

Iniciativa Exascale

J. F. Estupiñan, J. L. Gomez

Universidad Industrial de Santander UIS

Bucaramanga, Colombia

jfestupinan@hotmail.com, Joseluisjlgc95@gmail.com

Resumen- En este artículo se aborda la iniciativa Exascale, un proyecto surgido en los Estados Unidos de América para generar un sistema de computación funcional al menos cincuenta veces más veloz que el más poderoso de la actualidad. Además, se analiza el desarrollo y las condiciones dadas para la investigación en dicho país, a fin de comprender qué factores sería necesario cambiar, mejorar o implementar en términos de ingeniería e investigación a nivel nacional para poder participar de una iniciativa de este tamaño.

Abstract- This article addresses the Exascale Initiative, a project which emerged in the United States of America whose purpose is to produce a functional computing system at least fifty times faster than the most powerful one up-to-date. In addition, this paper makes an analysis of United States' development and conditions for research, in order to be able to comprehend which factors would it be necessary to change, enhance or implement in matters of engineering or research in the country so that it can participate in an initiative of this proportions.

Palabras clave- ECP, Exascale, Petascale, Flops, ley de Moore.

I. INTRODUCCIÓN

La capacidad de cómputo es una herramienta muy importante para el desarrollo de la investigación, dado que es precisamente en ella en donde se plantean las problemáticas de mayor complejidad, las cuales representan un mayor reto para esta área de la tecnología.

En la computación actual, las supercomputadoras resuelven problemas en el orden de los petaFlops, es decir, realizan 10^{15} operaciones de punto flotante por segundo. Sin embargo, y dada la vertiginosa evolución de la tecnología, con estos avances también han surgido aplicaciones que demandan estos recursos e incluso más.

Por esta razón, surge el Exascale Computing Project, un proyecto colaborativo de la Oficina de Ciencia del Departamento de Energía de los Estados Unidos (DOE-SC) y la Administración Nacional de Seguridad Nuclear (NNSA), establecido con el fin de contribuir al desarrollo de un sistema de cómputo del orden de los exaFlops, es decir, 10^{18} operaciones de punto flotante por segundo, 1000 veces más que la capacidad actual.

I. EL PROYECTO EXASCALE

A. La Iniciativa Exascale

El proyecto Exascale (ECP) fue establecido en el año 2015, y su principal meta es optimizar la computación de alto nivel en Estados Unidos. Actualmente, existen 20 sistemas de computación petascale, lo que quiere decir que, si se consiguiera lo perseguido por el ECP, el resultado sería un sistema de computación 50 veces más poderoso que toda la pila de supercomputadores junta, que permitiría abordar problemas que en la actualidad son demasiado complejos para resolver. Cabe mencionar, por ejemplo, que con la potencia de computación actual, un supercomputador japonés ha logrado hacer simulaciones del 1% de la actividad cerebral de un ser humano (1 segundo de actividad en 40 minutos), lo que significa que, teóricamente, con la llegada de exascale se podrá simular toda la actividad de un cerebro humano.

Sin embargo, un sistema de tales proporciones sobrepasa cualquier desarrollo hardware o software hecho hasta la fecha, por lo cual el ECP tiene diversos propósitos, entre los cuales están la construcción de un sistema exascale integrado, compuesto tanto de hardware como software especializado para el objetivo, y la programación de aplicaciones que permitan aprovechar el potencial de dichos sistemas.

ECP Timeline

The Project has three phases:

- Phase 1 – R&D before DOE facilities exascale systems RFP in 2019
- Phase 2 – Exascale architectures and NRE are known. Targeted development
- Phase 3 – Exascale systems delivered. Meet Mission Challenges

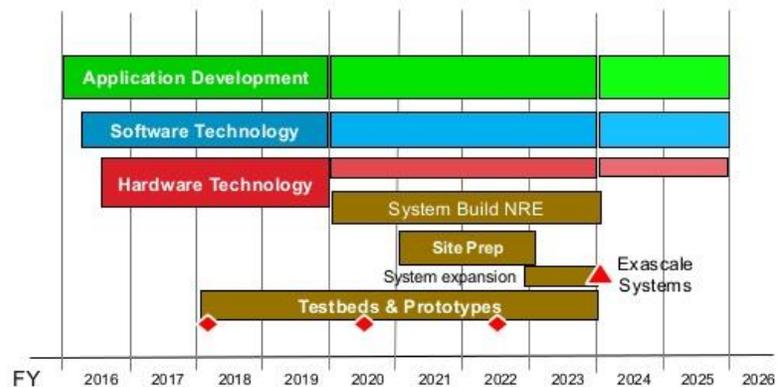


Figura 1: Cronograma del ECP.

B. *Acerca del Departamento de Energía*

El Departamento de Energía (The Department of Energy, DOE) de estados unidos ha logrado a través de los años elaborar capacidades de computación de vanguardia para el desarrollo científico y la seguridad nacional, fundada en 1977 debido a la crisis energética de 1973, al poner las diferentes agencias federales que trataban con posturas en la política energética bajo un solo departamento.

El incremento de la necesidad de computación exascale lo ha convertido en uno de los proyectos de DOE. Sin embargo, previos enfoques para el diseño de HPC software y hardware muestran que las limitaciones tecnologías actuales prohíben la producción de las capacidades requeridas para exascale.

Para poder superar estos desafíos DOE comenzó los proyectos FastForward y DesignForward en una colaboración con los líderes de la industria computacional. FastForward se centra en componentes de tecnología en las áreas de memoria, procesador y tecnología de almacenamiento para a largo plazo sentar las bases en los avances requeridos para cumplir los requerimientos de la misión de DOE. DesignForward se enfoca en diseño de sistemas totalmente integrados, y la interconexión que es requerida que permita sistemas eficientes y escalables para casos de extrema escala computacional.

Con Un periodo de funcionamiento de dos años el objetivo de los programas es iniciar una asociación con múltiples compañías con en el propósito de comenzar proyectos que aceleren la investigación y desarrollo de componentes necesarios para la elaboración de computación exascale. Se empezaron reconociendo el hecho de que el mercado de computación logre innovaciones en dirección que no sean de utilidad para los logros de DOE. Los proyectos buscan fundar múltiples innovaciones tecnológicas que comiencen producción en un lapso de 5 a 10 años.

1) *FastForward*: En el 2011, la solicitud de propuestas de FastForward solicitaba propuestas innovadoras en la investigación y desarrollo de procesadores, memoria, almacenamiento y I/O que mejoren el desempeño por watt, productividad y la fiabilidad de aplicaciones de extrema escala. El objetivo es crear un impacto a largo plazo en la elaboración de sistemas exascale en la siguiente década.

La colaboración pública-privada del programa apoyará el desarrollo de tecnologías que reduzcan las barreras económicas y de fabricación no solo de los futuros sistemas exascale, pero que ciertas innovaciones logren ser implementadas en sistemas de menor escala, asegurando que los gasto de DOE fomente el desarrollo comercial de múltiples sistemas futuros.

Intel Fast Forward 1 Extension – Storage & I/O

Investigators

- Eric Barton (PI - Technical lead)

DOE representatives

- Gary Grider, LANL
- Rob Ross, ANL
- Mark Gary, LLNL

Total Funding

- Approximately \$2M

Contract Period of Performance

- July 2015 – June 2017

Engagement with applications

- Application porting & workshops
- Task-based programming models

HDF5

- Improved user productivity
- Stored metadata indices & accelerated query
- Improved asynchronous error handling

DAOS

- Ultra fine-grain / fragmented, low latency I/O
 - Mercury – OS bypass communications
 - Persistent Memory – OS bypass storage
- Enable more programming & workflow models
 - Improved transaction model
 - Generalized key-value storage
 - Replication & erasure coding
 - Consistent stack on all storage tiers

Containerized Global Namespaces

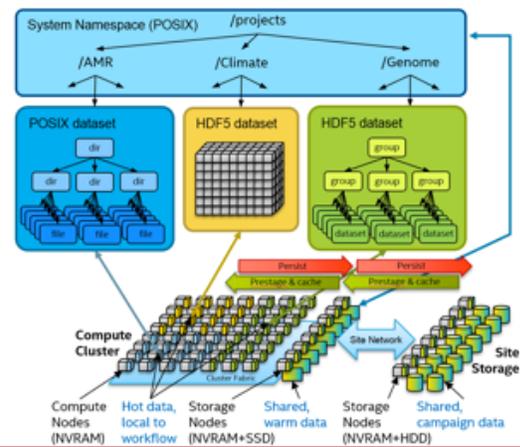


Figura 2: Proyecto FastForward.

2) *DesignForward*: Si FastForward se centra en la elaboración de hardware DesignForward funciona de manera similar pero con un enfoque en el desarrollo de aplicaciones de escala extrema, donde computación, movimiento de datos, integridad y el tiempo total de solución tienen un papel importante en los proyectos, es importante mencionar que los contratos premiados en el programa pueden ser elegidos para fundaciones adicionales para acelerar la investigación de la tecnología elaborada esto si el congreso aprueba la fundación para este propósito.

Las propuestas presentadas al proyecto deben de explicar cómo lo propuesto va a impactar los objetivos de DOE en aplicaciones de escala extrema, así como a la comunidad más amplia de HPC y exhibir el potencial de la tecnología para usos por fuera del mercado de supercomputadoras de DOE, como son los centros de datos, la nube y aplicaciones HPC de rango medio. Importante aclarar que DOE no está interesado en fundar desarrollos de capacidades a corto plazo que ya existen en mapas de evolución de tecnología.

3) *FastForward II*: La siguiente fase del programa FastForward sigue los mismos parámetros, pero hay un enfoque en la arquitectura de nodo y tecnología de memoria. Con los resultados de los proyectos siendo producidos en un periodo de tiempo de 2020-2023.

DOE se asegura de suministrar la información necesaria sobre los proyectos de la anterior fase para que la elaboración de los nuevos proyectos los utilice como bases y logren la integración de ellos.

Entre los proyectos se encuentran: AAR Memory Technology, AAR Node Architectures, Cray Node Architecture, IBM Memory Technology, Intel Node Architecture, NVIDIA Node Architecture.

4) *Designforward II*: Los sistemas de integración han sido el enfoque de la nueva fase que aún se encuentra vigente, aceptando nuevas propuestas que cumplan con lo requerido dicho anteriormente.

Proyectos: AAR system integration, Cray system integration, IBM system integration, Intel System integration.

5) *Pathforward*: Por último, se encuentra pathforward la fase final de DOE, el RFL de esta última fase se encuentra en línea desde junio del 2016 en la siguiente dirección, <https://asc.llnl.gov/pathforward/>. La página de Exascale Initiative no presenta ninguna información sobre los proyectos que fueron o no aceptados, pero si presenta la información y respuestas a preguntas frecuentes que pueden ser útiles para la presentación de una propuesta junto con la forma de contactar DOE.

Lo que busca DOE con pathforward es:

- I. Un incremento sustancial en la competitividad de sistemas exascale HPC, donde el desempeño de aplicaciones es el criterio más importante.
- II. Mejorar la confianza de los oferentes en el valor y factibilidad de las opciones agresivas de avances tecnológicos en las que estén dispuestos a proponer para la adquisición de sistemas exascale.
- III. Identificar la opción tecnología más prometedora que será incluida en las propuestas del 2019 para los sistemas exascale HPC.

C. Desarrollo de sistemas exascale

Como se mencionó anteriormente, el primer gran objetivo del ECP se subdivide a su vez en dos partes. La primera de ellas es la fabricación de componentes hardware que faciliten los recursos que un sistema exascale pueda llegar a requerir, y que permitan además mejorar el rendimiento de aplicaciones tradicionales de investigación e ingeniería como sistemas de análisis de datos. En el área de tecnología del hardware, el ECP busca establecer las bases para que exista diversidad en la arquitectura a nivel de computadores de alto rendimiento.

Pero evidentemente con este hardware debe también diseñarse su respectivo software, lo cual es la segunda parte del primer objetivo. Aquí, el ECP busca desarrollar un set de herramientas software intuitivo que permita a los desarrolladores producir aplicaciones que no solo aprovechen de la mejor

manera los recursos del sistema, sino que además tengan portabilidad entre diferentes arquitecturas software.

D. Desarrollo de aplicaciones exascale

El área de desarrollo del ECP busca proporcionar aplicaciones producto de la investigación que le den sentido a la construcción de un sistema exascale. Esto quiere decir, aplicaciones que realmente exploten el potencial de estos sistemas para dar respuesta a problemas críticos en materia de seguridad, información, simulación de procesos y, más globalmente, investigación. El ECP ya ha identificado muchas áreas críticas en distintos campos de investigación donde las aplicaciones exascale tendrían un impacto significativo.

E. Áreas de investigación

El ECP tiene pensado desarrollar su sistema para investigación en diversas áreas como la astrofísica, para la investigación acerca del origen de los elementos químicos; la química, para el desarrollo de cultivos más resistentes y el diseño de catalizadores de biocombustible; la meteorología, el diseño de motores de alta eficiencia y baja combustión, el estudio de la materia oscura y la energía oscura, la energía nuclear, la energía eólica, la medicina orientada al cáncer, entre otras. Estos campos tienen en común que requieren el movimiento de enormes volúmenes de información para la resolución algunos de sus problemas, y esto puede representar grandes costos de tiempo en análisis de datos.

F. Posposición del proyecto

Desde su inicio, el ECP ha tenido que pasar por varios procesos de dilación para la entrega de su sistema. Inicialmente este se había concebido para el año 2018, fecha que fue cambiada para el 2020, y luego nuevamente aplazada para entre el 2023 y el 2025.

Hay una variedad de que factores han impedido la consecución de los objetivos para los plazos marcados previamente, los cuales van desde la programación y el consumo de energía hasta el diseño del hardware, pues como ya se ha mencionado, este proyecto requiere componentes especializados dadas las enormes prestaciones que ofrece, pues estas sobrepasan radicalmente los lineamientos de crecimiento computacional que trazó la ley de Moore.

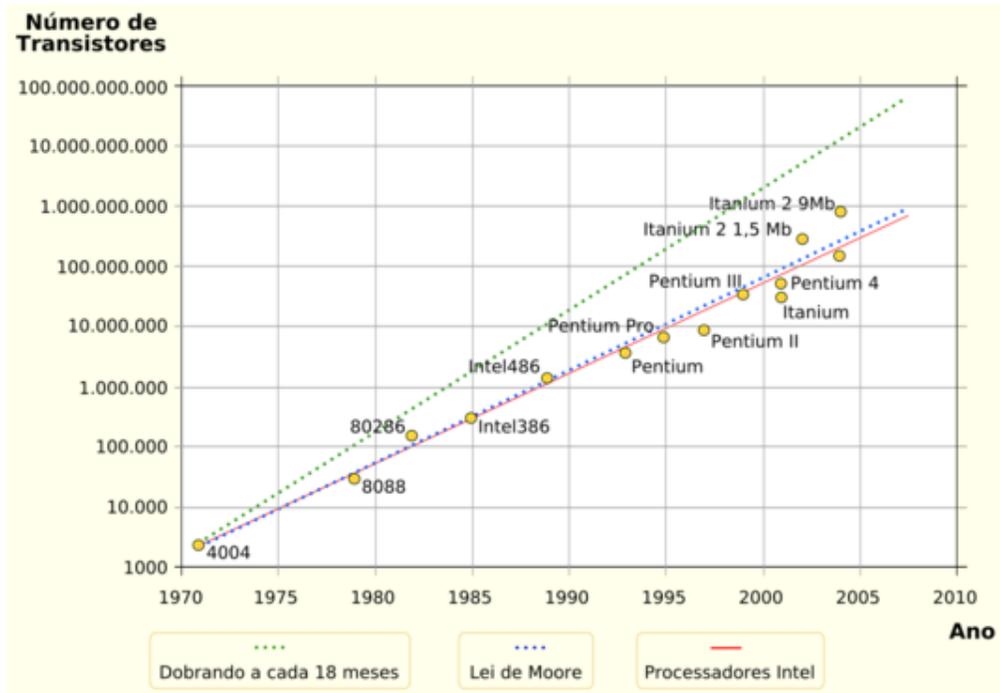


Figura 3: Ley de Moore comparada con datos históricos.

Más detalladamente a nivel de hardware, algunos de los grandes retos a superar por parte del equipo ECP son los siguientes:

1) *Paralelismo*: Para poder ejecutar un trillón de operaciones por segundo, el sistema exascale deberá poseer la capacidad de distribuir eficientemente tareas entre diferentes hilos en un mismo nodo, además de tener una mayor cantidad de nodos en un sistema, a fin de reducir un problema complejo a varios problemas más simples que puedan ser resueltos simultáneamente. Esto representa un desafío en los departamentos de algoritmos, sistemas operativos y compiladores.

2) *Memoria y almacenamiento*: La memoria es un reto enorme en la construcción del sistema, además del almacenamiento en disco, el cual podría no proveer el rendimiento adecuado, por lo que se están considerando nuevas alternativas en materia de tecnologías de memoria, como la STT-RAM, PC-RAM O R-RAM, que podrían solventar algunos de los problemas inherentes al almacenamiento en memoria.

3) *Confiabledad*: Existe un concepto llamado ‘Error silencioso,’ que hace referencia a errores ocurridos a nivel del hardware o del software de bajo nivel. En las arquitecturas de hardware actuales este tema es tenido en cuenta, pero no con la prioridad suficiente como lo requiere un sistema de la complejidad de exascale. Dicho esto, es necesario que la arquitectura exascale se construya desde cero, para facilitar la detección y corrección de fallas que puedan encaminar a errores silenciosos.

4) *Consumo de energía:* Los computadores de alto rendimiento consumen altos niveles de energía en la transmisión de datos en memoria y disco. Actualmente, los costos por el consumo de energía de los sistemas petascale ya representan una suma importante del TCO de dichos sistemas. Es lógico pensar que los sistemas exascale tendrán un consumo de energía aún mayor, dada la cantidad de hardware al cual será necesario alimentar para el correcto funcionamiento de estos sistemas. Por esta razón, se hace necesario diseñar hardware de consumo eficiente para mitigar el impacto económico producto de este aspecto.

El reto clave en este problema es encontrar una forma de minimizar el consumo de energía por transferencia de datos. A nivel de circuitos, la solución requiere que el diseño de estos se enfoque primordialmente hacia la eficiencia energética, dejando de lado (en cierta medida) la velocidad en la mayoría de dichos diseños.

Las anteriores son algunos de los retos importantes para la consecución de los logros planteados por el ECP, además de algunos otros como la interconexión de redes, los sistemas operativos, el debugging, la eficiencia de los algoritmos utilizados, la optimización de los compiladores, entre otros. Adicionalmente, consecuencia de las aplicaciones innovadoras que el ECP pretende diseñar, surgen también problemas en las áreas de ciencia de datos y aprendizaje automático los cuales deben ser resueltos para su correcta implementación.

Por esta razón, ya hay iniciativas externas al ECP, como por ejemplo Mont-Blanc, que pretenden afrontar el reto de diseñar el hardware que cumpla con los requerimientos para los sistemas exascale.

CONCLUSIÓN

Si Colombia desea integrarse a la iniciativa de exascale está en la completa capacidad de lograrlo, los RFL de DOE permiten que entidades extranjeras sean fundadas por el programa, pero para poder lograrlo necesitará de una propuesta que logre persuadir a DOE, los méritos que se deben de cumplir para poder ser fundados incluyen, el desarrollo e investigación de tecnología a largo plazo que no se encuentre en las rutas de tecnología actuales, la demostración de que la elaboración de esta tecnología no solo beneficiara a los sistemas exascale sino también al mercado más amplio de computación, algo a tener en cuenta es que DOE está buscando software y uno de los parámetros es que cada propuesta de relación al software debe de ser acompañada por un diseño completo de un sistema exascale.

La búsqueda de propuestas innovadoras es la clave para ser aprobados en el programa si se logrará fomentar esta noción a través de compañías, universidades y/o centros de investigación además de una

pequeña fundación pueda ser usada para la elaboración de los requerimientos de DOE, la noción de participar en el programa podría aumentar.

REFERENCIAS

- [1] *Proyecto de Computación Exascale*. Disponible en: <https://exascaleproject.org/>
- [2] (The Next Platform) N. Hemsoth. (2016, abril). “*Exascale Timeline Pushed to 2023: What’s missing in Supercomputing?*”. Disponible en: <https://www.nextplatform.com/2016/04/27/exascale-timeline-pushed-2023-whats-missing-supercomputing/>
- [3] C. Stunkel. (2010, abril). “*Exascale: Paralelism gone wild!*”. Disponible en: <http://www.ipdps.org/ipdps2010/ipdps2010-slides/keynote/2010%2004%20TCPP%20Exascale%20FINAL%20clean.pdf>
- [4] FastForward program. Disponible en: <http://www.exascaleinitiative.org/fast-forward>
- [5] (Lawrence Livermore National Laboratory) S. Hemmert et al. “*Exascale Hardware Architectures Working Group*.” Disponible en: <https://asc.llnl.gov/content/assets/docs/exascale-hwaWG.pdf>