

Tru64 UNIX Best Practice

Creating Multiple Indexes and Search Pages Using Ht://Dig

June 2003

Product Version: **Internet Express Version 6.1**

This Best Practice describes how to create multiple indexes and search pages using the ht://Dig search tool provided by *HP* Internet Express.

Hewlett-Packard Company
Palo Alto, California

Contents

Creating Multiple Indexes and Search Pages Using ht://Dig

Is This Best Practice Right for You?	1
Before You Begin	2
Applying the Best Practice	2
Create Configuration Files for Multiple Indexes	3
Create a Search Page Form	3
Verifying Success	4
Troubleshooting	4
Alternative Practices	4
Comments and Questions	4
Legal Notice	5

Creating Multiple Indexes and Search Pages Using ht://Dig

Internet Express is a collection of popular Open Source and other Internet software on CD-ROM. Internet Express gives you the latest versions of the ht://Dig search tool. Using ht://Dig, you can create multiple search pages and search multiple indexes (databases).

See the Tru64 UNIX Best Practices Web page for more information about Best Practices documentation:

http://www.tru64unix.compaq.com/docs/best_practices

Is This Best Practice Right for You?

Not all Best Practices apply to all configurations, so you must be sure this Best Practice is appropriate for your system and circumstances. To use this Best Practice, you must meet the requirements described in the following table.

Requirement	Description
Operating System	Tru64 UNIX Version 5.1 or higher
Product Version	Internet Express Version 6.1 or higher
System Configuration	High-speed network interface. Consider multiple network interfaces for very high volumes.

If you do not meet the previous requirements, see *Alternative Practices* for information.

Before You Begin

Before you apply the Best Practice for Creating Multiple Indexes and Search Pages Using ht://Dig, you must understand some background information and perform some preliminary tasks:

1. If you have not already done so, register the IP address and any IP aliases for the system and configure the network. The Internet Express installation procedure will lead you through the network configuration if necessary.
2. Obtain the Internet Express kit — HP includes the Internet Express CD-ROM with Tru64 UNIX AlphaServer systems. If you need the Internet Express CD-ROMs, you can contact your HP representative. The part number for the Internet Express Kit is QB-3NCAA-SA.
3. Install the Internet Express Administration utility subset (IAEADM) and ht://Dig subset (IAEHTDIG). Follow the instructions in the *Internet Express for Tru64 UNIX Installation Guide*.

The Internet Express installation procedure automatically selects prerequisite subsets for installation. For ht://Dig, which requires the Apache Web Server, it selects the Apache Web Server subset (IAEAPCH). When you install the Apache Web Server, be sure to choose to configure a public Web server when asked.

4. Use the Internet Express Administration utility ht://Dig page to create a single search index. When you click on the htdig symlink button, the Administration utility makes the ht://Dig files available from the document root. Complete the process by following the steps for creating a search index in the *Internet Express for Tru64 UNIX Administration Guide*.
5. Install and configure a Web server and then test that the single search index works properly. Follow the instructions for searching the index in the *Internet Express for Tru64 UNIX Administration Guide*.

Applying the Best Practice

Before you apply this Best Practice, be sure to follow the recommendations in *Before You Begin*.

The Internet Express CD-ROM contains the ht://Dig search tool, as well as other Internet software. This Best Practice recommends that you:

1. Obtain the Internet Express kit and install the Internet Express Administration utility subset (IAEADM) and ht://Dig subset (IAEHTDIG). (See *Before You Begin*).

2. *Create Configuration Files for Multiple Indexes* — Edit the `ht://Dig` configuration file to create a new configuration file for each search index.
3. *Create a Search Page Form*.

Create Configuration Files for Multiple Indexes

Create a configuration file and directory for each index you want to create:

1. Make a copy of the `/usr/internet/www/conf/htdig.conf` file, and save it under a new name (for example, `new-htdig-config.conf`).
2. In the `new-htdig-config.conf` file, change the `database_dir` attribute to point to a new directory (for example, `htdig-new-directory`):

```
database_dir: /usr/internet/www/var/htdig-new-directory
```

3. In the `new-htdig-config.conf` file, change the `start_url` attribute to point to the URL for the content to be indexed. For example:

```
start_url: http://xyz.company.com/content/
```

4. From the command line, create the search index (database) using the `rundig` command with the `-c` option:

```
# /usr/internet/www/bin/rundig -c new-htdig-config.conf
```

Create a Search Page Form

After creating a configuration file for multiple search indexes (see *Create Configuration Files for Multiple Indexes*), create a form for each search page:

1. Copy the `/usr/internet/www/htdocs/htdig/search.html` file. This file will serve as a template for the new search page form.
2. In the `search.html`, change the following line so that the `value` parameter defines your new configuration file (for example, `new-htdig-config.conf`):

```
<input type="hidden" name="config" value="new-htdig-config">
```

3. Save the new file and exit the editor. Make sure that your new search page is accessible from the Web server.

Verifying Success

After you apply the Best Practice for Creating Multiple Indexes and Search Pages Using ht://Dig, you can verify whether it was successful.

- You successfully installed the ht://Dig, Apache Web Server, and Administration utility from Internet Express.
- You successfully created multiple search indexes and search pages.
- You were able to search multiple indexes using the new search pages.

If the Best Practice was not successful, see *Troubleshooting* for information about identifying and solving problems.

Troubleshooting

If you determine that the Best Practice was not successful, as described in *Verifying Success*, use the following table to identify and solve problems.

Problem	Possible Solutions
An Internet Express subset did not install.	Review the installation steps described in the <i>Internet Express for Tru64 UNIX Installation Guide</i> .
You could not search an index using the new search page.	Create another search index with the <code>rundig</code> command, using the <code>-v</code> option to capture the output in a log file: <pre># /usr/internet/www/bin/rundig -v -c new-htdig-config.conf</pre> Review the log file for any new problems.

Alternative Practices

Although this Best Practice is the recommended method for configuring the IMAP mail server, you can use an alternative method if your system does not meet the requirements described in *Is This Best Practice Right for You?*.

Install the binary kit or install sources copied from

<http://www.htdig.org>

Comments and Questions

We value your comments and questions on the information in this document. Please mail your comments to us at this address:

Legal Notice

UNIX® and The Open Group™ are trademarks of The Open Group in the U.S. and/or other countries. All other product names mentioned herein may be trademarks of their respective owners.

Confidential computer software. Valid license from HP required for possession, use, or copying. Consistent with FAR 12.2.11 and 12.212, Commercial Computer Software, Computer Software Documentation, and Technical Data for Commercial Items are licensed to the U.S. Government under vendor's standard commercial license.

Neither HP nor any of its subsidiaries shall be liable for technical or editorial errors or omissions contained herein. The information is provided "as is" without any warranty of any kind and is subject to change without notice. The warranties of HP products are set forth in express limited warranty statements accompanying such products. Nothing herein should be construed as constituting an additional warranty.