

Uso de modelos PGAS para computación de Alto Rendimiento en clusters Beowulf

D'Elía, Dalcin, Sarraf, López, Battaglia, Rodríguez, Sonzogni

Centro de Investigación de Métodos Computacionales (CIMEC) (new name! ☺)

Univ. Nac. del Litoral (UNL), CONICET, Santa Fe

Grupo de Inves. en Mét. Numéricos en Ing. (GIMNI),

Univ. Tecn. Nac., Fac. Reg. Santa Fe

Dpto de Mecánica Aplicada, Fac. de Ing.

Univ. Nac. del Comahue, Neuquén

<http://www.cimec.org.ar/>

VI Latin American Symposium on High Performance Computing HPCLatAm 2013,
Mendoza, Argentina

Symposium: July 29-30, School: July 22 to July 26, 2013

CAI+D 2009-III-4-2; PIP 112-20111-00978; PICT 2492-10; PIRSES-246977



- Clusters: Coyote, Bora, Aquiles y un 4to en proceso de compra;
- Todos son de tipo Beowulf, de 64 bits, con arranque disk-less;
- Ganglia (Coyote y Aquiles):
<http://200.9.237.240:9000/ganglia>
- Gabinetes tower con ventilación reforzada en estanterías rodantes;
- UPS (estables si los fan control de las CPU están off);
- Salas con AA y con ventiladores para circulación forzada.



Cluster Coyote (porque nunca alcanzará a Road Runner...):

- 222 núcleos (33 nodos + 1 servidor);
- nodos 1-8: Xeon E5420 de 2.5 Ghz (2 x 4 núcleos);
- nodos 9-24: Xeon W3690 de 3.47 Ghz (1 x 6 núcleos);
- nodos 25-33: Xeon E5-1660 de 3.30 GHz (1 x 6 núcleos);
- 1 switch de Gigabit ethernet (1 Gbit/s) con 52 puertos;
- GNU-Linux, distribución Fedora 17;
- Adquirido con fondos de dos PICT y fondos propios de la UNL (valor aprox USD\$80,000).



Cluster Bora:

- Disponible desde 2013.
- Segregado, porque está orientado a software más especializado.
- Descripción:
 - 48 núcleos;
 - 7 nodos + 1 servidor;
 - Todos son procesadores I7 (☺), de 1 x 6 núcleos;
 - 1 switch de Gigabit ethernet (1 Gbit/s);
 - GNU–Linux, distribución Fedora 17;
 - Adquirido con fondos ?? (valor aprox USD?).



- (Viejo) **cluster Aquiles** disponible desde 2006, compartido por un consorcio entre 4 grupos de investigación de la región:
 - **Grupo de Física del Instituto de Desarrollo Tecnológico para la Industria Química (INTEC)**
 - **Instituto de Matemática Aplicada del Litoral (IMAL)**
 - **Centro de Meteorología de la Facultad de Ingeniería y Ciencias Hídricas (FICH) de la Universidad Nacional del Litoral (UNL)**
 - **CIMEC**: responsable de su administración.
- 82 nodos Pentium IV (*extended*), 3.0 GHz, 2 GB por nodo
- Red ethernet de 1 GBit/sec, con 2 switches 3COM de 48 puertos c/u, apilados por 1 stacking-cable de 10 Gbit/sec
- GNU-Linux, Fedora 9 (update demorado)
- Valor al momento de compra USD\$100,000.



- Lenguajes de programación basados en el modelo *Partitioned Global Address Space* (PGAS), otro modelo de programación paralela en computadoras con memoria distribuida.
- El modelo PGAS provee comunicación *one-side*, y es la base de, entre otros, los lenguajes de programación multiparadigma tales como UPC (*Unified Parallel C*) y Coarray Fortran (CAF o Fortran 2008).
- e.g. Intel tiene previsto extender los coarrays en sus compiladores Fortran para las Xeon-Phi.



- PGAS runtimes y librerías instaladas en Coyote:
 - *Global-Address Space Networking (GASNet)* runtimes
 - *Aggregate Remote Memory Copy Interface (ARMCI)*
 - Librería *Global Arrays (GA)*
 - Librería *UPCBLAS* (para cómputo paralelo matricial en UPC)
- Leng. de progr. basados en PGAS instalados en Coyote:
 - *Unified Parallel C (UPC)*: compiladores Berkeley UPC y GUPC;
 - *Coarray Fortran* (o *Fortran 2008*): compilador OpenUH.
- Los compiladores Berkeley UPC y Open-UH, así como la librería GA han sido instalados en el *cluster* Coyote, y estan en etapa experimental.
- Todos estos fueron construídos sobre las capas GASNet y ARMCI, además de la ubicua librería *Message Passing Interface (MPI)*, a través de la distribución OpenMPI.



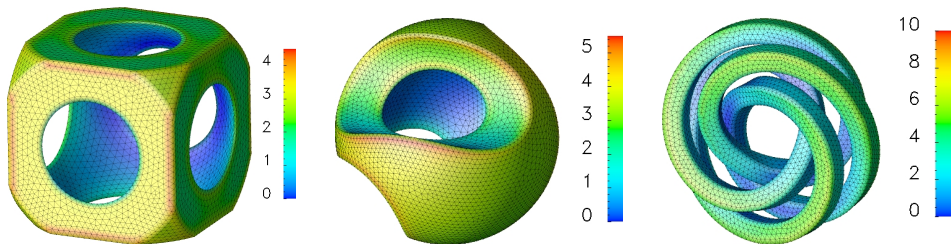


Figura: Flujo reptante por BEM, GBEM y HBEM alrededor de geometrías intrincadas o con aristas. *hollow cube* (izq.), *sculpted sphere* (cen.) (NETGEN), *square knots* (der.) Tracción adimensional τ_1 bajo flujo paralelo $\mathbf{U}_\infty = (10^{-3}, 0, 0)$ m/s.

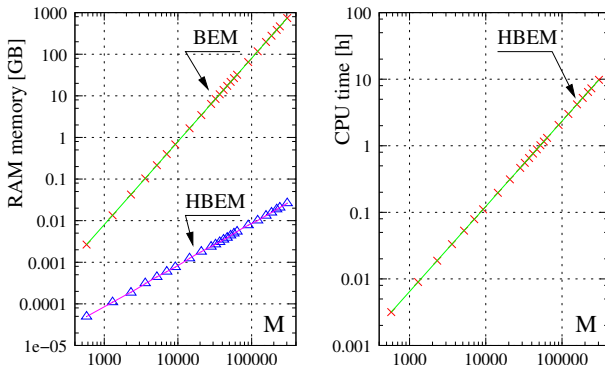


Figura: Costo computacional en función del número de grados de libertad M , en la esfera unitaria con refinamientos uniformes. Memoria RAM con BEM versus HBEM (izq.), y tiempo de CPU con HBEM usando GMRES (der.).