



PBS Pro on the Cray X1 Platform

Michael Karo
June 2, 2003



Scheduling Hierarchy and Scope

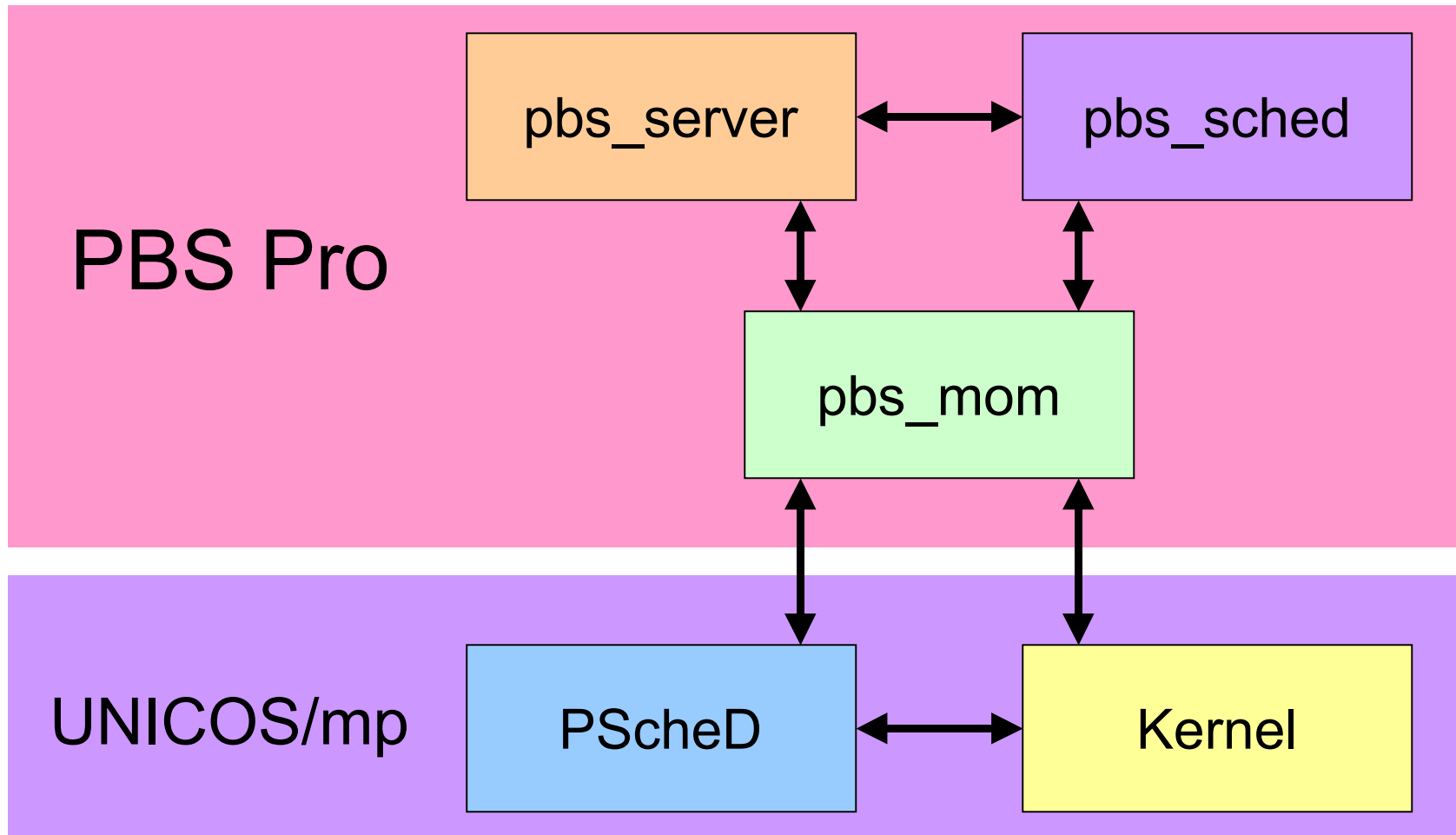


Name	Scope	Example
Grid	Global	Globus [†]
Batch	Organizational, Departmental, or Cluster	PBS Pro
Placement	Single System, Multinode	PScheD
Process	Single Node, Multiprocessor	UNICOS/mp

[†] Not presently a Cray supported product.



Functional Organization

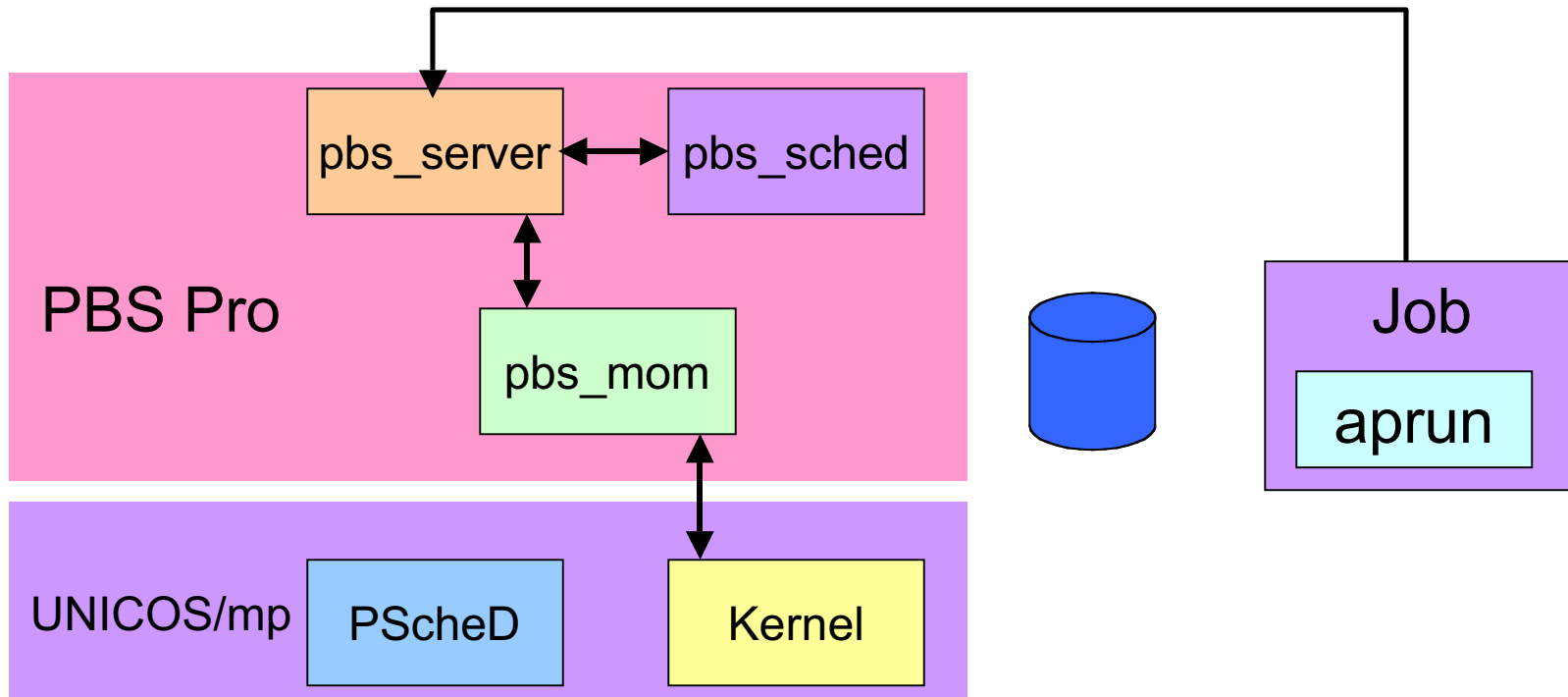




Job Flow



Step 1: Job Submission

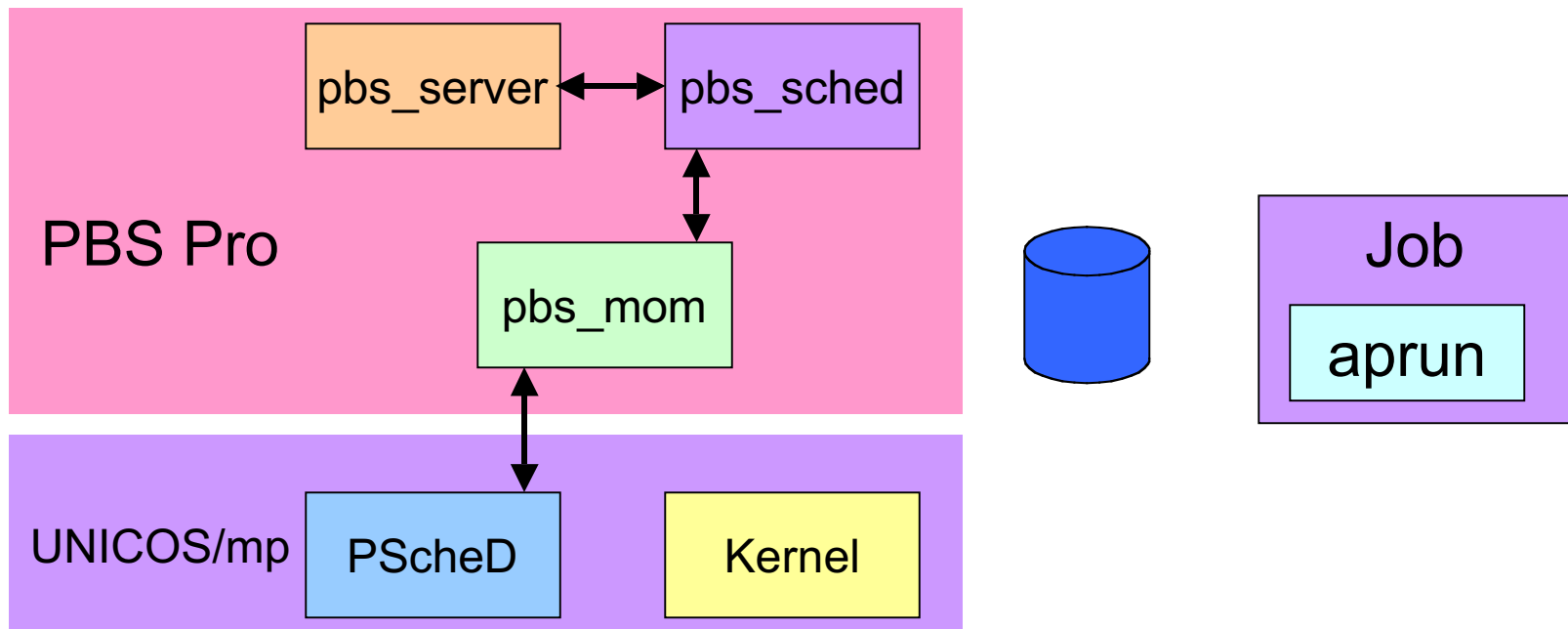




Job Flow

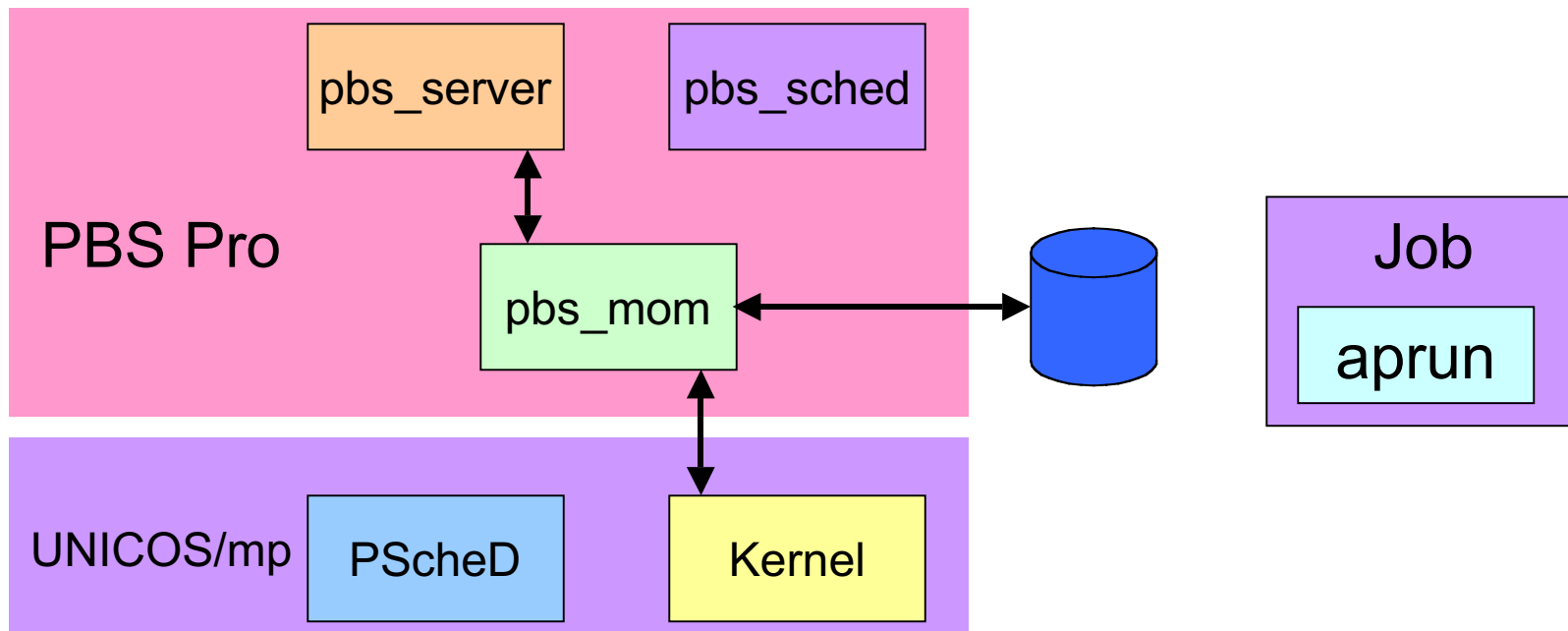


Step 2: Job Scheduling





Step 3: Support Node Initiation

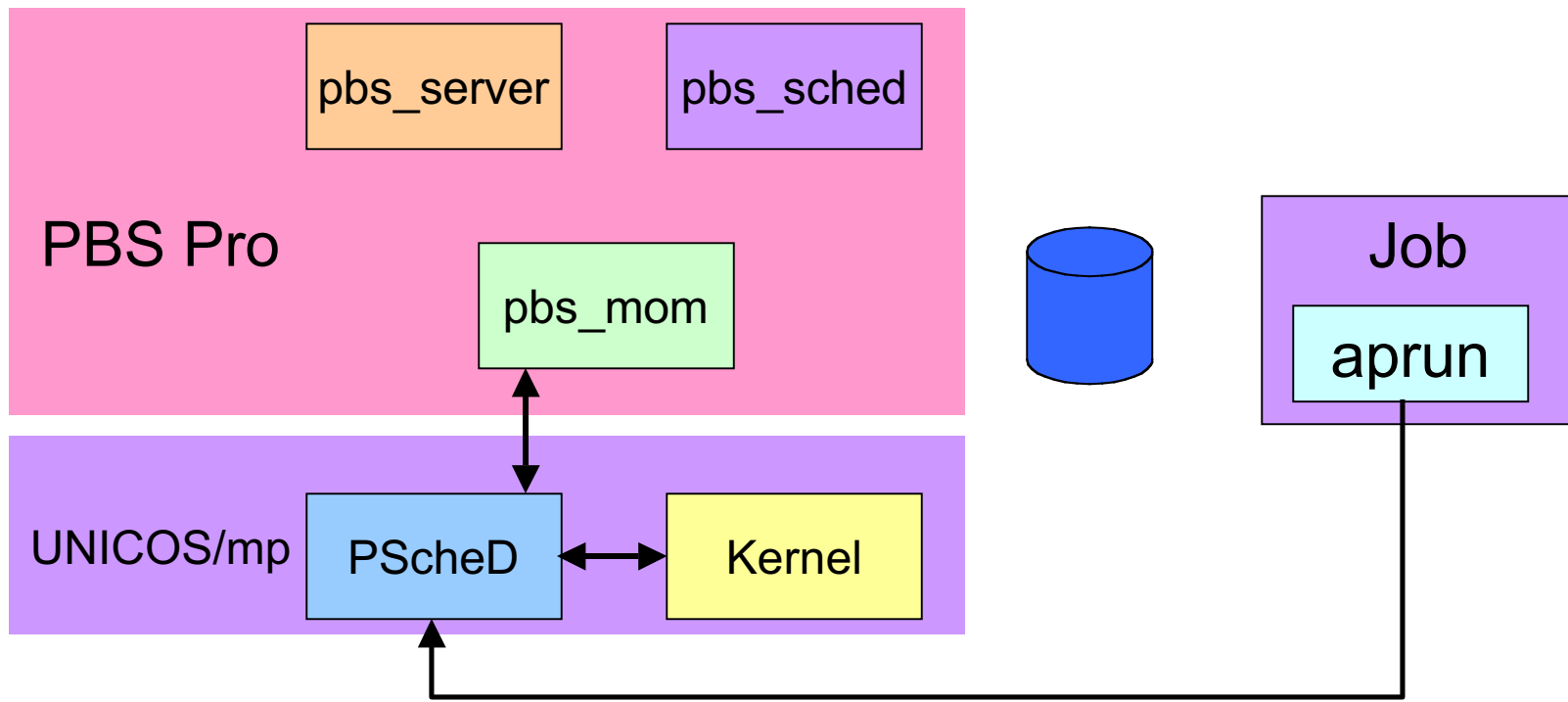




Job Flow

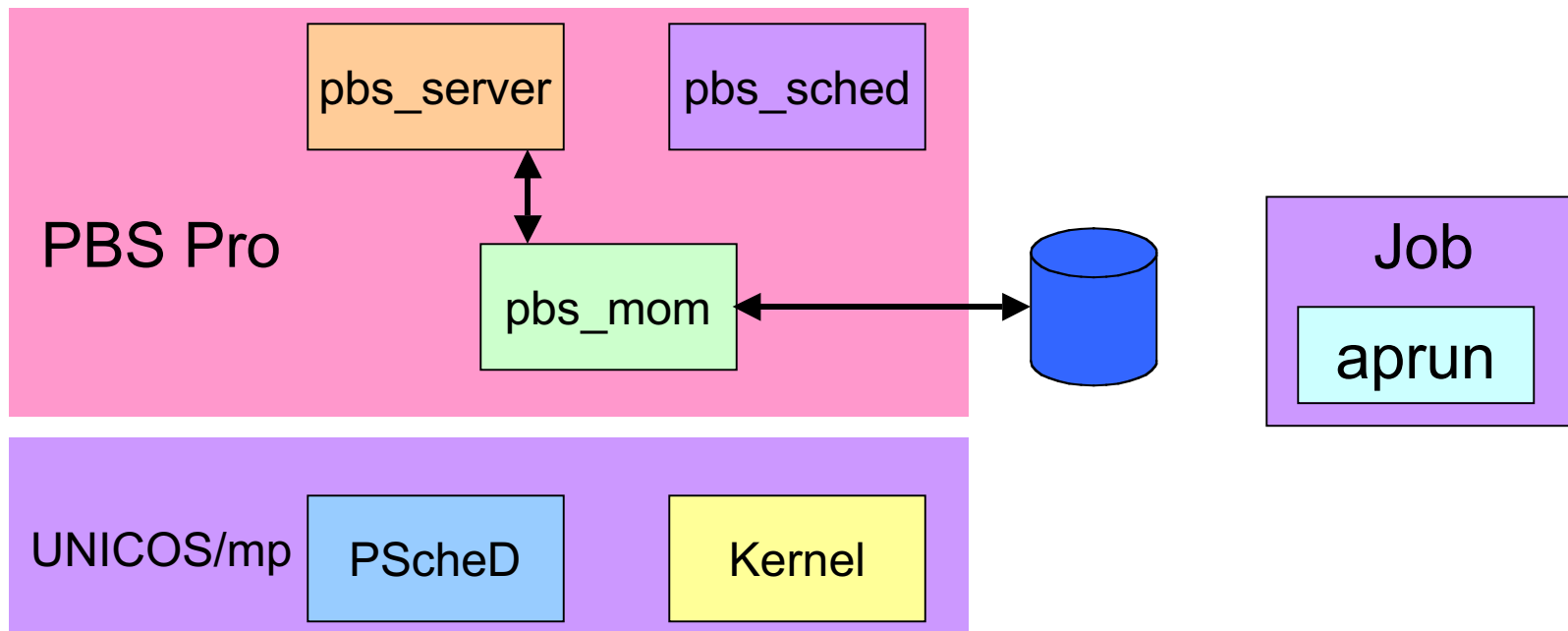


Step 4: Application Node Placement



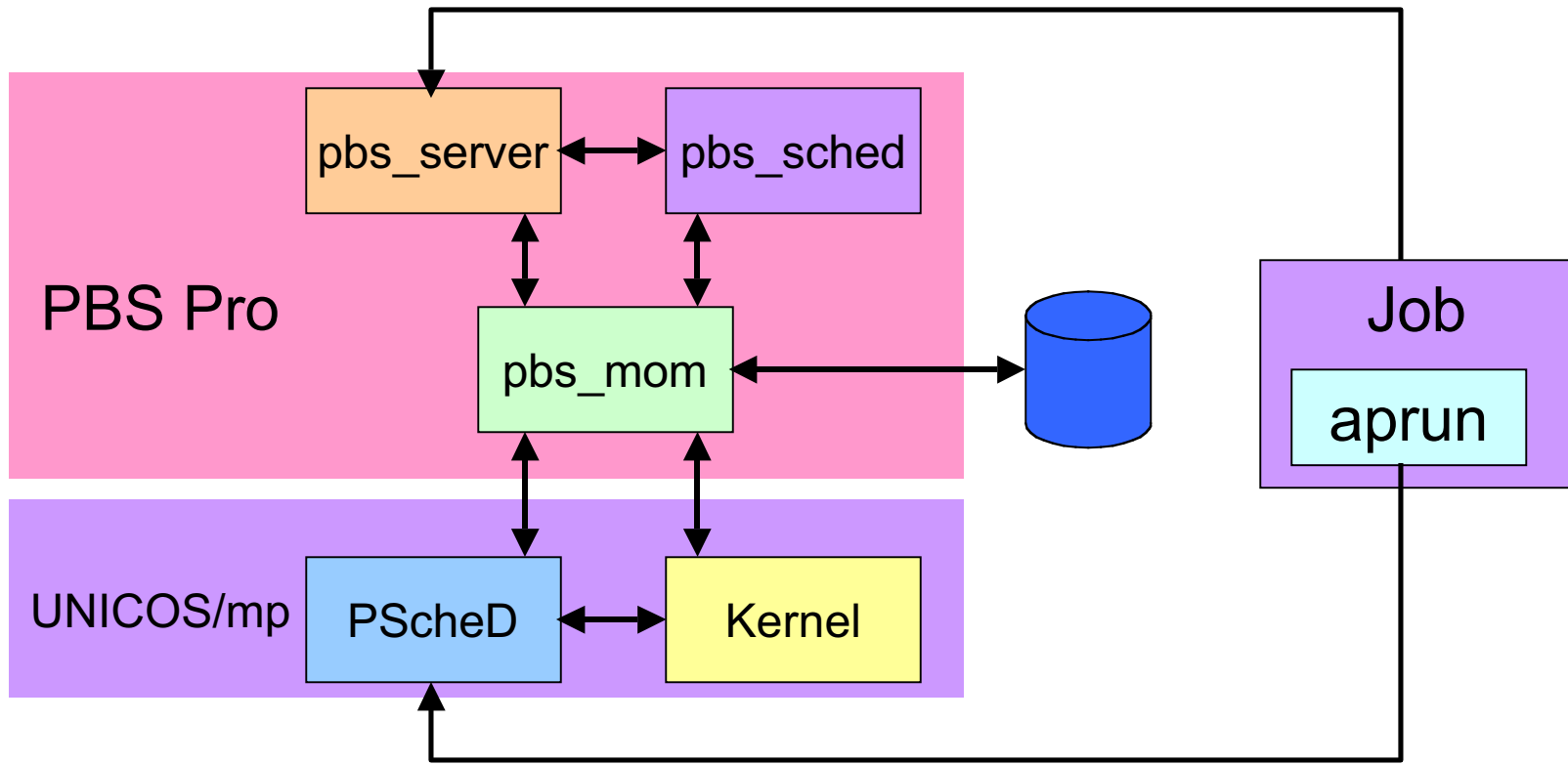


Step 5: Job Termination





The Big Picture





UNICOS/mp Specific Job Resource Limits



<p>mppe</p>	<p>Maximum number of MSP processing elements that may be used by a single process running on application nodes.</p>
<p>mppssp</p>	<p>Maximum number of SSP processing elements that may be used by a single process running on application nodes.</p>
<p>mppfile</p>	<p>Maximum size of any single file that each process in the job may create while running on application nodes.</p>
<p>pmppt</p>	<p>Maximum amount of CPU time that each process in the job may use while running on application nodes.</p>
<p>pmppmem</p>	<p>Maximum resident memory segment size that each process in the job may allocate while running on application nodes.</p>
<p>pmppvmem</p>	<p>Maximum amount of virtual memory that each process in the job may allocate while running on application nodes.</p>



Checkpoint/Restart



- Manual CPR control via qhold/qrls
- Periodic checkpoint supported
- External checkpoint utility employed
- pbs_mom configuration attributes supported
 - **\$checkpoint_path** – defines directory location where checkpoint/restart files are stored
 - **\$checkpoint_utility** – defines the full path to the executable that is called to perform checkpoint/restart functions (/usr/bin/cpr by default)



Scheduling Features/Concerns



- Oversubscription
- Preemption
- Interactive Jobs
- Reservations
- Peer Scheduling
- Load Balancing
- Fairshare

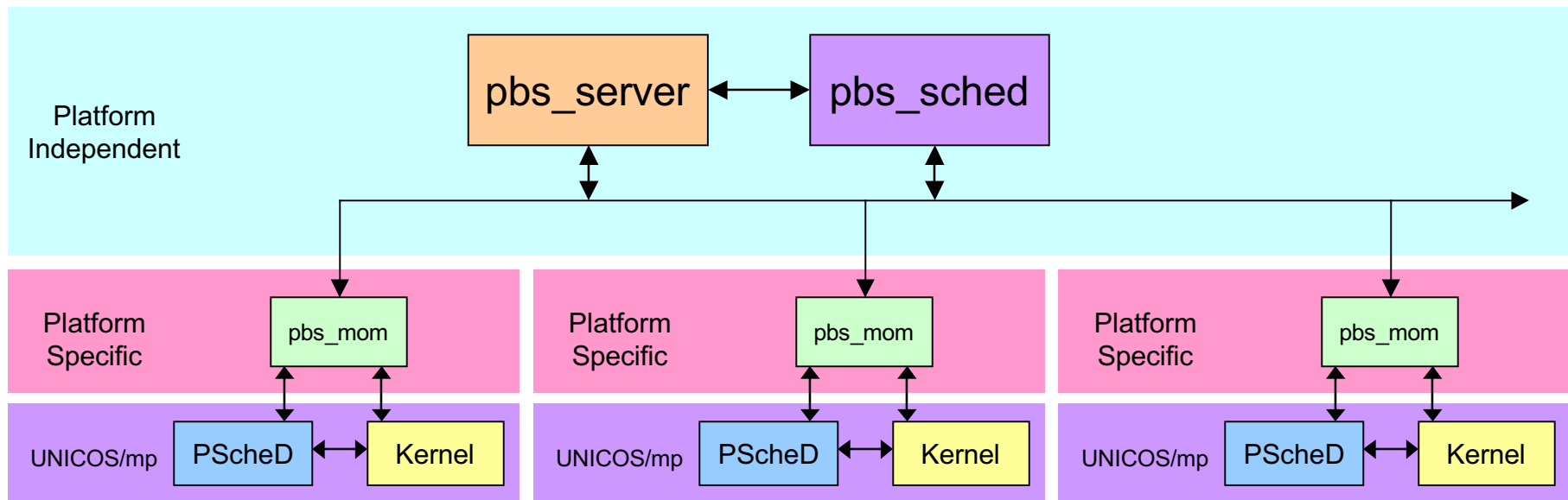


Future: Homogeneous Cluster Integration



PBS Manages:

- Multiple Cray X1 resource sets (a.k.a. partitions)
- Job dependency and synchronization



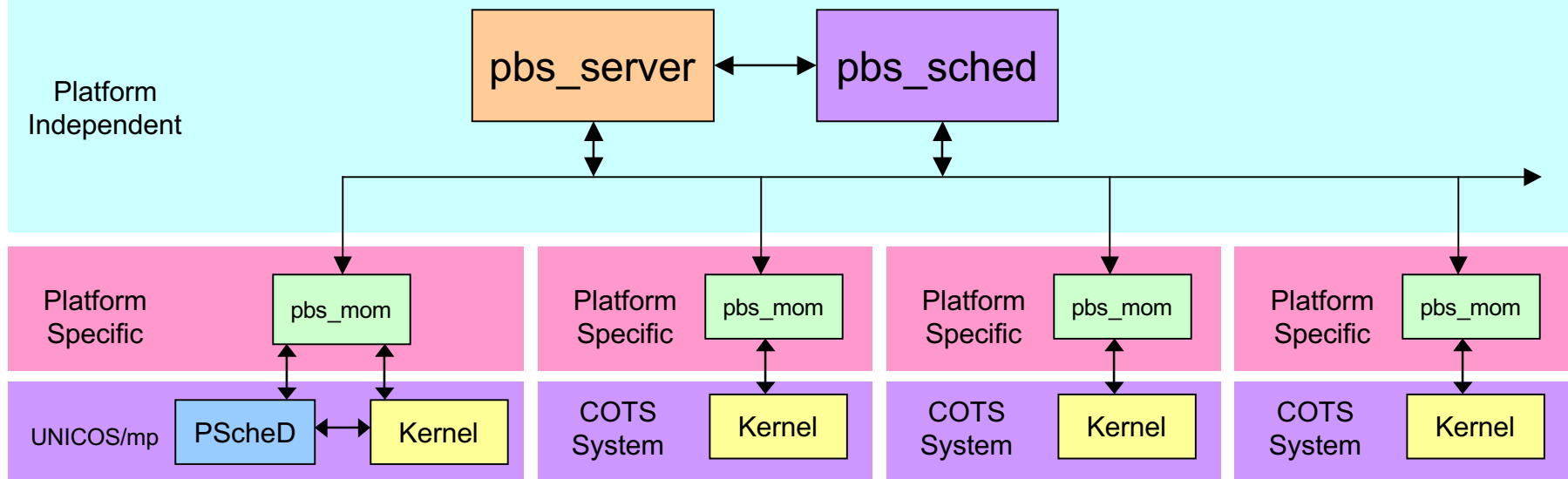


Future: Heterogeneous Cluster Integration



PBS Manages:

- Heterogeneous resource sets
- Job dependency and synchronization





Future: Computational Grid Support



- Investigating support for
 - Globus 3.0
 - UNICORE
- Enhanced accessibility
- Grid Services
 - GASS
 - GRAM
 - GIIS

Globus
GRAM

