
SIMULACIÓN DE LA PROTEÍNA SPIKE COVID19 EN LABORATORIOS DE INFORMÁTICA EDUCATIVA

*Simulation of SPIKE COVID19 protein in educational computer
laboratories*

*Simulação da proteína SPIKE COVID19 em laboratórios de
computadores educacionais*

Erwin Blanco¹, Fabiola Sáez-Delgado¹ & Lorena Blanco¹

¹ Facultad de Educación. Universidad Católica de la Santísima Concepción.
Concepción-Chile. Correo: eblanco@magisteredu.ucsc.cl, fsaez@ucsc.cl,
lblanco@ucsc.cl

Fecha de recepción: 13 de junio de 2020

Fecha de aceptación: 30 de junio de 2020

RESUMEN

INTRODUCCIÓN: Las proteínas son una de las moléculas orgánicas que cumplen funciones vitales en la mantención de la vida y reproducción celular, su fabricación es un proceso complejo gobernado por una secuencia de plegamientos aún desconocida. En el contexto del Covid19, la iniciativa Folding@home llevó a cabo un proyecto de computación distribuida que permite la simulación del proceso de plegamiento de la proteína Spike del Covid19, cuya función es acoplarse al receptor ACE2 de las células animales y así penetrar a la célula y utilizar su maquinaria para reproducirse. **OBJETIVO:** Probar que el tiempo computacional ocioso disponible en los laboratorios de informática educacionales, puede ser usado para la simulación del plegamiento de proteínas. **MÉTODO:** Esta fue una investigación descriptiva donde se instaló un software cliente en computadores de gama baja que continuamente enviaron unidades de trabajo de plegamiento proteico a un servidor central. **RESULTADOS:** Tras 90 días de trabajo, un clúster de 27 PCs finalizaron 1993 unidades de trabajo de simulación de la proteína Spike. **DISCUSIÓN Y CONCLUSIONES:** A pesar de que el piloteo fue un éxito, se advierte que el software cliente debe ser optimizado para sacar el máximo de provecho a los diferentes procesadores y sistemas operativos con los cuales es compatible el software de computación distribuida proveído por Folding@home.

Palabras claves: Computación distribuida, Folding@home, plegamiento de proteínas, Linux, proteína Spike Covid19, informática educativa.

ABSTRACT

INTRODUCTION: Proteins are one of the organic molecules that have vital functions in the maintenance of cell life and reproduction, their manufacture is a complex process governed by a sequence of folding still unknown. In the context of Covid19, the Folding@home initiative carried out a distributed computing project that allows the simulation of the folding process of the Spike protein of Covid19, whose function is to couple to the ACE2 receptor of animal cells and thus penetrate the cell and use its machinery to reproduce. **OBJECTIVE:** To prove that the idle computational time available in educational computer labs can be used to simulate protein folding. **METHOD:** This is descriptive research where client software is installed on low-end computers that continuously send protein folding work units to a central server. **RESULTS:** After 90 days of work, a cluster of 27 PCs completed 1993 units of simulation of the Spike protein. **DISCUSSION AND CONCLUSIONS:** Although the piloting was successful, the client software must be optimized to take full advantage of the different processors and operating systems with which the distributed computing software provided by Folding@home is compatible.

Keywords: Distributed computing, Folding@home, protein folding, Linux, Spike Covid19 protein, educational computing.

RESUMO

INTRODUÇÃO: As proteínas são uma das moléculas orgânicas que têm funções vitais na manutenção da vida e reprodução celular, sendo sua fabricação um processo complexo, governado por uma sequência de dobragem ainda desconhecida. No contexto da Covid19, a iniciativa Folding@home realizou um projeto de computação distribuída que permite a simulação do processo de dobragem da proteína Spike da Covid19, cuja função é acoplar ao receptor ACE2 das células animais e, assim, penetrar na célula e use suas máquinas para se reproduzir. **OBJETIVO:** Provar que o tempo de inatividade disponível em laboratórios de informática educacional pode ser usado para simular o dobramento de proteínas. **MÉTODO:** Esta é uma pesquisa descritiva, na qual o software cliente é instalado em computadores de baixo custo que enviam continuamente unidades de trabalho de dobra de proteínas para um servidor central. **RESULTADOS:** Após 90 dias de trabalho, um cluster de 27 PCs completou 1993 unidades de simulação da proteína Spike. **DISCUSSÃO E CONCLUSÕES:** Embora o teste tenha sido bem-sucedido, o software cliente deve ser otimizado para aproveitar ao máximo os diferentes processadores e sistemas operacionais com os quais o software de computação distribuído fornecido pelo Folding@home é compatível.

Palavras-chave: Computação distribuída, Folding@home, dobragem de proteínas, Linux, proteína Spike Covid19, computação educacional.

INTRODUCCIÓN

La maquinaria biológica base de todos los seres vivos es la célula, en sus variantes vegetal y animal. Su funcionamiento está basado en la interacción de varios sistemas que soportan funciones muy complejas, entre ellas una enzima llamada RNA polimerasa que es una proteína encargada de cortar ADN en 2 partes y así crear una hebra de ADN mensajero. Otra estructura clave de la célula es el ribosoma, la función del ribosoma es leer la secuencia de instrucciones que porta el ADN mensajero, con estas instrucciones se crearan incontables nuevas proteínas a través del ensamblaje

de 20 aminoácidos [1]. Los aminoácidos son moléculas orgánicas del grupo amino y grupo carboxilo, es decir son compuestas de carbono, hidrógeno, oxígeno y nitrógeno.

La creación de nuevas proteínas es un proceso intrincado que no termina instantáneamente tras el ensamblaje de aminoácidos en el ribosoma, una vez encadenados una serie de aminoácidos claves, estos comienzan a plegarse en un orden regular aún desconocido [2] hasta alcanzar una estructura tridimensional final que les permitirá cumplir la función para la cual fueron codificados según la secuencia de ADN de cada especie.

Las proteínas son clave para el sostén de la vida, entre las funciones de las proteínas podemos citar las siguientes: (a) de catálisis, es decir iniciar y controlar reacciones químicas; (b) reguladoras, por ejemplo, la proteína insulina regula la glucosa en sangre; (c) de defensa, las inmunoglobulinas son proteínas que sirven de anticuerpos de primera línea [3]. Las funciones de las proteínas son tan variadas y extensas que la vida, tal como la conocemos, sería imposible sin ellas.

Tal como se mencionó anteriormente, la creación de nuevas proteínas es el resultado de la combinación de aminoácidos encadenados que se pliegan de forma regular hasta un estado tridimensional final que les permita ser funcionales, este proceso de plegamiento no está exento de fallas, durante el plegamiento se producen aberraciones que evitan que una proteína llegue a su estadio final, lo que desemboca en la producción de una proteína no terminada que será disfuncional a sus propósitos. Estas fallas de ser constantes conllevan a resultados biológicos desastrosos en los seres vivos, tales como el Alzheimer [4], la enfermedad de Huntington [5], inmunodeficiencias, parkinson, fibrosis quística entre muchas otras.

El plegamiento de las proteínas sigue una serie de algoritmos o pasos regulares para llegar a su estadio final correcto. Por ahora, se desconoce la serie de movimientos regulares que se ejecutan entre los aminoácidos desde el inicio de la concatenación hasta la producción de una proteína funcional. Son tantas las combinaciones posibles entre aminoácidos y tantas las posibilidades de plegamiento, que no se podría crear un modelo matemático simple que pueda predecir la función y estructura final de una proteína a partir de una combinación de aminoácidos específicos.

En la actualidad, existe un proyecto de computación distribuida llamado Folding@home, cuyo ánimo es simular las dinámicas de las proteínas, incluyendo el proceso de plegamiento proteico y los movimientos de las proteínas implicadas en una variedad de enfermedades tales como el cáncer y enfermedades degenerativas [6]. Debido a que las proteínas son una base fundamental para la vida, su estudio también trae beneficios para el tratamiento de enfermedades infecciosas puesto que los virus y bacterias también basan su funcionamiento y ataque en proteínas, es esta la razón por la cual el proyecto Folding@home invierte parte de su simulación computacional en la comprensión de proteínas específicas que, por ejemplo, permitan generar tratamientos específicos para bloquear la producción o funcionamiento de proteínas virales.

En el contexto de la pandemia por la infección por Coronavirus SARS- CoV-2, el proyecto Folding@home ha volcado gran parte de sus recursos computacionales a simular la estructura tridimensional de la glicoproteína Spike [7], esta proteína que se encuentra en la superficie del Covid19, tiene por función acoplarse a la enzima ACE2 que es una proteína que se encuentra en la superficie de las células humanas. El acoplamiento de la proteína Spike a la ACE2 permite que esta última actúe como una

cerradura que se abre [8], lo que desemboca en que el Covid19 acceda a la célula humana, comience la replicación de nuevos virus y finalmente gatille una infección general, multiorgánica con posibilidad de deceso.

La comprensión de cómo la proteína Spike se abre y se acopla a la proteína ACE2 podría ser extremadamente útil. Cada paso que ejecuta Spike podría permitir potencialmente encontrar un punto débil que permita atacar al virus de forma terapéutica [9]. Desafortunadamente, a la fecha, no hay forma de observar las transiciones que sufren las proteínas Spike durante su acoplamiento, al menos no con las técnicas experimentales existentes. Para superar esta limitación concreta, Folding@home está utilizando simulaciones computacionales que permitan construir un modelo confiable de la estructura tridimensional y los pasos que utiliza la proteína Spike para acoplarse a ACE2 (Figura 1).

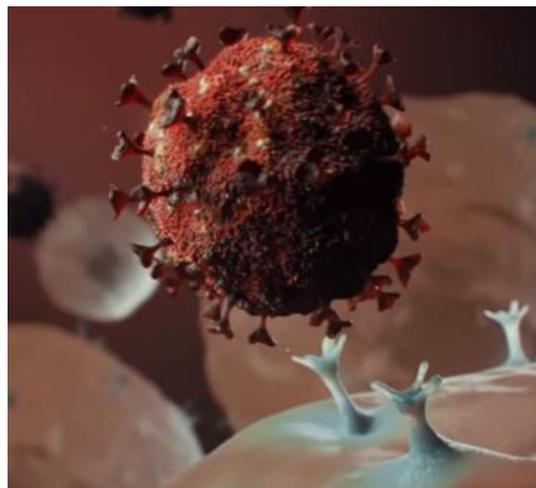


Figura 1. Virus SARS-CoV-2 intentando acoplar una proteína Spike a la proteína ACE2 lo que le permitirá cruzar la membrana celular y comenzar la infección.

Fuente: <https://bit.ly/3e9wP7Q>

La simulación de proteínas mediante modelos computacionales es un proceso costosísimo por la descomunal cantidad de operaciones lógico-matemáticas que demanda a los procesadores, esto implica que, para obtener resultados en ventanas de tiempo de aceptables, se necesita correr esas simulaciones en supercomputadoras con enormes capacidades de cálculo, de otra forma, las simulaciones tomarían cientos de años en completarse [7]. A la fecha, presionados por la pandemia virulenta del SARS-CoV-2, el proyecto Folding@home ha recurrido a la estrategia de aumentar su capacidad de cómputo gracias a un enfoque de computación distribuida que consiste en liberar un software cliente para que cada computador personal o servidor pueda simular segmentos de la estructura proteica de Spike o la dinámica de su apertura y ensamble. Una vez que un computador ha descargado, procesado y terminado su carga de trabajo, envía sus resultados a los servidores centrales de Folding@home quienes ensamblan los resultados de cientos de miles de computadoras en un modelo final que sea usable gratuitamente por investigadores [10], universidades o la industria farmacéutica para crear tratamientos farmacológicos que bloqueen la proteína Spike e inhiban su acoplamiento a ACE2.

Blanco, Sáez & Blanco. Simulación de la proteína Spike COVID19 en laboratorios.

Número Especial Desafíos Humanos ante el COVID-19

Enero – Junio. 2020

<http://dx.doi.org/10.33210/ca.v9i2.326>



Debido a que Folding@home es un proyecto sin fines de lucro, los tiempos computacionales que utilizan son donados voluntariamente por miles de usuarios alrededor del mundo. Bajo este paradigma, el Departamento de Educación de la Comuna de Ránquil, Chile, decidió unirse a esta iniciativa y donar tiempo computacional ocioso disponible en los laboratorios de informática escolar de tres escuelas bajo su jurisdicción. Los laboratorios se encuentran sin uso debido al cierre de escuelas por la pandemia, sin embargo, permanecen encendidos 24 horas al día simulando proteínas en unidades de trabajo parceladas que son enviadas de vuelta día a día al proyecto Folding@home para su divulgación.

El objetivo de esta investigación fue pilotear, a baja escala, pero a largo plazo, un proyecto de computación distribuida orientada a la salud a través del uso de tiempo computacional ocioso en contextos educacionales. Los resultados que se obtuvieron permitirán a profesionales de la educación, instituciones públicas, privadas y laboratorios de investigación, instalar softwares de computación distribuida orientados a la salud o a cualquier otra disciplina del conocimiento cuyo funcionamiento se base en esta tecnología.

MÉTODO

Materiales

La preparación previa para la simulación de proteínas del SARS-CoV-2, consistió en elegir los materiales necesarios para llevar a cabo este proceso. Específicamente, los materiales fueron los siguientes: 3 laboratorios de informática educativa pertenecientes a 3 escuelas rurales de la comuna de Ránquil, Región de Ñuble, Chile. Los laboratorios contaban con internet de banda ancha. Los equipos seleccionados fueron 27 computadores de escritorio de al menos 2 núcleos (64 bits) y 2 GB de memoria RAM. De los 27 equipos, 24 equipos corrían Linux Mint 19.X y 3 equipos corrían Windows 7. Se descargaron los softwares cliente y control de 64 bits desde la página oficial del proyecto Folding@home. Los informes de la cantidad de datos procesados por cada computador se apilaron en la página <https://bit.ly/2Z72QJr>. El identificador único como grupo de trabajo fue el 242919. Cada unidad de trabajo completada satisfactoriamente por cada equipo se anotó de forma acumulativa en el sitio web ya mencionado.

Consideraciones éticas

Respecto de las consideraciones éticas, en este estudio no se trabajó con seres humanos, sino con datos simulados en computadoras y dispositivos de propiedad estatal, por tanto, se solicitó a la máxima autoridad de Educación, correspondiente al Director de Administración Educacional Municipal, la autorización para desarrollar este estudio, asimismo, se solicitó autorización de los directores de los respectivos establecimientos. Por otra parte, se contó con la autorización y cuenta de servicio disponible proveída por Folding@home para registrar índices de cargas de trabajo realizadas por los equipos utilizados en este investigación.

Procedimiento

El primer paso antes de comenzar la donación de tiempo computacional, consistió en solicitar autorización al jefe del Departamento de Educación y los directores de cada establecimiento, quienes aprobaron en comenzar a ejecutar simulaciones en 3 escuelas de su jurisdicción a modo de pilotaje. El segundo paso

Blanco, Sáez & Blanco. Simulación de la proteína Spike COVID19 en laboratorios.

Número Especial Desafíos Humanos ante el COVID-19

Enero – Junio. 2020

<http://dx.doi.org/10.33210/ca.v9i2.326>



Compartir



Compartir

consistió en catastrar el estado de software y hardware de los equipos a utilizar, el tercer paso consistió en reinstalar Linux en aquellos computadores que presentaron problemas de sistema operativo, del total de 27 computadores, 15 computadores presentaban fallas de sistema operativo Windows 7 y 10, lo que obligó a instalar Linux en ellos. Una vez los equipos estaban en condiciones operativas, se procedió a instalar el software cliente de nombre fahclient_7.6.13_amd64.deb que permite descargar cargas de trabajo y llevar a cabo simulaciones de proteínas. Durante su instalación en la plataforma Windows, fue necesario activar el uso de redes públicas y privadas para dejar que el software pudiera descargar y enviar datos sobre Internet sin ser bloqueado por el cortafuegos. El segundo paso consistió en instalar el software de control para el cliente de nombre fahcontrol_7.6.13-1_all.deb (Figura 2).



Figura 2. Opciones de aplicaciones disponibles para la simulación de plegamiento proteico disponibles para GNU-Linux.

Fuente: <https://bit.ly/2ZKSrST>

Una vez finalizada la instalación, se procedió a dar ajustes finos a través del software de control (Figura 3), tales ajustes consistieron en establecer el poder de plegamiento (la opción bajo medio y alto están disponibles) y permitir la simulación sólo en ausencia de trabajo por parte de los usuarios, es decir, cuando nadie hiciera uso de los equipos. Finalmente, se les asignó a todos los computadores el identificador 242919 más un nombre de equipo propio, este identificador permitió revisar en un sitio web los avances de todos los equipos apilados en un solo equipo de trabajo.

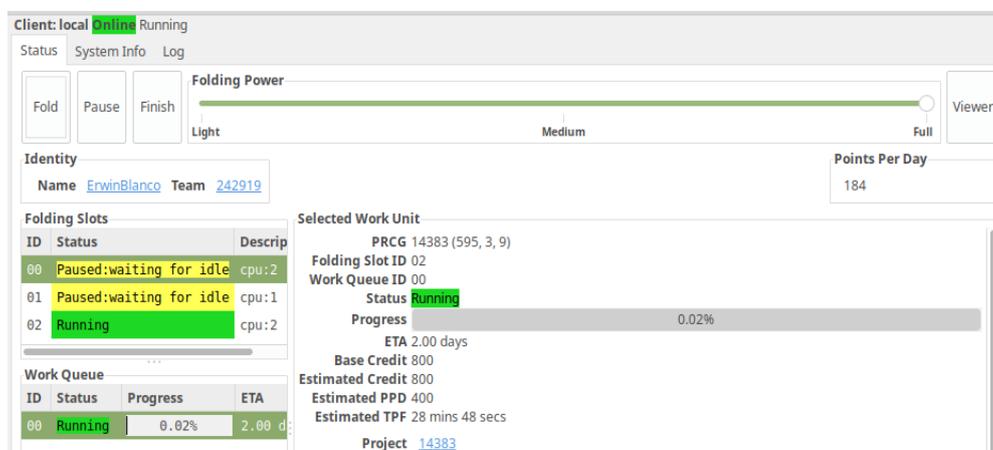


Figura 3. Ajustes finos disponibles para el cliente Folding@home.

Fuente: Interfaz de ajustes finos disponibles en el panel de control de Folding@home.

Después de varios días de simulación ininterrumpida, nos percatamos que los computadores de una escuela dejaron de enviar informes, el problema radicó en que al haber cortes de energía eléctrica, el equipo asumía que no debía encenderse nuevamente al volver el suministro eléctrico. Tal situación fue remediada en prácticamente todos los equipos, para ello se accedió a la BIOS y se configuró el comportamiento del equipo como After Power Loss=ON, instrucción que un PC interpreta como encendido una vez que vuelva el suministro eléctrico. Solo 1 computador de los 27 no poseía tal capacidad.

Los equipos fueron monitoreados continuamente a través del sitio web <https://bit.ly/2Ced7dV>. Este sitio web proveído por Folding@home permitió monitorear los avances de cada equipo día a día en cuanto a su rendimiento individual y colectivo sobre las simulaciones de la proteína Spike.

Análisis de datos

Los datos de doblamiento proteico se extrajeron desde la página web creada para el grupo de trabajo 242919. Una vez volcados los datos en una matriz CSV, se realizó un análisis descriptivo que incluyó la obtención de medidas de tendencia central y dispersión. Este estudio se llevó a cabo a través del análisis de datos en software estadístico RStudio versión 4.0.

RESULTADOS

Los equipos encendidos y configurados comenzaron a reportar unidades de trabajo completas en menos de 12 horas desde el comienzo de la simulación de proteínas. A la fecha del 6 de Julio del 2020, la mayoría de los equipos siguen trabajando de forma ininterrumpida. En la figura 4, se muestra un ejemplo de los reportes de cada computador apilados como un grupo único.



Figura 4. Sitio web proveído por Folding@home para apilamiento de reportes de progreso sobre unidades de trabajo asignadas a cada computador corriendo el software cliente.

Fuente: <https://bit.ly/2Ce2mbk>

Considerando la fecha de cierre de esta investigación y sobre un periodo de 3 meses, se han completado 1993 unidades de trabajo (Tabla 1.), con un promedio de 73 unidades de trabajo por cada equipo y una desviación estándar de 59, La unidad mínima de trabajo es 11 unidades y la unidad máxima fueron 190 unidades de trabajo. Cada unidad de trabajo se refiere a los datos enviados por el servidor de Folding@home a un cliente, que fueron modelados en un computador y que han sido satisfactoriamente enviados de vuelta al servidor.

Tabla 1. Estadígrafos descriptivos del rendimiento grupal del clúster de cómputo educacional en la simulación de la proteína Spike

	Unidades de trabajo
Validos	27
Perdidos	0
Media	73.81
Mediana	48.00
Moda	^a 11.00
Desviación estándar	59.23
Varianza	3509.15
Simetría	0.81
Desviación estándar de simetría	0.44
Test de normalidad Shapiro Wilk	0.864
Valor P de Shapiro Wilk	0.002
Mínimo	11.00
Máximo	190.00
Suma	1993.00

^a Existe más de una moda, pero solo 1 es reportada

DISCUSIÓN Y CONCLUSIONES

Las cargas de trabajo de simulación de proteínas ejecutadas durante el periodo de tres meses alcanzaron las 1993 unidades. Se puede advertir que algunos equipos procesaron casi 3 veces más que la media ($X=48$), esta situación se debe a la gran diferencia de procesamiento que existe entre equipos. Algunos equipos, los más modernos, contaban con procesadores de 4 núcleos, mientras la mayoría de los equipos contaba con procesadores de dos núcleos y 10 años de antigüedad. La desviación estándar de cargas de trabajo fue de $DE=59$, lo que incluso supera la media de $X=48$. Una situación no prevista se suscitó al advertir que equipos con las mismas características, misma marca y sistema operativo tuvieron rendimientos

diversos, esto se explica en el hecho documentado que los procesadores en su fabricación sufren aberraciones moleculares en su estructura de superconductores y transistores [11]. Otra situación no prevista se produjo cerca del cierre de esta investigación, al revisar los reportes de rendimiento, un laboratorio completo dejó de enviar datos al servidor de Folding@home, al asistir al laboratorio se constató que los equipos seguían encendidos y esperando conexión para el envío de datos, el problema resultó ser un switch de red que dejó de funcionar debido a la falta de mantención preventiva que experimentan los laboratorios computacionales escolares.

Esta experiencia de uso de tiempo computacional educacional ocioso para la simulación de plegamiento molecular es única a nivel país, puesto que no se han documentado ni publicado otras experiencias similares al menos en contextos educacionales en Chile. A diferencia de nuestro país, Brasil ha puesto a disposición su poder computacional público a través de la donación de tiempo computacional de los servidores de la petrolera Petrobras [12], lo que demuestra que grandes capacidades de cómputo públicas y privadas se encuentran ociosas y que podrían ser utilizadas para resolver problemas complejos.

La simulación del plegamiento proteico a través de modelos matemático-computacionales es una tecnología de punta que se encuentra aún en desarrollo y que ya ha tenido frutos publicados en revistas científicas de alto impacto, por ejemplo, recientemente se descubrió a través de modelamiento molecular que el Covid19 posee una nueva forma de ingreso a través de la interacción con la proteína humana GPR78 [13]. Los avances de la modelación molecular son promisorios, sin embargo, la necesidad de descomunales niveles de procesamiento hace inviable para cualquier compañía invertir en arriendo de tiempo computacional para simular moléculas aun en situaciones de pandemia. Es por estas consideraciones que el enfoque de computación distribuida desarrollado por Folding@home es el más acertado y eficiente por el momento. Es necesario que iniciativas como esta tengan un mayor impacto y colaboración de la sociedad civil debido a que las capacidades de procesamiento de miles de computadores de escritorio superan largamente las capacidades de los centros de cómputo dedicados a cálculos complejos.

Esperamos que en el futuro esta experiencia no sea un hecho aislado, sino que se convierta en una propuesta útil que permita al Ministerio de Educación ampliar la donación de tiempo computacional educacional a una región entera y después replicar la experiencia a otras regiones del país.

En las futuras implementaciones de Folding@home, existen dos problemas pendientes que se deberán solucionar, primero, específicamente en Linux, el cliente o el sistema operativo no pudieron poner en pausa el plegamiento de proteínas cuando un usuario hacía uso de los equipos, esta situación es problemática puesto que el plegamiento hace uso extensivo de los tiempos de procesador, en síntesis, el plegamiento no debe trabajar en forma paralela al usuario, sólo debe iniciarse cuando el usuario no hace uso del equipo. Al cierre de la investigación no pudimos detectar si el problema es del sistema operativo o del cliente de Folding@home. El segundo problema tiene que ver con eficiencia y gasto energético, advertimos que algunos equipos producían simulaciones a muy baja velocidad, pero con igual o superior

Blanco, Sáez & Blanco. Simulación de la proteína Spike COVID19 en laboratorios.

Número Especial Desafíos Humanos ante el COVID-19

Enero – Junio. 2020

<http://dx.doi.org/10.33210/ca.v9i2.326>



consumo de energía eléctrica a otros equipos que producían más simulaciones. Este problema de eficiencia versus consumo energético ya ha sido acusado por implementadores de servicios basados en cloud computing [14], y por tanto se debe establecer cuáles características técnicas son las mínimas para decidir si un equipo es viable para simular proteínas versus su consumo energético.

FUENTE DE FINANCIAMIENTO

La ejecución de esta investigación no fue apoyada ni solventada económicamente por ninguna persona jurídica ni natural aparte de los investigadores autores de este manuscrito.

DECLARACIÓN DE CONFLICTO DE INTERÉS

Los autores y coautores de esta investigación declaran no tener conflicto de interés alguno sobre los hallazgos de esta investigación.

APORTE DEL ARTÍCULO EN LA LÍNEA DE INVESTIGACIÓN

La computación distribuida como arquitectura de modelamiento de sistemas complejos no es el enfoque tradicional que usan las empresas o laboratorios, ellos comúnmente compran tiempo computacional de supercomputadoras para modelar problemas tales como el clima, trayectoria de cohetes, encriptación, etc. En el caso de Folding@home, ellos han apostado por segmentar cargas de trabajo en millones de partes y asignarlas a cientos de miles de computadores personales. Este approach lógico y funcional está sin embargo supeditado a la voluntad de cada usuario individual que escasa vez tiene las competencias para configurar el software y mantenerlo corriendo, es esta brecha de conocimientos una limitante enorme para el usuario no avanzado ni especializado. Les queda la misión a los líderes de Folding@home el mejorar y simplificar aún más la instalación y configuración del software para que pueda adaptarse automáticamente según los recursos de cada equipo.

A través del piloteo de esta experiencia se ha tratado de apuntar a un segmento que cómputo que actualmente está con recursos ociosos: Los laboratorios de computación educacionales. Estos clústeres de equipos presentan un reservorio de poder computacional subutilizado durante el año escolar y sin uso durante la pandemia que pueden ser configurados rápidamente con Linux para servir de centros de cómputo para resolver problemas tan importantes como la simulación de moléculas orgánicas que nos permitan descubrir tratamientos farmacológicos frente a una gama amplia de enfermedades, en nuestro caso en particular, descubrir una forma de interferir con la proteína Spike que actúa como llave molecular para infectar células animales con el SARS-CoV-2.

DECLARACIÓN DE CONTRIBUCIÓN DE CADA AUTOR

Erwin Blanco San Martín tuvo la tarea de instalar el software de simulación de proteínas, crear un grupo de trabajo en la página de Folding@home e instalar Linux en los equipos fuera de servicio.

Fabiola Sáez, condujo la recolección de datos y la creación de estadígrafos, además llevo a cabo la revisión de estilo y formato del manuscrito.

Lorena Blanco San Martín, estuvo a cargo de la creación del marco bibliográfico, la recolección de datos y la edición del manuscrito.

AGRADECIMIENTOS

Agradecemos al jefe del Departamento de Educación de la Comuna de Ránquil, don Benjamín Maureira Álvarez por creer en el futuro de la ciencia molecular modelada y facilitar los laboratorios educacionales donde se llevó a cabo esta investigación.

REFERENCIAS

- [1] Li-Chan, E. C. Y., & Lacroix, I. M. E. (2018). Properties of proteins in food systems: an introduction. In *Proteins in food processing* (pp. 1-25). Woodhead Publishing.
- [2] Englander, S. W., & Mayne, L. (2017). The case for defined protein folding pathways. *Proceedings of the National Academy of Sciences*, 114(31), 8253-8258.
- [3] Schroeder Jr, H. W., & Cavacini, L. (2010). Structure and function of immunoglobulins. *Journal of Allergy and Clinical Immunology*, 125(2), S41-S52.
- [4] Penke, B., Bogár, F., & Fülöp, L. (2016). Protein folding and misfolding, endoplasmic reticulum stress in neurodegenerative diseases: in trace of novel drug targets. *Current Protein and Peptide Science*, 17(2), 169-182.
- [5] Ghosh, R., & Tabrizi, S. J. (2018). Huntington disease. In *Handbook of clinical neurology* (Vol. 147, pp. 255-278). Elsevier.
- [6] Curtis, V. (2018). Patterns of participation and motivation in Folding@ home: The contribution of hardware enthusiasts and overclockers. *Citizen Science: Theory and Practice*, 3(1), 1-14.
- [7] Zimmerman, M. I., Porter, J. R., Ward, M. D., Singh, S., Vithani, N., Meller, A., ... & Hurley, M. F. (2020). Citizen Scientists Create an Exascale Computer to Combat COVID-19. *bioRxiv*.
- [8] Chen, Y., Guo, Y., Pan, Y., & Zhao, Z. J. (2020). Structure analysis of the receptor binding of 2019-nCoV. *Biochemical and biophysical research communications*.
- [9] Drew, E. D., & Janes, R. W. (2020). Identification of a druggable binding pocket in the spike protein reveals a key site for existing drugs potentially capable of combating Covid-19 infectivity. *BMC Molecular and Cell Biology*, 21(1), 1-13.
- [10] Strait, J. E. (2019, March 29). Bowman leading international supercomputing project. Retrieved July 08, 2020, from <https://medicine.wustl.edu/news/bowman-leading-international-supercomputing-project/>
- [11] Mittal, S. (2016). A survey of techniques for approximate computing. *ACM Computing Surveys (CSUR)*, 48(4), 1-33.

[12] Green car congress, P. (2020, March 26). Petrobras directing supercomputer capacity to Folding@home Project effort on coronavirus. Retrieved July 13, 2020, from <https://www.greencarcongress.com/2020/03/20200326-petrobras.html>

[13] Ibrahim, I. M., Abdelmalek, D. H., Elshahat, M. E., & Elfiky, A. A. (2020). COVID-19 spike-host cell receptor GRP78 binding site prediction. *Journal of Infection*.

[14] Naiouf, M., Chichizola, F., & Rucci, E. (Eds.). (2019). *Cloud Computing and Big Data: 7th Conference, JCC&BD 2019, La Plata, Buenos Aires, Argentina, June 24–28, 2019, Revised Selected Papers* (Vol. 1050). Springer.

NOTA BIOGRÁFICA



Erwin Blanco San Martín. **ORCID iD**  <https://orcid.org/0000-0001-9219-2705>
Es investigador de la Universidad Católica de la SSMA Concepción. Obtuvo su licenciatura en Pedagogía Media en Inglés. Su línea de investigación es en informática educativa y neurociencias cognitivas. Actualmente es estudiante de postgrado del programa de Magíster en Educación de la Universidad Católica de la SSMA Concepción, de la ciudad Concepción, Chile.



Fabiola Sáez-Delgado. **ORCID iD**  <https://orcid.org/0000-0002-7993-5356> . PhD en Psicología por la Universidad de Concepción, Chile. Su línea de investigación son las variables cognitivas motivacionales de los procesos de enseñanza y aprendizaje; también las variables de salud mental en comunidades educativas. Actualmente es académica investigadora del Departamento Fundamentos de la Pedagogía de la Facultad de Educación de la Universidad Católica de la Santísima Concepción, Chile.



Lorena Blanco san Martín. **ORCID ID** <https://orcid.org/0000-0002-9584-899X>
Es investigadora de la Universidad de Concepción. Obtuvo su licenciatura en la Universidad Católica de la SSMA. Concepción, tiene una maestría en Lingüística Aplicada. Su línea de investigación es corpus del texto y análisis lingüístico computacional. Actualmente es investigadora en la universidad de Concepción, de la ciudad Concepción, Chile.



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/> or send a letter to Creative Commons, PO Box 1866, Mountain View, CA 94042, USA.

