THUNDERSTONE

Thunderstone Search Appliance

The Thunderstone™ Search Appliance is the fastest and easiest way to add high quality search results to your web site, whether external or internal.

The Appliance comes completely pre-configured. Simply plug it in and begin using the web-based menus. A business administrator can extensively customize its operation. The end-user view is simple and intuitive. The relevance ranking of search results is state-of-the-art.

Integration

The Appliance can serve search results directly to the user, using predefined templates with your look and feel added, or use an XSL stylesheet to customize the user interface completely, or integrate it into your portal using XML.

Performance

The Thunderstone Search Appliance is designed for performance. It can handle over 1,000 queries a minute, or more than a million per day. Its innovative Adaptive Indexing process conserves network resources by automatically determining how often each item needs to be updated.

Flexibility

Create multiple independent search collections on the Appliance. Each can consist of one or more websites. You can create unique collections of web information for your users or for competitive intelligence by indexing third-party information, such as government or industry web sites anywhere on the internet.

Designate specific pages to be highlighted and appear first in response to any query. The Appliance makes sure you and your visitors can find the data on your site!

Completeness

The Appliance can deal with most types of content on your network. It searches PDF files and highlight the matches in Acrobat Reader, as well as more than 100 other file formats that you might

have on your intranet. It indexes content and follows links generated by JavaScript. It indexes database content whether or not the database is web accessible.

The Appliance helps you monitor and protect your content. Its crawl reports are invaluable for checking your web site for broken links, which pages are accessible via links and other problems. If desired, the Appliance keeps a record of every search and the results viewed.

Security

There are many controls to address most security concerns. With document-level security, each user only sees results they are authorized to see. You can delegate administration selectively so that sub-administrators have access only to specified settings and collections.

Warranty

With two-year warranty, support and upgrades included, and an optional maintenance program thereafter, we ensure your peace of mind. In the rare event of a hardware failure, a replacement machine is cross shipped. If required, two Appliances may be set up in a mirrored configuration for redundancy.

Thunderstone

Thunderstone Software LLC is a pioneer of search engine technology, providing text retrieval products to industry, government and educational institutions since 1981. Thunderstone distinguishes itself through:

- **Performance**. Since our software's inception on mainframes and early PCs, performance has been critical. With today's large Internet user base, performance remains essential.
- **Experience**. From the early 1980s to today, Thunderstone has been solving information retrieval problems. Put that experience to work for you.
- **Unparalleled support**. We want your project to succeed. We take pride in having the most sophisticated and responsive technical support organization in the industry.

Thunderstone Search Appliance
Specifications

Capacities

Queries (sustained peak)	1,000/minute
Number of documents or web pages	250K, 500K, 1M, 3M, 6M, 9M, 12M, 15M, larger sizes custom config.
Number of collections	Unlimited
Number of web sites	Unlimited

Crawling

Protocols supported	http: https: ftp: gopher: file: javascript:
File system crawling	Yes, webserver not required
JavaScript/Flash	Code executed, embedded links followed
Cookie management	Automatic
Fastest page refresh	One minute
Re-crawl scheduling	Adaptive auto-schedule
Protected page crawling	Basic authentication, Integrated Windows Authentication, form-based login, SSL. Multiple passwords supported.
Add page latency	Immediate
Delete page latency	Immediate
Exclusions	Specify by URL, html tag, or field contents
Categorization	Define by URL pattern
Split large documents	By page or character count

Content Types

File formats	PDF, MS-Office, most word processors, many other formats
JavaScript/Flash	Text and links indexed
Databases indexed (whether or not web accessible)	Oracle, Microsoft SQL Server, PostgreSQL, Sybase, Texis
Binary file indexing	Finds and indexes ASCII content within arbitrary binary files.
Duplicate detection	Automatically removes pages, or just headers/footers
Foreign languages	Automatic UTF-8
Real-time data feeds	Defined XML format

Search Options

Spelling checker	Suggestions based on indexed text
Find similar	Automatic query-by-example link
Wildcard	Yes, * character
Thesaurus	Built-in
Stemming	Built-in
Advanced search	Control proximity, word forms, ranking

Results Display

Results ordering	Date or relevance. Five user-controlled relevance "knobs".
Predefined results	"Best bets" (admin-defined) forced to top
Hit highlighting	Query words in context
Results formatting options	HTML template, XSL, or XML
Cached page text	Yes, with highlighting
PDF features	Highlighting, jump to page in Acrobat Reader

Administration

Administrator interface	Web-based
Look-and-feel	Totally customizable
Admin authorization	Multiple administrators with distinct permissions
User authorization	Checked for each item of search results
Reporting	All queries; top queries; top words; no-hit queries; click-overs; error logs
List crawled URLs	Yes, ordered by depth, URL, date, or size
Maintenance & upgrades	Remote, automatic (OS and application)
Redundancy	Automatic data duplication (requires 2 Appliances)
Warranty	2 years included; additional years optional

Physical

Form factor	1U/2U rack mount
Disks	RAID array
Power Supply	Redundant 450 W/700W
Cooling Requirement	1500 BTU/2500 BTU

For more information: call +1 216-820-2200, email info@thunderstone.com, or visit www.thunderstone.com