

# Scaling RRDtool without breaking the Bank

Tobias Oetiker

OETIKER+PARTNER AG

SwiNOG 18

# about rrdtool

- ▶ a database with controlled memory loss
- ▶ data is only an approximation of reality
- ▶ quick setup and use
- ▶ graphics included

# designed for speed

10:45:00

| HEADER                      |
|-----------------------------|
| RRA 0<br>5 Minutes Average  |
|                             |
| RRA 1<br>60 Minutes Average |
|                             |
| RRA 2<br>60 Minutes Maximum |
|                             |
| RRA 3<br>10 Minutes Average |
|                             |

10:50:00

| HEADER                      |
|-----------------------------|
| RRA 0<br>5 Minutes Average  |
|                             |
| RRA 1<br>60 Minutes Average |
|                             |
| RRA 2<br>60 Minutes Maximum |
|                             |
| RRA 3<br>10 Minutes Average |
|                             |

10:55:00

| HEADER                      |
|-----------------------------|
| RRA 0<br>5 Minutes Average  |
|                             |
| RRA 1<br>60 Minutes Average |
|                             |
| RRA 2<br>60 Minutes Maximum |
|                             |
| RRA 3<br>10 Minutes Average |
|                             |

11:00:00

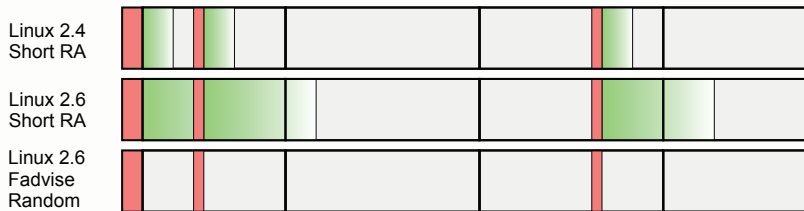
| HEADER                      |
|-----------------------------|
| RRA 0<br>5 Minutes Average  |
|                             |
| RRA 1<br>60 Minutes Average |
|                             |
| RRA 2<br>60 Minutes Maximum |
|                             |
| RRA 3<br>10 Minutes Average |
|                             |

# diy performance improvements

- ▶ rrd is disk bound
- ▶ striping
- ▶ nvram
- ▶ tmpfs and copy

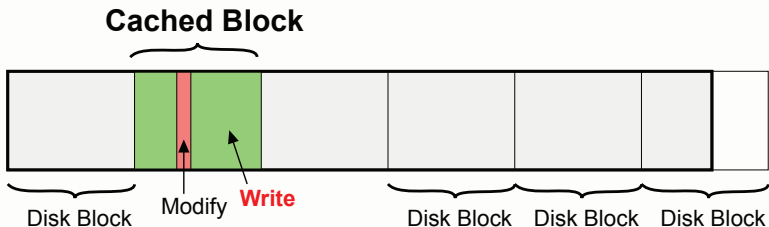
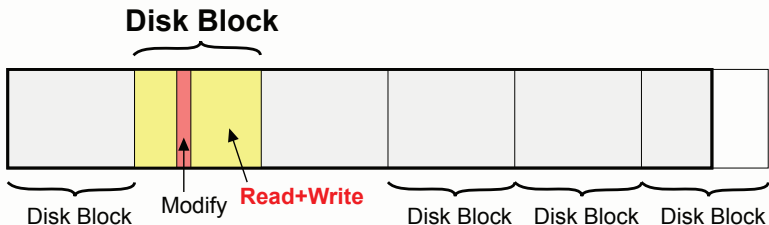
# scientific approach

- ▶ Dave Plonka, LISA'07
- ▶ vm optimization kills rrd
- ▶ fix read-ahead with fadvise random
- ▶ batch updates of a single rrd
- ▶ running 320k RRD updates every 5 Minutes



# buffer cache is king

- ▶ empirical evidence shows cache helps
- ▶ no-cache: double the work
- ▶ worse: writes gets blocked by read



# active buffer cache management with fadvise

- ▶ fadvise and madvise
- ▶ RANDOM - no read ahead
- ▶ DONTNEED - drop synced data from cache
- ▶ highly implementation dependent
- ▶ linux  $\geq$  2.6.18 gets it right
- ▶ keep only the hot blocks

**Next Data Access**

**Hot Blocks**

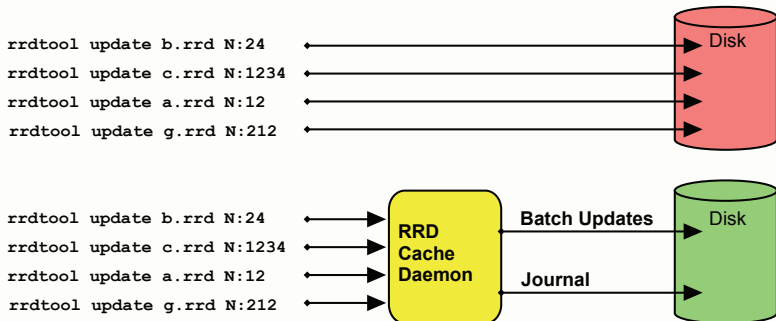


# features of RRDtool 1.3

- ▶ active buffer cache management with fadvise and madvise
- ▶ memory mapped io (bernhard fischer)
- ▶ holt winters with moving baseline (evan miller)
- ▶ cairo and pango for graphics
- ▶ pdf, svg, eps output
- ▶ inline text formatting
- ▶ anti-aliasing controllable

# rrd cache daemon

- ▶ multiple updates to the same rrd are the fastest
- ▶ cache daemon batches updates
- ▶ journal replay for crash case
- ▶ remote updates (no auth)



# features of RRDtool 1.4

- ▶ rrd cache daemon (florian forster and kevin brintnall)
- ▶ libdbi integration for instant db access (martin sperl)
- ▶ graph prediction functions (martin sperl)
- ▶ graph legend placement (melchior rabe)
- ▶ inline text formatting

beta in April.

# features of RRDtool 1.5

- ▶ portable data format
- ▶ remote graphing through rrd cache daemon
- ▶ ...

?

Tobi Oetiker <tobi@oetiker.ch>