

Accelerate Insights

Converging Cloud and HPC

Gary Brown
HPC Product Manager
Adaptive Computing

IDC HPC User Forum
September 2014

Adaptive Computing Highlights

- **Innovating world-class HPC, Big Data, and Cloud solutions**
- **Over a decade of experience**
- **Pioneering in Scheduling and Optimization**
- **70+ patents issued or pending**
- **Backed by top-tier investors**
- **Numerous customers in the Top 500 and Fortune 500**
 - Top systems including: Blue Waters, Titan, Cielo, Hopper
 - Major multi-nationals including: DOW, ExxonMobil, & Boeing
 - Largest provider of HPC workload management software to HPC sites*
 - Global partnerships include Intel, Cray, HP, IBM, SGI, & Microsoft

Moab
CLOUD | BIG DATA | HPC



Cloud System
Management
Innovator

Accelerating Insights with Big Workflow

- **Unify data center resources**
 - As a single, adaptive ecosystem
 - Technical computing (HPC & big data)
 - Public and private cloud
 - Bare metal and virtual machines
- **Optimize the analysis process**
 - Increase throughput and productivity
 - Ensure SLAs, maximize uptime
 - Reduce cost, complexity and errors
- **Guarantee service to the business**
 - Policies that model your organization
 - Prove services were delivered
 - Job completion in spite of failures
 - Verify resources were allocated fairly



HPC Bursting

Virtual Cluster and Job Isolation Use Case

- **Medical research**

- Health information privacy regulations

- **Hospital Isolation**

- One physical data center shared among multiple hospitals
- Hospitals have minimum compute node requirement
- Dynamically partition data center into independent “virtual clusters”

- **Patient Isolation**

- One patient – one job
- Isolate patient job from all other patient jobs
- Run 1 patient job in 1 VM on 1 compute node for complete isolation

Virtual Cluster and Job Isolation Use Case

- **One physical HPC cluster**
 - Multiple “virtual clusters”
 - Virtual clusters completely isolated using own scheduler
 - One partition of “unused” compute nodes for on-demand allocation to “virtual clusters”
- **Virtual clusters expand and shrink resources**
 - Request compute nodes from “unused” partition
 - Provision, use, and destroy one VM per compute node
 - Return compute nodes to “unused” partition
 - Triggered by job backlog “pressure”
- **HPC Bursting**

Canadian Hospital Consortium

Pilot project with two hospitals and one consortium cluster with three partitions

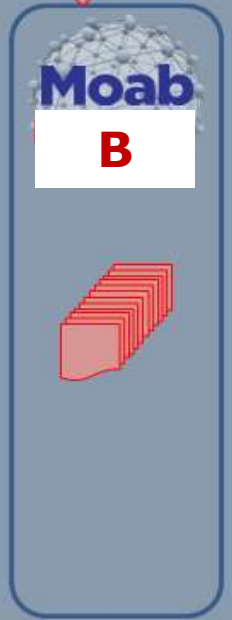
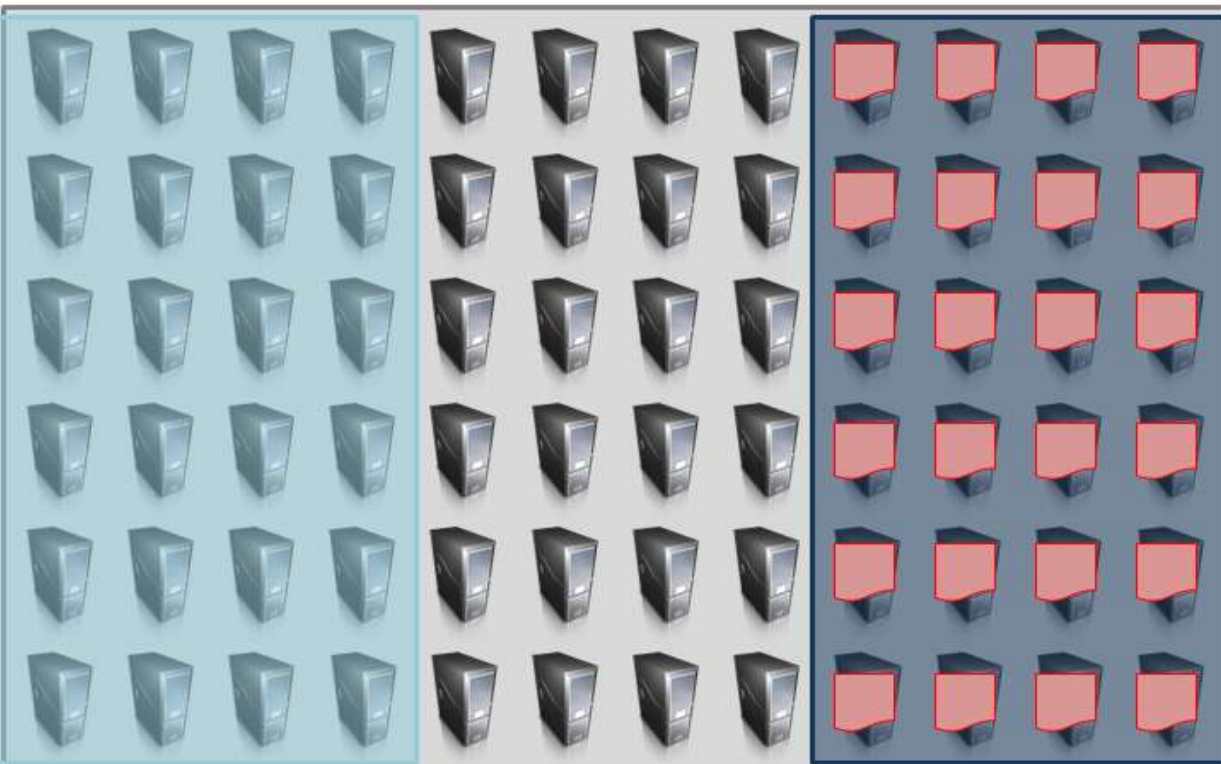
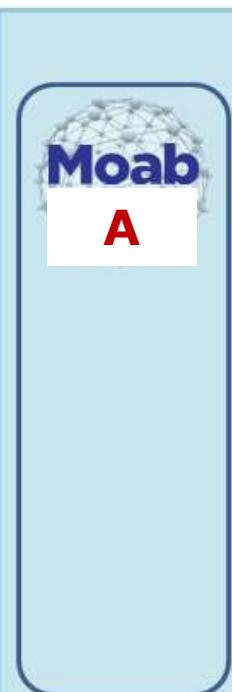


Two partitions for two hospitals (A & B) and one "free" partition (HPC)



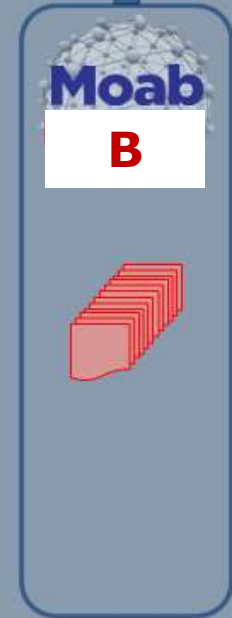
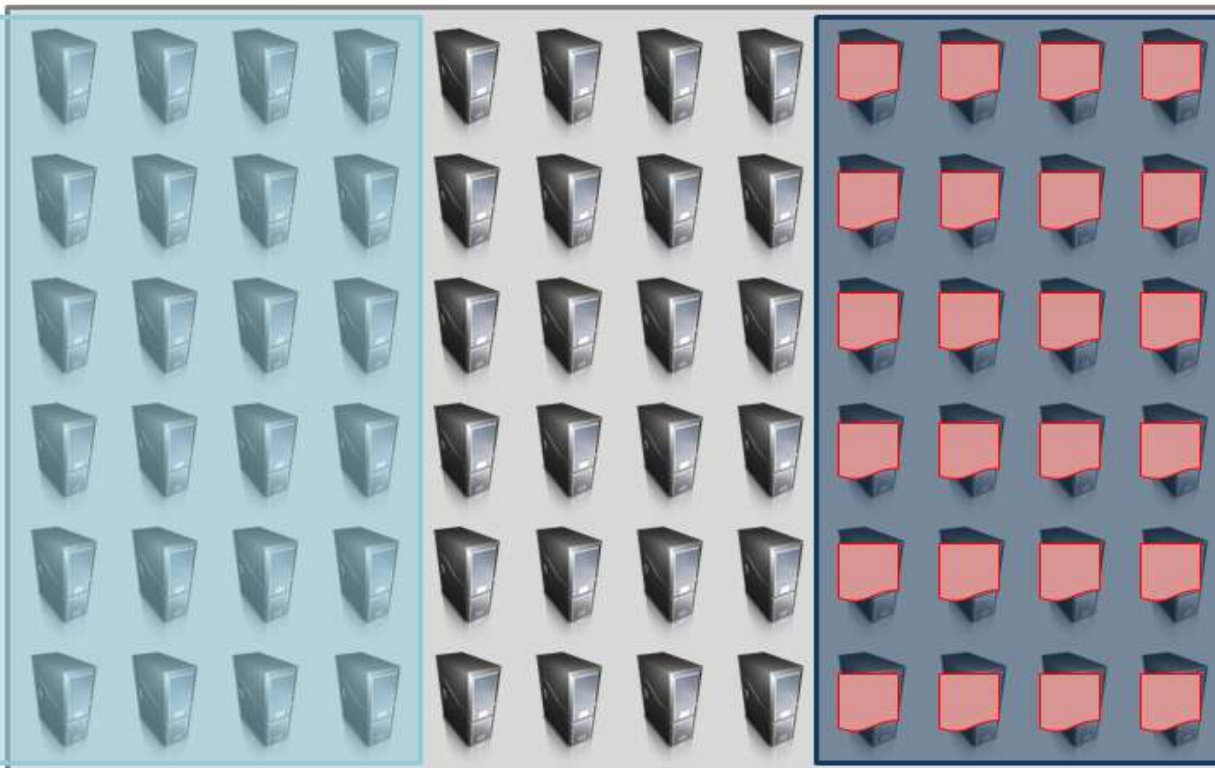
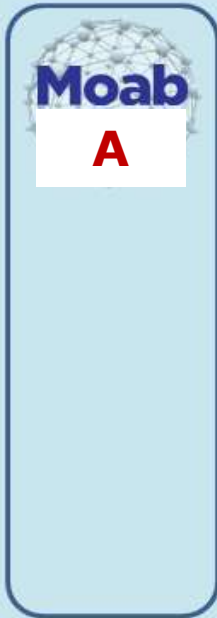
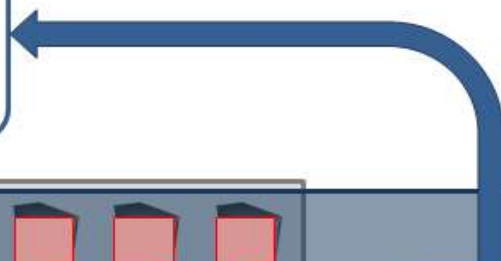
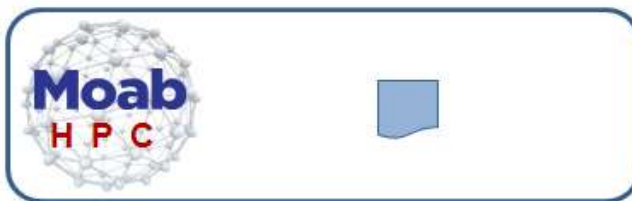
Hospital Consortium

Users submit jobs to B Moab until job backlog present



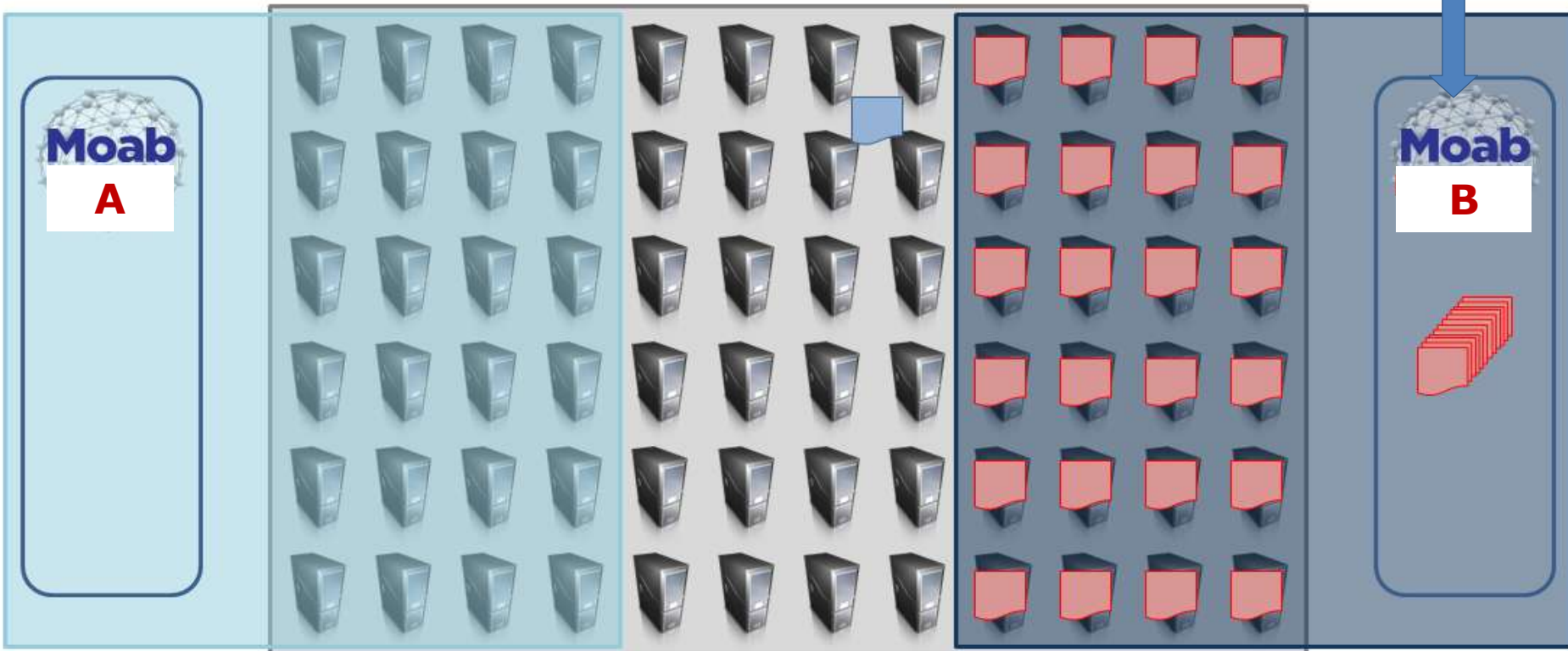
Hospital Consortium

Job backlog triggers
B Moab to request
additional resources
of HPC Moab



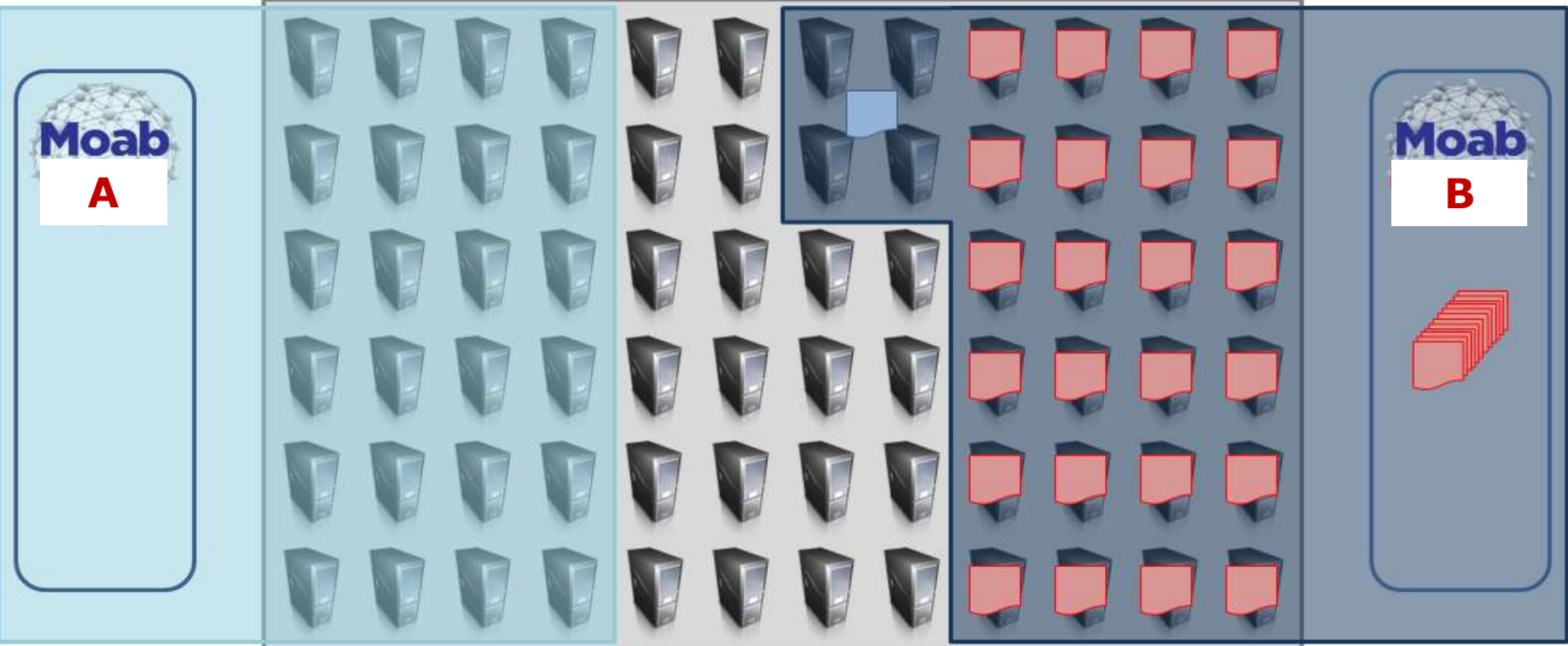
Hospital Consortium

HPC Moab allocates unused nodes and informs B's Moab of resources



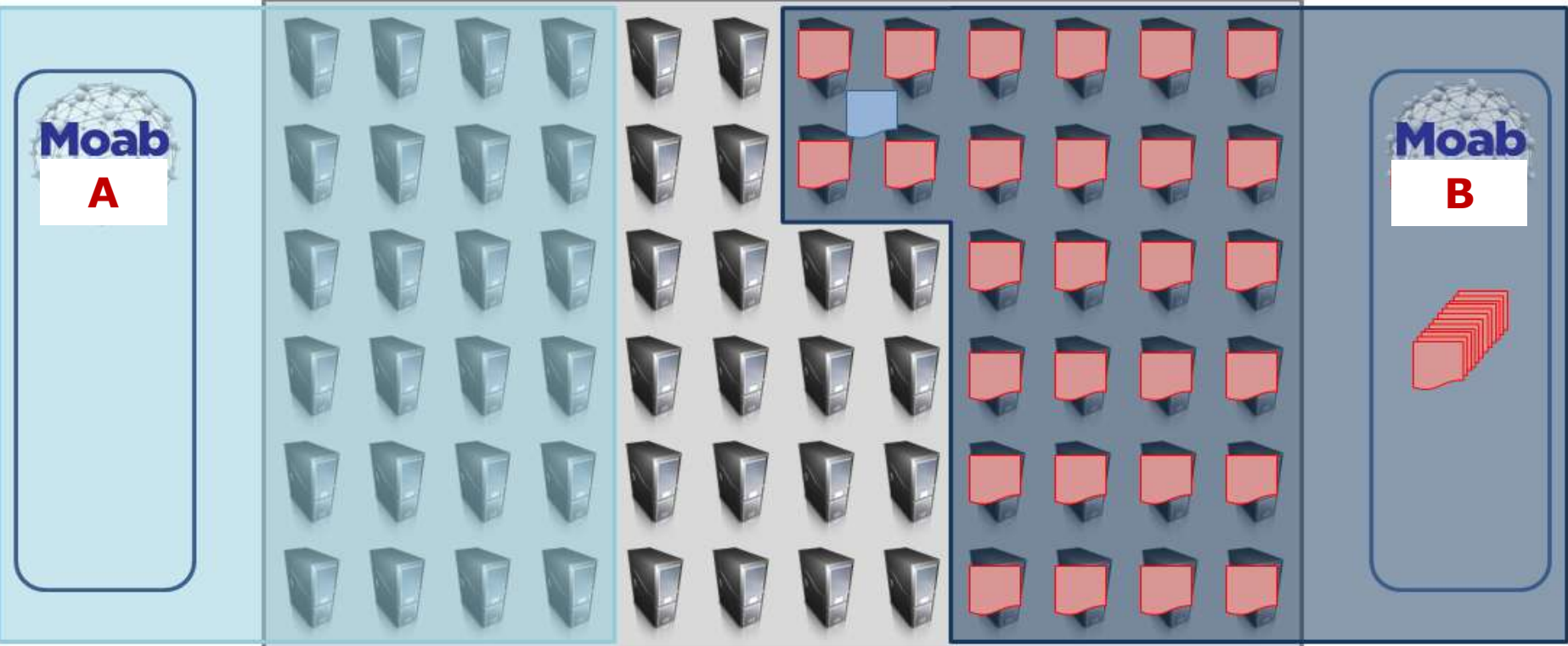
Hospital Consortium

**B Moab adds
allocated nodes to its
cluster**



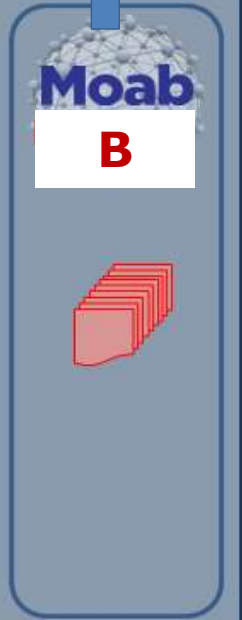
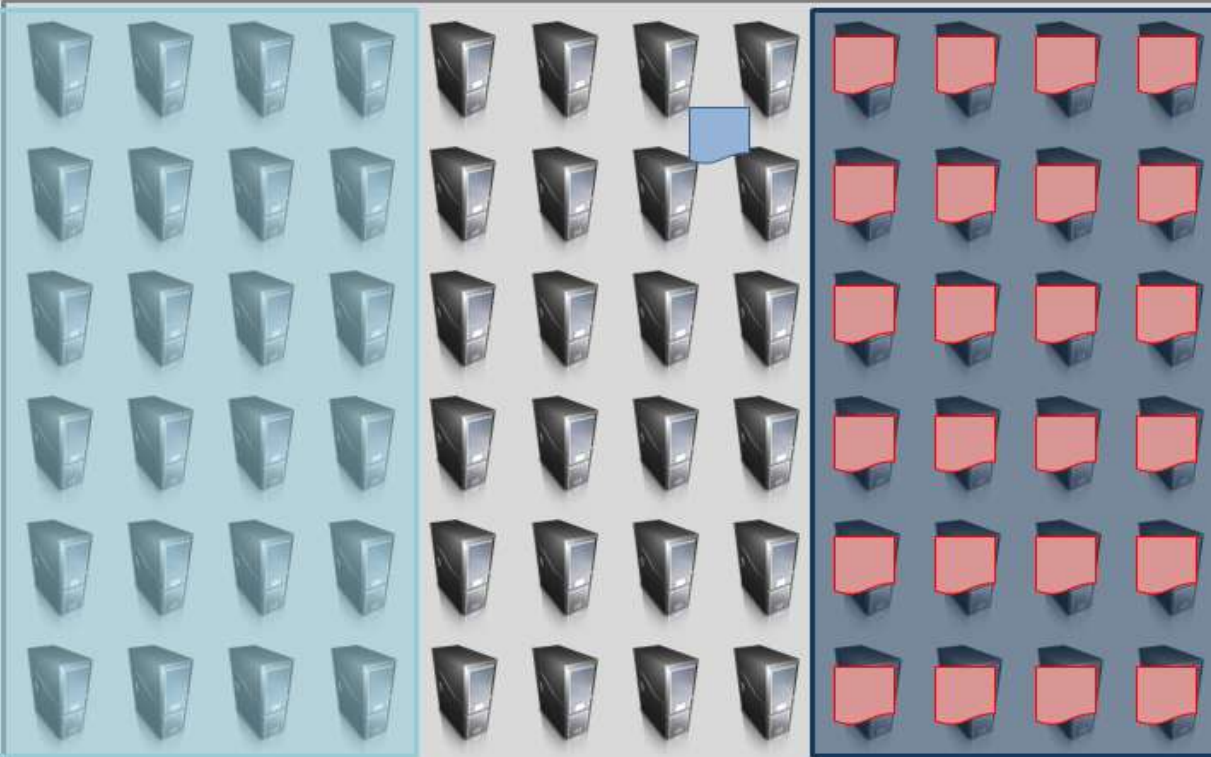
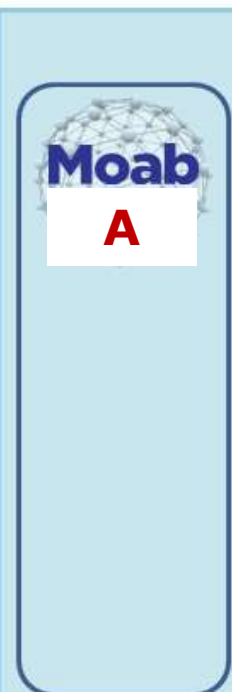
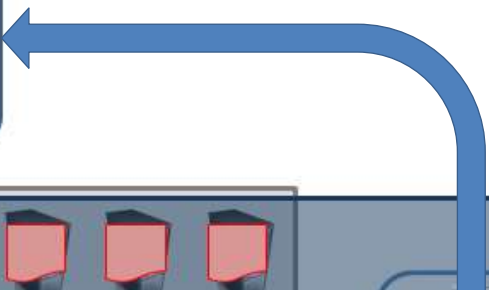
Hospital Consortium

B Moab runs jobs on new additional nodes



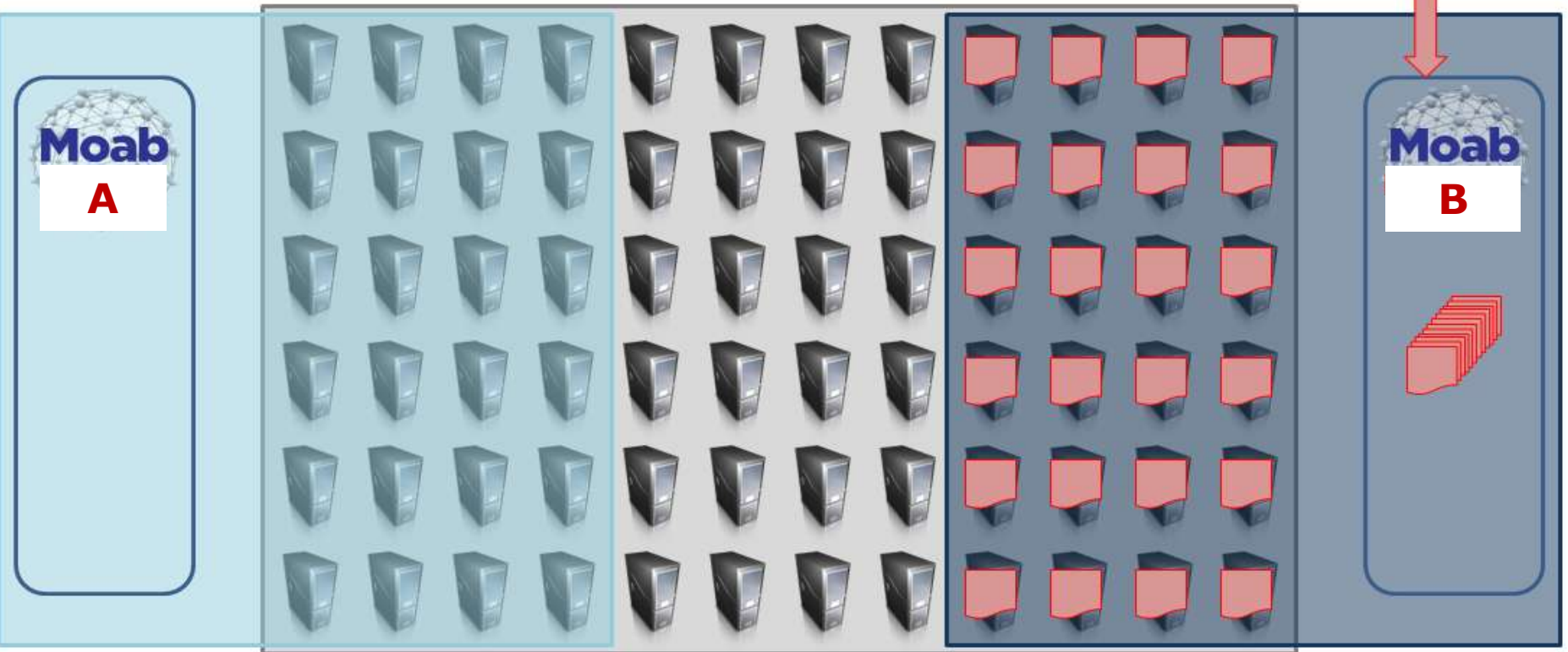
Hospital Consortium

When job backlog eases, B Moab returns nodes to HPC Moab



Hospital Consortium

Users submit more jobs to B Moab, repeating cycle



Virtual Cluster and Job Isolation Use Case

- **Use case can extend to cloud resources**
- **Future possible use of OpenStack for PM and VM provisioning**

High-Throughput Computing (HTC)

Nitro Task Manager

- **Distributed Scheduling Approach**
 - Combines small, homogeneous jobs into a “batch of tasks”
 - Creates policies for entire batch
- **Schedules batch as one job**
 - Incurs scheduling overhead only once
- **Launches Nitro job with task file**
 - Nitro “coordinator” launches tasks on “workers”
 - Tasks execute on Nitro job’s allocated resources
- ***Independent Product***
 - Submit as job to any batch scheduler
 - Run standalone



Nitro Task Manager

- **100x faster throughput for small, short, serial jobs**
- **Early-access Version Benchmark**
 - 13,500 tasks/second submission rate
 - Sustained in 20-node EC2 cluster
 - 10 million tasks <12.5 minutes
 - Ramps higher for larger systems

