FINALE RESEARCH REPORT:

SUPERCOMPUTERS

Prepared by

Caleb Fontaine - 100832588, Sebastian Carneiro - 100798391, Haekang Song - 100625189

Prepared for

Bill Kapralos

April 16th 2022

Abstract:

Compared to the long life of normal computers, supercomputers have not been around as long and the majority of the world knows very little about supercomputers compared to their knowledge of everyday computers. However this does not mean that supercomputers did not have a great effect throughout their lifetime and in the future. Supercomputers work differently than normal computers, they are used to make big calculations that normal computers couldn't handle. Then these calculations are used for multiple different aspects in our life (space, math, technology, etc.) This report covers multiple resources and other sources of information to give us a better understanding of supercomputers in terms of what they are, what they do, their history, etc. The findings in this report are that supercomputers are used to help the progression of technology to improve our overall lives.

Introduction:

The topic of the report is on supercomputers. Supercomputers are man made machines that are able to perform gargantuan calculations at very high speeds without the need of human input. They are used to make most of the bigger calculations that we need for our developpement, whether that may be for space travel/exploration, nuclear weapons and reactors, the climate, weather, etc. The difference between a normal computer and a supercomputer is that supercomputers are high level performance machines compared to a general-purpose machine that is a computer. Supercomputers calculate their performance in terms of FLOPS (floating-point operations per second) while normal computers use MIPS (million instructions per seconds). While the first supercomputer was built in 1964, by the United States, the development and the use of supercomputers expanded significantly around the 1970s due to nuclear testing and normal computers didn't have the power or capacity to run the calculations. With the advance of time and with the development of technology, supercomputers have advanced way beyond what they started off as. Even with all these advancements, supercomputers still have problems that are ligering till this day. One of these problems was the sheer amount of power that is needed to be able to run a supercomputer. With all of its components, supercomputers need a lot of energy (electricity) to be able to run properly and since they need to be active 24 hours a day everyday, you can imagine how much energy they can consume. Another recurring problem about supercomputers is the space they need. Supercomputers are very big machines that take a lot of space, they can cover a whole floor in a hotel with the amount of machine put together and everything else with it to make them able to run (power, etc.). With these problems comes the cost of manufacturing and building supercomputers. Since supercomputers are made with multiple little circuit boards and with a lot of wiring, the cost of these machines is extremely high and this is without going into details about everything else they need for these machines that also costs a lot.

For this report, we have made multiple searches with multiple different criterias to find more specific information about supercomputers. The search for supercomputers was done around the beginning of March and ended towards the beginning of April. We used multiple search engines such as Google, OTU Campus Library, etc. When conducting our research, when looking for information about the history of supercomputers we looked for websites or articles that contained information that was dated for minimum 1960 (around the years where the first supercomputer was built) and when looking for future trends for the supercomputers, we looked for websites or articles that contained information that was dated for as recent as possible (2021-2022) to get the most recent information on the topic. As for the core of our research, we were a bit biased toward information/research that was completed recently compared to information that is a bit more dated, but we still took a look at the information of the older documents to see if there was any new information that we could use. For the search terms that we used, we simply used the term 'supercomputers' for our general searches, but for our more specific search such as the history of supercomputers and the future of supercomputers we had more detailed searches. For the history of supercomers we searched up the name of the first supercomputer that was invented, 'CDC6600 supercomputer' and for the search for the future of supercomputers we simply put 'future of supercomputers'. As an added step in our information search, we tried to only take our information from sites that seemed reliable. Sites that had their own references, sites that are more well known for their information (i.e. IEEE, IBM, ACM, etc.).

SUPERCOMPUTERS

What is a super computer?

A 'supercomputer' is a computer built to perform large-scale operations at high speed. While today's computers are getting closer to internet terminals, game consoles, and office machines. They are being used everywhere by applying their decent computing power, but a supercomputer is a computer that is made only for real calculations.

In the past, supercomputers used a dedicated CPU with a special architecture and communication mechanism specialized for computation. However, there is a limit to integrating performance into one CPU. Thus, since 2005, the performance of a single CPU core has reached a plateau. There was a limit with just one CPU, so the industry turned to 'parallel processing' with the idea that they can use multiple CPUs. With the simple idea that the higher the number of cores, the better the computational power. In order to make a better supercomputer, the number of cores has been developed rapidly. As a result, supercomputers now reach more than a million cores.

Figure1:
Fugaku Supercomputer

For example, Fugaku of Japan, which ranked 1st in the supercomputer ranking in 2022, has 7.63 million CPU cores. However, as the number of cores increased, problems began to arise such as electricity cost and heat control. For example, Fugaku supercomputer power consumption is 2,989,923kW. This is almost the same as a small town's yearly power consumption. Moreover, to reduce the heating problem, the air conditioners, which also use a lot of electricity, need to run all day to prevent the CPU cores from overheating. Therefore, scientists started to use low voltage processors to reduce the electricity cost. So, supercomputers also need to be managed by their energy consumption to have a good balance of performance and cost (cost-effectiveness). Moreover, the engineers thought that if only CPUs had a limit for their performance, it would not be enough to satisfy the balance of cost and performance. So they paid attention to GPGPU (General-Purpose Graphics Processing Unit) for their calculations. The calculations were performed by computational processors such as GPUs, and the CPU was controlling and managing these computations. For example, AlphaGo, which became popular in 2016, used the parallel processing of the GPU structure to utilize the maximized computational ability of the GPUs. However, since non-CPU have difficulties with programming, their application fields are limited to machine learning.

Nowadays supercomputers have thousands of nodes associated with them and each node is connected through a high-speed communication network. That connection is very important because the data communication speed between nodes has a huge influence on the overall computational performance of the supercomputers. When testing the performance of a supercomputer (with the LINPACK benchmark), there will be two types of performance recorded: The theoretical maximum performance (Rpeak) and the actual performance (Rmax). They are separated because the performance of all cores cannot be fully exhibited due to communication problems. For example, Fugaku ,1st ranked supercomputer, got 537.212Pflop/s in Rpeak but got only 442.010Pflop/s in Rmax. Most modern supercomputers have a computing power of petaflops. Peta is 10^15 and this means Petaflops can perform 10^15th floating-point calculations per second. That is, 1000 trillion calculations per second. So, a Fugaku's Rmax is 442.010Pflop and this means it has a capacity of 442,010 trillion calculations per second.

In addition, in 1993 there were several different types of OS in supercomputers such as Windows, Unix and Cray OS. But now, most supercomputers use Linux.

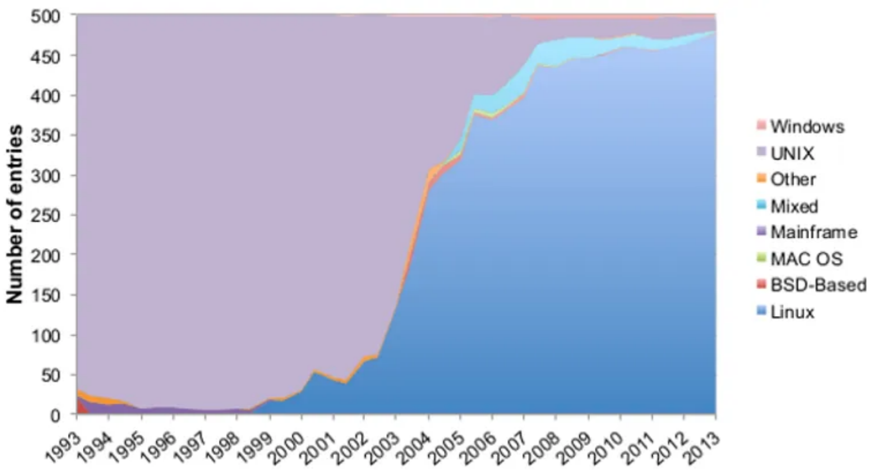


Figure2: Growth of Linux on Supercomputers

Supercomputer history

The first supercomputer was CDC6600 which was released by the CDC of the United States in 1964. The calculation speed of the CDC 6600 was around 3MFlop/s.

Figure3: Image of the first supercomputer 'CDC6600'

The command processing structure back then, was similar to the current RISC (Reduced Instruction Set Computer) method with processor parallelization and wiring optimization. More than 100 CDC6600 were sold at that time, and more than half of them were used for nuclear testing. Considering the fact that many supercomputers are still used intensively in special fields that require high computational power, such as weapons development, space industries and weather forecasting, it was obvious that they used supercomputers for nuclear tests during the cold war.

There were a lot of companies that started to develop supercomputers. However, the long depression damaged the global economy in the mid-1990s, thus, the market shrank. Several manufacturers could not afford the huge cost of developing supercomputers and went bankrupt. Therefore, there are only a few brands that are still developing supercomputers. The major brands are IBM, Fujitsu and Cray. These companies are the ones developing, and incorporating supercomputers as leaders in the industry. Of course there are other companies in the industry, but these companies are the most well known as of now and have made major impacts over not only the entire history of computers, but including supercomputers, then and now.

First let's take a look at IBM, where they came from and what they have done. IBM or International Business Machines Corporation was first founded in 1911 as the Computing-Tabulating-Recording Company and then changed names to IBM later on in 1924. But to begin, IBM started as a merge of three companies, and sold an assortment of things like "deli slicers and commercial scales, to tabulators and punch cards" (Potter, 2012). But since this IBM has come a long way to where they are today. Over the years since, IBM has had a pretty big focus and funding for supercomputers. This all really started

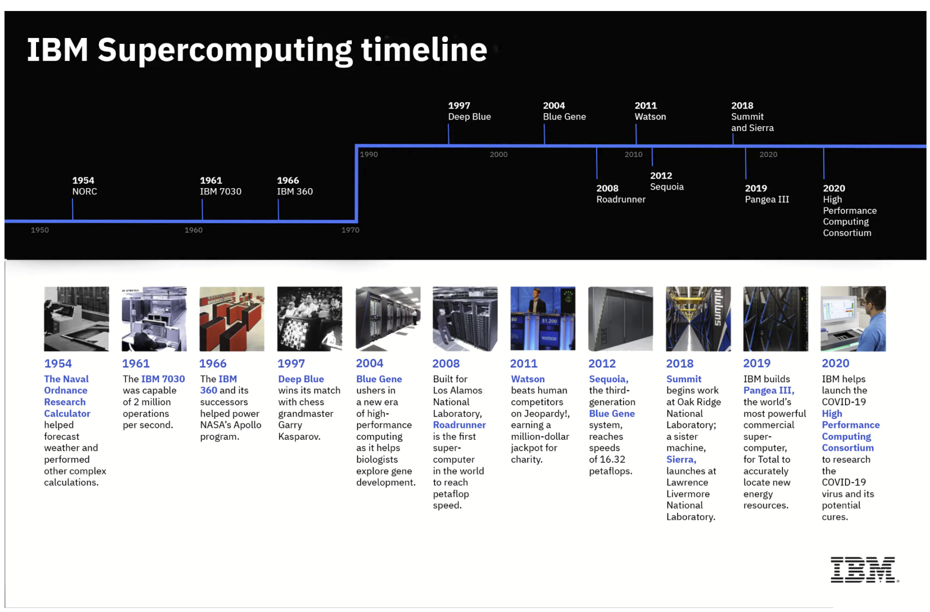


Figure 4: IBM Timeline

when ‘The Colossus’, “the first functional, electronic, digital computer” (IBM, *what is supercomputing?*) was invented, along with the fact that the US military was putting lots of funding into the research and creation of very advanced technology for military use was the push into the ‘world of supercomputing’. And as IBM says, “The term supercomputer came into use in the early 1960s, when IBM rolled out the IBM 7030 Stretch, and Sperry Rand unveiled the UNIVAC LARC” (IBM, *what is supercomputing?*). These two machines were purposefully created to be faster and ‘more powerful’ than the commercial machines at the time, thus the term ‘supercomputer’. The IBM 7030 at the time had the best data processing system which “was capable of 2 million operations per second” (IBM, *supercomputing* timeline) and the was best on the market, paving the way for future development. For a look at IBMs supercomputer history look at figure 4.The IBM 7030 was used at the time for scientific research and military use. Compared to today IBMs supercomputer has come a long way. After the 7030 there were lots of new developments.

IBM supercomputers went from being able to do large calculations to being able to beat a chess grandmaster, win Jeopardy, to being used today to help fight against Covid-19. IBM also now holds the second and third place for the current fastest supercomputers. Second being Summit at Oak Ridge National Laboratory (ORNL) used for “neutron science, high-performance computing, advanced materials, biology and environmental science, nuclear science and engineering, isotopes, and national security research to benefit science and society” (ORNL, *Solving the big problems: ORNL*). The third-place supercomputer, Sierra, is used at Lawrence Livermore National Laboratory (LLNL) to “assess the performance of nuclear weapon systems as well as nuclear weapon science and engineering calculations” (LLNL, *Sierra*), along with “detection of nuclear non-proliferation, and rapid development of cancer therapeutics” (LLNL, *Sierra-Fact-Sheet*). The current uses of IBMs supercomputers haven’t changed much over time as they are still primarily used to do massive calculations and help in the scientific fields. Despite being extremely fast supercomputers they are still not the fastest in the world.

As mentioned above, the fastest supercomputer on the planet belongs to Fujitsu with their supercomputer Fugaku. Fugaku was completed on March 9th, 2021 and has since been put to use to help with current issues like Covid-19. But how has Fujitsu come to this point in development? Fujitsu was founded in 1935 and like IBM initially started as a "joint venture between Furukawa Electric Co., Ltd. And Siemens Ag" (Fujitsu, *History of Fujitsu*) back in 1923 to push forward the production of generators and electric motors in Japan. From there Fujitsu moved to "importing and selling telephones and "step-by-step" automatic switching systems made by Siemens AG" (Fujitsu, *History of Fujitsu*) to eventually creating and selling their own. It wasn't until the 1950s where Fujitsu started in the computer business. Fujitsu wanted to create a computer using relays despite most computers at the time using vacuum tubes, and in 1954 the FACOM100 did just that. But it wasn't until 1977 when Fujitsu had developed their own supercomputer, the FACOM230-75 APU. This supercomputer had "an APU (Array Processing Unit) … structures with asymmetrical multi-processors… and could efficiently describe vector processing." (Fujitsu, *Product Milestones*) and was first implemented at the National Aerospace Laboratory of Japan. Since then They have continued to make and develop their technology and come up with faster and faster supercomputers. For instance in 1992 Fujitsu had the VPP500 which used a "vector parallel system" (Fujitsu, *Product Milestones*), and at the time was the fastest computer in the world. Fujitsu has taken this title multiple times. The next being with the "K computer" and most recently their Fugaku supercomputer, which still reigns at the top for now. The main use for Fugaku as of now has been to try and predict and understand what will happen with the current pandemic. It is also used to "contribute to the development of science and technology and the realization of a safe and secure society" (Fujitsu, 2021), and to help predict weather and climate effects like tsunamis.

Figure 5: Seymour Cray

Another company in the supercomputer industry today is Cray. Cray is a company named after its creator Seymour Cray, but Seymour Cray didn't just start working on supercomputers here, he already had. Cray first started work at Engineering Research Associates, from there he left to help create Control Data Corporation (CDC). It was here where Cray made what many consider to be the 'first' supercomputer. In 1964 Cray and the CDC has completed the CDC 6600, the fastest computer at the time capable of 3 million FLOPS. The CDC 6600 remained at the top, "it wasn't surpassed until the CDC 7600 in 1969 – another Cray design." (Computer History Museum, *CDC 6600's Five Year Reign*). But although making good progress Seymour Cray wanted to work for a less bureaucratic company and left the CDC to for Cray Research Inc., where they would be "the world leader in supercomputing for more than two decades." (Computer History Museum, *Cray Research, Inc.*). At his new company the first supercomputer they released was the Cray-1 which came out in 19756, and like many of the supercomputers I've talked about it was also the fastest when released. Cray Research continues to pump out supercomputers and remains very prominent in the processing world with many of their supercomputers taking the top spot for being the fastest. Now, Cray Research is now a subsidiary of Hewlett-Packard (HP) but still continues in this field. With their most recent supercomputer being the HPE Cray EX which is 6 billion times faster than the Cray-1. HP and Cray like the rest are now pushing towards consistently achieving the exaflop barrier or a quintillion floating-point operations per second. Being able to compute at these speeds could help in research drastically.

Supercomputers have mainly used vector-type processors. The vector processor is composed of an array of data structures and has the characteristic of being able to operate multiple data simultaneously with one instruction. In addition, it shows great performance in the iterative and arithmetic operation area which was suitable for supercomputing. However, in terms of economics, the price is very high compared to the scalar type processor. A scalar processor is a simple processor that can process only one data at a time, but it is cheap and very versatile, so most PCs use the scalar processor. Gradually, the vector-type processor began to disappear from the supercomputer market where parallel computing has become mainstream and the vector-type processor rose to glory.

To understand the architecture of modern supercomputers, we must know about parallel computing. Parallel computing is dividing large or complex problems into smaller ones and processing them at the same time. It’s the same as having a problem that would take an hour with one computer but it could be solved in 10 minutes by dividing the work into six computers. For this reason, it is common for high-performance supercomputers to use more cores to split the work and make it faster. However, as the number of cores increases, the processing speed does not increase infinitely. It states that once hardware parallelism is achieved above a certain level, the level of parallelism in the software will affect the speed of the process to a certain point. In the supercomputer architecture, supercomputers consist of several nodes. Each node is a computer with an independent processor and memory. These nodes are racked together with network switches that connect the racks to each other for all the calculations, and it's also attached to a storage node for storing calculation results. In the past, a symmetrical multi processing (SMP) method was used in supercomputers which involved multiple processors in a node sharing a single memory. Nowadays, the massively parallel processing (MPP) method is used where an individual memory is allocated to each processor. Each node completes the assigned command and sends the result through the network switch to the IO Node (input/output node) which is in the same rack. The collected calculation results are stored in a storage node through the network switch that connects the racks. A supercomputer that connects a large number of nodes can achieve great performance, but it requires a large amount of space and has high power consumption.

Nowadays, most supercomputers are used in the science and technology industry. There are weather forecasting, large-scale simulations of space astronomy, polymer characterization, and physics calculations that use supercomputers for their calculations. Moreover, global tech-companies such as Google, Apple, and Amazon are also operating supercomputers and large-scale data centers (IDCs) to provide services that require large-scale computing resources. In addition, individual customers can experience the supercomputer performance by borrowing computing resources through large-scale B2C cloud services from Google, Amazon, and Microsoft. Of course, cloud computing performance cannot reach the real supercomputer level of speed, but it has pretty decent performance compared to the regular computers. Although this is B2C services, Individuals do not use it because they do not need supercomputer-level computing power. Therefore, Research institutes and industries often use cloud supercomputing for cost efficiency.

Supercomputer Architecture

Supercomputers have been around for a long time and are still being developed to this day. But what separates them from each other, then and now? We will take a look at the differences between supercomputers within their own company and to the likes of others regarding their architecture.

Before we start it is key to understand what computer architecture is. Computer architecture is “a specification describing how hardware and software technologies interact to create a computer platform or system” (*What is computer architecture?* 2018). Along with this computer architecture has 3 components, System Design, Instruction Set Architecture, and Microarchitecture. System design is the hardware components that go into making a computer, with things like the Central Processing Unit (CPU) and Graphics Processing Unit (GPU), memory controllers, data processor, so on and so forth. Next the Instruction Set Architecture or ISA, this is the “functions and capabilities of the CPU. It is the embedded programming language and defines what programming it can perform or process.” (*What is computer architecture?* 2021, Sunderland), and includes things like Operating System (OS) along with data formats, addressing modes and more. Lastly, Microarchitecture is the computer ‘organization’, and explains data processing and storage elements and how it can be incorporated into ISA. These are the fundamentals to computer architecture. Now that we understand what computer architecture is, we can take a look at the architecture used in supercomputers.

IBM has been making supercomputers for arguably the longest time as the IBM 7030 Stretch came out in 1961. Let’s take a look as to how it worked. First, we’ll go through the System design (hardware) components that were used to bring the 7030 together. The 7030 had a type 7101 CPU with a 64-bit processor, which “contained approximately 169, 000 transistors on SMS cards, and had a power consumption of 21.6KW.” (*IBM stretch,* 2009). The 7030 was able to do 1.2 million instructions per second. At the time there were no GPUs. but they did use ‘high-speed’ magnetic core memory which allowed for data to be retrieved in 2.18 millionths of a second, and the 7030 had 2048 kilobytes of memory total. For Instruction Set Architecture, the 7030 used Master Control Program (MCP) which is now a term for OS, there was even a FORTRAN compiler created to work with the 7030, and there was also the STRETCH Assembly Program (STRAP). The stretch was capable of handling either 32-bit or 64-bit instructions. The Stretch could also use ‘four levels of lookahead’. Lookahead would be like “a prefetch and store buffer” (Smotherman, 2010), this means that the computer could ‘fetch’ instructions from their storage and move it to a faster memory location. In the 7030 Stretch the instructions followed a process. The instructions would go “through 2 processing elements: an indexing and instruction unit that fetched, predecoded, and partially executed the instruction stream, and an arithmetic unit that executed the remainder of the instructions.” (Smotherman, 2010). Of course there was more to the IBM 7030, even the registers were split in a way that the indexing and instruction unit had a set of registers, and the rest for the arithmetic section. “The indexing and instruction unit directly executed indexing instructions, and prepared arithmetic instructions by calculating addresses and starting memory fetches.” (Smotherman, 2010). The 7030 also was able to use parallelism, meaning it could ‘decode instructions in parallel with execution’. Once instructions were either ‘fully or partially executed’ they were moved into a “form of buffering called a “lookahead " " (Smotherman, 2010). This would allow for the fully executed indexing instructions to go into one of the four levels of lookahead with multiple pieces of data allowing for the system to ‘rollback’ the information in case of an error when ‘looking ahead’. For partially executed arithmetic instructions would also be placed into a lookahead section and “would wait for the completion of its memory operand fetch” (Smotherman, 2010). For Microarchitecture the 7030 “contained many aggressive organization techniques… predecoding, memory operand prefetch, out-of-order prediction, branch misprediction, and precise interrupts”. (Hardy, *The IBM 7030 (STRETCH)*) to take a look at one of them out-of-order prediction is when a computer executes an instruction that is not ‘next in line’ so long as it does not need to wait for the previous instruction to be completed, allowing for simultaneous completion of instructions and improved speed. To relate this to the real world, think about when you are doing house chores. When you are doing laundry it takes a while, first the clothes go into the wash, then into the dryer. But in between the two is maybe 40 minutes. This means that you cannot proceed with the dryer until the wash cycle is complete. But in the meantime, you can do any other chores, vacuuming, dusting, etc., and once the wash is done you can switch it over to the dryer. This is what is being done when out-of-order processing is occurring. There are many other things that relate to computer architecture, but this is just to go through some.

Now to see how the IBM 7030 Stretch compares to today's supercomputers, let's compare it to the IBM Summit. The Summit "links more than 27,000 NVIDIA Volta GPUs with more than 9,000 IBM Power9 CPUs" (Oar Ridge National Laboratory, *Summit*), making up its 4,608 nodes, whereas when the 7030 was being made nodes and GPUs hadn't even existed at the time. To see what a node in Summit looks like look at figure 6. If we look at how many transistors Summit has, we can see just how far supercomputers have come. The IBM 7030 in total had only 169,000 transistors. A single CPU chip in the Summit contains 8 billion, and that multiplied by the 9,216 CPUs it has is 73,728,000,000,000 or approximately 73 trillion transistors. This is approximately 436 million times more than what the 7030 had. The Power9 CPU supports hardware threads, "each of the Power9's physical cores have 4 'slices'. These slices provide simultaneous Multi Threading (SMT) support within the core" (Oak Ridge

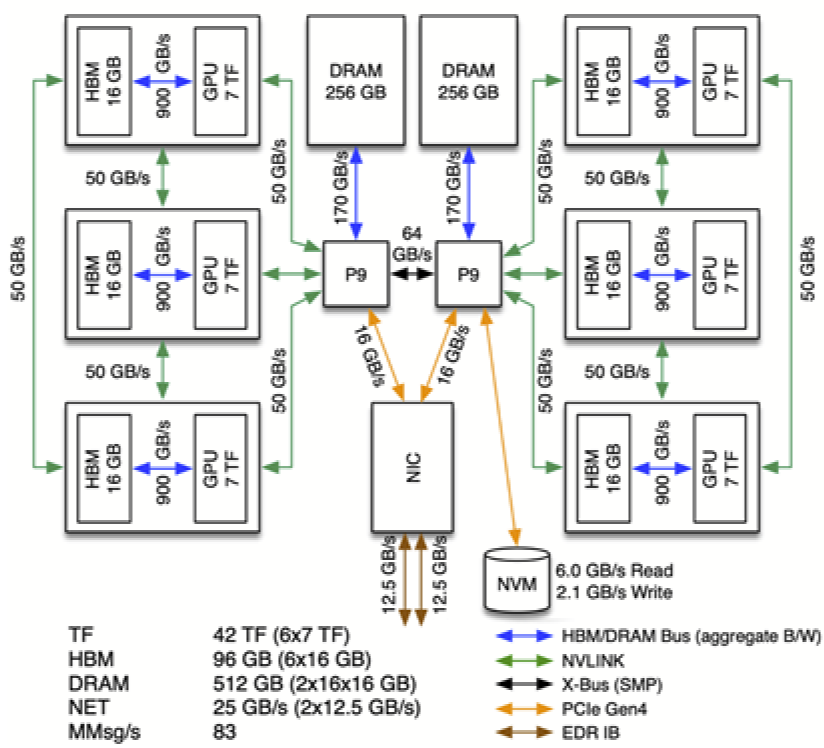


Figure 6: IBM Summit Node

National Laboratory, *Summit user guide*). These ‘slices’ work independently meaning that 4 executions can happen at the same time per core, with each CPU having 21 cores. To depict the 21 cores and their ‘slices’ look to figure 7. MPS also exists in Summits system. MPS or Multi-Process Service works with something called CUDA. CUDA (Computer Unified Device Architecture) is “parallel computing platform and programming model created by Nvidia” (Oh, 2012). Using CUDA with MPS “enables multiple processes to concurrently share the resources on a single GPU” (Oak Ridge National Laboratory, *Summit user guide*), which can result in better GPU utilization. The GPUs also supports individual thread scheduling and contain Tensor cores, all improvements to speed and efficiency compared to the 7030 Stretch. There are also major advancements in memory with Summit having a filesystem capable of 250PB, with 2.5TB/s write speeds. To add on to that, each node has 512GB DDR4,

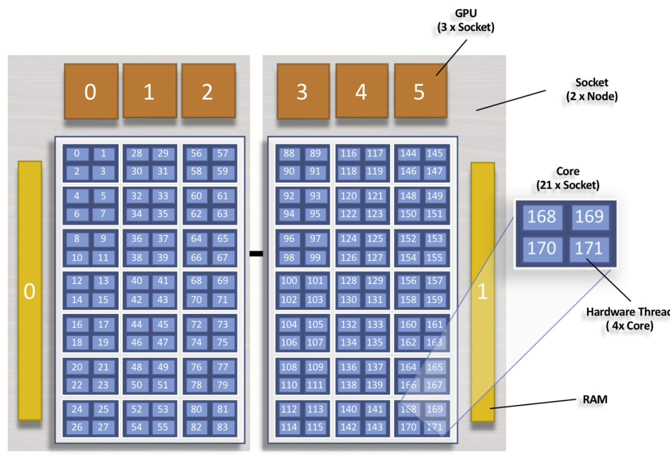


Figure 7: IBM CPU Cores

96GB HBM2 (High Bandwidth Memory) and 1600GB of Non-Volatile Memory, giving the entire system more than 10PB of all these memory types combined. There is also more to nodes than what is in them (CPU, GPU…). Each node is mainly similar in terms of how hardware, but there are actually three different types of nodes. There are Login, Launch, and Compute nodes. Login nodes are where “write/edit/compile your code, manage data, submit jobs, etc.” (Oak Ridge National Laboratory, *Summit user guide*). Launch nodes are when your code starts going, all of those commands will be done on launch nodes. Compute nodes make up the majority and is “where your parallel job executes” (Oak Ridge National Laboratory, *Summit user guide*). All of the nodes are connected via ‘dual-rail EDR InfiniBand network’. For more simple stats on Summit, Summit is capable of doing 200PF’s or 200 quadrillion floating point operations per second, and its original Operation System is the Red Hat Enterprise Linux (RHEL) version 7.4, which has since been updated to RHEL version 8.2. The Summit is capable of doing “astrophysics, materials science, systems biology, cancer research, and AI/ML” (Fisher, 2018), AI being Artificial Intelligence, and ML being Machine Learning. These two applications on their own are extremely demanding and are proof in their own to the capabilities of the newer generation of supercomputers being used to tackle problems that could only be imagined when the IBM 7030 Stretch came out. The amount of progress that has occurred from the 1960s to now is insane and we can only look forward to this continuation.

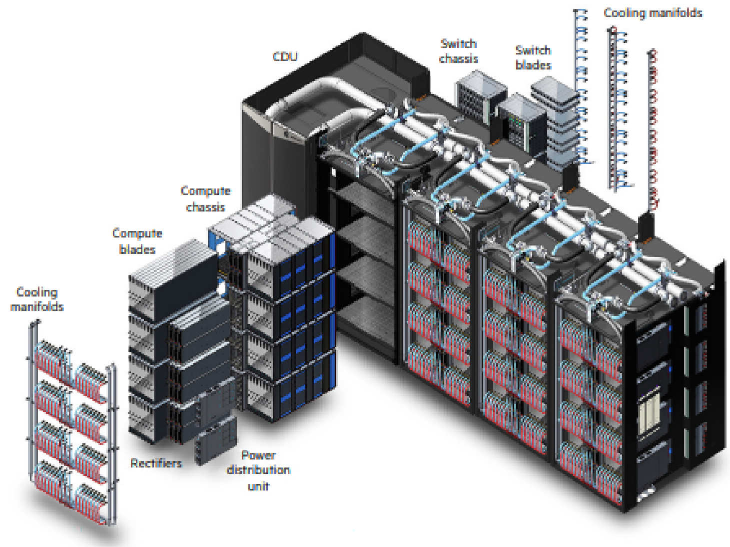
What is the future of supercomputers?

The supercomputer industry is already preparing for the ExaFlops, which can perform more than 100 operations per second beyond peta. If the performance of supercomputers improves, the accuracy of large-scale calculations and simulations will increase with it. The "Data" is one of the main key points in the 21st century technology race. It can change depending on who obtains more quality data and derives more valuable results from it.

The race for the best supercomputers is still going. One company who is trying to jump ahead in the supercomputer industry is Cray with their HPE Cray supercomputer. As the demands for supercomputers keep increasing and the need for more powerful supercomputers, due to the increase of calculation, the results needed and the ever changing needs of today's society. In the next decade, we will probably see a big change in the architecture of supercomputers and Cray is already trying to implement it in their HPE Cray supercomputer. With the bigger demands of calculation coming in, the huge advancement in the artificial intelligence (AI) industry and for any other big projects where supercomputers are needed, the architecture needed an update to keep up with everything. This new architecture has many new features to help and support the high energy demande of the CPUs and GPUs in the supercomputers will also keep the cost of management at a minimum. Some of these advancement are that they are reducing the amount of intercabling required by the supercomputers to lower the cost by buying and less expensive electrical wires and to also decrease the amount of space that these wires took to increase the amount of CPUs and GPUs that they can fit in their supercomputers. Another step that they decided to take into improving their architecture, is that they are changing their cooling system to a liquid cooling system. This will not only once again deduce the cost of cabling from the previous cooling system, but it will once again reduce the amount of space that it takes leaving more room for more CPUs and GPUs in their system. Not only is this architecture improving on what it had before, but with the amount of space and money that they are saving from what they had before, they leave a lot of room open for any new changes or demands that are needed as they move forward into the future, whether that be the next-generation of GPUs and CPUs new wiring, new cooling systems etc.

Figure 8:

HPE Cray EX supercomputer



For the HPE Cray EX supercomputer architecture, each cabinet can hold up to 8 compute chassis which means 64 compute blades (8 per compute chassis) and 512 CPUs (8 per compute blades) as seen in the figure 8. The compute chassis is what holds the compute blades and everything else all together (like a box inside the cabinet), the compute blades are the ones that hold all the components needed for the supercomputer to run and make their calculations. It holds: the processors, the connectors, the circuit boards, the cooling system and the power system. For each of the compute chassis, there is a need for a switch chassis (which is connected to it). The switch chassis is what holds the switch blades and everything else related to it. The switch blades hold the HPE Slingshot fabric (High speed cables), the circuit boards, the connectors to the compute blades, the cooling system and the power system needed. The switch blades use is to make the connection between all of the compute blades so that they can communicate with each and gather all the information from all of the compute blades at a speed of 12.8 Tb/s. All of these components work together to create a single cabinet and supercomputers have hundreds of cabinets that are all linked together meaning thousands of CPUs and GPUs working together to accomplish their task.

Conclusion:

Supercomputers were created for the purpose of handling long and complicated calculations that could not be calculated with normal computers. Supercomputers got a big growth in development with the development of nuclear weapons, plants and reactors. This report talked about the definition of what a supercomputer is. From what it's made out of, to how it's built, what's its use etc. This report also talks about the history of supercomputers, from the first supercomputers to the latest of supercomputers. It also contained information on the different companies that have worked on supercomputers that have also helped in its development to what it is today. The architecture of the supercomputers is also included in this report, talking about the details of supercomputers such as the system that it uses, how it works all together and how much it can actually do. To finish off, this report talks about the future of supercomputers and where it's going. For the most part, the evolution of supercomputers consists of making them more powerful and making them able to keep up with the development of other technologies while working on improving its flaws and problems. Some of these problems that are still occurring in supercomputers but are being worked on are its cost (multiple pricy components), its energy consumption (they run for 24 hours a day consuming a lot of electricity) and the space that it takes (they need a lot of room to be able to fit all the machinery and components). Overall, supercomputers have made a lot of development over the years with the development of other technologies surrounding it. They are complex machines that are doing most of the large-scale calculations and handle the results that are hard for normal machines to match their work. Personally we think that supercomputers are a must for our future since they help us a lot but with its development, we see that it's getting hard to improve on the size and the cost of it while keeping the same system in place. Thus maybe we will see big changes in the architecture of it for the better of supercomputers.

References

Work cited

13 fastest supercomputers in the World: In 2022. RankRed. (2022, February 3). Retrieved from https://www.rankred.com/fastest-supercomputers-in-the-world/

Advantages and disadvantages of SMP and MPP. actorsfit. (n.d.). Retrieved from https://blog.actorsfit.com/a?ID=00400-c9607d3f-5c1a-496b-ad6c-589d872dd675

Amodei, D. (2021, June 21). *Ai and Compute*. OpenAI. Retrieved from https://openai.com/blog/ai-and-compute/

Anthony, S. C. (2012, April 12). *The history of Supercomputers*. ExtremeTech. Retrieved from https://www.extremetech.com/extreme/125271-the-history-of-supercomputers

Arm. (n.d.). *What is Instruction Set Architecture (ISA)?* Arm. Retrieved from https://www.arm.com/glossary/isa

Biography. IEEE Computer Society Seymour R Cray Comments. (n.d.). Retrieved from https://www.computer.org/profiles/seymour-cray

Brock, D. (2019, September 24). *Seymour Cray: The man who brought style to supercomputers*. CHM. Retrieved from https://computerhistory.org/blog/seymour-cray-the-man-who-brought-style-to-supercomputers/

Computational and information systems lab. CDC 6600 | Computational and Information Systems Lab. (n.d.). Retrieved from https://www2.cisl.ucar.edu/supercomputer/cdc6600

Computer History Museum. (n.d.). *Artifact details*. 7030 Engineering Console (IBM Stretch) | 102716145 | Computer History Museum. Retrieved from https://www.computerhistory.org/collections/catalog/102716145

Computer History Museum. CDC 6600's Five Year Reign. (n.d.). Retrieved from https://www.computerhistory.org/revolution/supercomputers/10/33

Computer History Museum. (n.d.). *1944: Timeline of Computer History: Computer History Museum*. 1944 | Timeline of Computer History | Computer History Museum. Retrieved from https://www.computerhistory.org/timeline/1944/

Computer History Museum. (n.d.). Cray Research, Inc. Retrieved from https://www.computerhistory.org/revolution/supercomputers/10/35

Computer Science Degree Hub. (2018, October 8). *What is computer architecture?* Computer Science Degree Hub. Retrieved from https://www.computersciencedegreehub.com/faq/what-is-computer-architecture/

Cray history - supercomputers inspired by Curiosity - Seymour cray. HPE. (n.d.). Retrieved from https://www.hpe.com/us/en/compute/hpc/cray.html

Encyclopedia Britannica, inc. (n.d.). *IBM*. Encyclopedia Britannica. Retrieved from https://www.britannica.com/topic/International-Business-Machines-Corporation

Feldman, M. (2016, August 24). *IBM Power9*. TOP500. Retrieved from https://www.top500.org/news/ibm-ups-its-game-with-the-power9-processor/

Fisher, Y. (2018, August 20). *Exploring summit: The newest intelligent supercomputer*. Red Hat - We make open source technologies for the enterprise. Retrieved from https://www.redhat.com/en/blog/exploring-summit-newest-intelligent-supercomputer

F, M. (2022, March 11). *What is the difference between vector and scalar processors?* EasyTechJunkie. Retrieved from https://www.easytechjunkie.com/what-is-the-difference-between-vector-and-scalar-processors.htm

Fujitsu. (2021, November 16). *Japan's Fugaku retains title as World's fastest supercomputer for fourth consecutive term*. Fujitsu Global. Retrieved from https://www.fujitsu.com/global/about/resources/news/press-releases/2021/1116-01.html

Fujitsu. (n.d.). *Company Milestones*. Fujitsu Global. Retrieved from https://www.fujitsu.com/global/about/corporate/history/company-milestones/history.html#anc-04

Fujitsu. (n.d.). *History of Fujitsu*. Fujitsu Global. Retrieved from https://www.fujitsu.com/global/about/corporate/history/

Fujitsu. (n.d.). *Product Milestones*. Fujitsu Global. Retrieved April from https://www.fujitsu.com/global/about/corporate/history/products/index.html

Hardy, N. (n.d.). *The IBM 7030 (STRETCH)*. The IBM 7030 (stretch). Retrieved from https://computer-history.info/Page4.dir/pages/IBM.7030.Stretch.dir/

Hewlett Packard Enterprise Development LP. (2020, August). *HPE Cray ex liquid-cooled cabinet for large-scale systems*. Hewlett Packard Enterprise (HPE). Retrieved from https://www.hpe.com/us/en/collaterals/collateral.a50002389enw.HPE-Cray-EX-Liquid-Cooled-Cabinet-for-Large-Scale-Systems-brochure.html?rpv=cpf&parentPage=%2Fus%2Fen%2Fproducts%2Fcompute%2Fhpc%2Fsupercomputing%2Fcray-exascale-supercomputer

Hirapara, K. (n.d.). *Supercomputer vs. the cloud - vtech solution, Inc.*. Supercomputer vs. The Cloud. Retrieved from https://www.vtechsolution.com/supercomputer-vs-the-cloud/

Hp. (2020, January 9). *History of supercomputing: HP® Tech takes*. History of Supercomputing | HP® Tech Takes. Retrieved from https://www.hp.com/us-en/shop/tech-takes/history-of-supercomputing

HP. (n.d.). *Cray history - supercomputers - Seymour Cray*. Cray History - Supercomputers - Seymour Cray | HPE Canada. Retrieved from https://www.hpe.com/ca/en/compute/hpc/cray.html

Hp. (2021, May 31). *The fastest supercomputers ever built and WHO BUILT THEM: HP® Tech takes*. The Fastest Supercomputers Ever Built and Who Built Them | HP® Tech Takes. Retrieved from https://www.hp.com/us-en/shop/tech-takes/fastest-supercomputers-ever-built

https://futureoflife.org/wp-content/uploads/2015/10/FLI_logo-1.png. (2016, March 8). *AlphaGo and Ai Progress*. Future of Life Institute. Retrieved from https://futureoflife.org/2016/03/08/alphago-and-ai-progress/

IBM stretch (aka IBM 7030 Data Processing System). (2009, June 17). Retrieved from http://www.brouhaha.com/~eric/retrocomputing/ibm/stretch/

IBM. (n.d.). *7030 data processing system*. IBM Archives: 7030 Data Processing System. Retrieved from https://www.ibm.com/ibm/history/exhibits/mainframe/mainframe_PP7030.html

IBM. (2018, June 6). *Meet two of the most powerful supercomputers on the planet*. IBM. Retrieved from https://www.ibm.com/thought-leadership/summit-supercomputer/

IBM. (n.d.). *What is supercomputing?* IBM. Retrieved from https://www.ibm.com/topics/supercomputing

IBM. (n.d.). IBM Supercomputing Timeline. IBM. Retrieved from https://1.cms.s81c.com/sites/default/files/2020-04-22/0d867600-8499-11ea-9a7f-a81d902bd496.jpg?_ga=2.149489462.1009109276.1649513244-181999856.1649513244.

Indiana University. In high-performance computing, what are Rmax and Rpeak? (n.d.). Retrieved from https://kb.iu.edu/d/bbzo

Layton, J. (n.d.). *A brief history of supercomputers " Admin Magazine*. ADMIN Magazine. Retrieved from https://www.admin-magazine.com/HPC/Articles/A-Brief-History-of-Supercomputers

Lawrence Livermore National Laboratory. (n.d.). Sierra-Fact-Sheet. (PDF)

Lawrence Livermore National Laboratory. (n.d.). *Sierra*. Sierra | HPC @ LLNL. Retrieved from https://hpc.llnl.gov/hardware/compute-platforms/sierra

Moss, S. (2022, February 28). *China may already have two exascale supercomputers*. All Content RSS. Retrieved from https://www.datacenterdynamics.com/en/news/china-may-already-have-two-exascale-supercomputers/

Oak Ridge National Laboratory. (n.d.). *Solving the big problems: ORNL*. Solving the Big Problems. Retrieved from https://www.ornl.gov/content/solving-big-problems

Oak Ridge National Laboratory. (n.d.). *Summit*. Oak Ridge Leadership Computing Facility. Retrieved from https://www.olcf.ornl.gov/summit/

Oak Ridge National Laboratory. (n.d.). *Summit user guide*. Summit User Guide - OLCF User Documentation. Retrieved from https://docs.olcf.ornl.gov/systems/summit_user_guide.html#system-overview

O'Flaherty, D. (2021, March 8). *The fastest storage for the fastest system: Summit*. Servers & Storage. Retrieved from https://www.ibm.com/blogs/systems/fastest-storage-fastest-system-summit/

Oh, F. (2012, September 10). *What is Cuda: Nvidia official blog*. NVIDIA Blog. Retrieved from https://blogs.nvidia.com/blog/2012/09/10/what-is-cuda-2/

Out-of-order execution. (n.d.). Retrieved from https://www.cs.uaf.edu/2011/spring/cs641/proj1/vsanditi/

Potter, J. (2012, January 9). International Business Machines Corporation (IBM). Retrieved from https://www.thecanadianencyclopedia.ca/en/article/international-business-machines-corporation-ibm

Prakash, A. (2020, November 20). *Linux runs all of the top 500 fastest supercomputers*. It's FOSS. Retrieved from https://itsfoss.com/linux-runs-top-supercomputers/

Rosenfield, M. (2019, February 8). *Summit: Introducing the world's smartest, most powerful supercomputer*. IBM Research Blog. Retrieved from https://www.ibm.com/blogs/research/2018/06/summit/

Roundy, J. (2021, September 2). *Learn the difference between SMP vs. MPP*. SearchDataCenter. Retrieved from https://www.techtarget.com/searchdatacenter/tip/Learn-the-difference-between-SMP-vs-MPP

Scalar processor. Scalar Processor - an overview | ScienceDirect Topics. (n.d.). Retrieved from https://www.sciencedirect.com/topics/computer-science/scalar-processor

Schneck, P. B. (1987, January 1). *The IBM 7030: Stretch*. SpringerLink. Retrieved from https://link.springer.com/chapter/10.1007/978-1-4615-7957-1_2

Science Direct. (n.d.). *Computer architecture*. Computer Architecture - an overview | ScienceDirect Topics. Retrieved from https://www.sciencedirect.com/topics/computer-science/computer-architecture

Smotherman, M. (2010, July). *IBM stretch (7030) -- aggressive uniprocessor parallelism*. Organization Sketch of IBM Stretch -- Mark Smotherman. Retrieved from https://people.cs.clemson.edu/~mark/stretch.html

Stepania, D. (2019, January 29). *Supercomputing vs. cloud computing*. LinkedIn. Retrieved from https://www.linkedin.com/pulse/supercomputing-vs-cloud-computing-david-stepania

Supercomputer Fugaku - supercomputer Fugaku, A64FX 48C 2.2ghz, tofu interconnect D. TOP500. (n.d.). Retrieved from https://www.top500.org/system/179807/

Supercomputer types: Different types of supercomputer. EDUCBA. (2021, June 30). Retrieved from https://www.educba.com/supercomputer-types/?source=leftnav

Suse. (2018, March 26). *What is a supercomputer?: Answer from SUSE defines*. SUSE Defines. Retrieved from https://www.suse.com/suse-defines/definition/supercomputer/

Team, S. Q. L. S. (2021, January 18). *Transitioning from SMP to MPP, the why and the how*. Microsoft SQL Server Blog. Retrieved from https://cloudblogs.microsoft.com/sqlserver/2014/07/30/transitioning-from-smp-to-mpp-the-why-and-the-how/

Techopedia. (2017, February 15). *What is Microarchitecture (μarch)? - definition from Techopedia*. Techopedia.com. Retrieved from https://www.techopedia.com/definition/4502/microarchitecture-arch

Techopedia. (2017, June 20). *What is computer architecture? - definition from Techopedia*. Techopedia.com. Retrieved from https://www.techopedia.com/definition/26757/computer-architecture

Tsukimori, O. (2021, January 8). *Japan's Fugaku supercomputer is tackling some of the world's biggest problems*. The Japan Times. Retrieved from https://www.japantimes.co.jp/news/2021/01/07/business/tech/japans-fugaku-supercomputer/

University of Sunderland. (2021, August 26). *What is computer architecture?* University of Sunderland. Retrieved from https://online.sunderland.ac.uk/what-is-computer-architecture/

Vaughan-Nichols, S. (2017, November 27). *A super-fast history of supercomputers: CDC 6600 to Today*. HPE. Retrieved from https://www.hpe.com/us/en/insights/articles/a-super-fast-history-of-supercomputers-from-the-cdc-6600-to-the-sunway-taihulight-1711.html

Vector Processor. Vector Processor - an overview | ScienceDirect Topics. (n.d.). Retrieved from https://www.sciencedirect.com/topics/computer-science/vector-processor

Woodford, C. (2021, July 23). *Supercomputers - A simple introduction and history*. Explain that Stuff. Retrieved from https://www.explainthatstuff.com/how-supercomputers-work.html

Image work cited

Craft, L. (2020, June 26). *The world's fastest supercomputer has been chosen, and it's hunting for coronavirus cures*. CBS News. Retrieved from https://www.cbsnews.com/news/fastest-supercomputer-fugaku-in-japan-busy-on-coronavirus-treatment-weather-forecasting-today/

Hewlett Packard Enterprise Development LP. (2020, August). *HPE Cray ex liquid-cooled cabinet for large-scale systems*. Hewlett Packard Enterprise (HPE). Retrieved from https://www.hpe.com/us/en/collaterals/collateral.a50002389enw.HPE-Cray-EX-Liquid-Cooled-Cabinet-for-Large-Scale-Systems-brochure.html?rpv=cpf&parentPage=%2Fus%2Fen%2Fproducts%2Fcompute%2Fhpc%2Fsupercomputing%2Fcray-exascale-supercomputer

IBM. (n.d.). IBM Supercomputing Timeline. IBM. Retrieved from https://1.cms.s81c.com/sites/default/files/2020-04-22/0d867600-8499-11ea-9a7f-a81d902bd496.jpg?_ga=2.149489462.1009109276.1649513244-181999856.1649513244.

NIHF inductee Seymour Cray invented the Supercomputer. NIHF Inductee Seymour Cray Invented the Supercomputer. (n.d.). Retrieved from https://www.invent.org/inductees/seymour-cray

Oak Ridge National Laboratory. (n.d.). *Summit user guide*. Summit User Guide - OLCF User Documentation. Retrieved from https://docs.olcf.ornl.gov/systems/summit_user_guide.html#burst-buffer

Prakash, A. (2020, November 20). *Linux runs all of the top 500 fastest supercomputers*. It's FOSS. Retrieved from https://itsfoss.com/linux-runs-top-supercomputers/

Wikimedia Foundation. (2021, December 12). *CDC 6600*. Wikipedia. Retrieved from https://en.wikipedia.org/wiki/CDC_6600

Work Distribution

Sebastian Carneiro: 100798391, Thursday lecture

- 33.33%
- The companies Are Working on Supercomputers
- Supercomputer Architecture
- References (organizing)
- Grammar and review

Caleb Fontaine: 100832588, Thursday lecture

- 33.33%
- Abstract
- Introduction
- The future of supercomputer
- Conclusion
- Grammar and review
- Page structure

Haekang Song: 100625189, Wednesday lecture

- 33.33%
- Introduction to supercomputer
- History of supercomputer
- Review