A Solution to Optimize Pattern Matching Performance Using Sensory Networks HyperScan™ Software and Intel® Xeon® Processor 5500 Series

The multi-threaded architecture of Sensory Networks HyperScan™ software enables multiple instances of the pattern-matching executable, perhaps utilizing different signature sets, to run simultaneously. It should be used on a multi-core hardware platform in order to maximize performance.

But how many cores are needed? What are the minimum platform requirements that will enable equipment designers to forego dedicated hardware acceleration? What kind of extra, performance-enhancing features are needed to optimize pattern matching and deliver better overall data security?

In an effort to answer these questions and identify the ideal platform for their software, Sensory Networks prepared a series of benchmark tests using different hardware platforms. They selected multi-core Intel® architecture platforms for all of the tests because it was available in multiple variants for number of cores available, performance levels, and feature sets. That coupled with micro-architectural level innovations, rich instruction set architecture, and a fully developed software toolkit enabled Sensory Networks to perform the necessary tests evenly across multiple systems.

While the HyperScan software performed well on Intel’s original quad-core platform, it scaled significantly when run on the new Intel® Xeon® processor 5500 series (formerly code-named “Nehalem”). In fact, the tests showed the new platform performing 22 percent faster (clock-for-clock) than the previous generation quad-core processor, and an 18 percent aggregate improvement in throughput performance.

This solution study explains the benchmark tests and describes how Sensory Networks HyperScan software running on a platform with the Intel Xeon processor 5500 series can deliver the necessary performance to maintain line rates when running security applications.
**Background: Security Challenges**

Application data continues to grow on a massive scale, as does the importance of that data to its users. At the same time, the task of protecting that data is increasingly complex and challenging. Threats today come in many varieties and seek to infiltrate systems through any means possible. The entire network operation is vulnerable, including the infrastructure, applications, data, and all of the traffic that comes through.

Deep packet inspection—the process of checking all data against known signature sets that signal the presence of malware—provides the most effective means for protecting data and thwarting attacks. This action must be performed on all traffic, even though it is a process-intensive task that consumes massive compute resources, IT personnel and infrastructure.

The challenge for IT managers, equipment designers and security software vendors is to develop solutions that can maintain line rates while performing deep packet inspection on the ever-increasing volume of network traffic.

**An Idea Worth Testing**

Sensory Networks developed its software-based pattern matching solution with the goal of maximizing performance. It is multi-threaded by design, enabling simultaneous scanning of multiple patterns. That means Sensory Networks HyperScan pattern matching software can more quickly process a larger volume of network traffic than many traditional linear solutions.

But it is a very process-intensive program that requires some form of dedicated hardware so that it doesn’t interfere with routine tasks. Although the traditional means of adding hardware acceleration or dedicated security appliances can provide the necessary performance, this solution tends to be costly and inefficient.

Sensory Networks suggests there is a more efficient solution: its multi-threaded software running on a high-performance multi-core hardware platform. This, so that developers can dedicate one or more entire cores to the job of packet inspection and therefore deploy more sophisticated network security applications on a single system. As posited, this theory implies that such a multi-threaded software/multi-core hardware solution can consolidate and simplify hardware design and reduce the number of tools needed for implementing deep packet inspection solutions. The result: reduced costs for both IT managers and developers.

To test this hypothesis, Sensory Networks sought a multi-core hardware architecture capable of running its pattern matching software applications. Their engineers chose multi-core Intel architecture because it offers a number of variants in terms of core count, performance levels, feature sets, and built-in security support.

The theory to be tested: Running the Sensory Networks HyperScan software on an Intel architecture-based system speeds delivery of security applications to the network and helps maintain line rates. The combined software/hardware solution can then obviate the need for specialized acceleration hardware and create more efficiency in the network.

But obvious questions remained:

1. How many cores were needed to provide the optimum performance for the Sensory Networks software?
2. Can performance-enhancing technologies in the hardware improve software performance without rewriting code?
3. Is there a point at which the software ceases to benefit from added cores?

**Benchmark Test Configuration**

Sensory Networks conducted performance benchmarking tests for its HyperScan library running on multiple, multi-core Intel architecture platforms, including the new Intel Xeon processor 5500 series platform at 2.93 GHz. The test configuration included the HyperScan software library and a small benchmarking application installed on a dual-socket (eight-core) Intel platform. The test workload was comprised of HTTP-based test traffic using a complete set of IPS signatures sourced from a leading security equipment vendor. (See signature set code descriptions in Table 1.)

<table>
<thead>
<tr>
<th>Signature Set Code</th>
<th>Streaming</th>
<th>Signature Set Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vendor A, http to-client traffic 1</td>
<td>Y</td>
<td>69 complex</td>
</tr>
<tr>
<td>Vendor A, http to-client traffic 2</td>
<td>Y</td>
<td>142 complex</td>
</tr>
<tr>
<td>Vendor A, http to-server traffic 1</td>
<td>Y</td>
<td>43 complex</td>
</tr>
<tr>
<td>Vendor A, http to-server traffic 2</td>
<td>Y</td>
<td>235 complex</td>
</tr>
<tr>
<td>Vendor A, http to-server URI-only traffic 1</td>
<td>N</td>
<td>13K medium</td>
</tr>
<tr>
<td>Vendor A, http to-server URI-only traffic 2</td>
<td>N</td>
<td>8K medium</td>
</tr>
<tr>
<td>Vendor B, http to-client traffic</td>
<td>Y</td>
<td>67 complex</td>
</tr>
</tbody>
</table>

Table 1: Signature Set Codes used in the Sensory Networks benchmark tests.

Captured HTTP traffic (from a PCAP capture file) was passed through the HyperScan software library. This data was scanned packet-by-packet, simulating the behavior of a real network application such as an IPS or a web proxy appliance. Data was matched in ‘streaming mode’ for cases where the threats might be spread across multiple packets, and in ‘non-streaming mode’ for threats that would be contained within a single chunk of data.

Patterns provided from the security vendors included multiple variants of both ‘to-client’/’to-server’ and URI signatures. All signature sets were compiled into their runtime database within four seconds. It should be noted that the benchmarking application measured the time spent scanning. All data used for inspection/pattern matching was resident in memory to perform the tests, thus no I/O interfaces were required.
Benchmark Results

Throughput Performance

Total HyperScan throughput performance scales from 9Gbps (1 thread) to over 73Gbps (16 threads) depending on the signature workload for the platform used in this benchmarking. When used with Sensory’s HyperScan software on a dual-processor platform, the Intel Xeon 5500 processor performs 22% faster clock-for-clock and slightly better near-linear scaling than a dual-processor, quad-core Intel Xeon 5400 processor.

Performance Scalability

HyperScan achieves near-perfect linear scalability from 1-8 threads, and an 18% improvement in aggregate throughput performance by leveraging the Intel® Hyper-Threading Technology available on the Intel Xeon 5500 series platform. This additional gain in performance is directly attributed to the hyper-threading capabilities, an architecture that increases application parallelism. It essentially gives the software 16 logical cores on which to run simultaneous threads.

The ‘net’ performance improvement that may be realized in a real world deployment could be much higher due to the fact that HyperScan is a very compute-intensive application that makes high use of available CPU resources – interleaving 8 additional HyperScan threads with the 8 original threads can be representative of a ‘worst case’ environment due to the contention for CPU resources, as compared to interleaving threads that perform a mixture of HyperScan and IPS activity.

Sensory Networks HyperScan Software – A Deeper Look

Sensory Networks’ HyperScan software library is highly portable, OS independent and easy to integrate into network security solutions. With rich regular expression (PCRE) support, the engine supports several specialized content scanning engines. Each engine is unique and tailored for a specific task to deliver excellent content scanning performance for a wide variety of signatures, expressions and applications. Further specialization allows HyperScan to take advantage of specific features found on most Intel processors, such as vector extensions.

HyperScan linearly scales throughput performance along with the number of CPU cores to provide a pattern-matching solution that leverages cache memory and is capable of scanning tens of thousands of patterns simultaneously. HyperScan compiles sets of patterns into an optimized database and then compares them to data that arrives in large discrete buffers or as packetized data streams. (Both streams of packets and complete objects can be scanned in this way.) All regular expressions in the signature set are simultaneously scanned and matches are returned to the application as they are found.

Product Highlights Include:

- **Full-featured regular expression matching**
  - Support the range of ‘classic’ regular expression constructs as well as enhancements (bounded repeats, anchors, etc.)
  - Orders of magnitude higher performance than libPCRE
  - Match thousands of patterns in parallel
  - Support ‘streaming’ (match regular expressions across packets without holding on to old packet data)
High performance
- Up to 20Gbps/core, (pattern dependent)
- Low latency (especially compared to hardware)

Low overhead
- Compile time, bytecode size, stream state size are competitive

Easy to integrate
- HyperScan is compatible on all generations of Intel architecture, making it a safe investment
- Validated to run on several operating systems
- Scales to run on multiple cores and utilizes Intel Hyper-Threading technology

Multi-Core Intel® Architecture – A Scalable Roadmap
Intel has a long track record of delivering high-performance multi-core processors for use in a variety of infrastructure environments. In fact, Intel architecture is considered one of the most popular server architectures in the world and is known for reliability, scalability, and advanced technologies that enhance performance.

The newest processor family in the multi-core Intel architecture roadmap is the Intel Xeon processor 5500 series – formerly codenamed “Nehalem.” It can automatically adjust performance and power usage in real time to meet the computing requirements of the workload.

Servers based on the Intel Xeon processor 5500 series deliver greater performance while reducing energy consumption, footprint and operating costs – making them ideal for use in demanding data centers where large data volumes require strong support for deep packet inspection solutions. Among the important new features in Intel’s new multi-core processor family:

- **Automated Energy Efficiency** – scales energy usage to the workload to achieve optimal performance/watt and reduce operating costs.

Flexible virtualization – offers best-in-class performance and manageability in virtualized environments to improve IT infrastructure and reduce costs.

Performance-Boosting Technologies – that optimize and increase performance to match the workload, especially for multi-threaded applications. These technologies include:

- **Intel® Turbo Boost Technology** automatically kicks in when needed for demanding workloads and allows the processor to run faster than the marked frequency if the part is operating below power, temperature and current limits. Maximum Turbo Boost frequency is dependent on the number of active cores.

- **Intel Hyper-Threading Technology** delivers greater throughput and responsiveness for multi-threaded applications through parallel processing of multiple instructions – even on a single core. Four cores, therefore, can appear and run as eight cores to an application.

Conclusion
Networking security equipment vendors are under pressure to speed up development cycles and spin new products that address emerging threats and the growing demand for performance. They are looking for agile platforms that cost-effectively provide predictable performance, scalability and high levels of flexibility. Meanwhile, IT departments need high-performance networking solutions that operate for less cost within a smaller footprint. Together, these two factors are making multi-core processing the platform of choice for networking and security vendors.

Multi-threaded software solutions in combination with multi-core hardware platforms enable vendors to easily integrate, accelerate and enhance network throughput while improving data center efficiency. The combination of Intel architecture and Sensory Networks’ scalable high-performance libraries delivers a high-performance, optimized solution for multi-gigabit network security tools that can truly scale into the next generation network.

For more information
http://www.sensorynetworks.com
http://www.intel.com/technology