Intel® SOA Expressway Performance Comparison to IBM* DataPower XI50

Intel SOA Expressway easily outperforms IBM* DataPower XI50 for real-world workloads at a fraction of the cost.

White Paper

Intel® SOA Expressway

Performance Comparison to IBM* DataPower XI50

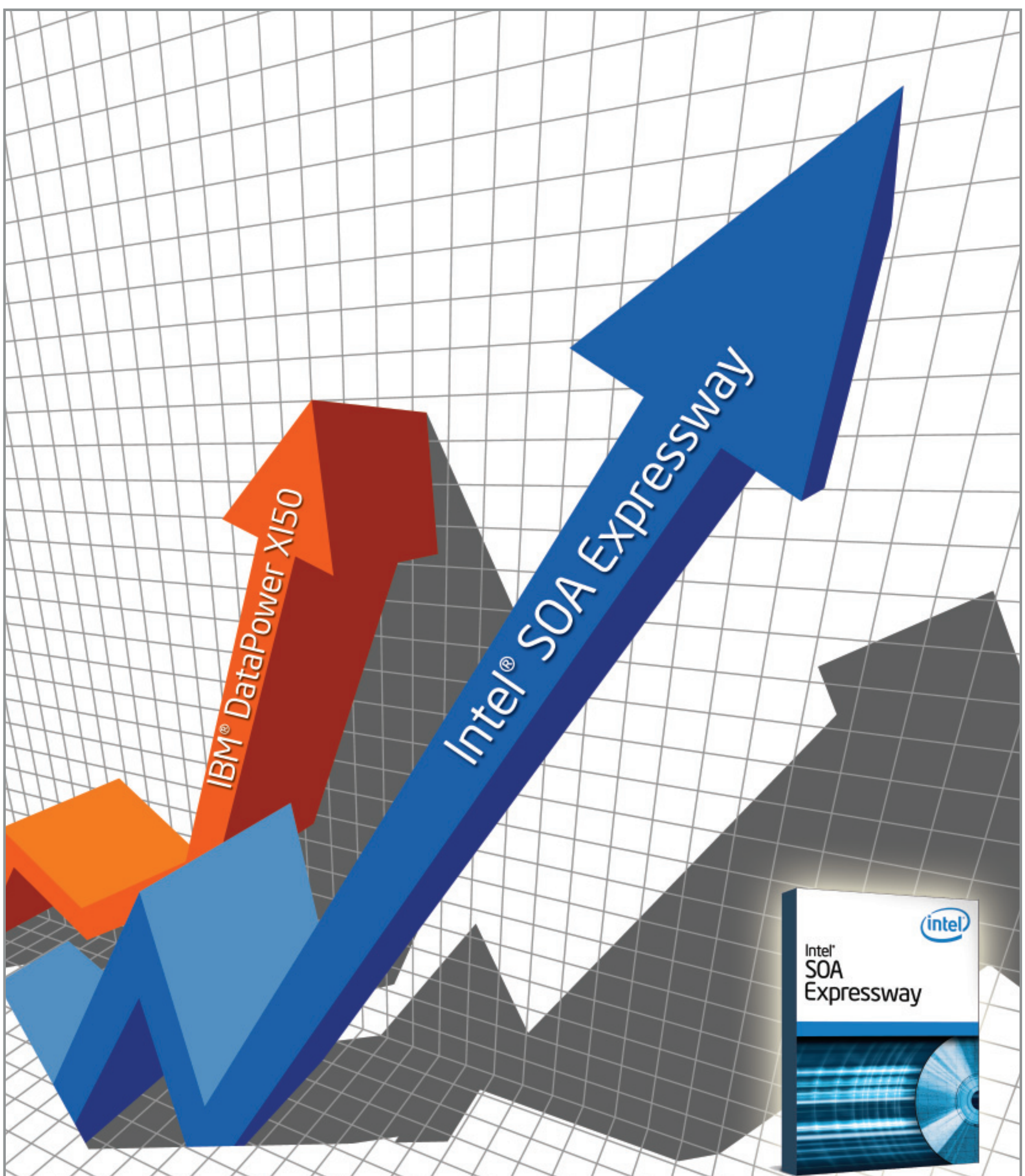


Table of Contents

Executive Summary

Hardware appliances such as the IBM* DataPower XI50 are typically deployed in datacenters to solve integration problems around SOA security, governance and mediation. While these types of highly specialized hardware appliances may solve a few specific problems, they are slow to evolve, difficult to repurpose, and incur higher maintenance and operational costs as compared to general purpose server software solutions. Furthermore, hardware appliances are at odds with major cost-saving datacenter trends such as multi-core and virtualization. As financial constraints on IT become more severe, new products are needed to avoid the compromise inherent in a hardware appliance between performance, and the flexibility and extensibility otherwise available in software solutions. In this new environment, a multi-core optimized software appliance or service router (SR) is superior in offering high performing governance, integration, and mediation capabilities at a lower total cost.

Intel's Software & Services group has released Intel® SOA Expressway (SOAE), a highly tuned, multi-core optimized software service router that provides the full SOA governance, integration and mediation functionality with an Eclipse based user interface. It makes SOA deployment on multi-core environment significantly more scalable and manageable. Just like hardware appliances, Intel's SOA Expressway is deployable in minutes rather than days and requires zero coding for common service intermediary patterns. SOA Expressway offers mediation capabilities that can deal with heterogeneous service and legacy environments and different messaging patterns supported in service interactions. SOA Expressway also offers common shared security enforcement for XML threat protection (XML Firewall), authentication, authorization, access control and auditing. It can be deployed with its Eclipse based governance infrastructure or can be used in conjunction with other design-time SOA governance solutions from major software providers.

Intel SOA Expressway provides superior performance for high service mediation and governance use cases outpacing IBM DataPower in a direct "apples-to-apples" comparison by 1.5X to 8X at a fraction of the total cost.

In this paper we provide:

- A description of the SOA Expressway product and architectural elements and how they are continuously taking advantage of Intel processors, platforms and technologies in delivering unparalleled performance.
- Results demonstrating the superior performance of SOA Expressway as compared to IBM DataPower XI50 for service mediation and service governance use cases.
- A cost comparison showing the advantage of deploying Intel SOA Expressway as compared to the IBM DataPower XI50.

1. SOA Expressway Overview

Intel SOA Expressway is a software-appliance designed to simplify, accelerate, and secure the Enterprise SOA architecture. It expedites SOA deployments by addressing common SOA bottlenecks – it accelerates, secures, integrates and routes Web Services, XML and legacy data in a single, easy to manage form factor.

- SOA without Compromise: Simplify SOA infrastructure with specialized software-appliances with no compromise on extensibility or virtualization.
- Cost Reduction: Address common SOA bottlenecks on standard Intel Multi-Core servers – accelerate, secure, integrate and route Web services, XML and legacy data without the need for purpose-built appliances or multiple application servers.
- Simplify, Simplify, Simplify: Eliminate disruption to existing IT infrastructure caused by rigid governance, operational overhead and lack of virtualization typical of hardware appliances. SOA Expressway contains everything needed for a high performance services mediation layer in one software form-factor.
- ISV and End-User Ready: Available both as a software solution through Intel and its resellers as well as a white-box for independent software vendors (ISV)

The technology behind Intel's SOA Expressway is based on 8-years of research and development both within Intel and through the acquisition of Sarvega, Inc and has been field-tested in some of largest security and performance-sensitive deployments of distributed SOA design patterns.

2. Performance Architecture

2.1 Intel® Multi-Core Optimizations

Intel SOA Expressway provides a software engine that is highly optimized for Intel Multi-Core architectures, including Intel® Core™2 and Intel® Core™ i7 based servers. SOA Expressway helps bridge the gap between multi-core optimizations and business software by providing a runtime layer that efficiently scales with the Intel processor roadmap, bringing Moore's law to business computing without requiring special programming.

2.2 Efficient XML Processing

Intel SOAE provides support for XML acceleration by building a number of highly optimized XML processing engines using an efficient binary representation called Event-Stream-II (ESII). ESII scales effectively with Intel Multi-Core architectures and platform enhancements without any special hardware. With this set of XML processing engines, SOAE executes XML operations such as XPath evaluation, XSLT, XML Schema validation, XML-Content attack filtering, XML content-based routing, and XML/WS-Security (including canonicalization) at wire speed, removing traditional XML processing bottlenecks.

2.3 Efficient Service Mediation Engine

While XML optimizations form an important cornerstone of SOA Expressway's performance characteristics, they are only a portion of the performance equation. Aside from point optimizations such as faster transforms (XSLT) or faster schema validation (XSD), the real problem is finding a good way to mediate between the services that form the underlying components of the larger service or distributed application. For instance, any gain in XML acceleration may be offset by inefficiency in the many layers involved in a large service application with respect to application data I/O, either between local or remote services. To solve this overhead problem, SOA Expressway offers a light-weight service mediation engine that interprets message policies in byte-code form with a zero-copy data interface. This design allows for efficient scaling across multiple threads and allows SOA Expressway to offer predictable, linear scalability for thousands of incoming connections.

3. Performance Testing

3.1 Methodology

Our testing methodology involves use cases with real-world applicability rather than synthetic benchmark that only highlight a few peak numbers. As such, we have designed an environment to tests three aspects of each service intermediary: (a) Messages per Second, (b) Latency, and (c) CPU Utilization. The first metric measures how many transactions the intermediary can process per second and is sometimes abbreviated MPS or TPS. The second metric is the time interval in which a client request is received and responded, and the third metric is an indicator of how much headroom is remaining on the system – this is commonly presented as a percentage as either CPU idle or CPU used.

While a very high peak number may look impressive, what really matters is the interplay between these three metrics to achieve a balanced, predictable and stable behavior under heavy load. It is not enough to look at a single metric without considering the other metrics. For instance, the total throughput (TPS) must be qualified against CPU utilization – if the CPU utilization is low, it may be possible to increase throughput further by pushing more clients threads with no degradation in perceived latency. Also, if CPU utilization is high or nearly maxed out we would expect to see predictable latency (the time taken to complete a request) to increase as the load increases. Similarly, CPU utilization can be compared against latency. If two products have similar latencies between requests and one has more CPU headroom, the product with more headroom may be able to handle increased throughput while maintaining client service-levels.

To achieve this balance, the test environment was designed to minimize network overhead while at the same time ramping up the number of requesting threads in a gradual manner to determine the optimal balance for all three metrics.

3.2 Test Environment

The test environment was constructed as follows:

Figure 1: Environment Topology

The service intermediary represents the product under test, which was either the IBM DataPower XI50[1] or one of the versions of Intel SOA Expressway (SOAE) as described in the following Table 1:

Table 1

Intel® SOAE Version	CPU Configuration	Memory	Manufacturer	Operating System
V2.0	Dual Socket Intel® Xeon® CPU X5560 @ 2.8Ghz	6x2GB 1333 DDR3 RDIMMs (12G total)	Supermicro	Red Hat Linux AS4
V2.0	Dual Socket Intel® Xeon® CPU E5450 @ 3.00GHz	4x2GB FBD PC2-5300 (8G total)	HP ProLiant DL360 G5	Red Hat Linux AS4

The tests were run on an isolated network using a 1Gbit network switch connecting the client machine, intermediary, and back-end server. The client is a custom-built generic HTTP client designed to scale to a large number of worker threads and the back-end server is a custom-built generic HTTP web server also designed for low latency and high volumes. Both the client and server are designed to take minimal memory and CPU to allow the service intermediary to scale and not become the bottleneck. For example, the web server simulates a service response and returns a predefined canned response – no actual work or service call is executing on the backend server.

[1] IBM® DataPower XI50 Firmware Rev: XI50.3.6.1.5, Build: 156262

[2] For tests that required a private key, we ensured the key was loaded into the HSM (Hardware Security Module) to obtain the best performance rather than flash memory.

3.2 Test Execution and Setup

In order to gauge the behavior of each system under load, we used ramping style tests that measured the latency and throughput at various concurrent client thread levels spanning from low single digits to a maximum of 450 threads. The IBM DataPower appliance we configured a single XML Firewall[2] application with the application probe disabled. For Intel SOA Expressway we configured a workflow application using the Intel® Services Designer GUI. No special tuning was done on the Intel SOA Expressway and in both cases each product used a single NIC interface.

The next two sections describe the processing steps for the two main cases: service mediation and service governance.

3.2.1 Service Mediation

The service mediation use case involves a SOAP v1.1 purchase order message sent over HTTP that undergoes XML schema validation, XSL transformation and endpoint routing. The transformation alters the value a single node (shipping address) in the purchase order to simulate service mediation from one format to another. The response from the back-end server is then schema validated for grammar correctness and returned to the calling client. Figure 2 shows a graphical illustration of the service mediation pipeline.

3.2.2 Service Governance

The service governance use case involves applying a security policy at runtime on the response leg. In this case, the same SOAP v1.1 purchase order was sent to the intermediary for schema validation. The response from the back-end server includes a schema validation step, a WS-Security signature over the SOAP Body and then a WS-Security encryption step over the SOAP body. The signature algorithm used was 1024-bit RSA w/SHA-1 and the encryption algorithm was 3DES-CBC with a 1024 bit RSA key-wrap. Figure 3 shows a graphical illustration of the service governance pipeline.

Figure 2: Service Mediation Workflow

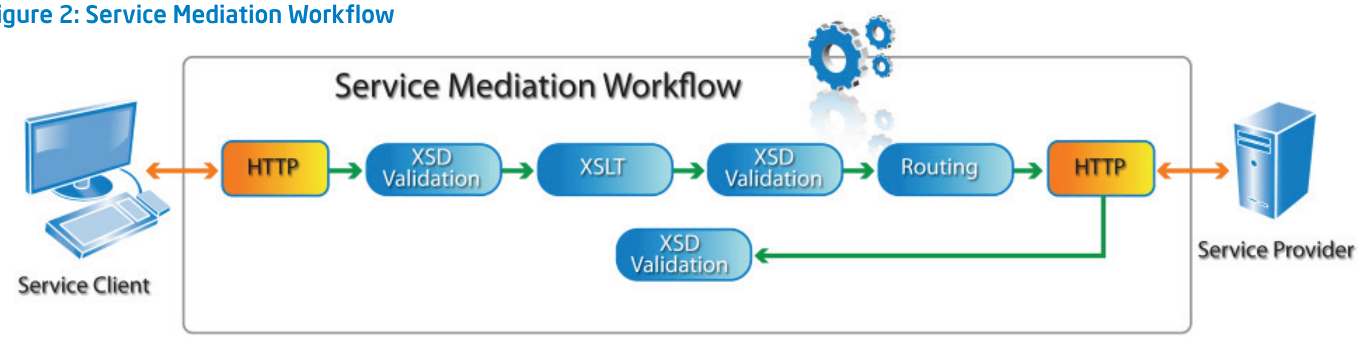


Figure 3: Service Governance Workflow

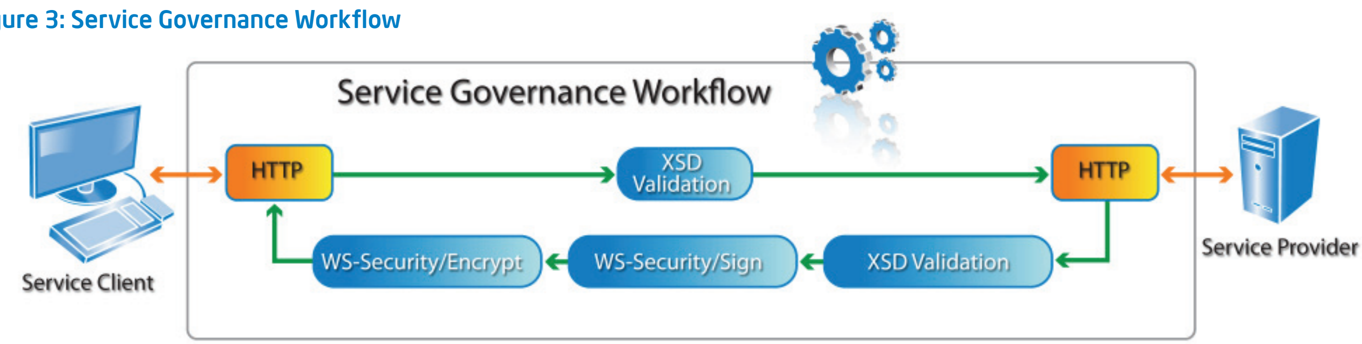
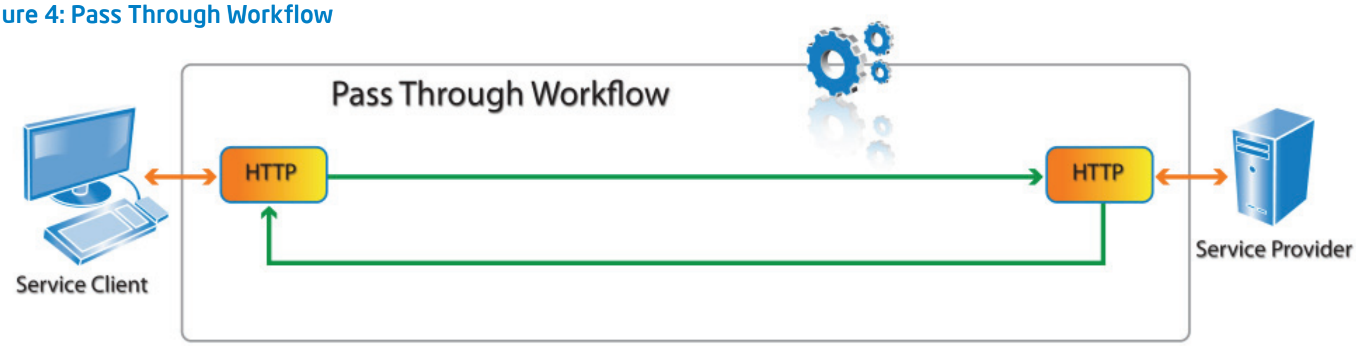


Figure 4: Pass Through Workflow



3.2.3 Traffic Pass-Through

In addition to the two aforementioned tests, a third "pass-through" or service virtualization test (see Figure 4) was also conducted to determine peak pass-through rates. In a pass-through test the job of the service intermediary is to act as a pure HTTP message proxy. That is, it receives the SOAP v1.1 request and passes it along to the back-end server with no additional processing. To configure a pass-through test the IBM DataPower was configured using a manual front-side HTTP handler (no WSDL) and the Intel SOA Expressway was configured with a receive, invoke and reply policy pipeline. A pass-through test is a good baseline to determine the upper bound on the HTTP and network processing before additional processing steps are applied. In addition to basic service firewall protection, the pass-through scenario is useful in determining the potential of the connection handling and processing inherent in each product.

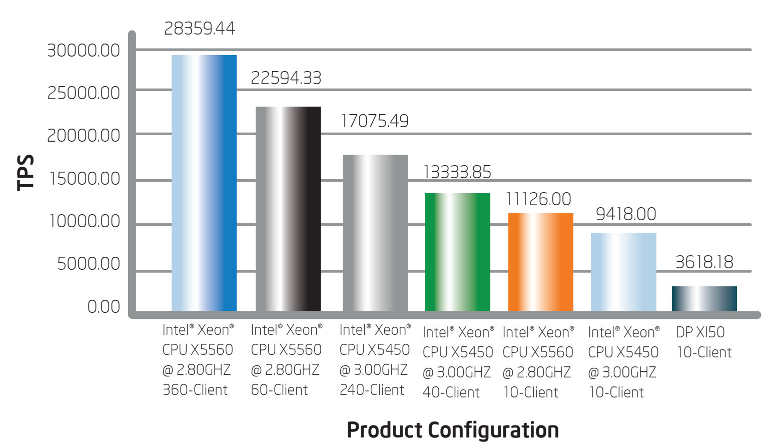
4. Test Results

The next three sections describe the throughput, latency and CPU utilization results for the pass-through test.

4.1 Traffic Pass-Through

Figure 5 demonstrate the baseline proxy performance for each product at a different number of client threads. In all cases, the Intel SOA Expressway was able to handle considerably more pass-through traffic. In the extreme case of the Intel® Xeon® X5560 platform, SOA Expressway was able to handle more than 8 times the throughput as compared to the IBM DataPower XI50. It is natural to ask why IBM DataPower was only tested with 10 client threads in this case. We observed that the CPU utilization (shown in Figure 6) was already at 90% and furthermore, errors were observed with thread counts higher than ten[3].

Figure 5: Pass-Through TPS
Intel® SOA Expressway vs. IBM* DataPower XI50



[3] We found problems with higher numbers of threads with persistent connections enabled. Disabling persistent connections allowed the IBM DataPower XI50 product to process a larger number of worker threads without errors, but at a reduced throughput rate.

Next we look at the latency between requests in the pass-through case. As shown in Figure 6.

Figure 6: Latency
Intel® SOA Expressway vs. IBM* DataPower XI50

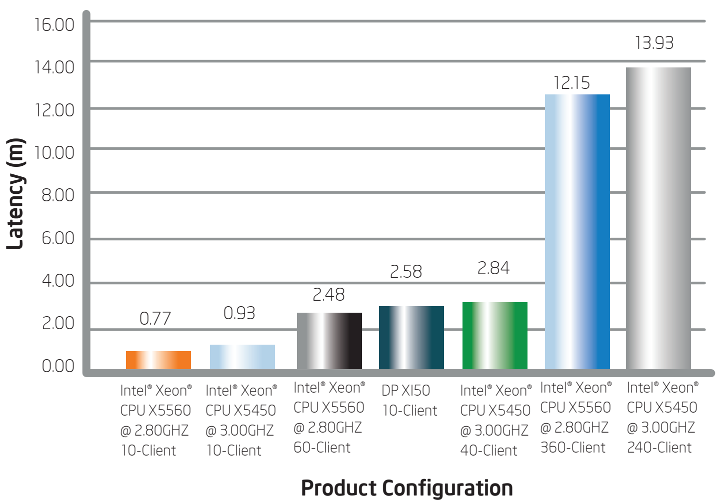


Figure 7: CPU Utilization
Intel® SOA Expressway vs. IBM* DataPower XI50

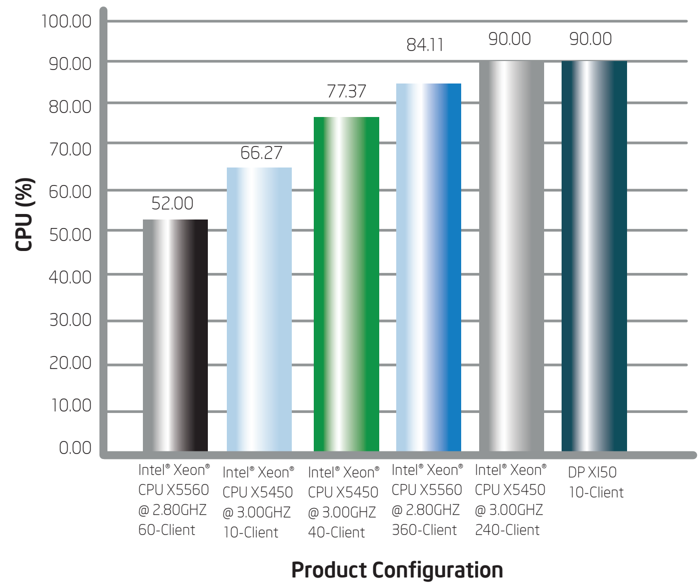


Figure 6 shows the latency comparisons for the baseline pass-through case. In these graphs the latency is the wait time as perceived by each worker thread. Here we can see that once again, the IBM DataPower XI50 is only able to scale to 10 threads while SOA Expressway on the Intel Xeon X5560 is handling 6 times the traffic at a lower latency. Intel SOA Expressway also shows the lowest latency in the 10 client test as well, demonstrating over one-third the wait time as compared to the IBM DataPower XI50. The next data point is CPU utilization for each product, which is shown as follows in Figure 7:

Figure 7 shows how much CPU is utilized during the pass-through tests. Here we can see that the available headroom on the IBM DataPower XI50 product is nearly exhausted at 10 clients and 3600 TPS while Intel SOA Expressway still has nearly half its potential headroom available in the case of the Intel Xeon 5560 Platform at 60 clients. Overall, the DataPower product shows the lowest available headroom at the lowest number of threads, indicating it is nearly maxed out. In the next section we will look at how each product scales over time for both the service mediation and service governance use cases.

4.2 Service Mediation Test Results

The following figures compare the performance of the IBM DataPower XI50 to Intel SOA Expressway running on both Intel platforms for service mediation. Each section shows a plot over time comparing the TPS, latency and idle CPU as the number of client threads increase.

Figure 8 shows the throughput measured messages per second (TPS) as the number of client threads increases. Here we can see SOA Expressway on the E5450 platform hitting a peak throughput of about 6200 TPS and about 9600 TPS on the X5560 platform before leveling off. The leveling off behavior indicates that throughput levels are being maintained at the expense of latency (see Figure 8). That is, despite the number of clients, SOA Expressway is behaving predictably. This contrasts the IBM DataPower XI50, which hits its peak at about ~2600 TPS before degrading to just under 1000 TPS with heavy load. In this case, the IBM DataPower XI50 is unable to scale to the higher thread levels[4].

Figure 9 shows the latency in milliseconds between client requests for each product. Here we can see that SOA Expressway has linear behavior with respect to the time between requests, even under heavy loads. SOA Expressway's slight initial advantage in latency quickly expands over IBM DataPower XI50 as the client threads increase. The IBM DataPower XI50 begins with linear behavior, but this turns erratic at about 200 threads, indicating difficulty in scaling stably at higher thread levels. It is also important to note that on the faster Intel platform (X5560) the latency slope is more gradual, indicating that SOA Expressway is scaling with high predictability and stability based on platform improvements alone.

Figure 10 shows the percentage of available CPU as the number of client threads increase. Here we can see that Intel SOA Expressway (E5450) and IBM DataPower XI50 are about on-par for available CPU (around 10%) for the first 200 threads. As the threads increase, however, IBM DataPower begins to have scalability problems. Even though more of the CPU becomes available it is unable to process the increased number of client requests and maintain throughput levels (See Figures 8 and 9). For SOA Expressway on the Intel X5560 platform we can see a large increase in available CPU once again due to Intel platform improvements.

Figure 8: Throughput for SOA Expressway vs. IBM* DataPower XI50

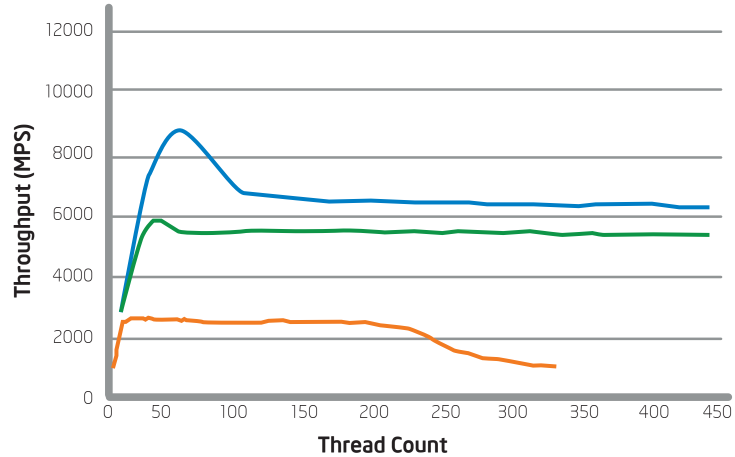


Figure 9: Latency for SOA Expressway vs IBM* DataPower XI50

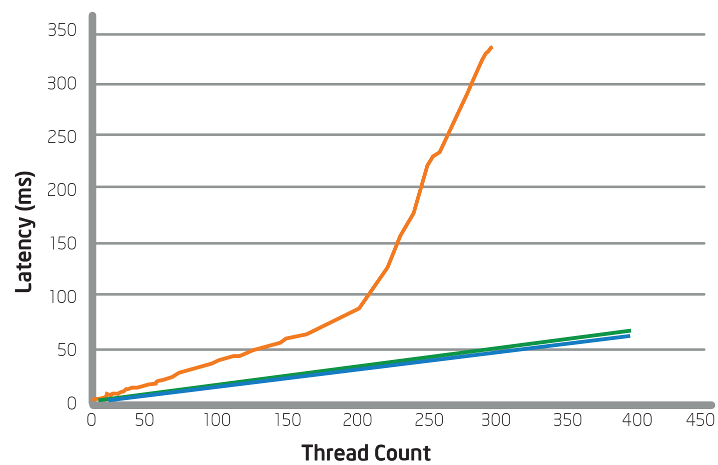
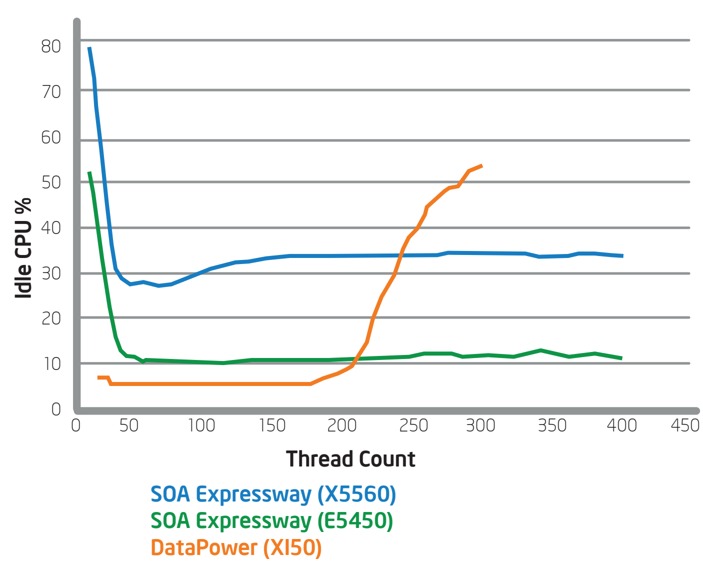


Figure 10: Idle CPU for SOA Expressway vs IBM* DataPower XI50



[4] Processing errors were reported by the IBM DataPower XI50 for the service mediation case when thread levels reached 230. Intel SOA Expressway produced zero errors.

4.3 Service Governance Test Results

The following figures compare the performance of the IBM DataPower XI50 to Intel SOA Expressway running on both Intel platforms for service governance. Each section shows a plot over time comparing the TPS, latency and idle CPU as the number of client threads increase. It should be noted that the service governance scenario is considerably more taxing to each system due to the addition of security operations. While the absolute throughput numbers are expected to be lower for all the systems under test, the overall behavior under heavy client load should remain consistent with the service mediation case.

Figure 11 shows the throughput measured in messages per second (TPS) as the number of client threads increase. Here we can see SOA Expressway on the E5450 platform hitting a peak throughput of about 2030 TPS and about 3700 TPS on the X5560 platform before leveling off. This contrasts the IBM DataPower XI50 which achieves a peak throughput of 1280 TPS before leveling off and then eventually degrading under heavy load[5]. It should be noted that we obtained the lower bound performance advantage[6] of about 1.5X by considering the sustained throughput of SOA Expressway on the E5450 platform as compared to IBM DataPower XI50.

Figure 12 shows the latency in milliseconds between client requests for each product. Here we can see that SOA Expressway has linear behavior with respect to the time between requests, even at heavy loads. In the latency comparison between SOA Expressway on the E5450 and X5560 it is also important to see that the difference in slope is larger for the service governance case as compared to the latency slope shown in Figure 8. This suggests that platform improvements in the X5560 are more applicable to the service governance case and that Intel SOA Expressway is once again successfully scaling on the latest Intel platforms.

Figure 13 shows the percentage of available CPU as the number of client threads increase. While the DataPower XI50 has more available CPU than Intel SOA Expressway on the E5450, it is unable to make use of the extra processing power to scale its throughput or reduce its latency (see Figures 11 and 12). Moreover, the spike in free CPU for DataPower is due to excessive errors which begin at thread 169. In this case, the DataPower XI50 is unable to handle the increased number of threads and begins to lose requests; in other words, the spike in idle CPU is a symptom of overload rather than an indicator of increased headroom. Also, in the case of Intel SOA Expressway on the X5560 we can see a dramatic increase in the amount of idle CPU, indicating once again that Intel SOA Expressway scales on the latest Intel platforms.

[5] Processing errors were reported by the IBM DataPower XI50 for the service mediation case when thread levels reached 169. Intel SOA Expressway reported zero errors.

[6] If we assume an average 2000 TPS for SOA Expressway on the E540 as compared to 1280 TPS for the IBM DataPower XI50 the ratio is approximately 1.55X

Figure 11: Throughput for SOA Expressway vs IBM* DataPower XI50

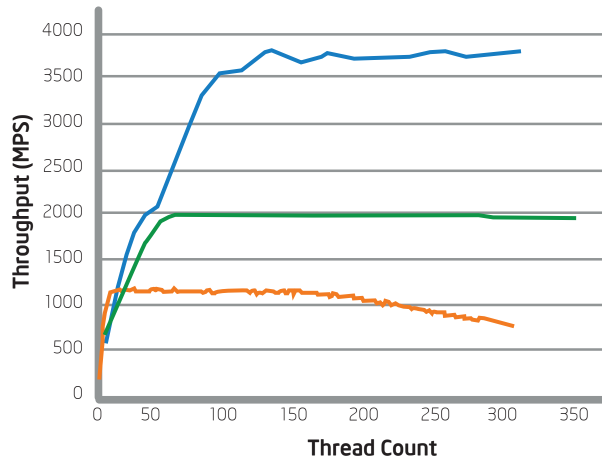


Figure 12: Latency for SOA Expressway vs IBM* DataPower XI50

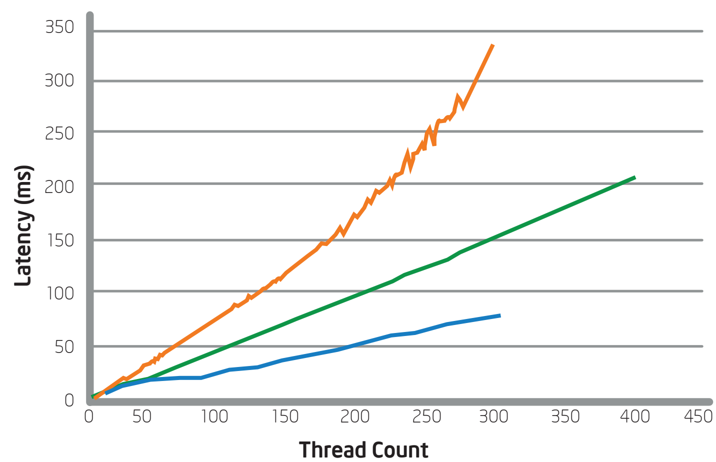
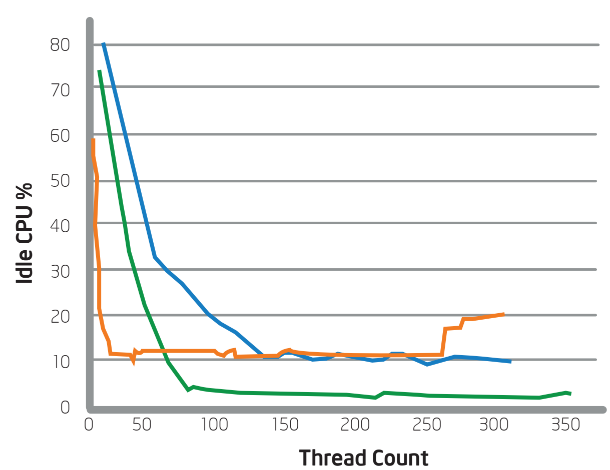


Figure 13: Idle CPU for SOA Expressway vs IBM* DataPower XI50



SOA Expressway (X5560)
SOA Expressway (E5450)
DataPower (XI50)

5. Cost Comparison

The following table shows an initial upfront cost comparison for Intel SOA Expressway versus IBM DataPower XI50 for a typical deployment consisting of 2 units deployed at 2 different physical locations, each with a single hot-standby unit (4 units total).

Table 2: Initial Cost Comparison

Product	Per-Unit Price	Per-Unit Maintenance & Support	Total Upfront Cost (4 units)
IBM* Websphere DataPower XI50[7] Total Price	$100,000	$9,750	**$439,000**
Intel® SOA Expressway Production Units[8]	$64,000	$12,600	
Intel® SOA Expressway Hot Standby Units	$34,000	$5,800	
Intel® SOA Expressway Total Price			**$232,800**

While discounting may apply in both cases, Intel SOA Expressway's list price is only about 50% of the IBM DataPower XI50, despite its considerable performance advantage. Moreover, since Intel SOA Expressway runs on standard Intel servers, it can be deployed on existing server hardware in the datacenter rather than requiring the acquisition of purpose-built appliances. Additionally, the Intel SOA Expressway soft appliance form factor allows datacenter to update server hardware independently riding on Intel's hardware performance evolution while maintaining system architecture continuity.

Conclusion

This detailed "apples-to-apples" comparison demonstrates that the Intel SOA Expressway soft appliance running on Intel Multi-Core servers easily outperforms a custom-built hardware appliance on real-world use cases. Not only can Intel SOA Expressway achieve higher "peak" throughput, it also demonstrates predictability, linear scalability, and the ability to take advantage of Intel microprocessor advancements and apply them directly to business processing without any specialized coding or custom configurations. We demonstrated a lower bound performance metric of 1.5X on the service governance use case and an upper bound throughput metric of 8X on the pass-through use case, with an average advantage of 2X to 5X improvement at different thread levels. This clear advantage coupled with the lower cost and increased flexibility of a software solution makes Intel SOA Expressway a compelling choice to increase performance and reduce costs and complexity for service integration in the datacenter.

[7]Source: IBM Websphere DataPower SOA Appliance Announcement, March 2006 (http://www.ibm.com/common/ssi/rep_ca/3/897/ENUS106-253/ENUS106253.PDF)

[8]This price includes the price of a suitably configured server (Intel Xeon E5450 class) as well as a 3 year enterprise Linux license with support.

About SOA Expressway

SOA Expressway is a soft-appliance deployed to address common XML and SOA problem areas such as acceleration, security, service mediation and service governance. SOA Expressway is available for any organization deploying services (SOA), hosted services (SaaS) or Web 2.0 (RIA). SOA Expressway is available for standard operating systems such as Windows and Linux and requires no special custom hardware other than standard OEM servers.

For more information:
http://www.intel.com/software/soae

Catch the "Practical Approaches to Service Delivery Webinar Series"
www.intel.com/software/soae/webinars

Americas: 1-630-627-3370
Europe, Middle East, Africa: 44 (0)7919 303 236
All other Geographies: 1-905-681-8768

E-mail: intelsoainfo@intel.com

Performance Test Details

- Intel® SOA Expressway v2.0 on Intel Xeon
  - Dual Processor Intel(R) Xeon(R) CPU X5560 @ 2.80GHz Nehalem-EP
  - Manufacturer: Supermicro
  - Memory: 6x2GB 1333 DDR3 RDIMMs (12G total)
  - Operating System: Red Hat Linux AS 4
- Intel® SOA Expressway v2.0 on HP ProLiant
  - Platform: HP ProLiant DL360 G5
  - Dual Processor
  - Manufacturer: HP
  - Memory: 4x2GB FBD PC2-5300 (8G total)
- IBM* DataPower XI50
  - Firmware Rev: XI50.3.6.1.5
  - Build: 156262
- Client: High performance generic HTTP Linux client
- Server: Standard low-latency HTTP web server
  - 4X Quadcore: Genuine Intel® CPU @ 2.40GHz
  - Memory: 16G
  - Red Hat Linux AS4
- Default settings were used for both Intel® SOA Expressway and the IBM* DataPower XI50

Printed in USA Please Recycle SOAE-CompareWhitePaper-001US