Abstract:

It is widely accepted that future extreme-scale parallel computing systems will require alternative methods to enable applications to maintain current levels of uninterrupted execution. As the component count of future extreme-scale systems continues to grow, the likelihood of a failure impacting an application as well as other software components grows as well. Current methods of providing resiliency for applications, such as checkpoint/restart, may likely become ineffective, largely due to the overhead required to checkpoint, restart, and recover lost work. As such, the research community is pursuing several alternative approaches aimed at providing the ability for an application to survive in the face of failures and to continue to make efficient computational progress. This talk will examine resiliency from a system software perspective, examining the opportunities and challenges for application/architecture co-design. We will also present recent work at Sandia studying the feasibility of redundant computation.