Leaping Shadows: Adaptive and Power-aware Resilience for Extreme-scale Systems

Xiaolong Cui[*], Taieb Znati, Rami Melhem
University of Pittsburgh
Pittsburgh, PA
{mclarencui, znati, melhem}@cs.pitt.edu

Tariq Alturkestani[*†]
KAUST
Thuwal, Saudi Arabia
tariq.alturkestani@kaust.edu.sa

Esteban Meneses
Costa Rica Institute of Technology
Cartago, Costa Rica
emeneses@ic-itcr.ac.cr

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 space redundancy for recovery. Rollback-recovery, which relies on 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 space redundancy and executes multiple instances of the same task in parallel to guarantee completion without delay. This solution, however, requires a significant increase in hardware resources and increases the power 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.

The proposed approach to resiliency, referred to as **Leaping Shadows**, 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 basic tenet of Leaping Shadows is the concept of shadowing, where each process is associated with a "shadow" (process) that dynamically adjusts execution rate. The shadows initially execute at a reduced rate to minimize redundancy cost. Upon failure of a main process, its associated shadow increases its rate to complete the task, thereby reducing the impact of such a failure on the progress of other tasks. The dynamic and differential control of execution rates provides a unique opportunity that allows the lagging shadows to benefit from the faster main processes, without incurring extra overhead. Specifically, when a failure occurs, Leaping Shadows takes advantage of the recovery time and leaps forward the shadows by copying states from the main processes. Consequently, the high probability that shadows never have to complete the full task, coupled with the fact that they initially only consume a minimal amount of power, dramatically increases a power-constrained systems tolerance to failure.

The proposed solution differs from existing approaches 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.

[*]Student
[†]Presenter