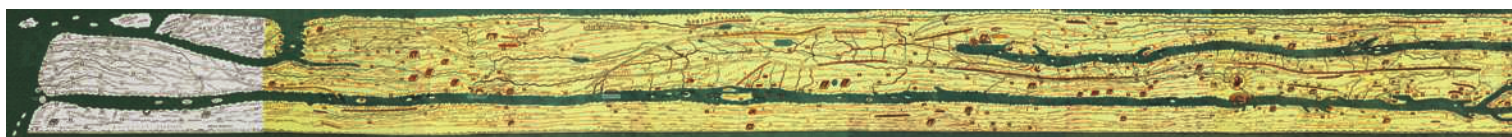


Tan Humano como la *Visualización*, tan Divino como la Memoria

El tema de la visualización de datos es muy amplio— demasiado, tal vez. Para tener claro qué ángulo podemos darle, pues, vamos a buscar una definición que se ajuste al ámbito en que nuestra revista pretende incidir.

¿Qué es la visualización?

Las definiciones para este tema que podemos encontrar en la red dan a entender una cierta desesperanza por parte de sus autores por aterrizar un concepto tan amplio, tan carente de límites. Por ejemplo, de la Wikipedia:

“Visualización es la generación de una imagen mental o una imagen real de algo abstracto o invisible.”

Me quedo con la impresión de que la definición presentada merece cierta visualización —para pasar de algo abstracto a una imagen mental o real. Cuando vamos restringiendo el ámbito de ese “algo” abstracto, la definición comienza a tener sentido. Por ejemplo, también de Wikipedia, el inicio de la definición de “visualización científica” dice:

“La visualización científica es la transformación de datos científicos y abstractos en imágenes.”

Pero seguimos en lo mismo, ya que siguiendo esta lógica podríamos factorizar el término “científico”, y volvemos al punto de partida. Podríamos decir que la visualización hierática es “la transformación de datos hieráticos y abstractos en imágenes”, y no por eso hemos de comprender mejor el sistema de escritura empleado en el antiguo Egipto, ¿o sí?

Puede ser que no vayamos por el camino correcto. Va de nuevo: ¿Y si olvidamos que esta es una revista sobre desarrollo de software? Podríamos entonces dedicarnos a discurrir por las interminables infografías que podemos encontrar en sitios como Medialab Prado (<http://visualizar.org>), <http://visualizar.info>, <http://DataVis.ca> y tantos más. A fin de cuentas (cito de Medialab Prado):

“La visualización de datos es una disciplina transversal que utiliza el inmenso poder de comunicación de las imágenes para explicar de manera comprensible las relaciones de significado, causa y dependencia que se pueden encontrar entre las grandes masas abstractas de información que generan los procesos científicos y sociales.”

El desarrollo histórico

Y sí, casi cualquier representación que podamos hacer de un proceso, o de un hecho en el tiempo, puede tomarse como una forma de visualización. Podríamos incluso mencionar que la visualización

es lo que hizo evolucionar a los protohumanos: ¿Qué no es la pintura rupestre una visualización, una forma de recordarle a los colegas el despliegue de valor que tuvieron en la cacería del otro día?

O una de las principales herramientas que nos fueron llevando hacia la abstracción: ¿No es una de las aplicaciones más claras de visualización la representación de un espacio físico en un espacio menor? (un mapa, pues). Uno de mis ejemplos favoritos que muestra cómo la abstracción se va haciendo necesaria y útil es la Tabula Peutingeriana [1], un mapa distorsionado para reducir el espacio poco importante (el mar) y enfatizar en el espacio de interés (dando dimensiones exageradas a la península itálica, por ejemplo) para presentar un mapa de carreteras del imperio romano. Una vez más, cito un texto de la Wikipedia:

“Existen dudas acerca de si se debería llamar “mapa” a la Tabula Peutingeriana, ya que no intenta representar las formas reales del terreno. Se cree que la Tabula se basa en «itinerarios», o listas de destinos a lo largo de las calzadas romanas. Los viajeros no poseían nada tan sofisticado como un mapa, pero necesitaban saber lo que tenían por delante en la calzada y a cuanto quedaba.”

Y sí, a diferencia de los mapas, la Tabula incorpora la cuantificación semántica de cada porción del territorio —sobrepone de una forma maravillosa una imagen bidimensional con la densidad relativa del interés que despierta cada parte del mapa. Un usuario de este mapa debe saber que no está a escala, que el Mediterráneo es más que un conjunto de ríos alrededor de la tierra, que los ángulos en los caminos no serán como están descritos —pero podría seguir cada uno de los pueblos a su paso, y tomar como referencia las características que más llamaron la atención del autor. Ver Figura 1.

Y es precisamente lo que hicieron los autores de esta tábula lo que nos toca emular hoy en día, en la época del big data y de los procesadores que miden su eficiencia en miles de millones de operaciones por segundo. Pero podemos aprovechar aún más esta ventaja que tenemos si seguimos las enseñanzas de Charles Minard: Encontrar abstracciones y mecanismos que nos permitan comprimir un conjunto numérico multidimensional en una proyección de menor dimensionalidad, sin pérdida relevante de información.

¿A qué me refiero con esto? La famosa “carta figurativa de las pérdidas sucesivas en hombres”[2] registra el avance geográfico de las tropas de Napoleón en su malograda campaña a Rusia (y su todavía más impresionante retorno ya como ejército fracasado), referenciado contra las gélidas temperaturas del invierno ruso. La gráfica incluso indica algunos hitos geográficos (los principales ríos y ciudades) como referencia. Bueno, vamos entonces entrando en materia: Lo que nos permite hacer la visualización es presentar grandes cantidades de información numérica, que podría parecer incomprensible al humano, de modo que su significado nos resulte de natural comprensión. Después de una muy larga



Gunnar Wolf es administrador de sistemas para el Instituto de Investigaciones Económicas de la UNAM y desarrollador del proyecto Debian GNU/Linux.
<http://gwolf.org>

Figura 1. Tabula Peutingeriana.



introducción, vamos al tema principal que yo quería abordar cuando comencé con la escritura de esta columna: un programa en específico para visualizar elementos de la administración de sistemas, Munin[3].

Munin

Munin es una herramienta libre, diseñada para ayudar a administradores de sistemas y de redes a analizar los distintos aspectos de sus distintos equipos, servidores y subsistemas. Su diseño es muy sencillo, y permite comenzar a ser de utilidad tan pronto es instalado, y al mismo tiempo, está pensado para ser arbitrariamente extensible. Munin sigue de cerca la línea de análisis histórico que presento, y eso le ganó bastante simpatía por mi parte: Cuenta la mitología noruega que Munin y Hugin son los dos cuervos emisarios del dios Odín. Ellos vuelan todos los días por la Tierra (Midgard), viendo y recordando todo lo que ocurre, para contárselo. Munin representa a la memoria, y Hugin al pensamiento.

Munin es una excelente herramienta de visualización, en el sentido previamente descrito: presenta números aparentemente inconexos e incomprensibles de forma que, viendo un conjunto de gráficas, resulta muy simple no sólo el relacionar a los patrones presentados visualmente, sino que expresa ciertos subsistemas complejos de maneras tan obvias que termina ilustrando no sólo el qué hace, sino el cómo funciona.

Para ilustrar a qué me refero, presento aquí la gráfica de uso de memoria de uno de los servidores que administro. Ver Figura 2. Es común que nos preguntemos cuánta memoria disponible tiene nuestro servidor —en un sistema Linux estándar, podríamos ingenuamente responder con los números que nos genera el comando free, e incluso sería frecuente que nos restringamos a las tres primeras columnas:

	total	used	free	shared	buffers	cached
Mem:	3374916	2723708	651208	0	304720	1498288
-/+ buffers/cache:		920700	2454216			
Swap:	3903784	191292	3712492			

El tener gráficamente la proporción de áreas de memoria en cada uno de los estados presentados, sin embargo, nos lleva a entender mucho mejor lo que ocurre: ¿Por qué es que tengo casi la cuarta parte de la memoria marcada como no utilizada (amarillo claro) y al mismo tiempo estoy utilizando swap? ¿Ah! Con la gráfica queda claro que una porción no trivial de la memoria se mantiene disponible, por lo visto, para dar plasticidad al sistema, para permitir que los procesos entren y salgan de ejecución. Y para eso se emplea también buena parte de los buffers, que son presentados en color violeta. El uso duro real de la memoria está representado por la sección más baja que ocupa un espacio significativo, las aplicaciones (verde). Ver Figura 2.

No sé si la fidelidad de la impresión permita apreciar las líneas que aparecen superpuestas a las áreas: Las líneas que marcan los niveles de

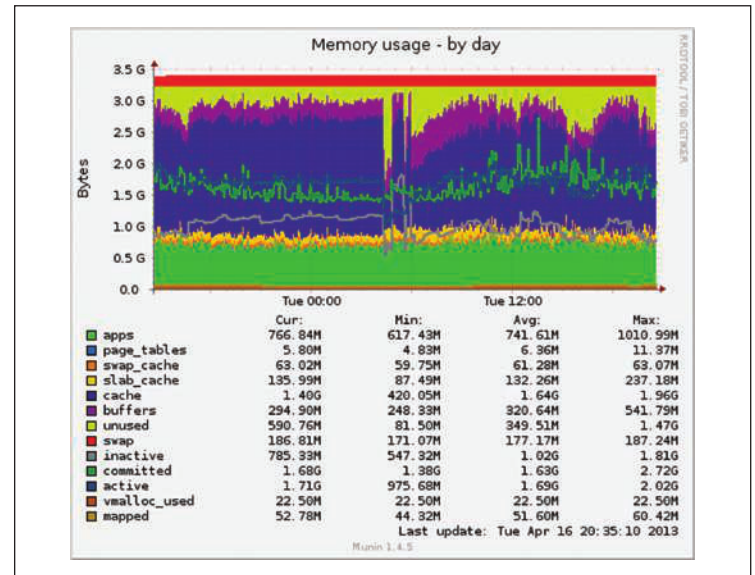


Figura 2. Munin: Visualización de uso de memoria.

memoria activa, comprometida e inactiva. Estos tres niveles sirven mucho mejor como termómetro para entender el nivel de estrés que está el sistema enfrentando en un momento dado. Las dos primeras líneas oscilan, en el periodo de un día que les presento, entre el 50 y el 70% de utilización, con un pico mínimo de cerca del 40% (para la memoria activa) y un pico máximo de cerca del 85% (para la memoria comprometida).

No entro en detalles intentando explicar qué significa cada cresta o valle, y a qué eventos del día podrían estar respondiendo. Mi intención es presentar una herramienta libre y sencilla que me ha enseñado mucho y se ha convertido en un aliado permanente en mi tarea de administración de sistemas.

Si les interesa encontrar mayor información acerca de Munin, les invito a leer un artículo que publiqué al respecto [4], así como una presentación donde entro en más detalle en ciertos aspectos de la administración de sistemas[5]. También dejo la referencia al sitio oficial de Munin, donde podrán encontrar todos los ejemplos, documentación e información que puedan requerir al respecto. [SC](#)

>> Por Gunnar Wolf

Referencias

- [1] https://es.wikipedia.org/wiki/Tabula_Peutingeriana
- [2] https://en.wikipedia.org/wiki/Charles_Joseph_Minard
- [3] <http://munin-monitoring.org>
- [4] Monitoreo de PostgreSQL con Munin, Revista Cubana de Ciencias Informáticas, 2011; <http://ru.iiec.unam.mx/116/>
- [5] Monitoreo de redes con Munin, FONASOL 2008, <http://gwolf.org/soft/munin>