

Initiate® Software Performance Tech Brief

RESULTS AT A GLANCE

- Initiate® software achieved a rate of 2200 transactions per second (TPS) in a data set of 150 million records
- Average latency ranged from 34 to 42 milliseconds
- Results demonstrate linear scalability of approximately 370 TPS per CPU
- Initiate provides performance, scalability and accuracy, regardless of implementation style: registry, hybrid or transactional data hub

Initiate Systems completed extensive testing that conclusively demonstrates its master data management (MDM) software delivers superior performance and the highest number of transactions per second (TPS), as compared to other MDM software tested in similar published studies.

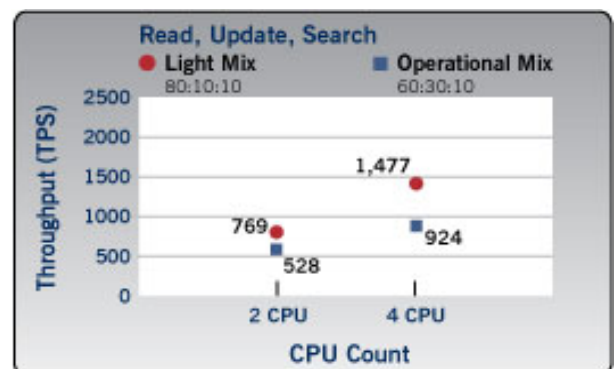
Why Performance and Accuracy are Important

Data volumes are increasing exponentially within global organizations today. To gain competitive advantage in the marketplace, companies are gathering more data from inside and outside their firewall than ever, and this trend will only accelerate. As IT departments wrestle with this issue, business users demand that the information they need be accurate, and be delivered in time to support critical processes. When examining solutions to meet these needs, companies should strive to meet all the performance, scale and accuracy requirements. Sacrificing in any of these areas can exacerbate an organization's challenges, instead of solving them.

During testing at the IBM Innovation Center in Chicago, Initiate® software yielded 2200 TPS, with a per-core average rate of 370 TPS. The tests were executed with an environment of moderately sized hardware and a data set with over 150 million customer records. As expected with any data-intensive application, the system performance scaled linearly when adding CPU and I/O capability to the database server. However, unlike as recorded by other data matching solutions in similar published studies, additional application CPUs were not necessary to drive these impressive performance numbers.

But what do all these technical numbers mean in the real world? Let's say you work in a major corporation and you need to compare 10 million new records against a known database each day. Processing needs to be completed overnight, within an 8-hour window. This requires a processing rate of 347 transactions per second.

Additionally, using standard server platforms, Initiate can sustain up to 2200 TPS in an 8-core configuration, depending on the transaction mix. This level of performance and scalability provides virtually any customer with the current and future capacity to address their business needs.



In addition to outperforming other products in TPS, Initiate Master Data Service also delivers the most accurate results. The software's acclaimed matching technology can evaluate hundreds of records from various data sources in milliseconds, returning only those with the highest probability of matching the request. Through this combination of speed and precision, Initiate offers unparalleled value for customers with stringent accuracy and performance requirements, even in very large data sets.

Approach

Initiate® software’s patented approach supports an organization’s performance, scale and accuracy requirements by casting a wide net and searching across all possible records that might match a request from a client application. Initiate defines search differently than most applications, which either assume single identifiers for lookup or only evaluate a handful of potential matches. Instead, Initiate software searches across and assesses hundreds of records in milliseconds. This ensures that all records that could possibly match a client request are evaluated, and those with the highest probability of matching are returned.

Initiate can support customer performance, scale and accuracy requirements regardless of implementation style, whether it is a registry, hybrid or transactional data hub. In addition, as data are updated by clients or other systems, Initiate Master Data Service re-evaluates each record to determine whether data are associated correctly with other stored records. This methodology enables every client request to receive the most current and trusted data, based on the requirements of the requesting system.

Transaction Type	Light Mix	Operational Mix
Read	80	60
Write	10	30
Search	10	10
Peak Throughput	2200 TPS, avg latency 34ms	1000 TPS, avg latency 42ms
Per core average TPS	385	260

Test environment

Hardware. The environment consisted of an application server and a database server. Both were installed on 8-core POWER5+ systems with 64GB of memory. The core count was scaled dynamically using IBM’s hardware management console. In addition, the environment included two local unit numbers (LUNs) on dedicated 1.1TB capacity RAID 10 arrays, one for data with a 32-spindle configuration and one for logs with a 16-spindle configuration.

Software. The application and database servers ran on AIX 5.3 with Initiate Master Data Service™ 8.0 installed on the application server. The 150-million record data set was generated with demographic data consistent with US census frequencies and loaded into an Oracle 10g database instance. To create realistic customer records, additional attributes and payload data were added to each record.

Workload. Two variations of GET:PUT:SEARCH (read, update, search) data interaction mixes were tested, an operational mix and a light mix. The operational mix testing was based on a 60:30:10 ratio, which is the transaction ratio observed in most actual customer environments. For comparison, a less realistic light mix testing was based on an 80:10:10 ratio.

For both interaction mixes, Apache JMeter was used to generate a consistently-random workload and to collect the key metrics, interaction throughput and latency. The tests were conducted in an isolated environment with reboots of all systems between test execution. In addition, instrumentation was implemented to monitor overall system health, including the periodic review of I/O, operating system (OS), database, network and application statistics.

Conclusion

The results from the IBM Innovation Center tests prove that Initiate software delivers superior performance, even with a minimal hardware configuration, as compared to other MDM software tested in similar published studies. This performance, combined with Initiate software's acclaimed accuracy, quick deployment requirements and flexibility in deployment styles, make Initiate an outstanding value in master data management.

For more information on how software from Initiate Systems can help with your master data management efforts, visit www.Initiate.com.