

# apache+perl+BerkeleyDBで簡易リダイレクト

2011/09/04 作成freestyle wiki から dokuwiki に移行する際にやった事。

2013/04/06

こちらのドキュメントは管理都合によりwww関連ドキュメントのネームスペースから移動しました。

## 概要

wikiシステムの更新でURLの指定方法がガラリと変わりました。

旧URLから新URLへのリダイレクトを行うにあたり.htaccessへの ReDirectディレクティブ記述では難しい場合がありCGIで置き換えする方法をとりました。

その時の備忘録です。

## URL変換

新旧URLは以下ようになります。

旧URL	<code>http://hgotoh.jp/wiki/wiki.cgi?page=[ページ名(URLエンコードされたもの)]</code>
新URL	<code>http://hgotoh.jp/wiki/doku.php/[documents/カテゴリ/ページ名(カテゴリ+番号)]</code>

旧URLではページ名にはEUC-JPな漢字を使いまくっていたため、見た目にも長くて大変なURLになっていました。これを、新URLではページ名にカテゴリ+番号を適用し、URLをすっきりさせます。

また、旧URLは検索エンジンのインデックスにすでに登録されているし、他の人のサイトで使われていたりします。利便性やケアを考えると、旧URLのアクセスでも新URLへリダイレクトしてあげるほうが混乱も無くよいかと思えます。

リダイレクトする場合、ページ名が同じでベースのURLが異なるだけなら .htaccess の ReDirectディレクティブを記述すればよいのですが、完全に不一致なURL間の変換になるので、.htaccess へフルパスで記述しなければならなくなります。面倒です。

## CGI

旧URLで使用していたwikiシステムの wiki.cgi を廃止し、新たに新URLへのリダイレクトプログラムとして作成します。具体的にはpageパラメタで指定された旧ページ名を新ページ名に変換し、新URLを生成、このURLへSTATUS 301でリダイレクトさせます。

[wiki.cgi](#)

```
use BerkeleyDB;

my $DB = new BerkeleyDB::Hash -
    Filename=>"/home/hogehoge/redirect.db" ;
```

```
my $page      = $ENV{'QUERY_STRING'};
my $newpage   = "";

print "Status: 301 Moved Permanently\n";

if ( $page eq "" )
{
    print "Location: http://hgotoh.jp/wiki/\n\n";
}
else
{
    $page =~ s/page=(.+[^\&]).*/$1/;

    if ( $DB->db_get($page, $newpage) == 0 )
    {
        print "Location: http://hgotoh.jp/wiki/doku.php$newpage\n\n";
    }
    else
    {
        print "Location: http://hgotoh.jp/wiki/olddoclist.html\n\n";
    }
}

undef $DB;
```

バークレイ /home/hogehoge/redirect.db に登録されている単純な Key - Value の一覧から Keyを旧ページ名として検索を行い、検索できればValueを新ページ名とし新URLに適用、遷移させます。検索できなければ、あらかじめ準備したエラー画面へ遷移させます。

スクリプト中の

```
print "Status: 301 Moved Permanently\n";
```

はCGIからHTTPのステータスを制御させるためにStatusヘッダを出力させます。このヘッダで指定されたステータスをapacheがクライアントに返します。Statusヘッダ自体はクライアントへ送りません。

## 一覧表の作成

バークレイ /home/hogehoge/redirect.db に登録するスクリプトは以下のようになります。

[redirectdb.pl](#)

```
use BerkeleyDB;

my $DB = new BerkeleyDB::Hash -Filename=>"/home/hogehoge/redirect.db",
    -Flags=>DB_CREATE;

$DB->db_put("RSS%A3%B1%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss01")
```

```
;  
$DB->db_put("RSS%A3%B2%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss02")  
;  
$DB->db_put("RSS%A3%B3%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss03")  
;  
$DB->db_put("RSS%A3%B4%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss04")  
;  
$DB->db_put("RSS%A3%B5%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss05")  
;  
$DB->db_put("RSS%A3%B6%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss06")  
;  
$DB->db_put("RSS%A3%B7%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss07")  
;  
$DB->db_put("RSS%A3%B8%A5%A6%A5%AA%A5%C3%A5%C1", "/documents/rss/rss08")  
;  
$DB->db_put("FrontPage", "/start");  
  
undef $DB;
```

旧ページ名「RSSウオッチ」はURLエンコードされると  
RSS%A3%B1%A5%A6%A5%AA%A5%C3%A5%C1になります。  
これに対応するのが新ページ名「/documents/rss/rss01」です。

必要なページの数だけ、エントリを増やしていきます。

## URLエンコードの罨

URLエンコードは本来一意に決まるはずなのですが、その解釈の仕方でずいぶん流儀があるようです。

旧ページ名の「JDK16のコンパイル on FreeBSD/AMD64 8.0 RELEASE」ですが、現在確認されているだけで以下の三通りがあります。

旧ページ名	URLエンコードされた旧ページ名
JDK16のコンパ イル on FreeBSD/AMD64	JDK16%A4%CE%A5%B3%A5%F3%A5%D1%A5%A4%A5%EB+on+FreeBSD%2FAMD64+8%2E0+RELEASE
8.0 RELEASE	JDK16%C2%A4%CE%A5%C2%B3%C2%A5%C3%B3%C2%A5%D1%A5%C2%A4%C2%A5%C3%AB%20on%20FreeBSD/AMD64%208.0%20RELEASE
	JDK16%A4%CE%A5%B3%A5%F3%A5%D1%A5%A4%A5%EB+on+FreeBSD%2FAMD64+8.0+RELEASE

スペースは“+”になるはずなのですが、%20で扱っていたり、変な変換がされたコード列が使われていたり.... こういう場合はエントリを三つ登録しなければいけません。

[技術資料](#), [Perl](#), [URL](#), [リダイレクト](#)

From:  
<https://wiki.hgotoh.jp/> - 努力したWiki

Permanent link:  
<https://wiki.hgotoh.jp/documents/perl/perl-011>

Last update: **2024/11/01 16:30**

