

Arquitectura para Consulta de Imágenes Médicas Anotadas en un GRID: Servicios Basados en Ontologías

Jesús Campos¹, Yudith Cardinale¹ y Emilio Hernández¹

¹Universidad Simón Bolívar, Venezuela

RESUMEN

En este artículo se propone una arquitectura para el almacenamiento de imágenes médicas anotadas ontológicamente dentro de un GRID. De las anotaciones se podrá extraer información que será representada ontológicamente utilizando RDF. Se propone la implementación de servicios en el GRID asociados al desarrollo de metodologías para anotar ontológicamente fuentes de datos contenidas en los Elementos de Almacenamiento (“Storage Elements”) de un Grid de Datos, así como el desarrollo de servicios de consultas que respondan al criterio de las perspectivas de conocimiento del usuario. Se plantea el almacenamiento de las imágenes y la integración de los servicios considerando las facilidades de gLite, un middleware de GRID.

ABSTRACT

This article proposes an architecture for storing ontologically annotated medical images in a GRID. We will be able to extract information from the annotations, which are represented ontologically using RDF. We propose to implement GRID services associated with the development of methodologies for ontological annotations of data contained in a Data Grid (specifically on the Storage Elements) as well as the development of query services based on the knowledge perspective of the users. We consider issues related to image storage and service integration, using the facilities of the gLite GRID middleware.

Keywords: Grid Computing, Healthgrid, Ontology, Medical Image, gLite, AMGA.

1. Introducción

Grandes volúmenes de datos, tanto estructurados como no estructurados, exigen novedosos esquemas que permitan la consulta y el procesamiento de los mismos. Concretamente, tanto el área de la salud como la biomedicina han incrementado significativamente el uso y la gestión de datos, incluyendo cada vez más las imágenes (estáticas o videos) asociadas a cada uno de los pacientes que forman parte de las bases de datos clínicas.

Ya desde el año 2002 hay reportes de la necesidad de tener una enorme capacidad de almacenamiento y procesamiento de datos en el área de salud [MGVG]. Esto implica tener que tomar decisiones acertadas en cuanto al manejo de las fuentes de datos, con el objeto de ser utilizadas no solamente en la atención de los pacientes, en ambientes clínicos, sino también para ser debidamente compartidas en el sector de la investigación científica.

Justamente, y como parte del estudio que motiva el presente trabajo, se identifican al menos dos tipos de servicios que requieren acceder apropiadamente a estos repositorios de datos. Estos dos tipos de servicio están orientados a la función clínica y a la función de investigación. No sólo el personal médico, sino también los investigadores deberían acceder, consultar y obtener resultados a partir del cruce y barrido de estas enormes bases de datos, con las debidas

restricciones de acceso y protección de la privacidad de los datos de los pacientes. Para ello se debe contar con una infraestructura que garantice la adecuada y segura administración y conservación de los datos, con especial énfasis en el anonimato de éstos, dependiendo de los privilegios de acceso que tenga el usuario.

Pensar en formas de gestión y tratamiento de datos que requieren grandes cantidades de espacio disponible y un importante número de recursos de procesamiento, implica tomar en cuenta un modelo que proyecte el crecimiento del almacenamiento de datos de salud en el entorno intra e inter-hospitalario. Por tanto este modelo debe considerar la automatización y los procesos propios que deben ocurrir en cada uno de los centros de salud, pero también debe tomar en cuenta la interacción entre estos centros [Her]. Esta interacción debe garantizar un alto nivel de escalamiento para apoyar su capacidad de crecimiento, ofreciendo una estructura segura, estable y heterogénea. Sobre todo, esta interacción debe permitir la posibilidad de resguardar estos datos clínicos cuando tengan otros usos, tales como en la investigación y en la producción de conocimiento.

Para obtener resultados a partir de esta enorme cantidad de datos, dependiendo de la complejidad que pueda resultar la consulta deseada, se hace imperativa su manipulación dentro de un entorno apropiado. Cualquier Nivel de Virtualización de Datos (NVD) y recursos, como el que ofrece

“Cloud Computing” [Bla] o “Grid Computing” [FK99], puede ser utilizado, y en cada caso las soluciones a los problemas planteados deben abordarse con estrategias específicas. En este caso planteamos una arquitectura en un ambiente de “Grid Computing”, el cual no solamente contribuye a cumplir con los requisitos antes mencionados, sino que también permite disminuir los tiempos de cálculo y de procesamiento referidos a las consultas que se derivarán de las necesidades de los investigadores. Por otro lado, para facilitar la interacción de los investigadores con las fuentes de datos, se identifica la necesidad de hacer uso de motores ontológicos que ayuden a estructurar el conocimiento, inferido tras las anotaciones ontológicas de las imágenes médicas -transformadas en datos- que puedan ser gestionados por sistemas manejadores de bases de datos convencionales.

En el artículo se plantea una arquitectura que permitirá extraer información y conocimiento a partir de las anotaciones realizadas sobre las imágenes médicas, considerando para ello, al menos, dos servicios integrados al Grid. De las anotaciones se podrá extraer información que será representada ontológicamente utilizando RDF. Los servicios a considerar están asociados a (1) el desarrollo de metodologías para anotar ontológicamente fuentes de datos contenidas en el Grid de Datos, así mismo como para (2) el desarrollo de servicios de consulta que respondan al criterio de “Perspectivas del Conocimiento” [CH06] desde las interfaces de los usuarios.

Es importante destacar que se observan estudios que han focalizado su atención en la relación entre los componentes de Ontología y Salud, como también con lo que tiene que ver con la relación existente entre el Grid y Salud; ambos casos pudiesen ser considerados muy cercanos al proyecto que aquí se plantea, sin embargo no logran integrar todos los componentes como un todo. Concretamente los casos de proyectos o investigaciones tales como: Mammogrid [Mam] el cual es orientado a facilitar la gestión de imágenes médicas en el Grid, Trencadis [Tre] el cual es una arquitectura de software para compartir imágenes médicas, Health Ontology Generator [YSN09] el que propone el desarrollo de una aplicación para extraer y manipular datos desde una ontología usando formatos convencionales. Todo lo anterior, por citar al menos estos tres ejemplos, corresponde a una revisión que supera al menos 43 proyectos debidamente cotejados y analizados, donde los aportes relacionados con ontología-grid-salud no se han explotados significativamente, y donde los resultados están por obtenerse.

El presente artículo se organiza de la siguiente manera: En la sección 2 se identifican, de manera general, los componentes asociados a la solución que se espera lograr al implementar el diseño de esta arquitectura. En la sección 3 se sugieren los servicios que deben implementarse para sacar provecho de las ventajas del uso del Grid, concretamente cuando se accede a grandes fuentes de datos. En la sección 4 se propone el diseño de una arquitectura del Grid, el cual muestra como elemento novedoso la presencia del “Ontology Element”, así mismo se describe la plataforma donde reposará la implementación del diseño que aquí se propone. En la sección 5 se muestran los resultados

preliminares y el trabajo en curso. Finalmente, la sección 6 concluye con los aportes y lo que podría derivarse del presente trabajo.

2. Componentes identificados

El aprovechamiento de la arquitectura a proponerse adapta el concepto del manejo de supercomputadores dentro de niveles de virtualización de datos, tales como Grid, Cloud o Sky Computing [KTMF09]. Un nivel de virtualización de datos incorpora modelos distintos para el acceso a los datos, que podrían estar definidos como una base de datos centralizada, aunque su ubicación física puede estar distribuida. A los efectos de los resultados que se esperan lograr, se ha identificado al Grid Computing como el nivel de virtualización deseado para implementar la solución.

2.1 Componentes en el GRID

El uso del Grid en esta arquitectura debe ser visto como una alternativa de costo compartido que sustituye la presencia de computadores de gran escala y accede a recursos, geográficamente dispersos, tanto de almacenamiento, como de procesamiento de datos [FK99]. En nuestro caso, se hace hincapié en aprovechar este “middleware” para respaldar y manipular los datos aquí mencionados. El middleware utilizado para tal fin es “gLite”, por ser la Universidad Simón Bolívar signataria del proyecto EELA (*E-Infrastructure shared between Europe and Latin America*) [E-In]. Este software de Grid se utiliza para proveer entornos de cálculo intensivo con enormes cantidades de datos, tanto en áreas aplicadas como de investigación.

gLite es el middleware para “grid computing” utilizado por los experimentos del LHC (*Large Hadron Collider*) del CERN (*Conseil Européen pour la Recherche Nucléaire*) [Cern] y de una gran variedad de uso en los campos científicos. Nace de la colaboración de diversos centros de investigación -académicos e industriales- como parte del proyecto EGEE (*Enabling Grids for E-science*) [EGEE]. Proporciona un conjunto de servicios para la construcción de una infraestructura de red de producción de aplicaciones para el entorno referido a Computación de Alto Rendimiento.

gLite provee un marco para la creación de aplicaciones Grid, aprovechando las ventajas del poder de la computación distribuida y recursos de almacenamiento a través de Internet. Los servicios de gLite son habituales en más de 250 Centros de Informática y utilizada por más de 15.000 investigadores en Europa y en todo el mundo (Taiwán, América Latina, etc.). Auspiciado por el Proyecto EGEE y su derivado EELA, su principal objetivo es permitir a los investigadores hacerse de una infraestructura geográficamente distribuida, disponible las 24 horas/día. EGEE se focaliza en mantener y desarrollar gLite, así como de gestionarlo en el aprovechamiento de la gran infraestructura operacional, para el beneficio de una vasta y diversa comunidad investigativa [Cern].

A los efectos del presente trabajo, gLite es el middleware donde finalmente se colocará en producción el resultado de la presente propuesta.

Básicamente el gLite organiza el uso de recurso con base en Organizaciones Virtuales para acceder a recursos a través de estos elementos:

- Worker Nodes -Nodos de trabajo
- Computing Elements -Elementos de Gestión de Cómputo
- Storages Elements -Elementos de Almacenamiento
- User Interfaces -Interfases de Usuarios

A pesar del esfuerzo de virtualización y transparencia, se recomienda que los datos deben estar los más cercanos posibles entre sí, como también deben estar muy cerca de los recursos de cómputo. Dependiendo de dónde deben quedar alojados estos datos en los “Storage Elements”, el objetivo es que se debe minimizar el retraso producido por efecto de la distancia y de las piezas de tecnologías de comunicación que deben estar entre ellos.

2.2 Ontologías como recursos para estructurar los datos

Las imágenes médicas, conjuntamente con sus respectivas anotaciones, representan el núcleo de los datos a ser utilizados. Sin embargo, estos datos deben ser cruzados con otras fuentes de datos –ontológicos, por ejemplo- que serán de utilidad en los procesos de búsqueda y/o consulta, exigidos por los especialistas o investigadores.

Las representaciones integradas de los datos se encuentran expresadas por intermedio de ontologías, las cuales se entienden como una “especificación de las entidades, sus atributos y las relaciones entre las entidades en un dominio del discurso” [RSN07]. El dominio en cuestión contiene una representación de los conceptos asociados a las imágenes médicas y sus derivaciones, tales como: partes y regiones del cuerpo humano, enfermedades, observaciones clínicas, por citar los más importantes.

A partir de estos conceptos, y como parte de la estructura de la ontología seleccionada, se determinan los elementos que permiten relacionar tales conceptos por intermedio de los predicados. Estos predicados juegan un rol vinculante entre los conceptos, de manera que los datos están fuertemente ensamblados entre sí: sinónimos que enriquecen el concepto del mismo, abreviaciones y acrónimos. A su vez, estos datos, o conceptos, están regidos por las relaciones existentes entre ellos. Lo anterior permite usar la ontología como un controlador de la terminología [RSN07].

Igualmente, se recomienda que las ontologías usadas coexistan con los datos almacenados en el Grid, siendo el caso las imágenes médicas. No necesariamente las ontologías residirían en los mismos nodos donde se encuentran los datos clínicos, sino que físicamente deben implementarse en lugares desde donde el acceso a los datos se realiza de manera segura y rápida.

2.3 Formas convencionales de administración de datos

También se observa que, a pesar de las ventajas que ofrecen las ontologías dentro de los servicios a ser imple-

mentados, con el objeto de sacar provecho de las anotaciones existentes en las imágenes médicas, los procesos de búsqueda y barrido sobre estas anotaciones no suelen realizarse sobre las imágenes almacenadas, si son consideradas como objetos dentro de los registros correspondientes.

Por el contrario, estos procesos se suelen realizar desde manejadores convencionales de bases de datos relacionales, los cuales ofrecen formas de indexar los datos que permiten realizar cruces y búsquedas de datos de manera eficiente. Estas tablas pueden ser gestionadas por manejadores de bases de datos conocidos, preferiblemente en software libre –por ejemplo, PostgreSQL- o en el caso de gLite bajo el uso de ciertos servicios de almacenamiento, tal es el caso de AMGA, el cual se encuentra de manera nativa en este middleware.

Para cualquiera de los casos que se elijan, dentro de esta arquitectura debe implementarse un servicio que, accediendo a los “Storage Elements” del Grid, pueda extraer datos y monitorear el comportamiento de los grandes repositorios de imágenes médicas sin tener que alterar la estructura y el contenido de ésta. De este modo los datos pertinentes pueden ser indexados apropiadamente y en tiempo real.

Particularmente, el componente donde se almacenan los índices resultantes no necesariamente debe estar cercano a los nodos donde se encuentran las imágenes médicas, pero sí muy próximo a las aplicaciones que ejecutan los servicios sugeridos, por las mismas razones antes mencionadas de baja dilación.

Finalmente, se saca provecho del hecho que estas imágenes médicas deben estar asociadas a tablas que están vinculadas a aplicaciones especializadas en la gestión de procesos clínicos. Esta extensión de datos amplía el marco de los datos a ser utilizados por los investigadores, siempre y cuando los mismos estén debidamente protegidos bajo los criterios que garanticen la anonimidad y privacidad de éstos.

2.4 Gestión de aplicaciones

Las aplicaciones, desde la óptica del Grid, pueden provenir y ser ejecutadas remotamente –fuera del Grid-, o disparadas desde un lugar del Grid por intermedio de un “Job”, que pudiese ser provisto remotamente. En el primer caso, el “Job” lleva empotrado el contenido del código que se espera ejecutar, en el segundo la aplicación reside en algún lugar del Grid.

En este trabajo se considerará el segundo caso, para lo cual se observa que la ubicación natural de estas aplicaciones, dentro del Grid, debe ser en algún nodo específico dentro del “Storage Element”. Estos servicios hibernarán en los nodos, antes referidos, hasta que sean invocados por el “Job” correspondiente, o pueden estar ejecutándose indistintamente –dependiendo del “Proxy” asignado-, para cumplir tareas de monitoreo, extracción, transformación y carga de datos externos al interior del Grid, por ejemplo.

3. Diseño de los servicios

Se plantea integrar dos servicios a la arquitectura del middleware de gLite. Ambos servicios requieren de una importante porción de datos que estén presentes en sus correspondientes nodos del “Storage Elements”, tanto a nivel de las imágenes médicas -y sus datos correspondientes-, como también en lo referente a la ontología.

El primero de estos servicios está asociado al desarrollo de Metodologías para anotar ontológicamente fuentes de datos contenidas en el Grid de Datos, a saber, en los elementos de almacenamiento o “Storage Elements”. El segundo tiene que ver con el desarrollo de “Servicios de Consultas” que respondan al criterio de las “Perspectivas de Conocimiento” [CH06] desde las interfaces de los usuarios (“User Interfaces”).

3.1 Metodologías para la anotación ontológica

Este servicio tiene por objeto ofrecer el desarrollo de una metodología que permita extraer información de los datos que se encuentran en las anotaciones realizadas sobre las imágenes médicas. Por otro lado, el aprovechamiento de estas técnicas para extraer información a partir del desarrollo de esta metodología, también será de gran utilidad para estructurar la consulta que realicen los investigadores, a partir de formulaciones redactadas en lenguaje natural, para que puedan ser interpretadas por el servicio definido.

La anotación ontológica se utiliza para generar información útil a partir de los datos, de modo que se puedan realizar manipulaciones al nivel de representación de conocimientos. Se relaciona con el área de solicitud por contenido (“*Content Querying*”), pero se define la solicitud en dos niveles o fases:

- El primer nivel realiza la anotación ontológica de los datos, creando servicios de consulta que no tendrán que procesar los “datos crudos”.
- Posteriormente se pueden realizar las solicitudes a través de los servicios creados.

Esta metodología permite hacer solicitudes en un tiempo muy corto sobre grandes bases de datos, habiendo previamente hecho un proceso de anotación que barre todos los datos y que puede ser relativamente largo.

3.2 Perspectiva del conocimiento

Este servicio tiene por objeto ofrecer “ventanas” para realizar consultas atendiendo las particularidades requeridas por cada uno de los investigadores, los cuales accederán a las mismas fuentes de datos (tanto de las imágenes médicas anotadas y de las ontologías, como hasta de datos clínicos, si así fuese requerido), clasificadas según la necesidad de cada investigador. Por tanto, el perfil del investigador permitirá procesar y filtrar los datos relevantes a sus intereses, a partir del conjunto del universo de los datos para convertirlos en subconjuntos/subclases de éstos y que son relevantes a las exigencias de los investigadores en cuestión.

De lo anterior se puede inferir que cada usuario (investigador) podrá observar una porción de los datos, para obtener, de sus consultas, la respuesta deseada. Este principio que permite visualizar una región de datos específicos que responda a los cuestionamientos del usuario, es lo que se conoce como “Perspectiva” y lo que se extrae del resultado de tales consultas no es más que información valiosa, dato con significado o “Conocimiento” [CH06].

Se utiliza el concepto de “Perspectiva de Conocimiento” para denotar la acción de aplicar conjeturas o teorías particulares a la interpretación de datos e información. En este escenario los NVDs proveen los niveles de seguridad y virtualización necesarios para permitir a los usuarios crear nuevos servicios de datos, en el lado del servidor de datos. Estos servicios nuevos están basados en la perspectiva de conocimiento del usuario.

Un ejemplo de ello puede ser expuesto como sigue: Sea el caso de dos usuarios -investigadores- que se especializan en diferentes áreas. Tomemos como primer caso un investigador que se especializa en cáncer en los huesos y como segundo caso a un investigador que se especializa en fractura de huesos. A los efectos de lo que se espera ofrecer en el servicio que se expone en este trabajo, se permite crear a ambos investigadores diferentes perspectivas de consulta sobre el mismo universo de los datos, posiblemente usando ontologías diferentes o similares, logrando extraer para cada uno de los usuarios una porción o un subconjunto/subclases de datos.

Esto último es equivalente a poder visualizar una parte de los datos comunes, pero con una particular acepción: las conjeturas obtenidas a partir de inferencia que se pueda lograr, de los mismos datos, sólo responderán a los requerimientos que se plantean cada uno de los usuarios. Todo esto se logra con la manera en que se filtraron los datos y dependiendo del uso que hicieron de la ontología correspondiente.

4. Diseño de la arquitectura propuesta

La arquitectura propuesta (Figura1) aprovecha los componentes fundamentales de la arquitectura del Grid, ya que respeta la distribución de los elementos del mismo, pero incorpora dentro de esta plataforma nodos de datos que agilizan las búsquedas y el proceso de anotación ontológica de las imágenes médicas.

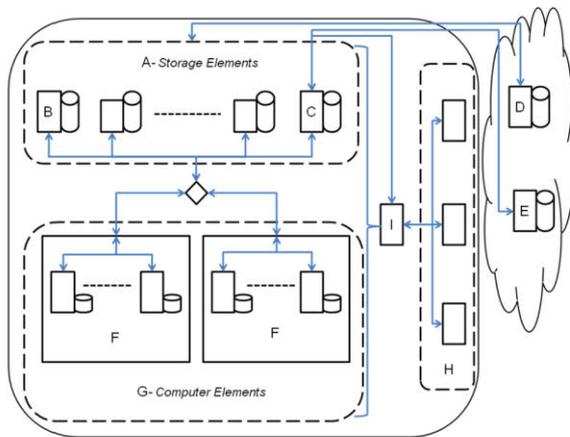


Figura 1. Representación del “Grid Computing” que satisface los requerimientos planteados en los Servicios que se desean implementar, está compuesto por: A) Nodos del Storage Elements, B) Nodo de Base de Datos de Anotaciones de Imágenes Indexadas, C) Nodo Réplica Repositorio de la Ontología, D) Fuentes de Datos Clínicas, E) Fuentes de Datos Ontológicos, F) Nodos de Trabajo –Worked Node-, G) Computer Elements, H) User Interfaces, I) Servicios de Información –Workload Management Systems-.

El diseño contempla incluir un componente de “Base de Datos” o “Data Base Element” (B) dentro del “Storage Element” (A), el cual tiene por objeto indexar las imágenes médicas que estarán también residiendo en el “Storage Element” (A) y relacionarlas con su correspondiente información clínica (D). Estas tablas relacionales deben actualizarse con cierta regularidad, en la medida que se reciban nuevas imágenes médicas en el nodo que le corresponda.

Las garantías de privacidad y anonimidad de los datos clínicos deben implementarse como un paso previo a la indexación de los datos. Esto es: la indexación debe aplicarse sobre “data sets” debidamente anonimizados, previamente extraídos de los datos clínicos originales. Por tanto, la anonimidad debe responder a una política de acceso definida, posiblemente, por las autoridades de salud, de acuerdo a la legislación existente en cada escenario donde se implemente este sistema.

Por otro lado, se incluye un nodo que contendrá la ontología relacionada con el área de salud, que hemos denominado “Ontology Element” (C). Este nodo estará actualizando su dominio del conocimiento a partir de sincronizaciones con la ontología seleccionada (E) que son administradas por comunidades en el entorno de la Web. El motivo de tener dentro de Grid la ontología en el “Ontology Element” se debe básicamente a dos razones.

La primera de ellas se debe a la proximidad que debe tener la ontología con las anotaciones, a los efectos de minimizar los tiempos de intercambio para el análisis. Y el otro motivo se refiere a que las ontologías están publicadas en ambientes autónomos, los cuales pueden estar eventualmente desconectados. Si bien es cierto que se pudiese estar utilizando una ontología, que en algún instante no esté lo suficientemente actualizada, siempre será una versión más cercana de la realidad, hasta que se logre sincronizar nue-

vamente el “Ontology Element” [CH07] con la ontología referente.

A los efectos de la propuesta que se realiza en este diseño, aparecen dos importantes contribuciones dentro del “Storage Element” (Figura 2).

- Una de las contribuciones tiene que ver con la posibilidad de definir uno o varios nodos que se especialicen en hacer gestión de datos a partir de la indexación de datos –haciendo uso de manejadores de bases de datos relacionales-, con la idea de hacer más eficiente las búsquedas de los datos anotados que están vinculados a sus respectivas imágenes médicas, lo que hemos denominado “Data Base Elements”.
- La otra contribución tiene que ver con el diseño e implementación de un nodo relacionado con la gestión de las ontologías relacionadas con el área de salud, permitiendo la actualización de estas ontologías por intermedio de un agente y garantizando la seguridad en la transacción desde la nube de la Web, de manera de tener acceso a las ontologías directamente dentro de la estructura del Grid. A estos nodos los denominamos “Ontology Elements – O.E.”.

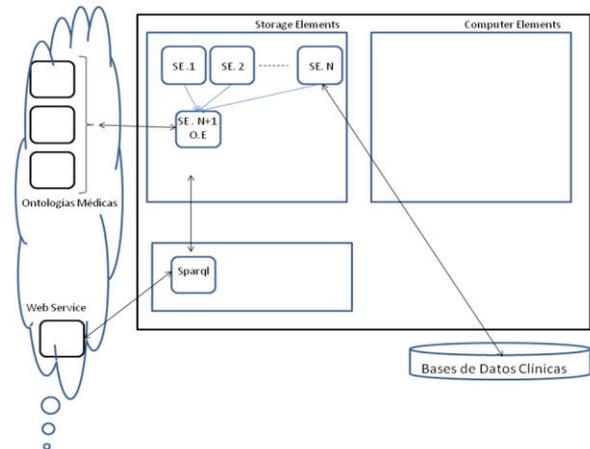


Figura 2. Representación del “Grid Computing” muestra los vínculos entre cada uno de los participantes del Storage Elements y las fuentes de datos que los alimentan.

5. Resultados preliminares y trabajo en curso

A pesar de que lo expuesto en el presente artículo es una propuesta del diseño de una arquitectura a ser desarrollada, se han implementado servicios de prueba como los descritos anteriormente (equivalentes a “Database Elements” y “Ontology Elements”) en servidores aislados y en clusters de computadoras, con resultados satisfactorios desde el punto de vista funcional. Requiere de mejoras desde el punto de vista de rendimiento, lo cual se espera mejorar en ambientes de cómputo de alto rendimiento. La implementación de estos mismos servicios a lo interno de un GRID de tipo gLite es un trabajo en desarrollo.

Para sacar provecho de las ventajas del Grid, se considera la utilización de recursos, inicialmente, ubicados en la Universidad Simón Bolívar, concretamente en el Laboratorio de Computación de Alto Rendimiento –CAR-, localiza-

dos en Edificio Física I. El mismo está compuesto por dos “Cluster”, todos con salida vía “Internet 2”, los cuales se describen en la Tabla 1 y en la Tabla 2:

Tabla 1. Plataforma de equipos -Universidad Simón Bolívar.-

Hardware	Servers/Nodos	Storage
MIRANDA	64/128	100 TB
CHUAO	6/24	0 TB

Tabla 2. Software habilitado –Universidad Simón Bolívar-

Tipo	Nombre
O.S.Title	Scientific Linux 5
Middleware	gLite 3.1
Tools	Java, C, C++, Python
DB	MySQL, Postgres, AMGA

Paralelamente se cuenta con la debida autenticación para acceder a la infraestructura del EGEE.

En la fase posterior de pruebas, se ejecutarán “Jobs” en una Organización Virtual dentro de EGEE, debidamente habilitado para tal fin, con el objeto de operar una producción de esta solución en dicho Grid Distribuido, el cual cuenta con más de 200 “sites” alrededor del mundo, más de 30.000 CPU’s y superior a los 20 Petabytes de almacenamiento, en híbridas configuraciones. Esta plataforma ya está previamente configurada y respeta mayormente la configuración de la Tabla 2.

En cuanto a la ontología, se extraerán datos de MESH [Bio] a ser publicados en el nodo de ontología dentro de la arquitectura propuesta, tomando en cuenta a GALEN [RRP96]. Finalmente, los datos relacionados a las imágenes médicas deben ser extraídos de centros de salud, bien sea nacionales o internacionales.

6. Conclusiones

En el presente trabajo se describe una arquitectura dentro de un Grid de Datos que soporta servicios asociados a consulta de datos de salud, no sólo con el fin clínico sino con fines de hacer investigación con esos datos. Se plantea una arquitectura que permite extraer información y conocimiento a partir de las anotaciones realizadas sobre las imágenes médicas, considerando para ello, al menos, dos servicios integrados al Grid.

Los servicios están asociados al desarrollo de metodologías para anotar ontológicamente fuentes de datos contenidas en el Grid de Datos y el desarrollo de servicios de consulta que respondan a un criterio conocido como “Perspectivas de Conocimiento”, desde las interfaces de los usuarios. Todo bajo restricciones de acceso que garanticen la privacidad de los datos de los pacientes.

Se han realizado pruebas preliminares de estos servicios fuera del ambiente Grid. El trabajo en curso apunta a realizar la implementación real en un ambiente Grid en fase de producción de estos servicios.

Referencias

- [MGVG] Müller H., García A., Vallée J.P., Geissbuhler A., Grid Computing at the University Hospitals of Geneva, University Hospitals of Geneva, Division of Medical Informatics, Rue Micheli-du-Crest 24, 1211 Geneva 14, Switzerland.
- [Her] Hernández E., Grid Access Systems, Universidad Simón Bolívar, Departamento de Computación, Caracas, Venezuela., <http://www.saber.ula.ve/bitstream/123456789/16014/1/ehernandez.pdf>.
- [RSN07] Rubin D.L., Shah N.S., Noy N.F., Biomedical ontologies: a functional perspective, BRIEFINGS IN BIOINFORMATICS. VOL 9. NO 1. 75-90, Advance Access publication December 12, 2007.
- [Bio] Biomedical Knowledge Engineering Lab, <http://bike.snu.ac.kr/sites/default/files/meshont.o.owl>.
- [RRP96] Rector A., Rogers J.E., Pole P., The GALEN High Level Ontology, Medical Informatics Group, Department of Computer Science, University of Manchester, Manchester, UK, M13 9PL, 1996.
- [CH07] Campos J., Hernández E., “Gestión de Imágenes Médicas Usando Grid Computing Modelo para el Apoyo a la Investigación en Venezuela”, IV Latin American Congress on Biomedical Engineering 2007, Bioengineering Solutions for Latin America Health, IFMBE Proceedings, Volume 18, Pág. 928-931, 2007.
- [Cer] CERN, <https://edms.cern.ch/file/722398/gLite-3-UserGuide.html>.
- [FK99] Foster I., Kesselman C., The Grid Blueprint for a New Computing Infrastructure, Morgan & Kaufmann, 1999.
- [CH06] Cadenas L.E., Hernandez E., Artificial Intelligence in Theory and Practice, IFIP International Federation for Information Processing, 2006, Volume 217/2006, 189-198.
- [E-In] E-Infrastructure shared between Europe and Latin America, <http://www.eu-eela.org>.
- [EGEE] EGEE, <http://www.eu-egee.org/>.
- [Bla] Blake Y.W., Service-Oriented Computing and Cloud Computing: Challenges and Opportunities, Univ. of Notre Dame, Notre Dame, IN, USA, IEEE, ISSN: 1089-7801, On page(s): 72-75.
- [KTMF09] Keahey K., Tsugawa M., Matsunaga A., Fortes J., Sky Computing, IEEE Internet Computing archive, Volume 13, Issue 5, September 2009, IEEE Educational Activities Department Piscataway, NJ, USA. <http://www.cems.uwe.ac.uk/cccs/project.php?name=mammogrid>.
- [Mam] <http://www.cems.uwe.ac.uk/cccs/project.php?name=mammogrid>.
- [Tre] <http://www.grycap.upv.es/cvimo/trencadis.html>.

- [YSN09] Yip Chi Kiong, Sellappan Palaniappan, Nor Adnan Yahaya, "Health Ontology Generator: Design And Implementation", IJCSNS International Journal of Computer Science and 104 Network Security, VOL.9 No.2, February 2009.