

producción *grid*

Un caso de adaptación del *GRID* a las necesidades de la empresa



Por Andrés Gómez Tato (CESGA), Juan Nouche (DYGRA Films)
Pablo Rey Mayo (CESGA) y Lucía Vallés (DYGRA Films)

La producción audiovisual en formato digital requiere ingentes cantidades de tiempo de procesado y de almacenamiento de datos. Además, presenta la característica de la independencia del procesado de cada fotograma o escena, que permite el reparto del trabajo en entornos distribuidos. Asimismo, la realización de cada proyecto (largometraje, anuncio, etc.) es frecuentemente una colaboración entre diferentes empresas (se forma el grupo liderado por una empresa o productora para la realización de un único proyecto). Por todo ello, es un caso que encaja perfectamente en los escenarios para los cuales se han desarrollado las tecnologías *GRID* (1).

Sin embargo, también presenta características propias nuevas en esos escenarios: orientación comercial; protección del resultado; organización interna jerárquica; la forma de producción a través de trabajo colaborativo multiempresa, incluso con "freelan-

ce"; la utilización mayoritaria de aplicaciones basadas en Microsoft Windows para los puestos de trabajo y servidores de gestión; la escasez de ancho de banda frente a las disponibles en entornos de investigación y el tamaño de los ficheros utilizados. Todo ello hace que la aplicación de tecnologías *GRID* sea más compleja que en el caso de instituciones académicas.

En este artículo se hace una descripción de la experiencia realizada por DYGRA y el Centro de Supercomputación de Galicia sobre la utilización de la tecnología *GRID* al sector de producción de películas animadas. En un primer lugar se introduce brevemente el concepto de *GRID*. Se describe a continuación el proyecto realizado, con las necesidades y soluciones planteadas. Para finalizar, se enumerarán las conclusiones obtenidas durante el proyecto sobre la utilización de esta tecnología en el sector.

El *GRID*

Hace unos años, cuando se desarrollaba un proyecto de software, en la etapa del diseño, se pensaba en un entorno de ejecución centralizado dentro de una misma organización, excepto en casos muy particulares (por ejemplo, EDI). Con la irrupción de Internet y el *World Wide Web*, esta forma de ejecución empezó a cambiar, ofreciendo las empresas servicios a sus clientes o proveedores a través de páginas *WEB*. Actualmente, la tendencia es desarrollar o prestar servicios de forma colaborativa, por ejemplo, utilizando servicios *WEB*. La tecnología *GRID* es un paso más en ese camino, facilitando la prestación de servicios de forma distribuida, utilizando los recursos de varias organizaciones. En este nuevo concepto se pueden definir las organizaciones virtuales como aquel conjunto de recursos y servicios pertenecientes a varias organizaciones físicas que se ponen en común para la





Nombre	Tamaño (MB)	Tiempo Transmisión (s)
pele_sc005_pn01_anim_FX_luz_RM_RM.mb	131	2096
pele_sc013_pn08_iluminacion_personajes_fx_RM.mb	128	2048
pele_sc013_pn12_iluminacion_personajes_RM.mb	94	1504
pele_sc018_pn06_iluminacion_personajes_RM_mostaza.mb	88	1408
pele_sc004_pn10_iluminacion_fondo_RM.mb	81	1296
pele_sc004_pn10_iluminacion_personajes_fx_RM.mb	69	1104
pele_sc004_pn12_iluminacion_personajes_RM.mb	68	1088
pele_sc013_pn10_iluminacion_personajes_fx_RM.mb	65	1040
pele_sc013_pn09_iluminacion_personajes_RM.mb	58	928
pele_sc001_pn33_iluminacion_personajes_FX_RM.mb	47	752
pele_sc013_pn12_iluminacion_fondo_RM.mb	44	704
pele_sc013_pn09_iluminacion_fondo_RM.mb	42	672

TABLA 1. Tamaños típicos de algunos ficheros de entrada junto con el tiempo de transmisión

realización de un objetivo.

Esta nueva forma de trabajo aporta grandes ventajas: la gran cantidad de recursos disponibles que anteriormente no se podían utilizar al estar dispersos entre diferentes organizaciones; la generación de nuevos servicios de valor añadido a través de la conjunción de la experiencia y conocimiento de varias organizaciones; la puesta en valor de recursos infrautilizados; etc. Ahora con la tecnología GRID se puede hacer uso de ellos, pero teniendo un coste a cambio: la pérdida del control, total o parcial, de los recursos utilizados. Así, por ejemplo, la pérdida de control implica que en cualquier momento estos trabajos pueden ser interrumpidos por causas ajenas al usuario, por lo que el programador debe tener en cuenta estas posibles interrupciones para en el caso de que se produzcan se pueda reaccionar en consecuencia.

El proyecto ProducciónGrid

ProducciónGrid es un caso práctico de aplicación de las tecnologías GRID - en

este caso Globus Toolkit (GT3.22) (2), basada en la especificación Open Grid Service Infrastructure (OGSI3) (3) - a empresas de producción audiovisual que buscaba valorar la realización de una intranet/extranet empresarial que permita utilizar los recursos distribuidos de diferentes organizaciones para la producción de películas de animación, formando lo que se denomina una "organización virtual". Este proyecto fue realizado en colaboración entre DYGRA y el Centro de Supercomputación de Galicia dentro del Plan Gallego de I+D+i.

El diseño de la solución se enmarcó dentro de las siguientes premisas:

1. El aspecto artístico del producto final (la película, el anuncio, etc.) tiene prioridad sobre la calidad técnica o científica. Así, la solución propuesta ha de ser una herramienta para mejorar el trabajo de los animadores.
2. La empresa matriz de animación es la encargada de la gestión del proyecto audiovisual. Un proyecto tiene varios actores: la propia empresa de animación y sus empleados (denominados usuarios internos) y las empre-

sas o contratados externos para la realización de tareas específicas del proyecto (denominados usuarios externos).

3. La seguridad del proyecto es fundamental. Nadie externo a la empresa debe tener acceso al resultado global del proyecto, dado el tremendo impacto económico que puede suponer una copia no autorizada al mismo.

4. Pueden existir varios proyectos simultáneos que pueden tener su propia organización virtual. Sin embargo, los diferentes proyectos tienen una gestión común de prioridades en el tiempo.

5. Cada empresa participante en un proyecto ha de mantener la independencia (dentro de los límites del contrato que se haya establecido) de los servicios que presta. Así, el CESGA, como proveedor de horas de cálculo, no debería instalar servicios adicionales no requeridos para la prestación del servicio.

6. La complicación de utilización del GRID ha de estar oculta al usuario final. Se buscará únicamente suminis-



trar información y servicios que simplifiquen el trabajo y la utilización del sistema.

Dentro de estas limitaciones autoimpuestas, las necesidades de la empresa que se han detectado que pueden ser solucionadas utilizando tecnología GRID son las siguientes:

1. Control de acceso a los ficheros del proyecto. Se pueden incluir sistemas de control de acceso avanzados utilizando tecnologías GRID para, sobre todo, dar acceso a trabajadores externos a las partes del producto que tienen encomendadas.

2. Visualización remota. Durante el proceso de producción, es necesario realizar previsualizaciones del resultado para la toma de decisiones. Estas visualizaciones han de poder realizarse desde cualquier parte del mundo, por lo que es necesario incluir la seguridad necesaria. Esta necesidad también se puede abordar utilizando tecnologías actuales (como VPNs) existentes en el mercado.

3. Ejecución paramétrica. Existen determinados productos comerciales que se utilizan para la realización de simulaciones físicas que se incluyen en una película (como el agua de un arroyo o la caída de una cuerda). El resultado final obtenido para esas simulaciones depende de los parámetros de entrada, aunque el resultado final seleccionado vendrá dado por cuestiones artísticas. De este modo, es posible utilizar el GRID para realizar diferentes simulaciones dentro de un rango de parámetros para que el animador elija entre las salidas, aquella que le convence más desde el punto de vista artístico.

4. Renderización. El *render* es un proceso complejo de elevado cálculo imprescindible en una producción audiovisual digital que consiste en la aplicación de texturas, luces, etc. El *render* es una parte fundamental de la producción y, habitualmente, ya se utilizan servicios de terceros.

Durante el proceso producción existen diversos tipos de renderizado:

* **Render de iluminación:** Necesario

para una parte del proceso de trabajo en que los usuarios necesitan ver el resultado que les permita tomar decisiones y así poder seguir avanzando. Se ha de ejecutar en el menor tiempo posible, ya que el usuario puede estar esperando por el resultado. Es, por tanto, un proceso interactivo o casi interactivo. En este caso, se hace necesario dividir verticalmente el fotograma en tantas subpartes como nodos se dediquen, de tal forma que el tiempo de renderizado de un plano sea el menor posible. Además, cada parte se ha de ejecutar de forma inmediata, no pudiendo esperar en cola para su ejecución.

* **Render Final:** Partiendo de los planos iluminados, se obtienen los fotogramas finales de las capas en las que se divide cada plano. Es un proceso que no necesita interactividad, solo control del estado de su ejecución.

* **Render Postproducción:** Se usa el fichero de composición para sustituir los elementos provisionales por sus fotogramas finales por capas y obtener de esta manera los fotogramas y videos finales que en su momento se llevarán al proceso último de filmado y sonorizado para obtener el producto final.

Dadas las características de cada uno de estos tipos de render, las conexiones de red existentes abordables por una empresa mediana y sus posibilidades de adaptación a un sistema basado en tecnologías GRID, se llegó a la conclusión de que solamente el render final y el de iluminación son apropiados para ser integrados en un sistema GRID.

El *render* de Postproducción no es adecuado para su utilización en GRID, debido a la necesidad de mover una gran cantidad de información para un proceso relativamente corto en tiempo de computación ya que, salvo que se aplique algún tipo de efecto que realice un gran cálculo (por ejemplo, un efecto de ondas), *renderizar* una secuencia de postproducción puede llevar solo unos pocos minutos. Con lo cual el tiempo necesario para mover

todos los ficheros necesarios es muchísimo mayor que el tiempo de *renderizado*.

El *render* de iluminación, aunque en una escala menor, presenta problemas similares al *render* de postproducción: es necesario mover gran cantidad de datos para la realización de un proceso de computación relativamente corto (si es que se quiere cumplir con las necesidades del usuario final). En la **TABLA 1** se observan los tamaños típicos de algunos ficheros de entrada así como el tiempo mínimo necesario para la transmisión de los mismos suponiendo una línea de 512Kbit/s de subida. Se puede ver que los tiempos necesarios para el envío de la información son, actualmente, muy superiores a los tiempos de renderizado esperados una vez divididos los fotogramas (no superior a los 600s).

Implementación

La gestión del proyecto audiovisual (unidad de trabajo de la Organización Virtual) está basado en un portal Web, utilizado como gestor de trabajos, realizada sobre Microsoft .NET. Este portal tiene acceso directo a los recursos locales de computación (denominados *Renderfarm*). Además, existen servidores de almacenamiento en donde se guarda la información necesaria (escenas, texturas, etc.) y los fotogramas resultantes. Todos los usuarios, tanto internos como externos, accederán a la información del proyecto a través del WEB, si bien los usuarios externos podrán contar en el futuro con sistemas de autenticación fuertes.

El acceso al GRID se realizará a través de este portal de forma transparente para el usuario. DYGRA, organización productora que presta el **servicio de producción**, realiza y gestiona el proyecto, distribuyendo en los diferentes GRIDs de computación contratados la renderización (en este caso de la prueba, el Centro de Supercomputación de Galicia, CESGA, que simplemente presta un servicio de ejecución remota) y asegurando su ejecución (es decir, vigila la correcta



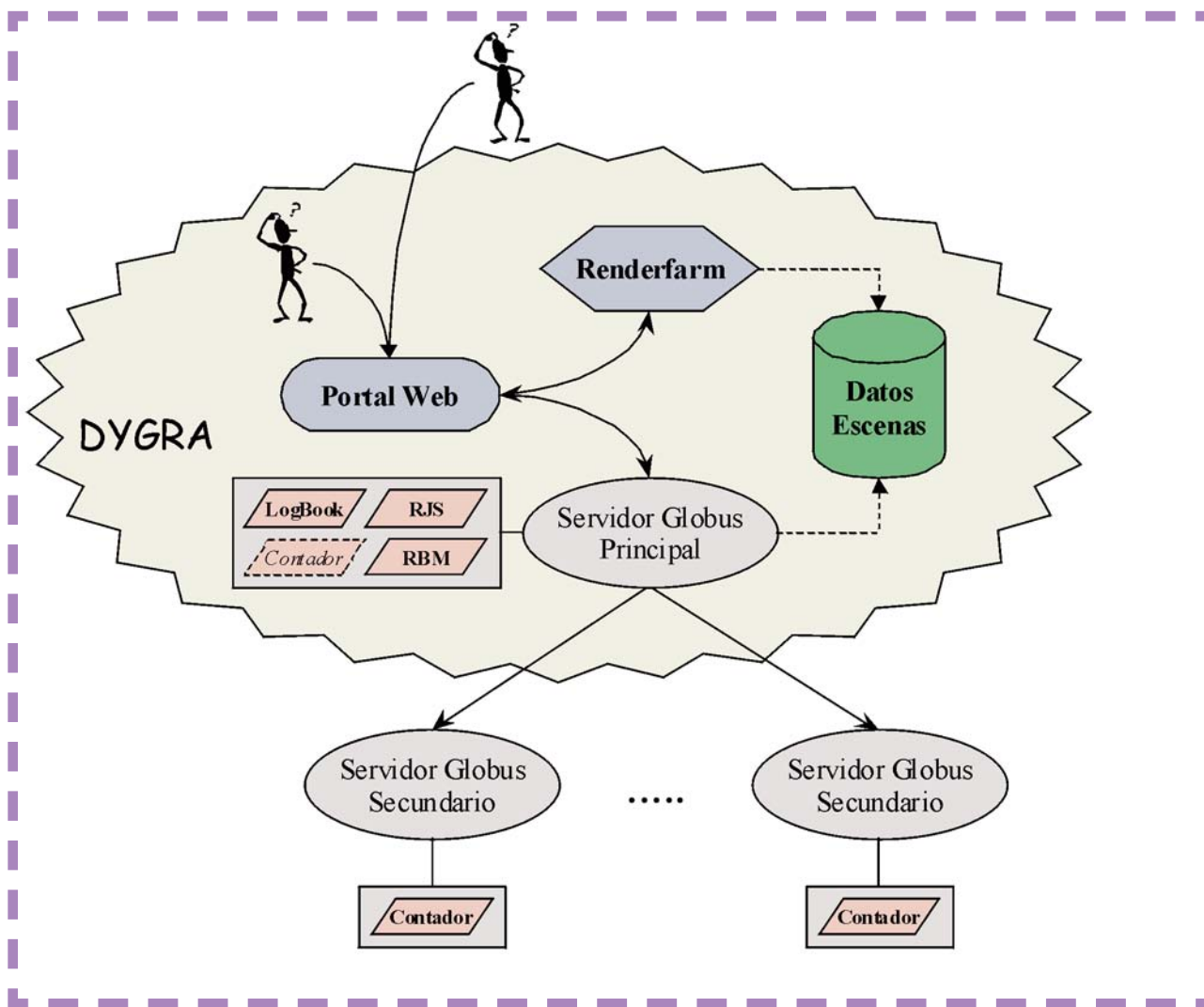


FIGURA 1. Estructura del sistema

ejecución de la petición). La selección de recursos a utilizar en cada caso, vendrá dada fundamentalmente por reglas de negocio (coste, prioridad) orquestadas a través de un Gestor de Recursos (denominado *ResourceBroker*). El sistema GRID creado consta de un **servidor principal**, que es el encargado de interactuar con el portal Web y controlar todo el proceso de *renderización* dentro del GRID, y de uno o varios **servidores secundarios**, los cuales llevan a cabo la *renderización* de los trabajos que le sean asignados. Estos ser-

vidores secundarios pueden ser a su vez infraestructuras GRID de computación que reconozcan las credenciales (identificaciones) asignadas por el servidor principal.

El servidor principal consta de los siguientes servicios:

1. Render Job Submittion (RJS). Se encarga de todo lo relacionado con la gestión de los trabajos (envío, cancelación, estado, ...). Así, para cada solicitud se crea una instancia (hilo). Esta instancia se encarga de buscar los ficheros necesarios para la solicitud y

empaquetarlos, obtener del RBM la selección de recursos para ejecutar el trabajo, contactar con éstos para arrancar la petición remota, enviar los ficheros necesarios, controlar el estado del trabajo (y rearmarlo si es preciso) y descargar el resultado a la finalización de la *renderización*. Todas estas operaciones se han de realizar en nombre del usuario solicitante, para la cual se utiliza el concepto de delegación de identidad: el servicio asume durante su ejecución la identidad del solicitante. Este servicio es indepen-



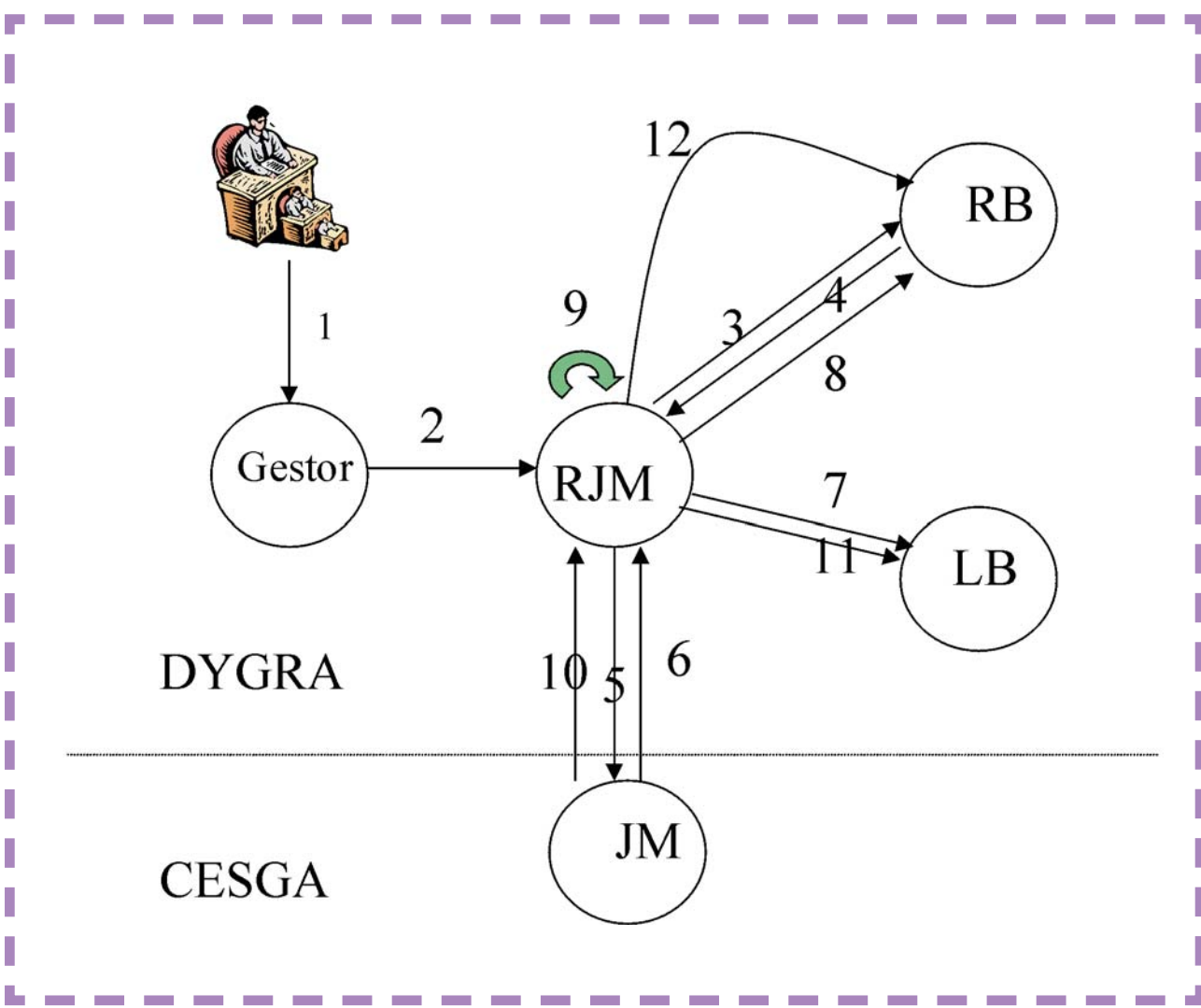


FIGURA 2. Flujo de llamadas entre los servicios.

diente del software de render utilizado, soportando en la actualidad tanto Maya como RenderMan, aunque está preparado para la adopción de otros software similares.

2. Resource Broker Manager (RBM). Distribuye la realización de trabajos en función de las necesidades y las reglas de negocio de la empresa. Su función es localizar los servicios disponibles y seleccionar aquellos que se ajusten en ese momento a las reglas de la empresa (coste, ocupación, etc.).

3. LogBook. Mantiene la información

de estado de cada una de las peticiones realizadas, para recuperarla en caso de caída del sistema y para la información del gestor.

Todo el sistema fue probado con éxito en el mes de octubre de 2004, utilizando los servicios de computación del CESGA y de DYGRA como recursos de computación.

Conclusiones

Durante la ejecución del proyecto se han detectado varias situaciones no resueltas satisfactoriamente en GRID

cuando se quiere utilizar en Pequeñas y Medianas Empresas y en sistemas reales de telecomunicaciones (que tengan un coste asumible por éstas, claro). En concreto, es necesario abordar más intensamente los siguientes puntos:

1. Jerarquía: Los sistemas Grid están fundamentalmente diseñados para la utilización individual final, es decir, cada usuario reconocido puede hacer uso de los recursos obtenidos libremente, estando simplemente limitado por la disponibilidad de los mismos.





Sin embargo, en un sistema empresarial las prioridades pueden cambiar con el tiempo, por lo que es posible que se plantee la necesidad de ajustar el consumo de recursos momentáneamente. Esto implica que, por ejemplo, un administrador puede parar o suspender trabajos ya iniciados. Dentro de la organización eso es posible hacerlo, pero los sistemas actuales de GRID no siempre contemplan la posibilidad de administración jerárquica de peticiones entre organizaciones.

2. Herramientas de desarrollo. Las herramientas de desarrollo para servicios Grid todavía son muy inmaduras, inadecuadas para implementar soluciones comerciales adecuadas. Es necesario contar con mejores herramientas de desarrollo multiplataforma.

3. Comunicación entre Windows y GRID. Aunque se está en fase de estandarización y utilizan el mismo concepto de servicios WEB, todavía la comunicación entre .NET y Globus no está completamente resuelta a nivel de seguridad. Es posible que este hándicap se resuelva en un espacio corto de tiempo con el estándar WS-Security.

4. Disponibilidad barata de ancho de banda. La utilización del GRID de la forma propuesta requiere un considerable ancho de banda entre el servidor principal y los secundarios para transmitir los grandes ficheros necesarios para cada ejecución (de más de 100MB típicamente), en ambos sentidos. Desde luego, ADSL no es la solución para estas PYMEs, debido sobre todo a su bajo ancho de banda de subida.

5. Licencias de software. La estructura actual de licencias está ampliamente identificada como un problema para el desarrollo del GRID, ya que no son suficientemente flexibles desde el punto de vista técnico y legal (en muchos casos impide la ejecución fuera de la empresa o lugar de instalación). Para que un esquema GRID funcione adecuadamente, es necesario desarrollar esquemas de licencias adaptados a la tecnología tanto desde el punto de vista legal como técnico.



FIGURA 3. El sistema puede que se utilice para utilizarlo en parte del renderizado de la película *El sueño de una noche de san Juan*

Lo deseable sería que ante una petición de trabajo se solicitara a un servicio GRID licencia para la ejecución del software y que la licencia no tuviera que estar previamente instalada en el proveedor de CPU.

6. Instalación automática del software. Dado que el usuario final puede necesitar varios tipos de software, aunque el proveedor actualmente lo tenga instalado, sería preferible que para cada ejecución se realizara una instalación automática del software demandado, de tal forma que el nodo de cálculo pudiera quedar limpio a la finalización del trabajo.

7. Cifrado de los ficheros adaptado a GRID. Dadas las características de la información a tratar (extremadamente confidencial) sería deseable que los ficheros de datos transmitidos entre las partes fueran cifrados de tal forma que solo los programas autorizados a leerlos pudieran tener acceso.

Aunque todavía quedan problemas

por solucionar, la solución desarrollada está preparada para su utilización en entornos de producción. De hecho, se está evaluando la posibilidad de utilizarla en parte del renderizado final de la última película de DYGRA Films, *El sueño de una noche de San Juan*. ❌

Notas

- 1 I. Foster, C. Kesselman. *The Grid: Blueprint for a New Computing Infrastructure*, Morgan-Kaufman, 1999 y *The Anatomy of the Grid: Enabling Scalable Virtual Organizations*. I. Foster, C. Kesselman, S. Tuecke. International J. Supercomputer Applications, 15(3), 2001.
- 2 <http://www-unix.globus.org/toolkit/>
- 3 *Open Grid Services Infrastructure (OGSI) Version 1.0*. S. Tuecke, K. Czajkowski, I. Foster, J. Frey, S. Graham, C. Kesselman, T. Maguire, T. Sandholm, P. Vanderbilt, D. Snelling; Global Grid Forum Draft Recommendation, 6/27/2003.

