

OPTIMIZACIÓN DE LA EJECUCIÓN DEL MODELO WRF (WEATHER RESEARCH AND FORECAST) RESPECTO AL NÚMERO DE NÚCLEOS UTILIZADOS EN EL CLÚSTER COMPUTACIONAL DE ALTO RENDIMIENTO DEL IMARPE

OPTIMIZATION OF EXECUTION OF THE WRF (WEATHER RESEARCH AND FORECAST) MODEL REGARDING THE NUMBER OF PROCESSORS USED IN IMARPE'S HIGH-PERFORMANCE COMPUTING (HPC) CLUSTER

Jorge Ramos-Flores¹ Adolfo Chamorro¹
 Carlos Quispe-Ccalluari¹ Jorge Tam¹

RESUMEN

RAMOS-FLORES, J., CHAMORRO, A., QUISPE-CALLUARI, C., TAM, J. (2022). Optimización de la ejecución del modelo WRF (Weather Research and Forecast) respecto al número de núcleos utilizados en el clúster computacional de alto rendimiento del IMARPE. *Inf Inst Mar Perú*. 49(2): 193-196.- El presente artículo tiene como propósito mostrar los resultados del trabajo realizado para la determinación del número de núcleos sobre los cuales la ejecución del modelo *Weather Research and Forecast* (WRF) encuentra su mejor tiempo de ejecución, para un dominio y resolución de grilla determinado, cuando es ejecutado en el clúster computacional de alto rendimiento del Instituto del Mar del Perú (IMARPE). Se encontró que el número óptimo fue de 192 núcleos usando el Clúster Computacional de Alto Rendimiento "Norte Humboldt" del LMOECC- IMARPE.

PALABRAS CLAVE: WRF, optimización de ejecución, clúster computacional

ABSTRACT

RAMOS-FLORES, J., CHAMORRO, A., QUISPE-CALLUARI, C., TAM, J. (2022). Optimization of the execution of the WRF (Weather Research and Forecast) model regarding the number of processors used in IMARPE's high-performance computing (HPC) cluster. *Inf Inst Mar Peru*. 49(2): 193-196.- Our aim is to present the results of the work done to determine the number of cores on which the Weather Research and Forecast (WRF) model reaches its best execution time, for a given domain and grid resolution, when it is run on the IMARPE's HPC cluster. We found that the optimal number was 192 cores using the "Norte Humboldt" HPC cluster of the IMARPE's Oceanographic, Ecosystem, and Climate Change Modeling Laboratory.

KEYWORDS: WRF, execution optimization, computational cluster

1. INTRODUCCIÓN

En el 2017 se instaló un Clúster Computacional de Alto Rendimiento (CCAR) en el Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC) del Instituto del Mar del Perú (IMARPE), que fue adquirido en el marco del Proyecto "Adaptación al cambio climático del Sector Pesquero y el Ecosistema Marino-Costero del Perú" financiado por el Banco Interamericano de Desarrollo (BID) (GUTIÉRREZ *et al.*, 2019). El CCAR fue denominado "Norte Humboldt" y está dedicado a ejecutar diversos modelos de simulación para el desarrollo de escenarios de cambio climático en el mar peruano.

Con el fin de optimizar la ejecución de los modelos en práctica en el LMOECC, es necesario conocer el número mínimo de núcleos con el que se obtiene la mayor velocidad de procesamiento

para un determinado dominio y resolución de grilla. Por tal motivo, el objetivo del presente trabajo fue optimizar la ejecución del modelo WRF determinando el número de núcleos necesarios y suficientes para la realización del mismo sobre el CCAR "Norte Humboldt" del LMOECC-IMARPE.

2. MATERIAL Y MÉTODOS

Modelo WRF

El modelo *Weather Research and Forecasting* (WRF) versión 3.3.1 (SKAMAROCK *et al.*, 2008), fue desarrollado en un proyecto de colaboración entre diversas entidades norteamericanas (NCAR-UCAR, 2020), como el Centro Nacional para la investigación Atmosférica (NCAR), la Administración Nacional Oceánica y Atmosférica (NOAA), los Centros Nacionales para Predicción Medioambiental

¹ Instituto del Mar del Perú (IMARPE), Dirección General de Investigaciones en Oceanografía y Cambio Climático (DGIOCC), jramos@imarpe.gob.pe

(NCEP), la Agencia de Meteorología de Fuerza Aérea estadounidense (AFWA), el Laboratorio de Investigación Naval (NRL), la Universidad de Oklahoma (OU), y la Administración de Aviación Federal (FAA). WRF es un modelo de mesoescala de última generación, que es utilizado tanto para la investigación como para pronóstico operativo, diseñado y escrito para funcionar mediante computación paralela en clústeres computacionales de alto rendimiento (CCAR). El código de WRF está escrito en Fortran90 y se puede compilar de manera serial, en paralelo (MPI) y en modo mixto (*OpenMP* y MPI), eligiendo la opción adecuada en el proceso de configuración.

En el LMOECC, el modelo WRF es utilizado para estudiar procesos de interacción océano-atmósfera y simular vientos a alta resolución espacial, necesarios para forzar los modelos oceánicos que simulan las condiciones del mar.

Clúster computacional de alto Rendimiento “Norte Humboldt”

El Clúster Computacional de Alto Rendimiento (CCAR, en inglés: *High Performance Computing Cluster, HPCC*) instalado en el LMOECC-IMARPE denominado “Norte Humboldt” cuenta con 12 nodos computacionales (servidores *Blade Dell*), cada nodo contiene 2 procesadores Intel Xeon E5-2680 de 2.50 GHz, con 12 núcleos por cada procesador, lo que resulta en 288 núcleos. El CCAR tiene una memoria RAM de 128 GB por servidor, lo que hace 1,536 TB de memoria principal, un dispositivo de almacenamiento de 400 TB y sus nodos computacionales están interconectados mediante conectividad *infiniband* FDR de 56 Gbps y alcanza una capacidad teórica pico de ~12 TFLOPS (Fig. 1).

Configuración del modelo WRF

El dominio del modelo se extendió entre 100°W - 59°W y entre 10°N - 30°S, con una resolución horizontal de 0,25° y con 174x165 puntos de grilla (Fig. 2). Se usó la configuración del modelo desarrollado por CHAMORRO *et al.* (2018) para el Sistema de Afloramiento Peruano.

El modelo WRF fue compilado utilizando los compiladores *icc* e *ifort* de Intel versión 17.04 con la opción de memoria distribuida (*dmpar*). La librería *netCDF* con la que el modelo fue



Figura 1.- Clúster computacional de alto rendimiento implementado en el Laboratorio de Modelado Oceanográfico, Ecosistémico y del Cambio Climático (LMOECC)

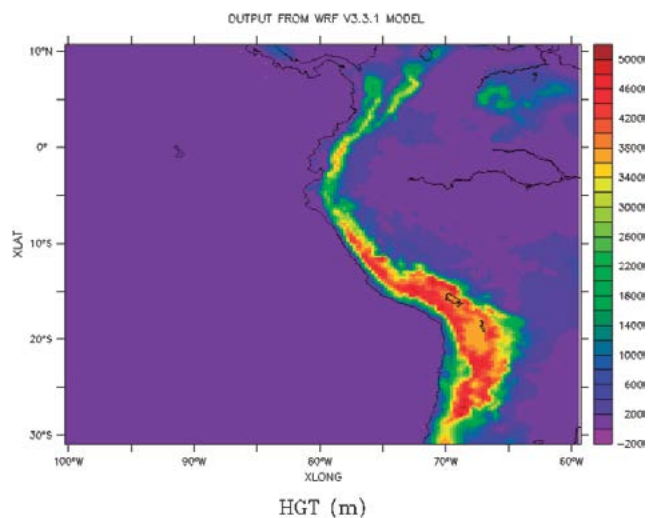


Figura 2.- Dominio del modelo WRF usado para evaluar la eficiencia del CCAR

enlazado fue un compendio clásico realizado también con los compiladores de Intel *icc*, *icpc*, *ifort* (UCP, 2019) indicando el *flag disable-netcdf-4*. En la paralelización, el dominio geográfico del modelo se descompone en “n” subdominios rectangulares de igual tamaño, tal que “n” es el número de núcleos disponibles para procesamiento. Por defecto, el número de subdominios por longitud es aproximadamente igual al número de subdominios por latitud.

3. RESULTADOS

Determinación del número óptimo de núcleos

Para optimizar la aplicación del modelo WRF en el CCAR del LMOECC-IMARPE, se realizó una serie de ejecuciones del modelo, cada una de 1 mes de simulación, y se programó un *script* en *bash* para medir con exactitud cuánto tiempo demora cada simulación, cambiando el número de núcleos (de 24 a 288, cada 24 núcleos). En la figura 3 se presenta el tiempo de ejecución de 1 mes de simulación vs. el número de núcleos activos en el CCAR. Los datos se ajustan a un modelo potencial ($r = 0,988$, $p < 0,001$):

$$TE = 5568,3 NC^{-0,911}$$

Donde: TE: tiempo de ejecución del modelo WRF, NC: número de núcleos.

El modelo indica que el tiempo de ejecución disminuye rápidamente hasta llegar a los 144 núcleos, después de este punto el tiempo disminuye muy poco logrando ser casi estable entre los 192 y 216, luego vuelve a disminuir ligeramente hasta los 240 y finalmente se podría decir que permanece estable desde este punto hasta los 288 núcleos.

Por tales motivos, se considera que la ejecución óptima del clúster computacional de alto rendimiento se obtiene con 192 núcleos. El uso de un mayor número de

núcleos no generaría una reducción sustancial del tiempo de ejecución, por lo que no sería aconsejable tomar más de 192 núcleos si se quiere lograr una óptima utilización de los recursos computacionales y así poder utilizar los núcleos restantes para otros modelos (e.g. ROMS-PISCES) que podrían ejecutarse de manera simultánea.

4. CONCLUSIONES Y RECOMENDACIONES

El número óptimo para realizar simulaciones con el modelo WRF fue de 192 núcleos usando el Clúster Computacional de Alto Rendimiento "Norte Humboldt" del LMOECC-IMARPE, para un dominio entre 100°W – 59°W y entre 10°N – 30°S y una resolución espacial de 0,25°.

Se recomienda realizar pruebas adicionales para evaluar la eficiencia con respecto a la forma de descomposición del dominio WRF (número de subdominios por longitud y latitud), para así determinar si es posible bajar los tiempos de ejecución obtenidos con la descomposición automática del dominio.

También sería interesante, en evaluaciones futuras, comparar opciones de compilación que utilicen el modo mixto versus el MPI directo, así como ajustar el uso de caché en tiempo de ejecución y las estrategias de E/S óptimas (PORTER & ASHWORTH, 2010).

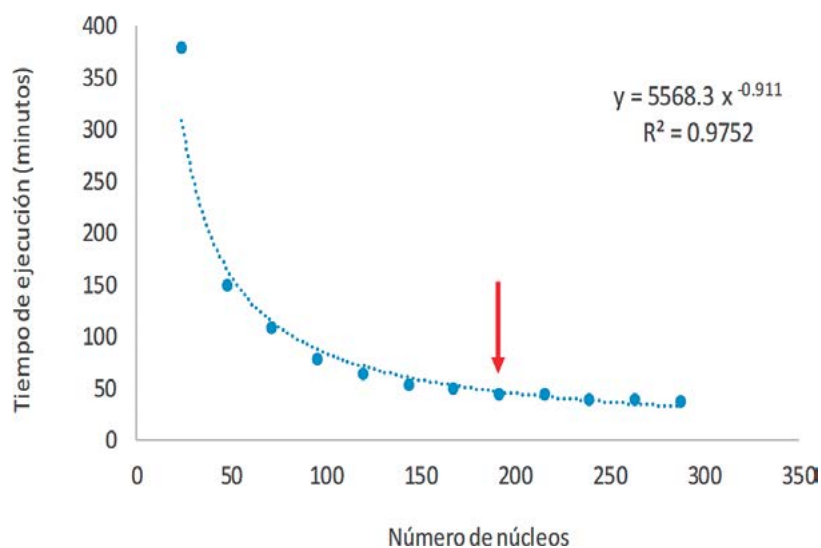


Figura 3.- Tiempo de ejecución del modelo WRF versus número de núcleos utilizados, indicando el número óptimo de núcleos (flecha roja)

Asimismo, se recomienda realizar pruebas similares para optimizar la ejecución del modelo oceánico ROMS o los modelos acoplados ROMS-PISCES, ROMS-WRF, y así poder determinar la cantidad de núcleos adecuados a utilizar con cada uno de ellos.

5. REFERENCIAS

- CHAMORRO, A., ECHEVIN, V., COLAS, F. *et al.* (2018). Mechanisms of the intensification of the upwelling-favorable winds during El Niño 1997–1998 in the Peruvian upwelling system. *Clim Dyn* 51, 3717–3733. <https://doi.org/10.1007/s00382-018-4106-6>
- GUTIÉRREZ, D., TAM, J., REGUERO, B. G., RAMOS CASTILLEJOS, J., OLIVEROS, R., CHAMORRO, A., GÉVAUDAN, M., ESPINOZA, D., COLAS, F., ECHEVIN, V., CORREA, D., DOMÍNGUEZ, N., ZAVALA, R., GONZALES, N., RAMOS, J., GRADOS, D., ROMERO, C. Y. (2019). Fortalecimiento del conocimiento actual sobre los impactos del cambio climático en la pesquería peruana. In: Zavala, R. *et al.* (eds.). *Avances del Perú en la adaptación al cambio climático del sector pesquero y del ecosistema marino-costero*. Monografía del BID, Serie IDB-MG-679, Lima, 125 p.
- NCAR-UCAR. (2020). The Weather Research and Forecasting Model. En línea: <https://www.mmm.ucar.edu/weather-research-and-forecasting-model>.
- PORTER, A., ASHWORTH, M. (2010). Configuring and Optimizing the Weather Research and Forecast Model On the Cray XT.
- SKAMAROCK, W. C., KLEMP, J. B., DUDHIA, J., GILL, D. O., BARKER, D. M., DUDA, M. G., HUANG, X. Y., WANG, W., POWERS, J. G. (2008). A Description of the Advanced Research WRF Version 3. NCAR Technical Note, NCAR/TN-475+STR. Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder.
- UCP. (2019). Network Common Data Form (NetCDF). En línea: <https://www.unidata.ucar.edu/software/netcdf/>