Analyzing Off-Normals In Large Distributed Control Systems Using Deep Packet Inspection and Data Mining Techniques

M.Fedorov, G.Brunton, C.Estes, J.Fisher, C.Marshall, E.Stout

Network packet inspection using port mirroring provides the ultimate tool for understanding complex behaviors in large distributed control systems. The timestamped captures of network packets embody the full spectrum of protocol layers and uncover intricate and surprising interactions. No other tool is capable of penetrating through the layers of software and hardware abstractions to allow the researcher to analyze an integrated system composed of various operating systems, closed-source embedded controllers, software libraries and middleware. Being completely passive, the packet inspection does not modify the timings or behaviors. The completeness and fine resolution of the network captures present an analysis challenge, due to huge data volumes and difficulty of determining what constitutes the signal and noise in each situation. We discuss the development of a deep packet inspection toolchain and application of the R language for data mining and visualization. We present case studies demonstrating off-normal analysis in a distributed real-time control system. In each case, the toolkit pinpointed the problem root cause which had escaped traditional software debugging techniques.

Benefits and Challenges of Network Packet Captures

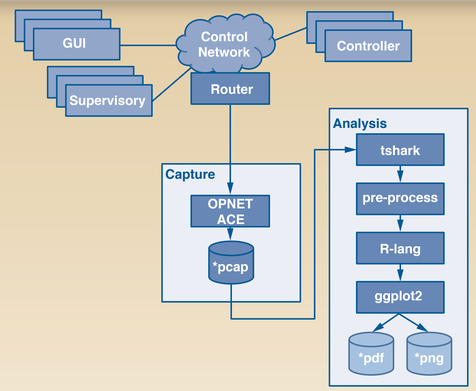
Pros:

- Works across server and embedded platforms
- Captures variety of protocols
- Cuts across network layers
- Passive, does not modify timings or behaviors
- Millisecond-scale accuracy of event timestamps
- Supported by high-end network routers
- Open-source tools are available (tcpdump)
- Commercial tools are available (OPNET)

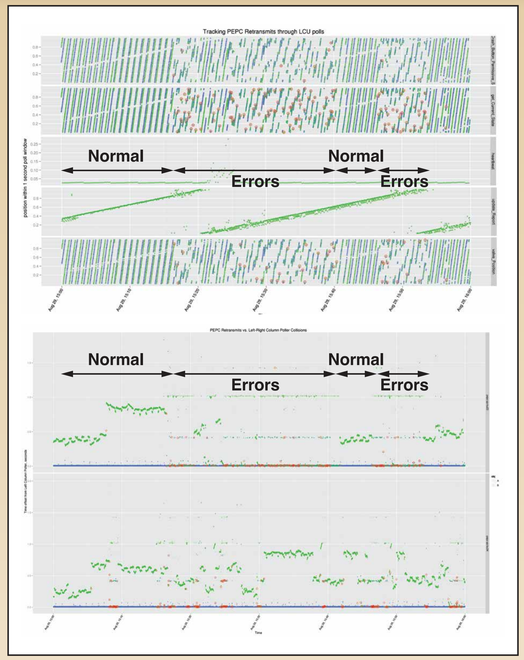
Cons:

- Specialized network hookup is required
- Dedicated capture host is needed
- Large fast storage is required, GBs/host
- Visualization is slow with WireShark or OPNET
- Hard to identify interesting events

Deep Packet Inspection Toolkit

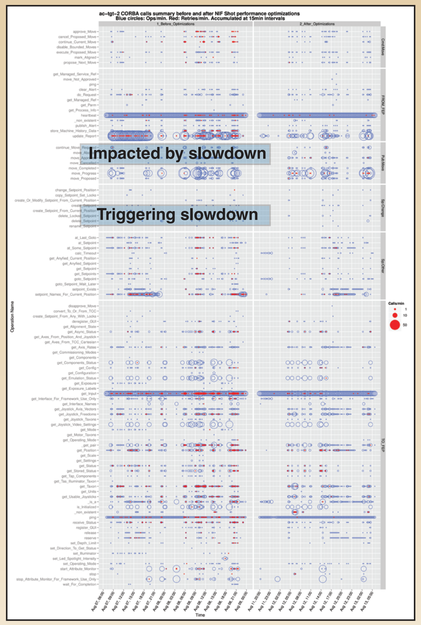


Unwinding Concurrent Interactions



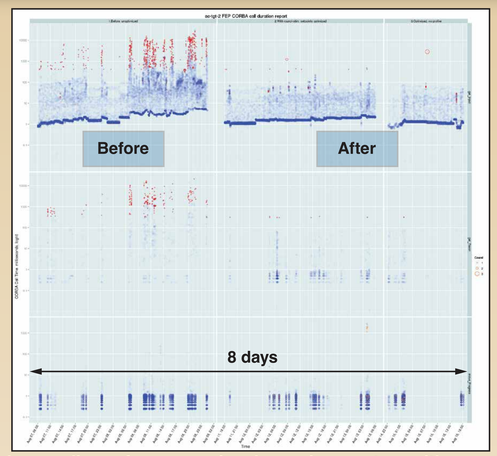
- The Rigid Timeframe Diagram supported initial analysis of periodic interactions
- The Floating Timeframe Diagram with time reference anchored to one of the pollers identified the thread collision problem
- The issue was resolved by interleaving the poll periods

Identification of Distributed Event Patterns



- The API Breadth and Volume Diagram provided an efficient tool for identifying event patterns
- Visualization covered millions of distributed operations over days of operations
- Volume of high-frequency events was accurately represented
- Single and low-frequency events were not lost
- Additional attributes (e.g. TCP retransmits) were overlaid and color-coded

Troubleshooting Sporadic System Slowdowns



- Analyzed sporadic performance slowdowns in a system with embedded controller
- The controller performed correctly all the time, but with intervals of extremely slow performance
- During the slowdowns, the network response times were 10x-100x slower than usual
- Increased TCP retransmissions were seen during the slowdowns
- This visualization helped to identify the root cause, enabled resolution and verified the fix

Summary

- Developed a data analysis and visualization toolkit for network packet captures
- This toolkit effectively processed raw network captures of several gigabytes per host
- The R language was used to generate informative and intuitive visualizations
- The toolkit was used to understand and address performance and timing issues in a large distributed control system

P279394

LLNL-POST-644258

Lawrence Livermore National Laboratory

National Ignition Facility • Lawrence Livermore National Laboratory • Operated by the US Department of Energy
This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

NIF