

## INSTALLATION DE MNOGOSEARCH SUR DEBIAN 4.0 (ETCH)



Sauf mention contraire, le contenu de ce document est sous [contrat Creative Commons](#).



I . Pré-requis .....	3
I . 1 . Configuration des sources apt.....	3
I . 2 . Vérification de la présence de tous les outils nécessaires.....	3
I . 3 . Création de la base de donnée pour mnogosearch.....	4
II . Installation de l'applicatif mnogosearch.....	5
II . 1 . Installation logicielle.....	5
II . 2 . Configuration de mnogosearch.....	11
II . 2 . 1 . Paramètres obligatoires.....	11
II . 2 . 2 . Exemple de configuration.....	11
II . 3 . Lancement de l'indexation.....	31
II . 4 . Automatisation de l'indexation.....	31
III . Installation de l'extension mnogosearch pour php .....	32
III . 1 . Récupération des sources de php.....	32
III . 2 . Récupération des sources de l'extension mnogosearch pour php .....	32
III . 3 . Modification des options de compilation du package debian php .....	32
III . 4 . Compilation de l'extension.....	46
III . 5 . Installation de la nouvelle version php supportant l'extension mnogosearch.....	46
III . 6 . Copie de l'extension dans le répertoire des modules php .....	46
III . 7 . Ajout de la directive de chargement automatique du module .....	46
III . 8 . Redémarrage d'apache .....	46

## I . Pré-requis

### I . 1 . Configuration des sources apt

Dans `/etc/apt/source.list` , vérifier que vous avez bien a minima les lignes suivantes :

```
deb http://ftp.fr.debian.org/debian/ etch main
deb-src http://ftp.fr.debian.org/debian/ etch main

deb http://security.debian.org/ etch/updates main contrib
deb-src http://security.debian.org/ etch/updates main contrib
```

### I . 2 . Vérification de la présence de tous les outils nécessaires

Le serveur doit avoir les dernières informations de packages à jour :

```
apt-get update
```

Le serveur doit avoir les dernières versions de logiciels à jour :

```
apt-get upgrade
```

Le serveur doit bien avoir les logiciels nécessaires installés :

```
apt-get install apache2 mysql-server php5 php5-mysql apache2-prefork-dev
```

Le serveur doit avoir tout ce qu'il faut pour construire l'extension mnogosearch pour php

```
apt-get build-dep php5
```

### ***1.3. Création de la base de donnée pour mnogosearch***

```
mysql -u root -p  
Enter password:  
mysql> create database mnogosearch;  
mysql> exit
```

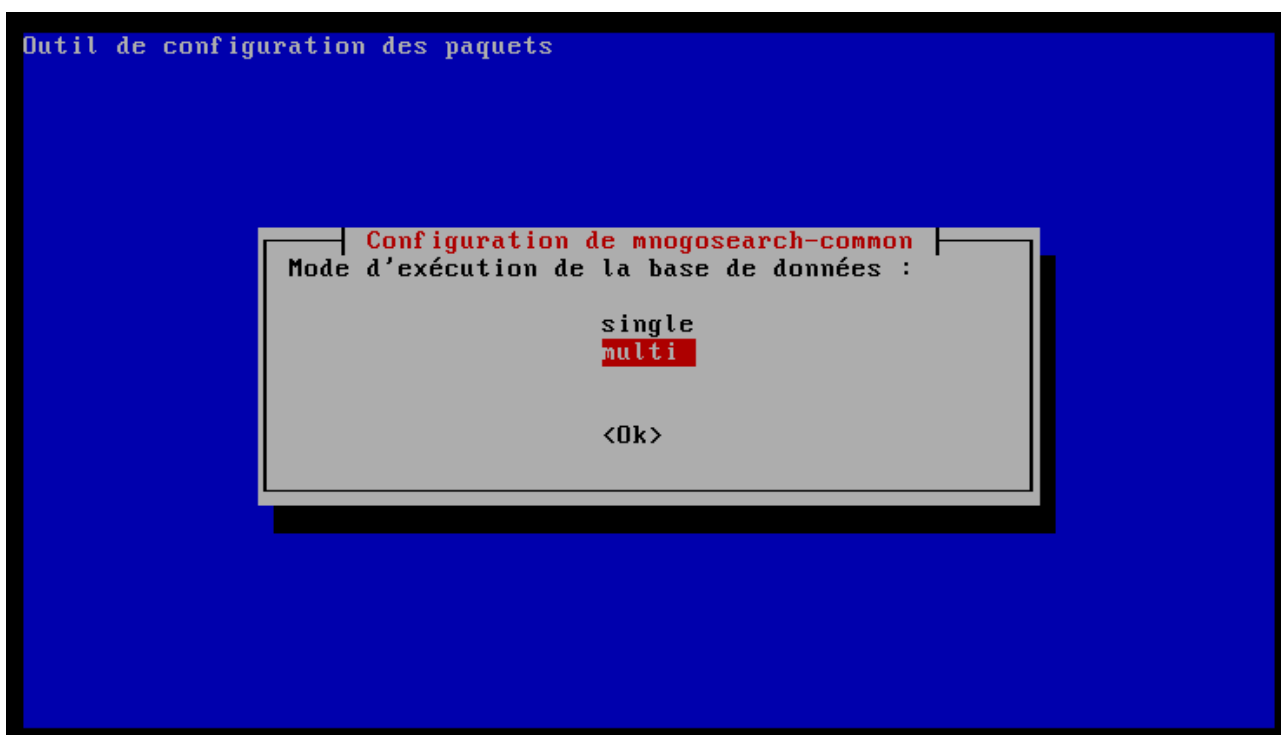
## II . Installation de l'applicatif mnogosearch

### II . 1 . Installation logicielle

Tout d'abord lancer la commande suivante :

```
apt-get install mnogosearch-mysql mnogosearch-common mnogosearch-dev
```

Comme mode d'exécution, dans un soucis de performance, nous allons choisir « multi » :



Nous préférons utiliser mnogosearch avec une base de donnée mysql :



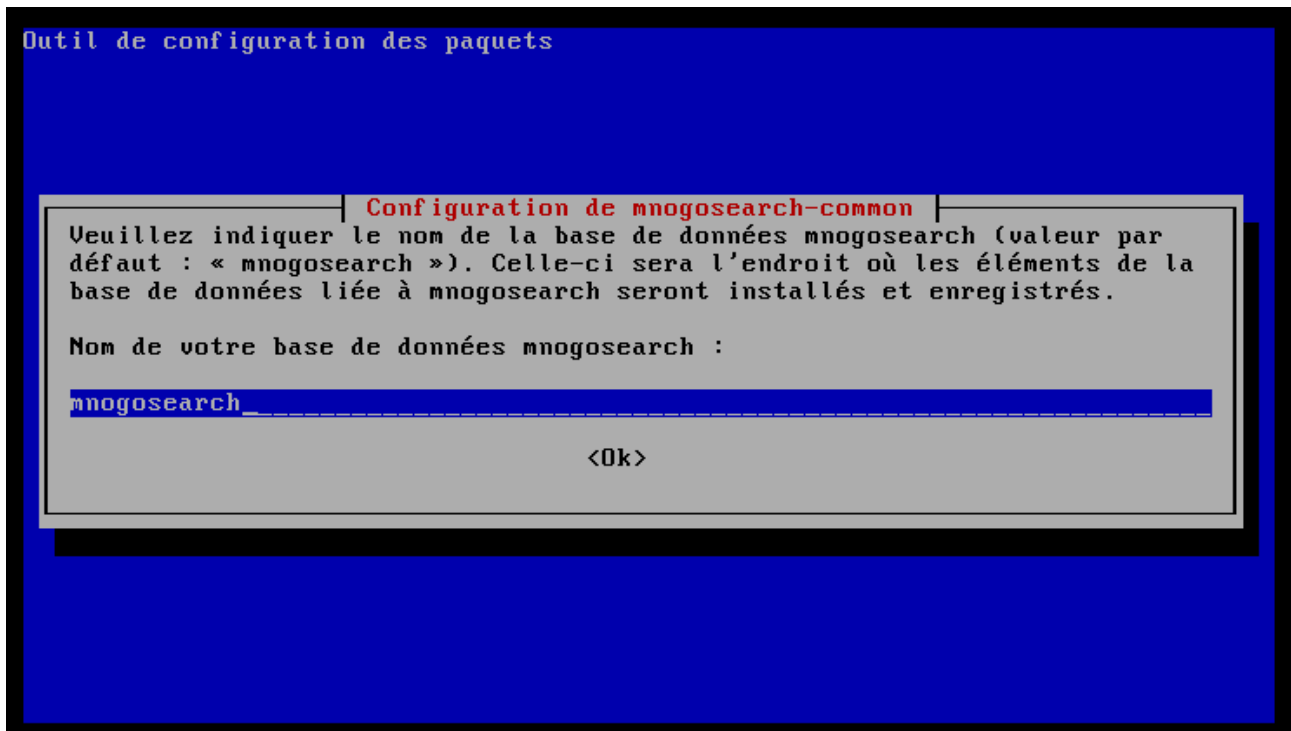
Nous indiquons le nom d'utilisateur qui sera créé pour l'accès à la base de donnée mnogosearch :



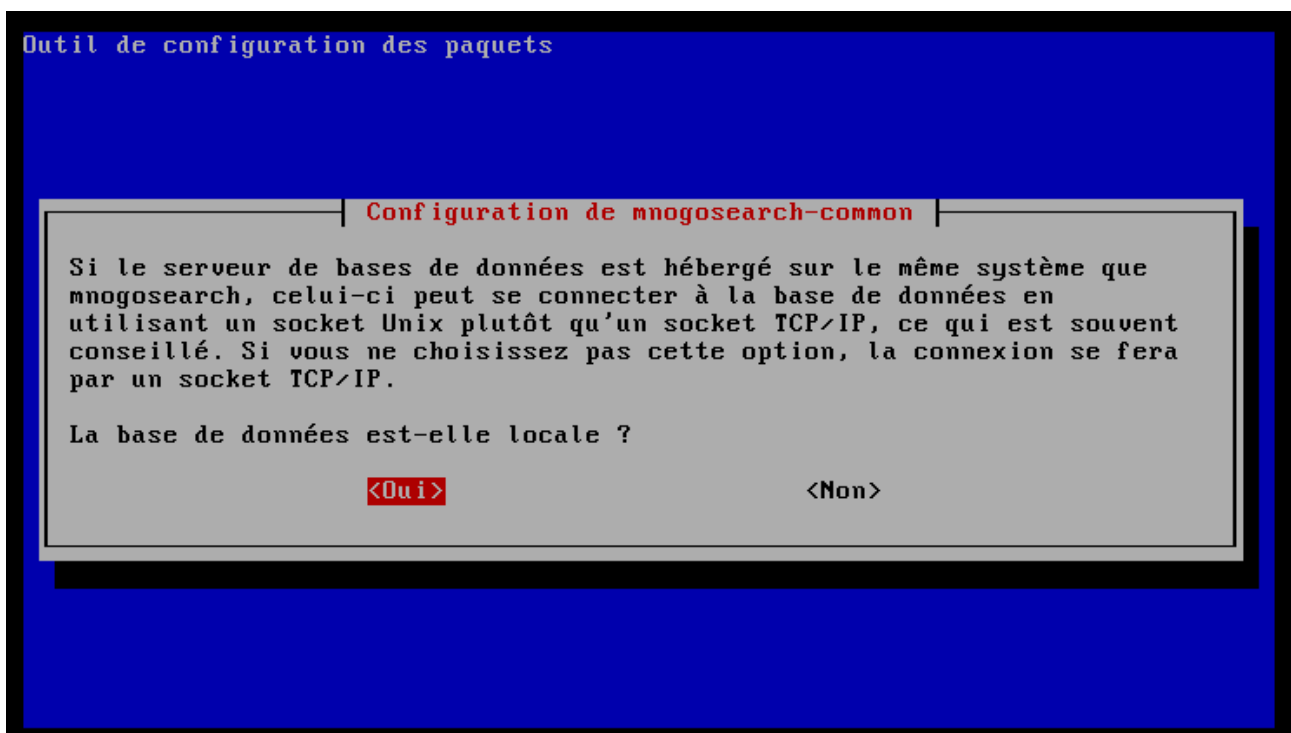
Nous indiquons le mot de passe à définir pour l'utilisateur qui sera créé pour l'accès à la base de donnée mnogosearch :



Nous indiquons le nom de la base de donnée précédemment créé :



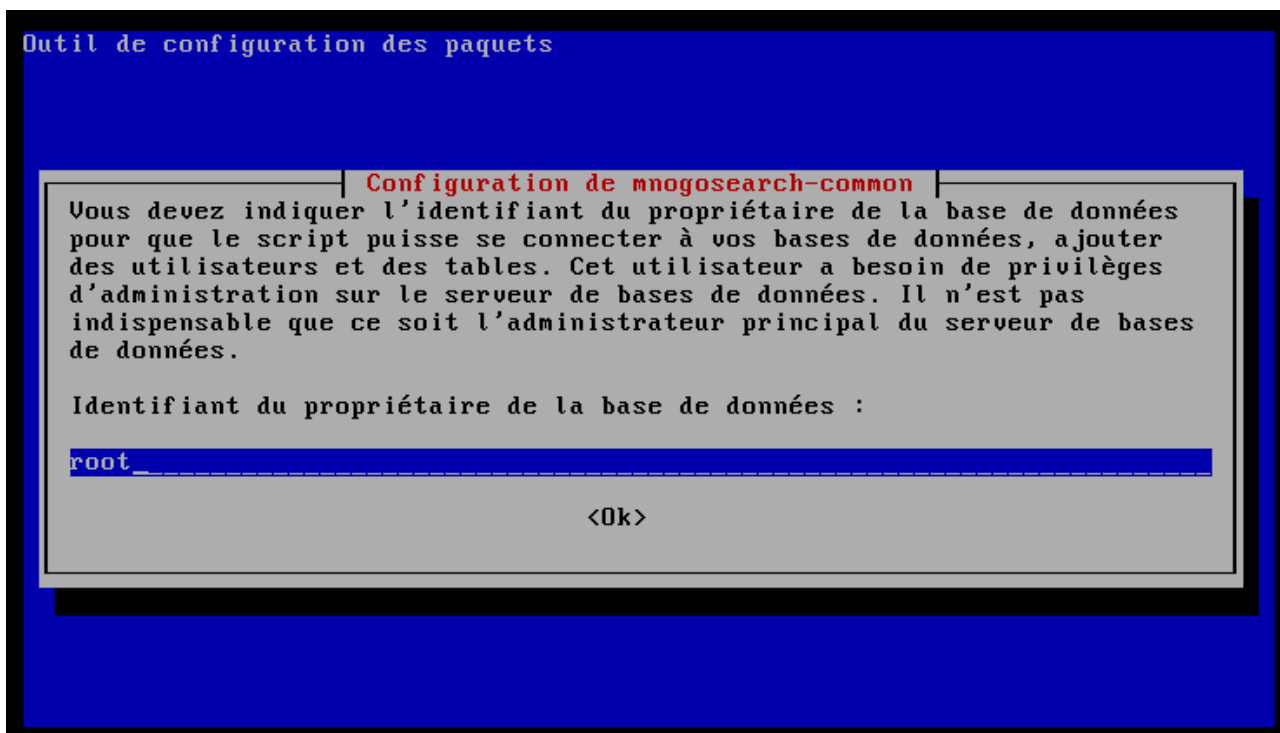
Normalement, la base de donnée tourne sur le même serveur que le serveur web :



Nous voulons que la base soit créée :



Le login administrateur de la base de donnée :



Le mot de passe de l'administrateur de la base de donnée :



Puis le serveur va créer toute la structure de la base pour aboutir à :

```
T NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFA (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFB (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFC (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFD (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFE (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>'CREATE TABLE dictFF (url_id int(11) DEFAULT '0' NOT NULL, secno tinyint unsigned DEFAULT '0' NOT NULL, word varchar(255) DEFAULT '' NOT NULL, intag blob NOT NULL, KEY url_id (url_id), KEY word_url (word))'
SQL>
269 queries sent, 269 succeeded, 0 failed

Paramétrage de mnogosearch-mysql (3.2.37-3.3) ...

debian:~# _
```

Le logiciel mnogosearch est maintenant installé.

## II . 2 . Configuration de mnogosearch

Toute la configuration se situe dans le fichier `/etc/mnogosearch/indexer.conf`

### II . 2 . 1 . Paramètres obligatoires

```
DBAddr mysql://login:password@localhost/mnogosearch/?dbmode=multi
Server http://monserveur.gouv.fr/
```

### II . 2 . 2 . Exemple de configuration

```
#!/usr/local/mnogosearch/sbin/indexer -d

#####
# This is a sample indexer config file.
# To start using it please edit and rename to indexer.conf.
# You can also make this file executable and run it directly.
# You may want to keep the original indexer.conf-dist for future references.
# Use '#' to comment out lines.
# All command names are case insensitive (DBAddr=DBADDR=dbaddr).
# You may use '\' character to prolong current command to next line
# when it is required.
#
# You may include another configuration file in any place of the indexer.conf
# using "Include <filename>" command.
# Absolute path if <filename> starts with "/":
#include /usr/local/mnogosearch/etc/incl.conf
# Relative path else:
#include incl.conf
#####

#####
# Section 1.
# Global parameters.

#####
# DBAddr <URL-style database description>
# Options (type, host, database name, port, user and password)
# to connect to SQL database.
# Should be used before any other commands.
```

```
# Has global effect for whole config file.
# Format:
#DBAddr <DBType>:[//[DBUser[:DBPass]@]DBHost[:DBPort]]/DBName/[?dbmode=mode]
#
# ODBC notes:
#     Use DBName to specify ODBC data source name (DSN)
#     DBHost does not matter, use "localhost".
#
# Currently supported DBType values are
# mysql, pgsq, mssql, oracle, ibase, db2, mimer, sqlite.
#
# MySQL users can specify path to Unix socket when connecting to localhost:
# mysql://foo:bar@localhost/mnogosearch/?socket=/tmp/mysql.sock
#
# If you are using PostgreSQL and do not specify hostname,
#     e.g. pgsq://user:password@/dbname/
# then PostgreSQL will not work via TCP, but will use Unix socket.
#
# You may also select database mode of word storage.
# When "single" is specified, all words are stored in the same table.
# If "multi" is selected, words will be located in different tables.
# "multi" mode is usually faster but requires more tables.
# Default mode is "single".
```

```
DBAddr mysql://mnogosearch:password@localhost/mnogosearch/?dbmode=multi
```

```
#####
```

```
# VarDir /usr/local/mnogosearch/var
# You may choose alternative working directory for
# search results cache:
#
#VarDir /usr/local/mnogosearch/var
```

```
#####
```

```
# NewsExtensions yes/no
# Whether to enable news extensions.
# Default value is no.
#NewsExtensions no
```

```
#####
```

```
#SyslogFacility <facility>
# This is used if indexer was compiled with syslog support and if you
# don't like the default value. Argument is the same as used in syslog.conf
# file. For list of possible facilities see syslog.conf(5)
#SyslogFacility local7
```

```
#####  
# LocalCharset <charset>  
# Defines the charset which will be used to store data in the database.  
# All other character sets will be converted into the given charset.  
# Take a look into mnoGoSearch documentation for detailed explanation  
# how to choose a LocalCharset depending on languages used on your site(s).  
# This command should be used once and takes global effect for the config file.  
# Only most popular charsets used in Internet are written here.  
# Take a look into the documentation to check the whole list of  
# supported charsets.  
# Default LocalCharset is iso-8859-1 (latin1).  
#  
# Western Europe: German, Finnish, French, Swedish  
#LocalCharset iso-8859-1  
#LocalCharset windows-1252  
  
# Central Europe: Czech, Slovenian, Slovak, Hungarian, Polish  
#LocalCharset iso-8859-2  
#LocalCharset windows-1250  
  
# Baltic: Lithuanian, Estonian, Latvian  
#LocalCharset iso-8859-4  
#LocalCharset iso-8859-13  
#LocalCharset windows-1257  
  
# Cyrillic: Russian, Serbian, Ukrainian, Belarussian, Macedonian, Bulgarian  
#LocalCharset koi8-r  
#LocalCharset iso-8859-5  
#LocalCharset x-mac-cyrillic  
#LocalCharset windows-1251  
  
# Arabic  
#LocalCharset iso-8859-6  
#LocalCharset windows-1256  
  
# Greek  
#LocalCharset iso-8859-7  
#LocalCharset windows-1253  
  
# Hebrew  
#LocalCharset iso-8859-8  
#LocalCharset windows-1255  
  
# Turkish  
#LocalCharset iso-8859-9  
#LocalCharset windows-1254  
  
# Vietnamese  
#LocalCharset viscii  
#LocalCharset windows-1258  
  
# Chinese  
#LocalCharset gb2312
```

```
#LocalCharset BIG5

# Korean
#LocalCharset EUC-KR

# Japanese
#LocalCharset Shift-JIS

# Full UNICODE
LocalCharset UTF-8

#####
#ForceIISCharset1251 yes/no
#This option is useful for users which deals with Cyrillic content and broken
#(or misconfigured?) Microsoft IIS web servers, which tends to not report
#charset correctly. This is really dirty hack, but if this option is turned on
#it is assumed that all servers which reports as 'Microsoft' or 'IIS' have
#content in Windows-1251 charset.
#This command should be used only once in configuration file and takes global
#effect.
#Default: no
#ForceIISCharset1251 no

#####
#CrossWords yes/no
# Whether to build CrossWords index
# Default value is no
#CrossWords no

#####
# StopwordFile <filename>
# Load stop words from the given text file. You may specify either absolute
# file name or a name relative to mnoGoSearch /etc directory. You may use
# several StopwordFile commands.
#
#StopwordFile stopwords/en.sl

Include stopwords.conf

#####
# LangMapFile <filename>
# Load language map for charset and language guesser from the given file.
# You may specify either an absolute file name or a name relative
# to mnoGoSearch /etc directory. You may use several LangMapFile commands.
#
#LangMapFile langmap/en.ascii.lm

Include langmap.conf
```

```
#####  
# Word lengths. You may change default length range of words  
# stored in the database. By default, words with the length in the  
# range from 1 to 32 are stored.  
#  
#MinWordLength 1  
#MaxWordLength 32  
  
#####  
# MaxDocSize bytes  
# Default value 1048576 (1 Mb)  
# Takes global effect for whole config file  
MaxDocSize 30485760  
  
#####  
# URLSelectCacheSize num  
# Default value 128  
# Select <num> targets to index at once.  
#URLSelectCacheSize 1024  
  
#####  
# WordCacheSize bytes  
# Default value 8388608 (8 Mb)  
# Defines maximal in-memory words cache size.  
# Note: cache is allocated for every DBAddr, so if you have 3 DBAddr  
# commands and WordCacheSize is 10Mb, it can take up to 30Mb of memory.  
#WordCacheSize 8388608  
  
#####  
# HTTPHeader <header>  
# You may add your desired headers in indexer HTTP request.  
# You should not use "If-Modified-Since", "Accept-Charset" headers,  
# these headers are composed by indexer itself.  
# "User-Agent: mnoGoSearch/version" is sent too, but you may override it.  
# Command has global effect for all configuration file.  
#  
#HTTPHeader "User-Agent: My_Own_Agent"  
#HTTPHeader "Accept-Language: ru, en"  
#HTTPHeader "From: webmaster@mysite.com"  
  
#####  
# UseCookie yes|no  
# Whether to accept per-session cookies.  
# The default value is no.  
#UseCookie no  
  
# flush server.active to inactive for all server table records  
# before loading new  
#FlushServerTable  
  
#####
```

```
# ServerTable <table_addr>
# Load servers with all their parameters from the table specified in argument.
# Check an example of tables server and srvinfo structure in
# create/(your_database)/create.txt
#
#ServerTable mysql://user:pass@host/dbname/tablename

#####
# LoadChineseList <charset> <dictionaryfilename>
# Load Chinese word frequency list.
# By default GB2312 charset and mandarin.freq dictionary is used.
#LoadChineseList

#####
# LoadThaiList <charset> <dictionaryfilename>
# Load Thai word frequency list
# By default tis-620 and thai.freq dictionary is used.
#LoadThaiList

#####
# Section 2.
# URL control configuration.

#####
#Allow [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ... ]
# Use this to allow URLs that match (doesn't match) the given argument.
# First three optional parameters describe the type of comparison.
# Default values are Match, NoCase, String.
# Use "NoCase" or "Case" values to choose case insensitive or case sensitive
# comparison.
# Use "Regex" to choose regular expression comparison.
# Use "String" to choose string with wildcards comparison.
# Wildcards are '*' for any number of characters and '?' for one character.
# Note that '?' and '*' have special meaning in "String" match type. Please use
# "Regex" to describe documents with '?' and '*' signs in URL.
# "String" match is much faster than "Regex". Use "String" where it
# is possible.
# You may use several arguments for one 'Allow' command.
# You may use this command any times.
# Takes global effect for config file.
# Note that mnoGoSearch automatically adds one "Allow regex .*"
# command after reading config file. It means that allowed everything
# that is not disallowed.
# Examples
# Allow everything:
#Allow *
# Allow everything but .php .cgi .pl extensions case insensitively using
# regex:
#Allow NoMatch Regex /\.php$|\.cgi$|\.pl$
# Allow .HTM extension case sensitively:
#Allow Case *.HTM
```

```
#####
#Disallow [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ... ]
# Use this to disallow URLs that match (doesn't match) given argument.
# The meaning of first three optional parameters is exactly the same
# with "Allow" command.
# You can use several arguments for one 'Disallow' command.
# Takes global effect for config file.
#
# Examples:
# Disallow URLs that are not in udm.net domains using "string" match:
#Disallow NoMatch *.udm.net/*
# Disallow any except known extensions and directory index using "regex" match:
#Disallow NoMatch Regex \/$|\.htm$|\.html$|\.shtml$|\.phtml$|\.php$|\.txt$
# Exclude cgi-bin and non-parsed-headers using "string" match:
#Disallow */cgi-bin/* *.cgi */nph-*
# Exclude anything with '?' sign in URL. Note that '?' sign has a
# special meaning in "string" match, so we have to use "regex" match here:
#Disallow Regex \?

# Exclude some known extensions using fast "String" match:
Disallow *.b *.sh *.md5 *.rpm
Disallow *.arj *.tar *.zip *.tgz *.gz *.z *.bz2
Disallow *.lha *.lzh *.rar *.zoo *.ha *.tar.Z
Disallow *.gif *.jpg *.jpeg *.bmp *.tiff *.tif *.xpm *.xbm *.pcx
Disallow *.vdo *.mpeg *.mpe *.mpg *.avi *.movie *.mov *.wmv
Disallow *.mid *.mp3 *.rm *.ram *.wav *.aiff *.ra
Disallow *.vrml *.wrl *.png *.ico *.psd *.dat
Disallow *.exe *.com *.cab *.dll *.bin *.class *.ex_
#Disallow *.tex *.texi *.xls *.doc *.texinfo
#Disallow *.rtf *.pdf *.cdf *.ps
Disallow *.ai *.eps *.ppt *.hqx
Disallow *.cpt *.bms *.oda *.tcl
Disallow *.o *.a *.la *.so
Disallow *.pat *.pm *.m4 *.am *.css
Disallow *.map *.aif *.sit *.sea
Disallow *.m3u *.qt

# Exclude Apache directory list in different sort order using "string" match:
Disallow *D=A *D=D *M=A *M=D *N=A *N=D *S=A *S=D

# More complicated case. RAR .r00-.r99, ARJ a00-a99 files
# and UNIX shared libraries. We use "Regex" match type here:
Disallow Regex \.r[0-9][0-9]$ \.a[0-9][0-9]$ \.so\.[0-9]$

#####
#CheckOnly [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ... ]
# The meaning of first three optional parameters is exactly the same
# with "Allow" command.
# Indexer will use HEAD instead of GET HTTP method for URLs that
```

```
# match/do not match given regular expressions. It means that the file
# will be checked only for being existing and will not be downloaded.
# Useful for zip,exe,arj and other binary files.
# Note that you can disallow those files with commands given below.
# You may use several arguments for one "CheckOnly" commands.
# Useful for example for searching through the URL names rather than
# the contents (a la FTP-search).
# Takes global effect for config file.
#
# Check some known non-text extensions using "string" match:
#CheckOnly *.b      *.sh      *.md5
#CheckOnly *.arj    *.tar    *.zip    *.tgz    *.gz
#CheckOnly *.lha    *.lzh    *.rar    *.zoo    *.tar*.Z
#CheckOnly *.gif    *.jpg    *.jpeg   *.bmp    *.tiff
#CheckOnly *.vdo    *.mpeg   *.mpe    *.mpg    *.avi    *.movie
#CheckOnly *.mid    *.mp3    *.rm     *.ram    *.wav    *.aiff
#CheckOnly *.vrml   *.wrl    *.png
#CheckOnly *.exe    *.cab    *.dll    *.bin    *.class
#CheckOnly *.tex    *.texi   *.xls    *.doc    *.texinfo
#CheckOnly *.rtf    *.pdf    *.cdf    *.ps
#CheckOnly *.ai     *.eps    *.ppt    *.hqx
#CheckOnly *.cpt    *.bms    *.oda    *.tcl
#CheckOnly *.rpm    *.m3u    *.qt     *.mov
#CheckOnly *.map    *.aif    *.sit    *.sea
#
# or check ANY except known text extensions using "regex" match:
#CheckOnly NoMatch Regex \/$|\.html$|\.shtml$|\.phtml$|\.php$|\.txt$

#####
#HrefOnly [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ... ]
# The meaning of first three optional parameters is exactly the same
# with "Allow" command.
#
# Use this to scan a HTML page for "href" tags but not to index the contents
# of the page with an URLs that match (doesn't match) given argument.
# Commands have global effect for all configuration file.
#
# When indexing large mail list archives for example, the index and thread
# index pages (like mail.10.html, thread.21.html, etc.) should be scanned
# for links but shouldn't be indexed:
#
#HrefOnly */mail*.html */thread*.html

#####
#CheckMp3 [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ...]
# The meaning of first three optional parameters is exactly the same
# with "Allow" command.
# If an URL matches given rules, indexer will download only a little part
# of the document and try to find MP3 tags in it. On success, indexer
# will parse MP3 tags, else it will download whole document then parse
# it as usual.
```

```
# Notes:
# This works only with those servers which support HTTP/1.1 protocol.
# It is used "Range: bytes" header to download mp3 tag.
#CheckMp3 *.bin *.mp3

#####
#CheckMP3Only [Match|NoMatch] [NoCase|Case] [String|Regex] <arg> [<arg> ...]
# The meaning of first three optional parameters is exactly the same
# with "Allow" command.
# If an URL matches given rules, indexer, like in the case CheckMP3 command,
# will download only a little part of the document and try to find MP3 tags.
# On success, indexer will parse MP3 tags, else it will NOT download whole
# document.
#CheckMP3Only *.bin *.mp3

# How to combine Allow, Disallow, CheckOnly, HrefOnly commands.
#
# indexer compares URLs against all these command arguments in the
# order of their appearance in indexer.conf file.
# If indexer finds that URL matches some rule, it will make a decision of what
# to do with this URL, allow it, disallow it or use HEAD instead
# of the GET method. So, you may use different Allow, Disallow,
# CheckOnly, HrefOnly commands order.
# If no one of these commands are given, mnoGoSearch will allow everything
# by default.
#
# There are many possible combinations. Samples of two of them are here:
#
# Sample of first useful combination.
# Disallow known non-text extensions (zip,wav etc),
# then allow everything else. This sample is uncommented above (note that
# there is actually no "Allow *" command, it is added automatically after
# indexer.conf loading).
#
# Sample of second combination.
# Allow some known text extensions (html, txt) and directory index ( / ),
# then disallow everything else:
#
#Allow .html .txt */
#Disallow *

# HoldBadHrefs <time>
# Default value is 0.
# How much time to hold URLs with erroneous status before deleting them
# from the database. For example, if host is down, indexer will not delete
# pages from this site immediately and search will use previous content
# of these pages.. However if site doesn't respond for a month, probably
# it's time to remove these pages from the database.
# For <time> format see description of Period command.
#HoldBadHrefs 30d
```

```
#####
# Section 3.
# Mime types and external parsers.

#####
#UseRemoteContentType yes/no
# Default value is yes.
# This command specifies if the indexer should get content type
# from HTTP server headers (yes) or from it's AddType settings (no).
# If set to 'no' and the indexer could not determine content-type
# by using its AddType settings, then it will use HTTP header.
# Default: yes
#UseRemoteContentType yes

#####
#AddType [String|Regex] [Case|NoCase] <mime type> <arg> [<arg>...]
# This command associates filename extensions (for services
# that don't automatically include them) with their mime types.
# Currently "file:" protocol uses these commands.
# Use optional first two parameter to choose comparison type.
# Default type is "String" "NoCase" (case sensitive string match with
# '?' and '*' wildcards for one and several characters correspondingly).
#
AddType image/x-xpixmap *.xpm
AddType image/x-xbitmap *.xbm
AddType image/gif *.gif

AddType text/plain *.txt *.pl *.js *.h *.c *.pm *.e
AddType text/html *.html *.htm

AddType text/rtf *.rtf
AddType application/pdf *.pdf
AddType application/msword *.doc
AddType application/vnd.ms-excel *.xls
AddType text/x-postscript *.ps

# You may also use quotes in mime type definition
# for example to specify charset. e.g. Russian webmasters
# often use *.htm extension for windows-1251 documents and
# *.html for UNIX koi8-r documents:
#
#AddType "text/html; charset=koi8-r" *.html
#AddType "text/html; charset=windows-1251" *.htm
#
# More complicated example for rar .r00-r.99 using "Regex" match:
#AddType Regex application/rar \.r[0-9][0-9]$
#
# Default unknown type for other extensions:
AddType application/unknown *.*
```

```
# Mime <from_mime> <to_mime> <command line>
#
# This is used to add support for parsing documents with mime types other
# than text/plain and text/html. It can be done via external parser (which
# must provide output in plain or html text) or just by substituting mime
# type so indexer will understand it.
#
# <from_mime> and <to_mime> are standard mime types
# <to_mime> is either text/plain or text/html
#
# Optional charset parameter used to change charset if needed
#
# Command line may have $1 parameter which stands for temporary file name.
# Some parsers can not operate on stdin, so indexer creates temporary file
# for parser and it's name passed instead of $1. Take a look into documentation
# for other parser types and parsers usage explanation.
# Examples:
#
#      from_mime          to_mime[charset]          [command line]
#      [ $1 ]
#
Mime application/msword      "text/plain; charset=utf-8" "catdoc -a -dutf-8 $1"
#Mime application/x-troff-man text/plain          "deroff"
#Mime text/x-postscript      text/plain          "ps2ascii"
Mime application/pdf         text/plain          "pdftotext $1 -"
#Mime application/vnd.ms-excel text/plain          "xls2csv $1"
#Mime "text/rtf*"            text/html           "rthc --use-stdout $1
2>/dev/null"

# Use ParserTimeout to specify amount of time for parser execution
# to avoid possible indexer hang.
ParserTimeout 300

#####
# Section 4.
# Aliases configuration.

#####
#Alias <master> <mirror>
# You can use this command for example to organize search through
# master site by indexing a mirror site. It is also useful to
# index your site from local file system.
# mnoGoSearch will display URLs from <master> while searching
# but go to the <mirror> while indexing.
# This command has global indexer.conf file effect.
# You may use several aliases in one indexer.conf.
#Alias http://www.mysql.com/ http://mysql.udm.net/
#Alias http://www.site.com/ file:///usr/local/apache/htdocs/
```

```
#####
#AliasProg <command line>
# AliasProg is an external program that can be called, that takes a URL,
# and returns the appropriate alias to stdout. Use $1 to pass a URL. This
# command has global effect for whole indexer.conf.
# Example:
#AliasProg "echo $1 | /usr/local/mysql/bin/replace http://localhost/
file:///home/httpd/"

#####
# Section 5.
# Servers configuration.

#####
#Period <time>
# Default value is 7d (i.e. 7days).
# Set reindex period.
# <time> is in the form 'xxxA[yyyB[zzzC]]'
# (Spaces are allowed between xxx and A and yyy and so on)
#   there xxx, yyy, zzz are numbers (can be negative!)
#       A, B, C can be one of the following:
#           s - second
#           M - minute
#           h - hour
#           d - day
#           m - month
#           y - year
#       (these letters are the same as in strftime/strptime functions)
#
# Examples:
# 15s - 15 seconds
# 4h30M - 4 hours and 30 minutes
# 1y6m-15d - 1 year and six month minus 15 days
# 1h-10M+1s - 1 hour minus 10 minutes plus 1 second
#
# If you specify only number without any character, it is assumed
# that time is given in seconds.
#
# Can be set many times before "Server" command and
# takes effect till the end of config file or till next Period command.
#Period 7d

#####
#Tag <string>
# Use this field for your own purposes. For example for grouping
# some servers into one group, etc... During search you'll be able to
# limit URLs to be searched through by their tags.
# Can be set multiple times before "Server" command and
# takes effect till the end of config file or till next Tag command.
```

```
# Default values is an empty sting

#####
#Category <string>
#You may distribute documents between nested categories. Category
#is a string in hex number notation. You may have up to 6 levels with
#256 members per level. Empty category means the root of category tree.
#Take a look into doc/categories.xml for more information.
#This command means a category on first level:
#Category AA
#This command means a category on 5th level:
#Category FFAABBCCDD

#####
#DefaultLang <string>
#Default language for server. Can be used if you need language
#restriction while doing search.
#Default value is empty.
#DefaultLang en

#####
#VaryLang <string>
#Default value is empty, i.e. don't vary language.
#Specify languages list for multilingual servers indexing
#VaryLang "ru en fr de"

#####
#MaxHops <number>
# Maximum way in "mouse clicks" from start url.
# Default value is 256.
# Can be set multiple times before "Server" command and
# takes effect till the end of config file or till next MaxHops command.
#MaxHops 256

#####
#MaxNetErrors <number>
# Maximum network errors for each server.
# Default value is 16. Use 0 for unlimited errors number.
# If there too many network errors on some server
# (server is down, host unreachable, etc) indexer will try to do
# not more then 'number' attempts to connect to this server.
# Takes effect till the end of config file or till next MaxNetErrors command.
#MaxNetErrors 16

#####
#ReadTimeOut <time>
# Connect timeout and stalled connections timeout.
```

```
# For <time> format see description of Period above.
# Default value is 30 seconds.
# Can be set any times before "Server" command and
# takes effect till the end of config file or till next ReadTimeOut command.
#ReadTimeOut 30s

#####
#DocTimeOut <time>
# Maximum amount of time indexer spends for one document downloading.
# For <time> format see description of Period above.
# Default value is 90 seconds.
# Can be set any times before "Server" command and
# takes effect till the end of config file or till next DocTimeOut command.
#DocTimeOut 1m30s

#####
#NetErrorDelayTime <time>
# Specify document processing delay time if network error has occurred.
# For <time> format see description of Period above.
# Default value is one day
#NetErrorDelayTime 1d

#####
#Robots yes/no
# Allows/disallows using robots.txt and <META NAME="robots">
# exclusions. Use "no", for example for link validation of your server(s).
# Command may be used several times before "Server" command and
# takes effect till the end of config file or till next Robots command.
# Default value is "yes".
#Robots yes

#####
#DetectClones yes/no
# Allow/disallow clone detection and eliminating. If allowed, indexer will
# detect the same documents under different location, such as
# mirrors, and will index only one document from the group of
# such equal documents. "DetectClones yes" also allows to reduce space usage.
# Default value is "yes".
#DetectClones yes

#####
# Document sections.
#
# Format is:
#
# Section <string> <number> <maxlen> [clone] [sep] [{expr} {repl}]
#
# where <string> is a section name and <number> is section ID
```

```
# between 0 and 255. Use 0 if you don't want to index some of
# these sections. It is better to use different sections IDs
# for different documents parts. In this case during search
# time you'll be able to give different weight to each part
# or even disallow some sections at a search time.
# <maxlen> argument contains a maximum length of section
# which will be stored in database.
# "clone" is an optional parameter describing whether this
# section should affect clone detection. It can
# be "DetectClone" or "cdon", or "NoDetectClone" or "cdoff".
# By default, url.* section values are not taken in account
# for clone detection, while any other sections take part
# in clone detection.
# "sep" is an optional argument to specify a separator between
# parts of the same section. It is a space character by default.
# "expr" and "repl" can be used to extract user defined sections,
# for example pieces of text between the given tags. "expr" is
# a regular expression, "repl" is a replacement with $1, $2, etc
# meta-characters designating matches "expr" matches.

# Standard HTML sections: body, title

Section body                1          256
Section title              2          128

# META tags
# For example <META NAME="KEYWORDS" CONTENT="xxxx">
#

Section meta.keywords      3          128
Section meta.description   4          128

# HTTP headers example, let's store "Server" HTTP header
#
#
#Section header.server     5          64

# Document's URL parts

Section url.file           6          0
Section url.path           7          0
Section url.host           8          0
Section url.proto          9          0

# CrossWords

Section crosswords         10         0

#
# If you use CachedCopy for smart excerpts (see below),
# please keep Charset section active.
#
```

```
Section Charset                11      32
Section Content-Type          12      64
Section Content-Language      13      16

# Uncomment the following lines if you want tag attributes
# to be indexed

#Section attribute.alt        14      128
#Section attribute.label     15      128
#Section attribute.summary   16      128
#Section attribute.title     17      128
#Section attribute.face      27       0

# Uncomment the following lines if you want use NewsExtensions
# You may add any Newsgroups header to be indexed and stored in urlinfo table

#Section References          18       0
#Section Message-ID        19       0
#Section Parent-ID         20       0

# Uncomment the following lines if you want index MP3 tags.
#Section MP3.Song           21      128
#Section MP3.Album          22      128
#Section MP3.Artist         23      128
#Section MP3.Year           24      128

# Comment this line out if you don't want to store "cached copies"
# to generate smart excerpts at search time.
# Don't forget to keep "Charset" section active if you use cached copies.
# NOTE: 3.2.18 has limits for CachedCopy size, 32000 for Ibase and
# 15000 for Mimer. Other databases do not have limits.
# If indexer fails with 'string too long' error message then reduce
# this number. This will be fixed in the future versions.
#
Section CachedCopy           25 64000

# A user defined section example.
# Extract text between <h1> and </h1> tags:
#Section h1                  26 128 "<h1>(.*</h1>" $1

#####
#<IndexIf|NoIndexIf> [Match|NoMatch] [NoCase|Case] [String|Regex] <Section>
<arg> [<arg> ... ]
# Use this to allow documents, which sections match given argument,
# to be indexed (IndexIf) or not indexed (NoIndexIf).
# First three optional parameters describe the type of comparison.
# Default values are Match, NoCase, String.
# See also Allow/Disallow.
# You may use several arguments for one 'IndexIf/NoIndexIf' commands.
# You may use this command any times.
# Takes global effect for config file.
# Example1: Don't index a document if its Body section contains "porno"
```

```
#NoIndexIf Body *porno*
#
# Example2: Index only those documents with Title section containing
"reference"
#IndexIf Title *reference*
#NoIndexIf Title *

#####
#Index yes/no
# Prevent indexer from storing words into database.
# Useful for example for link validation.
# Can be set multiple times before "Server" command and
# takes effect till the end of config file or till next Index command.
# Default value is "yes".
#Index yes

#####
#RemoteCharset <charset>
#<RemoteCharset> is default character set for the server in next "Server"
# command(s).
# This is required only for "bad" servers that do not send information
# about charset in header: "Content-type: text/html; charset=some_charset"
# and do not have <META NAME="Content" Content="text/html;
charset="some_charset">
# Can be set before every "Server" command and
# takes effect till the end of config file or till next RemoteCharset command.
# Default value is iso-8859-1 (latin1).
#RemoteCharset iso-8859-1

#####
#ProxyAuthBasic login:passwd
# Use HTTP proxy basic authorization
# Can be used before every "Server" command and
# takes effect only for next one "Server" command!
# It should be also before "Proxy" command.
# Examples:
#ProxyAuthBasic somebody:something

#####
#Proxy your.proxy.host[:port]
# Use proxy rather than connect directly
#One can index ftp servers when using proxy
#Default port value if not specified is 3128 (Squid)
#If proxy host is not specified direct connect will be used.
#Can be set before every "Server" command and
# takes effect till the end of config file or till next Proxy command.
#If no one "Proxy" command specified indexer will use direct connect.
#
# Examples:
# Proxy on atoll.anywhere.com, port 3128:
```

```
#Proxy atoll.anywhere.com
#
# Proxy on lota.anywhere.com, port 8090:
#Proxy lota.anywhere.com:8090
#
# Disable proxy (direct connect):
#Proxy

#####
#AuthBasic login:passwd
# Use basic HTTP authorization
# Can be set before every "Server" command and
# takes effect only for next one Server command!
# Examples:
#AuthBasic somebody:something
#
# If you have password protected directory(ies), but whole server is open,use:
#AuthBasic login1:passwd1
#Server http://my.server.com/my/secure/directory1/
#AuthBasic login2:passwd2
#Server http://my.server.com/my/secure/directory2/
#Server http://my.server.com/

#####
# Mirroring parameters commands.
#
# You may specify a path to root directory to enable sites mirroring
#MirrorRoot /path/to/mirror
#
# You may specify as well root directory of mirrored document's headers
# indexer will store HTTP headers to local disk too.
#MirrorHeadersRoot /path/to/headers
#
# MirrorPeriod <time>
# You may specify period during which earlier mirrored files
# will be used while indexing instead of real downloading.
# It is very useful when you do some experiments with mnoGoSearch
# indexing the same hosts and do not want much traffic from/to Internet.
# If MirrorHeadersRoot is not specified and headers are not stored
# to local disk then default Content-Type's given in AddType commands
# will be used.
# Default value of the MirrorPeriod is -1, which means
# "do not use mirrored files".
#
# For <time> format see Period command description above.
#
# The command below will force using local copies for one day:
#MirrorPeriod 1d

#####
```

```
#ServerWeight <number>
# Server weight for Popularity Rank calculation.
# Default value is 1.
#ServerWeight 1

#####
#PopRankSkipSameSite yes|no
# Skip links from same site for Popularity Rank calculation.
# Default value is "no".
#PopRankSkipSameSite yes

#####
#PopRankFeedBack yes|no
# Calculate sites wights before Popularity Rank calculation.
# Default value is "no".
#PopRankSkipSameSite yes

#####
#Server [Method] [SubSection] <URL> [alias]
# This is the main command of the indexer.conf file. It's used
# to describe web-space you want to index. It also inserts
# given URL into database to use it as a start point.
# You may use "Server" command as many times as a number of different
# servers or their parts you want to index.
#
# "Method" is an optional parameter which can take on of the following values:
# Allow, Disallow, CheckOnly, HrefOnly, CheckMP3, CheckMP3Only, Skip.
#
# "SubSection" is an optional parameter to specify server's subsection,
# i.e. a part of Server command argument.
# It can take the following values:
# "page" describes web space which consists of one page with address <URL>.
# "path" describes all documents which are under the same path with <URL>.
# "site" describes all documents from the same host with <URL>.
# "world" means "any document".
# Default value is "path".
#
# To index whole server "localhost":
#Server http://localhost/
#
# You can also specify some path to index subdirectory only:
#Server http://localhost/subdir/
#
# To specify the only one page:
#Server page http://localhost/path/main.html
#
# To index whole server but giving non-root page as a start point:
#Server site http://localhost/path/main.html
#
#
```

```
# You can also specify optional parameter "alias". This example will
# index server "http://www.mnogosearch.org/" directly from disk instead of
# fetching from HTTP server:
#Server http://www.mnogosearch.org/ file:///home/httpd/www.mnogosearch.org/
#

Server http://WEB_SITE_URL

#####
#Realm [Method] [CmpType] [Match|NoMatch] <arg> [alias]
# It works almost like "Server" command but takes a regular expression or
# a string wildcards as it's argument and do not insert any URL into
# database for indexing. To insert URLs into database use URL command (see
# below).
#
# "Method" is an optional parameter which can take one of the following
# values: Allow, Disallow, CheckOnly, HrefOnly, CheckMP3, CheckMP3Only
# with Allow as a default value.
#
# "CmpType" is an optional parameter to specify comparison type and can
# take either String or RegExp value with String as a default value.
#
# For example, if you want to index all HTTP sites in ".ru" domain, use:
#Realm http://*.ru/*
# The same using "Regex" match:
#Realm Regexp ^http://.*\.ru/
# Another example. Use this command to index everything without .com domain:
#Realm NoMatch http://*.com/*
#
# Optional "alias" argument allows to provide very complicated URL rewrite
# more powerful than other aliasing mechanism. Take a look into alias.txt
# for "alias" argument usage explanation.

#####
#URL http://localhost/path/to/page.html
# This command inserts given URL into database. This is useful to add
# several entry points to one server. Has no effect if an URL is already
# in the database. When inserting indexer does not executes any checks
# and this URL may be deleted at first indexing attempt if URL has no
# correspondent Server command or is disallowed by rules given in
# Allow/Disallow commands.
#
#This command will add /main/index.html page:
#URL http://localhost/main/index.html
```



### ***II . 3 . Lancement de l'indexation***

```
/usr/sbin/indexer -Ecreate
```

### ***II . 4 . Automatisation de l'indexation***

```
echo '#!/bin/sh' > /etc/cron.daily/mnogosearch  
echo '/usr/sbin/indexer' >> /etc/cron.daily/mnogosearch  
chmod u+x /etc/cron.daily/mnogosearch
```

### III . Installation de l'extension mnogosearch pour php

#### III . 1 . Récupération des sources de php

```
cd /usr/src/  
apt-get source php5
```

#### III . 2 . Récupération des sources de l'extension mnogosearch pour php

```
cd /usr/src/php5-5.2.0  
wget http://www.mnogosearch.org/Download/php/mnogosearch-php-extension-1.96.tar.gz  
tar -zxvf mnogosearch-php-extension-1.96.tar.gz  
mv 1.96 ext/mnogosearch  
rm mnogosearch-php-extension-1.96.tar.gz
```

#### III . 3 . Modification des options de compilation du package debian php

Éditer le fichier `/usr/src/php5-5.2.0/debian/rules` pour obtenir ( en rouge la ligne ajoutée ) :

```
#!/usr/bin/make -f  
# Sample debian/rules that uses debhelper.  
# GNU copyright 1997 by Joey Hess.  
#  
# This version is for a hypothetical package that builds an  
# architecture-dependant package, as well as an architecture-independent  
# package.  
  
# Uncomment this to turn on verbose mode.  
#export DH_VERBOSE=1  
  
# This is the debhelper compatability version to use.  
export DH_COMPAT=4  
  
# This has to be exported to make some magic below work.  
export DH_OPTIONS
```

```
DEB_HOST_GNU_TYPE   ?= $(shell dpkg-architecture -qDEB_HOST_GNU_TYPE)
DEB_BUILD_GNU_TYPE  ?= $(shell dpkg-architecture -qDEB_BUILD_GNU_TYPE)
DEB_BUILD_ARCH      ?= $(shell dpkg-architecture -qDEB_BUILD_ARCH)

PHP5_HOST_GNU_TYPE   = $(shell echo $(DEB_HOST_GNU_TYPE) | sed 's/-gnu$$/')
PHP5_BUILD_GNU_TYPE  = $(shell echo $(DEB_BUILD_GNU_TYPE) | sed 's/-gnu$$/')

PHP5_SOURCE_VERSION = $(shell dpkg-parsechangelog | grep ^Version | sed "s/Version: //")
PHP5_UPSTREAM_VERSION = $(shell echo $(PHP5_SOURCE_VERSION) | sed -e "s/-.*/" -e "s/.*://")
PHP5_DEBIAN_REVISION = $(shell echo $(PHP5_SOURCE_VERSION) | sed "s/.*-//")

APACHE_VERSION = $(shell dpkg -s apache-dev | grep ^Version | cut -d\ -f2 | cut -d- -f1)

PROG_SENDMAIL = /usr/sbin/sendmail
CFLAGS = -O2 -Wall -fsigned-char -fno-strict-aliasing
# LFS support
CFLAGS += $(shell getconf LFS_CFLAGS)

# Enable IEEE-conformant floating point math on alphas (not the default)
ifeq (alpha-linux,$(PHP5_HOST_GNU_TYPE))
    CFLAGS += -mieee
endif

ifeq ($(PHP5_HOST_GNU_TYPE), $(findstring $(PHP5_HOST_GNU_TYPE), ia64-linux
powerpc64-linux))
    CFLAGS += -g
else
    CFLAGS += -gstabs
endif

ifneq (nostrip, $(findstring nostrip, $(DEB_BUILD_OPTIONS)))
    install_strip = -s
endif

# Old magic.mime location:
ifeq ($(wildcard /usr/share/misc/file/magic.mime), /usr/share/misc/file/magic.mime)
MAGIC_MIME = /usr/share/misc/file/magic.mime
endif
# New magic.mime location:
ifeq ($(wildcard /usr/share/file/magic.mime), /usr/share/file/magic.mime)
MAGIC_MIME = /usr/share/file/magic.mime
endif
```

```
# only enable interbase support for i386/amd64
ifneq (,$(filter $(DEB_BUILD_ARCH),amd64 i386))
    IBFLAGS=--with-interbase=shared,/usr \
        --with-pdo-firebird=shared,/usr
endif

COMMON_CONFIG=--build=$(PHP5_BUILD_GNU_TYPE)-gnu \
    --host=$(PHP5_HOST_GNU_TYPE)-gnu \
    --mandir=/usr/share/man \
    --enable-memory-limit \
    --disable-debug \
    --with-regex=php \
    --disable-rpath \
    --disable-static \
    --with-pic \
    --with-layout=GNU \
    --with-pear=/usr/share/php \
    --enable-calendar \
    --enable-sysvsem \
    --enable-sysvshm \
    --enable-sysvmsg \
    --enable-track-vars \
    --enable-trans-sid \
    --enable-bcmath \
    --with-bz2 \
    --enable-ctype \
    --with-db4 \
    --without-gdbm \
    --with-iconv \
    --enable-exif \
    --enable-filepro \
    --enable-ftp \
    --with-gettext \
    --enable-mbstring \
    --with-pcre-regex=/usr \
    --enable-shmop \
    --enable-sockets \
    --enable-wddx \
    --with-libxml-dir=/usr \
    --with-zlib \
    --with-kerberos=/usr \
    --with-openssl=/usr \
    --enable-dbx \
    --enable-soap \
    --enable-zip \
```

```
--with-mime-magic=$(MAGIC_MIME) \  
--with-mnogosearch=shared,/usr \  
--with-exec-dir=/usr/lib/php5/libexec  
  
BUILTIN_EXTENSION_CHECK=$$e=get_loaded_extensions(); natcasesort($$e); \  
  $$s="The following extensions are built in:"; \  
  foreach($$e as $$i) { $$s .= " $$i"; } \  
  echo("php:Extensions=" . wordwrap($$s . ".\n", 75, "\${Newline} "));  
  
patch: patch-stamp  
patch-stamp:  
  dh_testdir  
  for patch in debian/patches/*.patch; do \  
    echo '->'`basename $$patch`; \  
    if ! patch -p1 --ignore-whitespace --dry-run < $$patch; \  
    then \  
      exit 1; \  
    fi; \  
    patch -p1 --ignore-whitespace < $$patch; \  
  done  
  sed -i -e 's/EXTRA_VERSION=""/EXTRA_VERSION="-${PHP5_DEBIAN_REVISION}"/'  
configure.in  
  rm -f aclocal.m4 config.sub config.guess ltmain.sh  
  ./buildconf --force  
  touch patch-stamp  
  
unpatch:  
  dh_testdir  
  sed -i -e 's/EXTRA_VERSION="-${PHP5_DEBIAN_REVISION}"/EXTRA_VERSION=""/'  
configure.in  
  if [ -f patch-stamp ]; then \  
    for patch in `ls debian/patches/*.patch | sort -r`; do \  
      patch -p1 -R --ignore-whitespace < $$patch; \  
    done; \  
    rm -f aclocal.m4 config.sub config.guess ltmain.sh; \  
    ./buildconf --force; \  
  fi  
  rm -f patch-stamp  
  
build: build-apache-stamp build-apache2-stamp build-cgi-stamp build-cli-stamp build-pear-stamp  
build-apache-stamp: configure-apache-stamp  
  dh_testdir  
  # Add here commands to compile the package.  
  cd apache-build && $(MAKE)
```

```
touch build-apache-stamp

build-apache2-stamp: configure-apache2-stamp
dh_testdir
cd apache2-build && $(MAKE)

touch build-apache2-stamp

build-cli-stamp: configure-cli-stamp
dh_testdir
cd cli-build && $(MAKE)

touch build-cli-stamp

build-cgi-stamp: configure-cgi-stamp
dh_testdir
cd cgi-build && $(MAKE) && mv sapi/cgi/php sapi/cgi/cgi-bin.php5

# Dirty hack to not rebuild everything twice
cd cgi-build/main && \
    sed -i -e 's/FORCE_CGI_REDIRECT 1/FORCE_CGI_REDIRECT 0/' \
        -e 's/DISCARD_PATH 0/DISCARD_PATH 1/' php_config.h && \
    sed -i -e 's/--enable-force-cgi-redirect/--enable-discard-path/' build-defs.h && \
    touch ../../ext/standard/info.c && \
    touch ../../sapi/cgi/cgi_main.c

cd cgi-build && $(MAKE) && mv sapi/cgi/php sapi/cgi/usr.bin.php5-cgi

touch build-cgi-stamp

build-pear-stamp: build-cgi-stamp
dh_testdir
-mkdir pear-build
cd cgi-build && make install-pear PHP_PEAR_PHP_BIN=/usr/bin/php
PHP_PEAR_INSTALL_DIR=/usr/share/php PHP_PEAR_SYSCONF_DIR=/etc/pear
PHP_PEAR_SIG_BIN=/usr/bin/gpg INSTALL_ROOT=$(CURDIR)/pear-build
    sed -i -e 's/-d output_buffering=1 -d open_basedir="" -d safe_mode=0/-d output_buffering=1
-d open_basedir="" -d safe_mode=0 -d memory_limit="-1"/' \
        $(CURDIR)/pear-build/usr/bin/pear && \
    sed -i -e 's/-d output_buffering=1 -d safe_mode=0/-d output_buffering=1 -d open_basedir=""
-d safe_mode=0 -d memory_limit="-1"/' \
        $(CURDIR)/pear-build/usr/bin/pecl && \
    sed -i -e 's/-d memory_limit="-1"/' \
        -e 's/-d output_buffering=1 -d open_basedir="" -d safe_mode=0/-d output_buffering=1 -d
```

```
open_basedir="" -d safe_mode=0 -d memory_limit="-1"/ \
    $(CURDIR)/pear-build/usr/bin/peardev
touch build-pear-stamp

configure: configure-apache-stamp configure-apache2-stamp configure-cli-stamp configure-cgi-
stamp
configure-apache-stamp: patch-stamp
    dh_testdir
    if [ -d apache-build ]; then rm -rf apache-build; fi
    -mkdir apache-build
    cd apache-build && \
    CFLAGS="$(CFLAGS)" PROG_SENDMAIL="$(PROG_SENDMAIL)" ../configure \
        --prefix=/usr --with-apxs=/usr/bin/apxs \
        --with-config-file-path=/etc/php5/apache \
        --with-config-file-scan-dir=/etc/php5/apache/conf.d \
        $(COMMON_CONFIG) \
        --without-mm \
        --disable-pdo \
        --without-mysql --without-sybase-ct --without-sqlite
    cd apache-build && \
    cp ../Zend/zend_ini_scanner.c ../Zend/zend_language_scanner.c \
        ../Zend/zend_ini_parser.h ../Zend/zend_language_parser.h \
        ../Zend/zend_ini_parser.c ../Zend/zend_language_parser.c \
        Zend/

    touch configure-apache-stamp

configure-apache2-stamp: patch-stamp
    dh_testdir
    if [ -d apache2-build ]; then rm -rf apache2-build; fi
    -mkdir apache2-build
    cd apache2-build && \
    CFLAGS="$(CFLAGS)" PROG_SENDMAIL="$(PROG_SENDMAIL)" ../configure \
        --prefix=/usr --with-apxs2=/usr/bin/apxs2 \
        --with-config-file-path=/etc/php5/apache2 \
        --with-config-file-scan-dir=/etc/php5/apache2/conf.d \
        $(COMMON_CONFIG) \
        --without-mm \
        --with-curl=shared,/usr \
        --with-zlib-dir=/usr \
        --with-gd=shared,/usr --enable-gd-native-ttf \
        --with-jpeg-dir=shared,/usr \
        --with-xpm-dir=shared,/usr/X11R6 \
        --with-png-dir=shared,/usr \
        --with-freetype-dir=shared,/usr \
```

```
--with-ldap=shared,/usr \  
--with-ldap-ssl \  
$(IBFLAGS) \  
--with-ttf=shared,/usr \  
--with-t1lib=shared,/usr \  
--with-ldap=shared,/usr \  
--with-mcrypt=shared,/usr \  
--with-mhash=shared,/usr \  
--with-mysql=shared,/usr \  
--with-mysqli=shared,/usr/bin/mysql_config \  
--with-pspell=shared,/usr \  
--with-unixODBC=shared,/usr \  
--with-recode=shared,/usr \  
--with-xsl=shared,/usr \  
--with-snmp=shared,/usr \  
--with-sqlite=shared,/usr \  
--with-sybase-ct=shared,/usr \  
--with-tidy=shared,/usr \  
--with-xmlrpc=shared \  
--with-pgsql=shared,/usr PGSQL_INCLUDE=`pg_config --includedir` \  
--enable-pdo=shared \  
--without-pdo-dblib \  
--with-pdo-mysql=shared,/usr \  
--with-pdo-odbc=shared,unixODBC,/usr \  
--with-pdo-pgsql=shared,/usr/bin/pg_config \  
--with-pdo-sqlite=shared,/usr  
cd apache2-build && \  
cp ../Zend/zend_ini_scanner.c ../Zend/zend_language_scanner.c \  
../Zend/zend_ini_parser.h ../Zend/zend_language_parser.h \  
../Zend/zend_ini_parser.c ../Zend/zend_language_parser.c \  
Zend/  
touch configure-apache2-stamp  
  
configure-cgi-stamp: patch-stamp  
dh_testdir  
if [ -d cgi-build ]; then rm -rf cgi-build; fi  
-mkdir cgi-build  
cd cgi-build && \  
CFLAGS="$(CFLAGS)" PROG_SENDMAIL="$(PROG_SENDMAIL)" ../configure \  
--prefix=/usr --enable-force-cgi-redirect --enable-fastcgi \  
--with-config-file-path=/etc/php5/cgi \  
--with-config-file-scan-dir=/etc/php5/cgi/conf.d \  
$(COMMON_CONFIG) \  
--without-mm \  
--disable-pdo \  

```

```
--without-mysql --without-sybase-ct --without-sqlite
cd cgi-build && \
cp ../Zend/zend_ini_scanner.c ../Zend/zend_language_scanner.c \
  ../Zend/zend_ini_parser.h ../Zend/zend_language_parser.h \
  ../Zend/zend_ini_parser.c ../Zend/zend_language_parser.c \
  Zend/
touch configure-cgi-stamp

configure-cli-stamp: patch-stamp
dh_testdir
if [ -d cli-build ]; then rm -rf cli-build; fi
-mkdir cli-build
cd cli-build && \
CFLAGS="$(CFLAGS)" PROG_SENDMAIL="$(PROG_SENDMAIL)" ../configure \
  --prefix=/usr --disable-cgi \
  --with-config-file-path=/etc/php5/cli \
  --with-config-file-scan-dir=/etc/php5/cli/conf.d \
  $(COMMON_CONFIG) \
  --without-mm \
  --disable-pdo \
  --without-mysql --without-sybase-ct --without-sqlite \
  --enable-pcntl \
  --with-ncurses=/usr
cd cli-build && \
cp ../Zend/zend_ini_scanner.c ../Zend/zend_language_scanner.c \
  ../Zend/zend_ini_parser.h ../Zend/zend_language_parser.h \
  ../Zend/zend_ini_parser.c ../Zend/zend_language_parser.c \
  Zend/
touch configure-cli-stamp

clean: unpatch
dh_testdir
dh_testroot
rm -f configure-apache-stamp build-apache-stamp
rm -f configure-apache2-stamp build-apache2-stamp
rm -f configure-cgi-stamp build-cgi-stamp
rm -f configure-cli-stamp build-cli-stamp
rm -f build-pear-stamp
rm -f install-stamp
rm -rf apache-build
rm -rf apache2-build
rm -rf cgi-build
rm -rf cli-build
rm -rf pear-build
rm -f debian/copyright
```

```
dh_clean
# clean up autogenerated cruft
cat debian/modulelist | while read package extname dsoname; do \
    rm -f debian/php5-$$package.postinst; \
done
for sapi in libapache-mod-php5 libapache2-mod-php5 php5-cgi php5-cli; do \
    for cruft in postrm links; do \
        rm -f debian/$$ {sapi} .$$ {cruft}; \
    done; \
done

install: DH_OPTIONS=
install: build
    dh_testdir
    dh_testroot
    dh_clean -k
    dh_installdirs

    chmod 01733 debian/php5-common/var/lib/php5

# Add here commands to install the package into debian/php5.
# install apache DSO module
cp apache-build/.libs/libphp5.so debian/libapache-mod-php5/apxs -q LIBEXECDIR`
cp debian/libapache-mod-php5.500mod_php5.info \
    debian/libapache-mod-php5/apxs -q LIBEXECDIR`/500mod_php5.info
cp debian/libapache-mod-php5.conf \
    debian/libapache-mod-php5/etc/apache/conf.d/php5.conf

# install apache2 DSO module
cp apache2-build/.libs/libphp5.so \
    debian/libapache2-mod-php5/apxs2 -q LIBEXECDIR`/
cp debian/libapache2-mod-php5.load \
    debian/libapache2-mod-php5/etc/apache2/mods-available/php5.load
cp debian/libapache2-mod-php5.conf \
    debian/libapache2-mod-php5/etc/apache2/mods-available/php5.conf

# sanitize php.ini file
cat php.ini-dist | tr "\t" " " > debian/php5-common/usr/share/php5/php.ini-dist
cat php.ini-dist | tr "\t" " " | sed -e'/memory_limit =/ s/8M/32M/g' > debian/php5-common/usr/
share/php5/php.ini-dist.cli
cat php.ini-dist | tr "\t" " " > debian/php5-common/usr/share/doc/php5-
common/examples/php.ini-dist
cat php.ini-recommended | tr "\t" " " > debian/php5-common/usr/share/doc/php5-
common/examples/php.ini-recommended
cat php.ini-paranoid | tr "\t" " " > debian/php5-common/usr/share/doc/php5-
```

```
common/examples/php.ini-paranoid
```

```
# install the apache modules' files
cd apache2-build && make install-headers install-build install-modules install-programs
INSTALL_ROOT=$(CURDIR)/debian/libapache2-mod-php5
# remove netware and win32 headers that we don't want
cd debian/libapache2-mod-php5/usr/include/php5/ && \
    rm -f TSRM/readdir.h \
        TSRM/tsrm_config.{nw,w32}.h \
        TSRM/tsrm_{nw,win32}.h \
        Zend/zend_config.{nw,w32}.h \
        main/config.{nw,w32}.h \
        main/win95nt.h

# install PEAR
cp -r pear-build/* debian/php-pear/

# install extensions
ext=`./debian/libapache2-mod-php5/usr/bin/php-config --extension-dir`; \
for i in libapache2-mod-php5 libapache2-mod-php5 php5-cgi php5-cli; do \
    mkdir -p debian/$$i/$${ext}; \
done; \
cat debian/modulelist debian/extramodulelist | while read package extname dsoname; do \
    if [ -z "$$dsoname" ]; then dsoname=$$package; fi; \
    mkdir -p debian/php5-$$package$${ext}; \
    chrpath debian/libapache2-mod-php5/$${ext}/$$dsoname.so; \
    chrpath -d debian/libapache2-mod-php5/$${ext}/$$dsoname.so; \
    install ${{install_strip}} -m 644 -o root -g root \
        debian/libapache2-mod-php5/$${ext}/$$dsoname.so \
        debian/php5-$$package$${ext}/$$dsoname.so; \
    rm debian/libapache2-mod-php5/$${ext}/$$dsoname.so; \
done

# install CGI
cp cgi-build/sapi/cgi/cgi-bin.php5 debian/php5-cgi/usr/lib/cgi-bin/php5
cp cgi-build/sapi/cgi/usr.bin.php5-cgi debian/php5-cgi/usr/bin/php5-cgi
cp cli-build/sapi/cli/php.1 debian/php5-cgi/usr/share/man/man1/php5-cgi.1

# install CLI
cp cli-build/sapi/cli/php debian/php5-cli/usr/bin/php5
cp cli-build/sapi/cli/php.1 debian/php5-cli/usr/share/man/man1/php5.1

# move and install -dev files
dh_movefiles --sourcedir=debian/libapache2-mod-php5
rm -rf debian/libapache2-mod-php5/usr/lib/php5/build/ \
```

```
    debian/libapache2-mod-php5/usr/include/ \
    debian/libapache2-mod-php5/usr/bin/
for i in Makefile.global acinclude.m4 mkdep.awk phpize.m4 scan_makefile_in.awk; do \
    chmod 644 debian/php5-dev/usr/lib/php5/build/$$i; \
done
# shipping duplicate files from other packages is hell for security audits
rm debian/php5-dev/usr/lib/php5/build/config.guess && \
    ln -s ../../share/misc/config.guess debian/php5-dev/usr/lib/php5/build/config.guess
rm debian/php5-dev/usr/lib/php5/build/config.sub && \
    ln -s ../../share/misc/config.sub debian/php5-dev/usr/lib/php5/build/config.sub
rm debian/php5-dev/usr/lib/php5/build/libtool.m4 && \
    ln -s ../../share/libtool/libtool.m4 debian/php5-dev/usr/lib/php5/build/libtool.m4
rm debian/php5-dev/usr/lib/php5/build/ltmain.sh && \
    ln -s ../../share/libtool/ltmain.sh debian/php5-dev/usr/lib/php5/build/ltmain.sh
rm debian/php5-dev/usr/lib/php5/build/shtool && \
    ln -s ../../bin/shtool debian/php5-dev/usr/lib/php5/build/shtool
# make php-dev stuff versioned
for i in php-config phpize; do \
    mv debian/php5-dev/usr/bin/$$i debian/php5-dev/usr/bin/"$$i"5; \
    mv debian/php5-dev/usr/share/man/man1/"$$i".1 debian/php5-
dev/usr/share/man/man1/"$$i"5.1; \
done

# install common files
install -m755 debian/maxlifetime debian/php5-common/usr/lib/php5

touch install-stamp

# Build architecture-independent files here.
# Pass -i to all debhelper commands in this target to reduce clutter.
binary-indep: DH_OPTIONS=-i
binary-indep: build install
    # Need this version of debhelper for DH_OPTIONS to work.
    dh_testdir
    dh_testroot
    cat debian/copyright.header LICENSE Zend/LICENSE > debian/copyright

    dh_installdocs

for package in php5 php-pear; do \
    rm -rf debian/$$package/usr/share/doc/$$package; \
    ln -s php5-common debian/$$package/usr/share/doc/$$package; \
done

dh_link
```

```
dh_compress -Xphp.ini
dh_fixperms
dh_installdeb
dh_gencontrol
dh_md5sums
dh_builddeb
```

```
# Build architecture-dependent files here.
```

```
binary-arch: build install
```

```
# Need this version of debhelper for DH_OPTIONS to work.
```

```
dh_testdir
```

```
dh_testroot
```

```
# Do this first so we don't overwrite any debhelper-generated files
```

```
#
```

```
# generate the maintscripts for various php
```

```
# modules from the templates.
```

```
cat debian/modulelist | while read package extname dsoname; do \
    if [ -z "$$dsoname" ]; then dsoname=$$package; fi; \
    sed -e"s/@extname@/$$extname/g; s/@dsoname@/$$dsoname/g" \
        < debian/php5-module.postinst \
        > debian/php5-$$package.postinst; \
done
```

```
# generate the config snippets for various php
```

```
# modules from the templates.
```

```
cat debian/modulelist debian/extramodulelist | while read package extname dsoname; do \
    if [ -z "$$dsoname" ]; then dsoname=$$package; fi; \
    mkdir -p debian/php5-$$package/etc/php5/conf.d; \
    sed -e"s/@extname@/$$extname/g; s/@dsoname@/$$dsoname/g" \
        < debian/php5-module.ini \
        > debian/php5-$$package/etc/php5/conf.d/$$dsoname.ini; \
done
```

```
# likewise, for the different sapi implementations
```

```
for tmpl in postrm links; do \
```

```
    for sapi in apache apache2 cgi cli; do \
```

```
        sed -e "s/@sapi@/$$sapi/g" \
```

```
        < debian/php5-sapi.$$tmpl \
```

```
        > debian/php5-$$sapi.$$tmpl; \
```

```
    done; \
```

```
    mv debian/php5-apache.$$tmpl debian/libapache-mod-php5.$$tmpl; \
```

```
    mv debian/php5-apache2.$$tmpl debian/libapache2-mod-php5.$$tmpl; \
```

```
done
```

```
cat debian/copyright.header LICENSE Zend/LICENSE > debian/copyright
```

```
dh_installdocs -s

cat debian/modulelist | while read package extname dsoname; do \
    rm -rf debian/php5-$$package/usr/share/doc/php5-$$package; \
    ln -s php5-common debian/php5-$$package/usr/share/doc/php5-$$package; \
done

for package in php5-dev php5-cgi php5-cli libapache-mod-php5 libapache2-mod-php5; do \
    rm -rf debian/$$package/usr/share/doc/$$package; \
    ln -s php5-common debian/$$package/usr/share/doc/$$package; \
done
dh_installcron -pphp5-common --name=php5
dh_installchangelogs -pphp5-common NEWS
dh_strip -s
dh_link -s
dh_compress -s -Xphp.ini
dh_fixperms -s -X /var/lib/php5
mkdir -p debian/php5-common/usr/share/lintian/overrides
echo "php5-common: non-standard-dir-perm var/lib/php5/ 1733 != 0755" >> debian/php5-
common/usr/share/lintian/overrides/php5-common
mkdir -p debian/php5-common/usr/share/linda/overrides
echo "Tag: non-standard-dir-perm" >> debian/php5-common/usr/share/linda/overrides/php5-
common
echo "Data: /var/lib/php5.*" >> debian/php5-common/usr/share/linda/overrides/php5-
common
dh_installdeb -s
dh_shlibdeps -s

phpapi=`./debian/php5-dev/usr/bin/php-config5 --phpapi`; \
for i in libapache-mod-php5 libapache2-mod-php5 php5-cgi php5-cli; do \
    echo "php:Provides=phpapi-$$ {phpapi}" >> debian/$$i.substvars; \
done; \
cat debian/modulelist | while read package extname dsoname; do \
    echo "php:Depends=phpapi-$$ {phpapi}" >> debian/php5-$$package.substvars; \
done

for i in cgi cli; do \
    "$$i"-build/sapi/cli/php -n -r '$(BUILTIN_EXTENSION_CHECK)' \
    >> debian/php5- "$$i".substvars; \
done
for i in apache apache2; do \
    "$$i"-build/sapi/cli/php -n -r '$(BUILTIN_EXTENSION_CHECK)' \
    >> debian/lib "$$i"-mod-php5.substvars; \
done
```

```
echo "apache:Depends=apache-common (>= $(APACHE_VERSION))" >>debian/libapache-  
mod-php5.substvars  
echo "apache2:Depends=apache2-mpm-prefork (>> 2.0.52) | apache2-mpm-itk, apache2.2-  
common" >>debian/libapache2-mod-php5.substvars  
dh_gencontrol -s  
dh_md5sums -s  
dh_builddeb -s  
  
binary: binary-arch binary-indep  
.PHONY: build clean binary-indep binary-arch binary install configure
```

Éditer le fichier `/usr/src/php5-5.2.0/debian/extramodulelist` pour obtenir ( en rouge la ligne ajoutée ) :

```
mysql MySQL mysqli  
mysql MySQL pdo_mysql  
interbase InterBase/Firebird pdo_firebird  
common PDO pdo  
odbc ODBC pdo_odbc  
pgsql PostgreSQL pdo_pgsql  
sqlite SQLite pdo_sqlite  
mnogosearch mnoGoSearch mnogosearch
```

### **III . 4 . Compilation de l'extension**

```
cd /usr/src/php5-5.2.0  
dpkg-buildpackage
```

### **III . 5 . Installation de la nouvelle version php supportant l'extension mnogosearch**

```
cd /usr/src/  
dpkg -i php5_5.2.0-8+etch13_all.deb  
dpkg -i php5-common_5.2.0-8+etch13_i386.deb  
dpkg -i php5-gd_5.2.0-8+etch13_i386.deb  
dpkg -i php5-mcrypt_5.2.0-8+etch13_i386.deb  
dpkg -i php5-mysql_5.2.0-8+etch13_i386.deb  
dpkg -i php5-odbc_5.2.0-8+etch13_i386.deb
```

### **III . 6 . Copie de l'extension dans le répertoire des modules php**

```
cp /usr/src/php5-5.2.0/apache2-build/ext/mnogosearch/.libs/mnogosearch.so  
/usr/lib/php5/20060613+ifs/
```

### **III . 7 . Ajout de la directive de chargement automatique du module**

```
echo "extension=mnogosearch.so" > /etc/php5/conf.d/mnogosearch.ini
```

### **III . 8 . Redémarrage d'apache**

```
apache2ctl graceful
```



**BYSOFT**

26 Rue Bancel

77000 MELUN

Tel : +33 1 64 81 18 48

Fax : +33 1 60 66 87 72