

Encontrando un bosón de Higgs en un pajar



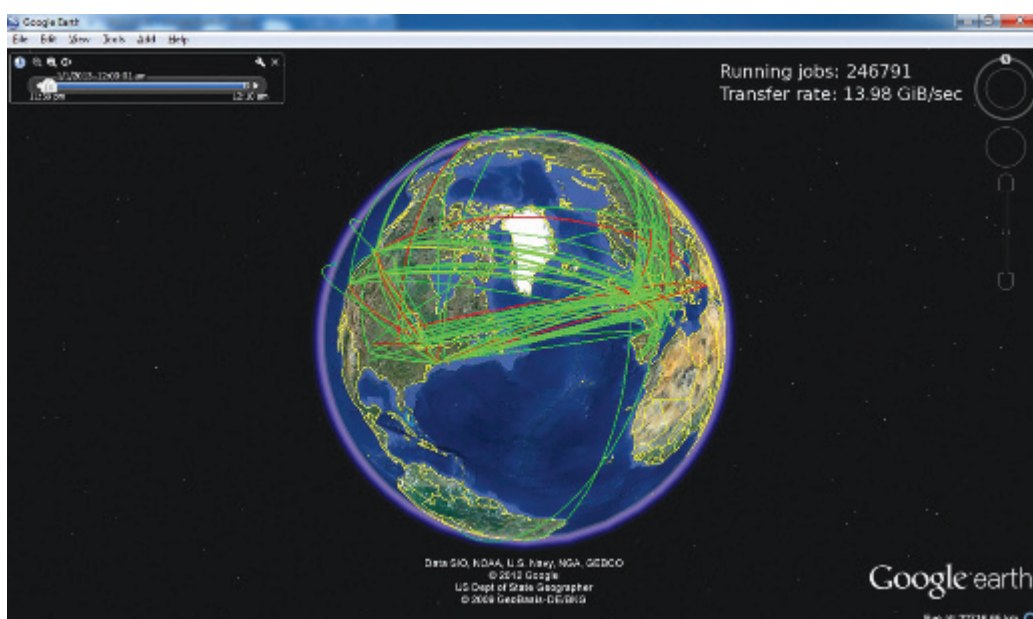
por Andrew Purcell¹ y Alberto Pace²

Como se ha descrito en el artículo precedente, la comunidad meteorológica está mejorando rápidamente su capacidad de gestionar y distribuir las crecientes cantidades de datos e información que generan los Miembros de la OMM y otras organizaciones. Como era de esperar, otras comunidades técnicas y científicas también están lidiando con el problema que representa encontrar la mejor manera de organizar cantidades ingentes de datos. Al igual que la OMM, están afrontando este reto aprovechando la ventaja que supone el auge de Internet, la potencia en aumento de los ordenadores y la creciente sofisticación de los programas informáticos.

El siguiente artículo describe la estrategia que está adoptando la Organización Europea para la Investigación Nuclear, conocida por su acrónimo francés como CERN y situada a apenas unos pocos kilómetros de la sede central de la OMM en Ginebra. Aunque la naturaleza de las necesidades de gestión de los datos del CERN es en muchos aspectos diferente a la de la OMM, su estrategia ofrece una interesante comparación con el enfoque utilizado por el Sistema de información de la OMM.

Cuando las partículas colisionan, los datos explotan. El gran colisionador de hadrones (GCH) del CERN produce aproximadamente un millón de gigabytes de datos por segundo. Utilizando sofisticados sistemas de selección los investigadores del CERN son capaces de filtrar todos los datos salvo los más interesantes, pero esto todavía deja a la organización con más de 25 millones de gigabytes de datos que ha de gestionar anualmente (lo

- 1 Editor de *International Science Grid This Week*, radicado en el CERN.
- 2 Jefe del grupo de Servicios de datos y almacenamiento, departamento de Tecnología de la información, CERN.



La red nunca duerme: esta imagen muestra la actividad el 1 de enero de 2013, justo después de medianoche, con casi 250 000 tareas en marcha.

que equivale a más de cinco millones de DVD). Analizar y entender estos datos de manera que tengan sentido es un reto tremendo que requiere un esfuerzo global y cooperativo. Aunque la escala del desafío que encara el CERN pueda ser mayor que el de otras organizaciones, se pueden aprender importantes lecciones sobre cómo el CERN gestiona sus datos.

El cálculo científico moderno se puede dividir en tres grandes pilares:

- **Procesamiento** – Para transformar y analizar los datos relacionados con un determinado proyecto de investigación se pueden necesitar muchos núcleos de cálculo. El análisis de datos del GCH, por ejemplo, requiere una potencia de cálculo aproximadamente equivalente a 300 000 de los más rápidos núcleos de proceso de datos de los ordenadores actuales.
- **Interconexión** – Una red rápida, incluyendo un gran ancho de banda para Internet, es vital para que los centros de investigación y los laboratorios geográficamente dispersos puedan acceder a la infraestructura de cálculo.
- **Tratamiento de los datos** – Los datos tienen que almacenarse, trasladarse en cualquier momento allá donde se disponga de recursos de procesamiento y distribuirse a un gran número de interesados (universidades, laboratorios de investigación, etc.). La conservación a largo plazo de los datos es con frecuencia también muy importante.

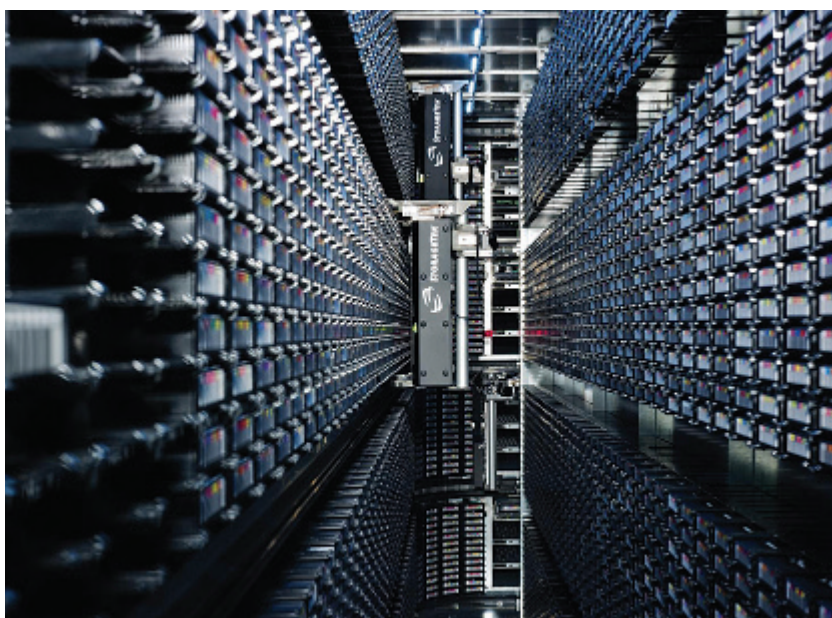
Aunque los ordenadores son herramientas indispensables en casi todos los campos de la ciencia, nunca han

sido tan fundamentales para la investigación como hoy en día. Históricamente el cálculo científico ha estado enfocado principalmente al procesamiento de la información, pero en los últimos años se ha apreciado una evolución significativa en tecnologías relacionadas con el almacenamiento, el procesamiento, la distribución y la conservación a largo plazo de los datos. El CERN ha contribuido a esta evolución tanto como se ha aprovechado de ella, logrando ser un excelente ejemplo sobre cómo tratar con éxito grandes cantidades de datos científicos.

Un ordenador mundial

El centro de datos del CERN tiene una potencia de cálculo de 3,5 megavatios y se vanagloria de un impactante número de 88 000 núcleos de procesamiento. Sin embargo, la organización solo ofrece el 15 por ciento de la capacidad de cálculo requerida para procesar la montaña de datos generados por el GCH, incluso después de que el filtrado mediante complicados algoritmos haya desechado todos menos alrededor del uno por ciento de los datos.

Desde finales de la década de 1990 ya quedó claro que la cantidad de datos que se esperaba del GCH excedía con mucho la capacidad de cálculo del CERN en solitario. Esta es la razón por la que este organismo inició en 2001 el proyecto de la Red mundial de cálculo para el GCH (WLCG) que es una solución de cálculo distribuido que conecta los procesadores de datos y las instalaciones de almacenamiento de más de 150 lugares en cerca de 40 países de todo el mundo. Iniciado en 2003, el servicio de la WLCG fue desarrollándose gradualmente mediante una serie de crecientes retos por lo que



CERN

El manejo de cartuchos de cintas magnéticas está ahora totalmente automatizado, de manera que están colocadas en cámaras de donde se mueven entre los estantes de almacenamiento y los lectores de cintas mediante brazos mecánicos. Más de 100 petabytes de datos están permanentemente archivados, lo que equivale a 700 años de películas de alta definición total.

respecta a su rendimiento, antes de ser inaugurado oficialmente en 2008. Hoy en día es utilizado por unos 10 000 físicos y, de media, se superan ampliamente las 250 000 tareas realizadas de forma simultánea.

La fuerza en números

La WLCG tiene una estructura en niveles con la copia primaria de todos los datos almacenada en el centro de datos del CERN, que suele denominarse Nivel 0. A partir de aquí el CERN envía una segunda copia de los datos a los 11 centros principales de datos en todo el mundo que conjuntamente forman el primer nivel o Nivel 1. Los centros de datos de la Federación de Rusia y de la República de Corea están adaptándose para incorporarse próximamente con lo que el número total de lugares del Nivel 1 de la WLCG aumentará a 13.

Los lugares del Nivel 1 tienen la responsabilidad de la custodia a largo plazo de los datos y proporcionan un segundo nivel de procesamiento de los datos. Cada lugar del Nivel 1 se enlaza con un determinado número de lugares de Nivel 2, normalmente situados en la misma región geográfica. En total hay alrededor de 140 lugares de Nivel 2 que normalmente son departamentos universitarios o laboratorios de física. Aunque los lugares de Nivel 2 suelen ser pequeños, actualmente distribuyen regularmente más de la mitad de los recursos totales, y es en estos lugares donde se realiza el análisis físico real.

Antes de este año el CERN y el Centro de investigación Wigner para la física inauguraron una ampliación del centro de datos del CERN en Budapest (Hungría). Alrededor de 500 servidores, 20 000 núcleos de cálculo y 5,5 petabytes de almacenamiento están ya operativos en ese lugar. Dos circuitos exclusivos y redundantes de 100 gigabits por segundo conectan los dos lugares que están operativos desde febrero de 2013 y constituyen uno de los primeros enlaces transnacionales a esta distancia. La capacidad en Wigner está administrada a distancia desde el CERN y amplía de forma sustancial las capacidades de la WLCG de Nivel 0.

Distribución de los datos

Con una solución para el cálculo muy distribuida, el factor más importante que limita el procesamiento es que los recursos de CPU puede que no estén próximos al lugar en el que se almacenan los datos. Esto conduce a la transferencia de grandes cantidades de datos, a menudo a través de Internet académico o a través de cables privados de fibra óptica, lo que a su vez conlleva que haya momentos durante los cuales algunos recursos de CPU están inactivos. En consecuencia, es imprescindible una buena estrategia en el emplazamiento de los datos si se desea maximizar la disponibilidad de los recursos de cálculo en cualquier momento.

En general la eficacia está muy influida por las estrategias de almacenamiento de datos, además de por la velocidad de transferencia de los mismos desde los

depósitos desconectados de Internet hasta los que están conectados, de los lugares remotos a los locales, de los servidores de datos a los discos locales, y de estos últimos a las memorias de los ordenadores locales. Del mismo modo, las estrategias utilizadas en la corrección de errores y en la duplicación de los datos, pueden tener importantes consecuencias tanto en la disponibilidad como en la fiabilidad de los datos. Por estas razones, es importante que en los cálculos científicos a gran escala exista una buena gestión de datos: la arquitectura, las normas y los procedimientos que gestionan todas las actividades de los datos y su ciclo de vida.

En los grandes proyectos de investigación los flujos de trabajo de los datos pueden ser extremadamente complejos: los datos se pueden duplicar en múltiples lugares para reducir la probabilidad de pérdida en caso de incidentes operativos importantes, para aumentar el número de núcleos de CPU capaces de procesarlos en paralelo o para ampliar la disponibilidad de producción para las transferencias hechas por encargo (los datos en paralelo sirven de múltiples copias).

Equipos múltiples

Los experimentos físicos del GCH en el CERN proporcionan un buen ejemplo de tales actividades en la gestión de los datos. Los datos generados por estos experimentos se trasladan regularmente de un equipo de almacenamiento a otro, ya sea para minimizar costes o porque la calidad del servicio varía entre diferentes equipos. Un equipo de almacenamiento basado en conductores de estado sólido con réplicas de datos múltiples e interconexión rápida puede, por ejemplo, utilizarse para almacenar datos brutos de valor elevado que necesitan ser analizados. Una vez que se ha completado el procesamiento de los datos, estos se pueden trasladar a depósitos de bajo coste y latencia alta en los que la fiabilidad a largo plazo es lo más importante. Del mismo modo, algunos resultados intermedios resultantes de los análisis, que pueden volver a calcularse si se pierden, probablemente se trasladarán a un equipo de almacenamiento temporal poco fiable, de latencia baja, alto rendimiento y bajo coste, equipo conocido como "espacio instantáneo".

El modelo de datos de los experimentos del GCH es, por supuesto, mucho más complejo que este resumen aproximado y hay una necesidad obvia de tener equipos de datos con diferentes calidades de servicio (en términos de fiabilidad, rendimiento y coste). También se necesita una variedad de equipos para transferir y procesar estos datos entre los equipos.

La necesidad de una arquitectura con múltiples equipos para gestionar eficientemente complejos flujos de trabajo de datos en la investigación científica se pone de manifiesto por el hecho de que incluso el ordenador más simple tiene diferentes tipos de almacenamiento: niveles 1 y 2 de almacenamiento, memoria RAM, disco duro. El funcionamiento de un ordenador a partir de un único

sistema de almacenamiento sería una simplificación que conduciría a ineficiencias.

¿Compartiendo equipos?

Esta necesidad de disponer de múltiples equipos con una calidad de servicio específica es probablemente la principal diferencia entre lo que podría llamarse “almacenamiento en la nube” y “grandes archivos de datos”, dos enfoques para gestionar grandes cantidades de datos.

- En el modelo de almacenamiento en la nube todos los datos se almacenan en un equipo uniforme, enorme y plano. Como solo hay un equipo que contiene todos los datos (normalmente con tres copias difundidas entre diversos lugares) no hay necesidad de mover los datos. Este enfoque de equipo único implica una calidad uniforme del servicio con la consecuencia de que esto rápidamente deja de ser óptimo (y, por lo tanto, poco económico) desde el momento en el que la cantidad de datos crece más allá del punto a partir del cual el almacenamiento es una parte considerable de los costes del proyecto científico completo.
- Por el contrario, el enfoque de grandes archivos de datos va más allá de un único equipo de almacenamiento, lo que introduce los conceptos de flujos de trabajo de los datos, tiempo de vida de los mismos, movimiento de datos, colocación y una estructura de almacenamiento basada en diversos equipos con diferentes calidades de servicio, con los que los costes de almacenamiento se pueden reducir significativamente y, con suerte, optimizarse.

Cualidades de los medios

Cualquiera que utilice un ordenador en casa o en el lugar de trabajo tiene una clara idea de qué se puede esperar de los discos duros. El sentido común dice que son

realmente rápidos y baratos. La mayoría de los usuarios de ordenadores portátiles o de sobremesa también dirían que son verdaderamente fiables porque nunca han sufrido pérdidas de datos debido a fallos en el disco, aunque probablemente son conscientes de que eso pudiera ocurrir. Sin embargo, en el centro de datos del CERN se tiene la oportunidad de medir la calidad de los diferentes medios de forma sistemática sobre un volumen muy grande de datos con lo que se ha llegado a alguna conclusión sorprendente.

Se ha descubierto que los discos duros son una solución rentable (0,03 euros por gigabyte) para el almacenamiento en línea (sin tener en cuenta el coste del consumo eléctrico 24 horas al día y 7 días a la semana) y que su funcionamiento es aceptable (100 megabytes por segundo en la lectura y una transmisión simple en la escritura con unos pocos milisegundos en el tiempo de búsqueda). Sin embargo, esta fiabilidad es demasiado baja para asegurar cualquier servicio aceptable sin pérdida de datos: el centro de datos del CERN tiene 80 000 discos y experimenta una tasa típica de fallos de alrededor de 5 discos de media por cada día, lo que significa una pérdida diaria de alrededor de 10 000 gigabytes y es el equivalente de alrededor de 5 000 buzones de correo electrónico de tamaño medio por día. Evidentemente esto es inaceptable y es por lo que almacenar una única copia de datos en un disco sencillo no es una estrategia viable para un servicio de almacenamiento.

Almacenamiento duradero a largo plazo

La situación que se midió para las cintas es bastante diferente. Las cintas vienen en cartuchos de 4 a 5 000 gigabytes y tienen un coste (0,02 € por gigabyte) semejante a los discos. Su principal inconveniente es la alta latencia (su elevado tiempo de acceso), de manera que puede llevar un par de minutos rebobinar una cinta, desmontarla, montar otro cartucho y avanzar hasta el



CERN

El centro de datos del CERN alberga servidores y sistemas de almacenamiento de datos no solo para el Nivel 0 de la WLGC y para otros análisis físicos, sino también para sistemas que son críticos en el funcionamiento diario del laboratorio.

lugar donde están los datos que se necesitan leer. Sin embargo, a pesar de su fama de ser lentas, una vez que una cinta ha sido montada y colocada, los datos realmente se pueden escribir o leer a velocidades que normalmente son el doble de las que se obtienen con discos duros. Además, las cintas tienen separadas las cabezas de lectura y escritura, permitiendo que los datos se lean “sobre la marcha” justo después de que se hayan escrito, proporcionando una mejora adicional con un factor igual a dos desde el punto de vista de velocidad de transferencia cuando es necesaria la verificación de datos. A diferencia de los discos duros, las cintas no consumen energía cuando no se leen o no se escriben.

Otra diferencia entre los discos y las cintas aparece cuando comparamos la fiabilidad. Cuando una cinta falla, la cantidad de datos perdidos se limita a una zona localizada en la misma quedando el resto legible. Estos “incidentes de las cintas” generan pérdidas de datos que van de unos pocos cientos de megabytes a unos pocos gigabytes, que vienen a representar tres órdenes de magnitud menos que la pérdida de datos por un fallo en un disco. Esta cifra se confirma en el centro de datos del CERN que, con más de 50 000 cintas, sufre una pérdida de datos en las cintas de apenas unos pocos cientos de gigabytes al año; algo que puede compararse con los pocos cientos de terabytes de datos perdidos debidos a fallos en discos cada año.

Los fallos en las cintas también están mucho menos correlacionados entre ellos: la probabilidad de fallo en una operación de lectura en una cinta A es bastante independiente de la probabilidad de fallo en una cinta B. Con los discos duros hay mayor correlación cuando están en los mismos servidores o bajo el mismo controlador: la probabilidad de un segundo fallo después de que se haya producido uno primero es mucho mayor que la probabilidad de un fallo inicial. La independencia de los fallos es una estrategia clave para aumentar la fiabilidad del servicio y esta independencia a menudo se pierde si se utilizan en la infraestructura servidores con un gran número de discos.

Conflictos e integridad de los datos

El centro de datos del CERN ha registrado hasta ahora más de 100 petabytes (igual a 100 millones de gigabytes) de datos físicos procedentes del GCH, lo que es aproximadamente equivalente a 700 años de películas de alta definición total. Las colisiones de partículas del GCH han generado alrededor de 75 petabytes de estos datos solamente en los últimos tres años. En el CERN, la mayor parte de los datos (alrededor de 88 petabytes) se archivan en cintas. En total, la organización tiene ocho bibliotecas de cintas robotizadas distribuidas en dos edificios y cada biblioteca de cintas puede contener hasta 14 000 cartuchos de cintas. El CERN actualmente tiene alrededor de 52 000 cartuchos de cintas. Los restantes datos físicos (13 petabytes) se almacenan en un sistema de discos duros optimizados para análisis rápidos y que son accesibles para muchos usuarios de forma

simultánea. Los datos se almacenan en más de 17 000 discos vinculados a 800 servidores.

En un sistema de múltiples equipos, como el que utiliza el CERN, es importante asegurar que no haya discrepancias entre las múltiples copias de los datos duplicados en diferentes equipos. En un mundo perfecto donde los programas informáticos siempre se comportan como se han diseñado, la duplicación basada en la información de los metadatos debería ser suficiente para asegurar un adecuado nivel de consistencia. Sin embargo, si aparecen conflictos entre conjuntos de datos en diferentes equipos, es importante tener a punto una estrategia bien definida para gestionarlos.

La seguridad es otro aspecto vital a la hora de asegurar la integridad de los datos a largo plazo. Cada petición que el servicio de almacenamiento ejecuta necesita ser adjudicada a una determinada identidad y esta identidad necesita poder ser asignada a una persona. La seguridad puede involucrar, por ejemplo, el cifrado de los datos o el establecimiento de límites en cuanto a los datos a los que determinados usuarios pueden acceder o modificar. Sin embargo, la alta seguridad requiere alcanzar un equilibrio entre el funcionamiento, la posibilidad de crecimiento y los costes.

Conclusiones

Este artículo muestra que aunque puede ser complicado distribuir petabytes de datos científicos 24 horas al día y 7 días a la semana es, sin duda, posible. Los componentes pertenecientes a estos grandes sistemas fallarán inevitablemente, de manera que es importante desarrollar sistemas resistentes. Merece la pena tener presente que llevó más de una década construir la WLCG. Durante su existencia, las redes han demostrado ser mucho más fiables y asequibles de lo que se suponía. Entidades mundiales federadas también han sido decisivas para el éxito de la WLCG, ya que los investigadores normalmente cambian de organizaciones pero quieren acceder a “sus” datos.

El CERN está a la vanguardia en la gestión de conjuntos de datos extremadamente grandes, aportando soluciones sofisticadas para el procesamiento, la distribución y el análisis de los mismos. La organización está ahora abordando retos relacionados con la conservación de datos a largo plazo. La conservación y la sostenibilidad de dichos datos requieren esfuerzos permanentes, incluso después de que cesen los fondos para los experimentos, con medidas para garantizarlos que hoy en día se están convirtiendo, de manera habitual, en un requisito de las agencias de financiación. Esta es una importante y bienvenida tendencia ya que se tiene el deber crucial de conservar con fiabilidad y eficacia y para las futuras generaciones los datos producidos por los experimentos del CERN.

Puede conocer más acerca del modo en que el CERN gestiona sus datos en este vídeo: <http://cds.cern.ch/record/1541893>