
KubeCon
CloudNativeCon
North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022
TiKV
TiKV
HELM

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

Accelerates Image Distribution In Multi-Cluster With Dragonfly

Wenbo Qi – Ant Group

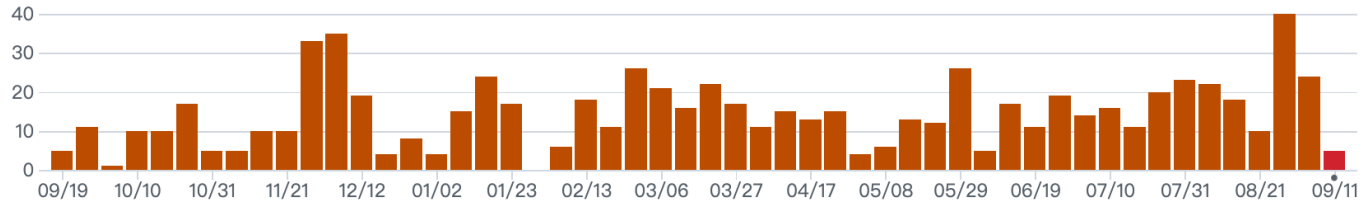
KubeCon | CloudNativeCon
North America 2022
BUILDING FOR THE ROAD AHEAD
DETROIT 2022

Wenbo Qi
Software Engineer,
Ant Group

Introduction

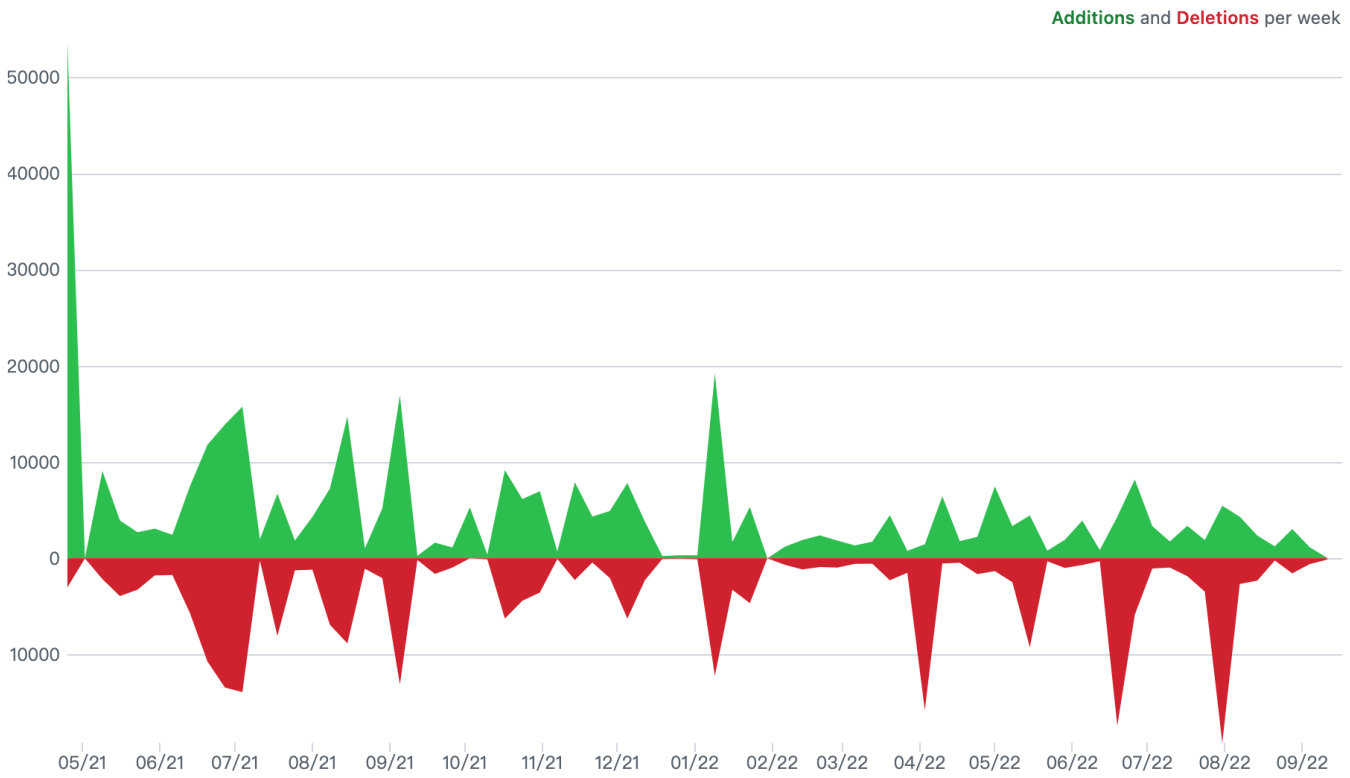
Introduction:

Dragonfly is an **intelligent P2P based image and file distribution system**, it also provides a variety of enterprise-level (efficiency, stability, safety, low-cost) product features.



Milestone:

1. Dragonfly was accepted to CNCF on **11/15/2018** and it is a CNCF **Incubating** project.
2. Dragonfly **1.X** has been upgraded to **2.0** on **9/9/2021**.
3. Dragonfly has released **126+** releases on **9/9/2022**.



Contributor:

From Ant Group, Alibaba, ByteDance, GitLab, Meituan, Xiaomi, Inspur, Shanghai Jiao Tong University, etc.

Discussion Group:

DingTalk: **23304666**
Slack Channel: **#dragonfly**
Twitter: **dragonfly_oss**
Discussion Group: **dragonfly-discuss@googlegroups.com**
Developer Group: **dragonfly-developers@googlegroups.com**

Architecture

Manager:

Maintain the **relationship** between each P2P cluster, dynamic configuration management and RBAC. It also includes a **front-end console**, which is convenient for users to visually operate the cluster.

Scheduler:

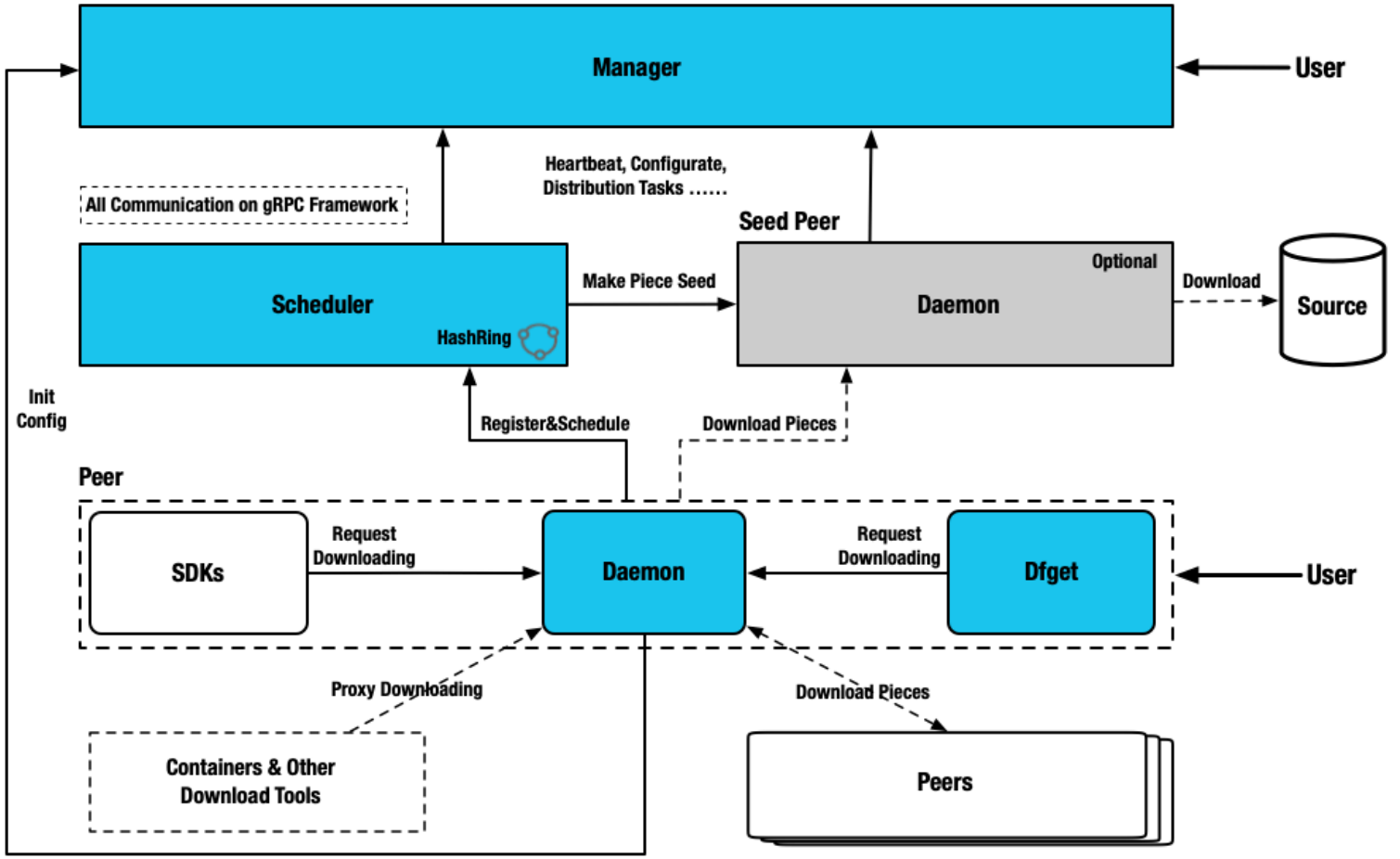
Select the **optimal download parent** peer for the download peer. Exceptions control peer back-to-source.

Seed Peer:

Dfdaemon turns on the Seed Peer mode can be used as a **back-to-source download peer** in a P2P cluster, which is the **root peer** for download in the entire cluster.

Peer:

Deploy with dfdaemon, based on the C/S architecture, it provides the **dfget** command download tool, and the **dfget daemon** running daemon to provide task download capabilities.



Feature Upgrades

Stability:

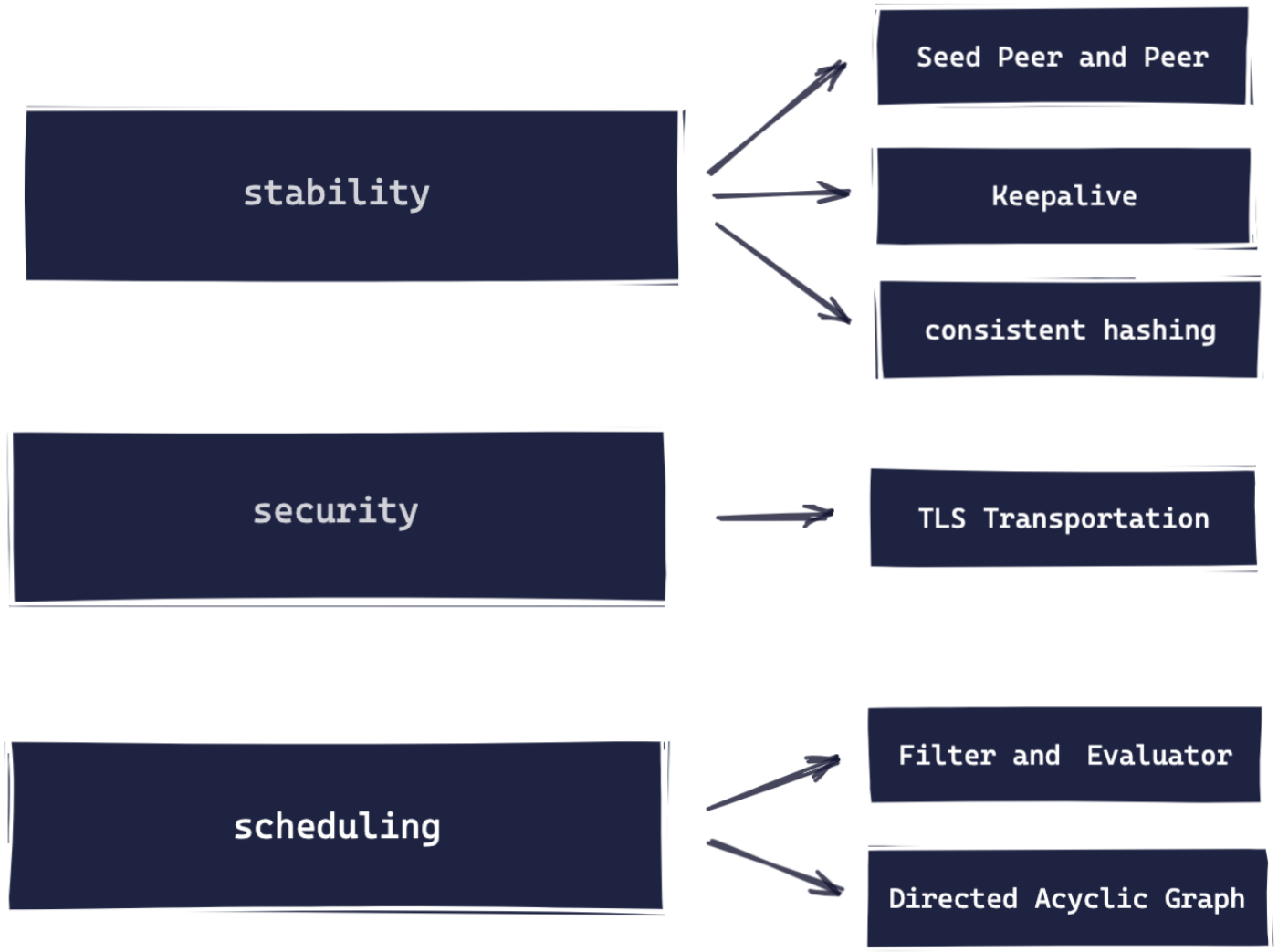
Use **Seed Peer** instead of **CDN** service. **Keepalives** are maintained between each services and managers. Improve task hit rate with **consistent hashing**, and resolve addresses when grpc connections fail.

Security:

Manager issues each service **certificate** as a CA, and transfers task metedata and data based on **TLS**.

Scheduling:

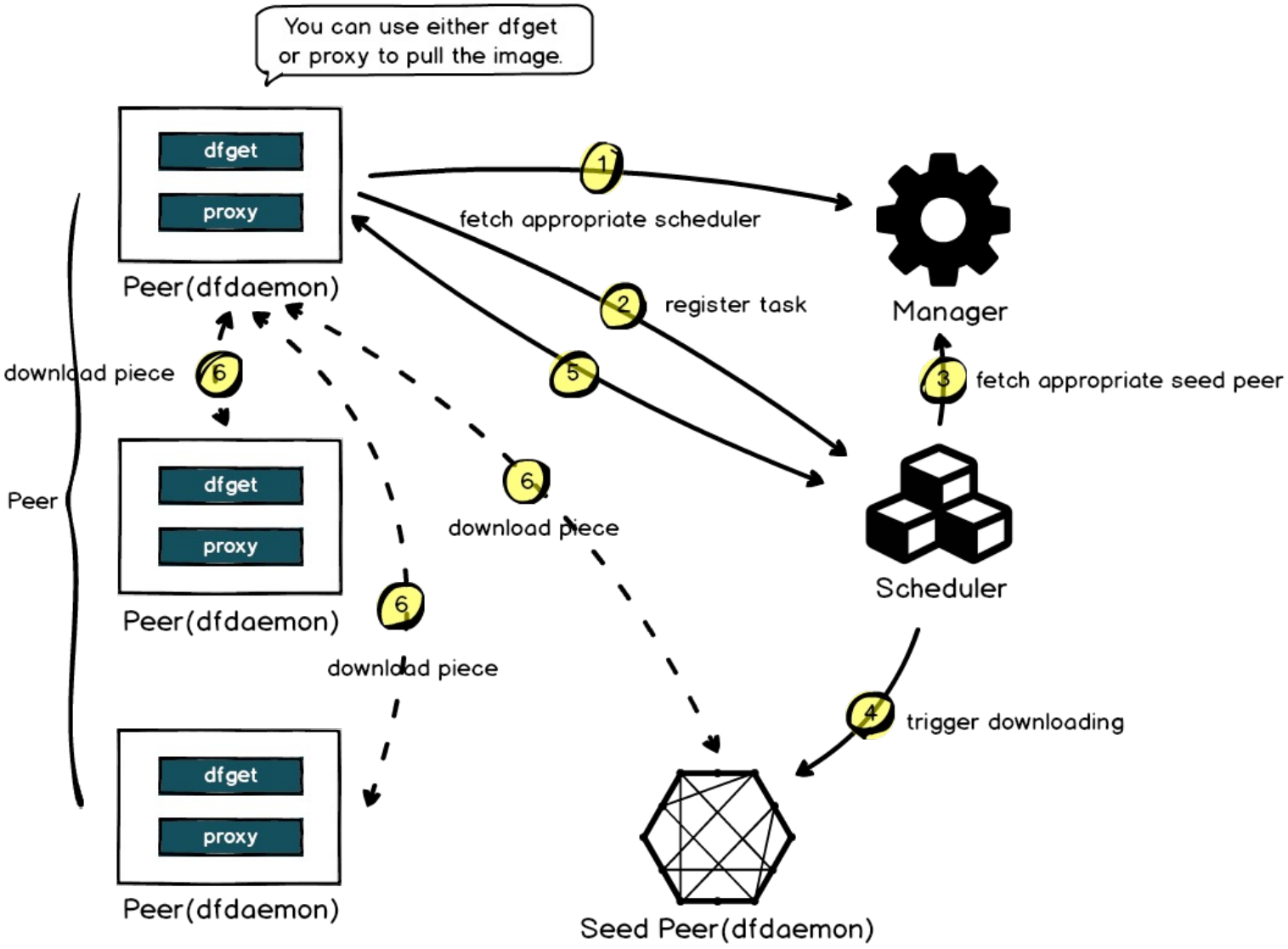
Scheduling through two steps of **filter** and **evaluator**, select a set of optimal parents for the current peer. Build a p2p network model based on **DAG**.



Seed Peer

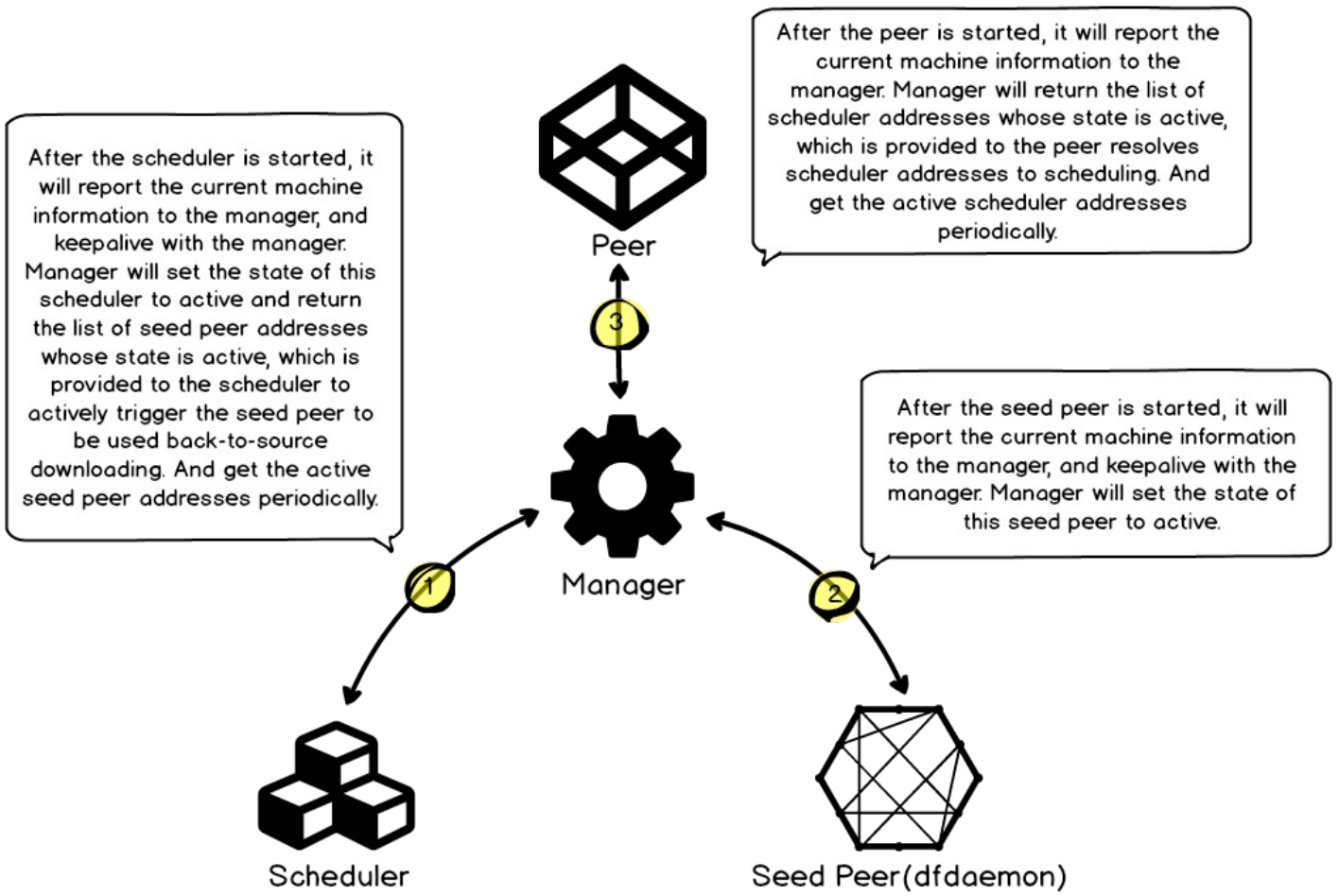
Seed Peer:

Use **Seed Peer** instead of **CDN** service, because CDN was a separate module before, and it can be **actively triggered back-to-source downloading**. So the current version removes CDN and Dfdameon adds triggered back-to-source downloading API. It reduces maintenance costs for CDN modules.



Keepalive

Dynamically update the list of available addresses



Service Type	Address	State
Scheduler	scheduler-1.com	active
Scheduler	scheduler-2.com	inactive
Scheduler	scheduler-3.com	inactive
Seed Peer	seed-peer-1.com	active
Seed Peer	seed-peer-2.com	inactive
Seed Peer	seed-peer-3.com	inactive

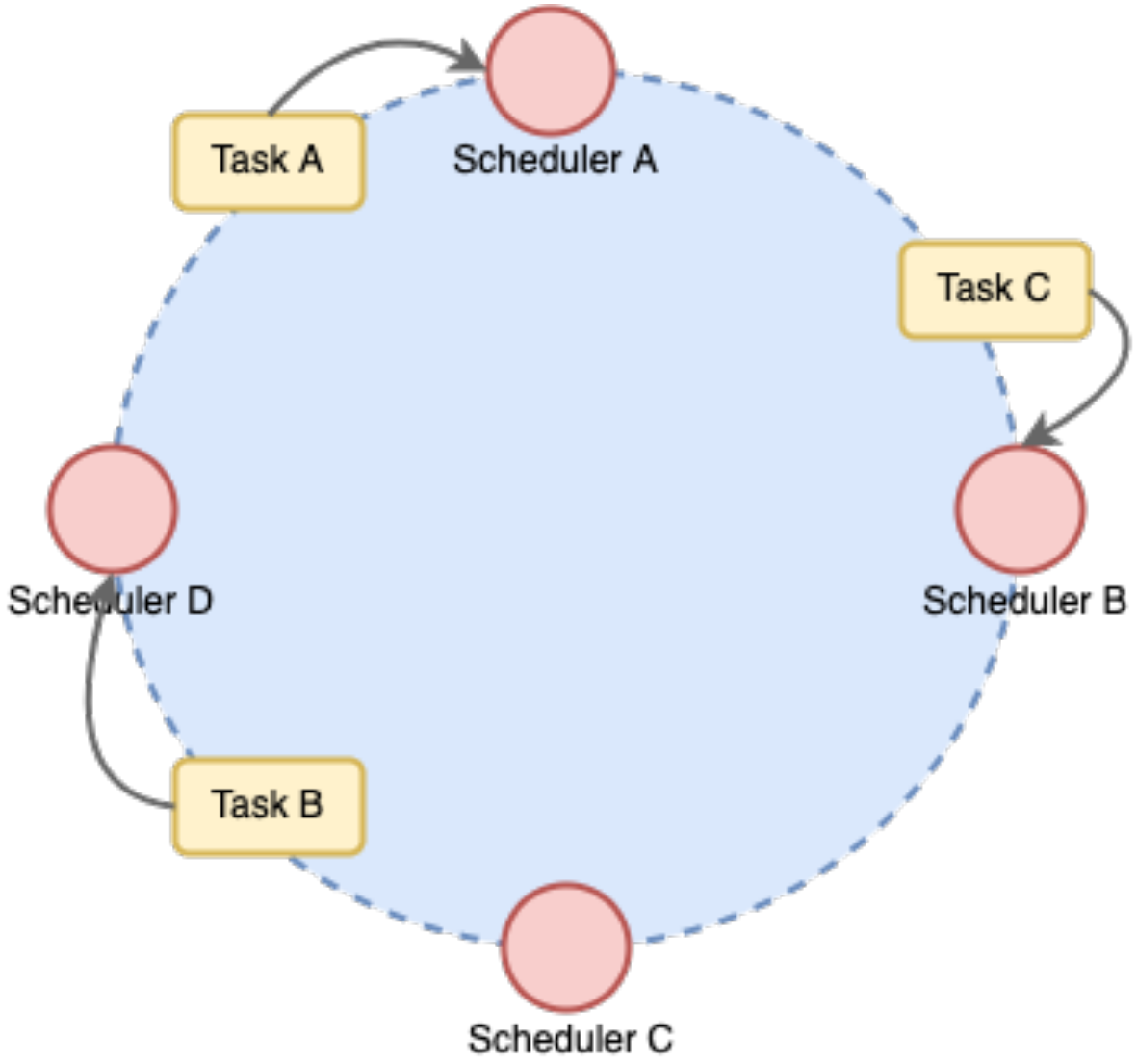
Consistent Hashing

Balancer:

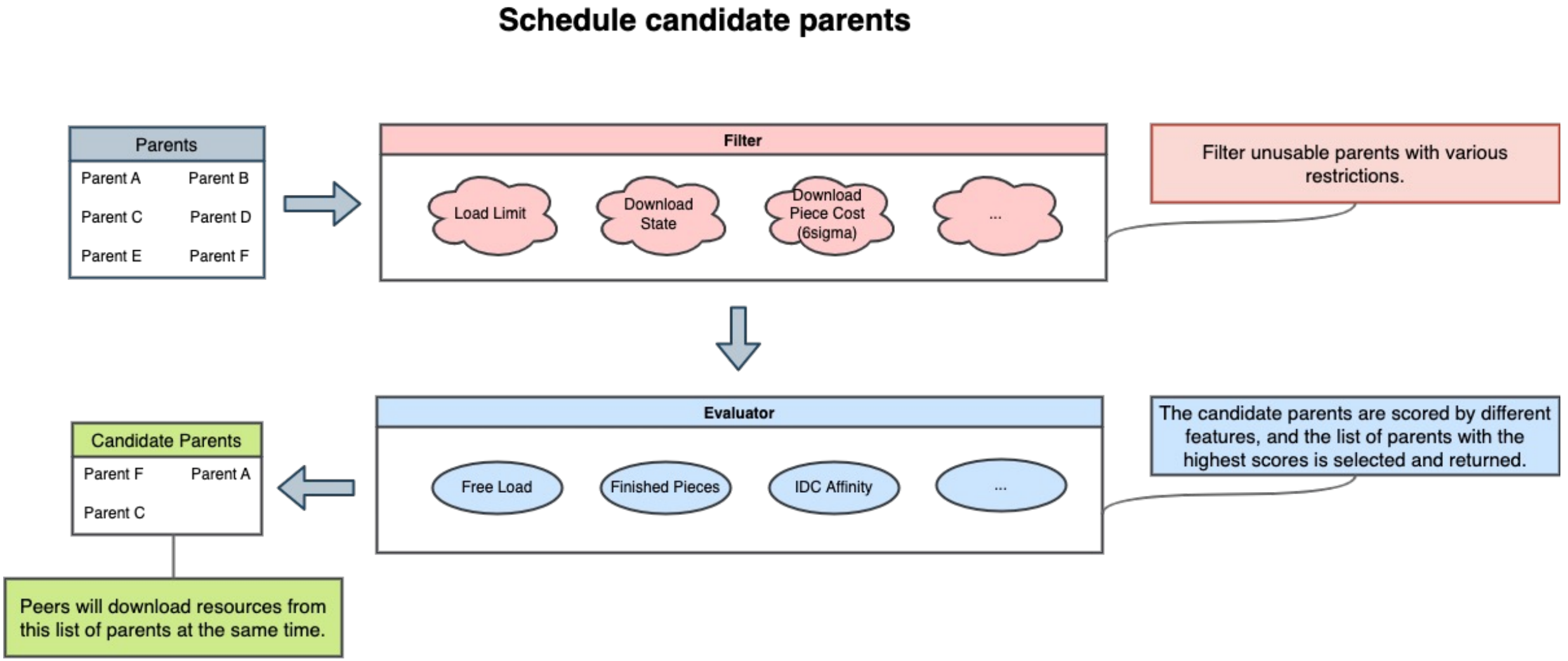
Consistent hashing is accomplished by reimplementing the **balancer.PickerBuilder** interface for grpc's base balancer.

Resolver:

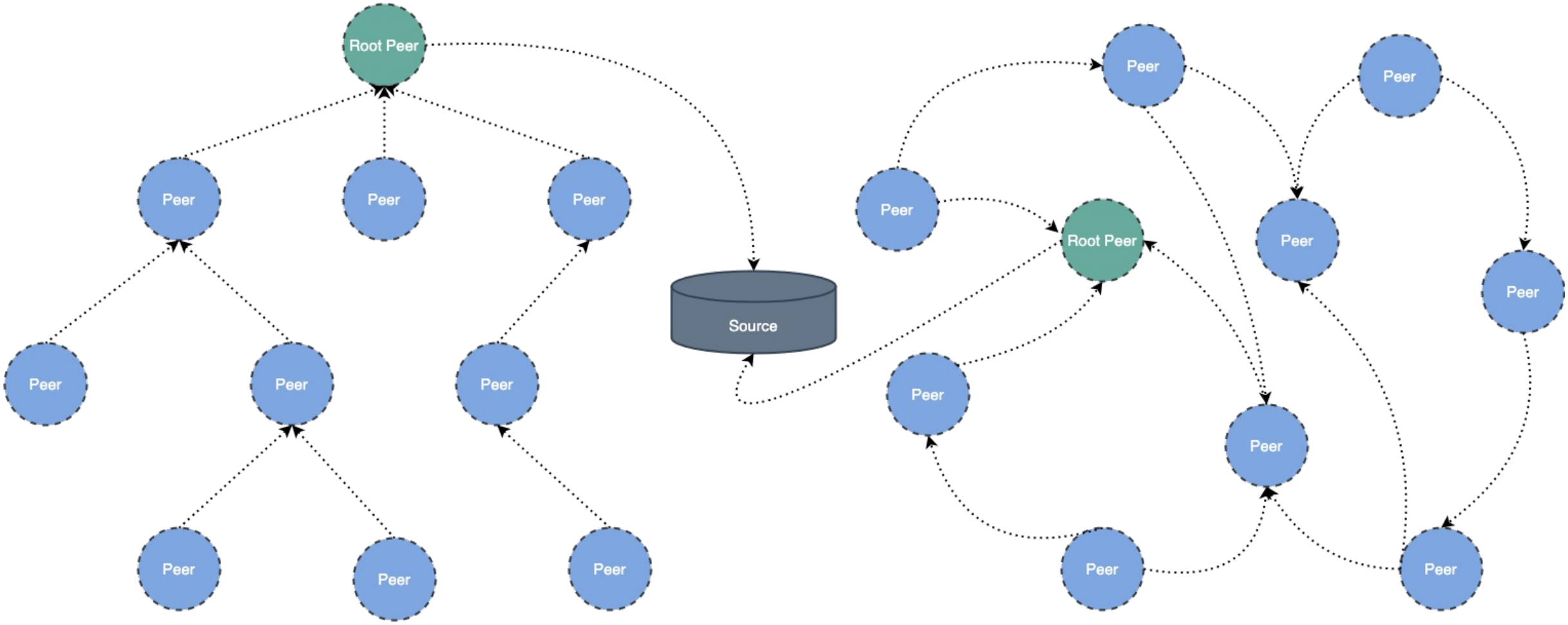
By reimplementing grpc's **resolver.Builder** interface, when there is a GRPC exception, the resolver will refresh the available addresses and update the hashring.



Filter And Evaluator



Directed Acyclic Graph



Download from a parent(v2.0.2)

Download from multiple parents(v2.0.3+)

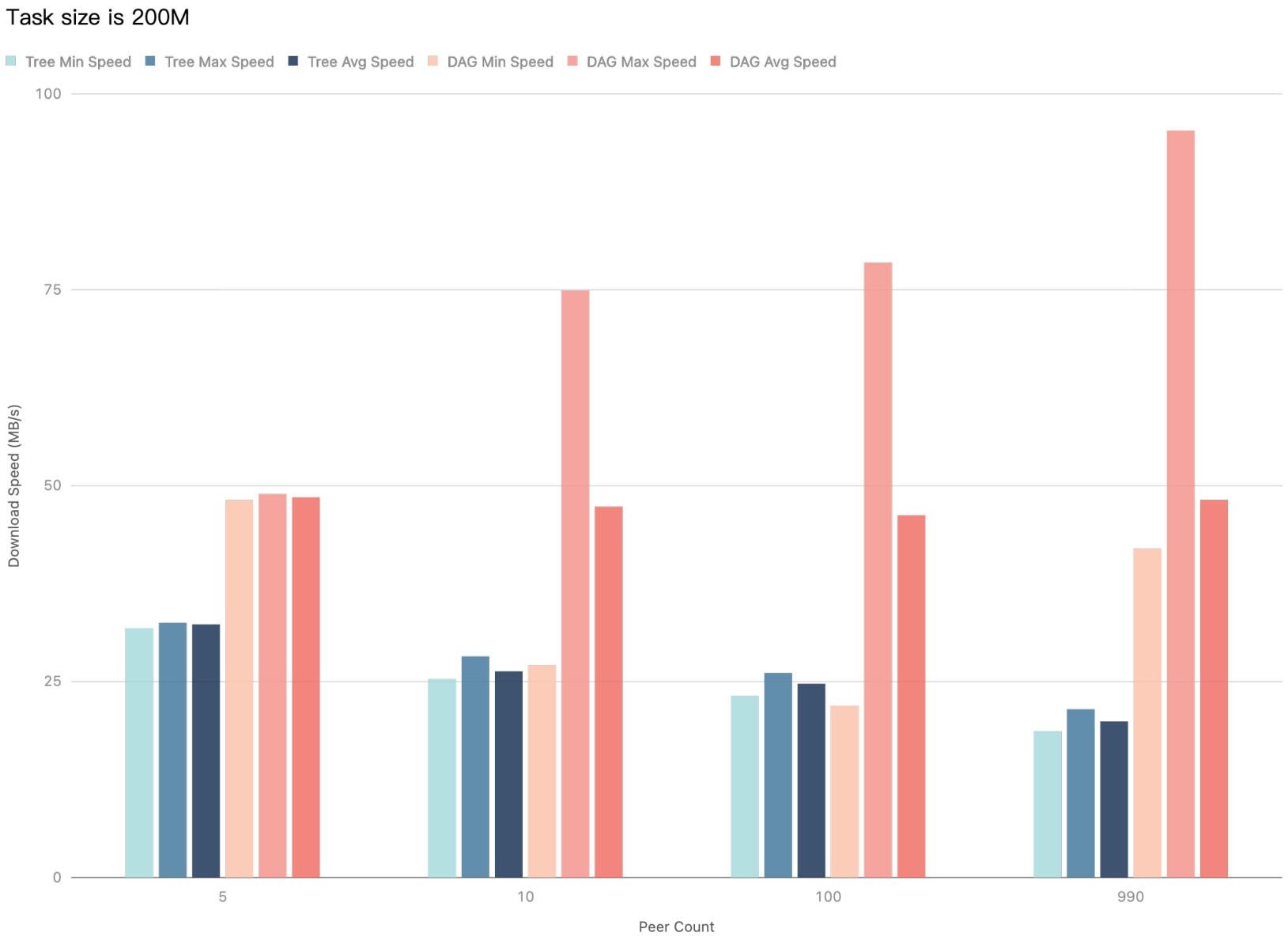
Directed Acyclic Graph

Environment:

Downstream bandwidth is **100MB/s**.
Upstream bandwidth is **100MB/s**.
Peers are deployed in **different IDCs**, so the bandwidth between peers is **not fixed**. Peer is 4C4G ECS and load limit is 8, Seed Peer is 4C4G ECS and load limit is 2. Scheduler has only 8C8G ECS.

Performance Testing:

The **first concurrent download task** in the P2P cluster does **not hit any cache**, and there is **only one peer back-to-source** download task. Control the number of different peers and download **200M** file.



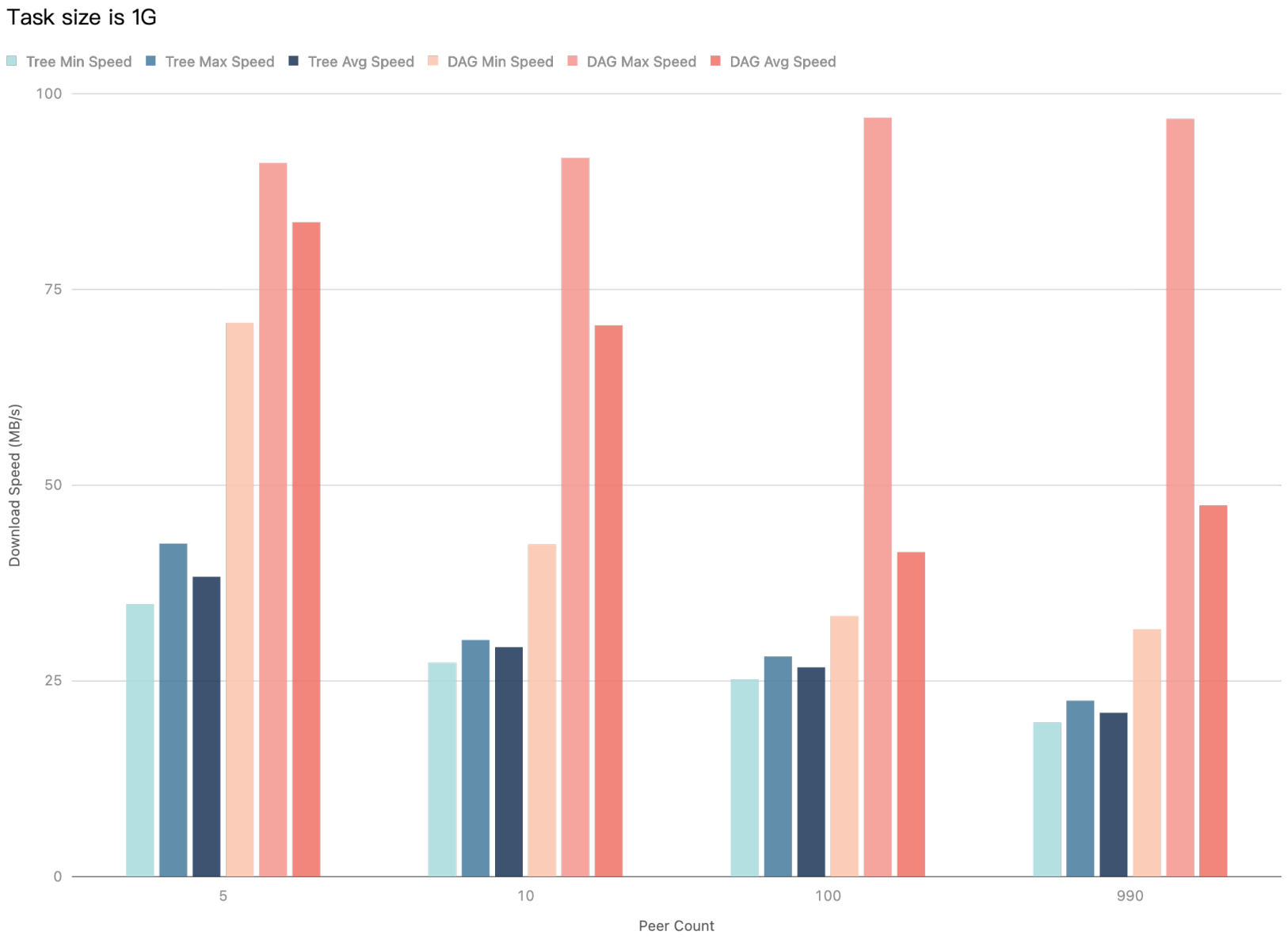
Directed Acyclic Graph

Environment:

Downstream bandwidth is **100MB/s**. Upstream bandwidth is **100MB/s**. Peers are deployed in **different IDCs**, so the bandwidth between peers is **not fixed**. Peer is 4C4G ECS and load limit is 8, Seed Peer is 4C4G ECS and load limit is 2. Scheduler has only 8C8G ECS.

Performance Testing:

The **first concurrent download task** in the P2P cluster does **not hit any cache**, and there is **only one peer back-to-source** download task. Control the number of different peers and download **1G** file.



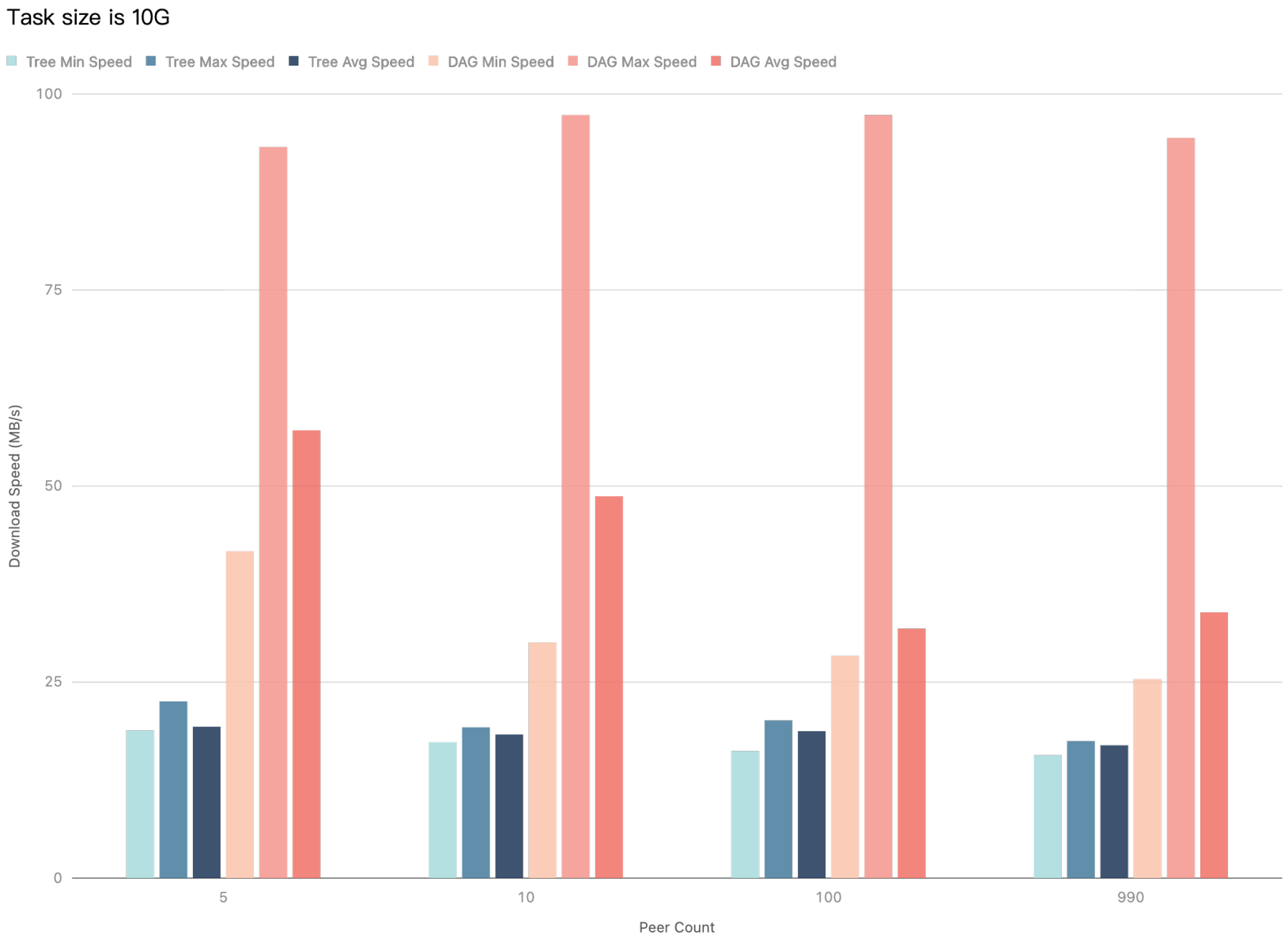
Directed Acyclic Graph

Environment:

Downstream bandwidth is **100MB/s**.
Upstream bandwidth is **100MB/s**.
Peers are deployed in **different IDCs**, so the bandwidth between peers is **not fixed**. Peer is 4C4G ECS and load limit is 8, Seed Peer is 4C4G ECS and load limit is 2. Scheduler has only 8C8G ECS.

Performance Testing:

The **first concurrent download task** in the P2P cluster does **not hit any cache**, and there is **only one peer back-to-source** download task. Control the number of different peers and download **10G** file.



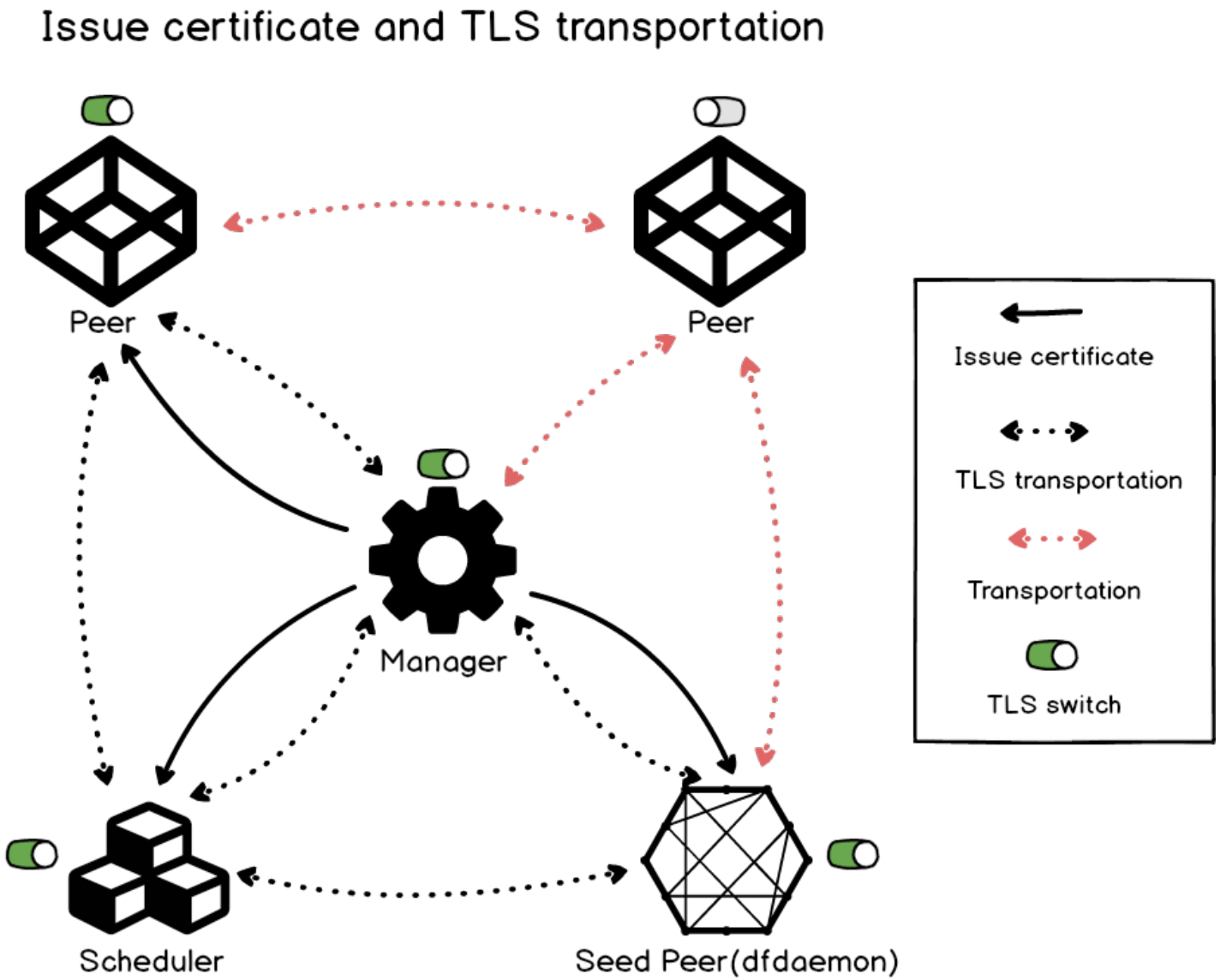
TLS Transportation

Manager:

Manger is dragonfly's CA and **issues certificates** for each service. Each service will call the manager api to get the **certificate chain** when it is started.

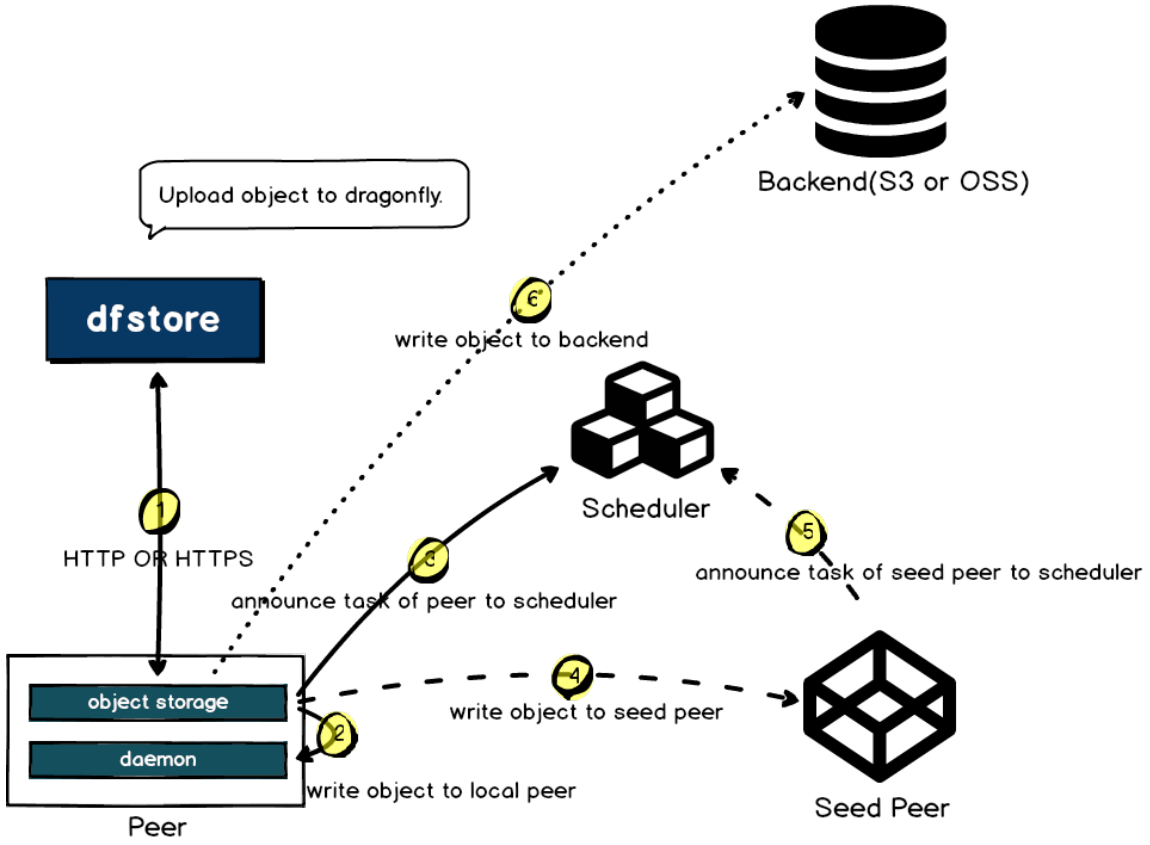
TLS Transportation:

If TLS transportation is enabled, all data transfers will be over TLS, and the GRPC service will use **mutual TLS**.

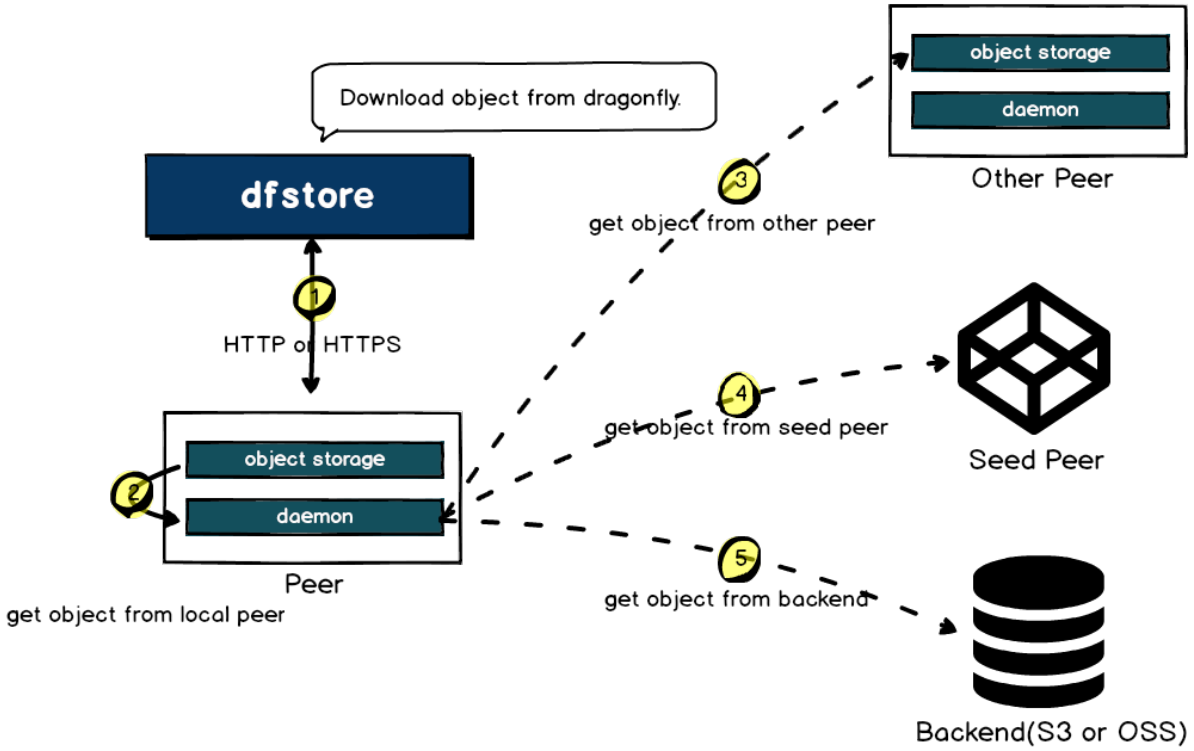


Dfstore

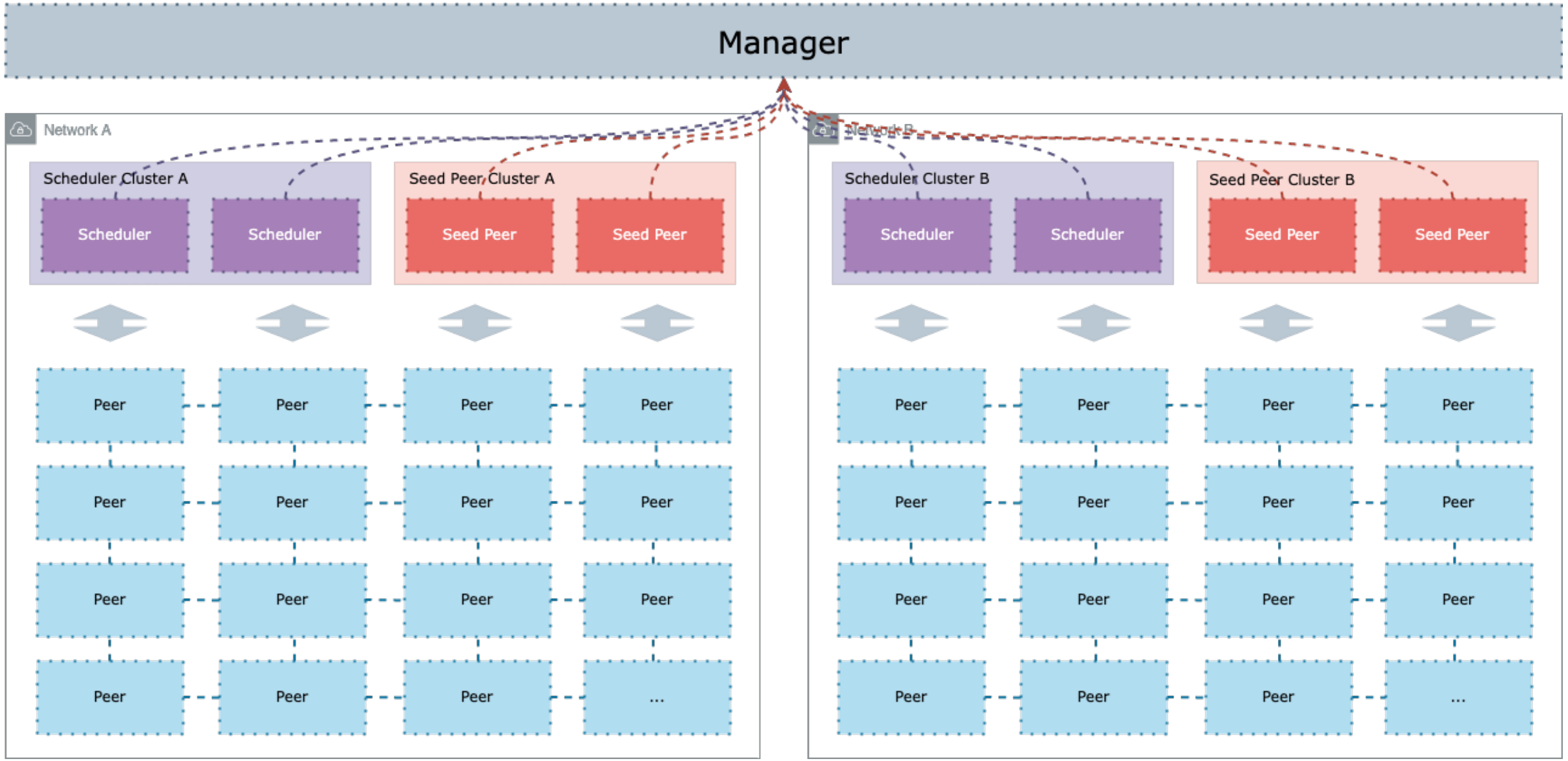
Upload Object



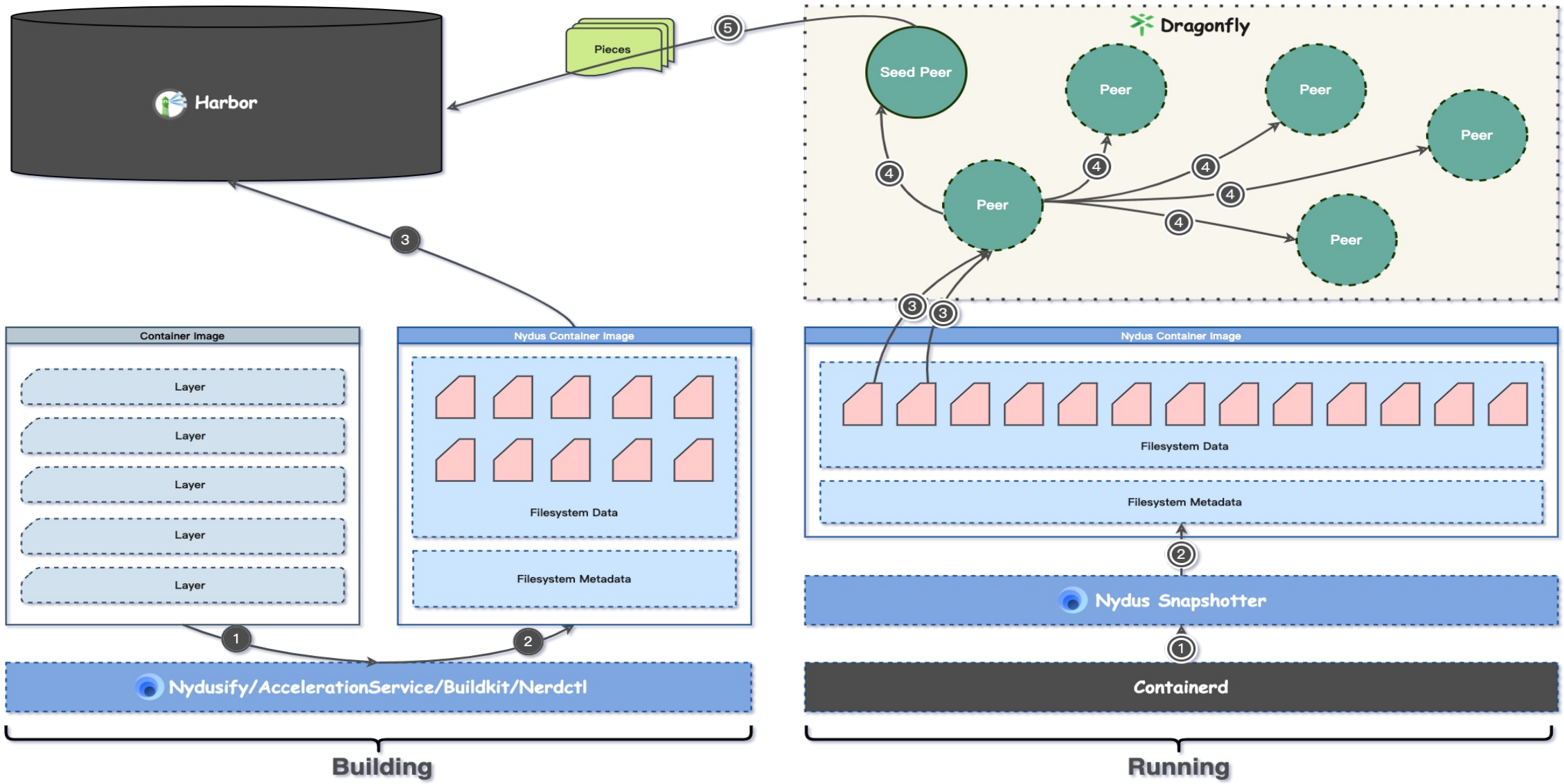
Download Object



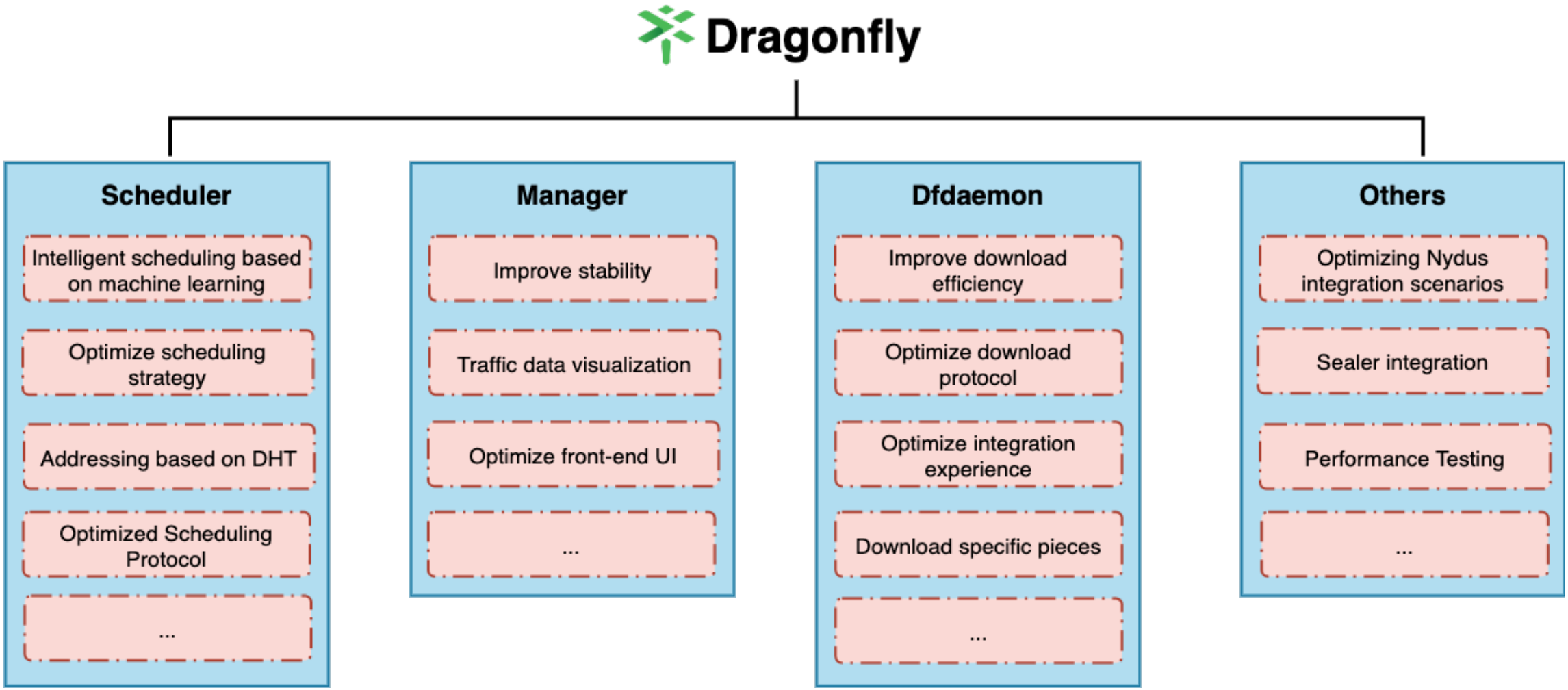
Multi-Cluster Dragonfly



Acceleration Framework For Image



Future


Dragonfly
Scheduler
Intelligent scheduling based on machine learning
Optimize scheduling strategy
Addressing based on DHT
Optimized Scheduling Protocol
...
Manager
Improve stability
Traffic data visualization
Optimize front-end UI
...
Dfdaemon
Improve download efficiency
Optimize download protocol
Optimize integration experience
Download specific pieces
...
Others
Optimizing Nydus integration scenarios
Sealer integration
Performance Testing
...

KubeCon CloudNativeCon North America 2022

BUILDING FOR THE ROAD AHEAD

DETROIT 2022

THANK YOU!

Session QR Codes will be sent via email before the event

Github: https://github.com/dragonflyoss/Dragonfly2/

Website: https://d7y.io/

Slack Channel: https://cloud-native.slack.com/messages/dragonfly/

Twitter: @dragonfly_oss

Discussion Group: dragonfly-discuss@googlegroups.com

DingTalk