

- [GOTO 2019 • An Introduction to Systems & Service Monitoring with Prometheus • Julius Volz](#) (<https://youtu.be/5O1djJ13gRU>)
 - <https://youtu.be/5O1djJ13gRU> (<https://youtu.be/5O1djJ13gRU>)
- [Prometheus: Deep-Dive](#) (<https://www.youtube.com/watch?v=9GMWvFcQjYI>)
 - <https://www.youtube.com/watch?v=9GMWvFcQjYI> (<https://www.youtube.com/watch?v=9GMWvFcQjYI>)

```
In [22]: 1 files = !ls -haltr *.png # sort screenshot in reverse date order
```

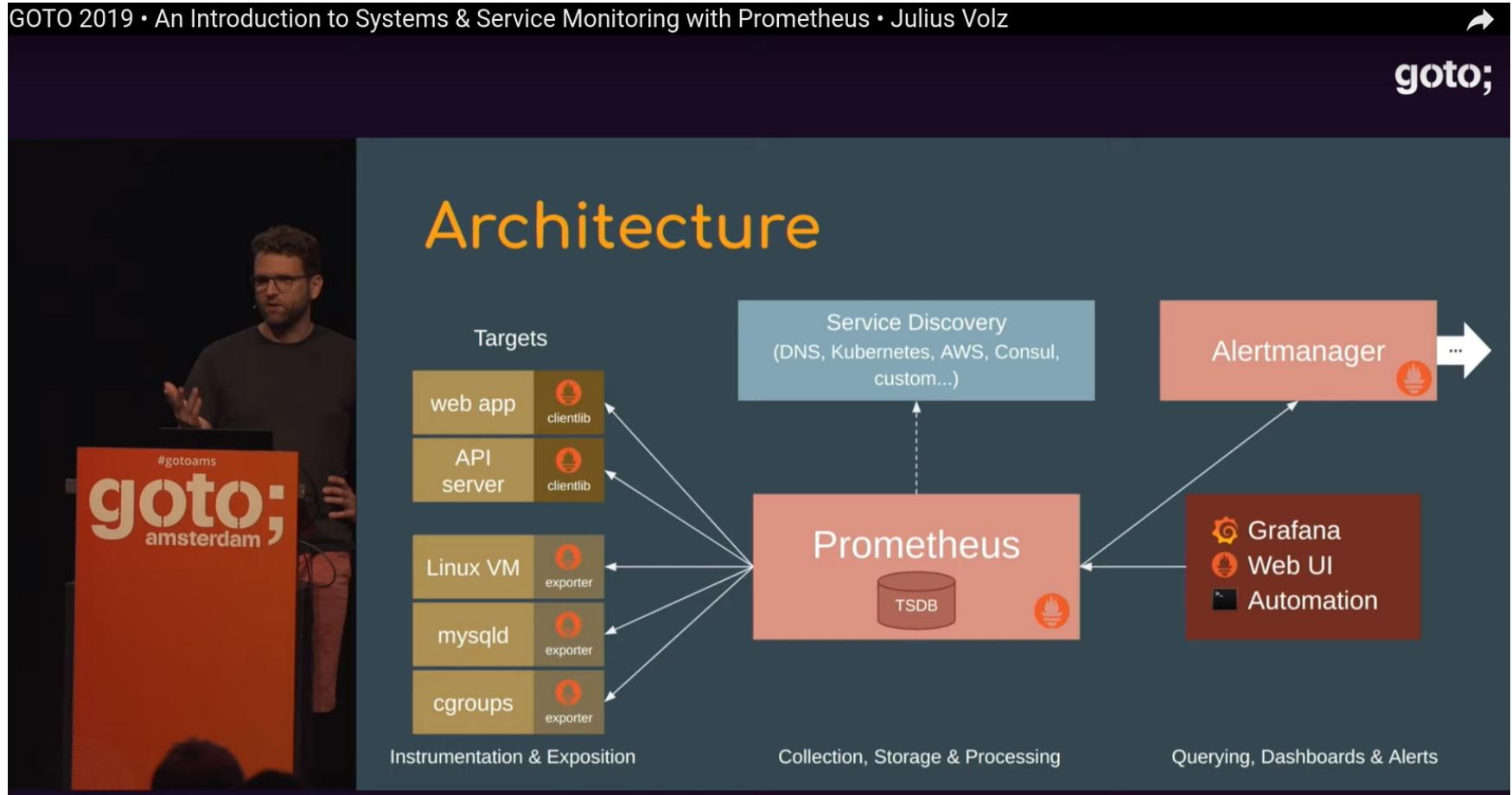
```
In [23]: 1 from IPython.core.display import HTML
```

In [24]:

```
1 img_html = ""
2 for i in f.split(' ')[-1] for f in files]:
3     img_html += f"<img src={i}>"
4 HTML(img_html)
```

Out[24]:

GOTO 2019 • An Introduction to Systems & Service Monitoring with Prometheus • Julius Volz



Favorite Features

- Dimensional data model
- Powerful query language
- Simple & efficient server
- Service discovery integration

Data Model

What is a time series?

<identifier> → [(t₀, v₀), (t₁, v₁), ...]

What is this?

int64

float64



Data Model

What identifies a time series?

```
http_requests_total{job="nginx",instance="1.2.3.4:80",path="/home",status="200"}
```



- Flexible
- No hierarchy
- Explicit dimensions

Querying

PromQL

- New query language
- Great for time series computations
- Not SQL-style

Querying

All partitions in my entire infrastructure with more than 100GB capacity that are not mounted on root?

```
node_filesystem_bytes_total{mountpoint!="/" } / 1e9 > 100
```

```
{device="sda1", mountpoint="/home", instance="10.0.0.1"} 118.8
{device="sda1", mountpoint="/home", instance="10.0.0.2"} 118.8
{device="sdb1", mountpoint="/data", instance="10.0.0.2"} 451.2
{device="xdvc", mountpoint="/mnt", instance="10.0.0.3"} 320.0
```

Querying

What's the ratio of request errors across all service instances?

```
sum by(path) (rate(http_requests_total{status="500"}[5m]))  
/ sum by(path) (rate(http_requests_total[5m]))
```

| | |
|--------------------------------|--------|
| {path="/status"} | 0.0039 |
| {path="/"} | 0.0011 |
| {path="/api/v1/topics/:topic"} | 0.087 |
| {path="/api/v1/topics"} | 0.0342 |

Querying

99th percentile request latency across all instances?

```
histogram_quantile(0.99,  
    sum without(instance) (rate(request_latency_seconds_bucket[5m]))  
)
```

| | |
|---|-------|
| {path="/status", method="GET"} | 0.012 |
| {path="/", method="GET"} | 0.43 |
| {path="/api/v1/topics/:topic", method="POST"} | 1.31 |
| {path="/api/v1/topics", method="GET"} | 0.192 |

Efficiency

Local storage is scalable enough for many orgs:

- 1 million+ samples/s
- Millions of series
- 1-2 bytes per sample

Good for keeping a few weeks or months of data.

Expression browser

Prometheus Alerts Graph Status Help

```
sort_desc(sum(bazooka_instance_memory_limit_bytes - bazooka_instance_memory_usage_bytes) by (app, proc)) / 1024 / 1024 / 1024
```

Execute

Graph Console

| Element | Value |
|--|------------|
| {app="harsh-dagger",proc="api"} | 132.720802 |
| {app="quality-locomotive",proc="web"} | 89.547081 |
| {app="husky-long-oyster",proc="web"} | 68.982738 |
| {app="vital-albatross",proc="api"} | 48.033772 |
| {app="autopsy-gutsy",proc="widget"} | 47.410583 |
| {app="western-python",proc="cruncher"} | 40.126926 |
| {app="harsh-dagger",proc="api"} | 28.527714 |
| {app="outstanding-dagger",proc="api"} | 26.119423 |
| {app="gruesome-waterbird",proc="web"} | 17.666714 |
| {app="gutsy-square",proc="public"} | 15.296242 |
| {app="harsh-dagger",proc="web"} | 14.738327 |
| {app="northern-electron",proc="api"} | 13.349815 |

Built-in graphing



Dashboarding via Grafana



Alerting

generate an alert for each path with an error rate of >5%

```
alert: Many500Errors
expr: |
  (
    sum by(path) (rate(http_requests_total{status="500"}[5m]))
    /
    sum by(path) (rate(http_requests_total[5m]))
  ) * 100 > 5
for: 5m
labels:
  severity: "critical"
annotations:
  summary: "Many 500 errors for path {{$labels.path}} (${{value}}%)"
```

Dynamic Environments

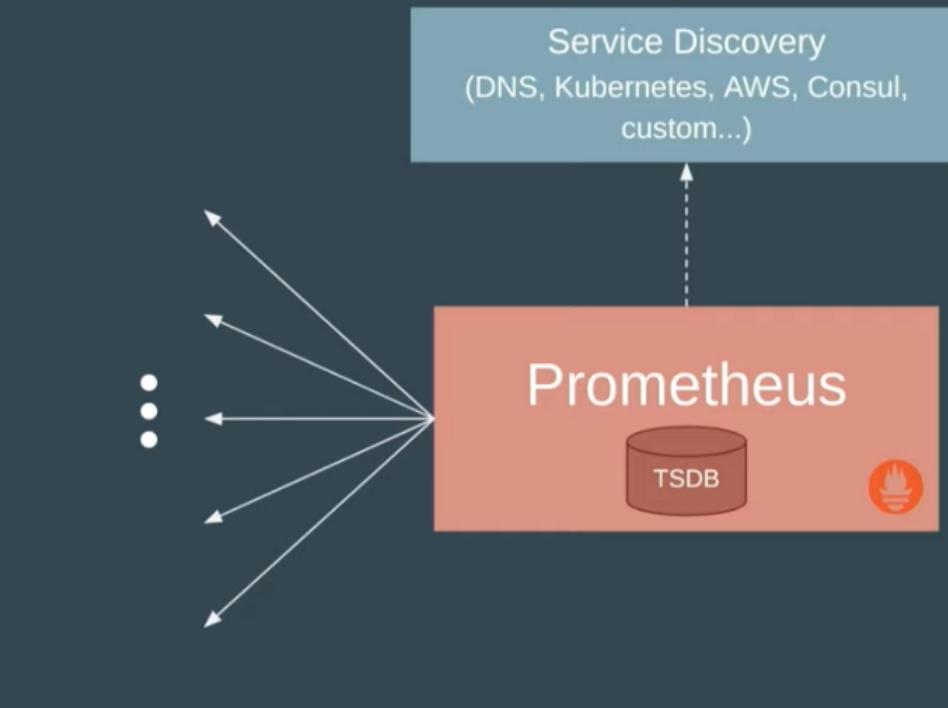
...pose new challenges:

- Dynamic VMs
- Cluster schedulers
- Microservices

→ many services, dynamic hosts, and ports

How to make sense of this all?

Service Discovery Integration



Answers three questions:

- what **should** be there?
- how do I reach it?
- what is it? (metadata)

Conclusion

Prometheus helps you make sense of complex dynamic environments thanks to its:

- Dimensional data model
- Powerful query language
- Simplicity + efficiency
- Service discovery integration

In []:

1