Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle *Since 1965*

Resilient Message Passing Interface for Fault-tolerant Runtimes

CHALLENGE

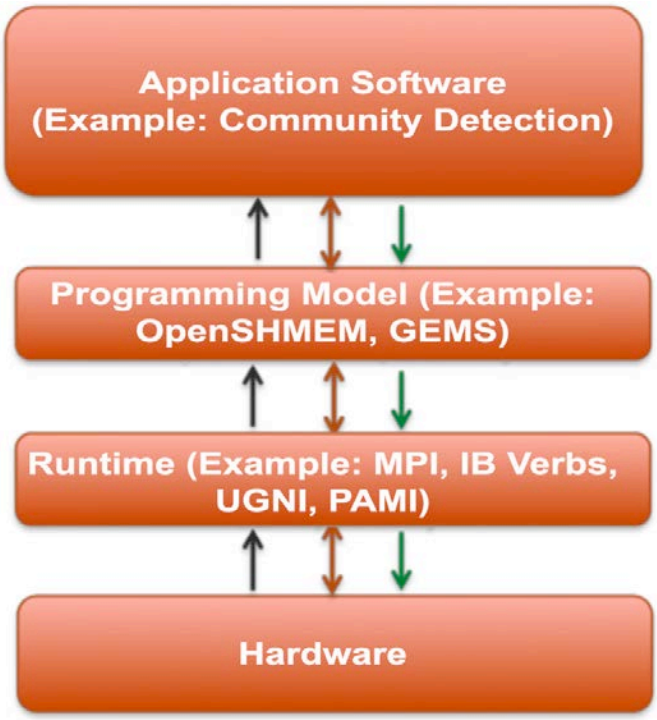
Faults are normal in large-scale systems. These systems suffer from a variety of faults, such as permanent, transient, and intermittent, potentially resulting in application error. The majority of research/software development in fault tolerance has focused on linear algebra operations for faster checkpoint/recovery. Yet, these approaches are insufficient when considering graph query systems such as the Graph Engine for Multithreaded Systems (GEMS), which is the primary focus of this project. Specifically, we observe the following challenges in developing fault-tolerant GEMS:

- The ability to develop data structures that are amenable to asynchronous check-pointing
- Minimizing the cost of recovery (re-establishing the data structures and associated metadata)
- Considering the limitations of existing runtimes, such as the fault-tolerant implementations of message passing interface (MPI) for practical deployments.

CURRENT PRACTICE

Several solutions may be considered for fault tolerance. Map-reduce-based solutions are effective in fault tolerance by automatic check-pointing and replay of the computations. However, recovery cost and limitations of functional programming models are insufficient for graph algorithms, including the one considered in this project. Customized fault-tolerant solutions, primarily

Design fault-tolerant MPI runtimes that provide high-fidelity fault detection and fault tolerance for compute-intensive and graph algorithms.



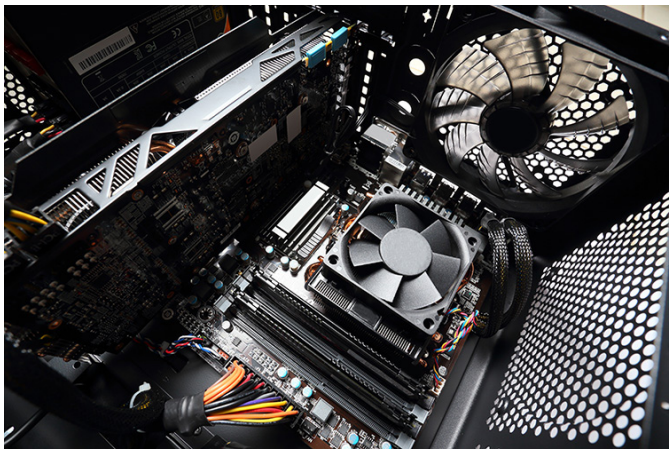
restricted to linear algebra, are targeted toward matrix operations, and are not necessarily applicable to graph query systems (although a few sparse-matrix-based techniques may be used for building customized fault-tolerant graph algorithms). Specific to runtimes, the MPI forum has considered adding extensions for fault-tolerant communication, with user-level fault mitigation (ULFM) as the leading example. However, these techniques have not been integrated in the specification and, as a result, are much less supported by vendors.

TECHNICAL APPROACH

During the project's first year, we considered several approaches for building fault-tolerant runtimes, which would be used as a building block for designing fault-tolerant GEMS. As a by-product of Year 1, we have enhanced an existing version of ULFM (a fault tolerance proposal in consideration with the MPI forum) that can be used for recovery for GEMS and specifically enhanced for high-performance interconnects.

The second year of the project is focused on designing fault-tolerant GEMS-2.0, a revolutionary software architecture for handling graph queries on distributed systems. We are exploring the following techniques for handling queries, which may run for arbitrary lengths of time on very large datasets:

- Techniques to efficiently checkpoint GEMS tables and graph data structures using a Lustre file system and solid-state drives (SSDs)



- Techniques to recover from intermediate results on lost compute nodes, minimizing the recovery time
- Techniques to rectify the incorrect index translation in global address data structures
- Integration of fault-tolerant communication subsystems with fault-tolerant GEMS.

For each of these explorations, there are several choices, such as default (which would recover the precise result, albeit at a higher upfront cost), and approximation algorithms, which would consider best effort techniques for recovering intermediate results.

IMPACT

We consider the impact of the proposed approaches to be significant to handle faults in large systems seamlessly. As a result, an analyst could execute queries on GEMS without worrying about faults. For upcoming years, the project will consider multi-bit transient faults and their potential impact on application errors.

Contacts

Abhinav Vishnu
Principal Investigator
(509) 372-4794
Abhinav.Vishnu@pnnl.gov

John R. Johnson
Program Director
(509) 375-2651
John.Johnson@pnnl.gov

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965