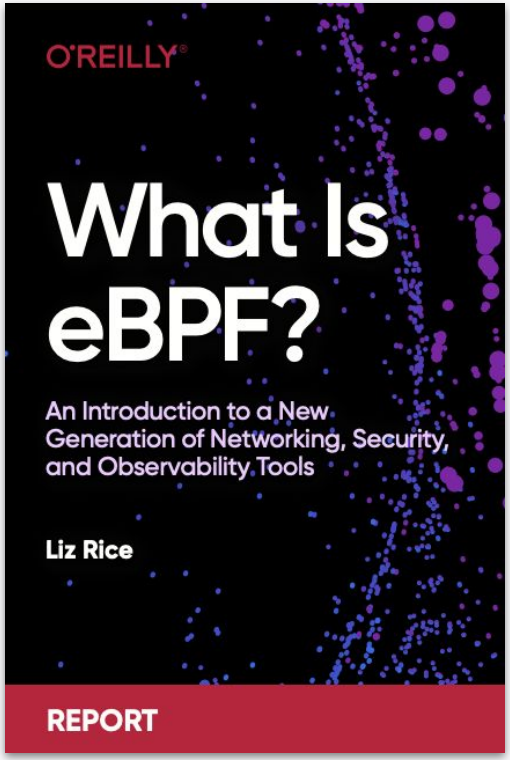
ISOVALENT

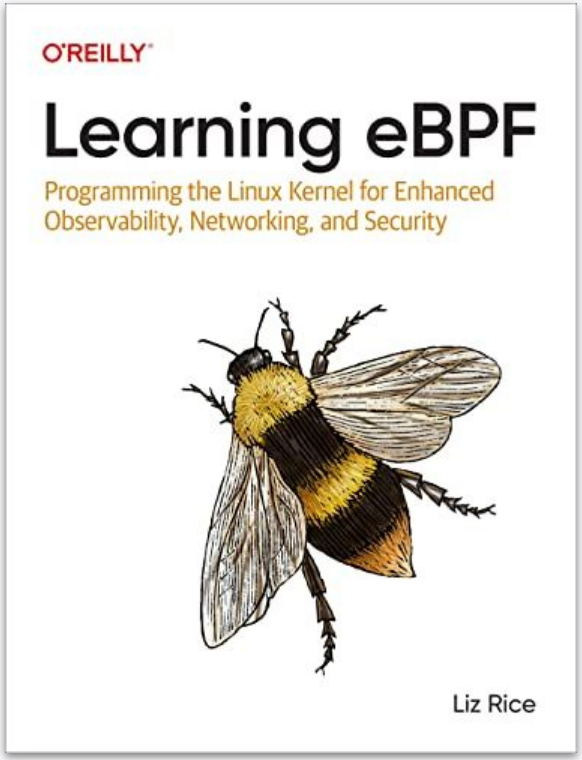
eBPF Superpowers for SRE

Liz Rice | @lizrice
Chief Open Source Officer, Isovalent
Emeritus Chair, CNCF Technical Oversight Committee | CNCF & OpenUK boards

ISOVALENT

@lizrice


O'REILLY
What Is eBPF?
An Introduction to a New Generation of Networking, Security, and Observability Tools
Liz Rice
REPORT


O'REILLY
Learning eBPF
Programming the Linux Kernel for Enhanced Observability, Networking, and Security
Liz Rice

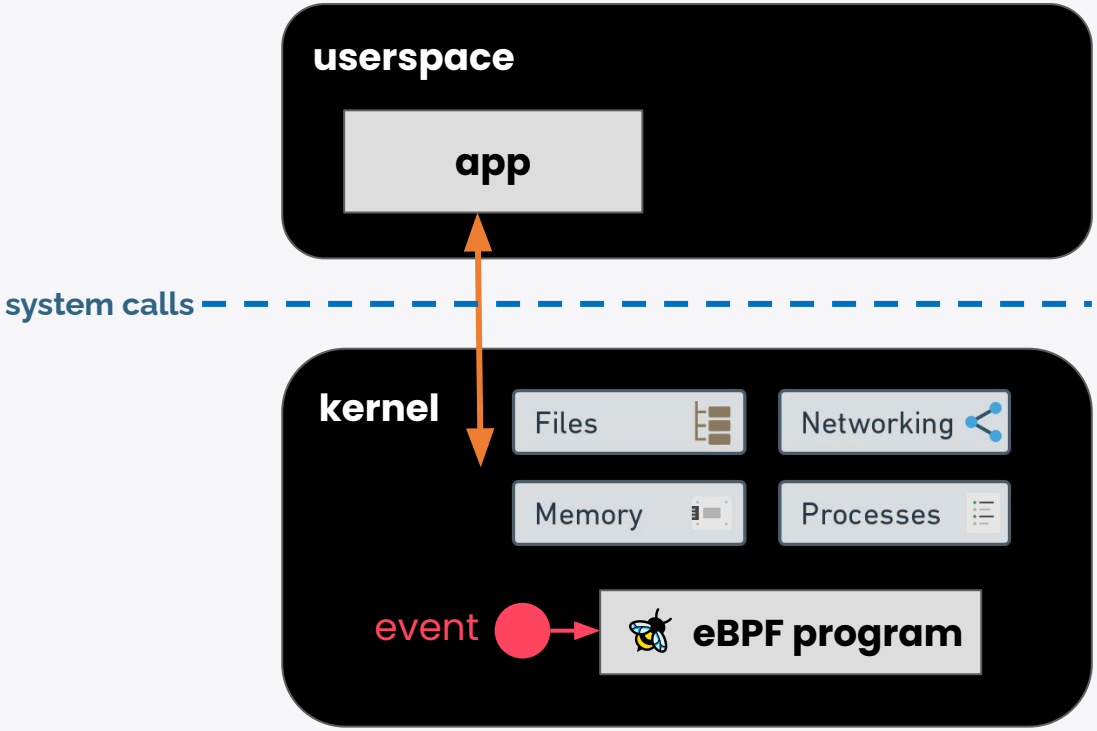
ISOVALENT

What is  eBPF ?

Makes the kernel programmable

ISOVALENT

Run custom code in the kernel



@lizrice

ISOVALENT

eBPF Hello World

```
SEC("kprobe/sys_execve")
int hello(void *ctx)
{
    bpf_printk("I'm alive!");
    return 0;
}
```

Info about process that called execve syscall

```
$ sudo ./hello
  bash-20241   [004] d... 84210.752785: 0: I'm alive!
  bash-20242   [004] d... 84216.321993: 0: I'm alive!
  bash-20243   [004] d... 84225.858880: 0: I'm alive!
```

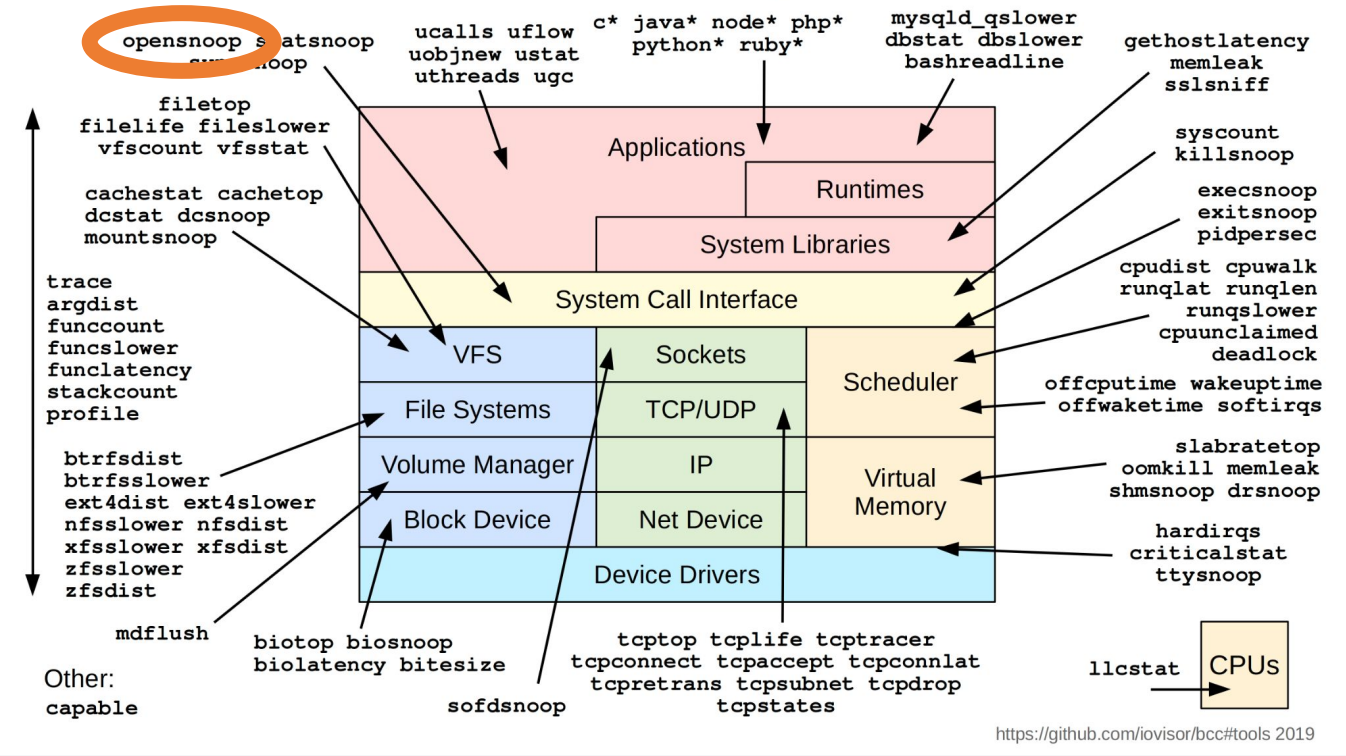
ISOVALENT

SRE practical use case #1
Event tracing

@lizrice

ISOVALENT

eBPF tracing tools from iovisor/bcc



https://github.com/iovisor/bcc#tools 2019

@lizrice

ISOVALENT

eBPF tracing - opensnoop

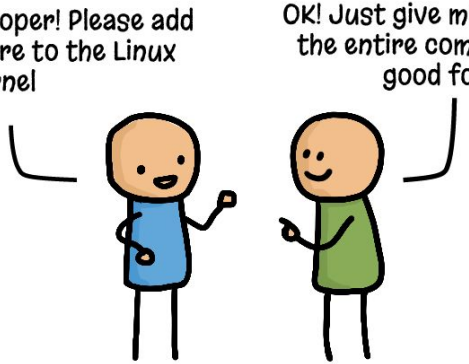
```
~/bcc/libbpf-tools$ sudo ./opensnoop
PID    COMM               FD ERR PATH
5040   node               21   0 /proc/5132/cmdline
5040   node               21   0 /proc/6460/cmdline
5040   node               21   0 /proc/6460/cmdline
6461   opensnoop          18   0 /etc/localtime
5040   node               21   0 /proc/5132/cmdline
5040   node               21   0 /proc/6460/cmdline
5060   node               23   0 /home/liz/.vscode-server/data/User/workspaceStorage/48b53
5040   node               21   0 /proc/5132/cmdline
5040   node               21   0 /proc/6460/cmdline
5040   node               21   0 /proc/5132/cmdline
5040   node               21   0 /proc/6460/cmdline
…
```

ISOVALENT

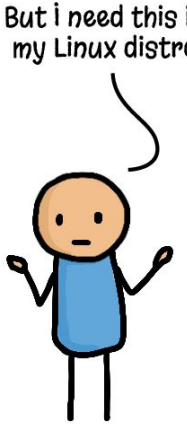
Dynamically change kernel behaviour

@lizrice

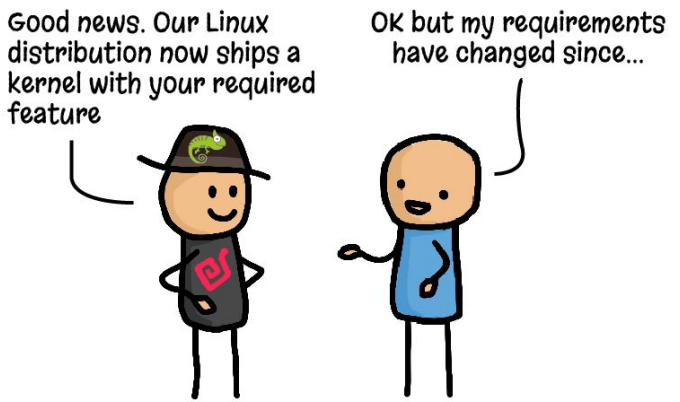
Application Developer:



1 year later...



5 years later...

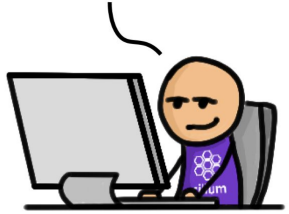


Application Developer:

i want this new feature to observe my app

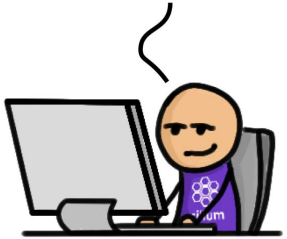
eBPF Developer:

OK! The kernel can't do this so let me quickly solve this with eBPF.



A couple of days later...

Here is a release of our eBPF project that has this feature now. BTW, you don't have to reboot your machine.

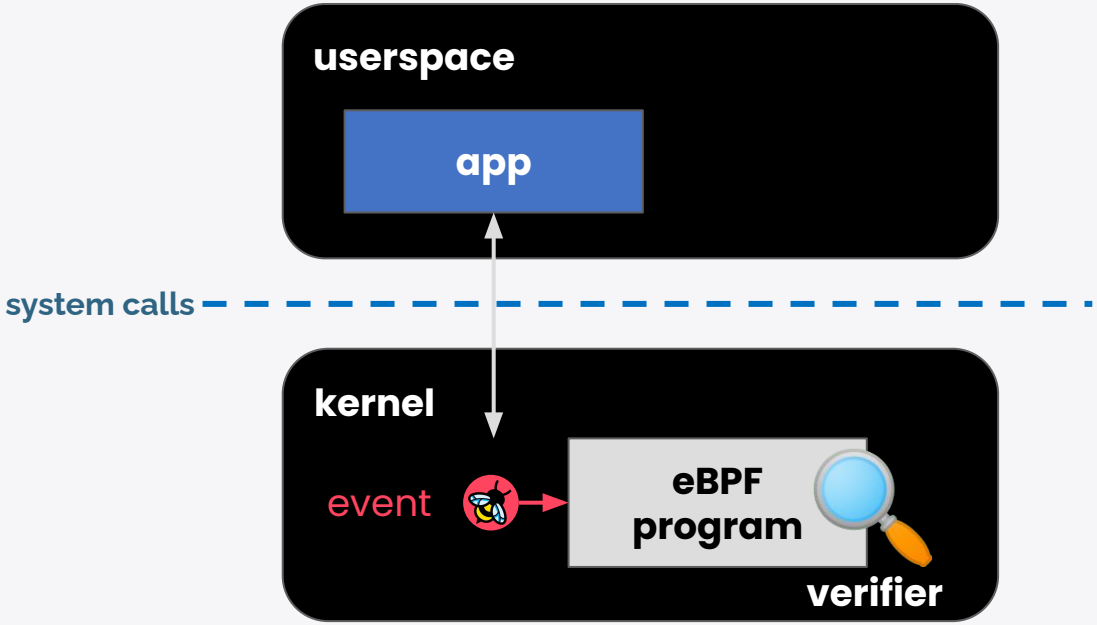


ISOVALENT

Couldn’t I do this with kernel modules?

@lizrice

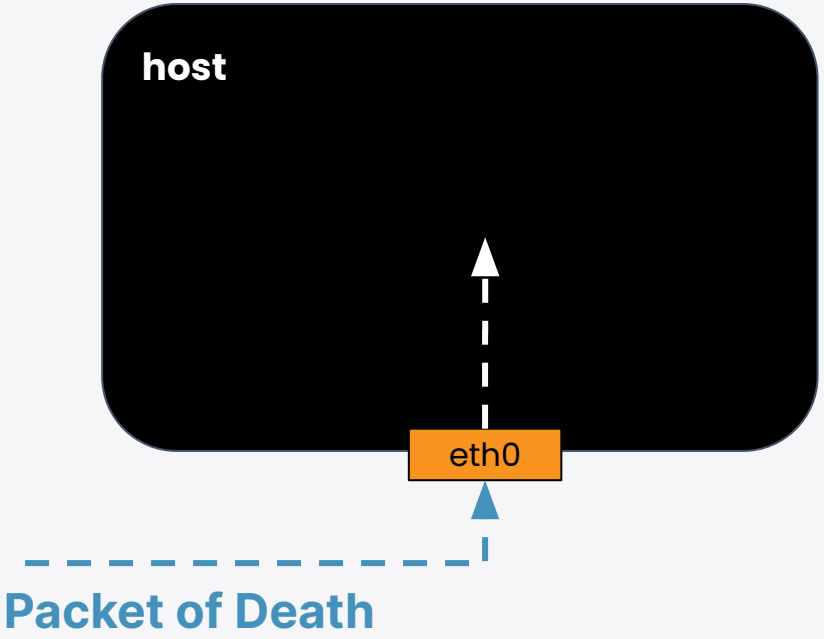
eBPF verification ensures program safety

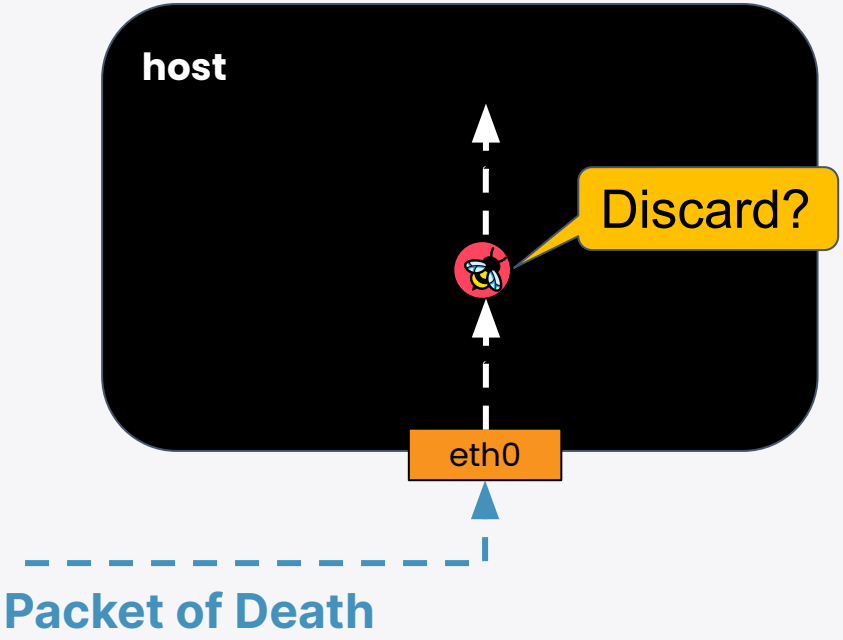


ISOVALENT

SRE practical use case #2
Kernel vulnerability mitigation

@lizrice


host
eth0
Packet of Death


host
Discard?
eth0
Packet of Death

ISOVALENT

eBPF Packet Drop

```
SEC("xdp/bye")
int goodbye_ping(struct xdp_md *ctx)
{
    ...
    if (iph->protocol == IPPROTO_ICMP)
       return XDP_DROP;

    return XDP_PASS;
}
```

ISOVALENT

eBPF with Kubernetes

@lizrice

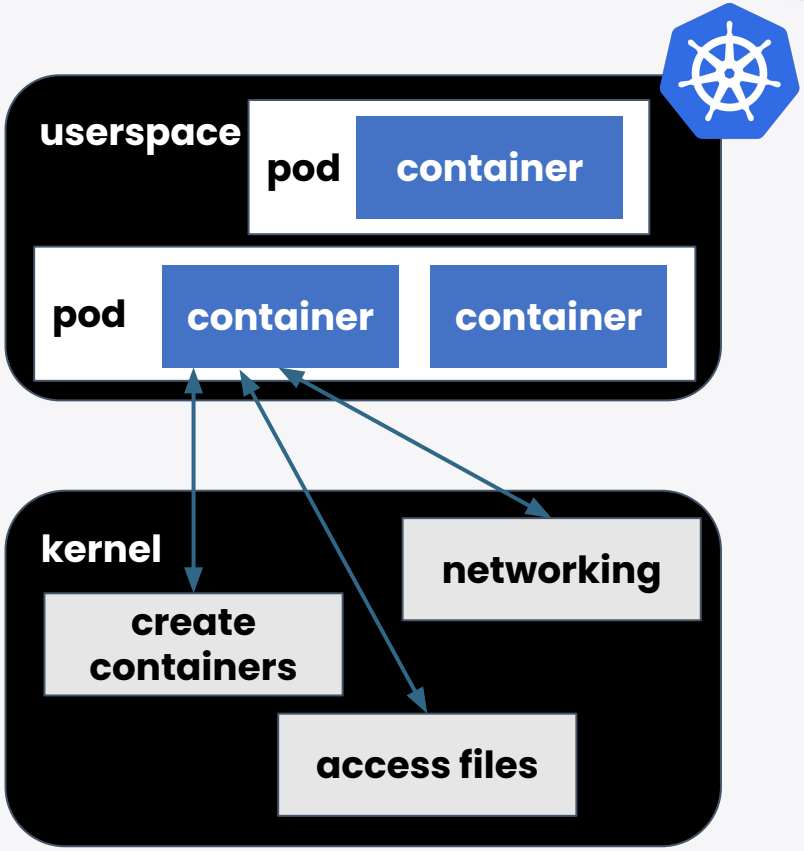
ISOVALENT

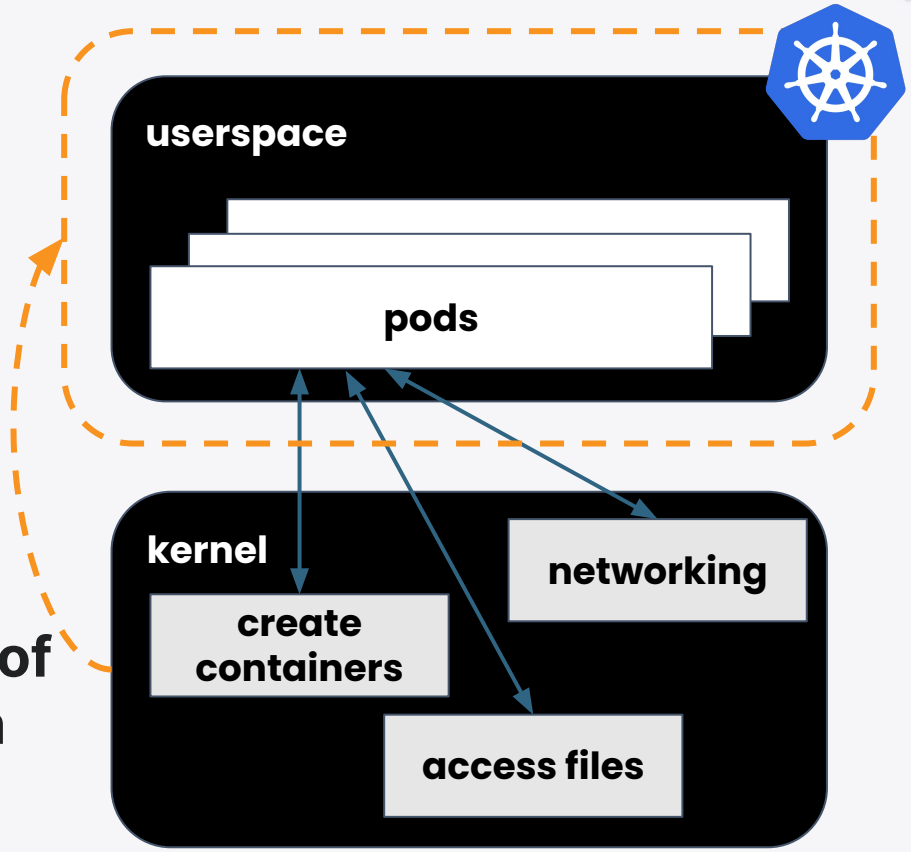
eBPF started a **whole new infrastructure movement** in the cloud native space

- Daniel Borkmann

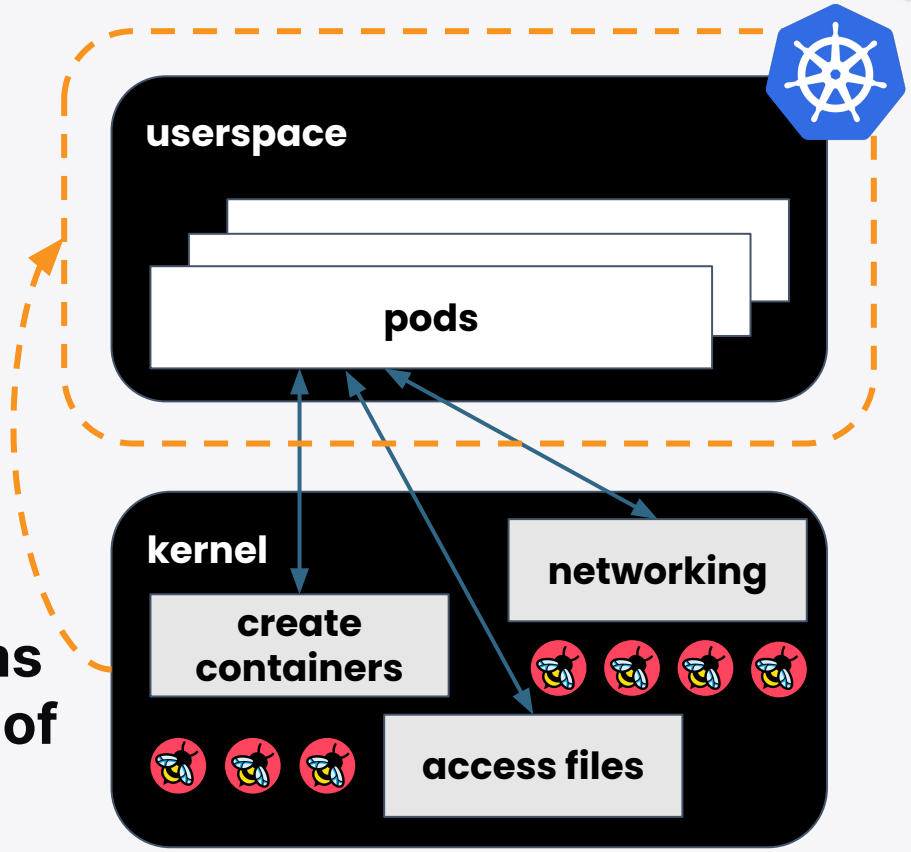
@lizrice

One kernel per host





Kernel aware of everything on the host



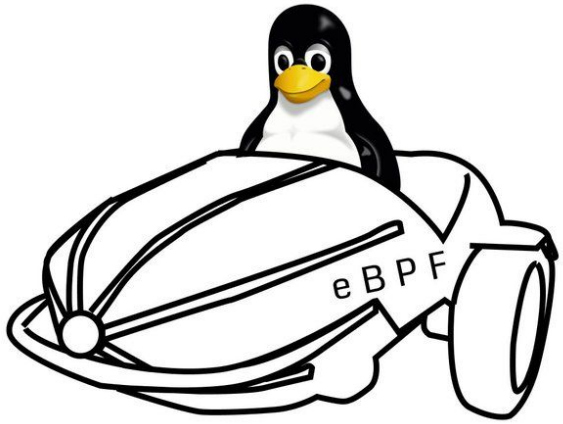
eBPF programs can be aware of everything

ISOVALENT

eBPF tools instrument the system without any app or config changes

@lizrice

My other sidecar is a **kernel**

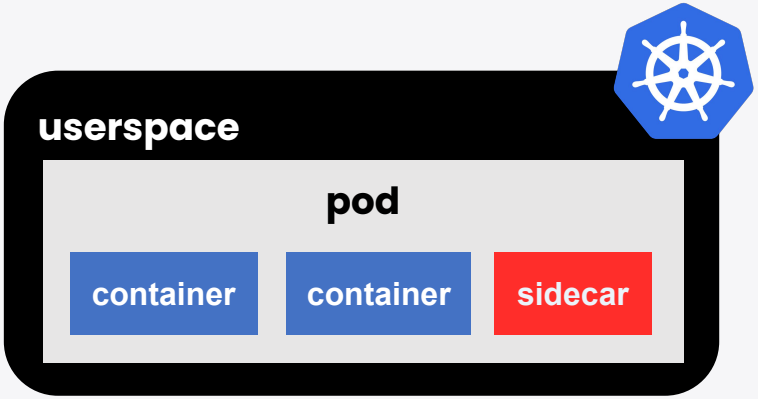


"Get in loser. We're going tracing"

- Nathan LeClaire @dotpem

ISOVALENT

A sidecar has a view across one pod

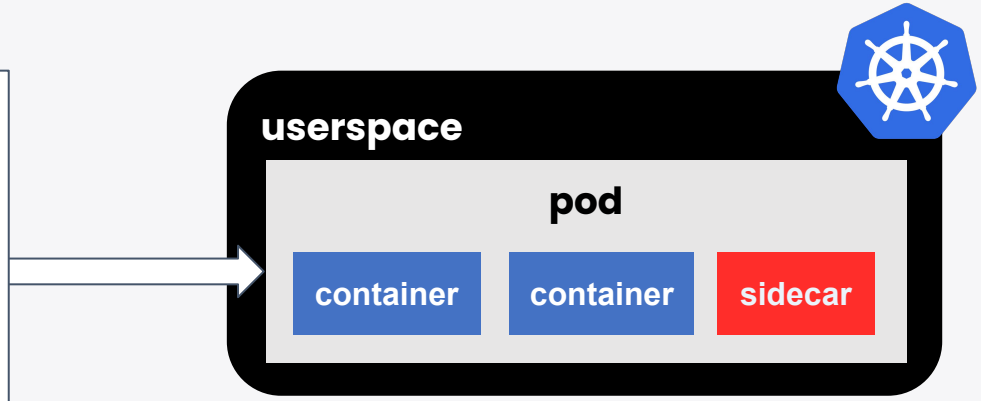


ISOVALENT

Sidecars need YAML

my-app.yaml

```
containers:
  -   name: my-app
      ...
  -   name: my-app-init
      …
  -   name: my-sidecar
      ...
```

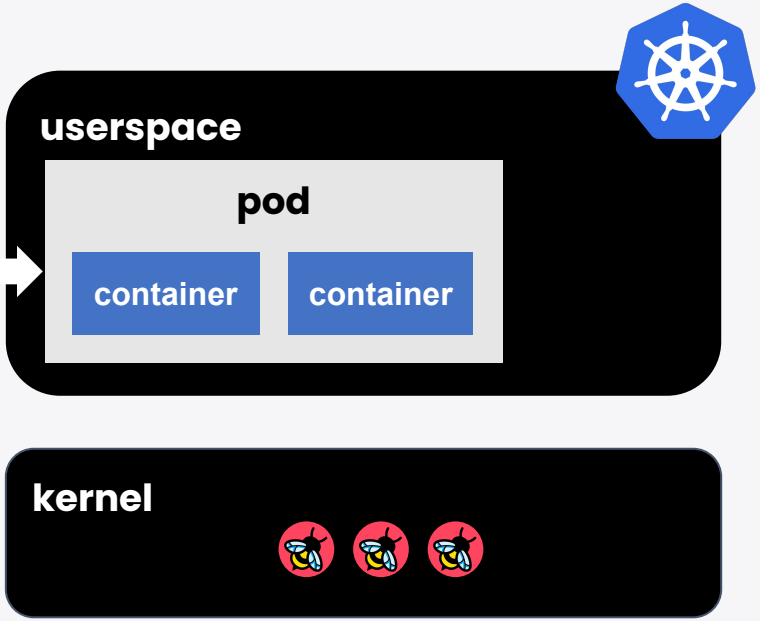


ISOVALENT

eBPF does not need any app changes

my-app.yaml

```
containers:
  -   name: my-app
      ...
  -   name: my-app-init
      …
```

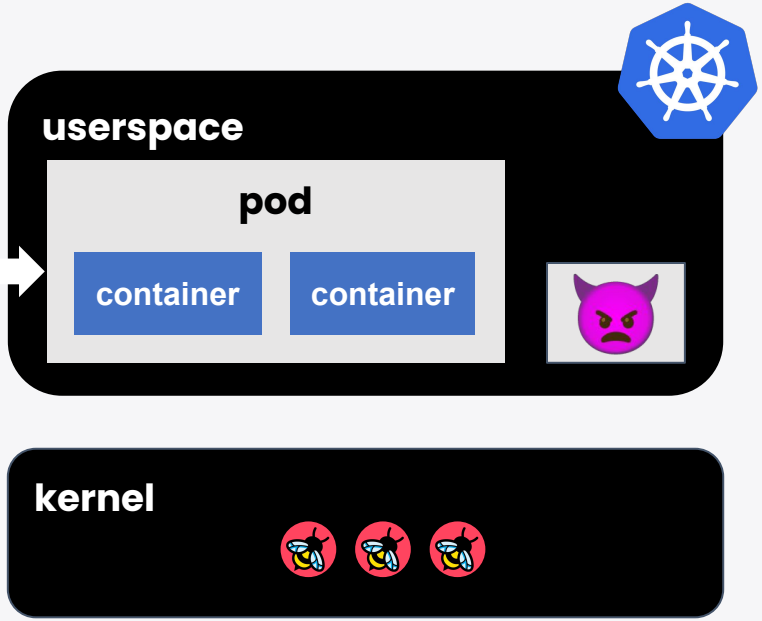


ISOVALENT

eBPF can see ALL activity on the node

my-app.yaml

```
containers:
  -   name: my-app
      ...
  -   name: my-app-init
      …
```



ISOVALENT

SRE practical use case #3
Cloud native observability tools

@lizrice

ISOVALENT

Observability is the fundamental basis for all SRE

- Mario Biemans

ISOVALENT

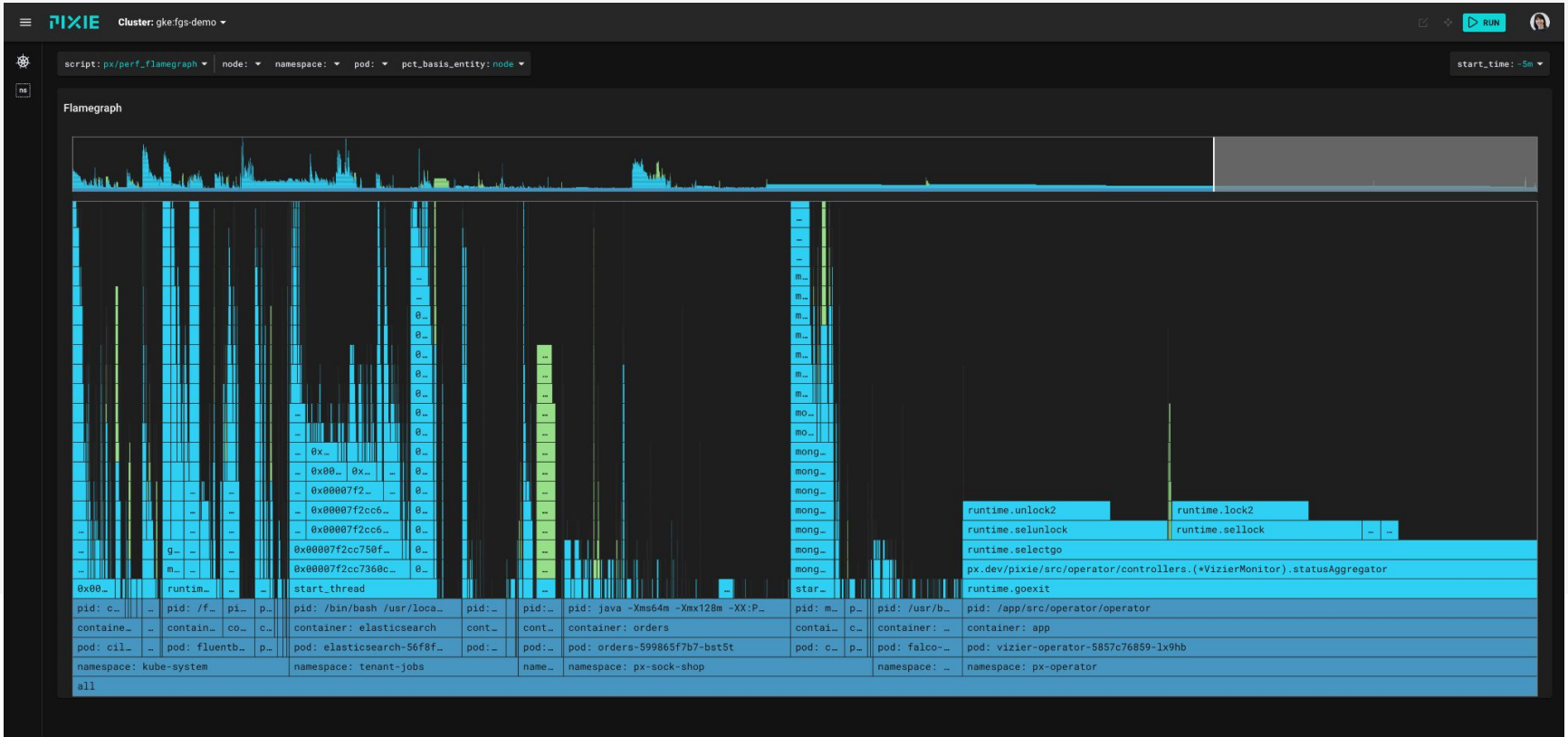
INSPEKTOR GADGET tracing on Kubernetes

```
$ kubectl gadget trace open
NODE                 NAMESPACE    POD     CONTAINER   PID    COMM   FD  ERR PATH
kind-2-control-plane default      xwing   spaceship   361876 vi     3     0 /etc/passwd
```

Kubernetes info

ISOVALENT

PIXIE flame graph

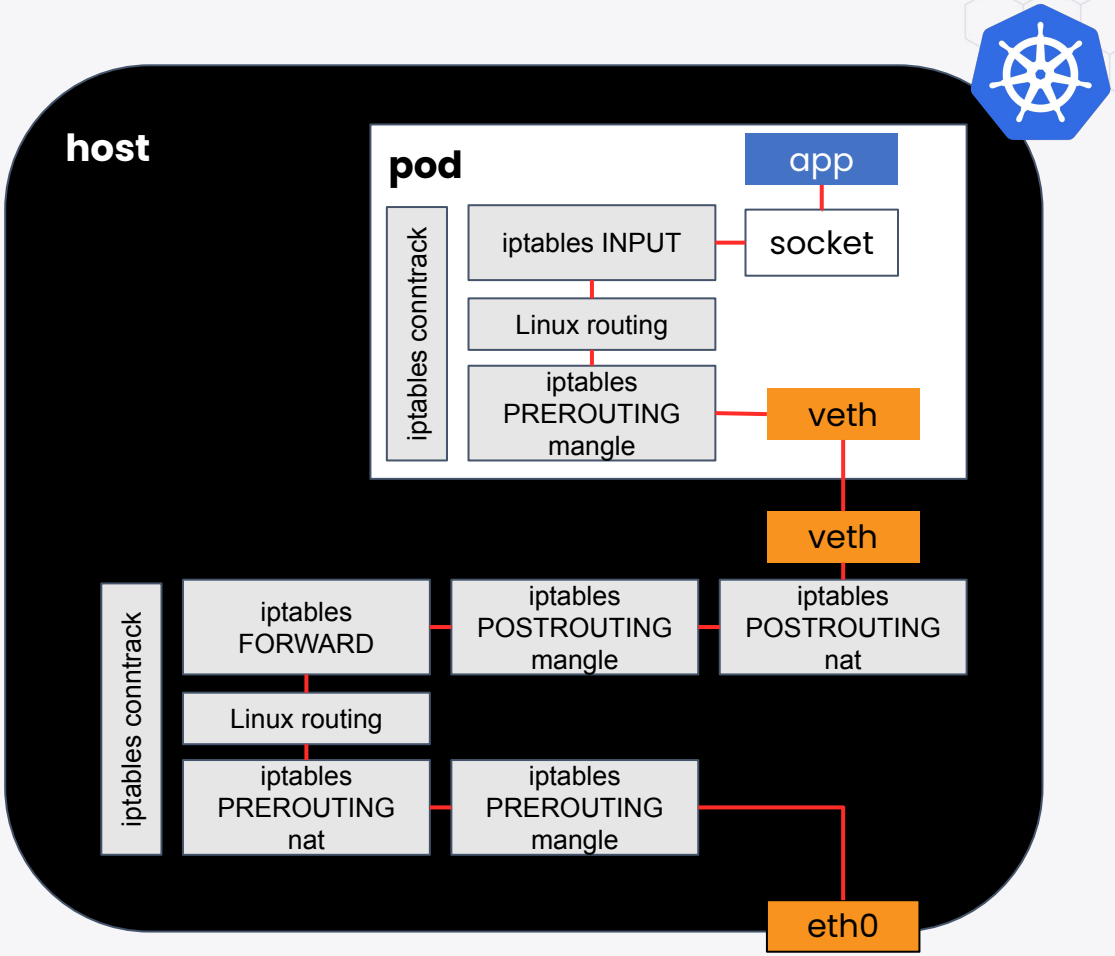


ISOVALENT

SRE practical use case #4
High performance networking

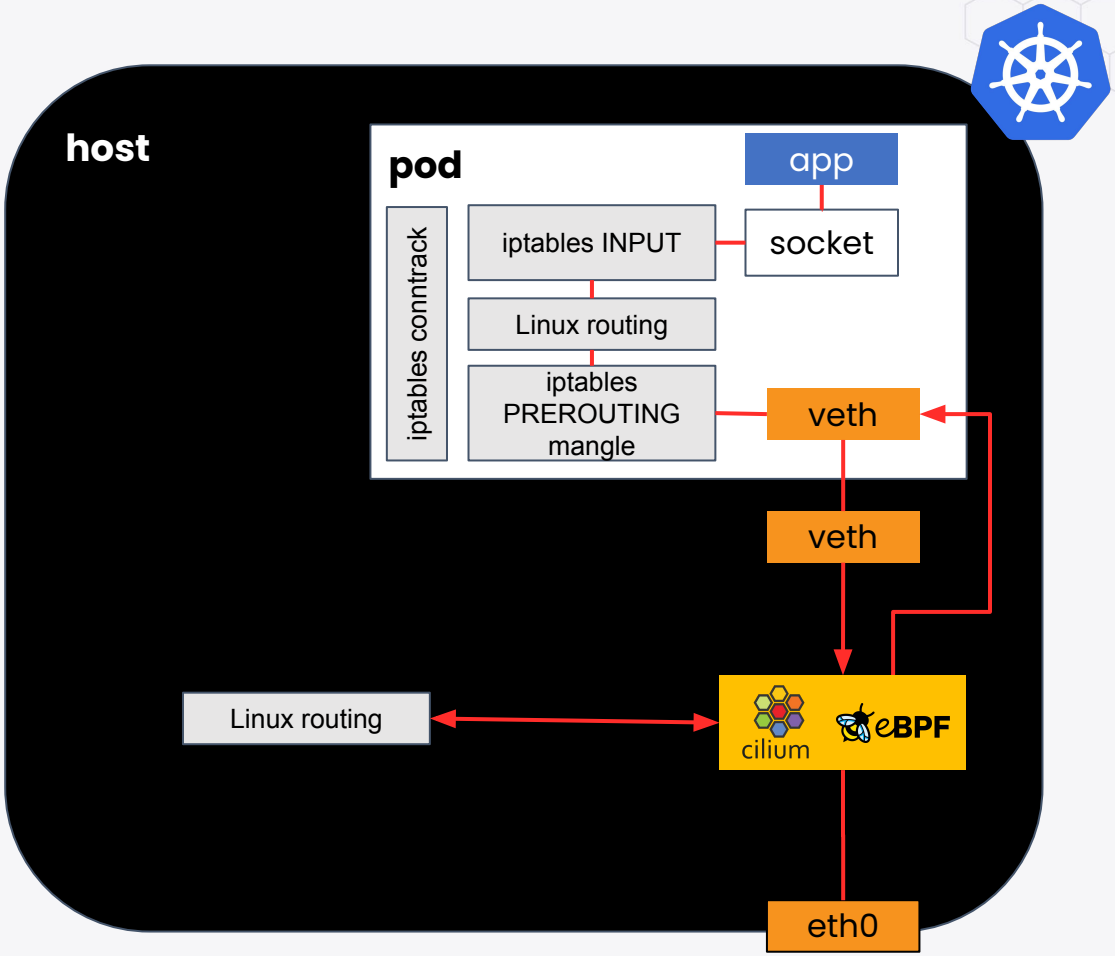
@lizrice

ISOVALENT



@lizrice

ISOVALENT

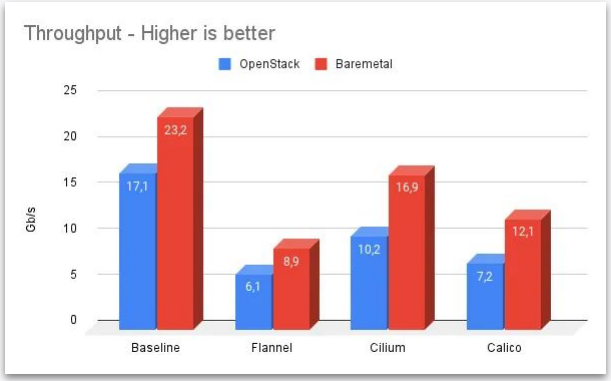


@lizrice

ISOVALENT

Unleashing the Power of Cilium CNI to Propel Trendyol's Performance Up to 40%!

Trendyol implemented Cilium as the default CNI for the Kubernetes Cluster starting from version 1.26. Discover our journey.

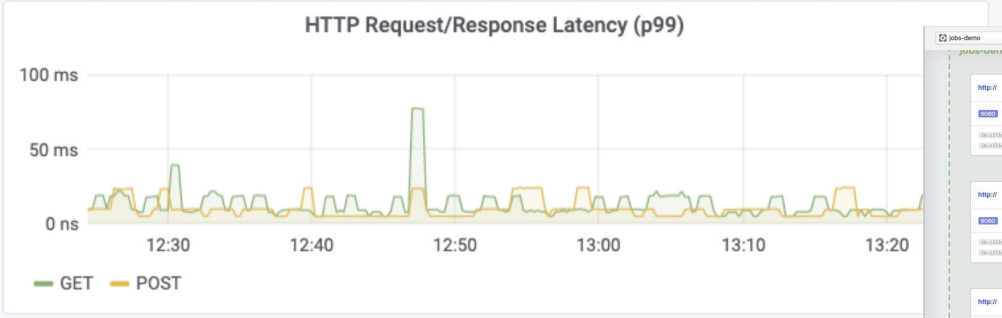


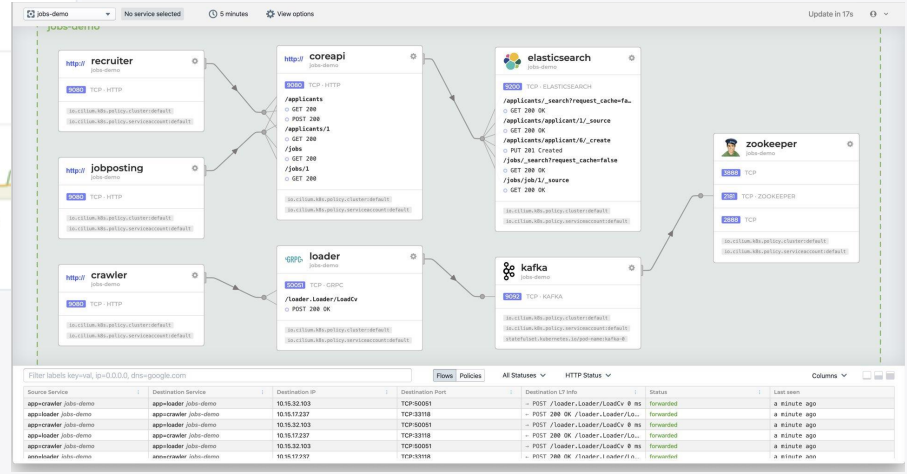
Unleashing the Power of Cilium CNI to Propel Trendyol's Performance Up to 40%! 20 July 2023

@lizrice

ISOVALENT

Cilium Hubble observability





- Network flow logs
- Prometheus metrics
- Service map
- L3/4 & L7 (HTTP, DNS, Kafka, ...)
- Aware of Kubernetes identities

@lizrice

ISOVALENT

SRE practical use case #5
Cloud native security tools

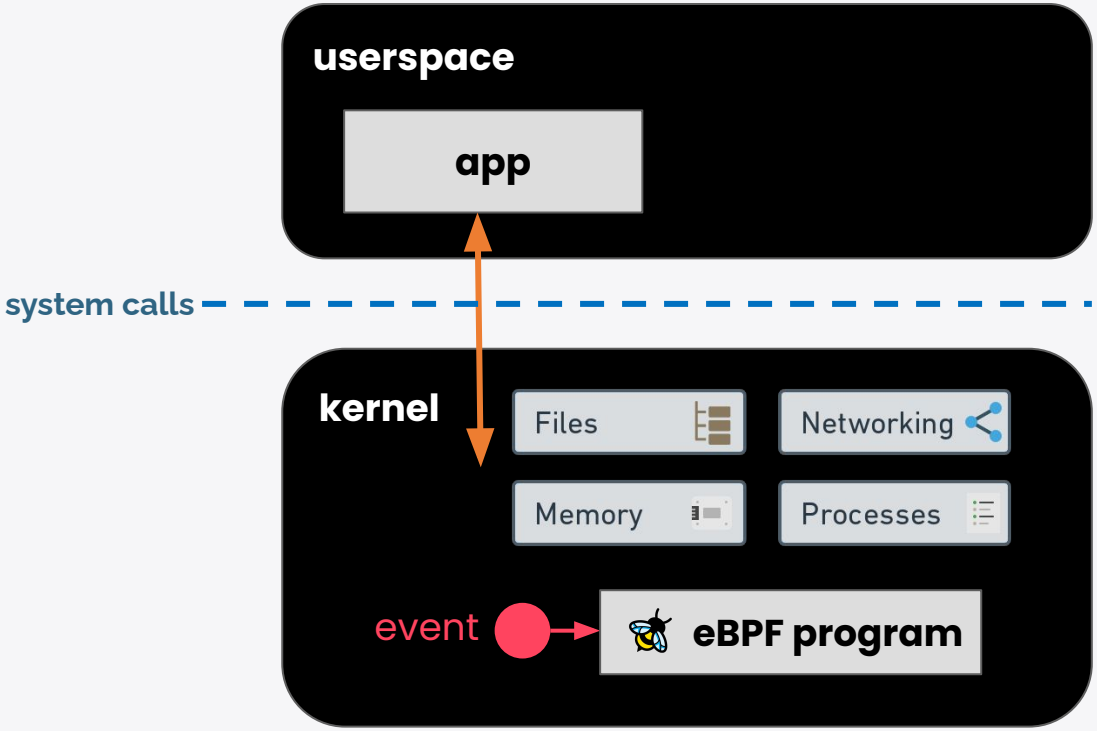
@lizrice

68% of SREs say they expect their role in **security** to become even more central

- Dynatrace State of SRE report 2022

ISOVALENT

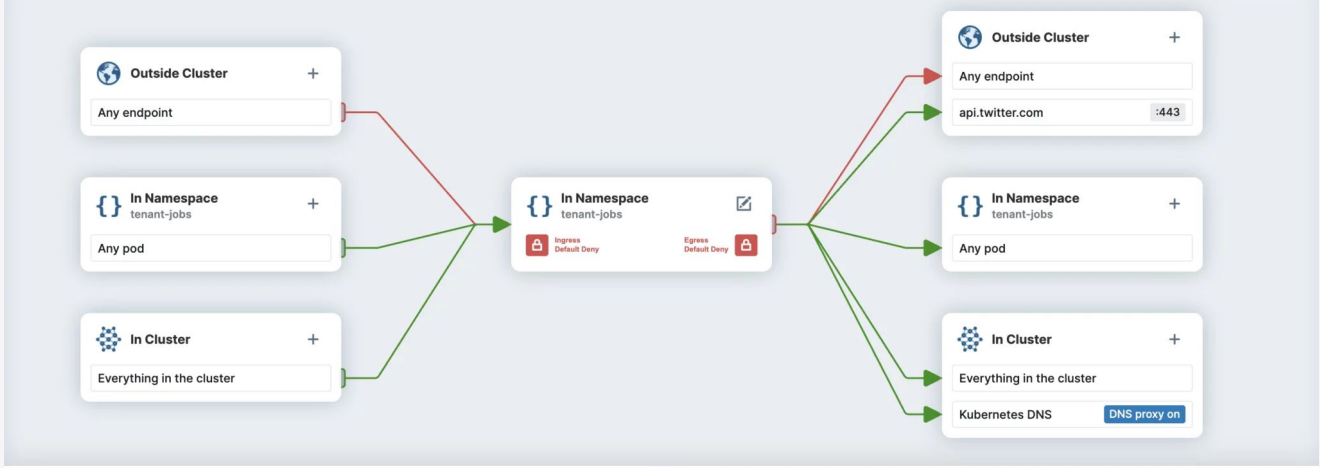
eBPF for security-relevant events



@lizrice

ISOVALENT

Cilium network policy → eBPF programs drop packets



Kubernetes | Cilium | Download

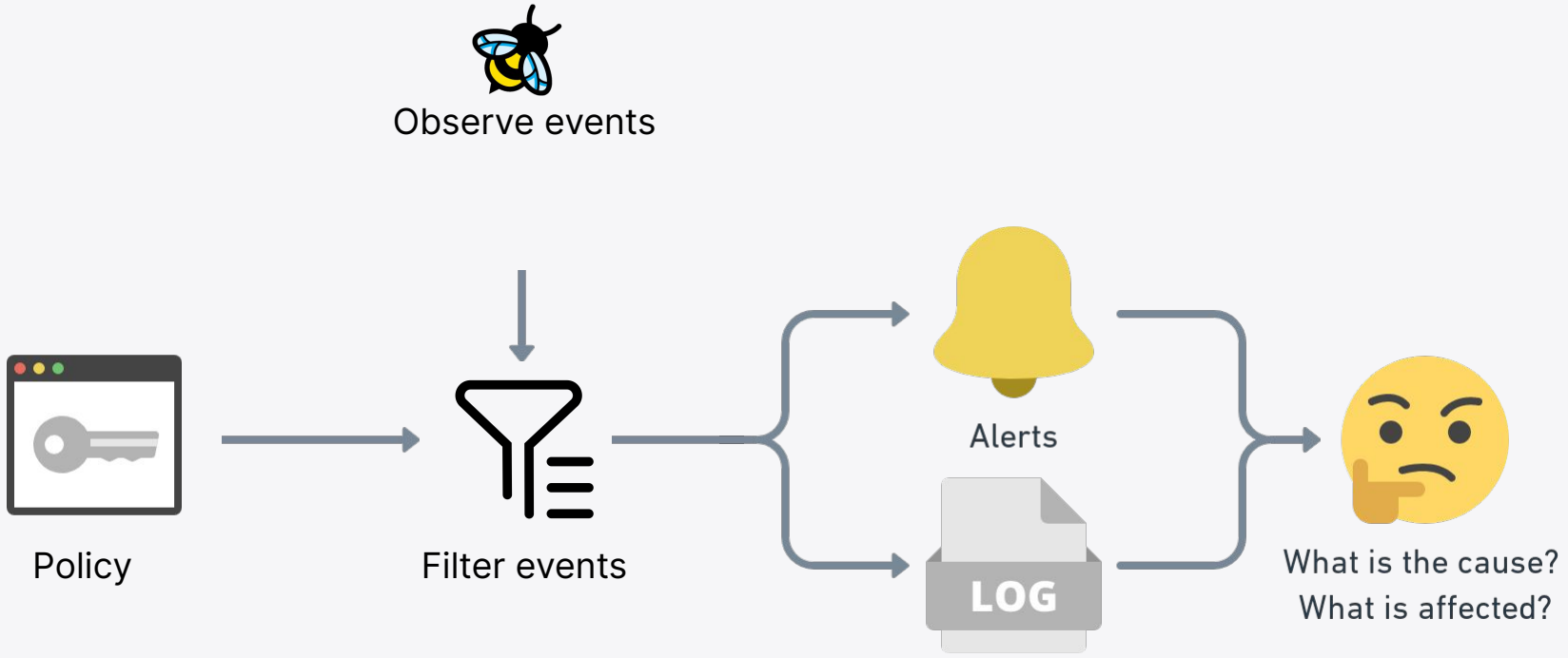
```
21        - ports:
22            - port: "53"
23              protocol: ANY
24          rules:
25            dns:
26              - matchPattern: "*"
27      - toFQDNs:
28          - matchName: api.twitter.com
29        toPorts:
30          - ports:
31              - port: "443"
```

Source Identity	Destination Identity	Verdict
crawler *tenant-jobs*	api.twitter.com	forwarded
crawler *tenant-jobs*	elasticsearch *tenant-jobs*	forwarded
crawler *tenant-jobs*	elasticsearch *tenant-jobs*	forwarded
crawler *tenant-jobs*	api.twitter.com	forwarded
loader *tenant-jobs*	kafka *tenant-jobs*	forwarded
coreapi *tenant-jobs*	elasticsearch *tenant-jobs*	forwarded
jobposting *tenant-jobs*	coreapi *tenant-jobs*	forwarded
jobposting *tenant-jobs*	coreapi *tenant-jobs*	forwarded
coreapi *tenant-jobs*	elasticsearch *tenant-jobs*	forwarded
kafka *tenant-jobs*	zookeeper *tenant-jobs*	forwarded
coreapi *tenant-jobs*	elasticsearch *tenant-jobs*	forwarded
recruiter *tenant-jobs*	coreapi *tenant-jobs*	forwarded
recruiter *tenant-jobs*	coreapi *tenant-jobs*	forwarded

@lizrice

ISOVALENT

Security observability



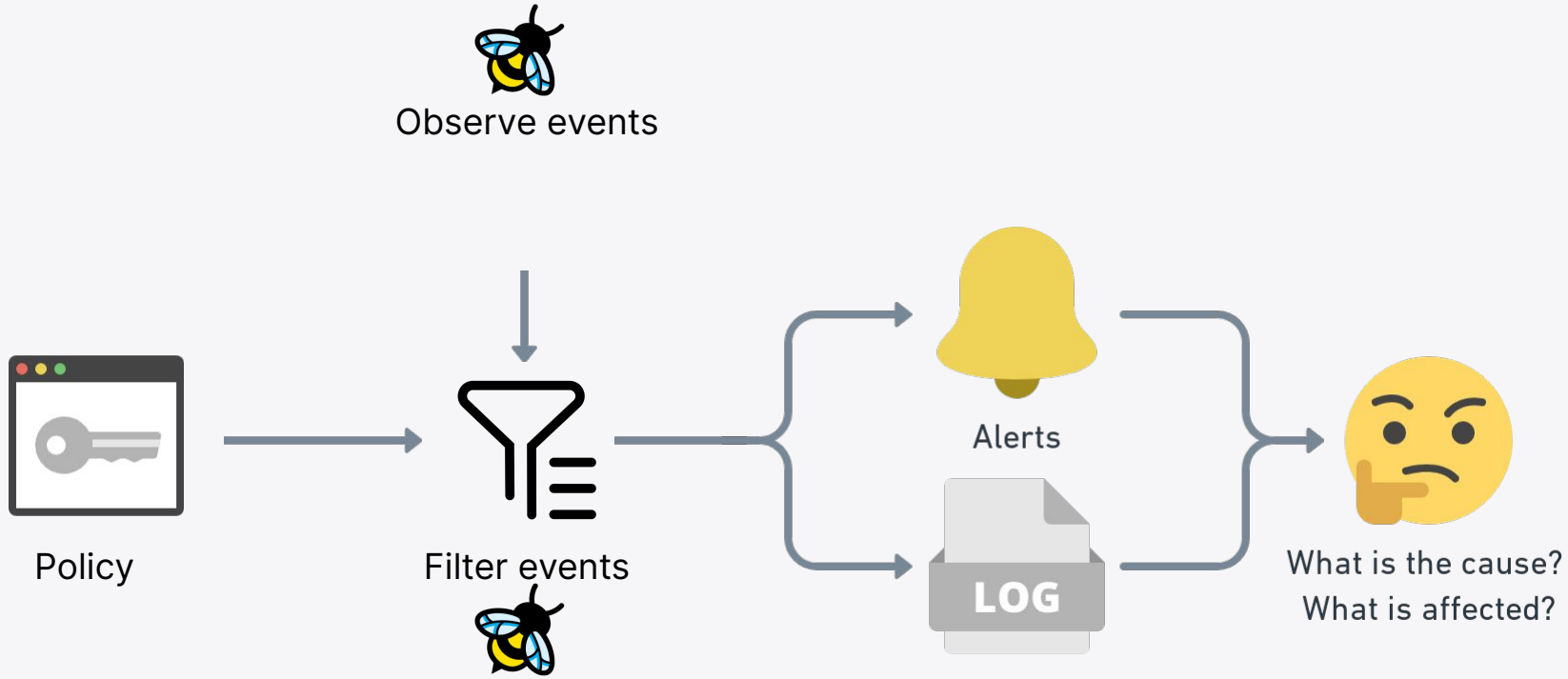
@lizrice

Taking eBPF observability tools as-is and using them for security monitoring would be like **driving your car into the ocean and expecting it to float**

- Brendan Gregg

ISOVALENT

Security observability



ISOVALENT

Cilium Tetragon observes security events

```
$ kubectl logs tetragon-74ffc -c export-stdout -f | tetra getevents -o compact
🚀 process default/xwing /usr/bin/vi /etc/passwd
📬 open    default/xwing /usr/bin/vi /etc/passwd
📪 close   default/xwing /usr/bin/vi
📬 open    default/xwing /usr/bin/vi /etc/passwd
📝 write   default/xwing /usr/bin/vi /etc/passwd 1275 bytes
📪 close   default/xwing /usr/bin/vi
💥 exit    default/xwing /usr/bin/vi /etc/passwd 0
```

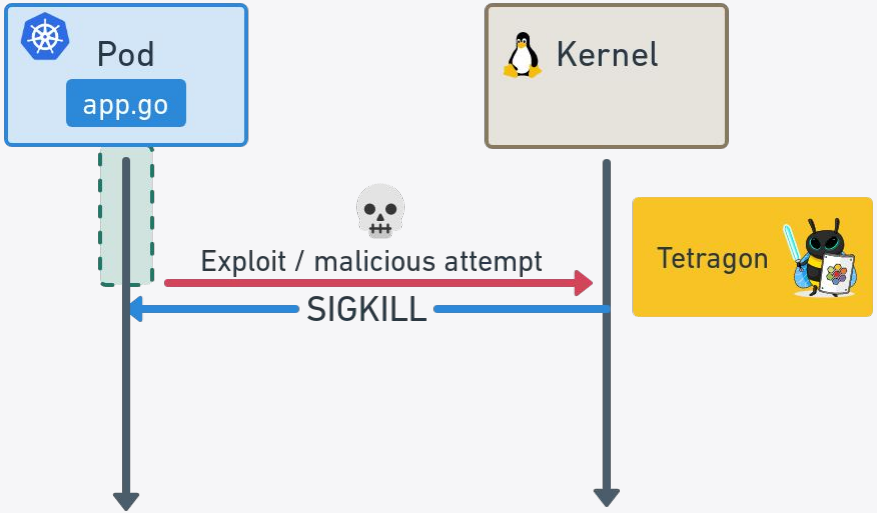
Policy events

Kubernetes info

@lizrice

ISOVALENT

Preventative actions from kernel

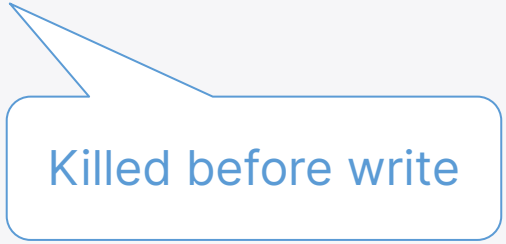


@lizrice

ISOVALENT

Cilium Tetragon preventative security

```
$ kubectl logs tetragon-74ffc -c export-stdout -f | tetra getevents -o compact
🚀 process default/xwing /usr/bin/vi /etc/passwd
📬 open    default/xwing /usr/bin/vi /etc/passwd
📪 close   default/xwing /usr/bin/vi
📬 open    default/xwing /usr/bin/vi /etc/passwd
📝 write   default/xwing /usr/bin/vi /etc/passwd 1269 bytes
💥 exit    default/xwing /usr/bin/vi /etc/passwd SIGKILL
```



ISOVALENT

eBPF enables powerful Cloud Native tools

- High performance observability, networking and security
- Dynamic instrumentation - zero app modifications
- Contextual information, Kubernetes identity-aware

ISOVALENT

eBPF in the CNCF

@lizrice

ISOVALENT

Thank you

cilium/cilium

@ciliumproject

cilium.io

@lizrice