

Tenant_A

Tenant_B

Tenant_C

W_A

W_B

W_C

App₁

App₂

App₃

App₁

App₂

App₃

App₁

App₂

App₃

Emulated
GPGPU Server

Node0

Node1

NodeN

gMap

gMap

gMap

CPU0

CPU1

CPU0

CPU1

CPU0

CPU1

CPU2

CPU3

CPU2

CPU3

CPU2

CPU3

Interposer Library

Interposer Library

Interposer Library

GPGPU Request

(1)

(0)

Device Selection call

(N)

GPU Affinity Mapper/ Load Balancer

Feedback

Context
Packer

Context
Packer

Context
Packer

GPU Scheduler

GPU Scheduler

GPU Scheduler

GPU₀

GPU₁

...

GPU_N

gPool