

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 386 134**

21 Número de solicitud: 201001371

51 Int. Cl.:

**G06F 9/44** (2006.01)

**G06F 21/00** (2013.01)

12

PATENTE DE INVENCION

B1

22 Fecha de presentación:

**22.10.2010**

43 Fecha de publicación de la solicitud:

**09.08.2012**

Fecha de la concesión:

**16.04.2013**

45 Fecha de publicación de la concesión:

**26.04.2013**

73 Titular/es:

**UNIVERSIDAD DE SEVILLA (100.0%)  
OTRI-PABELLON DE BRASIL, PO. DE LAS  
DELICIAS S/N  
41012 SEVILLA (Sevilla) ES**

72 Inventor/es:

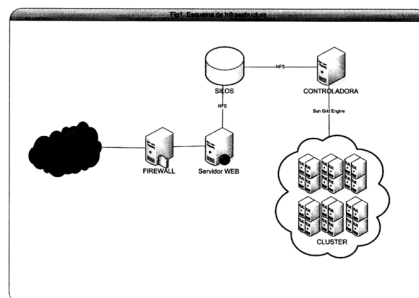
**ORTEGA RAMIREZ , Juan A. ;  
SILVA GALLEGO , Ana ;  
GONZALEZ ABRIL , Luis ;  
VELASCO MORENTE , Francisco ;  
TORRES VALDERRAMA , Jesus ;  
ESCALONA CUARESMA , M<sup>a</sup> Jose ;  
ALVAREZ GARCIA , Juan Antonio ;  
ALVAREZ DE LA CONCEPCION , Miguel Angel ;  
FERNANDEZ MONTES , Alejandro ;  
FUENTES BRENES , Daniel y  
ANGULO BAHON, Cecilio**

54 Título: **SCI (SIMPLE CLUSTER INTERFACE): ARQUITECTURA PARA LA GESTION DE TAREAS DE USUARIO EN UN CLUSTER A TRAVES DE LA WEB**

57 Resumen:

La arquitectura presentada en esta patente, llamada Simple Cluster Interface (SCI), permite al usuario hacer uso de los recursos de computación a través de la web, eliminando la necesidad de introducir comandos por consola. La facilidad que da esta arquitectura, extiende el uso de recursos de computación, hasta ahora elitistas, a un mayor número de investigadores sin la condición de tener un perfil técnico avanzado, contando con un protocolo de seguridad en el intercambio de ficheros basado en la generación de un mapa XML.

Figura 1:



ES 2 386 134 B1

## DESCRIPCIÓN

SCI (Simple Cluster Interface) : Arquitectura para la gestión de tareas de usuario en un clúster a través de la web

### 5 Objeto de la invención

Simple Cluster Interface (SCI) es un sistema permite al usuario hacer uso de los recursos a través de la web eliminando así la necesidad de introducir comandos por consola, acercando de esta manera los servicios de Supercomputación a un mayor número de usuarios sin la necesidad de tener un perfil técnico. A su vez la  
10 gestión de los datos de usuario se realizan mediante un protocolo propio que garantiza la seguridad de los datos almacenados.

### Estado de la técnica

Simple Cluster Interface (CLG), permite al usuario hacer uso de los recursos a  
15 través de la web eliminando así la necesidad de introducir comandos por consola, acercando de esta manera estos servicios a un mayor numero de usuarios ya que se suprime la necesidad de que este tenga un perfil técnico.

Sin embargo, con el acceso al sistema a través de la web los problemas de seguridad se agravan. Por tanto, la seguridad es un punto importante a tratar en  
20 sistemas de estas características ya que su gran potencia de computo lo hace una herramienta muy poderosa para usos ilícitos, como puede ser, por ejemplo, la descriptación de claves. Para evitar este problema, SCI engloba una arquitectura lógica y física que permite al usuario gestionar sus tareas, el envío y recepción de trabajos de computo, desacoplando la arquitectura física del clúster  
25 de la interfaz web de usuario, acabando así con los problemas de seguridad.

SCI pertenece a una categoría de programas que permiten el acceso mediante una interface Web a sistemas de cálculo de alto rendimiento. En la actualidad existen otros productos con una funcionalidad similar. De los productos  
30 existentes, unos tienen como finalidad permitir a los usuarios interactuar con sistemas de Grid Computing para el envío de trabajos sencillos, autenticación y selección de recursos. Un ejemplo de esto es DIRAC\*. Otros productos similares están contruidos para proporcionar al usuario un entorno de trabajo específico para una aplicación. En los productos existentes ocurre que, o bien pretenden ser de uso general, con lo cual pierden funcionalidades o bien dan al usuario mucha  
35 funcionalidad a costa de perder generalidad. SCI, por el contrario, es un producto

que ofrece un buen equilibrio entre posibilidad de uso general para muchas tareas diferentes y funciones ofrecidas al usuario.

Tiene, además, otras características en su diseño que lo diferencian de los sistemas existentes:

- 5       ● Está pensado para su uso en sistemas de clusters para cálculo de alto rendimiento.
- Su diseño interno se ha mantenido con un mínimo de complejidad y requisitos. En particular, SCI no necesita que el usuario instale ningún tipo de componente extra en su ordenador (Java, Flash, etc.).
- 10     ● El entorno de trabajo es de tipo escritorio con el que el usuario está familiarizado, por lo que SCI tiene una curva de aprendizaje reducida.

\*DIRAC: Es una herramienta gráfica para la gestión de una infraestructura distribuida bajo middleware GRID, usado para el sistema de producción de LHCb Monte Carlo. Su arquitectura cliente/servidor está basada en elementos de computación distribuidos entre los distintos centros que colaboran con el proyecto, mediante una base de datos realiza el control de los elementos disponibles, tiene también un catalogo de datos (sistema de información) , y un repositorio software. Para el envío de trabajos utiliza el protocolo XML-RPC, el agente comprueba las necesidades software del trabajo, si es necesario y lo asigna a una cola. Una vez que ha sido procesado el agente transfiere los resultados y actualiza el catálogo de datos del sistema. DIRAC contiene una serie de requisitos para su instalación y manejo que lo diferencian de SCI. Dejando atrás la particularidad que hacen novedoso a esta invención.

25 Destacamos además, las siguientes patentes relacionadas:

- W020971DE, Procedimiento, ordenador servidor y sistema para el control de acceso a los datos.
- WO2009134819 (A1) , System and method for programmactic management of distributed computing.
- 30 – WO2007036602 (A1), Security of virtual computing platforms.
- US2008271020 (A1), System and method for working in a virtualized computing enviroment through secure access.
- US2008178179 (A1), System and method for automating and scheduling remote data transfer and computation for high performance computing.

- WO2008038277 (A2), A system and method for secure web brouing using server-based computing configuration.
- US2006167674 (A1), Provisioning computing services via an on-line networked computing enviroment.
- 5 - TW224899 (B) Dynamic binding and fail-over of comparable web service instances in a services grid.
- EP1475938 (A2), Web access to secure data.
- US2004103339 (A1), Policy enabled grid architecture.

10

### **Descripción de las figuras**

En la **figura 1** se presenta un esquema básico de conexión física entre los distintos elementos de la arquitectura, donde se observan la máquina controladora  
15 donde se encuentra el sistema de colas, y es la que intercambia información con los elementos formados por los nodos (que ejecutarán las tareas) y el servidor web que ofrece al usuario la interfaz web.

En la **figura 2** se presenta arquitectura lógica del sistema de espacio compartido ,  
20 la zona de intercambio entre los dos elementos, servidor web y controladora del sistema de colas. Se trata de un espacio exportado mediante NFS de la controladora al servidor web, mediante el cuál se intercambia la información.

### **Descripción de la invención**

Simple Cluster Interface (SCI) es un sistema permite al usuario hacer uso de los recursos a través de la web eliminando así la necesidad de introducir comandos por consola, acercando de esta manera los servicios de Supercomputación a un mayor número de usuarios sin la necesidad de tener un perfil técnico. A su vez la  
30 gestión de los datos de usuario se realizan mediante un protocolo propio que garantiza la seguridad de los datos almacenados.

Se basa en la comunicación de las máquinas que forman el clúster. Cada usuario dispone de un espacio privado con permisos para gestionarlo. El mecanismo de comunicación obliga a utilizar un protocolo estricto de intercambio de información

donde cualquier acción no contemplada en dicho protocolo está totalmente prohibida, creando de esta forma un blindaje sobre el servidor web, que es la máquina con riesgo de ser atacada.

5 El protocolo está basado en la creación de ficheros de intercambio, que manejan los datos. Crean un mapa de los datos del usuario, antes de acceder a ellos, realizando las modificaciones y peticiones de acceso sobre los mismos. De esta forma, nunca podrían realizarse modificaciones de forma directa. El fichero de intercambio se genera de forma automática, generado en lenguaje XML.

10 Este sistema garantiza que un ataque con éxito sobre el servidor web no puede ser aprovechado para hacerse con el control de los recursos de supercomputación ni la integridad de los datos del usuario. SCI conecta a gestor de colas con una interfaz web sin violar la seguridad de un clúster de supercomputación.

15 Entre el servidor web y los recursos de supercomputación existe un espacio compartido, que a su vez se divide en espacio compartido de entrada y espacio compartido de salida. En entrada se almacenan los ficheros que se envían a los recursos de supercomputación. En salida se almacenan los ficheros que se envían al servidor web desde los recursos de supercomputación.

## 20 Arquitectura Física

Los diferentes componentes que integran la arquitectura física de SCI son:

- Un servidor web, la máquina controladora del clúster y los nodos de computo.

## Seguridad

25 SCI, el servidor web y la máquina controladora están físicamente separadas, pero se habilita dentro de la máquina controladora un espacio el cual es visible por el servidor web y su función es ejercer como zona de intercambio entre ellas, a diferencia de otros sistemas patentados, este espacio montado por red en ambos servidores que hará de intercambio como en cualquier sistema pero  
30 estableciendo permisos restringidos en ambos servidores, el que lee los datos sólo tendrá acceso en lectura, y el que escribe los datos solo tendrá acceso en lectura. Este protocolo lo exponemos a continuación.

Esta zona de intercambio esta exportada desde la máquina controladora mediante el protocolo NFS (Network File System) a la máquina que alberga el  
35 servidor web. Los permisos en esta zona son muy restrictivos en escritura y

lectura, tan solo el UID (User IDentifier) bajo el que se ejecuta el servicio web tiene, dependiendo del caso, acceso a leer o a escribir. Este usuario solo existe en la máquina servidor web, por lo que únicamente son posibles las lecturas o escrituras procedentes de esta máquina, por ello, en el caso que se produzca una

5 intrusión en el servidor web, un atacante no podrá salir de la zona de intercambio, puesto que el UID del servidor web no tiene acceso al resto de la máquina controladora. Además, ni siquiera el usuario root de la máquina servidor web tiene acceso a ese espacio compartido, por lo que los temidos escalados de privilegios no tienen efecto.

10 Se cumplen los principios de buenas practicas en seguridad:

- Principio de la mínima exposición : Solo el servicio web es accesible por los usuarios.
  - Principio de defensa en profundidad: Espacio de disco común con permisos restrictivos y firewall de aplicación, mediante la aplicación
- 15 Modsecurity: OpenSource Web Application Firewall (<http://www.modsecurity.org/>), y Kernel con seguridad reforzada, mediante la aplicación SELinux: Security-Enhanced Linux (<http://selinux.sourceforge.net>).

## 20 Arquitectura Lógica

La arquitectura lógica de SCI indica como se realiza la comunicación entre las máquinas que formaran el clúster. El usuario dispone de un espacio privado dentro del

25 clúster y de los permisos suficientes para gestionarlo. Desde el punto de vista de la máquina controladora, el usuario tiene habilitada una cuenta personal y privada, con el espacio demandado donde albergar sus ficheros privados necesarios para el desarrollo de sus propósitos, así como los datos y el software pertinente. Por otro lado, el interfaz web permite al usuario visualizar este espacio personal de

30 ficheros de la forma mas cómoda y transparente posible dentro de la arquitectura subyacente.

Este mecanismo de comunicación obliga a utilizar un protocolo estricto de intercambio de información donde cualquier acción no contemplada en dicho

35 protocolo esta totalmente prohibida, creando de esta forma un blindaje sobre el servidor web, que es la máquina con riesgo de ser atacada. Un hecho importante

de esta arquitectura es que se garantiza que un ataque con éxito sobre dicho servidor web no puede ser aprovechado para hacerse con el control de la controladora del clúster, ni puede corromper los ficheros de la misma.

5 La máquina controladora comunica a la máquina del servidor web toda la información del espacio del usuario demandado, a través del espacio compartido por ambas máquinas. Además la controladora permite que la máquina servidora pueda usar este espacio personal para representar la distribución de ficheros del usuario. Es muy importante tener en cuenta que en el espacio compartido no se encuentran los ficheros del usuario, sino la información que representan estos (un  
10 mapa generado en XML), proporcionando así la seguridad e integridad necesarias sobre estos ficheros privados.

El espacio compartido en salida, igualmente, es el espacio que usa la máquina controladora para dejar los informes sobre los usuarios, el estado del  
15 clúster y los resultados de cálculo. Este espacio solo es escribible por la máquina generadora de dichos informes, dando únicamente permiso de lectura a la máquina servidor web que necesita capturar dicha información, asegurando así la integridad de los informes en caso de intrusión o sabotaje de la máquina servidor web. En este espacio, además de los informes generados por el sistema de colas  
20 instalado en la máquina controladora, también existe un espacio por cada usuario donde se albergan los informes sobre el espacio personal que posee cada usuario anteriormente comentado.

Los informes sobre el estado del clúster se generan periódicamente, mediante procesos automatizados por SCI. Los informes sobre el espacio de  
25 personal de cada usuario, solo se regeneran cuando se realiza algún cambio sobre dicho espacio.

De esta manera se consigue que el interfaz web de usuario pueda consultar los datos en cualquier momento, de forma muy eficiente debido a que se encuentra en un sistema de ficheros propio a la máquina (espacio compartido  
30 mediante NFS) y eliminando la sobrecarga de conexiones demandadas por los usuarios, entre el servidor web y la controladora.

Lo novedoso de esta idea no es la automatización, sino el conjunto hardware/software en la gestión de los recursos de un clúster de  
supercomputación. El conjunto está compuesto por: la interfaz que interactúa con  
35 el usuario, sencilla e intuitiva. Un servidor de salto al clúster y un espacio

compartido, protegido con el protocolo SCI que consisten en la creación de un fichero XML que describe un mapa de los contenidos de los espacios personales de usuario, en el cuál se realizarán los cambios de entrada/salida al clúster de recursos, protegiendo el contenido y la potencia de los recursos.

5

### **Modo de realización de la invención**

Para el acceso a la arquitectura el usuario solicita el alta en la misma, una vez autorizado para el acceso, la autenticación de usuarios a través de la interfaz web para el acceso a los recursos del clúster facilita al usuario el trabajo con el clúster.

10

Con el fin de que la información del usuario del clúster. sea confidencial e inalterable, se llevarán a cabo una serie de procesos basados en PKI. Una vez que el usuario se haya autenticado en la web mediante su certificado, podrá enviar tareas al clúster. las cuales suelen llevar asociadas ficheros de entrada que el usuario debe enviar al sistema de cálculo. Para esta tarea, que se realizará de forma transparente, el usuario recibirá la clave pública de la máquina controladora del clúster, que usará para cifrar los datos que desee enviar. Además, firmará estos datos digitalmente con su propia clave privada. De esta manera, se consigue que los datos del usuario sean únicamente legibles por la máquina controladora, gracias al cifrado, además de garantizar que los datos que se van a procesar son exactamente los que el usuario ha enviado, gracias a la firma digital.

15

20

Una vez la transferencia por la red de los datos cifrados y firmados ha concluido, la máquina controladora descifra los datos y comprueba la firma del usuario. Si todo ha ido bien, la tarea será enviada a los nodos de cálculo para su proceso.

25

Cuando un nodo de cálculo finalice su tarea, devolverá los resultados a la máquina controladora que, usando la clave pública del usuario que envió la tarea y su propia clave privada, cifrará y firmará los resultados, de manera que únicamente el usuario propietario de los resultados pueda leerlos.

30

Un usuario de SCI no necesitan especificar al interfaz web cómo se han de ejecutar sus tareas, ni cuantos nodos usará, etc. El interfaz sólo necesitará conocer el binario y los posibles argumentos de entrada, el resto de información necesaria para ejecutar la tarea, la intentará averiguar el propio SCI.



El sistema recibe un binario procedente del usuario y una lista de argumentos de entrada que puede ser tanto ficheros como datos explícitos. Una vez recabada toda la información, el interfaz comprobará contra qué librerías está compilado el binario y con dicha información podrá decidir si es una tarea MPI o una tarea JAVA etc . Además, según el tipo de tarea, el clúster tiene configurados por defecto el uso de un número de procesadores óptimo, aunque esto puede ser configurado por el usuario.

Una vez la interfaz analizada la tarea, construirá un informe que presentará al usuario, ofreciendo la posibilidad de modificarlo a usuarios experimentados.

10

## **Reivindicaciones**

1 - Sistema de información para la gestión de trabajos de supercomputación y computación paralela caracterizado porque comprende una máquina firewall, una máquina que ejercerá de servidor Web, una máquina controladora de nodos de supercomputación y un conjunto de máquinas (o nodos) de supercomputación que forman el cluster.

La sucesión de pasos que un usuario debe realizar para poder ejecutar una tarea de supercomputación al cluster consiste en:

- a) Realizar login en el servidor Web. Este login es seguro mediante el firewall que controla el acceso.
- b) Una vez que el usuario está autenticado de manera segura, el servidor Web le presenta una interfaz similar a la de cualquier ordenador personal.
- c) El usuario incluye el software en binario que quiere lanzar en el cluster a través de dicha interfaz.
- d) El servidor Web tras analizar el binario, paraleliza al máximo su código y se comunica con la máquina controladora indicando el número de nodos mínimos que se deben utilizar para ejecutarlo.
- e) La máquina controladora envía a los nodos necesarios el código a ejecutar y tras la finalización de cada uno de estos, devuelve el resultado al usuario.

Figura 1:

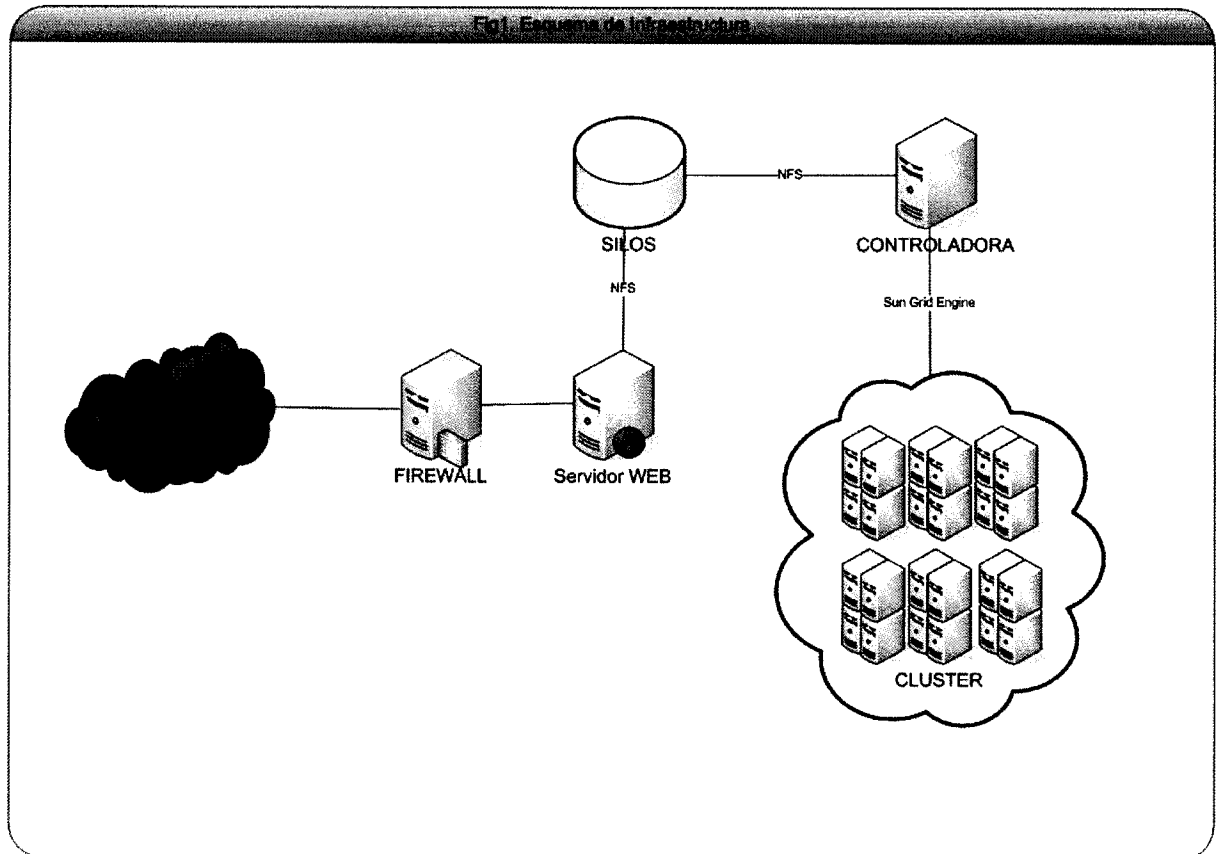
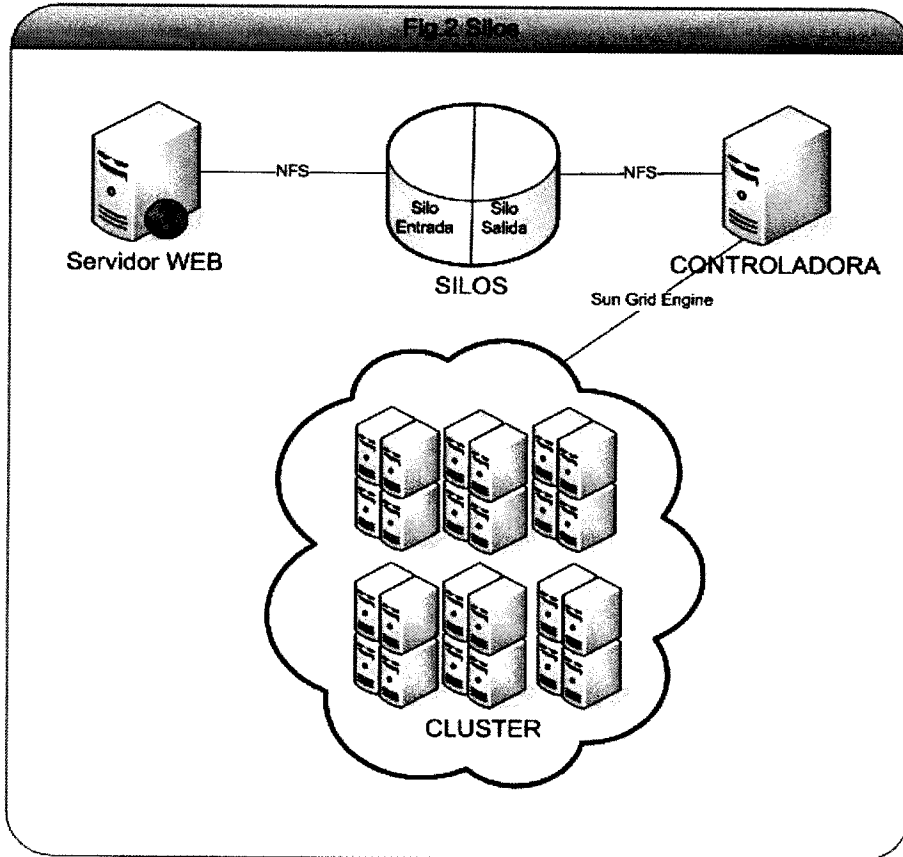


Figura 2:





OFICINA ESPAÑOLA  
DE PATENTES Y MARCAS

ESPAÑA

②① N.º solicitud: 201001371

②② Fecha de presentación de la solicitud: 22.10.2010

③② Fecha de prioridad:

## INFORME SOBRE EL ESTADO DE LA TÉCNICA

⑤① Int. Cl.: **G06F9/44** (2006.01)  
G06F21/00 (2006.01)

### DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
X	CANTÓN, J. et al: "Interfaz web para gestionar los recursos de supercomputación". RedIRIS:boletín de la Red Nacional de I+D RedIRIS, ISSN 1139-207X, Nº. 85-86, 2009 , págs. 61-66.URL: <a href="http://www.rediris.es/difusion/publicaciones/boletin/85-86/ponencias85-8.pdf">http://www.rediris.es/difusion/publicaciones/boletin/85-86/ponencias85-8.pdf</a> . Todo el documento.	1

#### Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

#### El presente informe ha sido realizado

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
27.07.2012

Examinador  
M. L. Alvarez Moreno

Página  
1/4

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

G06F

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, Inspec, Internet

Fecha de Realización de la Opinión Escrita: 27.07.2012

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones	<b>SI</b>
	Reivindicaciones 1	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones	<b>SI</b>
	Reivindicaciones 1	<b>NO</b>

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

**Base de la Opinión.-**

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

**1. Documentos considerados.-**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
	CANTÓN, J. et al: "Interfaz web para gestionar los recursos de supercomputación". Red IRIS:boletín de la Red Nacional de I+D RedIRIS, ISSN 1139-207X, Nº. 85-86, 2009 , págs. 61-66.URL: <a href="http://www.rediris.es/difusion/publicaciones/boletin/85-86/ponencias85-8.pdf">http://www.rediris.es/difusion/publicaciones/boletin/85-86/ponencias85-8.pdf</a> . Todo el documento.	

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

El objeto de la invención recogido en la reivindicación 1 ha sido divulgado de forma idéntica en el documento D01. El documento D01 muestra un sistema que permite al usuario hacer uso de los recursos de computación a través de la web (resumen). Su contenido coincide plenamente con el contenido de la descripción aportada en la solicitud. La arquitectura física del sistema [Apartado 2. Arquitectura Física; Figura 2] está compuesta por los nodos de supercomputación que forman el cluster, la máquina controladora de los mismos, y el servidor Web encargado de interaccionar tanto con el usuario y como con la zona de intercambio que la máquina controladora pone a su disposición. El usuario interactúa con el sistema [Apartado 3. Arquitectura lógica; Apartado 4. Interacción con el usuario; Apartado 5. Interfaz Web] previa autenticación segura a través del interfaz web. Una vez autenticado, el servidor Web le presenta la interfaz apropiada para la interacción apropiada con el sistema. El usuario introduce el binario que quiere enviar al cluster. El sistema procesa la información y decide cómo debe ejecutar las tareas y el número de nodos a usar. Tras la finalización de las tareas se devuelven los resultados al usuario. El documento D01 muestra la existencia de los elementos definidos en la reivindicación 1 y la realización de las mismas acciones.

A la vista del documento D01 la reivindicación 1 carece de novedad según el artículo 6 de la Ley de Patentes.