A word on averages...
The simplest thing...

...is sometimes also wrong

Which of the following distributions has average=500?
The simplest thing...

...is sometimes also wrong

Which of the following distributions has average=500? All!
But we can fix it by adding standard deviation, right?

- Most useful if your data is close to a normal distribution
What if my data looks different?

- Long tail? Use median and percentiles!
Example of plot with median

- Plotting an average with stdev might be meaningless in this case...
How do I get data to compute percentiles?

- In the middleware: process the logs by creating “buckets” for the different response times
- Memaslap: this information is already given to you

Total Statistics (3656 events)

- Min: 128
- Max: 19450
- Avg: 1366
- Geo: 834.20
- Std: 2058.84

Log2 Dist:

- 4: 0 0 0 1
- 8: 72 932 1675 599
- 12: 75 236 62 4

*an approximation, but good enough