Resilience Design Patterns - A Structured Approach to Resilience at Extreme Scale

Achievement: Developed a structured approach for managing high-performance (HPC) resilience using design patterns. Developed a complete specification of resilience design patterns, including a pattern catalog, classification scheme and a design framework to build resilience solutions using the patterns.

Significance and Impact: The resilience design patterns specification provides HPC system architects and application developers with a collection of reusable design elements, and a discipline to combine an essential set of design patterns into productive and efficient resilience solutions. The structured approach enables the systematic improvement of resilience in HPC systems, keeping scientific applications running to a correct solution in a timely and efficient manner in spite of frequent faults, errors, and failures.

Research Details:
• Identified and evaluated repeatedly occurring resilience problems and solutions in HPC hardware/software.
• Developed a catalog of design patterns based on well-understood HPC resilience solutions.
• Codified patterns in a layered hierarchy that classifies the patterns and conveys their relationships.

Sponsor/Facility: This work was performed at Oak Ridge National Laboratory (ORNL). It was sponsored by the US Department of Energy Office of Science Early Career Research Program.

PI and affiliation: Christian Engelmann – Computer Science and Mathematics Division (CSMD), Oak Ridge National Laboratory (ORNL)

Team: Saurabh Hukerikar and Christian Engelmann

Publications:

Overview:
A design pattern describes a generalizable solution to a recurring problem that occurs within a well-defined context. Patterns are often derived from best practices used by designers and they contain essential elements of the problems and their solutions. They provide designers with a template on how to solve a resilience problem that may be used in many different situations. The patterns in the catalog provide solutions for detection, containment and recovery for faults, errors and failures. This work also defined a conceptual framework based on the notion of design spaces that guides hardware and software designers and architects, as well as application developers, in navigating the complexities of developing effective resilience solutions. The framework enables designers to combine the patterns and refine their interactions to create alternative resilience solutions for a specific problem, each with different efficiency and complexity characteristics.