

“Breve Historia de la evolución de HPC en una Institución de Investigación en la Argentina”

Roberto Bevilacqua¹
robevi@gmail.com

¹Ex-Prof. Asociado Regular UBA (doc. 1977-2013)-Prof. Titular Ordinario UNSAM (Inv-Doc.1995-2018)-Prof.Titular AH simple (desde 2019)-Director Área Informática ECyT-UNSAM (1995-2018)-Ex-Jefe Departamento Computación de Alta Prestación GTIC-CAC-CNEA (Mayo 2011-Mayo 2017)- Asesor Tecnológico del Departamento de Computación de Alta Prestación - Gerencia de Tecnología Informática y de las Comunicaciones (DCAP-GTIC-CNEA) (1978-2018)

Resumen: La presente ponencia es una breve historia de la evolución de High Performance Computing (HPC) en una Institución de Investigación en la Argentina. Además de la reseña histórica, se pretende mostrar, que en los proyectos importantes son necesarios dos componentes: el conocer y el interés de los gobiernos en que se realicen.

Palabras Claves: HPC – Historia – Gobierno

I. Un poco de historia

En sus inicios (09/1977) el Centro de Cómputos (CC) de la CNEA contaba con una computadora IBM 370 modelo 158 U35 con 1MB de memoria discos IBM 3350 de 300 MB c/u, llegándose en 1985 a tener una capacidad de 4,2GB en disco. Su capacidad de procesamiento era de 1 MIPS (Mega Instrucciones Por Segundo) y su sistema operativo fue el VM/370 (Virtual Machine), contando con 20 terminales IBM 3277 dependientes funcionalmente del procesador central (que más tarde se los identificó vulgarmente como monitores bobos). En su momento de apogeo el CC dispuso de 250 terminales de trabajo en línea, conectadas en su mayoría en forma remota mediante un controlador de comunicaciones IBM 3705 que empleaba 22 líneas punto a punto (una conexión directa entre dos puntos sin capacidad de switcheo) y operaba bajo un protocolo BSC (Binario Sincrónico con Corrección), y a

partir del cual se armó un entorno de Teleproceso o procesamiento a distancia: o sea, no había procesamiento local, ya que todo era resuelto por el computador central. Estas terminales podían realizar una serie de actividades valiéndose de la ayuda de un procesador externo denominado Unidad de Control y sin la intervención del procesador central, a saber: manejo completo de pantalla, refrescos y eco. Para el momento, este equipamiento (hardware IBM 370 y sistema operativo VM /370), implicaba un importante avance tecnológico, pues su combinación de hardware y software permitía la utilización de lo que se conoció como paradigma de memoria virtual, consistente en simular mayores espacios de memoria y múltiples espacios de direcciones (o sea hacerle creer a los procesos/programas que la memoria central era mucho más grande que la que realmente se disponía) y en la utilización fuerte de la multitarea o multiprogramación. Asimismo, y por primera vez en el país, fue empleado el término máquina virtual (mención que aún hoy perdura) que concibe el uso simultáneo de sistemas operativos diferentes en un mismo equipo físico. Así, al coexistir varias máquinas virtuales en un mismo aparato concreto, se comenzaron a utilizar los CMS (Conversational Monitoring Systems) que contenían una alta calidad de lenguaje interactivo y de compiladores, además de un excelente intérprete, el EXEC (más tarde conocido como REXX), para generar prototipos en forma rápida (prototipación) que favorezcan el control de entornos, funciones similares a las que cumplen los scripts y awk de las actuales versiones de UNIX. Todo esto controlado por medio de un soporte software de comunicación al hardware llamado Control Program (CP), que actualmente sería conocido como un Microkernel.

Otras Virtual Machine operaban con un sistema operativo denominado OS/VS1, el cual contenía aplicaciones de tipo batch o en lote, junto a otras aplicaciones transaccionales como el CICS (Conversational Information Customer System) y bases de datos como ADABAS (de software AG) que ya era multithreading. Esta conjunción de máquina y software eficaz permitió que un grupo considerable de científicos, técnicos y administrativos se beneficiaran por la tecnología de la computación. Este tipo de elementos electrónicos pertenecen a un conjunto de invenciones muy avanzadas para la época. Debemos recordar que para esos momentos, más precisamente entre los años 1965 y 1974, se producía la tercera generación de computadoras. Ya a mediados de los años sesenta se logró la invención del circuito integrado (encapsulado de material semiconductor que contiene diminutos circuitos electrónicos), y hacia mediados de esa década IBM anunciaba la construcción de computadoras compuestas por este tipo de circuitos: la serie 360.

Por su parte y en lo que atañe al Centro de Cómputos, la documentación de operaciones y los registros de inventarios informan que en 1982 existían más de 250 usuarios en línea, generándose la saturación del sistema. Esto motivó a que se adicionara a la IBM /370 una BASF (HITACHI) 7/68 con memoria de 8MB de RAM. Más tarde se transforma en una HITACHI 7/69, con lo que se alcanzan los 3 MIPS de rendimiento, y hacia principios de los años 90 se reemplaza la IBM /370 por una HITACHI 7/77, con lo cual se alcanzan los 8 MIPS de rendimiento en total. A fines de 1989, y luego de varios años de trabajo ininterrumpido bajo modalidad independiente, se instaló el software de conexión RSCS (Remote Spool Communications Subsystem) y el PVM (Passthrough Virtual Machine): el primero de los cuales efectuaba la

transferencia de archivos entre máquinas, mientras que el segundo permitía un login remoto y el trabajo en forma interactiva en cualquiera de las dos máquinas y desde cualquier terminal de usuario. Con el incipiente comienzo de sistemas UNIX en la CNEA (años 1993/4), este mismo software pudo establecer comunicación entre mainframes y los sistemas UNIX; además, por medio de una placa especial de comunicación (CLAW) que se instalaba en uno de los sistemas UNIX fue posible transferir archivos y acceder a los sistemas en forma interactiva y bidireccional entre este sistema y los propios del mainframe. Lo interesante de esta comunicación fue que los sistemas UNIX identificaban a la placa CLAW como una placa-tarjeta de comunicaciones, mientras que los mainframes la interpretaban como una de unidad de control. Tal condición permitió que el Ing. Andrés Alonso y el Lic. Roberto Bevilacqua (años 1993/4) pudieran diseñar e implementar el primer bridge entre BITNet y UUCP (ver más adelante) de la Argentina, aunque probado durante un largo periodo, nunca pudo ser puesto en operación.

Desde el año 1987 se instaló un servicio gráfico de alta resolución, con monitores vectoriales Tektronix de 20" y plotters (trazadores de inyección de tinta que imprimen en forma lineal) de gran tamaño. La alta calidad del servicio tecnológico ofrecido por el repercutió en un vertiginoso aumento de su demanda, a punto tal que muchas veces se cubrían necesidades tecnológicas que paulatinamente se implementarían en el ámbito de la informática, siendo el Lic. Héctor Schilman el abanderado de estas aplicaciones. Gracias a una conferencia sobre BITNet (Because Is Time to Net, contemporánea de la comunicación UNIX to UNIXrotocol), que se desarrolló en el edificio TANDAR del Centro Atómico Constituyentes, técnicos y administradores de sistemas del Centro de Cálculo Científico (CCC) se percataron que estaban utilizando BITNet sin saberlo, y en forma autista. Esto despertó un fuerte interés en la comunidad científica, en particular a quienes habían pasado temporadas en el exterior y ya habían utilizado este paradigma de comunicación en otros países. BITNet consistía básicamente en el envío y recepción de correo (al presente e-mails), y en el uso de mensajes en línea (hoy en día chat). Frente a esta experiencia se solicitó a la Dra. Emma Pérez Ferreira (Presidente de la Institución) la autorización correspondiente para efectuar una secuencia de comunicaciones nocturnas -por línea telefónica desde su despacho y con acceso vía enlace de radio entre Sede y CAC- con la Universidad de Santiago de Chile, inaugurándose así la conexión BITNet, antecesora de Internet. Este servicio se mantuvo en actividad hasta el año 1992, ofreciendo el servicio de comunicación internacional a la Universidad Nacional de La Plata (por su intermedio también a la Universidad Nacional del Centro de la Provincia de Buenos Aires), a la Universidad Tecnológica Nacional, al Centro de Tecnología de Comunicación y Sistemas (CTCS-UBA) y a la Facultad de Ingeniería de la Universidad Nacional de la República Oriental del Uruguay.

Años más tarde (1995) comienza a estudiarse la posibilidad de incorporar un servidor numérico de alta capacidad, hecho que se concreta gracias a la intervención conjunta del Dr. Audero y del Dr. Pochettino. Tras la instalación del software correspondiente se inician las aperturas de cuentas de usuarios (abril de 1997). Este servidor es un equipamiento SUN Enterprise 6000 de 16 procesadores RISC de 160 MHz (de frecuencia de procesador de tecnología RISC), con un rendimiento de 5

MFLOPS (Mega instrucciones de Punto Flotante Por Segundo – Mega FLOating Point per Second), 8GB de memoria RAM y 40 GB en discos SCSI y un robot automatizado para realizar los back-ups (copias de seguridad de información) llamado FENIX; en 2002 se le adiciona un servidor NFS de 80 GB en disco. Debido a la capacidad de procesamiento de este equipo multiprocesador, y en particular al gran tamaño (en esos tiempos) de la memoria RAM, muchos investigadores de CNEA pudieron desarrollar, probar y utilizar sus modelos computacionales, sin encontrarse frente a la necesidad de recurrir a soporte informático vía remota al exterior, o incluso de trasladarse al exterior para encontrar recursos y equipamientos similares: lo que significó un salto cuantitativo y cualitativo en el campo de la computación. Entendemos que este hito es el que permitió a CNEA, luego, a partir de 2003, comenzar a diseñar otro salto informático, el de la supercomputación.

II. La era de los Clusters en CNEA

La incorporación de FENIX, junto al comienzo de una administración de redes más adecuada y el resurgir de las actividades propias del antiguo CCC, todas ellas propiciadas por el Dr. Miguel Audero, permitieron que a partir del año 1995 y en lo sucesivo, la actividad del CCC fuera nuevamente valorada y considerada por todos los usuarios científicos y administrativos.

En torno a los clusters, tema muchas veces discutido, evaluado e incluso avalado por varias pruebas realizadas con máquinas pequeñas, puede decirse que este periodo casi tiene fecha de comienzo cuando el Grupo de Sistemas Operativos (Roberto Bevilacqua; Andrés Alonso y Gastón Aguilera) presenta la ponencia Proyecto de Cluster en el Taller de Materiales Avanzados del Centro Atómico Constituyentes – CNEA (MACAC, 15 y 16 de mayo de 2003). A partir de ese momento varios investigadores se contactan con el Grupo y finalmente se recibe el ofrecimiento de presentarse al PME2003 con la dirección del Dr. Pablo Fainstein y la codirección de la Dra. Ana María Llois. Con el entusiasmo del Lic. Guillermo Vallone, actual Gerente de la GTIC, se logra concretar el proyecto del cluster Sheldon en 2007 -nombre seleccionado en homenaje al personaje de la saga Fundación, de Isaac Asimov-, el cual alcanza un rendimiento operativo promedio de 290 GFLOPS (no homologado). (puesto 23ro. LARTop50 2014). Un poco más tarde, y bajo la dirección de la Dra. Llois, se accede al PME2006, lo cual permite que a partir del 29 de septiembre de 2010 se puedan ofrecer los servicios de ISAAC (Information Systems And Advanced Computing), esta vez en referencia directa tanto a Newton y a Asimov, y cuyo rendimiento pico es de 5,5 TFLOPS (no homologado). (Puesto 5to. LARTop50 2013, puesto 12vo. LARTop50 2014).

III. La era de los Clusters en CNEA (II)

El cálculo y las simulaciones numéricas se han transformado gradualmente en herramientas de trabajo indispensables dentro de casi todas las ramas de las ciencias. Su rango de aplicación cubre actualmente un amplísimo espectro, desde el estudio de propiedades estructurales, de transporte y dinámicas de distintos tipos de sistemas, la simulación de reacciones químicas, biológicas y catalíticas, hasta el diseño de nuevos materiales e incluso, la adquisición y procesamiento de gran cantidad de datos en astrofísica y física de la atmósfera, entre otros.

La era de los clusters es casi un eufemismo para indicar que la computación ha entrado en una etapa en la cual el procesamiento, de aplicaciones y del software de base, deja de ser secuencial, para entrar en la era de la concurrencia, que incluye paralelismo, simultaneidad y pipeline. Dado que las computadoras hogareñas y los teléfonos celulares, ya tienen más de un núcleo, es más que evidente que se debe mirar con entusiasmo todo aquello que sea paralelo. El concepto de concurrencia está pensado para que las ejecuciones dejen de ser secuenciales y se ejecuten en alguno de los tres modos que implica la concurrencia; es el afán de resolver las cosas más rápido, y para lograrlo se dividen los problemas en partes que sean independientes pero que a su vez sean cooperantes... o sea: los paradigmas son dividir y cooperar.

Un cluster, si se mira de una manera sencilla, es un conjunto de máquinas que resuelven un problema, dividido en partes, pero que intercambian información para cooperar en busca de un resultado en el menor tiempo operable. La forma de dividirlo es por medio de concurrencia sobre el dominio de datos, es decir, el mismo problema se ejecuta en varias máquinas sobre conjunto de datos distintos o sobre dominio de procesamiento... o sea piezas de software distintas que operan los mismos datos o distintos pero que en todos los casos cooperan para obtener un resultado. Fundamentalmente se requieren éstos para procesar información y obtener resultados en la forma más rápida posible. O sea: El objetivo de un cluster es que el conjunto de nodos, conectados por medio de un enlace, ejecuten subtareas (de una tarea ya dada) con el objetivo de aprovechar aquellas acciones que son concurrentes para superponerlas en tiempo y espacio de tal manera de acortar los tiempos de procesamiento de la tarea original, si esta fuese ejecutada secuencialmente. Esta forma de trabajo debe permitir que el usuario vea al cluster como un sistema único, y que el conjunto de subtareas, que son cooperantes para resolver la tarea, ejecutadas en nodos distintos, sumen las potencias de los nodos hasta obtener una capacidad de procesamiento como si esta tarea fuese ejecutada en un solo nodo igual a la capacidad de la suma de las capacidades de los nodos intervinientes. Está claro que esta capacidad depende fuertemente de la potencia de procesamiento de cada nodo, de la velocidad del enlace, de la forma de administrar los recursos y del algoritmo utilizado para resolver el problema que la tarea representa (abstrae) en su forma computacional. Es así como aparecen cuatro áreas distintas pero interrelacionadas:

- Estructura de Sistemas de Hardware
- Recursos de administración y manejo de entornos
- Bibliotecas y herramientas para programación distribuida
- Algoritmos Paralelos

La Estructura de Sistemas de Hardware comprende todos los aspectos de los componentes de hardware de los nodos, los enlaces y controladores o switches que componen la red, así como la topología de la red. Los recursos de administración y manejo de entornos son un conjunto de herramientas software que gobiernan todas las fases de la operación del sistema, desde la instalación, configuración, manejo de tareas, monitoreo del sistema, diagnóstico de fallas, hasta las tareas de mantenimiento. Las bibliotecas y herramientas de programación distribuida determinan el paradigma con el que el usuario coordina los recursos distribuidos para ejecutar, en forma concurrente y cooperativa, los componentes lógicos que constituyen la aplicación paralela. Los algoritmos paralelos serán los que provean los modelos que permitan al usuario explotar el paralelismo intrínseco de su problema, para obtener en la práctica un rendimiento efectivo.

El GRID y/o el CLOUD son una tecnología que, en principio y en forma grosera y de la misma manera que un cluster es un conjunto de equipos cooperantes, podría considerarse como un conjunto de clusters cooperantes. De alguna manera lo es, pero, y debido a la tecnología actual, no es la solución para los grandes problemas de procesamiento de datos que requieran una fuerza muy grande de cálculo. De todas maneras es una abstracción superior al cluster, pues permite efectivamente la cooperación en acciones computacionales que no requieran esa fuerza de cálculo que dan los HPC. La abstracción está dada por medio de una serie de capas de software que permiten utilizar piezas de software, perfectamente ya armadas, desde lugares sencillos, como por ejemplo desde una página WEB que esté conectada a un portal. O sea: el equipamiento que existe actualmente para realizar tales tareas se puede dividir entre supercomputadoras (mainframes) y clusters de computadoras. En los últimos años se ha demostrado que, a rendimientos equivalentes, las supercomputadoras no ofrecen una adecuada escalabilidad y protección de la inversión. Luego de un tiempo corto (en términos relativos a la inversión), una supercomputadora debe ser retirada de servicio sin poder aprovechar el equipamiento remanente y sin haber amortizado realmente su costo. Un cluster de computadoras de alto rendimiento permite un reemplazo sostenido de sus nodos los cuales pueden ser utilizados en aplicaciones menos demandantes, ya sea en oficinas, laboratorios o aulas, permitiendo el máximo aprovechamiento de la inversión inicial. Por otro lado, más recientemente se está implementando en muchos centros de investigación de avanzada un tipo de entorno de cálculo que promete contribuir al cambio en la forma de resolver los problemas complejos. Este entorno permite agregar y compartir recursos computacionales, a través de fronteras institucionales y geográficas, construyendo redes de clusters que operan bajo Linux, por ejemplo, y que están interconectadas vía Internet. Este entorno de cálculo utiliza interfaces ya disponibles, tales como MPICH-G2, para

ejecutar programas MPI. Tanto los clusters como así también la ampliación del concepto a Grid/CLOUD de clusters y computadoras aumentan, a bajo costo, la capacidad de cálculo y funcionalidad de distintos grupo que comparten recursos computacionales. La ventaja principal de los clusters de computadoras de alto rendimiento, accesible a un número importante de investigadores, consiste en que tiende a poner a su alcance una facilidad de cómputo con capacidad pico superior a aquella que los grupos pueden alcanzar individualmente, minimizándose los tiempos de CPU ociosos.

Además la tecnología nueva requiere de redes de alta velocidad, de muchos componentes de software para administrar y permitir el uso de los recursos de computación, tales como planificadores de carga, distribución de datos, distribución de procesos, disponibilidad de objetos de computación en cualquier punto de las NUBES, manteniendo la mayor transparencia posible sobre la ubicación de los recursos para los procesos y usuarios finales. Está claro que esto requiere de profundas investigaciones sobre muchos aspectos de la combinación de la llamada Tecnología de la Información y Comunicaciones (TIC) como protocolos de comunicaciones, políticas de asignación, políticas de distribución y políticas de comunicaciones, abriendo nuevos frentes de investigación en términos de programación paralela y distribuida, aplicables éstas al conjunto de e-science, a las aplicaciones de modelo cliente servidor y de e-learning, teleinmersión, e incluso de bases de datos; además a controles de procesos y de minería de la información.

La importancia de este tema ya se nota por la inclusión de los mismos en el Plan Estratégico de CNEA 2010-2019. Desde la GTIC, en concordancia con el mismo, se ha creado el Departamento de Computación de Alta Prestación (DCAP), de reciente creación, en de mayo 2011. El DCAP tiene tres Programas troncales que son el de HPC, el de Servicios de Datos y Aplicaciones (Data-Storage, Data-Grid, Cloud Computing) y el de Comunicación Avanzada (Redes de Baja Latencia, Redes Avanzadas, Redes Ad-Hoc). (PE 2015/2019)

Actualmente, y dado que la infraestructura es muy importante para avanzar este tipo de tecnología, se encuentra en marcha una importante licitación sobre equipamiento de Refrigeración para HPC, cuya característica principal es la de no enfriar ambiente, sino controlar la temperatura en forma directa sobre los equipos, y cuya característica principal es que éste está dimensionado para poder refrigerar 3 veces más que los equipos actuales. Los diseños y especificaciones (en borrador) para la incorporación de equipamiento nuevo y más moderno, que permita aumentar la capacidad de cálculo en un orden de magnitud, también están esbozados con tecnología de principios de este año... y hago la advertencia de “borrador y fecha”, pues al momento de disponer de los recursos necesarios, es siempre imprescindible una actualización de los mismos. Además, debe destacarse que el interés de esta clase de equipamiento no sólo es propio de la CNEA, sino también alcanza a niveles del gobierno nacional, y eso queda demostrado en la creación del Sistema Nacional de Computación de Alto Desempeño, proyecto dependiente de la Secretaría de Articulación Científico Tecnológica del ex-Ministerio de Ciencia, Tecnología e Innovación Productiva. Allí, en el Consejo Consultor, la CNEA está representada, y los objetivos que persigue este proyecto giran en torno a la comunicación y

cooperación de Centros Nacionales como el de la DCAP-GTIC-CNEA y a la incorporación de un equipo de proporciones importantes a nivel nacional, para todo el sistema científico tecnológico nacional.

IV. Conclusiones

Creemos que la hipótesis de la Introducción ha sido claramente explicada y que se debe continuar interesando a las más altas autoridades del país para lograr que se involucren realmente en el desarrollo de estas capacidades y potencialidades con que la Argentina ya cuenta.



SALA CNEA-GTIC-DCAP

De izquierda a derecha: Leonardo Domínguez - Agustín Nieto - Gastón Aguilera - Roberto Bevilacqua - Roque Iozzo - Andrés Alonso. Todos especialistas en Computación de Alto Rendimiento (HPC).