Energy-aware Resilience Approaches for Large-scale Systems

Xiaolong Cui
mclarencui@cs.pitt.edu

1 Introduction

As our reliance on IT continues to increase, the complexity and urgency of the problems our society will face in the future will increase much faster than are our abilities to understand and deal with them. Future IT systems are likely to exhibit a level of interconnected complexity that makes it prone to failure and exceptional behaviors. The high risk of relying on IT systems that are failure-prone calls for new approaches to enhance their performance and resiliency to failure. My research focuses on novel and energy-aware computational models that tolerate failures for large-scale distributed systems, including HPC supercomputers and cloud data centers.

Current fault-tolerance approaches are designed to deal with fail-stop errors, and rely on either time or hardware redundancy for recovery. Checkpoint/restart, which uses time redundancy, requires full or partial re-execution when failure occurs. Such an approach can incur a significant delay, and high energy costs due to extended execution time. On the other hand, Process Replication exploits hardware redundancy and executes multiple instances of the same task in parallel to guarantee completion without delay. This solution, however, requires additional hardware resources and increases the energy consumption proportionally.

My proposed approaches to resiliency go beyond adapting or optimizing well known and proven techniques, and explore radical methodologies to fault tolerance in large-scale computing infrastructures. The proposed solutions differ in the type of faults they manage, their design, fault tolerance approach and the fault tolerance protocol they use. It is not just a scale up of "point" solutions, but an exploration of innovative and scalable fault tolerance frameworks. When integrated, it will lead to efficient solutions for a "tunable" resiliency that takes into consideration the nature of the data and the requirements of the application.

2 Methodology

The basic tenet of the proposed fault tolerance framework, referred to as Shadow Computing, is to associate with each main execution instance (a process or a thread) a suite of "shadow instances", whose size depends on the criticality of the application and its performance requirements. To mitigate correlated failures, the main and shadow instances execute on separate computing nodes.

The novelty of Shadow Computing lies in its differentiation of the execution rates. Specifically, it executes the main instance at the rate required for response time constraint, while slowing down the shadows for energy saving, thereby enabling a parameterized trade-off between response time and energy consumption. Shadow Computing is a generalization of existing fault tolerance approaches. Specifically, if there is enough laxity in response time constraint, Shadow Computing would start the shadows only after the main instance fails, mimicking Re-execution. If the target response time is stringent, however, the shadows would execute simultaneously with the main at the maximal rate, mimicking Process Replication. The flexibility of Shadow Replication provides a spectrum of fault tolerance strategies that strike a balance between completion time and energy saving.

2.1 Execution rate control

One challenge in Shadow Computing is how to control the execution rates. So far, we have explored two methods to differentiate the execution rates among the instances. The first approach is to use Dynamic Voltage and Frequency Scaling (DVFS), of which the CPU frequency (execution rate) scales linearly with the supply voltage and the power consumption has a super-linear relationship with the frequency [1,2]. Our second approach is to use time sharing to reduce the effective execution rate while keeping the computing nodes running at maximal frequency. Since this approach collocates multiple execution instances on a same node, it reduces the number of machines to be used and the energy consumption correspondingly [3].

2.2 Modeling and Simulation

The other challenge resides in determining jointly the execution rates of all processes, with the objective to minimize energy while satisfying QoS requirements. To achieve this, I propose two analytical models, corresponding to the two rate-control approaches above. The models consider the time and energy needed for a job, under different system specifics and failure distributions. From the analytical models an optimization problem is formulated and solved to derive the optimal execution rates.

To verify the correctness of the analytical models, I build an event-driven simulator that simulates the behaviors of Shadow Computing under different system specifics, task characteristics, and failure distributions. It can report all necessary statistics, such as number of failures encountered, time to completion, and energy consumption. The statistics can then be used to compare with the results from the analytical model.

2.3 Results

Several important parameters are identified that impact the energy consumption of Shadow Computing. Correspondingly, I conduct a series of sensitivity studies where Shadow Computing is compared to state-of-the-art approaches. The results from both the analytical model and simulator show that Shadow Computing can achieve significant energy savings, without violating the QoS constraints. Specifically, Shadow Computing can achieve 15%-30% energy savings under normal configurations. Furthermore, Shadow Computing would converge to Process Replication, when target response time is stringent, and to Re-execution when target response time is relaxed or when failure is unlikely.

3 Conclusion

My current research is enabling new insights into the multi-faceted and challenging resiliency problem in large-scale distributed systems. The goal is to investigate radical approaches to the design of scalable and energy efficient fault tolerant approaches that go beyond state-of-the-art algorithms. Throughout my design, the interplay between resiliency, performance and energy consumption is analyzed carefully to determine the required levels of endurance and redundancy, in order to achieve a desired level of fault tolerance while maintaining a specified level of QoS.

Shadow Computing is a novel, scalable, and energy-aware computational model that achieves fault tolerance. Preliminary results show that Shadow Computing is able to achieve significant energy savings while satisfying QoS requirements, and adapt to various large-scale computing environments and requirements. I will continue to explore and optimize Shadow Computing in support of green and sustainable computing.

Acknowledgments

The author would like to thank Dr. Bryan Mills for his work in developing the initial analytical model (modeling of a single task) for Shadow Computing with DVFS. For clarification, the author performed all the other work presented in this paper.

References

[1] X. Cui, B. Mills, T. Znati, and R. Melhem, "Shadow replication: An energy-aware, fault-tolerant computational model for green cloud computing," *Energies*, vol. 7, no. 8, pp. 5151–5176, 2014.

[2] X. Cui, B. Mills, R. Melhem, and T. Znati, "Shadows on the cloud: An energy-aware, profit maximizing resilience framework for cloud computing." in *CLOSER 2014*, Barcelona, Spain.

[3] X. Cui, R. Melhem, and T. Znati, "Lazy shadowing: An adaptive, power-aware, resilience framework for extreme-scale computing." in *ICS 2015*, Newport beach, CA [under review].