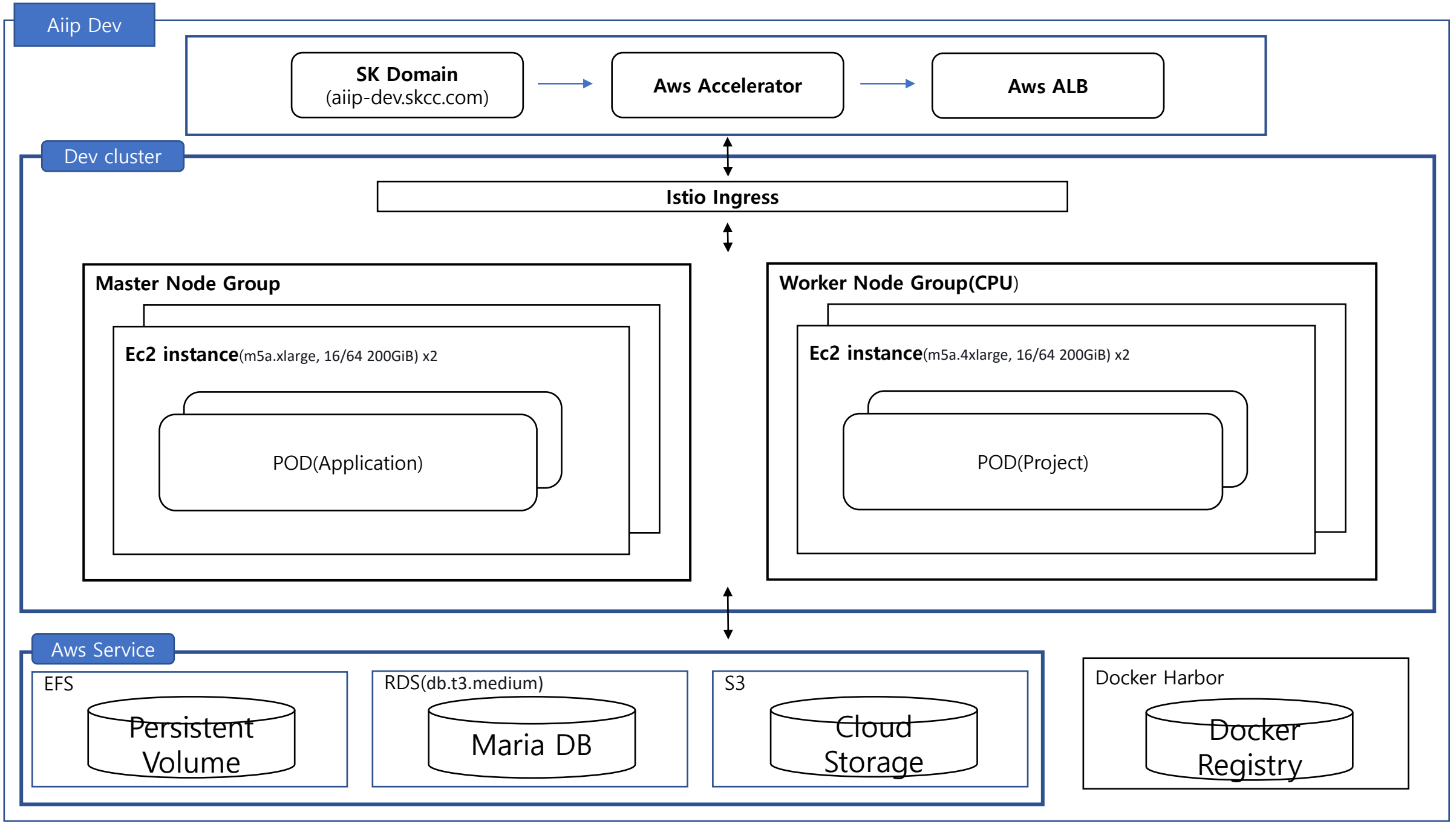
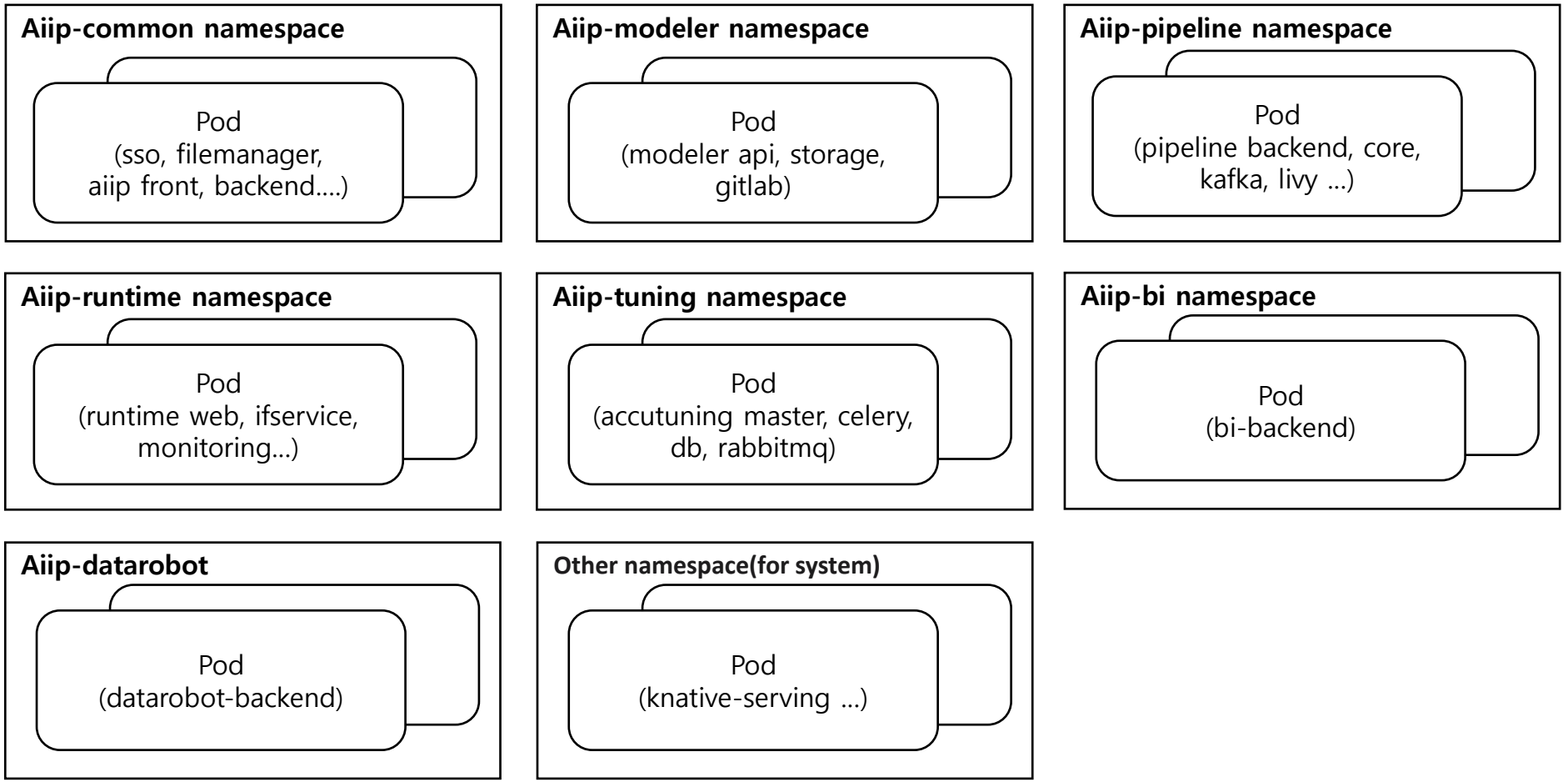
Accutuning - AWS


Aiip Dev
SK Domain
(aiip-dev.skcc.com)
Aws Accelerator
Aws ALB
Dev cluster
Istio Ingress
Master Node Group
Ec2 instance(m5a.xlarge, 16/64 200GiB) x2
POD(Application)
Worker Node Group(CPU)
Ec2 instance(m5a.4xlarge, 16/64 200GiB) x2
POD(Project)
Aws Service
EFS
Persistent Volume
RDS(db.t3.medium)
Maria DB
S3
Cloud Storage
Docker Harbor
Docker Registry

Master Node Group

* Running 상태로 계속해서 사용되어야 하는 Pod(Application 전용)는 Master Node에 할당된다

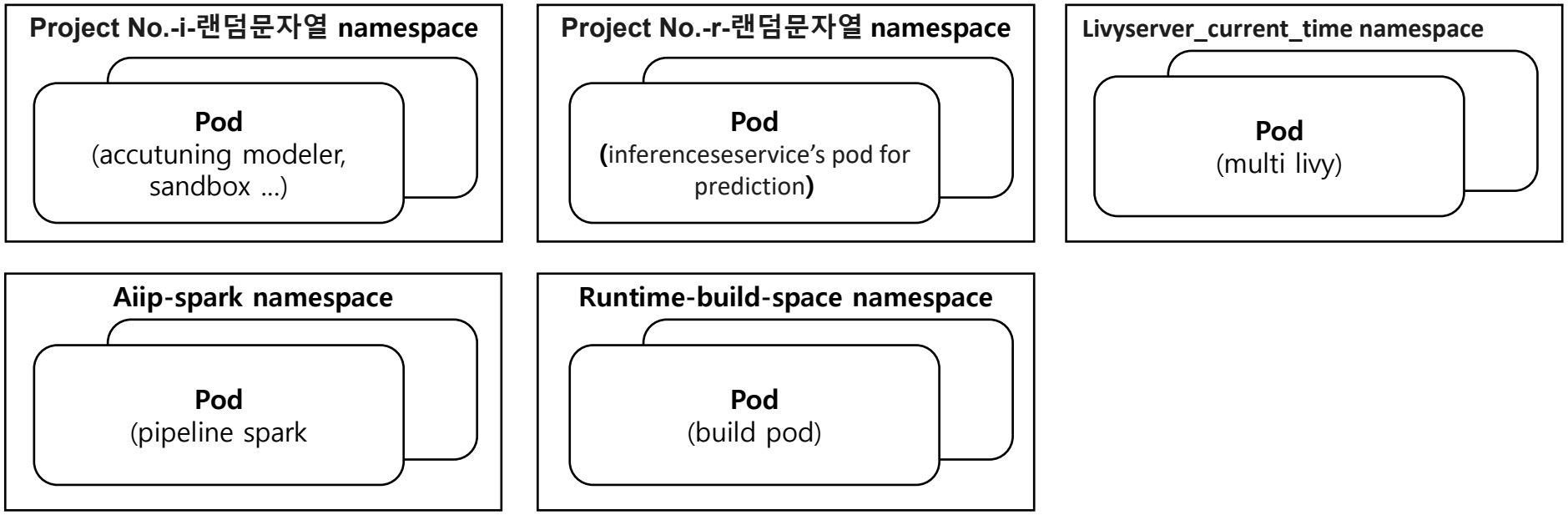
EC2 Instance



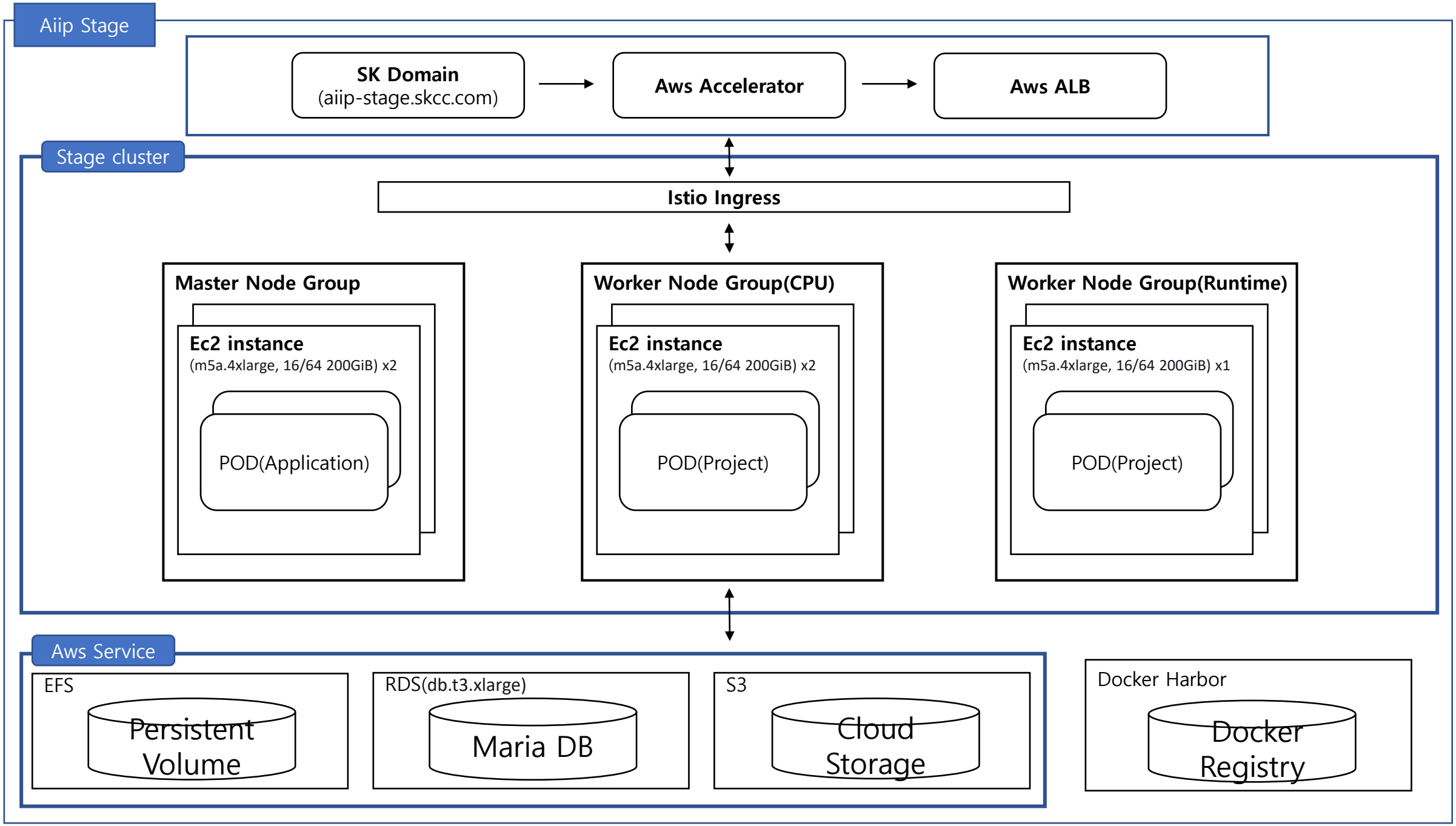
Worker Node Group

* Project 생성시 사용되었다가 반납되는 Pod(Project 전용)는 worker node에 할당된다

EC2 Instance



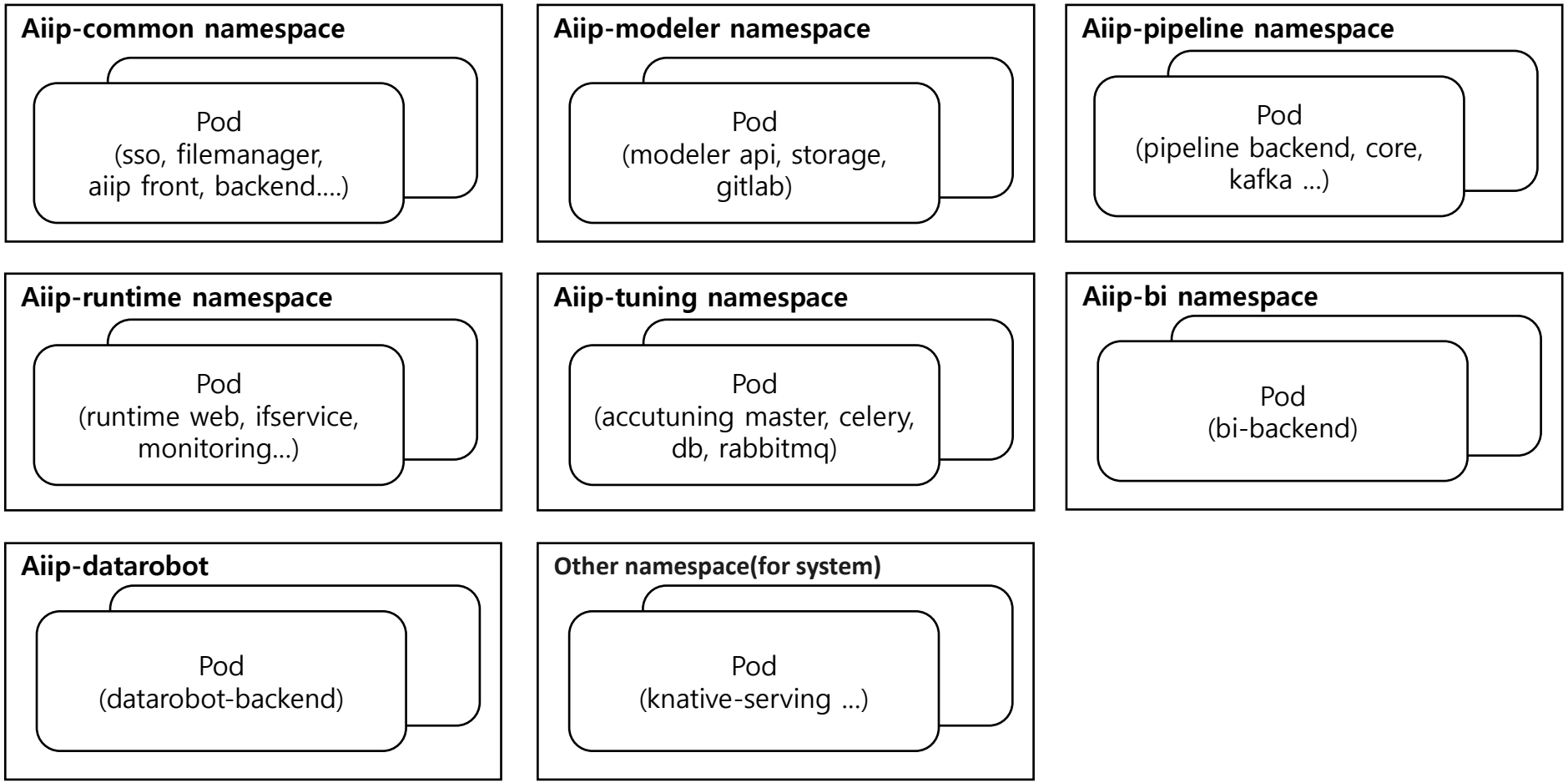
Accuinsight+ 3.0 – Stage Cluster


Aiip Stage
SK Domain
(aiip-stage.skcc.com)
Aws Accelerator
Aws ALB
Stage cluster
Istio Ingress
Master Node Group
Ec2 instance
(m5a.4xlarge, 16/64 200GiB) x2
POD(Application)
Worker Node Group(CPU)
Ec2 instance
(m5a.4xlarge, 16/64 200GiB) x2
POD(Project)
Worker Node Group(Runtime)
Ec2 instance
(m5a.4xlarge, 16/64 200GiB) x1
POD(Project)
Aws Service
EFS
Persistent Volume
RDS(db.t3.xlarge)
Maria DB
S3
Cloud Storage
Docker Harbor
Docker Registry

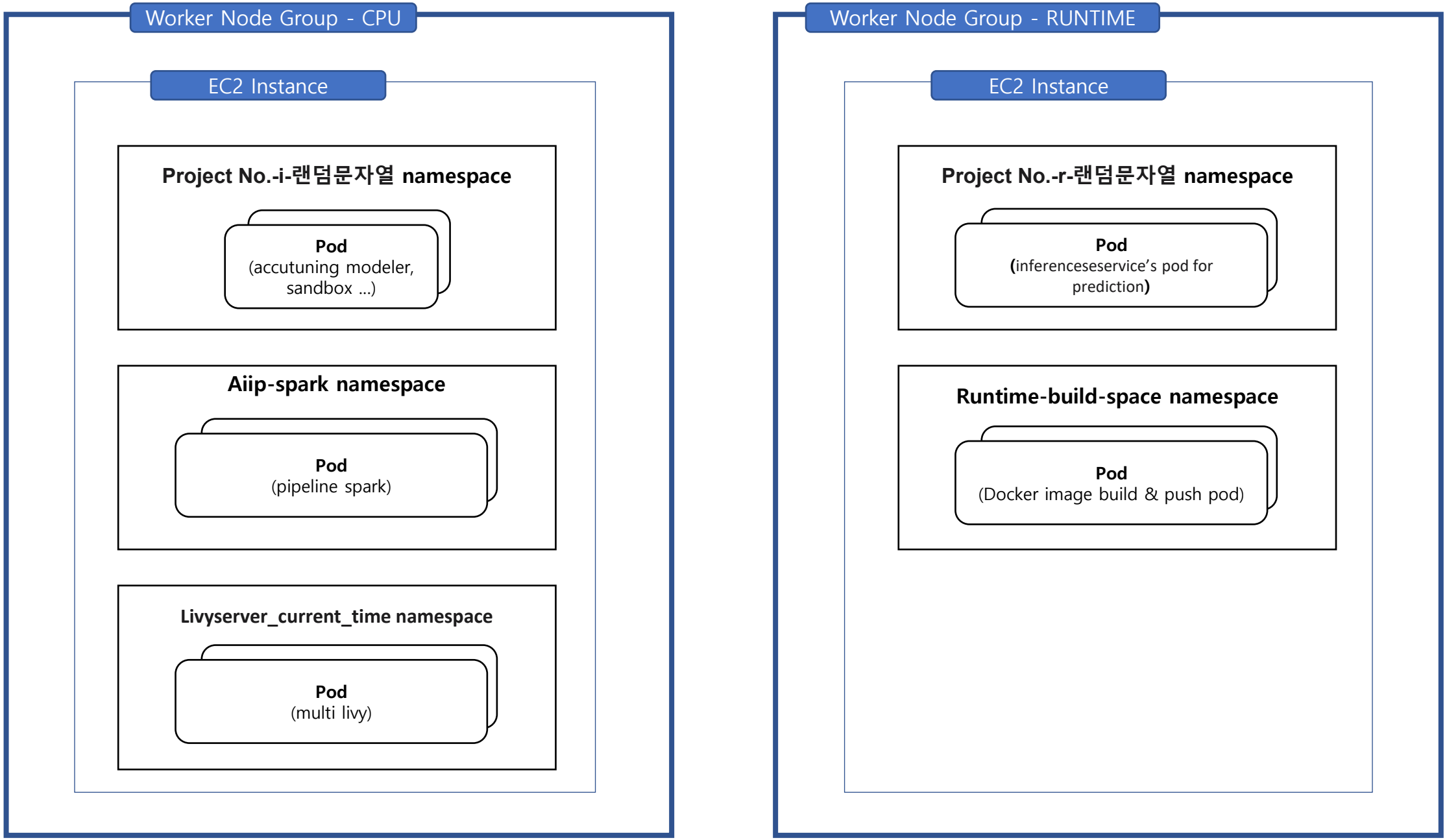
Master Node Group

* Running 상태로 계속해서 사용되어야 하는 Pod(Application 전용)는 Master Node에 할당된다

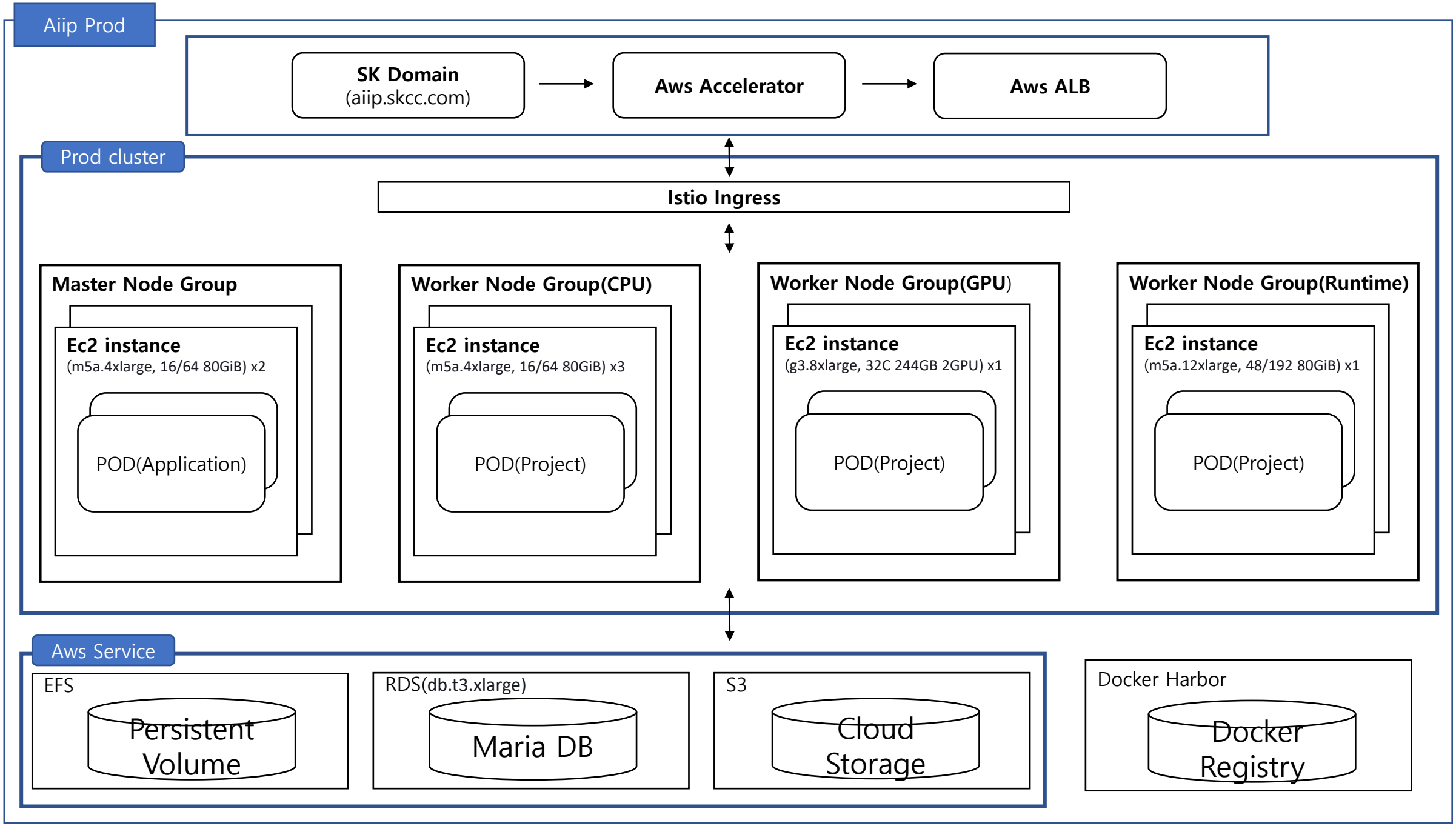
EC2 Instance



* Project 생성시 사용되었다가 반납되는 Pod(Project 전용)는 worker node에 할당된다.



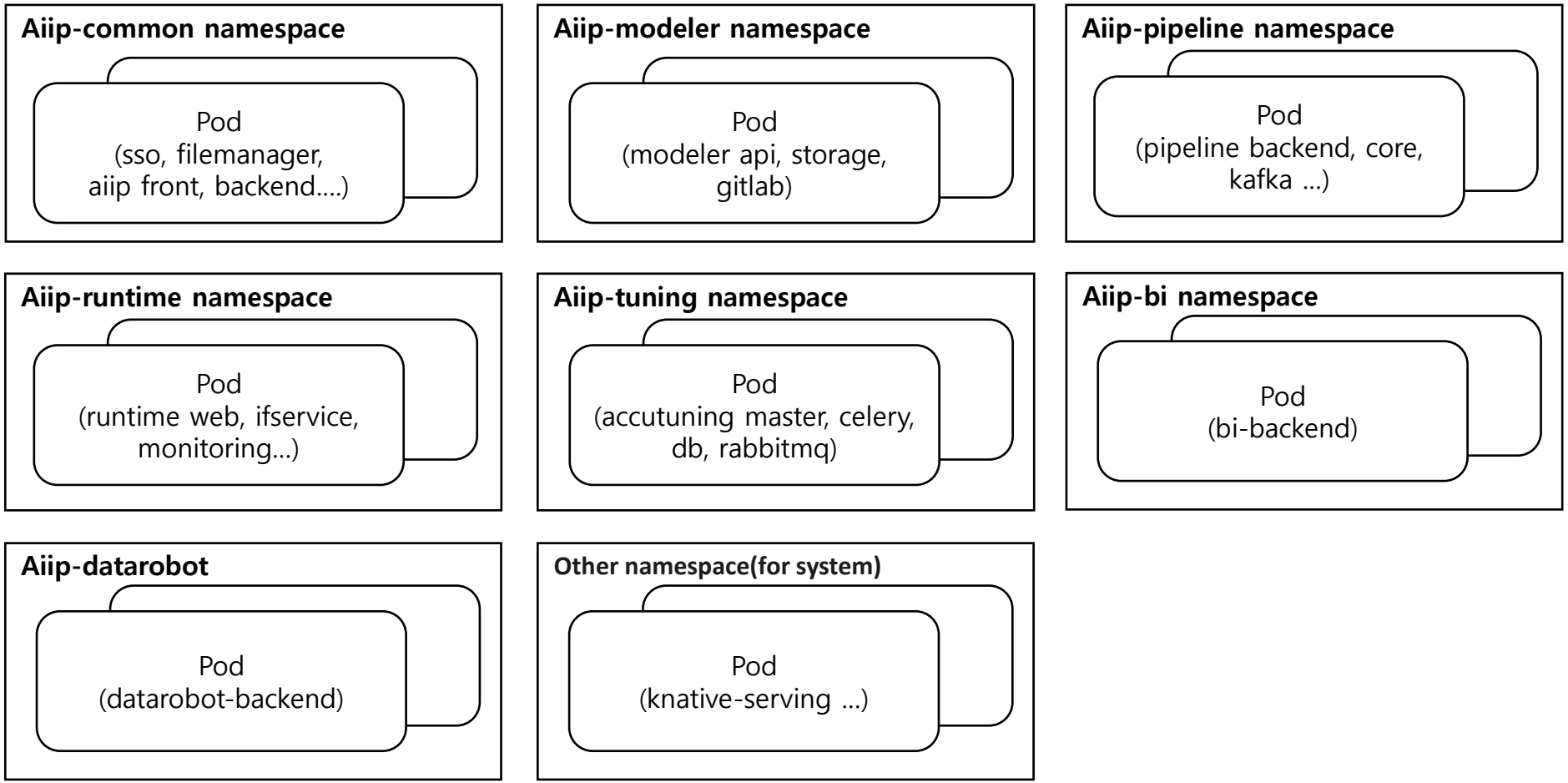
Accuinsight+ 3.0 – Prod Cluster


Aiip Prod
SK Domain
(aiip.skcc.com)
Aws Accelerator
Aws ALB
Prod cluster
Istio Ingress
Master Node Group
Ec2 instance
(m5a.4xlarge, 16/64 80GiB) x2
POD(Application)
Worker Node Group(CPU)
Ec2 instance
(m5a.4xlarge, 16/64 80GiB) x3
POD(Project)
Worker Node Group(GPU)
Ec2 instance
(g3.8xlarge, 32C 244GB 2GPU) x1
POD(Project)
Worker Node Group(Runtime)
Ec2 instance
(m5a.12xlarge, 48/192 80GiB) x1
POD(Project)
Aws Service
EFS
Persistent Volume
RDS(db.t3.xlarge)
Maria DB
S3
Cloud Storage
Docker Harbor
Docker Registry

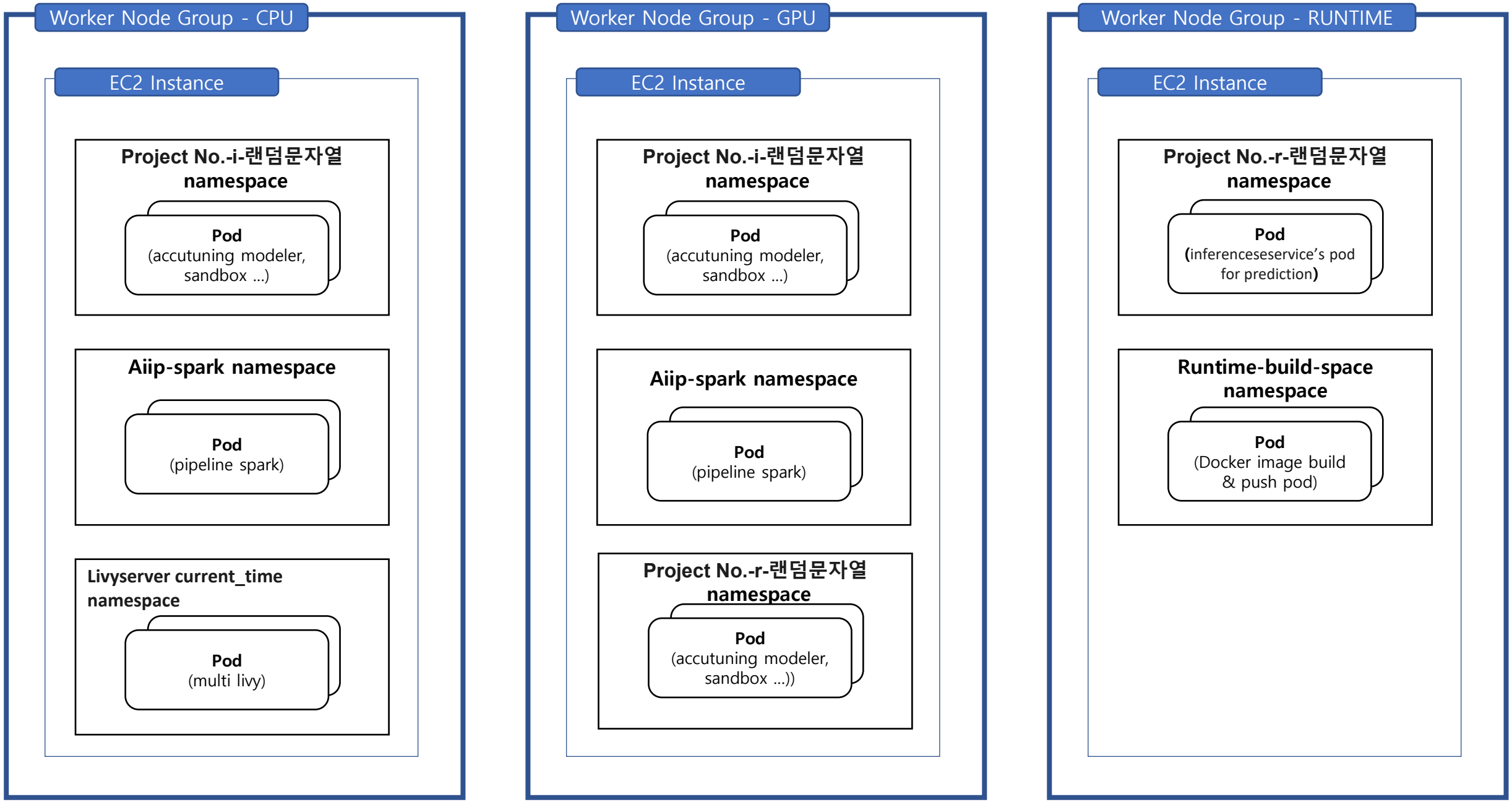
Master Node Group

* Running 상태로 계속해서 사용되어야 하는 Pod(Application 전용)는 Master Node에 할당된다

EC2 Instance



* Project 생성시 사용되었다가 반납되는 Pod(Project 전용)는 worker node에 할당된다 (GPU를 사용하도록 설정되었을 경우, GPU Node에 할당된다)



* Prod - Infra Layer