Using server performance statistics to make consortium fees fairer

Clare Whittaker

IGeLU 2007
Brno, 3 September.

A script which captures system load
Analysing the data

A new price structure

How we used server performance statistics to make Consortium membership fees fairer

Dire performance problems

• Circ hanging & system slowness coincided with a new site joining the Consortium
• more unplanned downtime in 6 months than in the previous 5 years

We tracked down a random memory segment error

• Was it ad-hoc SQL?
• Were our sizing metrics wrong?

There were a number of triggers…

Production Infrastructure autumn 2004

Database Server

Application Server

Academic Hub

Special Libraries’ Hub

Mirrors

Founded 1999 as Scottish Voyager Consortium
• Financial savings
• Operational benefits
• Cross-sectoral synergy
• A successful model
• Organic growth

Attribution:
http://www.flickr.com/photos/14516334@N00/337095799/

Attribution:
http://www.flickr.com/photos/danakay/126738338/
Analysing load in the SDLC environment

- news
- unix commands e.g. ps-ef grep [database name]
- Inhouse script

A script to snapshot system load
Captures hourly a count of svr processes
By service
By database
Allows us to monitor peaks by database (institution)

Perl script (1)

```perl
open (PS, "/usr/svclib/ps -sxuwvv") || die ("ps failed\n");
$envs = split ($/); for ($i = 0; $i <= $#$envs; $i++) {
    print $i . "\n";
}
$psflags = "-P -f -a -i -v -o " ;
if ($i == 0) {
    chomp $psflags;
}
print "ps -e -f -a -i -v -o " ;
```

Perl script (2)

```perl
# The last argument in $args will be the config file - the # third dir (eg /m1/voyager/dbname/foo) will tell you the database name
# and so the institution
$sDir = (split ' ', $args[2])
$envs = split ($/); for ($i = 0; $i <= $#$envs; $i++) {
    print $i . "\n";
}
$psflags = "-P -f -a -i -v -o " ;
if ($i == 0) {
    chomp $psflags;
}
print "ps -e -f -a -i -v -o " ;
```

Perl script (3)

```perl
# Record some vague performance stats
# print "Voyager usage totals by database 
" #print "$/sadusage($command)++; 
$cmdusages[$command]++; 
$cmdusage[$command]++; 
```

Analysing the data

Tools already in use for processing logs

- Standard unix tools
- Perl
- Logsurfer

Why we chose Excel as our analysis tool

- Needed to extract the information quickly
- Started by looking at a subset of the log data
- Existing familiarity with Excel

Reprise: the information we use

- We take an hourly snapshot of processes running against each institution’s Voyager database
- We total the hourly count “total (by database)” for each institution
- This gives a weekly figure which we can use as an indicator of system use

What we expected to find (1)

Were the founder members still equal partners?

What we expected to find (2)

Were some universities hammering the system more than others?
Institutional breakdown of process use

Bib count per institution

Academic and special libraries compared

Seasonal patterns

Split hub architecture helped our performance problems

MEANWHILE...
Patterns of server use by different library types

A new price structure

Our old charging model needed to be updated

- Was based on Endeavor’s pricing model at the point when a library joined SDLC
- Took no account of system use or relative size of an institution in the Consortium

Indicators we use to compare Consortium members

- database size (number of bibs)
- number of staff user licenses
- peak number of connections to the OPAC
- number of Voyager server processes by database (i.e. library)

“% of Consortium” expresses each library’s relative size in the Consortium as a percentage which can be applied to the total annual running costs of the Consortium across both hubs.

“% of Consortium” is applied to

- Sun server costs
- Facilities management
- Consortium support costs
- Consortium management costs
New charging system in place for Financial Year 2005-6

- Consortium Board approval
- Member agreement
- Income remained the same
- Costs more fairly distributed

The future …

- Refine the script
- Better tools for analysing the data
- More refined analysis
- Costing model subject to further revision

Questions?

Thank you!

Clare Whittaker
clare.whittaker@ed.ac.uk
http://www.nls.uk/sdlc