



Tuning your Verity Search Engine

If you're a Verity customer, you'll want to get optimal performance from your collections. Raritan Technologies specializes in helping you to get the best results from your current infrastructure. We'll help you by using several strategies to tune your engine and then our "Verity Stress Tester" will let you quantify your results.

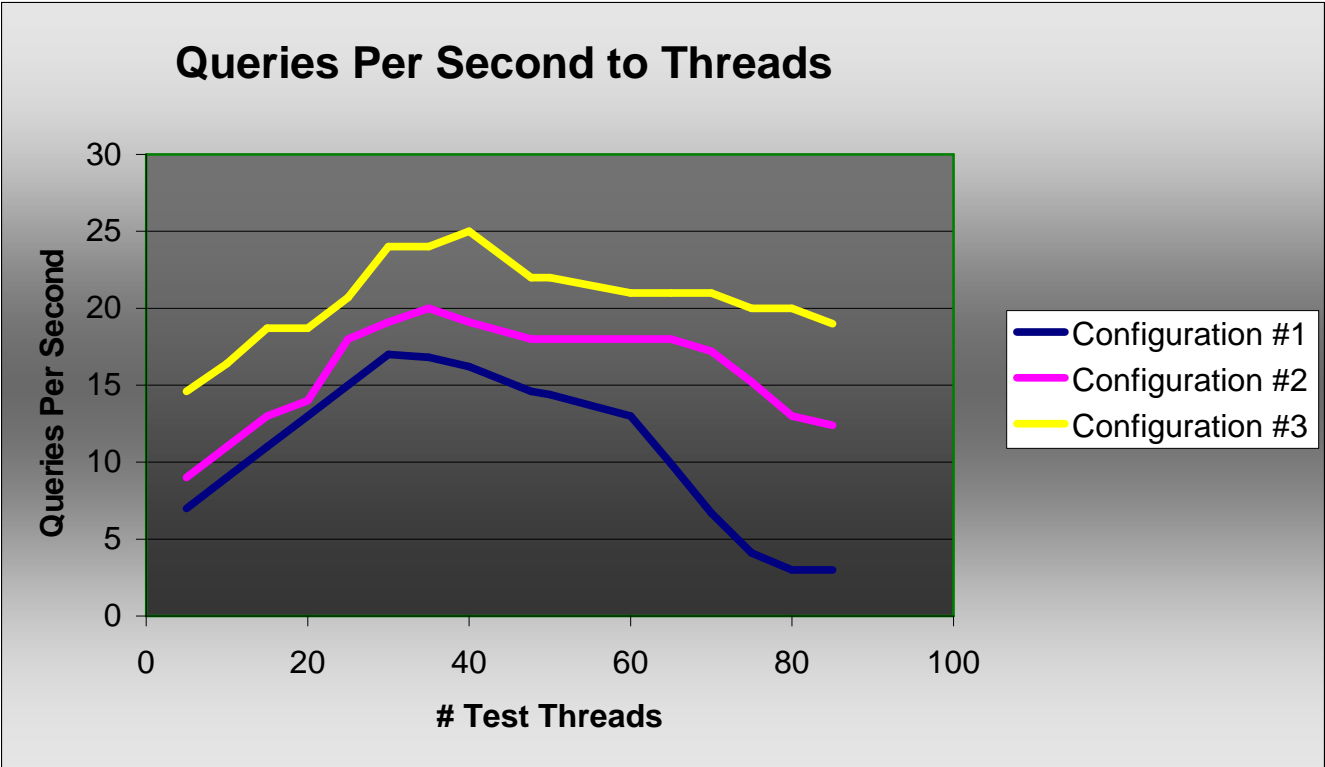
Your Strategic Options

There are many different strategies that you can employ to achieve peak performance from your Verity applications and collections. We'll help you navigate through these choices and choose the ones which are most appropriate for your needs:

- Topics – These pre-compiled queries allow you to access your result sets in a fraction of the time.
- Parametric Indexes – Use these in-memory indexes to navigate efficiently through your data.
- Source Queries – Often used in conjunction with Topics, source queries perform a pre-search with no scoring or ranking to reduce the set of records to search in the final query. Source queries can also be used to implement application security.
- Query tuning – Get the results that you want by using the most efficient Verity searches. Modify your queries to use Full Text queries, Indexed Topic searches, zones instead of fields for quicker results
- Query re-writing – Sometimes called Query Cooking, this strategy allows you to modify user's original query to still get the desired results but use a more efficient query. You can also extend Verity's language this way.
- 80/20 Rule – Use this Verity Best Practice to allocate 80% of your resources to the most often run queries.
- Collection Re-design – Determine if you can re-structure your collections so that you can use more efficient Verity operators to get your results faster.
- Collection Split – Sometimes collections grow and can span too many physical partitions to be read efficiently. When this happens a collection split can allow you to access numerous physical collections as one logical one.
- Collection Replication – If a collection (or set of collections) is accessed by large volumes of queries that cannot be optimized, collection replication is a strategy that can be used to bring in additional hardware to resolve this problem.
- Optimize – Verity has some built in tools, which enable you to optimize a collection that has started to fragment.
- Use of "Spanning Word Lists" can make certain collections perform better.

Quantify your gains

Our stress-testing tool allows you to gather detailed performance statistics for each different configuration. You can easily plug in one or more Verity Collections against which the tests are performed. This java-based utility takes as input sets of user queries (and source queries if desired) and returns detailed performance statistics (queries per second, run time for each query in milliseconds and documents found for each query). The utility allows you to manipulate the number of concurrent threads or searches, to realistically model your usage patterns. The data returned is compatible with Microsoft Excel and is easily imported into graphs that clearly show if a tuning strategy has merit to follow up or not.



Raritan Technologies is a full service consulting company specializing in search applications. We provide development, custom software and database solutions. We can provide custom solutions using industry standard software, whether you are a Fortune 500 company or a small independent organization.

The services we provide have many different titles - application development, web development, IT consulting, B2B development. Our specialty is the development of browser based user interfaces to very large and diverse document databases. We are experts in the **Verity** product line, including Federated Technology, Collection and Application Tuning and Development, Taxonomy Integration and Parametric Technology. We also work with document storage technology, including Documentum, Oracle and other relational databases.