STAC-M1 v2 (beta) Methodology

[STAC-M1 v2 is currently in beta testing. It was first used in a project to test the SR Labs MIPS system. The following is an excerpt from the recent STAC Report on two SR Labs "stacks under test" (SUTs): MIPS110612a and MIPS110612b, entitled "SR Labs MIPS 3.6.7 Feed Handler Library on Intel X5687 (Westmere) and Intel X5698 (Everest) Processors with Myricom 10Gbps NIC and DBL 2.0 in the IBM x3650 M3". This report is available in the STAC Vault.]

This project used benchmarks from the draft STAC-M1 v2 Benchmark specifications. STAC-M1 provides a framework for consistent, reproducible testing of low-latency feed-handling solutions. The test specs make no assumption about the architecture of the stack under test (SUT)—whether FPGA vs CPU, in-process library vs distributed system, Ethernet vs InfiniBand, etc.

The STAC-M1 specifications draw from the input of leading trading firms and vendors on the STAC Benchmark Council (www.STACresearch.com/council). STAC-M1 v2 is currently in a beta stage, in preparation for submission to the Council for approval.

In a STAC-M1 test, the inputs are recorded raw exchange messages played back via hardware at multipliers of their original rate. The outputs are market data updates normalized to a specified structure in the main memory of a server. The tests yield numerous metrics related to latency, throughput, power consumption, and other key performance indicators. Each test sequence yields statistics on the one-way latency from the time that an exchange message is available to the SUT to the time that normalized data is available to a Consumer application.

Addenda can be attached to STAC-M1 to define variants of the tests specific to particular datafeeds. The test harness in this project uses a 1 Gbps UDP-based NASDAQ TotalView ITCH 4.1 data replay as its source, which it plays back at multiples of the recorded data rate while preserving the natural 'micro-burstiness" of the traffic. The harness controls the watchlists, field types, and other application particulars in accordance with usage patterns that trading firms have indicated are common for automated equities trading in the US. While these particulars vary from application to application in the real world, STAC-M1 tends toward the patterns that yield the most conservative performance measurements.

STAC-M1 v2 requires integration with a latency-monitoring system, which is responsible for capturing data on the wire and in the application (through an implementation of the STAC Observer API), synchronizing these, and correlating observations for each message to compute pair-wise latencies. The STAC-M1 v2 harness is open for integration with any latency-monitoring solution that can meet its requirements.

Figure 1 illustrates the test harness used in this project. Recorded data plays at various rates, and the SUT is responsible for delivering normalized data to the test client. Update latency is shown as Δt, which is from the availability of raw exchange data on the wire to the availability of consumable, normalized data in the test client.

Copyright © 2011 Securities Technology Analysis Center LLC

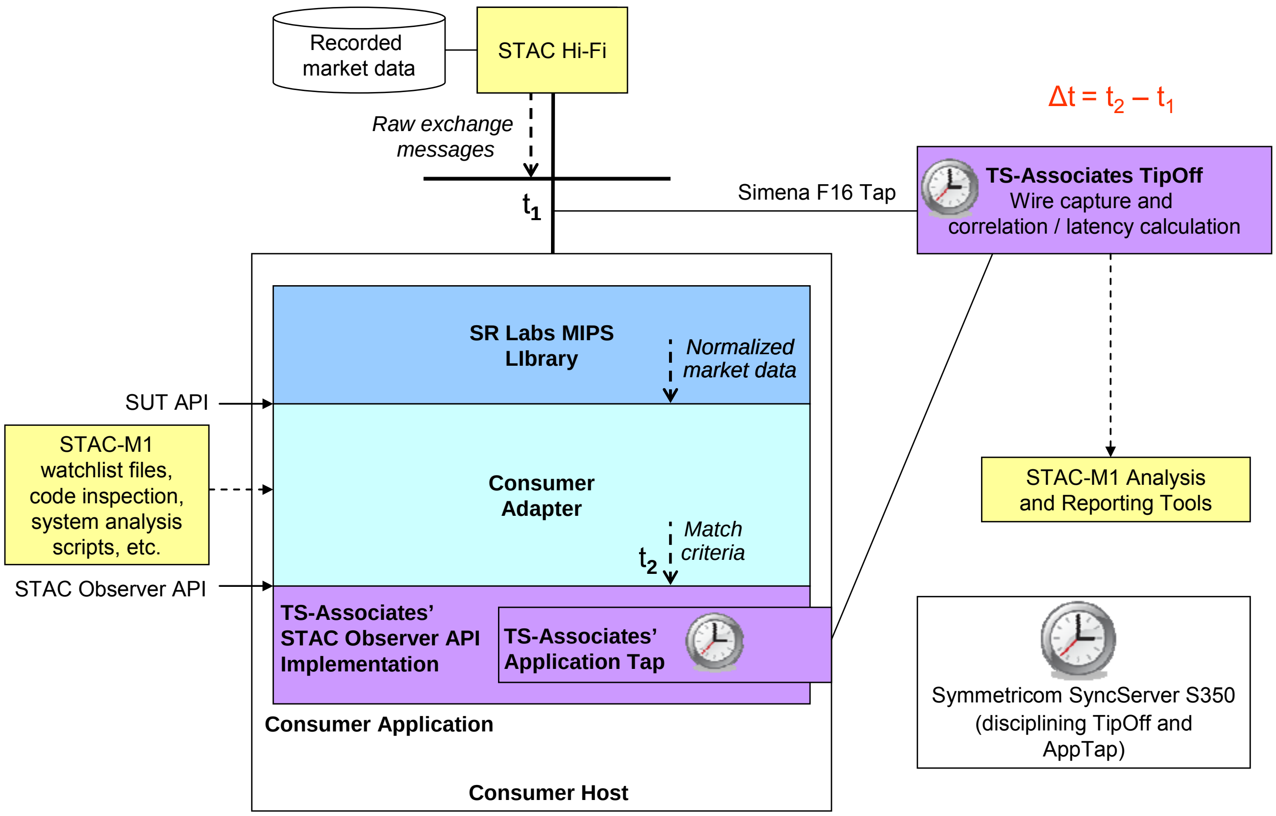

Figure 1 – The test setup used in this project

Components in either shade of blue were supplied by SR Labs. These are the MIPS library and the Consumer adapter, which binds the MIPS library to the STAC Observer API.

Components in yellow were supplied by STAC. These components are responsible for providing a repeatable workload to the SUT, validating required SUT functionality, and analyzing and reporting results that conform to the benchmark specs. Input data comes from STAC Hi-Fi, which leverages proprietary hardware to replay pre- recorded data with precise fidelity to the timing of the original datafeed stream. The STAC-M1 Observer API specification/header file specifies how the Consumer adapter interfaces with the code responsible for timestamping and persisting observations.

Components in violet were supplied by TS-Assocates, the provider of the monitoring solution for this project. The monitoring solution correlates wire-capture observations at the head end with in- process observations from a library embedded in the Consumer, calculates per-update latencies, and supplies the un-summarized latency information to STAC tools which analyze and report results in conformance with the benchmark specs. The TS-Associates products in this harness are TipOff and the Application Tap.

TS-Associates supplied the following information about these products

> *The Application Tap® is a PCI Express (PCIe) card that enables the implementation of precisely timed software instrumentation with minimal performance overhead. By combining a user mode, kernel bypass API, with an on-board FPGA co-processor and high-precision clock, the Application Tap delivers hitherto unavailable insights into the performance dynamics of real time applications. This unique software precision instrumentation tool is suitable both as an enabler of production systems monitoring, and also for a range of software performance optimisation tasks such as deterministic profiling.*

Copyright © 2011 Securities Technology Analysis Center LLC

The Application Tap solves the challenge of minimally invasive software instrumentation by providing a user mode API that enables software events within applications to be instrumented with minimal overhead. The Application Tap API is integrated with custom developed firmware running on a state of the art FPGA co-processor using OS kernel by-pass techniques. Instrumentation code added to applications is able to pass instrumentation metadata to the Application Tap which time stamps events to 10ns resolution, and forwards the instrumentation metadata to an external monitoring appliance, such as TipOff®, using an on-board gigabit network interface.

TipOff® is a real time middleware analysis and passive latency monitoring appliance that delivers Precision Instrumentation across all standards based and vendor proprietary middleware stacks. TipOff supports passive latency monitoring of market data and transaction flows. By decoding packet streams in real time through all layers from network up to application message content, TipOff can monitor both packet and message latency, accounting for retransmission latency in the reliability layer. A single multi-hop flow may comprise several different middleware stacks or protocols. TipOff supports flows that undergo transformations through order aggregation, trading engines and execution venues. TipOff is able to republish latency and other middleware statistics as a real time meta-data feed enabling trading applications to share a common view of data latency by leveraging their existing market data API connection. TipOff supports a range of both push and pull API for integration with other monitoring tools and products.

The Application Tap is fully integrated with TipOff. This enables correlation of precisely timed software events with network events, providing hop by hop latency visibility, not just at the server level, but right down to the software component level - processes, threads, even code blocks. As legacy distributed systems consolidate down to multi-core servers, and traditional network tap based latency monitoring solutions lose visibility, the Application Tap becomes an essential tool in the precision instrumentation repertoire.

A Simena F16 Fiber Optic Tap monitors the network link into the SUT to observe inbound messages, passing them to TipOff. In the Consumer Host, a TS-A provided implementation of the STAC Observer API accepts identifying information ("match criteria") from the Consumer adapter, provides them to the Application Tap, which timestamps them and transmits the information to TipOff. TipOff then correlates the wire-based and API-based observations to calculate latency on a per-message basis.

Simena supplied the following information about the monitoring switch:

Passive, easy to install, fault tolerant and transparent to the network, the F16 Fiber Optic Tap offers an efficient and cost effective solution for IDS, SNMP, RMON, and Protocol analysis. Since it is transparent to the network there is no point of failure, no network downtime, and 100% security. The F16 operates at maximum wire speed, monitoring half/full duplex traffic and is capable of monitoring and copying all seven (7) protocol layers and any packet size, including errors.

All Simena Fiber Taps are "all-optical", data rate independent and designed for use in 10/100/1000 and 10Gb Ethernet, ATM, SONET, FDDI, etc networks. High quality, low loss Zirconia sleeve adapters are used in both Multi- and Single-Mode models. The F16 has three (3) duplex LC connectors in each network link – two (2) for the network and one (1) for the analyzer. The unit is compact, housed in a rigid metal enclosure, and can be ordered in Network/Tap ratios of 50/50, 60/40, 70/30, 80/20, and 90/10.

Symmetricom supplied the following information about the time server:

The Symmetricom SyncServer® S350 GPS Network Time Server supporting PTP and NTP is the ideal solution for synchronizing the time on servers for high frequency trading systems or large and expanding IT enterprises. All S350 SyncServers are factory ready for high accuracy, hardware based PTP time stamping. When enabled, the PTP Grandmaster functions are very easy to configure via the web interface, and the PTP protocol begins immediate operation. The PTP features are a superset of the default profile as *defined in the IEEE 1588 2008 standard. Additional control is provided over key PTP protocol parameters to accommodate unique network topologies and optimize PTP slave accuracy requirements.*

The S350 SyncServer actively monitors PTP slave access and Grandmaster performance thereby removing any uncertainty about slave connectivity and grandmaster loading. Grandmaster performance metrics provide essential insights into PTP activity load on the SyncServer. Performance charts available from the web interface provide visual insights into the history of key PTP time stamping factors over time.

PTP slave accuracy testing is an essential part of any high-accuracy PTP deployment. It's a fact that network induced delays degrade PTP slave accuracy. 1PPS Time Interval measurements built into the S350 will measure in real-time the accuracy of hardware based PTP slaves compared to the S350 clock. The 1PPS output from the slave is connected to the S350 and from there a full range of slave accuracy statistics are computed and displayed via the web interface in real-time.

Real-time charting of slave 1PPS accuracy via time history or histogram is also available via the web interface. While measurements are being made the user configurable charts will continuously present updated results. The SyncServer S350 in particular provides a unique and expanded set of testing and analysis capabilities of hardware-based PTP slaves. In applications where slave accuracy is critical, the ability to measure that accuracy from the source is essential.

Visit www.symmetricom.com for more information.