

Proyecto final, Aplicaciones para Comunicaciones en Red.

Aplicación distribuida para búsqueda de archivos

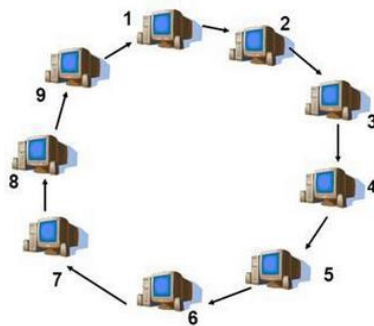
Objetivo:

El estudiante, implementará una aplicación en java para la búsqueda de archivos en una topología de anillo mediante uso de llamadas a procedimientos remotos, sockets de flujo y multicast, con la ayuda de hilos usando una arquitectura P2P.

Instrucciones:

Se deberá de implementar una aplicación con arquitectura P2P para la búsqueda y descarga de archivos.

Se creará una red a nivel de capa de aplicación con una topología de anillo, donde cada nodo tendrá interacción únicamente con su nodo sucesor y predecesor.



La aplicación deberá permitir las búsquedas sucesivas donde, cada nodo solo le pregunta al nodo siguiente y responde únicamente al nodo anterior.

Si algún nodo tiene el archivo solicitado, continuara con la búsqueda hasta cubrir toda la topología y esperara la respuesta. En caso de que más de un nodo tenga el archivo solicitado, se le avisara a su nodo predecesor hasta que la información llegue al nodo que hizo la pregunta original informando sobre todos los nodos que tengan el recurso.

Una vez obtenido esto, el nodo que solicito el recurso se encargara de recuperarlo desde cada uno de los nodos que lo contienen por partes, es decir, si el archivo se encontró en 3 nodos, le pedirá una tercera parte del archivo a cada nodo al mismo tiempo para recuperar el archivo original.

Para asegurar que se trata del mismo archivo, deberá de coincidir tanto el nombre como el MD5.

Funcionamiento

La aplicación, en el momento de activarse, solicitará un número de puerto, el cual aparecerá en el nombre de la ventana donde se encuentre la interfaz gráfica.

Cada aplicación levantada tendrá relacionada una carpeta donde se contendrán los archivos que maneja el nodo, con el mismo nombre que el número de puerto que se le paso.

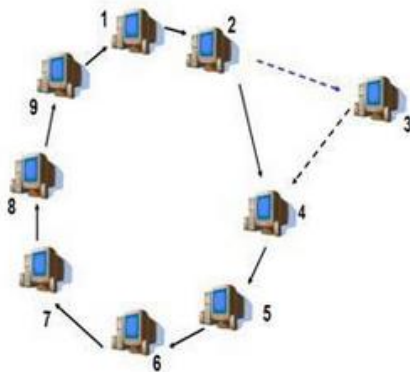
Posteriormente, se desplegará la interfaz gráfica que, después de un tiempo no mayor a 30 segundos detectara los distintos nodos que estén corriendo, y que deberán aparecer en la lista de nodos activos.

Al realizarse la búsqueda de un archivo inexistente, mandará un mensaje de archivo no encontrado, en caso contrario se descargará en la carpeta del nodo que lo solicito.

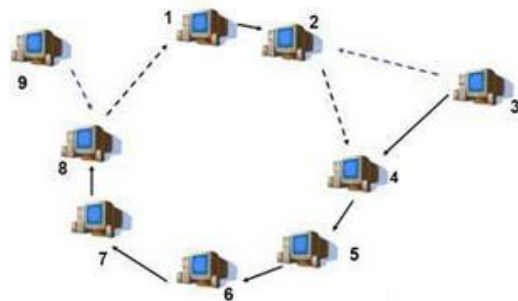
Si el archivo se encuentra en varios nodos, recuperara de forma simultánea una parte equitativa de cada uno de los nodos donde esté, de forma directa, es decir, en este caso no se trabajara con la topología de anillo.

También implementara un servicio que permita el ingreso y abandono de la red para nodos. Es decir, cada nodo que este corriendo deberá de ser capaz de ajustar la topología al nuevo número de nodos que estén ejecutándose.

Ingreso y abandono de la Red:



Simultaneo...



Desarrollo del proyecto

La aplicación para búsqueda y descarga de archivos consta de la implementación de los siguientes hilos:

- Un servidor multicast para el anuncio del servicio de búsqueda de archivos.
- Un cliente multicast para generar la lista de servidores disponibles.
- Un servidor RMI para la búsqueda de archivos.
- Un cliente RMI que invoque el servicio de búsqueda en el nodo siguiente de la topología.
- Un servidor de sockets de flujo para el envío de archivos, o parte de ellos.
- Un cliente de sockets de flujo para la descarga de archivos divididos.
- Una interfaz gráfica para integrar todos los elementos necesarios.

A continuación, se describirá la funcionalidad que deberá implementar para cada una de ellas.

Servidor Multicast: Este servidor básicamente se encargará de anunciar cada 5 segundos el puerto de servicio del servidor de flujo. La dirección de grupo multicas que usará será la “228.1.1.1”, en el puerto 8000.

Cliente Multicast: El cliente deberá unirse a la dirección de grupo “228.1.1.1” y leer todos los datagramas que lleguen. De ellos extraerá la dirección IP del anunciante, y el número de puerto del servidor de flujo del servidor anunciante (note que es diferente al número de puerto del cliente multicast). El nodo definirá un identificador único (ID), usando la combinación de su dirección IP y el número de puerto de su servidor de datagramas.

Una vez conocidos todos los identificadores, entablará sus búsquedas solamente hacia el identificador de valor mayor inmediato superior al suyo y solo responderá búsquedas del inmediato inferior. Considerando esto de forma anular, es decir, si se trata del identificador mayor, sus búsquedas se harán hacia el identificador más pequeño y viceversa.

Se debe de implementar un método para agregar nodos del anillo, así como en el caso de que algún nodo se deshabilite se ajuste el anillo al nuevo grupo de servidores

Cada entrada de la lista de servidores activos deberá tener la siguiente información:

- IP del servidor.
- Puerto de servicio del servidor de datagramas.
- Temporizador decreciente

El temporizador de cada entrada será inicializado con un valor de 16 segundos y será decrementado cada segundo. Si el contador de alguna de las entradas de la lista llega a 0 sin que antes haya llegado un datagrama anunciando dicho servidor, esta entrada será eliminada de la lista. Si llega un datagrama anunciando la dirección de una entrada ya existente en la lista, el contador de dicha entrada será reiniciado a 16 segundos nuevamente.

Servidor RMI: Una vez hecha la solicitud del cliente, deberá buscar en una carpeta específica el nombre del archivo que se solicita. Aun si encuentra el archivo en su carpeta o no, le pedirá al cliente RMI que siga con la búsqueda con el siguiente nodo hasta completar la búsqueda en todo el anillo.

En caso de que localice el archivo solicitado en uno o varios nodos, avisará siguiente la regla de la topología de anillo al nodo desde donde se realizó la búsqueda incluyendo la información de la o las direcciones IP y número de puerto del servidor de sockets de flujo donde puede solicitarlo, así como el MD5 de cada archivo.

Cliente RMI: Se implementará un servicio de búsqueda de archivos sucesiva donde cada nodo solo le pregunta al nodo siguiente hasta cubrir toda la topología, enviará como parámetro la cadena del archivo solicitado y esperará como respuesta una lista serializada con la dirección o direcciones IP y el número de puerto de los servidores de sockets de flujo en donde puede pedir el archivo, así como el MD5 de cada uno de ellos. En caso de que el archivo no se localice en ningún nodo, regresará una lista vacía para que se informe al usuario que el archivo no existe.

En caso de que el archivo esté en su propia carpeta, el nodo que solicita la búsqueda lo anunciará y no realizará la búsqueda.

El nodo donde se realizó la búsqueda, una vez obtenida la lista de nodos realizará primeramente una validación del MD5 y verificará que sean iguales, en caso de que no sea así, se le preguntará al usuario cual prefiere bajar indicando el número de veces que se repite el MD5 para cada caso.

Una vez definido la o las fuentes del archivo, usando el cliente de llamadas a procedimiento remoto se descargará una parte del archivo en todas las fuentes encontradas de forma simultánea, es decir, si el archivo a descargar se encuentra en 2 nodos, el cliente de llamadas a procedimiento remoto bajará el 50% del archivo de un nodo y el otro 50% del otro al mismo tiempo, en caso de que sean 3 nodos se bajará un tercio del archivo de cada uno y así sucesivamente.

Servidor de sockets de flujo: Este servidor únicamente se encargará de recibir peticiones de descarga y enviar los archivos o parte de ellos, y se levantará en un puerto igual al número de puerto solicitado al arrancar el nodo.

Cliente de sockets de flujo: El cliente se encargará de solicitar al servidor de flujo el archivo a ser descargado, ya sea en su totalidad o el fragmento de archivo que desea descargar.

Interfaz de usuario: Desde aquí integrará todos los servicios de una aplicación de búsqueda y descarga de archivos. El usuario tendrá la siguiente información en pantalla:

- Lista de servidores activos, donde se pueda ver la dirección IP, el número de puerto y el tiempo que le queda antes de eliminarlo de la topología.

- Campo de mensajes donde se anunciarán los siguientes eventos:

- El nodo anterior y siguiente, cada vez que cambie.
- Cuando el nodo anterior le pregunta sobre un archivo, incluyendo su nombre, y si lo localizo o no en su carpeta, incluyendo el identificador del nodo.
- Cuando pregunta al nodo siguiente en caso de que no lo localice en su carpeta y la respuesta que recibe, incluyendo el identificador del nodo.
- Cuando cargue o descargue un archivo con su nombre y el progreso de la descarga para cada uno de los nodos donde se localizó, ya sea como cliente o como servidor.
- El valor inicial y final de bites enviados cuando envíe un archivo completo o la parte correspondiente.
- Cuando el archivo no sea localizado en ninguna carpeta.

Campos donde se muestren los identificadores de los nodos anterior y siguiente de la topología de anillo.

Caja de texto para escribir el nombre del archivo a buscar, con botón para realizar la búsqueda.

Orden de la revisión

Para la revisión del funcionamiento de la aplicación, los mensajes indicados en el campo de mensajes deberán de aparecer como se indica, en caso contrario no se continuará con la revisión del proyecto.

1. Primeramente, se solicitará que se levanten entre 3 y 6 nodos y se le asignaran números de puerto no consecutivos ni ordenados, los cuales deberán de colocarse en topología de anillo, revisado con los campos correspondientes en la interfaz gráfica, tanto de los nodos siguiente y anterior como en el campo de mensajes.
2. Se solicitará que se levanten entre 3 y 6 nodos más asignando números de puerto no ordenados, y la topología debe adaptarse a la nueva configuración. Esto se verificará con los campos correspondientes mencionados en el párrafo anterior.
3. Se eliminará entre 1 y 3 nodos y la topología deberá de adaptarse de la misma forma que se a indicado previamente. En caso de que no ocurra no se podrá seguir con la revisión.
4. Se pedirá que se vuelvan a crear de 1 a 3 nodos y deberá de mantener la topología de anillo.
5. Se buscará un archivo con el mismo MD5 o equivalente que existe en más de una carpeta de nodo, los archivos a buscar serán del tipo pdf o zip.
6. Se buscará un archivo que no existe en ninguna carpeta.

7. Se revisará en el código que los clientes y servidores implementados correspondan a lo solicitados en el proyecto, en caso de no ser así, no tendrá ninguno de los puntos indicados en la rúbrica.

Rúbrica

Condición	Calificación		
	No funciona	Funciono parcialmente	Funciono correctamente
La aplicación solicita el número de puerto inicial y aparece en el nombre de la ventana que se despliega.	0	1	2 puntos
La venta despliega todos los objetos que se solicitan.	0	1	2 puntos
Se ejecutan de 3 a 6 nodos			
Los nodos aparecen con su ID en la lista de nodos activos, sin duplicarse y cada uno con un número de puerto diferente y el tiempo restante de vida.	0	1	3 puntos
Cada ventana despliega el ID del nodo precedente y subsecuente correctamente.	0	1	3 puntos
El nodo con ID más grande tiene como nodo subsecuente el ID más pequeño y viceversa.	0	1	3 puntos
Se ejecutan entre 3 y 6 nodos más			
El resto de los nodos activos se ajustan a las nuevas condiciones de nodos	0	2	10 puntos
Se eliminará entre 1 y 3 nodos al azar, y se esperaran 20 segundos.			
El resto de los nodos activos se ajustan a las nuevas condiciones de nodos	0	2	5 puntos
Se crean entre 1 y 3 nodos nuevos			
La topología se adapta a esta nueva condición.	0	2	5 puntos
Desde un nodo al azar, se solicita un archivo que se encuentre en 2 o 3 nodos diferentes (con el mismo MD5 o equivalente).			
Los nodos intermedios dejan rastro de que la búsqueda paso por su nodo.	0	1	2 puntos
La búsqueda cubre el nodo completo	0	2	4 puntos
Los nodos intermedios dejan rastro de que la respuesta paso por su nodo.	0	1	2 puntos
Se recupera el archivo simultáneamente de los nodos donde se encuentra	0	7	20 puntos

El archivo se recupera de forma correcta.	0	0	5 puntos
Desde un nodo al azar, se solicitará un archivo que se encuentre en 2 o 3 nodos diferentes y en otro nodo un archivo diferente se le cambiará el nombre al solicitado (MD5 diferente).			
La aplicación detecta la diferencia y pregunta al usuario sobre cual archivo bajar.	0	7	15 puntos
Una vez seleccionado, el archivo se baja únicamente de los nodos donde se indicó.	0	5	10 puntos
Desde un nodo al azar se solicita un archivo inexistente.			
Todos los nodos muestran evidencia del paso de la respuesta negativa por su nodo.	0	1	2 puntos
El nodo que realizó la búsqueda anuncia que el archivo no se encontró.	0	1	2 puntos
El puerto del servidor de llamadas a procedimientos remotos es el número de puerto solicitado al inicio más 100	0	0	5 puntos
TOTAL			100 puntos