

etcd的高可用实践

光音网络 王鹏(Tad)

Agenda

- etcd是什么
- Raft的原理
- 如何安装etcd
- 构建高可用etcd
- 我们踩过的坑

What's etcd?

- A Highly-Available Key Value Store for Shared Configuration and Service Discovery

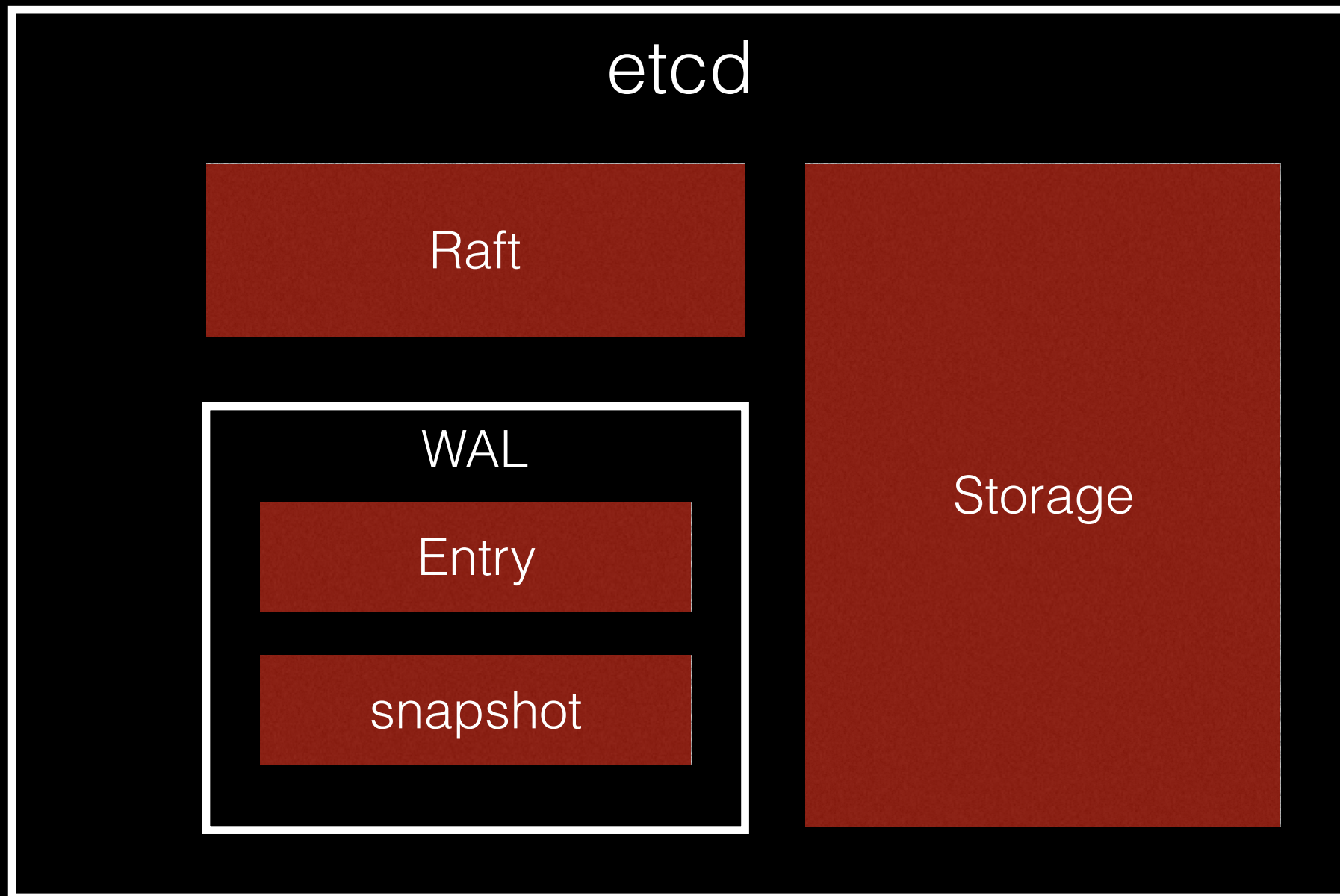
Features

- Raft
- TTL
- Watching a prefix
- Directories
- TTL
- RESTful API

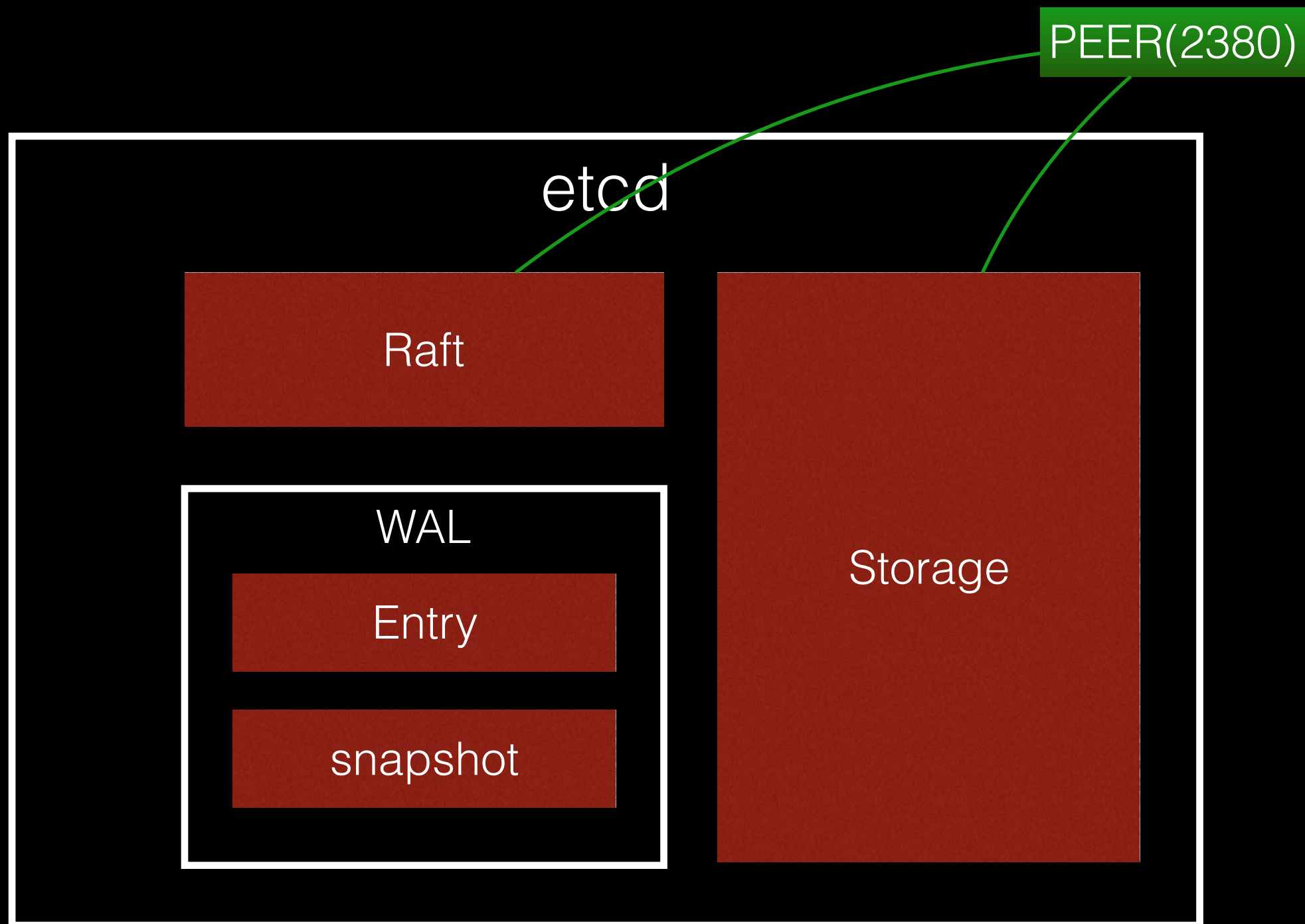
Operations

- Get
- Set
- Watch
- Recursive
- Directory
- TTL

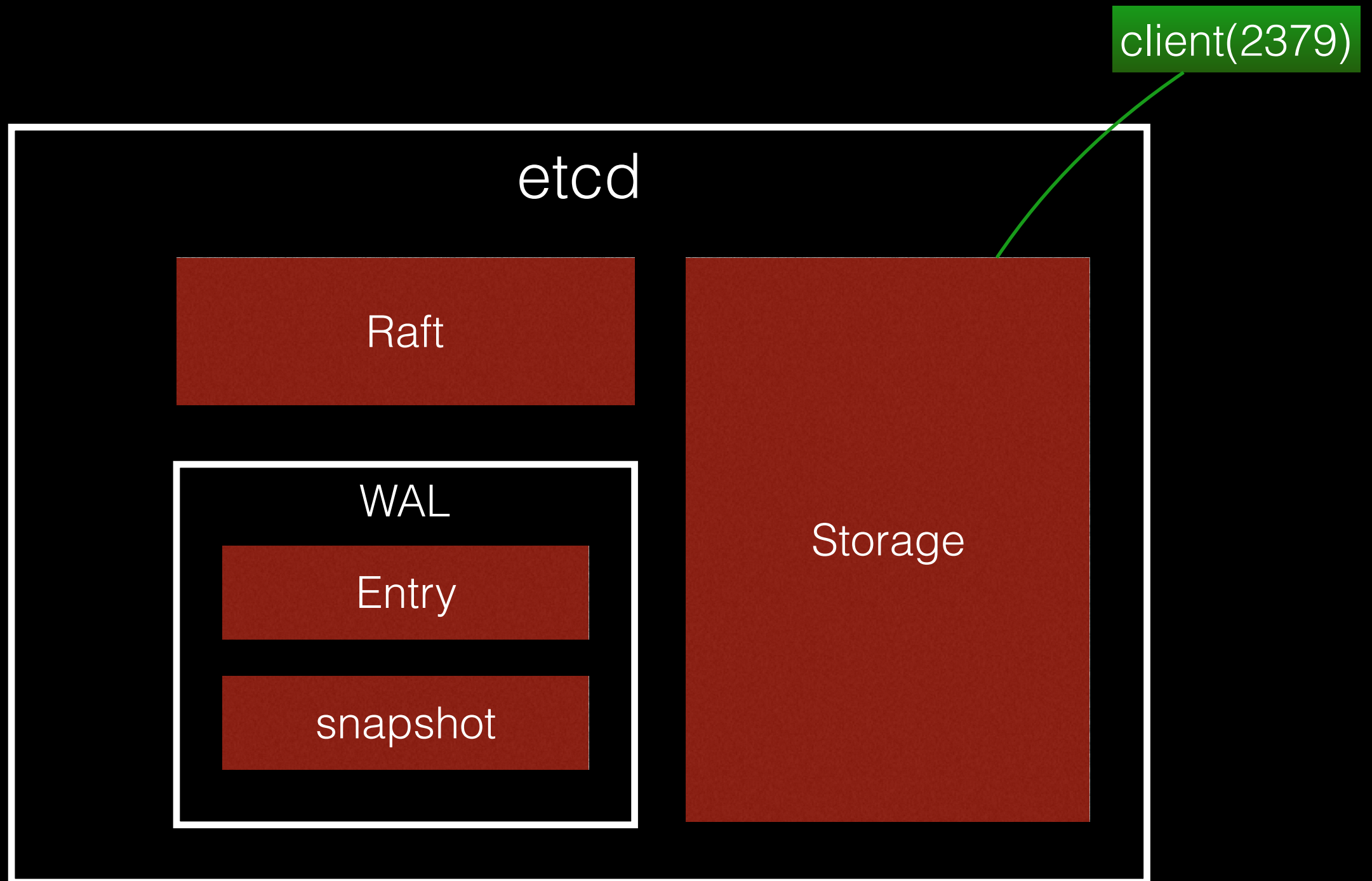
Framework



Framework



Framework



Raft



Raft—致性算法

- DEMO
- <http://thesecretlivesofdata.com/raft/#home>



如何安装etcd

- Static
- Public Discovery Service
- DNS Discovery

如何安装etcd

- name:
- advertise-client-urls:
- listen-peer-urls:
- listen-client-urls:
- initial-advertise-peer-urls:
- initial-cluster:
- data-dir:

如何安装etcd

- DEMO
- <https://github.com/Goyoo/etcd-demo>



DEMO

- Single
- 5 Nodes Cluster
- 1 proxy-mode



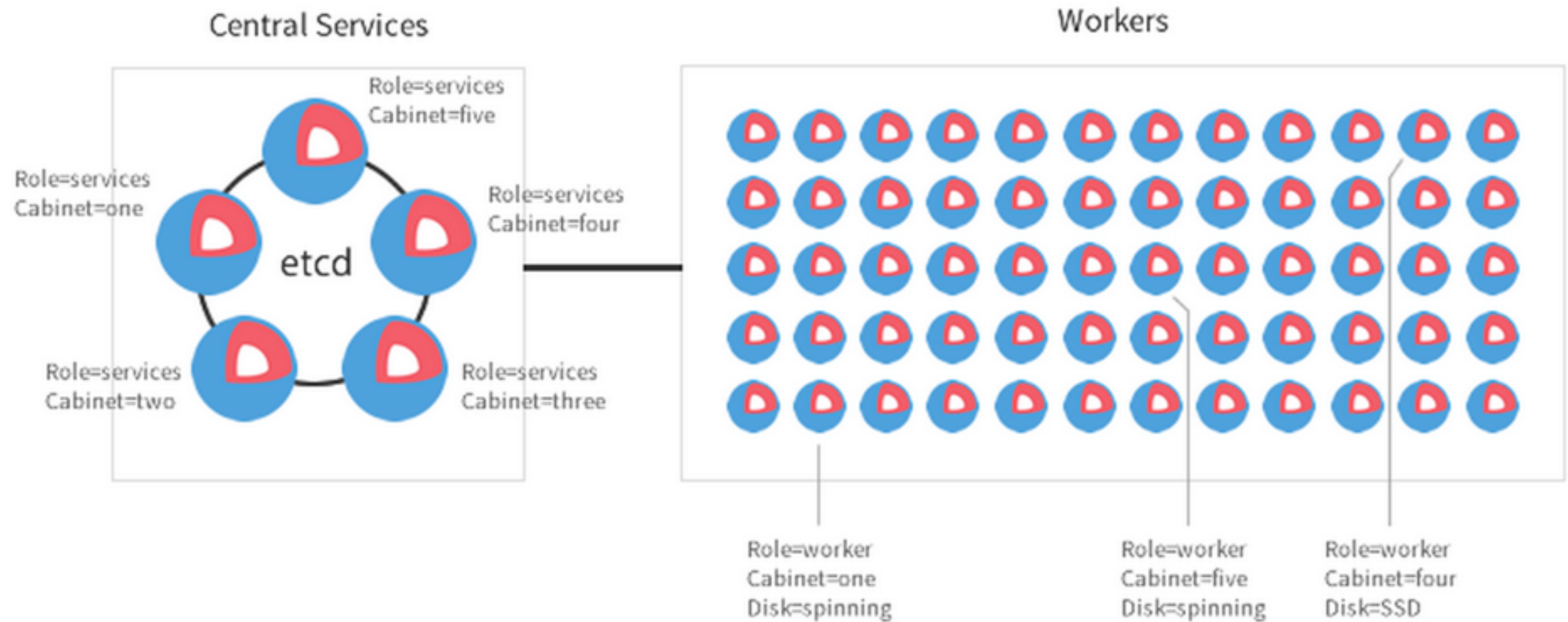
High-Available

- 5+ node
- Heartbeat/election timeout
- GOMAXPROC

High-Available

- Backup / Restore
- Add / Remove Node

Production Cluster with Central Services



Production Cluster

Tuning

- Metrics: <https://coreos.com/etcd/docs/latest/metrics.html>
- etcd-top: <https://github.com/mesosphere/etcd-top>

Benchmark (v2.2.0)

Reading Single Key

KEY SIZE IN BYTES	NUMBER OF CLIENTS	TARGET ETCD SERVER	READ QPS	90TH PERCENTILE LATENCY (MS)
64	1	leader only	2804 (-5%)	0.4 (+0%)
64	64	leader only	17816 (+0%)	5.7 (-6%)
64	256	leader only	18667 (-6%)	20.4 (+2%)
256	1	leader only	2181 (-15%)	0.5 (+25%)
256	64	leader only	17435 (-7%)	6.0 (+9%)
256	256	leader only	18180 (-8%)	21.3 (+3%)
64	64	all servers	46965 (-4%)	2.1 (+0%)
64	256	all servers	55286 (-6%)	7.4 (+6%)
256	64	all servers	46603 (-6%)	2.1 (+5%)
256	256	all servers	55291 (-6%)	7.3 (+4%)

Benchmark (v2.2.0)

Writing Single Key

KEY SIZE IN BYTES	NUMBER OF CLIENTS	TARGET ETCD SERVER	WRITE QPS	90TH PERCENTILE LATENCY (MS)
64	1	leader only	76 (+22%)	19.4 (-15%)
64	64	leader only	2461 (+45%)	31.8 (-32%)
64	256	leader only	4275 (+1%)	69.6 (-10%)
256	1	leader only	64 (+20%)	16.7 (-30%)
256	64	leader only	2385 (+30%)	31.5 (-19%)
256	256	leader only	4353 (-3%)	74.0 (+9%)
64	64	all servers	2005 (+81%)	49.8 (-55%)
64	256	all servers	4868 (+35%)	81.5 (-40%)
256	64	all servers	1925 (+72%)	47.7 (-59%)
256	256	all servers	4975 (+36%)	70.3 (-36%)

FROM: <https://coreos.com/etcd/docs/latest/benchmarks/etcd-2-2-0-rc-benchmarks.html>

Benchmark (v2.2.0)

Memory Maximal Usage

VALUE BYTES	KEY NUMBER	DATA SIZE(MB)	MAX RSS(MB)	MAX RSS/DATA RATE ON LEADER
128	50000	6	433	72x
128	100000	12	659	54x
128	200000	24	1466	61x
1024	50000	48	1253	26x
1024	100000	96	2344	24x
1024	200000	192	4361	22x

Benchmark (v2.2.0)

Data Size Threshold

VALUE BYTES	KEY NUMBER LIMITATION	SUGGESTED DATA SIZE THRESHOLD(MB)	CONSUMED RSS(MB)
128	400K	48	2400
1024	300K	292	6500

我们踩过的坑

- 不建议运行在Docker里
- Kubernetes升级时，注意etcd里的版本等信息
- 线上使用前，熟练增删节点的操作
- 避免频繁rget操作
- proxy-mode时，切换集群的问题

Thank You!

- We Are Hiring!
- www.goyoo.com

