

#### INSTITUTO POLITÉCNICO NACIONAL ESCUELA SUPERIOR DE CÓMPUTO

ESCOM

**Desarrollo de Sistemas Distribuidos** 

# Tarea 3

# Multiplicación de matrices distribuida utilizando paso de mensajes

**PROFESOR:** Pineda Guerrero Carlos

**EQUIPO 12** 

• Ramírez Galindo Karina

• Toledo Espinosa Cristina Aline

• Vázquez Hernández Alan Mauricio

GRUPO: 4CV11

Fecha de realización: 10 - marzo - 2022

Fecha de entrega: 10 – marzo - 2022

## Contenido

Introducción	
Matrices	
Multiplicación de matrices en forma distribuida	1
Desarrollo	3
Funcionamiento	7
Creación de máquinas virtuales	7
Configuración de las máquinas virtuales	16
Compilación y ejecución del programa	22
	23
Conclusiones	24
Ramírez Galindo Karina	24
Toledo Espinosa Cristina Aline	24
Vázquez Hernández Alan Mauricio	24
Referencias	25

#### Introducción

#### **Matrices**

Una matriz es un arreglo rectangular de números en renglones y columnas. A cada número en una matriz se le conoce como un elemento de matriz o entrada. [1]

Por ejemplo, la matriz A tiene 2 renglones y 3 columnas. El elemento  $a_{2,1}$  es la entrada en el  $2^{do}$  renglón y la  $1^{ra}$  columna de la matriz A, es decir, 5, como se muestra en la *Figura 1*.

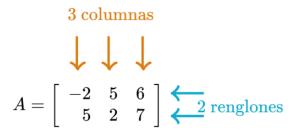


Figura 1 Ejemplo de una matriz

#### Multiplicación de matrices en forma distribuida

Suponga que divide la matriz A en las matrices  $A_1$  y  $A_2$ . El tamaño de las matrices  $A_1$  y  $A_2$  es N/2 renglones y N columnas.

La matriz B en las matrices  $B_1$  y  $B_2$ . El tamaño de las matrices  $B_1$  y  $B_2$  es N renglones y N/2 columnas.

Entonces la matriz  $C = A \times B$  se compone de las matrices  $C_1$ ,  $C_2$ ,  $C_3$  y  $C_4$ , tal como se observa en la *Figura 2*:

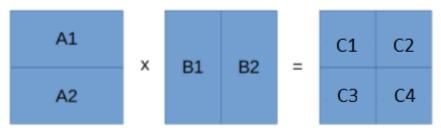


Figura 2 Representación de la multiplicación de matrices

Donde:

$$C_1 = A_1 \times B_1$$

$$C_2 = A_1 \times B_2$$

$$C_3 = A_2 \times B_1$$

$$C_4 = A_2 \times B_2$$

Debido a que las matrices se guardan en memoria por renglones, es más eficiente transponer la matriz B y dividirla como se muestra en la *Figura 3*:

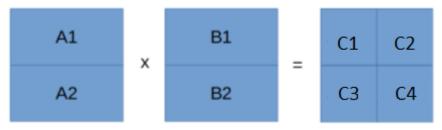


Figura 3 Representación de la multiplicación de matrices

En este caso el producto se realizará renglón por renglón, entonces:

$$C_1 = A_1 \times B_1$$
  
 $C_2 = A_1 \times B_2$   
 $C_3 = A_2 \times B_1$   
 $C_4 = A_2 \times B_2$ 

Ahora supongamos que tenemos cuatro nodos identificados con los números 0, 1, 2 y 3, el nodo 0 actuará como cliente y los nodos 1, 2 y 3 como servidores (véase *Figura 4*). [2]

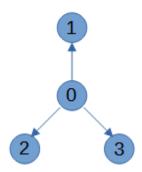


Figura 4 Topología ejemplo

#### Desarrollo

En esta tarea se desarrolla un solo programa en Java, el cual calcula el producto de dos matrices cuadradas en forma distribuida sobre cuatro nodos.

- Sean A, B y C matrices cuadradas con elementos de tipo double, N renglones y N columnas, N par y  $C = A \times B$ .
- Se ejecutan dos casos:
  - ✓ N=8, desplegar las matrices A, B y C y el checksum de la matriz C.
  - √ N=1000, desplegar el checksum de la matriz C.
- El checksum de la matriz C se calcula como la suma de todos elementos de la matriz. El checksum debe ser de tipo double.

$$checksum \ = \ \sum C[i][j] \ , i = 0, \ldots, N-1, j = 0, \ldots, N-1.$$

Se inicializan las matrices de la siguiente manera:

$$A[i][j] = i + 5 * j$$
  
 $B[i][j] = 5 * i - j$ 

Donde A[i][j] y B[i][j] son los elementos Ai,j y Bi,j respectivamente.

El programa se ejecuta en cuatro máquinas virtuales con Ubuntu (1 CPU, 1GB de RAM y disco HDD estándar), en cada máquina virtual se pasa como parámetro al programa el número de nodo a saber: 0, 1, 2 y 3.

El nombre de cada máquina virtual cumple con lo siguiente:

```
"Tarea-3-12-NumeroDeNodo"
```

Para multiplicar las matrices A y B se realizó lo siguiente:

#### Nodo 0 (cliente):

1. Se inicializan las matrices A y B como se menciona anteriormente (véase *Figura* 5).

```
for (int <u>i</u> = 0; <u>i</u> < size; <u>i</u>++)
    for (int j = 0; j < size; j++) {
        a[<u>i</u>][j] = <u>i</u> + 5 * j;
        b[<u>i</u>][j] = 5 * <u>i</u> - j;
}
```

Figura 5 inicialización de las matrices A y B

 Para facilitar la obtención de la multiplicación de matrices A y B se crea un método para transponer la matriz B y además de simplificar la operación permite también la partición de las matrices a como se desea (véase *Figuras* 2 y 3).

```
public static void transpose(double[][] matrix, int size) {
    double temp;
    for(int i = 0; i < size; i++)
        for(int j = 0; j < i; j++) {
            temp = matrix[i][j];
            matrix[i][j] = matrix[j][i];
            matrix[j][i] = temp;
    }
}</pre>
```

Figura 6 Método transpose()

Como se puede observar en la *Figura* 6, la matriz se transpone y queda guardada en la misma variable.

3. Se solicita enviar al nodo 1 la matriz  $A_1$  y  $B_1$ . al nodo 2 la matriz  $A_1$  y  $B_2$ , finalmente al nodo 3 la matriz  $A_2$  y  $B_1$ , para ello se transpone primero B y después se dividen ambas como se muestra en la *Figura 7* de la siguiente manera:

```
for(int i = 0; i < a1.length;i++) {
    a1[i] = a[i];
    a2[i] = a[i + size/2];
    b1[i] = b[i];
    b2[i] = b[i + size/2];
}</pre>
```

Figura 7 Partición de las matrices

Una vez divididas se crean los hilos con sus respectivos pedazos de matriz para ser enviados a los nodos en el orden previamente mencionado (véase *Figura* 8).

```
serverTasks.add(createTask( node: 1,a1,b1));
serverTasks.add(createTask( node: 2,a1,b2));
serverTasks.add(createTask( node: 3,a2,b1));
```

Figura 8 Creación de las tareas para los nodos 1-3

4. Una vez que los nodos reciben el fragmento de matriz con el que deberán trabajar se hace simplemente una multiplicación renglón a renglón gracias al cambio que sufrió previamente B. (Véase Figura 9)

```
 \begin{array}{lll} & \text{public static double[][] multiply(double[][] a, double[][] b) } \{ \\ & \text{double[][] c = new double[a.length][b.length];} \\ & \text{for (int $\underline{i} = 0$; $\underline{i} < c.length; $\underline{i}$++)} \\ & & \text{for (int $\underline{i} = 0$; $\underline{i} < c[\underline{i}].length; $\underline{i}$++)} \\ & & \text{for (int $\underline{k} = 0$; $\underline{k} < a[0].length; $\underline{k}$++)} \\ & & & c[\underline{i}][\underline{i}] += a[\underline{i}][\underline{k}] * b[\underline{j}][\underline{k}]; \\ & \text{return c;} \} \\ \end{array}
```

Figura 9 Método multiply()

 Una vez que se reparten las tareas a los nodos, al terminar dichos nodos regresan un fragmento de la matriz resultante C que deberá unir el nodo 0. (véase Figura 10)

```
public static void fillMatrix(int node, double[][] matrix, double[][] result) {
   int rowShift = node > 2 ? size/2: 0; //row shift only occurs on node 2 & 0
   int columnShift = (node + 1) % 2 * size/2; //column shift only occurs on node 3 & 0
   for(int i = 0; i < result.length; i++)
        for(int j = 0; j < result[i].length; j++) {
            matrix[i + rowShift][j + columnShift] = result[i][j];
        }
}</pre>
```

Figura 10 Método fillMatrix()

6. Finalmente se solicita el cálculo del checksum de la matriz resultante C una vez que está ya ha sido completada. (*Figura* 11)

Figura 11 Método getChecksum()

#### Nodo 1, 2 y 3 (servidores):

Una vez que los nodos reciben las matrices con las que deberán trabajar únicamente hacen la multiplicación fila por fila.

**NOTA:** Para establecer la conexión en las máquinas virtuales de Azure se implementa únicamente un array de strings con las 3 dirección es ip y los nodos se conectan uno a uno según su correspondiente ip a través del mismo puerto. (véase figura 12)

```
while(true) { //Waits until the connection with the server is established
    try {
        socket = new Socket( ip[node-1] , port 50001);
        System.out.println("Connection to the node " + node + " accepted");
        break;
    } catch (IOException e) {
            Thread.sleep( millis: 100);
    }
}
```

Figura 12 Algoritmo para crear reintentos de conexión

#### **Funcionamiento**

#### Creación de máquinas virtuales

Para esta práctica se hace uso de 4 máquinas virtuales del nodo 0 al 3, creadas en Azure, A continuación, se explica a detalle la creación de la máquina virtual para el nodo 0.

#### Creación paso a paso:

Ingresar al portal de Azure en la siguiente URL:

https://azure.microsoft.com/es-mx/features/azure-portal/

- 1. Dar click al botón "Iniciar sesión".
- 2. En el portal de Azure seleccionar "Máquinas virtuales".
- 3. En la figura 13 se observa que aún no se tiene una máquina virtual, se creara una nueva dando click en el botón "+Crear".

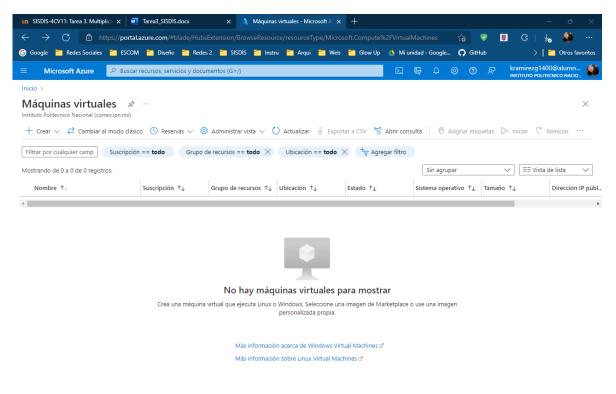


Figura 13 Creación de una máquina virtual para el nodo 0

- 4. Seleccionar la opción de "Máguina virtual"
- 5. Se crea un grupo de recursos, en este caso llamado "Tarea3Equipo12"
- 6. El nombre de la máquina virtual para el nodo 0 es "Tarea-3-12-0"

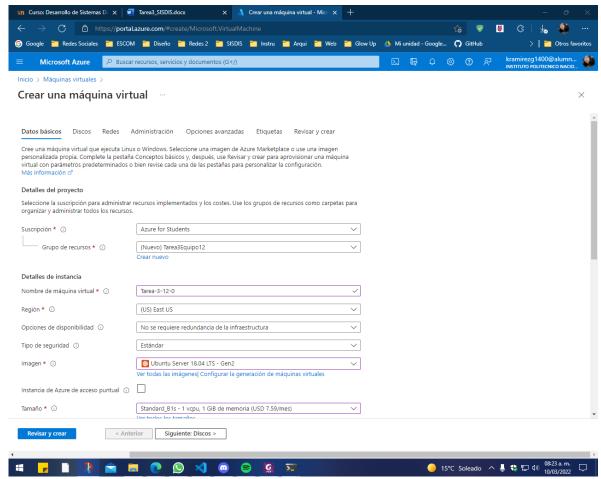


Figura 14 Creación de una máquina virtual para el nodo 0

- 7. Seleccionar la región donde se creará la máguina virtual
- 8. Seleccionar la imagen, en este caso se elige Ubuntu Server 18.04 LTS.
- Dar click en "Seleccionar tamaño" de la máquina virtual, en este caso se selecciona una máquina virtual con 1 GB de memoria RAM. Dar click en el botón "Seleccionar".
- 10. En tipo de autenticación seleccionamos "Contraseña".
- 11. Ingresamos el nombre del usuario, en este caso: "Equipo12"
- 12. Ingresamos la contraseña y confirmamos la contraseña. La contraseña debe tener al menos 12 caracteres, debe al menos una letra minúscula, una letra mayúscula, un dígito y un carácter especial.
- 13. En las "Reglas de puerto de entrada" se deberá dejar abierto el puerto 22 para utilizar SSH (la terminal de secure shell). (véase la Figura 15)

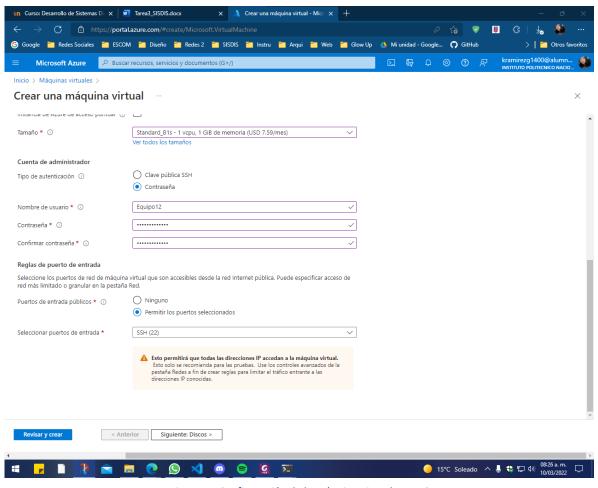


Figura 15 Configuración de la máquina virtual parte 2

- 14. Dar click en el botón "Siguiente: Discos>"
- 15. Seleccionar el tipo de disco de sistema operativo, en este caso se ocupa HDD estándar. (véase Figura 16)
- 16. Dar click en el botón "Siguiente: Redes>" (véase la Figura 17)
- 17. Dar click en el botón "Siguiente: Administración>"
- 18. En el campo "Diagnóstico de arranque" seleccionar "Desactivado".
- 19. Dar click en el botón "Revisar y crear". (véase la Figura 18)
- 20. Dar click en el botón "Crear". (véase la Figura 19)
- 21. Dar click a la campana de notificaciones (barra superior de la pantalla) para verificar que la máquina virtual se haya creado. (véase la figura 20)

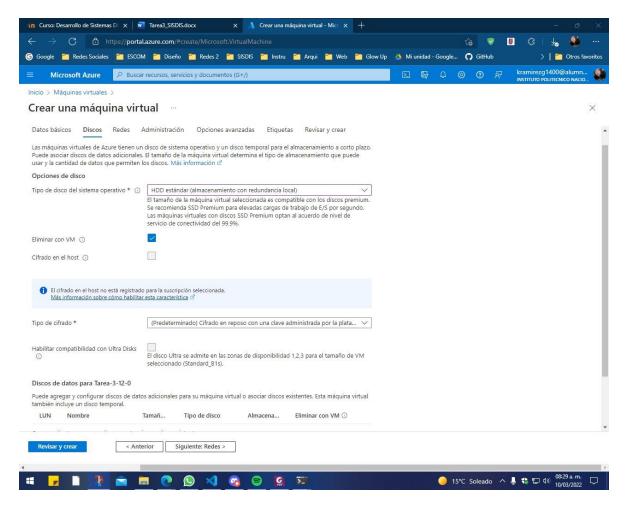


Figura 16 Configuración de la máquina virtual parte 3

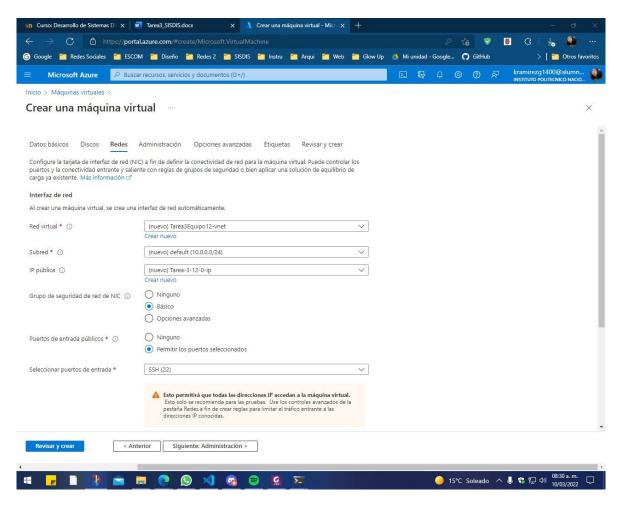


Figura 17 Configuración de la máquina virtual (Redes)

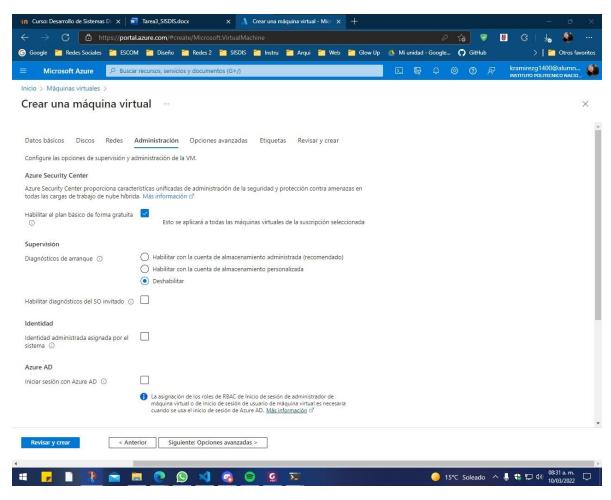


Figura 18 Configuración de la máquina virtual (Administración)

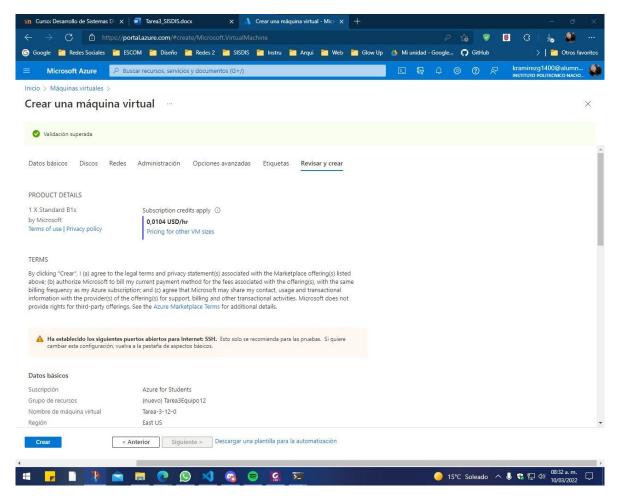


Figura 19 "Crear" la máquina virtual

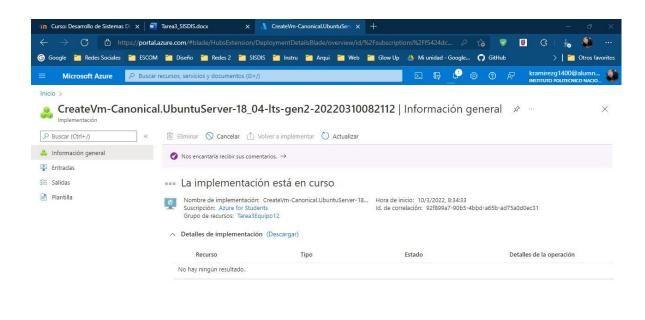




Figura 20 Máquina virtual para el nodo 0 creada

22. Dar click en el botón "Ir al recurso". En la página de puede ver la dirección IP pública de la máquina virtual. Esta dirección puede cambiar cada vez que se apague y se encienda la máquina virtual. (véase la figura 21)

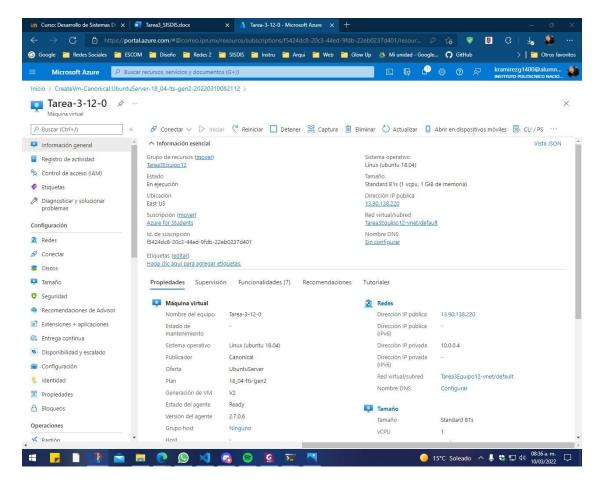


Figura 21 Datos de la máquina virtual del nodo 0

#### Configuración de las máquinas virtuales

- Para conectarse a la máquina virtual se utiliza el programa ssh disponible en Windows, Linux y MacOS.
- Mediante ssh y por medio de Windows con la siguiente línea se establece la conexión:

#### Equipo12@13.90.138.220

Y se introduce la contraseña de autenticación de Azure (véase la figura 22)

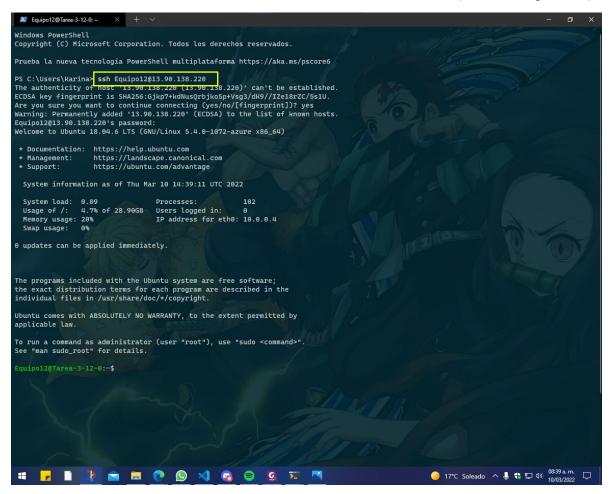


Figura 22 Acceso a la máquina virtual por SSH

Hay que comprobar que esté instalado java, de lo contrario, se debe a instalar (véase las figuras 23 y 24).

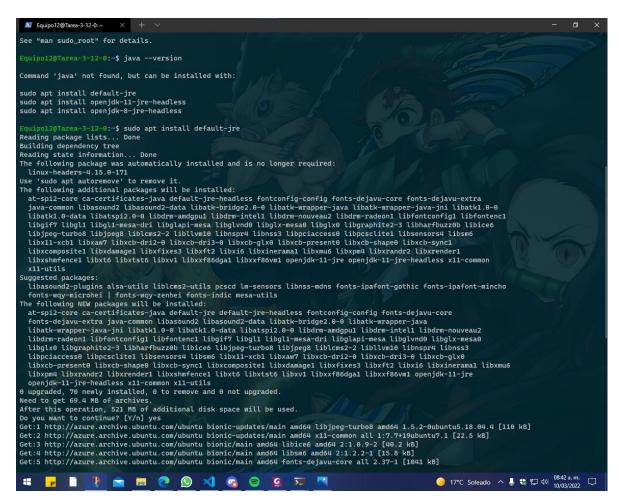


Figura 23 Instalación del jre

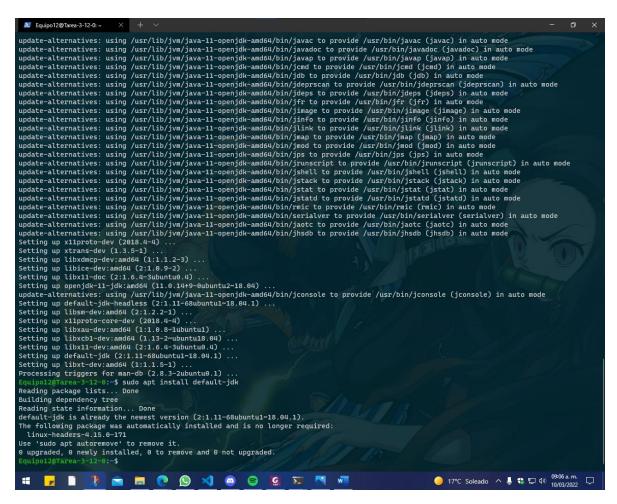
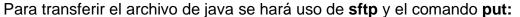


Figura 24 Instalación del jdk



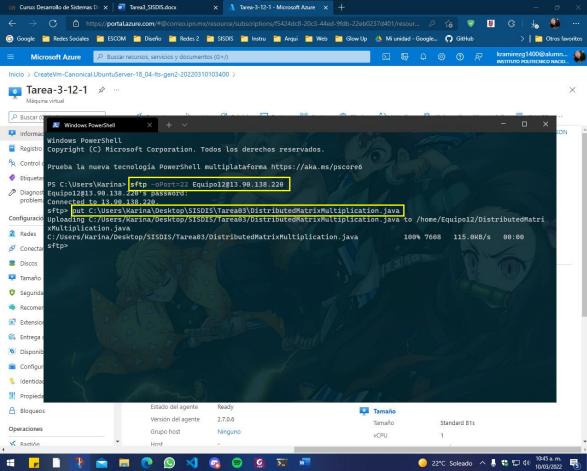


Figura 25 Transferencia del archivo .java a la máquina virtual

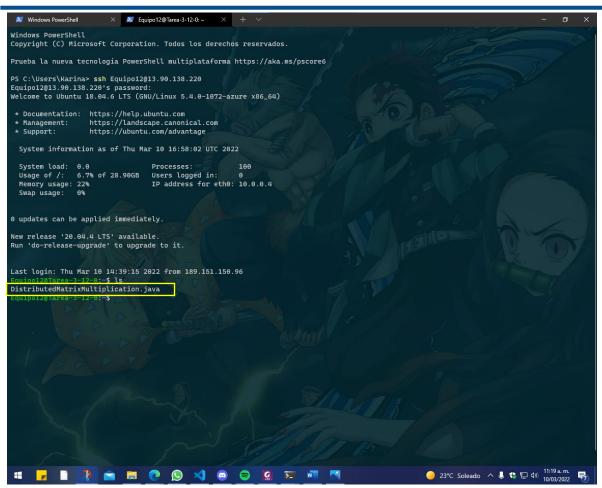


Figura 26 Comprobación de transferencia del archivo a la máquina virtual

Se agrega la regla de seguridad de entrada para permitir peticiones del protocolo TCP por el puerto 5000, esto se repite en las 4 máquinas virtuales (figura 27).

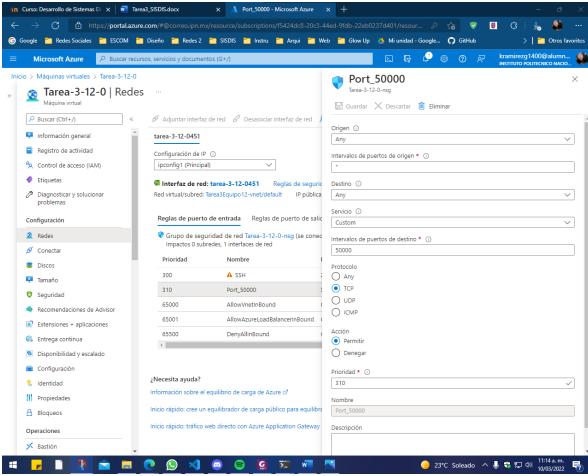


Figura 27 Regla de puerto de entrada

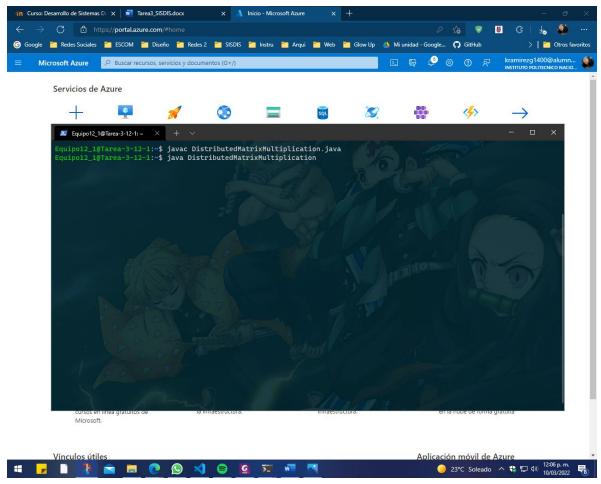


Figura 28 Ejecución del programa DistributedMatrixMultiplication.java

#### Compilación y ejecución del programa

En la *Figura 28* se observa la compilación del programa sin errores. Después se ejecuta el programa agregando como parámetros el número de nodo y N, para el desarrollo de esta práctica existen 2 casos:

- N=8
   Como se observa en la Figura 29, en el nodo 0 se despliegan las matrices A,
   B y C dando como resultado Checksum=217728.0
- N=1000

En la Figura 30, se observa el resultado Checksum=8.07133725E15

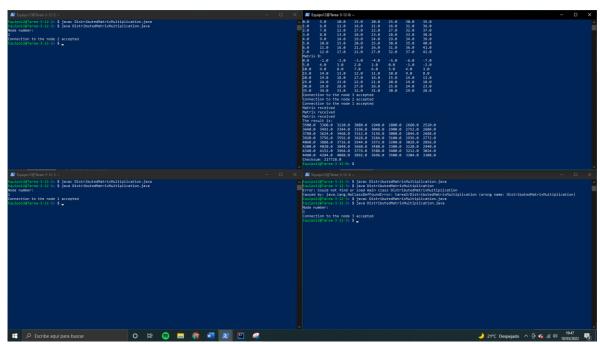


Figura 29 Ejecución del programa para N=8

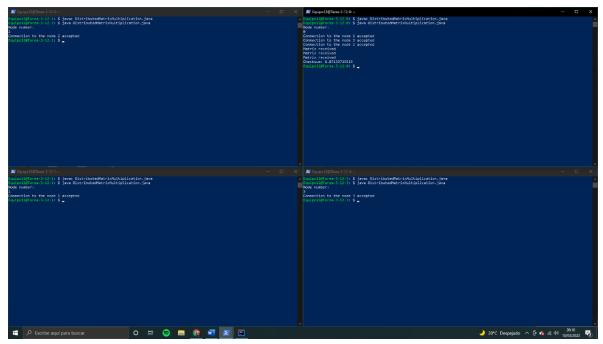


Figura 30 Ejecución del programa para N=100

#### **Conclusiones**

#### Ramírez Galindo Karina

Esta práctica fue de gran ayuda para tener un mejor acercamiento a un sistema distribuido, ya que no solo se hizo la ejecución de manera local como en los anteriores programas, esta vez se tuvo un acercamiento a computo en la nube al momento de hacer que el programa se ejecute en máquinas virtuales creadas a través de Azure. Es importante considerar los atributos de cada máquina, como el sistema operativo, memoria, entre otros, para que pueda haber conexión entre ellas al momento de ejecutar el programa y así aprovechar todas las bondades que ofrece Azure en cuestiones de escalabilidad, seguridad y demás características útiles para el desarrollo de nuevas aplicaciones.

Se usaron protocolos de transferencia de archivos, ente otros conceptos vistos en clase, lo que hizo una experiencia completa del uso de esta nueva tecnología.

#### Toledo Espinosa Cristina Aline

En esta práctica pude notar las bondades de contar con servidores (haciendo uso de las máquinas virtuales para dicho rol) en la nube y cuan fácil es levantar una máquina virtual con ayuda de Azure, esta herramienta nos permite ejecutar el programa desde diferentes maquinas, un nodo por máquina, lo que me permitió observar mejor el funcionamiento de la distribución de una tarea y cuán importante es que su diseño sea correcto para un optima ejecución. Un problema que tuve al querer copiar el programa a mis máquinas virtuales fue que no permitió utilizar los comandos sftp por lo que tuve que ayudarme de la herramienta "FileZilla". Otro problema fue la utilización de puertos para la conexión ya que para ejecutar el caso 1 con N=8 tuve que hacer uso de uno distinto a cuando N=1000.

#### Vázquez Hernández Alan Mauricio

Resolver un problema mediante el uso de cómputo distribuido requiere de dos tareas muy importantes: la primera es resolver el problema que se nos plantea, y la segunda es replantear el problema mediante el uso de cómputo distribuido, considerando los recursos que se utilizarán y necesitarán para que el sistema funcione de forma adecuada. Durante esta práctica, se comprendió que es importante tener claro cómo es que se distribuirá el cálculo de un programa para evitar resultados erróneos. Por otra parte, el establecimiento de las máquinas virtuales para probar el programa fue una de las actividades más importantes, ya que fue posible realizar un primer acercamiento con el cómputo en la nube, con lo que en posteriores ocasiones se podrá evitar una cantidad de errores considerables.

### Referencias

- [1] «Matrices,» [En línea]. Available: http://hyperphysics.phy-astr.gsu.edu/hbasees/Math/matrix.html.
- [2] P. G. Carlos, «Desarrollo de sistemas Distrubuidos 4CV11 Plataforma Educativa Moodle,» [En línea]. Available: https://m4gm.com/moodle.