(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 700 929**

(51) Int. CI.:

H04L 12/24 (2006.01)
H04L 29/08 (2006.01)
G06F 9/54 (2006.01)
G06F 17/30 (2006.01)
H04L 12/26 (2006.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(96) Fecha de presentación y número de la solicitud europea: **30.03.2015 E 15161764 (4)**

(97) Fecha y número de publicación de la concesión europea: **10.10.2018 EP 2955875**

(54) Título: **Servidor, cliente y sistema de gestión de una red de interconexión**

(30) Prioridad:

13.06.2014 FR 1455446

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
20.02.2019

(73) Titular/es:

BULL S.A.S. (100.0%)
Rue Jean Jaurès, B.P. 68
78340 Les Clayes-sous-Bois, FR

(72) Inventor/es:

VIGNERAS, PIERRE y
MIQUEE, SÉBASTIEN

(74) Agente/Representante:

STEPHANN, Valérie

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Servidor, cliente y sistema de gestión de una red de interconexión

La invención se refiere a un sistema de gestión para la gestión de una red de interconexión.

Se entiende, en este caso, por “red de interconexión”, cualquier red de cálculo dedicada (tal como una red InfiniBand), o más generalmente cualquier colección de elementos informáticos, principalmente procesadores distribuidos, con enlaces de comunicación física entre ellos.

Un sistema de gestión, que utiliza una red secundaria, pudiendo ser de tipo Ethernet, se configura para gestionar esta red de cálculo en un modo fuera de banda (out-of-band).

Sin embargo, con el crecimiento del tamaño de los súper-calculadores, las topologías de las redes de cálculo de alto rendimiento se hacen más complejas en tanto que se densifican. De ahí que la gestión fuera de banda, por medio de una red de gestión dedicada también denominada “red secundaria” de la red de interconexión de un súper-calculador necesita más de un componente para:

- reaccionar a los eventos que pudieran producirse en esta red de interconexión (averías de cables, de puertos, de conmutadores o de nodos, por ejemplo);
- verificar las eventuales modificaciones de la topología vinculadas a cambios de cables o de conmutadores (conexión, conformidad con las topologías teóricas);
- calcular las nuevas tablas de enrutado en función de las modificaciones y de las actualizaciones cargadas en los conmutadores.

El documento US 2003/140132 A1 describe un sistema de gestión de red que contiene, además de una base de datos de configuración de un dispositivo de red administrado, un servidor y un cliente que comunican con el dispositivo de red administrado, pudiendo recibir dicho dispositivo las solicitudes de cambio de configuración, no solamente por parte del sistema de gestión, a través de una primera interfaz, sino también por parte de otra entidad, a través de una segunda interfaz. Visto que los datos de configuraciones almacenados en la base de datos pueden convertirse en obsoletos debido a las demandas de su modificación de configuración recibidas en la segunda interfaz, para actualizar las informaciones en la base de datos, el sistema de gestión envía una solicitud de configuración de descarga al dispositivo de red, que responde transfiriendo un fichero de configuración que contiene las informaciones de configuración en curso de ejecución al sistema de gestión. El sistema de gestión utiliza el contenido del archivo de configuración descargado para actualizar las informaciones en la base de datos. Este proceso de sincronización puede iniciarse por el cliente de gestión de la red en el sistema de gestión en respuesta a una entrada de un usuario pero puede iniciarse igualmente a través de un mensaje del dispositivo de red durante la puesta en tensión del dispositivo.

Además, por razones de cambio de escala y de robustez, cada una de las tareas anteriores se ha de realizar por procesos separados, potencialmente desde máquinas dedicadas diferentes. Con este fin, un mecanismo de comunicación eficaz permite a estos componentes compartir un estado global común mediante un diálogo. Este mecanismo debe asegurarse por el sistema de gestión. Los mensajes se intercambian en la red secundaria (o red de gestión) de la red de interconexión del súper-calculador. El sistema de gestión permite a los procesos a cargo de la gestión de la red de interconexión, comunicar y compartir un estado global, del que un subconjunto representa el estado de la red de interconexión del súper-calculador (es decir el conjunto de los estados de los equipos que componen la red de interconexión del súper-calculador).

Con el fin de asegurar la comunicación entre procesos en ejecución en máquinas diferentes de la red de gestión, esta red de gestión debe, principalmente,

- poder soportar un gran número de procesos (varios centenares incluso varios millares);
- permitir a un proceso unirse o salirse, en cualquier momento, del sistema de gestión (por razones de avería, o de actualización, por ejemplo);
- mantener un estado global actualizado de la red de interconexión del súper-calculador;
- mantener un estado global coherente a largo plazo (“eventual consistency” en terminología anglosajona) de la red de interconexión del súper-calculador;
- representar este estado global en la forma de una tabla de tipo “clave-valor”;
- compartir este estado global entre los diferentes procesos interconectados del sistema de gestión;
- permitir en cualquier momento y en cada uno de los procesos del sistema de gestión actualizar un subconjunto incluso la integridad del estado global de la red de interconexión;
- gestionar la frecuencia de las actualizaciones del estado global (preferentemente, como máximo algunos centenares por segundo);
- limitar el acceso al estado global, no debiendo/deseando ciertos procesos ver más que un subconjunto del estado global;

- poder tener en la memoria activa el estado global incluso si puede tener un número relativamente grande de entradas (generalmente, varios centenares de millares).

En este sentido, existen sistemas de compartición de estados tales como las tablas de la función hash distribuidas. Sin embargo, ninguna de las soluciones existentes satisface el conjunto de las exigencias anteriores a las que un sistema de gestión debe responder. En este caso, una tabla de función hash distribuida no puede ofrecer las dos últimas restricciones enunciadas anteriormente.

Un objeto de la presente invención es proponer un mecanismo de comunicación entre procesos en la forma de un sistema de gestión que responda a las exigencias antes mencionadas.

Otro proyecto de la presente invención es proponer una arquitectura de comunicación de tipo cliente/servidor para interconectar procesos distribuidos.

Otro objeto de la presente invención es proponer un sistema de gestión de una red de interconexión basada en una comunicación entre procesos.

Otro objeto de la presente invención es proponer un mecanismo de comunicación entre procesos asíncrono y desconectado.

Otro proyecto de la presente invención es proponer un método de gestión de la red de interconexión de un súper-calculador.

Con estos fines, la invención se refiere, según un primer aspecto, a un servidor de un sistema de gestión de una red de interconexión, comprendiendo este servidor

- una estructura de datos asociativa clave-valor configurada para almacenar un estado global de la red de interconexión;
- las interfaces de conexión siguientes

    ○ una interfaz del servidor de configuración configurada para comunicar, en respuesta a una solicitud de configuración procedente de un cliente para unirse al sistema de gestión, la configuración de dicho sistema de gestión a este cliente;
    ○ una interfaz del servidor de publicación configurada para difundir, a un cliente conectado al sistema de gestión, una actualización en la estructura de datos asociativa clave-valor;
    ○ una interfaz del servidor de recopilación configurada para actualizar, en respuesta a una solicitud de actualización procedente de un proceso cliente conectado al sistema de gestión, la estructura de datos asociativa clave-valor;
    ○ una interfaz del servidor de instantánea configurada para comunicar, en respuesta a la solicitud de instantánea procedente de un proceso cliente conectado al sistema de gestión, una instantánea de un subconjunto del estado global de este proceso cliente, estando almacenado este estado global en la estructura de datos asociativa clave-valor.

El servidor de un sistema de gestión de una red de interconexión presenta, según diversas realizaciones, los siguientes caracteres, combinados si procede:

- la interfaz del servidor de configuración se configura para comunicar la dirección de la interfaz del servidor de publicación, la dirección de la interfaz del servidor de recopilación, y la dirección de la interfaz del servidor de instantánea;
- la estructura de datos asociativa clave-valor es una tabla de función hash;
- la implementación de este servidor se basa, para reaccionar a la solicitud de configuración, en la solicitud de actualización o en la solicitud de instantánea, sobre el paradigma de evento de interfaz del servidor de conexión, siendo esta interfaz del servidor de conexión la interfaz del servidor de configuración, la interfaz del servidor de recopilación o la interfaz del servidor de instantánea.

La invención se refiere, según un segundo aspecto, a un cliente de un sistema de gestión de una red de interconexión, comprendiendo este cliente

- un proceso cliente;
- una interfaz cliente de instantánea configurada para recuperar, desde el servidor del sistema de gestión, una instantánea de un subconjunto del estado global de la red de interconexión;
- una interfaz cliente de abono configurada para recibir, desde el servidor del sistema de gestión, una actualización del estado global de la red de interconexión;
- una interfaz cliente de configuración configurada para recuperar, desde el servidor del sistema de gestión, la configuración del sistema de gestión;

- una interfaz cliente de publicación de manera que el proceso cliente pueda publicar una actualización del estado global de la red de interconexión, siendo compartido y almacenado este estado global de la red de interconexión por el servidor del sistema de gestión.

El cliente de un sistema de gestión de una red de interconexión comprende, además, un proceso operativo, estando provisto este proceso operativo de una interfaz cliente de publicación asociada a este proceso operativo de manera que este proceso operativo pueda publicar una actualización del estado global de la red de interconexión.

Ventajosamente, el dato publicado por el proceso cliente es un mensaje en la forma de un mensaje “clave-valor”.

La invención se refiere, según un tercer aspecto, a un sistema de gestión para la gestión de una red de interconexión que comprende el servidor y el cliente presentados anteriormente.

La invención se refiere, según un cuarto aspecto, a un súper-calculador que comprende una red de interconexión y el sistema de gestión citado anteriormente.

El súper-calculador comprende, además,

- un gestor de topología configurado para verificar la topología de la red de interconexión;
- un módulo de supervisión configurado para recuperar una alarma enviada por un equipo de red de la red de interconexión, y para actualizar, en consecuencia, el sistema de gestión;
- un calculador de enrutado configurado para calcular una tabla de enrutado para un dato recibido desde el sistema de gestión.

Surgirán otros objetos y ventajas de la invención a la luz de la descripción de los modos de realización, realizada a continuación con referencia a los dibujos adjuntos en los que:

- la figura 1 ilustra un servidor de un sistema de gestión según un modo de realización;
- la figura 2 ilustra un cliente de un sistema de gestión según un modo de realización;
- la figura 3 ilustra la implementación de un sistema de gestión según un modo de realización.

El sistema de gestión de una red de interconexión se basa en una arquitectura de red de tipo cliente/servidor que implementa paradigmas de comunicación diferentes en función de las interfaces de conexión.

Con referencia a la figura 1, el lado del servidor (o la implementación del servidor) de esta arquitectura de red del sistema de gestión comprende un servidor 10 de la red de gestión.

El servidor 10 se configura para conservar el estado global de la red de interconexión. Este estado global comprende informaciones que se refieren a la red de interconexión del súper-calculador. Este estado global se almacena en una estructura de datos 2 asociativa clave-valor. Esta estructura de datos 2 se almacena en una memoria activa del servidor 10.

La estructura de datos 2 asociativa clave-valor es un contenedor de datos, preferentemente local en el servidor 10, que posee un protocolo particular para añadir, retirar y buscar elementos. Esta estructura de datos 2 asociativa clave-valor asocia una clave a un valor. Debe garantizarse por los procesos emisores la unicidad de las claves. Si la actualización se realiza sobre una clave ya existente, el antiguo valor es borrado por el nuevo.

En un modo de realización, la estructura de datos 2 asociativa clave-valor es una tabla asociativa, también llamada tabla de función hash o en inglés “hashmap”, que tiene una función de asociación o de función hash predefinida. Ventajosamente, esta estructura particular de datos permite un acceso rápido a un valor en función de la clave.

El servidor 10 de la red de gestión comprende, además,

- una interfaz del servidor de configuración 3 configurada para gestionar la(s) solicitud(es) de configuración procedente(s) de un cliente que desea unirse al sistema de gestión;
- una interfaz del servidor de publicación 4 configurada para difundir, a los clientes conectados al sistema de gestión, actualizaciones que tienen lugar en la estructura de datos 2 asociativa clave-valor, es decir una entrada de la estructura de datos 2 asociativa clave-valor;
- una interfaz del servidor de recopilación 5 configurada para gestionar las solicitudes de actualización (de las actualizaciones, generalmente, en la forma de un mensaje clave-valor) procedentes de los procesos cliente conectados al sistema de gestión;
- una interfaz del servidor de instantánea 6 (o de “snapshot” en terminología anglosajona) configurada para enviar el estado actual, es decir una instantánea del conjunto o solamente de un subconjunto de la estructura de datos 2 asociativa clave-valor (pudiendo hacerse el filtrado a nivel del servidor 10, con el fin de limitar la utilización de la banda pasante), en respuesta a una solicitud de instantánea procedente de un proceso cliente conectado al sistema de gestión.

La interfaz del servidor de configuración 3 permite comunicar la configuración del sistema de gestión a un cliente que desee unirse a este sistema de gestión de la red de interconexión. La configuración enviada comprende, principalmente, las direcciones (o URL por “Uniform Resource Locator”) de las otras interfaces de conexión citadas anteriormente, a saber la interfaz del servidor de publicación 4, la interfaz del servidor de recopilación 5, y la interfaz del servidor de instantánea 6.

Preferentemente, las interfaces del servidor de conexión 3-6 son unos conectores ZeroMQ (el documento en el enlace siguiente: http://zeromq.org/intro:read-the-manual especifica los conectores ZeroMQ o ZMQ). Este modo de realización permite, en efecto, evitar los problemas del paradigma solicitud/respuesta.

Se ha de observar que en la implementación del servidor, representada en la figura 1, del sistema de gestión

- la interfaz del servidor de configuración 3 no interactúa con la estructura de datos 2 asociativa clave-valor;
- la interfaz del servidor de recopilación 5 interactúa con esta estructura de datos 2 asociativa clave-valor en modo escritura para almacenar en ella entradas clave-valor;
- la interfaz del servidor de publicación 4 y la interfaz del servidor de instantánea 6 interactúan con la estructura de datos 2 asociativa clave-valor en modo lectura.

Esta implementación servidor del sistema de gestión puede realizarse en lenguaje C, Python, o cualquier otro lenguaje de programación apropiado, así como que puede estar en un modo mono-hilo o multi-hilo.
En una implementación mono-hilo de servidor 10, este último ejecuta el bloque siguiente, basándose en el paradigma de eventos de interfaz servidor de conexión 3, 5, 6,

```
EN TANTO que Verdad
    SEGÚN “interfaz del servidor de conexión”
        CASO “interfaz del servidor de configuración”
            enviar los datos de configuración a través de la interfaz del servidor de configuración
        CASO “interfaz del servidor de recopilación”
            enviar la actualización a través de la interfaz del servidor de publicación
            actualizar la estructura de datos 2
        CASO “interfaz del servidor de instantánea”
            enviar una instantánea través de la interfaz del servidor de instantánea
        CASO “el retardo de una temporización ha expirado”
            enviar un mensaje de presencia
    FIN
FIN
```

Ventajosamente, esta implementación basada en el paradigma de eventos de interfaz del servidor de conexión 3, 5, 6 permite reaccionar rápidamente ante una acción después de un proceso cliente (solicitud o actualización). La instrucción condicional “SEGÚN” gestiona las tres interfaces del servidor de conexión (de configuración 3, de recopilación 5 y de instantánea 6) y un tiempo de espera (una temporización) predefinida. Devuelve una lista que comprende la o las interfaces de conexión activadas por un evento de una interfaz del servidor de conexión 3, 5, 6. Si no se produce ningún evento antes de que expire el retardo de temporización, el servidor 10 envía a los clientes del sistema de gestión un mensaje de presencia (designado por “Heartbeat message” en terminología anglosajona).

El orden particular en el que se tratan las interfaces del servidor de conexión 3, 5, 6 permite

- en primer lugar, a un cliente conectarse al servidor 10 y configurarse con ayuda de las direcciones (URL) de las otras interfaces del servidor de conexión 4-6 que se le comunican a través de la interfaz del servidor de configuración 3;
- en segundo lugar, si el cliente envía automáticamente (o, en terminología anglosajona, “push”) una actualización, tiene en cuenta esta actualización desde que sea posible por el servidor 10 y la publica ante todos los clientes (comprendidos en ellos aquellos que están en el origen de actualización), pudiendo activar esta actualización acciones en otros clientes del sistema de gestión;
- en tercer lugar, a una solicitud de una instantánea (en terminología anglosajona, de un “snapshot”) de ser tratada en último lugar porque le lleva, en general, el tiempo más largo para ejecutarse.

Ventajosamente, en una implementación mono-hilo del servidor 10, no es necesario un mecanismo de enclavamiento para evitar accesos simultáneos y concurrentes a la estructura de datos 2 y/o la utilización simultánea de una interfaz del servidor de conexión 3-6.

El servidor 10 del sistema de gestión recibe y trata entradas de tipo clave-valor que pueden interpretarse como comandos. Estas entradas pueden proceder de un cliente del sistema de gestión. Ventajosamente, esto permite tener entradas almacenadas, principalmente cuando la estructura de datos 2 se describe en archivos del sistema (el caso de un sistema basado en Unix por ejemplo).

La figura 2 ilustra una implementación cliente o el “lado cliente” de la arquitectura de cliente-servidor del sistema de gestión de una red de interconexión.

Se entiende, en este caso, por un cliente 20 del sistema de gestión un proceso cliente 21 que se configura, en general, para soportar al menos un proceso operativo 31 para la ejecución de un código operativo 30.

El cliente 20 del sistema de gestión de una red de interconexión, comprende

- una interfaz cliente de instantánea 26 configurada para recuperar una instantánea del estado global de la red de interconexión, o de un subconjunto de esta;
- una interfaz cliente de abono 24 configurada para recibir actualizaciones desde el servidor del sistema de gestión;
- una interfaz cliente de configuración 23 configurada para recuperar la configuración del servidor del sistema de gestión, en particular las direcciones (URL) de las interfaces del servidor de conexión del servidor;
- una pluralidad de interfaces cliente de publicación 25 de manera que el proceso cliente 21 y cualquier otro proceso operativo 31 de un código operativo 30 pueda publicar, por medio de entidades “claves-valor”, las actualizaciones de (o, más en general, un dato que se refiere a) sus estados respectivos a través de una interfaz cliente de publicación 25 apropiada. Se ha de observar que un código operativo 30 puede comprender, para sus propios fines, más de un proceso operativo 31 al que se asocia una interfaz cliente de publicación 25;
- una interfaz cliente de comunicación 32 entre el proceso operativo 31 del código operativo 30 y el proceso cliente 21.

El proceso cliente 21 gestiona las comunicaciones entrantes del cliente 20 del sistema de gestión. Para ello, este proceso cliente 21 dispone de sus propias interfaces cliente de entrada 24 y 26, es decir la de la instantánea 26 y la del abono 24.

La interfaz cliente de publicación 25 permite a un proceso operativo 31 actualizar el estado global de la red de interconexión que se comparte y almacena en el lado del servidor.

En la figura 3, el cliente 20 se conecta al servidor 10 del sistema de gestión.

Con el fin de conectarse al servidor 10, el cliente 20 se conecta, inicialmente, a la interfaz del servidor de configuración 3 para recibir el conjunto de los parámetros de configuración del sistema de gestión, entre ellos particularmente las direcciones (URL) de las otras interfaces del servidor de conexión 4-6 del servidor 10. En un modo de realización, a partir de la recepción en la interfaz del servidor de configuración 3 de una solicitud de configuración (por ejemplo, de tipo CONFIG?), el servidor 10 (ver el enlace entre las interfaces de conexión 3 y 23)

- comunica a través de la interfaz del servidor de configuración 3, en respuesta a esta solicitud, las direcciones (URL) de las otras tres interfaces de conexión 4-6; y
- envía, a través de esta interfaz del servidor de configuración 3, un mensaje que indica el final de la etapa de configuración (de tipo, por ejemplo, Fin_CONFIG).

Con la recepción de los parámetros de configuración, el cliente 20 decide recuperar o no una instantánea (un “snapshot”) del estado global actual de la red de interconexión (ver el enlace entre las interfaces de conexión 6 y 26). Si es este el caso,

- el cliente 20 requiere, de la interfaz servidor de instantánea 6, una instantánea del estado global de la red de interconexión que se almacena en la estructura de datos 2 asociativa clave-valor;
- el servidor 10 espera que el cliente 20 reciba las jerarquías a las que desea abonarse (estas jerarquías pueden, principalmente, ser diferentes de las especificadas durante una conexión en la interfaz servidor de publicación 4 del servidor 10). A continuación, el servidor 10 envía al cliente 20, utilizando la interfaz servidor de instantánea 6, una instantánea del estado de las jerarquías requeridas;
- el servidor 10 envía el número de revisión actual de la estructura de datos 2 asociativa clave-valor.

Se deduce que es el cliente 20 el encargado de conectarse a la interfaz del servidor de publicación 4 (véase el enlace entre las interfaces de conexión 4 y 24) para recibir actualizaciones, y a la interfaz servidor de recopilación 5 (véase el enlace entre las interfaces de conexión 5 y 25) para enviar eventualmente actualizaciones (o nuevas entradas).

Preferentemente, las actualizaciones publicadas por el servidor 10 están fechadas. Ventajosamente, el fechado (o “timestamping” en inglés) del lado del servidor garantiza la coherencia a largo plazo (“eventual consistency”) del estado global en el seno de cada cliente 20. Se observa que un cliente 20 puede ignorar actualizaciones enviadas durante la recuperación de una instantánea cuando estas son anteriores (gracias al fechado) a la instantánea recibida.

Aparte del tiempo de conexión de un cliente 20, el servidor 10 se configura para reaccionar a cada solicitud de actualización y enviar, regularmente, mensajes de presencia (“heartbeat”).

Con la recepción de una solicitud de actualización, o más generalmente de un comando, a través de la interfaz del servidor de recopilación 5,

- si este comando no se refiere al servidor 10 o si no es un comando especial (tal como un comando de parada “shutdown”), se transmite a los otros clientes a través de la interfaz del servidor de publicación 4;
- el servidor 10 aplica a la estructura de datos 2 el contenido de esta solicitud de actualización. En el caso de un mensaje de comando del servidor, este mensaje se compara inicialmente con el comando “vaciar” que requiere la supresión de todo el contenido de la estructura de datos 2.

Utilizando los campos especiales de datos de entrada, tales como “vaciar” o “ttl” por “time to live”, el servidor 10 puede suprimir este dato de entrada de la estructura de datos 2 local. En caso contrario, si no existe ningún dato de entrada, entonces se añade o actualiza en la estructura de datos 2.

Como ejemplos de solicitudes que el servidor 10 puede recibir, se citan

- .bbctrl/_server_kill_: requerir del servidor 10 salir (este mensaje no se difunde por el servidor 10);
- .bbcrtl/_clear_all_: requerir del servidor 10 y de todos los clientes suprimir el conjunto de la estructura de datos (difundiéndose este mensaje por el servidor 10);
- .bbctrl/_hugz_: un mensaje de presencia por parte del servidor.

La entidad de base, en este caso, es un mensaje basado sobre el paradigma “clave-valor”. Esta entidad proporciona, además del par clave-valor, otras informaciones tales como el número de la secuencia, el identificador universal único de esta entidad, la identidad de su emisor, o la vida útil de la entidad (“Time to Live”).

Estas informaciones pueden enviarse en una o más de diversas trama(s) de datos. Preferentemente, estas informaciones se envían en una primera trama que contiene la clave para el mecanismo de suscripción, y una segunda trama que contiene el resto de las informaciones (datos, identificador del emisor, número de secuencia, por ejemplo).

Preferentemente, se calcula un tiempo de reemplazo dinámico antes de que el detector de la interfaz del servidor de recopilación 5 comience un periodo de inactividad.

Ventajosamente, la implementación en servidor 10 de la red de gestión permite

- proporcionar los datos de configuración a los clientes;
- proporcionar instantáneas del estado global de la red de interconexión a los clientes del sistema de gestión nuevamente conectados; y
- transferir actualizaciones en el estado global de la red de interconexión a todos los clientes conectados al sistema de gestión.

En un modo de realización, el sistema de gestión se utiliza para la gestión de la red de interconexión (o de cálculo) de un súper-calculador. Para ello, pueden requerirse los módulos siguientes

- un gestor de topología responsable de la verificación de la topología (el estado de los equipos por ejemplo o el esquema de conexión de los equipos) de la red de interconexión;
- un módulo de supervisión responsable de la recuperación de las alarmas enviadas por los diferentes equipos de la red del súper-calculador (principalmente, los conmutadores), correlacionándolas con el fin de actualizar el sistema de gestión con datos pertinentes (actualización de los estados por ejemplo);
- un calculador de enrutado responsable del cálculo de las tablas de enrutado para unos datos recibidos desde el sistema de gestión.

Cuando las componentes adicionales anteriores arrancan, el servidor 10 de la red de gestión ya está presente de manera que el conjunto de los datos publicados por el gestor de topología pueda recuperarse por los otros clientes del servidor 10.

Ventajosamente, las actualizaciones de los estados se realizan de manera unitaria, y después de la agregación, se envía una clave especial, llamada elemento activador. Es con la recepción de esta última cuando el calculador de enrutado activa su fase de cálculo de las nuevas tablas de enrutado. De ese modo, es posible diferenciar fácil y rápidamente los diferentes tipos de mensajes enviados por el servidor del sistema de gestión

- las actualizaciones del estado global de la red de interconexión;
- los elementos activadores que permiten activar una acción.

Resultando de ahí, ventajosamente, una agregación de los eventos con el fin de ganar en eficacia. En particular, el calculador de enrutado no calcula las tablas de enrutado para cada cambio de estado en un único equipo.

En este sentido, un mecanismo a base de prefijos de las claves permite no solamente a cada cliente abonarse a un subconjunto del estado global, sino también hacer esta diferenciación. Además, este mecanismo permite añadir en el sistema de gestión otros tipos de información que se encuentran en los subconjuntos específicos. A título de ejemplo, pueden utilizarse los prefijos siguientes en un sistema de gestión:

- @interconnect/: conjunto de las entradas que se refieren al estado de los equipos de la red de interconexión;
- !activadores/: conjunto de las entradas correspondientes a los elementos activadores;
- /mngt/: conjunto de las entradas que contienen informaciones de conexión a los de diversos equipos de la red de interconexión;
- .bbctrl/: conjunto de las entradas que permiten controlar el sistema de gestión, tanto del lado del cliente como del servidor (por ejemplo un comando que permita borrar toda la tabla de entradas).

Ventajosamente, los diferentes modos de realización descritos anteriormente utilizan paradigmas de comunicación diferentes en función de las interfaces de conexión. Por ejemplo, la interfaz del servidor de publicación 4 del servidor 10 funciona en modo “difusión” y puede utilizar un protocolo de “multiemisión” (simétrico con la interfaz del servidor de abono 24 del lado del cliente 20). Además, un cliente 20 puede abonarse a uno o varios prefijos de claves, lo que le permite no recuperar más que un subconjunto del estado global almacenado en la estructura de datos 2.

La invención se define por la extensión de las reivindicaciones.

REIVINDICACIONES

1. Un servidor (10) de un sistema de gestión de una red de interconexión, comprendiendo este servidor

- una estructura de datos (2) asociativa clave-valor configurada para almacenar un estado global de la red de interconexión;
- las interfaces de conexión (3-6) siguientes

  ○ una interfaz del servidor de configuración (3) configurada para comunicar, en respuesta a una solicitud de configuración procedente de un cliente (20) para unirse al sistema de gestión, la configuración de dicho sistema de gestión a este cliente (20);
  ○ una interfaz del servidor de publicación (4) configurada para difundir, a un cliente (20) conectado al sistema de gestión, una actualización en la estructura de datos (2) asociativa clave-valor;
  ○ una interfaz del servidor de recopilación (5) configurada para actualizar, en respuesta a una solicitud de actualización procedente de un proceso cliente (21) comprendido en un cliente conectado al sistema de gestión, la estructura de datos (2) asociativa clave-valor;
  ○ una interfaz del servidor de instantánea (6) configurada para comunicar, en respuesta a la solicitud de instantánea procedente de un proceso cliente (21) conectado al sistema de gestión, una instantánea de un subconjunto del estado global de este proceso cliente (21), estando almacenado este estado global en la estructura de datos (2) asociativa clave-valor.

comprendiendo la configuración de dicho sistema de gestión la dirección de la interfaz del servidor de publicación (4), la dirección de la interfaz del servidor de recopilación (5), y la dirección de la interfaz del servidor de instantánea (6).

2. El servidor (10) de la reivindicación 1, en el que la estructura de datos (2) asociativa clave-valor es una tabla de función hash.

3. El servidor de la reivindicación 1 o 2, en el que la implementación de este servidor se basa, para reaccionar a la solicitud de configuración, en la solicitud de actualización o en la solicitud de instantánea, sobre el paradigma de evento de interfaz del servidor de conexión, siendo esta interfaz del servidor de conexión la interfaz del servidor de configuración (3), la interfaz del servidor de recopilación (5) o la interfaz del servidor de instantánea (6).

4. Un cliente (20) de un sistema de gestión de una red de interconexión, comprendiendo este cliente

- un proceso cliente (21);
- una interfaz cliente de configuración (23) configurada para recuperar, desde una primera interfaz (3) de un servidor del sistema de gestión, la configuración del sistema de gestión para unirse a este sistema de gestión;
- una interfaz cliente de abono (24) configurada para recibir, desde una segunda interfaz (4) del servidor del sistema de gestión, una actualización del estado global de la red de interconexión;
- una interfaz cliente de publicación (25) de manera que el proceso cliente (21) pueda publicar con destino en una tercera interfaz (5) del servidor del sistema de gestión una actualización del estado global de la red de interconexión, siendo compartido y almacenado este estado global de la red de interconexión por el servidor del sistema de gestión.
- una interfaz cliente de instantánea (26) configurada para recuperar, desde una cuarta interfaz (26) del servidor del sistema de gestión, una instantánea de un subconjunto del estado global de la red de interconexión,

comprendiendo la configuración de dicho sistema de gestión la dirección de la segunda interfaz (4), la dirección de la tercera interfaz (5), y la dirección de la cuarta interfaz (6).

5. El cliente de la reivindicación anterior, que comprende además un proceso operativo, estando provisto este proceso operativo de una interfaz cliente de publicación (25) asociada a este proceso operativo de manera que este proceso operativo pueda publicar una actualización del estado global de la red de interconexión.

6. El cliente de la reivindicación 4 o 5, en el que la actualización publicada es un mensaje en la forma de un mensaje “clave-valor”.

7. Un sistema de gestión para la gestión de una red de interconexión que comprende un servidor según una cualquiera de las reivindicaciones 1 a 3 y un cliente según una cualquiera de las reivindicaciones 4 a 6.

8. Un súper-calculador que comprende una red de interconexión y un sistema de gestión de dicha red de interconexión según la reivindicación 7.

9. El súper-calculador de la reivindicación anterior que comprende además,

- un gestor de topología configurado para verificar la topología de la red de interconexión;
- un módulo de supervisión configurado para recuperar una alarma enviada por un equipo de red de la red de interconexión, y para actualizar, en consecuencia, el sistema de gestión;
- un calculador de enrutado configurado para calcular una tabla de enrutado para un dato recibido desde el sistema de gestión.

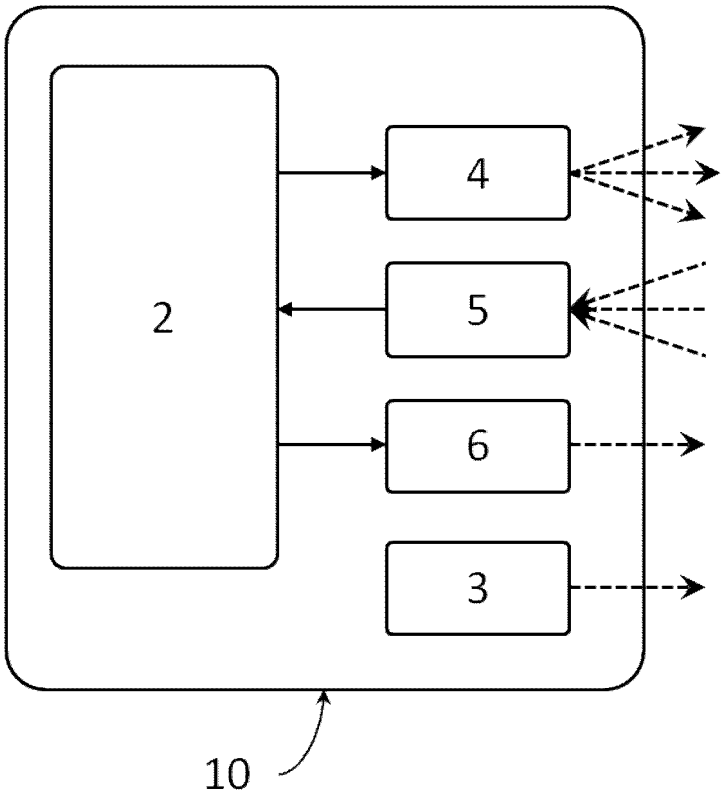

FIG.1

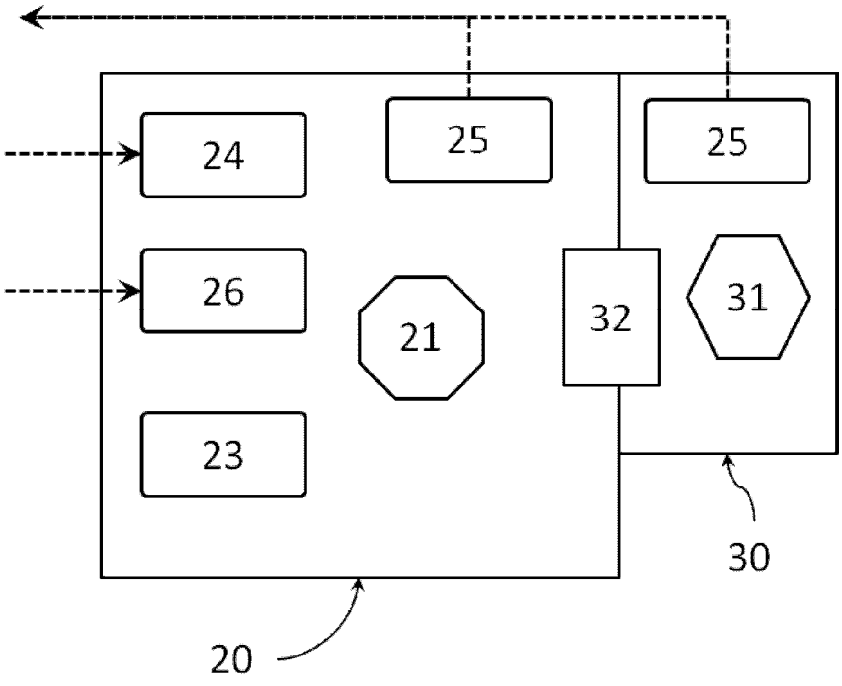

FIG.2

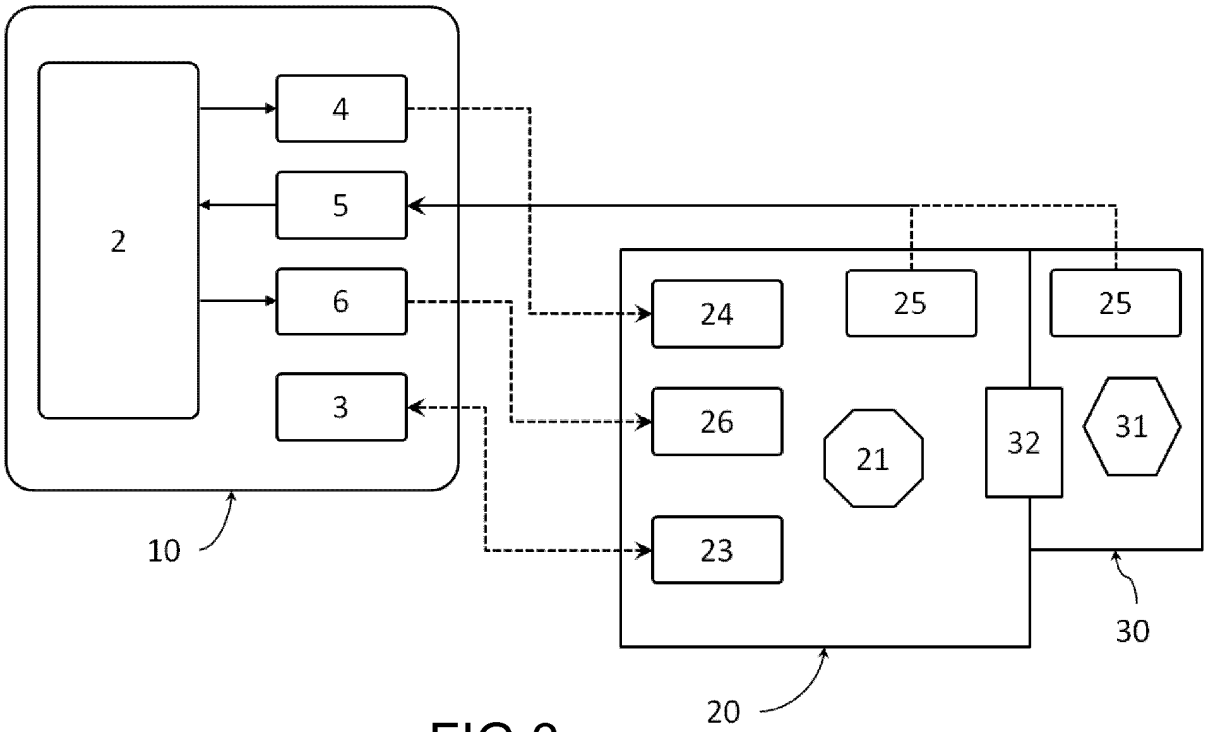

FIG.3