Application Buffer-Cache Management for Performance: Running the World's Largest MRTG

LISA '07, November 14, 2007

David Plonka, Archit Gupta, and Dale Carder
{plonka,archit}@cs.wisc.edu, dwcarder@doit.wisc.edu
Outline

• Background on MRTG and RRDTool
• Problem & Motivation
• Investigative Tools & Methodology
• Approach 1: Application-level Buffering
• Approach 2: Application Advice
• Scalability
• Analysis: Model & Simulation
• Summary and Contributions
• “Questions?”
What is MRTG?

- Daily, 5 min averages
- Weekly, 30 min ave.
- Monthly, 2 hour ave.
- Yearly, 1 day averages
Our Network

UW Madison - Existing fiber
Last Updated: 3/25/06 SJB
What is RRDTOOL?

# MRTG daemon
readConfiguration();
do {
  # 1. POLL TARGETS via SNMP
  readTargets();
  # 2. WRITE TARGETS to RRDs
  foreach my $target (@targets) {
    RRDs::update(...);
  }
  sleep(...); # balance of 5 mins
} while (1); # forever
What's the Problem?
What's the Problem?
What's the Problem?
RRDCache: App-level Buffering

Every 5 minutes, append time & data to journal. RRDCache Library also rotates journal files hourly.

MRTG → RRDCache::Update → RRDCache Library
MRTG → RRDCache::Update → RRDCache Library

Once an hour, flush ramdisk cache to Disk Array.

Ramdisk
rrdcache journal file
rrdcache journal file
rrdcache journal file

RRDCache Library

RRDCachewritter

Disk Array
file_1.rrd
file_2.rrd

Client wants real-time data.
RRDCache Library forces a flush of cached data to disk for that rrd through RRDs::update. Then RRDCache Library calls RRDs::Graph to read data back.

Client wants trending data, doesn’t need last hour.
Client uses rrdtool normally.

gengraph.cgi

rrdttool
RRDCache Performance: Good!

MRTG

CPU

Disk I/O
Methodology & Tools

- **sar**
  - System activity reports (CPU, disk, etc.)
  - Can visualize data with new ksar tool

- **fincore**
  - A user command that exposes the cache “footprint” of a set of files; shows which file blocks are in core

- **fadvise**
  - A user command that can forcibly evict a file's pages from buffer-cache
  - Invaluable for controlled experimentation
160K files x 25 pages x 4KB = ~16GB!

160K files x 4 pages x 4KB = ~3GB!
Simulation: Hot Pages

- AVERAGE data area
- MAX data area
- 5 min
- 30 min
- 2 hour
- 1 day
- End-Of-File

Update Frequency

4KB Block/Page Number

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

Meta-data
What's the Solution?

`fadvise(RANDOM);`
fadvice(RANDOM): Before & After
Analytical Model for Average RRD update Page Fault Rate

\[ r = \frac{m}{t} \]

\[ r = \frac{\delta \sum_{i=1}^{n} \frac{1}{x_i}}{\text{minimum}(uT, x_s T)} \]
fadvise(RANDOM) Before & After

![Graph showing the proportion of RRD files cached before and after the use of fadvise(RANDOM). The graph compares the cumulative distribution function (CDF) with and without fadvise(RANDOM). The CDF with fadvise is shown in green and the CDF without fadvise is shown in red. The x-axis represents the blocks cached (4KB pages in core), and the y-axis represents the proportion of RRD files.]
MRTG Performance: Before
MRTG Performance: After. Great!
Scalability: `fadvise(RANDOM)`

Time of Day (UTC)

Seconds Elapsed

Number of Targets (RRD Files)

- total time (poll + write)
- poll time
Scalability: `fadvise(RANDOM)`

![Graph showing CPU utilization over time and number of targets.](image)
Scalability: RRDCache + `fadvise`
Scalability: RRDCache + fadvise
Contributions

- A **method and user tools** to examine buffer-cache and readahead behaviors:
  - `fincore`
  - `fadvise`

- An **analytical model** and simulation of page fault behavior for RRD files:
  - Useful to size memory and to determine system capacity

- **RRDTool performance optimizations**:
  - Application-level buffering: RRDCache
  - Application advice: `fadvise(RANDOM)`
  - Result: approx. triples system capacity
Thanks!

- Thanks to:
  - Hideko Mills
  - Robert Plankers
  - Kevin Kettner II
  - Michael Swift
  - Tobi Oetiker

  - Released rrdtool-1.2.24 yesterday (Nov 13, 2007) with our `fadvise(RANDOM)` patch (Patches available for 1.0.49 and 1.2.23: See “Conclusions” in paper.)

- “Questions?”