

Computación en Clúster

Bibliografía:

Rajkumar Buyya, “High Performance Cluster Computing”, vol. 1, Prentice Hall, 1999.

Objetivos

- Proporcionar una visión general de la arquitectura, los elementos hardware y software de un sistema clúster.

Contenidos

1. Introducción.
2. Arquitectura.
3. Tipos de clúster.
4. Componentes.
5. Middleware.
6. Manejo y planificación de recursos.
7. Recursos básicos de programación.

1. Introducción

“Hay tres maneras de mejorar el rendimiento:

Trabajar duro,

Trabajar más inteligentemente, y

Conseguir ayuda.”

G. Pfister, “In Search of Clusters”, 1998.

1. Introducción

- Un clúster. Es un tipo de arquitectura escalable para el procesamiento distribuido.
- Arquitecturas para el procesamiento paralelo
 - Procesadores para el cómputo paralelo masivo (MPP).
 - Multiprocesadores simétricos (SMP).
 - Sistemas de acceso no uniforme a memoria de caché coherente (CC-NUMA).
 - Sistemas distribuidos.
 - **Sistemas Clúster.**

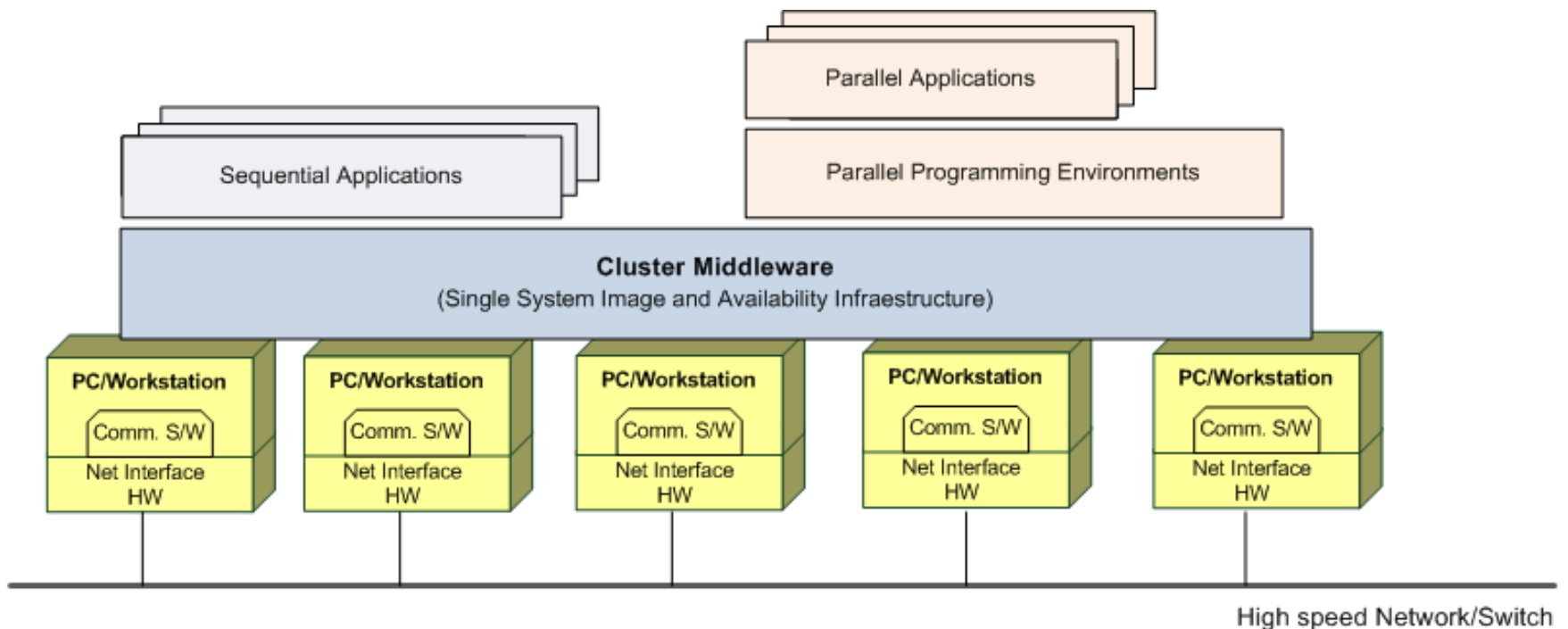
1. Introducción

- Sinónimos del término clúster:
 - Red de estaciones de trabajo.
 - Clúster de estaciones de trabajo.
- Principal motivación: supercomputación de bajo coste.
 - Basado en tecnología de uso popular, y por tanto atractiva desde el punto de vista de su coste.
 - Escalabilidad.

2. Arquitectura

- Elementos de un entorno clúster:
 - Nodos (computadores).
 - Sistemas operativos.
 - Infraestructura de red.
 - Protocolos de comunicación y servicios.
 - Middleware.
 - Entorno de desarrollo de aplicaciones.
 - Aplicaciones.

2. Arquitectura



Fuente: "High Performance Cluster Computing", vol. 1, Prentice Hall

3. Tipos de clúster

- En función del objetivo perseguido:
 - Alto rendimiento.
 - Alta disponibilidad.
 - Almacenamiento distribuido.
 - De balanceo de carga.
- En función de la configuración del nodo:
 - Homogéneo.
 - Heterogéneo.

3. Tipos de clúster

- En función del número de nodos:
 - Grupo: 2-29.
 - Departamentales: 10-100.
 - Corporativos: más de 100.
 - Nacionales: (agrupaciones de clúster de alguno de los tipos anteriores vía WAN/Internet).
 - Internacionales: (agrupaciones de clúster de alguno de los tipos anteriores, vía WAN/Internet, de más de 1000 nodos).

4. Componentes

- Nodos:
 - CPU.
 - Memoria.
 - Bus del sistema.
 - Subsistema de E/S.
- Interconexión:
 - Ethernet, Fast Ethernet y Gigabit Ethernet.
 - Asynchronous Transfer Mode (ATM).
 - Scalable Coherent Interface (SCI).
 - Myrinet
- Sistema operativo:
 - Linux.
 - Microsoft Windows.

5. Middleware

- Objetivo. Cumplir con el requerimiento de Imagen Única del sistema: *“a partir de un conjunto de nodos interconectados, proporcionar una imagen unificada de los recursos del sistema (SSI)”*.
 - *Transparencia.*
 - *Mejora de la disponibilidad.*
 - *Rendimiento escalable.*

5. Middleware

- Middleware es un módulo software estructurado en dos capas:
 - Infraestructura SSI.
 - Un único punto de entrada.
 - Jerarquía única de archivos.
 - Un único punto de control y gestión del clúster.
 - Una única red virtual.
 - Un único espacio de memoria.
 - Un único sistema de gestión de tareas.
 - Una única interfaz de usuario.
 - Infraestructura de disponibilidad del sistema.
 - Espacio de E/S único.
 - Un único espacio de procesos.
 - Mecanismo de salvaguardado de contexto.
 - Mecanismo de migración de procesos.

6. Manejo y planificación de recursos

- En un entorno clúster el manejo y la planificación de recursos (RMS) tiene por objeto hacer un uso eficiente de los recursos disponibles con el fin de cumplir los objetivos del clúster (alto rendimiento, alta disponibilidad, balanceo de carga, almacenamiento distribuido)
- ¿Cómo? Gestionando los procesos entre los nodos del clúster, con un mínimo impacto de cara a los usuarios. Dependiendo del objetivo perseguido, se requerirán alguna(s) de las siguientes funcionalidades:
 - Migración de procesos.
 - Monitorización de contextos.
 - Aprovechamiento de los ciclos ociosos.
 - Tolerancia a fallos.
 - Balanceado de carga.
 - Planificación mediante múltiples colas de proceso

7. Recursos básicos para la programación y administración

- Entornos de programación estándares para entornos distribuidos.
 - Uso de hilos.
 - IEEE POSIX threads interace (pthreads).
 - Sistemas de mensajes.
 - Parallel Virtual Machine (PVM).
 - Message Passing Interface (MPI).
 - Sistemas de memoria compartida y distribuida.
 - Sistemas software de memoria compartida.
 - TreadMarks.
 - Linda.
 - Sistemas hardware de memoria compartida.
 - DASH.
 - Merlin.

7. Recursos básicos para la programación y administración

- Depuradores
 - High Performance Debugging Forum (HPDF).
 - TotalView
 - C, C++,F77,F99,HPF.
 - MPI, PVM.
 - IBM AIX, SGI IRIX, SunOS, etc.

7. Recursos básicos para la programación y administración

- Herramientas de análisis de rendimiento.
 - Funciones para monitorizar el rendimiento.
 - Librería “run-time” .
 - Herramientas para procesar y visualizar los datos de rendimiento.

7. Recursos básicos para la programación y administración

- Administración
 - NOW.
 - SMILE (K-CAP).
 - PARMON (Grandes configuraciones).
 - **Orden pcs (Red Hat Enterprise High Availability Add-On).**
 - Servicio pcsd Web UI (Red Hat Enterprise High Availability Add-On).