

Presented October 11, 2011 at NYExUG Meeting
Last Updated on October 14, 2010



Exchange Server 2010 Search Functionality in Technical Detail & Comparison to Exchange 2003/2007

** throughout presentation indicates
additional information on reference page in
back of presentation*

Ben Serebin
Ehlo & Network Consultant
REEF Solutions (www.reefsolutions.com)
If you can't get enough Exchange & technology
read my blog <http://blog.reefsolutions.com>

About Ben Serebin

- Working in the IT sector since 1996
- Specialty is Exchange Server, Spam Filtering, DNS, & Wireless.
 - Recently Completed Projects Q3 2011: Re-designing/architecting over 25TB of iSCSI storage for improving efficiency (performance & capacity), performance testing of NFS vs iSCSI, & implementation of VMware resource clustering for load balancing of CPU/Memory/storage.
- Upcoming Projects: Designing & implementing a 4 node cluster of Windows RDS with RemoteFX functionality, deploying KVM over IP for cluster of ESXi hardware, & minimal downtime of VM migration for replacement of Sun/Oracle hardware.
- Current Environment: Running Exchange 2007 & 2010 Server on Windows 2008 & 2008 R2 on ESXi 4 Cluster, server computing environment is entirely virtual, Blackberry Enterprise Server Express 5. Current handheld email device is a BlackBerry Bold 9650.

Agenda for Presentation

- Exchange 2010 Search Methods
- Exchange 2010 Search Highlights - Major
- Exchange 2010 Search Highlights - Minor
- Exchange 2010 Search Highlight in Detail
- Exchange 2007 Search Highlights - Major
- Exchange 2007 Search Highlights – Minor (Clustering)
- Exchange 2003 Search “Highlights” (aka warnings)
- Exchange 2010 Search Demo
- Exchange 2007 Search Demo
- Exchange 2003 Search Demo

1. Exchange Search (new content indexing) – 2007/2010) - FASTEST
2. Store Search (2003 approach) - slow
3. Outlook Online Search with Windows Desktop Search - faster
4. Outlook Cached Search – better than slow
5. Outlook Cached Search with Windows Desktop Search - fast

IMPORTANT – Why is my search slow?

How Outlook searches VARIOUS based on Outlook version, cached/online, OS running Outlook, # of messages embedded, OWA version, and more.

BEST SEARCHING PERFORMANCE – Windows Vista/7 plus Outlook 2007/2010 in online mode.
Cached mode results in PC based searching (slower).

Can OWA 2003 search within attachments? No.

How about OWA 2007?
2010? Yes.



Exchange 2010 Search Highlights (Major League)

- 2010 – overall performance improvements are very significant. Indexing speed is lightning fast.
- 2010 – default install loads search filters (Office file types, text, html attachments). 53 file types! PDF not one. You must load a 3rd party tool (Adobe or FoxIt [$\sim 39x$ faster than Adobe*]). Technically they are “iFilters”.
- 2007/2010 – Always up to date vs crawl mode. For 2010 new messages are typically indexed within 10 seconds of arrival. 2007 is 3x that.

- 2007/2010 - Content indexing is enabled by default on mailbox databases.
- 2010 – auto throttle when I/O or CPU utilization passes a threshold.
- 2010 – Multi-mailbox Searching via end user rights. Uses same index catalog & web based.
- 2010 – DAG based search index adds only about 10% more storage. 2007 CCR also supports single search index catalog.
- 2010 SP1 - Ability to “dedup” results from a multi-mailbox search.
- 2010 - Ability to disable searching per database or server.

- Microsoft recommended way is SharePoint.
- But, searching Exchange 2007/2010 Public Folder Databases *is* possible if a # of criteria are met:
 1. Outlook is in Cached Exchange Mode
 2. Only on Public Folder Favorites
 3. Public Folders must be enabled to Download as Favorites.

By design, Exchange 2007 only indexes mailbox store databases, while both previous versions Exchange 2003 and 2000 also indexed public folder databases. By design, the current method of indexing Exchange 2007 public folder databases is to use SharePoint 2007. Similar guidance by Microsoft suggests similar for 2010.

- 2007/2010 have smaller search index catalog. Typically only 5-10% of the database size.
- New messages appears typically within 30 seconds versus 2003's hours or daily (typically config).
- Database storage capacity should allow for 20% overhead size due to indexing updates. It's short term storage needs, but this space is needed. Process for merge is called "Master Index Merge".
- Search performance improved by implementation of auto-throttle mechanism.

Exchange 2007 Searching – Additional Points + Clustering

- Search Filters for Office 2007 file types are NOT installed by default.
- If search process fails on an attachment, message body will **NOT** be indexed. 2003/2010 do not have “feature”.
- Introduced, ability to exclude commonly used words (aka noise words – e.g. “a” “the” “for”) from the search catalog (based on SharePoint technology) to keep the index relevant, smaller, and provide fast results. Not enabled by default, but recommend this be implemented to improve performance. Noise word lists are per language. *

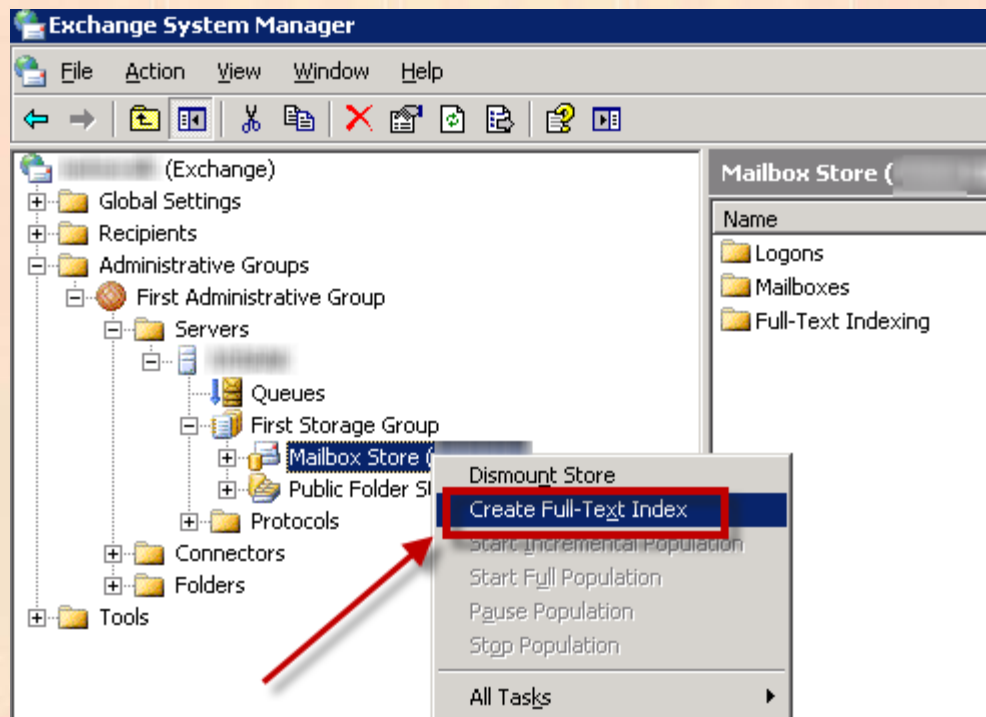
Clustering Exchange? Search Index Storage Needs.

- CCR - 2 search index catalogs
- LCR - 1 search index catalog
- SSC - 1 search index catalog
- Catalog index is typically 5-10% of the database size.
- Database storage capacity should allow for 20% overhead size due to indexing updates. It's short term storage needs, but this space is needed.

Exchange 2003 Searching

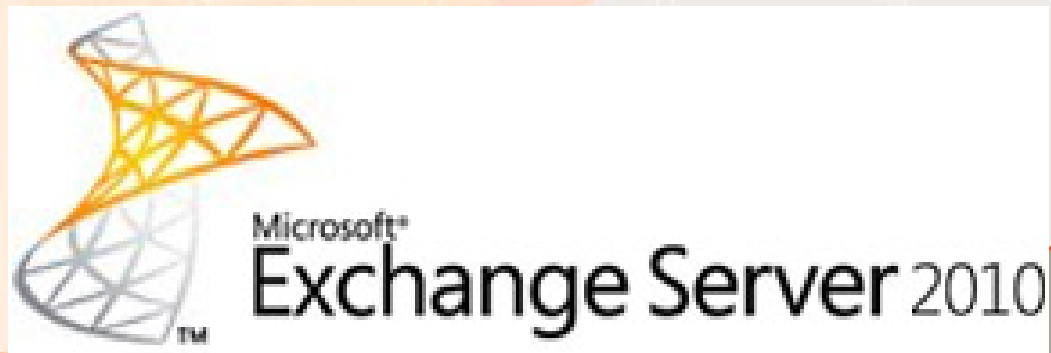
Highlights

- Full Index Searching is disabled due to the performance overhead.
- Search is controlled on database level (e.g. Mailbox Store and/or Public Folder store)
- To enable? Right-click on database level, and select “Create Full-Text Index”.*
- Search filter failures did not stop message body indexes.
- Catalog index was typically 35-45% of the database size.





Time for generating thousands of emails in minutes for showing a comparison of search speeds for various Exchange versions. Free tool used is “Server Traffic Test” from <http://www.icewarp.com/downloads/tools/>



References Details

Content Indexing PDF files for Exchange 2007/2010 – Recommend – FoxIt iFilter [\$] compared to Adobe Acrobat 9 [free] (2007 version works for 2010).

<http://blogs.msdn.com/b/opal/archive/2010/02/09/pdf-ifilter-test-with-sharepoint-2010.aspx>

Database diagram for how Content Indexing works in Exchange 2007 – by Microsoft Exchange Team

<http://blogs.technet.com/b/exchange/archive/2009/09/17/3408316.aspx>

Exchange 2003 - HOW TO: Configure Full-Text Indexing on Mailbox Stores

<http://support.microsoft.com/kb/822932>

Noise Words for 2007/2010 – see section titled “Noise Words in Exchange 2007”

<http://blogs.technet.com/b/exchange/archive/2009/09/17/3408316.aspx>

Exchange 2010 – “C:\Program Files\Microsoft\Exchange Server\V14\Bin\FTERef”

- search appears to be based on Office SharePoint Portal Server 2004

Thank you for attending the October 2011 NY Exchange User Group Meeting.

Raffles – do not forget, if you bring in new members, you 2x, 3x, or more your chances to win raffle prizes.

TechHit.com (Outlook add-ons) – raffling off any one of their products (SimplyFile [intelligent filing], EZ-Detach [easily detach multi-emails of attachments], MessageSave [backup, archive, save msgs], AutoRead [mark as read/remove new mail icon], & QuickJump [quickly open/move between Windows folders]).



Now for Question of the Month.....