

文章

[Michael Lei](#) · 五月 24, 2021 阅读大约需 3 分钟

将 pButtons 数据提取到 csv 文件以便绘制图表

本帖的目的是回答一个问题。在本系列的第二篇帖子中，我包括了从 pButtons 提取的性能数据的图表。有人在线下问我，有没有比剪切/粘贴更快的方法从 pButtons .html文件中提取 mgstat 等指标，以便在 Excel 中绘图。

参见：- [第 2 部分 - 研究收集的指标](#)

pButtons 将其收集的数据编译成一个 html 文件，以便发送给 WRC 和查看整理的数据。不过，尤其是对于长时间（如 24 小时）进行收集的 pButtons 来说，一些基于时间的数据（如 mgstat、vmstat 等）以图形方式查看更容易查找趋势或模式。

我知道把 pButtons 数据压缩到一个 html 文件再花时间去解压听起来很疯狂，但请记住，pButtons 是 WRC 用来获取许多系统指标的视图以帮助解决性能问题的工具。系统级指标和 Caché 指标可以单独运行，但对我来说，在这个系列中使用 pButtons 捕获和分析性能指标是很方便的，因为我知道所有 Caché

安装都会有一个副本，或者可以下载副本，而且所有基本指标都可以放在一个文件中供不同的操作系统使用。能够每天通过一个简单的例程来捕获这些指标也是很方便的，前提是没有以任何其他方式收集数据。

2017 年 2 月。我用 Python 重写了本文中的脚本，并添加了包括交互式 html 在内的图表。我认为 Python 实用工具有用得多。请参见 [Yape - 另一个 pButtons 提取程序（以及自动创建图表）](#)

将 pButtons 性能指标提取到 csv 文件

由于我使用 Apple 笔记本电脑和 Unix 操作系统，所以很自然地写了一个快速 shell 脚本来提取数据到 csv 文件。以下脚本从 pButtons .html 文件中提取 mgstat、vmstat 或 Windows 性能监视器数据。下面的示例使用了大多数 *nix 系统都已安装的 Perl，但也可以使用其他脚本语言或在 Windows 上使用 powershell。

我将展示如何进行提取，有了这些信息，您就可以使用您喜欢的工具来执行同样操作。关键是 html 文件中有标记来分隔指标。例如，mgstat 用括号括起：

```
<!-- beg_mgstat -->
```

和

```
<!-- end_mgstat -->
```

在 mgstat 部分中还有一些其他描述符信息，后面是 mgstat 输出的标题行。vmstat 和 winperfmon 的标记类似。

这个简单的脚本只是查找开始标记，然后输出从标题行到结束标记之前的行的所有内容。

```
#!/usr/bin/perl
```

```
# extract_pButtons.pl - Simple extractor for pButtons
```

将 pButtons 数据提取到 csv 文件以便绘制图表

Published on InterSystems Developer Community (<https://community.intersystems.com>)

```
# usage: ./extract_pButtons.pl <input pButtons> <search start> <search first line out
put>

# pButtons has the following markers in the html source
# Metrics                               Parameters to pass
# -----                               -----
# mgstat                               mgstat Date
# windows performance monitor         win_perfmon Time
# vmstat                               vmstat fre

# usage example - Search for mgstat and redirect to .csv file
# ./extract_pButtons.pl DB1_20160211_0001_24Hour_5Sec.html mgstat Date > myMgstatOutp
ut.csv

# usage example - Process a set of html files
# for i in $(ls *.html); do ./extract_pButtons.pl ${i} vmstat fre > ${i}.vmstat.csv ;
done

# usage example - Pipeline to add commas
# ./extract_pButtons.pl P570A_CACHE_20150418_0030_day.html vmstat fre | ./make_csv.pl
>P570A_CACHE_20150418_0030_day.html.vmstat.csv

$filename=$ARGV[0];
$string=$ARGV[1];
$firstLine=$ARGV[2];

$searchBeg="beg_".$string;
$search2=$firstLine;
$foundEnd="end_".$string;

$foundString=0;
$printIt=0;
$break=0;

open HTMLFILEIN, "<".$filename or die "Bad input file";

while (<HTMLFILEIN>) {

    if (/ $searchBeg/) {
        $foundString=1;
    }

    # Look for first actual line - use something on header line
    if (($foundString==1) && (/ $search2/)) {
        $printIt=1;
    }

    # No more data
    if (/ $foundEnd/) {
        $break=1;
    }

    if ($break==0) {

        if ($printIt==1) {
            print;
        }
    }
}
```

```
}
```

```
close HTMLFILEIN;
```

如脚本开头的 # 注释所示，extractpButtons.pl 既可以将数据输出到屏幕，也可以将输出重定向到 csv 文件，或者在较长的工作流程中使用流水线，例如输出到绘图实用工具。我使用开源的 gnuplot，但 excel 也没问题。

向空格分隔的文本文件添加逗号

以下 perl 短脚本可以方便地将 vmstat 的输出或其他文本文件转换为逗号分隔的文件，以便进行处理。

```
#!/usr/bin/perl

# Convert space delimited text file to csv

# Usage example 1:
# Will create backup file vmstat.csv.bak and original file called vmstat.csv will be
# updated
# ./make_csv.pl freecnt.csv

# Usage example 2:
# No backup, original vmstat.txt file stays same, new output csv file
# ./make_csv.pl < vmstat.txt >freecnt.csv

use strict;

# create .bak backup file for each change
$^I = ".bak";

while (<>) {
    # remove leading blanks; substitute 1 or more blanks for a single comma
    s/^ +//; s/ +/,/g;
    print;
}
```

总结

我鼓励您查看 pButtons 文件的 .html 源代码，以了解其包含的内容。其中不仅有系统指标。例如，您将在 .html 文件的顶部看到 pButtons 运行的命令列表以及版本信息。

如果您使用 Windows 脚本提取或绘制数据，或者有更好或不同的工作流程，我鼓励您发帖与开发者社区分享。

[#CSV](#) [#性能](#) [#系统管理](#) [#InterSystems IRIS](#)

源

URL:

<https://cn.community.intersystems.com/post/%E5%B0%86-pbuttons-%E6%95%B0%E6%8D%AE%E6%8F%90%E5%8F%96%E5%88%B0-csv-%E6%96%87%E4%BB%B6%E4%BB%A5%E4%BE%BF%E7%BB%98%E5%88%B>

[6%E5%9B%BE%E8%A1%A8](#)