

# What makes Prometheus a “next generation” monitoring system?



Björn “Beorn” Rabenstein, Production Engineer, SoundCloud Ltd.



1

Sales pitch.

It's how "they" did it...

O'REILLY®



# Site Reliability Engineering

HOW GOOGLE RUNS PRODUCTION SYSTEMS

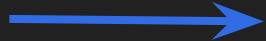
Edited by Betsy Beyer, Chris Jones,  
Jennifer Petoff & Niall Murphy

#GIFEE

Google Infrastructure  
For Everybody Else

# Flow of inspiration – from 1st to 2nd systems

Borg



Borgmon



2005

2015

It's an "industry standard"...



CLOUD N.  
COMPUTI  
FOUNDAT

Consolidate monitoring and alerting

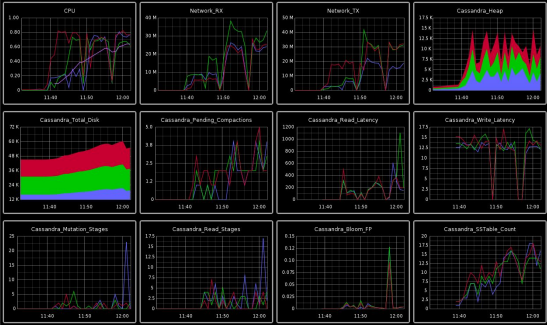
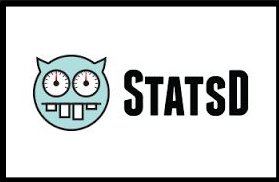


Application

Host (OS, Hardware)

Network

Application



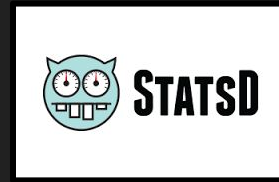
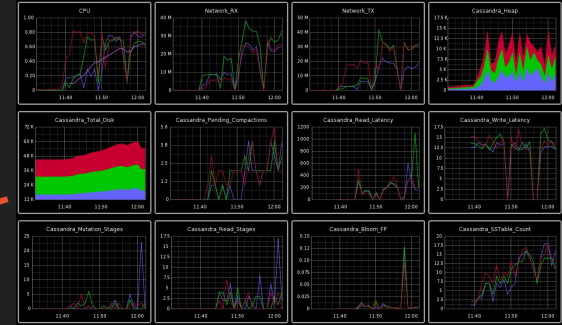
**Nagios®**

Host (OS, Hardware)



Network





**Nagios®**

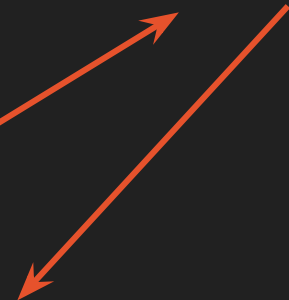
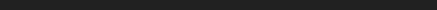
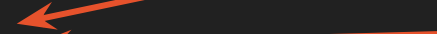
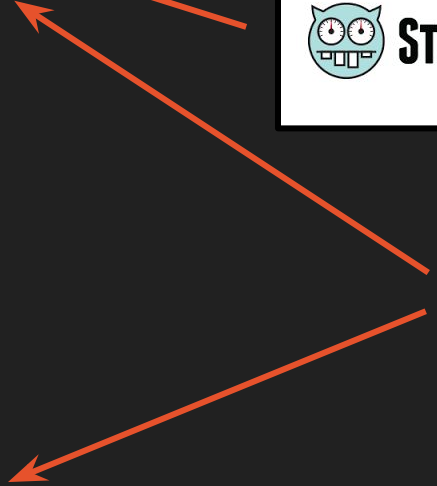


Application

Container

Host (OS, Hardware)

Network



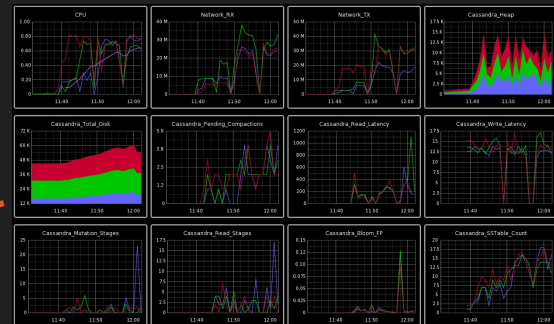
Application

Container

Orchestration 

Host (OS, Hardware)

Network



**Nagios®**



Application

Container

Orchestration

Host (OS, Hardware)

Network



**Nagios**



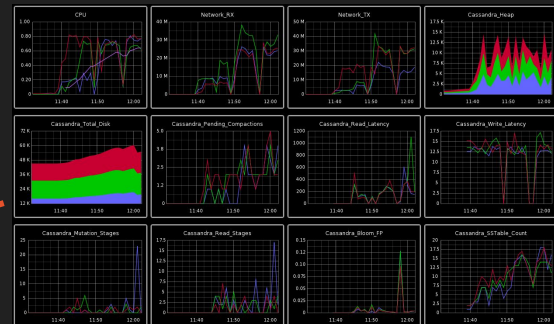
Application

Container

Orchestration

Host (OS, Hardware)

Network



Heapster

**Nagios**



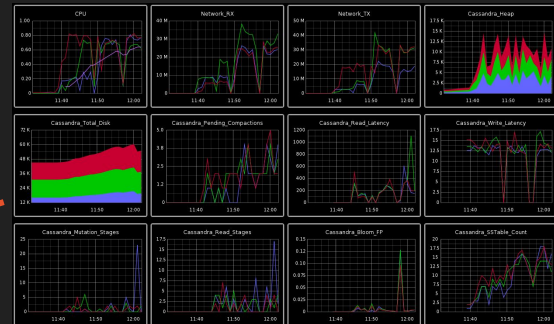
Application

Container

Orchestration

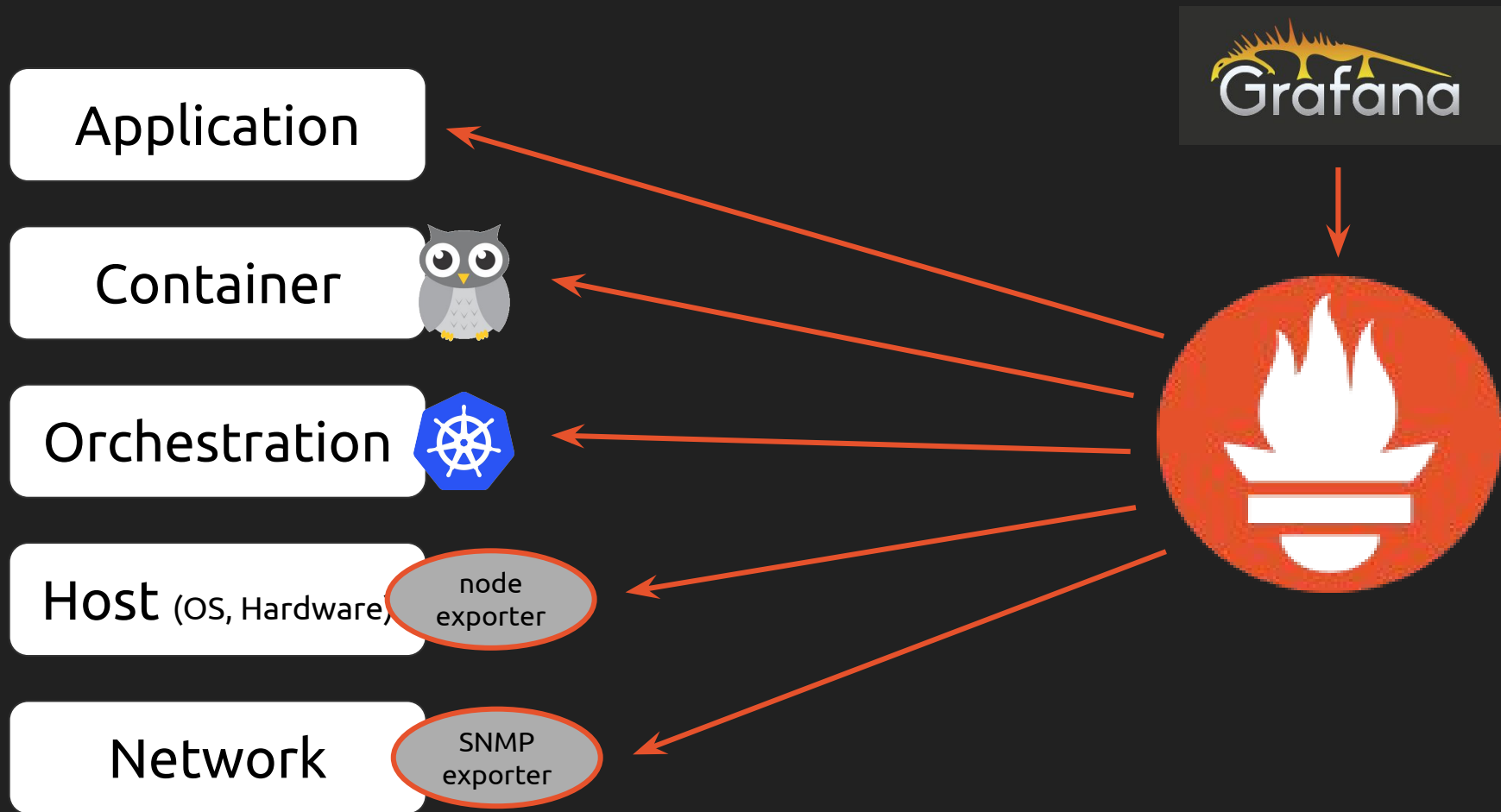
Host (OS, Hardware)

Network



Heapster







Popular against all odds...

## Prometheus - Wikipedia

<https://en.wikipedia.org/wiki/Prometheus> ▼

**Prometheus** is a Titan in Greek mythology, best known as the deity in Greek mythology who was the creator of mankind and its greatest benefactor, who stole fire ...

[Prometheus](#) · [Theft of fire](#) · [Theogony](#) · [Prometheus \(2012 film\)](#)

## Prometheus (2012 film) - Wikipedia

[https://en.wikipedia.org/wiki/Prometheus\\_\(2012\\_film\)](https://en.wikipedia.org/wiki/Prometheus_(2012_film)) ▼

**Prometheus** is a 2012 science fiction film directed by Ridley Scott, written by Jon Spaihts and Damon Lindelof and starring Noomi Rapace, Michael Fassbender, ...

[Ovipositor](#) · [Iridescence](#) · [Damon Lindelof](#) · [Humanoid](#)

## Prometheus (2012) - IMDb

[www.imdb.com/title/tt1446714/](http://www.imdb.com/title/tt1446714/) ▼

★ ★ ★ ★ ★ Rating: 7/10 - 469,170 votes

Adventure - Following clues to the origin of mankind a team journey across the universe and find a structure on a distant moon containing a monolithic statue of a ...

## Prometheus - Greek Mythology

[www.greekmythology.com/Titans/Prometheus/prometheus.html](http://www.greekmythology.com/Titans/Prometheus/prometheus.html) ▼

**Prometheus** was one of the Titans, son of Iapetus (also a Titan) and Clymene, an Oceanid. His brothers were Epimetheus, Atlas and Menoetius. The name d...

## Top stories



It's the match everyone wanted – god of tennis meets Prometheus | Greg Jericho

[The Guardian](#) · 1 day ago



'Arrow' Season 5 Spoilers News: Prometheus' Diabolical Side To Be Displayed

[Christian Post](#) · 4 hours ago



'Prometheus' Deleted Scene Is Way More Brutal Than The Theatrical Cut

[moviepilot.com](#) · 2 days ago

➔ [More for prometheus](#)

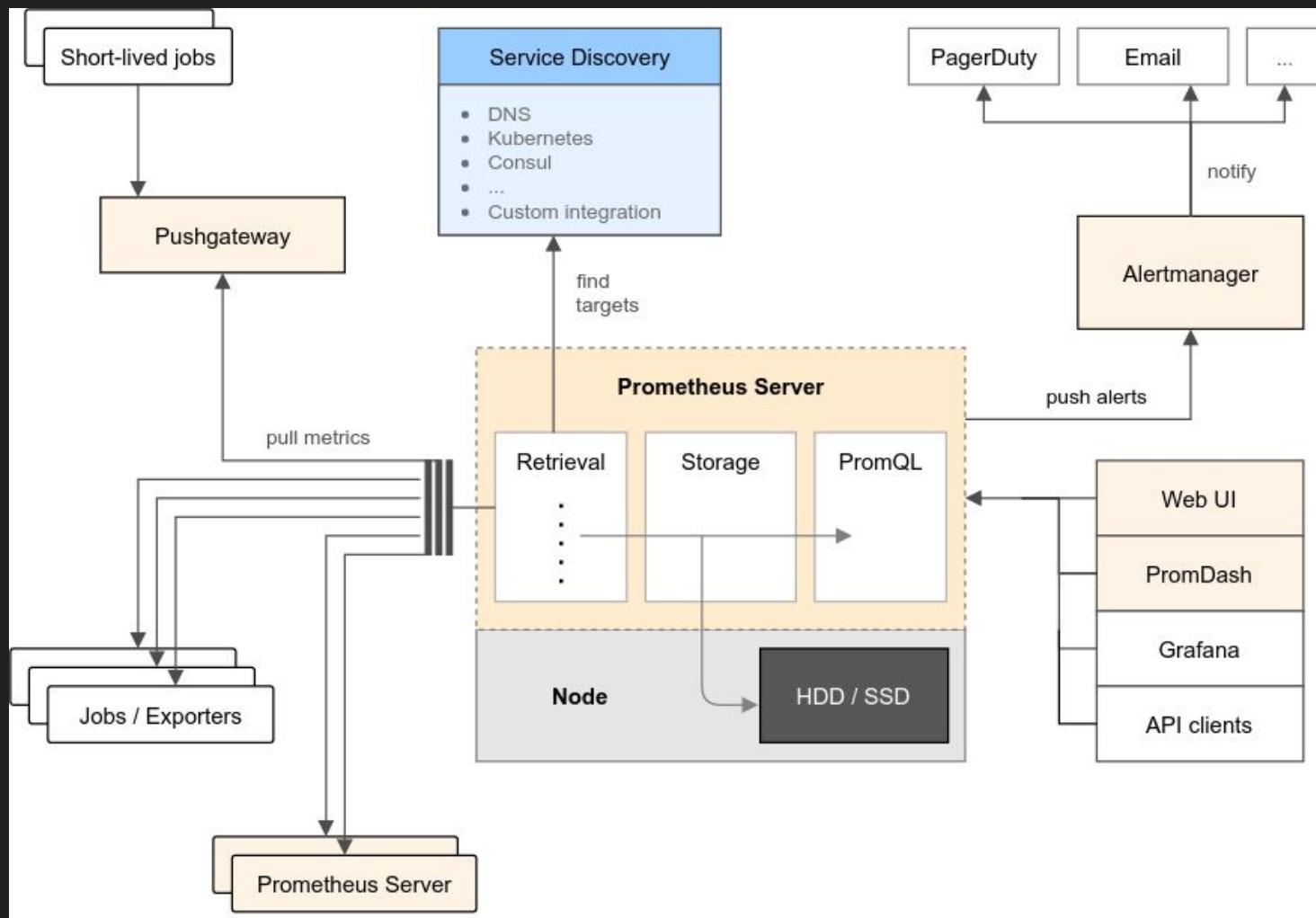
## Prometheus - Monitoring system & time series database

<https://prometheus.io/> ▼

An open-source monitoring system with a dimensional data model, flexible query language, efficient time series database and modern alerting approach.

2

It's a whole ecosystem!



Best 15m introduction talk:  
Brian Brazil at FOSDEM 2016

# Best talk about Promethean philosophy: Julius Volz at PromCon 2016

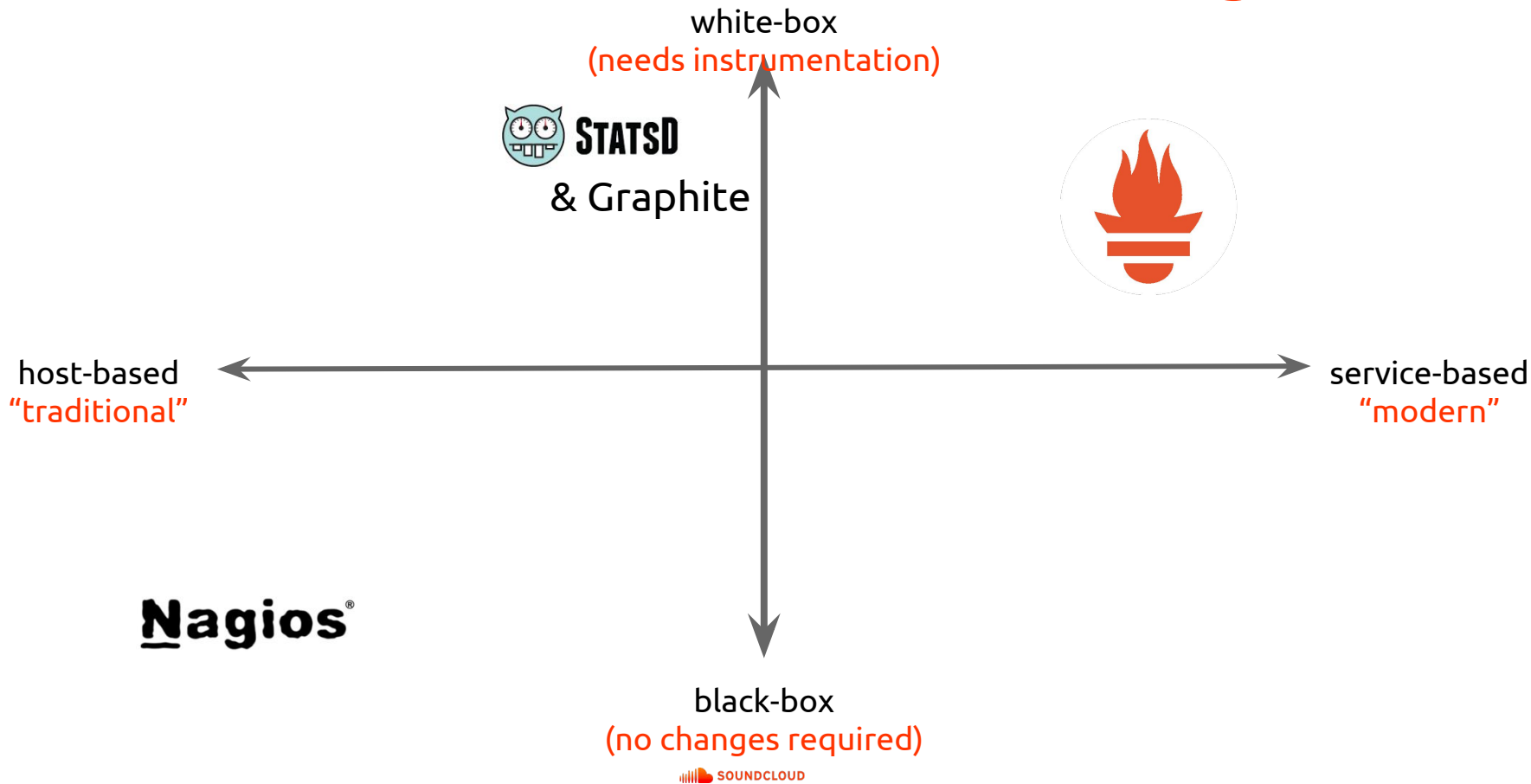
“Prometheus is an opinionated monitoring system that chooses to do a lot of things differently from traditional monitoring systems. This leads to a culture clash for those used to other approaches, and raises questions as to why we didn't take a seemingly better approach.”

3

Blackbox vs. Whitebox.

Host vs. Service.

# Dimensions of monitoring





# Challenges of service-based monitoring

- **Way more targets to monitor.** Scale and semantics
  - A service has many instances.
  - A hosts can run many instances of different services.
- **And they constantly change.**
  - Deploys, rescheduling unhealthy instances ...
- **Need a fleet-wide 3,048m view.**
  - What's my overall 99th percentile latency?
- **Still need to be able to drill down for troubleshooting.**
  - Which instance causes those errors I'm seeing?
- **Meaningful alerting.**
  - Symptom-based alerting for pages, cause-based alerting for warnings.
  - See Rob Ewaschuk's *My philosophy on alerting* <https://goo.gl/2vrpSO>

# Black-box vs. white-box

Black-box is just perfect for symptom-based alerting, isn't it?

We combine *heavy use of white-box* monitoring with *modest but critical uses of black-box* monitoring. The simplest way to think about black-box monitoring versus white-box monitoring is that black-box monitoring is *symptom-oriented and represents active—not predicted—problems*. [...]

For paging, black-box monitoring has the key benefit of forcing discipline to only nag a human when a problem is both *already ongoing and contributing to real symptoms*. On the other hand, for *not-yet-occurring but imminent* problems, black-box monitoring is *fairly useless*.

*Chapter 6: Monitoring Distributed Systems*

## Pros & cons

### Black-box:

- End-to-end test “as the user sees it”.
- Probes may be different from current user traffic.
- Tail latency and rare failures only visible over a long time.

### White-box:

- Reported latency serving the frontend might be a lie, but reported latency of requests to the backend is “live-traffic probing”.
- Must resist temptation to alert on countless internal details.
- Indispensable to detect imminent problems and to investigate causes.

# Imminent problems

White-box and time-series based monitoring FTW.

- Loss of redundancy (going from  $N+1$  to  $N+0$ ).
- More complex reasoning based on insights into a system.
- “Nearly full” scenarios.

[...] the idea of treating time-series data as a data source for generating alerts is now accessible to everyone through those open source tools like Prometheus, Riemann, Heka, and Bosun [...]

*Chapter 10: Practical Alerting from Time-Series Data*

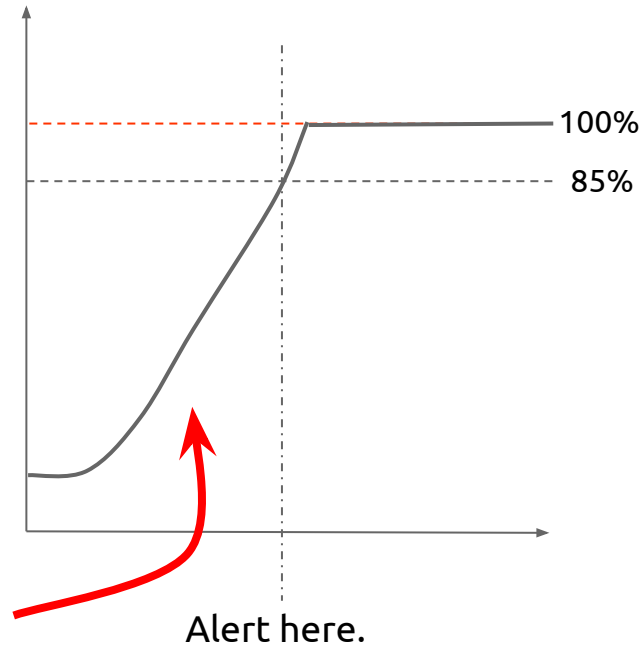
4

Time series.

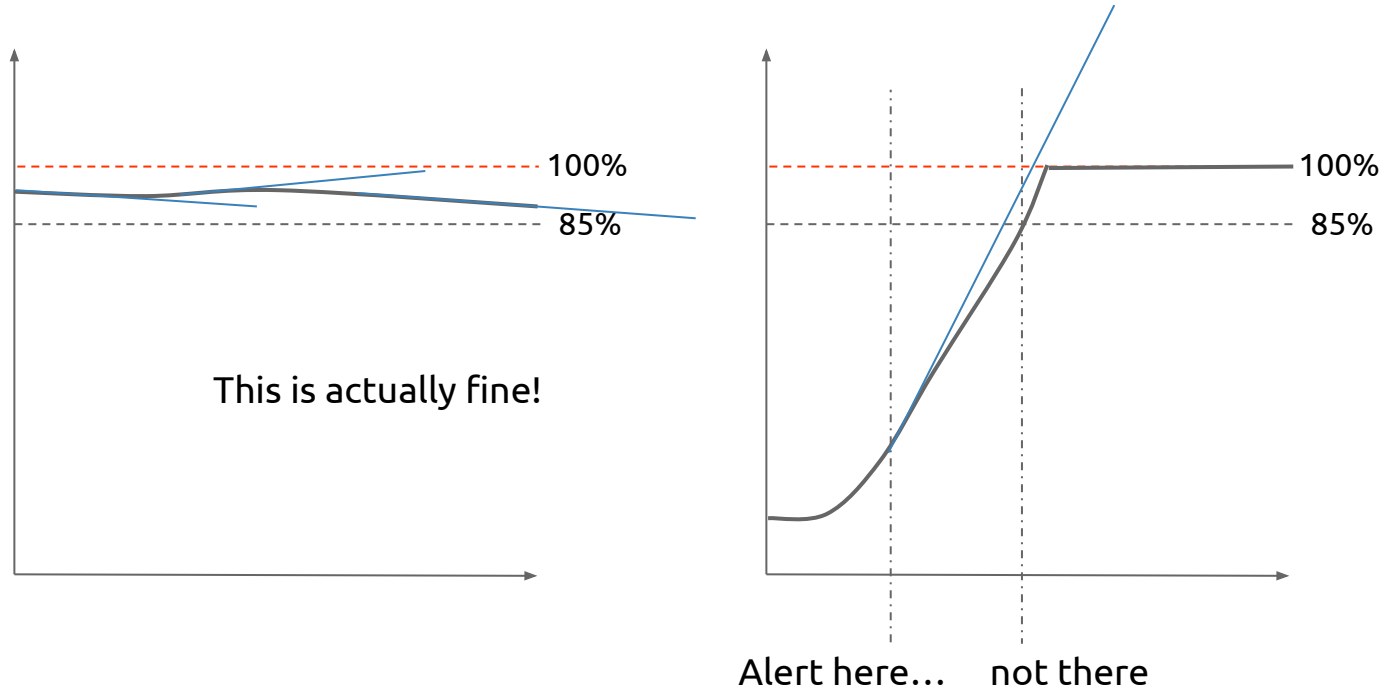
## Static disk-full alert (e.g. Nagios)



This is fine!?!



## Time-series based disk-full alert (e.g. Prometheus)



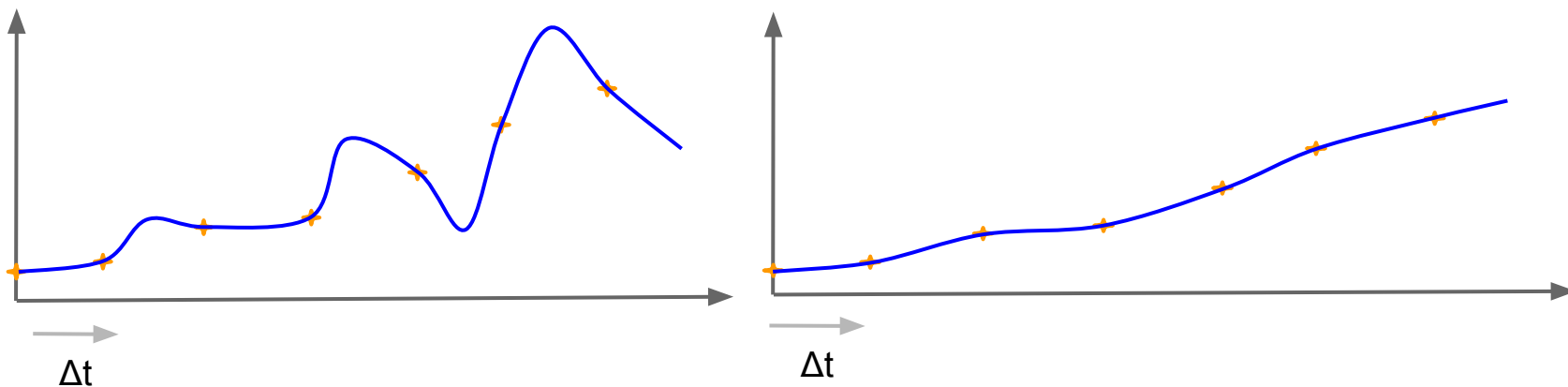
5

Counter vs Gauge.



# Why gauges suck and counters rule

Graphs shamelessly stolen from Jamie Wilkinson




Look up “Nyquist-Shannon sampling theorem”.

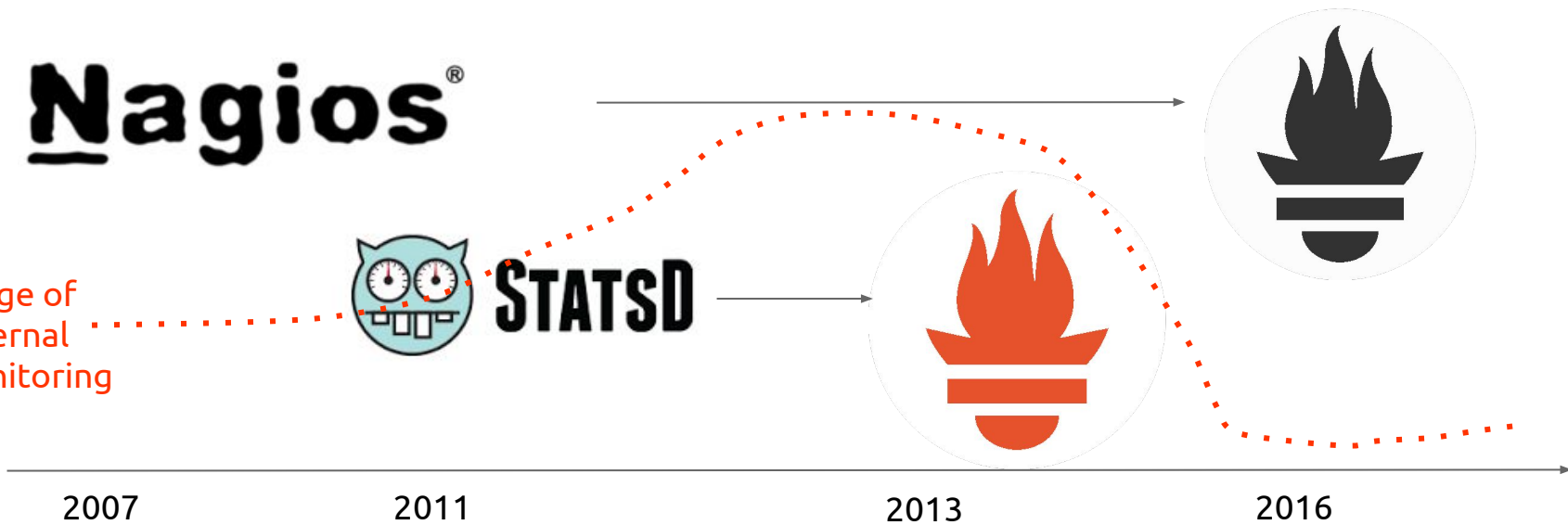
6

Labels are the new  
hierarchies.

# Nagios®

Usage of  
external  
monitoring

 **STATSD**



We need monitoring systems that allow us to alert for high-level service objectives, but retain the granularity to inspect individual components as needed.

*Chapter 10: Practical Alerting from Time-Series Data*

```
ampelmann.ams3.ip-10-12-11-17.api-mobile.5xx.requests
```

```
sum(ampelmann.ams3.ip-10-12-11-17.*.*.requests)
```

```
sum(ampelmann.*.*.api-mobile.5xx.requests)
```

```
ampelmann_requests_total{pool="ams3",instance="10.12.11.17:80",  
                           backend="api-mobile",code="5xx"}
```

```
sum(ampelmann_requests_total{pool="ams3",instance="10.12.11.17:80"})
```

```
sum(ampelmann_requests_total{backend="api-mobile",code="5xx"})
```

app\_env\_zone:request\_duration:99perc\_5m{app="api",zone="ams",env="production"} 0.243  
app\_env\_zone:request\_duration:99perc\_5m{app="api",zone="ams",env="canary"} 0.219



AlertManager BOT 2:35 AM

[FIRING:1] APILatencyHigh (production api-team critical)

api app in env production has 99th percentile latency of 243ms, SLO is 240ms. [Runbook](#)

