

## Use of Common Technologies between XT and Black Widow

CUG 2006



This Presentation May Contain Some Preliminary Information, Subject To Change



# Agenda

- System Architecture Directions
- Software Development and Customer Rationales
- Software Stack
- Status

# Hardware Commonality

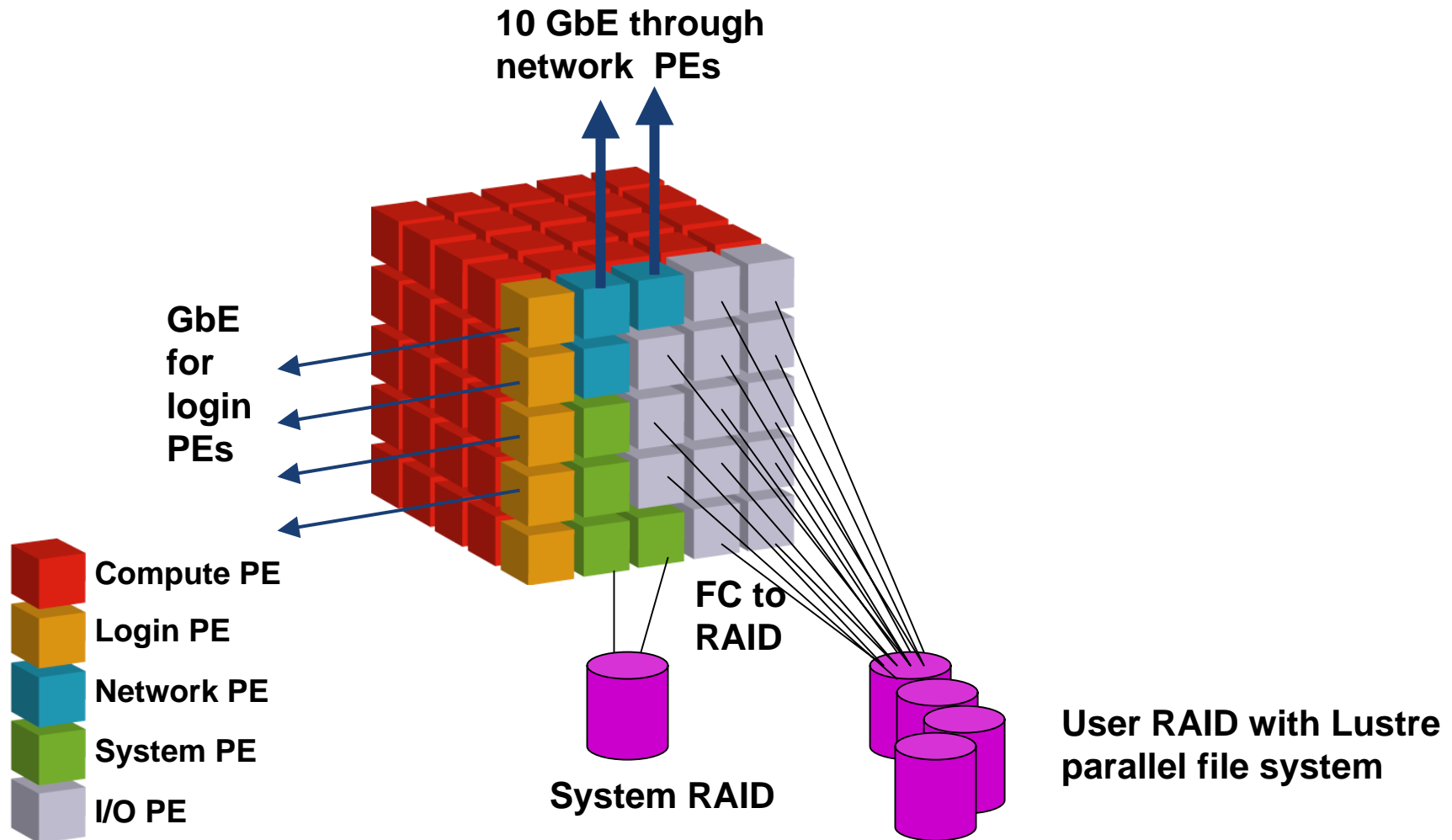
## 2005

Component	X1E	XD1	XT3	Eldorado
Processor	MSP	Opteron	Opteron	Threadstorm
Interconnect	2-D Torus	RapidArray Direct or Fat Tree	SeaStar 2D-Mesh or 3D-Torus	SeaStar 3D-Torus
External Disk	FC RAID	FC RAID	FC RAID	FC RAID
Memory	RDRAM(custom)	SDRAM	SDRAM	SDRAM
I/O Subsystem	Custom	HT to PCI-x	HT to PCI-x	Same as XT3
System Mgmt	Custom	AM	CRMS	CRMS

## 2007

Component	BlackWidow	Hood	Eldorado
Processor	SSP VCPU	Opteron	Threadstorm
Interconnect	Fat Tree	SeaStar 2D-Mesh or 3D-Torus	SeaStar 3D-Torus
Disk	FC RAID	FC RAID	FC RAID
Memory	DDRII (custom)	SDRAM	SDRAM
I/O SubSystem	Hood	HT to PCI-x/e	Hood
System Mgmt	HSS	HSS/CRMS	HSS/CRMS

# Rainer – XT3 Architecture



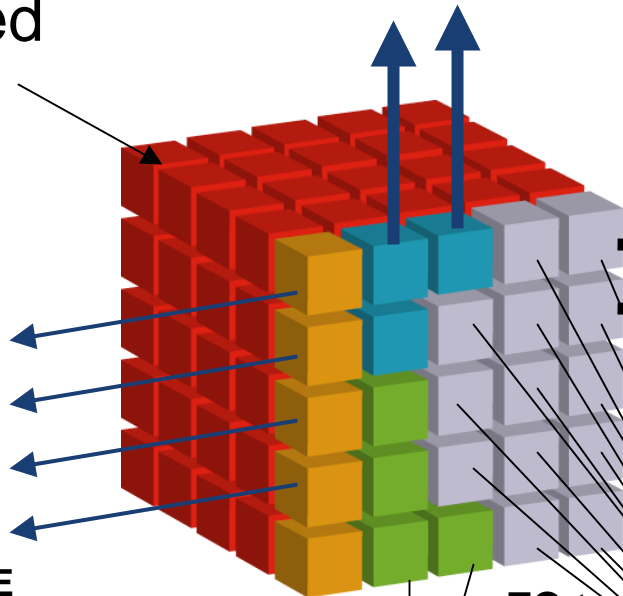
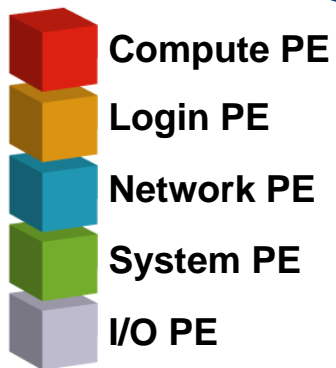
# Rainer and Black Widow

Compute PEs  
not required  
for BW

10 GbE through  
network PEs

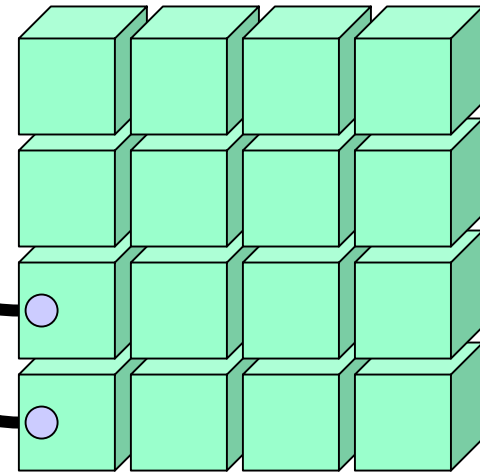
## Cray BlackWidow System

GbE  
for  
login  
PEs



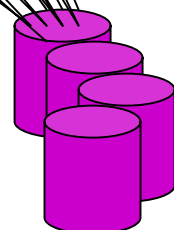
FC to  
RAID

System RAID

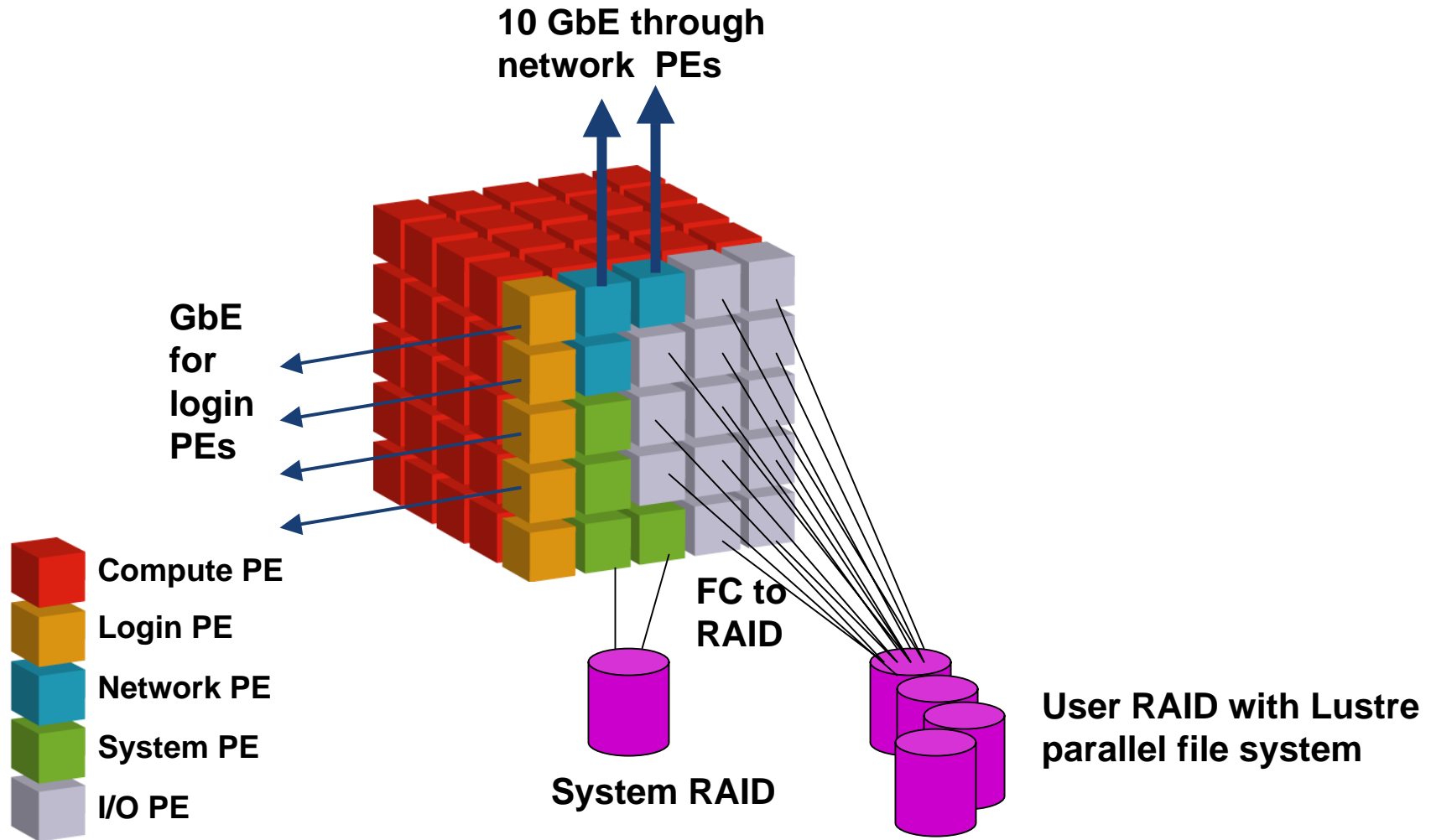


○ StarGate blade for  
BlackWidow system

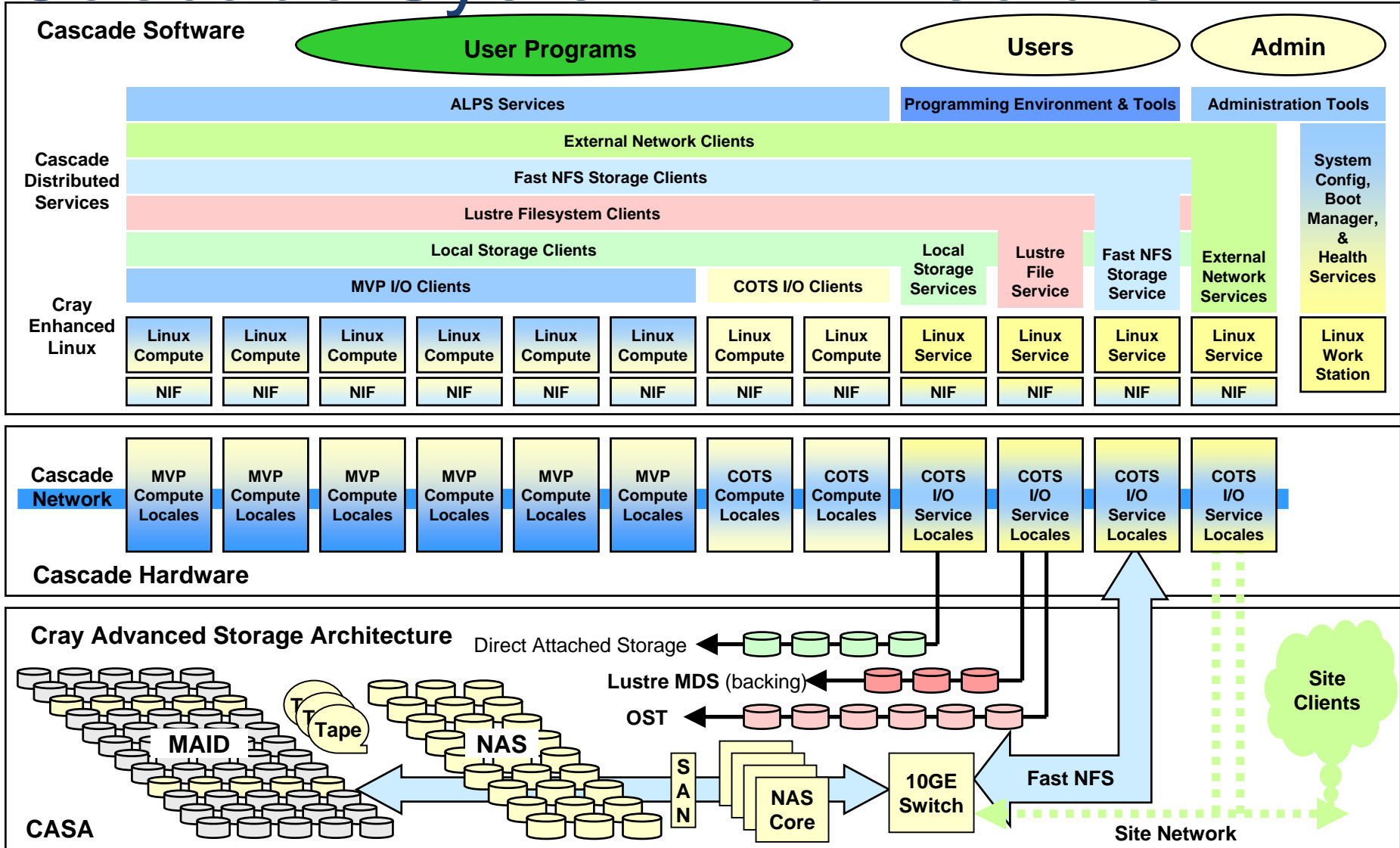
User RAID with Lustre  
parallel file system



# Rainer – Eldorado Architecture



# Cascade System Architecture



# Software Development and Customer Rationales



# Software Development Rationale

- Inside Software Development - Common Software for any active development is a big goal -
  - A common base will allow more developers to work on the code and more runtime on the code at more sites
  - Common processes and tools will allow development to be more agile
  - Efforts underway to set up common repositories and make development processes common
- Heterogeneous Systems will require a higher level of integration and sharing of technologies
- Common base is required to integrate future Scalar products with Cascade's vector and threaded capabilities

**Goal is to leverage all program efforts into a superior integrated whole**

# Customer Rationale

- For the User and Administrator – fewer and more common interfaces
- More confidence from the testing and usage of software
- Easier to add and support mixed Cray environments
- Leads to the future heterogeneous systems

Used Software is a positive

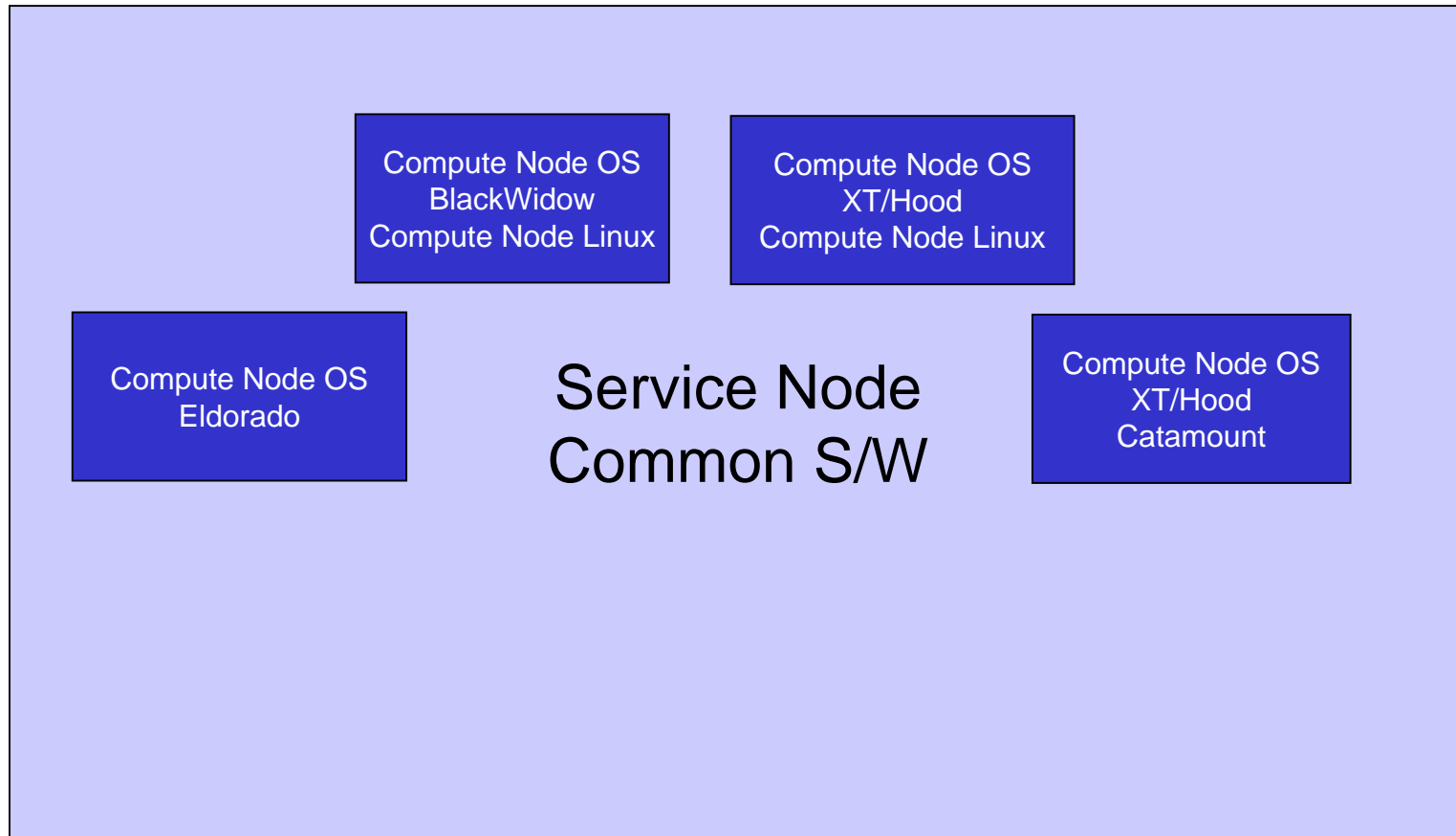
# Software Stack

# Software Commonality

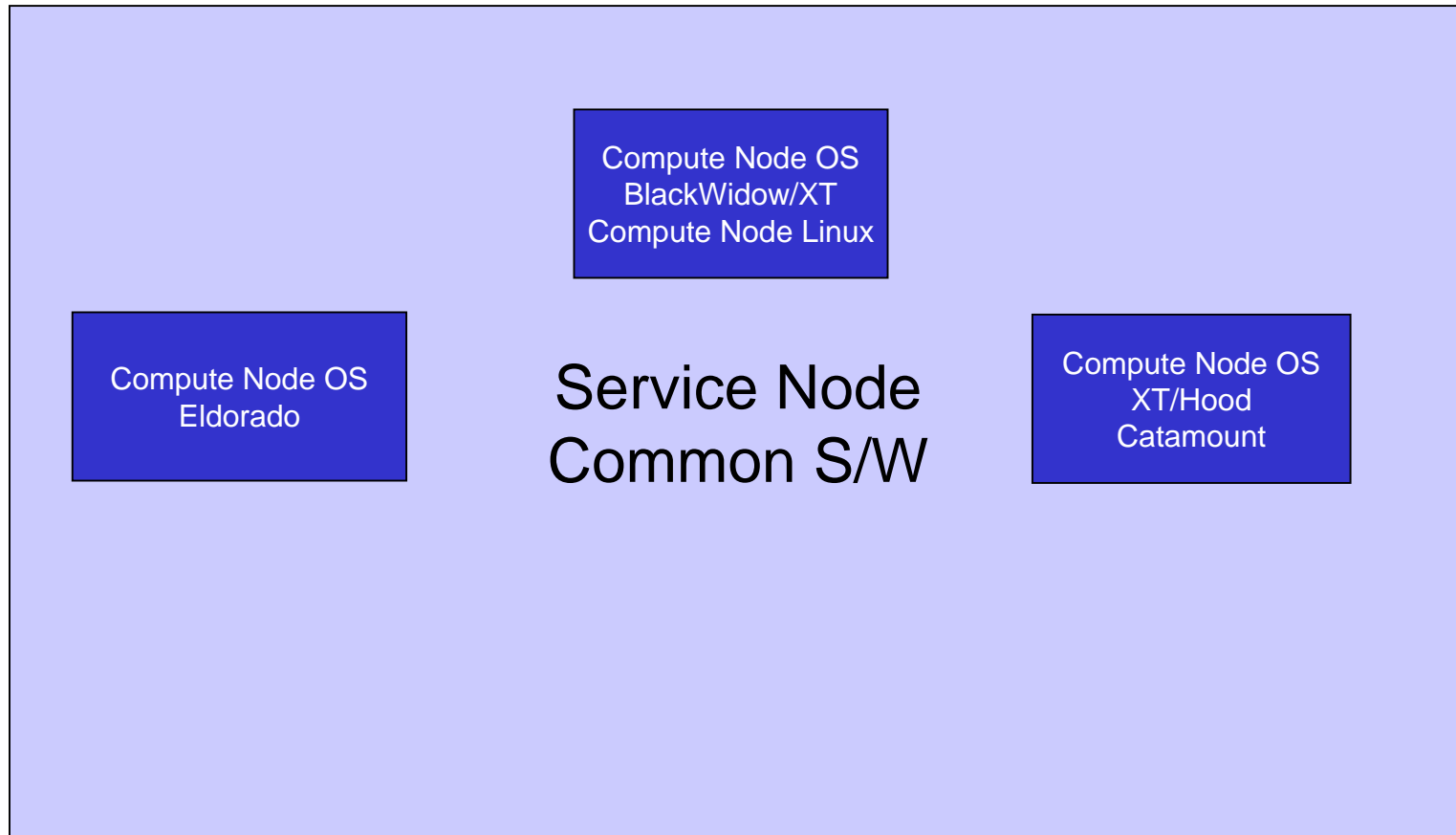
Component	X1E	XD1	XT3	Eldorado
OS	UNICOS/mp IRIX-based	Linux	UNICOS/lc Catamount (compute) Linux (SIO)	UNICOS/mt MTX OS-based (compute) Linux (SIO)
System Mgmt	SSI; IRIX Sys Mgmt	Active Manager	CRMS + Linux mgmt	CRMS with Eldo extensions
File Systems	XFS; ADIC StorNext	Lustre, NFS	Lustre	Lustre
Workload Manager	PBS Pro; LSF	PBS Pro; LSF	PBS Pro	MTA2 tools; MTX schedulers
Compilers	Cray Fortran, C/C++, UPC, Co-Array Fortran	Fortan, C/C++ (PGI, Pathscale, Absoft) Java	Fortran, C/C++ (PGI)	Eldo Compiler - C/C++
Programming Env	MPI 1.2, SHMEM, OpenMP, CrayPAT, Cray Apprentice2	MPI 1.2, MPI-IO, SHMEM, OpenMP, CrayPAT, Apprentice2	MPI 2.0, SHMEM, CrayPAT, Cray Apprentice2	MTA-2 Tools

Component	BlackWidow	Hood		Eldorado
OS	Linux	Linux	Catamount	UNICOS/mt MTX OS-based (compute) Linux (SIO)
System Mgmt	Mazama	Mazama	Mazama/CPA	Mazama with Eldo extensions
File Systems	Lustre	Lustre		Lustre
Workload Manager	PBS Pro; LSF	PBS Pro; LSF		MTA2 tools; MTX schedulers
Compilers	Cray Fortran, C/C++, UPC, Co-Array Fortran	Fortran, C/C++ (PGI, PathScale)		Eldo Compiler - C/C++
Programming Env	MPI 2.0, SHMEM, OpenMP, CrayPAT, Cray Apprentice2	MPI 2.0, OpenMP, SHMEM, CrayPAT, Cray Apprentice2		MTA Tools, Cray Apprentice2

# Common Software Example



# Common Software Example

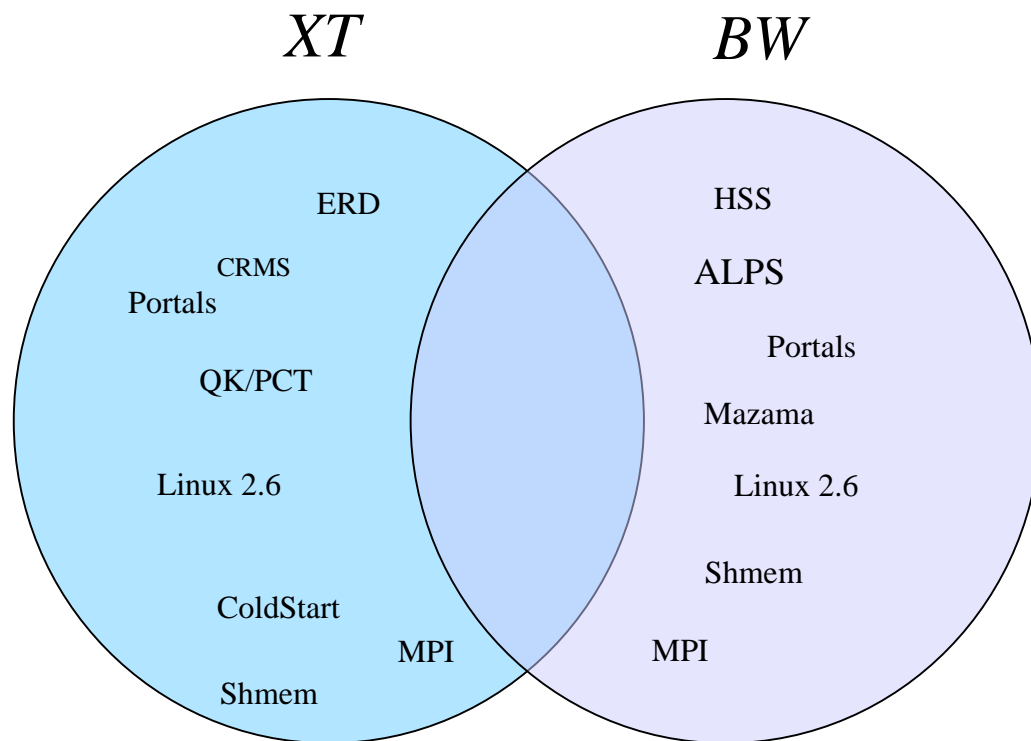


# Common Software - differences

- Common Node SW
  - Common services/code aka distribution
  - Kernel common source
- Architecture specific OS/Drivers/services
  - XT LWK node OS
  - BW LWK node OS
  - Eldorado LWK OS
  - Catamount OS
- Application architecture differences
  - Kernel drivers
    - SeaStar
    - StarGate
  - PE environments (Vector/Scalar)
    - Compilers
    - Libraries

# Existing Software XT/BW

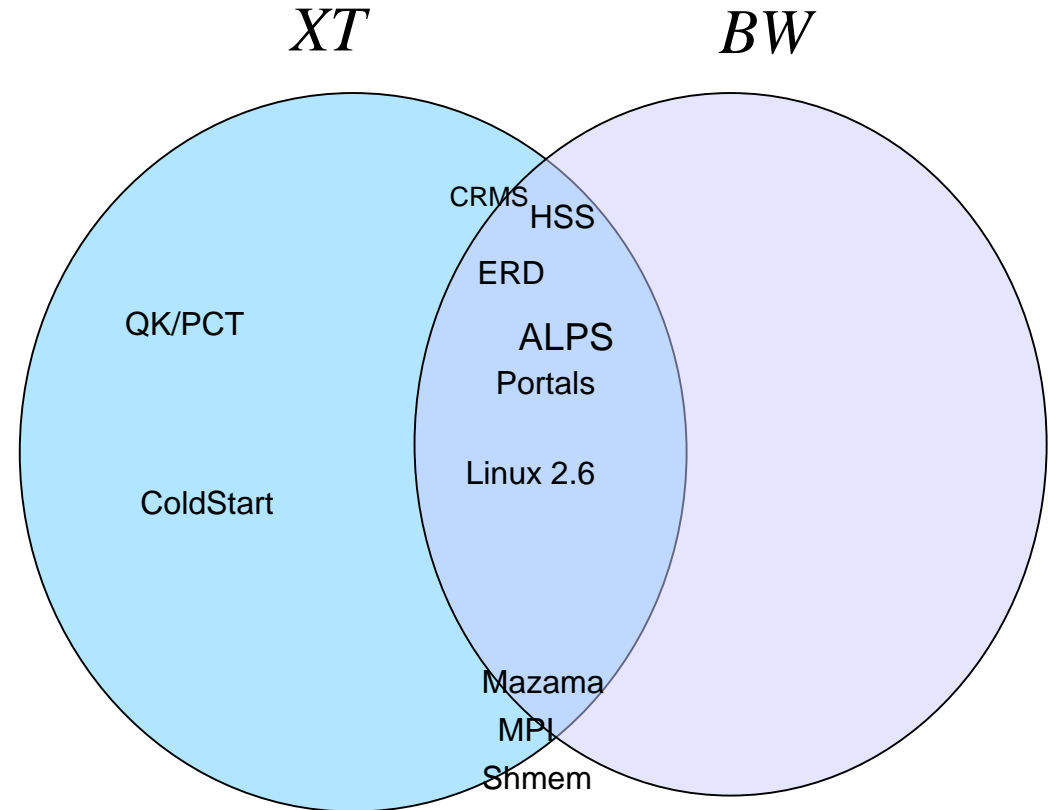
- XT has 140,000 C files
- XT proven at scale.
- XT proven software stack
- BW using newer components (kernel, libraries)
- BW not released yet, zero legacy support



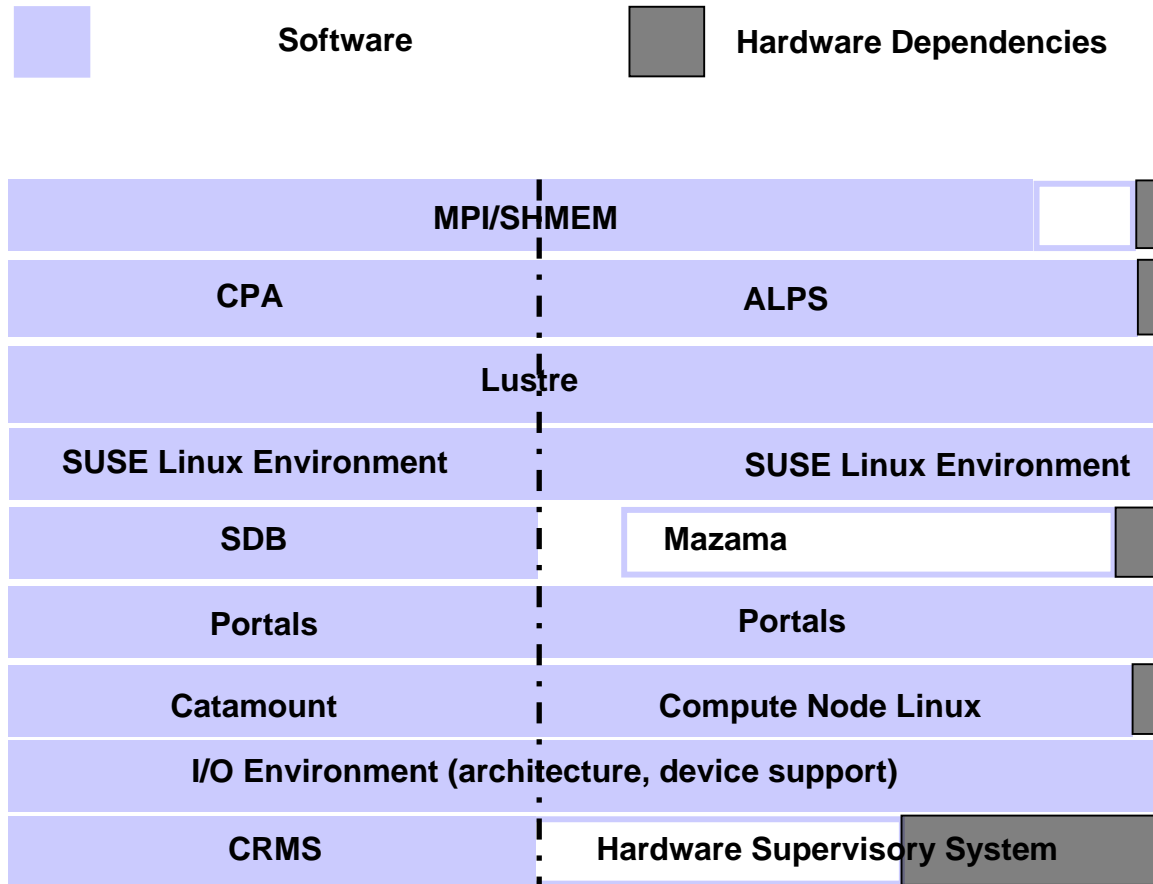


# Merging Software XT/BW

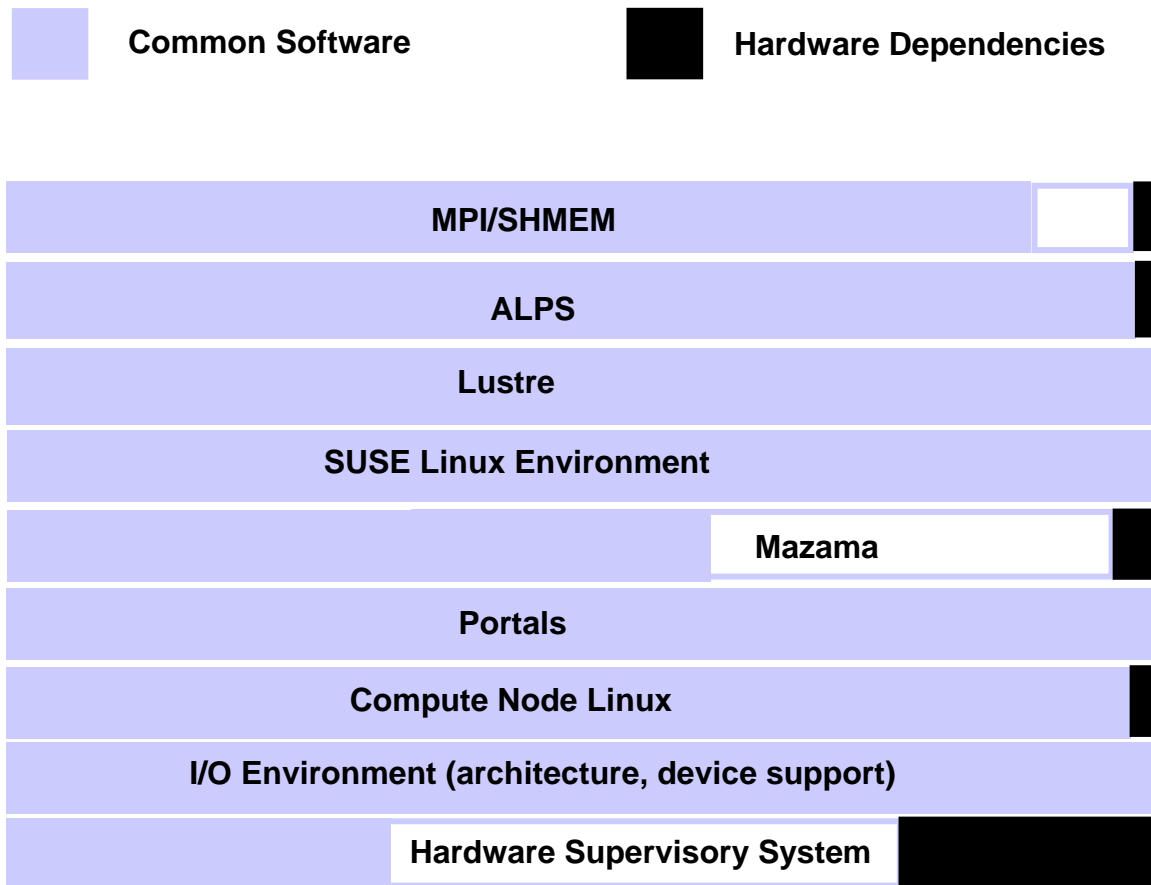
- Common software ported to both platforms
- Some differences because of machine dependent support



# BW and XT3/Hood Commonality



# BW and XT3/Hood Commonality



# Status

- Work on Common Software Repositories is underway
- Common Development Processes are also underway (differences have been reduced in past year)
- Common tools and integrated work schedules are being discussed