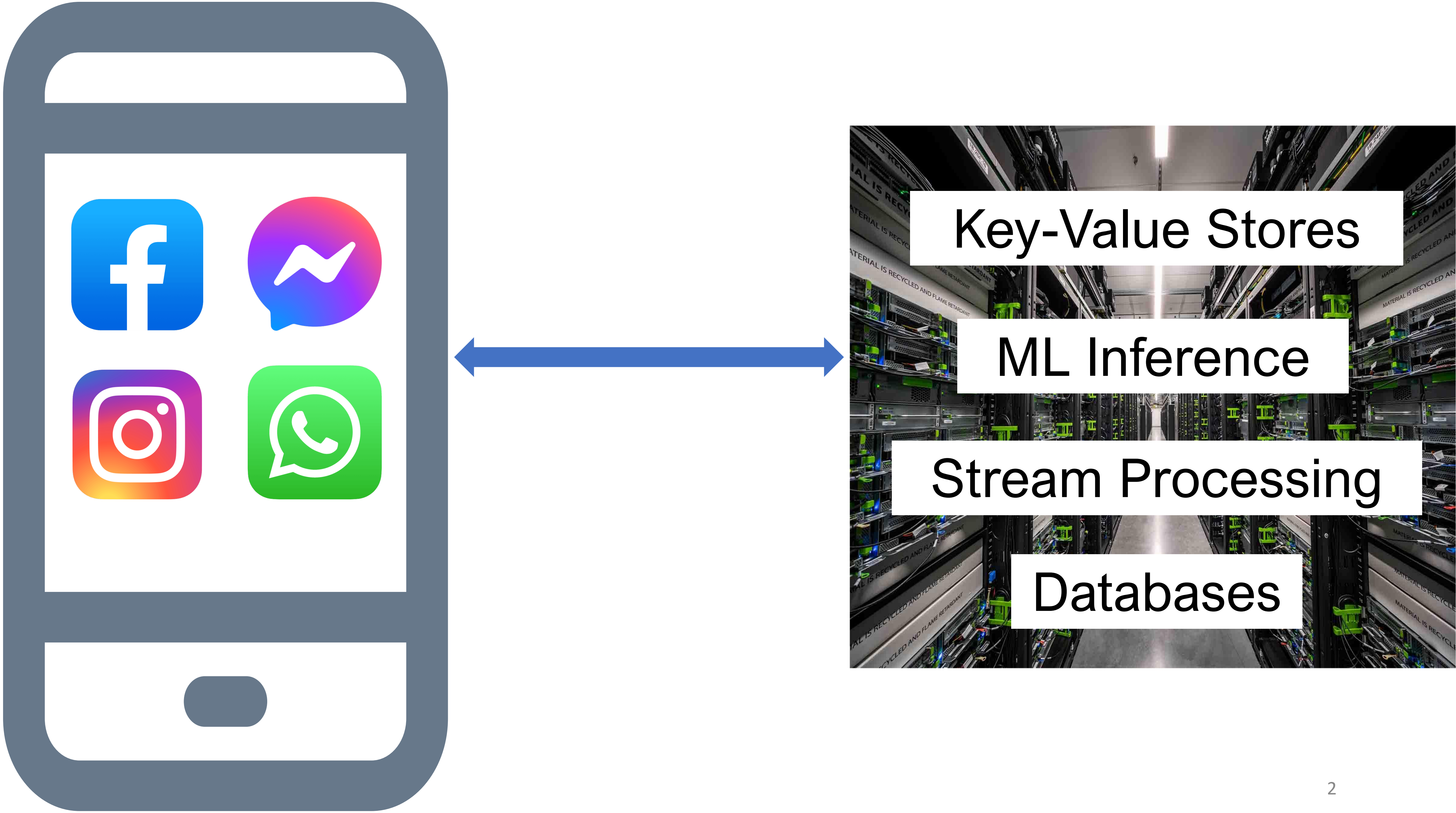
RAS: Continuously Optimized Region-Wide Datacenter Resource Allocation

Andrew Newell, **Dimitrios Skarlatos**‡, Jingyuan Fan, Pavan Kumar, Maxim Khutornenko, Mayank Pundir, Yirui Zhang, Mingjun Zhang, Yuanlai Liu, Linh Le, Brendon Daugherty, Apurva Samudra, Prashasti Baid, James Kneeland, Igor Kabiljo, Dmitry Shchukin, Andre Rodrigues, Scott Michelson, Ben Christensen, Kaushik Veeraraghavan, Chunqiang Tang

FACEBOOK Infrastructure

‡ Carnegie Mellon University Computer Science Department



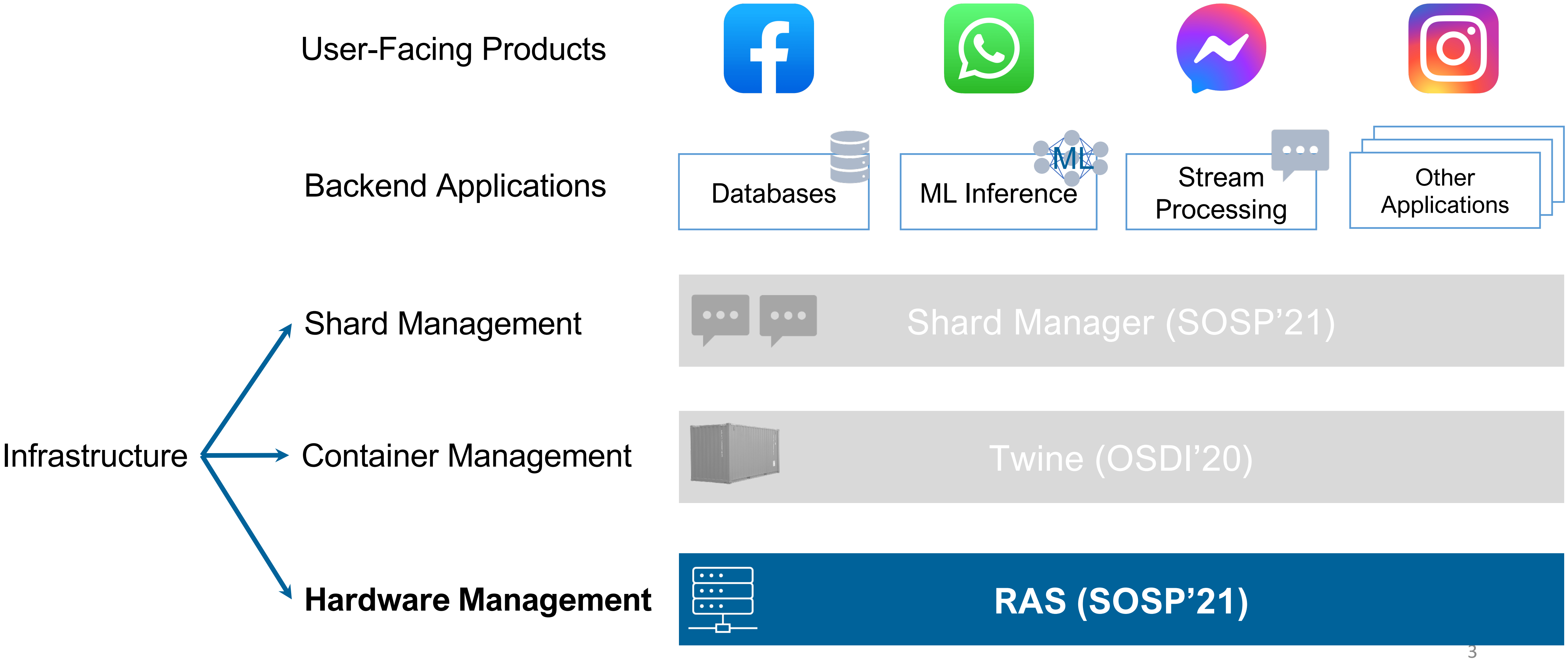
Key-Value Stores

ML Inference

Stream Processing

Databases

RAS in the Software Stack



User-Facing Products

Backend Applications

Databases

ML Inference

Stream Processing

Other Applications

Infrastructure

Shard Management

Container Management

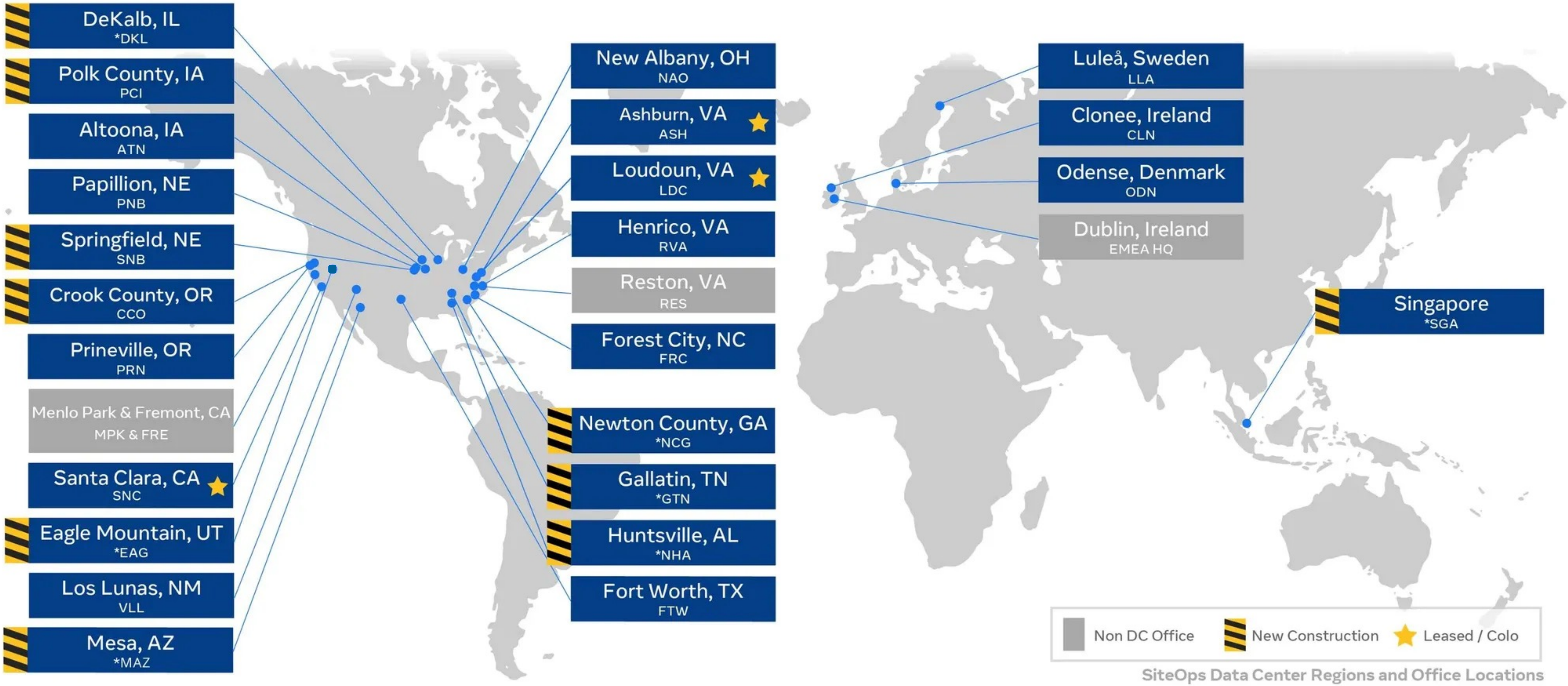
Hardware Management

Shard Manager (SOSP’21)

Twine (OSDI’20)

RAS (SOSP’21)

Facebook Datacenter Regions



▨ DeKalb, IL *DKL

▨ Polk County, IA PCI

Altoona, IA ATN

Papillion, NE PNB

▨ Springfield, NE SNB

▨ Crook County, OR CCO

Prineville, OR PRN

■ Menlo Park & Fremont, CA MPK & FRE

Santa Clara, CA SNC ★

▨ Eagle Mountain, UT *EAG

Los Lunas, NM VLL

▨ Mesa, AZ *MAZ

New Albany, OH NAO

Ashburn, VA ASH ★

Loudoun, VA LDC ★

Henrico, VA RVA

■ Reston, VA RES

Forest City, NC FRC

▨ Newton County, GA *NCG

▨ Gallatin, TN *GTN

▨ Huntsville, AL *NHA

Fort Worth, TX FTW

Luleå, Sweden LLA

Clonee, Ireland CLN

Odense, Denmark ODN

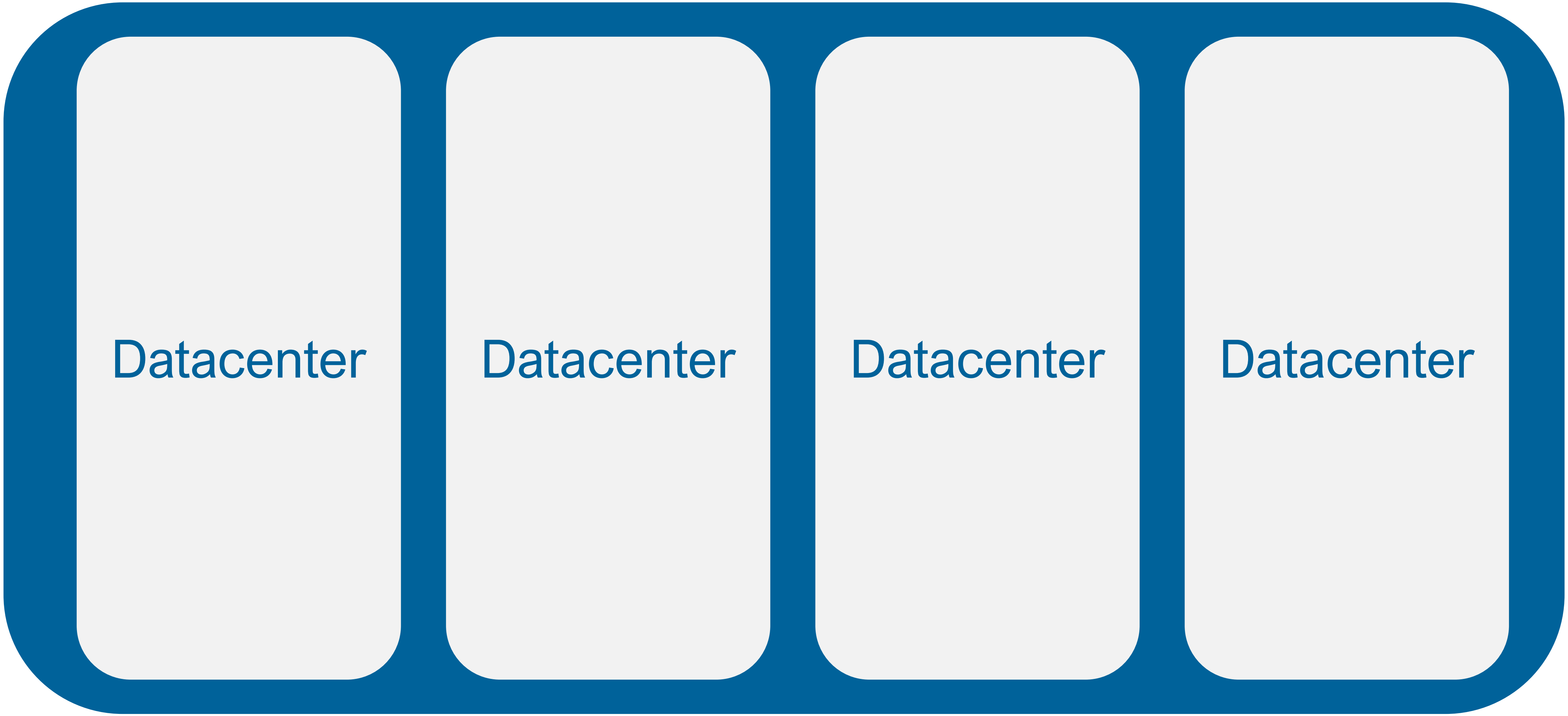
■ Dublin, Ireland EMEA HQ

▨ Singapore *SGA

■ Non DC Office ▨ New Construction ★ Leased / Colo

SiteOps Data Center Regions and Office Locations

Datacenter Region



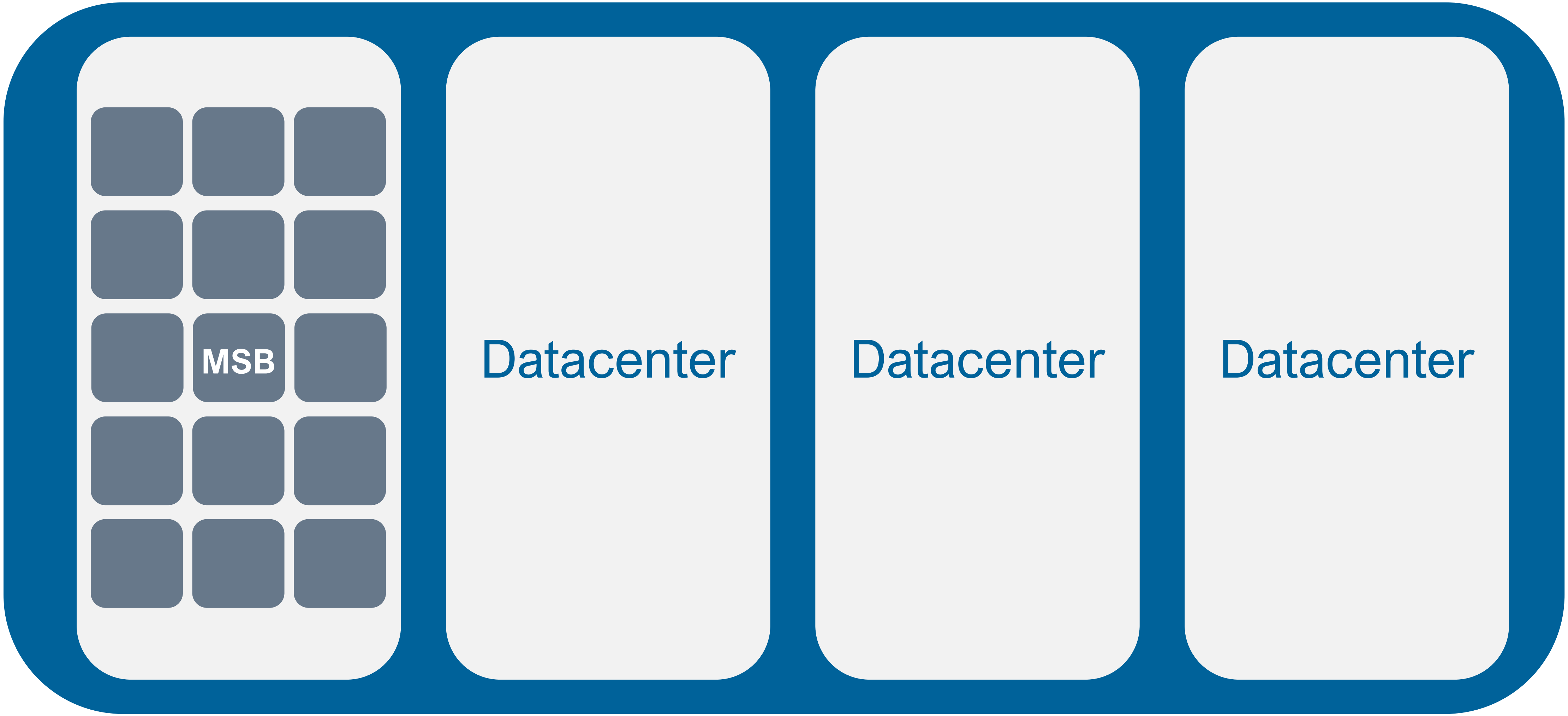
Datacenter

Datacenter

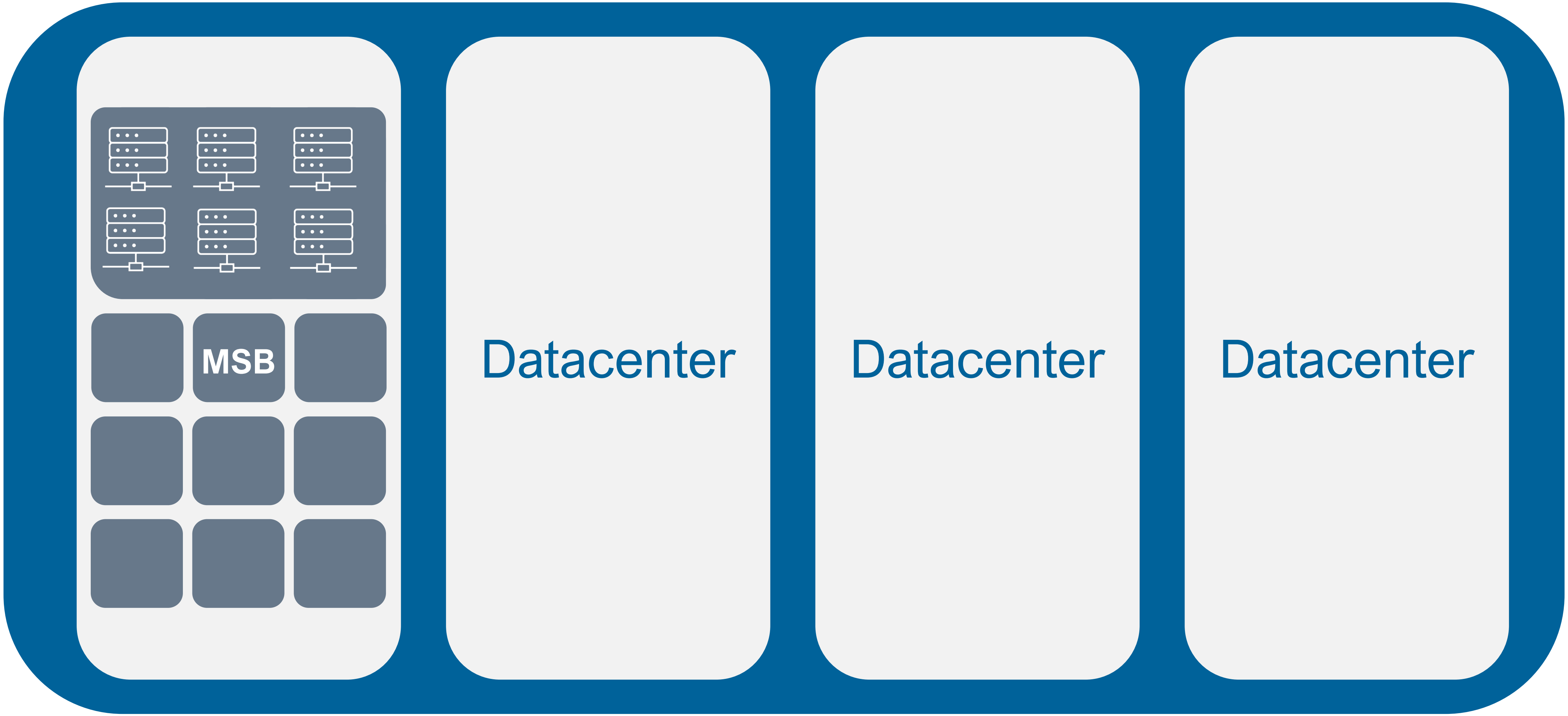
Datacenter

Datacenter

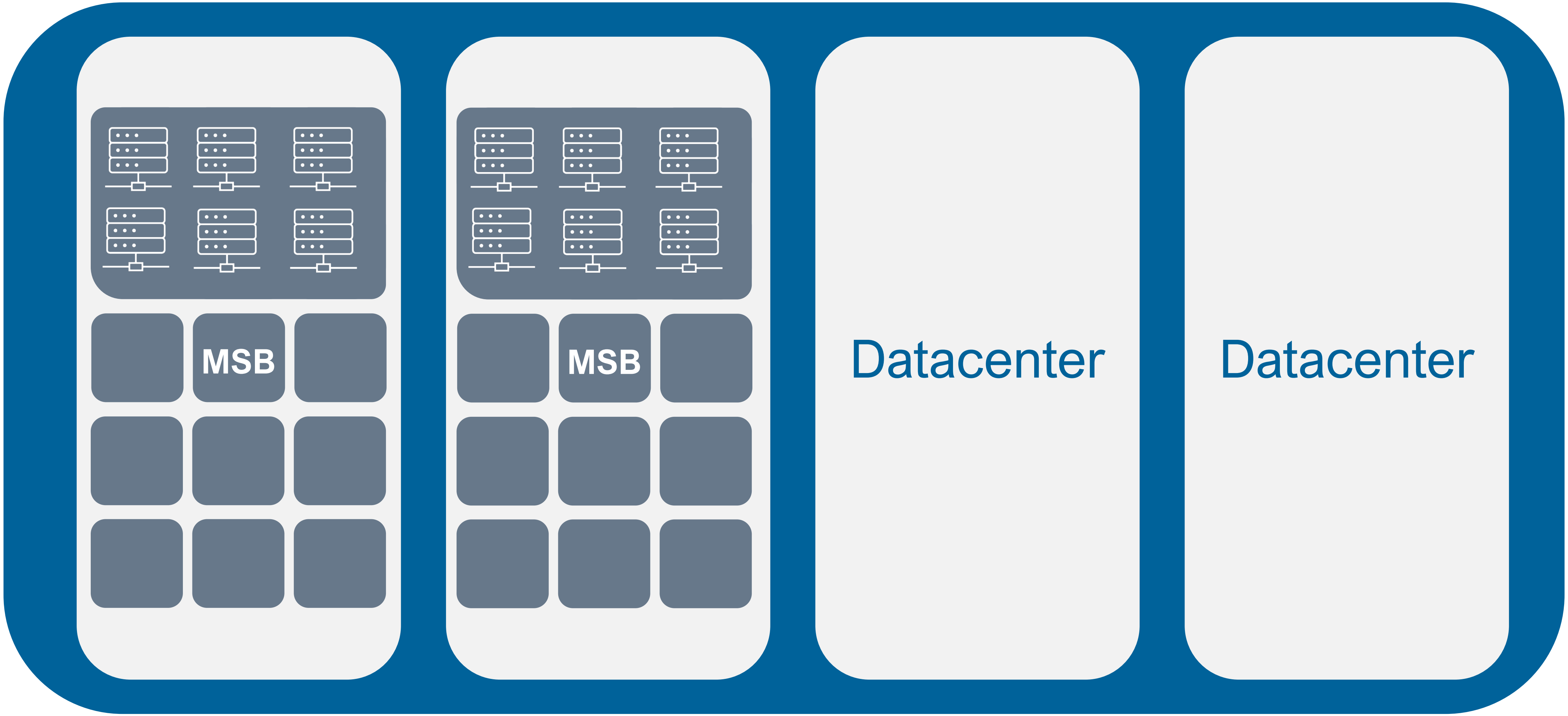
Failure Domains-Main Switch Board (MSB)



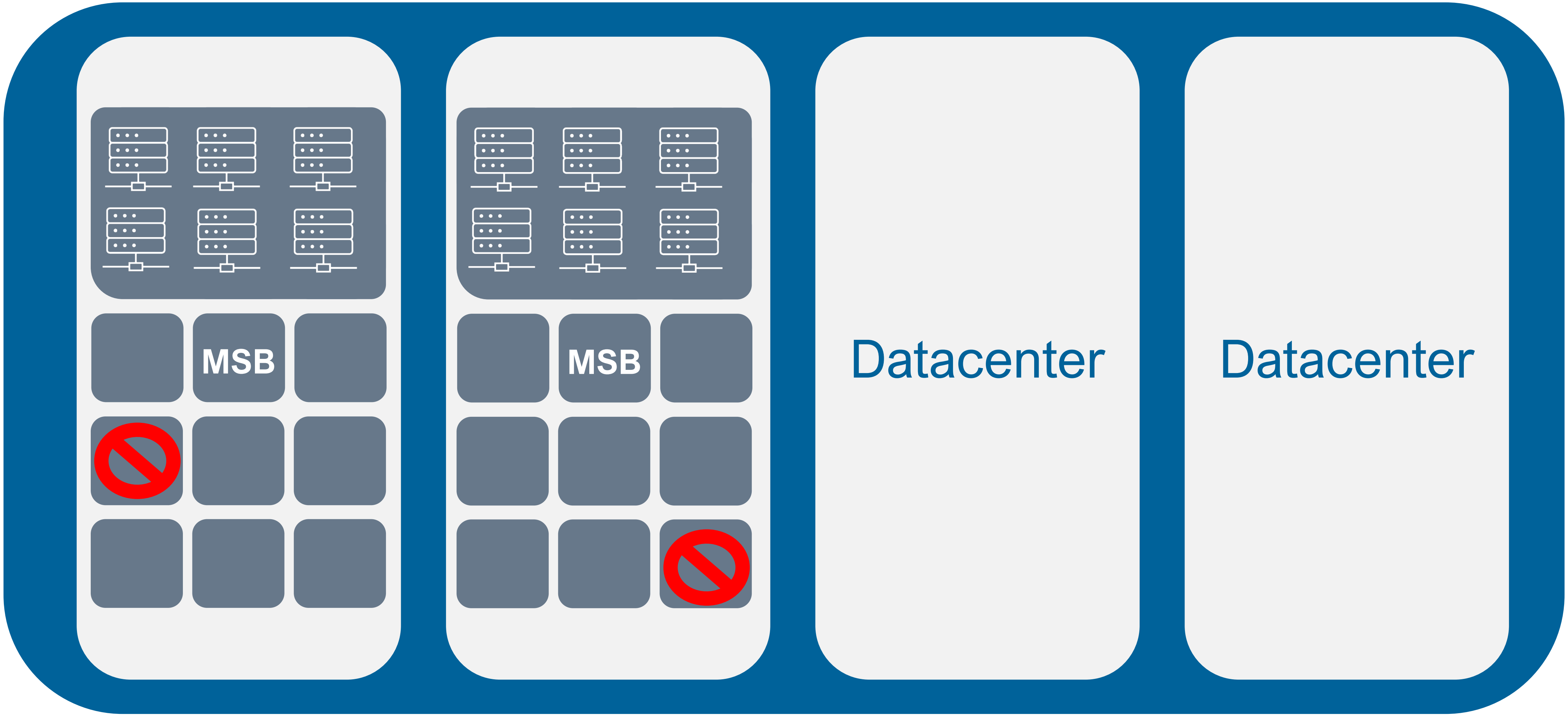
Failure Domains-Main Switch Board (MSB)



Failure Domains-Main Switch Board (MSB)

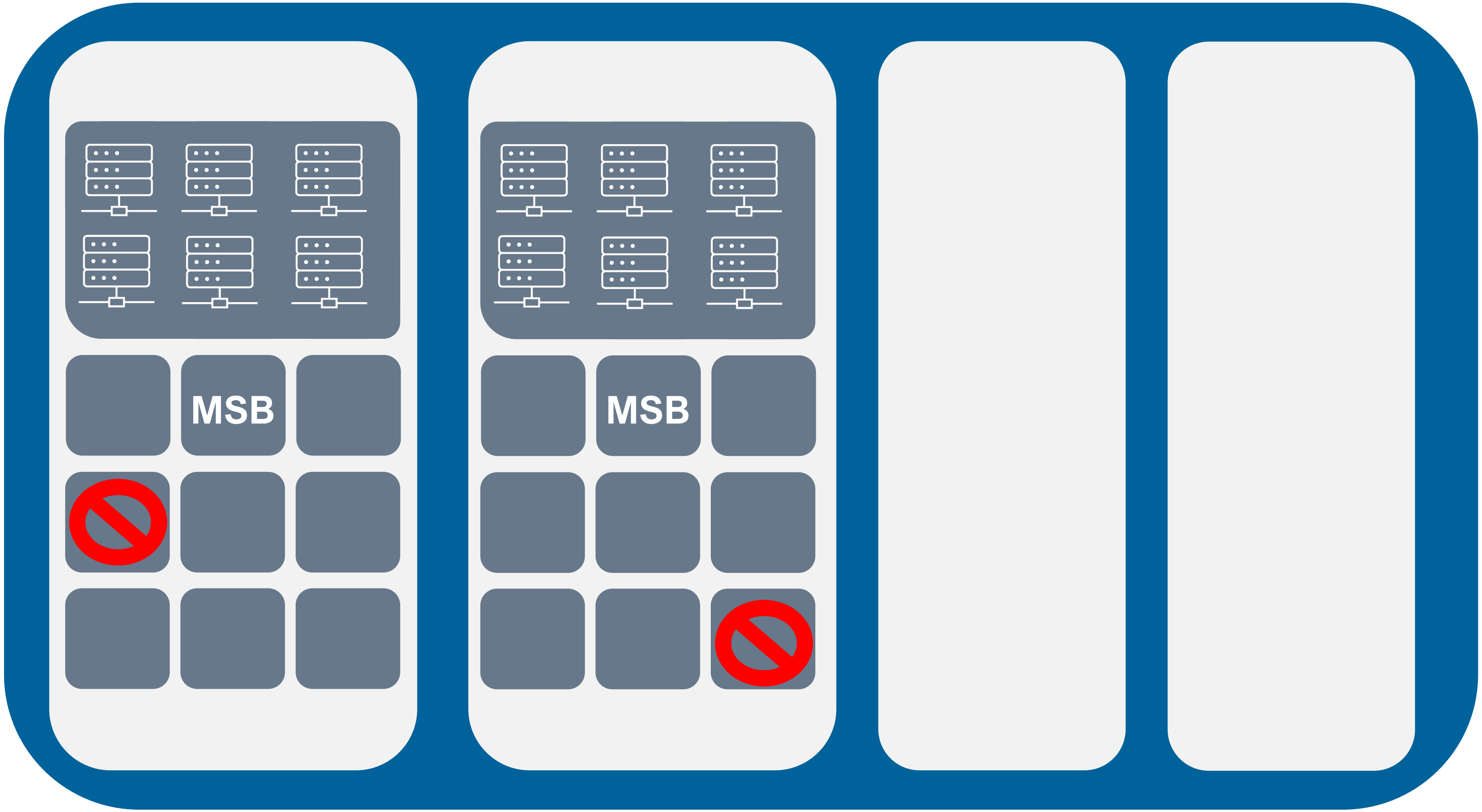


Unplanned Events → Large-Scale Failures



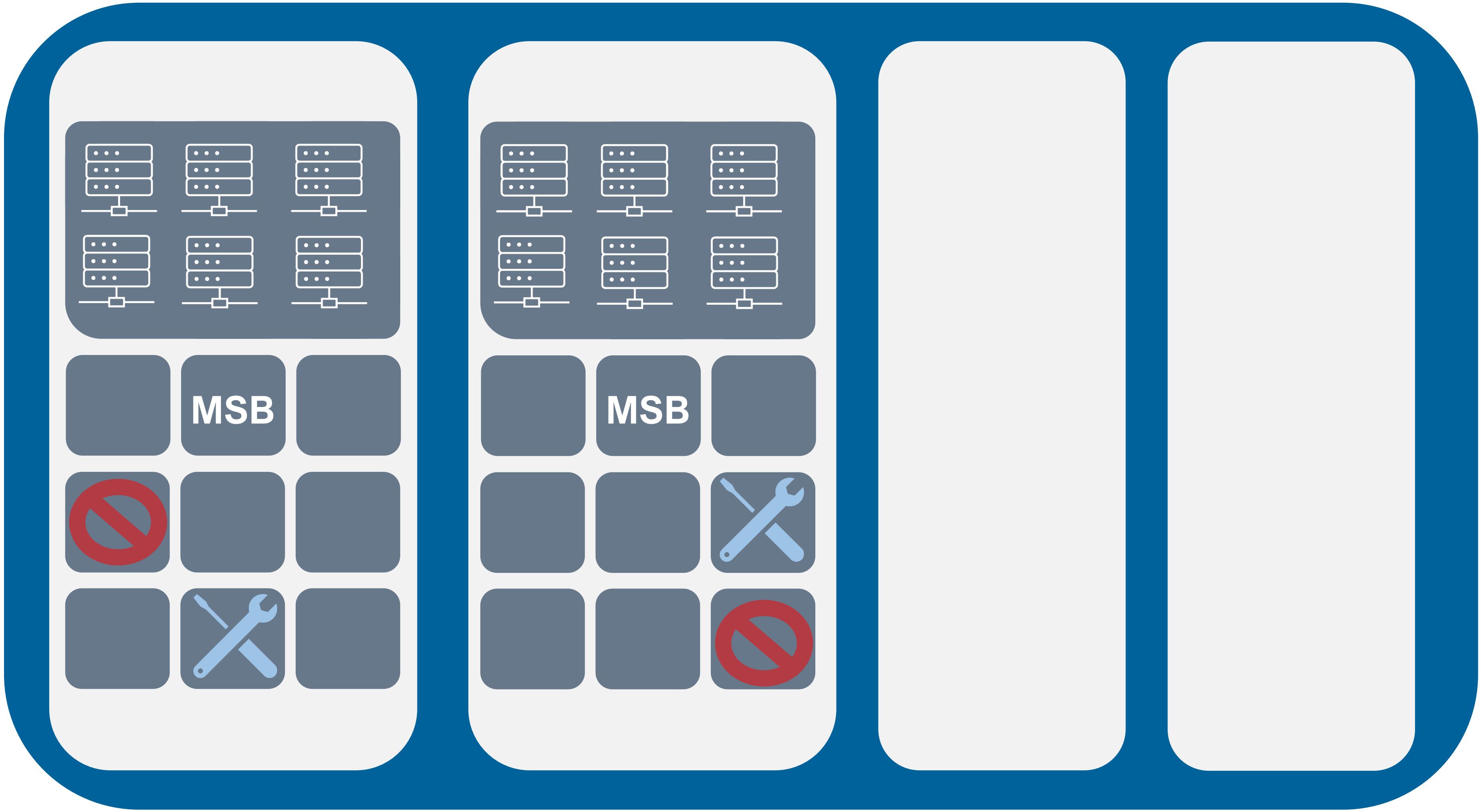
Unplanned Events → Large-Scale Failures

Large-scale Failures



Planned Events → Datacenter Maintenance

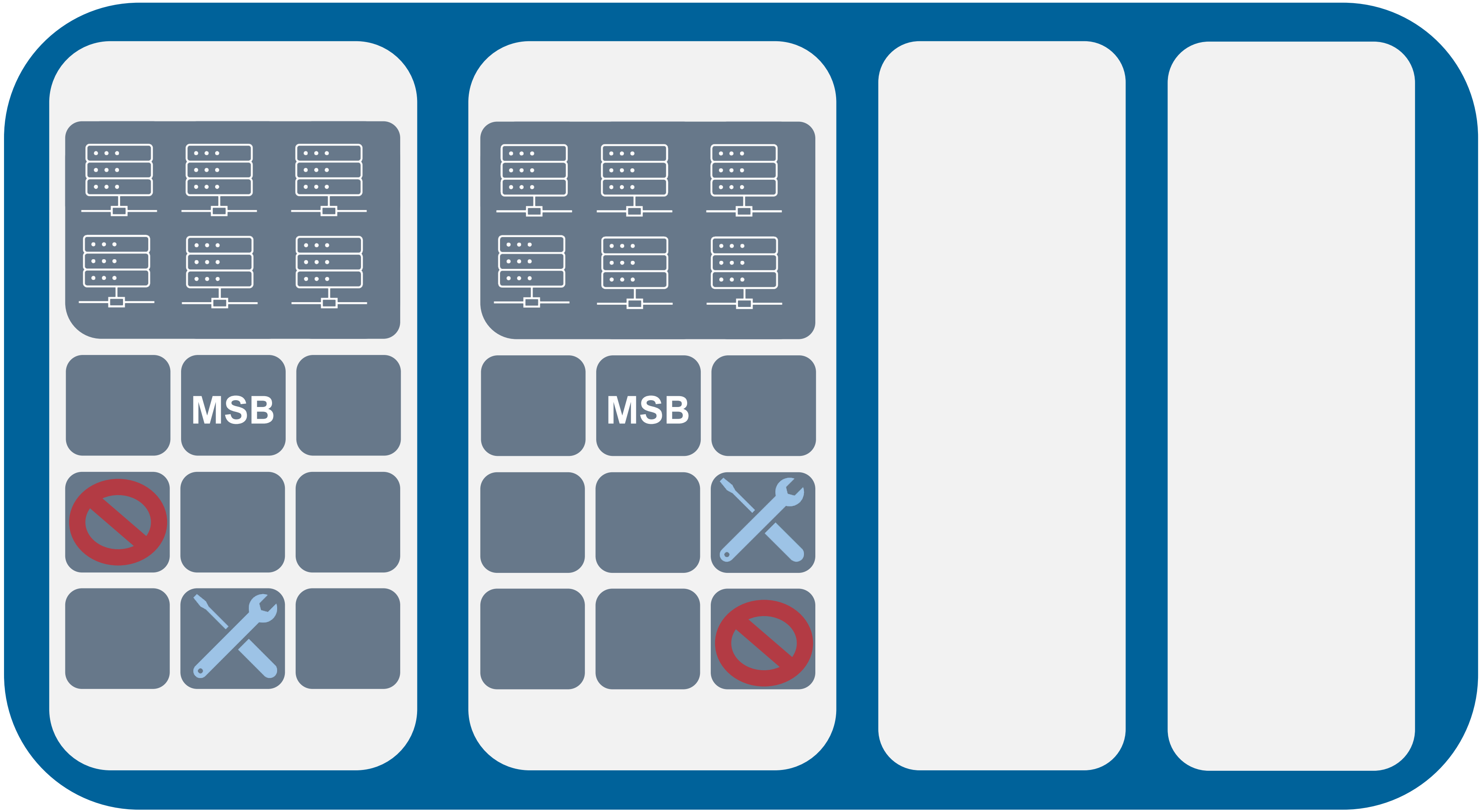
Large-scale Failures



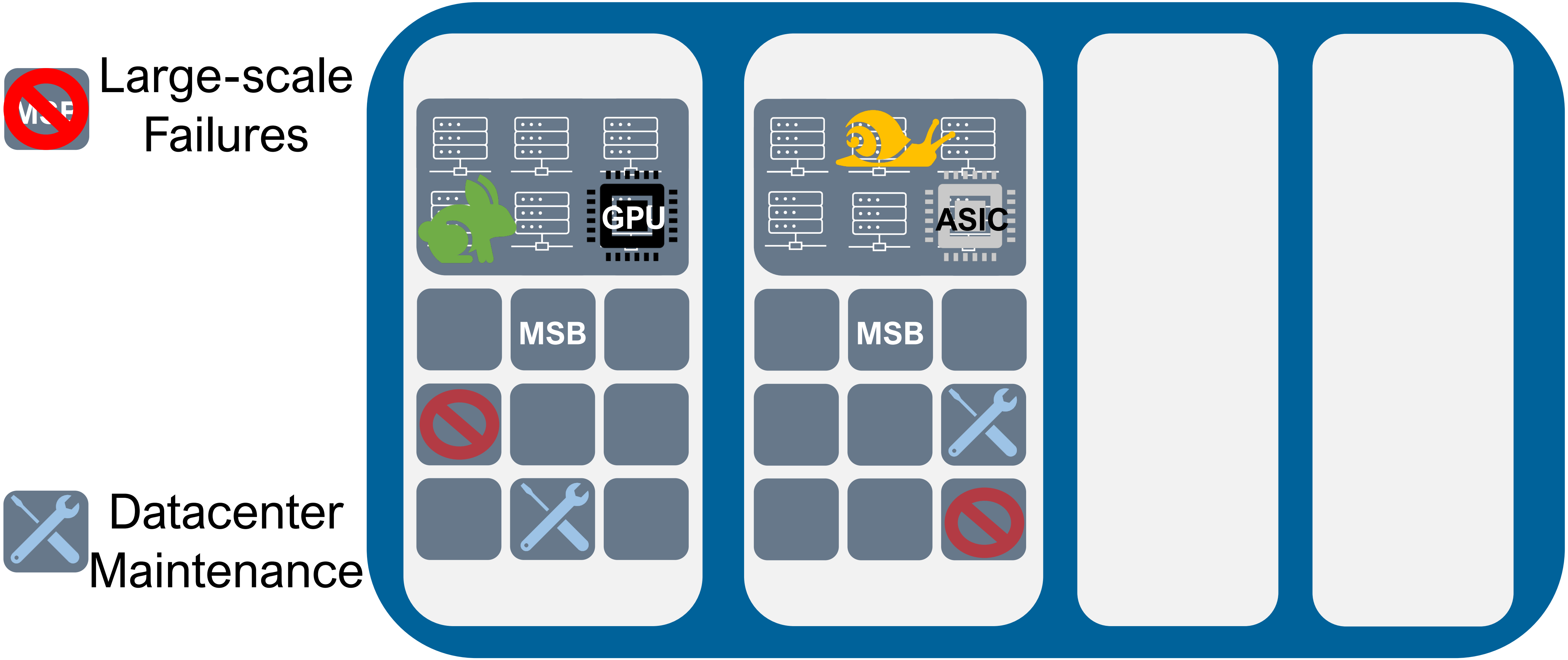
Planned Events → Datacenter Maintenance

Large-scale Failures

Datacenter Maintenance



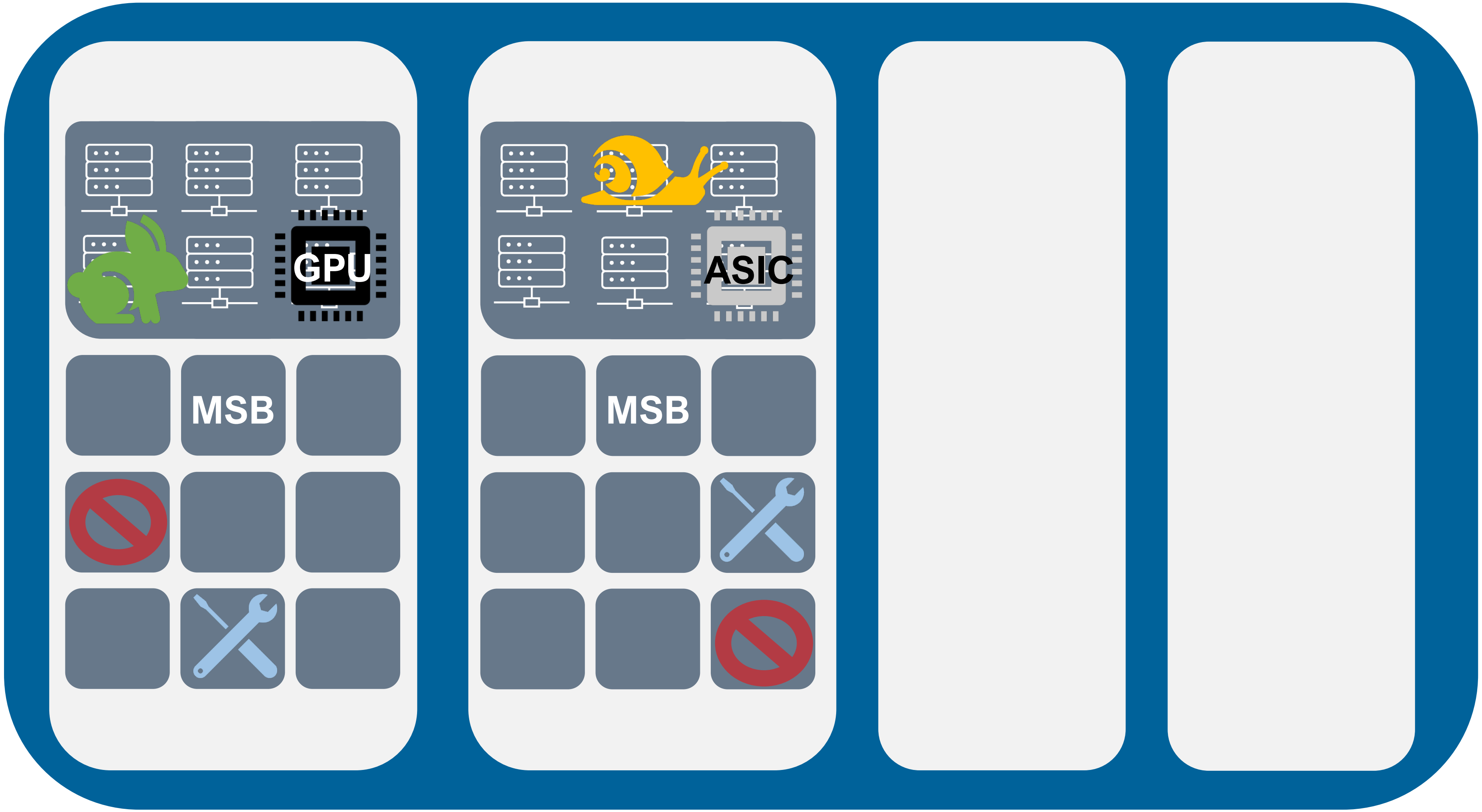
Heterogenous Hardware



Heterogenous Hardware

Large-scale Failures

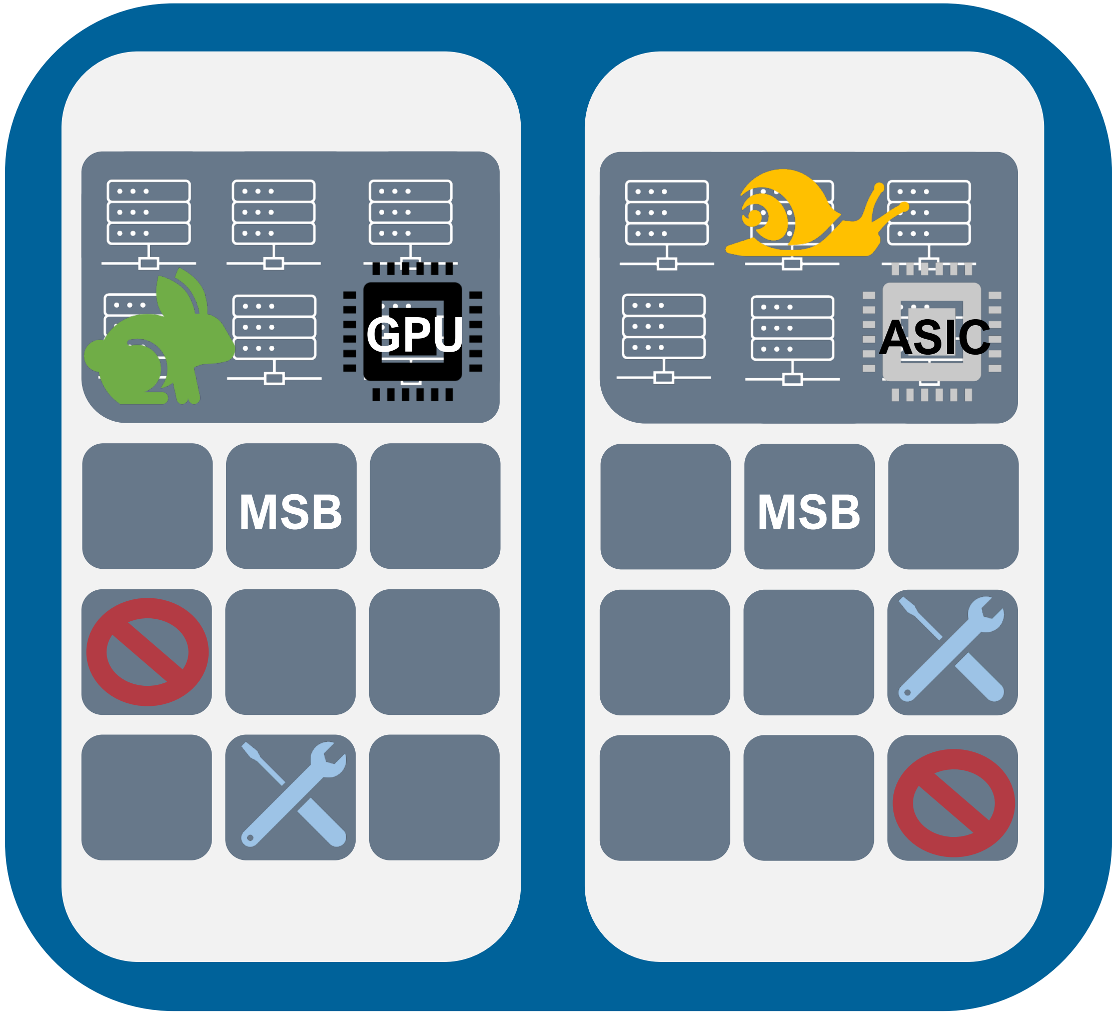
Datacenter Maintenance



Heterogenous Hardware

Large-scale Failures

Datacenter Maintenance

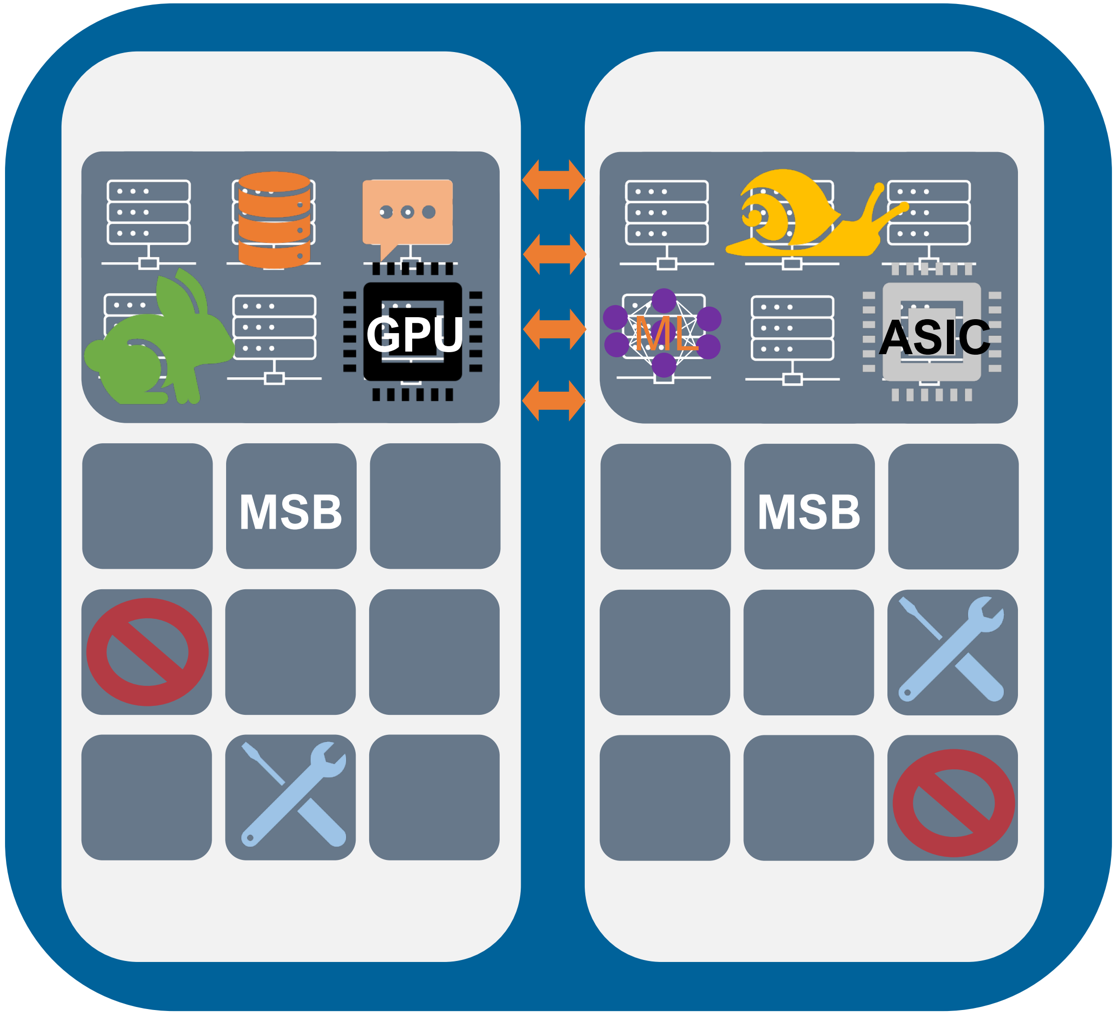


Heterogenous Hardware

Workload Constraints

Large-scale Failures

Datacenter Maintenance

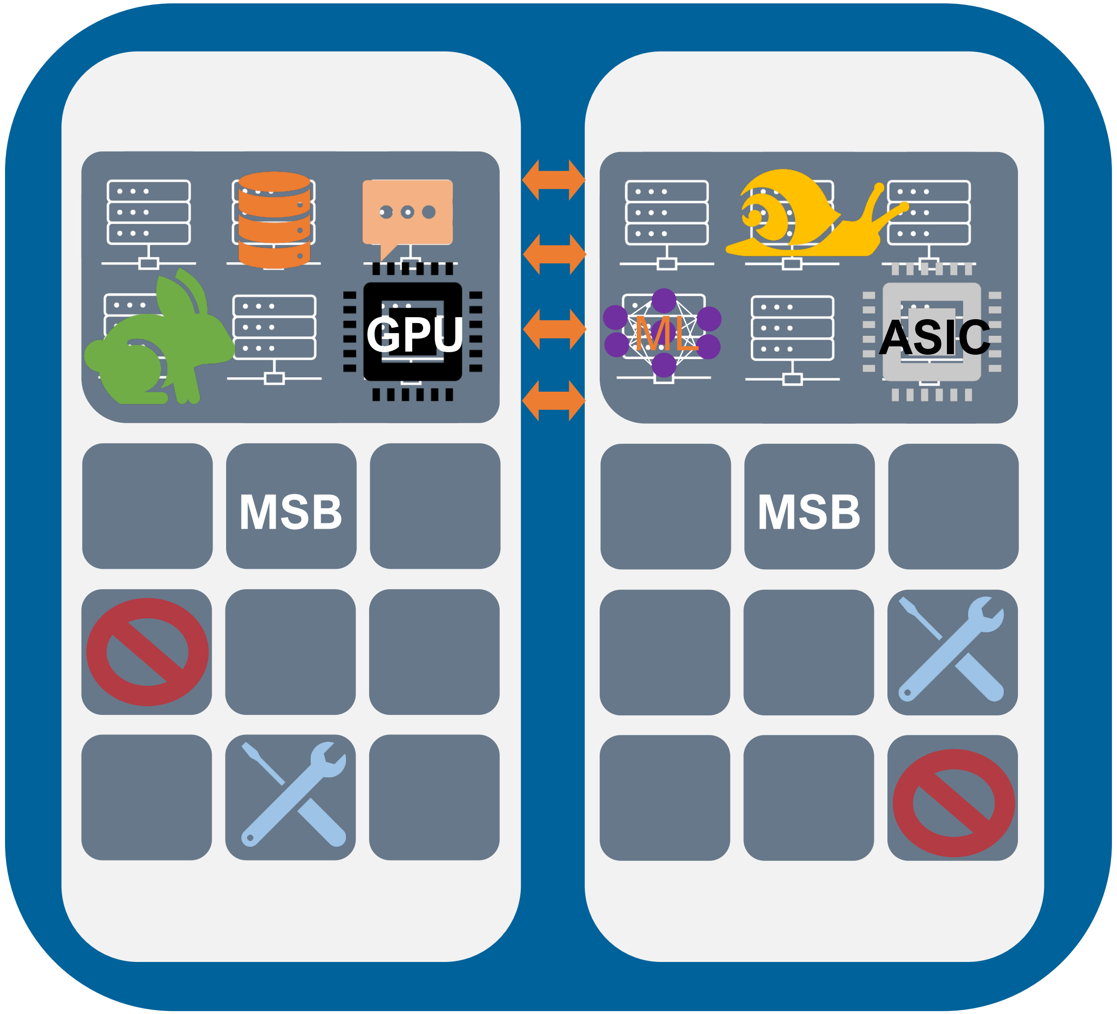


Heterogenous Hardware

Workload Constraints

Large-scale Failures

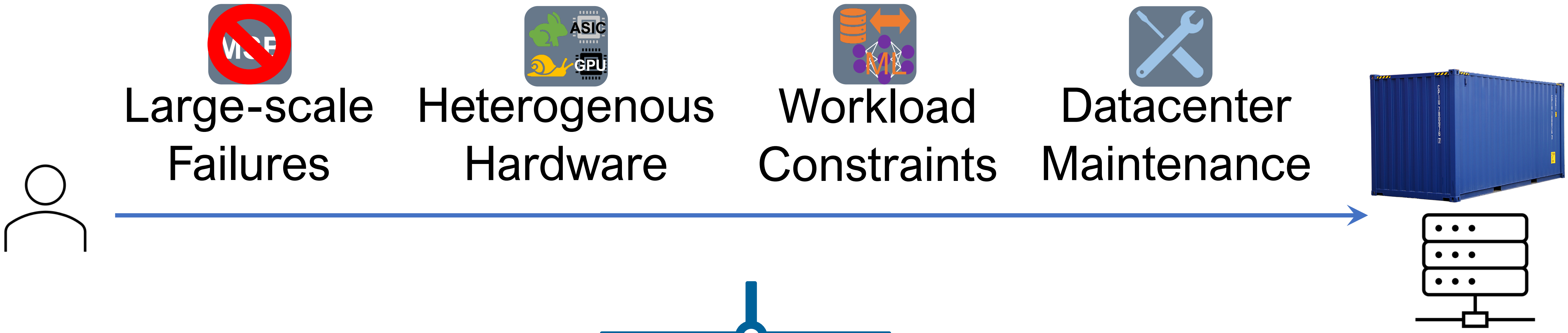
Datacenter Maintenance



Heterogenous Hardware

Workload Constraints

Datacenter Resource Allocation



Large-scale Failures

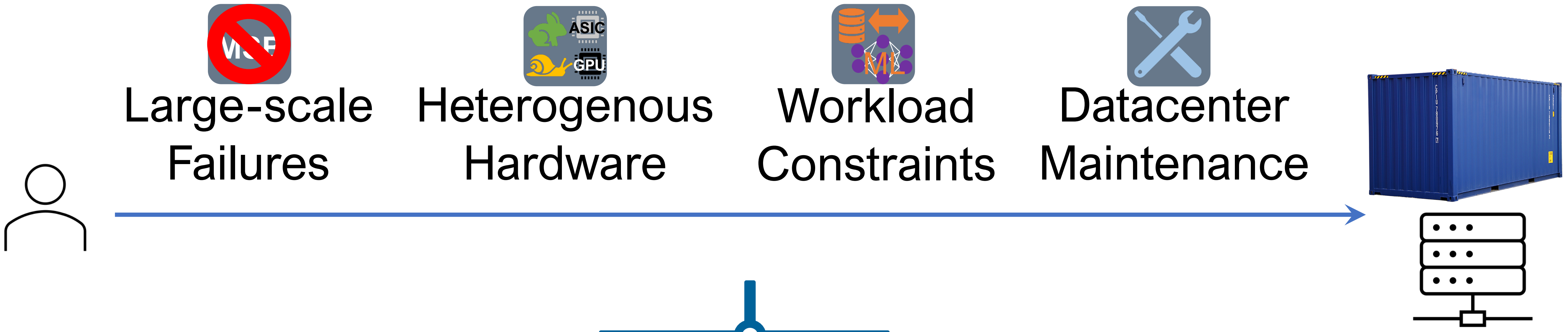
Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

The Resource Allocation Scale

Datacenter Resource Allocation



Large-scale Failures

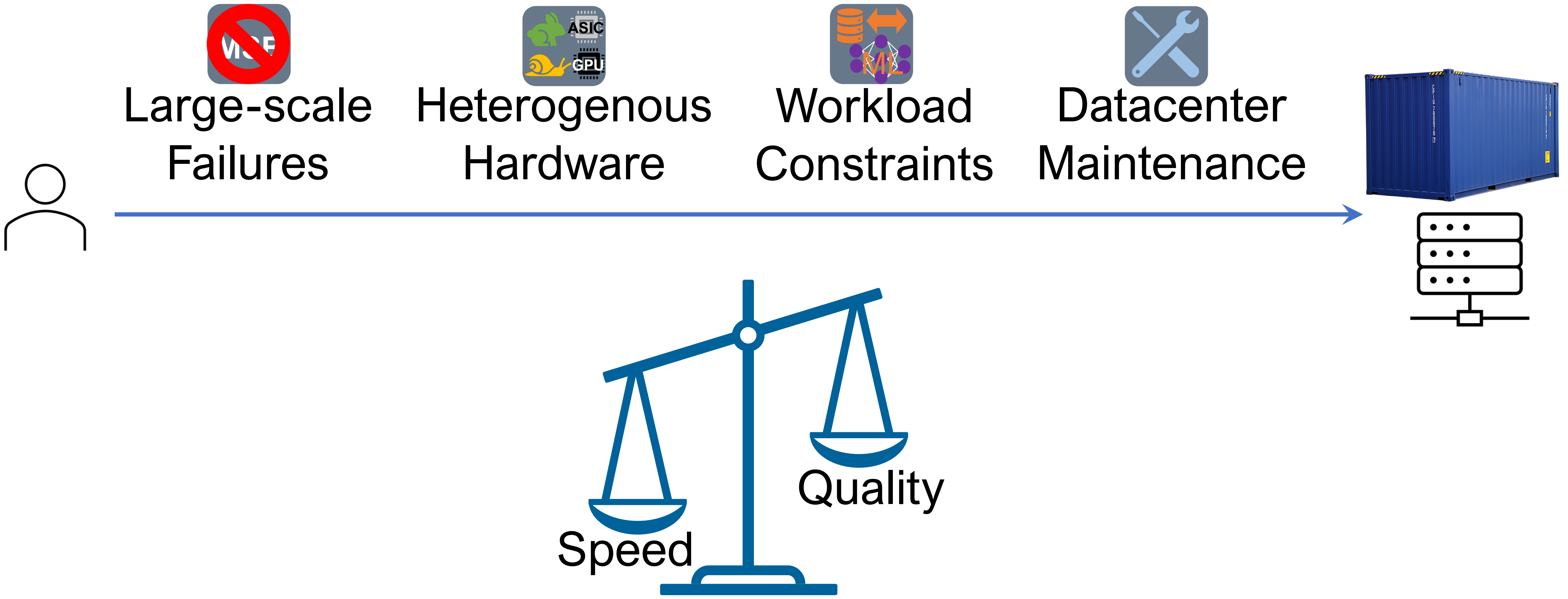
Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

The Resource Allocation Scale

Trade-off Between Speed and Quality



Large-scale Failures

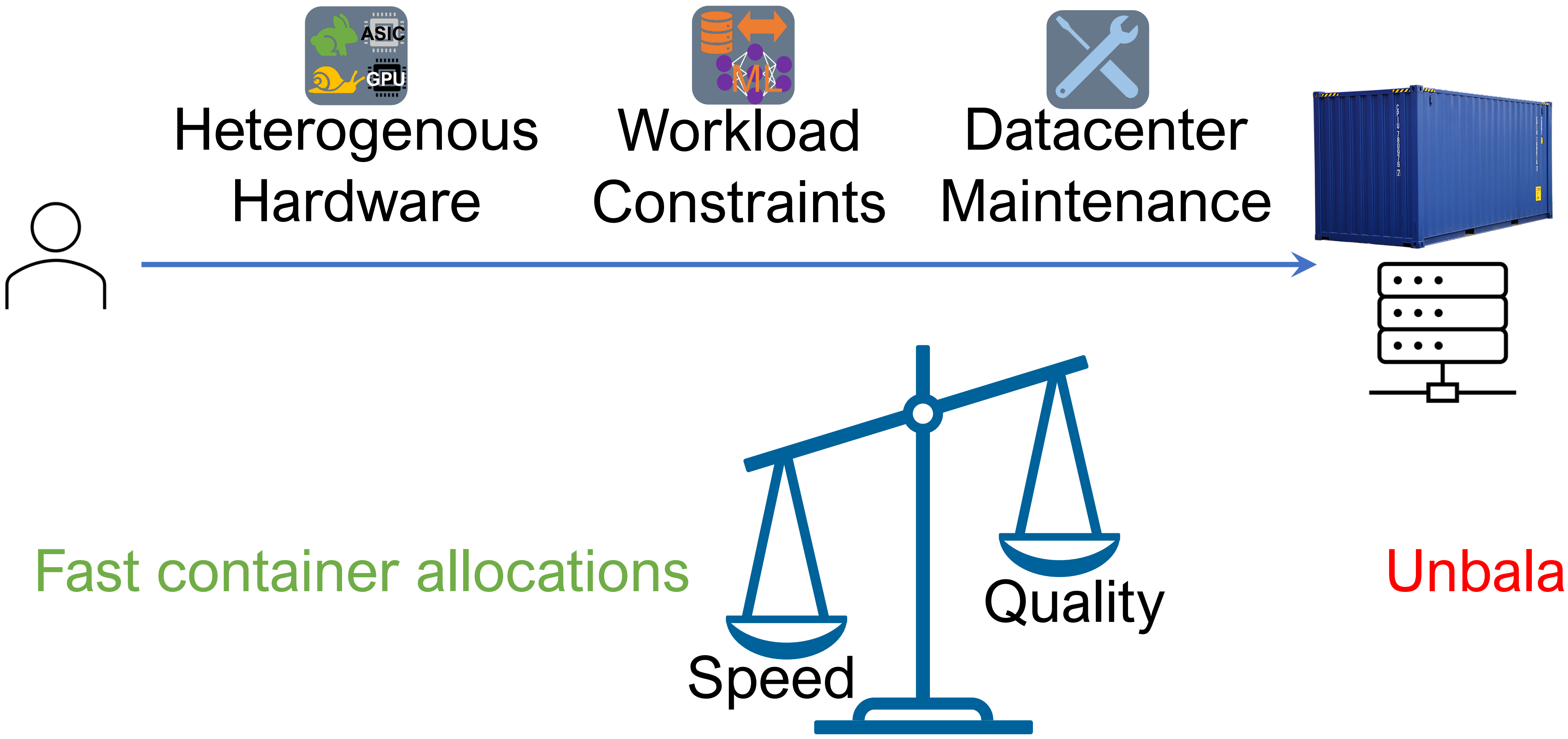
Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

The Resource Allocation Scale

Trade-off Between Speed and Quality



Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

Fast container allocations

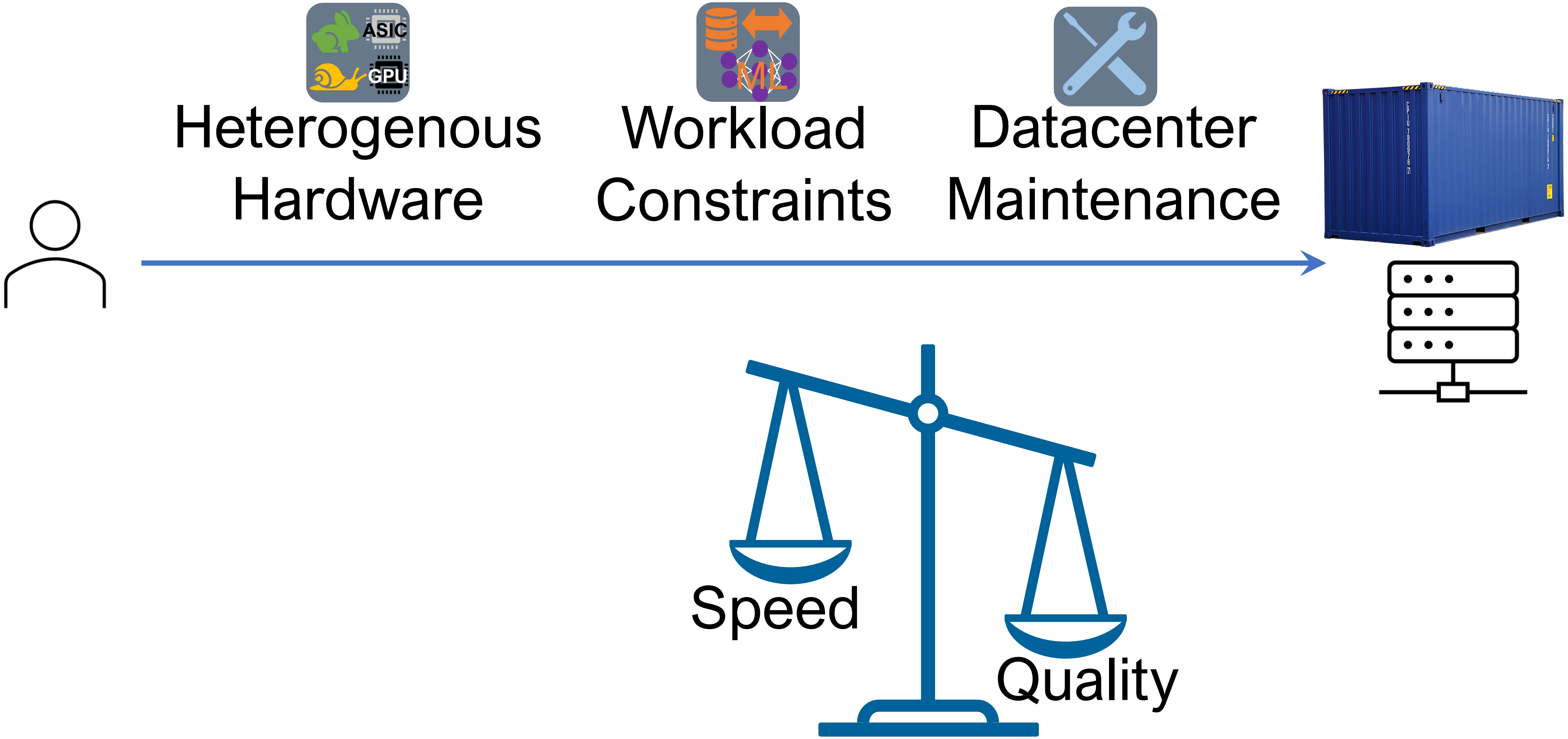
Speed

Quality

Unbalanced spread

The Resource Allocation Scale

Trade-off Between Speed and Quality



Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

Speed

Quality

The Resource Allocation Scale

Trade-off Between Speed and Quality

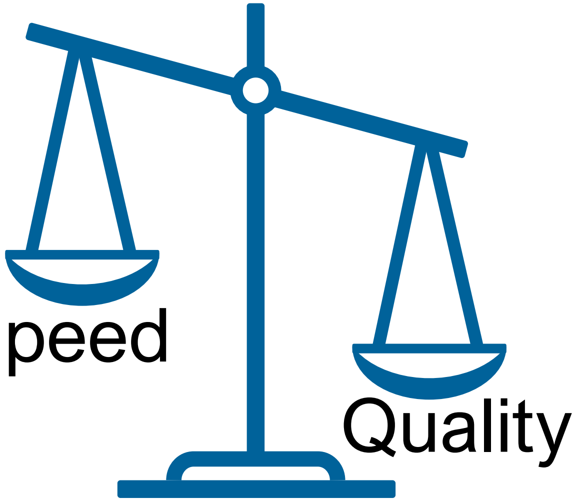
Large-scale Failures

Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

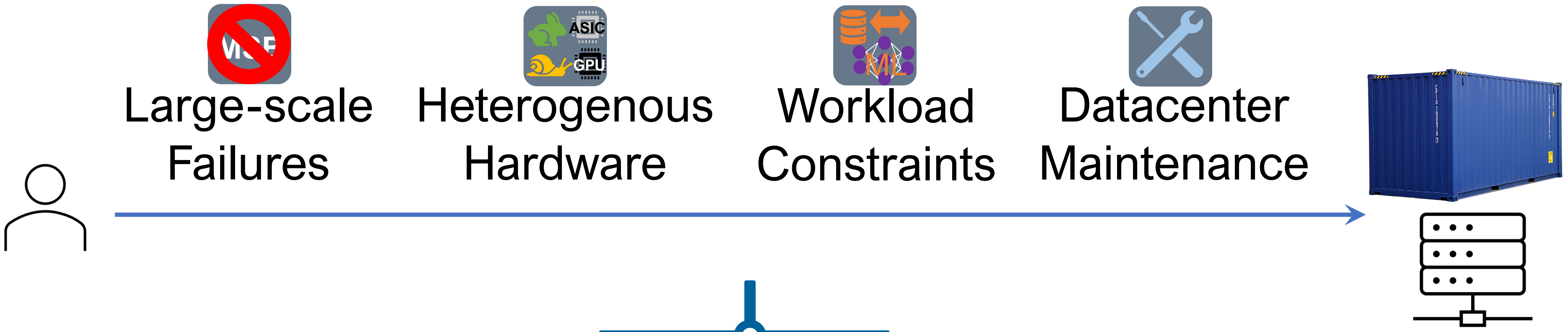
Slow container allocations



Balanced spread

The Resource Allocation Scale

Why not both?



Large-scale Failures

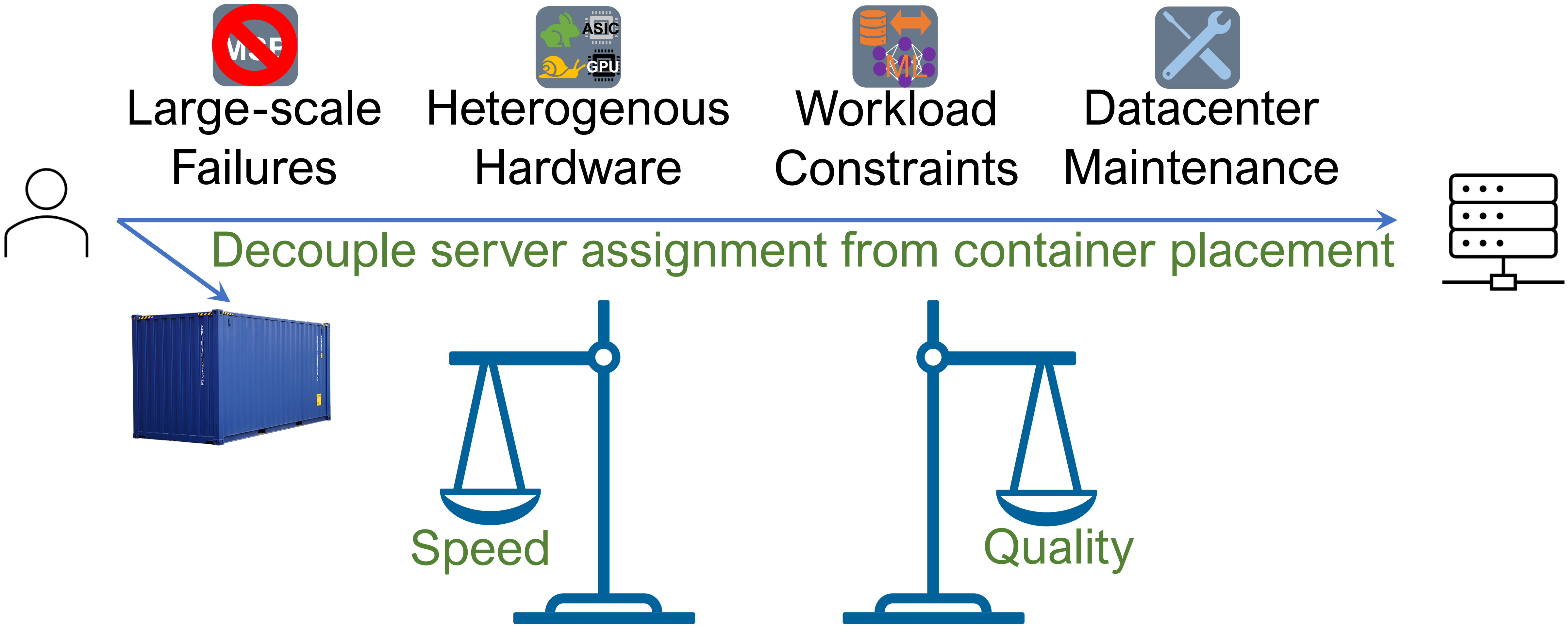
Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

The Resource Allocation Scale

RAS Reservation Abstraction



Large-scale Failures

Heterogenous Hardware

Workload Constraints

Datacenter Maintenance

Decouple server assignment from container placement

Speed

Quality

The Resource Allocation Scale

Our Solution:
Continuously Optimized Resource Allocation

Reservations as a capacity abstraction

Provide guaranteed capacity to services

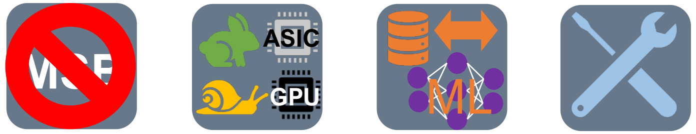
RAS breaks resource allocation into a two-level problem

1. Server-to-reservation assignments
2. Container placement

continuous MIP re-evaluation off the critical-path

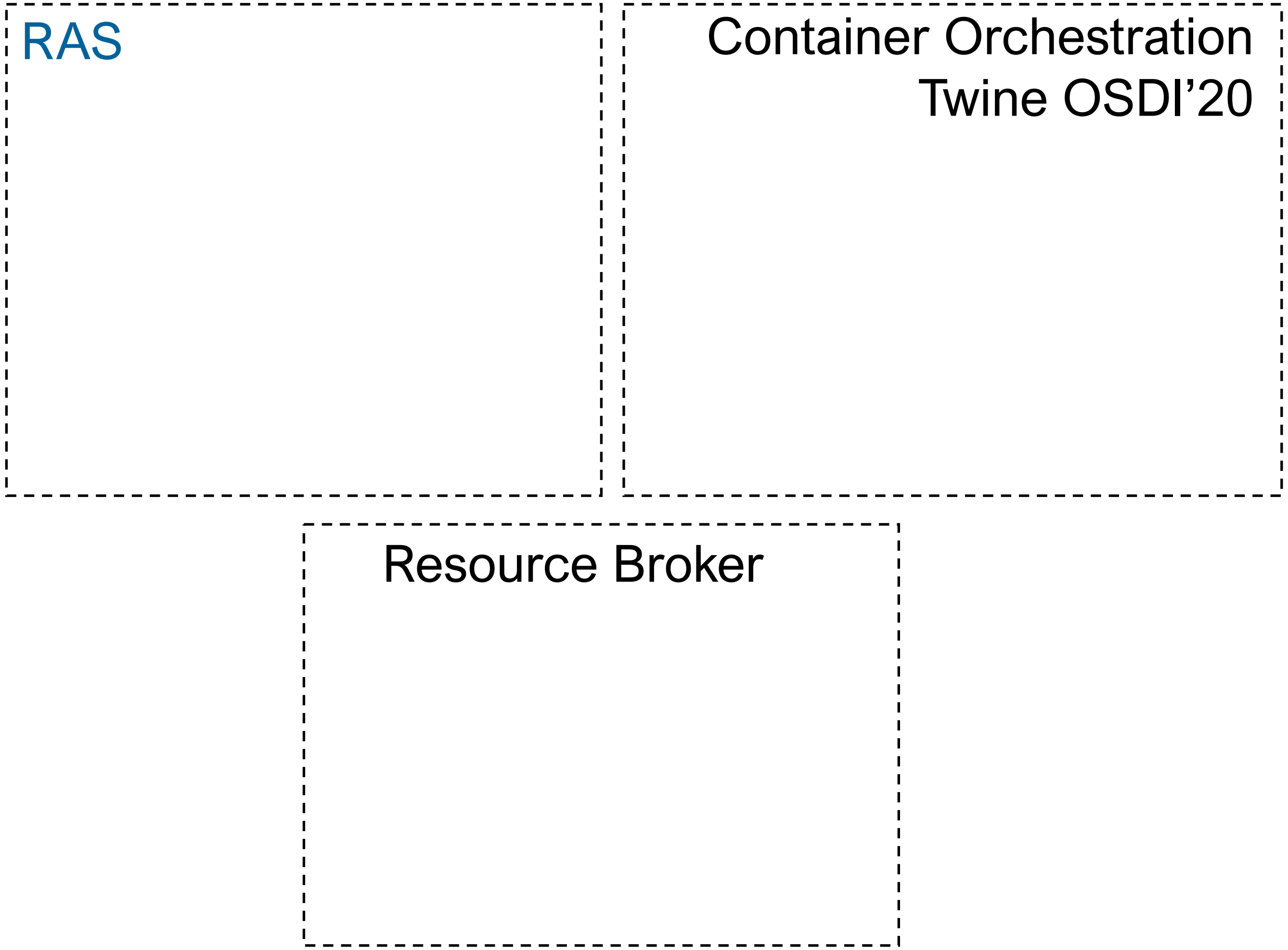
RAS optimizes reservations region-wide:

- Large-scale failures
- Heterogenous hardware
- Workload constraints
- Datacenter maintenance



RAS manages capacity across the entire Facebook fleet!

RAS Operation

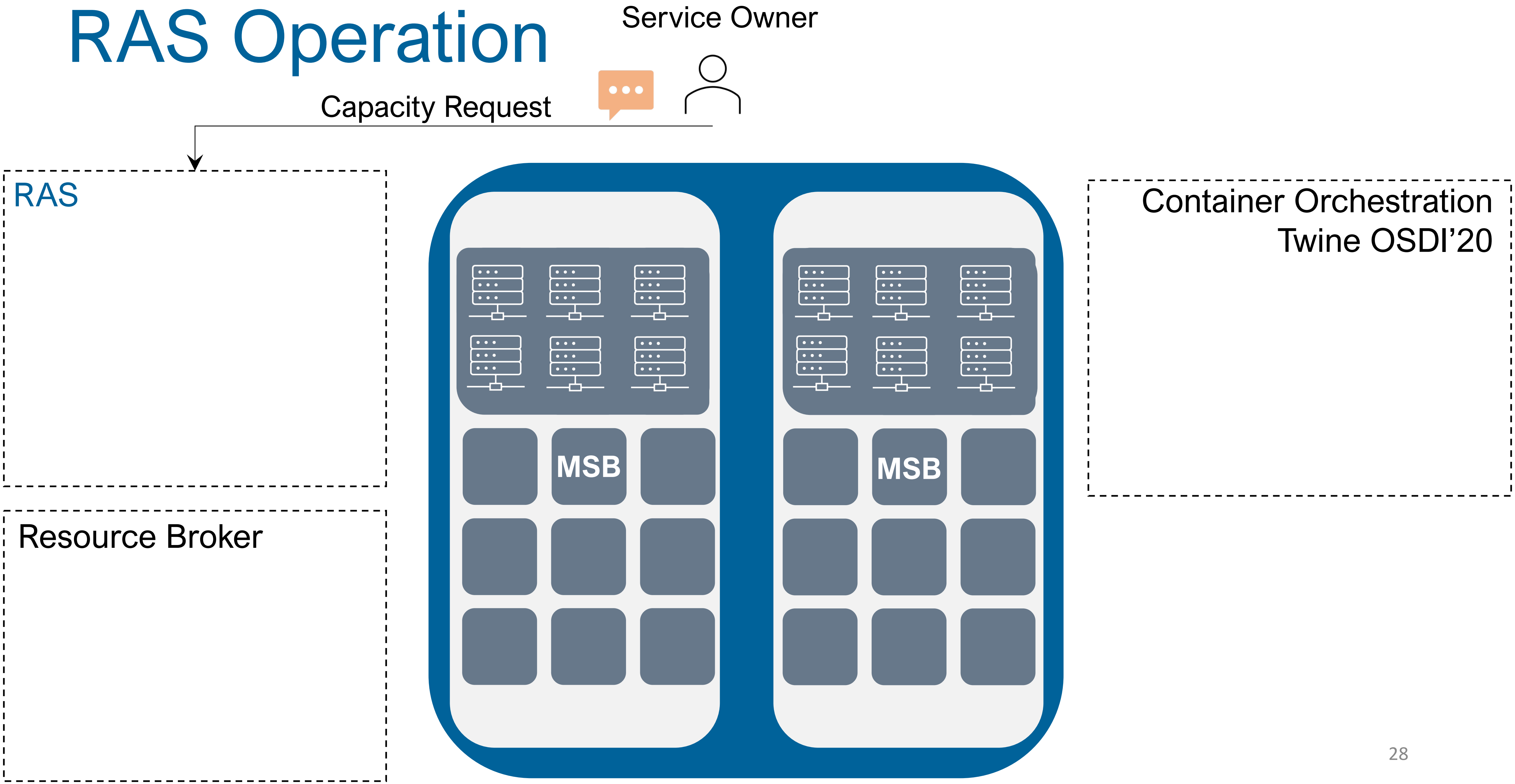


RAS

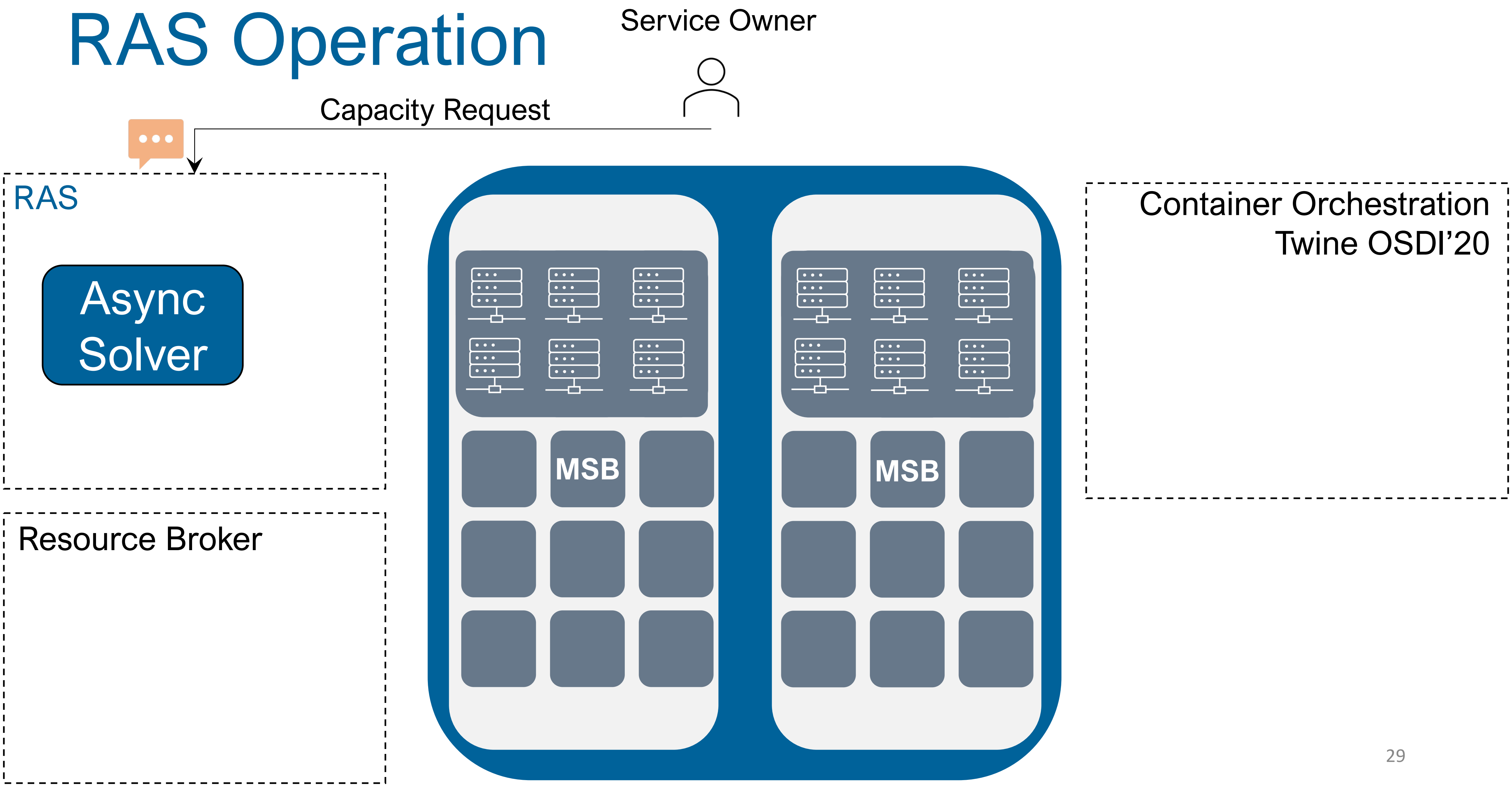
Container Orchestration
Twine OSDI’20

Resource Broker

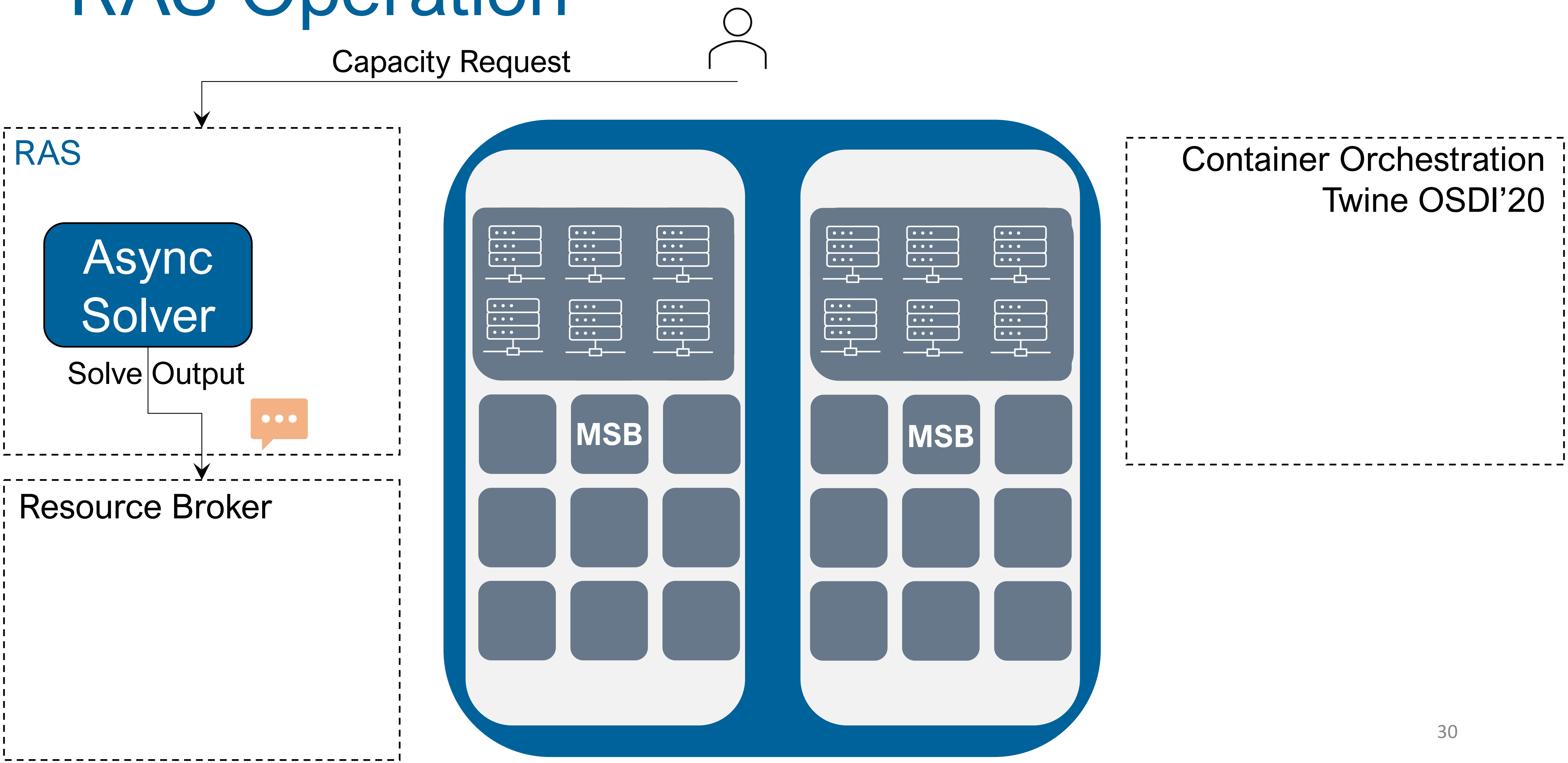
RAS Operation



RAS Operation



RAS Operation



Service Owner

Capacity Request

RAS

Async Solver

Solve Output

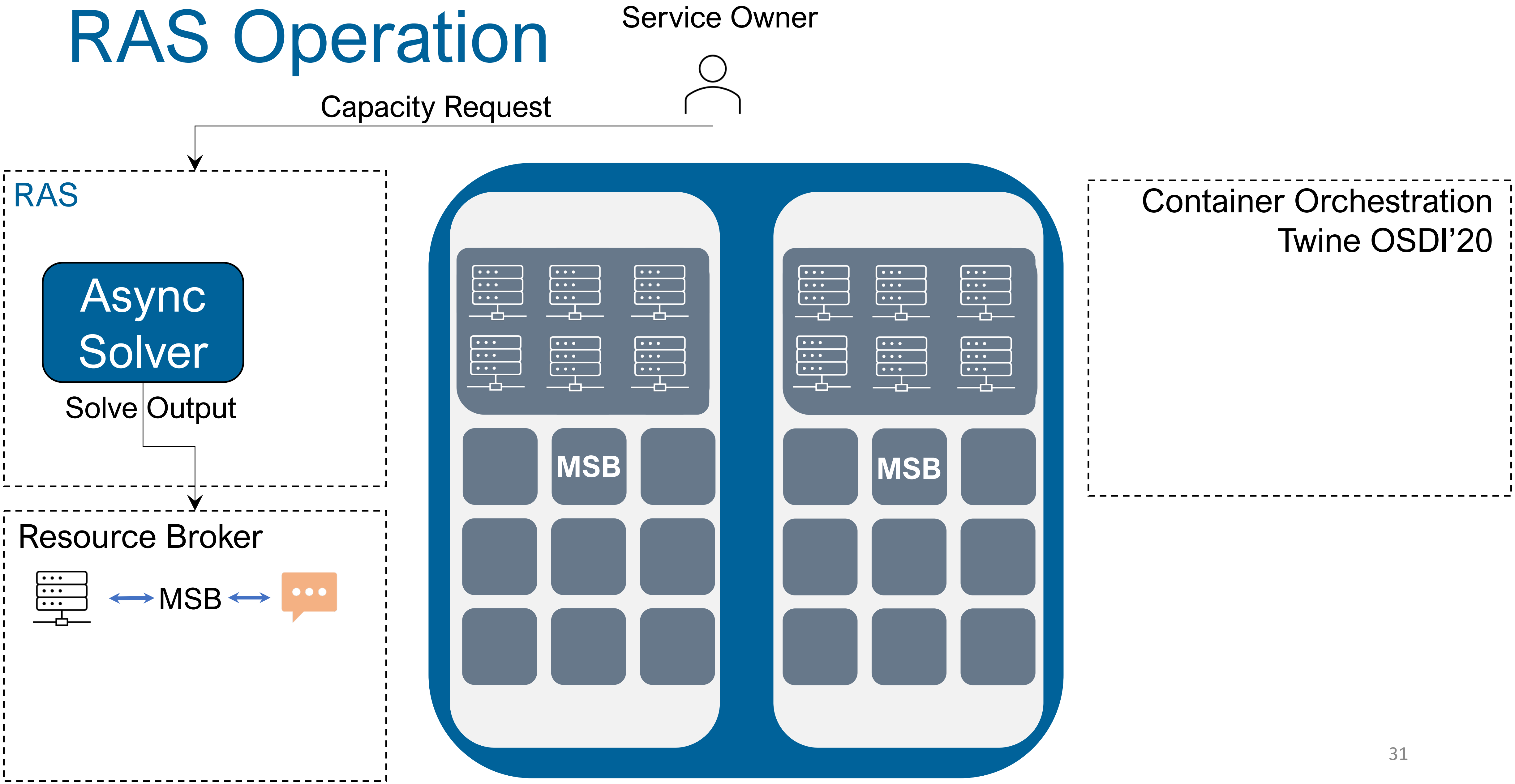
Resource Broker

MSB

MSB

Container Orchestration
Twine OSDI’20

RAS Operation



Service Owner

Capacity Request

RAS

Async Solver

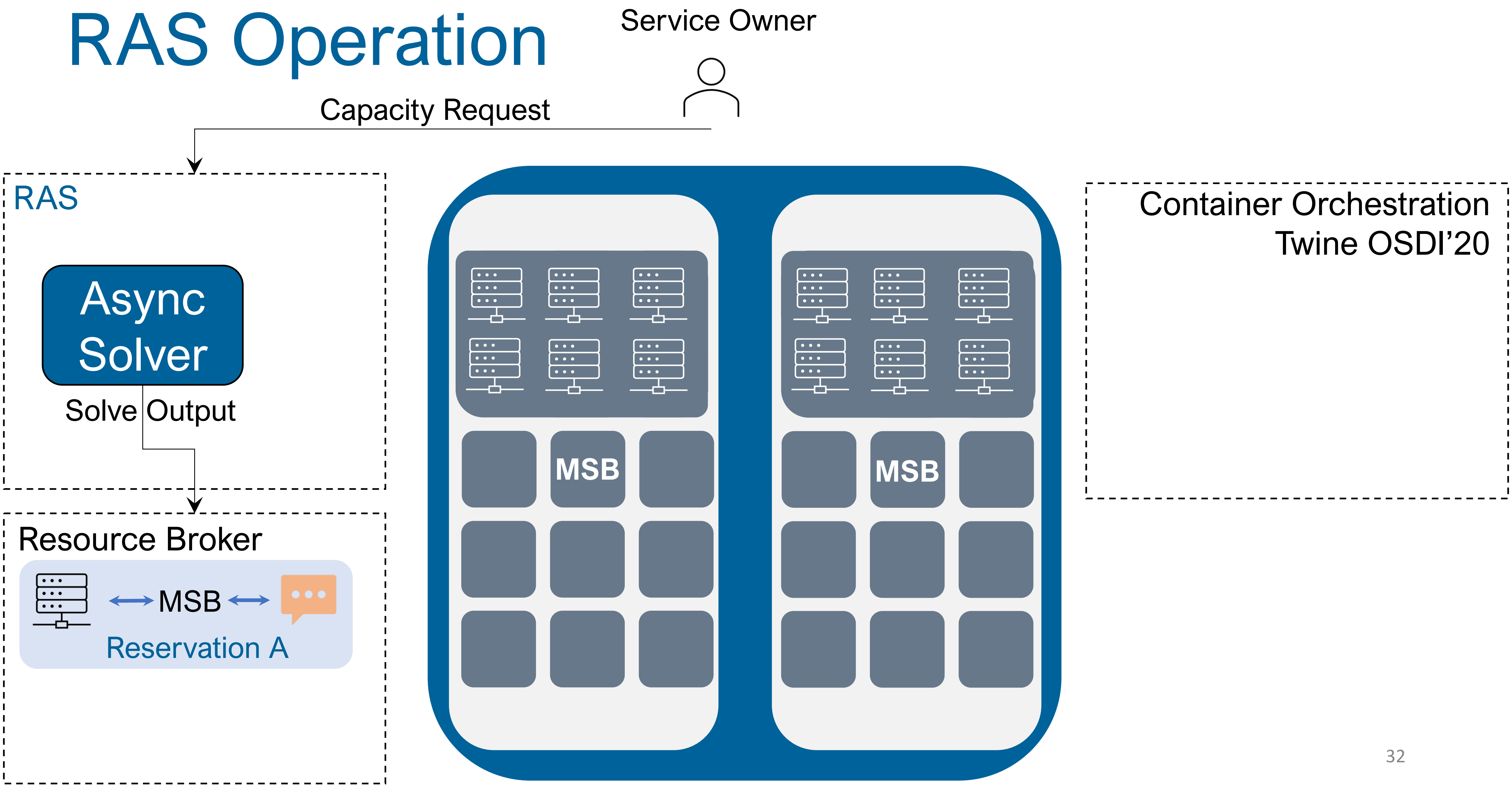
Solve Output

Resource Broker

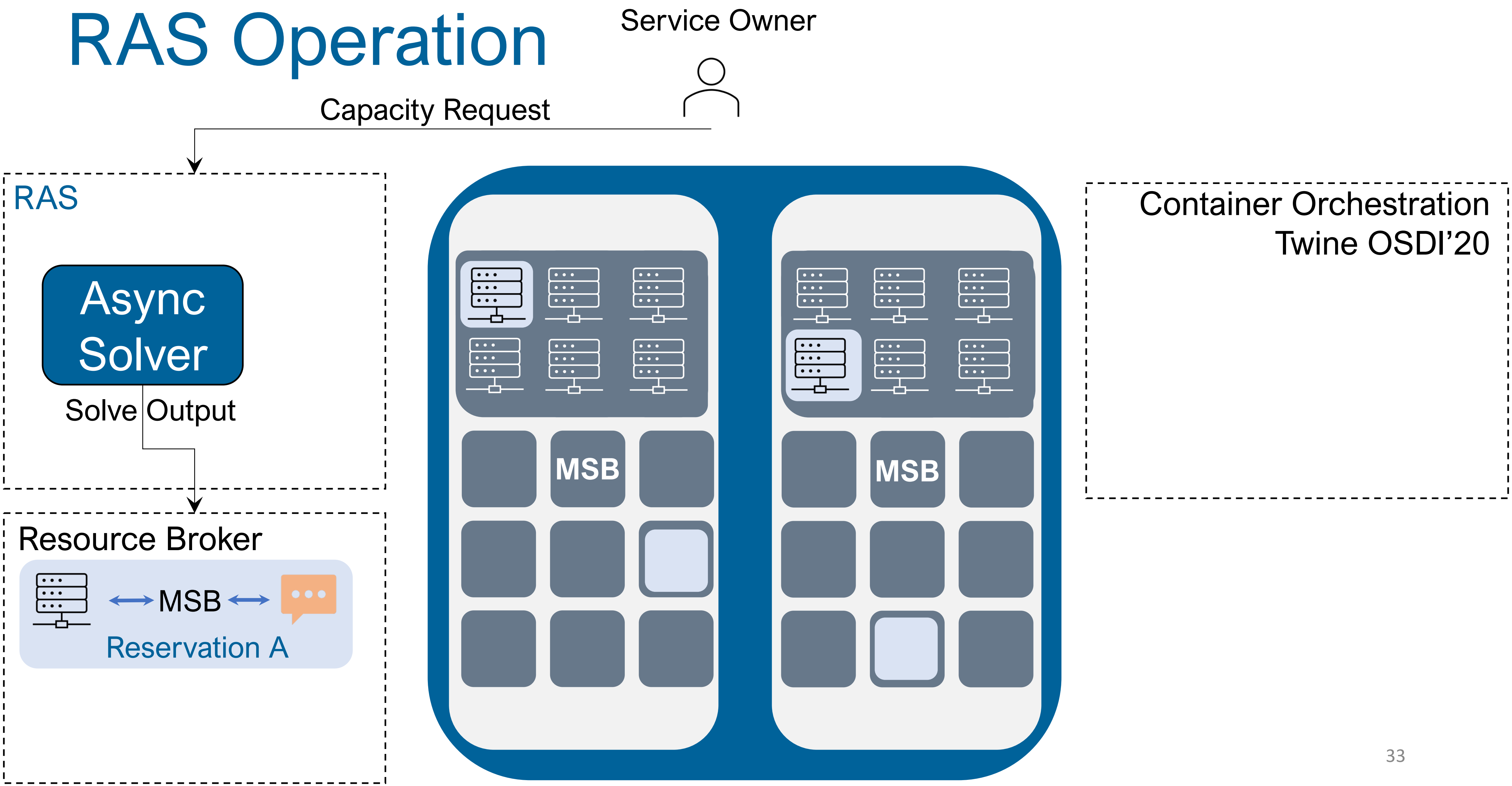
↔ MSB ↔

Container Orchestration
Twine OSDI’20

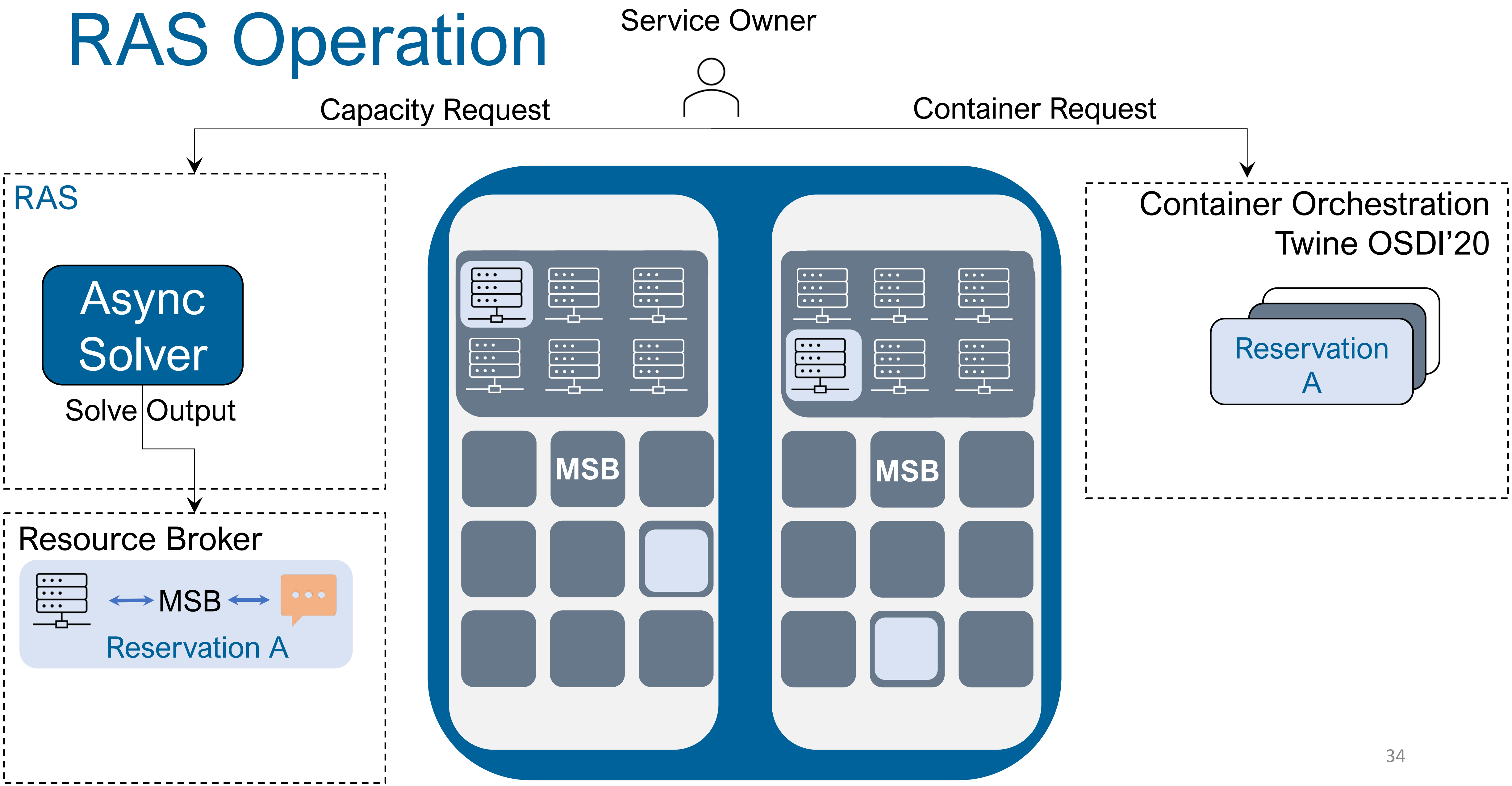
RAS Operation



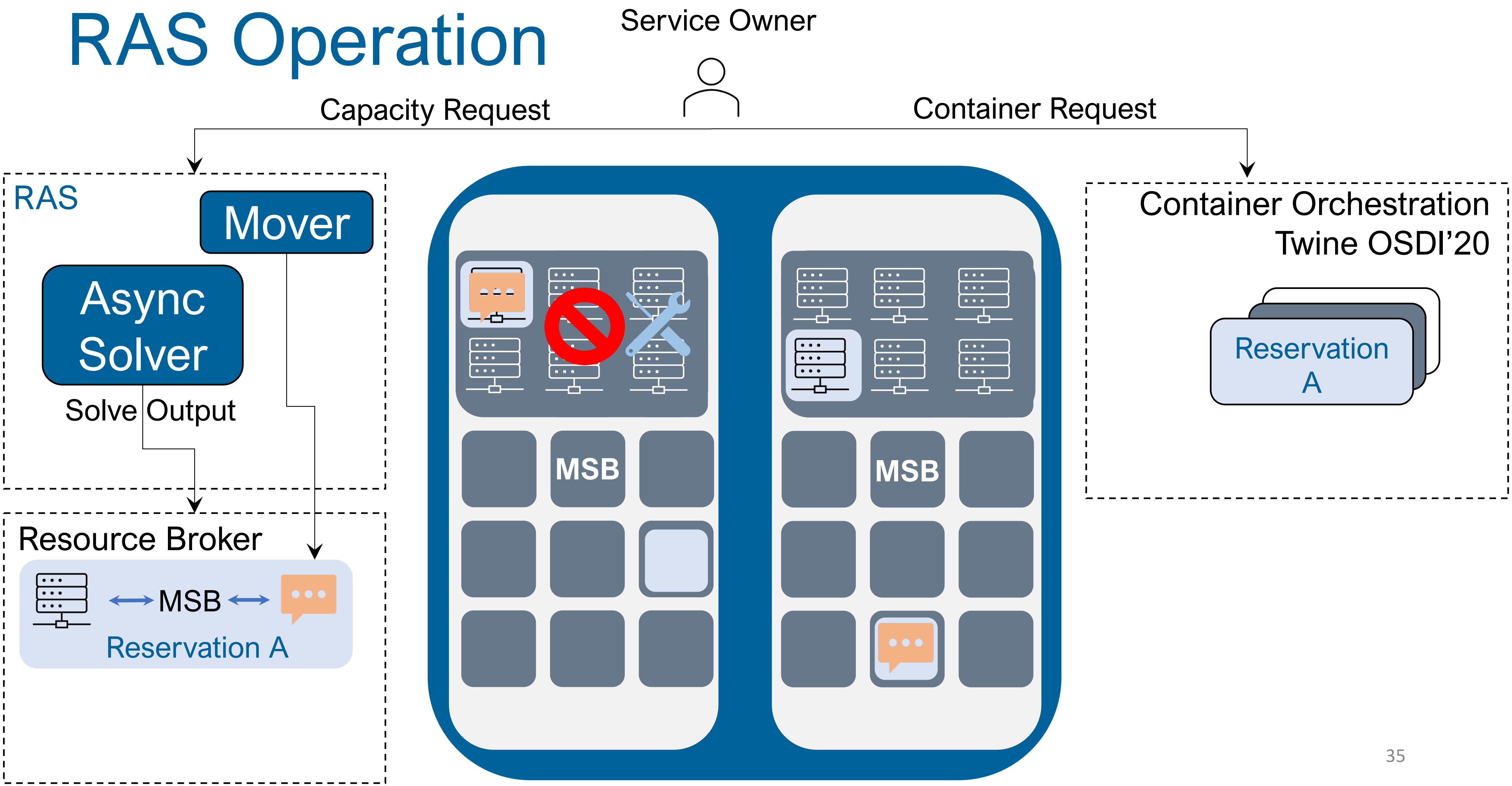
RAS Operation



RAS Operation



RAS Operation



Service Owner

Capacity Request

Container Request

RAS

Mover

Async Solver

Solve Output

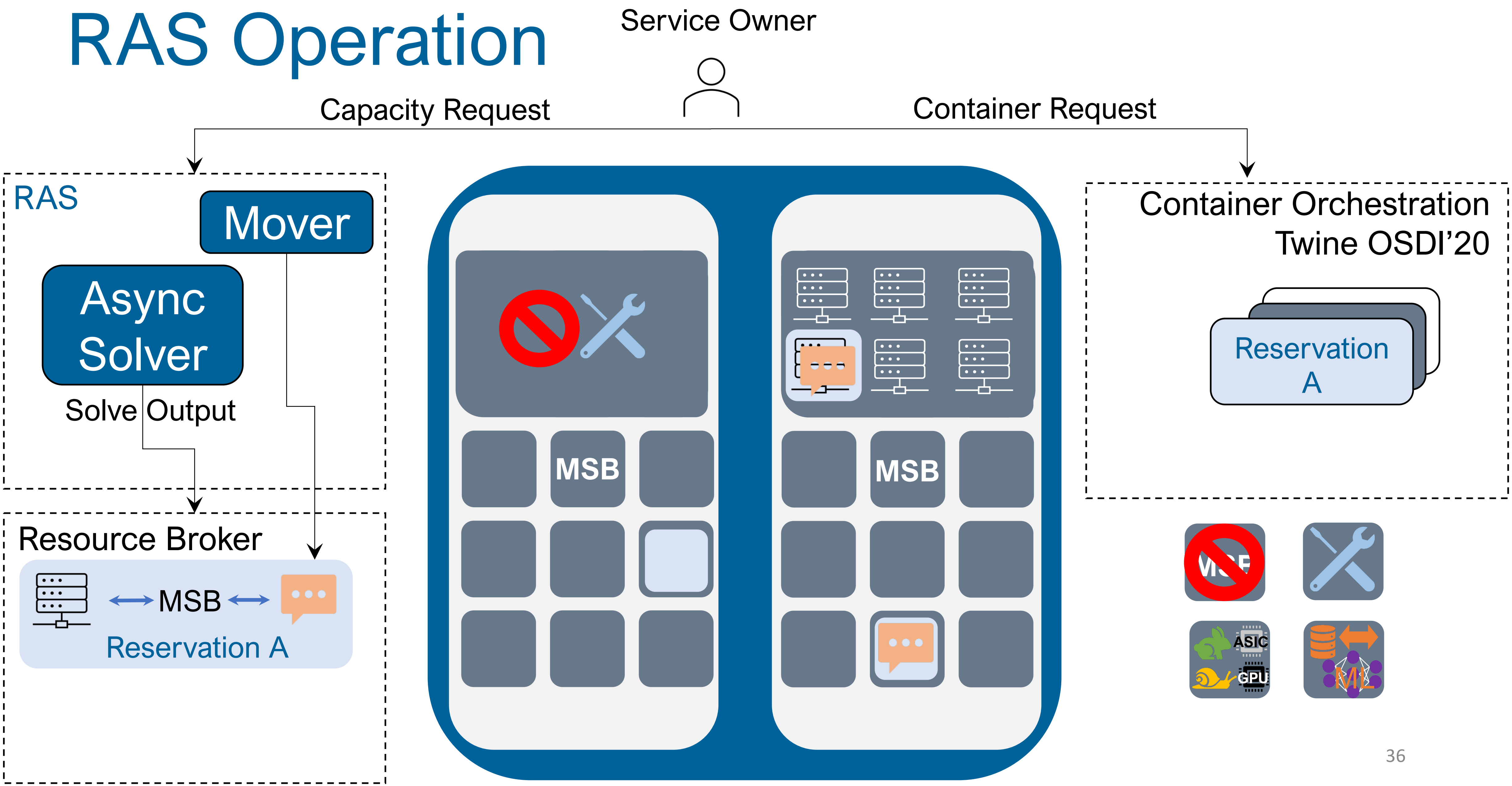
Resource Broker

MSB

Reservation A

Container Orchestration
Twine OSDI'20

RAS Operation



Service Owner

Capacity Request

Container Request

RAS

Mover

Async Solver

Solve Output

Resource Broker

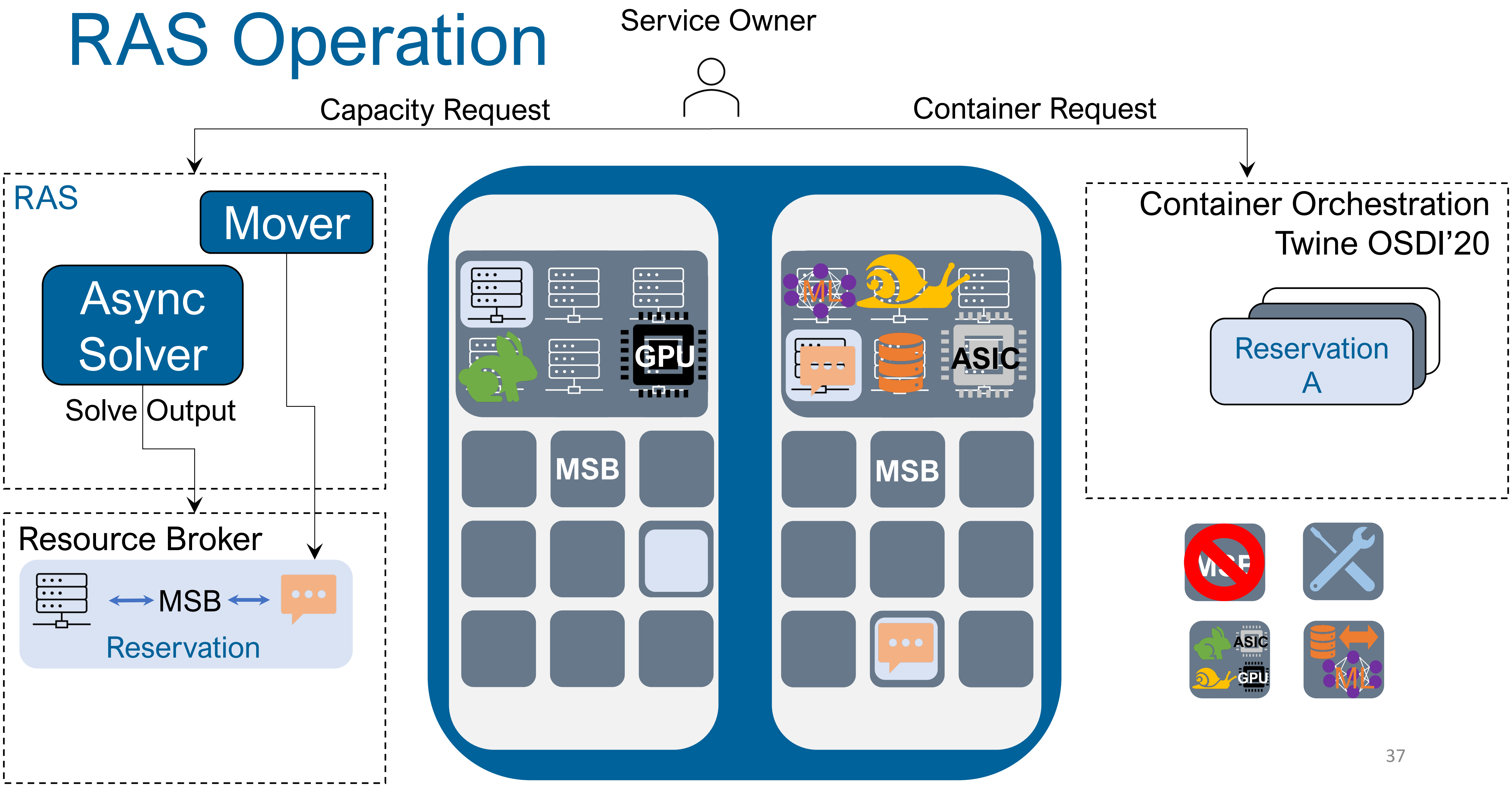
MSB

Reservation A

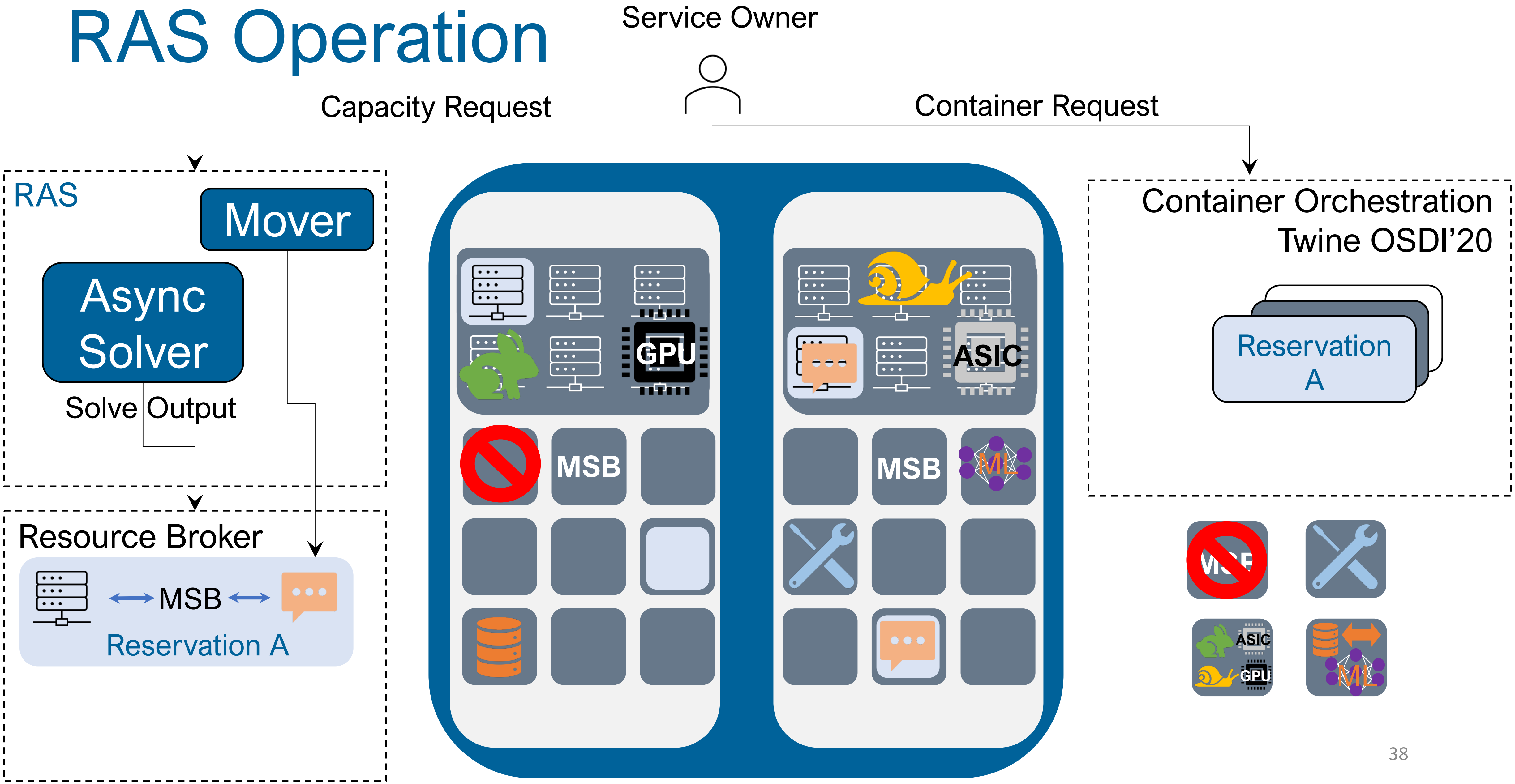
Container Orchestration
Twine OSDI’20

Reservation A

RAS Operation



RAS Operation



Service Owner

Capacity Request

Container Request

RAS

Mover

Async Solver

Solve Output

Resource Broker

MSB

Reservation A

Container Orchestration
Twine OSDI'20

Reservation A

GPU

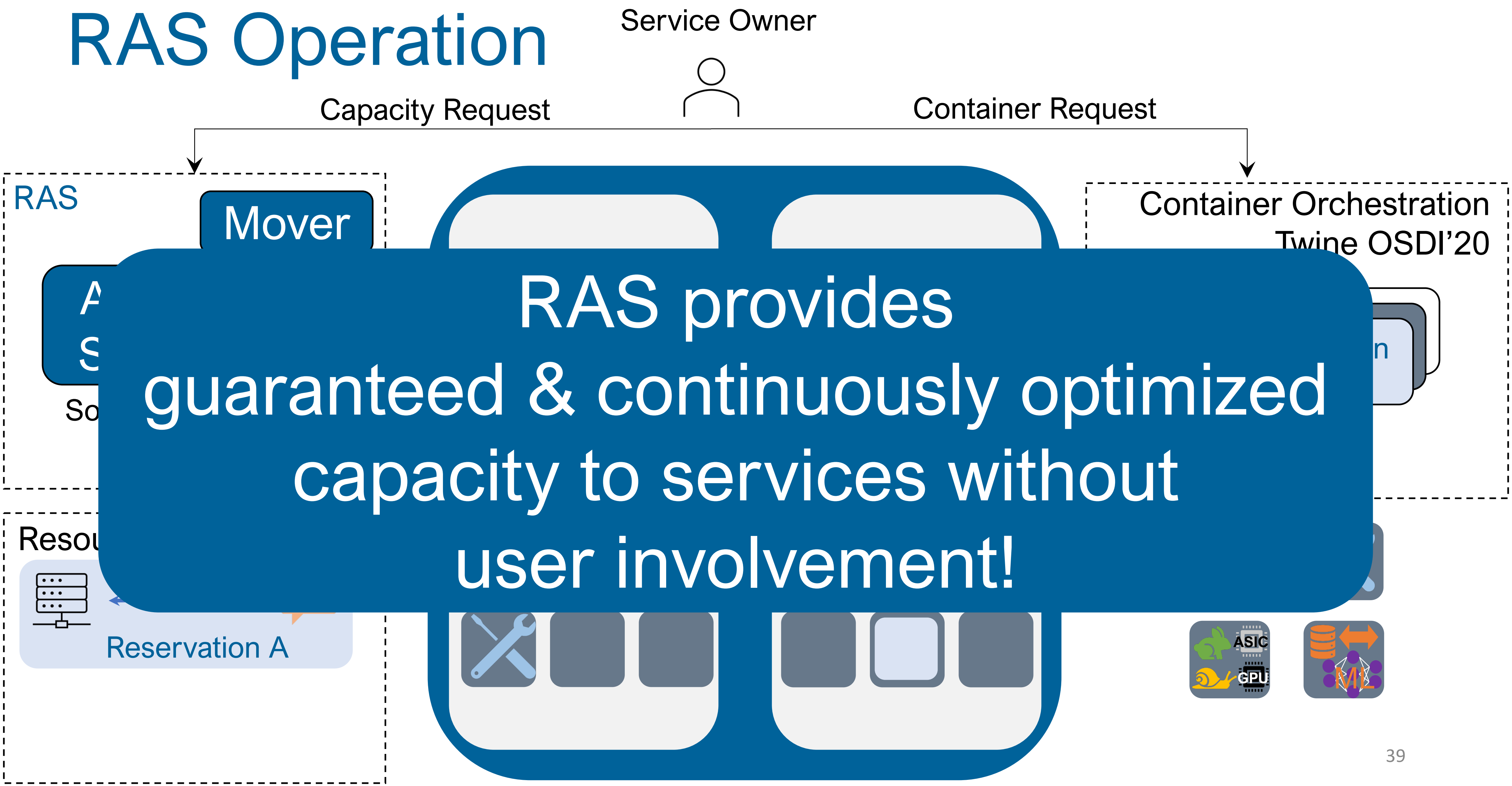
ASIC

MSB

MSB

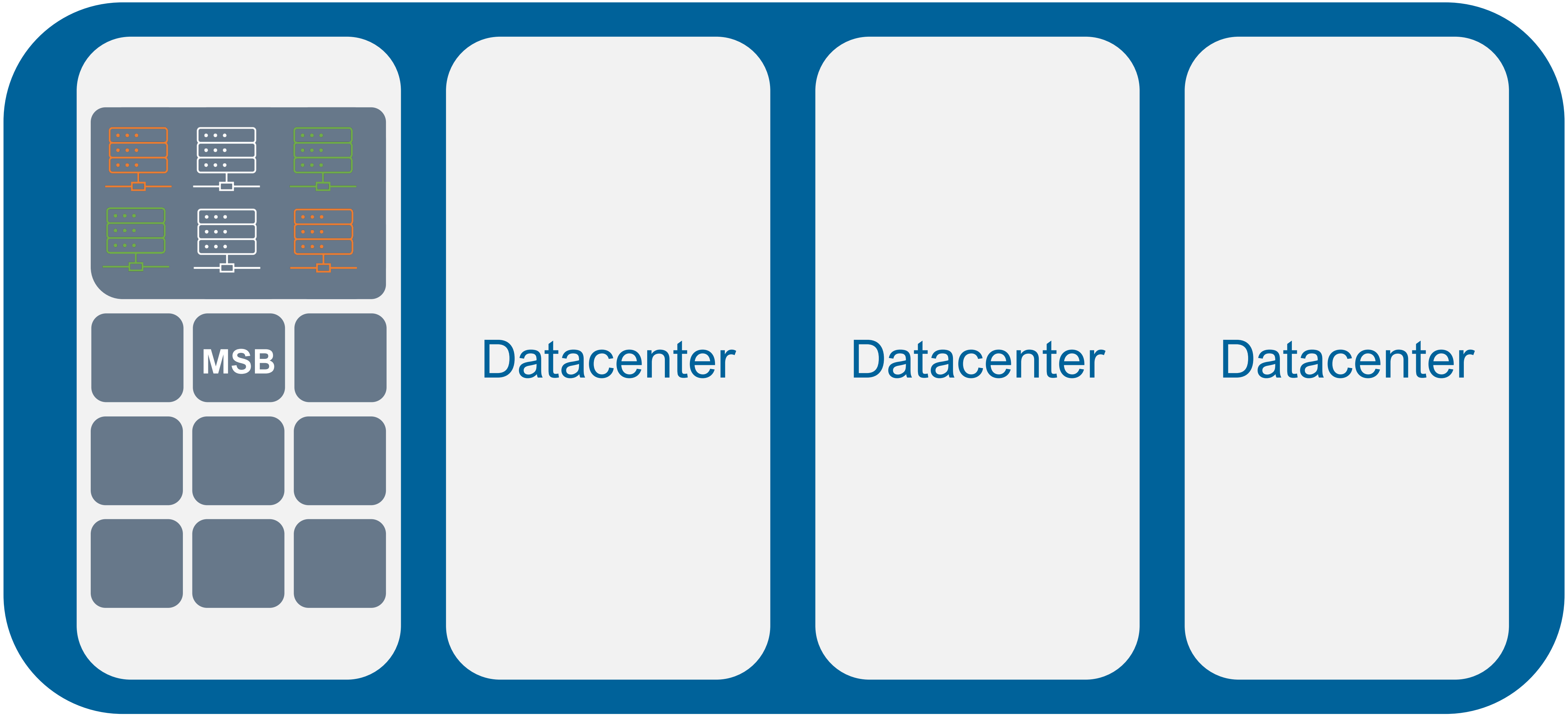
ML

RAS Operation

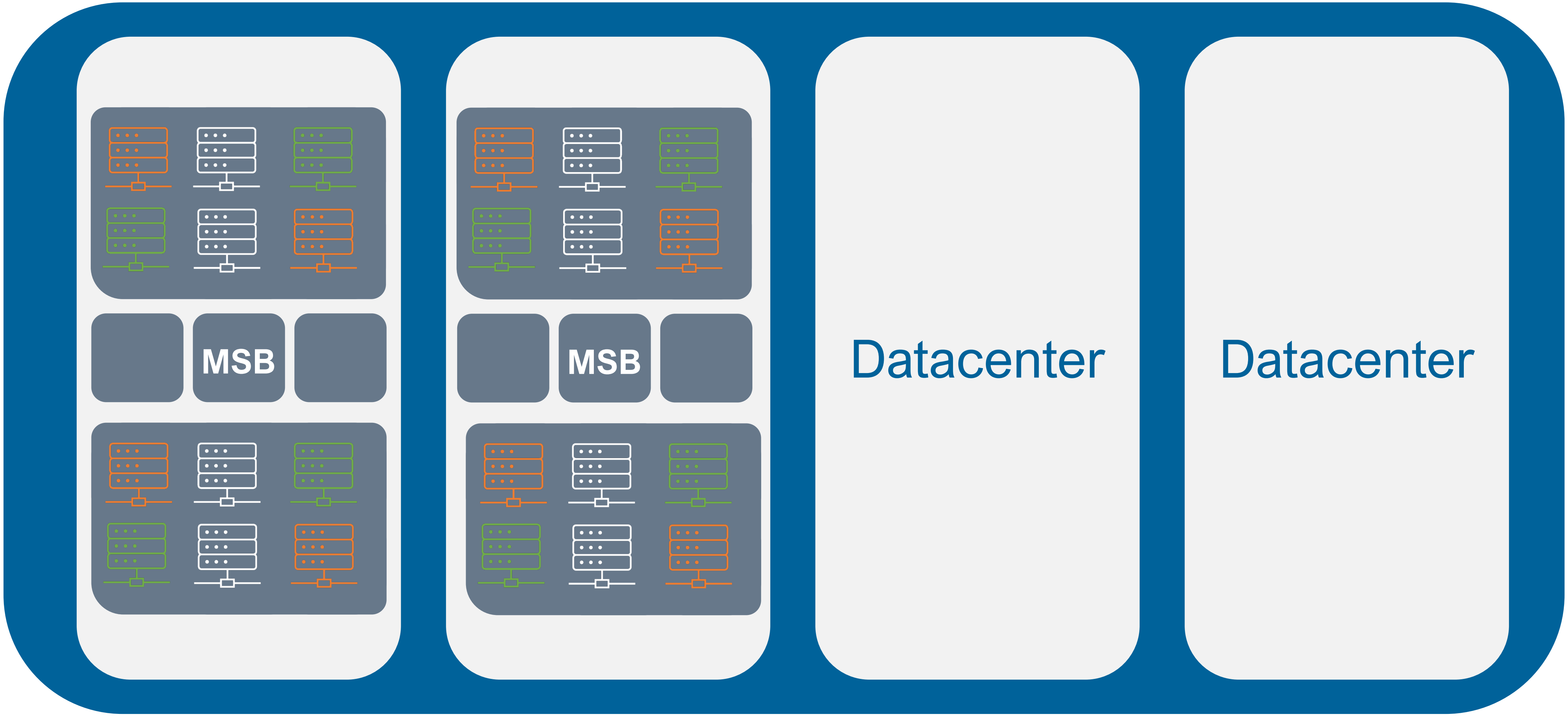


RAS provides guaranteed & continuously optimized capacity to services without user involvement!

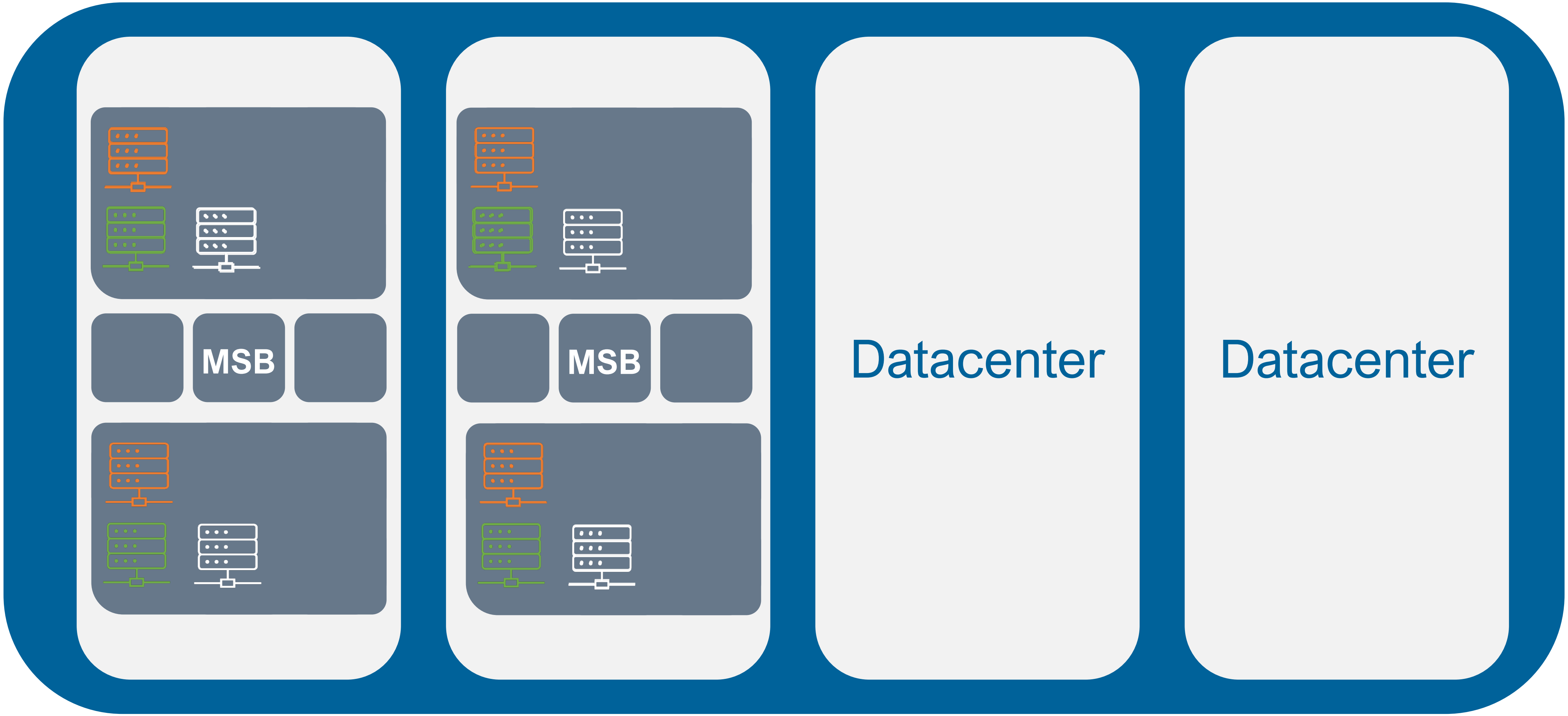
Exploiting Server Symmetry



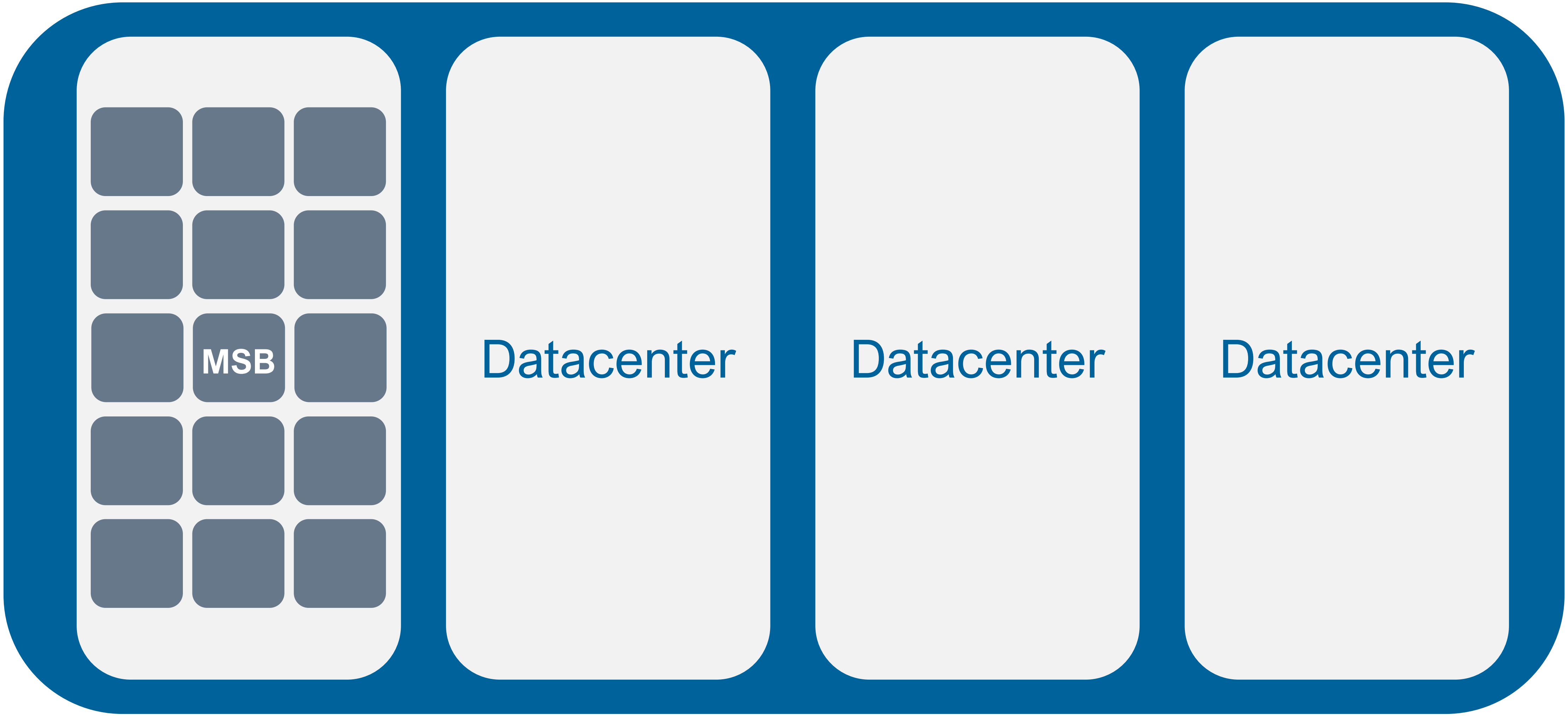
Exploiting Server Symmetry



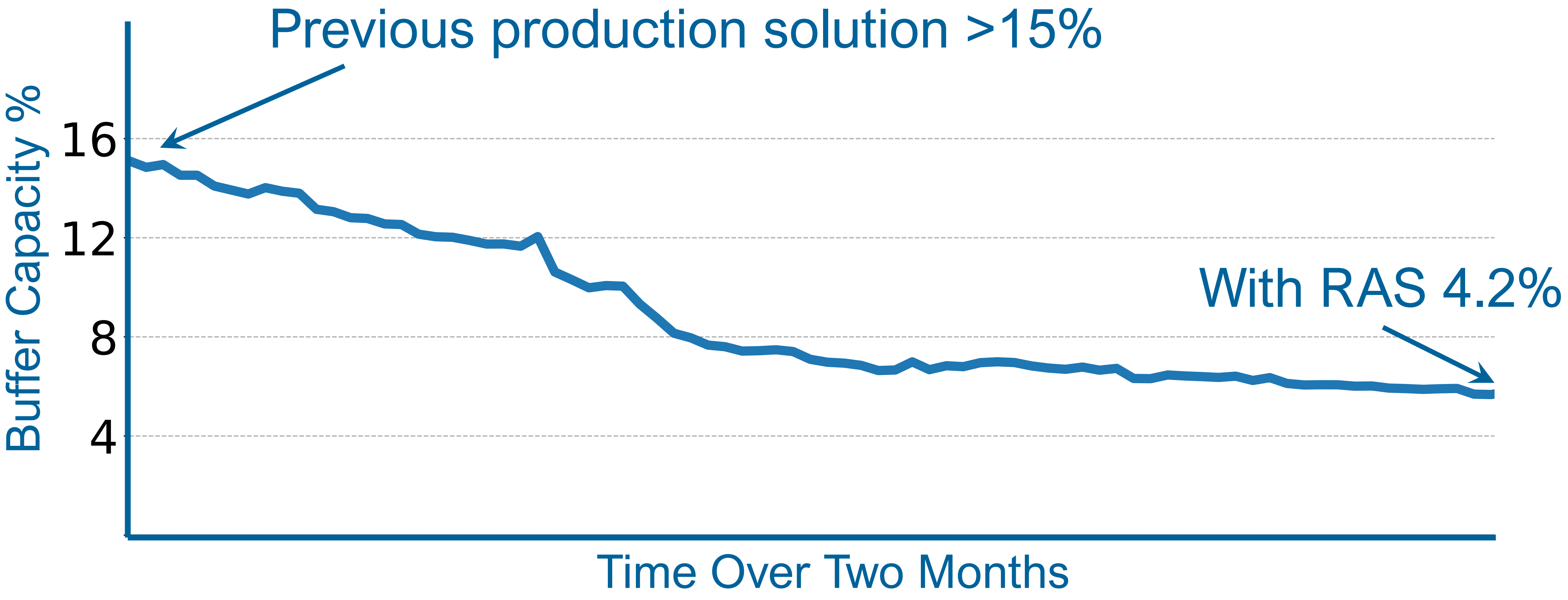
Exploiting Server Symmetry



Exploiting Server Symmetry



RAS Evaluation



More in The Paper

Resource Management Realities

Elastic Reservations

Detailed MIP Formulation

Additional Evaluation

- MSB Spread, Power, Network, Churns and more!

Lessons Learned

Challenges and Ongoing work

Takeaways:
Continuously Optimized Resource Allocation

Reservations → A new abstraction for guaranteed capacity allocation

Decouple server-to-reservation assignments from container placement

Formulate capacity as a MIP across 1M+ servers in production FB regions

RAS optimizes reservations region-wide:

- Large-scale failures
- Heterogenous hardware
- Workload constraints
- Datacenter maintenance

RAS manages capacity across the entire Facebook fleet!