Honeypot Data

Process calls from Ubuntu and Hackers alike…

Kate Highnam

What is a Honeypot?

A pleasant looking trap for unpleasant people

- Vulnerable system/infrastructure open to exploitation to expose/observe hackers/scripts/malware in the wild
- Virtual or real
- Entire system or network structure

Honeypots

High-interaction, virtual, … so sweet

Virtual Honeypots

High-interaction, virtual, … so sweet

Kernel-level Calls

Virtual Honeypots

High-interaction, virtual, … so sweet

Kernel-level Calls

Network Activity

Virtual Honeypots

High-interaction, virtual, … so sweet

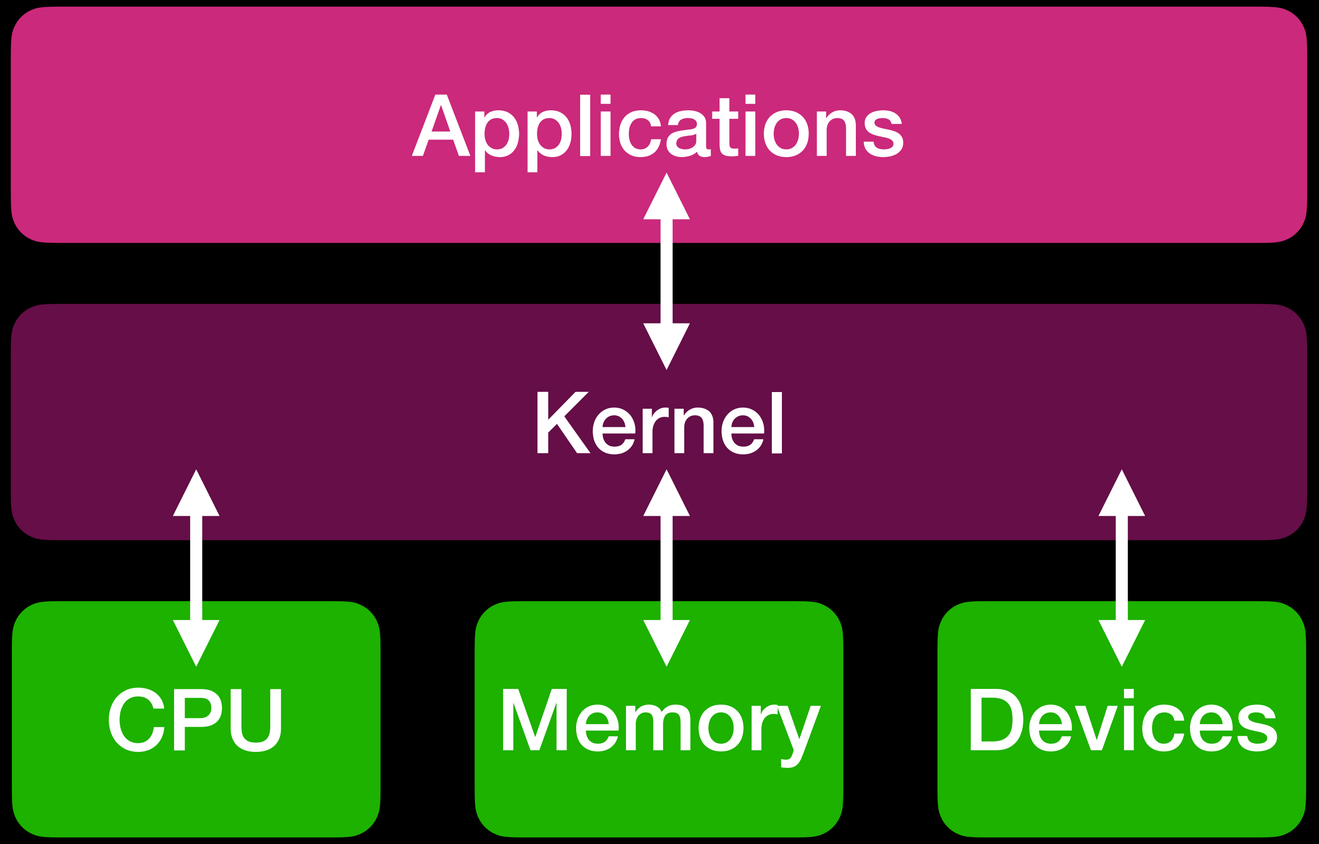
Kernel-level Calls

Network Activity

Virtual Honeypots

What is a kernel-level call?

Kernel-level Calls

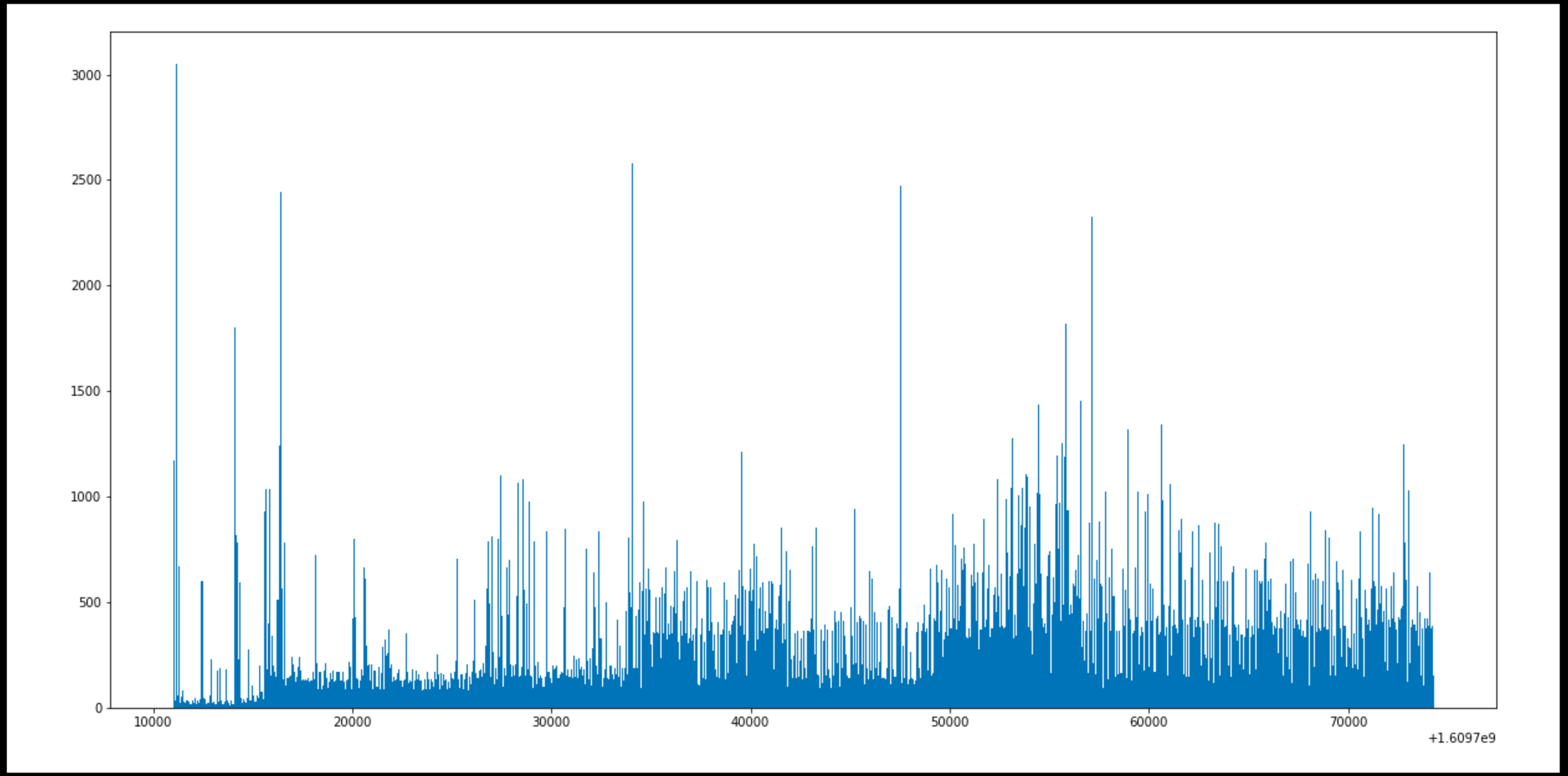


Dataset

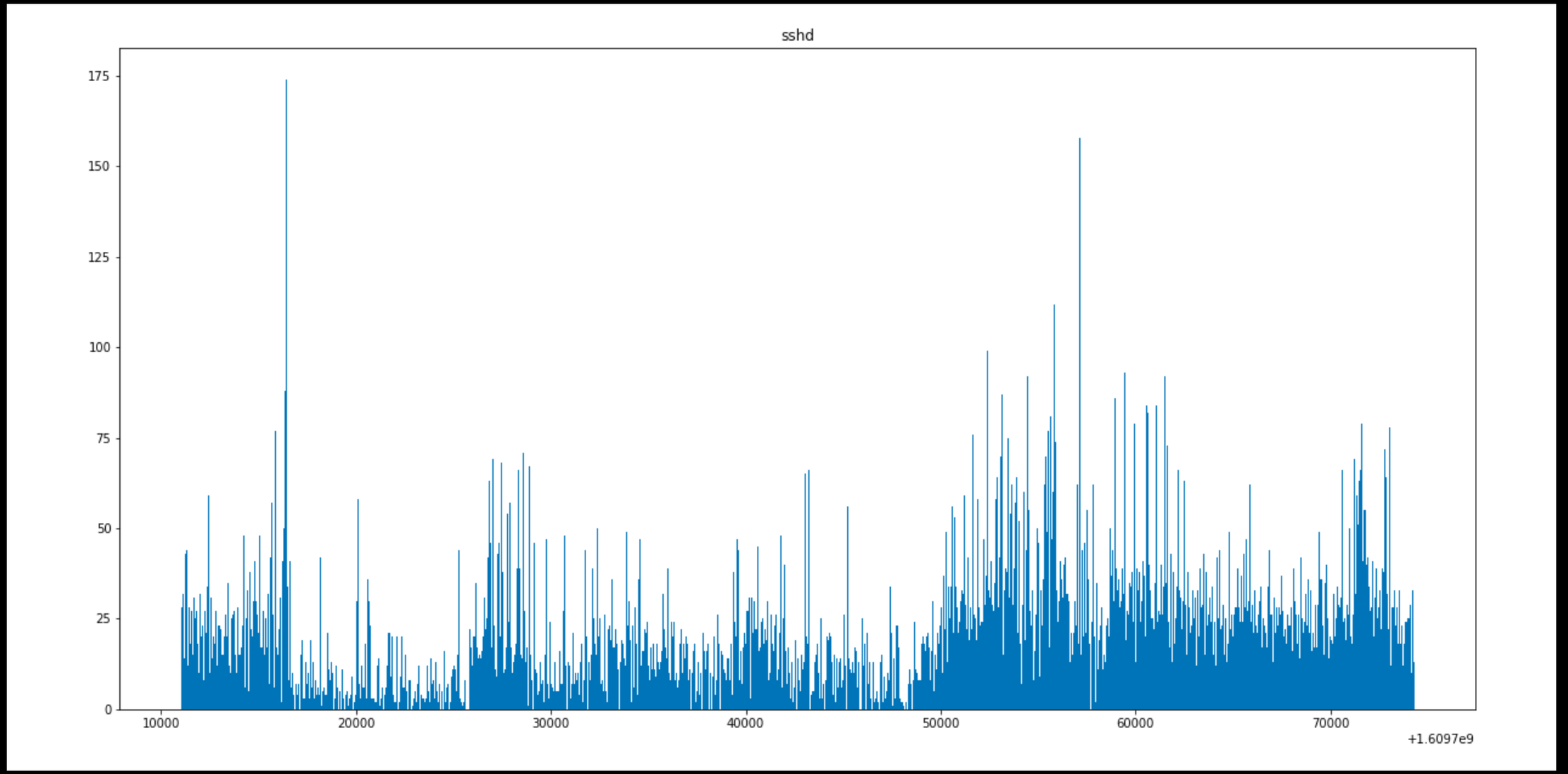
Process level calls within an Ubuntu 18.04 environment with Systems DNS turned off…

```
on_message_userdata, b'{"event_type": 0,
   "start_time": 1631926664960114,
   "event_time": 1631926665301538,
   "pid": 7419,
   "tid": 7429,
   "ppid": 3161,
   "exit_code": 0,
   "sig_info": 0,
   "loginuid": 4294967295, // not necessarily unique to each user, but there are multiple
   "fileumask": "0o22", // file permissions used
   "task": "curl", // string name of process task
   "age": 0.000341424 // how long it took to complete
}'
```

Timeline


3000
2500
2000
1500
1000
500
0
10000
20000
30000
40000
50000
60000
70000
+1.6097e9

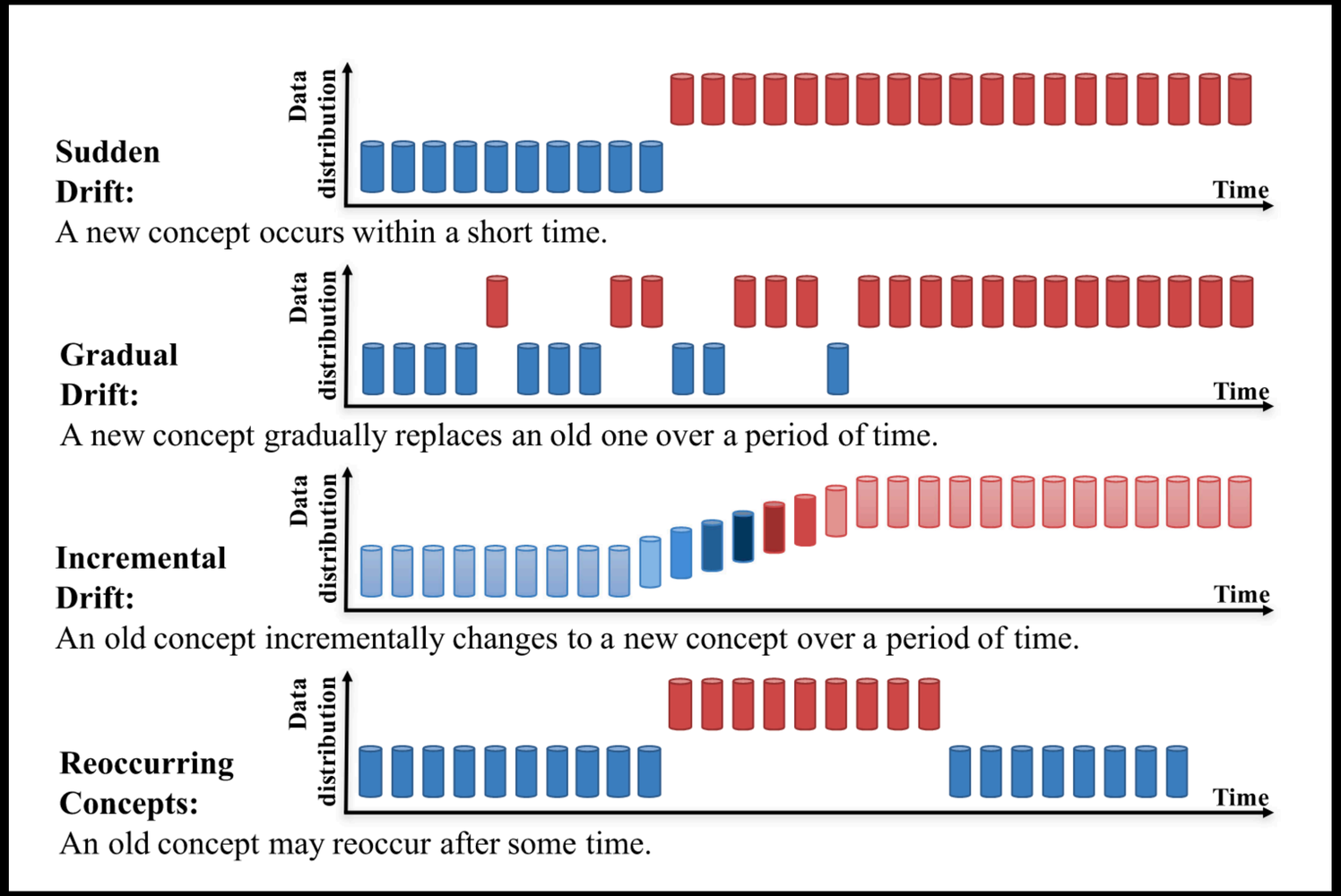
Timeline


sshd
175
150
125
100
75
50
25
0
10000
20000
30000
40000
50000
60000
70000
+1.6097e9

Concept Drift

When the underlying data-generating distribution changes

$$\exists t : P_t \neq P_{t+1}$$



Lu, J., Liu, A., Dong, F., Gu, F., Gama, J., and Zhang, G. (2018). Learning under concept drift: A review. IEEE Transactions on Knowledge and Data Engineering, 31(12):2346–2363.