

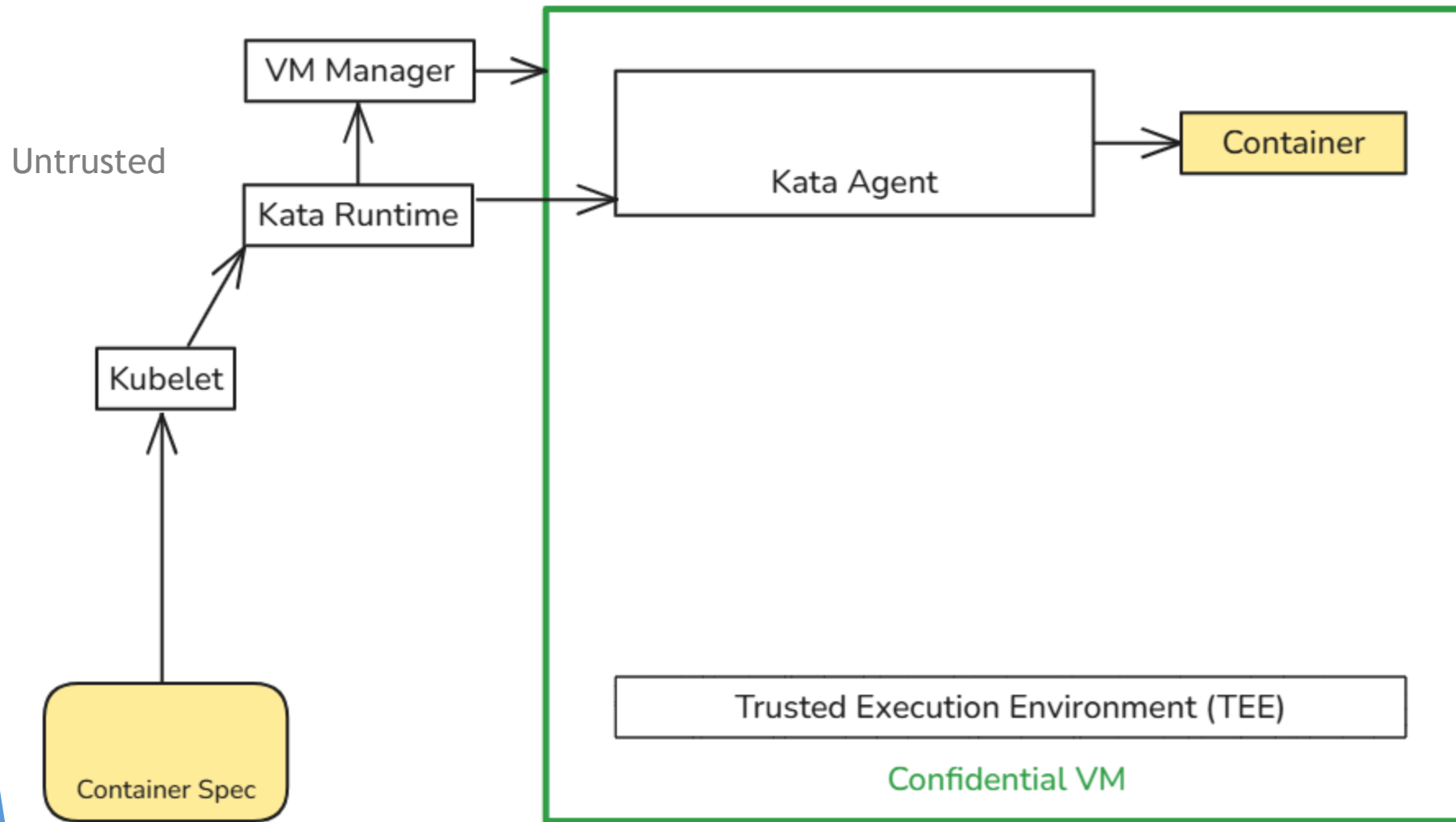
Trust No One: Secure Storage for Confidential Containers

Aurélien Bombo

Confidential Containers | Kata Containers | Microsoft

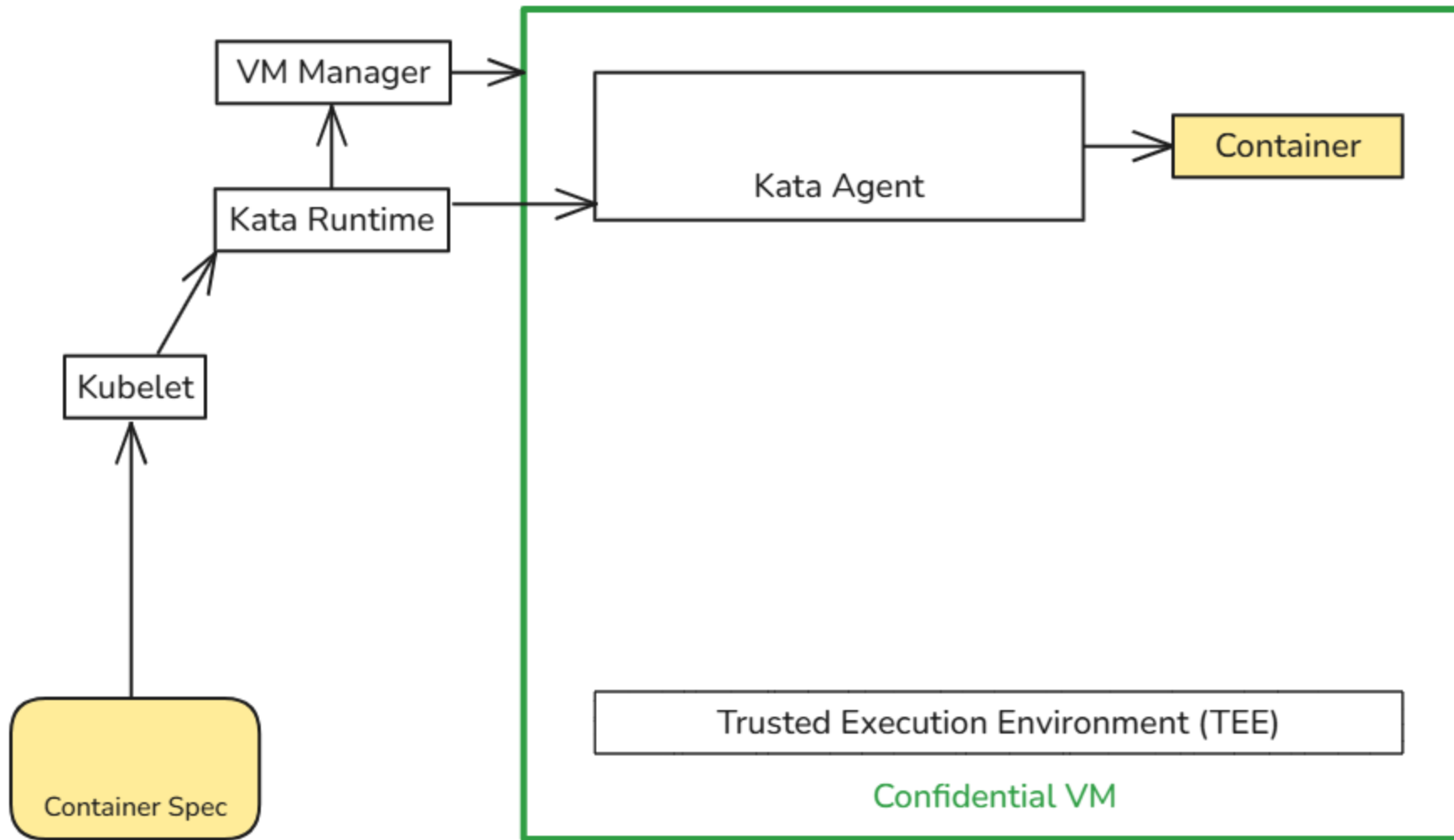
What is a confidential container?

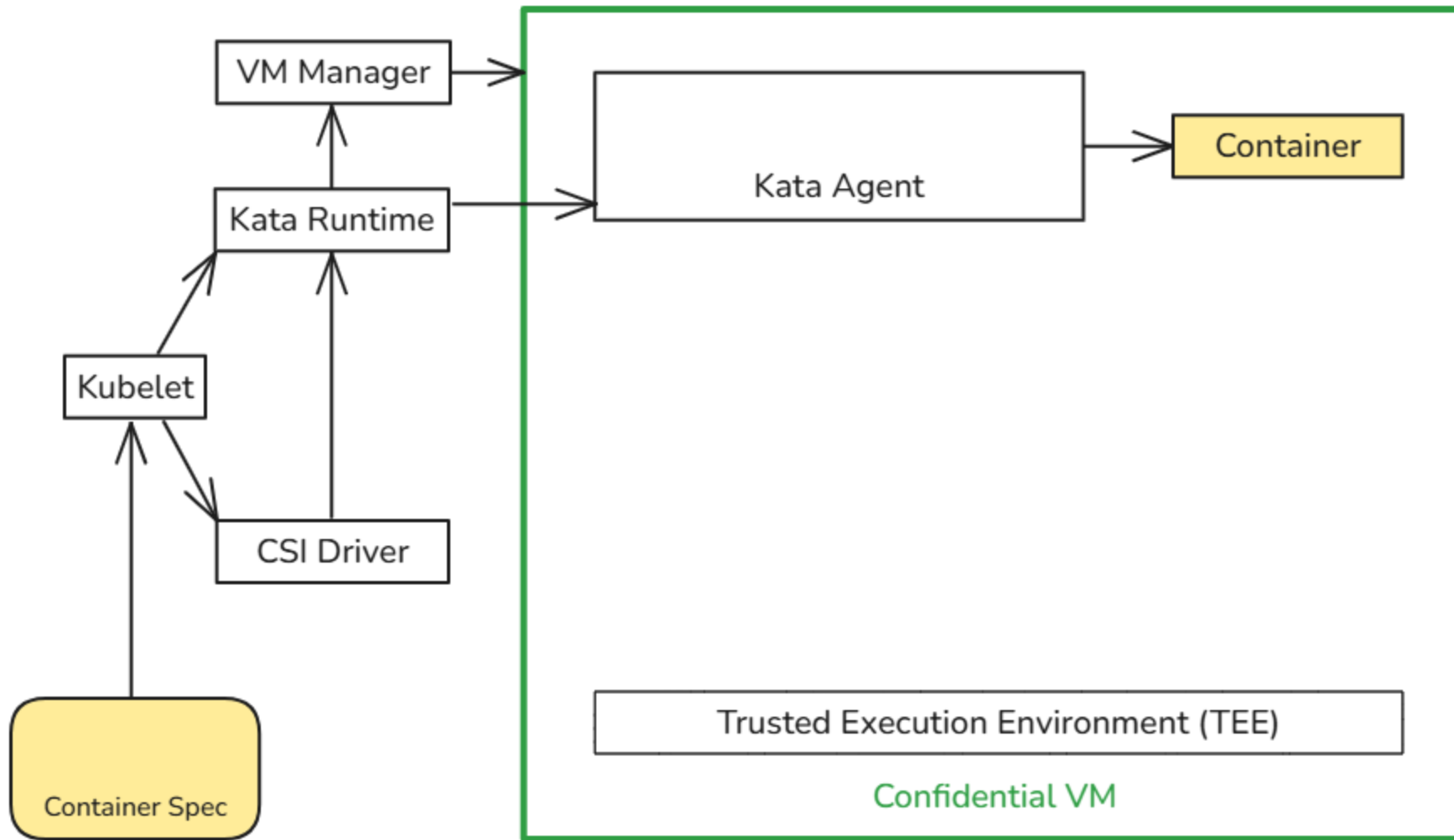
- ▶ runC containers
 - ▶ Shared kernel
- ▶ Kata Containers
 - ▶ Virtualization
- ▶ **Confidential Containers (CoCo)**
 - ▶ Builds on top of Kata Containers
 - ▶ Trusted Execution Environment (TEE)
 - ▶ Remote Attestation
 - ▶ Security Policy
- ▶ **Kubernetes**
 - ▶ Container Orchestrator
 - ▶ Storage



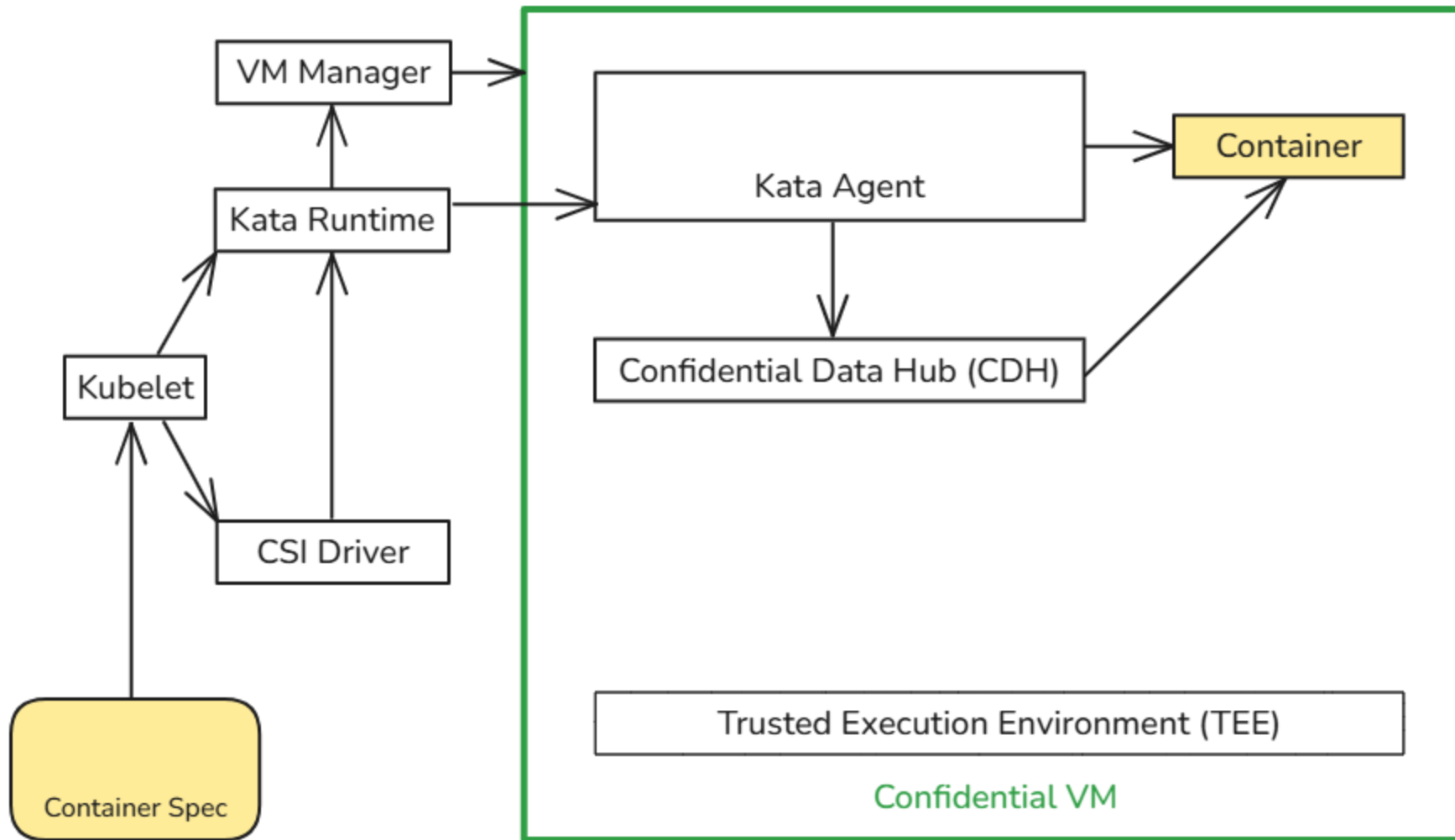
Ephemeral storage

- ▶ Data is short-lived, but does not fit in memory
 - ▶ Sharing between containers inside VM
- ▶ Goals
 - ▶ Confidentiality
 - ▶ Kubernetes integration
- ▶ Design
 - ▶ Create block device on host and pass to VM
 - ▶ Encrypt & format device inside VM
- ▶ Challenge
 - ▶ Secure VM boundary: security policy

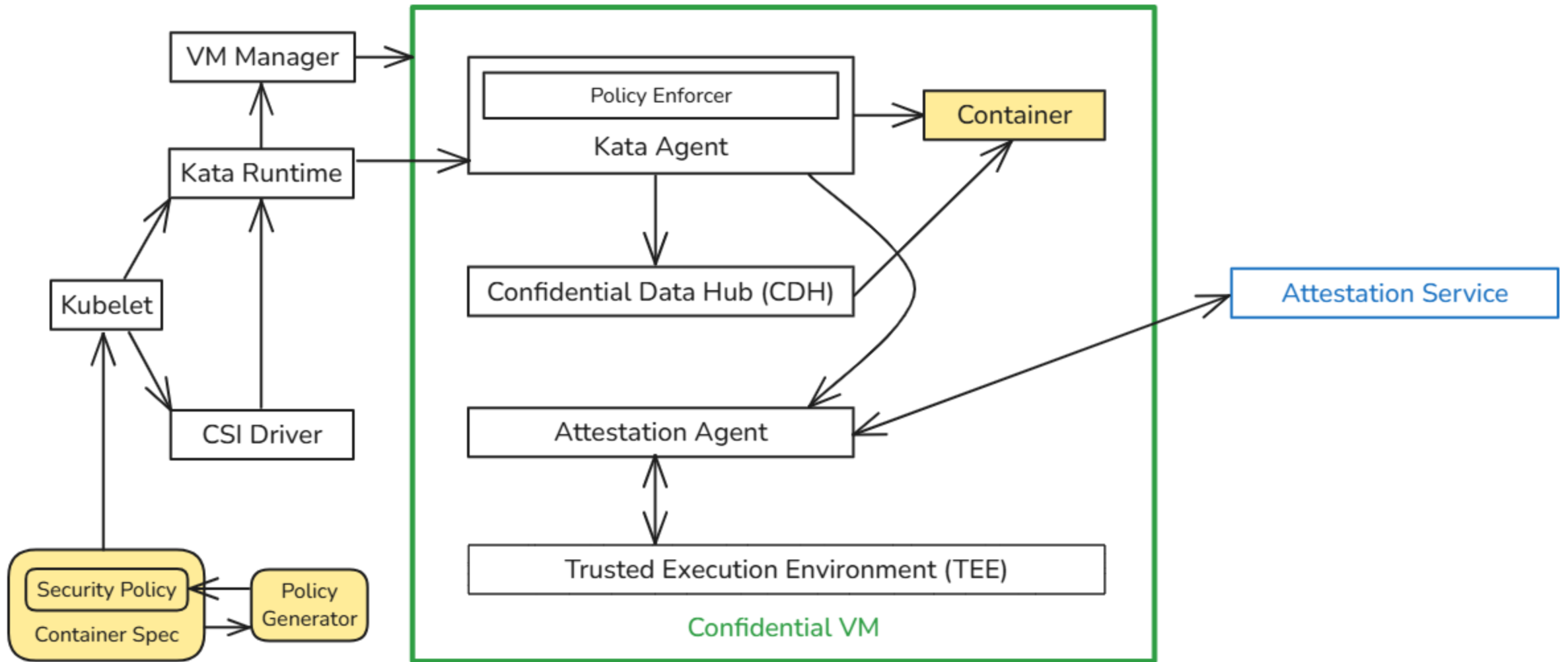




virto-blk



CDH uses
dm-crypt & dm-integrity




```
$ vim my-app.yaml
```

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmIuaHQgKGmpIDIwMjMgTWljcm9zb2Z0IENvcnBvcnF0aW9uCiMKIy
8 spec:
9   volumes:
10     - name: encrypted
11       ephemeral:
12         volumeClaimTemplate:
13           spec:
14             accessModes: [ReadWriteOncePod]
15             storageClassName: coco-local-csi
16             resources:
17               requests:
18                 storage: 10G
19   containers:
20     - volumeMounts:
21       - name: encrypted
22         mountPath: /mnt/encrypted
23       image: busybox:latest
24       name: my-container
25       command: ["sleep", "infinity"]
26 runtimeClassName: kata
~
~
~
```

```

1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmIuaHQgKGmpIDIwMjMgTWljcm9zb2Z0IENvcnBvcnF0aW9uCiMKIy
8 spec:
9   volumes:
10    - name: encrypted
11      ephemeral:
12        volumeClaimTemplate:
13          spec:
14            accessModes: [ReadWriteOncePod]
15            storageClassName: coco-local-csi
16            resources:
17              requests:
18                storage: 10G
19   containers:
20    - volumeMounts:
21      - name: encrypted
22        mountPath: /mnt/encrypted
23      image: busybox:latest
24      name: my-container
25      command: ["sleep", "infinity"]
26 runtimeClassName: kata

```

```

~
~
~

```

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmIuaHQgKGmpIDIwMjMgTWljcm9zb2Z0IENvcnBvcnF0aW9uCiMKIy
8 spec:
9   volumes:
10     - name: encrypted
11       ephemeral:
12         volumeClaimTemplate:
13           spec:
14             accessModes: [ReadWriteOncePod]
15             storageClassName: coco-local-csi
16             resources:
17               requests:
18                 storage: 10G
19   containers:
20     - volumeMounts:
21       - name: encrypted
22         mountPath: /mnt/encrypted
23       image: busybox:latest
24       name: my-container
25       command: ["sleep", "infinity"]
26 runtimeClassName: kata
~
~
~
```

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmI naHQgKGMpIDIwMjMgTWljcm9zb2Z0IENvcnBvc mF0aW9uCiMKIy
8 spec:
9   volumes:
10     - name: encrypted
11       ephemeral:
12         volumeClaimTemplate:
13           spec:
14             accessModes: [ReadWriteOncePod]
15             storageClassName: coco-local-csi
16             resources:
17               requests:
18                 storage: 10G
19   containers:
20     - volumeMounts:
21       - name: encrypted
22         mountPath: /mnt/encrypted
23       image: busybox:latest
24       name: my-container
25       command: ["sleep", "infinity"]
26 runtimeClassName: kata
```

~
~
~

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmI naHQgKG MpIDIwMjMgTWl jcm9zb2Z0IENvcnBvc mF0aW9uCiMKIy
8 spec:
9   volumes:
10     - name: encrypted
11       ephemeral:
12         volumeClaimTemplate:
13           spec:
14             accessModes: [ReadWriteOncePod]
15             storageClassName: coco-local-csi
16             resources:
17               requests:
18                 storage: 10G
19   containers:
20     - volumeMounts:
21       - name: encrypted
22         mountPath: /mnt/encrypted
23       image: busybox:latest
24       name: my-container
25       command: ["sleep", "infinity"]
26 runtimeClassName: kata
```

~
~
~

```
$ vim my-app.yaml  
$ |
```

```
$ vim my-app.yaml  
$ kubectl apply -f my-app.yaml|
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4            8.9G         24.0K         8.4G         0% /mnt/encrypted
/ # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4            8.9G         24.0K         8.4G         0% /mnt/encrypted
/ # cd /mnt/encrypted
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l|
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root      16384 Jan 28 01:42 lost+found
/mnt/encrypted # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgZN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root      16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgZN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root      16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
/mnt/encrypted # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgZN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root      16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
/mnt/encrypted # ls -l
```

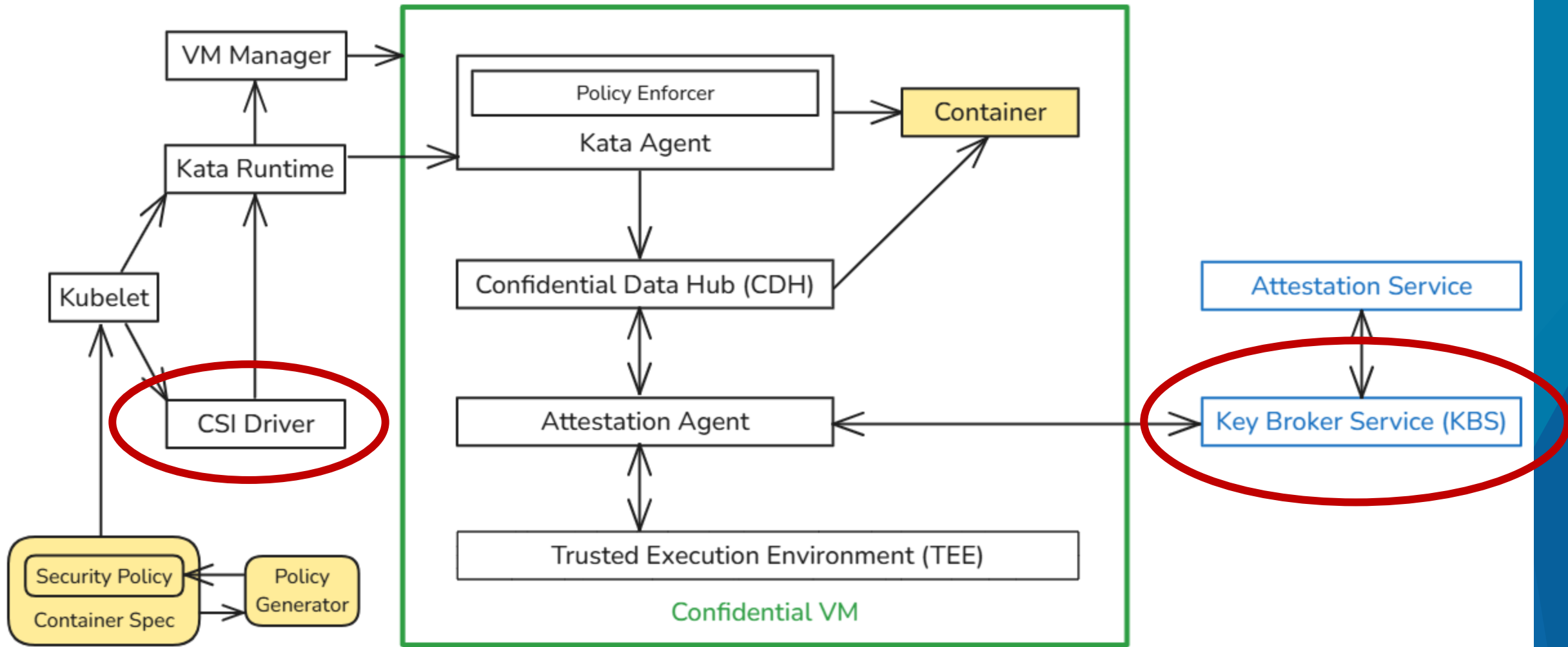
```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgZN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----    2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
/mnt/encrypted # ls -l
total 20
-rw-r--r--    1 root    root           14 Jan 28 01:46 ai_model.py
drwx-----    2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
/mnt/encrypted # ls -l
total 20
-rw-r--r--   1 root    root           14 Jan 28 01:46 ai_model.py
drwx-----  2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # cat ai_model.py |
```

```
$ vim my-app.yaml
$ kubectl apply -f my-app.yaml
pod/my-app created
$ kubectl exec -it my-app -- sh
/ #
/ # df -hTP | grep encrypted
/dev/mapper/encrypted_disk_4BgzN          ext4          8.9G      24.0K      8.4G      0% /mnt/encrypted
/ # cd /mnt/encrypted
/mnt/encrypted # ls -l
total 16
drwx-----  2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # echo 'print("true")' > ai_model.py
/mnt/encrypted # ls -l
total 20
-rw-r--r--   1 root    root           14 Jan 28 01:46 ai_model.py
drwx-----  2 root    root          16384 Jan 28 01:42 lost+found
/mnt/encrypted # cat ai_model.py
print("true")
/mnt/encrypted # |
```

Persistent storage

- ▶ Final design TBD
- ▶ Builds on ephemeral storage
- ▶ Key differences
 - ▶ CSI driver
 - ▶ Key Broker Service

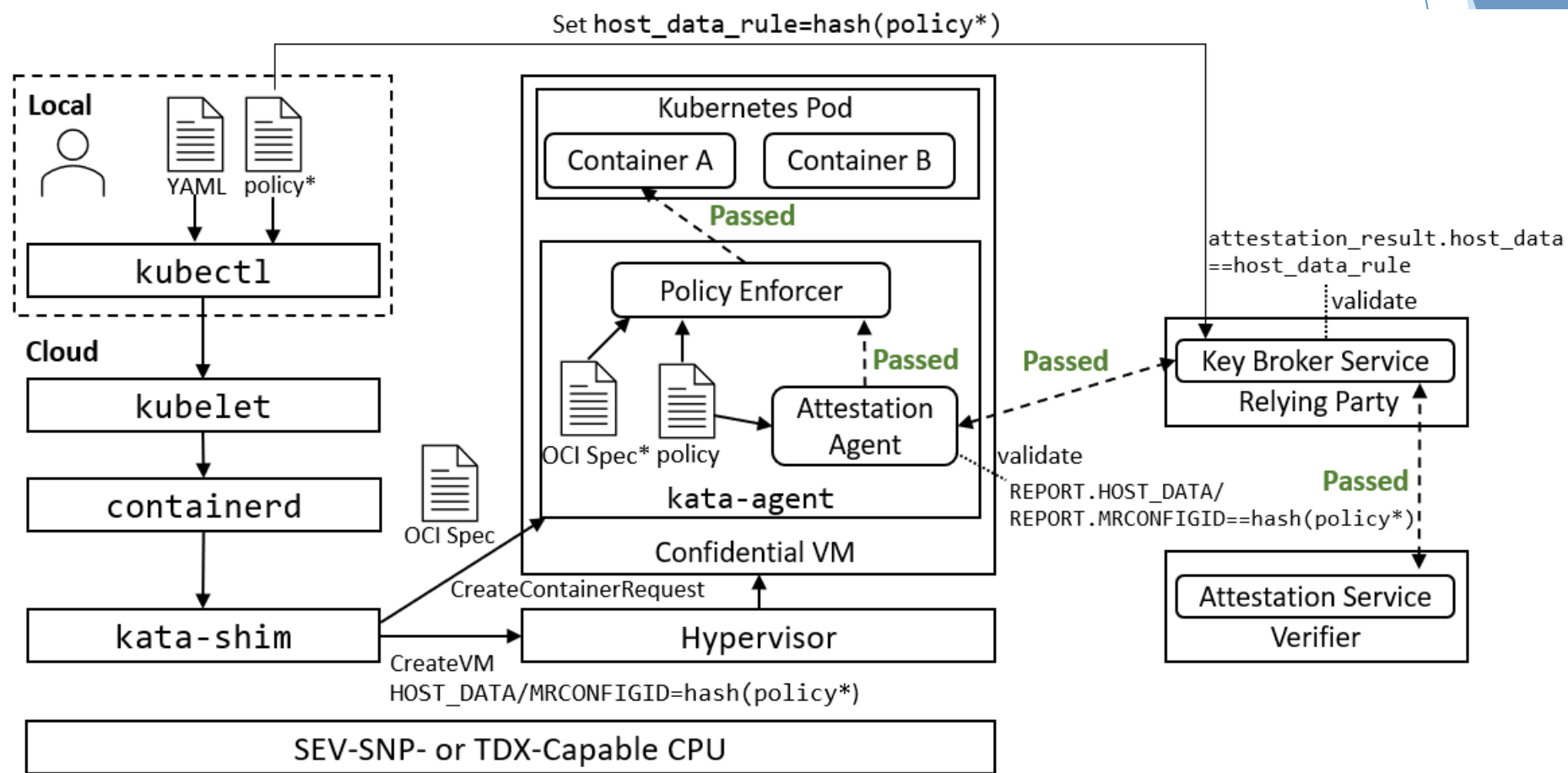


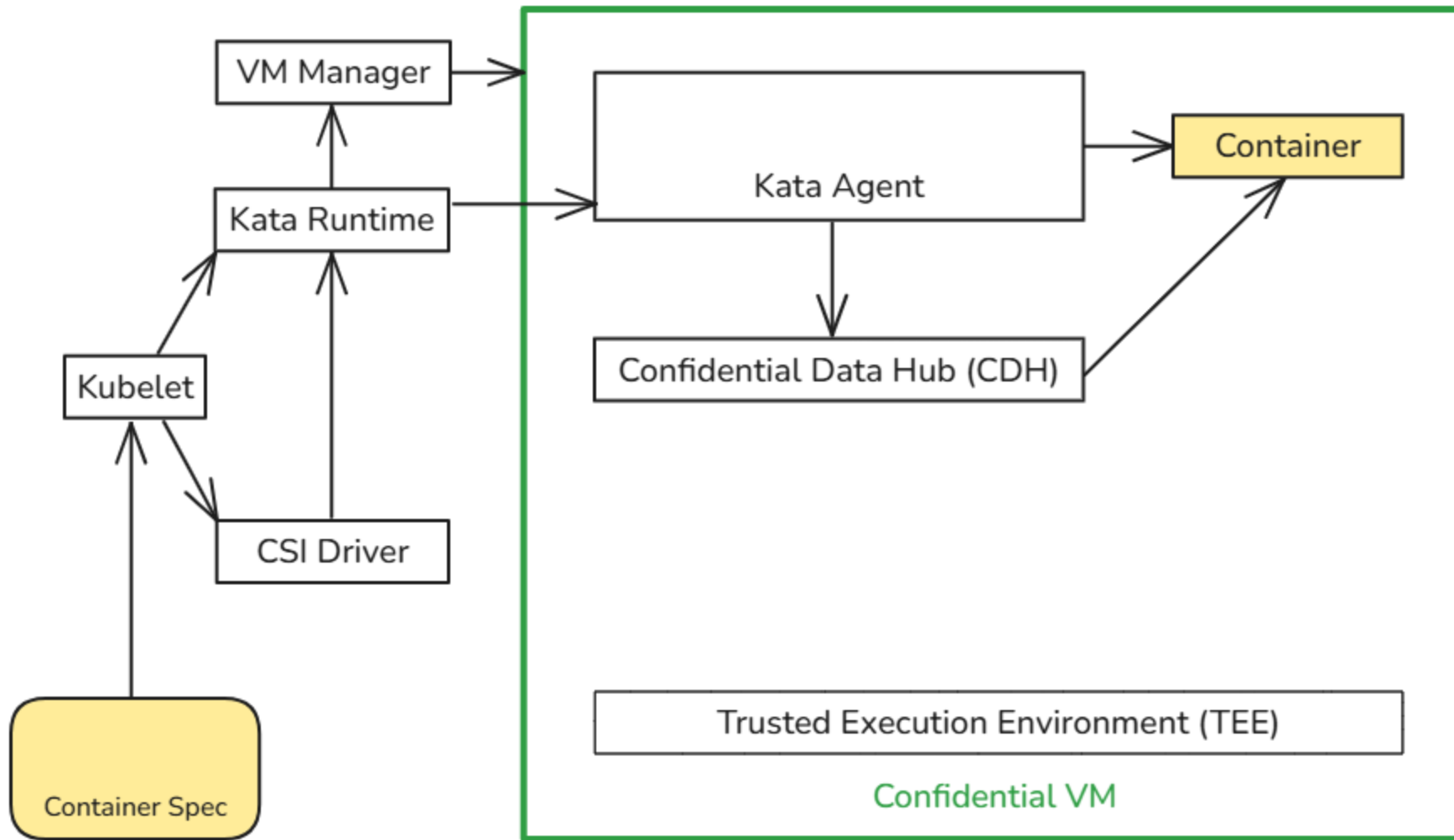
Questions?

Links

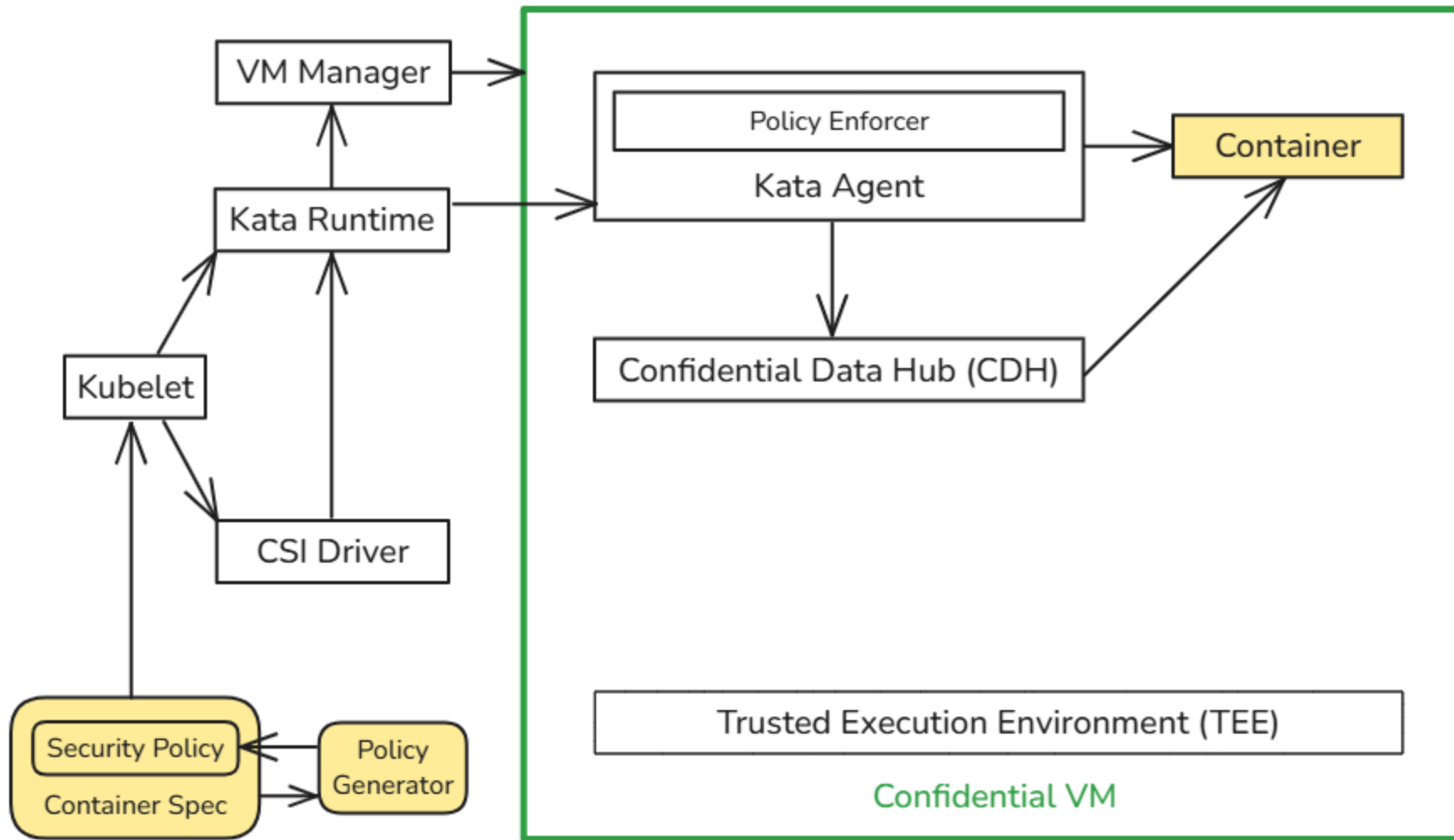
- ▶ PR (& proposal) to implement confidential ephemeral storage
 - ▶ github.com/kata-containers/kata-containers/pull/10559
- ▶ Confidential Containers website
 - ▶ confidentialcontainers.org
- ▶ Confidential Containers code
 - ▶ github.com/kata-containers/kata-containers
 - ▶ github.com/confidential-containers/guest-components

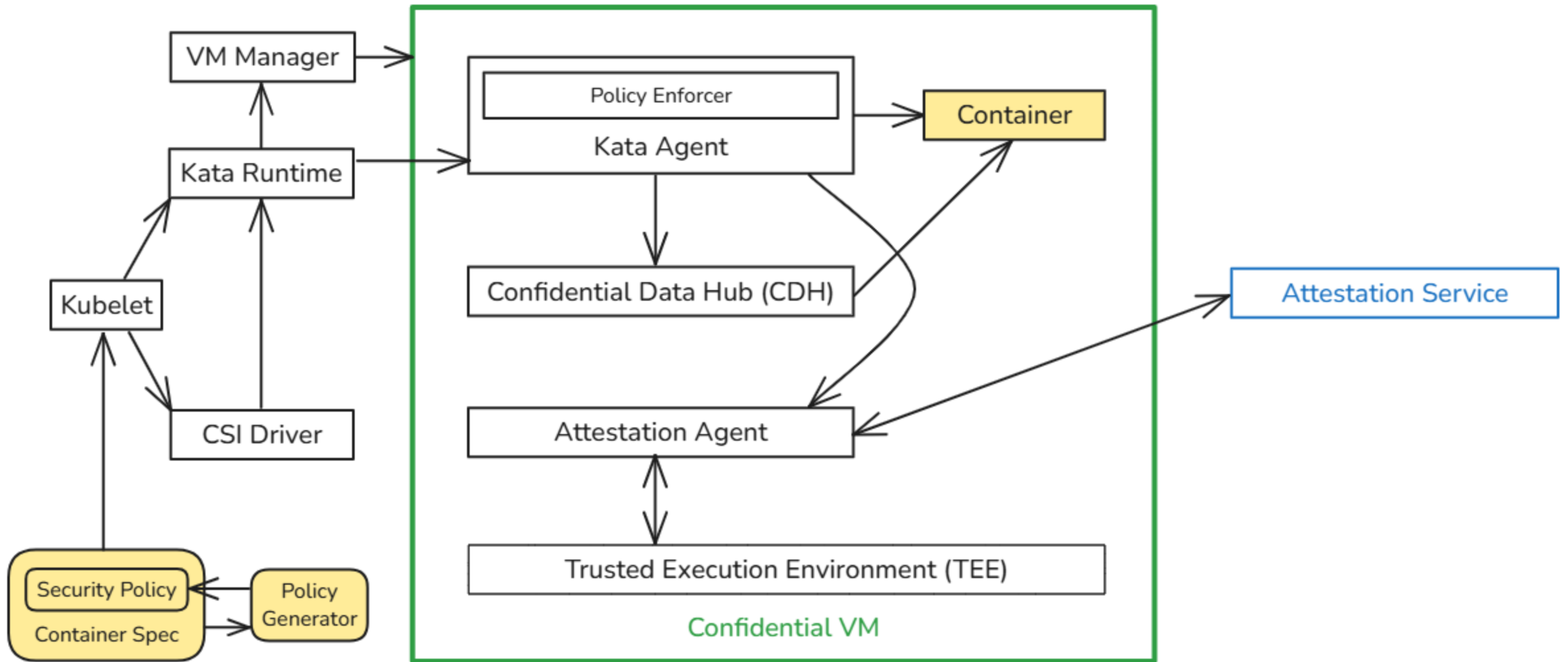
Policy validation



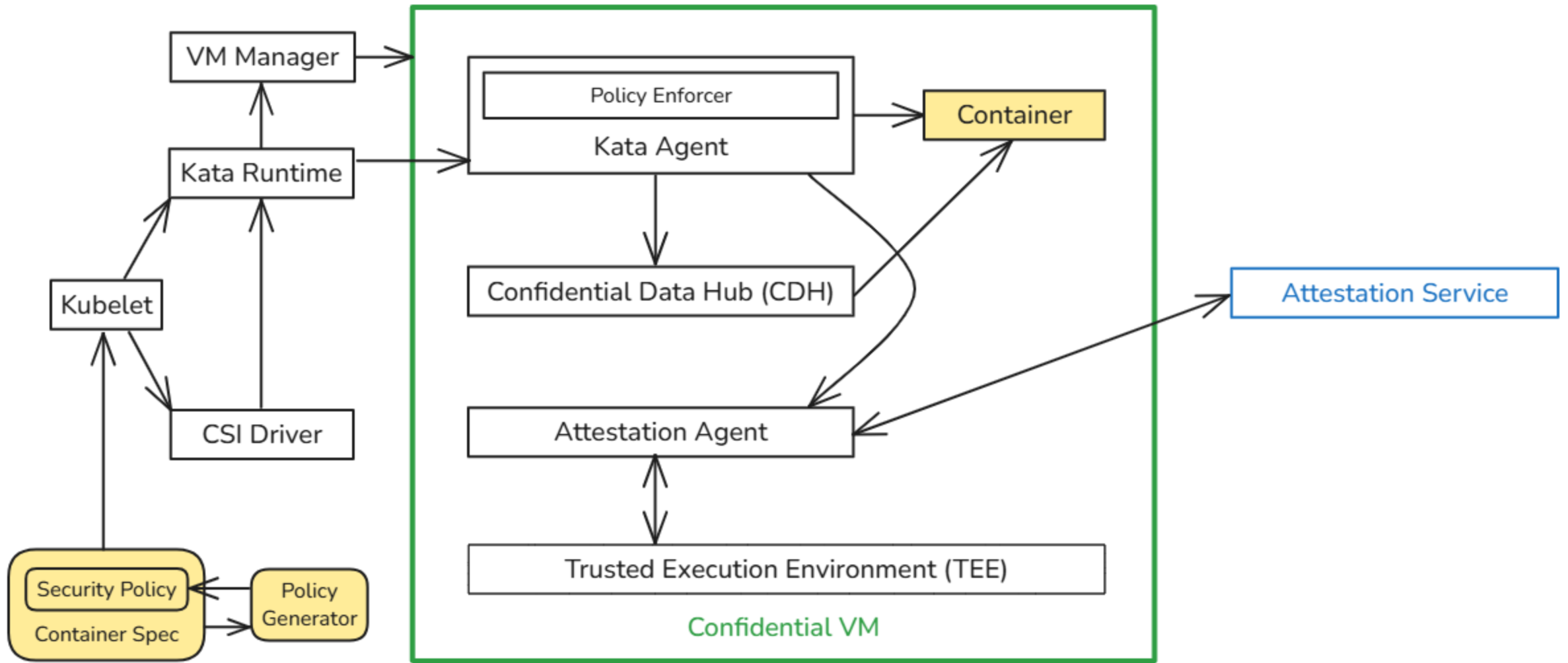


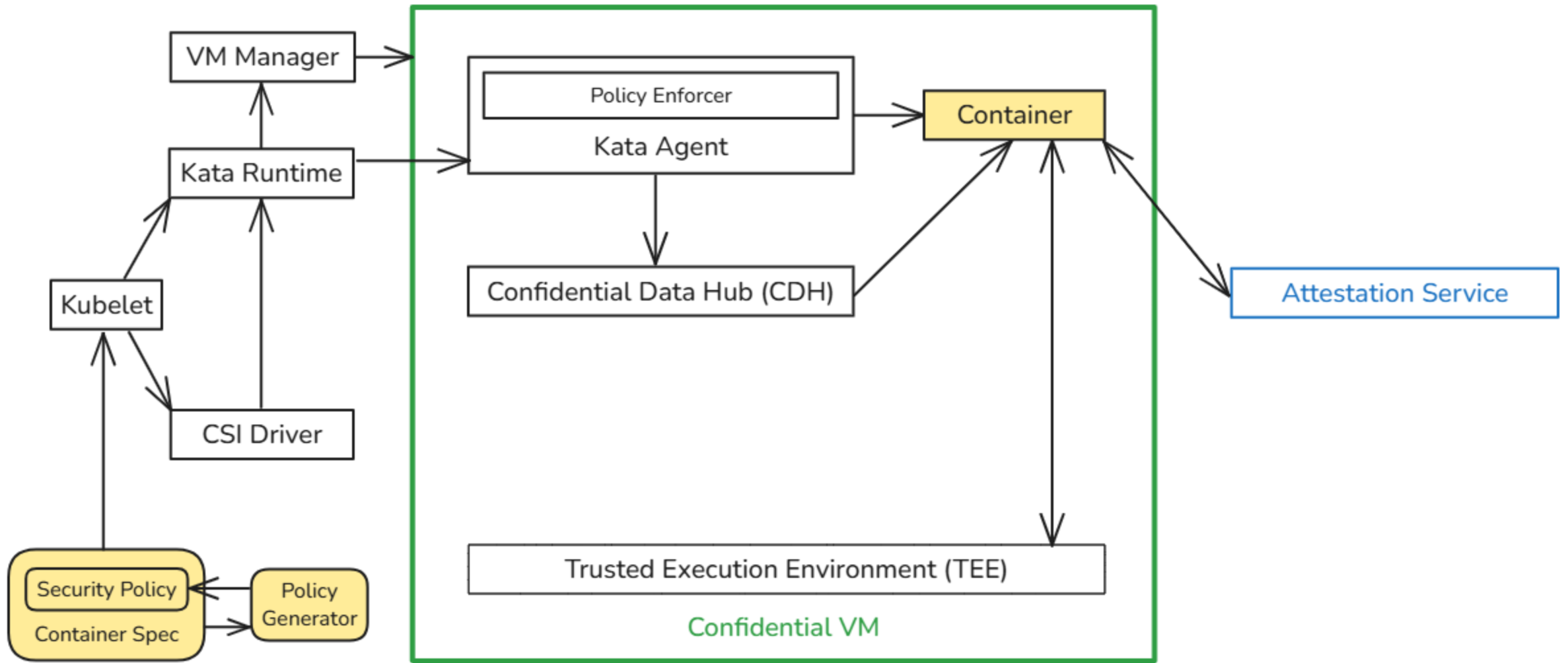
CDH uses
dm-crypt & dm-integrity



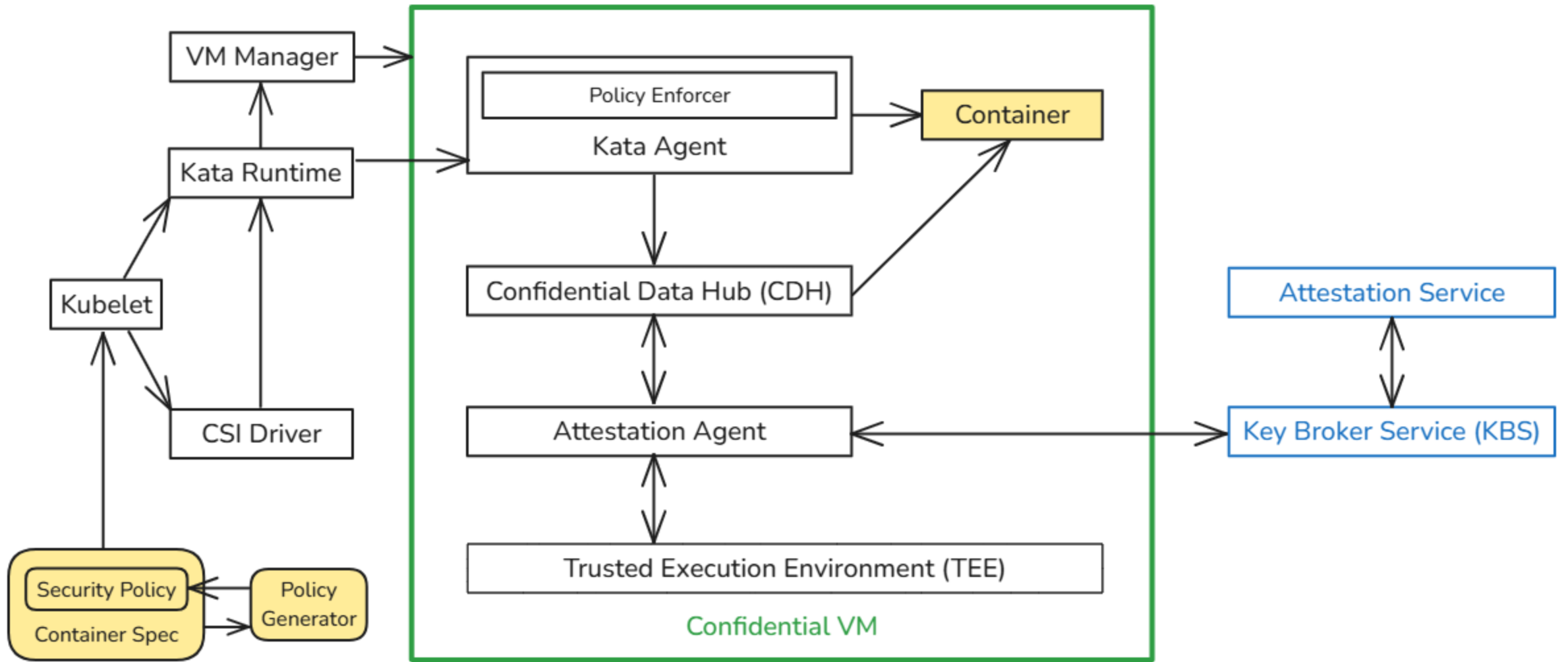


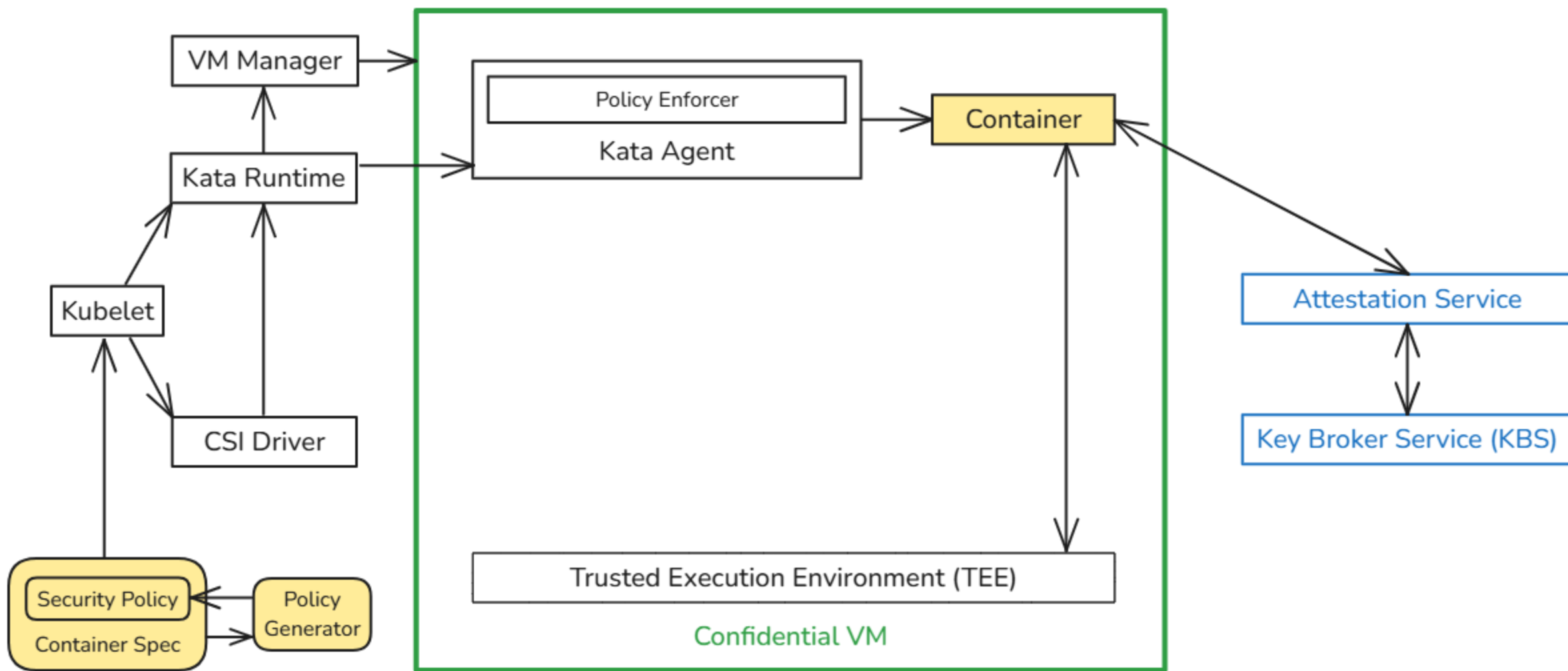
Ephemeral models

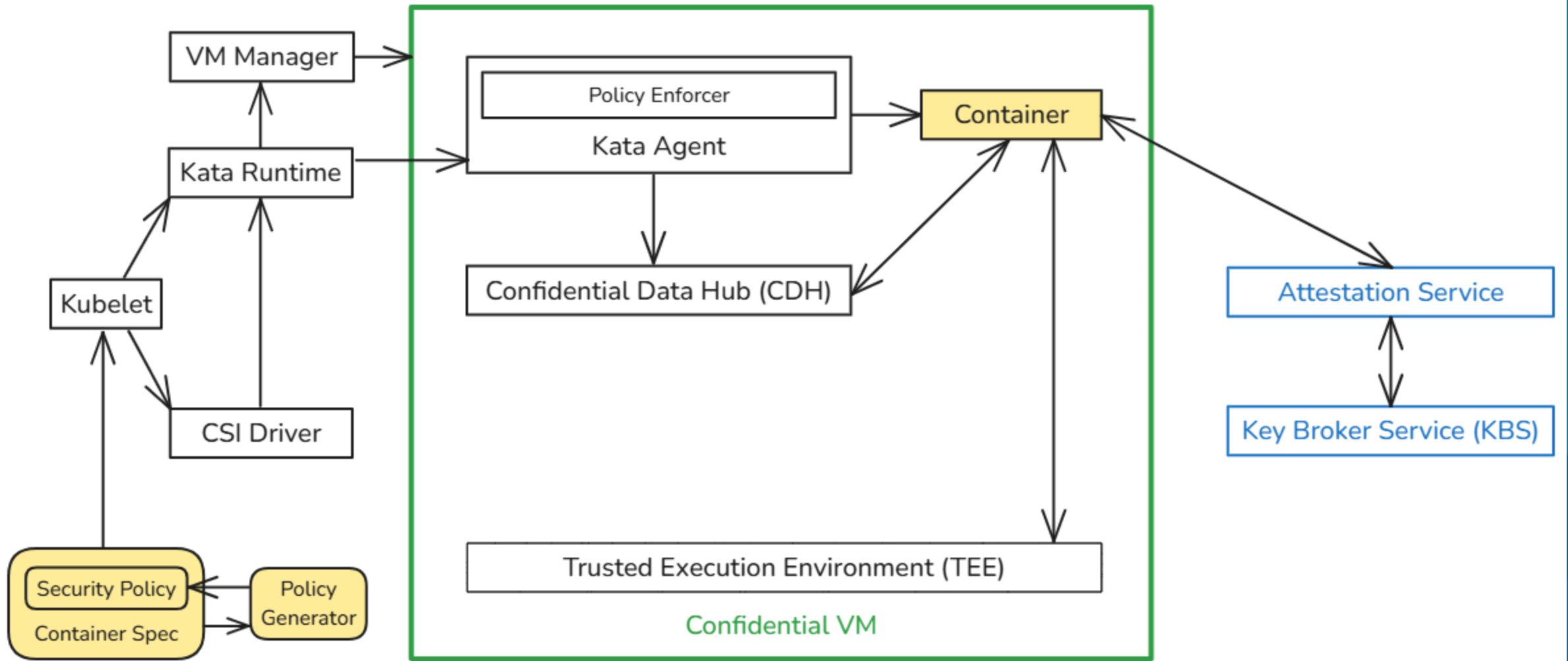




Persistent models







Verifying encryption settings

```
root@localhost:/# |
```

```
|
```

```
root@localhost:/# lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE  MOUNTPOINT
vda                                254:0    0   9.3G  0 disk
└─ encrypted_disk_4BgzN_dif
    └─ encrypted_disk_4BgzN 253:1    0   9.2G  0 crypt
        └─ encrypted_disk_4BgzN 253:2    0   9.2G  0 crypt /run/kata-containers/sandbox/
pmem0                             259:0    0  254M  1 disk
├─ pmem0p1                       259:1    0  250M  1 part
└─ pmem0p2                       259:2    0    3M   1 part
root@localhost:/# |
```

```
root@localhost:/# lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE  MOUNTPOINT
vda                                254:0    0   9.3G  0 disk
└─encrypted_disk_4BgzN_dif
    encrypted_disk_4BgzN           253:1    0   9.2G  0 crypt
    └─encrypted_disk_4BgzN         253:2    0   9.2G  0 crypt /run/kata-containers/sandbox/
pmem0                             259:0    0  254M  1 disk
├─pmem0p1                         259:1    0  250M  1 part
└─pmem0p2                         259:2    0    3M   1 part
root@localhost:/# cryptsetup -v status encrypted_disk_4BgzN
```

```
root@localhost:/# lsblk
NAME                                MAJ:MIN RM  SIZE RO TYPE  MOUNTPOINT
vda                                254:0    0   9.3G  0 disk
└─encrypted_disk_4BgzN_dif
    encrypted_disk_4BgzN            253:1    0   9.2G  0 crypt
    └─encrypted_disk_4BgzN          253:2    0   9.2G  0 crypt /run/kata-containers/sandbox/
pmem0                             259:0    0  254M  1 disk
└─pmem0p1                         259:1    0  250M  1 part
└─pmem0p2                         259:2    0    3M   1 part
root@localhost:/# cryptsetup -v status encrypted_disk_4BgzN
/dev/mapper/encrypted_disk_4BgzN is active and is in use.
type:          LUKS2
cipher:        aes-xts-plain64
keysize:       768 bits
key location:  keyring
integrity:     hmac(sha256)
integrity keysize: 256 bits
device:        /dev/vda
sector size:   4096
offset:        0 sectors
size:          19217256 sectors
mode:          read/write
Command successful.
root@localhost:/# |
```

Generating security policy


```
$ vim my-app.yaml|
```

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6 spec:
7   volumes:
8     - name: encrypted
9       ephemeral:
10        volumeClaimTemplate:
11          spec:
12            accessModes: [ReadWriteOncePod]
13            storageClassName: coco-local-csi
14            resources:
15              requests:
16                storage: 10G
17   containers:
18     - volumeMounts:
19       - name: encrypted
20         mountPath: /mnt/encrypted
21       image: busybox:latest
22       name: my-container
23       command: ["sleep", "infinity"]
24   runtimeClassName: kata
```

~
~
~
~
~
~
~

```
$ vim my-app.yaml  
$ RUST_LOG=INFO ./genpolicy -y my-app.yaml |
```

```
$ vim my-app.yaml
$ RUST_LOG=INFO ./genpolicy -y my-app.yaml
[2025-01-28T03:15:46Z INFO genpolicy::registry] =====
[2025-01-28T03:15:46Z INFO genpolicy::registry] Pulling manifest and config for busybox:latest
[2025-01-28T03:15:47Z INFO genpolicy::registry] Pulling layer "sha256:9c0abc9c5bd3a7854141800ba1f4a227baa88b
11b49d8207eadc483c3f2496de"
[2025-01-28T03:15:47Z INFO genpolicy::registry] Decompressing layer
[2025-01-28T03:15:47Z INFO genpolicy::registry] Adding tarfs index to layer
[2025-01-28T03:15:47Z INFO genpolicy::registry] Calculating dm-verity root hash
[2025-01-28T03:15:47Z INFO genpolicy::registry] dm-verity root hash: 6d7f699c55ff95c54a823b6ebbb3d27373e1b15
4f97439a081019d1b78b1a943
[2025-01-28T03:15:47Z INFO genpolicy::registry] =====
[2025-01-28T03:15:47Z INFO genpolicy::registry] Pulling manifest and config for mcr.microsoft.com/oss/kubern
etes/pause:3.6
[2025-01-28T03:15:47Z INFO genpolicy::registry] Pulling layer "sha256:5720cd9c19ca69b58202945924c37c9bd7b287
ce1a88882098fd59a4292e7cd9"
[2025-01-28T03:15:47Z INFO genpolicy::registry] Decompressing layer
[2025-01-28T03:15:47Z INFO genpolicy::registry] Adding tarfs index to layer
[2025-01-28T03:15:47Z INFO genpolicy::registry] Calculating dm-verity root hash
[2025-01-28T03:15:47Z INFO genpolicy::registry] dm-verity root hash: 817250f1a3e336da76f5bd3fa784e1b26d959b9
c131876815ba2604048b70c18
[2025-01-28T03:15:47Z INFO genpolicy] Success!
$ |
```



```
$ vim my-app.yaml|
```

```
1 |---
2 kind: Pod
3 apiVersion: v1
4 metadata:
5   name: my-app
6   annotations:
7     io.katacontainers.config.agent.policy: IyBDb3B5cmIuaHQgKGmpIDIwMjMgTWljcm9zb2Z0IENvcnBvcnF0aW9uCiMKIy
8 spec:
9   volumes:
10     - name: encrypted
11       ephemeral:
12         volumeClaimTemplate:
13           spec:
14             accessModes: [ReadWriteOncePod]
15             storageClassName: coco-local-csi
16             resources:
17               requests:
18                 storage: 10G
19   containers:
20     - volumeMounts:
21       - name: encrypted
22         mountPath: /mnt/encrypted
23       image: busybox:latest
24       name: my-container
25       command: ["sleep", "infinity"]
26 runtimeClassName: kata
```

~
~
~