



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 711**

51 Int. Cl.:  
**H04L 12/24** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **09159875 .5**

96 Fecha de presentación : **11.05.2009**

97 Número de publicación de la solicitud: **2134033**

97 Fecha de publicación de la solicitud: **16.12.2009**

54 Título: **Mecanismo de tolerancia a fallos optimizado para red entre pares.**

30 Prioridad: **11.06.2008 FR 08 53867**

45 Fecha de publicación de la mención BOPI:  
**12.04.2011**

45 Fecha de la publicación del folleto de la patente:  
**12.04.2011**

73 Titular/es: **ALCATEL LUCENT**  
**54, rue la Boétie**  
**75008 Paris, FR**

72 Inventor/es: **Tombroff, Dimitri y**  
**De Rop, Pierre**

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 356 711 T3

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).

## DESCRIPCION

La presente invención concierne al despliegue de una red entre pares o punto a punto en un conjunto de equipos de procesamiento. Más en particular, concierne a la utilización de estas redes entre pares para aplicaciones de telecomunicaciones.

5 Se conoce la distribución de determinadas aplicaciones de telecomunicaciones, en particular en un conjunto de nodos de procesamiento. De esta manera, cada nodo no procesa más que una parte de las peticiones dirigidas a la aplicación y es posible dimensionar, de manera ocasionalmente dinámica, el número de nodos en función de los recursos necesarios para procesar esas peticiones. Semejante arquitectura permite asimismo con facilidad hacer la aplicación tolerante a fallos, debido a la redundancia inherente a los nodos entre sí.

10 No obstante, para determinadas aplicaciones se plantea el problema de la localización de información contextual en el seno de este conjunto de nodos. En efecto, ciertas aplicaciones como un elemento de señalización pueden necesitar la memorización de un contexto entre dos mensajes que a él son dirigidos. Este contexto puede permitir procesar un mensaje siguiente de manera apropiada. Tal es el caso de un elemento de señalización SIP o «proxy SIP», que tiene que procesar una petición en función del estado de la sesión SIP (por «Session Initiation Protocol»).

15 Para este tipo de aplicaciones, es necesario, por una parte, guardar esa información contextual, pero también poder localizarla en el momento oportuno.

Un posible acercamiento consiste en disponer de una base de datos centralizada a la que tiene acceso cada uno de los nodos y que puede guardar y recuperar la información contextual de las sesiones que estos gestionan.

20 Sin embargo, cuando esta base de datos centralizada se implementa en un disco duro o en un soporte físico equivalente, el tiempo de acceso a la información adquiere un carácter penalizador y hace esta solución inadecuada para las aplicaciones que demandan muy cortos tiempos de respuesta, como las aplicaciones de telecomunicaciones.

25 Implementar la base de datos en memoria permite eludir este problema del tiempo de acceso pero la necesidad de incorporar redundancia para repartir la carga de los accesos y para satisfacer las restricciones de tolerancia a fallos torna complejo el sistema. Se trata de hecho de constituir una red de bases distinta del conjunto de nodos de procesamiento. Tal acercamiento, además de que no parece intelectualmente muy satisfactorio, origina problemas de configuración y no es ni flexible ni sencillo de gestionar.

30 Las redes entre pares permiten, mediante la utilización de una tabla hash distribuida, resolver estos problemas de una manera automática y transparente para el desarrollador de la aplicación y para los dispositivos que, externos a la aplicación, tengan que comunicarse con ella.

La figura 1 esquematiza una red entre pares o punto a punto compuesta por N nodos X1, X2, X3, X4... XN.

35 La información contextual que ha de memorizarse es asociada con claves que se distribuyen entre este conjunto de nodos. Generalmente, esta distribución se efectúa mediante una función hash que permite proyectar el espacio de las claves hacia el espacio de los nodos, obteniendo al propio tiempo un buen reparto de la carga entre esos nodos.

A partir de la clave, se puede dar con la información contextual. La aplicación de la misma función hash permite determinar el nodo asociado a esa clave y recuperar seguidamente la información almacenada en ese nodo y asociada a la clave.

40 Con objeto de hacer el sistema tolerante a fallos, generalmente está previsto que cada asociación entre una clave e información contextual se halle replicada en un segundo nodo. De esta manera, si el primer nodo deja de funcionar, la información contextual puede ser recuperada desde ese segundo nodo.

45 Un algoritmo simple y comúnmente utilizado para determinar qué nodo tiene que guardar la copia de la información contextual consiste en elegir el sucesor, es decir, el nodo siguiente en el orden de la red entre pares. El interés de tal elección es que, en caso de avería de un nodo, la localización de la copia de la información contextual es inmediata. Se evita así un período de incertidumbre a consecuencia de una avería, que precisa de un procedimiento suplementario para gestionar correctamente este período sin incurrir en el riesgo de responder de manera errónea a una petición.

50 Así, si la función hash aplicada a la clave buscada designa al nodo X2 y este último ya no es accesible o ya no contiene la información a consecuencia de un mal funcionamiento, el sistema determina automáticamente que en el nodo X3 se encuentra una copia de la información contextual buscada.

Estas implementaciones de una red entre pares están explicadas con más detalle en el artículo «Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications» de Ian Stoica, Robert Morris, David Karger, M. Frans Kaashoek y Hari Balakrishnan, ACM SIGCOMM 2001, San Diego, CA, Agosto de 2001, pp. 149-160, así como en la

solicitud de patente estadounidense US2007/230482.

5 Con todo, en la práctica, el problema técnico de la tolerancia a fallos sigue planteándose. En efecto, las redes entre pares se despliegan en la práctica en redes de equipos de procesamiento. Al ser estos equipos cada vez más potentes, es ventajoso desplegar varios nodos en un mismo equipo. En el ejemplo de la figura 1, los nodos X2 y X3 se hallan situados en el equipo M2, los nodos X1 y XN se hallan situados en el equipo M1 y el nodo X4 se halla situado en el equipo M3.

En consecuencia, si un equipo de procesamiento experimenta un mal funcionamiento, los nodos presentes pueden verse afectados y dejar de funcionar. Así, si el equipo de procesamiento M2 sufre una avería, los dos nodos X2 y X3 dejan de funcionar.

10 Ahora bien, según el algoritmo clásico consistente en localizar la copia de la información en el nodo sucesor, la información guardada en el nodo X2 está duplicada en el nodo X3. En consecuencia, pese al mecanismo de tolerancia a averías del estado de la técnica, la información contextual asociada al nodo ya no está disponible.

15 Por lo tanto, el estado de la técnica es inadecuado para procurar una suficiente tolerancia a averías. La finalidad de la presente invención es subsanar esta carencia mejorando la gestión de una red entre pares o punto a punto.

Para tal fin, la invención tiene por principal objeto un dispositivo de gestión de una red entre pares compuesta por un conjunto de nodos repartidos entre un conjunto de equipos de procesamiento y organizada en una forma circular, de modo que cada nodo posee un único nodo sucesor. El dispositivo de gestión incluye medios de admisión (o componente de admisión) para insertar un nuevo nodo en el seno de la red entre pares.

20 Éste se caracteriza porque los medios de admisión están previstos además para determinar un par de nodos adyacentes desplegados en un mismo equipo de procesamiento y para insertar ese nuevo nodo entre los dos nodos del par así determinado.

25 Los medios de admisión pueden ser, por ejemplo, un componente de admisión puesto en práctica por un dispositivo distinto de dichos nodos, o un componente de admisión implementado por la totalidad o parte de los nodos de la red entre pares.

30 La invención tiene asimismo por objeto un elemento de comunicación que incluye al menos una interfaz con una red de comunicaciones que contiene otros elementos de comunicación, y una red entre pares para el procesamiento de peticiones provenientes de esos otros elementos de comunicación. La red entre pares está constituida por nodos repartidos entre un conjunto de equipos de procesamiento y organizados en una forma circular, de modo que cada nodo posee un único nodo sucesor. Esta red entre pares está asociada además con un dispositivo de gestión que incluye medios de admisión para insertar un nuevo nodo en el seno de la red entre pares.

35 El elemento de comunicación según la invención está previsto para que se guarde en los nodos información contextual utilizada para el procesamiento de las peticiones y se caracteriza porque los medios de admisión están además previstos para determinar un par de nodos adyacentes desplegados en un mismo equipo de procesamiento y para insertar ese nuevo nodo entre los dos nodos del par de que se trate.

La información contextual puede hallarse duplicada en los nodos sucesores.

Los medios de admisión pueden ser, por ejemplo, un componente de admisión puesto en práctica por un dispositivo distinto de dichos nodos, o un componente de admisión implementado por la totalidad o parte de los nodos de la red.

40 La interfaz con una red de comunicaciones puede ser apta para transmitir mensajes de señalización acordes con el protocolo SIP.

La invención se pondrá más claramente de manifiesto a partir de la descripción que sigue con referencia a las figuras que se acompañan.

45 La figura 1 anteriormente comentada, esquematiza una red entre pares o punto a punto, conforme al estado de la técnica.

La figura 2 esquematiza una red entre pares o punto a punto desplegada en un conjunto de equipos de procesamiento, de acuerdo con la invención.

Las figuras 3a y 3b ilustran el reparto de la información guardada en la red entre pares o punto a punto.

50 La figura 4 ilustra un elemento de comunicación según la invención.

Se describirá en primer lugar la red entre pares y el dispositivo de gestión y, seguidamente, su aplicación en el seno de un elemento de comunicación.

5 La figura 2 representa una red entre pares desplegada en una red constituida por 3 equipos de procesamiento M1, M2, M3. La red entre pares está constituida por un conjunto de nodos N1, N2, N3 y N4 organizados en una forma circular, de modo que cada nodo posee un único nodo sucesor (y un único nodo antecesor).

Esta red entre pares lleva agregado un dispositivo de gestión que incluye en particular unos medios de admisión. Estos medios de admisión son los responsables de la inserción de un nuevo nodo en el seno de la red entre pares. Una propiedad esencial de las redes entre pares es efectivamente que son evolutivas, en el sentido de que se pueden añadir o retirar nodos en todo momento y de que la red tiene que estar prevista para autoorganizarse.

10 Si se asume que, en un instante, la red está constituida por los nodos N1, N2, N3, la admisión del nuevo nodo N4 comprende la determinación de la ubicación del nodo entrante en el seno de la red ya constituida, así como la determinación de un nuevo reparto de las claves en el conjunto de los nodos. Este reparto permite en particular repartir la carga obteniendo provecho de la aportación de recursos de procesamiento y de almacenamiento de los nuevos nodos.

15 Las figuras 3a y 3b ilustran más claramente la determinación de un nuevo reparto de las claves.

Cada línea horizontal representa el espacio de las claves. Sobre estas líneas se colocan los nodos presentes en la red. Cada uno de estos nodos va asociado a un subconjunto de claves, correspondiente a las claves que guarda.

20 La figura 3a corresponde a una primera situación en la que la red se compone de los nodos N1, N2 y N3. Cada uno de los nodos guarda un subconjunto correspondiente a la tercera parte del conjunto total de las claves. Este subconjunto se corresponde, en la figura 3a, con la parte de la línea horizontal situada en su izquierda y acotada por el nodo precedente o por el inicio de la línea.

Cuando ha de añadirse a la red entre pares el nodo N4, los medios de admisión determinan la ubicación que éste tiene que adoptar en el seno de la red, así como el reparto de las claves.

25 La localización del despliegue de un nodo en un equipo de procesamiento M1, M2, M3, M4 pueden desencadenarla los medios de admisión, pero de la determinación de esta localización clásicamente se hacen cargo otros módulos del sistema de gestión. Por ejemplo, se puede prever que esta localización sea determinada por el sistema operativo subyacente, sin que los medios de admisión tengan medios para controlarla.

30 Los medios de admisión son un componente de admisión que puede ser puesto en práctica por un dispositivo distinto de los nodos de la red entre pares. Pero también puede ser implementado por la totalidad o parte de estos nodos. Así, un nuevo nodo entrante puede dirigirse bien a cualquier nodo miembro ya de la red, o bien a algunos de esos nodos que desempeñan una función particular de admisión.

De acuerdo con la invención, el componente de admisión está previsto para [determinar un par de nodos adyacentes desplegados en un mismo equipo] determinar un par de nodos adyacentes desplegados en un mismo equipo de procesamiento.

35 Dos nodos son adyacentes si uno es el sucesor del otro en el seno de la red entre pares. Por ejemplo, en el caso de la red conformada por los nodos. Si se asume que, en un momento dado, la red está constituida por los nodos N1, N2, N3, los nodos N1 y N2 son adyacentes, los nodos N2 y N3 son adyacentes y los nodos N1 y N3 son adyacentes. Los nodos N2 y N3 forman el único par de nodos adyacentes desplegados en un mismo equipo de procesamiento, M2.

40 El componente de admisión está previsto para insertar entonces el nuevo nodo N4 entre los dos nodos de este par N2, N3.

45 La nueva red entre pares forma un círculo que contiene los nodos ordenados N1, N2, N4 y N3. Así, el par de nodos adyacentes N2, N3 se rompe y ya no quedan pares de nodos adyacentes desplegados en un mismo equipo de procesamiento. De manera más general, la admisión de un nuevo nodo permite disminuir en una unidad el número de estos pares.

En caso de que no existieran pares de este tipo, la composición de admisión puede funcionar según el estado de la técnica y, por ejemplo, insertar el nuevo nodo entre el último nodo y el primer nodo del círculo conformado por la red.

50 En la figura 3b figura el nodo N4, añadido entre los nodos N2 y N3. La red entre pares se autoconfigura de modo que el subconjunto de claves anteriormente asociado al nodo N3 queda dividido en dos partes, asociándose la parte más a la izquierda al nuevo nodo N4 y manteniéndose asociada al nodo N3 la parte más a la derecha. Estas dos partes son sensiblemente de idéntico tamaño.

La parte de claves recién asociada al nodo N4 es transferida desde el nodo N3 hacia este nuevo nodo. Con estas claves se transfiere asimismo la información que llevan asociada. Esta información puede ser información contextual en el caso de una aplicación en un elemento de comunicación, como se describirá posteriormente.

5 Para asegurar la tolerancia a fallos de la red entre pares, la información (contextual) se duplica en el nodo sucesor. El nodo sucesor es el nodo siguiente en el orden circular de la red N1, N2, N4, N3.

En virtud de la manera en que se insertan los nuevos nodos en la red entre pares, el sucesor de cada nodo no se halla situado en el mismo equipo de procesamiento que este último.

Así, el sucesor del nodo N2 es en adelante el nodo N4 que está desplegado en el equipo de procesamiento M3.

10 Si el equipo de procesamiento M2 experimenta un mal funcionamiento, el nodo N4, situado en el equipo M3, puede facilitar la copia de la información del nodo N2 y el nodo N1, situado en el equipo M1, puede facilitar la copia de la información del nodo N3.

Queda por tanto resuelto el problema planteado por los nodos N2 y N3 en el estado de la red que antecede a la admisión del nodo N4.

15 La figura 4 ilustra la aplicación de una red entre pares según la invención en un elemento de comunicación. Este último puede ser un servidor de aplicaciones en una arquitectura de comunicaciones, tal como una arquitectura IMS («IP Multimedia Subsystem»). También puede ser un elemento de señalización, tal como un proxy acorde con el protocolo SIP («Session Initiation Protocol») según lo especificado por el RFC 3261 del IETF, o una función CSCF («Call Session Control Function») en el seno de una arquitectura IMS.

20 Este elemento de comunicación E<sub>2</sub> posee medios de recepción de una petición Req proveniente de otro elemento E<sub>1</sub> a través de una red de comunicaciones N<sub>TEL</sub>. La petición Req es procesada por un módulo de reparto de carga LB, destinado a determinar, para cada primera petición de una sesión de comunicación, qué nodo N1, N2, N3, N4, N5 tiene que procesar la petición. Estos nodos conforman una red entre pares N<sub>p2p</sub> tal como se ha descrito anteriormente.

25 El procesamiento de la petición puede producir la generación de información contextual que puede ser útil o necesaria para el procesamiento de otra petición perteneciente a la misma sesión. Se la asocia, pues, a una clave y se guarda en el nodo correspondiente a esa clave.

30 La asociación se efectúa mediante una función hash aplicada a un identificador de la sesión (por ejemplo, direcciones IP del elemento de comunicación emisor E<sub>1</sub>, cabecera «call ID» de la petición SIP, etc.). El resultado de esta función hash puede dar directamente el número del nodo módulo el número de nodos en la red. Estos distintos mecanismos son conocidos en el estado de la técnica y no se describen en detalle en esta solicitud de patente. Para obtener algunos de estos detalles, se puede consultar el artículo anteriormente mencionado y referente al mecanismo «Chord».

35 En el ejemplo de la figura 4, la información contextual C se guarda en el nodo N<sub>2</sub>. Ésta se duplica en su sucesor, el nodo N<sub>3</sub>. Como por la presente se ha visto anteriormente, el módulo de admisión del dispositivo de gestión de la red entre pares N<sub>p2p</sub> ha hecho que los nodos adyacentes N<sub>2</sub> y N<sub>3</sub> no se desplieguen en un mismo equipo de procesamiento.

De esta manera, las próximas peticiones pertenecientes a la misma sesión podrán ser procesadas recuperando la información [con]textual presente en el nodo N<sub>2</sub> y, en caso de avería de este último, en el sucesor, el nodo N<sub>3</sub>.

40 Al no hallarse los dos N<sub>2</sub> y N<sub>3</sub> en el mismo equipo de procesamiento, es muy escasa la probabilidad de que estos sean defectuosos simultáneamente.

La invención permite por tanto resolver el problema técnico, sin cambiar básicamente los mecanismos conocidos de gestión de la red entre pares.

## REIVINDICACIONES

- 5 1. Dispositivo de gestión de una red entre pares o punto a punto compuesta por un conjunto de nodos (N1, N2, N3, N4) repartidos entre un conjunto de equipos de procesamiento (M1, M2, M3) y organizada en una forma circular, de modo que cada nodo posee un único nodo sucesor, incluyendo dicho dispositivo de gestión unos medios de admisión para insertar un nuevo nodo en el seno de dicha red entre pares y estando **caracterizado porque** dichos medios de admisión están previstos además para determinar un par de nodos adyacentes desplegados en un mismo equipo de procesamiento y para insertar dicho nuevo nodo entre los dos nodos de dicho par.
- 10 2. Dispositivo de gestión según la reivindicación precedente, en el que dichos medios de admisión son un componente de admisión puesto en práctica por un dispositivo distinto de dichos nodos.
- 10 3. Dispositivo de gestión según la reivindicación precedente, en el que dichos medios de admisión son un componente de admisión implementado por la totalidad o parte de dichos nodos.
- 15 4. Elemento de comunicación (E<sub>2</sub>) que incluye al menos una interfaz con una red de comunicaciones (N<sub>TEL</sub>) que contiene otros elementos de comunicación (E<sub>1</sub>), y una red entre pares (N<sub>p2p</sub>) para el procesamiento de peticiones (Req) provenientes de dichos otros elementos de comunicación, estando constituida dicha red entre pares por nodos (N<sub>1</sub>, N<sub>2</sub>, N<sub>3</sub>, N<sub>4</sub>, N<sub>5</sub>) repartidos entre un conjunto de equipos de procesamiento y organizados en una forma circular, de modo que cada nodo posee un único nodo sucesor e incluyendo un dispositivo de gestión que comprende medios de admisión para insertar un nuevo nodo en el seno de dicha red entre pares, cuyo elemento de comunicación está previsto para que se guarde en dichos nodos información contextual utilizada para el procesamiento de dichas peticiones y **se caracteriza porque** dichos medios de admisión están además previstos para determinar un par de nodos adyacentes desplegados en un mismo equipo de procesamiento y para insertar dicho nuevo nodo entre los dos nodos de dicho par.
- 20 5. Elemento de comunicación según la reivindicación precedente, en el que dicha información contextual está duplicada en los nodos sucesores.
6. Elemento de comunicación según una de las reivindicaciones 4 ó 5, en el que dichos medios de admisión son un componente de admisión puesto en práctica por un dispositivo distinto de dichos nodos.
- 25 7. Elemento de comunicación según una de las reivindicaciones 4 ó 5, en el que dichos medios de admisión son un componente de admisión implementado por la totalidad o parte de dichos nodos.
8. Elemento de comunicación según una de las reivindicaciones 4 a 7, en el que dicha interfaz con una red de comunicaciones es apta para transmitir mensajes de señalización acordes con el protocolo SIP.

FIG. 1

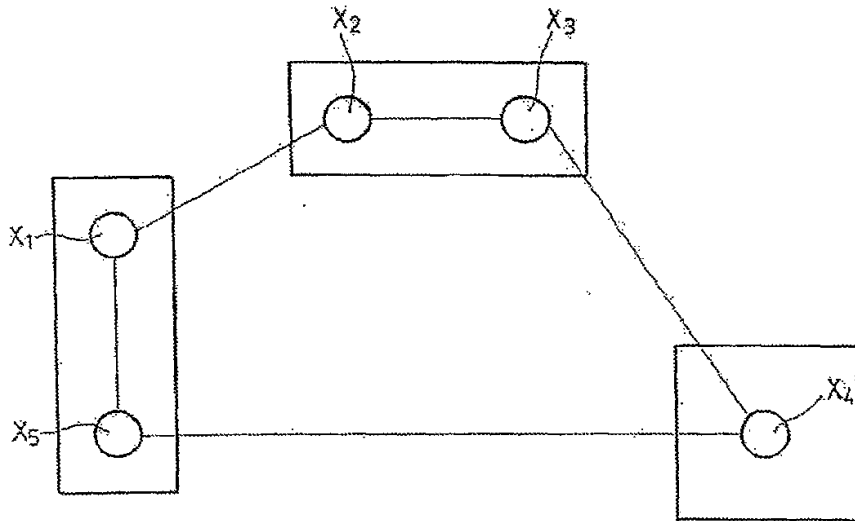
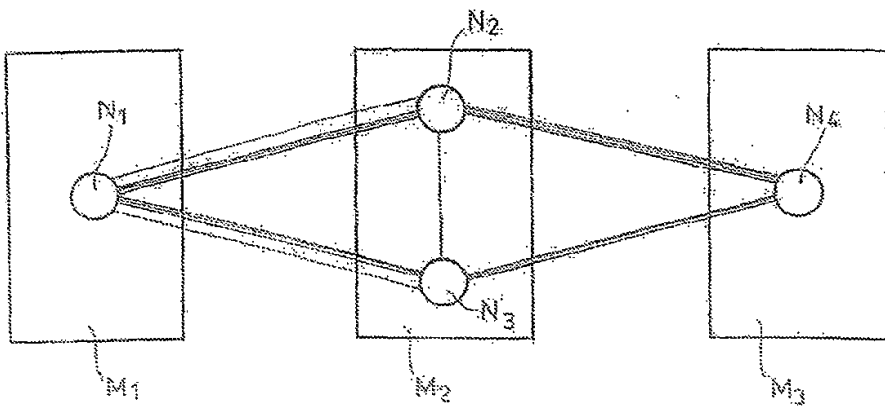
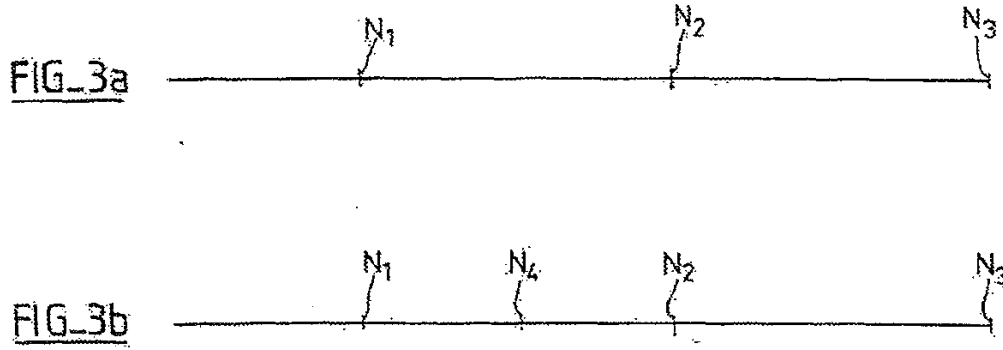


FIG. 2





FIG\_4

