

Ambiente de computación distribuido

Resumen

En este documento se describen los objetivos y las instrucciones para lograr el uso de herramientas de un ambiente de computación distribuida. Además, se presentan las condiciones de entrega del informe, contenido y formato del informe.

Tabla de contenido

| | |
|----------------------------------|----------|
| Objetivos | 1 |
| Descripción del ejercicio | 2 |
| Entregables | 4 |
| Número de integrantes | 4 |
| Fecha de entrega | 4 |

Objetivos

A continuación se enumeran algunos de los objetivos que el estudiante deberá alcanzar y posteriormente evidenciar a través del informe final y la sustentación del ejercicio.

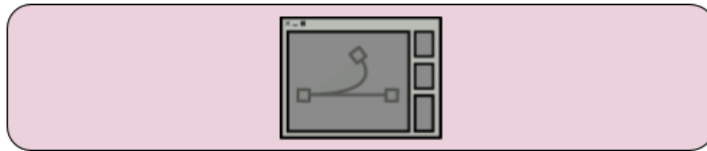
- El estudiante debe demostrar capacidad en la utilización de herramientas de virtualización, e.g. VirtualBox o KVM/QEMU.
- El estudiante debe ser capaz de instalar un sistema operativo tipo Unix/Linux en una máquina virtual.
- El estudiante estará en la capacidad de instalar y poner a punto [al menos] un servicio de red.
- El estudiante estará en la capacidad de hacer el despliegue de una aplicación que se ejecute sobre un ambiente de computación distribuida.

Descripción del ejercicio

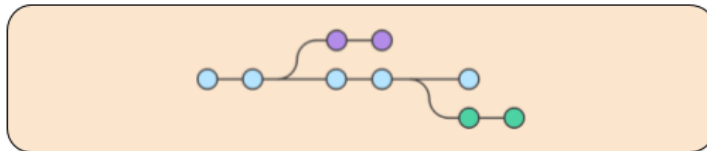
En este ejercicio, el/la estudiante llevará a cabo el despliegue de un ambiente computacional distribuido sobre el cual se ejecutará una aplicación de forma distribuida. Ejemplo de una aplicación distribuida es una que se encargue de renderizar una serie de imágenes y generar un video.

El ejercicio se puede pensar que tiene tres etapas:

- Despliegue del cluster computacional
- Despliegue del *framework* para el desarrollo de aplicaciones
- Desarrollo de la aplicación



Aplicación



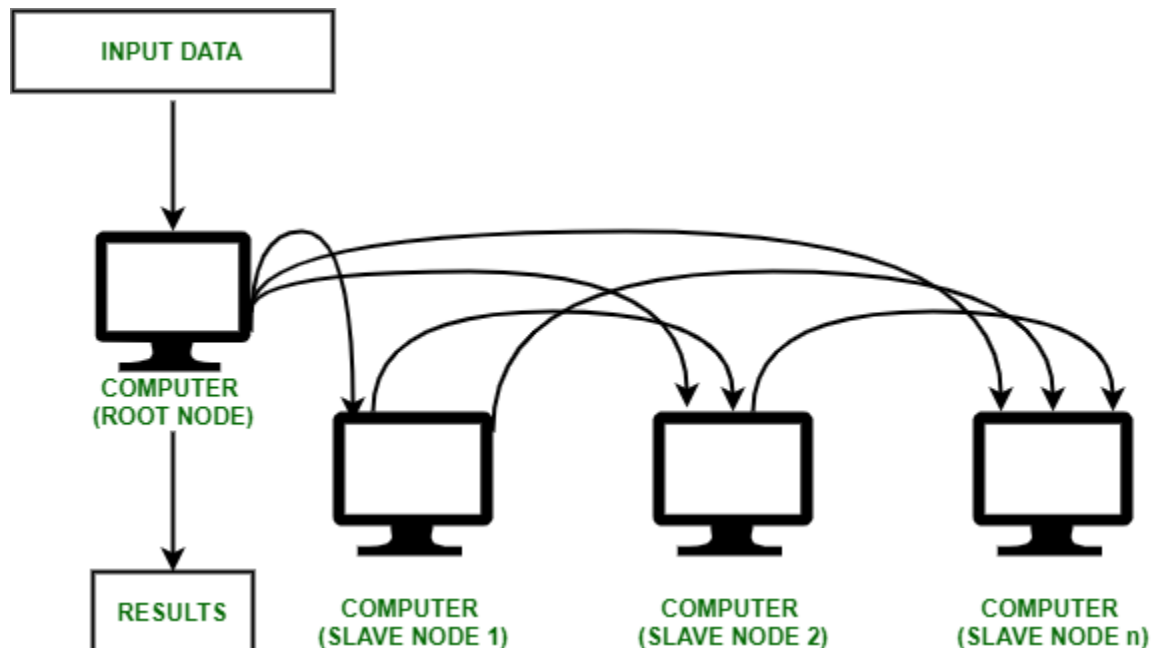
Framework para desarrollo y ejecución de aplicaciones distribuidas



Cluster computacional de máquinas virtuales

Descripción gráfica del ejercicio a implementar. Íconos tomados de: <https://openclipart.org/>

Para el cluster computacional el/la estudiante deberá tener al menos dos máquinas virtuales que formen parte del cluster. La/el estudiante deberá hacer el despliegue de una infraestructura computacional como la que se muestra a continuación.



Entorno de computación distribuida maestro/esclavo, tomado de:

<https://www.geeksforgeeks.org/an-overview-of-cluster-computing/>

Esta infraestructura se conoce como una infraestructura computacional **maestro/esclavo**. En ella se pueden observar varios nodos que cumplen dos roles, un rol maestro y un rol esclavo o trabajador. El rol maestro tiene como propósito recibir la tarea o aplicación que se va a ejecutar, los datos a procesar, distribuir el procesamiento de los datos entre los esclavos o trabajadores, coleccionar los datos de cada uno de los nodos esclavos o trabajadores y finalmente presentar los datos coleccionados.

El/la estudiante se debe encargar de hacer el despliegue (instalación y configuración) de los diferentes roles en máquinas virtuales.

Un punto que es importante es que estos nodos se conectarán a través de una red virtual provista por la herramienta de virtualización. En VirtualBox el tipo de red que usarán las máquinas virtuales es la conocida como [host-only adapter](#). En KVM se crea un interface de red tipo bridge y a través de esa interface de red es que se logra la conexión entre las diferentes máquinas virtuales que se ejecutan bajo un mismo equipo anfitrión¹.

La segunda etapa consiste en el despliegue de un *framework* que facilite el desarrollo de aplicaciones cuya ejecución se espera sea de manera paralela y/o distribuida. Para este ejercicio se usará la herramienta llamada [Ray](#) la cual permite el desarrollo de aplicaciones en Python que pueden ejecutarse de manera paralela o de forma distribuida. Ray es capaz de llevar a cabo la ejecución de programas escritos en Python de manera paralela sobre equipos de cómputo que tengan procesadores multi-núcleos. Así mismo, para llevar a cabo ejecuciones de manera distribuida puede hacerlo sobre infraestructuras de cómputo *on-premise*² o en entornos de nube pública (e.g. AWS, GCP, Azure, entre otros).

Para este ejercicio el/la estudiante llevará a cabo la ejecución de Ray sobre un ambiente *on-premise* sobre máquinas virtuales.

Finalmente, la/el estudiante deberá llevar a cabo la ejecución de una aplicación en Python bien sea desarrollada por el mismo o tomada de una tercera parte. Si la aplicación es desarrollada por el/la estudiante y dependiendo de la complejidad/utilidad de la misma se tendrán puntos a favor. Si la aplicación es tomada o basada de una tercera fuente se deben indicar en el informe de donde se tomó dicha aplicación.

Entregables

El informe del ejercicio se debe entregar en documento en PDF, **No se aceptarán archivos de otro tipo.**

El documento deberá indicar claramente el nombre de los integrantes y presentar una tabla de contenido. El documento deberá presentar un informe en el cual se describa como se llevaron y alcanzaron cada uno de los objetivos planteados en el presente documento. Así mismo es muy importante documentar los problemas, desafíos que se presentaron durante el desarrollo del

¹ <https://superuser.com/questions/549350/host-only-like-interface-in-kvm>

² *On-premise* hace referencia a recursos computacionales propios del usuario o la organización.

ejercicio y mostrar cómo estos desafíos se superaron. Indicar las fuentes de información en las que se apoyaron para llevar a cabo el desarrollo del ejercicio es parte importante del informe.

El documento deberá proveer **enlaces a los recursos** que sustentan el logro de los objetivos planteados por la tarea:

- Presentación en ~~claro~~ ^{Clase} o video de Youtube, no superior a 20 minutos, y que sirva como una demostración de que el ejercicio funciona. Es importante que se evidencie también la participación de todos los miembros del grupo en el video. Pueden dividirse la presentación del video en segmentos donde cada miembro del grupo haga su participación. El video debe salir sin errores, es decir, que debe llevarse una presentación altamente depurada y ensayada.
- Repositorio en GitHub donde se muestre la documentación técnica que presenta los insumos que se usaron para lograr los objetivos planteados por la tarea. Aquí van los códigos de la aplicación. También puede incluir algún tipo de script que le haya permitido automatizar una o varias partes del desarrollo del ejercicio.

Número de integrantes

Los ejercicios se deben llevar a cabo en grupo. El número máximo de integrantes es de tres (3).

Fecha de entrega

Según como se establezca en BrightSpace