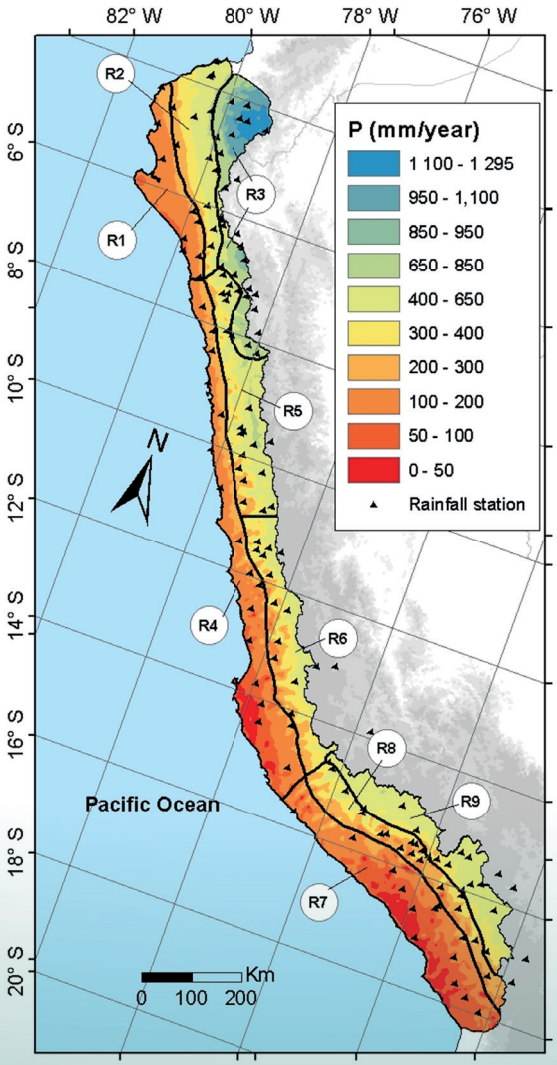
PERÚ Ministerio del Ambiente | Instituto Geofísico del Perú - IGP

Programa Presupuestal por Resultados Nº 68 “Reducción de vulnerabilidad y atención de emergencias por desastres”. Producto: “Estudios para la estimación del riesgo de desastres”

“Generación de información y monitoreo del Fenómeno El Niño”

Boletín Técnico

Sobre la distribución de las lluvias en la vertiente del Pacífico peruano y su relación con El Niño



PERÚ PROGRESO PARA TODOS

Vol. 4 Nº 1 Enero 2017

Análisis de Rendimiento del HPC-Linux-Clúster usando el método Benchmark WRF

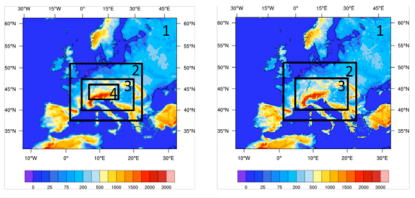
Berlín Segura, Ivonne Montes y Kobi Mosquera
Instituto Geofísico del Perú

El Instituto Geofísico del Perú (IGP) ha implementado una infraestructura computacional de última generación dedicada al cálculo científico intensivo denominada "Sistema Computacional de Alto Rendimiento para la Simulación de Fluidos Geofísicos" (en adelante HPC-Linux-Clúster), la cual es ahora una de las herramientas computacionales más grandes y poderosas a disposición de la comunidad científica peruana; adquirida gracias al Convenio de Subvención Nº101-2014-FONDECYT, a los proyectos de colaboración SPIRALES2012 IRD-IGP, Manglares IGP (IDRC) y al Programa Presupuestal por Resultados 068 (PPR).

Medir el rendimiento de este tipo de infraestructura requiere utilizar técnicas de *"benchmark"*, que se define en informática, como herramientas estandarizadas para evaluar el desempeño del sistema completo o de cada componente, medir los límites en la capacidad de transferencia en el sistema o entre conexiones, establecer comparativos con otros sistemas o arquitecturas computacionales, así como determinar o elegir los sistemas más adecuados para ciertas aplicaciones. El presente estudio propone analizar el rendimiento del HPC-Linux-Cluster mediante el *benchmark* de tipo aplicativo que consiste en utilizar un modelo numérico complejo, en este caso el modelo atmosférico regional WRF (Weather Research and Forecasting Model) y seguir la metodología de Arnold et al. (2011); la cual permite medir el rendimiento del WRF y probar la escalabilidad de diferentes plataformas. En este caso, el método además permite calcular el rendimiento (rapidez y eficiencia) del HPC-Linux-Cluster y compararlo con otras plataformas de Europa y USA.

Metodología Arnold et al. (2011): *Benchmark* aplicativo usando WRF

El método consiste en utilizar 3 configuraciones de múltiples dominios sobre el modelo WRF (desarrollado para la investigación y el pronóstico del tiempo) que abarcan la región de Europa (Fig. 1), implementados en otras cuatro diferentes plataformas computacionales a nivel internacional (Vienna Scientific Cluster – VSC, Pacific Area Climate Monitoring and Analysis Network – PACMAN, Arctic Region Supercomputing Center University of Alaska – CHUGACH, National Institute for Computational Sciences University of Tennessee – KRAKEN ; más detalles en la Tabla 1). La primera configuración, 4dbasic, consiste en configurar 4 dominios (21.6, 7.2, 2.4 y 0.8 km de resolución) anidados (One-way) con 40 niveles verticales. La segunda, 4basiclev, tiene la misma configuración que la primera pero con 63 niveles verticales. La tercera, 3dhrlev, se configura para 3 dominios (7.2, 2.4 y 0.8 Km) teniendo mayor resolución horizontal y vertical. Estas configuraciones presentan varios desafíos para el modelo WRF, entre ellas tenemos el coste computacional debido a la gran cantidad de puntos de grilla y la estabilidad del modelo.



Dominios	4dbasic Grilla(xyz)	4basiclev Grilla(xyz)	R(km)
1	196x167x40	196x167x63	21.6
2	274x217x40	274x217x63	7.2
3	592x355x40	592x355x63	2.4
4	1003x505x40	1003x505x63	0.8

Dominios	3dhrlev Grilla(xyz)	R(km)
1	585x495x63	7.2
2	823x652x63	2.4
3	1777x1066x63	0.8

Figura 1. Configuraciones de múltiples dominios: 4dbasic, 4basiclev y 3dhrlev para el benchmark usando el modelo WRF.

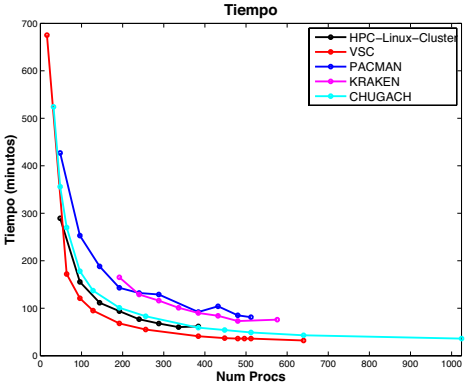
Siguiendo la metodología de Arnold et al. (2011) se realizaron 8 simulaciones por cada configuración, dando un total de 24 simulaciones; empleando de 2 a 16 nodos que alcanzan entre 48 y 384 procesadores. Para cada simulación se mide el tiempo de ejecución que lleva la simulación y se calcula las *medidas del rendimiento para aplicaciones paralelas*; las cuales son:

• Tiempo: es el tiempo de cálculo y el tiempo en crear los archivos de entrada y salida.

• Speed-up (*S*): expresada como ganancia de velocidad o rapidez, se define como la razón entre el tiempo de ejecución secuencial (*Ts*) con un procesador y el tiempo de ejecución paralelo (*Tp*) con múltiples procesadores.

• Eficiencia (*E*): es una medida del grado de aprovechamiento de los recursos computacionales. La eficiencia se define como la rapidez dividida por el número de procesadores (*p*).

• Escalabilidad: un sistema paralelo es escalable cuando tiene la capacidad de mantener la eficiencia con el aumento simultáneo de los procesadores y el tamaño del problema. La escalabilidad de un sistema paralelo es una medida de la capacidad de incrementar la rapidez con el número de procesadores, esto refleja la capacidad del sistema paralelo de utilizar eficazmente los recursos de procesamiento (Grama, 2003).



Análisis de Rendimiento del HPC-Linux-Clúster usando el método Benchmark WRF

Berlin Segura, Ivonne Montes y Kobi Mosquera
Instituto Geofísico del Perú

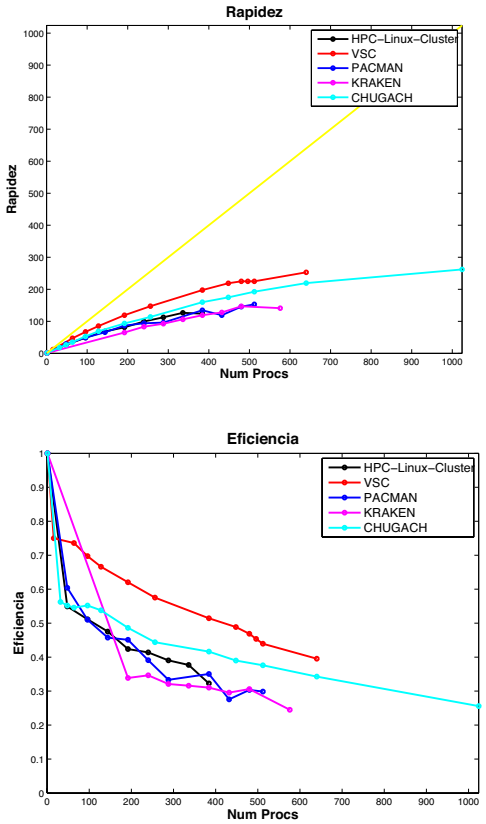


Figura 2: Tiempo (arriba), rapidez (centro) y eficiencia (abajo) para la configuración 4dbasic: plataformas HPC-Linux-Cluster (negro), VSC (rojo), CHUGACH (celeste), KRAKEN(magenta) y PACMAN(azul).

Resultados Preliminares

De la comparación entre las configuraciones sobre el HPC-Linux-Cluster se tiene que el menor tiempo fue alcanzado para el caso 4dbasic seguido por el 4basiclev y 3dhrlev, pero la diferencia entre las configuraciones es pequeña en la rapidez y la eficiencia.

Para la comparación del rendimiento del HPC-Linux-Cluster con las otras plataformas para las 3 configuraciones (4dbasic, 4basiclev y 3dhrlev) se obtuvo que, el menor tiempo, mayor rapidez y eficiencia fue alcanzado por VSC seguida por HPC-Linux-Cluster CHUGACH, KRAKEN y PACMAN. Aunque las plataformas son muy diferentes entre sí, se observa que el HPC-Linux-Cluster se aproxima a los resultados del CHUGACH (Fig. 2, mostrado solo para 4dbasic).

Un detalle observado es que el rendimiento del HPC-Linux-Cluster en las tres configuraciones se ve disminuido debido al tiempo que lleva en acceder a los archivos de entrada y salida (I/O ; lectura y escritura al disco). Por lo que, para aumentar la eficiencia se ha probado con cuatro métodos I/O: el primero crea archivos empleando por defecto un sólo procesador (I/O Netcdf en serie), en el segundo método los archivos son generados por cada procesador (I/O Netcdf en paralelo PNetcdf), en el tercero se definen los procesadores para el cálculo y para crear los archivos de entrada y salida (I/O Quilt) y, en el cuarto, se usa la combinación de Quilt con PNetcdf (I/O Quilt-PNetcdf); de los cuales se ha obtenido que el método I/O Pnetcdf alcanza el menor tiempo, mayor rapidez y eficiencia, luego le siguen los métodos I/O Quilt-Pnetcdf y I/O Quilt. Asimismo, al comparar los métodos I/O con 336CPU se tiene que el menor tiempo fue alcanzado por PNetcdf (33.88 min), luego tenemos a Quilt-PNetcdf, Quilt, Netcdf serie Intel y Netcdf serie GNU (60.25 min), donde el tiempo fue reducido a 26.37 min.

Como conclusión preliminar se puede afirmar que el HPC-Linux-Cluster tiene un desempeño comparable con otros sistemas similares a nivel internacional según el *benchmark* con el modelo WRF, el cual es escalable ya que la rapidez se incrementa con el número de procesadores, aunque el método escogido para el I/O tiene también un importante efecto. Esto indica que el HPC-Linux-Cluster es una herramienta computacional poderosa para apoyar las investigaciones científicas que se desarrollan en el IGP y la comunidad científica (ej., universidades).

Para acceder a nuestros recursos visite la web http://scah.igp.gob.pe/laboratorios/dfgc o escríbanos a dfgc@igp.gob.pe

Plataformas	Procesado	Conexión
HPC-linux-Cluster	Xeon 2.5GHz, 24 núcleos por Nodo, 128 GB de memoria, 5.3 GB por núcleo, 48-384 CPU	Infiniband QDR
VSCUniversity of Vienna	Sun Fire X2270, 8 núcleos por Nodo, 24 GB de memoria, 3 GB por núcleo, 16-640 CPU	Infiniband QDR
PACMAN at Arctic Region Su-percomputing Center (ARSC)University of Alaska	Penguin, Computing Cluster, 16 núcleos por No-do, 64 GB de memoria, 4 GB por núcleo, 48-512CPU	Mellanox QDR Infiniband
CHUGACH (ARSC)	Cray XE6, 16 núcleos por Nodo, 32 GB de memo-ria, 2 GB por núcleo, 32-1024CPU	Cray Gemini interconnect
KRAKEN The National Institute for Computational Sciences (NICS) University of Tennessee	Cray XT5, 12 núcleos por Nodo, 18 GB de memo-ria, 1.5 GB por núcleo, 192-576CPU	Cray SeaStar2+ interconnect

Tabla 1: Plataformas computacionales: HPC-Linux-cluster, VSC, PACMAN, CHU-GACH y KRAKEN

Referencias

Arnold D., D. Morton, I. Schicker, J. Zabloudil, O. Jorba, K. Harrison, G. Newby, and P. Seibert: "WRF benchmark for regional applications", in 12th Annual WRF Users' Workshop, Boulder, USA, 20-24 June, 2011.

Grama, Ananth, Anshul Gupta, George Karypis, and Vipin Kumar: "Introduction to Parallel Computing – 2nd Edition". Addison-Wesley, 2003.

Kaivalya M. Dixit: Overview of the SPEC Benchmarks. The Benchmark Handbook 1993

Skamarock, W. C., J. B. Klemp, J. Dudhia, D. O. Gill, D. M. Barker, M. Duda, X.-Y. Huang, W. Wang and J. G. Powers: "A Description of the Advanced Research WRF Version 3", NCAR Technical Note, 2008.

User-Oriented WRF Benchmarking Collection http://weather.arsc.edu/WRFBenchmarking/index.html

WRF V3 Parallel Benchmark Page http://www2.mmm.ucar.edu/wrf/WG2/bench/