Observability Through the Lens of Metrics and Events

Jack Neely jjneely@42lines.net
Breandan Dezendorf breandan@42lines.net

September 12, 2019

42 Lines, Inc.

What's a Metric?

Prometheus:

```
# HELP http_requests_total total HTTP hits
# TYPE http_requests_total counter
http_requests_total 34877
# HELP node_load1 1m load average.
# TYPE node_load1 gauge
node_load1 1.35
```

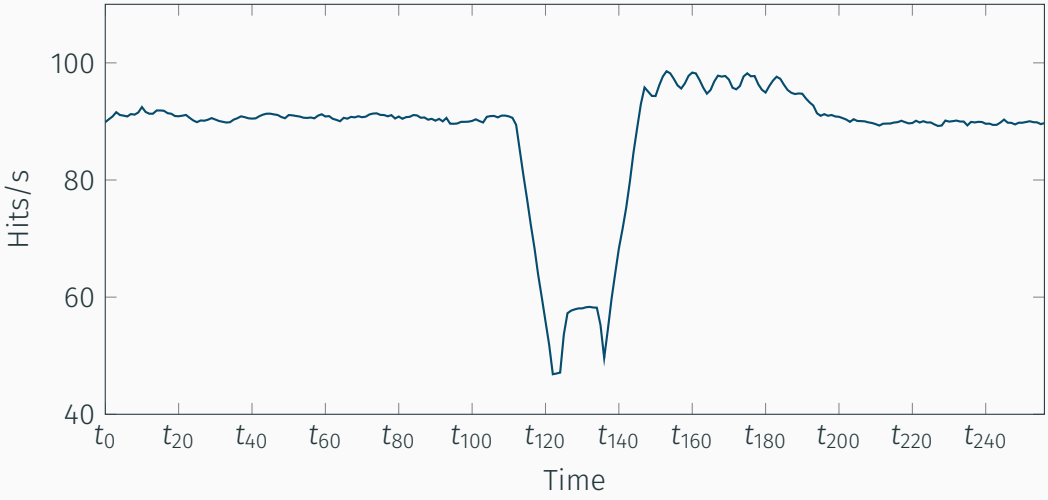
Graphite:

```
servers.A.http.hits 34877 1234567890
servers.A.collectd.load.load.shortterm 1.35 1234567890
```

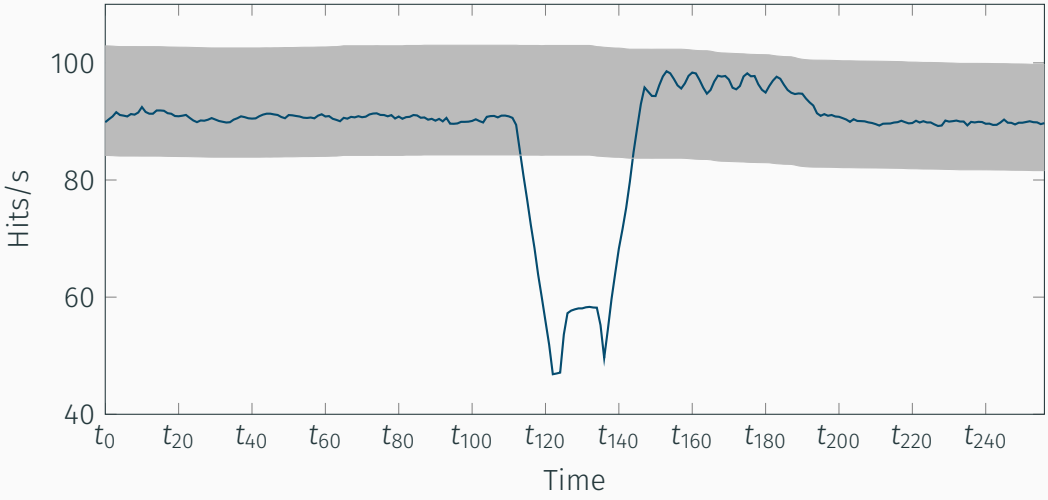
OpenTSDB:

```
put http.hits 1234567890 34877 host=A
put proc.loadavg.1min 1234567890 1.35 host=A
```

Time Series: sum(rate(http_requests_total[5m]))



Anomaly Detection: Moving Median 1 Hour Offset 10% Range



Basic Data Types

Gauge: Fluctuating Arbitrary Measurement

- Temperature
- Queue Size
- In Use File Descriptors

Counter: Continuously Incrementing (and Resetting)

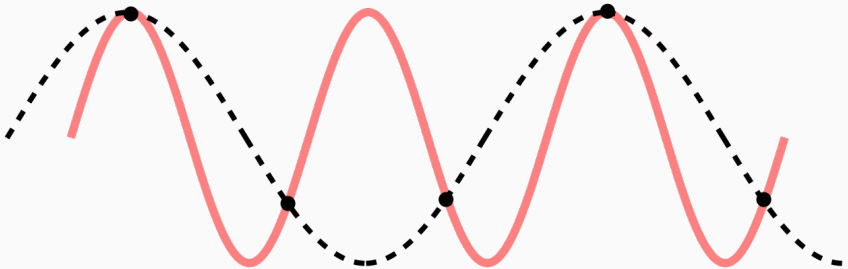
- Number of Bytes / Packets on a Network Interface
- Events

Stat: Summary Statistics for a Distribution of Events

- Duration of Event
- Size of Event

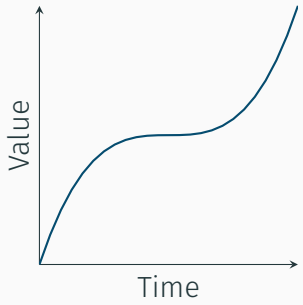
Theorem (Nyquist–Shannon Sampling)

If a function $x(t)$ contains no frequencies higher than B hertz, it is completely determined by giving its ordinates at a series of points spaced $1/2B$ seconds apart.



Definition (Monotonic Function)

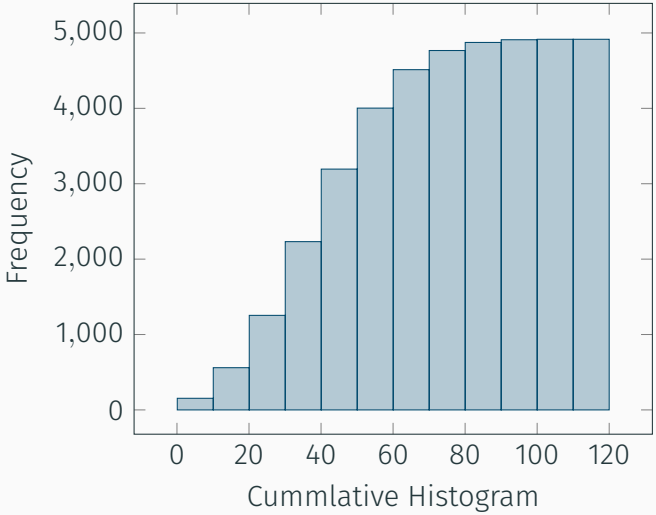
A function is called monotonically increasing if for all x and y such that $x \leq y$ one has $f(x) \leq f(y)$, so f preserves the order.



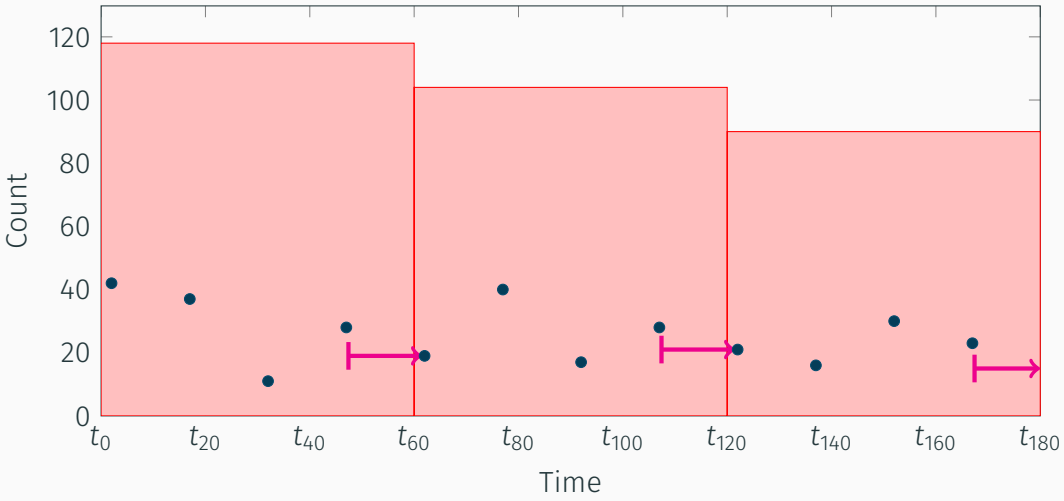
Options for Metrics of Events and Cardinality

StatsD – Generate Summary Metrics Per Interval

- Count
- Sum
- Min
- Max
- Percentiles



Data Alignment and Reporting



What's an Event?

Standard apache access log:

```
192.168.1.100 - bdezendorf [12/Sep/2019:31:32:000 -0000] "GET / HTTP/1.0" 200 2216
```

JSONLines formatted example of the same event:

```
{
  "timestamp":"2019-09-12T19:31:32.000Z",
  "http_size_bytes": 2216,
  "http_status_code": 200,
  "request": "GET / HTTP/1.0",
  "remote_address": "192.168.1.100",
  "remote_user": "bdezendorf"
}
```

Event Schema and Structure

Lucene isn't just for text anymore!

```
{
  "timestamp":"2019-09-12T19:31:32.000Z",
  "http_size_bytes": 2216,
  "http_status_code": 200,
  "request": "GET / HTTP/1.0",
  "remote_address": "192.168.1.100",
  "remote_user": "bdezendorf"
}
```

- Datetime
- IP Addresses
- Keyword
- Geopoint

How Do Events Compare to Metrics?

Disadvantages

- Events are roughly 100x more expensive than metrics
- Costs split evenly in memory, disk, CPU, and network I/O
- Guidelines for organization of data is hand wavy at best

Advantages

- Allows for ex post facto data exploration
- Finds needles in haystacks that p99 values can't

Guidelines for Using Metrics

- Count Performance
- Key Health Indicators
- Limited Debugging with Feature Flags
- Make a Plan for Cardinality
- Look for Histogram Support
- Have a “kit” for SOA Metrics
- Team or Application Namespaces
- Identify Important Aggregations and Ditch the Rest

Guidelines for Using Logs and Events

- Event size influences speed at every stage of the pipeline
- Do as little log processing as possible
- Sampling, rollup, and reporting

Use Hashing for Event Management

Generate request uuid at the loadbalancer

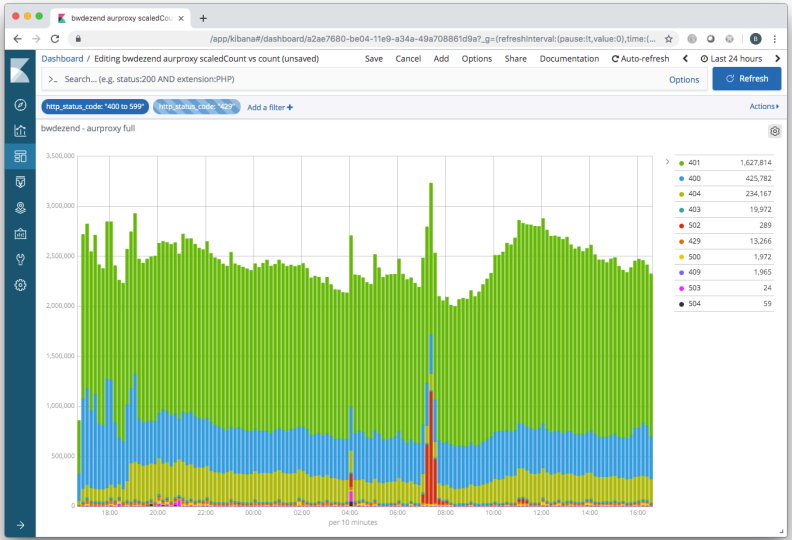
```
request_id:19CADF41-11BE-421D-89FA-52DBA8315A1A
```

MURMUR3 the id, divide by maxint

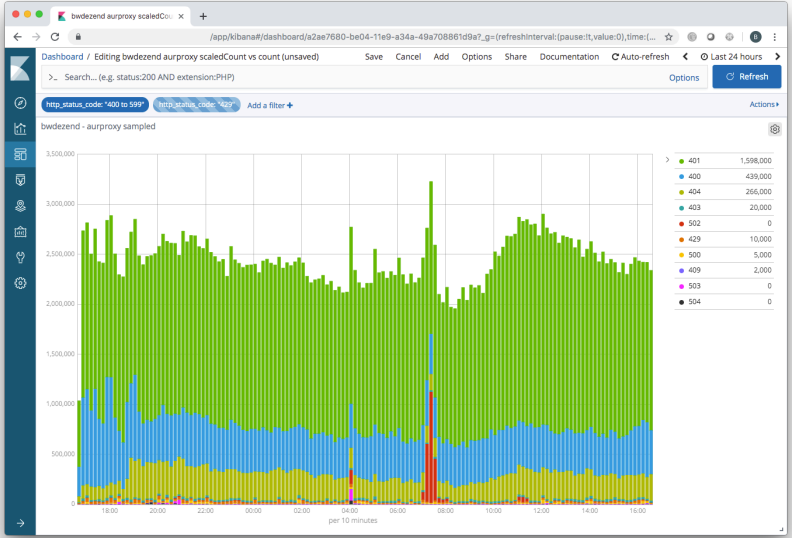
```
request_id_sampling_score:0.01799897874
```

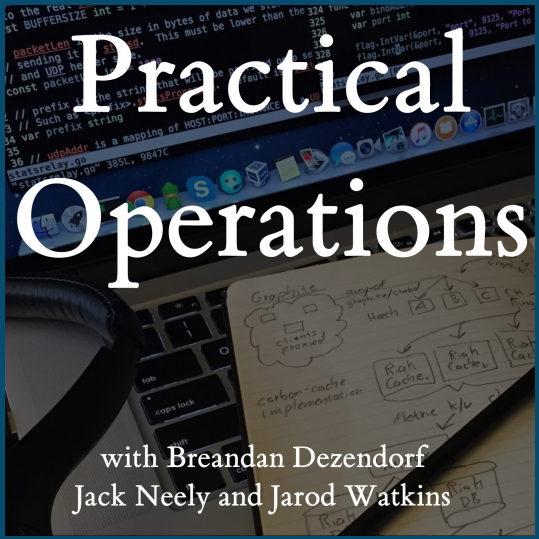
```
if ( [logstash][sampling_score] >= [logstash][long_tail_rate] ){
  mutate {
    replace => [ "[@metadata][retention]", "365" ]
  }
  ruby {
    code => "event.set('[logstash][scaledCount]',
    1 / ( 1 - event.get('[logstash][long_tail_rate]') ))"
  }
}
```

SAVE SAMPLING RATE TO RE-HYRDRATE DATA WITH LATER



SAVE SAMPLING RATE TO RE-HYRDRATE DATA WITH LATER





OPERATIONS.FM

Thank You!

Anomaly Detection Prometheus Examples

```
- record: job:http_requests:rate5m
  expr: sum(rate(http_requests_total[5m]))

- record: job:http_requests:rate5m_forecast
  expr: quantile_over_time(0.5, job:http_requests:rate5m[1h] offset 1h)

- alert: AnomalyFound
  expr: abs(job:http_requests:rate5m - job:http_requests:rate5m_forecast)
    / job:http_requests:rate5m_forecast > 0.1
  for: 3m
  labels:
    severity: page
  annotations:
    summary: Anomaly Found
    description: Red Alert
    runbook: http://wiki.example.com/AnomalyFound
```