

Project 'GAMERA'

(Semi-Powerful Console(Windows & Linux) Tools & gigabytes of English texts, downloadable from www.sanmayce.com)

WHERE THE WORD COUNTS



Caterpillar(LZSS-King of Brute-Force Heavy Sentence Dumpers, 32bit console application), revision 14+
Kazuya(LZ-Sovereign of Brute-Force Heavy Sentence Dumpers, 32bit console application), revision 17++
Salah-ed-din(GZ-Sultan of Brute-Force Heavy Sentence Dumpers, 32bit console application), revision 14++
Raccoondog(LZMA-Baron of Brute-Force Heavy Sentence Dumpers, 32bit console application), revision 17++
Yoshi(Filelist creator and more, 32bit console application), revision 06
Leprechaun(Fast and Greedy Word_Ripper, 32bit console application), revision 13++

WinRAR archive in eleven 624MB volumes • Required HDD space: 6.56 GB (*ready to go when extracted on D:*) • 2010 JUN 06

Kazuya delivers english sentences at 85-255MB/s speed(Obtained with Toshiba Satellite L305 (Intel Pentium(Merom-1M) T3400 2.16GHz))
Salah-ed-din delivers english sentences at 114-117MB/s speed(Obtained with Toshiba Satellite L305 (Intel Pentium(Merom-1M) T3400 2.16GHz))
Raccoondog delivers english sentences at 39MB/s speed(Obtained with Toshiba Satellite L305 (Intel Pentium(Merom-1M) T3400 2.16GHz))
Leprechaun rips 6,142,696++ words per second(Obtained with Toshiba Satellite L305 (Intel Pentium(Merom-1M) T3400 2.16GHz))

LBL stands for Line-By-Line(GRAMMATICAL ENGLISH LINES) i.e. sentences not merely CRLF or LF lines!
.LBL files are made from .TXT files which are made from respective .DOC, .RTF, .LIT, .PDF, .CHM, .HTM[L], .DJV[U] files;
Number and size of *.LBL files: 562,504 files(26GB or 27,991,747,152 bytes);
Lines and words in *.LBL files: 424,754,717 lines(with 4,582,451,898 words of them 9,177,221 distinct);

'Monstrous Dumpers' package, revision 13-

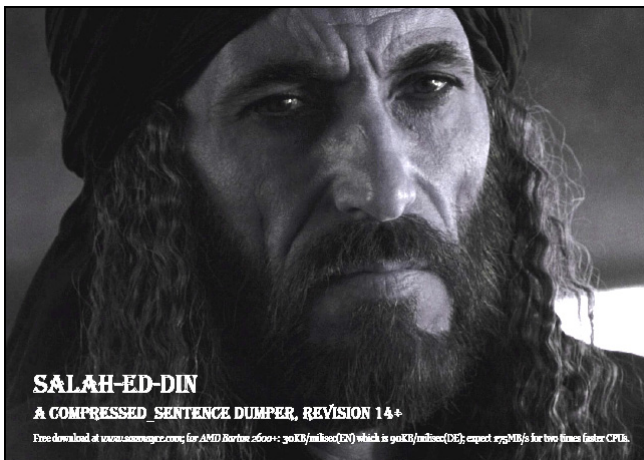
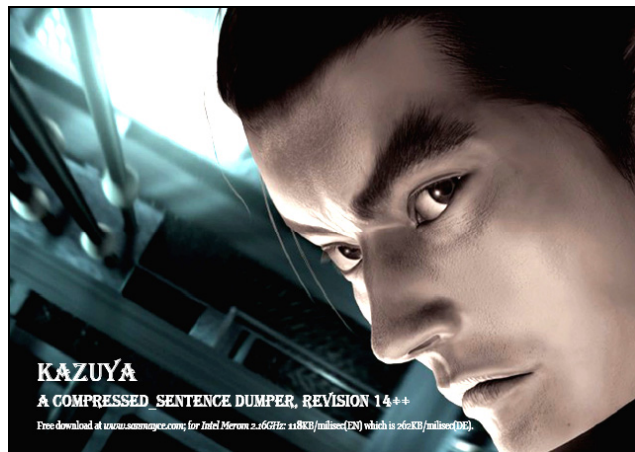
OR

How words can be mixed into sentences!?

With this package(the main part of project "GAMERA") you can make full-text(brute-force) requests into millions of lines(sentences). For example: make a search for **requests????????????into** to see whether that preposition has place near on right side of "requests". This package(a winrar archive) is intended as shareware and contains six very fast 32bit console text tools: **Caterpillar** (its rivals **Raccoondog**, **Salah-ed-din** and **Kazuya**), **Leprechaun**, **Yoshi** and of course 100++ million sentences(in English language) from various sources.

The package allows easily to create:

- a FILELIST(a text file with filenames);
- a WORDLIST(a text file with sorted distinct words);
- and as a main feature a text-pattern to be searched into LF(Unix)|CRLF(Windows) lines(or files) via filelist and to dump resultant hits(lines or filenames) into .HTML file.



Main features:

- 26GB english-ASCII-texts converted to .LBL(same as .TXT but each line is a sentence) format;
- File-by-file listings of all texts included:
 - 164,128 KAZE_G.S._Corpus_'.chm'_Caterpillar.html
 - 16,817 KAZE_G.S._Corpus_'.djb-1'_Caterpillar.html
 - 61,142 KAZE_G.S._Corpus_'.doc'_Caterpillar.html
 - 22,742,832 KAZE_G.S._Corpus_'.htm-1'_Caterpillar.html
 - 32,717 KAZE_G.S._Corpus_'.lit'_Caterpillar.html
 - 1,905,255 KAZE_G.S._Corpus_'.pdf'_Caterpillar.html
 - 30,535 KAZE_G.S._Corpus_'.rtf'_Caterpillar.html
 - 27,389,047 KAZE_G.S._Corpus_'.txt'_Caterpillar.html
- [Five(*,@,#,\$,%):Kazuya & Raccoondog] wildcards(*,?) available for patterns: very slow(pattern *underdog* took 490 seconds to look for into 400+ million sentences) but powerful;
- Fast(for laptops with a non-SSD disk)(90++MB/s) full-text traversing due to **zlib** used;
- Extra Fast(for laptops with a SSD disk)(200++MB/s) full-text traversing due to **quickLZ** used;
- Results are delivered as *screen output* immediately and as *pure HTML files* finally;
- 'Karp_Rabin_Kaze'(patterns *, **underdog** took 70 seconds to look for into 400+ million sentences) compared to 'strstr'^(85s) & 'Boyer-Moore-Horspool'^(86s) is $\frac{(85-70)}{70} * 100\% = 21\%$ faster when running on sentences.

Installation(i.e. extracting) notes:

- Unrar in D:\ if possible, "Caterpillar.lnk", "Go to PROMPT.lnk", "Raccoondog.lnk", "Salah-ed-din.lnk", "Kazuya.lnk" need manual adjustments if not D:\, 7GB must be free.
- To use "Caterpillar.lnk" and "Salah-ed-din.lnk" and "Kazuya.lnk" must run (R20.BAT) and (R2G.BAT) and (R2L.BAT) respectively.

Current revisions of tools:

- | | |
|-------------------------|-----------------------|
| - EXEs(Windows): | - ELFs(Linux): |
| Caterpillar r.14+ | Caterpillar r.14+ |
| Leprechaun r.13++ | Leprechaun r.13++ |
| Yoshi r.06 | Yoshi r.06 |
| Salah-ed-din r.14++ | Salah-ed-din r.14+ |
| Raccoondog r.17++ | |
| Kazuya r.17++ | |
- r.15 has an ability to search non-compressed files too!

Note1: Revisions 14++ are Experimental(but operational, not beta) Karp-Rabin function with my hash, see last page.
Note2: Predecessor of *Caterpillar*, *Salah-ed-din* & *Raccoondog* was **Kazuya**(with more functionality and critical parts written in 16bit assembler), someday I will resurrect him in 64bit.

Convert at will:

- Use G2R.BAT for .gz -> .lzma (1000+ minutes needed to convert, grmb1)
- Use R2G.BAT for .lzma -> .gz (11:05 PM - 12:18 AM i.e 73 minutes needed to convert)
- Use R2L.BAT for .lzma -> .Lasse (06:57 PM - 07:45 PM i.e 48 minutes needed to convert)
- Use R20.BAT for .lzma -> .Okumura (11:33 PM - 01:22 AM i.e 109 minutes needed to convert)

Some experience(Machine: Toshiba Satellite L305 - Intel Pentium Dual CPU T3400 @ 2.16GHz):

- **Caterpillar** uses **LZSS**(based on LZSS.C written by H.Okumura) compression;
24.9GB -> 11.4 GB (12,341,932,922 bytes);
delivering text at 82KB(*149KB when in system cache*)/clock i.e. 80MB/s;
suitable for FAST HDDS 80+MB/s.
- **Raccoondog** uses **LZMA**(based on LZMA SDK 4.65 written by I.Pavlov) compression;
24.9GB -> 5.5 GB (5,944,631,607 bytes);
delivering text at 40KB(*bottleneck is CPU power alone*)/clock i.e. 39MB/s;
suitable for flash cards like CFs, SDs.
- **Salah-ed-din** uses **GZ**(based on zlib 1.2.3 written by J.Gailly and M.Adler) compression;
24.9GB -> 8.4 GB (9,051,049,655 bytes);
delivering text at 117KB(*120KB when in system cache*)/clock i.e. 114MB/s;
suitable for FAST CPUs 3+GHZ.
- **Kazuya** uses **LZ**(based on quickLZ 1.4.0 written by Lasse Reinhold) compression;
24.9GB -> 10.6 GB (11,402,975,168 bytes);
delivering text at 88KB(*262KB(118KB(EN)) when in system cache*)/clock i.e. 85MB/s;
suitable for FAST SSDS 115+MB/s. Near future dreams: CPU(2x faster) and SSD(2x115Mb/s read) will give 2x255Mb/s.

Search|Seek|Find in order to Explore|Learn|Avoid Different Styles:

["Супруга съм на три деца. С чувекъ сбрахме пари и купихми триустаен партамент. Една вечер звъни вратата. Звънецъ чука. Отварям - НИНДЖА. И без да каже нищо, с карате в бъбреците. Дукат съ усета ми би два шамара с КРАК и един на детето в гръбначнийъ кош! От ударната вълна отлпихам на 20-30 метра. Абстрахираха децата. А чувекъ го нема. Ако общината в града не вземе спешни мерки, ще се самубеся илиъ ще изчезна безкрайно."]

/Интервю с ромка излъчено по КАНАЛ 1 за акция на НСБОП по залавяне на опасни рецидивисти в Пазарджишко./

Enjoy!
Sanmayce 'Kaze', 2009 Mar 13.

D:_KAZE_G.S._Corpus>yoshi

Yoshi(Filelist Creator), revision 06, written by Svalqyatchx,
in fact based on SWEEP.C from 'Open Watcom Project', thanks-thanks.

- Note1: So far, it works for current directory only.
Note2: Default method is depth-first traversal;
may use pipe 'Yoshi|sort' for breadth-first_like traversal results.
Note3: Make notice that '*.*(extensionfull only) is not equal to '*'(all);
one disadvantage is an inability to list only extensionless filenames.
Note4: Search is case-insensitive as-must.
Note5: This revision allows multiple '*', and meaning of masks is:
'?' - any character AND NOT EMPTY(default, for OR EMPTY see option -e);
'*' - any character(s) or empty.
Note6: What is a .LBL(LineByLine) file?
it is a bunch of GRAMMATICAL lines not mere LF or CRLF lines;
it contains not symbols under 32(except CR and LF) and above 127;
it contains not space symbol sequences.

Usage:

```
Yoshi [option(s)] [filename(s)]
option(s):
  -v          i.e. verbose mode; output goes to console;
  -f          i.e. fullpath mode for output;
  -e          i.e. treat '?' as any character OR EMPTY;
  -t          i.e. touch all encountered files;
  -2          i.e. convert all encountered .TXT files to .LBL files;
  -o<filename> i.e. output goes to file(in append mode).
filename(s):
  wildcards '*' and wildcards '?' are allowed i.e. "str*.c?";
  default filename is '*'; DO NOT FORGET TO PUT
  filename(s) WITH WILDCARD(S) INTO QUOTE MARKS!
```

Examples:

```
Yoshi -v -f -oCaterpillar_NON.lst "*.lbl" "*.txt" "*.htm" "*.html"
Yoshi -f -omyebooks.txt "*wiley*essential*.pdf" "*russian*.htm"
```

Yoshi: Total size of files: 00,027,750,342,332 bytes.
Yoshi: Total files: 000,000,001,088.
Yoshi: Total folders: 0,000,000,003.

D:_KAZE_G.S._Corpus>"Leprechaun_r13+_32bits.exe"

Leprechaun(Fast Greedy word-Ripper), revision 13+_, written by Svalqyatchx.
Leprechaun: 'Oh, well, didn't you hear? Bigger is good, but jumbo is dear.'
Kaze: Let's see what a 4-way hash + 6,602,752 Binary-Search-Trees can give us,
also the performance of a 4-way hash + 6,602,752 B-Trees of order 3.

'The Little Monster' short notes:

- Note1: I wish to thank to R.N. Horspool, Ranjan Sinha, Dmitry Shkarin,
Michael Abrash, J. Bentley, R. Sedgewick, Igor Pavlov, Lasse Reinhold
for sharing their knowledge to public.
Note2: Run it without parameters to get usage and short notes.
Note3: This simple amateurish(more over I am not versed well neither in C nor
in mathematics nor in english language, but I am persistent in INDEXING
GBs of english TEXTS) tool is written in ANSI C(at least its source is
compileable for CL(Windows) and GCC(Linux)), and its purpose is to
create a WordList for a group of files(given via filelist).
Its name comes(according to Heritage Dictionary) from 'low corpus' or
'little body', in fact from amazing movie saga 'Leprechaun 1-2-3-4-5-6'
starring by Warwick Davis.
Note4: Only words up to 31 chars are proceeded - the reason is 'DDT'(the
longest word in Heritage Dictionary 3rd edition) or
'dichlorodiphenyltrichloroethane'.
Note5: Cursor hiding in C - mission impossible for me.
Note6: By default(third parameter is 1023) allocated memory is 393MB.
Due to 'malloc()' limitation under WINDOWS, maximum value of third
parameter is 5174 which is 1988MB allocated block.
Note7: File Leprechaun.LOG is a log, where new statistics are appended.
Note8: Revision 12+ can handle files larger than 4GB.
Note9: Revision 12++ has a buffered 'fread()' - therefore I/O READ-BURST SPEED
is the first(worst) bottleneck, as a result r.12++ is much-much faster;
the second(worse) bottleneck: the linked lists - the b-trees
might be the answer; the third(bad) bottleneck: the amateurish author.
NoteA: Revision 12+++ has an improved(2 bits were used doltishly) main hash
function - therefore less collisions, for example:
for file 'wikipedia-de-html.tar' 42,291,855,360 bytes with
5,750,179,678 words of them 7,375,373 distinct attempts to Find/Put
a WORD into a linked list are 6,117,675,470(r.12++) and 5,845,989,790
(r.12+++); also two 'if' sections were moved because they were executed
unnecessarily many times.
NoteB: Revision 13 uses BSTs instead of LLS, that is Linked-Lists were
replaced by Binary-Search-Trees, as a result for 22,202,980 distinct
words(out of 35,271,297) r.12+++ needs 225,548,268 total attempts to
Find/Put WORDS into linked lists where r.13 needs 121,674,042 total
attempts to Find/Put WORDS into Binary-Search-Trees. But this is a
significant boost in performance only for wordlists of million words.
NoteC: Revision 13+ gives only more statistics. Future revisions could lessen
number of attempts to Find/Put WORDS into Binary-Search-Trees
furthermore by making them at some point Perfectly-Balanced. But
for huge amount(multi-(m)ibillion) of distinct words the b-tree family
must come in, until then this is the leprechaunish niche.
NoteD: Revision 13++ has a little fix(2 unnecessary ZEROings, when a new word
is inserted, were deleted) and a fixed bug(13+ adds stupidly the
highest BST to the wordlist). Also B-Tree of order 3 is added as a
searching method. Main goal of B-Tree is to reduce number of
comparisons but at nasty cost: a precious time wasted to construct it
and twice more memory, i.e. one step forward two backward: this tree is
more effective than BST in cases of 2++ billion/million
different/distinct words.
The improvement which comes from using B-Tree of order 3 is about 200%
much more pleasing than I expected, for wikipedia-en-html.tar.wrd with
12,561,874 distinct words Total Attempts to Find/Put WORDS into:
Binary-Search-Trees was 61,895,043 while for
B-trees order 3 was 19,295,791.
NoteE: For old r.12+ a USB connected HDD crippled test:
for 'H:\>Leprechaun.exe static.wikipedia.org_downloads_2008-06_en.lst
wikipedia-en-html.tar.wrd 5400'
where 223,674,511,360 wikipedia-en-html.tar
on laptop Toshiba Pentium T3400 2166 MHz with
Motherboard Name: Toshiba Satellite L305
CPU Type: Mobile DualCore Intel Pentium, 2166 MHz (13 x 167)
CPU Alias: Merom-1M
L1 Code Cache: 32 KB per core
L1 Data Cache: 32 KB per core
L2 Cache: 1 MB (On-Die, ECC, ASC, Full-Speed)

```

Bus Type: Dual DDR2 SDRAM
Bus width: 128-bit
Real Clock: 333 MHz (DDR)
Effective Clock: 666 MHz
EVEREST v5.00.1650 Memory Copy: 3725MB/s with timings 5-5-5-13
result is logged to 'Leprechaun.LOG':
Bytes per second performance: 20,658,955B/s
Words per second performance: 2,860,880W/s
Input File with a list of TEXTual Files:
static.wikipedia.org_downloads_2008-06_en.lst
Size of all TEXTual Files: 223,674,511,360
Word count: 30,974,750,142 of them 12,561,874 distinct
Number Of Files: 1
Number Of Lines: 2088618575
Allocated memory in MB: 1920
Words with length 01 occupy 0,033KB of 0,349KB given i.e. 09% utilization
Words with length 02 occupy 0,033KB of 0,349KB given i.e. 09% utilization
Words with length 03 occupy 0,037KB of 0,697KB given i.e. 05% utilization
Words with length 04 occupy 0,151KB of 0,871KB given i.e. 17% utilization
Words with length 05 occupy 0,744KB of 1,568KB given i.e. 47% utilization
Words with length 06 occupy 1,470KB of 3,136KB given i.e. 46% utilization
Words with length 07 occupy 2,605KB of 5,923KB given i.e. 43% utilization
Words with length 08 occupy 3,296KB of 6,968KB given i.e. 47% utilization
Words with length 09 occupy 3,714KB of 6,968KB given i.e. 53% utilization
Words with length 10 occupy 3,483KB of 6,968KB given i.e. 49% utilization
Words with length 11 occupy 3,235KB of 5,923KB given i.e. 54% utilization
Words with length 12 occupy 2,691KB of 4,181KB given i.e. 64% utilization
Words with length 13 occupy 2,230KB of 3,484KB given i.e. 64% utilization
Words with length 14 occupy 1,718KB of 3,484KB given i.e. 49% utilization
Words with length 15 occupy 1,357KB of 2,613KB given i.e. 51% utilization
Words with length 16 occupy 1,063KB of 2,613KB given i.e. 40% utilization
Words with length 17 occupy 0,814KB of 1,742KB given i.e. 46% utilization
Words with length 18 occupy 0,617KB of 1,742KB given i.e. 35% utilization
Words with length 19 occupy 0,485KB of 1,742KB given i.e. 27% utilization
Words with length 20 occupy 0,402KB of 1,742KB given i.e. 23% utilization
Words with length 21 occupy 0,327KB of 1,742KB given i.e. 18% utilization
Words with length 22 occupy 0,274KB of 1,742KB given i.e. 15% utilization
Words with length 23 occupy 0,224KB of 1,394KB given i.e. 16% utilization
Words with length 24 occupy 0,190KB of 1,394KB given i.e. 13% utilization
Words with length 25 occupy 0,162KB of 1,394KB given i.e. 11% utilization
Words with length 26 occupy 0,136KB of 1,220KB given i.e. 11% utilization
Words with length 27 occupy 0,119KB of 1,046KB given i.e. 11% utilization
Words with length 28 occupy 0,107KB of 0,871KB given i.e. 12% utilization
Words with length 29 occupy 0,091KB of 0,697KB given i.e. 13% utilization
Words with length 30 occupy 0,080KB of 0,523KB given i.e. 15% utilization
Words with length 31 occupy 0,076KB of 0,523KB given i.e. 14% utilization
Total pseudo(including hash table) memory utilization: 42%
Total real(wordlist's words VS allocated block) memory utilization: 60/1000
Used value for third parameter in KB: 5400
Use next time as third parameter: 3475-
Time for making unsorted wordlist: 10827 second(s)
Time for sorting unsorted wordlist: 10 second(s)

```

```

Usage: Leprechaun InFile OutFile [BufferSize] [SortMethod] [TreeMethod]
<InFile>: Input file with files for Leprechauning, in WINDOWS console
you can create it by 'E:\KAZEHOME>dir *.txt/s/b>Leprechaun.lst'
<OutFile>: Output WORDLIST(sorted since r.9, CRLF) file
<BufferSize>: Optional Dynamic RAM buffer in KB, default(and minimum
in the same time) is 1023, i.e. omit or specify greater one
<SortMethod>: Optional Sort Method, default is 'D',
A - InsertionSort
B - InsertionX26Sort
C - MultiKeyQuickSortSort by J. Bentley, R. Sedgewick
D - MultiKeyQuickSortX26Sort' by J. Bentley, R. Sedgewick
<TreeMethod>: Optional Tree Method, default is 'X',
X - Binary-Search-Trees
Y - B-Trees of order 3

```

Have a nice Leprechauning.
For contacts: sanmayce@hotmail.com
Sanmayce svalqatchx 'kaze', 2005 Feb 07(rev.13++: 2010 Apr 12).

D:_KAZE_G.S._Corpus>Caterpillar
Caterpillar(Sentence_Dumper), revision 14+, written by Svalqatchx,
in fact adapted from Haruhiko Okumura's excellent LZSS.C program.

How near are these words_forms to me: Masakari, Massacre, Steel-Coloss,
Monster-Truck, Dump-Mining-Truck, Caterpillar 797, Liebherr, Komatsu. They
resemble one thing: strong-devoid-of-ambition-power(i.e. a pure work/time).

'Caterpillar' is a simple pattern searcher(from 'Masakari' family tools)
into archived english-text files, designed to achieve up to 90% higher read
speed than the HDD READ BURST i.e. 'copy hugefile nul' gaining at same time
50% compression of searched data.

Its main feature is somewhat hidden nowadays, because of
pseudo-transparent decompression used, which leads to doubling(unreachable in
fact) uploaded data for search function(written by N. Horspool, thanks a lot)
due to LZSS algorithm implemented by H. Okumura(greetings to him). Okumura's
variant(HDD2RAM) which is much faster(!!!) and needs less memory than tuned
memory-to-memory decompression(RAM2RAM) variant. I am still stunned.

In few words: feeding search function is 100-% faster with very fast
CPU-Physical_RAM subsystems, in this way reducing the ugly penalty which comes
from reading a HDD. In numbers: me IDE HITACHI 7200rpm 2MB gives up to 60MB/s
READ BURST, 'Caterpillar' almost doubles(i.e. 120-MB/s) it in case of 3+++GHZ
CPU and 533+++MHZ RAM.

For windows 2003, VIA KT600, AMD XP 2500+(1836.12MHz=11x166.92MHz),
FSB 333.84MHz(2x166.92MHz), 512KB L2 cache, 1 DIMM DDR 512MB 333MHz(2x166MHz),
Caterpillar(in fact LZSS) decompresses 58,000KB per second i.e.
boost is negative: 60MB/s=61,440KB/s(READ BURST) is greater than 58,000KB/s.
But for two times faster CPU-RAM sub-system(SERVER) than described above OR
for two times slower HDD sub-system(LAPTOP) boost will be positive:
(1 - READ BURST SPEED / DECOMPRESSION SPEED) * READ BURST SPEED or
(1 - (61,440KB/s) / (2 * 58,000KB/s)) * 61,440KB/s = (0.471) * 61,440KB/s.

Since revision 5 'fread()' was changed with 'read()', for speed.

'The Monster-Dump-Truck' short notes:
Note1: Thanks a lot to N. Horspool, Dmitry Shkarin, H. Okumura, Igor Pavlov.
Note2: Run it without parameters to get usage and short notes.

Note3: Current revision searches only for case-sensitive and unexact matches.
 Note4: This simple amateurish (more over I am not versed well neither in C nor in mathematics nor in english language, but I am persistent in INDEXING GBS of english TEXTS) tool is written in ANSI C (at least its source is compileable for CL(windows) and not yet for GCC(Linux) because of 'O_BINARY in open(), gets(), getch(), kbhit()', and its purpose is to create a SentenceList for a group of compressed (with it) text files (LF and CRLF) given via filelist.
 Its name comes from a heavy-noprider-dumper-truck 'Caterpillar'.
 Note5: By default allocated memory is 95MB i.e. decoding is HDD2RAM.
 Note6: Disastrous performance in case 95MB|147MB not fully physical!
 Note7: For me digital library:
 where files are 54, Encoded 6,917,425,566, Decoded 14,419,485,826 with windows XP, VIA KT600, AMD XP 2500+(1836.12MHz=11x166.92MHz), FSB 333.84MHz(2x166.92MHz), 512KB L2 cache, DDR 512MB 333MHz(2x166MHz), IDE HDD Maxtor 80GB 7200 8MB and
 'D:\temp>dir E:\KAZEHOME\KAZUYA.O??\b>Caterpillar.lst'
 'D:\temp>Caterpillar Caterpillar.lst CaterpillarRAM2RAM.ini'
 result is: 282 seconds or 41000KB/s upload, 11000KB/s boost, 52000KB/s boosted upload, 56000KB/s decode.
 'D:\temp>Caterpillar Caterpillar.lst CaterpillarHDD2RAM.ini'
 result is: 142 seconds or 99000KB/s boosted upload!!!
 Note8: Matches(hits) containing neither '<' nor '>' are written to 'Caterpillar.hits.pattern?.html' file.
 Note9: Works both on UNIX(LF) and windows(CRLF) text files.
 NoteA: Never forget the importance of defragmented_AND_grouped files located at fastest area of disk - first partition is faster than second one, etc.
 NoteB: In ANSI, clock is defined as '#define CLOCKS_PER_SEC 1000'.
 NoteC: Since Caterpillar 13++:
 - limits (just skip longer ones) lines to 960 chars; OTHERWISE: HUGE TIME DELAYS due to recursive function;
 - shows hits to console too; MORE VIVID;
 NoteD: During execution hitting a 'Esc' causes termination (i.e. skipping rest).
 NoteE: At last NON-ENCODED regime has two modes: in addition to LINE (i.e. hits are lines) there is a FILE (i.e. hits are filenames) mode.
 NoteF: For all regimes files Caterpillar.HIT?.lst are created for each pattern (1,2,3 and 4) - containing hits filelist i.e. filenames containing HITS (either LINES or FILENAMES).

Below is LINE (default for DECODING ???2RAM regimes) mode pattern description:
 Pattern(s) note: You may specify (four times) a main-pattern (case insensitive with wildcards '*' i.e. any character(s) or empty and '?' i.e. any character or empty) with three nested-patterns (case sensitive and unexact), all four connected with AND.
 Due to different line endings (CRLF in windows; LF in UNIX) you must add a '?' wildcard in place of CR: for example in case of searching for '*.pdf' write '*.pdf?'.

Pattern(s) example: Pattern1: *take? *it*
 Pattern1_NestedPattern1: you
 Possible hit: ... your reason is so taken by It.

Usage: 'Caterpillar e file1 file2' encodes file1 into file2
 'Caterpillar d file2 file1' decodes file2 into file1
 'Caterpillar m ListOfFilesFile SolidSize'
 <ListOfFilesFile>: Files to be merged into Caterpillar.??? files
 <SolidSize>: Caterpillar.??? files size limit in MB.
 'Caterpillar ListOfFilesFile [OptionsFile]'
 <ListOfFilesFile>: Input file with files for Caterpillaring
 <OptionsFile>: Optional input file with options with following format:
 Optional line #1 contains method of decoding:
 'DECODING HDD2RAM' | 'DECODING RAM2RAM' | 'NON-ENCODED'
 'NON-ENCODED' allocates 95MB, size of biggest file must be lower;
 'DECODING HDD2RAM' needs less physical memory (95MB) but is faster!
 'DECODING RAM2RAM' needs more physical memory (147MB) but is slower!
 Optional line #2 contains terminal hits:
 '0' | 'long integer'
 '0' means all hits are needed
 'long integer' means reaching this value termination follows
 Optional line #3 contains Pattern1: 'string'
 if 'string' is specified then input from keyboard arise not
 if 'string' is not specified then input from keyboard arise
 Optional line #4 contains Pattern1_NestedPattern1: 'string'
 Optional line #5 contains Pattern1_NestedPattern2: 'string'
 Optional line #5 contains Pattern1_NestedPattern3: 'string'

Note1: One useful way to make 'ListOfFilesFile=Caterpillar_NON.lst' is next:
 D:\Caterpillar>copy con MAKE1st.bat
 @echo off
 dir Caterpillar_tree*.lbl /s/b>Caterpillar_NON.lst
 dir Caterpillar_tree*.txt /s/b>Caterpillar_NON.lst
 echo.
 F6

Have a nice Caterpillaring.
 For contacts: sanmayce@hotmail.com
 Sanmayce Svalqyatchx 'Kaze', 2009 Jan 29.

D:_KAZE_G.S._Corpus>Raccoondog.exe

LZMA Utility 4.65 : Igor Pavlov : Public domain : 2009-02-03

Usage: lzma <e|d> inputFile outputFile
 e: encode file
 d: decode file

D:_KAZE_G.S._Corpus>Raccoondog -SA4 Raccoondog.lst
 Raccoondog (LZMA Sentence_Dumper), revision 17++, written by Svalqyatchx,
 in fact adapted from Igor Pavlov's excellent LZMA 4.56 SDK.

Usage1: Raccoondog [-SA1|-SA2|-SA3|-SA4] filename
 Decodes all files from a list(filename)
 -SA1 : Brute_Force Search Algorithm
 -SA2 : Quick_Boyer_Moore Search Algorithm
 -SA3 : SMITH_Boyer_Moore Search Algorithm
 -SA4 : Karp_Rabin_Kaze Search Algorithm
 Default is HORSPOOL_Boyer_Moore Search Algorithm

Usage2: Raccoondog <e|d> inputFile outputFile
 e: encode file
 d: decode file

Example1: Raccoondog Raccoondog.lst
 Example2: Raccoondog -SA2 Raccoondog.lst
 Example3: Raccoondog e Caterpillar.001.txt Caterpillar.001.txt.lzma
 Note1: Benchmark:

Raccoondog(EN:8KB/clock, DE:39KB/clock) for 24.9GB(5.53GB LZMA) texts.
 Me machine is:
 Motherboard Name: Toshiba Satellite L305
 CPU Type: Mobile DualCore Intel Pentium, 2166 MHz (13 x 167)
 CPU Alias: Merom-1M
 L1 Code Cache: 32 KB per core
 L1 Data Cache: 32 KB per core
 L2 Cache: 1 MB (On-Die, ECC, ASC, Full-Speed)
 Bus Type: Dual DDR2 SDRAM
 Bus Width: 128-bit
 Real Clock: 333 MHz (DDR)
 Effective Clock: 666 MHz

Note2: Disastrous performance in case 128MB not fully physical!
 Note3: Matches(hits) are overwritten to Raccoondog.hits.Pattern?.html files.
 Note4: Works both on UNIX(LF) and windows(CRLF) text files.
 Note5: Never forget the importance of defragmented_AND_grouped files located at fastest area of disk - first partition is faster than second one, etc.
 Note6: In ANSI, cLock is defined as '#define CLOCKS_PER_SEC 1000'.
 Note7: Since Raccoondog 13++:
 - limits(just skip longer ones) lines to 960 chars; OTHERWISE: HUGE TIME DELAYS due to recursive function;
 - shows hits to console too; MORE VIVID;
 Note8: Since Raccoondog 14:
 - No deletion of input file after compressing/decompressing;
 Note9: During execution hitting a 'Esc' causes termination(i.e. skipping rest).
 NoteA: The two examples below show the need of one additional wildcard in order to match CR for Windows texts; end of line is LF(as in UNIX):
 Pattern(s) example: Pattern1: #####%
 Pattern1_NestedPattern1:
 Possible hit: NEW YORK
 Pattern(s) example: Pattern1: \$\$\$\$\$\$@
 Pattern1_NestedPattern1:
 Possible hit: Printing

Pattern(s) note: You may specify(four times) a main-pattern(case insensitive with wildcards '*' i.e. any character(s) or empty, also '@' i.e. any character or empty, also '#' i.e. any character and not empty, also '\$' i.e. any ALPHA character and not empty, also '%' i.e. any NON-ALPHA character and not empty) with three nested-patterns(case sensitive and unexact), all four connected with AND. Due to different line endings(CRLF in windows; LF in UNIX) you must add a '@' wildcard in place of CR: for example in case of searching for '*.pdf' write '*.pdf@'.

Pattern(s) example: Pattern1: %take@%\$\$\$@
 Pattern1_NestedPattern1:
 Possible hit: ... is taken by
 Possible hit: ... would take it
 Note: First % is to avoid e.g. 'mis' prefix
 Second % is to avoid e.g. 'ing' suffix

Master-pattern note: It is case insensitive with wildcards '*', '@', '#', '\$', '%' allowed. The purpose of this pattern is to decide whether a search for next patterns will be executed, it is applied on all lines i.e. the whole file. There must be at least one hit in order to execute search for next patterns.

Have a nice Raccoondoging.
 For contacts: sanmayce@hotmail.com
 Sanmayce Svalqyatchx 'Kaze', 2010 Jun 06.

Allocated memory for DEcoded file in MB: 256
 Size of input file with files for Raccoondoging: 9669

Input Master-pattern(hit only 'Enter' to skip):
 Input Pattern1(hit only 'Enter' to skip): *not anymore*
 - Input Pattern1_NestedPattern1(hit only 'Enter' to skip):
 Input Pattern2(hit only 'Enter' to skip):
 Processing .\Caterpillar.001.RAFT2.txt.lzma ...
 Doing DECODE from HDD to RAM ...
 Overall decode performance so far: 000,007KB/clock(EN) or 000,031KB/clock(DE)
 Doing SEARCH for Pattern1 at once and flushing hit-sentences ...
 000,000,001 It used to be just "the living room," but not anymore.
 000,000,002 But not anymore.
 Found 2 case-insensitive and unexact matches(hits), so far.
 'Esc' was pressed, so skip the rest files and quit!

Total Rough Upload and Decode time: 2,297 clocks
 Total Rough Search time: 1,719 clocks
 Total time: 4 seconds
 Total Lines encountered: 1,150,388
 Total Search(non-mask) function invocations: 0
 Total Search(MASK i.e. wildcard) function invocations: 1,149,075
 Total MASK i.e. wildcard invocations: 1,149,075
 Total MASK i.e. wildcard hits: 2
 Total MASK i.e. wildcard time: 1,434 clocks
 Total MASK i.e. wildcard performance: 46KB/clock
 Total BoyerMooreHorspool invocations: 0
 Total BoyerMooreHorspool(whole chunks, not lines) hits: 0
 Total BoyerMooreHorspool(whole chunks, not lines) time: 0 clocks
 Total KarpRabinKaze invocations: 0
 Total KarpRabinKaze(whole chunks, not lines) hits: 0
 Total KarpRabinKaze(whole chunks, not lines) time: 0 clocks
 Raccoondog: Done successfully.

D:_KAZE_G.S._Corpus>Salah-ed-din -SA4 Salah-ed-din.lst
 Salah-ed-din(Sentence_Dumper), revision 14++, written by Svalqyatchx,
 in fact adapted from Mark Adler's and Jean-loup Gailly's ZLIB package.

Usage1: Salah-ed-din [-SA1|-SA2|-SA3|-SA4] filename
 Decodes all files from a list(filename)
 -SA1 : Brute_Force Search Algorithm
 -SA2 : Quick_Boyer_Moore Search Algorithm
 -SA3 : SMITH_Boyer_Moore Search Algorithm
 -SA4 : Karp_Rabin_Kaze Search Algorithm
 Default is HORSPOOL_Boyer_Moore Search Algorithm
 Usage2: Salah-ed-din [-d] [-f] [-h] [-r] [-1 to -9] [files...]
 -d : decompress
 -f : compress with Z_FILTERED
 -h : compress with Z_HUFFMAN_ONLY
 -r : compress with Z_RLE
 -1 to -9 : compression level

Example1: Salah-ed-din Salah-ed-din.lst
Example2: Salah-ed-din -SA2 Salah-ed-din.lst
Example3: Salah-ed-din -f -6 Caterpillar.001.txt
Example4: Salah-ed-din -d Caterpillar.001.txt.gz

Note1: Benchmark:
Raccoondog(EN:39KB/clock, DE:117KB/clock) for 24.9GB(8.42GB GZ) texts.
Me machine is:
Motherboard Name: Toshiba Satellite L305
CPU Type: Mobile DualCore Intel Pentium, 2166 MHz (13 x 167)
CPU Alias: Merom-1M
L1 Code Cache: 32 KB per core
L1 Data Cache: 32 KB per core
L2 Cache: 1 MB (On-Die, ECC, ASC, Full-Speed)
Bus Type: Dual DDR2 SDRAM
Bus width: 128-bit
Real Clock: 333 MHz (DDR)
Effective Clock: 666 MHz

Note2: Disastrous performance in case 128MB not fully physical!
Note3: Matches(hits) are overwritten to Salah-ed-din.hits.Pattern?.html files.
Note4: Works both on UNIX(LF) and Windows(CRLF) text files.
Note5: Never forget the importance of defragmented_AND_grouped files located at fastest area of disk - first partition is faster than second one, etc.
Note6: In ANSI, clock is defined as '#define CLOCKS_PER_SEC 1000'.
Note7: Since Salah-ed-din 13++:
- limits(just skip longer ones) lines to 960 chars; OTHERWISE: HUGE TIME DELAYS due to recursive function;
- shows hits to console too; MORE VIVID;
Note8: Since Salah-ed-din 14:
- No deletion of input file after compressing/decompressing;
Note9: During execution hitting a 'Esc' causes termination(i.e. skipping rest).

Pattern(s) note: You may specify(four times) a main-pattern(case insensitive with wildcards '*' i.e. any character(s) or empty and '?' i.e. any character or empty) with three nested-patterns(case sensitive and unexact), all four connected with AND. Due to different line endings(CRLF in Windows; LF in UNIX) you must add a '?' wildcard in place of CR: for example in case of searching for '*.pdf' write '*.pdf?'.

Pattern(s) example: Pattern1: *take? *it*
Pattern1_NestedPattern1: you
Possible hit: ... your reason is so taken by It.

Have a nice Salah-ed-dining.
For contacts: sanmayce@hotmail.com
Sanmayce Svalqyatchx 'kaze', 2009 Feb 22.

Allocated memory for DEcoded file in MB: 96
Size of input file with files for Salah-ed-dining: 8680
Pattern(s) note: You may specify(four times) a main-pattern(case insensitive with wildcards '*' i.e. any character(s) or empty and '?' i.e. any character or empty) with three nested-patterns(case sensitive and unexact), all four connected with AND. Due to different line endings(CRLF in Windows; LF in UNIX) you must add a '?' wildcard in place of CR: for example in case of searching for '*.pdf' write '*.pdf?'.

Pattern(s) example: Pattern1: *take? *it*
Pattern1_NestedPattern1: you
Possible hit: ... your reason is so taken by It.

Input Pattern1(hit only 'Enter' to skip): *not anymore*
- Input Pattern1_NestedPattern1(hit only 'Enter' to skip):
Input Pattern2(hit only 'Enter' to skip):
Processing .\Caterpillar.001.RAFT3.txt.gz ...
Doing DECODE from HDD to RAM ...

Salah-ed-din decoded buffer size: 99,614,459
Overall decode performance so far: 000,033KB/clock(EN) or 000,102KB/clock(DE)
Doing SEARCH for Pattern1 at once and flushing hit-sentences ...
000,000,001 M: Not anymore.
000,000,002 M: Not anymore.
000,000,003 "Not anymore," Lidia replied.
000,000,004 Not anymore.
000,000,005 Not anymore.
000,000,006 "Not anymore," Lidia replied.
000,000,007 Not anymore.
000,000,008 Not anymore.
000,000,009 "Not anymore," Lidia replied.
000,000,010 Not anymore.
000,000,011 Not anymore.
Found 11 case-insensitive and unexact matches(hits), so far.
'Esc' was pressed, so skip the rest files and quit!

Total Rough Upload and Decode time: 953 clocks
Total Rough Search time: 1,907 clocks
Total time: 3 seconds
Total Lines encountered: 1,835,098
Total Search(non-mask) function invocations: 0
Total Search(MASK i.e. wildcard) function invocations: 1,834,650
Total Boyer-Moore-Horspool(whole chunks, not lines) hits: 0
Total Boyer-Moore-Horspool(whole chunks, not lines) time: 0 clocks
Total Karp_Rabin_Kaze(whole chunks, not lines) hits: 0
Total Karp_Rabin_Kaze(whole chunks, not lines) time: 0 clocks
Salah-ed-din: Done successfully.

D:_KAZE_G.S._Corpus>Kazuya.exe/?

Kazuya(LZ Sentence_Dumper), revision 17++, written by Svalqyatchx, in fact adapted from Lasse Reinhold's excellent QuickLZ 1.4.0 library, in fact adapted from Ariya Hidayat's sub-excellent FastLZ 0.1.0 library, in fact adapted from Markus F.X.J. Oberhumer's sub-excellent LZ0 2.03 library, in fact adapted from Haruhiko Okumura's sub-excellent LZSS 4/6/1989 library.

Usage1: Kazuya [-sa1|-sa2|-sa3|-sa4|-SA1|-SA2|-SA3|-SA4
|-SA1|-SA2|-SA3|-SA4|-Sa1|-Sa2|-Sa3|-Sa4|-krknd] filename
Decodes all files from a list(filename)
-sa1 : QuickLZ Decode + Brute_Force Search Algorithm
-sa2 : QuickLZ Decode + Quick_Boyer_Moore Search Algorithm
-sa3 : QuickLZ Decode + SMITH_Boyer_Moore Search Algorithm
-sa4 : QuickLZ Decode + Karp_Rabin_Kaze Search Algorithm
-SA1 : LZ0 Decode + Brute_Force Search Algorithm
-SA2 : LZ0 Decode + Quick_Boyer_Moore Search Algorithm
-SA3 : LZ0 Decode + SMITH_Boyer_Moore Search Algorithm
-SA4 : LZ0 Decode + Karp_Rabin_Kaze Search Algorithm


```

-SA1 : FastLZ Decode + Brute_Force Search Algorithm
-SA2 : FastLZ Decode + Quick_Boyer_Moore Search Algorithm
-SA3 : FastLZ Decode + SMITH_Boyer_Moore Search Algorithm
-SA4 : FastLZ Decode + Karp_Rabin_Kaze Search Algorithm
-Sa1 : OkumuraLZ Decode + Brute_Force Search Algorithm
-Sa2 : OkumuraLZ Decode + Quick_Boyer_Moore Search Algorithm
-Sa3 : OkumuraLZ Decode + SMITH_Boyer_Moore Search Algorithm
-Sa4 : OkumuraLZ Decode + Karp_Rabin_Kaze Search Algorithm
-krknd : Karp_Rabin_Kaze Search Algorithm, but with no additional
        chunk-searches overhead and without decompression of
        incoming files i.e. pure text is uploaded
Default is QuickLZ Decode + HORSPOOL_Boyer_Moore Search Algorithm
Usage2: Kazuya <e|d|E|D|A|R|a|r> inputFile outputFile
e: encode QuickLZ file
d: decode QuickLZ file
E: encode LZ0 file
D: decode LZ0 file
A: encode(archive) FastLZ file
R: decode(restore) FastLZ file
a: encode(archive) OkumuraLZ file
r: decode(restore) OkumuraLZ file
Example1: Kazuya Kazuya.lst
Example2: Kazuya -SA2 Kazuya.lst
Example3: Kazuya e Caterpillar.001.txt Caterpillar.001.txt.Lasse
Note1: Benchmark(HDD read speed is the nasty bottleneck):
      Kazuya(EN:37KB/clock, DE:88KB/clock) for 24.9GB(10.6GB Lasse) texts.
Me machine is:
Motherboard Name:          Toshiba Satellite L305
CPU Type:                  Mobile DualCore Intel Pentium, 2166 MHz (13 x 167)
CPU Alias:                  Merom-1M
L1 Code Cache:              32 KB per core
L1 Data Cache:              32 KB per core
L2 Cache:                   1 MB (On-Die, ECC, ASC, Full-Speed)
Bus Type:                   Dual DDR2 SDRAM
Bus Width:                  128-bit
Real Clock:                 333 MHz (DDR)
Effective Clock:            666 MHz
Note2: Disastrous performance in case 256MB not fully physical!
Note3: Matches(hits) are overwritten to Kazuya.hits.Pattern?.html files.
Note4: Works both on UNIX(LF) and windows(CRLF) text files.
Note5: Never forget the importance of defragmented_AND_grouped files located at
      fastest area of disk - first partition is faster than second one, etc.
Note6: In ANSI, clock is defined as '#define CLOCKS_PER_SEC 1000'.
Note7: Since Kazuya 13++:
      - limits(just skip longer ones) lines to 960 chars; OTHERWISE: HUGE TIME
        DELAYS due to recursive function;
      - shows hits to console too; MORE VIVID;
Note8: Since Kazuya 14:
      - No deletion of input file after compressing/decompressing;
Note9: During execution hitting a 'Esc' causes termination(i.e. skipping rest).
NoteA: This revision works with up to 127MB incoming(non-compressed) files,
      because it allocates 256MB of which one half is for incoming
      oher for outcoming file, i.e. decompression/compression is RAM to RAM.
NoteB: Charge(delivery) performance combines upload and decode performance.
NoteC: The two examples below show the need of one additional wildcard in
      order to match CR for windows texts; end of line is LF(as in UNIX):
Pattern(s) example: Pattern1: #####%
                    Pattern1_NestedPattern1:
                    Possible hit: NEW YORK
Pattern(s) example: Pattern1: $$$$$$$@
                    Pattern1_NestedPattern1:
                    Possible hit: Printing

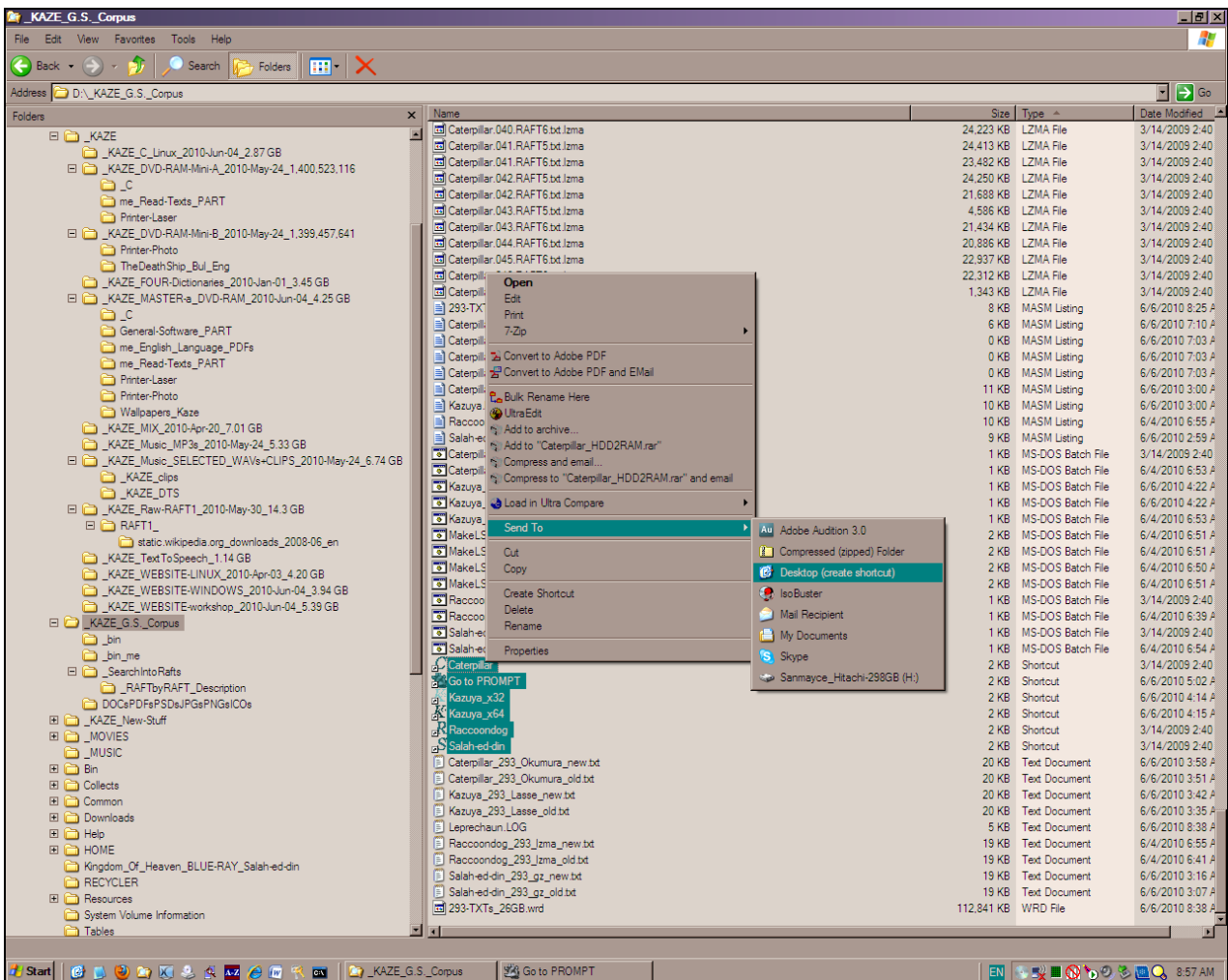
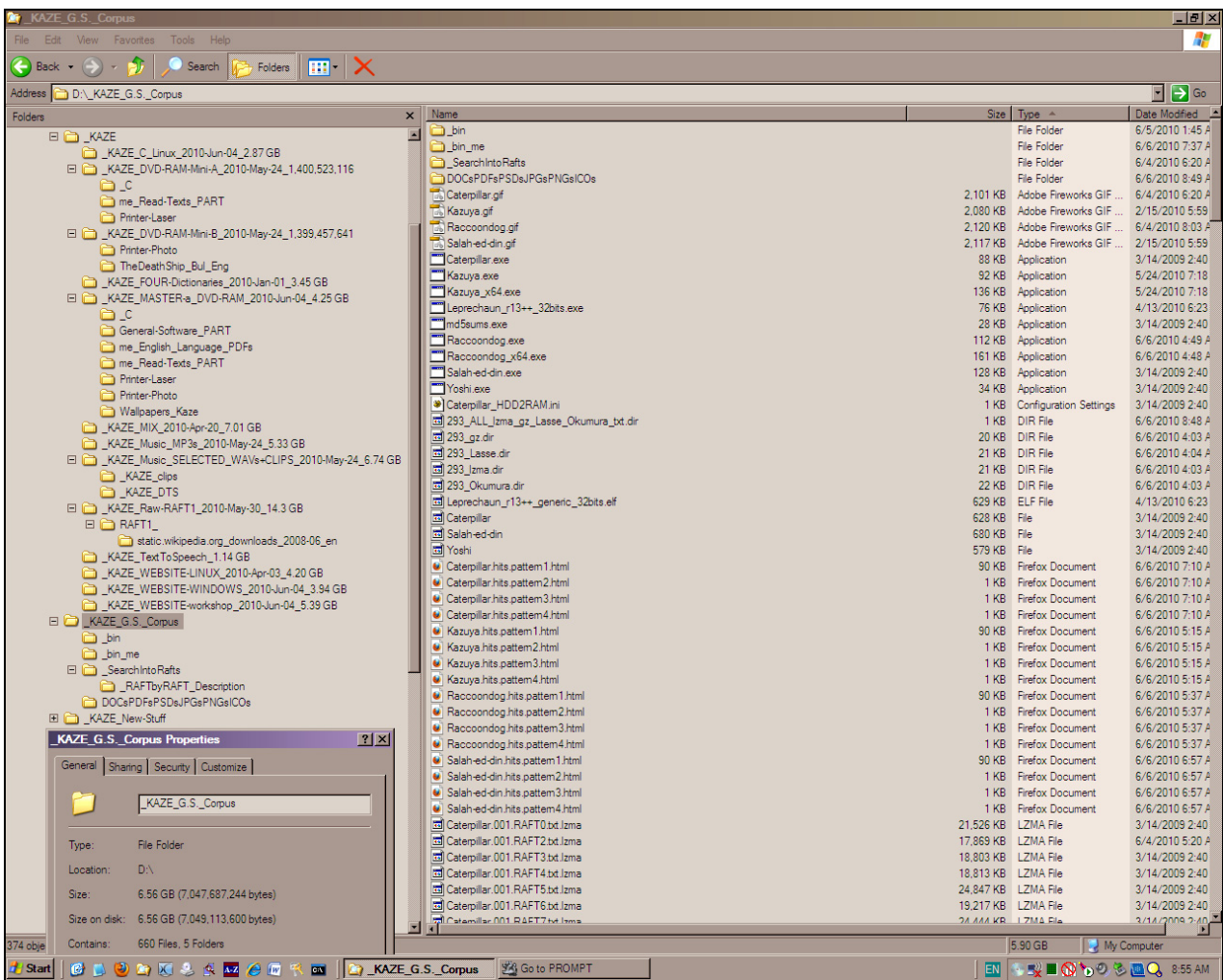
Pattern(s) note: You may specify(four times) a main-pattern(case insensitive
with wildcards '*' i.e. any character(s) or empty, also '@'
i.e. any character or empty, also '#' i.e. any character
and not empty, also '$' i.e. any ALPHA character
and not empty, also '%' i.e. any NON-ALPHA character
and not empty) with three nested-patterns(case
sensitive and unexact), all four connected with AND.
Due to different line endings(CRLF in windows; LF in UNIX)
you must add a '@' wildcard in place of CR: for example in
case of searching for '*.pdf' write '*.pdf@'.
Pattern(s) example: Pattern1: %*take@%$$$@
                    Pattern1_NestedPattern1:
                    Possible hit: ... is taken by
                    Possible hit: ... would take it
                    Note: First % is to avoid e.g. 'mis' prefix
                          Second % is to avoid e.g. 'ing' suffix
Master-pattern note: It is case insensitive with wildcards '*', '@', '#', '$', '%'
allowed. The purpose of this pattern is to
decide whether a search for next patterns will be
executed, it is applied on all lines i.e. the whole file.
There must be at least one hit in order to execute search
for next patterns.

```

Have a nice Kazuyaing.
For contacts: sanmayce@sanmayce.com
Sanmayce Svalqyatchx 'Kaze', 2010 May 24.

Allocated memory for DEcoded file in MB: 256
Kazuya: Can't open /? file.

D:_KAZE_G.S._Corpus>



```

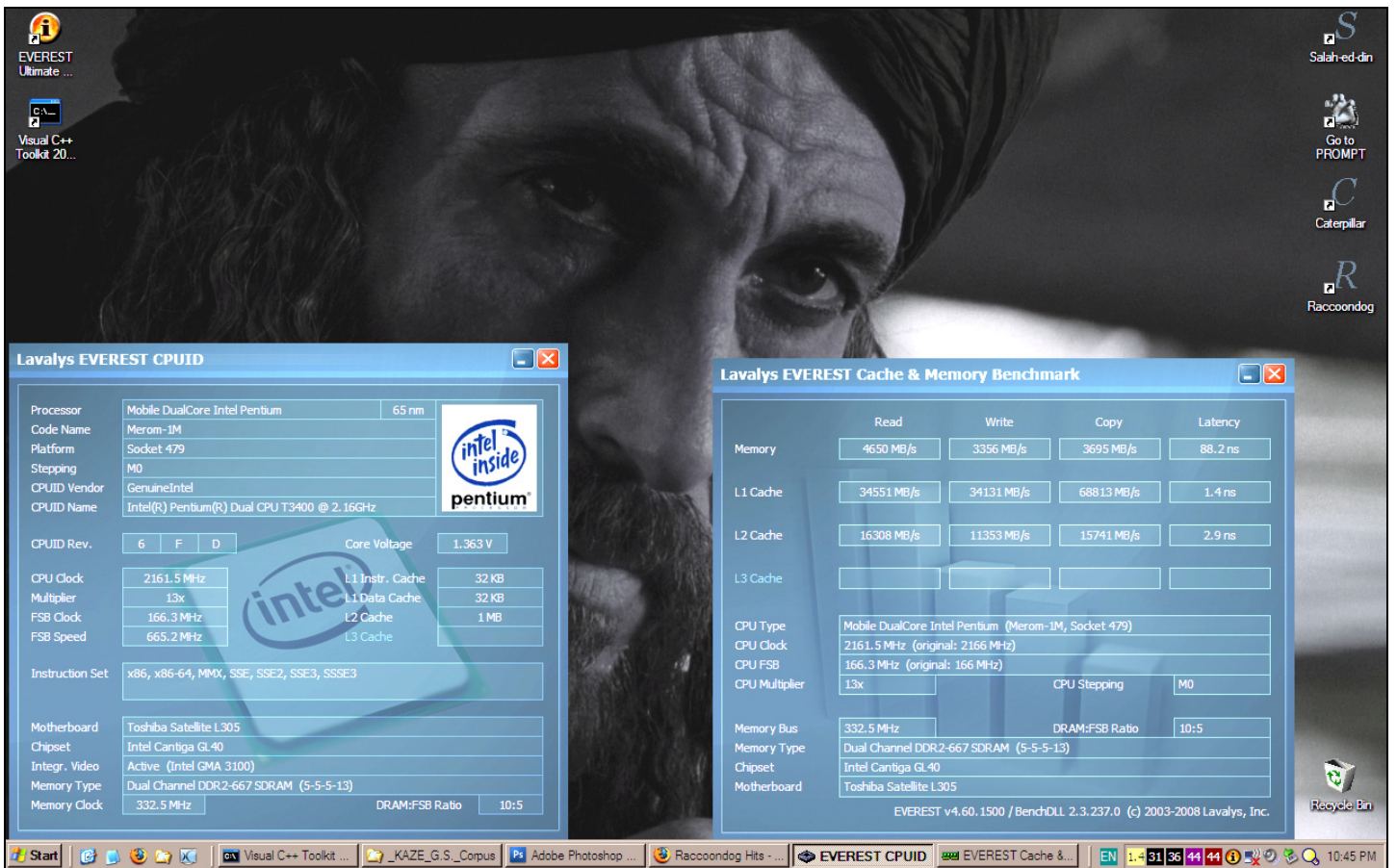
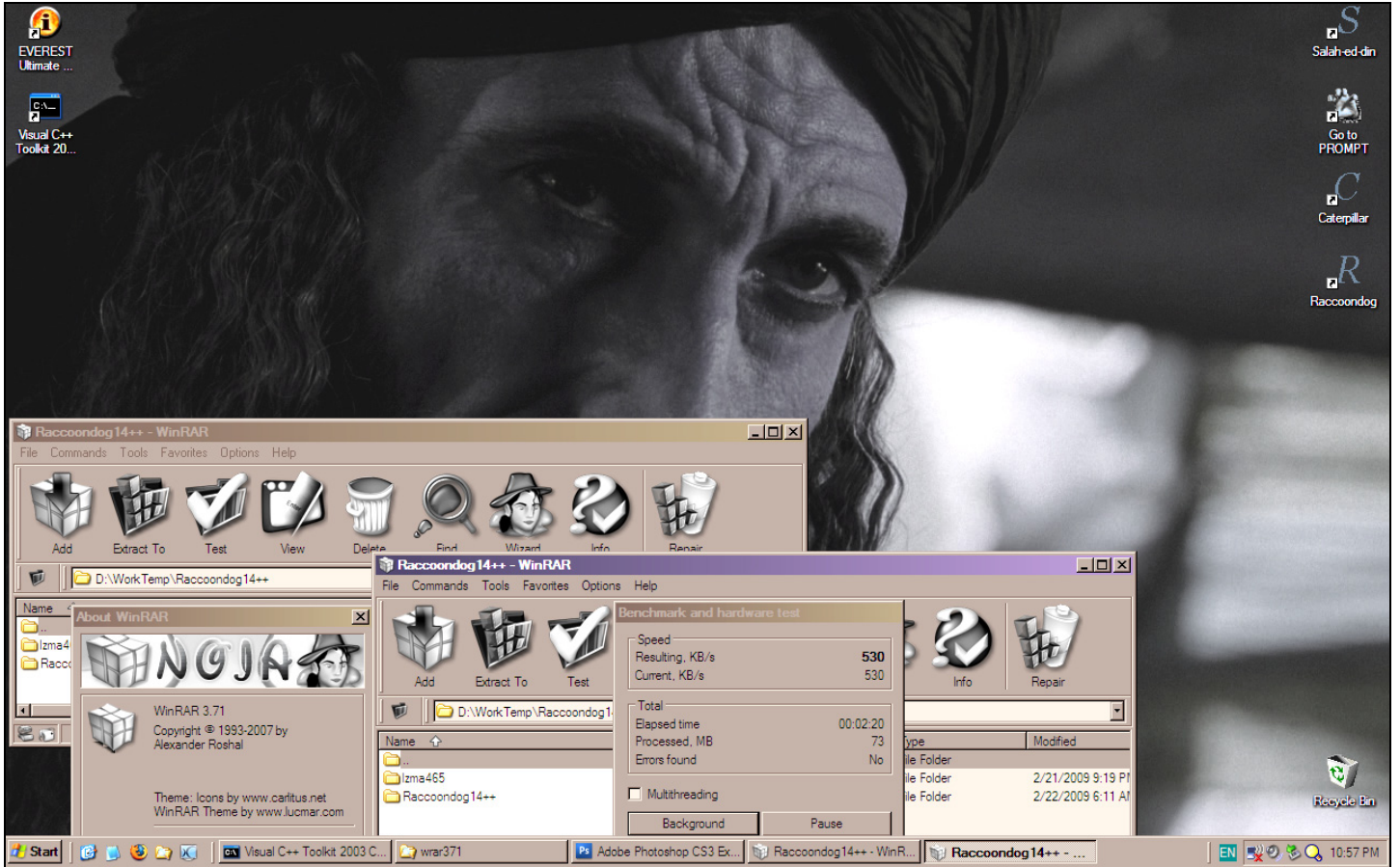
file:///D:/_KAZE_G.S_Corpus/_SearchIntoRafts/KAZE_G.S_Corpus_pdf_Caterpillar.html
Book of Lao Zi. Dao De ... Nifty console tools and L... Kazuya Hits Raccoondog Hits Salah-ed-din Hits Caterpillar Hits Caterpillar Hits
01/25/2009 05:02 AM 1,154,717 Evil Under the Sun By Agatha Christie.pdf
01/25/2009 05:02 AM 479,173 Five Little Pigs By Agatha Christie.pdf
01/25/2009 05:02 AM 473,007 Halloween Party By Agatha Christie.pdf
01/25/2009 05:02 AM 1,286,369 Hercule Poirot's Christmas By Agatha Christie.pdf
01/25/2009 05:02 AM 292,411 Hickory Dickory Death By Agatha Christie.pdf
01/25/2009 05:02 AM 668,622 Labours Of Hercules By Agatha Christie.pdf
01/25/2009 05:02 AM 1,236,545 Lord Edgware Dies By Agatha Christie.pdf
01/25/2009 05:02 AM 540,449 Mrs McGintys Dead By Agatha Christie.pdf
01/25/2009 05:02 AM 421,856 Murder At The Vicarage By Agatha Christie.pdf
01/25/2009 05:02 AM 1,254,714 Murder in Mesopotamia By Agatha Christie.pdf
01/25/2009 05:02 AM 386,129 Murder Is Easy By Agatha Christie.pdf
01/25/2009 05:02 AM 345,254 Murder Of Roger Ackroyd By Agatha Christie.pdf
01/25/2009 05:02 AM 334,816 Murder on the Links By Agatha Christie.pdf
01/25/2009 05:02 AM 770,946 Murder On The Orient Express By Agatha Christie.pdf
01/25/2009 05:02 AM 540,272 Mystery Of The Blue Train By Agatha Christie.pdf
01/25/2009 05:02 AM 434,242 N Or M By Agatha Christie.pdf
01/25/2009 05:02 AM 561,502 Nemesis By Agatha Christie.pdf
01/25/2009 05:02 AM 399,098 One Two Buckle My Shoe By Agatha Christie.pdf
01/25/2009 05:02 AM 794,314 Parker Pyne Investigates By Agatha Christie.pdf
01/25/2009 05:02 AM 436,599 Partners In Crime By Agatha Christie.pdf
01/25/2009 05:02 AM 443,521 Passenger To Frankfurt By Agatha Christie.pdf
01/25/2009 05:02 AM 596,994 Peril At End House By Agatha Christie.pdf
01/25/2009 05:02 AM 425,371 Poirot's Early Cases By Agatha Christie.pdf
01/25/2009 05:02 AM 791,030 Sad Cypress By Agatha Christie.pdf
01/25/2009 05:02 AM 598,974 Sittaford Mystery By Agatha Christie.pdf
01/25/2009 05:02 AM 304,772 Sleeping Murder By Agatha Christie.pdf
01/25/2009 05:02 AM 510,546 Sparkling Cyanide By Agatha Christie.pdf
01/25/2009 05:02 AM 488,436 Surprise Surprise.pdf
01/25/2009 05:02 AM 539,352 Taken At The Flood By Agatha Christie.pdf
01/25/2009 05:02 AM 466,081 The ABC Murders By Agatha Christie.pdf
01/25/2009 05:02 AM 308,229 The Body In The Library By Agatha Christie.pdf
01/25/2009 05:02 AM 362,240 The Burden By Agatha Christie.pdf
01/25/2009 05:02 AM 647,357 The Casebook Of Hercule Poirot.pdf
01/25/2009 05:02 AM 343,894 The Circular Staircase.pdf
01/25/2009 05:02 AM 497,478 The Clocks By Agatha Christie.pdf
01/25/2009 05:02 AM 579,351 The Hallow By Agatha Christie.pdf
01/25/2009 05:02 AM 330,173 The Man In Lower Ten.pdf
01/25/2009 05:02 AM 595,817 The Man In The Brown Suit By Agatha Christie.pdf
01/25/2009 05:02 AM 344,868 The mirror cracked from side to side By Agatha Christie.pdf
01/25/2009 05:02 AM 269,020 The Moving Finger By Agatha Christie.pdf
01/25/2009 05:02 AM 251,750 The Mysterious Affair At Styles By Agatha Christie.pdf
01/25/2009 05:02 AM 296,070 The Mysterious Mr Quin By Agatha Christie.pdf
01/25/2009 05:02 AM 440,713 The Regatta Mystery By Agatha Christie.pdf
01/25/2009 05:02 AM 327,430 The Secret Adversary By Agatha Christie.pdf
01/25/2009 05:02 AM 541,063 The Secret Of Chimneys By Agatha Christie.pdf
01/25/2009 05:02 AM 493,592 The Seven Dials Mystery By Agatha Christie.pdf
01/25/2009 05:02 AM 1,225,944 They Came to Baghdad By Agatha Christie.pdf
01/25/2009 05:02 AM 289,044 They Do It With Mirrors By Agatha Christie.pdf
01/25/2009 05:02 AM 304,417 Third Girl By Agatha Christie.pdf
01/25/2009 05:02 AM 472,152 Three Act Tragedy By Agatha Christie.pdf
01/25/2009 05:02 AM 1,174,993 Three Blind Mice and Other Stories - Agatha Christie - 1948.pdf
01/25/2009 05:02 AM 582,226 Towards Zero By Agatha Christie.pdf
01/25/2009 05:02 AM 357,435 Why Didn't They Ask Evans By Agatha Christie.pdf
002_512 [* .pdf?] [[] []] Caterpillar link to target file: D:\_KAZE_G.S_Corpus\_SearchIntoRafts\_RAFTbyRAFT_Description\RAFT0_og_on_s_dir
06/04/2010 04:49 AM 14,462,269 100 Questions and Answers About Alcoholism.pdf
06/04/2010 04:49 AM 12,167,723 210 Knots.pdf
06/04/2010 04:49 AM 52,613,250 Berkshire Encyclopedia Of World History Vol I - Abraham to Coal.pdf
06/04/2010 04:49 AM 43,800,823 Berkshire Encyclopedia Of World History Vol II - Cold War to Global Imperialism and Gender.pdf
06/04/2010 04:49 AM 51,772,828 Berkshire Encyclopedia Of World History Vol III - Global Migrations in Modern Time to Mysticism.pdf
06/04/2010 04:49 AM 52,200,347 Berkshire Encyclopedia Of World History Vol IV - Napoleon to Sun Yat sen.pdf

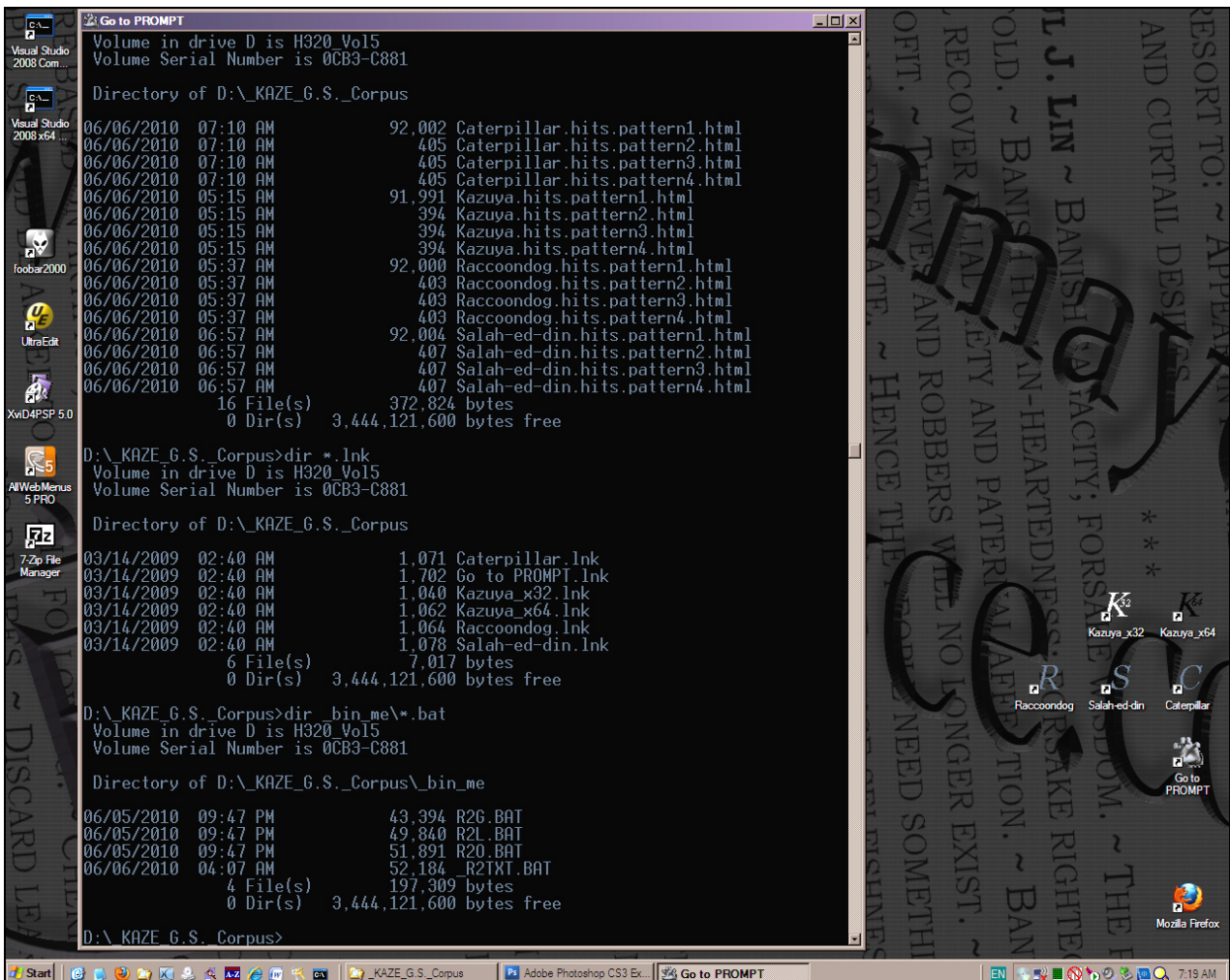
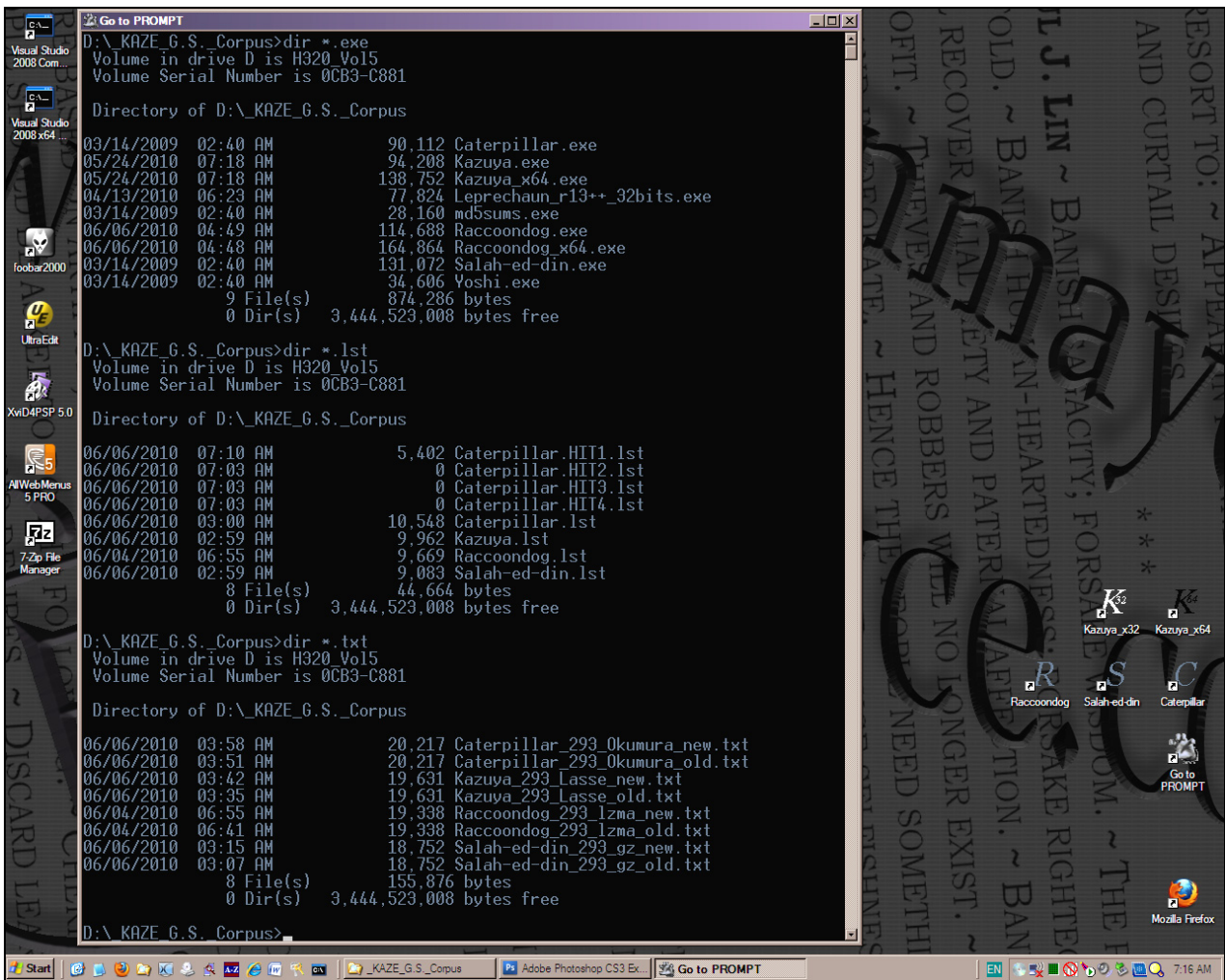
```

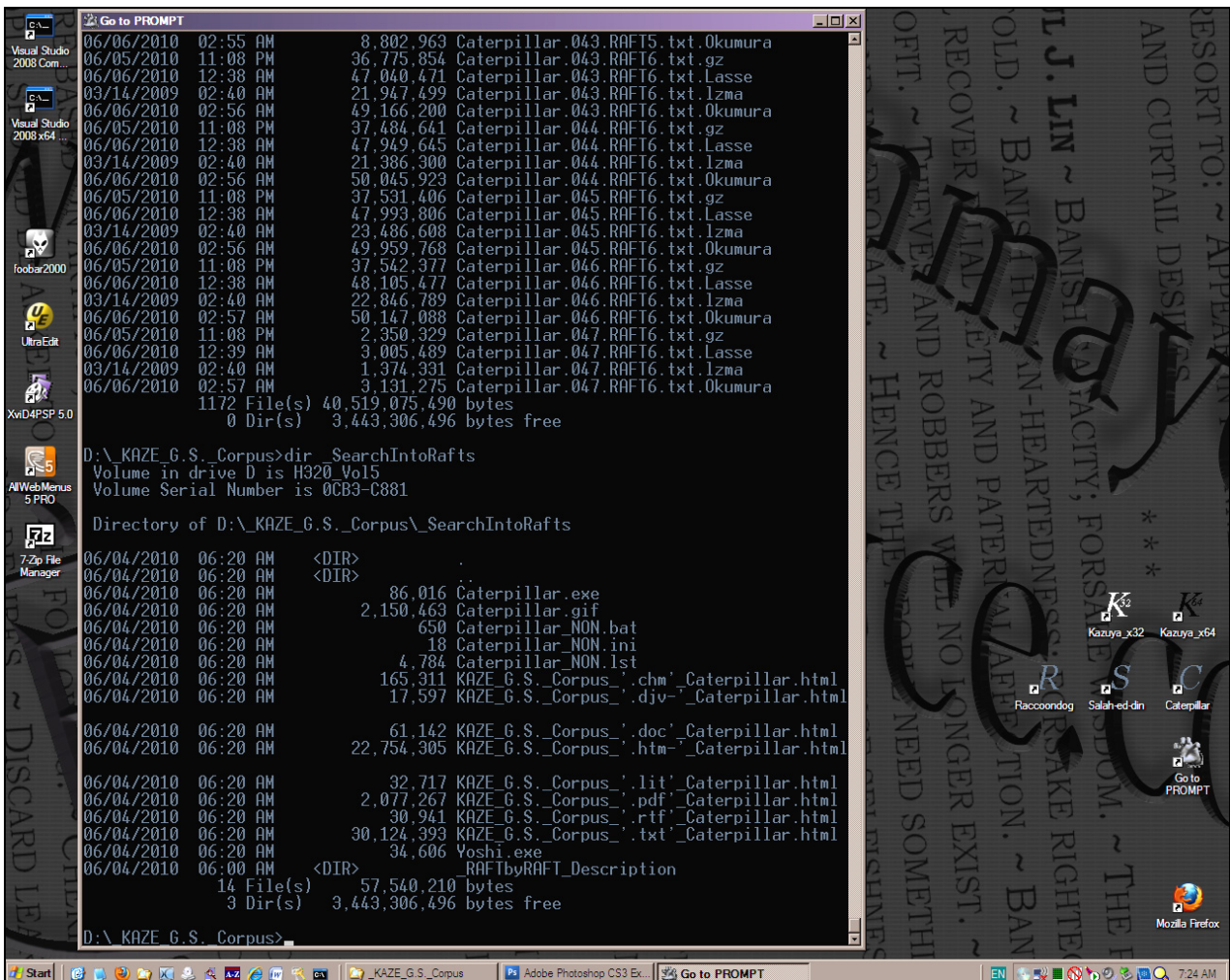
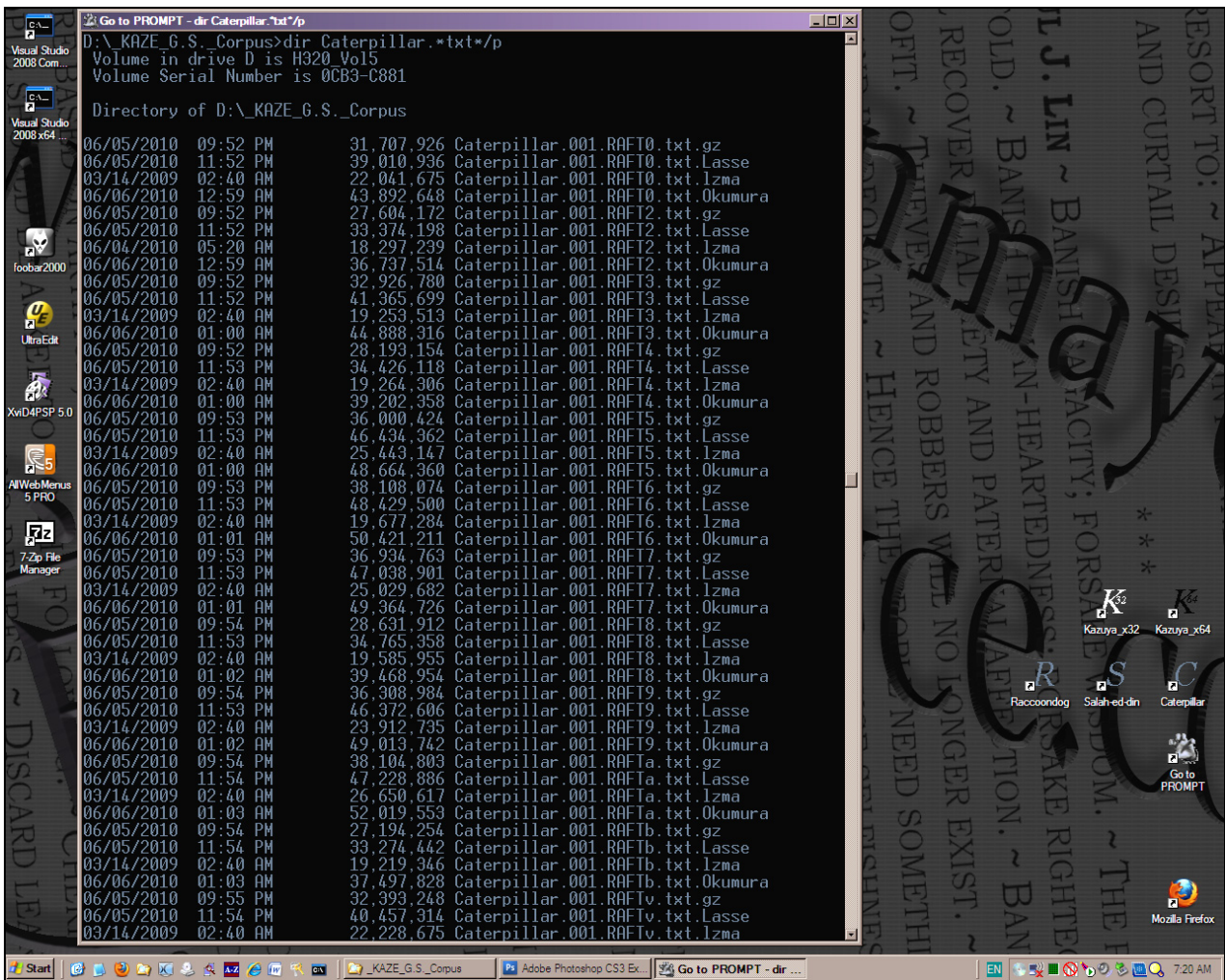
```

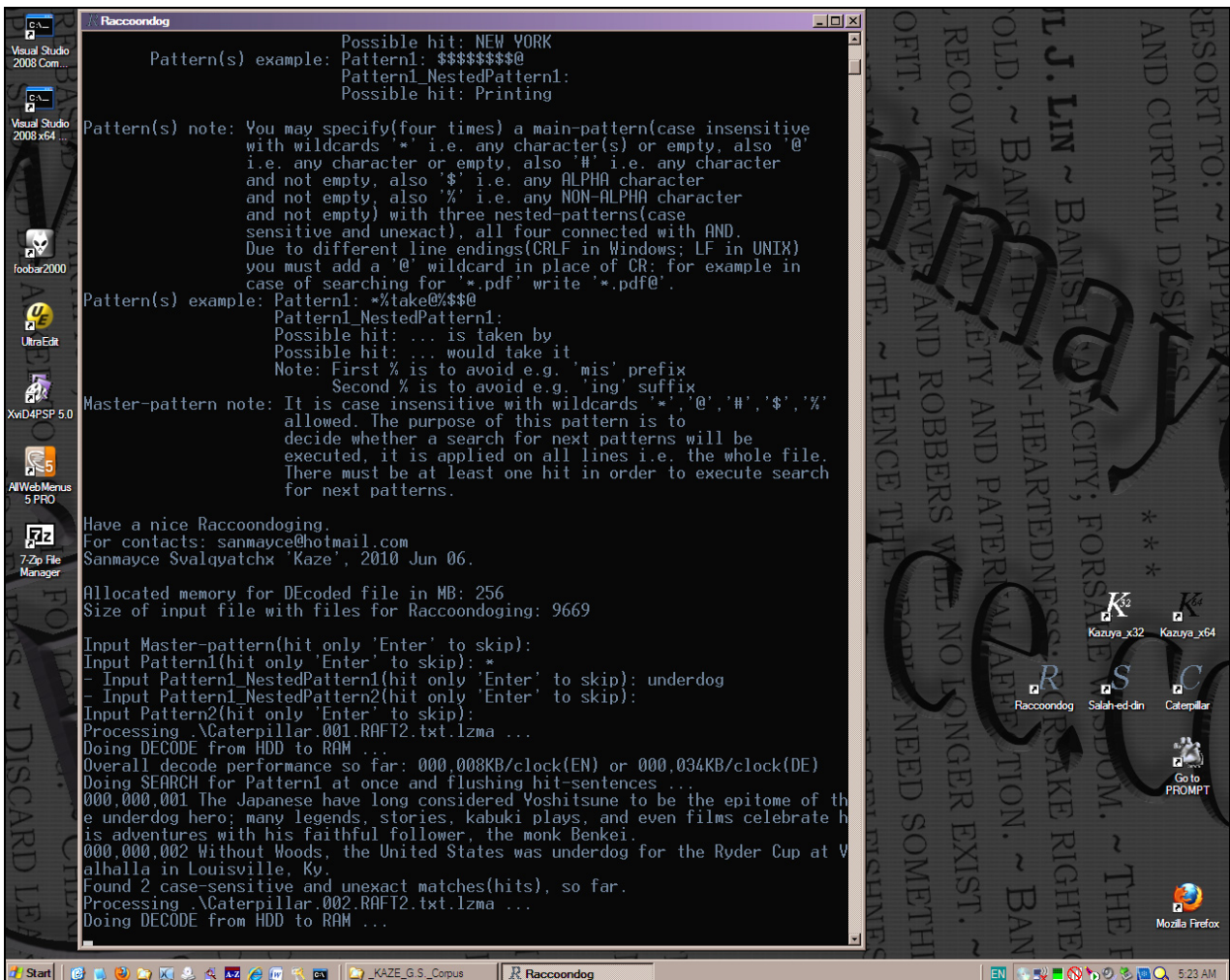
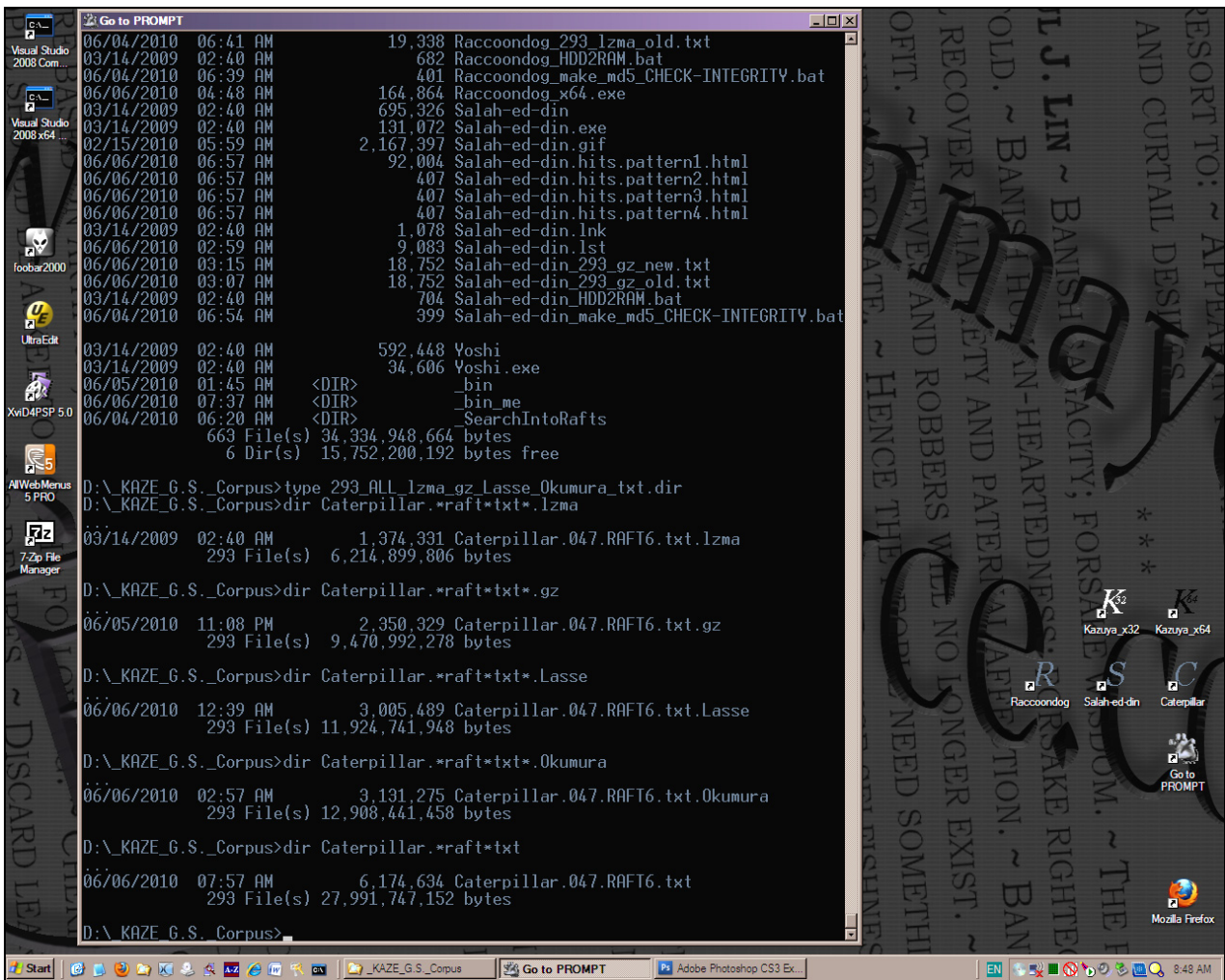
file:///D:/_KAZE_G.S_Corpus/_SearchIntoRafts/KAZE_G.S_Corpus_chm_Caterpillar.html
Book of Lao Zi. Da... Nifty console tools a... Kazuya Hits Raccoondog Hits Salah-ed-din Hits Caterpillar Hits Caterpillar Hits Caterpillar Hits
01/25/2009 05:02 AM 7,815,498 A Student's Guide to Textual Criticism of the Bible.chm
01/25/2009 05:02 AM 1,817,974 Death and the Afterlife - A Cultural Encyclopedia.chm
01/25/2009 05:02 AM 1,315,722 Textual Commentary on the Greek New Testament.chm
000_003 [* .chm?] [[] []] Caterpillar link to target file: D:\_KAZE_G.S_Corpus\_SearchIntoRafts\_RAFTbyRAFT_Description\RAFT0_og_
06/04/2010 04:49 AM 1,323,773 FatwaBase - v4.99.chm
06/04/2010 04:49 AM 2,317,291 Hisnul Muslim.chm
06/04/2010 04:49 AM 5,963,547 ibnkathir_english (TAFSIR).chm
06/04/2010 04:49 AM 1,512,921 Islam Against Terrorism - v1.33.chm
06/04/2010 04:49 AM 3,644,892 prophetsprayer.chm
06/04/2010 04:49 AM 31,516,757 guransahih.chm
06/04/2010 04:49 AM 7,941,208 Sahih Bukhari, Muslim, Muwatta _ Abu Dawood.chm
06/04/2010 04:49 AM 8,418,464 Sahih Bukhari.chm
06/04/2010 04:49 AM 2,501,422 sahih_bukhari_01.chm
06/04/2010 04:49 AM 2,244,116 sahih_muslim_01.chm
06/04/2010 04:49 AM 32,751,405 The Noble Quran.chm
06/04/2010 04:49 AM 3,644,892 The Prophet's Prayer - v3.20.chm
06/04/2010 04:49 AM 87,560,405 Complete Works of Osho.chm
000_013 [* .chm?] [[] []] Caterpillar link to target file: D:\_KAZE_G.S_Corpus\_SearchIntoRafts\_RAFTbyRAFT_Description\RAFT2_og_
12/02/2006 10:18 PM 26,168,854 3d Max User Reference.chm
12/02/2006 10:18 PM 37,108,777 60 Common Web Design Mistakes and How to Avoid Them.chm
12/02/2006 10:18 PM 1,526,125 70.214.examcram2.chm
12/02/2006 10:18 PM 2,092,162 A First Look at ADD.NET and System Xml 2.0.chm
12/02/2006 10:18 PM 6,323,265 A Guide to Constructing GUIs.chm
12/02/2006 10:18 PM 10,160,148 A Roadmap for Building a Linux File and Print Server.chm
12/02/2006 10:18 PM 7,129,066 A+ Certification Training Kit - Second Edition.chm
12/02/2006 10:18 PM 3,904,279 Absolute Beginner's Guide to Launching an eBay Business.chm
12/02/2006 10:18 PM 1,071,122 Absolute OpenBSD - UNIX for the Practical Paranoid.chm
12/02/2006 10:18 PM 1,577,580 ACADP9.CHM
12/02/2006 10:18 PM 3,516,968 ACMMAIN9.CHM
12/02/2006 10:18 PM 2,054,031 ACTIONSCRIPT FOR FLASH MX - THE DEFINITIVE GUIDE, 2ND EDITION.CHM
12/02/2006 10:18 PM 2,092,162 Addison.Wesley.A.First.Look.At.ADO.Dot.NET.And.System.Xml.v.2.0.eBook-LiB.chm
12/02/2006 10:18 PM 4,385,293 Addison.Wesley.Algorithms.In.Java.3rd.Ed.Part5.Graph.Algorithms.eBook-LiB.chm
12/02/2006 10:18 PM 2,057,225 Addison.Wesley.Parallel.And.Distributed.Programming.Using.Cpp.eBook-LiB.chm
12/02/2006 10:18 PM 1,565,802 Addison.Wesley.The.Art.Of.Unix.Programming.eBook-LiB.chm
12/02/2006 10:18 PM 7,493,194 Addison-Wesley - A Programmers Guide to .NET.chm
12/02/2006 10:18 PM 14,111,320 Administering and Securing the Apache Server.chm
12/02/2006 10:18 PM 1,384,762 ADO210.CHM
12/02/2006 10:18 PM 4,301,415 Advanced MS Visual Basic 6.0 Second Edition.chm
12/02/2006 10:18 PM 8,283,383 Advantage Database Server - The Official Guide.chm
12/02/2006 10:18 PM 11,260,817 Algorithms for Compiler Design.chm
12/02/2006 10:18 PM 4,385,293 Algorithms in Java, Third Edition, Part 5 - Graph Algorithms.chm
12/02/2006 10:18 PM 12,085,560 A-LIST.Publishing.Modern.Cryptography.eBook-LiB.chm
12/02/2006 10:18 PM 30,245,817 Anti-Hacker Tool Kit.chm
12/02/2006 10:18 PM 4,471,967 Applied Software Engineering Using Apache Jakarta Commons.chm
12/02/2006 10:18 PM 1,932,072 Applying.Enterprise.JavaBeans.2nd.Edition.chm
12/02/2006 10:18 PM 15,773,791 Architecting Portal Solutions.chm
12/02/2006 10:18 PM 29,025,837 ASP.NET Programming with Visual C# .NET 2003 Step by Step.chm
12/02/2006 10:18 PM 8,977,185 Assembly Language Step-by-Step - Programming with DOS and Linux.chm
12/02/2006 10:18 PM 1,015,798 ax_enu.chm
12/02/2006 10:18 PM 4,802,614 Beowulf.Cluster.Computing.With.Linux.Second.Edition.eBook-Li.chm
12/02/2006 10:18 PM 3,165,168 BeyondDreamweaver.chm
12/02/2006 10:18 PM 8,924,654 Building Dynamic Websites with Macromedia Studio MX 2004.chm

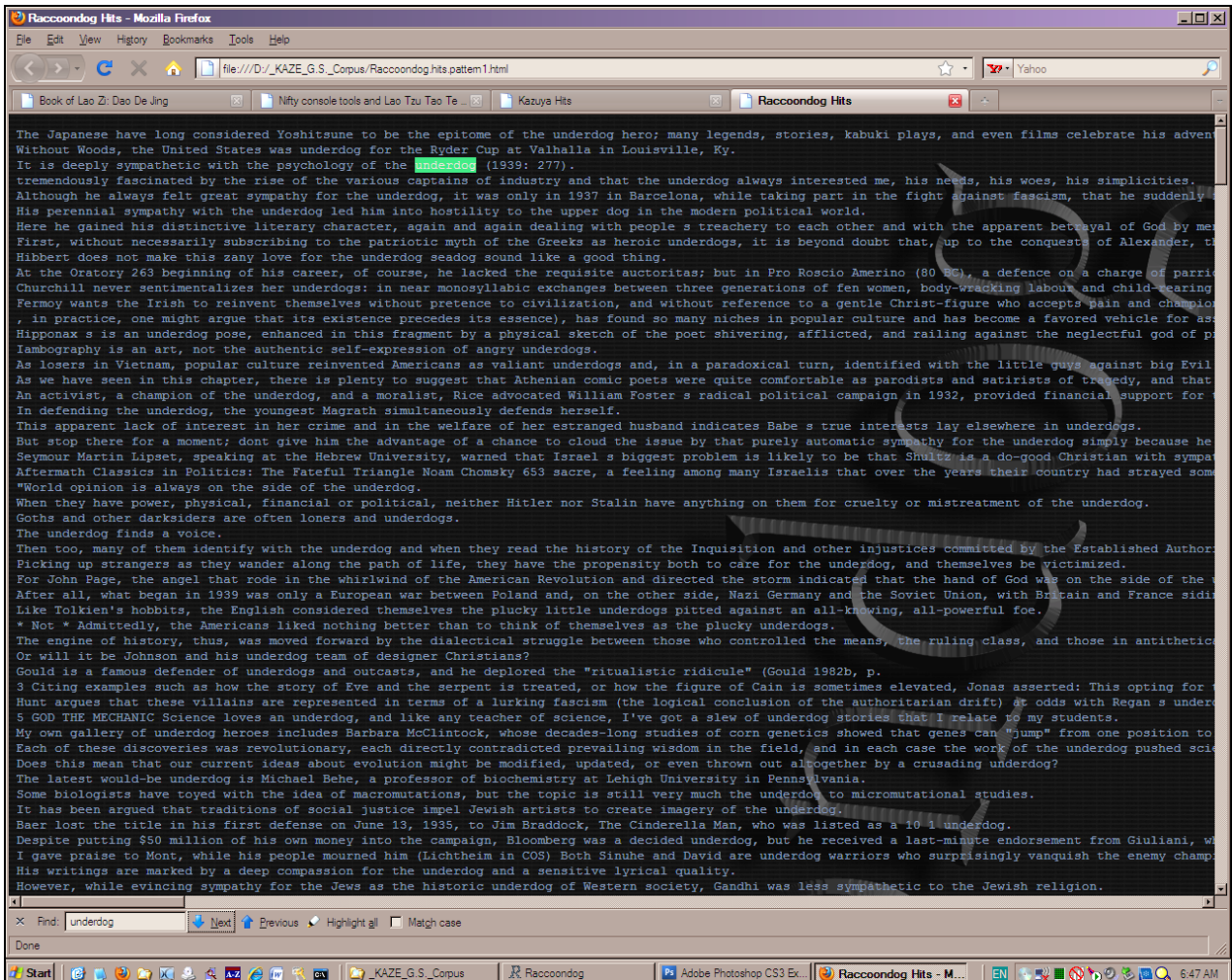
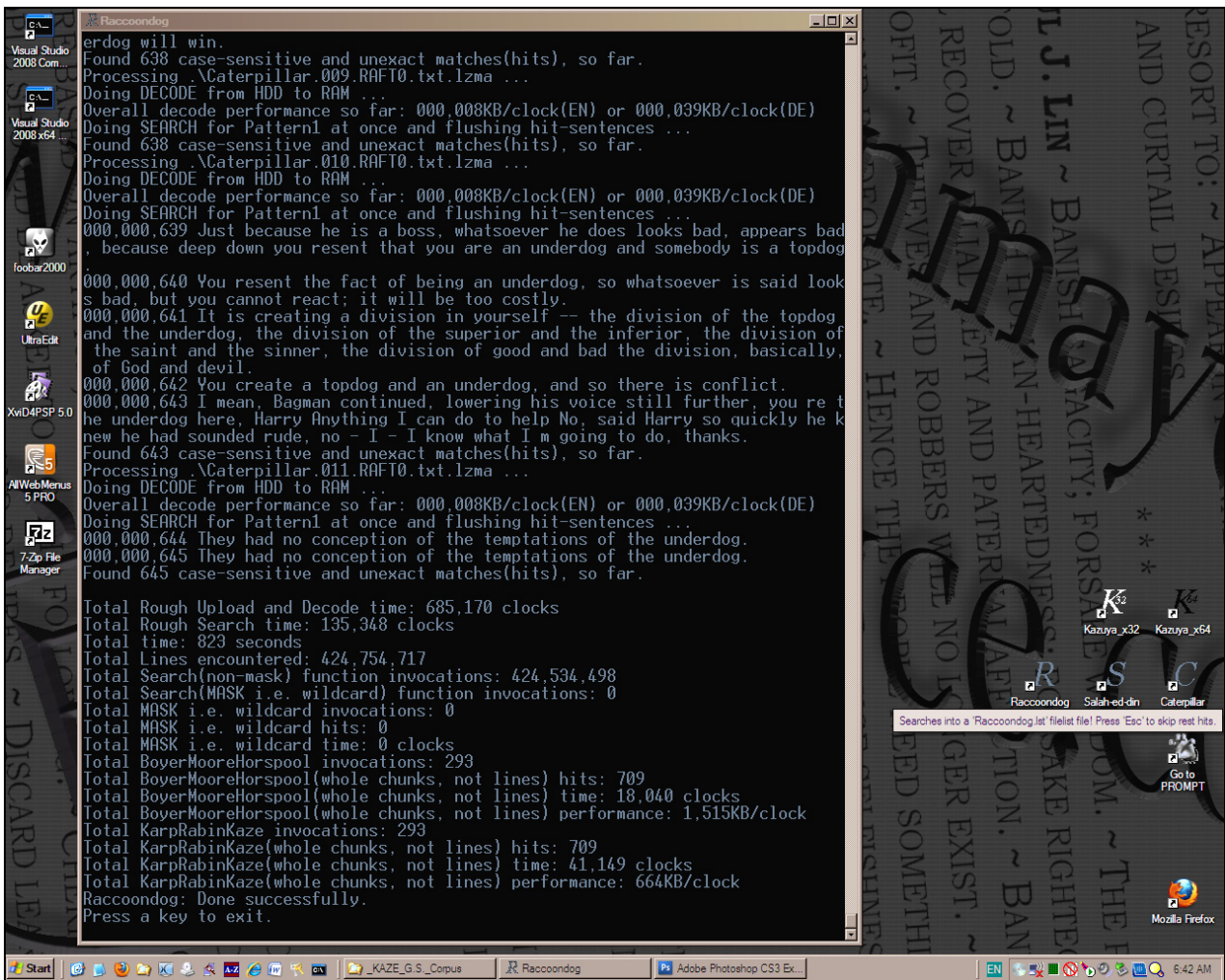
```

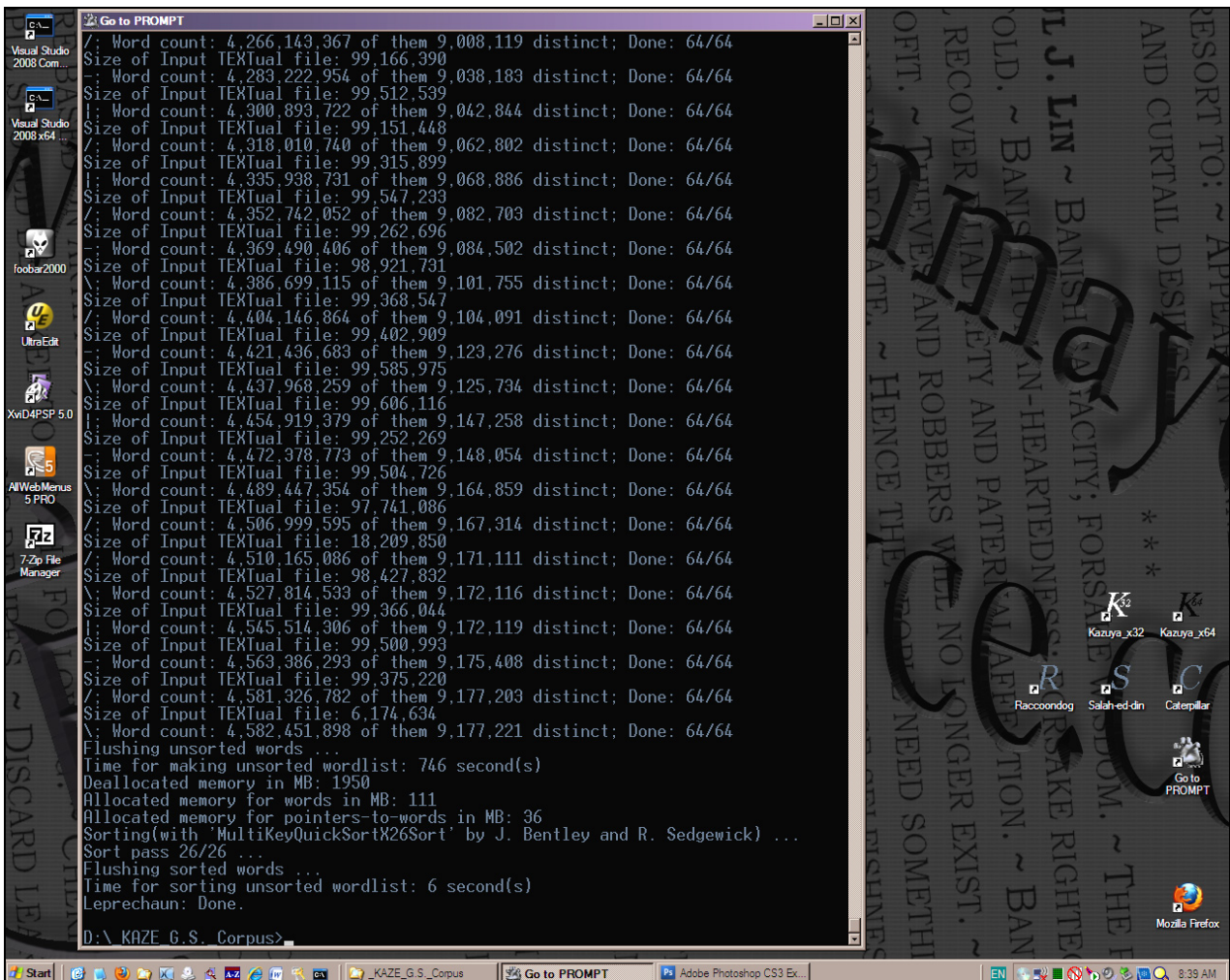
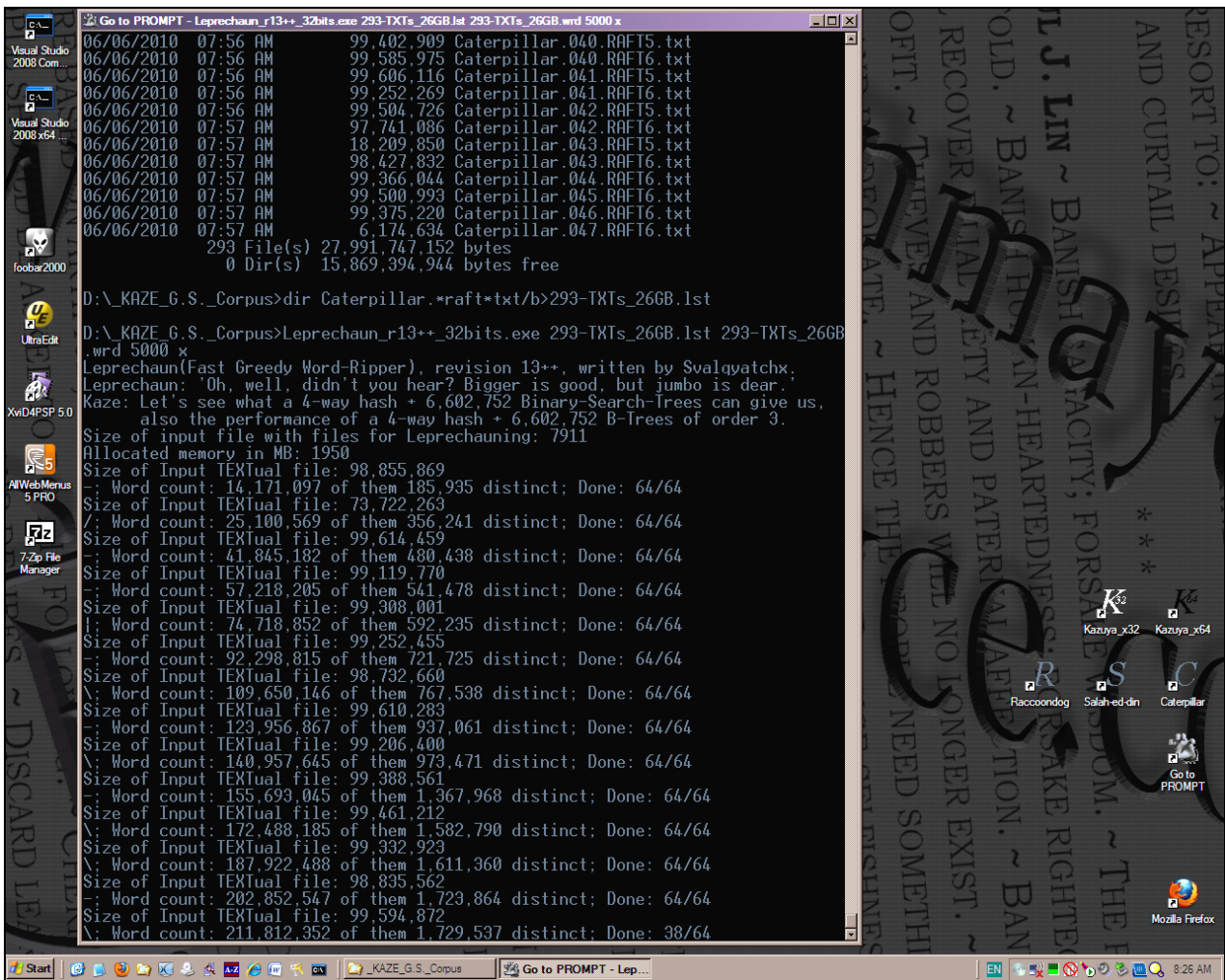












D:_KAZE_G.S._Corpus>type Leprechaun.LOG

Leprechaun report:

A(not always THE) Binary-Search-Tree with the longest path(height, PEAK, number of levels):

```
]syslade]
]swincian]
]swedloff]
]surtaxez[
]suddenne]
]stongrly]
]spellchk]
]spammail[
]shouzoug[
]shotaike[
]shitench[
]shahrani[
]sgcenari]
]sessionx[
]sedanais]
]schebaum[ ROOT
]scaunele]
]scappard]
]scachans[
]santinha]
]sankhaya[
]saleeite]
]saisihan]
```

Above Binary-Search-Tree with MaxPEAK = 13 has NODEs = 23 and LEAFs = 7

Legend:

At left side of the word - '[' means no left successor

At left side of the word - ']' means left successor exists

At right side of the word - '[' means no right successor

At right side of the word - ']' means right successor exists

Bytes per second performance: 37,522,449B/s

Words per second performance: **6,142,696W/s**

Input File with a list of TEXTual Files: 293-TXTs_26GB.lst

Size of all TEXTual Files: 27,991,747,152

Word count: 4,582,451,898 of them 9,177,221 distinct

Number Of Files: 293

Number Of Lines: 424754717

Allocated memory in MB: 1950

Number Of Trees(GREATER THE BETTER): **2855919**

Forest population(Hash Function Quality regarding Collisions i.e. Hash Table Utilization): **43%**

Number Of Hash Collisions(Distinct WORDs - Number Of Trees): **6321302**

Maximum Attempts to Find/Put a WORD into a Binary-Search-Tree: '13'

Total Attempts to Find/Put WORDs into Binary-Search-Trees: **4,746,283,042**

Total Number of LEAFs in Binary-Search-Trees(GREATER THE BETTER): 4,361,992

Perfectly-Balanced-Binary-Search-Tree for MaxNODEs = 34 must have PEAK = 6 = rounding down of integer (1+lb(34))

Binary-Search-Tree(1st out of 1) with MaxNODEs = 34 has PEAK = 11 and LEAFs = 11

Binary-Search-Tree(1st out of 2) with MaxPEAK = '13' has NODEs = 23 and LEAFs = 7

Binary-Search-Tree(1st out of 3) with MaxLEAFs = 12 has NODEs = 27 and PEAK = 8

```
words with length 01 occupy 0,033KB of 0,162KB given i.e. 19% utilization
words with length 02 occupy 0,033KB of 0,162KB given i.e. 19% utilization
words with length 03 occupy 0,040KB of 0,162KB given i.e. 24% utilization
words with length 04 occupy 0,158KB of 0,646KB given i.e. 24% utilization
words with length 05 occupy 0,487KB of 1,775KB given i.e. 27% utilization
words with length 06 occupy 0,991KB of 3,549KB given i.e. 27% utilization
words with length 07 occupy 1,431KB of 5,968KB given i.e. 23% utilization
words with length 08 occupy 1,803KB of 7,581KB given i.e. 23% utilization
words with length 09 occupy 1,643KB of 8,549KB given i.e. 19% utilization
words with length 10 occupy 1,546KB of 8,065KB given i.e. 19% utilization
words with length 11 occupy 1,317KB of 7,420KB given i.e. 17% utilization
words with length 12 occupy 1,131KB of 6,130KB given i.e. 18% utilization
words with length 13 occupy 0,945KB of 5,162KB given i.e. 18% utilization
words with length 14 occupy 0,796KB of 4,033KB given i.e. 19% utilization
words with length 15 occupy 0,662KB of 3,226KB given i.e. 20% utilization
words with length 16 occupy 0,561KB of 2,904KB given i.e. 19% utilization
words with length 17 occupy 0,461KB of 2,259KB given i.e. 20% utilization
words with length 18 occupy 0,394KB of 1,613KB given i.e. 24% utilization
words with length 19 occupy 0,335KB of 1,291KB given i.e. 25% utilization
words with length 20 occupy 0,297KB of 1,130KB given i.e. 26% utilization
words with length 21 occupy 0,266KB of 0,968KB given i.e. 27% utilization
words with length 22 occupy 0,248KB of 0,807KB given i.e. 30% utilization
words with length 23 occupy 0,222KB of 0,646KB given i.e. 34% utilization
words with length 24 occupy 0,210KB of 0,484KB given i.e. 43% utilization
words with length 25 occupy 0,194KB of 0,484KB given i.e. 40% utilization
words with length 26 occupy 0,178KB of 0,323KB given i.e. 55% utilization
words with length 27 occupy 0,164KB of 0,323KB given i.e. 50% utilization
words with length 28 occupy 0,160KB of 0,323KB given i.e. 49% utilization
words with length 29 occupy 0,150KB of 0,323KB given i.e. 46% utilization
words with length 30 occupy 0,138KB of 0,162KB given i.e. 85% utilization
words with length 31 occupy 0,134KB of 0,162KB given i.e. 82% utilization
```

Total pseudo(including hash table) memory utilization: 22%

Total real(wordlist's words vs allocated block) memory utilization: 47/1000

Used value for third parameter in KB: 5000

Use next time as third parameter: 4279-

Time for making unsorted wordlist: 746 second(s)

Time for sorting unsorted wordlist: 6 second(s)

```

#define ulPrime ((unsigned long) 0x00FF00F1)
#define ulBase ((unsigned long) 127)
// 9,223,372,036,854,775,807
// 74,051,159,531,521,793
// 257^7 = 19,031,147,999,601,100,801
// 257^8 = 8,594,754,748,609,397,887
// 127^9 = 362,033,331,456,891,249
// 57^10 = 665,416,609,183,179,841
// 13^16 = 244,140,625
// 5^12 = 815,730,721
// 13^8 =

long KarpRabinkazeHits (char * pbTarget,
char * pbPattern,
unsigned long cbTarget,
unsigned long cbPattern)
{
    unsigned int i;
    char * pbTargetMax = pbTarget + cbTarget;
    char * pbPatternMax = pbPattern + cbPattern;
    unsigned long ulBaseToPowerMod = 1;
    register unsigned long ulHashPattern = 0;
    unsigned long ulHashTarget = 0;
    long hits = 0;
//unsigned long count;
//char * buf1;
//char * buf2;

    if (cbPattern > cbTarget)
        return(0);

    // Compute the power of the left most character in base ulBase
    //for (i = 1; i < cbPattern; i++) ulBaseToPowerMod = (ulBase * ulBaseToPowerMod);

    // Calculate the hash function for the src (and the first dst)
    while (pbPattern < pbPatternMax)
    {
        // Below lines give 366KB/clock for 'underdog':
        //ulHashPattern = (ulHashPattern*ulBase + *pbPattern);
        //ulHashTarget = (ulHashTarget*ulBase + *pbTarget);
        pbPattern++;
        pbTarget++;
    }
    // Below lines give 436KB/clock for 'underdog' + requirement pattern to be 4 chars min.:
    //ulHashPattern = ( (* (long *) (pbPattern-cbPattern)) & 0xffffffff00 ) + *(pbPattern-1);
    //ulHashTarget = ( (* (long *) (pbTarget-cbPattern)) & 0xffffffff00 ) + *(pbTarget-1);
    // Below lines give 482KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
    //ulHashPattern = ( (* (unsigned short *) (pbPattern-cbPattern)) | *(pbPattern-1) );
    //ulHashTarget = ( (* (unsigned short *) (pbTarget-cbPattern)) | *(pbTarget-1) );
    // Below lines give 482KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
    //ulHashPattern = ( (* (unsigned short *) (pbPattern-cbPattern)) & 0xff00 ) + *(pbPattern-1);
    //ulHashTarget = ( (* (unsigned short *) (pbTarget-cbPattern)) & 0xff00 ) + *(pbTarget-1);
    // Below lines give 605KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
    //ulHashPattern = ( (* (unsigned short *) (pbPattern-cbPattern)) << 8 ) + *(pbPattern-1);
    //ulHashTarget = ( (* (unsigned short *) (pbTarget-cbPattern)) << 8 ) + *(pbTarget-1);
    // Below lines give 668KB/clock for 'underdog':
    ulHashPattern = ( (* (char *) (pbPattern-cbPattern)) << 8 ) + *(pbPattern-1);
    ulHashTarget = ( (* (char *) (pbTarget-cbPattern)) << 8 ) + *(pbTarget-1);

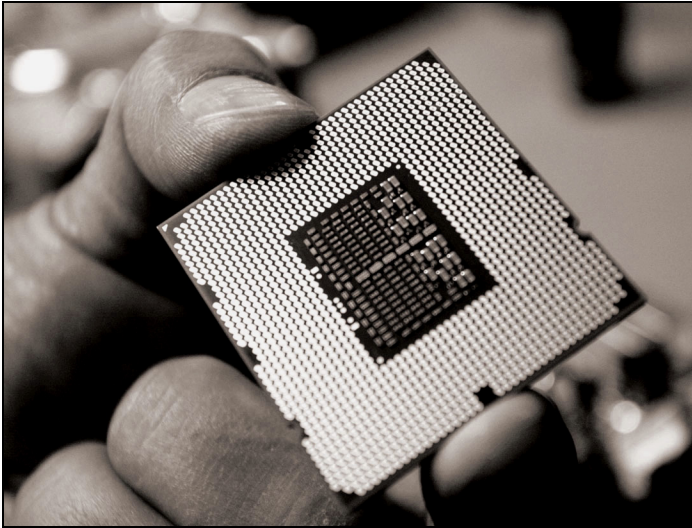
    // Dynamically produce hash values for the string as we go
    for ( ; ; )
    {
        if ( (ulHashPattern == ulHashTarget) && !memcmpKAZE(pbPattern-cbPattern, pbTarget-cbPattern, (unsigned int)cbPattern) )
        // if ( ulHashPattern == ulHashTarget ) {
        //
        // count = cbPattern;
        // buf1 = pbPattern-cbPattern;
        // buf2 = pbTarget-cbPattern;
        // while ( --count && *(char *)buf1 == *(char *)buf2 ) {
        //     buf1 = (char *)buf1 + 1;
        //     buf2 = (char *)buf2 + 1;
        // }
        //
        // if ( *((unsigned char *)buf1) - *((unsigned char *)buf2) == 0 ) hits++;
        // }
        hits++;
        //return((long)(pbTarget-cbPattern));

        if (pbTarget == pbTargetMax)
            return(hits);

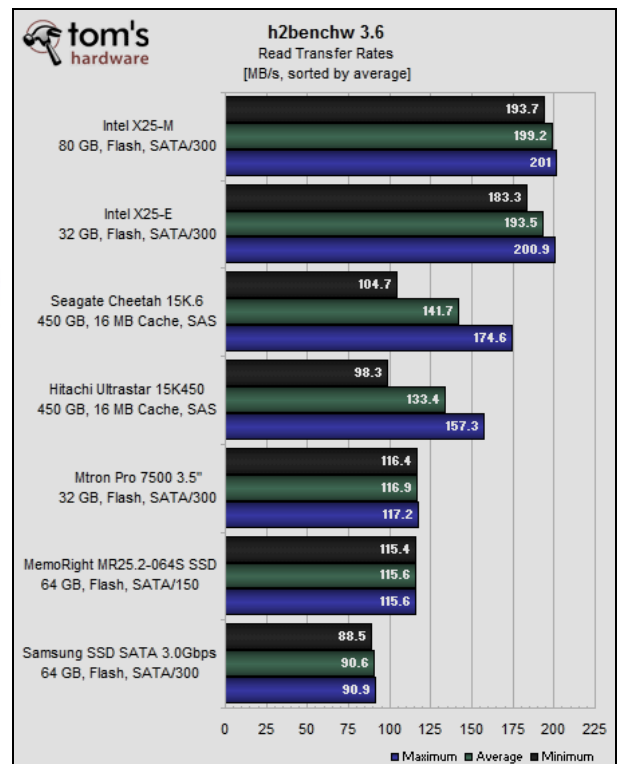
        // Below line gives 482KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
        //ulHashTarget = ( (* (unsigned short *) (pbTarget+1-cbPattern)) | *pbTarget );
        // Below line gives 436KB/clock for 'underdog' + requirement pattern to be 4 chars min.:
        //ulHashTarget = ( (* (long *) (pbTarget+1-cbPattern)) & 0xffffffff00 ) + *pbTarget;
//; Line 696
// movsx esi, BYTE PTR [ebx]
// mov ecx, DWORD PTR [edx+1]
// and ecx, -256 ; ffffffff00H
// add ecx, esi
// Below line gives 482KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
//ulHashTarget = ( (* (unsigned short *) (pbTarget+1-cbPattern)) & 0xff00 ) + *pbTarget;
//; Line 691
// movsx esi, BYTE PTR [ebx]
// xor ecx, ecx
// mov cx, WORD PTR [edx+1]
// and ecx, 65280 ; 0000ff00H
// add ecx, esi
// Below line gives 605KB/clock for 'underdog' + requirement pattern to be 2 chars min.:
//ulHashTarget = ( (* (unsigned short *) (pbTarget+1-cbPattern)) << 8 ) + *pbTarget;
// Below line gives 668KB/clock for 'underdog':
ulHashTarget = ( (* (char *) (pbTarget+1-cbPattern)) << 8 ) + *pbTarget;
//; Line 718
// movsx ecx, BYTE PTR [eax+1]
// movsx edx, BYTE PTR [ebp]
// shl ecx, 8
// add ecx, edx
// Below line gives 366KB/clock for 'underdog':
//ulHashTarget = (ulHashTarget - *(pbTarget-cbPattern)*ulBaseToPowerMod)*ulBase + *pbTarget;
pbTarget++;
    }
}

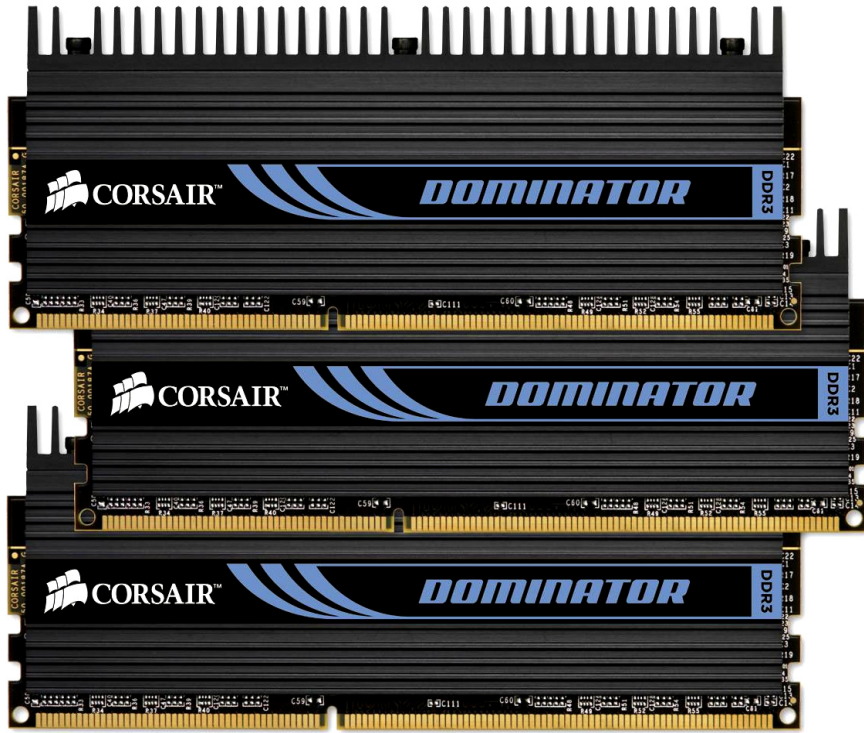
```

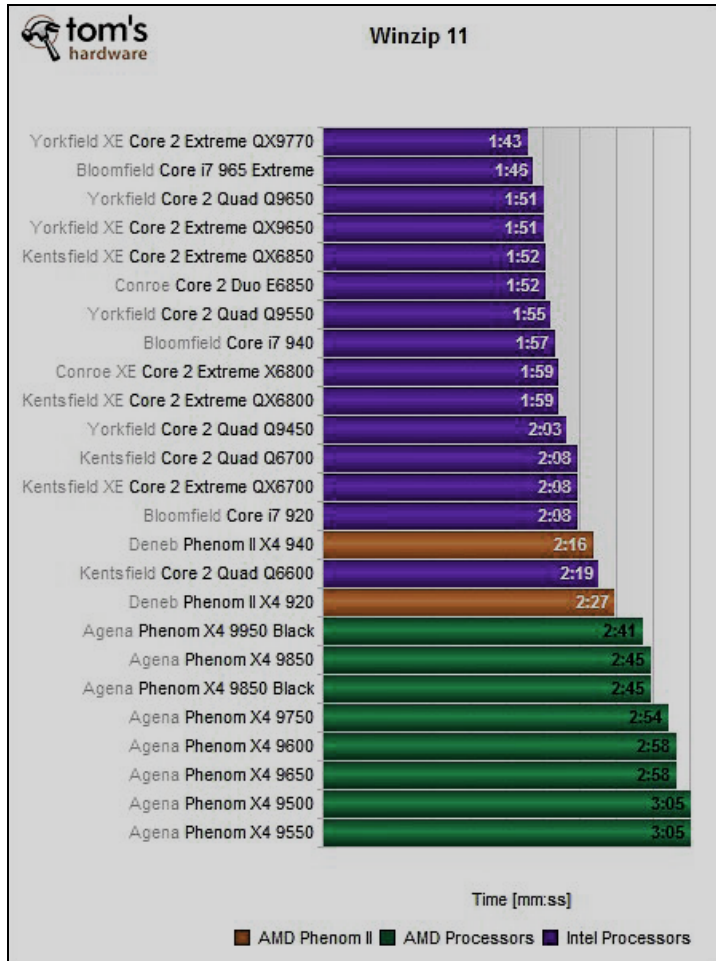
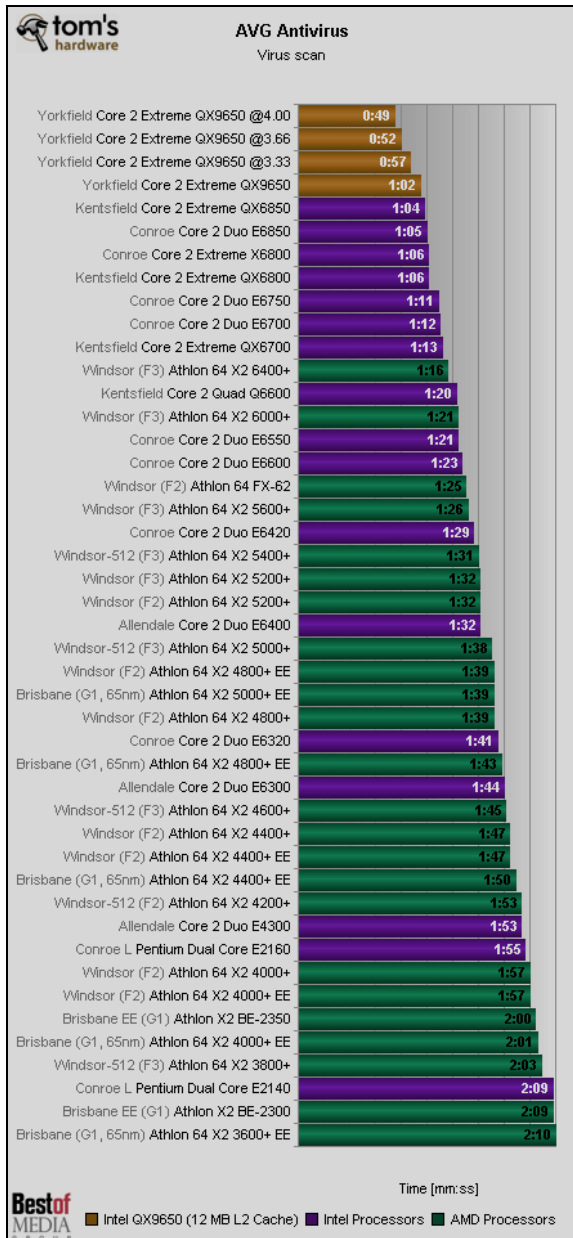
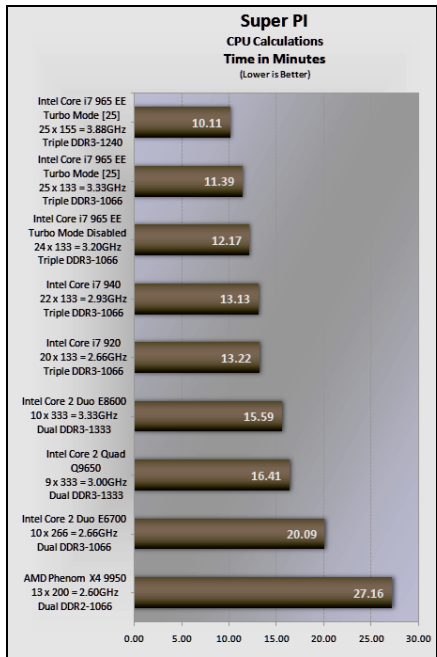
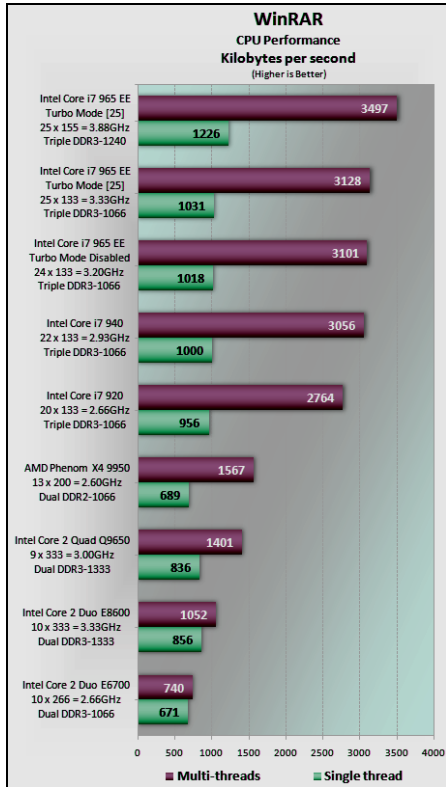
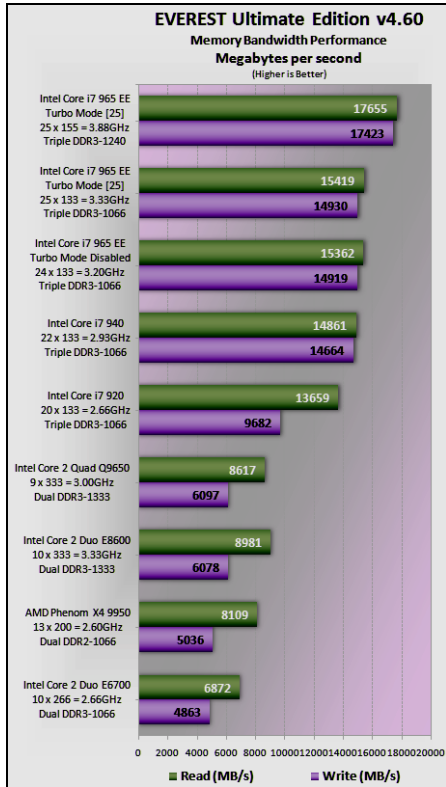
core i7 chip(731 million transistors) - the current dominator; plus heart-touching-data-storages:



SSD drives - simply the future:







CPU Queen

CPU	CPU Clock	Motherboard	Chipset	Memory	CL-RCD-RP-RAS	Score
4x Core i7 Extreme 965 HT	3333 MHz	Asus P6T Deluxe	X58	Triple DDR3-1333	9-9-9-24 CR1	30786
8x Xeon E5462	2800 MHz	Intel S5400SF	i5400	Quad DDR2-640FB	5-5-5-15	30472
4x Core 2 Extreme QX9650	3000 MHz	Gigabyte GA-EP35C-DS3R	P35	Dual DDR3-1066	8-8-8-20 CR2	21421
8x Xeon L5320	1866 MHz	Intel S5000VCL	i5000V	Dual DDR2-533FB	4-4-4-12	20452
4x Core 2 Extreme QX6700	2666 MHz	Intel D975XB2	i975X	Dual DDR2-667	5-5-5-15	19166
4x Phenom II X4 Black 940	3000 MHz	Asus M3N78-EM	GeForce8300 Int.	Ganged Dual DDR2-800	5-5-5-18 CR2	18636
4x Xeon 5140	2333 MHz	Intel S5000VSA	i5000V	Dual DDR2-667FB	5-5-5-15	16729
8x Opteron HE 2344	1700 MHz	Tyan Thunder n3600R	nForcePro-3600	Unganged Dual DDR2-667R	5-5-5-15 CR1	16146
4x Phenom X4 9500	2200 MHz	Asus M3A	AMD770	Ganged Dual DDR2-800	5-5-5-18 CR2	13693
2x Core 2 Duo E6700	2666 MHz	Abit AB9	P965	Dual DDR2-800	5-5-5-18 CR2	11406
2x Athlon64 X2 Black 6400+	3200 MHz	MSI K9N SLI Platinum	nForce570SLI	Dual DDR2-800	4-4-4-11 CR1	11169
2x Core 2 Duo P8400	2266 MHz	MSI MegaBook PR201	GM45 Int.	Dual DDR2-667	5-5-5-15	9578
2x Pentium T3400	2166 MHz	Toshiba Satellite L305	GL40 Int.	Dual DDR2-667	5-5-5-13	9145
2x Core Duo T2500	2000 MHz	Asus N4L-VM DH	i945GT Int.	Dual DDR2-667	5-5-5-15	7793
2x Core 2 Duo T5600	1833 MHz	Asus F3000Jc Notebook	i945PM	Dual DDR2-667	5-5-5-15	7717
2x Athlon64 X2 4000+	2100 MHz	ASRock ALiveNF7G-HDready	nForce7050-630a Int.	Dual DDR2-700	5-5-5-18 CR2	7280
2x Pentium EE 955 HT	3466 MHz	Intel D955XBK	i955X	Dual DDR2-667	4-4-4-11	7098
2x Xeon	3066 MHz	Asus PCH-DL	i875P + PAT	Dual DDR333	2-2-2-5	6188
2x Opteron 240	1400 MHz	MSI K8D Master3-133 FS	AMD8100	Dual DDR400R	3-4-4-8 CR1	4863
2x PIII-S	1266 MHz	MSI Pro266TD Master-LR	ApolloPro266TD	DDR266 SDRAM	2-3-3-6 CR2	4857
P4EE HT	3733 MHz	Intel SE7230NH1LX	iE7230	Dual DDR2-667	5-5-5-15	4210
Opteron 248	2200 MHz	MSI K8T Master1-FAR	K8T800	Dual DDR266R	2-3-3-6 CR1	3851
Atom 230	1600 MHz	Intel D945GCLF	i945GC	DDR2-533 SDRAM	4-4-4-12	3779

CPU AES

CPU	CPU Clock	Motherboard	Chipset	Memory	CL-RCD-RP-RAS	Score
8x Xeon E5462	2800 MHz	Intel S5400SF	i5400	Quad DDR2-640FB	5-5-5-15	41625
8x Xeon L5320	1866 MHz	Intel S5000VCL	i5000V	Dual DDR2-533FB	4-4-4-12	27698
4x Core i7 Extreme 965 HT	3333 MHz	Asus P6T Deluxe	X58	Triple DDR3-1333	9-9-9-24 CR1	26703
8x Opteron HE 2344	1700 MHz	Tyan Thunder n3600R	nForcePro-3600	Unganged Dual DDR2-667R	5-5-5-15 CR1	22599
4x Core 2 Extreme QX9650	3000 MHz	Gigabyte GA-EP35C-DS3R	P35	Dual DDR3-1066	8-8-8-20 CR2	22435
4x Phenom II X4 Black 940	3000 MHz	Asus M3N78-EM	GeForce8300 Int.	Ganged Dual DDR2-800	5-5-5-18 CR2	21658
4x Core 2 Extreme QX6700	2666 MHz	Intel D975XB2	i975X	Dual DDR2-667	5-5-5-15	19896
C7	1500 MHz	VIA EPIA EN	CN700 Int.	DDR2-533 SDRAM	4-4-4-12 CR2	17358
4x Xeon 5140	2333 MHz	Intel S5000VSA	i5000V	Dual DDR2-667FB	5-5-5-15	17320
4x Phenom X4 9500	2200 MHz	Asus M3A	AMD770	Ganged Dual DDR2-800	5-5-5-18 CR2	14802
2x Core 2 Duo E6700	2666 MHz	Abit AB9	P965	Dual DDR2-800	5-5-5-18 CR2	9969
2x Pentium EE 955 HT	3466 MHz	Intel D955XBK	i955X	Dual DDR2-667	4-4-4-11	8970
2x Core 2 Duo P8400	2266 MHz	MSI MegaBook PR201	GM45 Int.	Dual DDR2-667	5-5-5-15	8443
2x Athlon64 X2 Black 6400+	3200 MHz	MSI K9N SLI Platinum	nForce570SLI	Dual DDR2-800	4-4-4-11 CR1	8339
2x Pentium T3400	2166 MHz	Toshiba Satellite L305	GL40 Int.	Dual DDR2-667	5-5-5-13	7966
2x Core Duo T2500	2000 MHz	Asus N4L-VM DH	i945GT Int.	Dual DDR2-667	5-5-5-15	7109

CPU ZLib

CPU	CPU Clock	Motherboard	Chipset	Memory	CL-RCD-RP-RAS	Score
8x Xeon E5462	2800 MHz	Intel S5400SF	i5400	Quad DDR2-640FB	5-5-5-15	139481 KB/s
4x Core i7 Extreme 965 HT	3333 MHz	Asus P6T Deluxe	X58	Triple DDR3-1333	9-9-9-24 CR1	112330 KB/s
8x Xeon L5320	1866 MHz	Intel S5000VCL	i5000V	Dual DDR2-533FB	4-4-4-12	95887 KB/s
8x Opteron HE 2344	1700 MHz	Tyan Thunder n3600R	nForcePro-3600	Unganged Dual DDR2-667R	5-5-5-15 CR1	88845 KB/s
4x Phenom II X4 Black 940	3000 MHz	Asus M3N78-EM	GeForce8300 Int.	Ganged Dual DDR2-800	5-5-5-18 CR2	80081 KB/s
4x Core 2 Extreme QX9650	3000 MHz	Gigabyte GA-EP35C-DS3R	P35	Dual DDR3-1066	8-8-8-20 CR2	77167 KB/s
4x Core 2 Extreme QX6700	2666 MHz	Intel D975XB2	i975X	Dual DDR2-667	5-5-5-15	69756 KB/s
4x Xeon 5140	2333 MHz	Intel S5000VSA	i5000V	Dual DDR2-667FB	5-5-5-15	60995 KB/s
4x Phenom X4 9500	2200 MHz	Asus M3A	AMD770	Ganged Dual DDR2-800	5-5-5-18 CR2	58396 KB/s
2x Athlon64 X2 Black 6400+	3200 MHz	MSI K9N SLI Platinum	nForce570SLI	Dual DDR2-800	4-4-4-11 CR1	38059 KB/s
2x Core 2 Duo E6700	2666 MHz	Abit AB9	P965	Dual DDR2-800	5-5-5-18 CR2	35355 KB/s
2x Pentium EE 955 HT	3466 MHz	Intel D955XBK	i955X	Dual DDR2-667	4-4-4-11	29844 KB/s
2x Core 2 Duo P8400	2266 MHz	MSI MegaBook PR201	GM45 Int.	Dual DDR2-667	5-5-5-15	29223 KB/s
2x Pentium T3400	2166 MHz	Toshiba Satellite L305	GL40 Int.	Dual DDR2-667	5-5-5-13	28519 KB/s
2x Xeon	3066 MHz	Asus PCH-DL	i875P + PAT	Dual DDR333	2-2-2-5	25040 KB/s
2x Core 2 Duo T5600	1833 MHz	Asus F3000Jc Notebook	i945PM	Dual DDR2-667	5-5-5-15	23969 KB/s
2x Athlon64 X2 4000+	2100 MHz	ASRock ALiveNF7G-HDready	nForce7050-630a Int.	Dual DDR2-700	5-5-5-18 CR2	23852 KB/s



Why is Intel in the NAND Solution Business?

Normalized CPU Performance
 Normalized Media Access Time for 20K Read

Measured CPU performance scaling = 175x
 Measured HDD performance scaling = 1.3X since Jan'96

Intel® SSD
 Intel® Turbo Memory

Our Mission is to remove the IO bottleneck

Delivering on the full promise of SSDs

Intel® X18-M and X25-M Mainstream SATA SSD

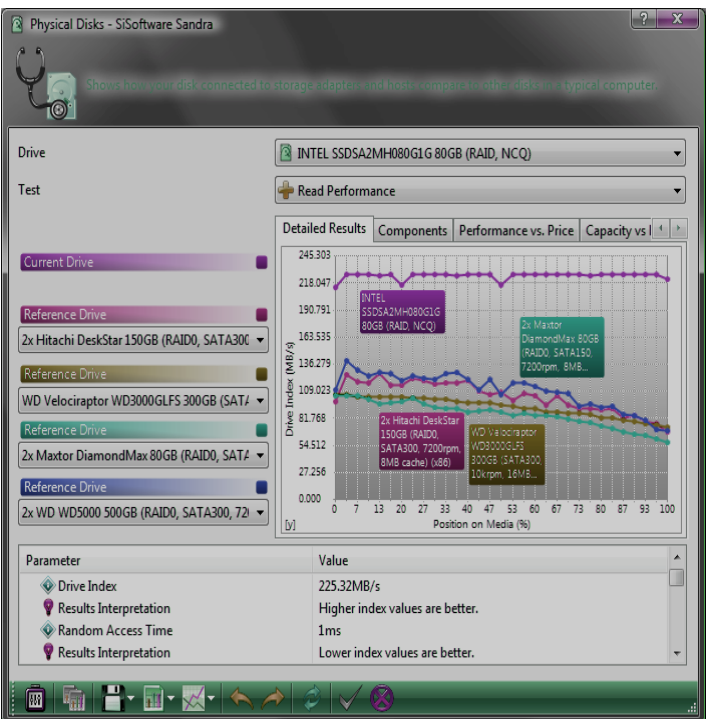
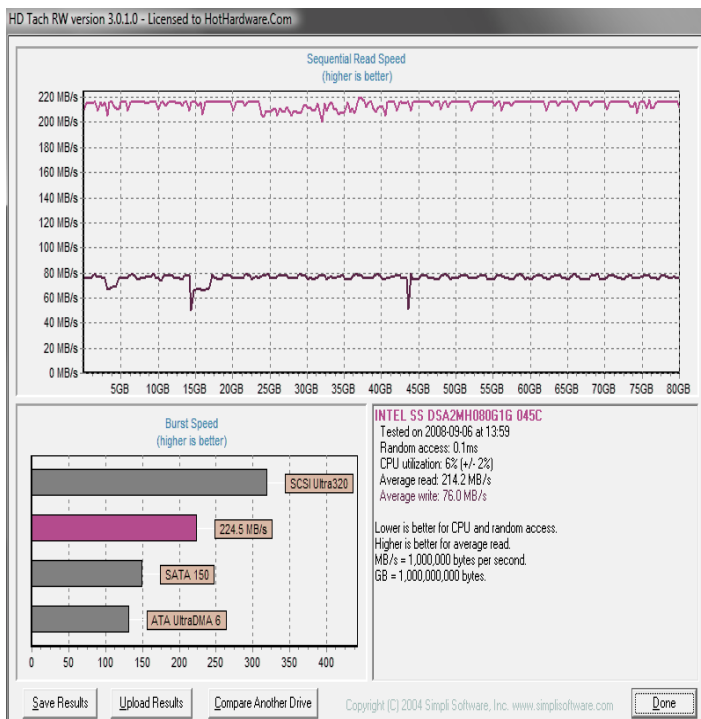
- Up to 250MB/s Sustained Read
- Up to 70MB/s Sustained Write
- 0.15 W Active Power, 0.06W Idle Power
- GB/day client workload → >100GB/day for 5 years (accepted is 20GB/day)

5+ Years Useful Life for Client PCs

Intel® X25-E Extreme SATA SSD

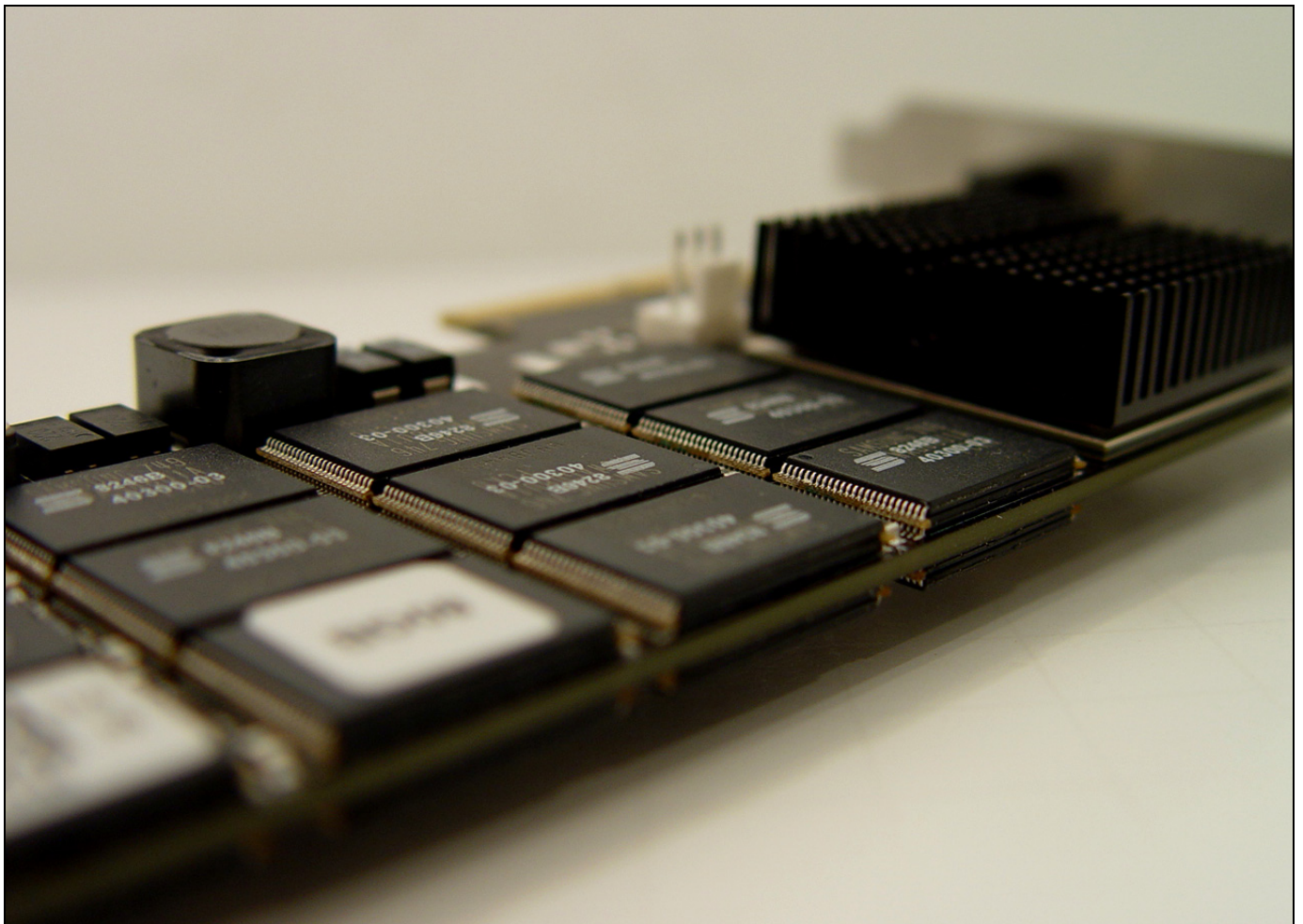
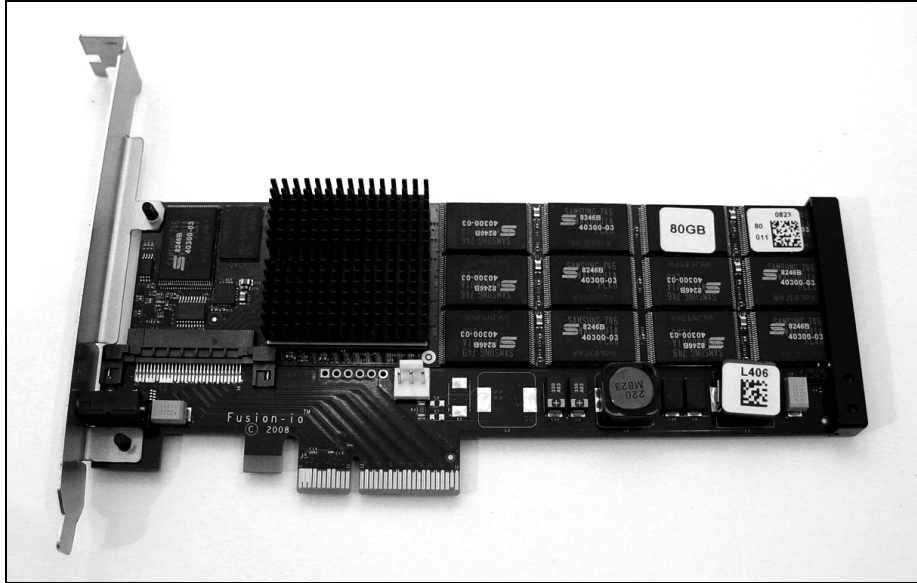
- 10 Channels Architecture with 50nm SLC ONFI 1.0 NAND
- Up to 250MB/s Sustained Read
- Up to 170MB/s Sustained Write
- 2.4 W Active Power, 0.06W Idle Power
- 2 Million Hours MTBF
- 35,000 IOPS (4KB Read), 3,300 IOPS (4KB Write)

Highest IOPS & Endurance to Replace many 15K RPM HDDs



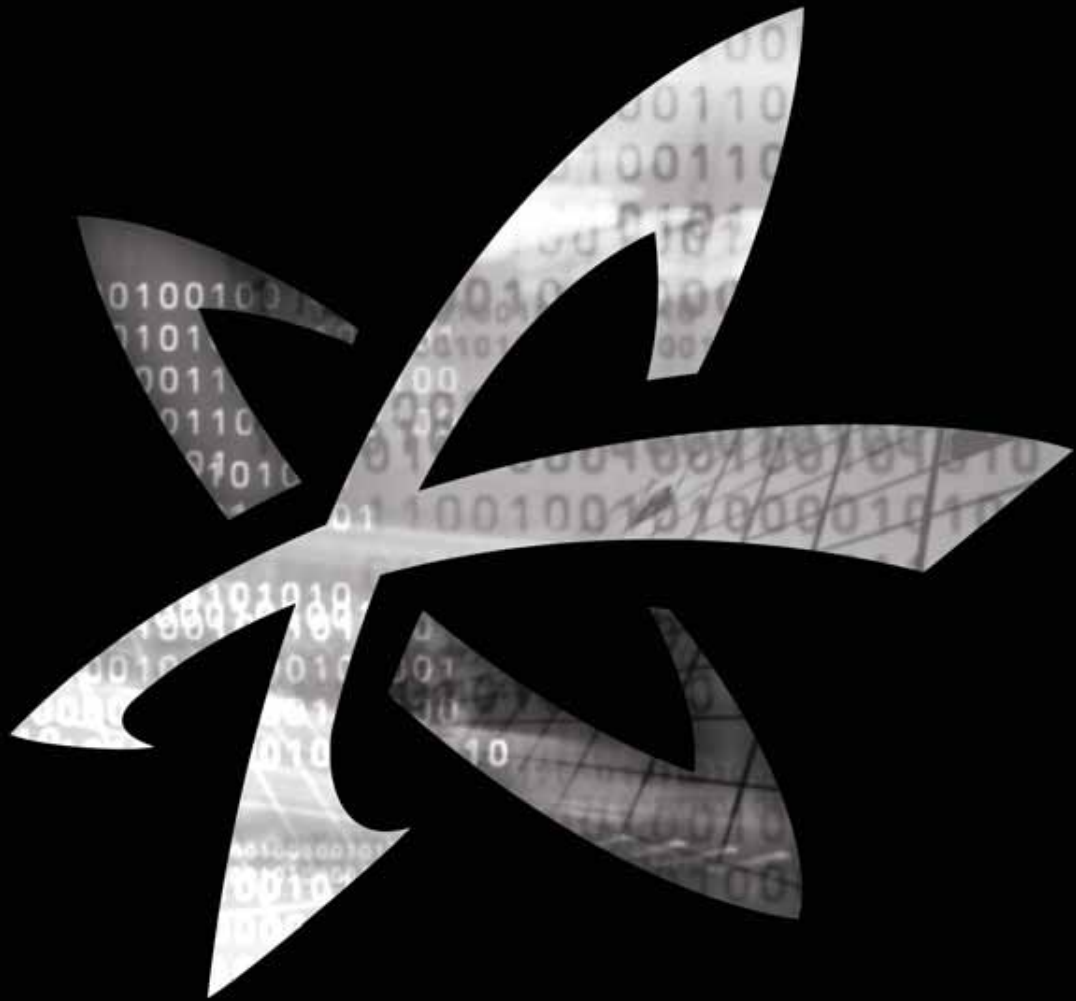
Fusion ioDrive SPECIFICATIONS:

NAND Type: Single Level Cell (SLC)
Read Bandwidth: 700 MB/s (random 16K)
Access Latency: 50µs
Bus Interface: PCI-Express x4
Operating Systems: Microsoft 64-Bit Windows(64-Bit Windows XP, Vista, Server 2003 & 2008)



* * *

with the **ioDrive Duo**, it is now possible for application, database and system administrators to get previously unheard-of levels of performance, protection and capacity utilization from a single server. Performance for multiple **ioDrive Duos** scales linearly, allowing any enterprise to scale performance to **six gigabytes per-second** (Gbytes/sec) of read bandwidth and over 500,000 read IOPS by using just four **ioDrive Duos**.



Fusion-io's Solid State Storage – A New Standard for Enterprise-Class Reliability

FUSION-io

Fusion-io's Solid State Storage – A New Standard for Enterprise-Class Reliability

Fusion-io offers solid state storage solutions based on NAND flash that provide a level of integrity and availability for mission-critical data that exceeds today's solid state storage solutions and significantly surpasses that of enterprise-class rotating magnetic storage devices.

With throughput and seek times many times faster than the fastest disk arrays, it is little wonder that enterprise data centers have been keen to include NAND flash as part of their server infrastructure. The primary reason NAND flash has not been widely adopted in the computer industry is its reputation for unreliability. There is a long-standing view that NAND flash storage works well for non-mission-critical applications, such as media storage devices (where the occasional bit error generally translates into a slight audio hiss or a stray errant pixel in a video), but cannot be relied upon for applications where a bit error could crash an operating system or compromise the integrity of critical data.

System architects face a number of storage-related challenges and NAND flash technology presents its own set of unique problems. But Fusion-io has developed patent-pending techniques to create NAND flash-based storage with reliability equal to or exceeding that of disk-based storage. This paper describes several inventions and advancements Fusion-io has introduced to ensure data is not corrupted or lost. Additionally, this paper discusses the probability of catastrophic storage device failure and how Fusion-io's architecture ensures predictable, controlled management of early device failure, long-term device attrition and data changes due to external and data transport interference.

NAND Flash

Flash memory chips are a non-volatile storage medium (i.e., they can retain their information even in the absence of power). The most common types of flash chips are silicon-based NOR and NAND, named after the types of logic gates used in their design. NAND flash, introduced in 1989, has become the most commonly used type of flash chip, due to its quicker write speed. Flash memory continues to grow in popularity as its price steadily declines, its storage capacity increases, and its physical size continues to decrease.

In Fusion-io's storage devices, NAND flash chips are stacked several at a time (to increase density), operated in parallel (to increase throughput) and mounted on a printed circuit board (PCB) that plugs into a PCI-Express (PCIe) slot on the server or in the CPU. The flash media is integrated with the controller onto a single PCI-Express card.

NAND flash, as a storage medium, offers a number of benefits in comparison to rotating magnetic storage devices (aka HDD, Hard Disk Drives). NAND flash has no moving parts and is therefore significantly less prone to shock or movement disturbance. It is a high speed solution in both latency and throughput. Temperature and humidity resistance mean that it can operate in a number of different environments. Finally, NAND flash consumes significantly less power than rotating magnetic storage devices, particularly when you take into account secondary power requirements for device cooling.



However, NAND flash does introduce a number of potential failure points including:

- Media – Media failures can occur on the NAND flash chips themselves.
- Transport – Transport errors can occur anywhere along the path carrying data from the CPU through to the NAND flash chips.
- Management – There is a small chance that management problems can occur within the logic of the device itself. The code that controls the operation can contain technical problems that can result in data failures.
- External – External problems can affect any part of the process.
- Device Failure – Catastrophic hardware failure can also occur. This includes the possibility of internal short circuits and open circuits within the memory array itself.

Protecting the Data

Implementing a variety of design and architectural strategies for protecting data integrity, Fusion-io's NAND flash devices greatly exceed the reliability of rotating magnetic media storage devices, while providing performance that is orders of magnitude better. Fusion-io protects your data at every step, ensuring that nothing is lost or corrupted in transit or on the media.

Data Integrity

Data integrity means having a high degree of confidence that what you put into a storage system is exactly what you get out when you request that data and it is the most important function of a storage system. While being moved from a computer's RAM or CPU to the Fusion-io device, several proven industry-standard approaches are used to ensure data integrity. The CPU, chipset, and RAM use SECDED (Single Error Correct Double Error Detect) or chipkill (method for on-the-fly replacement of a failed chip) to ensure accuracy. Once data is written to the storage medium, it is again checked for accuracy.

When data is read from the storage medium, error correction techniques are again employed to ensure that the data being retrieved is correct. The device can correct a substantial portion of the data being read. NAND's reputation for unreliability is based on studies that show potential data loss without utilizing error correction – or less correction than that employed by the Fusion-io device. Using the methods described here, Fusion-io devices can produce results that exceed target error probability by about four times. Fusion-io's devices also use a patent-pending approach when writing data, which allows the data's path to be reconstructed from information generated during the write process.

Data Availability

Data availability means having a high degree of confidence that data stored will not be lost, either while in transition to the storage device or after it has been written to the media.

Fusion-io employs a wide variety of techniques to overcome some of the common problems associated with data availability in general, and also addresses some that are particular to NAND flash as a storage medium. Generally speaking, NAND flash is substantially more reliable than rotating magnetic media. It eliminates the chance of mechanical failure

(the failure associated with moving parts). There is, however, a chance of bad chips and chip wear-out. Fusion-io mitigates this risk using a variety of approaches.

Fusion-io's redundant, patent-pending approach to writing data allows data to be rebuilt at a very high rate of speed, ensuring rapid data availability. Data is also regularly moved and checked for accuracy to ensure that it does not deteriorate on the flash chip. This also consolidates good data and reallocates space on the drive to ensure greater data availability. This system also spreads data evenly across the device, ensuring uniform wear across all chips.

Additionally, Fusion-io uses multiple error correction code (ECC) techniques to identify and correct faulty data. Using ECCs, the device controller can correct up to 11 missing or incorrect bits out of every 240 bytes. One of the biggest benefits of ECC routines is that they allow the device to predict the likelihood of failure on individual chips. When a particular area of a chip has passed a set unreliability threshold, its data can be moved and that area will be taken out of service. The controller continues to identify and remove bad blocks, regions of chips or even entire chips so that ordinary wear-out does not cause catastrophic failure rather a very predictable wear-out.

Device Longevity

The majority of this paper has concentrated on NAND flash in an enterprise-class storage device, and how to leverage its strengths while overcoming its weaknesses. NAND flash, however, is only part of a Fusion-io's storage device. The flash chips reside on a PCIe adapter card that has a number of other parts as well, all of which are susceptible to failure. The life of a NAND flash storage device can be estimated by examining the failure rate of its component parts. Wear-out is generally a function of having lost enough storage cells that both capacity and reliability drop below acceptable thresholds. This can be assessed by evaluating and keeping a record of the amount of errors detected at each physical location.

NAND flash wears out at a predictable rate as described by the formulas below. Effective use of wear-leveling strategies employed by Fusion-io can significantly improve the life expectancy of its drives. Please note that the formulas are applied to both MLC and SLC NAND-based non-volatile memory technologies. Single-Level Cell (SLC) NAND and Multi-Level Cell (MLC) NAND offer capabilities that serve two very different types of applications – respectively, those requiring high performance at an attractive cost-per-bit and those seeking even higher performance over time, that are less cost-sensitive:

Average-lifetime = lifetime / read-write- ratio

TYPE / WRITE DUTY	AVERAGE ESTIMATED LIFETIME FORMULA
SLC flash @ 40% write duty	25 calendar years
MLC flash @ 20% write duty	10 calendar years
MLC flash @ 40% write duty	5 calendar years

Average estimated lifetime based on Fusion-io lab testing

The read/write ratio is difficult to predict, and will vary considerably from environment to environment. As a point of reference, the International Disk-drive Equipment and Materials Association (IDEMA), an industry trade group that publishes storage device standards, recommends a read/write ratio of 60%/40% for its server-class device reliability testing (IDEMA Standards, Document R3-98).

Flashback Protection

Enterprises have long sought to take advantage of the speed, size, low-power and high-performance of NAND Flash because of its potential to change the way they manage large amounts of active data. The primary objection to NAND flash has been the reliability of the medium. Fusion-io has eliminated this barrier by inventing a revolutionary self-healing technology, known as Flashback Protection, in our controllers that instantaneously restores, corrects and resurrects lost data in the flash-based storage sub-system. Flashback Protection is accomplished by collectively using advanced bit error correction, proactive data integrity monitoring of stored data and the recent addition of a dedicated chip to repair failed devices.

Fusion-io is the first and only company to bring RAID-class redundancy and reliability using Flashback Protection down to the card level. The Flashback Protection system allows users to diagnose and correct system errors. Fusion-io integrates dedicated NAND flash chips, which offer information that enables the detection of single bit errors. This technique eliminates data loss due to chip failures and extends the usable lifetime of the NAND flash-based storage device. The NAND flash chips on Fusion-io's products contain an innovative storage architecture that enable it to deliver the performance, and now the reliability, of a storage area network (SAN) at a fraction of the power, size and cost of traditional disk arrays.

Controlled Predictable Usage Versus Catastrophic Failure

Among the greatest reliability benefits of the Fusion-io storage device is its ability to:

- Restore and Protect data
- Monitor and predict media wear-out
- Correct bad data as necessary
- Take blocks out of service when their failure rate becomes unacceptable
- Replace bad chips on-the-fly
- Move the data to a known good location (and update corresponding mapping information)

Data stored on the Fusion-io medium is double protected using both ECCs and parity data on the redundant chip. The net effect is that wear-out of the device, instead of being catastrophic, is predictable and incremental.



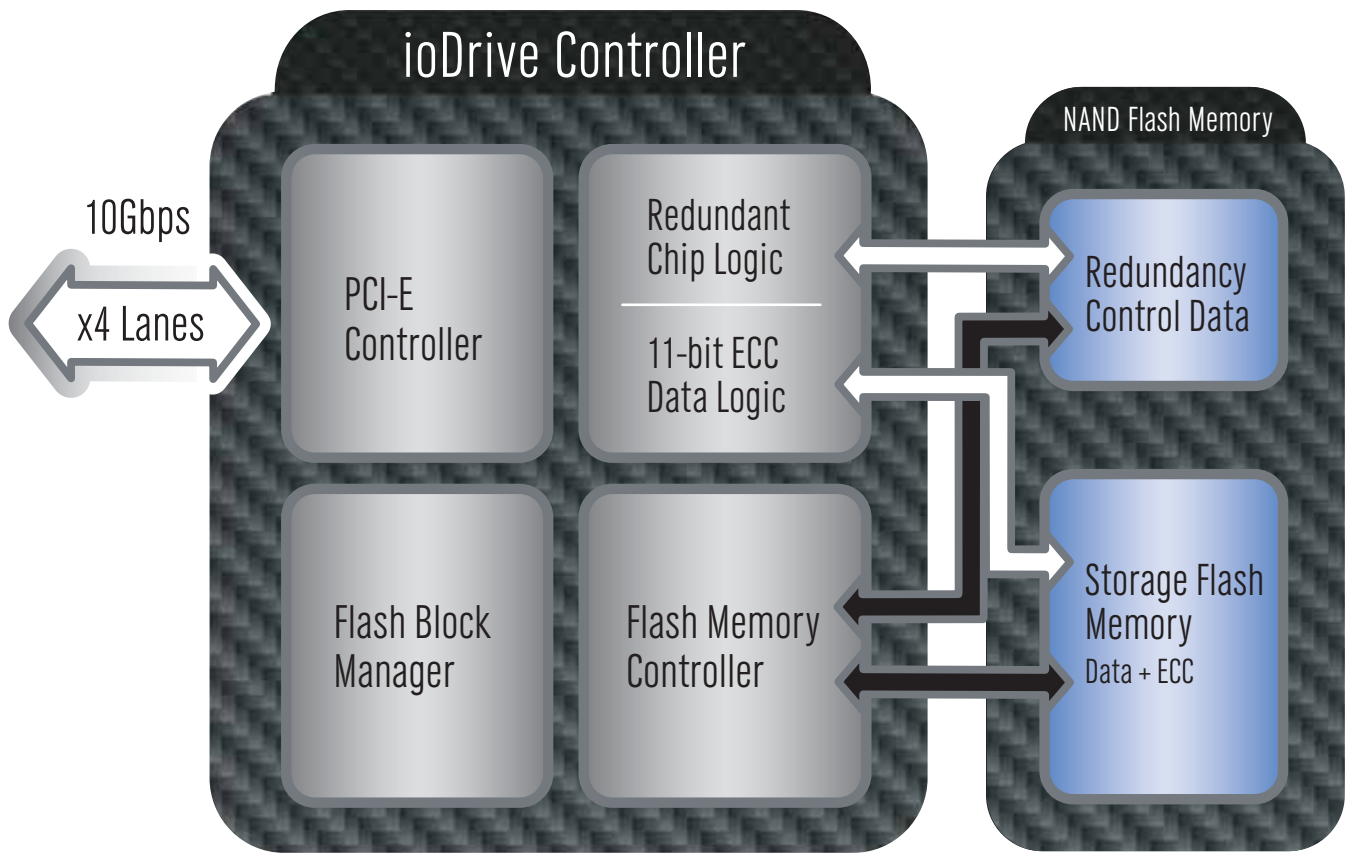
A Fusion-io device provides advanced warning prior to wear-out. Fusion-io supports today's monitoring management functions to measure and report on the device's status and usable life. In almost all cases, device upgrade is a smooth and predictable process, rather than an emergency situation.

Fusion-io protects your data at every stage of its path from your applications to the NAND flash storage medium, ensuring that nothing is lost or corrupted along the way or while the data is being stored. Data is checked multiple times, using several error detection methods. Once it reaches the storage medium, it is stored with robust error correction encoding that lets the flash device not only identify but correct bit errors. Fusion-io's data integrity design target is a 1 in 10^{30} probability of undetected bad data and a 1 in 10^{20} probability of uncorrectable data, as compared to a 1 in 10^{16} probability of undetected or uncorrectable errors for rotating magnetic storage devices.

Conclusion

Now with Fusion-io's comprehensive approach to data integrity, it is safe to exploit the exponential performance gains and many other benefits offered by NAND flash storage. The storage architecture pioneered by Fusion-io ensures predictable, controlled mitigation of early device failure, long-term device attrition and data changes due to external and data transport interference—issues that have up to now limited the adoption of NAND flash-based storage at the enterprise level. Fusion-io's NAND flash devices exceed the reliability of rotating magnetic media storage devices while providing an order of magnitude performance improvement.

Flash Back Block Diagram



Robert Brumfield
Fusion Public Relations
212.651.4215
robert.brumfield@fusionpr.com

Fusion-io Announces the ioDrive Duo—The World's Fastest and Most Innovative SSD

PCI Express, server-based solid-state storage offering sets a new standard for enterprise application-centric storage, with up to 640 gigabytes of capacity and 1.5 gigabytes per-second of sustained throughput

SALT LAKE CITY - March 11, 2009 - Fusion-io, the leader in solid-state architecture and high-performance I/O solutions, today announced the ioDrive Duo, which doubles the slot capacity of Fusion-io's successful PCI Express-based ioDrive storage solution. The new ioDrive Duo is the market's fastest and most innovative server-based solid-state storage solution.

With the ioDrive Duo, it is now possible for application, database and system administrators to get previously unheard-of levels of performance, protection and capacity utilization from a single server. Performance for multiple ioDrive Duos scales linearly, allowing any enterprise to scale performance to six gigabytes per-second (Gbytes/sec) of read bandwidth and over 500,000 read IOPS by using just four ioDrive Duos.

“Many database and system administrators are finding that SANs are too expensive and don't meet performance, protection and capacity utilization expectations,” said David Flynn, CTO of Fusion-io. “This is why more and more application vendors are moving toward application-centric solid-state storage. The ioDrive Duo offers the enterprise the advantages of application-centric storage without application-specific programming.”

ioDrive Duo Product Details

The following specifications describe the physical and performance characteristics of the ioDrive Duo.

Performance

Based on PCI Express x8 or PCI Express 2.0 x4 standards, which can sustain up to 20 gigabits per-second of raw throughput, the ioDrive Duo has more than enough bandwidth to obtain industry-leading performance from a single card. The ioDrive Duo can easily sustain 1.5 Gbytes/sec of read bandwidth and nearly 200,000 read IOPS. Its performance metrics are as follows:

- Sustained read bandwidth: 1500 MB/sec (32k packet size)
- Sustained write bandwidth: 1400 MB/sec (32k packet size)
- Read IOPS: 186,000 (4k packet size)
- Write IOPS: 167,000 (4k packet size)
- Latency < 50 µsec



Reliability

The ioDrive Duo offers unmatched solid-state protection for data integrity and reliability with triple redundancy for a single storage component.

- Multi-bit error detection and correction
- Patent-pending Flashback protection, offering chip-level N+1 redundancy and on-board self-healing so that no servicing is required
- Optional RAID-1 mirroring between two ioMemory modules on the same ioDrive Duo, offering complete redundancy on a single PCIe card

Capacity

The ioDrive Duo comes in the following capacities:

- 160 Gbytes
- 320 Gbytes
- 640 Gbytes
- 1.28 TB (second half of 2009)

The ioDrive Duo will be available in April 2009. To find out more about how this and Fusion-io's other enterprise solid-state storage products can benefit your organization, please visit www.fusionio.com.

About Fusion-io

Fusion-io is a leading provider of enterprise solid-state technology and high-performance I/O solutions. The company's solid-state storage technology closes the gap between processing power and storage needs delivering breakthrough performance at a fraction of the cost of traditional disk-based storage systems. The result is a world of possibilities for performance-starved applications.

ioDrive Duo



- > Sustain over a GB/sec of bandwidth
- > Easily RAID multiple ioDrive Duo's
- > OS support for Windows, Linux & Solaris

WWW.FUSIONIO.COM

ioDrive Duo Capacity	160GB	320GB	640GB
NAND Type	Single Level Cell (SLC)	Single Level Cell (SLC)	Multi Level Cell (MLC)
Write Bandwidth	1.1 GB/s (32k packet size)	1.4 GB/s (32k packet size)	1.0 GB/s (32k packet size)
Read Bandwidth	1.5 GB/s (32k packet size)	1.5 GB/s (32k packet size)	1.4 GB/s (32k packet size)
IOPS*	200,832 reads (4k packet size) 132,118 writes 4k packet size)	185,022 reads (4k packet size) 167,784 writes (4k packet size)	126,601 reads (4k packet size) 180,530 writes (4k packet size)
Access Latency	50µs Read	50µs Read	80µs Read
Bus Interface	PCI-Express x8 and PCI Express 2.0 x4	PCI-Express x8 and PCI Express 2.0 x4	PCI-Express x8 and PCI Express 2.0 x4
Weight	Less than 10 ounces	Less than 10 ounces	Less than 10 ounces
Operating Systems	Microsoft Windows**, Open Solaris 10 Solaris 10, RHEL 4 & 5; SLES 9 & 10	Microsoft Windows**, Open Solaris 10 Solaris 10, RHEL 4 & 5; SLES 9 & 10	Microsoft Windows**, Open Solaris 10 Solaris 10, RHEL 4 & 5; SLES 9 & 10
Wear Leveling and Sophisticated ECC (@ 5-TB write-erase / day)	24yrs	48yrs	16yrs

* Performance achieved using multiprocessor enterprise server ** 64-Bit Windows XP, Vista, Server 2003 & 2008

STANDARDS

Form Factor	Full height, 3/4 length PCI Express 2.0
Connectivity	PCI Express electromechanical spec 2.0
Power	PCI Express power spec 2.0

AGENCY

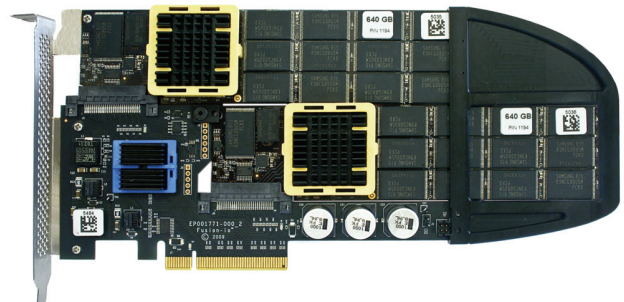
US / Canada	FCC Part 15, ICES-003, Class A
Europe	2004/108/EC EMC Directive CE Mark;
Japan	VCCI, Class A
Taiwan	BSMI, Class A
New Zealand /Australia	AS/NZS 3548 Class A
RoHS	R5 (Directive 2002/95/EC)

ENVIRONMENTAL SPECIFICATIONS

		Min	Max
Temperature (°C)*	Operational	0	55
	Non-operational	- 40	70
Air Flow (LFM)		300	
Humidity (%)	Non-condensing	5	95
	Operational		10,000
Altitude (ft)	Operational		30,000
	Non-operational		

* Temperature derated 1 C per 1000 ft elevation above sea level

100% Assembled in the U.S.A.



FUSION-io

©2009 Fusion-io, All Rights Reserved.



- > Less than 50 μ s latency
- > Easily RAID multiple ioDrives together
- > Managed like simple block storage

ioDrive Capacity	80GB	160GB	320GB
NAND Type	Single Level Cell (SLC)	Single Level Cell (SLC)	Multi Level Cell (MLC)
Write Bandwidth	550 MB/s (random 16K)	600 MB/s (random 16K)	500 MB/s (random 8K)
Read Bandwidth	700 MB/s (random 16K)	700 MB/s (random 16K)	700 MB/s (random 32K)
IOPS*	102,000 (random 4k reads) 91,000 (random 4k writes) 88,000 (70/30 random 4k mix)	104,400 (random 4k reads) 103,925 (random 4k writes) 95,000 (70/30 random 4k mix)	60,000 (random 4k reads) 79,000 (random 4k writes) 65,000 (70/30 random 4k mix)
Access Latency	50 μ s Read	50 μ s Read	80 μ s Read
Bus Interface	PCI-Express x4	PCI-Express x4	PCI-Express x4
Weight	Less than 2 ounces	Less than 2 ounces	Less than 2 ounces
Operating Systems	RHEL 4 & 5; SLES 9 & 10 Microsoft 64-Bit Windows**	RHEL 4 & 5; SLES 9 & 10 Microsoft 64-Bit Windows**	RHEL 4 & 5; SLES 9 & 10 Microsoft 64-Bit Windows**
Wear Leveling and Sophisticated ECC (@ 5-TB write-erase / day)	24yrs	48yrs	16yrs

* Performance data provided by Medusa Labs. ** 64-Bit Windows XP, Vista, Server 2003 & 2008

STANDARDS

Form Factor	Low profile PCI Express x4 slot (spec 1.1)
Connectivity	PCI Express x4 (electromechanical spec 1.1)
Power	PCI Express x4 (power spec 1.1)

ENVIRONMENTAL SPECIFICATIONS

		Min	Max
Temperature ($^{\circ}$ C)*	Operational	0	55
	Non-operational	- 40	70
Air Flow (LFM)		300	
Humidity (%)	Non-condensing	5	95
Altitude (ft)	Operational		10,000
	Non-operational		30,000

* Temperature derated 1 C per 1000 ft elevation above sea level

SAFETY

US / Canada	UL60950, CSA C22.2 No.60950-1-03
Europe	TUV EN60950-1:2001; 3N50825-1:

100% Assembled in the U.S.A.

AGENCY

US / Canada	FCC Part 15, ICES-003, Class A
Europe	2004/108/EC EMC Directive CE Mark;
Japan	VCCI, Class A
Taiwan	BSMI, Class A
New Zealand /Australia	AS/NZS 3548 Class A
RoHS	R5 (Directive 2002/95/EC)

