Less Alarming Alerts

/ Robert Treat

Former

WebDev SysAdmin DBA

I have now been promoted to where I can do the least damage

Now

CEO @OMNITI

Who Cares What Some Suite Thinks?

Phantom Pages

Benny

MyFirstPager

Multiple Rotations

always available, phone only

no pager for years

Phantom Pages

Hello / *@robtreat2*

I manage the SRE team at OmniTI

we manage multiple sites
24x7
millions of users

(omniti.com/is/hiring)

OmniTI

paging is useful

"broken systems should not be just another day at the office"

-- me

paging is useful

Who has ever gotten an alert and ignored it?

(/me looks at alert, says “oh, it’ll probably recover, no need to look further”)

paging is useful

How many alerts were received in the past week that were not actionable?

(no human action was required)

Why God Why?

paging CAN BE useful

Can We Fix It?

how to improve?

Can We Fix It?

hello@omniti.com

we offer operationally focused services to help build and manage your infrastructure

:-)

Terms

- **Metrics**
  - **(anything which can be measured)**

Terms

- Metrics
  - (anything which can be measured)
- **Graphs**
  - **(trending systems)**

Terms

- Metrics
  - (anything which can be measured)
- Graphs
  - (trending systems)
- **Notices**
  - **(notification of event; email)**

Terms

- Metrics
  - (anything which can be measured)
- Graphs
  - (trending systems)
- Notices
  - (notification of event; email)
- **ALERTS**
  - **(wake’n you up; pages)**

Terms

- Metrics
  - (anything which can be measured)
- Graphs
  - (trending systems)
- Notices
  - (notification of event; email)
- ALERTS
  - (wake’n you up; pages)

If you want to improve
your alerts

use systems thinking to reason about your
“system”

alerts should be seen as evidence that your system is behaving in a way outside of your existing understanding

If you want to improve your alerts

think in terms your business can get on board with

for every alert you receive

What is the business impact of this alert?

for every alert you receive

What is the remediation for this alert?

remediation:

- Summarize the problem
- What was done to solve the problem?
- Who was notified?
- Can this be prevented?

send the answer to these questions
to everyone on the team
every time

link to this documentation
from your alerting system

Onward and Upward

- Knowledge Transfer
- Gaps Exposed
- Patterns will emerge

you might be a bad alert

- cannot determine business impact
- no remediation necessary
- no one needs to be told
- work arounds are available

if you can't fix it, you don't need to wake up for it

if it can wait until morning,
you don't need to wake up
for it

Onward and Upward

in case of bad alert

- remove the alert

Onward and Upward

in case of bad alert

- remove the alert
- convert the alert to a notice

Onward and Upward

in case of bad alert

- remove the alert
- convert the alert to a notice
- implement fixes

pro tip:
never let anyone add an alert
unless they can answer these
questions first

Can We Really Do This?

this is partially an organizational issue

Can We Really Do This?

thought exercise:
if you launched a new web site today,
you really only need one alarm

Can We Really Do This?

"I don't care if my servers are on fire,
as long as I am still making money"

-- Kevin, actual OmniTI customer

This sounds good but...

Most SA/SRE types want to be pro-active, not re-active.

ie. they want to alert on leading indicators, not on problems

This sounds good but...

Carrie: I-I'm just making sure we don't get hit again.

Saul: Well, I'm glad someone's looking out for us, Carrie.

Carrie: I'm serious. I-I missed something once before, I won't... I can't let that happen again.

Saul: It was ten years ago. Everyone missed something that day.

Carrie: Yeah, everyone's not me.

site down: monitor was checking 200 response code.

failed to notice absence of response code.

easily fixed, but reactive

"root cause" ==> OOM

why don't we alert on OOM?

OOM does not consistently cause outages

too many false positives leads to ignoring alarms

Digression

Görges M, Markewitz BA, Westenskow DR
Improving Alarm Performance In The Medical Intensive Care Unit Using Delays and Clinical Context
http://www.ncbi.nlm.nih.gov/pubmed/19372334

"In an intensive care unit, alarms are used to call attention to a patient, to alert a change in the patient's physiology, or to warn of a failure in a medical device; however, up to 94% of the alarms are false."

Friendman, Naparstek, Taussing-Rubbo,
Alarmingly Useless, The Case For Banning Car Alarms In NYC
http://transalt.org/files/news/reports/caralarms/report.pdf

Blackstone, Buck, Hakim
Evaluation of alternative policies to combat false emergency calls
http://isc.temple.edu/economics/wkpapers/Pubs/FalsePolicy.pdf

Wickens, Rice, Keller, Hutchins, Hughes, Clayton
False Alerts in Air Traffic Control Conflict Alerting System: Is There A Cry Wolf Effect?
http://www.tc.faa.gov/LOGISTICS/grants/pdf/2007/07-G-002.pdf

OmniTI

AESOP

The Boy Who Cried Wolf

Based On A True Story

- send notice of OOM?
- fix the cause of OOM?
- make a useful alert?

Based On A True Story

useful alerting

- script that checks for OOM
- restart app server when found
- find offending process; kill it
- spin up new node; kill old node

in the event all of these fail, send an alert?

thought exercise:
if you launched a new web site today,
you really only need one alert

In Conclusion

if we need software that runs 24x7, we should design **resiliency** into our software, not *human intervention*

In Conclusion

thinking doesn't scale

especially at 2AM

In Conclusion

thanks!

more:

Surge 2016

http://surge.omniti.com

@robtreat2

@omniti