

Scans Searcher documentation

written by Thomas Lackó
last modification: 12 May 2006

Abstract

Scans Searcher is a scanner and searcher utility for local area networks, designed to search for files and directories matching the given criteria.

No server is required for any of the program's features. Searching even doesn't require a network connection. Scans Searcher has little space and memory requirements.



author: Thomas Lackó
versions 7...18s : design & graphics by DJ
e-mail: thomas.lacko@gmail.com
web: www.thomas.sk

© 2004–2006

Contents

Setup	2
Searcher.....	3
Page 'Search'	3
Page 'Results'	4
Page 'Statistics'	6
Page 'Options'	6
Scanner & Database	8
LAN Scanner.....	8
Lodzha's FastScan	9
Database builder	9
Example batch file	9
Version history.....	10
Searcher	10
LAN Scanner.....	13
Database builder	13
Database format.....	14
Future plans	14

Acknowledgements

I would like to thank the members of the "Sturak" college's intranet for the inspiration and their valuable comments and bugreports.

Licence agreement

This agreement is between the authors and the users of the Scans Searcher software. If you do not intend to honor this agreement, delete the software from your disk.

Scans Searcher is freeware. Anyone is free to copy and use this software free of charge for his non-commercial purpose. However, nobody may decompile any part of the code, nor analyze its internal structure. The user may not use or distribute the application for any unlawful purpose. The sale of copies of the product is strictly forbidden. No payment is allowed to get for using or distributing this software.

The authors are not responsible for any damages whatsoever, including loss of information, interruption of business, personal injury and/or any damage or consequential damage without limitation, incurred before, during or after the use of this product.

The user of Scans Searcher will indemnify, hold harmless, and defend the authors against lawsuits, claims, costs associated with defense or accusations that may result from the use of the software.

Short version. No income for you. No glory for you. Just use it and love it.

Setup

Software usage scheme

network
 ↓ *LAN scanner*
 scan.exe workdir\\$.dat
 text scans
 ↓ *database builder*
 MakeDB.exe workdir*.dat scans.db
 database
 ↓ *searcher*
 searcher.exe
 results

How to install

No installation process is required.
 The program may run directly from the network,
 if the database is also shared.

How to uninstall

Before deleting the files manually, you may want
 to remove saved settings from the registry. To do
 so, use the *delete* button on *Options* page.

Automatic scheduling

There is a known problem with user rights when using the WinXP scheduler. Running the scanner as administrator may fail to access some computers. Running the scanner as guest may not be able to access own disk. While no better solutions exist, use Lodzha's FastScan instead of LAN Scanner for automatic scheduling. (FastScan is not included in the original Scans Searcher package.)

System requirements

minimal:

- 100 MHz processor
- 32 MB RAM
- Windows-compatible OS

recommended:

- 1 GHz processor
- 128 MB RAM
- Windows 2K or newer

List of files

searching: *searcher.exe*
scans.db
 database creation: *scan.exe*
makedb.exe
 documentation: *scans.htm*
scans.pdf

 alternative *fastscan.exe*
 scanner: *mfc70.dll*
msvcr70.dll

Searcher

Searcher is the main program which searches for files using the given criteria. Instead of accessing the network directly, it uses a compressed database file. Thus, searching is completely independent of the network connection and can be performed even offline.

Usage: **searcher.exe [database file]**

By default, the database file is *scans.db* in the program's directory.

Pressing F1 at any time opens HTML help for Searcher, located in the program's directory.

F1 is the shortcut for the global popup menu's *help* command.

Page 'Search'

Search for strings

Words or phrases you want to search for in the directory and file names. Only those lines will be included in the result list which contain all the given strings, each one at any position (overlapping is allowed).

In the editbox, strings must be separated with single or more spaces. To include the space character in a phrase, surround that phrase by quotation marks "...". To include the quotation mark character, type it twice and surround the phrase in quotation marks. At most 31 strings can be specified.

Extensions

File types you want to search. Files whose extension is not in the list will be ignored. In the editbox, separate the extensions with single or more spaces. Do not include the period (.) to the extensions. Do not use wildcards. The number of specified extensions is not limited. To search all files, leave the editbox empty. Custom extension groups can be defined on the *Options* page.

Movie list

Creates list of movies. No settings have influence on movie list creation except the computer lifetime option. Directories are added to the list according to their name and the number of video and subtitle files they contain. (The method is quite precious but not sure.) Search results and movie list are listings of the same kind; thus, any of them will clear the contents of the *Results* page.

Name matching options

directory names only

Searches for the strings in directory names only. If extensions are specified, only the names of those directories will be searched which contain at least one file with a suitable extension.

whole words

Searches for the strings as whole words. The strings are found only at positions where they are surrounded by non-alphanumeric characters. In this context, native characters such as *é* are also treated as alphanumeric.

case sensitive

Makes the searching case sensitive. Searching for *hura* will not find the file that was scanned as *HurA* and vice-versa. Since FAT and NTFS file systems are case insensitive, this switch has no greater meaning, but it is supported.

native symbols

Makes the searching sensitive to interpunctual symbols. Searching for *hura* will not find the file named *hurá* and vice-versa.

Page ‘Results’**Menu button***create movie list*

Does the same as the *Movie list* button on *Search* page.

copy all to clipboard

Copies all lines to the clipboard. You can then paste the list in another application.

save to file

Saves the contents of the listbox to a text file. You will be asked to enter the filename each time you want to save the results.

font

Sets the font of the results window.

ignore ping result

For unknown reason, when searcher is running on specific machines, pinging may report all the computers in the network as offline. If checked, this option gives you a chance to run the *open* and *explore* commands even for offline items. However, if the computer is really not accessible, it may take several seconds to recover from the halted state. You should have no reason to use this option unless you often see all the computers in red.

Context menu*open*

Opens the file with the associated program. This command is only visible for files, not for directories. If *ignore ping result* is not checked, the command will not be executed for offline items.

explore

Opens the directory of the selected file using the default or custom browser. If *ignore ping result* is not checked, the command will not be executed for offline items.

copy directory

Copies the directory part of the selected line to the clipboard (without trailing backslash). You can then paste the directory name in another application.

copy string

Copies the selected line to the clipboard. You can then paste the filename in another application.

Line number reduction

To sufficiently reduce the number of lines in the list, some lines are grouped together and only the directory name is listed in the results. Such redundancy is made in two cases:

a) no strings occur deeper

This happens when all the requested strings appear in a directory name or its path, but none of them appear in any of the subdirectories nor in the files inside that directory. In such case, only the directory will be listed. Of course, every directory that is listed in the results contains at least one file that has the requested extension.

b) already in list

If a directory name with its path contains all the given strings and some of its subdirectories or files already appear in the results, the directory will not be listed again.

The purpose of line reduction is to avoid boring lists of files at unimportant depth. The principle of line reduction is to minimize the result while still holds that if a fully qualified filename contains all the strings and has an appropriate extension, either itself or one of its superdirectories is listed.

Coloured lines

The lines are drawn in multiple colours to indicate the status of the computer to which the line belongs. A check is made every 3–5 minutes. Possible states are:

pinging

The computer is being queued for the pinging thread. These lines are going to be repainted soon.

online

The file or directory is currently accessible and can be opened.

offline

The file or directory is currently not accessible and can not be opened.

special entry

The line contains a colon (:). This is a tricky way to include additional (given) lists in the database like FTP or CD/DVD scans.

If there are any problems with colouring or you just want everything in default colour, turn the checkbox off.

Page ‘Statistics’

Database properties

These are: the filename, the file size, the file’s last modification (when the database was built) and the database timestamp indicating the time of the latest scan.

Computers (bars)

computers / within x lifetime

The total number of computers and the number of computers that fall into the specified lifetime limit. Computers whose lifetime has expired will not be searched. See the *lifetime* option on *Options* page.

IP addresses known

Pinging through IP addresses is much faster than doing so using computer names. Knowing more addresses results in faster computer status detection. IP addresses are collected by the database builder and are stored in the database.

online – offline

The purpose of computer status detection is to mark those lines in the result list which contain links to currently accessible files and directories. Status detection is performed by a low-priority pinging thread that runs in the background unless the status of all computers are known. Pinging restarts about every 3–5 minutes.

Computer statistics table

The computer statistics table shows summary information about scanned data. Click on a header to sort lines by that column. Column meanings are described below:

<i>computer</i>	name of the computer
<i>status</i>	@ means online, — means offline
<i>days</i>	age of scans for the given computer
<i>IP address</i>	IP address saved in database, empty if unknown
<i>files</i>	total number of files
<i>dirs</i>	total number of directories
<i>max (~)</i>	maximal number of objects in a directory
<i>avg (~)</i>	average number of objects in a directory where objects are (<i>f</i>) files, (<i>d</i>) subdirectories*, (<i>*</i>) whatever

At runtime, the database is not refreshed. Only the online status of computers is checked periodically.

Page ‘Options’

Lifetime

Lifetime option serves to ignore computers with too old scans. If the time elapsed since the

* Zeros don't count when evaluating average number of subdirectories.

computer was scanned until the database was built is more than the specified lifetime, the computer will not be searched.

Custom browser syntax

User-defined syntax for the explore command. The target path will be inserted at the end unless %1 is used as the insertion point. The target always ends with a backslash. If the box is left blank, %SystemRoot%\Explorer.exe is used.

Predefined extension list

This is where you can edit predefined extension groups for the extensions entry on Search page. Each line contains one group in following form:

group name = ext1 ext2 ext3 ...

The left side is going to be shown in the combobox, while the right side will be pasted on request into the extensions input area. Spaces next to the equal sign are optional and don't count. Do not include the period (.) to the extensions. Do not use wildcards. Use empty right side for all files.

Keep in mind that without saving settings, any changes will be lost when the program quits.

View documentation

Opens the PDF documentation you are reading, located in the program's directory.

Window position / Settings

save settings (or window position) on exit

Saves actual settings (or window and splitter position) each time the program is being closed. Clearing this box will not remove already saved settings.

load now

Loads saved settings (or position) immediately. Anyway, settings are automatically loaded each time when the program starts.

save now

Saves current settings (or position) immediately, leaving the *save on exit* option unchanged.

restore default

Restores default settings (or position) by deleting related fields in the registry.

Registry

registry entry

Indicates the presence of the following registry key:

HKEY_CURRENT_USER\Software\Thomas Lackó\Scans Searcher

By default, registry key is not created unless you explicitly allow saving.

delete

Deletes the registry key and turns the two *save on exit* options off. Use this if you want to remove the software's hidden fragments from the system.

Scanner & Database

LAN Scanner

LAN Scanner collects information about shared files in a local area network. One text file is created for each computer.

Usage: **Scan.exe [option=value] [option=value] ...**
for which options are described below.

Example: **Scan.exe datadir\$.txt report=pe "format=f ** s dd-mm-y" fileattr+=h pause**
where *datadir* is the name of the directory to put the files in.

<i>option</i> (long / short)		<i>description</i>	<i>default value</i>
comp	c	computer name to scan (can use more times)	
complist	cc	computer list file name (can use more times) use ? to autodetect additional computers	?
write	w	output filename, " write= " may be omitted \$ will be replaced with the computer's name	\$.
format	f	file report format elements have to be delimited with non-letter symbols <i>f</i> : path+filename, <i>s</i> : size, <i>a</i> : attributes, <i>d</i> : day, <i>m</i> : month, <i>y</i> : year, <i>h</i> : hour, <i>n</i> : minutes, <i>x</i> : seconds more letters for more digits	f
dirformat	df	directory report format (-"-)	inherit: <i>format</i>
report	r	line report filter <i>f</i> : filenames, <i>p</i> : files with full path, <i>d</i> : all dirs, <i>e</i> : empty dirs	fe
filemask	fm	file name filter (classical wildcarded mask)	*.*
dirmask	dm	directory name filter (-"-)	*.*
fileattr+ fileattr-	fa+ fa-	file attribute filter (+ must / - mustn't be set) <i>a</i> : archive, <i>r</i> : read only, <i>h</i> : hidden, <i>s</i> : system, <i>l</i> : symbolic link	... h
dirattr+ dirattr-	da+ da-	directory attribute filter (-"-) (<i>dir</i> attribute handled automatically)	inherit: <i>fileattr</i>
threads	th	number of scanning threads (affects CPU and memory usage)	16
timeout	to	online detection timeout in msec	500
pause	p	delay after completion in seconds negative or missing value will wait for keypress	0

The parameter not containing equal sign but containing **\$** is treated as the *write* option.
With the exception of *comp* and *complist*, each option may be set no more than once.
Parameter *pause* (if set correctly) is applied always, even for error messages.

Lodzha's FastScan

Lodzha's Fast Network Scanner does not offer so wide set of options as LAN Scanner. However, its access method is more appropriate for automatic scheduling under WinXP. Unfortunately, no contact to the author is known.

Usage: **FastScan.exe <filemask> [computer list]**

where filemask is typically *.* and computer list is a text file containing names of computers, one in each line. The syntax is only guessed, no documentation is available. The following two calls should conceptually make the same result:

```
FastScan.exe aaa bbb  
Scan.exe write=$ filemask=aaa complist=bbb
```

Database builder

Database builder creates a compressed data file from text files, which will supply the searcher with data.

Usage: **MakeDB.exe <text scans [another another ...]> [option=value] [option=value] ...**

Example: **MakeDB.exe datadir*.txt c:\extra-data.txt scans.db ip=addresses.txt**

<i>option</i>	<i>description</i>	<i>default value</i>
db	Output filename. " db= " may be omitted if filename ends with ".db".	scans.db
ip	Address list. Text file containing "name IP-address" pairs. Will be created if doesn't exist. " ip= " may be omitted if filename contains string "ip" and ends with ".txt".	

Address list is only used to retrieve IP address of computers that are offline when the database is being built. The file will be created if doesn't exist. Under NT-based Windows, administrator rights are required to allow the database builder detect new or changed IP addresses.

Example batch file

Assume that the program files are in the same directory as the batch file and that there's a *datadir* directory where the text scans are to be placed.

Batch file for user's command

```
scan.exe datadir\$.txt  
makedb.exe datadir\*.txt scans.db ip.txt
```

Batch file for automatic scheduling

```
cd datadir  
..\fastscan.exe *.*  
cd ..  
makedb.exe datadir\*. scans.db ip.txt
```

Version history

Searcher

15 Feb 2004 Sun	0 400?	<ul style="list-style-type: none"> – database format 1 (list of fully qualified filenames) – searching for a substring in filenames, listing all matching files – choosable case sensitivity
16 Feb Mon	1 ...	<ul style="list-style-type: none"> – for directory name matches, the directory is listed only once (sufficiently reduces lines in the result) – listbox contents can be saved to file – whole words option – custom database allowed as command-line parameter
17 Feb Tue	2 ...	<ul style="list-style-type: none"> – info box shown while loading database (no progress bar yet) – whole words option treats central-european symbols as chars
17 Feb Tue	3 432 K	<ul style="list-style-type: none"> – filename suffix checking if any suffix is specified, all other files will be ignored
23 Feb Mon	4 ...	<ul style="list-style-type: none"> – database format 2 (tree structure) – native symbols option (interpunctual signs are now ignored by default) – improved method for line reduction – database timestamp is indicated as the time of the last scan instead of the database creation time
26 Feb Thu	5 ...	<ul style="list-style-type: none"> – superfast searching (pattern matching and string conversion in assembly, string variables optimized for speed) – computer names uppercase – database builder detects errors in scan files
2 Mar Tue	6 450 K	<ul style="list-style-type: none"> – more file extensions supported, predefined list of extension groups (built-in) – architecture split into: user interface + searcher engine
		Skinned version. Corporate author DJ for design and image editing.
8 Mar Mon	7 1100 K?	<ul style="list-style-type: none"> – new graphical design: image-skinned controls, unregular form under Win2K – <i>Search</i> and <i>Results</i> pages
18 Mar Thu	8 1118 K	<ul style="list-style-type: none"> – double click to open a file from results – minor visual modifications – predefined file types extended by <i>subtitles</i> – component visibility bug corrected for task switching
28 Mar Sun	9 1150 K	<ul style="list-style-type: none"> – rectangular window while moving with right button or resizing (unregular shape caused flipping on some computers) – double click now works on found directories as well – ping for double click to reduce reply time for offline comps – window movement bugfixes

15 Apr Thu	10 1112 K	<ul style="list-style-type: none"> – database format 3 (IP addresses) – recognizing online computers (coloured lines in result list) – ignoring computers with too old scans – setting window region partially through resource handles (unregular form also under Win98) – keeping the taskbar visible when maximized – never fullscreen – progress bar while loading database – other minor visual modifications
16 Apr Fri	11 1118 K	<ul style="list-style-type: none"> – <i>Statistics</i> page – general database info (solid background) – found and fixed the root of the flicker problems (window region through resource handles, no colour transparency)
18 Apr Sun	12 920 K	<ul style="list-style-type: none"> – more strings to search for – option to match directories only – statistics page has transparent background
19 Apr Mon		Documentation written and published in HTML format.
9 May Sun	13 935 K	<ul style="list-style-type: none"> – computers are pinged repeatedly in 5 minute intervals – context menu in result list, ability to copy a line to the clipboard – error text shown if the file couldn't be opened – fixed bug in searching whole words
9 May Sun	n/s n. 627 K s. 935 K	Project split into normal and skinned version. Functional (non-visual) new features evolving paralelly in both versions.
10 May Mon	14.n 628 K	– visual modifications, bars in statistics window
18 May Tue	15.n 628 K	– little visual modifications
23 May Sun	16 n. 680 K s. 992 K	<ul style="list-style-type: none"> – loading database in background – statistics table with max/avg values (analyzer starts on first request) – extensions can be surrounded by quotation marks – (16.s: visual modifications, bars in statistics window)
23 May Sun	17 n. 680 K s. 991 K	<ul style="list-style-type: none"> – faster searching (no memory allocation during extraction) – much faster statistics analyzer (starts in the background immediately after loading)
10 June Thu	18 – " –	<ul style="list-style-type: none"> – fixed thread synchronization bug (threw "invalid database" if scanning started too early)
		Skinned version is no longer evolved due to visualization problems.
29 Sep Wed	19 600 K	<ul style="list-style-type: none"> – menu button in results window, ability to copy all lines to the clipboard – shifted positions for popup menus so that <i>explore</i> and <i>copy all</i> can be clicked more comfortable – <i>view documentation</i> button

4 Oct Mon	20β 625 K	<ul style="list-style-type: none"> – movie list – specialized list of directories containing videos (double click and popup menu work as required) <i>version marked as beta due to low optimization</i> – visual modifications (some components moved)
21 Nov Sun	21 655 K	<ul style="list-style-type: none"> – options page, saving settings to registry – user-defined extension groups – movie list optimization improved – custom font for results – <i>font</i> command in menu – <i>copy directory</i> command in context menu – doubleclick has much faster response for offline computers – statistics table: <ul style="list-style-type: none"> - columns <i>online status</i>, <i>age of scan</i> and <i>IP address</i> - sorting by any column - fixed bug in displaying max/avg number of “whatever” – changed lifetime setting controls (trackbar instead of combobox, visual percent bar)
28 Jan 2005 Fri		LAN Scanner released. Scans Searcher no longer depends on external scanners. File <i>scans.exe</i> renamed to <i>searcher.exe</i> to avoid naming similarities.
25 Feb Fri	22 660 K	<ul style="list-style-type: none"> – faster executing and exploring on doubleclick (calling <i>CreateProcess</i> instead of <i>ShellExecute</i>) – shortcut for <i>explore</i>, <i>copy string</i> and <i>copy directory</i> commands – search button disabled while loading (searching too early might halt on some computers) – bugfix: right-clicking below the last line in results box threw an exception
11 Mar Fri	23 662 K	<ul style="list-style-type: none"> – custom browser for the <i>explore</i> command – age of scans shown in status bar (“<i>n days ago</i>” text) – window position saving is now optional – line colouring can be turned off – bugfix: searching could be launched by hitting the Enter key even in disabled state
29 Apr Fri	24 662 K	<ul style="list-style-type: none"> – <i>ignore ping result</i> option in menu – <i>copy all to clipboard</i> now uses CR+LF at line ends – bugfix: search button now enabled while loading, should not halt
1 Jun Wed	25 662 K	<ul style="list-style-type: none"> – special entries containing colon (:) in green
6 Jun Mon	26 667 K	<ul style="list-style-type: none"> – <i>clear results</i> menu item (just to offer one more command J) – computer online status updated when not responding – bugfix: left-clicking below the last line in movie-computers box threw an exception
20 Jan 2006 Fri		Documentation available in PDF format (copied from HTML). HTML file renamed and reduced to Searcher Help only.

8 Feb Wed	27 668 K	<ul style="list-style-type: none"> – pressing F1 opens HTML help – <i>view documentation</i> button opens PDF instead of HTML – safety bugfix: empty filename in database could throw an exception
--------------	--------------------	---

LAN Scanner

16 Dec 2004 Thu	β 153 K	<i>prototype</i> <ul style="list-style-type: none"> – multi-threaded scanner with plenty of command-line parameters – fully qualified filenames, no grouping
27 Jan 2005 Thu	1 155 K	<ul style="list-style-type: none"> – zero-sized files are not written – more detailed summary information
11 Feb 2006 Sat	2 158 K	<ul style="list-style-type: none"> – advanced command-line handling, incorrect parameters are reported – hidden files and directories are not reported by default – attribute options extended to “+” and “-” – <i>pause</i> option – bugfixes (some parameters had no influence)
12 May 2006 Fri	3 160 K	<ul style="list-style-type: none"> – grouped filenames with no path (default report filter) significantly reduces size of generated text files – fix: empty shared root folders were not reported

Database builder

15 Feb 2004 Sun	1 90 K	<ul style="list-style-type: none"> – database format 1
23 Feb Mon	2 95 K	<ul style="list-style-type: none"> – database format 2 – tree structure – better compression
15 Apr Thu	3 97 K	<ul style="list-style-type: none"> – database format 3 – IP address and timestamp for computers – memory requirements slightly reduced
17 May Mon	3.1 99 K	<ul style="list-style-type: none"> – using address list file to retrieve IP address of computers that are offline during the building process
28 Jan 2005 Fri	3.2 99 K	<ul style="list-style-type: none"> – bugfix: output file was always <i>scans.db</i> – case sensitive sorting (makes a very little higher compression)
17 Feb 2006 Fri	3.3 101 K	<ul style="list-style-type: none"> – advanced command-line handling, parameter syntax changed – empty lines and lines starting with colon (:) are safely ignored
3 Mar Thu	3.4 102 K	<ul style="list-style-type: none"> – more comfortable command-line parameters (auto-recognizing <i>db</i> & <i>ip</i>) – lines starting with semicolon (;) are treated as comments

12 May Fri	3.5 102 K	<ul style="list-style-type: none"> – entries with no path inherit path from previous line – forward slashes are converted to backslashes, safe double slashes in custom ftp lists – tab characters are cut from both sides
---------------	---------------------	---

Database format

Size ratio, memory requirement and building time statistics are valid for cca **100 computers**, total 100 MB of input text files, files per computer: 5000–10000, max. 70000, built on 2.4 GHz Pentium 4.

<i>version</i>	<i>release</i>	<i>structure</i>	<i>compression</i>	<i>ratio</i>	<i>building time</i>	<i>building mem.</i>
1	15 Feb 2004	list of full filenames	twoside differential	19%	1 min	8 MB
2	23 Feb 2004	directory tree for network	differential with extension list	16%	1½ min	70 MB
3	15 Apr 2004	directory tree for computers, IP addresses	differential with extension list	16%	5 min for 15% online	6 MB

Future plans

Complex changes

1. support for special entries (anything better than the actual „colon“ version)
2. approximative file sizes and dates
3. better database compression
4. scanning local disks

If ever, all of them should be solved at once by a new database builder.

Simple changes

1. doubleclick on movie title to open from any available computer
2. drag&drop to copy files
3. regular expressions
4. Linux-compatible version
5. better optimization for the movie list
6. ...file attributes
7. ...extracting text scans from database (reverse builder functionality)