Research Statement

Xiaolong Cui
Department of Computer Science
University of Pittsburgh
xic51@pitt.edu

1 Introduction

As our reliance on Information Technology (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 faults and exceptions. The high risk of relying on IT systems that are unreliable calls for new approaches to enhance their performance and resiliency to fault. Addressing this concern brings about unprecedented resiliency challenges, which put in question the ability of next generation IT infrastructure to continue operation in the presence of faults without compromising the requirements of compute- and data-intensive workloads.

Current fault-tolerance approaches rely on either time or hardware redundancy for recovery. Checkpoint/restart, which uses time redundancy, requires full or partial re-execution when fault 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.

It is without doubt that our understanding of how to build reliable systems out of unreliable components has led the development of robust and fairly reliable large-scale software and networking systems. The inherent instability of large-scale IT systems of the future in terms of the envisioned high-rate and diversity of faults, however, calls for a reconsideration of the fault tolerance problem as a whole. My proposed approach to resiliency goes beyond adapting or optimizing existing techniques, and explores radical methodologies to fault tolerance in large-scale computing environments, including both Cloud Computing and High Performance Computing. The proposed solutions differ in the type of faults they tolerate, their design, 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 Current progress

I have collaborated with Bryan Mills in coming up with and refining the initial idea of Shadow Replication. Then Bryan studies the application of Shadow Replication to a single task [1], while my focus is on large-scale distributed systems where a job is composed of multiple parallel tasks [2, 3]. Specifically, I model and study the performance of Shadow Replication in both Cloud Computing and High Performance Computing environments, and come up with techniques for optimization specific to each targeted environment.

I have accomplished the following goals:

- Design of a novel, scalable, and energy-aware fault tolerance framework, referred to as Shadow Replication;
- Development of a profit-based analytical model to explore the feasibility of Shadow Replication, and to determine the optimal execution rates to minimize energy and maintain quality of service (QoS);

- Design of a performance optimization technique, referred to as Leaping Shadows, for High Performance Computing jobs.
- A comprehensive evaluation using both the analytical model and simulator to analyze the energy savings achievable by Shadow Replication, compared to existing approaches.

2.1 Shadow Replication

The basic tenet of Shadow Replication is to associate with each main instance a suite of "shadows" whose size depends on the criticality of the application and its performance requirements. Each instance is executed using a process. If the main process fails, a shadow process can continue and finish the task.

The novelty of Shadow Replication lies in its differentiation of the execution rates. Specifically, it executes the main process 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. A closer look at the model reveals that Shadow Replication is a generalization of existing fault tolerance approaches. Specifically, if the QoS allows for flexible completion time, Shadow Replication would slow down the shadow processes and trade time redundancy for energy savings, mimicking Checkpoint/restart. If the target response time is stringent, however, the shadows would execute simultaneously with the main at high rates, mimicking Process Replication. The flexibility of Shadow Replication provides the basis for the design of a fault tolerance strategy that strikes a balance between task completion time and energy saving.

2.2 Analytical model and simulator

One challenge of Shadow Replication 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 an analytical model, from which an optimization problem is formulated to derive the optimal execution rates. The model considers the time and energy needed for a job, under different system specifics and fault distributions.

To verify the correctness of the analytical model, I build an event-driven simulator that simulates the behaviors of Shadow Replication under various configurations. It can report all necessary statistics, such as number of faults encountered, time to completion, and energy consumption. The statistics can then be used to compare with the results from the analytical model.

2.3 Preliminary results

Several important parameters are identified that impact the energy consumption of Shadow Replication. Correspondingly, I conduct a series of sensitivity studies where Shadow Replication is compared to state-of-the-art approaches. The results from both the analytical model and simulator show that Shadow Replication can achieve significant energy savings, without violating the QoS constraints. Specifically, Shadow Replication can achieve 15%-30% energy savings under normal configurations. Furthermore, Shadow Replication would converge to Process Replication, when target response time is stringent, and to Checkpoint/restart when target response time is relaxed or when fault is unlikely [3].

3 Future directions

Current results reveal that Shadow Replication is promising for significant energy saving within QoS constraints. The direct benefits include profit gains for cloud service providers and reduced CO_2 emission, making large-scale computing systems more environment-friendly and more sustainable. Inspired by that, I will be fully committed to solve some challenging questions in the next academic year.

The first plan is to further improve the efficiency of Shadow Replication for tightly-coupled jobs. My current design works well for loosely-coupled jobs, such as MapReduce jobs, where synchronization among tasks is minimized. In a tightly-coupled job, however, even a very short recovery time may be amplified by the frequent synchronizations, resulting in a delay in the job completion time. In order to minimize this effect and further improve performance, I plan to explore the potential benefits of a new approach, referred to as "Leaping Shadows". The idea is to take advantage of the recovery time and align the execution states of the slow shadow processes with their faster main processes to achieve forward progress. Remote Direct Memory Access (RDMA) is a possible way to implement Leaping Shadows, but how to efficiently use it needs further research. I plan to finish this work in three months.

The next research direction is to evaluate the feasibility and performance of using process collocation in Shadow Replication. My current work assumes Dynamic Voltage and Frequency Scaling (DVFS) in controlling the execution rates. The effectiveness of DVFS, however, may be markedly reduced in computational platforms that exhibit saturation of the processor clock frequencies or large static power consumption. An alternative is to collocate multiple processes on a single computing node, while keeping the node running at the maximum rate. Time sharing can then be used to achieve the desired execution rates.

The two alternatives are equivalent in terms of completion time, since they have the same effect on the execution rate control. In terms of energy, however, each of them has its own advantage. Process collocation requires less hardware resources and this reduces the energy linearly, while DVFS uses more hardwares but can reduce energy superlinearly. It needs further analysis to determine which alternative consumes less overall energy. Furthermore, more efforts are needed to study the potential issues with process collocation, such as correlated faults and collocation overhead. This will take approximately three months to complete.

The last and most challenging step is to build a prototype, in order to experimentally evaluate the performance of Shadow Replication using real life applications. This effort includes the design and implementation of a distributed software library that supports the main components of Shadow Replication, including process collocation, required consistency protocols, message logging and message forwarding protocols, and execution state transfer in support of Leaping Shadows. For Shadow Replication to be scalable and efficient, it is necessary to minimize its overhead to the normal execution of the running processes as well as to the operating system. This work will take approximately six months to accomplish.

4 Conclusion

My current research is enabling new insights into the multi-faceted and challenging resiliency problem in large-scale computing platforms. The goal is to investigate radical approaches to the design of scalable and energy efficient fault tolerant schemes that go beyond state-of-the-art algorithms. Throughout my design, the interplay between resiliency, performance and energy consumption will be 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.

To this end, I propose Shadow Replication as a novel, scalable, and energy-aware fault tolerance framework. My preliminary results predict that Shadow Replication is able to achieve significant energy savings while satisfying QoS requirements. I will continue to explore and optimize Shadow Replication in support of green and sustainable computing.

References

[1] B. Mills, T. Znati, and R. Melhem, "Shadow computing: An energy-aware fault tolerant computing model," in *Computing, Networking and Communications (ICNC), 2014 International Conference on*. IEEE, 2014, pp. 73–77.

[2] 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.

[3] 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.