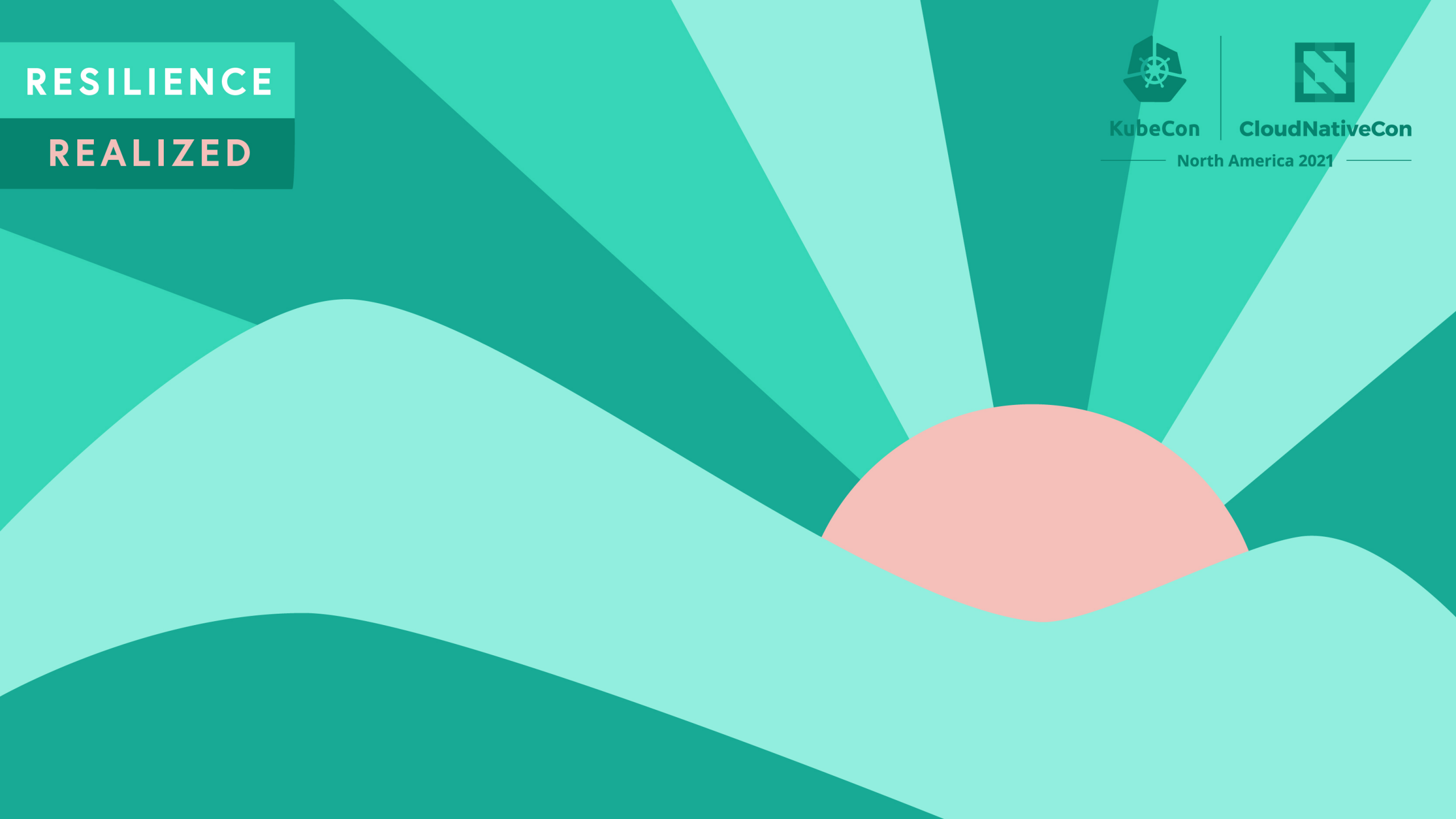
RESILIENCE REALIZED




KubeCon | CloudNativeCon

North America 2021

RESILIENCE
REALIZED

containerd introduction and deep dive

Mike Brown (IBM)
Phil Estes (AWS)
Derek McGowan (Apple)
Maksym Pavlenko (Apple)

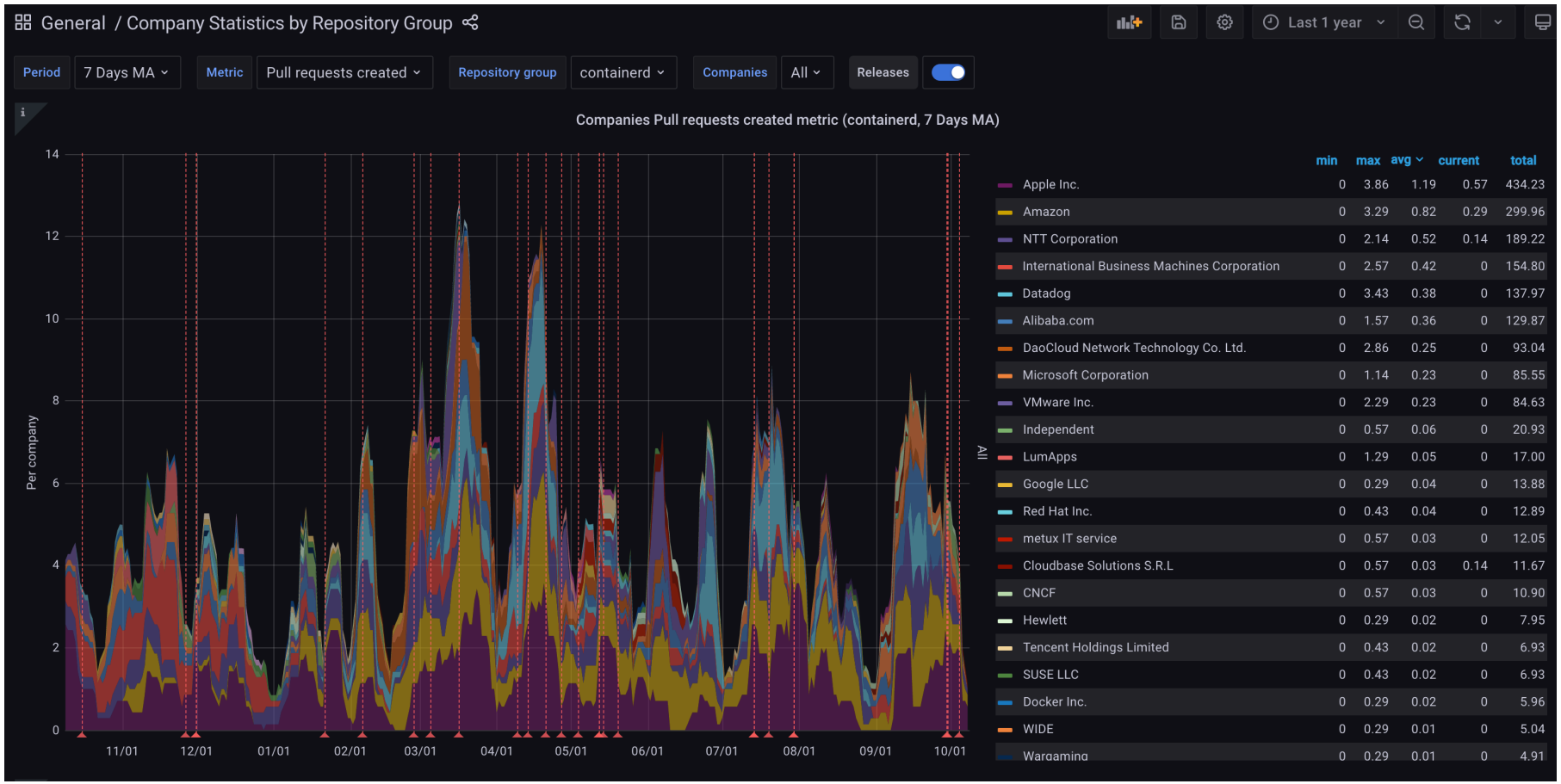
Status Update

- **2 years** since our last in-person update (KubeCon San Diego)
- Community has done **3 virtual** KubeCon/ CloudNativeCon updates

In the Last Two Years...

- **RELEASES**
  - **3 major release lines** (1.4.x, and 1.5.x; 1.6.0-beta.0 **yesterday**)
  - **29 total releases** across 3 supported release branches!
  - Releases **fully automated** through GitHub Actions (full CI migration)
  - **Windows** official release packaging; ARM64 now supported/tested
  - Security **disclosure/release process** implemented (and tested with 3 CVEs)
- **COMMUNITY**
  - Governance streamlined; **security advisor** role added
  - Many new reviewers from multiple employers
  - Over **370 unique code contributors**; 10% of those submitted 10 PRs or more

Community



	min	max	avg	current	total
Apple Inc.	0	3.86	1.19	0.57	434.23
Amazon	0	3.29	0.82	0.29	299.96
NTT Corporation	0	2.14	0.52	0.14	189.22
International Business Machines Corporation	0	2.57	0.42	0	154.80
Datadog	0	3.43	0.38	0	137.97
Alibaba.com	0	1.57	0.36	0	129.87
DaoCloud Network Technology Co. Ltd.	0	2.86	0.25	0	93.04
Microsoft Corporation	0	1.14	0.23	0	85.55
VMware Inc.	0	2.29	0.23	0	84.63
Independent	0	0.57	0.06	0	20.93
LumApps	0	1.29	0.05	0	17.00
Google LLC	0	0.29	0.04	0	13.88
Red Hat Inc.	0	0.43	0.04	0	12.89
metux IT service	0	0.57	0.03	0	12.05
Cloudbase Solutions S.R.L	0	0.57	0.03	0.14	11.67
CNCF	0	0.57	0.03	0	10.90
Hewlett	0	0.29	0.02	0	7.95
Tencent Holdings Limited	0	0.43	0.02	0	6.93
SUSE LLC	0	0.43	0.02	0	6.93
Docker Inc.	0	0.29	0.02	0	5.96
WIDE	0	0.29	0.01	0	5.04
Wargaming	0	0.29	0.01	0	4.91

- **67 unique companies** represented by contributors
- **Committers**: 13 members from 9 companies
- **Reviewers**: 14 members from 10 companies

Project Growth

- **NON-CORE SUBPROJECTS**
  - **stargz-snapshotter -** remote lazy-loading snapshotter
  - **imgcrypt** - container layer encryption support
  - **nerdctl** - more complete Docker client replacement using containerd. Offers packaged support for stargz, rootless, encryption, etc.
- **WHAT'S NEW**
  - CRIv1, more metrics, OpenTelemetry (initial support), shim plugins
  - Confidential Computing: Inclavare project using containerd
  - Upcoming: Sandbox API, Shim runtime/task/API rework, and more!

Parts of containerd

Client

- ctr
- nerdctl
- Go library

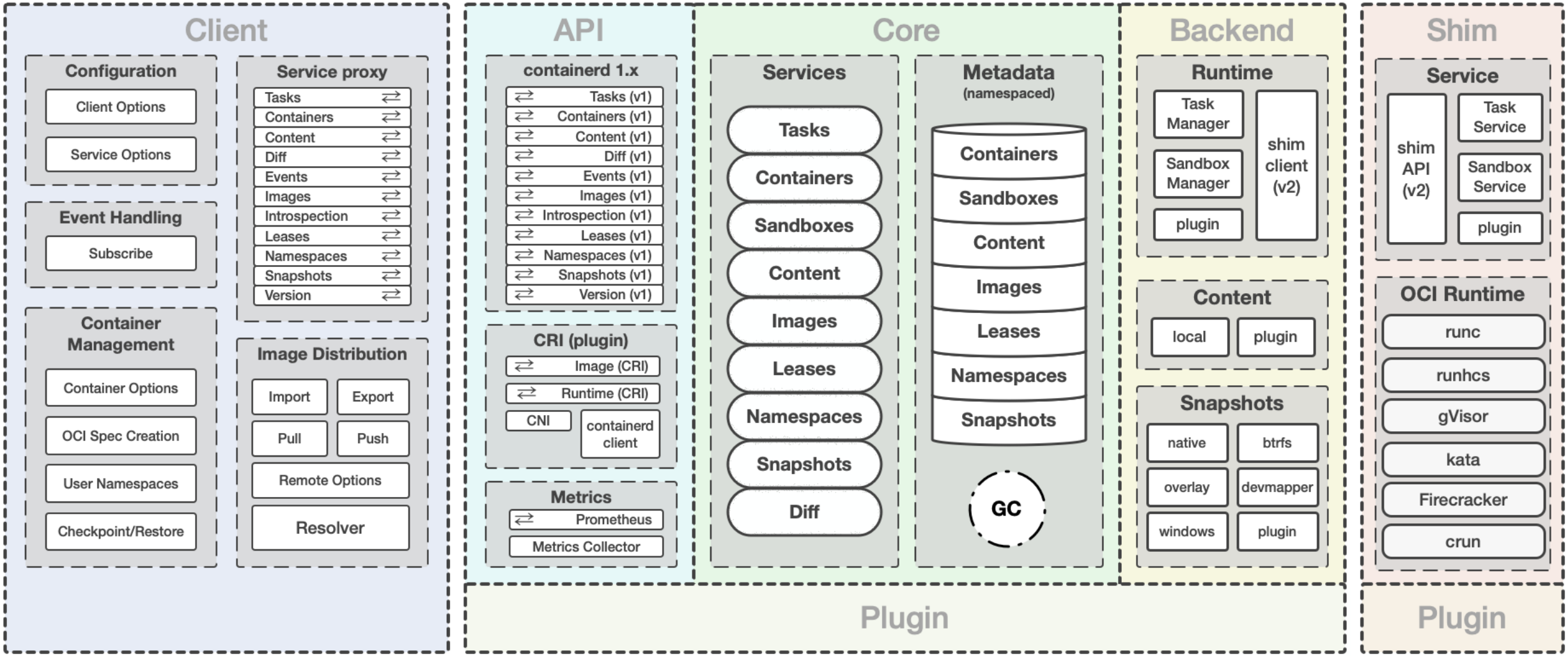
containerd daemon

- API server
- CRI plugin
- Resource managers
  - Data storage
  - Garbage Collection
  - Shim Management

containerd shim

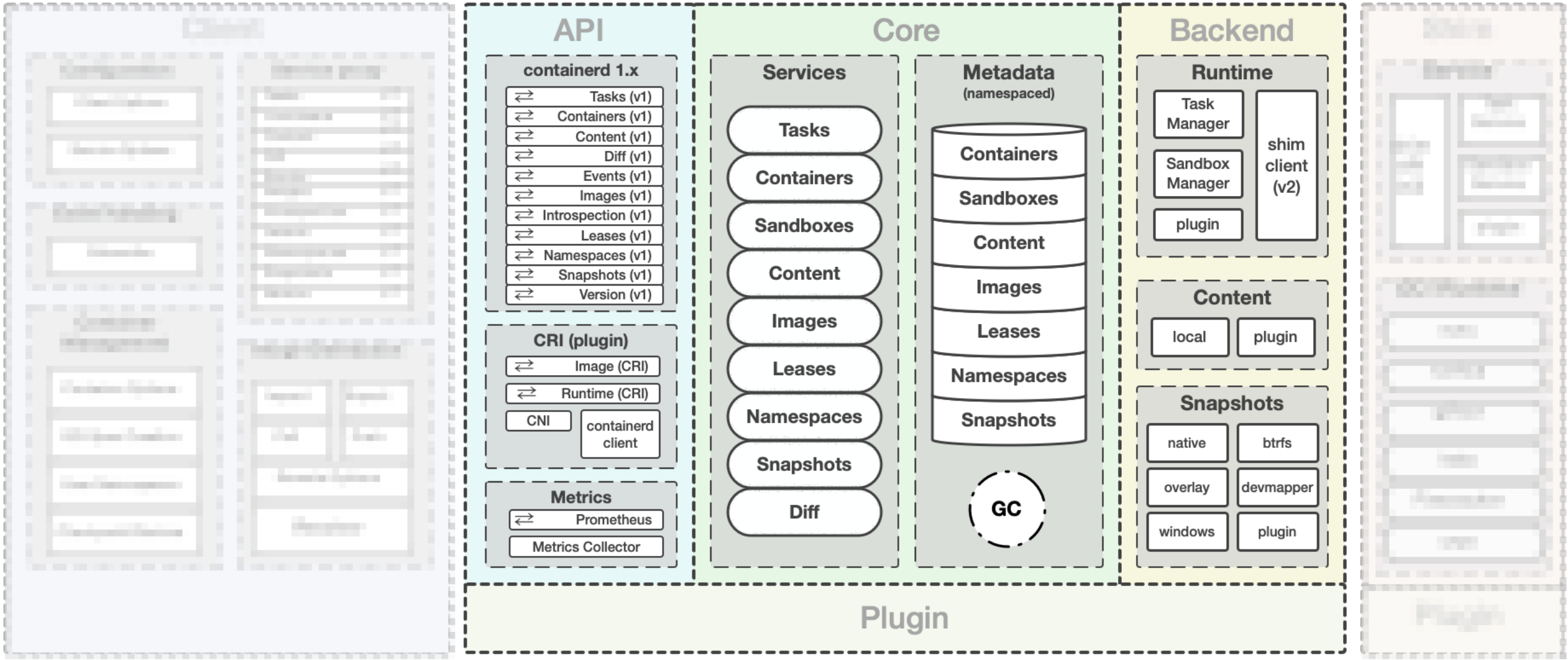
- Per container or pod instance
- Manages running processes

Architecture



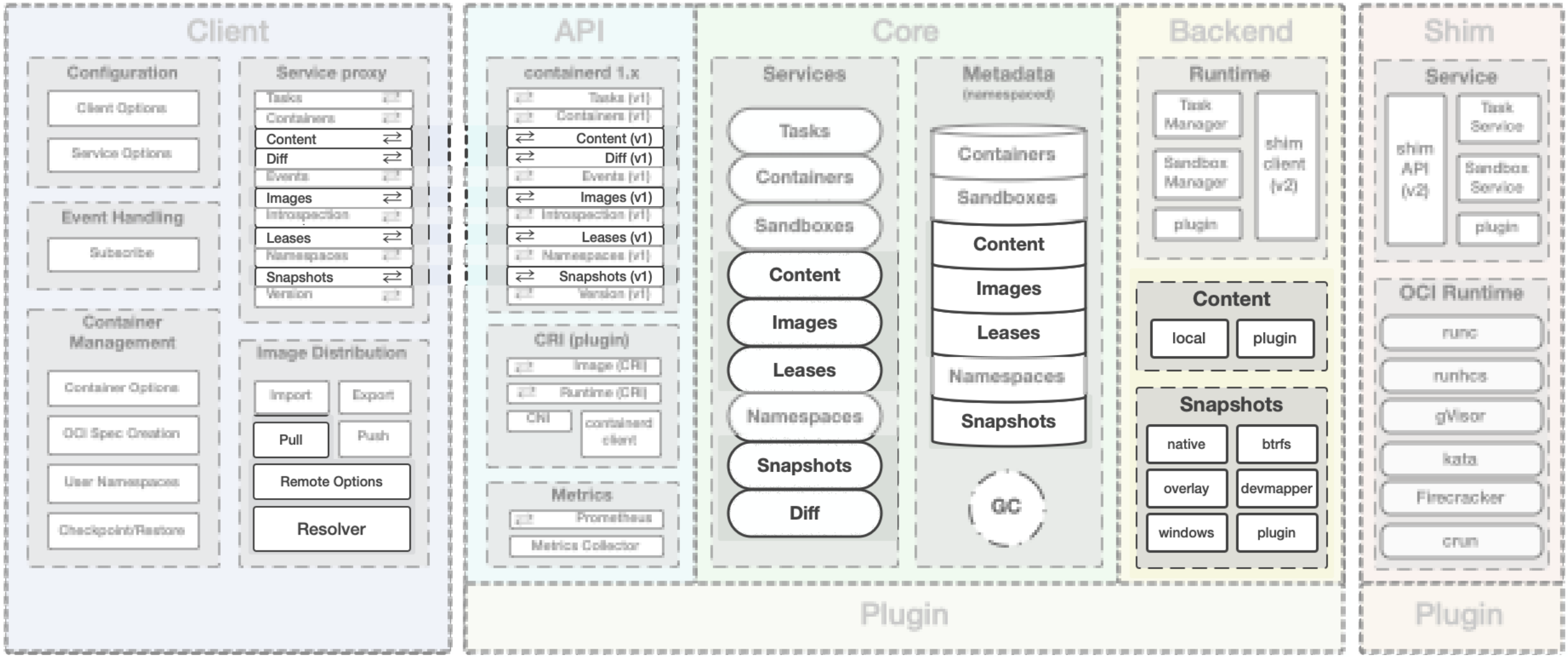
Architecture (Client)

Architecture (Core)

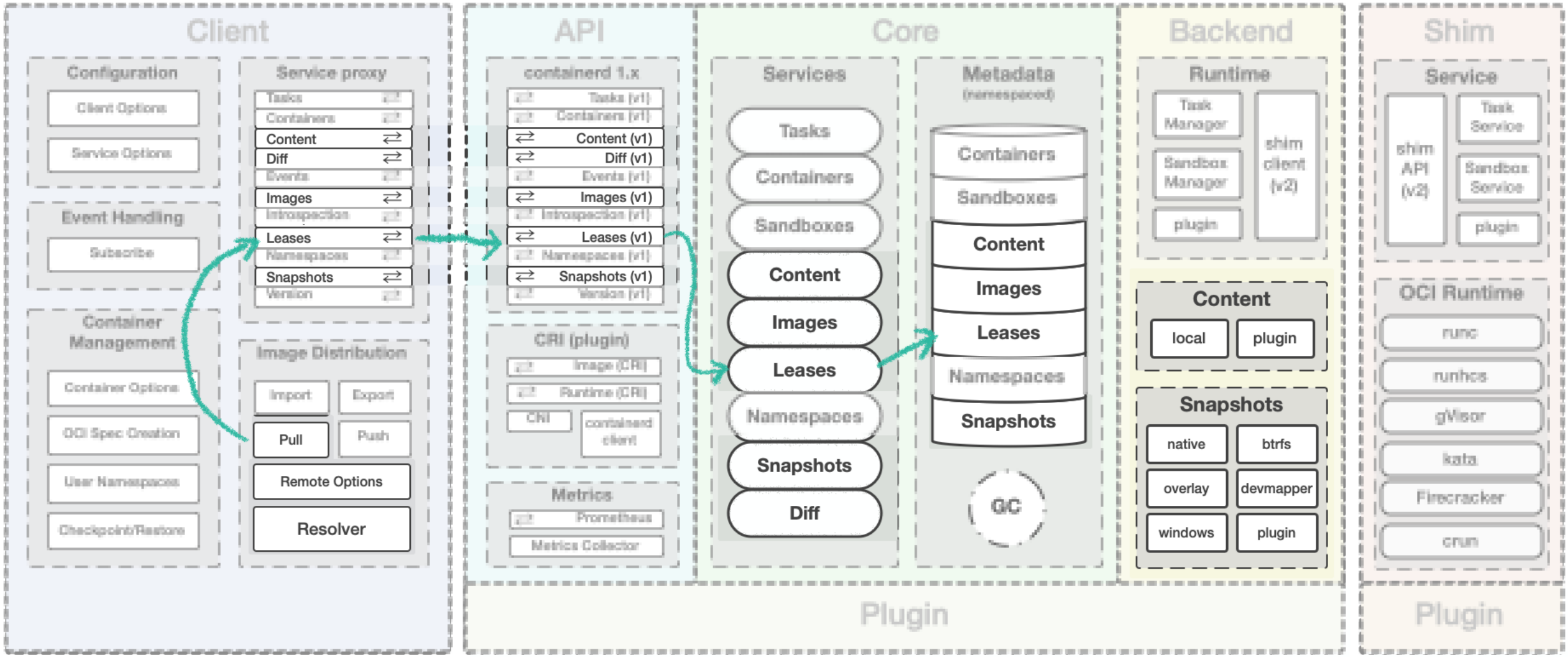


Architecture (Runtime/Shim)

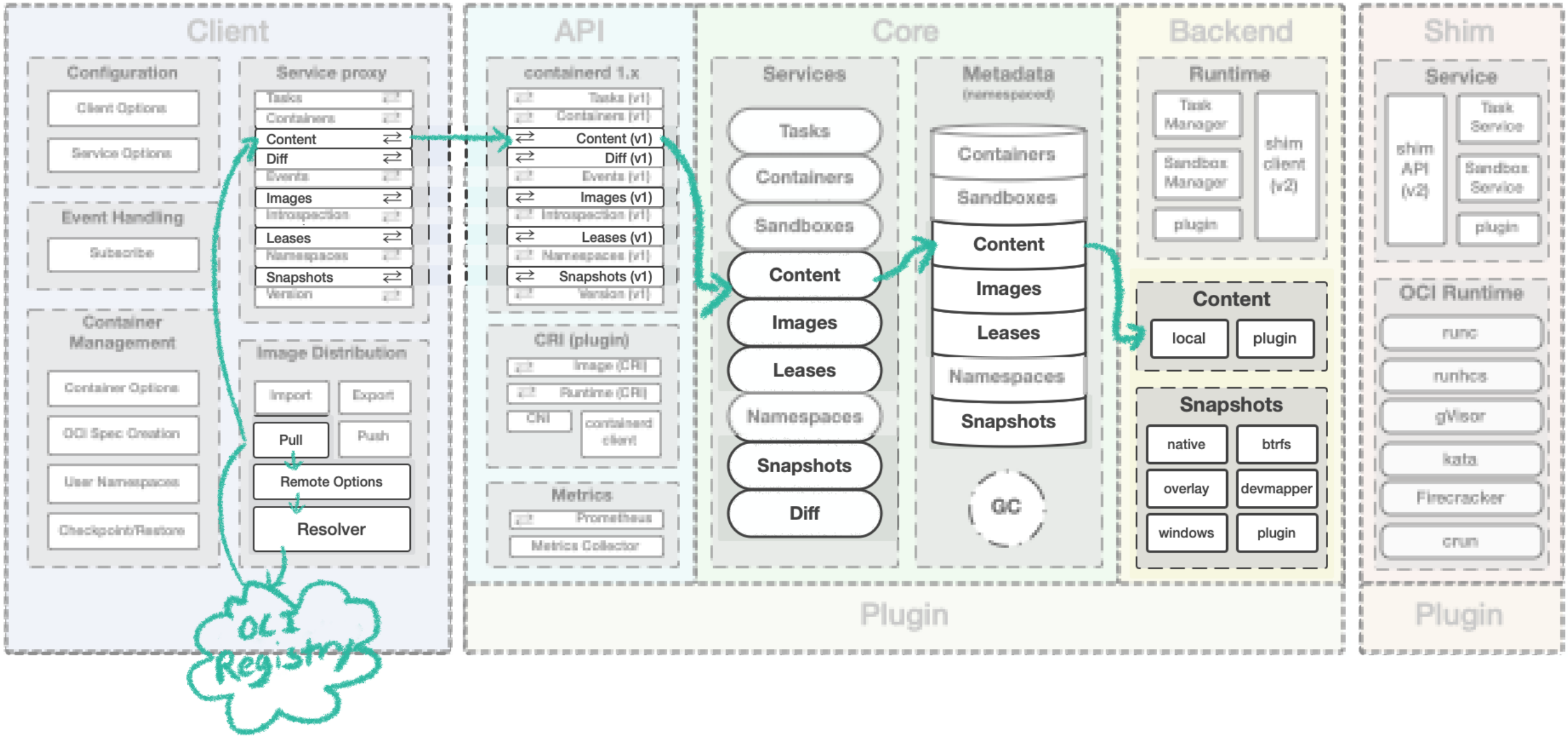
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

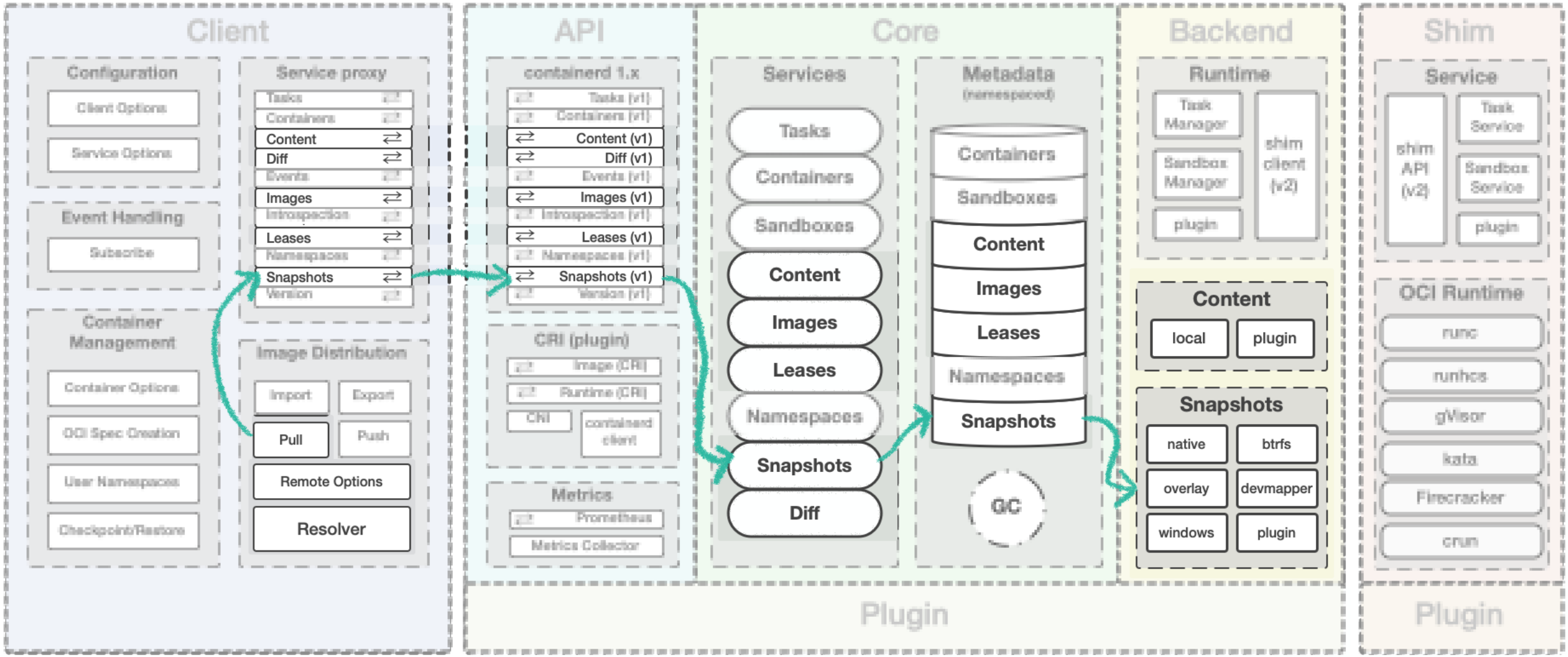
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

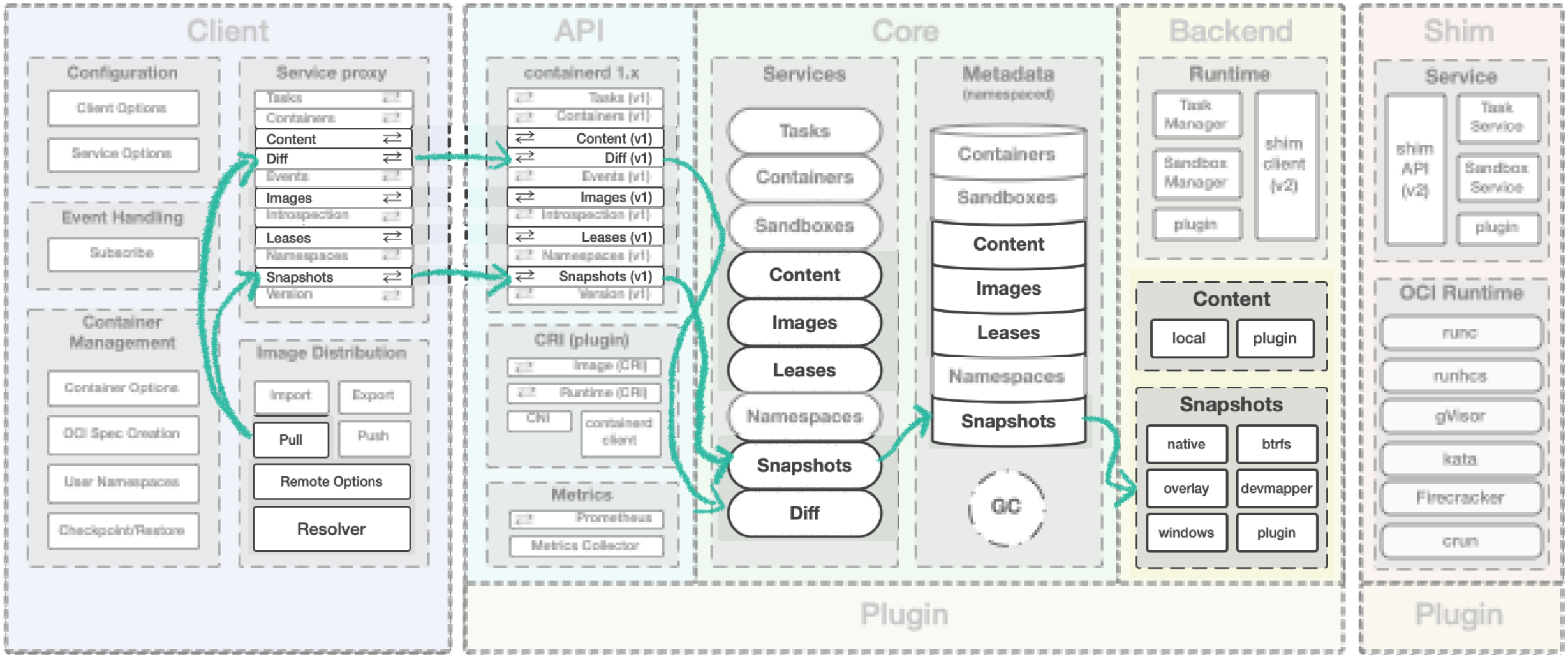
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
OCI Registry
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

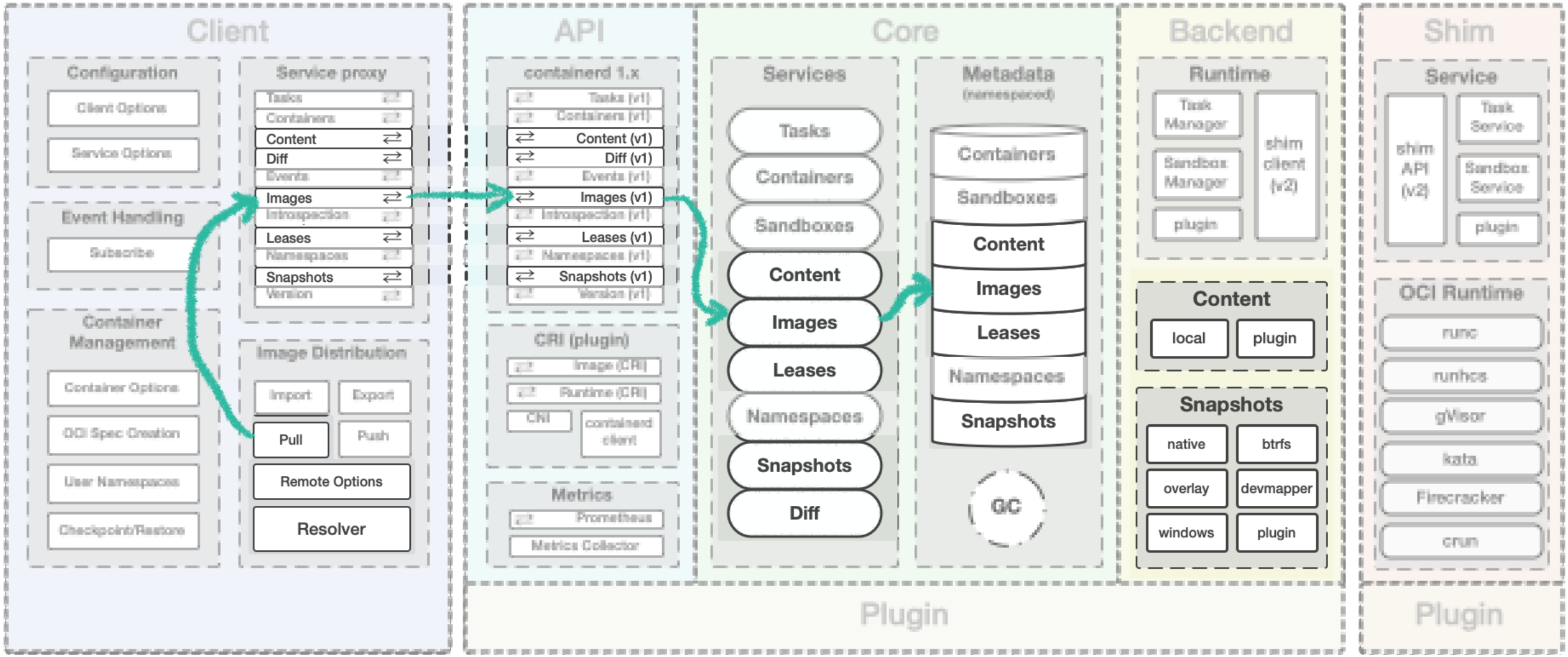
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

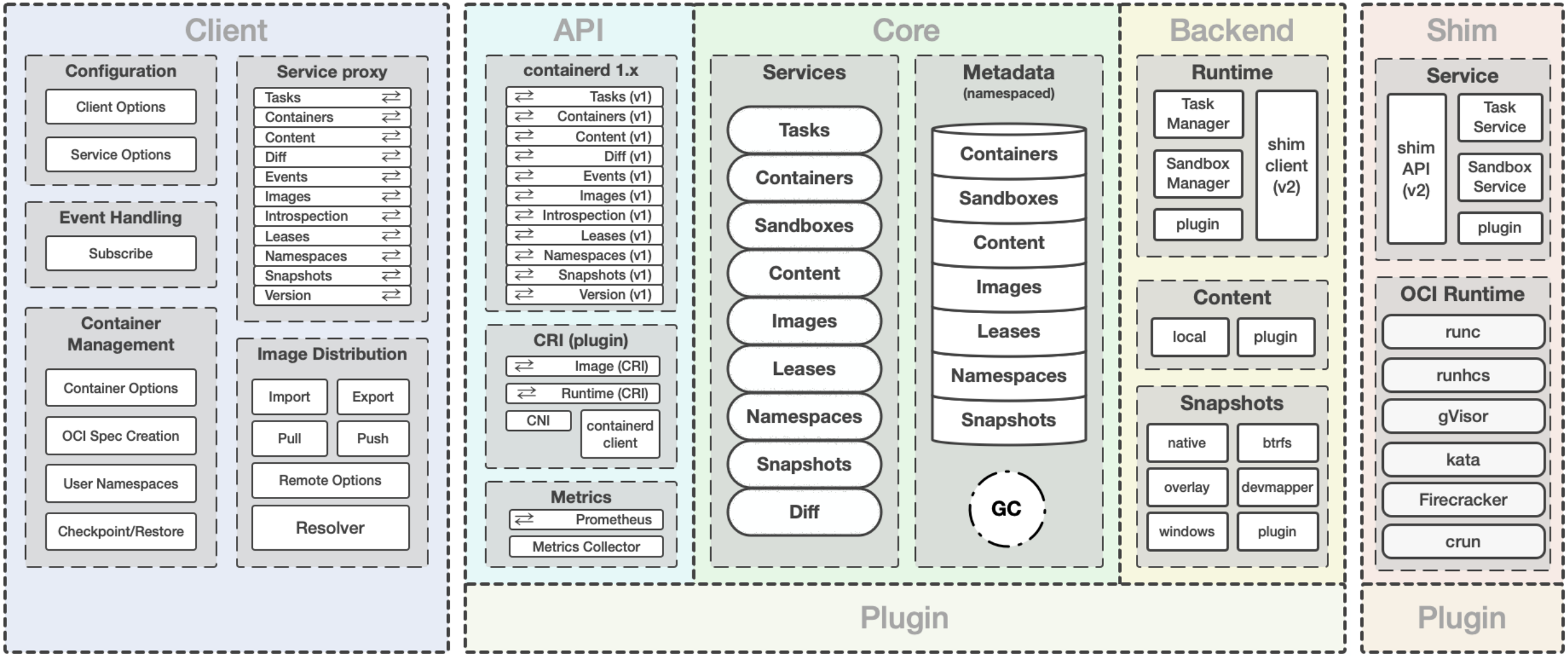
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

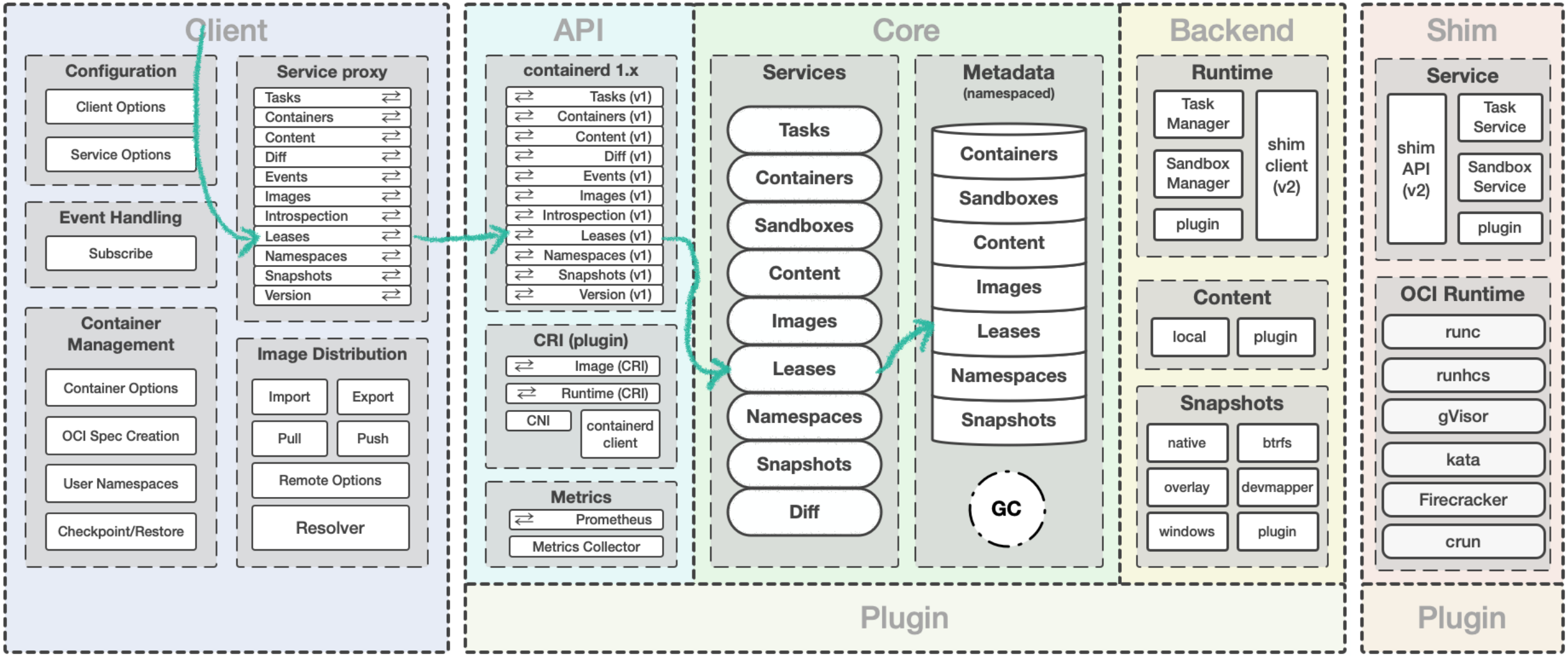
Pull Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CRI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

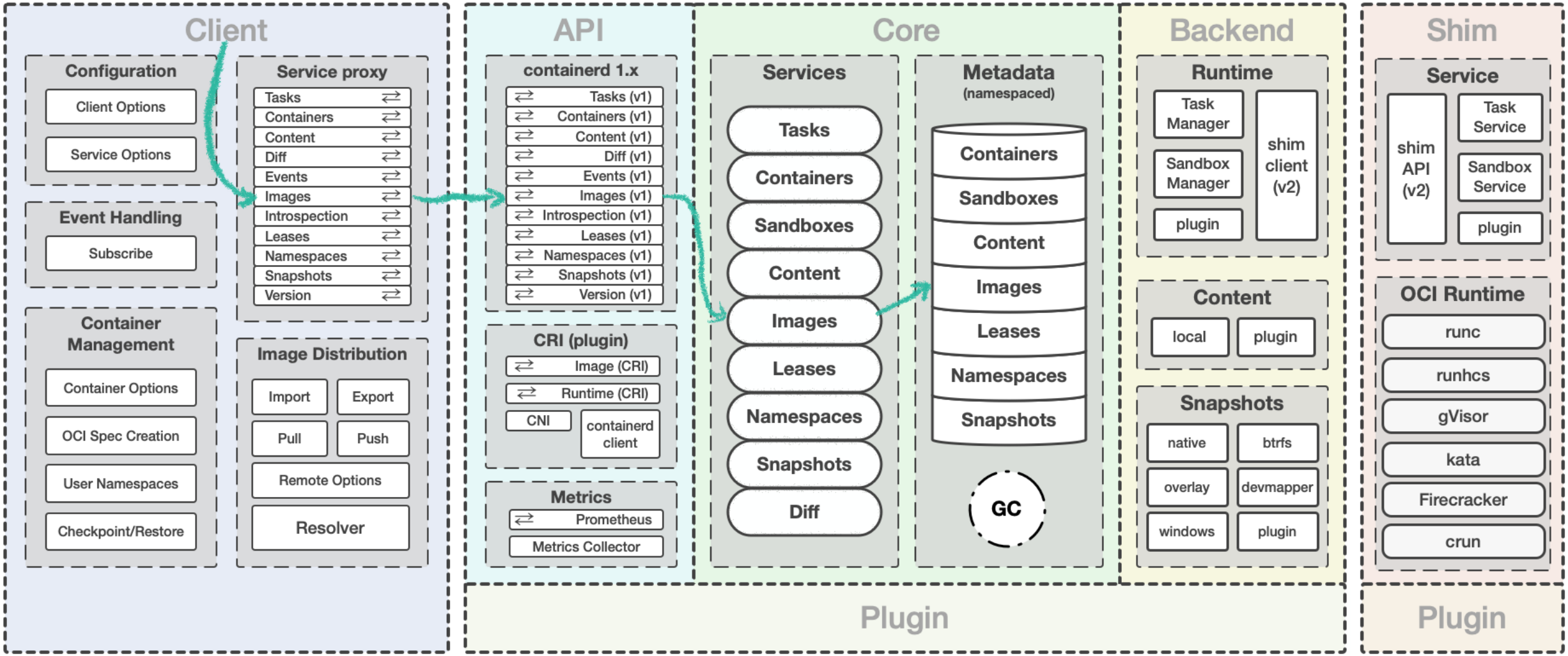
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

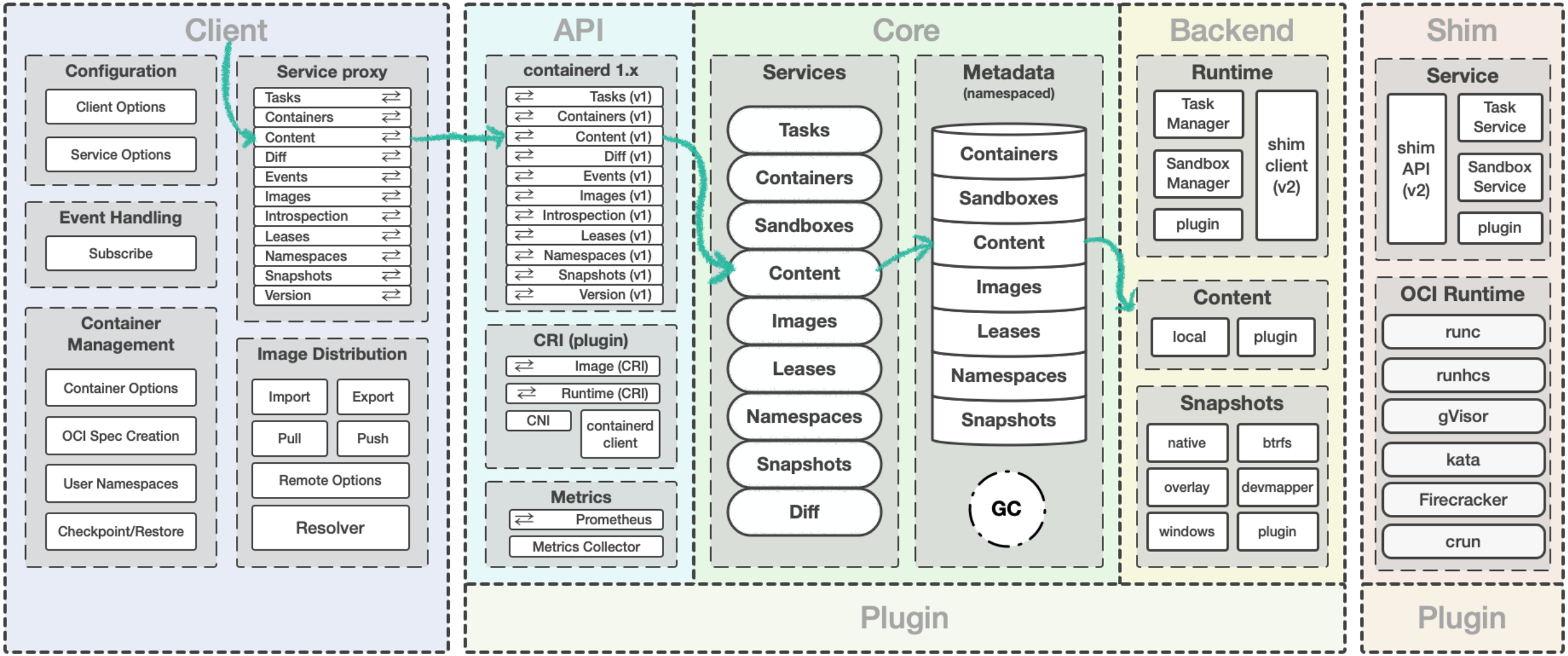
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

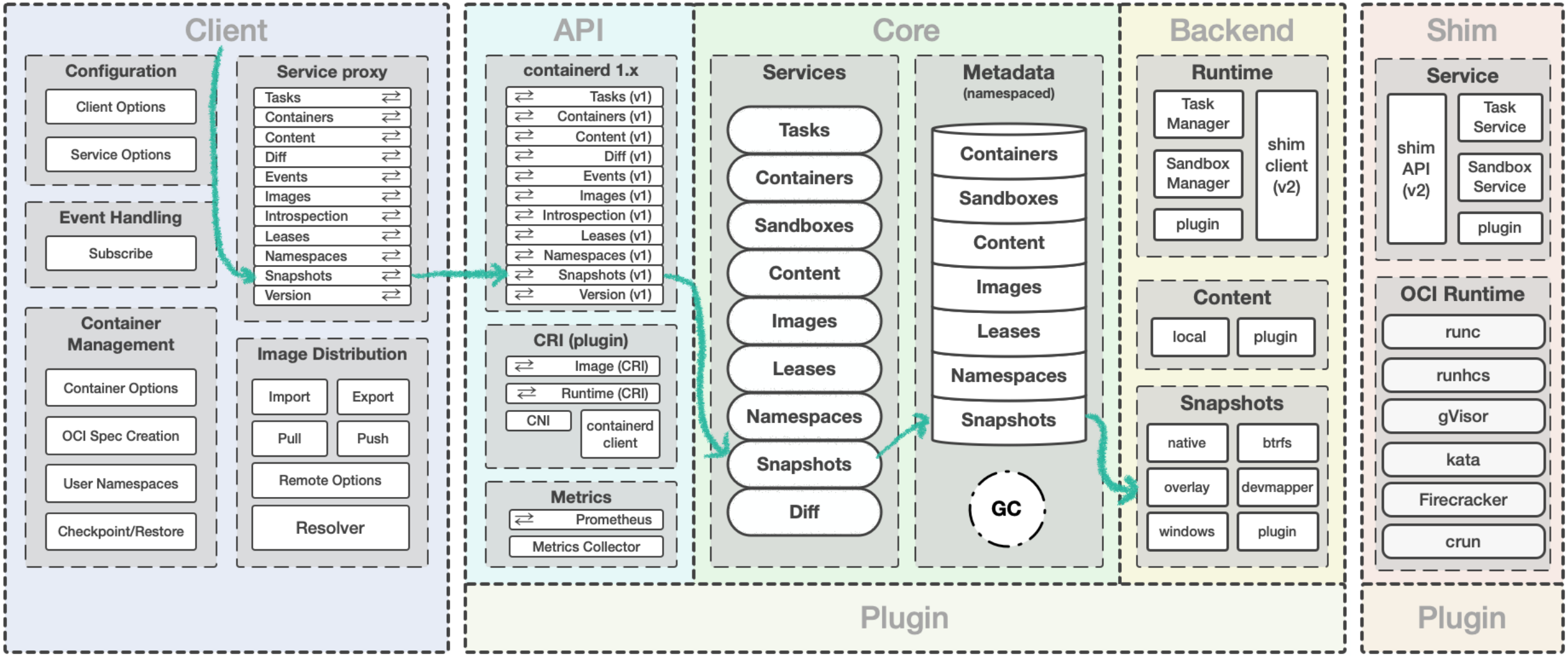
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

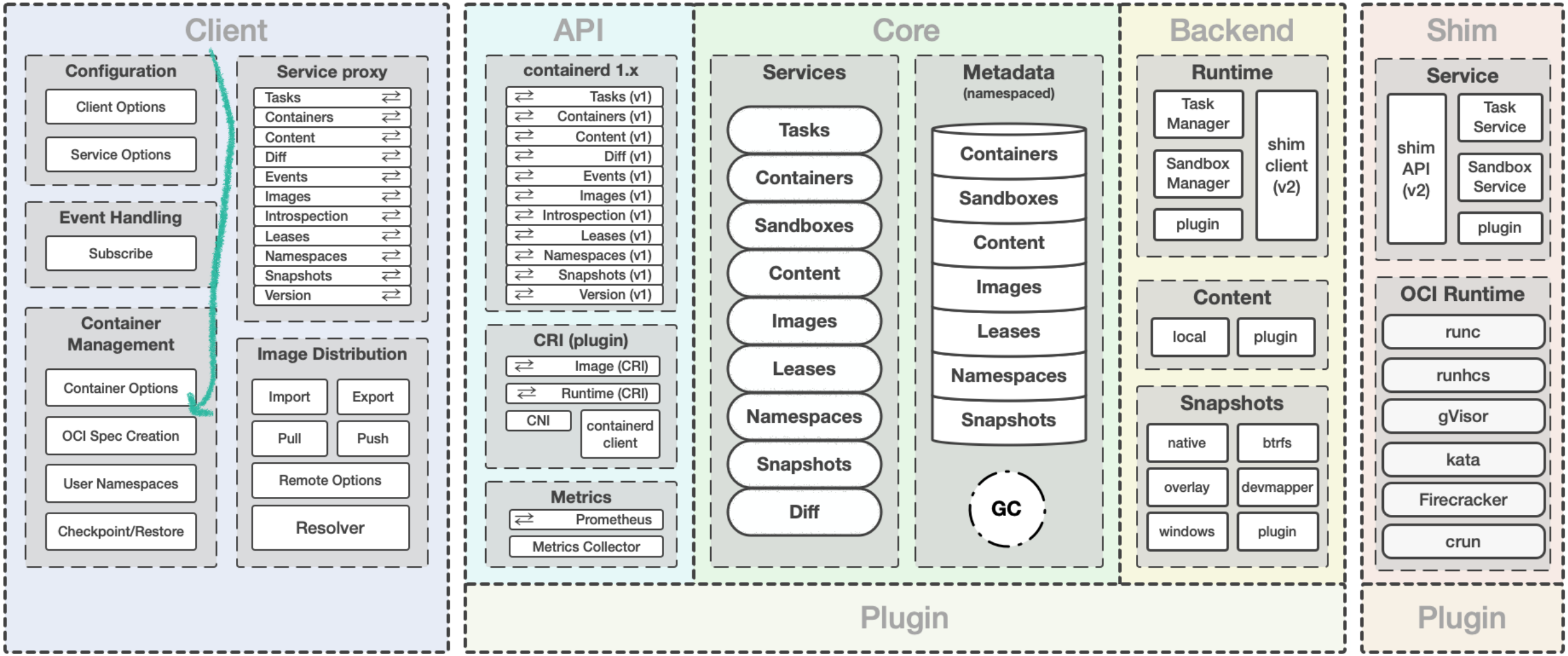
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

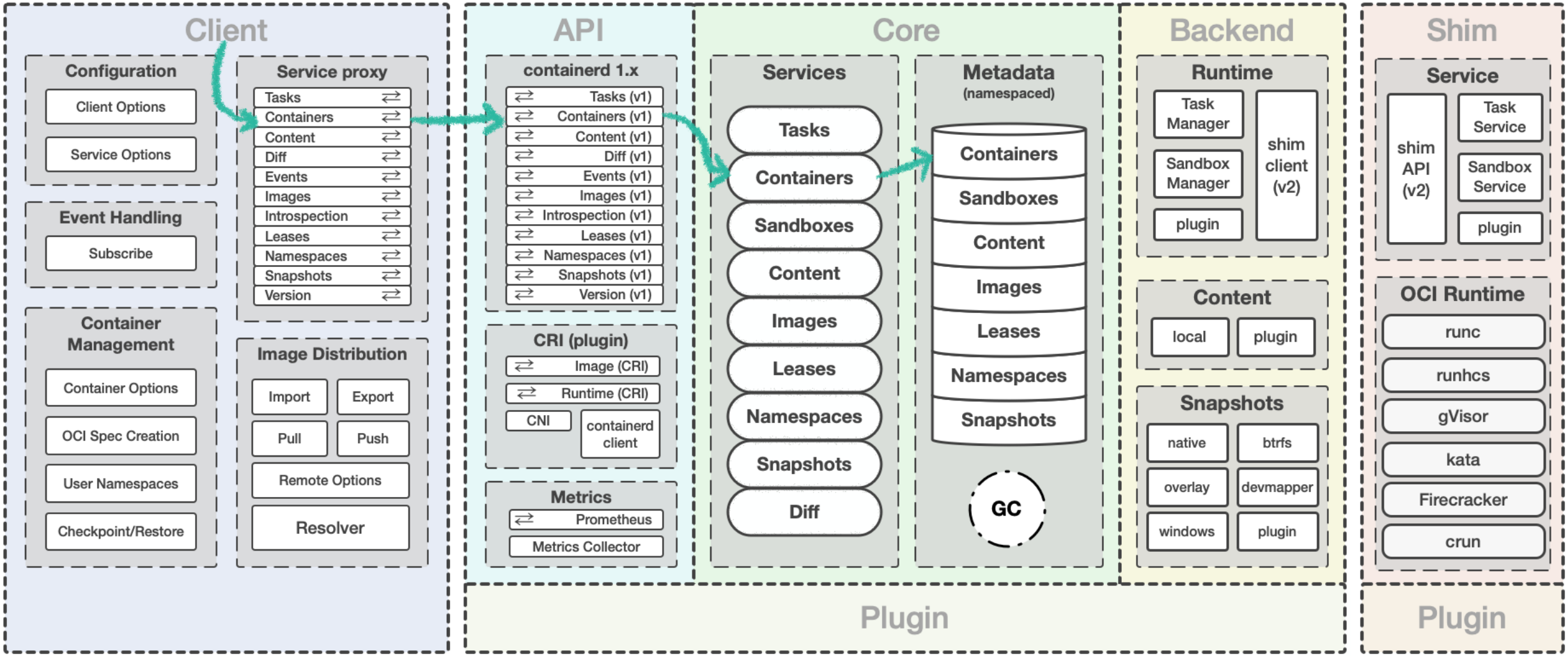
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

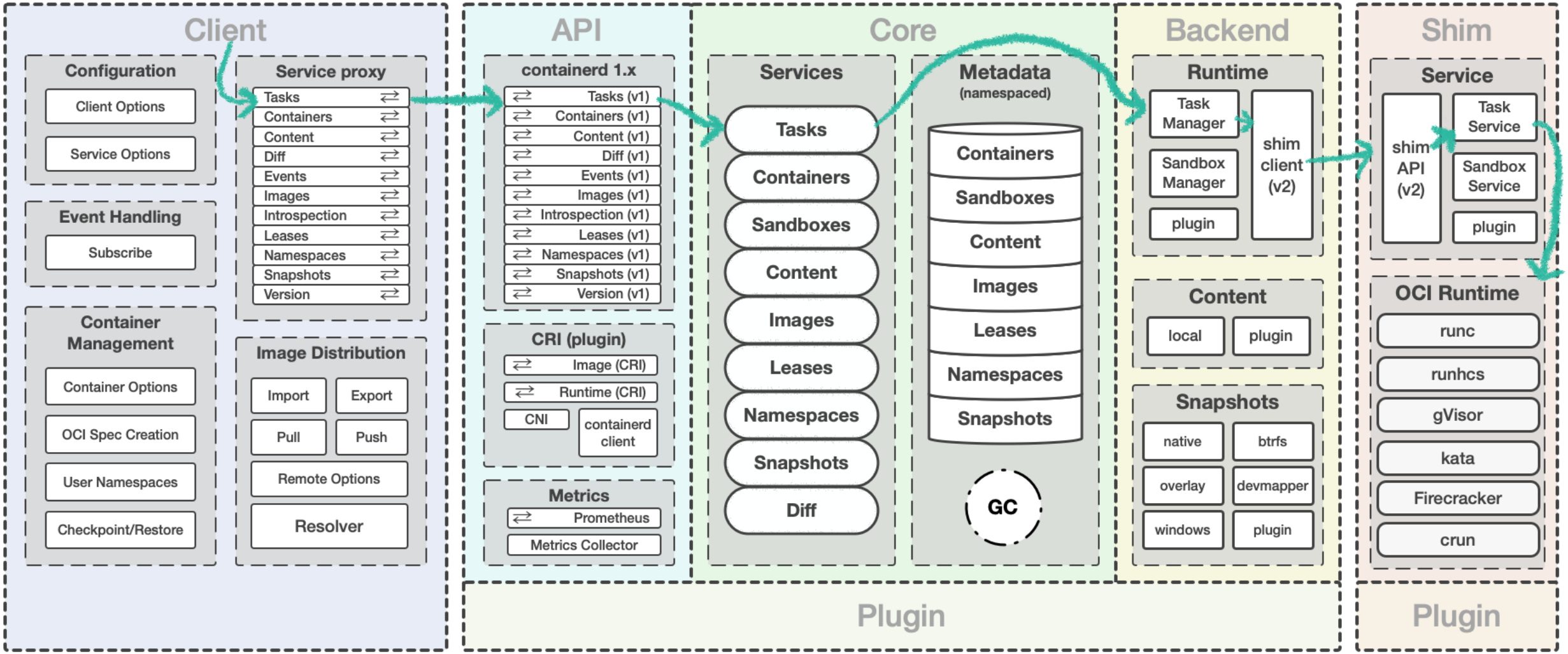
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

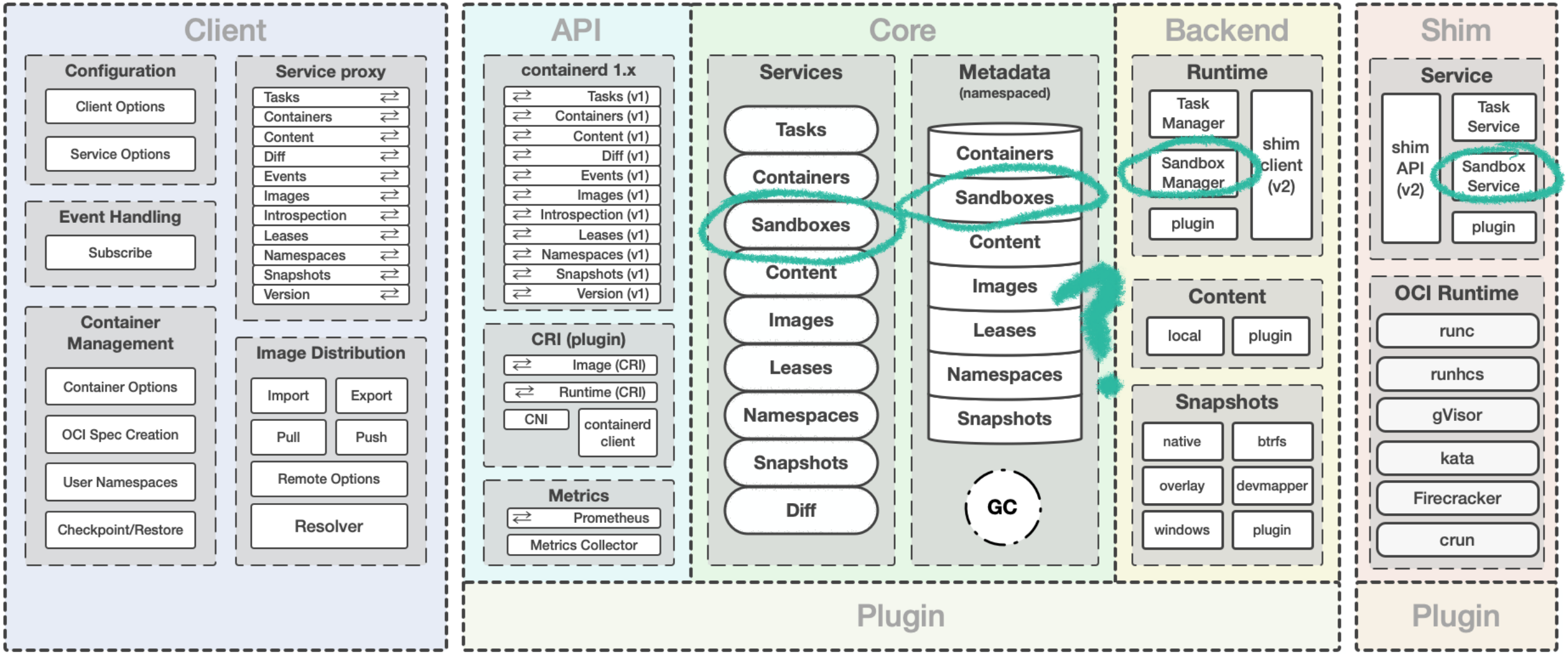
Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

Run Flow


Client
Configuration
Client Options
Service Options
Event Handling
Subscribe
Container Management
Container Options
OCI Spec Creation
User Namespaces
Checkpoint/Restore
Service proxy
Tasks
Containers
Content
Diff
Events
Images
Introspection
Leases
Namespaces
Snapshots
Version
Image Distribution
Import
Export
Pull
Push
Remote Options
Resolver
API
containerd 1.x
Tasks (v1)
Containers (v1)
Content (v1)
Diff (v1)
Events (v1)
Images (v1)
Introspection (v1)
Leases (v1)
Namespaces (v1)
Snapshots (v1)
Version (v1)
CRI (plugin)
Image (CRI)
Runtime (CRI)
CNI
containerd client
Metrics
Prometheus
Metrics Collector
Core
Services
Tasks
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
Diff
Metadata
(namespaced)
Containers
Sandboxes
Content
Images
Leases
Namespaces
Snapshots
GC
Backend
Runtime
Task Manager
Sandbox Manager
plugin
shim client (v2)
Content
local
plugin
Snapshots
native
btrfs
overlay
devmapper
windows
plugin
Shim
Service
shim API (v2)
Task Service
Sandbox Service
plugin
OCI Runtime
runc
runhcs
gVisor
kata
Firecracker
crun
Plugin
Plugin

CRI

- CRI arch refresher
  - Kubelet registers with the api-server to provide node level resource objects.
  - Kubelet uses the Container Runtime Interface (CRI) to communicate with a configured container runtime integration.
  - The container runtime services CRI requests to manage pods, containers, and container images.
  - Pod Spec may select which runtime engine the container runtime should use. (E.g. runc, crun, kata containers, gVisor, Firecracker, …)

CRI

- Kubelet has deprecated the internal docker shim, which was partially based on CRI - work underway for an external CRI docker shim
- Changes underway
  - Test bucket migration away from internal docker shim (e2e, e2e-node, … and critest(cri-tools)) - we need help here. **Shout out to the SIG-Node team for all their help.**
  - Simultaneous support for v1 and v1alpha CRI APIs (Kubelet and container runtimes with restart support for either Kubelet or the container runtime and through upgrades)

CRI

- Confidential Computing (CC) / Multi-Tenant Use Cases
  - Kubelet's Ensure Pod Container Images Process - ensures that a pod's container images are available on the node via CRI image services.
    - Pod pull image policies include "never pull", "pull if not present", and "always pull"
    - Pull image is performed based on policy, image status on the node, and if the registry requires authentication the POD's image pull secrets are used.
    - KEP #1608 Ensure Secret Pulled Images (phase 1) is scheduled for k8s v1.23 and will ensure that a pod using a container image should have access to the image.
  - Kubernetes has a per runtime handler which is pod specified/scoped for choosing runtime engine config (kata, gVisor, runc-config-a, runc-config-b, etc..)
    - We also need that runtime handler to be passed over CRI image services such that image metadata, container snapshots, and storage can be scoped similarly

CRI

- Faster cycle times for PODs
  - Probes
    - Kubernetes uses liveness, readiness and startup probes to help in the managing of containers.
    - These probes operate at the per/second(s) level
    - Certain workloads need us to move from per/second(s) probes to fine grained per/millisecond(s) probes KEP (wip) code (wip)
  - Kubelet currently uses a lazy update cycle process for determining state for PODs, Containers, and Container Images
    - Discussions are underway for introducing a KEP to optionally add a subscription service for POD/Container status changes and possibly also for container image/storage updates

Upcoming containerd 1.6+

- Big focus on shims
  - More runtime features
  - Better microVM support
  - New ways to manage images
- Revisit runtime for new use cases
  - How to support new use cases?
  - . . . and keep things flexible
  - . . . and keep things backward compatible

New shim features

- Sandbox API
  - Better microVM lifecycle/runtime support
- Confidential containers
  - Hide container image contents from the host
- Port forwarding
- New custom APIs?

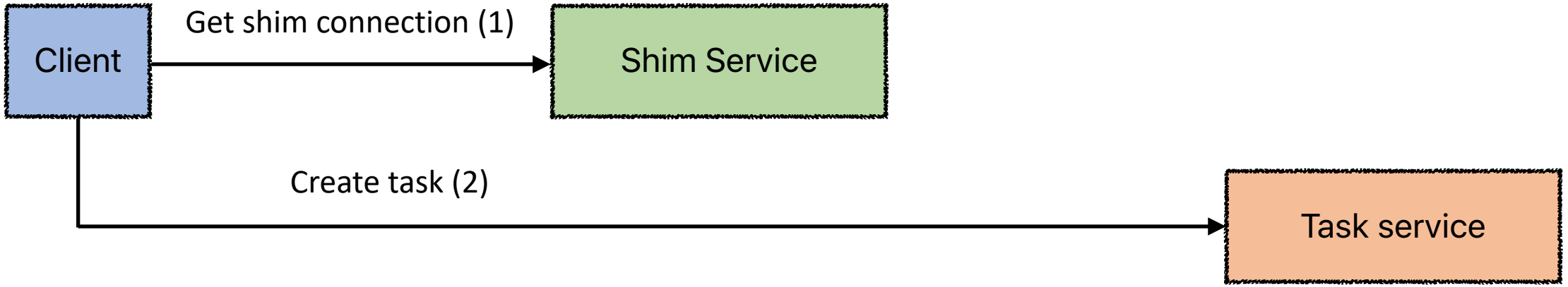
New shim runtime

- Plugin support
  - Extend shim with new plugins
- New Shim service
  - Manage shims independently from Tasks
- New client API to manage Tasks
  - Use new shim API, but keep backward compatibility
  - Support new shim features from client

Shim service call flow

1.5

1.6+



Shim service

- Manage shim lifecycle
  - Start
  - Stop
  - Clean
- Foundation for higher level services
  - Tasks
  - Sandbox

- Call flow defined by Client

```
type ShimService interface {
    StartShim(…) // call `shim --start`
    DeleteShim(…) // call `shim --delete`
    GetConn(shimId string) // get connection
}
```

Sandbox API

- Manage microVMs
  - Launch VM before Container
  - Resize VM instance
  - Suspend/resume
- Needs shim service as base
  - Different shim lifecycle
  - Driven by containerd client

Thank you!