

Poder computacional al alcance de todos: ¿Cómo aprovechar los beneficios de la tecnología de Grid Computing en Latinoamérica?

Mauricio PALETTA

Departamento de Ciencia y Tecnología - Universidad Nacional Experimental de Guayana
Ciudad Guayana, Estado Bolívar, 8050, Venezuela

RESUMEN

El objeto principal de este artículo es mostrar un análisis de cómo países en desarrollo, en particular latinoamericanos, contando con un amplio margen de poder computacional que está en manos de un mínimo sector y que además es subutilizado, puede ser aprovechado, principalmente por los sectores académicos y de investigación, haciendo uso de la tecnología de Grid Computing (GC). En este sentido, el artículo se enfoca en el siguiente dilema: “aprovechar el poder computacional que está al alcance de unos y que puede ser requerido por otros”. El artículo presenta además, el entorno introductorio de la tecnología de GC, el por qué de su creación e identificar claramente los beneficios que se buscan con su aplicación.

Palabras Claves: Grid Computing, organizaciones virtuales, Latinoamérica, poder computacional.

El objeto de este artículo es mostrar, en primer lugar, el entorno introductorio de la tecnología de GC, el por qué de su creación e identificar claramente los beneficios que se buscan con su aplicación. En segundo lugar, se presenta un análisis de cómo países en desarrollo, en particular latinoamericanos, contando con un amplio margen de poder computacional que está en manos de un mínimo sector y que además es subutilizado, pueden aprovecharse ampliamente de los beneficios que aporta esta tecnología, principalmente en lo que a las comunidades académicas se refiere.

El resto del artículo está organizado de la siguiente manera: la sección 2 presenta el entorno introductorio de la tecnología de GC; en la sección 3 se muestra un análisis de cómo se puede aprovechar esta tecnología en los países en desarrollo, presentando como ejemplo la situación particular de Ciudad Guayana en Venezuela. Finalmente, la sección 4 presenta las conclusiones del trabajo.

1. INTRODUCCIÓN

Es común ver en países en desarrollo como los de Latinoamérica, cómo se presentan situaciones en las cuales algunos sectores, particularmente académicos/científicos, requieren algún poder computacional que no puede ser dado por las instalaciones que están al alcance en ese momento. Basta citar por ejemplo, escasas horas de tiempo computacional que han sido asignadas a ciertos estudiantes en alguna universidad, inscritos en una asignatura que exija la elaboración de proyectos complejos de software que requieran más del tiempo/poder computacional asignado. La situación se torna interesante al conocer que en otros sectores cercanos (físicamente hablando), pudiera encontrarse un potencial computacional suficiente para cubrir estas necesidades. Luego, ¿cómo hacer para que este poder computacional pueda ser compartido en base a las necesidades actuales de algunos?

En los últimos años la tecnología de Grid Computing (GC) [1], [2] ha venido tomando fuerza como un intento de aprovechar el poder computacional que está al alcance de unos y que puede ser requerido por otros. Creada en principio para ser utilizada como una herramienta para el entorno científico / académico, hoy en día está también adoptando un comportamiento de índole comercial para beneficio de la industria, lo que está llevando a esta tecnología a ser actualmente un tema de gran interés científico en el área de la Informática. Prueba de ello lo demuestran la gran cantidad de proyectos, aplicaciones, organizaciones, alianzas, asociaciones y conferencias que están relacionadas con ella.

2. GRID COMPUTING

Introducción

La mejor forma de comprender la razón de ser inherente a la tecnología de GC es mostrando un escenario que sirva de ejemplo. En este sentido, supongamos se tiene un científico que, para probar los resultados de uno de sus trabajos, requiere computar el equivalente a 100.000 compuestos en una hora. En este contexto no es relevante conocer el por qué son 100.000, a qué se refiere cuando se habla de un compuesto y por qué debe ser en una hora; lo importante es que se sabe que el computador más potente que tiene disponible en su centro de investigación es capaz de computar un compuesto en 15 segundos, lo que implica que en una hora sólo será capaz de computar 240 de los 100.000 que necesita, lo cual no es aceptable para el éxito del experimento que está tratando de hacer. Por lo tanto el científico debe tratar de buscar una alternativa que lo saque de este problema y para ello se le presentan dos posibles opciones:

- 1) Buscar la posibilidad de usar una super-computadora; para ello debe lidiar con disponibilidad, costos, manejo de procesos y datos, etc. Si bien esta opción debe mejorar significativamente en la cantidad de cómputos que puede tener en una hora, no se puede garantizar que se alcance el tope de una hora requerido.
- 2) Seguir el principio “en la unión está la fuerza” y tratar de conseguir varias computadoras que, en conjunto, permitan lograr el objetivo que se quiere. Una alternativa se tiene por ejemplo, haciendo trabajar en paralelo 417 computadores similares al más potente que está en el centro de investigación. Pero, si se quiere ir por esta opción, deberá manejar adecuadamente la distribución de los procesos

(cómputos en este caso) y centralizar luego los resultados. En otras palabras, se requieren habilidades propias de la programación y/o computación distribuida.

En ambos casos se observa que hay que afrontar nuevas dificultades y contratiempos que pueden ir en contra del éxito de lograr un buen experimento. Lo ideal es que el centro de investigación en el cual se está formulando este escenario, pertenezca a algún “tipo de comunidad” que ponga al alcance de todos, el poder computacional que se tiene, de manera que todos en esa “comunidad” pueden beneficiarse mutuamente. Y, mejor aún, que este servicio sea transparente para todos los que pertenecen a la “comunidad”.

Lo relatado anteriormente es precisamente el objetivo principal que se busca con la tecnología de GC. Si bien es cierto el escenario antes relatado es ideal, hay muchos escenarios de la vida real que se pueden citar y que tienen aspectos comunes de la problemática presentada. Por ejemplo:

- 1) Científicos a lo largo del mundo que requieren analizar grandes volúmenes de data para compartir resultados en cuanto a temas como los cambios climáticos, el crecimiento de la población, etc.
- 2) Estudiantes y demás miembros de los centros educativos que requieren recursos computacionales para completar sus actividades académicas y de investigación.
- 3) Empresas que deben hacer simulaciones complejas sobre nuevos productos, análisis de mercados, etc.

Es claro que toda tecnología requiere de una serie de elementos concretos para su correcto aprovechamiento y GC no se escapa de ello. En este sentido, para lograr alcanzar el objetivo relatado anteriormente, se necesita alcanzar los siguientes ítems:

- 1) Tener acceso a los recursos computacionales por medio de una infraestructura de hardware y software debidamente preparada para ello.
- 2) Ser consistente (disponible todo el tiempo), seguro (en el uso de los recursos), global (alcanzable por cualquiera que se quiera unir) y económico.
- 3) La ausencia de una localidad y/o control central.

Algunos de estos elementos ya son presentes en otras tecnologías como el de las super-computadoras (super-computing), la computación basada en clusters (cluster-computing) [3] y la computación basada en intranet (intranet-computing). Las super-computadoras tienen el problema que no son escalables, son equipos muy costosos (adquisición y mantenimiento) y no garantizan los resultados de tiempo de respuesta deseados. Por su parte, la computación basada en clusters hace un uso extensivo de la comunicación y los costos de mantenimiento son muy altos. Por último, la computación basada en intranet no permite el manejo de recursos fuera del dominio de administración del sistema, por lo general hay incumplimiento de las políticas de seguridad y administración de recursos y los protocolos e interfaces básicos no son basados en sistemas abiertos.

Definición

GC es un término introducido por I. Foster y C. Kesselman a mediados de los 90's para denotar una infraestructura de computación distribuida propuesta para atender problemas avanzados de la ciencia y la ingeniería. La idea básica que se

propusieron estos autores es hacer computación distribuida lo más transparente posible para el usuario, inspirado en el servicio de suministro de energía eléctrica y el concepto de “Plug & Play”.

Una definición específica dada por Foster et al en [4] es la siguiente: “GC es una infraestructura de hardware y software que provee acceso fiable, coherente, penetrante y económico para lograr capacidades computacionales”. En otra definición más detallada dada por el mismo autor [5] y conocida como la lista de verificación de 3 puntos, se define a GC como un sistema que:

- 1) Coordina recursos que no están sujetos a un control centralizado.
- 2) Usa protocolos e interfaces estandarizados y abiertos.
- 3) Para ofrecer calidades de servicio no tradicionales.

En un sistema basado en GC se pueden identificar dos tipos diferentes de usuarios: 1) los consumidores que requieren la ejecución de tareas para una variedad de problemas de diversa complejidad y tamaño, se benefician por el manejo adecuado de los recursos y están interesados en minimizar gastos; 2) los proveedores que contribuyen con el aporte de los recursos holgados, se benefician al máximo de los recursos disponibles y están interesados en los requerimientos y oportunidades del mercado (maximizar la tasa de retorno en servicios).

Por otro lado, para poder estructurar adecuadamente un sistema basado en GC es necesario satisfacer los siguientes tres elementos que se mencionan a continuación:

- 1) Una arquitectura Grid: se refiere a la identificación de componentes fundamentales, la especificación del objeto y función de estos componentes y la especificación de cómo estos componentes van a interactuar y/o compartir información entre sí.
- 2) Una plataforma de software: identificación de protocolos comunes de comunicación, identificación/definición de los servicios estándares, contar con los mecanismos de integración con aplicaciones finales (API, SDK, etc.).
- 3) Recurso humano capacitado: para la realización de tareas administrativas de alto nivel (instalación de aplicaciones, manejo de políticas de seguridad, monitoreo, etc.) y el soporte técnico de la infraestructura (instalación del software básico, la administración de la red, el soporte a usuarios, el mantenimiento de un posible portal, etc.).

En lo que respecta a los recursos que se pueden compartir en GC, se tienen los siguientes: 1) computacionales: iniciar programas y monitorear y controlar la ejecución de los procesos; 2) almacenamiento: guardar y recuperar archivos; 3) red: manejar mecanismos que proveen control sobre recursos asignados en la transferencia; 4) repositorios: manejo de versiones de código y 5) catálogos: operaciones especiales sobre datos.

Organizaciones Virtuales

Además de haber introducido la tecnología, otro de los aportes de mayor importancia relacionado con GC y dado por Foster et al [1] es el que tiene que ver con el concepto de las Organizaciones Virtuales (VO – de sus siglas en inglés Virtual Organizations): El verdadero y específico problema que trata el concepto de GC es la coordinación de recursos compartidos y la resolución de problemas en VOs.

Las VOs pueden variar en propósito, alcance, tamaño, duración, estructura, comunidad y sociología. Una VO puede estar conformada, por ejemplo, por los centros educativos de una ciudad, una región, un país o un continente. Además, no se descarta la posibilidad de que varias VOs también se beneficien mutuamente. En este sentido, desde el punto de vista del tamaño o cobertura de una GC se presentan tres posibles casos o topologías (ver Figura 1):

- 1) Intragrid: sólo una VO; no hay ninguna integración entre “socios”.
- 2) Extragrid: múltiples VOs que manejan un mismo dominio de razón de ser (integración entre “socios”).
- 3) Intergrid: muchas VOs de diversa índole.

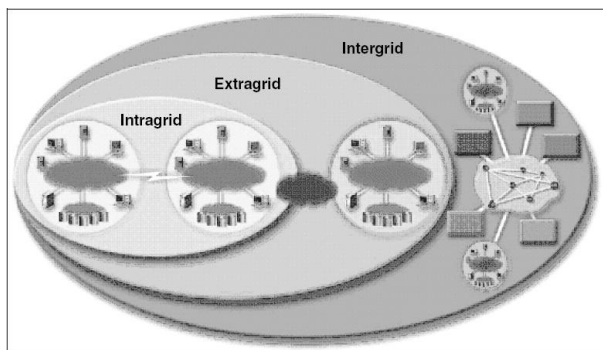


Figura 1. Topologías de Grid.

En la próxima sección se presenta el estado actual de la tecnología de GC en Latinoamérica y se muestra un análisis de la manera en la cual estos países se pueden aprovechar de los beneficios que la tecnología brinda.

3. ¿CÓMO APROVECHAR LOS BENEFICIOS DE LA TECNOLOGÍA DE GC EN LATINOAMÉRICA?

En esta sección se presenta, por una parte, el estado actual de la tecnología de GC en Latinoamérica y, por otra parte, un ejemplo donde se ilustra la manera en la cual esta tecnología puede aportar grandes beneficios a estos países en pleno desarrollo.

Estado actual de la tecnología de GC en Latinoamérica

No debería caer de sorpresa el hecho de que la tecnología de GC está en auge en países desarrollados, específicamente de Europa y Norteamérica, prueba de ello lo representan proyectos como CrossGrid¹, EU-DataGrid², IrisGrid³, Globus⁴ y otros. Sin embargo, en América latina este auge también está presente y, si bien está apenas iniciando, existen intentos de importancia como la alianza CLARA⁵ (Cooperación Latino Americana de Redes Avanzadas), una organización no gubernamental sin fines de lucro que estimula la cooperación regional en actividades educativas, científicas y culturales, además de promover la integración directa con las comunidades científicas de Europa. El objeto principal de esta organización es tratar de

¹ <http://www.eu-crossgrid.org>

² <http://www.eu-datagrid.org>

³ <http://www.rediris.es/irisgrid>

⁴ <http://www.globus.org>

⁵ <http://www.redclara.net/>

integrar una red regional de telecomunicaciones de la más avanzada tecnología para interconectar a las redes académicas nacionales de la región [6]. En estos momentos se pueden identificar los siguientes miembros.

- Red Teleinformática Académica de Argentina (RETINA).
- Agencia para el Desarrollo de la Sociedad de la Información en Bolivia (ADSIB).
- Red Nacional de Enseñanza e Investigación de Brasil (RNP).
- Red Nacional Académica de Tecnología Avanzada de Colombia (RENATA).
- Red Nacional de Investigación de Costa Rica (CR2Net).
- RedUniv de Cuba.
- Redes Universitarias Nacionales de Chile (REUNA y GREUNA).
- Consorcio Ecuatoriano para el Desarrollo de Internet Avanzado (CEDIA).
- Red Avanzada de Investigación, Ciencia y Educación Salvadoreña (RAICES).
- Red Avanzada Guatemalteca para la Investigación y Educación (RAGIE).
- Universidad Tecnológica Centroamericana en Honduras (UNITEC).
- Corporación Universitaria para el Desarrollo de Internet de México (CUDI).
- Red Nicaragüense de Internet Avanzada (RENIA).
- Red Científica y Tecnológica de Panamá (RedCyT).
- Arandu de Paraguay.
- Red Académica Peruana (RAAP).
- Red Académica Uruguaya (RAU)
- Red Académica de Centros de Investigación y Universidades Nacionales de Venezuela (REACCIUN).

En este mismo orden de ideas, uno de los proyectos de mayor importancia que en GC se adelanta actualmente en Latinoamérica es EELA⁶ (de sus siglas en inglés E-infrastructure Shared between Europe and Latin America). Su objetivo es llevar las e-Infraestructuras de los países latinoamericanos (actualmente participan 10 países con sus redes académicas y diferentes instituciones entre universidades y centros de investigación) al nivel de explotación de los países europeos [6].

Ahora bien, ¿por qué es importante para los países de Latinoamérica el tratar de evolucionar en la tecnología de GC? Está claro que parte de la respuesta a esta pregunta tiene que ver con la obtención de los beneficios que en general se puede obtener con esta tecnología:

- 1) Mejorar la relación eficiencia/costo.
- 2) Explotar convenientemente los recursos sobre-utilizados.
- 3) Permitir la colaboración.
- 4) Crear VOs en diferentes dominios.
- 5) Incrementar la capacidad de procesamiento paralelo.
- 6) Dar soporte a sistemas heterogéneos.
- 7) Proveer realidad/disponibilidad.
- 8) Reducir el tiempo de obtención de resultados.

Muchos de estos beneficios tienen mayor relevancia en países en desarrollo como los de América latina, específicamente en el área académica y científica, ya que:

⁶ http://www.eu-eela.org/eela_about.php

- 1) Ofrece poder computacional y otras capacidades relativas a la tecnología de información que suele ser insuficiente en centros de investigación e instituciones académicas.
- 2) Permite el intercambio de conocimiento entre organizaciones, particularmente con las que se encuentran en los países desarrollados.
- 3) Es una infraestructura ideal para implementar e-learning.
- 4) Es un campo amplio y joven para hacer investigación en esta área.

Un ejemplo

El ejemplo que se va a mostrar a continuación tiene que ver con Ciudad Guayana, una ciudad del sur-oeste de Venezuela de unos 750 mil habitantes. Esta ciudad es cuna de más de diez centros de educación superior, todos con un nivel de operación de bajo presupuesto. En este sentido, es común observar en estas instituciones serias deficiencias en la infraestructura computacional para actividades que están relacionadas con la academia y la investigación. Por otro lado, esta ciudad es también cuna de grandes empresas (promedio de 5.000 trabajadores), específicamente relacionadas con la generación eléctrica y la extracción y procesamiento de materia prima (hierro, aluminio, etc.).

Es un hecho que estas grandes empresas cuentan con una gran cantidad de poder computacional instalado, la mayoría de ella de reciente tecnología y dedicada a labores secretariales y administrativas y, más aún, cuyo uso se restringe a una cantidad reducida y específica de horas durante la semana. Cabe preguntarse entonces, ¿puede, la deficiencia computacional presente en la comunidad académica/científica de esta ciudad, ser cubierta o satisfecha por el poder computacional holgado de la industria instalada en la misma ciudad? La respuesta es sí, ¿cómo?, con tecnología de GC.

Si nos imaginamos el escenario de una VO formada entre los diferentes centros educativos (sea VO-edu esta organización) que, no sólo permitirá la compartición de recursos y conocimiento entre ellos, sino también, cada institución puede aprovechar todo el potencial computacional que existe dentro del centro educativo, independientemente si está destinado a labores académicas o administrativas. En este sentido, cabe resaltar que uno de las áreas de investigación inherentes a la tecnología de GC y del cual pueden beneficiarse ampliamente las instituciones académicas y de investigación, tiene que ver con el intercambio de conocimiento [7], [8], [9].

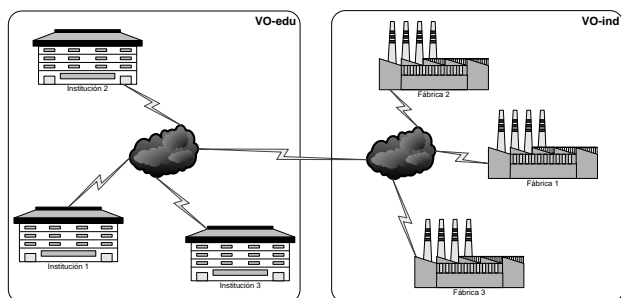


Figura 2. Un posible escenario de VOs.

En este mismo orden de ideas, nos podemos imaginar también a las empresas de Ciudad Guayana organizadas en una VO (sea VO-ind esta organización) para su beneficio mutuo; luego, es posible tener una VO de mayor jerarquía (posiblemente

relacionada con la ciudad misma) que permita la cooperación entre la VO-edu y la VO-ind. La Figura 2 presenta lo que puede ser un esquema de este escenario.

4. CONCLUSIONES

En base al desarrollo realizado en este artículo es posible emitir las siguientes conclusiones:

- 1) GC es una evolución de super-computing, cluster-computing e intranet-computing para hacer computación distribuida Plug & Play.
- 2) GC es un tema desafiante que tiene relación con otra gran cantidad de tecnologías de diversa índole.
- 3) ¿Es GC el futuro de Internet? En estos momentos es una pregunta abierta. El tiempo determinará la respuesta.
- 4) La tecnología de GC en Latinoamérica está apenas emergiendo con el intento de interconectar varios países de la región.
- 5) Dadas las características actuales del potencial computacional y de comunicaciones que se pueden encontrar en varias regiones de Latinoamérica, es factible y útil el buscar definir VOs para explotar la tecnología de GC y aprovechar sus beneficios, importante para el desarrollo creciente de estos países.

5. REFERENCIAS

- [1] I. Foster, C. Kesselman, S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations", **International Journal of Supercomputer applications and High Performance Computing**, Vol. 15, No. 3, ISSN:1094-3420, 2001, pp. 200-222.
- [2] R. Buyya, D. Abramson, J. Giddy, H. Stockinger, "Economic Models for Resource Management and Scheduling in Grid Computing", **Journal of Concurrency and Computation: Practice and Experience**, Vol. 14, Issue13-15, Wiley Press, 2002, pp. 1507-1542.
- [3] R. Buyya (ed.), **High Performance Cluster Computing: Systems and Architectures**, ISBN: 978-0130137845, Prentice Hall, 1999.
- [4] I. Foster, C. Kesselman (eds.), **The Grid: Blueprint for a New Computing Infrastructure**. Morgan Kaufmann, 1999.
- [5] I. Foster, "What is the Grid? A Three Point Checklist", **GRIDToday** (Daily news and information for the Global Grid Community), Vol. 1, No. 6, 2002.
- [6] W. Díaz Garzón, "Realidad de la Computación Grid en América Latina", **Sociedad de la Información**, ISSN: 1578-326x, No. 12, 2007.
- [7] F. Berman, "From TeraGrid to Knowledge Grid", **Communications of the ACM**, Vol.144, No. 11, 2001, pp. 27-28.
- [8] M. Cannataro, D. Talia, "Knowledge Grid: an architecture for distributed knowledge discovery", **Communications of the ACM**, Vol. 46, No. 1, 2003, pp. 89-93.
- [9] H. Zhuge, **The Knowledge Grid**, World Scientific, ISBN: 981-256-140-4, 2004.