ARQUITECTURA DE UN CENTRO DE CIENCIA BASADO EN CLUSTER PARA LA CREACION DE REDES CIENTIFICAS E INTERCAMBIO DE INFORMACION

Carlos Pérez Risquet, Romel Vázquez Rodríguez, Alberto Morell Pérez, Saumel Enríquez Caro

Centro de Estudios de Informática, Universidad Central de Las Villas
cperez@uclv.edu.cu

RESUMEN

El desarrollo acelerado de las nuevas tecnologías de la información ha motivado la generación de grandes volúmenes de datos en todas las ciencias, entre las que destacan la física, la biología, la medicina y las ciencias medio ambientales. El tamaño de estos datos generalmente excede la capacidad de almacenamiento y procesamiento de cualquier estación de trabajo. Estos avances compulsan el desarrollo de nuevas tecnologías, técnicas y escenarios para el tratamiento de los datos, entre los que se destacan los Centros de Ciencia, donde se almacenan y manipulan volúmenes de datos del orden de los peta bytes. Por otra parte, la Visualización Científica se ha desarrollado en los últimos años como un área importante de la Computación Gráfica, ofreciendo nuevas posibilidades para el análisis, comprensión y comunicación de grandes volúmenes de datos e información. En este trabajo se muestran las potencialidades de los Centros de Ciencia para el desarrollo de redes científicas e intercambio de información mediante la creación de espacios personales de trabajo, herramientas de colaboración y soluciones para la publicación de los resultados de la investigación. Se presenta una arquitectura general para la creación de un centro de ciencia sobre un cluster de computadoras que permite la integración de técnicas y herramientas de visualización para el análisis y comunicación de datos.
Palabras clave: Computación gráfica, visualización científica, centros de ciencia, redes científicas.

INTRODUCCIÓN

Mediante estaciones de medición de alta precisión, imágenes de satélite o supercomputadoras se generan a diario volúmenes de datos tan grandes y complejos, que no pueden ser analizados suficientemente bien en forma numérica [1]. Los volúmenes de datos se duplican aproximadamente cada año. En general se asume que de todos los datos generados solo una cuarta parte se almacena, y de éstos a su vez solo una cuarta parte realmente se analiza. Evidentemente se pierden datos valiosos, y de muchas informaciones importantes solamente se utiliza un pequeño porciento. Son necesarios nuevos métodos de análisis para examinar estos datos de forma efectiva y extraer de ellos información y conocimientos.

Estos volúmenes de datos del orden de los peta bytes requieren de nuevos escenarios de trabajo. En ocasiones no resulta posible copiar los archivos de datos en un servidor y procesarlos con los recursos propios disponibles. La magnitud de los mismos y la complejidad de los programas que los procesan es tal, que resulta más económico mover los programas hacia los datos y transmitir solo las peticiones y las respuestas, en lugar de mover datos y aplicaciones hacia los sistemas locales de los usuarios [2].

El hombre es capaz de asimilar una mayor cantidad de información cuando esta es presentada en forma visual y no numérica. La visualización científica se ha desarrollado en los últimos años como un área de investigación bien definida, que proporciona una alternativa eficaz de solución a problemas que difícilmente pueden enfrentarse con las técnicas tradicionales existentes, convirtiéndose en una potente herramienta para la extracción de información a partir de grandes volúmenes de datos. Por otra parte, centros de ciencia que brindan acceso tanto a los datos como a las aplicaciones que los analizan están emergiendo como estaciones de servicio para algunos dominios científicos. La integración de la visualización científica y los centros de ciencia se presenta en este trabajo como una nueva alternativa para el tratamiento de grandes volúmenes de datos.

En este trabajo se presentan las potencialidades de los centros de ciencia para el intercambio de información y el desarrollo de redes de colaboración científica. Se muestran las posibilidades de integración de la visualización científica en los centros de ciencia como una alternativa eficaz para el tratamiento de grandes volúmenes de datos. Se presenta además una arquitectura general para al creación de un centro de ciencia basado en cluster de computadoras y se discuten las ventajas y limitaciones de este enfoque.

Centros de ciencia

La demanda de herramientas y recursos computacionales para realizar análisis de datos científicos crece continuamente, incluso más rápido que los propios volúmenes de datos. Esto es consecuencia de 3 fenómenos [2]:

1. Los nuevos y sofisticados algoritmos consumen cada vez más instrucciones para analizar cada byte de dato.
2. Muchos algoritmos de análisis son superlineales, a menudo de orden O(N2) o O(N3).
3. El ancho de banda de entrada-salida no ha mantenido el mismo ritmo de avance que la capacidad de almacenamiento. En la última década, mientras la capacidad ha crecido más de 100 veces, el ancho de banda solo lo ha hecho 10 veces.

Estos 3 fenómenos provocan que el análisis de los datos se demore mucho más tiempo. Para mejorar estos problemas los científicos necesitarán mejores algoritmos de análisis, que puedan manejar inmensos conjuntos de datos con algoritmos aproximados (unos con tiempo de ejecución cercano al lineal), y serán necesarios algoritmos paralelos, que puedan utilizar muchos procesadores y discos para equiparar las demandas de densidad de CPU con el ancho

de banda.
Los centros de ciencia o centros de datos científicos, que están surgiendo como estaciones de servicio para algunos dominios científicos y suministran acceso tanto a los datos como a las aplicaciones que los analizan, se presentan como una vía para solucionar estas crecientes demandas de recursos computacionales.

Conjuntos de datos del orden de los peta bytes ocupan de 1000 a 10 000 discos, y requieren de miles de unidades de procesamiento. Un centro de ciencia es capaz de almacenar uno o más de estos grandes conjuntos de datos y ofertar aplicaciones que acceden a ellos y los analizan. Posee además, un equipo de personas que entiende los datos y está constantemente mejorándolos y agregando mas información [2].

Algunos ejemplos de estos centros de ciencia son: SDSS (Sloan Digital Sky Survey) en Fermilab (Laboratorio Nacional del Acelerador Fermi), BaBar en SLAC (Stanford Linear *Accelerator Center), BIRN (Biomedical Informatics research Network) en SDSC (San Diego Supercomputer Center), EntrezPubMedGenBank en NCBI (National Center for Biotechnology Information), y con muchos otros conjuntos de datos que abarcan otras* disciplinas.

El nuevo estilo de trabajo en estos dominios científicos es enviar solicitudes a las aplicaciones que están corriendo en el centro de datos, y recibir de vuelta las respuestas. Esto es mejor que hacer una copia de los datos hacia el servidor local para su posterior análisis con recursos propios.

Una tendencia actual que soportan muchos centros de ciencia es ofertar la posibilidad de almacenar espacios de trabajos personales (una base de datos personal) en el centro de datos, y guardar allí las respuestas a las consultas realizadas. Esto minimiza el movimiento de los datos, y permite la colaboración entre grupos de científicos que hacen análisis sobre temas similares. Estos espacios de trabajo son también una vía para la colaboración entre grupos de analizadores de datos. A largo plazo los espacios de trabajos personales de un centro de datos pueden convertirse en una forma para la publicación de datos, mostrando tanto los resultados científicos de un experimento o investigación, como los programas usados para generarlos.

Los centros de ciencia proporcionan herramientas y soporte para permitir la colaboración con otros centros de ciencia. Cuando un científico desea consultar datos de varios centros de ciencia, no quedaría otra opción que mover partes de los datos de un lugar a otro. Si esto se torna común, los centros de datos involucrados deben colaborar para suministrar la salva de los datos de manera mutua, ya que el tráfico entre ellos justifica disponer de las copias.

Muchos científicos prefieren hacer gran parte de su análisis en los centros de datos, ya que les ahorrará tener que administrar datos locales y granjas de computadoras. Algunos científicos pueden llevar extractos de datos a sus casas para hacer procesamiento, análisis y visualización

de manera local, por lo que los centros de datos deben brindar la posibilidad de que los investigadores hagan salvas de partes de los datos hacia sus estaciones de trabajo. Sin embargo todo el análisis será posible hacerlo en el mismo centro de datos utilizando espacios personales de trabajo que permitan a los investigadores almacenar los datos con los que trabajan en bases de datos personales.

Para almacenar y manipular conjuntos de datos del orden de los peta bytes se requiere de 1000 a 10000 discos y miles de nodos. En cualquier momento algunos de los discos de algunos de los nodos pudieran colapsar. Estos sistemas tienen que tener un mecanismo para evitar la pérdida de datos, y proporcionar buena disponibilidad incluso con menos de la configuración completa por lo que requieren un sistema de auto-sanación. Se hace necesario disponer de un sistema de replicación para distribuir los datos del centro de ciencia en distintos lugares geográficos.

Existe una tendencia de utilizar lo formatos de datos cientificos (HDF, NetCDF, FITS, ..) a centrarse principalmente en aspectos de los metadatos e intercambio de datos, y los sistemas tradicionales de gestión de datos (SQL y otros) son mas utilizados en la gestión y el análisis de grandes conjuntos de datos. Los sistemas de bases de datos tradicionales tienen varias virtudes como paralelismo automático, indexación, acceso no procedural, pero tienen que adoptar los tipos de datos de la comunidad científica y necesitan coexistir con los datos en los sistemas de archivos. Se espera que los sistemas de gestión bases de datos se extiendan para unificar las bases de datos con los lenguajes de programación de forma tal que se puedan incluir o enlazar nuevos tipos de objetos dentro de los sistemas gestores de bases de datos.

Para el análisis científico va a ser de gran importancia la utilización de 3 técnicas avanzadas: (1) Uso de grandes metadatos y metadatos estándares para facilitar el descubrimiento de los datos que existen, y que les sea cómodo tanto para las personas como para los programas entender los datos y se pueda saber fácilmente todos los procesos por los que han pasado los datos. (2) Buenas herramientas de análisis que les permitan a los científicos de forma fácil hacer las consultas, entender y visualizar las respuestas. (3) Permitir el acceso paralelo a los datos soportados por los nuevos esquemas indexados y los nuevos algoritmos que nos permitan interactuar y explorar conjuntos de datos del orden de los petabytes.

Como consecuencia, los centros de ciencia se mantendrán como la principal vía para la comunicación de datos e información entre las asociaciones mundiales y suministraran tanto los datos como la infraestructura para manipular, analizar, procesar, visualizar y publicar estos archivos de gran escala así como los algoritmos y herramientas.

Comparación de varios CC respecto a formas de intercambio de datos.

La mayoría de los centros de ciencia se relacionan y cooperan con muchos centros en lugares geográficos diferentes. Por ejemplo Babar tiene unos 25 sitios colaboradores, y NCBI tiene

también varios sitios compañeros que cooperan en varios aspectos relacionados con él. SDSS es parte del Observatorio Virtual Internacional.

Sloan Digital Sky Survey (SDSS):
SDSS SkyServer, es un proyecto para confeccionar un mapa de una gran parte del universo. SDSS posee una herramienta online llamada CasJobs que permite el acceso a la gran base de datos de este centro de ciencia a través de los catálogos científicos. Esta herramienta está diseñada para emular y mejorar el acceso libre a las consultas en un entorno web [4].

Algunas funciones de esta aplicación son:
- Permite ejecutar consultas de forma sincrónica y asincrónica, en forma de tareas rápidas o demoradas.
- Guarda un historial que registra las consultas y su estado.
- Posee una base de datos personal llamada “MyDB” en el lado del servidor que permite la creación de tablas, funciones y procedimientos. En esta base de datos personal se almacenan partes de los grandes volúmenes de datos del centro de ciencia que el usuario está investigando.
- Permite el compartimiento de datos entre los usuarios, a través un mecanismo de “grupos de usuarios”.
- Permite descargar en diferentes formatos los datos que están la base de datos personal MyDB.
- Presenta varias opciones de interfaz, incluyendo un navegador cliente, así como una herramienta de línea de comandos.

Este centro de ciencia posee además otro conjunto de herramientas dentro de las que se encuentran, la herramienta Famous Places que presenta una galería de imágenes de varias galaxias y algunas de las estrellas mas distantes que se han descubierto. La herramienta Get Images que permite descargar imágenes individuales con varios niveles de resolución. Otras herramientas no menos importantes son Scrolling Sky, Search, Object Crossid, Help y Download. Además SDSS presenta varias herramientas de visualización (Visual Tools) como son Finding Chart que devuelve una imagen jpg con varios parámetros aplicados a los datos, Navigate que permite la navegación interactiva por lo datos del universo y otras como Image List , Quick Look, y Explore, este ultimo permite explorar interactivamente varias propiedades de un objeto individual.

El Sloan Digital Sky Survey identificará la ubicación de más de 100 millones de galaxias. Además mediante la combinación de los datos que se encuentran en los catálogos o el reanálisis de cada uno de los píxeles de cada Imagen se pueden obtener nuevos objetos o generar estadísticas en forma de galaxias. Este último enfoque exige el flujo de imágenes desde su ubicación de almacenamiento a través de una plataforma de procesamiento. Para facilitar el análisis, las colecciones han sido replicadas en NSF Teragrid que proporciona tanto

los recursos de cómputo como de almacenamiento.

Sky Portals se utiliza para integrar la información de múltiples catálogos, en efecto, soportando la distribución de las consultas a través de los múltiples catálogos y haciendo acoples para generar la respuesta deseada a la consulta.

Los metadatos utilizados para describir los datos se basan en Descriptores uniformes de contenido para los catálogos de entradas. La astronomía ha creado una comunidad Estándar de formato de datos para las imágenes, llamado FITS estándar. Esto proporciona una manera de encapsular metadatos descriptivos acerca de la ubicación, la resolución, y el alcance de cada imagen, junto con los píxeles que componen la imagen.

Biomedical Informatics Research Network (BIRN):

BIRN es otro centro de ciencia que está compuesto por una gran cantidad de sitios, principalmente de Estados Unidos y Gran Bretaña, unidos todos con el objetivo de diseñar e implementar una arquitectura distribuida de recursos compartidos que sean usados por los investigadores de la Biomedicina para avanzar en el diagnostico y tratamiento de enfermedades.

La colaboración entre los diversos sitios que componen BIRN suministra al resto de la comunidad científica los siguientes aspectos [5]:

- Intercambio de datos: La colaboración entre los múltiples sitios permiten desarrollar infraestructuras para compartir los datos y definir políticas sobre los conjuntos de datos disponibles para la comunidad científica.
- Interoperatibilidad: La colaboración entre los diversos sitios puede motivar la interoperatibilidad de las herramientas de análisis desarrolladas en los diferentes laboratorios.
- Estándares de comparación: Esta colaboración puede llevar al desarrollo de estándares de los nuevos formatos de datos y protocolos de análisis. Estas ayudas mueven el campo hacia delante estableciendo las marcas del mismo campo de trabajo.
- Sinergia: La colaboración puede fomentar la sinergia que significa poder producir más con la unión de varios sitios que lo que pudiera producirse en cada uno por separado. La combinación de fuerzas de cada sitio puede suministrar recursos más robustos que los que están disponibles en cada sitio.

BIRN posee una gran cantidad de herramientas para:

- Adquisición, calibración y control de calidad
- Almacenamiento y administración de datos,
- Análisis y procesamiento,
- Visualización y construcción de mapas cerebrales, etc,
- Construcción de estándares, terminologías y ontologías
- Intercambio de datos y colaboración

Las comunidades de investigación biomédicas trabajan con grandes conjuntos de datos y a menudo altamente heterogéneos. BIRN provee una manera para integrar estos datos, permitiendo analizar los datos, distribuir la información para realizar nuevos descubrimientos. El BIRN Virtual Data Grid (BVDG) almacena los datos a través de un ambiente distribuido, donde estos son continuamente administrados y manipulados como un simple sistema virtual de archivos. Actualmente, dentro del BVDG existen más de tres millones de ficheros de datos, activos, que son accesibles directamente a través de un conjunto de interfaces estandarizadas.

A través de BIRN se puede compartir o publicar tanto datos como resultados de las investigaciones usando el BIRN Data Repository (BDR), el cual se rige por un conjunto de normas que tienen aspectos de publicación, garantía, seguridad y redistribución.

Los científicos pueden acceder desde cualquier ordenador que tenga acceso a Internet a través de un simple portal web que provee un conjunto de herramientas integradas, infraestructura, y servicios necesarios para realizar estudios de gran magnitud y en colaboración. Además este centro tiene la capacidad de autenticar todas las personas que acceden a las imágenes, la gestión de los controles de acceso a todos los archivos de datos independientemente del lugar donde los datos están almacenados, la gestión de los controles de acceso a los metadatos, para proporcionar pistas de auditoría, todos los accesos a los datos de forma independiente de la ubicación de almacenamiento, y el apoyo de cifrado de datos. En la práctica, la auditoría se utiliza para demostrar la cantidad de intercambio de datos entre los sitios de colaboración. La cantidad de datos remotos visitados por un investigador puede ser cuantificada, así como la cantidad de datos que el investigador distribuye a otros sitios.

BaBar:

La comunidad de Física de Altas Energías, tradicionalmente, ha venido acumulando experiencias en el almacenamiento y análisis de grandes volúmenes de datos, teniendo como uno de sus principales exponentes el centro de ciencia BaBar. Sólo la comunidad del experimento de BaBar aglutina a más de 600 físicos e ingenieros de 75 instituciones a escala mundial y ha logrado consolidar una de las bases de datos orientadas a objetos más grandes del mundo que tenia en noviembre del 2004 casi 900 TeraBytes. BaBar está orientado a la exploración de los principios físicos de la energía, la materia y la antimateria, utilizara recursos computacionales en modelo de GRID para aumentar la velocidad y la calidad en las investigaciones mundiales en este tema. El LHC (Large Hadron Collider) Computing Grid Project (LCG) ha adoptado las herramientas computacionales y ha creado grupos de evaluación y prueba conjunta con el Proyecto Europeo DataGrid. La organización para la operación del LCG es jerárquica y está compuesta por Tiers, donde la raíz de esta jerarquía está en el CERN (Tier 0), luego nodos regionales (Tier 1) en IN2P3 Lyon-Fancia, PIC Barcelona-España, CNAF en Boloña-Italia, PPARC en Rutherford-Inglaterra, por mencionar los principales nodos europeos; luego seguirían nodos menores de servicios. Cada uno de estos nodos deberá

realizar labores de limpieza de datos y proveer capacidades de cómputo para el análisis de esos resultados [3].

BaBar usa el Storage Resource Broker (SRB) desarrollado por el centro de súper computadoras de San Diego (SDSC) herramienta que suministra una interfaz uniforme para acceder a sistemas de almacenamientos heterogéneos (discos, cintas, bases de datos) que estén distribuidos en varios sitios, SRB además soporta herramientas para compartir archivos de forma colaborativa, permite el control de replicas por lo que es muy fácil duplicar una BD de un sitio para otro, y proporciona búsquedas por atributos de los datos y búsquedas por metadatos.

EntrezPubMedGenBank:
EntrezPubMedGenBank es un centro de ciencia que se dedica fundamentalmente a la Bioinformática, medicina y biotecnología también esta compuesto por un grupo sitios pero con objetivos diferentes [6]. Presenta un conjunto de servicios GRID que pueden ser accedidos a través de un cliente Desktop de ese GRID llamado Taverna, el cual se encuentra libre en Internet. Este cliente cuando se ejecuta, automáticamente reconoce todos los servicios que hay en los diferentes sitios del centro de ciencia y permite el uso de estos servicios mediante diagramas de workflow. Este centro brinda servicios relacionados con el genoma humano y la genética en general, también brinda servicios de proteínas, moléculas, encimas y otros servicios generales implementados sobre el protocolo SOAP (Simple Object Access Protocol). El intercambio de datos se realiza a través de los diferentes servicios.

El Centro Nacional para la Información Biotecnológica o National Center for Biotechnology Information (NCBI) es parte de La Biblioteca Nacional de Medicina de Estados Unidos y una rama de los Institutos Nacionales de Salud (National Institutes of Health o NIH). Está localizado en Bethesda, Maryland y fue fundado el 4 de noviembre de 1988 con la misión de ser una fuente importante de información de biología molecular. Almacena y actualiza constantemente varias bases de datos biológicas de acceso público. Entre las más conocidas y populares se encuentran las bases de datos de publicaciones científicas (PubMed), de secuencias de proteínas y ADN (GenBank), de estructuras tridimensionales de proteínas; y algunas otras no tan populares como OMIM (Online Mendelian Inheritance in Man).
El NCBI desarrolló Entrez como una herramienta para permitir a los usuarios interaccionar con estas bases de datos. Desde el punto de vista informático, Entrez es una 'interfaz de usuario', es decir, constituye el nexo entre el usuario y las bases de datos subyacentes. Como interfaz, Entrez permite al usuario realizar consultas simples y obtener resultados, aun desconociendo la arquitectura de las bases de datos. Todas las bases de datos del NCBI están disponibles en línea de manera gratuita, y pueden ser accedidas usando el buscador Entrez.

El sistema de búsqueda PubMed es otro proyecto desarrollado por (NCBI) en el National Library of Medicine (NLM). Permite el acceso a las base de datos bibliográficas de NLM: MEDLINE,

PreMEDLINE (citas enviadas por los editores antes de su publicación) y AIDS. Se pueden establecer alertas de correo electrónico automático para los nuevos artículos agregados a PubMed. Además tiene la opción de guardar las búsquedas, y una vez que ha guardado la búsqueda, puede escoger repetir la búsqueda en el futuro o bien que el sistema ejecute la búsqueda automáticamente y le envíe los resultados por correo electrónico.

Con el objetivo de mantener una buena escalabilidad, este centro fue diseñado para usar una arquitectura distribuida. El sistema consta de una base de datos central de autentificación o de entrada y múltiples bases de datos distribuidas por categorías. Principalmente 8 categorías disponibles. La base de datos central de autentificación provee el URL de cada una de estas ocho bases de datos. Las ocho instancias de la base de datos comparten un esquema idéntico de base de datos, dejando una implementación genérica para el servicio Web. Las bases de datos de categorías adicionales pueden fácilmente ser incorporadas en el sistema existente con una adición mínima en la base de datos central de autentificación, no requiriendo cambio en la capa de servicio Web.

La base de datos del GenBank crece de manera exponencial, este crecimiento es debido a la forma misma en que la base se actualiza. Son los mismos autores quienes se encargan de mantener la base al día, pero además de remisiones de autores, el GenBank se nutre también de las otras bases de datos existentes actualizando interactivamente sus ficheros.

Existen muchas interfaces de acceso al GenBank, pero sin duda alguna la más efectiva es Entrez. Esta es una interfase que brinda acceso a muchas de las bases de datos mantenidas por el NCBI. Haciendo una búsqueda a través del Entrez es posible llegar a múltiples fuentes de información acerca del mismo tema, por ejemplo es factible para una secuencia encontrar su listado de citaciones bibliográficas contenidas en el Medline.

Un servicio de NCBI (http://www.ncbi.nlm.nih.gov) que vale la pena resaltar: el mapa genético del el genoma humano, este es un compendio de aproximadamente 30,261 secuencias.

Arquitectura basada en cluster de computadoras

En este epígrafe mostramos los aspectos asumidos para el diseño de la arquitectura así como un diagrama de alto nivel explica la arquitectura de un centro de ciencia basado en cluster para la visualización científica.

Aspectos asumidos para el diseño.

El servicio más importantes que debe brindar nuestro centro es la visualización científica, primeramente debemos tener en cuenta que un CC para la visualización tiene más accesos de lectura de los datos que de escritura en los discos, algo que es común en los sistemas de visualización. El tiempo de respuesta de este servicio debe ser lo más rápido posible, debido a que cuando se visualiza un conjunto de datos es con el objetivo de interpretar la imagen, ajustar

algunos parámetros y volver a visualizar, por lo que es posible que un usuario tenga que realizar este proceso varias veces hasta obtener el resultado deseado, si no se logra un buen tiempo de respuesta esto podría causar que el usuario desista de utilizar el sistema.

El diseño debe contar con buena capacidad de procesamiento y capacidad de almacenamiento, los cuales son requerimientos de los CC, aspectos que utilizados de forma correcta en un cluster de computadoras con discos duros se pueden cumplir perfectamente. Además la escalabilidad del cluster permite adicionar más nodos y discos duros en caso que se necesite aumentar la capacidad de cómputo o de almacenamiento en cualquier momento.

Un aspecto fundamental es la conservación de la integridad de los datos almacenados en el Centro de Ciencia, por lo cual el sistema debe ser tolerante a fallos. En caso que ocurra algún problema de perdida de información, debemos tener mecanismos de auto recuperación, que permita mantener la integridad de los datos. El uso de la replicación es una alternativa; con los archivos replicados al ocurrir cualquier fallo en el sistema, ya sea la perdida de datos por cualquier motivo como la rotura de un disco duro, el sistema debe detectar estos problemas y si es posible solucionarlo de forma automática, reestableciendo los datos a partir de un disco espejo o en caso que no pueda resolver el problema, informarle al administrador el problema detectado.

Como ya se mencionó anteriormente un requisito importante de los CC es brindar la posibilidad a los usuarios de tener espacios personales de trabajo, donde puedan almacenar datos que con más frecuencia utilizan, y donde puedan guardar los resultados de sus investigaciones. A través de estos espacios de trabajos personales un usuario del centro puede compartir sus datos con otros usuarios. Por lo que el espacio de trabajo personal será un aspecto clave del diseño.

El manejo de Catálogos de metadatos en sistemas distribuidos, facilita mantener la ubicación de cada uno de los archivos en los diferentes nodos del cluster, así como otras informaciones importantes de los datos como fecha, el propietario y cualquier característica que sea de interés para los usuarios, sin entrar en detalles de los datos en sí.

Los metadatos descriptivos de los datos son necesarios para comprender los datos, localizar conjuntos de datos dentro de un archivo y saber por los distintos procesos que estos han pasado.

Aspectos del diseño de un CC soportados sobre la tecnología de cluster de computadoras.

En un cluster de computadoras se pueden implementar prácticamente todas las funcionalidades de los centros de ciencia, pero creemos necesario excluir algunas condiciones de heterogeneidad que poseen los CC para ajustarnos a las condiciones de un cluster de computadoras con similares configuraciones. Por ejemplo limitar las funcionalidades de un

midleware para que en lugar de manipular diferentes sistemas de archivos y sistemas operativos, solo se dedique a una configuración con condiciones homogéneas.

Con el objetivo de visualizar grandes volúmenes de datos de manera rápida y eficiente, es fundamental explotar la potencia de procesamiento del cluster, entiéndase gran volumen de datos, como un conjunto que no puede ser visualizado de manera eficiente por un solo ordenador. Es por eso que dentro del modulo de visualización tendremos que implementar técnicas de visualización distribuida, visualización paralela y renderizado paralelo. Debemos señalar que esta es una actividad bastante difícil y requiere un análisis profundo y que podemos contar con softwares de ayuda en la construcción de clusters para renderizado [7]. Dentro de ellos tenemos a las herramientas Chromium, ParaView, ambas de código abierto y el HDF5 paralelo que como su nombre indica permite la visualización en paralelo de archivos HDF5.

Es necesario mostrar dos variantes que se presentan en cuanto a la forma de almacenamiento cuando deseamos manipular datos científicos:

- Utilizar BD solo para almacenar los Catálogos de Metadatos y los archivos en formato de datos científicos (HDF, CDF, NetCDF, FITS) almacenarlos en los discos duros del cluster.

- Tener estructuras para almacenar objetos de formatos de datos científicos (HDF, CDF, NetCDF, FITS) en un Cluster de BD (Bases de datos objeto relacional, bases de datos orientadas a objetos, etc).

Según la primera variante, tenemos una base de datos para almacenar los Catálogos de Metadatos, que una vez que un usuario se autentifica le permite ubicar los datos que están replicados en lo distintitos discos del cluster. Con esta variante la BD de catálogos de metadatos tiene el control de todos los datos del sistema, en caso que algún dato sea movido de lugar, hay que actualizar el catalogo.

Además con la facilidad de brindar espacios de trabajo personales, cuando un usuario se autentifica, el catalogo debe tener el control sobre el cliente, es decir permitir que el usuario tenga la localización de sus datos, mientras que los grandes volúmenes de datos están almacenados en los discos duros del cluster, replicados entre varios nodos según la disponibilidad.

Los metadatos descriptivos estarían dentro de cada archivo. Con esta variante podemos utilizar las facilidades que brindan algunos formatos de datos de trabajar en paralelo como el HDF paralelo, disponible a partir de la versión HDF5.

En la segunda variante tenemos estructuras para almacenar objetos de formatos de datos

científicos como (CDF, NetCDF, FITS) en un cluster de base de datos, BD que pueden ser objeto relacional o orientadas a objetos, utilizando el gestor de BD que más nos convenga o varios gestores en caso que sea necesario.

En ésta variante el tratamiento de los catálogos de metadatos es similar a la primera. Una BD de catalogo de metadatos que puede estar replicada. Mientras que el tratamiento de los metadatos descriptivos es un poco más complicado, porque no podemos explotar las facilidades de los formatos de datos mencionados, que los mismos archivos tienen sus metadatos en la parte superior del archivo, una variante es tener un objeto en forma de tabla asociado a cada archivo también en forma de tabla que contenga sus metadatos descriptivos.

Si el trabajo con los metadatos descriptivos es un inconveniente para esta variante de almacenamiento, la posibilidad de brindar espacios de trabajo personales o BD personales es muy fácil. Simplemente creando para cada usuario una BD con estructura similar a las que almacenan los grandes conjuntos de datos, que seria su BD personal, donde puede almacenar los datos de mayor interés para él.

En un cluster al igual que en los centros que utilizan tecnología GRID podemos implementar buenos mecanismos de control de fallos para evitar la perdida de información y mantener la integridad de los datos. Estableciendo políticas de seguridad evitaremos cualquier problema que pueda causar un usuario ajeno al sistema, como la introducción de códigos malignos, por eso implementando un buen sistema de autentificación tendremos el control de los usuarios que acceden a los datos de nuestro centro. Utilizando la replicación de los datos, mantendremos la integridad de los mismos. Haciendo una suma de chequeo de cada conjunto podemos determinar si ocurrió algún cambio o modificación que requiera la recuperación o actualización y se recuperaría copiando desde una fuente similar. Vale destacar que debemos tener un mecanismo de actualización, que sea capaz de propagar las actualizaciones de los datos a todos sus conjuntos replicados, así como a las bases de datos personales.

El hecho de contar con un cluster de computadoras implica buena capacidad de procesamiento, pero para explotar al máximo la sinergia en el cluster necesitamos una buena programación de paso de mensajes, utilizando principalmente MPI, para lograr algoritmos capaces de repartir las tareas entre varios procesadores, así como balancear la carga a medida que los nodos se hacen disponibles después de terminar su tarea asignada.

El uso de las BD paralelas puede ser un buen mecanismo para aumentar la capacidad de cómputo.

Una característica importante de los Centros de Ciencia es la colaboración entre varios de ellos. El cluster de computadoras no interviene directamente en esta tarea, el midleware es el que debe encargarse de tener enlaces con otros centros existentes, por lo que esto no seria un

problema, bastaría con conectar varios centros y establecer los mecanismos de comunicación de datos.

Diseño del modelo de arquitectura para los servicios y aplicaciones brindados por un Centro de Ciencia.

El diseño de la arquitectura de nuestro Centro de Ciencia en un modelo lógico está formado por tres capas fundamentales, como se muestra en la figura:

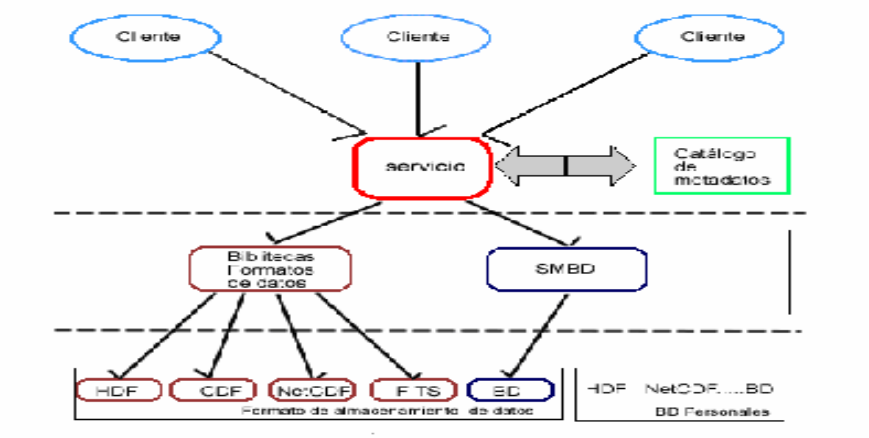

Figura Modelo Lógico del diseño del Centro de Ciencia

En la parte superior tenemos la capa cliente, en un nivel más bajo la capa manejadora de los datos y por ultimo la capa de almacenamiento.

Los usuarios del centro de datos desde su estación de trabajo o computadora personal se autentifican en el sistema mediante un servicio. Localizan los datos en el catálogo de metadatos y acceden a los servicios que brinda el centro. Una ves que están dentro del sistema, pueden acceder a los datos (nivel más bajo) a través de una capa intermedia donde están las bibliotecas de manejo de datos que permiten trabajar con los distintos formatos de datos científicos (HDF, NetCDF, FITS). En esta capa intermedia también se encuentran los sistemas manejadores de bases de datos, para trabajar con los datos que se encuentran almacenados en BD en el mismo nivel de almacenamiento. Ambos tanto bibliotecas de formatos de datos como sistemas manejadores de BD también son necesarios para acceder a los espacios de trabajo personales, que como se puede observar en la figura, están el mismo nivel de los datos.

A continuación se explican los tres niveles principales del diseño: Capa cliente, Capa manejadora de los datos y Capa de almacenamiento de los datos.

Capa Cliente

La capa superior o capa cliente, permite que múltiples usuarios se conecten y trabajen

simultáneamente en el sistema, accediendo a través de un servicio que les permite autentificarse mediante una interfaz común para todos. Una vez que un usuario se autentifica puede ver los servicios que están disponibles en ese momento o que sus privilegios le permitan acceder:

Entre los principales servicios para los cuales estará diseñado nuestro centro de ciencia se encuentra la visualización. Una vez que un usuario se autentifica puede visualizar datos directamente desde el repositorio maestro o desde su espacio personal, las herramientas de visualización también las brindaría el centro. Además el centro ofrece la posibilidad de realizar consultas complicadas a los datos, mediante el uso del cluster y de la replicación, este trabajo puede ser muy eficiente y reducir considerablemente el tiempo con respecto a un servidor de base de datos normal, principalmente por las ventajas que ofrecen las bases de datos paralelas.

En el sistema hay que manejar distintos grupos de usuarios: Grupo de administradores, grupo de usuarios normales y grupo de usuarios temporales.

Los administradores del sistema tienen el control total de Centro de Ciencia, ellos puedes acceder directamente para publicar, mover, actualizar o eliminar datos.

Los usuarios normales serían investigadores de algún área científica que necesiten acceder al centro para aprovechar las facilidades de los servicios que este brinde. Ellos se registrarían al sistema para consultar los datos que ya están almacenados en el centro. En caso que lo deseen pueden tener un espacio de trabajo personal, donde guardarían los datos de mayor interés o que usan con más frecuencia, en su espacio pueden guardar los resultados de su investigación y hacerlos públicos para el resto de los usuarios. También dentro de este tipo de usuarios puede que un grupo de ellos que investiguen el mismo tema quieran tener un espacio de trabajo común para colaborar entre ellos y evitar la redundancia. En caso que estos usuarios necesiten publicar nuevos datos que no se encuentren en el sistema, deben ser autorizados por los administradores del centro, que se encargarían de examinar esos datos. Los usuarios temporales serian usuarios que necesitan utilizar el sistema ocasionalmente, para realizar alguna búsqueda o obtener alguna imagen de un grupo de datos. Estos usuarios no tendrían espacios de trabajo personal, y su principal objetivo sería con fines educativos.

Otro elemento importante de este nivel es la BD de catálogo de metadatos. Como su nombre indica es una base de datos que almacena las direcciones de todos los recursos que están almacenados en el sistema, localización de los datos, espacios personales, y otras facilidades que ayude al cliente a encontrar lo mismo datos que recursos sin ser un experto en el funcionamiento del centro, ni conocer los detalles de los datos. Cuando un usuario se autentifica, casi siempre lo primero que realiza es consultar esta BD.

Capa Manejadora de Datos.

En la figura se puede observar la capa intermedia, que brinda las herramientas necesarias para poder manejar los distintos tipos de datos. Esta capa consta de dos partes bien definidas, una donde están las bibliotecas para el manejo de los distintos formatos de datos (HDF, NetCDF o otros), y la otra parte formada por los sistemas manejadores de bases de datos. Ambos recursos permiten el trabajo en paralelo con los datos y el acceso de múltiples clientes simultáneamente al repositorio de datos.

Las bibliotecas para el manejo de los datos (HDF, NetCDF), tienen el soporte que permite el trabajo en paralelo con los distintos archivos almacenados en el centro de ciencia. Se hace uso de las aplicaciones de paso de mensajes MPI para poder distribuir las tareas entre los nodos de procesamiento del cluster y establece el orden en que van a ser procesadas.

Estas bibliotecas de trabajo con los distintos formatos de datos permiten realizar las siguientes operaciones con los datos:

- Crear, abrir y cerrar un archivo.
- Crear, abrir y cerrar un conjunto de datos.
- Extender o reducir un conjunto de datos.
- Escribir o leer desde un conjunto de datos.

El trabajo con los datos puede de forma colectiva o independientemente. Una vez que un archivo es abierto por los procesos de comunicación:

- Todas las partes del archivo son accesibles por todos los procesos.
- Todos los objetos del archivo son accesibles por todos los procesos.
- Múltiples procesos pueden escribir en el mismo conjunto de datos.
- Cada proceso puede escribir en un conjunto de datos individual.

La comunicación mediante MPI (Message Passing Interface), permite la comunicación de un grupo de procesos entre si, esta comunicación tiene que realizarse en el orden correcto. Primeramente se abre un archivo paralelo con un comunicador, después obtiene el manejador del archivo que va a ser usado para los accesos posteriores.

En el caso de los Sistemas Manejadores de Bases de Datos (SMBD) es mucho más sencillo porque estos son mas robustos que las bibliotecas de los formatos de datos y la mayoría de ellos incorporan una gran cantidad de tareas que realizan sin que el usuario se tenga que preocupar. El usuario se tendría que concentrar en la formulación de las consultas.

Capa de Almacenamiento.

La capa de almacenamiento es la encargada de soportar los formatos de datos físicos y las bases de datos en si. Como se puede observar en la figura esta capa esta dividida en 2 secciones la de la izquierda significa que se tendrán formatos de datos científicos o BD para almacenar los grandes conjuntos de datos. La sección de la derecha significa que cada base de

datos personal pudiera en principios tener las mismas funcionalidades para los almacenar formatos de datos o BD pero en menor escala.

CONCLUSIONES

En este trabajo se han presentado las potencialidades de los Centros de Ciencia para el desarrollo de redes científicas e intercambio de información mediante la creación de espacios personales de trabajo, herramientas de colaboración y soluciones para la publicación de los resultados de la investigación.

Se presenta una arquitectura general para la creación de un centro de ciencia sobre un cluster de computadoras que permite la integración de técnicas y herramientas de visualización para el análisis y comunicación de datos.

La integración de la visualización científica con los centros de ciencia se propone como una tendencia que debe ganar auge en los próximos años.

REFERENCIAS BIBLIOGRÁFICAS

Frühauf, T.: “GraphischInteraktive Strömungsvisualisierung”, Springer Verlag, 1997.

Gray, J. et al: “Scientific Data Management in the Coming Decade”, CTWatch *Quarterly, Volume 1, Number 1, 2005.*

Sitio de Babar http://www-public.slac.stanford.edu/babar/

Sitio de SDSS http://skyserver.sdss.org/

Sitio de BIRN http://www.nbirn.net/

Sitio de GenBank http://www.ncbi.nlm.nih.gov/

Hansen, C. D. and C. R. Johnson. The Visualization Handbook, 2005.

Reagan, W. Moore: Digital Libraries and Data Intensive Computing. San Diego Supercomputer Center