

INFO273

Evaluación de Desempeño

Monitor de Recursos de Patagón (UACH)

Segunda entrega

30/06/2021

Felipe Ignacio Lagos Martínez

Sebastián Pacheco Cáceres

Contenido

Introducción.....	3
Explicación del Servidor y Monitor.....	4
Características del Servidor.....	4
Preguntas a investigar.....	5
Metodología.....	5
Primeros Resultados.....	5

Introducción

El Patagón es el nuevo supercomputador (y trabajando como servidor) adquirido por el Instituto de Informática de la *Universidad Austral de Chile*, del cual se encarga de utilizar, mantener y dirigir el profesor *Cristóbal Navarro*.

Para esta oportunidad nos hemos acercado a él con el propósito de hacer de este trabajo un poco más interesante. Su disposición fue inmediata y con mucho entusiasmo nos agendó una reunión para el próximo jueves 1 de Julio, por lo que, de antemano, esta 2da entrega puede estar sujeta a cambios.

Nuestro objetivo es *monitorear el Patagón*, en todos sus componentes en una primera instancia (ya sea *Porcentaje de Uso*, *Temperatura*, *Voltaje*, etc.) con el fin de determinar cuanta carga real genera *un usuario* que accede al servidor, lo que se denomina un **trabajo**, por lo que también necesitaremos datos del tráfico de usuarios. Todo lo anterior sería monitoreado desde el momento que tengamos acceso al servidor hasta lo más cercano posible de la fecha final de entrega.

Este mismo documento será modificado a medida que los plazos de entrega vayan cumpliéndose.

Explicación del Servidor y Monitor

El Patagón está basado en una estación NVIDIA DGX STATION A100, la cual tiene 3 *nodos* descritos a continuación:

Características del Servidor

- Nodo Master (o de conexión)
 - o 2x Intel Xeon Platinum 8260 (2.4Ghz, 24-cores, 35.75MB L3 Cache)
 - o 768GB RAM DDR4-3200Hz RDIMM
 - o 512 GB NVMe M.2
- Nodo Almacenamiento
 - o 2x Intel Xeon Silver 4210 (2.2Ghz, 10-cores, 14MB L3 cache)
 - o 192GB RAM DDR4-2666Hz
 - o 92TB Sata SSD RAID6
- Nodo de Cómputo
 - o 2x AMD EPYC 7742 CPU (2.6GHz, 64-cores, 256MB L3 cache)
 - o 8x Nvidia A100 GPUs. Cada GPU tiene:
 - 6,192 F32 Units
 - 3,456 FP64 Units
 - 422 Tensor Cores v3.0
 - 40 GB GPU Memory
 - 1,555 GB/s GPU Memory Bandwidth
 - 9.7 FP64 TFLOPS
 - 19.5 FP64 TFLOPS (tensor cores)
 - 19.5 FP32 TFLOPS
 - 156 TF32 TFLOPS (tensor cores)
- Red
 - o Mellanox Infiniband EDR Switch MSB7800-ES2F 100GB/s
 - o Cisco Gigabit Switch

Como se mencionó en la introducción, nuestra idea es monitorizar el servidor; ahora añadiendo los nodos anteriormente mencionados, podemos ampliar nuestro plan. La idea es monitorizar, en la medida de lo posible, *los 3 nodos*, con lo cual planeamos responder las siguientes preguntas (modificando la pregunta original y añadiendo más):

Preguntas a investigar

1. ¿Cuál es la carga promedio de un solo trabajo en los módulos CPU, GPU y Discos a lo largo de los 3 nodos del Patagón?
2. ¿Cuál es el Costo de Funcionamiento del Sistema Inactivo (sin trabajos)?
3. ¿Se pueden clasificar los trabajos existentes (diferenciar clases)?

Metodología

Para contestar las preguntas antes mencionadas, nuestra estrategia se puede separar en dos pasos:

- Rescatar datos directamente desde la [Página Oficial del Patagón](#)
- Ingresar al Patagón y, desde dentro, poder utilizar una serie de monitores, los cuales vimos en clases (por ejemplo: w).

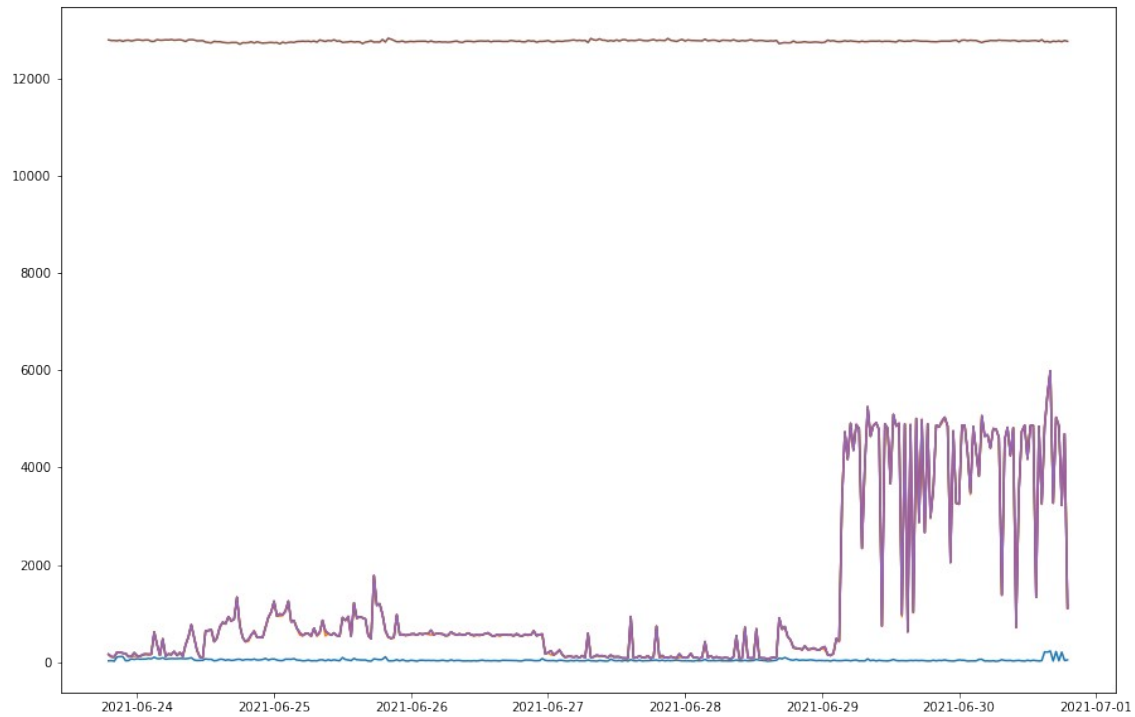
Esta metodología tiene también sus consideraciones, pues, por ejemplo: para poder responder a la pregunta 2 es necesario también tener en cuenta los datos de consumo eléctrico (Voltaje, Amperaje, Consumo en kW); datos que se encuentran disponibles en la página. Por otro lado, la ventaja de utilizar la web es que tenemos acceso a datos históricos, que en este caso abarcan el principio del funcionamiento “público” en abril del presente año.

Para poder trabajar con el Patagón primero debemos tener una reunión con el docente Cristóbal Navarro el jueves 1 de Julio, es decir, no se ha concretado antes de la fecha de entrega del presente informe.

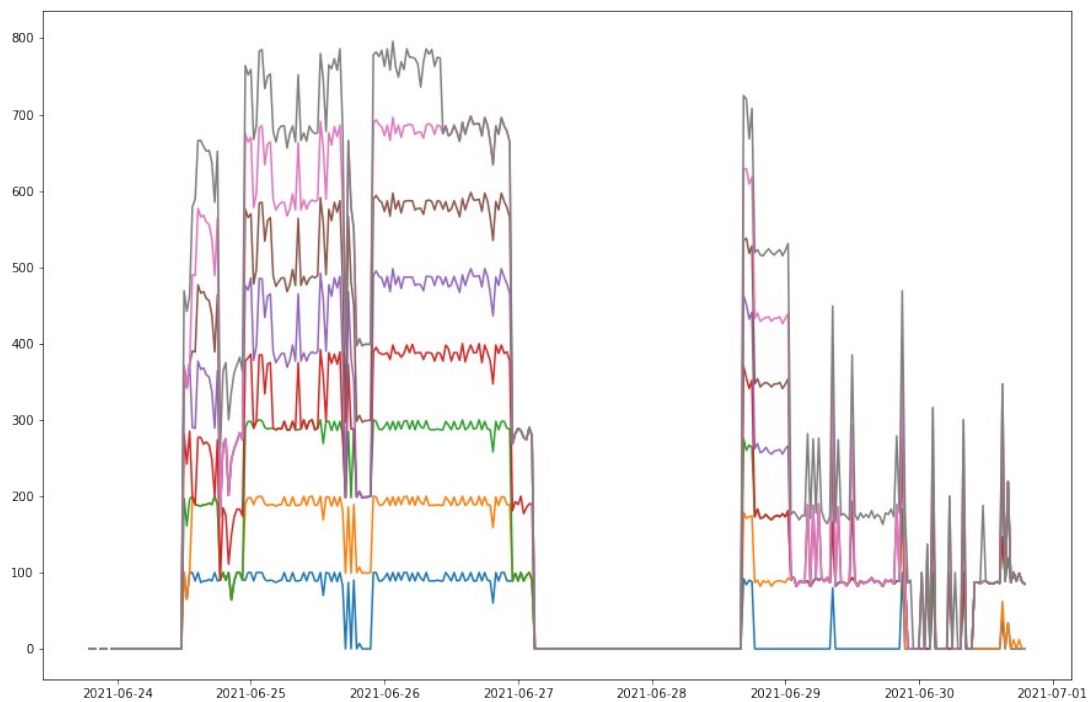
Primeros Resultados

De acuerdo con lo ya mencionado, los siguientes resultados se exponen utilizando los datos sacados de la web con un lapso de 7 días, y graficados en Python, con el uso de las librerías pandas y matplotlib.

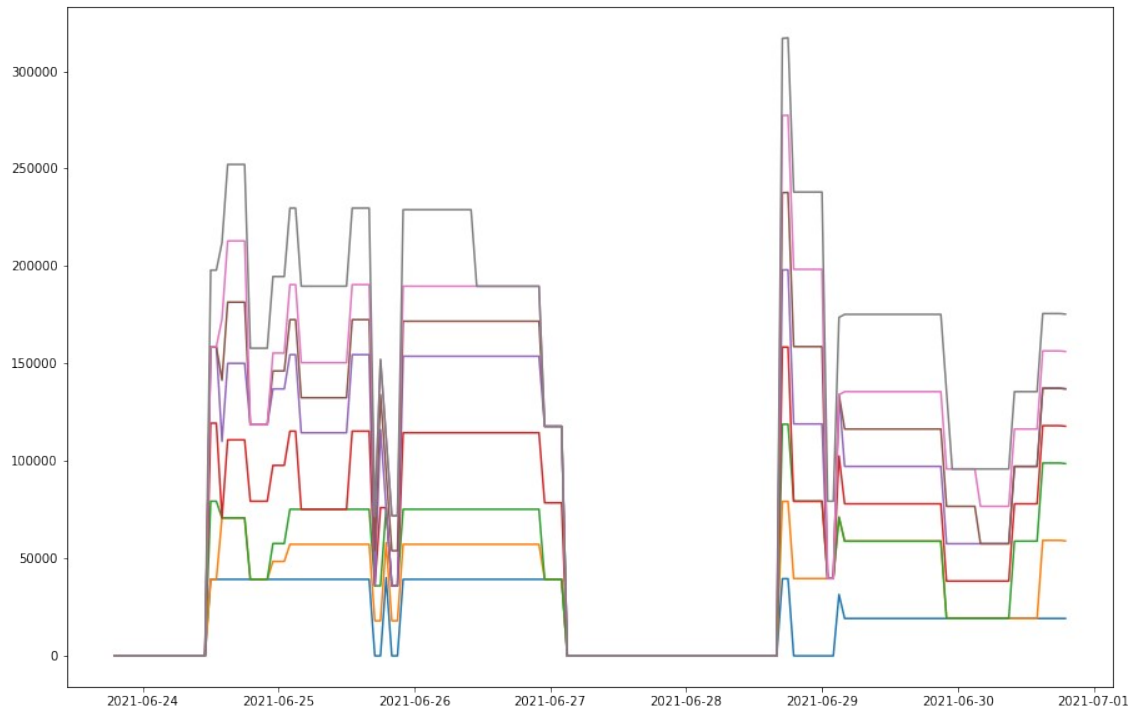
1. Utilización de CPU:



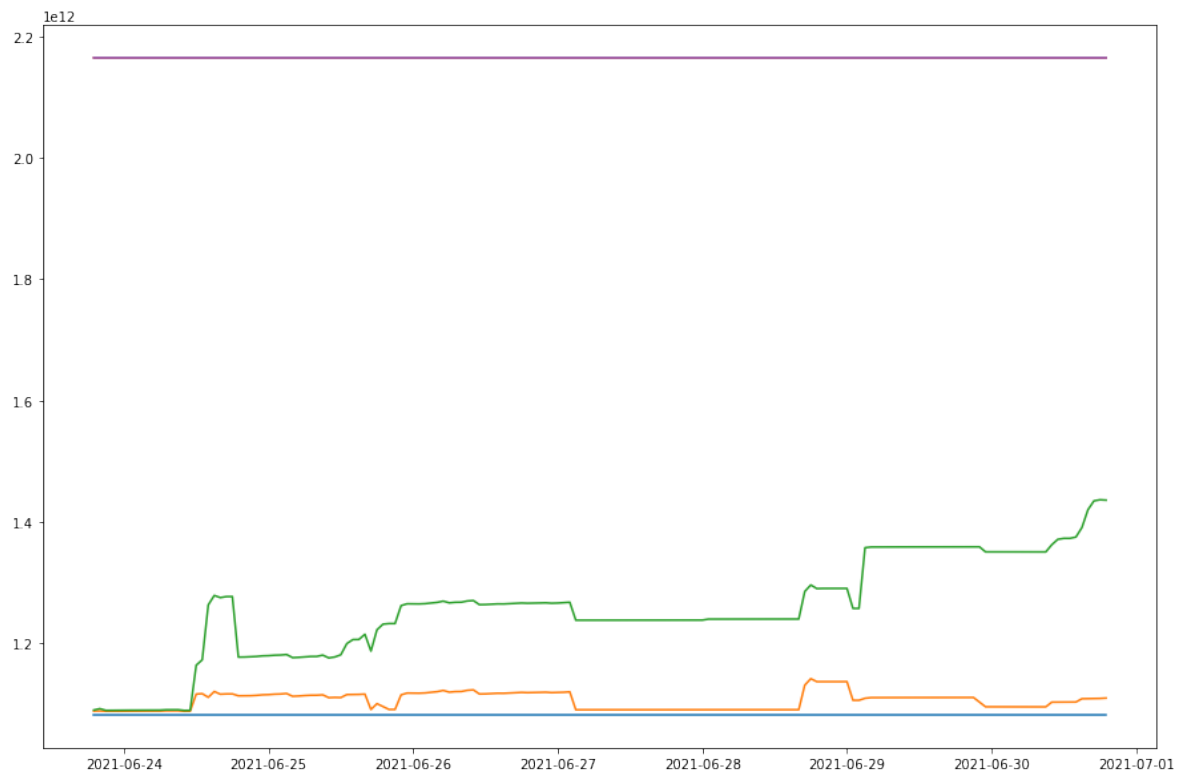
2. Utilización de GPU:



3. Uso Memoria GPU:



4. Uso de memoria RAM:



5. Uso de electricidad (kW):

