Clustered Logging with mod_log_spread

Theo Schlossmagle Sjesus@omniti.com



Theo Schlossmagle

Pricipal @ Omni TJ

open-source developer
mod_backhand
Wackamole
daíquírí
OpenSSH/SecurID
Spread
etc.

closed-source developer
 Ecelerity MTA
 Ecelerity Clustering

omniti



•Understanding the Problem Space

•A Survey of Technologies

Implementing Clustered Logging

Understanding New Possibilities

Enderstanding the Problem Space

The Purpose of Logging

•Journalling the fact that a transaction has taken place.

·Correlating a series of transactions into a session.

•An audit trail.

•Forensics.

 Activity analysis to understand current trends and predict the future.

Basic Expectations

Logs are reliable.

Events are logged in the order they occur. They can be partitioned by date.

They can be multiplexed and demultiplexed on demand.

Introducing Clustering

•Clustering: several machines acting together to provide a single service •Sessions may now be composed of a series transactions that occur on different machines. •Ordering is "harder" and more important.

Billiorporte inder france in an order of the france of the part of the second presence of t



participante propares the spen strume : there in the spen strume to be and the strume to be and the form the spen strume in the spen strume is the spen strume to be and the spen strume to be and the spen strume to be an allow the spen spent is an allow the spent of the spent of

trate are lover and hards query of lar shoes



Braditional Logging

Logs written locally on web servers
space must be allocated

Consolidation happens periodically
crashes will result in missing data
aggregators must preserve chronology
real-time metrics cannot be calculated

Monitors must run against log servers
monitors must tail log files
requires resources on the log servers



storage

Gick-stream løgger

Active Metwork Logging

Logs written directly to log servers
UDP is unreliable and thus not useful
TCP is a point-to-point protocol
Two log server mean double traffic
Add a monitor and that's triple!

Real-time metrics are possible
 monitors must tail log files still

 (or publishers must send directly to the monitors... yuck!)

Network Approach



Passive Network Logging

- Logs constructed from sniffed traffic
 The players no longer matter
 Web servers can be added easily
- Drops logs!
 When tested head-to-head with active logging frameworks we see loss
 Missing logs is unacceptable

Passive Logging

and from an Saute Mina offers Art o 4- ADACA ANDANAGA allow all the 101.0. A.A. omi stol HIND . VLUMA aporte change . o. Anto 62+ volke - : (PARO su BELL MULLE 30 Y.+ (7. son agod labod ?? MUMPERANTERS and of the form mod 1-5 in Morataro - ASSIARA

> A form differenter eftert ener barn ets ang hac belegier: B fora for forgage ichokarabi pibma accio che tariagi cegue rime tinga etter tarte un enamptara espiratori pibma accio che tariagi cegue rime tinga

A lapse in judgement

undred durfue . w. wash .

mod_log_spread Eogging

Logs are published over Spread
Efficient reliable network multicast
Preserves global ordering of logs

Multiple subscribers at no cost
well... almost zero

Extends well beyond Apache
 All logging (enterprise wide) can be utilize this publish/subscribe messaging bus

mod_log_spread



Clustered Logs Provide

ínstant aggregation
ordering
publish/subscríbe model
multiple subscríbers
multiple subscríbers
multiple subscríbers...

Multiple Subscriber Magic

 Data "feeds" •Write them to disk •Real-time analysis: popular pages •concurrent sessions •Who's online? •Understand load-balanced click streams

Implementing Olustered





•Spread •Apache 1.3 or 2.0 *mod_log_spread spreadlogd •A spread client API for your favorite language: •Perl, Python, C Java, Ruby, PHP, etc.

Install Spread http://www.spread.org/

A simple /etc/spread.conf:

DebugFlags = { EXIT CONFIGURATION }

EventLogFile = /var/log/spread/mainlog EventTimeStamp

10.225.209.255:4913 {
10.225.209.68
10.225.209.71
10.225.209.72
10.225.209.73
10.225.209.240
10.225.209.241

order matters
staging server

logging machines
monitoring machine

אומי אלגף יולט שור אותי באלי איר איראי איראי אומי אלגי איראי איר איראי באילי באיר אומי איר אאר איר מראי מראיר באביוי יולי איראי איראי מראיר איראי יולי איראי איראי איראי איראי אור איראי איראי איראי איראי יולי איראי איראי איראי איראי אור איראי איראי איראי איראי אור איראי איראי איראי איראי אוראי איראי איראי איראי אוראי איראי איראי איראי אוראי איראי איראי איראי אוראי איראי איראי אוראי איראי איראי אוראי איראי איראי איראי אוראי איראי איראי אוראי איראי איראי

> (Arota M. H. H. H. A. H. A. J. . (Arota H. H. H. H. A. J. . (Arota F. H. H. H. J. . (Arota F. H. H. H. J. . (Arota F. H. H. H. J. . (Arota F. H. J. . J. . (Arota F. H. J. J. . . (Croch M. A Gittogalored)

A.S. & obul . 8 pope de la . de .

Labort of a merico : le pole

[alado for the de lotal and

Install mod_log_spread http://www.backhand.org/

A simple httpd.conf:

LoadModule log_spread_module libexec/mod_log_spread.so AddModule mod_log_spread.c #AddModule mod_log_config.c SpreadDaemon 4913

LogFormat "%h %l %u %t \"%r\" %>s %b" common

VirtualHost slicksiteip:80> CustomLog \$slicksite common Contra proposition Never informer informer peder in anothe proposition in anothe proposition int anothe proposition

- 93 1 34 4

Anner . .

nom mai

P 2.5053

MANIERS

1 map

Verify it is working

; /opt/spread/bin/spuser -s 4913

User: connected to 4913 with private group #user#admin-va-1 User> j coolsite

Received REGULAR membership for group coolsite with 2 members, where I am member 1:

#user#admin-va-1 grp id is 182571332 1092928408 2

Due to the JOIN of #user#admin-va-1

User>

received RELIABLE message from #ap25454#admin-va-1, of type 1, (endian 0) to 1 groups (182 bytes): 68.55.183.91 - - (30/Oct/2004:11:48:51 -0400) "GET /~jesus/ HTTP/1.1" 200 57940 "-" "Mozilla/5.0 (Macintosh; U; PPC Mac OS X; en) AppleWebKit/125.5 (KHTML, like Gecko) Safari/125.9"

Install spreadlogd http://www.backhand.org/mod_log_spread/

A simple /etc/spreadlogd.conf:

BufferSize = 65536 Spread { Port = 4913 Log { Rewrite Timestamp = CommonLogFormat Group = "coolsite" File = /data/logs/apache/coolsite/common_log Log {

Rewrite Timestamp = CommonLogFormat Group = "slicksite" File = /data/logs/apache/slicksite/combined_log

Mallio . J. Je avollo precis : Proven pro. a destrogalerade pe = B . Hede some moder y . An . han bol . 12 IE . W. A. W. W. A. ST and paper reverse effer vorthe rever gens we mas . server and Lto out a pope ded a labor of

soullier sile source by the pit + for

- 1: A 8. Fu

+ · MOTAL

Spreadlogd: kung-fu(1)

```
BufferSize = 65536
PerILib /opt/spreadlogd/custom
PerIUse mylogger
Spread {
    Port = 4913
    Log {
        RewriteTimestamp = CommonLogFormat
        Group = "coolsite"
        PerILog mylogger::log
        File = /data/logs/apache/coolsite/common_log
    }
```

```
Log {
```

Rewrite Timestamp = CommonLogFormat Group = "slicksite" File = /data/logs/apache/slicksite/combined_log

Spreadlogd: kung-fu(2)

package mylogger;

use DBI; our \$dbh; our \$sth;

```
sub log($$$) {
    my $sender = shift;
    my $group = shift;
    my $message = shift;
    my ($user, $host) = ($sender = /#((^#)+)#((^#)+)/);
    chomp($message);
```

Understanding New Possibilities

un donne manificame persona est service entre entre entre service al ma dependent and present personale entre personal est service est personal de la constructione est personale entre person

Andregianeras Andrean Anton applant and have a there a stratege a the failer of the men all failer and the fame in Stratege to the stratege attended to the stratege a the failer of the of the

most ale prime toto more for trapane. pare fileme frelle grange ale and a file of the need

mit ymtay



•Logs are now streaming in real time •Real-time metrics •per server hit rates (traffic) •per server hits by response code •relative error serving rate •per server document size metrics ·detect unexpected bugs do to anomalous traffic

Track deeper data •user habits •length of visit online

•All this happens passively

Stupid Pet Bricks

0001	mls_mon: B	asic Metrics
_ Metrics		Codes Servers
Requests Bulk	275 1670444	HTTP Codes
Bandwidth	0.000 b/s	304 0.00/s 1 404 1.20/s 13
	74.294 Kb/s	
Spread D	aemon: 4913@10	0.77.52.67 Disconnect
Spread	Group: ardms	Leave
mls_mon: Connected.		

Gredit Where

Credit's Due

The John Hopkins University The Center for Networking and Distributed Systems

Omni TI Computer Consulting

Ehe Authors and Contributors of Spread: Yair Amir, Michal Miskin-Amir, Jonathan Stanton, Christin Nita-Rotaru, Eheo Schlossnagle, Dan Schoenblum, John Schultz, Ryan Caudy, Ben Eaurie, Daniel Rall, Marc Zyngier

The Authors of mod_log_spread and Tools: George Schlossnagle, Theo Schlossnagle, Jonathan Stanton, Yair Amir

