EXASCALE COMPUTING: CHALLENGES AND POTENTIAL SOLUTIONS

BART Number 208912

INTRODUCTION

The computational power of supercomputers have increased year-on-year. The massive computational power of supercomputers play a major role in the advancement of many scientific disciplines [1], particularly in case of Exascale computers, these would primarily improve scientific applications and prediction accuracy in domains such as weather forecasting, climate modelling, and personalised medicines [2].

Exascale computers are the class of computer systems that can calculate at least 10^{18} FLOPS [3]. The world's first Exascale computer, Frontier, that was announced in 2022, is the world's fastest supercomputer as per the June 2023 Top 500 List [4]. In addition to Frontier in the United States, OceanLight and Tianhe-3 are two other Exascale computers that are operational in China, with Aurora in the United States and Jupiter in Europe scheduled to launch soon [5].

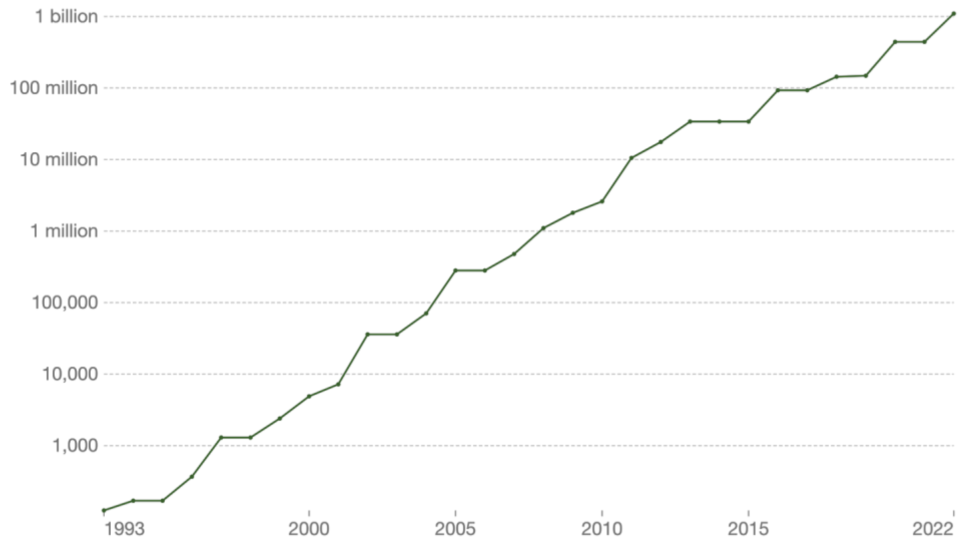

Figure 1: The increase of computational power in TFLOPS from the year 1993 to 2022 [4]

CHALLENGES

> *"Challenges [for Exascale computers] include a computational infrastructure that model the complexity of biological systems realistically and simulate them efficiently."*
>
> Kalpana S. Batti [5]

Many reports suggest that whilst Exascale computers have revolutionised analyses and simulations across scientific disciplines, there is lack of computational infrastructure for complex and stochastic systems.

The challenges that are posed by Exascale computers provided by Sterling *et al.* [6] and Turczyn [7] are listed as:

(i) **Energy and Power:** This is a factor considered as the maximum power that can feed a semiconductor die before it reaches a threshold of failure. A target goal on this factor is 20 MW or 50 Gflops/W.

(ii) **Hardware parallelism:** This pertains to having hundreds of millions of cores, each operating with 10-way parallelism like SIMD or vector.

(iii) **Software parallelism:** This pertains to application programs and algorithms using and exploiting more than a billionfold parallelism to take advantage of the hardware, including communication and secondary storage access.

(iv) **Overhead:** This pertains to the work required to manage the system and to control each task. This factor considers the granularity of the tasks, the available useful concurrency, and is essentially the source of loss of efficiency.

(v) **Latency:** With larger systems to avail global access of data and services increase, this would require even more parallelism to hide the latency (delay in communication).

(vi) **Reliability:** The potential reduction of the mean time to failure due to chances of single-point failures increasing could make Exascale computing impractical.

(vii) **Speed and Energy of data movement:** This is a significant computational problem, where the time taken to solution is dominated by the movement of data rather than number of computations. The impediment of getting data in and out of memory is referred to as the *memory wall*. This factor is identical to the latency factor described above.

(viii) **Fault tolerance:** Exascale systems were expected to become so large and so complex that computer failure rates would increase to the point of becoming insurmountable. This factor may sound similar to 'reliability'; However, it is slightly different when considering the *system resilience* to any fault that may arise.

In addition to these, the development of Exascale computing will demand advances in 3D packaging, high-speed electrical and optimal signalling and efficient in power conversion, and cooling and memory technology [8].

In this study, the 'reliability' factor is explored further and potential solutions to reduce this challenge towards Exascale are discussed.

THE RELIABILITY CHALLENGE

> *"Extrapolating from 2009 error rates, we predicted that Exascale system failures might happen faster than you could checkpoint a job. If the failure rate got this high, then traditional supercomputer applications would not be able to roll back to a checkpoint and make forward progress,"*
>
> - Al Geist, CTO, Oak Ridge National Laboratory [7]

The reliability of a system is referred to as the probability that it will function without failure under stated conditions for a specified amount of time [9]. Mean time to failure (MTTR) is a key reliability metric to assess this factor [10].

LANL has performed a study, where it covers 23,000 failures recorded on more than 20 different high-performance computing systems, to determine root cause analysis of failures recorded [11]. The root cause breakdown presented was:

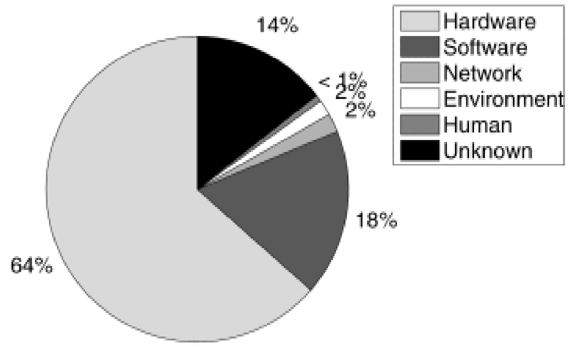

Figure 2: The breakdown of root causes, as presented by LANL [11]

They further analyse the likelihood of node failures for their system 20.

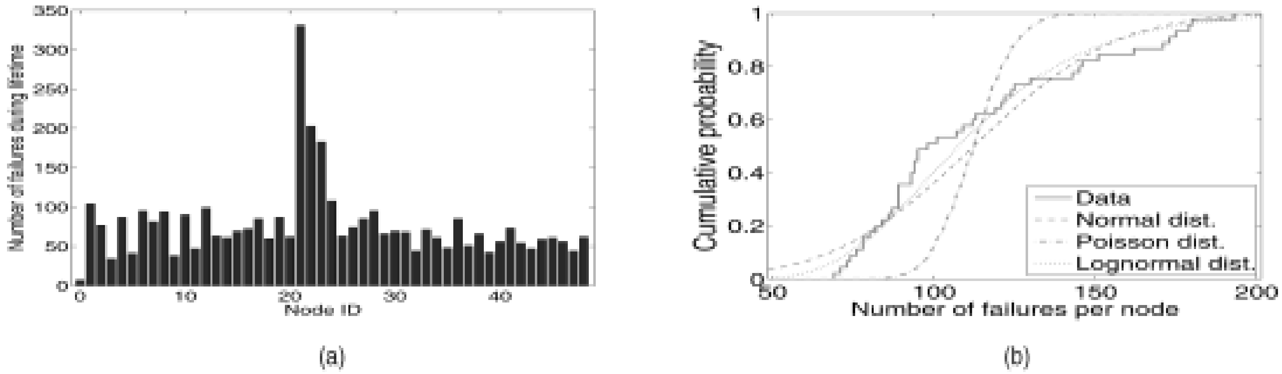

Figure 3: Extracted from the LANL report, (a): The number of failures occurred during a lifetime of a corresponding node for system 20; (b): A probability distribution of number of failures occurred per node [11]

From the above graph, it can be observed that nodes between 21-23 have a significant failure. Schroeder and Gibson [11] justify this significant hike due to these nodes running a different set of workloads, majorly involving visualisation and computation of a more varied and interactive nature, as compared to the other nodes from system 20.

Potential Solutions

A potential solution to reducing the checkpoint time below MTTR, which would help in overcoming reliability challenges. This is an example from Frontier's on-node, non-volatile memory, where checkpoint times were reduced from minutes to seconds, making the checkpoint time much shorter than MTTR [7]. However, Yang *et al.* [10] mention that this solution comes with a capital and time cost. As an improvement, they suggest an incorporation of reliability speedup factor, which is a function of number of nodes and capital costs. Their results demonstrate that when number of nodes increase, their costs flatline [10].

Another possible solution to resolving the reliability challenge is described by Chung *et al.* [12] is the introduction of Containment Domains (CD). CDs have weak transactional semantics and are nested to take advantage of the machine and application hierarchies and to enable hierarchical state preservation, restoration and recovery. These demonstrated superiority to checkpoint restart and redundant execution approaches [13].

One solution involving a task-based approach in overcoming reliability challenges: The codesign by Kerbyson *et al.* [14] utilizes a task-based approach that applies the Global Arrays (GA) programming model, in which a task is designed as a unit of computation with input, output, and dependencies to other tasks. A task's self-contained properties make one-sided communication a perfect model for describing data dependencies, simplifying the implementation of many applications [14].

Conclusion

It is apparent that there are multiple challenges faced in the development of Exascale computers, however, their applications are very promising. While there are multiple approaches to overcoming challenges, the fact that exascale are either currently operational or are scheduled to, there is various references [15], [16] to how these challenges were overcome to developing Exascale computers.

References

[1] S. Heldens, P. Hijma, B. Van Werkhoven, J. Maassen, A. S. Z. Belloum, and R. V. Van Nieuwpoort, 'The Landscape of Exascale Research', *ACM Comput Surv*, vol. 53, no. 2, 2021, doi: 10.1145/3372390.

[2] F. Gagliardi, M. Moreto, M. Olivieri, and M. Valero, 'The international race towards Exascale in Europe', *CCF Transactions on High Performance Computing*, vol. 1, no. 1. 2019. doi: 10.1007/s42514-019-00002-y.

[3] P. Kogge *et al.*, 'Exascale Computing Study: Technology Challenges in Achieving Exascale Systems', *DARPA*, Sep. 2008.

[4] www.top500.org, 'Top 500 - June 2023 highlights'. www.top500.org (accessed Jul. 09, 2023).

[5] C. Chang, V. L. Deringer, K. S. Katti, V. Van Speybroeck, and C. M. Wolverton, 'Simulations in the era of exascale computing', *Nat Rev Mater*, 2023, doi: 10.1038/s41578-023-00540-6.

[6] T. Sterling, M. Anderson, and M. Brodowicz, *High Performance Computing: Modern Systems and Practices*. 2017. doi: 10.1016/C2013-0-09704-6.

[7] C. Turczyn, 'Exascale Computing's four biggest challenges and how they were overcome', *Oak Ridge National Laboratory*, Oct. 18, 2021. https://www.olcf.ornl.gov/2021/10/18/exascale-computings-four-biggest-challenges-and-how-they-were-overcome/ (accessed Jul. 09, 2023).

[8] P. W. Coteus, J. U. Knickerbocker, C. H. Lam, and Y. A. Vlasov, 'Technologies for exascale systems', *IBM J Res Dev*, vol. 55, no. 5, 2011, doi: 10.1147/JRD.2011.2163967.

[9] J. Stearley, 'Defining and Measuring Supercomputer Reliability, Availability, and Serviceability', in *Proceedings of the Linux Clusters Institute Conference*, 2005.

[10] X. Yang, Z. Wang, J. Xue, and Y. Zhou, 'The reliability wall for exascale supercomputing', *IEEE Transactions on Computers*, vol. 61, no. 6, 2012, doi: 10.1109/TC.2011.106.

[11] B. Schroeder and G. Gibson, 'A large-scale study of failures in high-performance computing systems', *IEEE Trans Dependable Secure Comput*, vol. 7, no. 4, 2010, doi: 10.1109/TDSC.2009.4.

[12] J. Chung *et al.*, 'Containment Domains: A Scalable, Efficient and Flexible Resilience Scheme for Exascale Systems', *Sci Program*, vol. 21, no. 3–4, 2013, doi: 10.1155/2013/473915.

[13] A. Moody, G. Bronevetsky, K. Mohror, and B. Supinski, 'Design, Modeling, and Evaluation of a Scalable Multi-level Checkpointing System', in *2010 ACM/IEEE International Conference for High Performance Computing, Networking, Storage and Analysis, SC 2010*, Jul. 2010, pp. 1–11. doi: 10.1109/SC.2010.18.

[14] D. Kerbyson, A. Vishnu, K. Barker, and A. Hoisie, 'Codesign challenges for exascale systems: Performance, power, and reliability', *Computer (Long Beach Calif)*, vol. 44, no. 11, 2011, doi: 10.1109/MC.2011.298.

[15] J. Shalf, ‘The future of computing beyond Moore’s Law’, *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 378, no. 2166. 2020. doi: 10.1098/rsta.2019.0061.

[16] O. Terzo and J. Martinovič, *HPC, Big Data, and AI Convergence Towards Exascale: Challenge and Vision*. 2022. doi: 10.1201/9781003176664.