CS615 - Aspects of System Administration

Monitoring

Department of Computer Science
Stevens Institute of Technology
Jan Schaumann
jschauma@stevens.edu
https://www.cs.stevens.edu/~jschauma/615/

Problem Report

"Something's wrong."

Now what?

Problem Report

"The system feels slow."

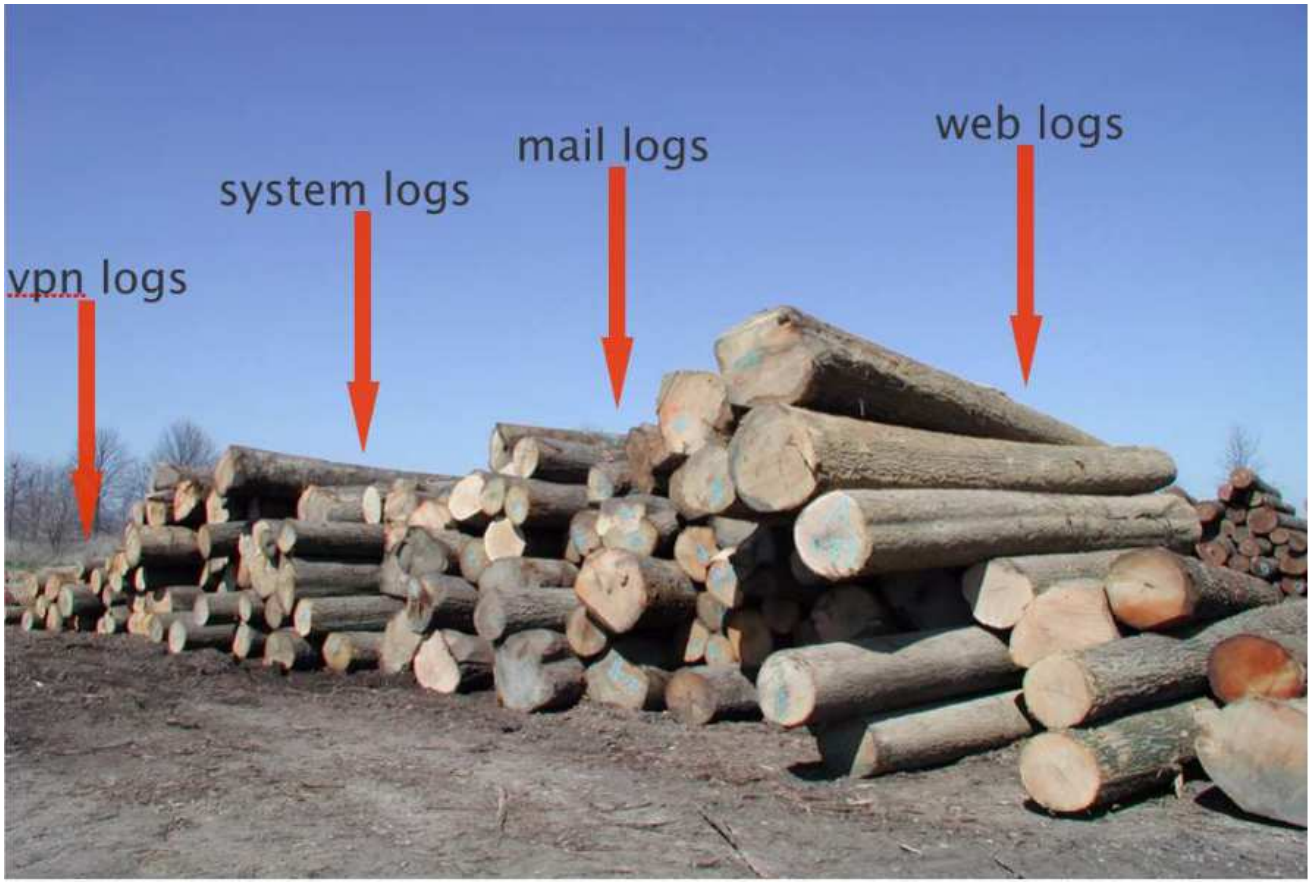
"I can't log in."

"My mail was not delivered."

"The site is down."

Now what?

To the logs!



Answers

"The system feels slow."

```
up 1318 days, 13:46, 1 user, load averages: 993.81, 272.91, 1012.18}
```

"I can't log in."

```
Apr 6 09:25:56 <auth.info>hostname sshd[1624]: Failed password for jdoe from
115.239.231.100 port 1047 ssh2}
```

"My mail was not delivered."

```
Apr 11 16:15:40 panix postfix/smtpd[7566]: connect from unknown[122.3.68.122]
Apr 11 16:15:41 panix postfix/smtpd[7566]: NOQUEUE: reject_warning: RCPT from
unknown[122.3.68.122]: 450 4.7.1 Client host rejected: cannot find your hostname,
[122.3.68.122]; from=<McneilRomany28@pldt.net> to=<jschauma@stevens.edu>
proto=ESMTP helo=<122.3.68.122.pldt.net>
```

Answers

“The site is down.”

```
94.242.252.41 - "" [11/Apr/2016:19:18:47 -0400] "GET /secret/ HTTP/1.1"
403 524 "-" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.9; rv:28.0)
Gecko/20100101 Firefox/28.0"
```

Answers

“The site is down.”

```
94.242.252.41 - "" [11/Apr/2016:19:18:47 -0400] "GET /secret/ HTTP/1.1"
403 524 "-" "Mozilla/5.0 (Macintosh; Intel Mac OS X 10.9; rv:28.0)
Gecko/20100101 Firefox/28.0"
```

Events

"Something's wrong." is just an *unexpected* or *undesirable* event.

Events

"Something's wrong." is just an *unexpected* or *undesirable* event.

Events happen all the time.

Events

"Something's wrong." is just an *unexpected* or *undesirable* event.

Events happen all the time.

Being able to identify *relevant* events allows you to diagnose, predict and even prevent *undesirable* events.

Events

In order to be able to identify an event as *unexpected*, you have to have *expected* events.

Expected Events

Know your applications.

Expected Events

Know your applications.

Know your users.

Expected Events

Know your applications.

Know your users.

Know your traffic patterns.

Expected Events

Know your applications.

Know your users.

Know your traffic patterns.

Know your systems.

Events and Metrics

```
$ dict event
  event
      n 1: something that happens at a given place and time
      2: a special set of circumstances; "in that event, the first
         possibility is excluded"; "it may rain in which case the
         picnic will be canceled" [syn: {event}, {case}]


$ dict metric
  metric
      3: a system of related measures that facilitates the
         quantification of some particular characteristic [syn:
         {system of measurement}, {metric}]
```

Events and Metrics

Event

Metric

You

Events and Metrics

Events

- may occur rarely / frequently / constantly
- can be collected in logs
- may be comprised of other events
- may be: 'something happened'
- may be: 'nothing (new) happened'

Metrics:

- correlation of related events
- may help identify outliers
- may trigger events
- may help make (automated or interactive) decisions

Collecting Data

Counters: easy, numeric data tracking individual events. Example: HTTP status codes

Timers: easy, numeric data tracking event duration. Example: Time to send all data for a successful HTTP request.

Thresholds: easy, numeric trigger for events; may itself trigger events or metrics. Example: more than N HTTP hits in X seconds yield 404.

Know Your Systems

Profile your application:

- execution time (for example: `time(1)`)
- data sources and destination affect execution
- `strace(1)` and friends for more detailed analysis

Understand your system performance:

- CPU load, memory (for example: `top(1)`, `vmstat(1)`)
- disk I/O (for example: `iostat(1)`)
- user activity (for example: `ac(1)`, `lsof(8)`, `sa(8)`)

Know Your Systems

Network statistics:

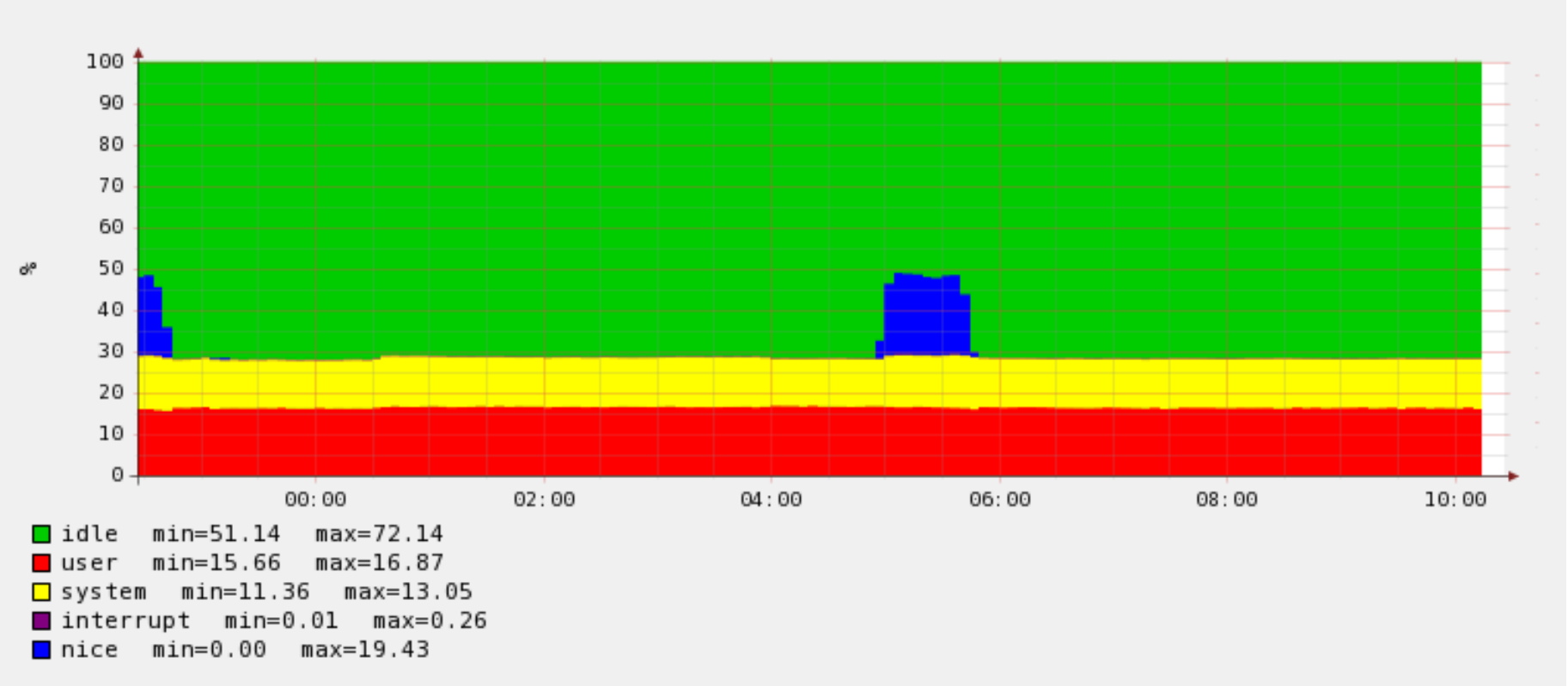
- ports and applications (for example: `lsof(8)`, `netstat(8)`)
- packets in and out
- connection origin
- *NetFlow* etc.

Context

Context lets you find *relevant* events in your haystack of metrics.

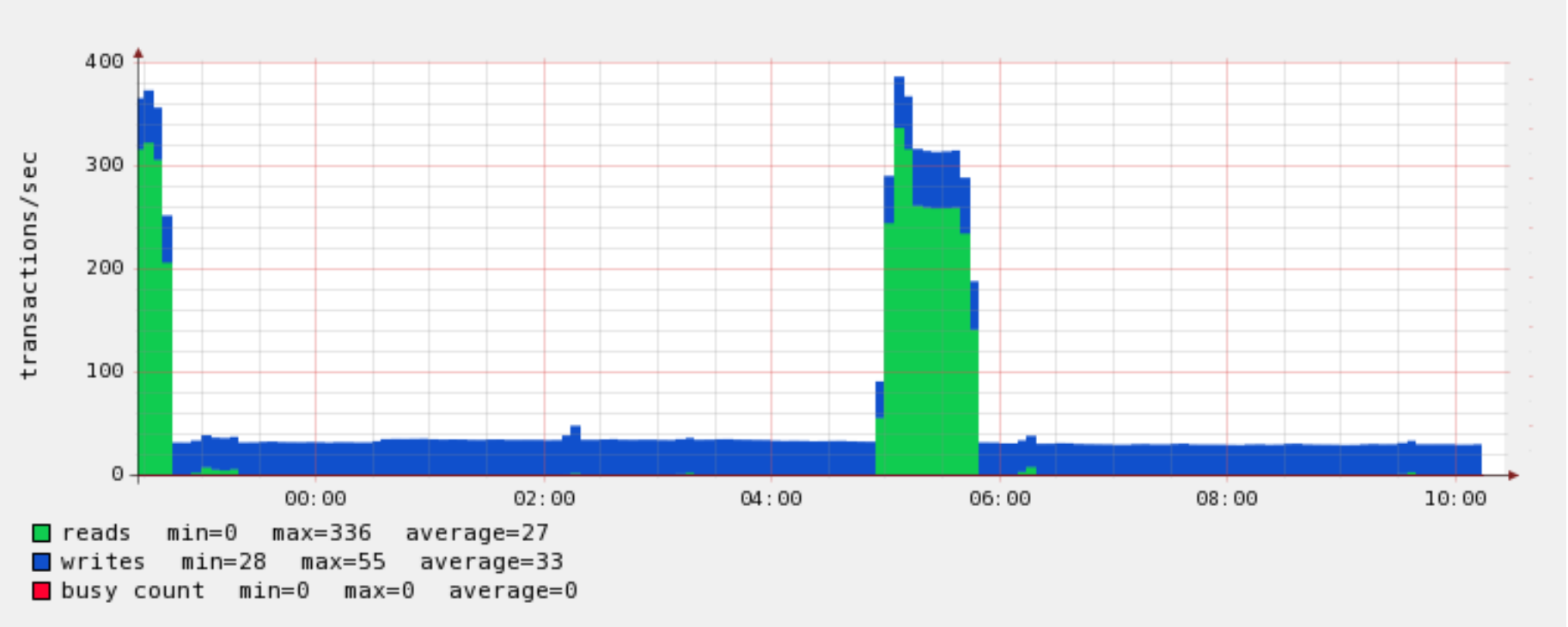
No context.

CPU load - 12 hours



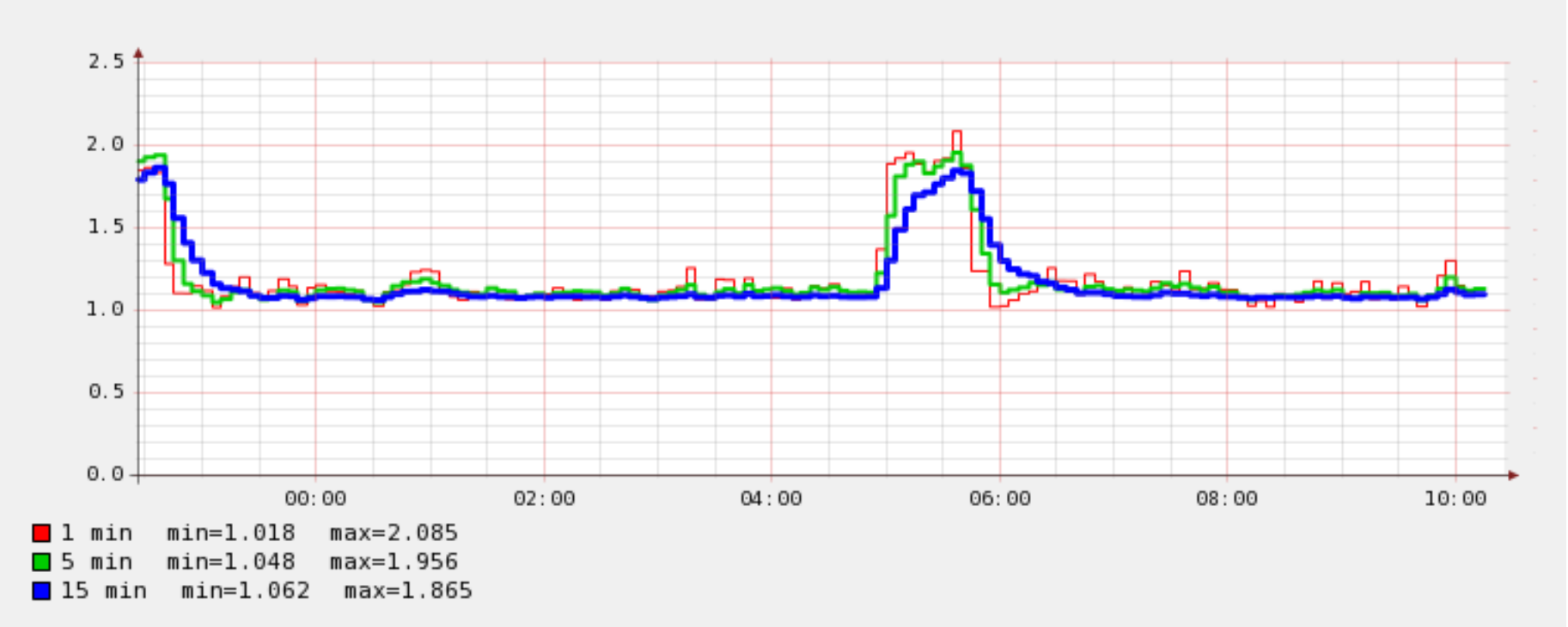
No context.

Disk I/O - 12 hours



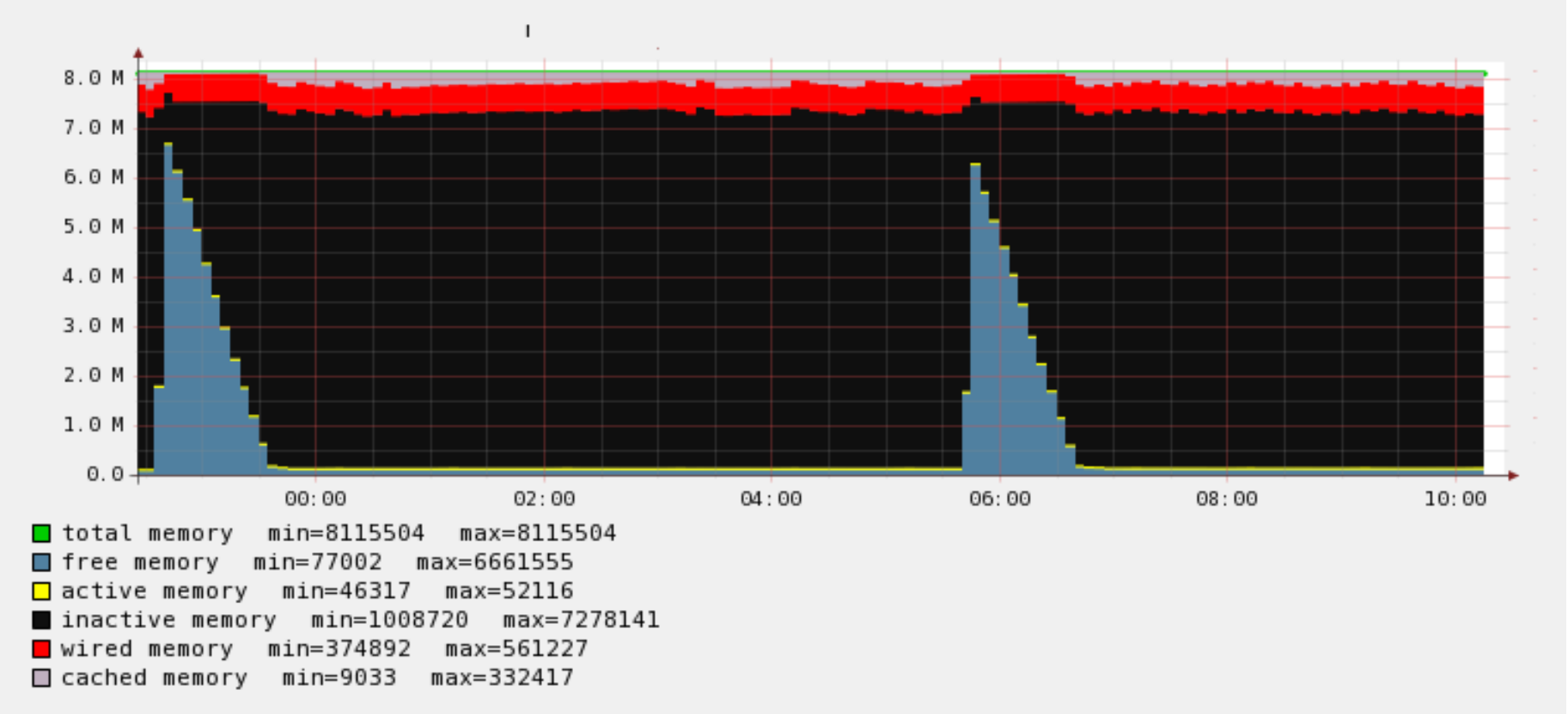
No context.

Load Average - 12 hours



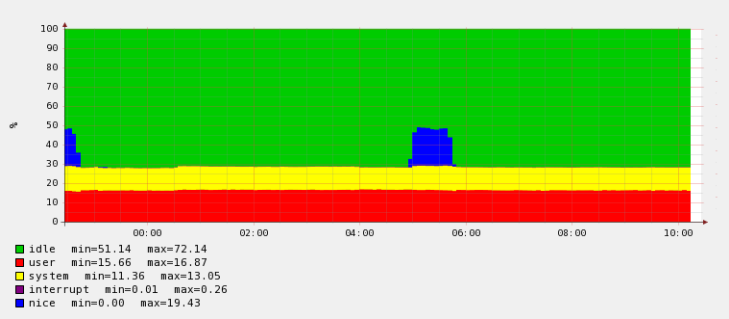
No context.

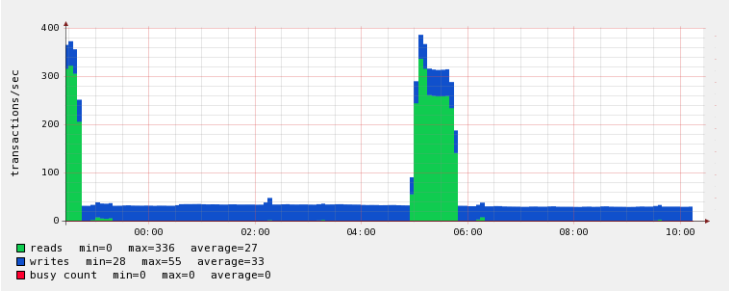
Memory - 12 hours

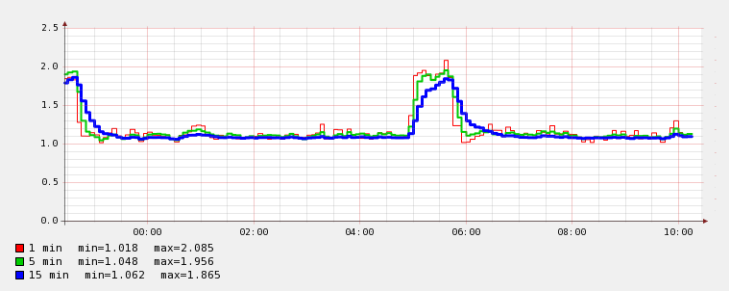


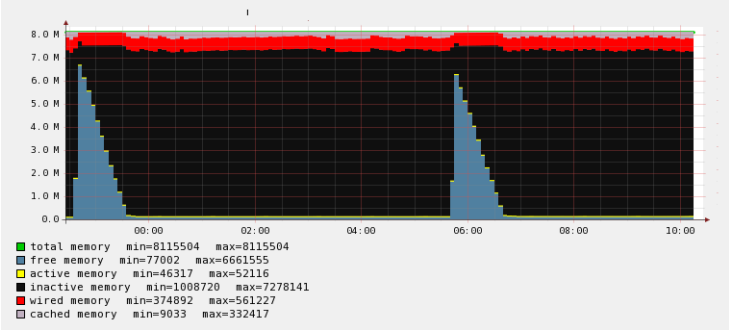
Some context.

12 hours



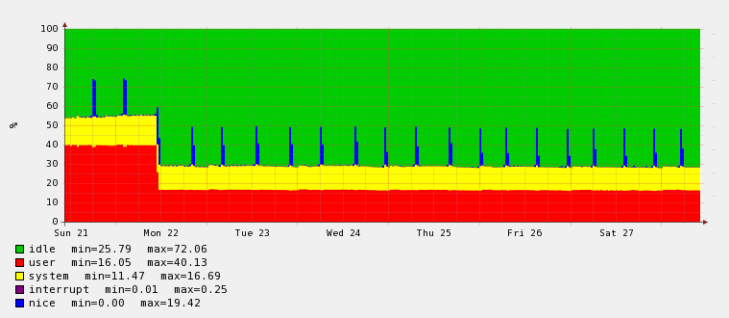


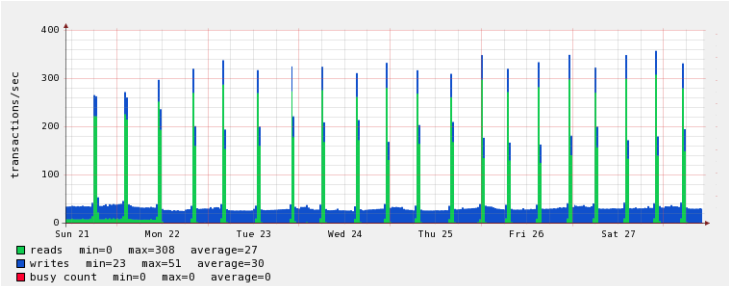


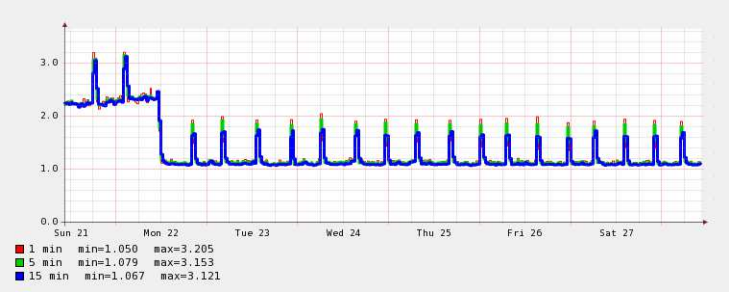


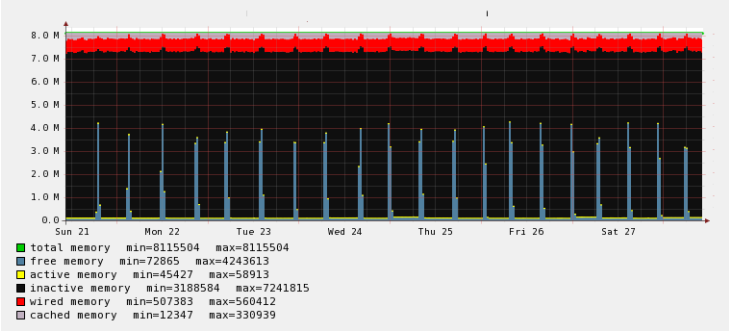
With context.

7 days



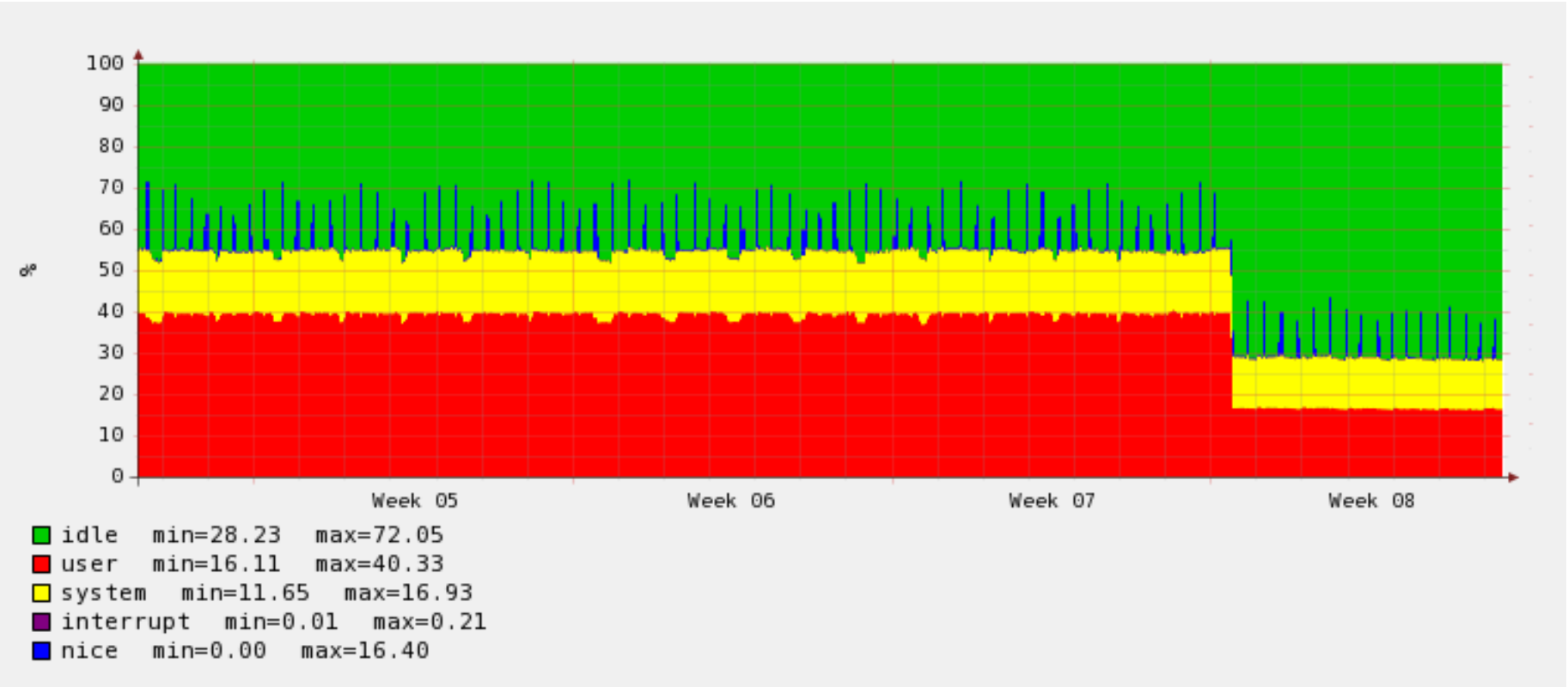






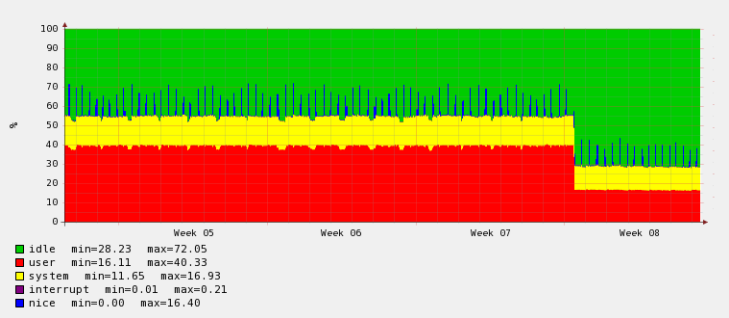
Know your systems.

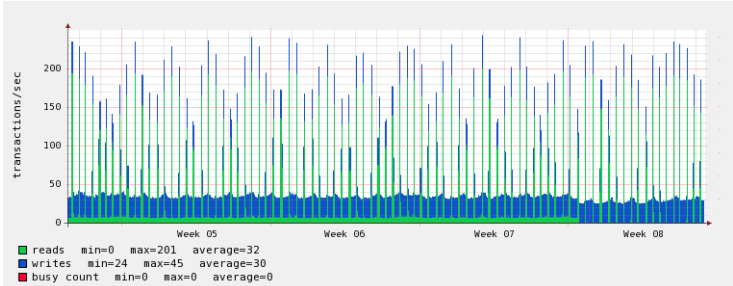
CPU load - 30 days

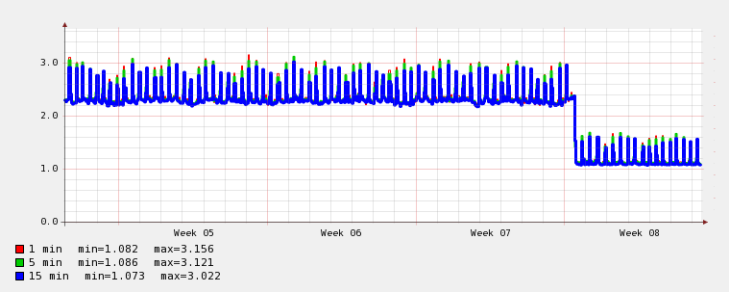


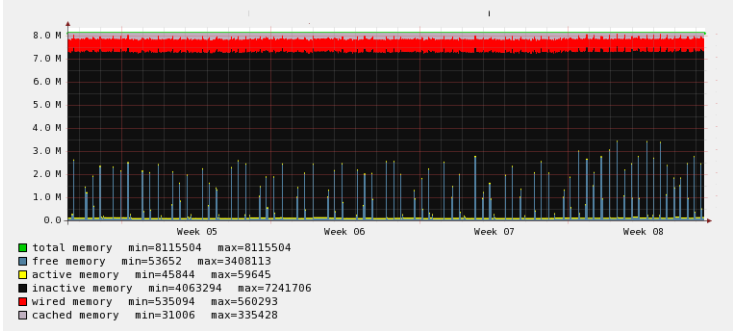
Know your systems.

30 days









Turn *events* into *metrics.*

- Log it!

- Export counters/timers from within your application.
- Process logs and produce counters/timers:

```
awk '{print $9}' /var/log/httpd/access.log | sort | uniq -c
```

- Graph it.
  `http://shouldigraphit.com/`

Monitoring/graphing

SNMP based:

- Cacti: `http://www.cacti.net/`
- MRTG: `http://oss.oetiker.ch/mrtg/`
- Observium: `http://demo.observium.org/`
- ...

Other / complementary:

- Ganglia: `http://monitor.millennium.berkeley.edu/`
- Munin: `http://munin.ping.uio.no/`
- Nagios: `http://nagioscore.demos.nagios.com/`
- Graphite: `http://graphite.wikidot.com/`

To the cloud!

There's a service for that. In the cloud.

Consider:

- support / convenience vs. do-it-yourself
- integration with your other services
- data confidentiality
- data lock-in (esp. when trending data over years)

Monitoring Pitfalls

Increasing the size of your haystack does not always help in finding the needle.

Monitoring Pitfalls

Increasing the size of your haystack does not always help in finding the needle.

Email is not a scalable network monitoring solution.

Monitoring Pitfalls

Increasing the size of your haystack does not always help in finding the needle.

Email is not a scalable network monitoring solution.

Absence of a signal can itself be a signal.

Monitoring Pitfalls

Increasing the size of your haystack does not always help in finding the needle.

Email is not a scalable network monitoring solution.

Absence of a signal can itself be a signal.

This list is incomplete.

Reading

Monitoring:

- https://www.slac.stanford.edu/xorg/nmtf/nmtf-tools.html
- http://www.datadoghq.com/
- https://www.newrelic.com/
- http://logstash.net/
- http://www.splunk.com/