the handful of players. Similarly, the very significant drop in profits for the system vendors reflects their loss of power due to the commoditization of that activity.

Disruption waves in the supercomputer supply chain

Supercomputer vendors such as Cray initially designed their own processors, but by the nineties, they relied increasingly on sourced processors as these had drastically improved their price and performance. Cheap PC processors were thus disrupting the custom-designed high performance processors that were used in supercomputers. When the system vendors stopped making processors, they focused on other proprietary parts such as the network links that connect different parts of the supercomputer (the so-called interconnects). However, this technology too has become increasingly commoditized, with standards such as 10 gigabit Ethernet and InfiniBand replacing proprietary IBM and Cray technology. System vendors have thus been downgraded to assemblers and unless they discover new niche markets or expand into more lucrative parts of the supercomputer supply chain, their future is looking less attractive.

The processor makers on the other hand are enjoying good times. While there is rivalry among the key players (Intel, AMD and NVIDIA), the reality is that there are very few companies that can design high-end processors.

One can therefore expect the processor makers to continue to capture a very substantial part of the profit in the chain. One can also expect the semiconductor manufacturers to grow their share of profits, particularly the top players such as Taiwan Semiconductor Manufacturing Company. Only a few of the manufacturers are sophisticated enough to be on the cutting edge in process technology, which is required for making very high performance processors.

On the horizon, however, is a new disruptive element -- cloud providers. These include major companies such as Amazon, Google and Microsoft, which rent remote access to their supercomputing clouds that are typically built using commodity clusters. This gives anyone the ability to utilize supercomputing at a very low price, and without owning a physical machine. While supercomputers have been demonstrated to still be more cost effective than cloud computing in government-level environments requiring high-end machines,[19] the ongoing price and feature war[20] between the cloud computing providers is making it increasingly difficult to justify owning a supercomputer for more common uses. We expect the cloud computing providers to grow significantly as more and more sites prefer to use cloud services rather than set aside a physical premise for their supercomputing sites which can cost over 150% of the actual machine. In 2014, 23.5% of supercomputer users were using cloud services to complement their internal supercomputers, up from 13.8% in 2011.[21]

How IBM is seeking to reshape supercomputing

As the system vendor business has become less hospitable – and less profitable – IBM may well be better off leaving the hardware market to Cray – that soon will have to deal with competition from lower-cost, system assemblers. In fact, in January of 2014, IBM sold its server business to Asian manufacturer Lenovo[22] and has taken several steps to shift focus to the more lucrative parts of the supercomputing system supply chain. First, it has extended its software platform for managing supercomputing systems to the management of clusters. IBM's management software, developed over decades while designing large-scale computers, used to be an important reason for buying a traditional IBM supercomputer. Clusters have typically been harder to manage since they use software and components built for lower-grade personal computers. However, with a broadening customer base that has less IT knowledge than before, system vendors have been put under pressure to provide comprehensive management software that makes it easier to use the computational power, lowers maintenance costs and keeps supercomputer utilization high. Starting about 2007, IBM made its management software simpler, compatible with any cluster and executed targeted acquisitions to expand its capabilities. For example, in 2012, IBM acquired Platform Computing[23], a company that made an important part of the supercomputer management software that comes with machines from several system vendors. Cray and HP, as large customers of Platform Computing, thus overnight became clients of their biggest competitor. Cray has since been trying to develop its own solutions through acquisitions and internal development, while HP still sells systems with IBM's software. While some alternatives to parts of IBM's offerings exist, competition is limited due to the complexity of developing the software and IBM's massive patent portfolio relating to supercomputer management. The management software also integrates with other IBM software and services, thereby acting as an entry point for upselling.

Second, IBM has shifted from selling supercomputing "machines" to "solutions". While building its services division over the past 20 years, IBM has gained expertise in nearly every conceivable domain that uses computing. The typical company that wants to utilize supercomputing thinks in terms of applications, not computers. Therefore, IBM has used its domain expertise to bundle software, hardware and technical consulting into ready-to-go solution packages for a wide range of supercomputing applications within domains such as biotechnology, financial services and energy. Smaller system vendors lack the knowledge to provide the complete solutions of IBM. For a customer, this means that buying a machine from a smaller vendor requires more IT expertise than buying a complete solution from IBM – and typically, lack of expertise is a big hurdle for adopting supercomputing.

Finally, IBM is trying to make sure that as its customers move to the cloud, they stick with IBM. A large gap in the offer of the major cloud providers is the lack of easily accessible domain-specific solutions. In June 2013, IBM acquired SoftLayer[24], a cloud computing provider. Since then, IBM has integrated cloud computing into its solutions packages. For example, in choosing a solution, a customer can (with IBM's help) choose to own its own machine, go for a hybrid solution that uses both a customer-owned machine and IBM's cloud, or exclusively utilize the cloud. IBM's cloud also provides high-speed interconnects and increased security features, which are typically important to supercomputing customers but lacking from the bigger cloud providers. By making sure that IBM is the first cloud provider a transitioning supercomputer customer uses, the customer can also be locked into IBM's cloud ecosystem.

Will IBM succeed in reshaping itself?

IBM's feeble stock market performance in 2013 and 2014 has led to scorching critique in the business press.[25] IBM's business performance indeed has sent conflicting signals to investors and some are becoming impatient with the seemingly slow response to the structural competitive and supply chain challenges it is facing. The changing fortunes and strategy of IBM in supercomputing demonstrate the significance of the challenge and the need to develop a new business model rather than tweak the strategy that worked well for them in the previous decade.

Disruption leading to resource obsolescence is one of the toughest challenges that a corporation can face as it undermines the very foundation of its long term-success. Examples of companies that could not overcome these challenges abound: Nokia, Research in Motion, Motorola, Kodak, DEC. Any change in business model entails the significant risks of leaving behind established competences and of being too slow or unsuccessful in building the new competences. However, while computing power has become a commodity, significant IT expertise is still required to provide solutions to very complex problems and IBM is making a credible commitment to reconfigure its resource base to be competitive in the emerging (super)computing environment. Time will tell whether they can do it fast enough to "be a leader out the other end, as we have been every other time", as CEO Rometty states.[26]. While IBM can bank on its very extensive experience in dealing with similar challenges in the past, ecosystem and supply chain dynamics are unique each time, and IBM will also need to contend with leaders such as Google, Amazon and smaller, new cloud computing companies such as Norwegian Cegal in the oil and gas industry, which are perhaps more nimble and eager than IBM is.

---

[1] We would like to thank Ken Keverian, Senior Vice President Corporate Strategy, IBM for his comments on a previous version of the paper.

[2] Cray Inc. Cray Q1-2014 Conference Call transcripts. http://seekingalpha.com/article/2175433-crays-ceo-discusses-q1-2014-results-earnings-call-transcript

[3] Knapp, A. Cray Signs A $54 Million Supercomputing Contract With South Korea. Forbes, June 25th, 2014. http://www.forbes.com/sites/alexknapp/2014/06/25/cray-signs-a-54-million-supercomputing-contract-with-south-korea/

[4] Cray Inc. Cray Awarded $174 Million Supercomputer Contract From the National Nuclear Security Administration. July 10th, 2014.

[5] IBM Corp. IBM 2013 Consolidated Financials. http://www.ibm.com/annualreport/2013/financial-highlights.html

[6] Joseph, E. (IDC). HPC Trends, Big Data and The Emerging Market For High Performance Data Analysis. Salishan, April, 2013.

[7] http://www.beowulf.org/overview/history.html; Thomas L. Sterling, Daniel Savarese, Donald J. Becker, John E. Dorband, Udaya A. Ranawake, Charles V. Packer: BEOWULF: A Parallel Workstation for Scientific Computation. ICPP (1) 1995: 11-14; Thomas Sterling, Ewing Lusk, and William Gropp (Eds.).:*Beowulf Cluster Computing with Linux* (2 ed.). 2003, MIT Press, Cambridge, MA, USA. See also Anne C. Elster, "High-Performance Computing: Past, Present and Future", PARA 2002, Espoo, Finland, June 15-18, 2002. Published in J. Fagerholm et al. (Eds.): PARA 2002, LNCS 2367, pp 433-444, 2002, Springer Verlag., and Anne C. Elster, Otto J. Anshus, Amund Tveit and Cyril Banino, "Recent Trends in Cluster Computing", presented in Mini Symposium "Cluster Computing", ParCo 2003, Dresden , Germany, September 2-5, 2003.

[8] Top 500 Project. Statistics over time. http://www.top500.org/statistics/overtime/

[9] Henderson, H. A to Z of Computer Scientists. Facts on File, 2002.

[10] Murray, C. The Supermen: The Story of Seymour Cray and the Technical Wizards Behind the Supercomputer. 1st.ed. Wiley, 1997.

[11] Top 500 Project. Development over time. http://top500.org/statistics/overtime/#.U5CQmxZRpBg

[12] Cray Inc. Cray Q1-2014 Conference Call transcripts. http://seekingalpha.com/article/2175433-crays-ceo-discusses-q1-2014-results-earnings-call-transcript

[13] Conway, S (IDC). Private communication. April, 2014

[14] Conway, S. (IDC). IDC Market Updates And Predictions For 2014. 52nd HPC User Forum Meeting, April, 2014.

[15] Clarke, P. AMD jumps into fabless chip company ranking. EETimes, January 19th, 2010. http://www.eetimes.com/document.asp?doc_id=1172794
[16] http://arm.com/products/processors/licensees.php
[17] LaPedus, M. Update: Intel to build fab for 14nm. EETimes, February 18th, 2011. http://www.eetimes.com/document.asp?doc_id=1258701
[18] Estimates are based on segment analysis for TSMS, ASI, UMC, SMIC and PowerChip (semiconductor manufacturers), AMD, NVIDIA and Intel (processor makers), and Cray, HP, Dell, SGI, Fujitsu, Bull and IBM (system vendors).
[19] Yelick, K., Coghlan, S., Draney, B., Canon, R.S. The Magellan Report on Cloud Computing for Science. US Department of Energy, Office of Advanced Scientific Computing Research (ASCR), 2011. http://science.energy.gov/~/media/ascr/pdf/program- documents/docs/Magellan_final_report.pdf
[20] Assay, M. Amazon Web Services leads war on Cloud Price Reductions. TechRepublic, 2014. http://www.techrepublic.com/article/amazon-web-services-lead-the-war-on-cloud- price-reductions/#.
[21] Conway, S. (IDC). IDC Market Updates And Predictions For 2014. 52nd HPC User Forum Meeting, April, 2014.
[22] IBM Corp. Lenovo Plans to Acquire IBM's x86 Server Business. Press release, January 23rd, 2014. https://www-03.ibm.com/press/us/en/pressrelease/43016.wss
[23] IBM Corp. IBM Closes on Acquisition of Platform Computing. January 9th, 2012.
[24] IBM Corp. IBM to Acquire SoftLayer to Accelerate Adoption of Cloud Computing in the Enterprise. June 4th, 2013
[25] Bloomberg, "Five Charts Show Why IBM is the Worst Dow Performer for Second Year", December 30 2014. http://www.bloomberg.com/news/2014-12-30/five-charts-show-why-ibm-is-worst-dow-performer-for-second-year.html; R. Waters and M. Dickson, "FT Interview: Big blues", Financial Times, 14 May 2014.
[26] R. Waters and M. Dickson, "FT Interview: Big blues", Financial Times, 14 May 2014