

HPC: Aplicaciones y Software Heredado

HPC: Aplicaciones y Software

Fernando G. Tinetti

III-LIDI, Fac. de Informática, UNLP

Comisión de Inv. Científicas, Bs. As.

fernando@info.unlp.edu.ar



Aplicaciones y Software

- Aplicaciones: *muchísimas*
 - Diferentes áreas: clima, materiales, señales, etc.
 - Un gran porcentaje: modelos numéricos
 - Otras: heurísticas, análisis de datos (GD), etc.
- Software: implementaciones
 - De alguna manera *traducen...* al hardware
 - Un gran porcentaje: desarrollado por científicos del área de aplicación
 - Modelos numéricos: muy bien *caracterizados*

¿HPC?

- Si ya hay software... ¿para qué HPC?
 - *if it ain't broke, don't fix it*
- Velocidad: ¿para qué y cómo?
 - Más precisión (Aplicación)
 - Otros modelos
 - Más rápido (HPC)
 - Más paralelo
- Modelos numéricos: muy bien *caracterizados*
 - Simplificaciones y errores
 - ¿Rendimiento? ==> factibilidad, “rapidez”

¿Qué involucra HPC?

- Optimización
 - Bibliotecas: “... ah, eso es una inversión de matrices...”
 - Conocer arquitecturas (SSE, superescalares, caches, ...)
- Paralelización
 - Múltiples Núcleos
 - Clusters
 - Aceleradores /as
 - Todo junto... ¿otros?
- Modelos
 - De programación: mensajes, mem. compartida, ...
 - De procesamiento: SPMD, MPMD, ...

¿Cómo se llega a HPC?

- Algoritmo numérico
- Requerimientos
 - De procesamiento u operaciones: análisis de alg.,
ej: $2/3n^3 + 2n^2$ flops
 - De memoria: matrices, vectores, estr. dinámicas,
ej: $(2n + 1)^3$ memoria
 - ¿E/S?
- Lo anterior sobre HPC: optimización y paralelización

Más Complicado: *Herencia*

- Software en producción
- Métodos de desarrollo anticuados, múltiples cambios
- Documentación desactualizada o inexistente
 - “nosotros solamente lo usamos”
- ¿Cómo optimizar y paralelizar?
 - Software de la época del “free lunch”
- ¿Cómo estimar costos de opt. y paralelización?

Optimización y Paralelización

- Optimización
 - Perfilado, Bibliotecas, Contadores de hardware
 - Dependiente de la arquitectura
- Paralelización
 - Múltiples núcleos, Clusters
 - Aceleradores: GPUs, Xeon Phi
 - Bibliotecas de cómputo paralelo
 - Lenguajes de programación
 - Redes de interconexión, bibliotecas de comunic.

Proyectos

- “Propios” o “de base”
 - Propuesta algorítmica para álgebra lineal
 - Propuesta de optimización de comunicaciones
 - Impacto de las propuestas sobre HPL (*detalle*)
- Sobre software heredado
 - Transformación de código para *mejorar*
 - Caracterización con métricas
 - Lejos de paralelizar a gran escala aún
 - Aunque algo paralelo logramos...

Proyectos

- Sobre aplicaciones (no *heredadas...*)
 - Storm Water Management Model (SWMM)
 - Resolución de atmósferas estelares peculiares con ecuaciones diferenciales
 - Identificación de objetos transneptunianos a partir de imágenes (Complejo Astronómico El Leoncito)
 - Objetos transneptunianos: N-Body+Heliospheric Model

Proyectos: *breakthrough*...

- "...proyectos que puedan en un corto plazo utilizar recursos de cómputo masivamente paralelos con potenciales breakthroughs científicos"
- "El software fundamental del que hace uso su proyecto ha sido desarrollado en su grupo?"
- "Posee Ud. evidencia demostrable del escaleo de su aplicación a sistemas medianos tales como arreglos de estaciones de múltiples CPU conectadas por redes Infiniband?"
- Ej: alineación de secuencias de ADN en Xeon Phi:
 - 240 hilos, 60 cores
 - **OpenMP**: aprox. 4 horas
 - **OpenMP** + Vectorización SIMD con OpenMP: aprox. 10 minutos
 - **OpenMP** + Vectorización SIMD provista por icc: aprox. 5 minutos

Muchas gracias