



Comunicación

335

TOMAS³: TOWARDS AND OPEN MANAGEMENT ARCHITECTURE FOR SYSTEMS, SOFTWARE AND SERVICES

Ismael Herrero Rodríguez

Analista
Ándago Ingeniería

M^a Isabel Martín, José Gato, Álvaro del Castillo, Rafael G. Leiva

Ándago Ingeniería

Marta Patiño, Ricardo Jimenez-Peris, Álvaro Orús, Jose Antonio Sánchez

Universidad Politécnica de Madrid

Palabras clave

Cluster, Grid, Open Source, Software Libre, Código Abierto, Administración, Management, CERN, TOMAS, Quattor, Linux, GNU, GNU/Linux, Configuration Management, Software Management, Hardware Management, Failure Tolerance, Authonomic Computing, Monitoring, Users Management, Services Management, WSDM, SOA, SOAP, WSDL, CIM, DTMF, OASIS, Ándago, UAM, UPM, Ministerio de Educación y Ciencia, Ministerio de Industria, Turismo y Comercio.

Resumen de su Comunicación

En el presente documento se describe TOMAS³, un estándar de código abierto para el despliegue, instalación, configuración, administración, monitorización y mantenimiento de medianas y grandes redes de ordenadores bajo el sistema operativo GNU/Linux. Las herramientas descritas permiten la total administración y operación de una red de ordenadores reduciendo sustancialmente los costes de operación.

TOMAS³: TOWARDS AND OPEN MANAGEMENT ARCHITECTURE FOR SYSTEMS, SOFTWARE AND SERVICES

1. Introducción

Ándago Ingeniería es una empresa del grupo Talde dedicada a servicios de consultoría que trabaja en proyectos de gran envergadura, utilizando para ello tecnologías de Software Libre. Recientemente Ándago Ingeniería ha decidido dotarse de un departamento de Investigación, Desarrollo e Innovación, que junto con la colaboración de varias empresas privadas e instituciones universitarias investiga en nuevas tecnologías para avanzar en el desarrollo del software y ofrecer productos punteros y de valor añadido a sus actuales y futuros clientes.

Uno de esos proyectos es la creación de un estándar y una solución completa y abierta para la gestión efectiva de grandes redes de ordenadores corriendo sistemas operativos *nix, principalmente Linux. Se trata del proyecto TOMAS³: Towards An Open Management Architecture for Systems, Software and Services. El proyecto TOMAS³ pretende fomentar el uso del software libre de dos formas: primero integrando en su arquitectura soluciones de código abierto, aunque no limitándose a éstas, y segundo basando la gestión del proyecto en herramientas de desarrollo colaborativo que permiten una libre y efectiva diseminación del conocimiento adquirido durante el desarrollo del proyecto a toda la comunidad.

Asimismo se pretende avanzar en el desarrollo de estándares utilizándolos para el diseño de la arquitectura de TOMAS³, como la modelización por medio de CIM [7] del DMTF [6] y la aplicación de servicios web via WSDM [8] de OASIS [9] entre varios otros estándares.

Además de ello la intención de Ándago respecto a TOMAS³ es la de liberar el resultado del proyecto en el dominio público a fin de implicar en su desarrollo a una comunidad de desarrolladores y usuarios interesados en el mismo para su permanente soporte, extensión y mejora, haciendo el resultado del proyecto útil y duradero en el tiempo. Actualmente tanto las fuentes como la documentación del proyecto son de acceso público, pero no está permitida su modificación. El equipo de TOMAS³ pretende presentar a la comunidad un proyecto terminado y utilizable antes de comenzar su distribución.

TOMAS³ es un proyecto PROFIT [4] del Ministerio de industria, turismo y comercio [2] y del Ministerio de educación y ciencia [3]. Conjuntamente a Ándago están participando en el desarrollo de TOMAS³ las siguientes entidades:

- el Centro de Referencia Linux [1] de la Universidad Autónoma de Madrid
- el Laboratorio de Sistemas Distribuidos [5] de la Universidad Politécnica de Madrid

que unen a los aportados por Ándago recursos de investigación.

Algunas empresas privadas también han mostrado su interés en el proyecto.

En este artículo se presentan los requisitos establecidos por el equipo de TOMAS³ para ser cumplidos por la aplicación.

2. Motivación

Una red de ordenadores tipo cluster puede contener una gran variedad de equipos, con distintas características hardware y con funciones muy dispares. Ejemplos de los equipos que podemos encontrar en este tipo de redes son:

- Servidores de ficheros.
- Servidores de correo.
- Nodos de cálculo donde se ejecutan trabajos.
- Nodos maestro que coordinan el trabajo de dichos nodos.
- Servidores de almacenamiento de datos, como servidores de disco (NFS, AFS, GridFTP, etc.), o equipos de almacenamiento en cinta.
- Servidores de instalación y repositorios de software.
- Servidores de información, como por ejemplo gestores de bases de datos.
- Equipos de infraestructura de red, tales como servidores DNS o servidores DHCP.
- Otros muchos tipos de servidores, como equipos de sincronización horaria, gestores de claves y credenciales, etc.

Gestionar una red de ordenadores de esta complejidad es una tarea ardua y difícil. El mantenimiento de la red no sólo implica la gestión de los componentes individuales (actualizaciones, parches, etc.), sino también la gestión de las interdependencias entre los distintos componentes, lo que exige un orden preciso de las tareas a realizar. Por ejemplo, parar un servidor NFS para mantenimiento implica una modificación previa de la configuración de todos los clientes para que utilicen un servidor alternativo.

Este caso tan trivial implica la necesidad de modelar las dependencias entre los distintos elementos de la red, en concreto, la dependencia entre los clientes y el servidor NFS.

También es importante que la ejecución de las tareas de administración tenga en cuenta la integridad de los trabajos. Por ejemplo, una operación de administración no debería abortar un trabajo o su entorno de ejecución a menos que haya buenas razones para ello (como por ejemplo una emergencia debido a un incidente de seguridad).

Distribuir la configuración entre las máquinas que componen una red también es un trabajo laborioso y delicado ya que puede ser diferente de una máquina a otra en un gran número de máquinas, y un mínimo error en una de ellas puede suponer una gran degradación o incluso la caída de todo el sistema. Por ello se hace necesario un sistema de gestión de la configuración que permita llevar un control preciso del estado de cada máquina.

Otra función compleja de administración es la monitorización de los equipos de la red, que debe ofrecer la posibilidad de escoger las métricas necesarias para poder obtener conocimiento del estado de la red y a su vez agrupar la información de varios ordenadores para tener una visión global del conjunto.

Una vez establecido un sistema fiable de monitorización también es deseable un sistema de alarmas que permitan la localización rápida de problemas, y más aún un sistema de tolerancia a fallos que permita ejecutar acciones que solucionen los problemas localizados de forma automática.

Normalmente, estas complejas operaciones de administración se realizan de manera manual, aunque se trata de procedimientos difíciles, caros, y muy propensos a errores, sobre todo en redes de gran tamaño. TOMAS³ viene a subsanar esta carencia.

Herramientas de Administración de Sistemas

Existen numerosas herramientas de ayuda para la instalación, configuración y mantenimiento de una red de ordenadores bajo Linux. En [15] podemos encontrar una comparativa de las principales herramientas existentes, y en [16] un análisis más detallado de aquellas herramientas que han sido específicamente diseñadas para la instalación y gestión de clusters bajo Linux. Algunas de ellas son muy potentes pero no ofrecen soluciones completas sino parciales para las múltiples tareas de administración descritas. Otras son suficientemente generales, pero son productos propietarios y cerrados que no permiten auditorías, ampliación ni mejoras de su código. Por todo ello es deseable una solución de código abierto.

Entre las herramientas de software libre que proporcionan soluciones integradas de administración se encuentra Quattor [10]. Quattor proporciona un sistema de configuración de las máquinas centralizado, un repositorio de software común y un sistema de despliegue automático, cumpliendo varios de los requisitos especificados para TOMAS³. Por ello el desarrollo de TOMAS³ partirá de los elementos proporcionados por Quattor adaptandolos a las nuevas necesidades y añadiéndole todo el resto de funcionalidades especificadas usando otras soluciones.

3. Descripción General del Sistema TOMAS³

TOMAS³ es tanto una aplicación middleware para la administración de medianos y grandes sistemas informáticos como la definición del conjunto de requisitos que ese sistema de administración centralizada debe cumplir. Además proporciona un framework para integrar las soluciones que ofrezcan las funcionalidades que cumplan dichos requisitos.

Por tanto TOMAS³ no es una aplicación monolítica sino una infraestructura donde se integran diversas soluciones. Además de dicho framework el proyecto TOMAS³ proporcionará un entorno completo que integrará un conjunto de soluciones seleccionadas y será perfectamente funcional.

La arquitectura global de TOMAS³ ofrece las siguientes funcionalidades:

- configuración: gestión y almacenamiento centralizado de la información de configuración de los equipos de la red, incluye información sobre el hardware, configuración del sistema y configuración de las aplicaciones,
- instalación: instalación inicial de los equipos, distribución e instalación de paquetes de software, y configuración y mantenimiento de los equipos de acuerdo a la información proporcionada por el subsistema de configuración,
- monitorización: recopilación, almacenamiento y consulta de la información sobre el estado actual de los equipos
- tolerancia a fallos: correlación de la información proporcionada por el subsistema de monitorización con una información patrón, y ejecución de acciones correctoras en caso de ser necesario,
- gestión de recursos: gestión de la distribución de la carga entre los elementos de computación de la red, proporcionando una capa de abstracción sobre los gestores locales.
- administración de servicios: Posibilidad de desplegar, arrancar, reiniciar y parar servicios de forma controlada y remota en una o múltiples máquinas simultáneamente.

- administración de software: Control del software desplegado en una o varias máquinas y la posibilidad de instalar o desinstalar software de forma remota llevando un inventario preciso del software contenido en cada máquina.
- administración de hardware: Inventario preciso y actualizado de los componentes hardware de una red o un cluster.
- administración de usuarios: Gestión de usuarios del sistema tanto para los sistemas administrados como para los propios recursos de administración de TOMAS³

Consideraciones Generales de Diseño

Las características más importantes del diseño de las herramientas de instalación y configuración de TOMAS³ son:

- Se trata de un sistema no intrusivo, es decir, no se sustituye ninguno de los elementos de instalación y configuración de Linux, sino que éstos son ampliados.
- Está diseñado para ser escalable, pudiendo administrar desde redes de pocos nodos hasta redes de decenas de miles. La optimización y distribución en el uso de los recursos es un pilar fundamental en el diseño de TOMAS³.
- Es un sistema altamente modular, con unos interfaces entre módulos bien definidos, lo que facilita la sustitución o actualización de módulos individuales, y el desarrollo de nuevas versiones del sistema para distintas distribuciones Linux, e incluso para otras versiones de Unix
- Se reutilizan tantos componentes externos como ha sido posible, lo que facilita el mantenimiento del software.
- Estándares de facto de la industria son integrados y reutilizados en la arquitectura, por ejemplo, protocolos estándares como HTTP, servicios web SOA y formatos como XML, son utilizados siempre que es posible.
- Se ha optado por un diseño distribuido donde la autonomía de los equipos de la red se mantiene tanto como sea posible, existiendo instancias locales en los equipos cliente de casi todos los subsistemas, realizando localmente todas las operaciones siempre que sea posible, asegurándose de esta manera la escalabilidad del sistema.
- Se proporciona un conjunto de librerías de base para acceder a cada subsistema vía funciones de control, y una librería de métodos comunes a todos los subsistemas.

Estado actual del proyecto

La primera versión utilizable de TOMAS³ está planificada para el 1 de junio de 2006. Actualmente las funcionalidades más desarrolladas son el sistema gestor de configuraciones y el de instalaciones, que comprende la instalación manual o automática de máquinas enteras y la instalación y actualización de paquetes de software. Dado que la incorporación de estas funcionalidades está más avanzada éstas se explicarán en más detalle, mientras que el resto, debido a que podrían sufrir ligeras modificaciones en sus requisitos, se explican de forma más breve.

4. Sistema gestor de configuraciones

El Sistema Gestor de Configuraciones proporciona un entorno para la gestión de la información de configuración de los equipos de una red. Por información de configuración entendemos cualquiera de los elementos de información necesarios para configurar de manera estática un equipo. No se incluye aquella información que es dinámica o que cambia, por ejemplo el contenido de una base de datos residente en el equipo, ni aquella información que es generada por la propia máquina, como por ejemplo la carga del sistema.

El Sistema Gestor de Configuraciones se compone de los siguientes elementos:

- un servidor central donde se almacenan todos los elementos de configuración de los equipos de la red, junto con un lenguaje de alto nivel específicamente diseñado para describir estos elementos de configuración,
- un gestor local de configuraciones residente en los equipos cliente, junto a una librería que permite a las aplicaciones acceder a su información de configuración, y
- un conjunto de módulos de acceso, que nos facilitan la consulta de la información almacenada en el servidor central.

5. Sistema Gestor de Instalaciones

El Sistema Gestor de Instalaciones o IMS (del inglés Installation Management Subsystem) proporciona las herramientas necesarias para instalar y configurar el sistema operativo y las aplicaciones en los equipos de la red. IMS nos permite gestionar la instalación y la actualización del sistema operativo y del software de aplicaciones, configurar correctamente los parámetros del sistema, y aplicar todas aquellas políticas y restricciones que hayamos definido para nuestra red.

Básicamente, el sistema se compone de los siguientes elementos:

- un gestor de la configuración de los equipos cliente, que se encarga de hacer efectivos los perfiles de configuración definidos en CDB,
- un repositorio de paquetes de software, que contiene todos los paquetes de software necesarios para instalar el sistema operativo y las aplicaciones,
- un gestor y distribuidor de paquetes de software, encargado de descargar del repositorio software los paquetes a instalar, y de instalarlos en el equipo, y
- un conjunto de herramientas para la instalación inicial y desatendida de los equipos.

Nótese que la configuración de los elementos que forman IMS también es gestionada por el Sistema Gestor de Configuraciones. A continuación se describen los elementos que forman IMS.

5.1. Distribución e Instalación de Paquetes de Software

El Sistema Gestor de Paquetes de Software o SPM (del inglés Software Package Management) es el responsable del almacenamiento y la gestión de los paquetes de software, así como de la distribución e instalación de los mismos en los equipos cliente.

El sistema SPM se compone de los siguientes elementos:

- un repositorio de software: donde se almacenan todos los paquetes de software a instalar en los clientes, tanto los paquetes correspondientes al sistema operativo, como aquellos que correspondan al software de aplicación,
- un gestor avanzado de paquetes (rpmt): una aplicación construida sobre la utilidad rpm, y que le añade la capacidad de realizar múltiples operaciones en múltiples paquetes de software en una única transacción,
- herramienta SPMA: (del inglés SPM Agent) encargada de calcular la lista de operaciones de instalación, desinstalación o actualización de paquetes necesarias para actualizar el equipo,
- componente SPM: integrado dentro del sistema NCM, se encarga de configurar y ejecutar la herramienta SPMA con la lista de paquetes a instalar en el equipo en base al contenido de su perfil,
- procedimientos CDB: un conjunto de procedimientos del lenguaje Pan pensados para facilitar la definición y validación de listas de paquetes.

Todos los paquetes de software a instalar en los equipos cliente son almacenados de manera centralizada en un repositorio de software. Este repositorio cuenta con una interfaz de usuario para facilitar su gestión, permitiendo a los administradores del sistema añadir, borrar, y consultar los paquetes disponibles. Los clientes pueden acceder al contenido del repositorio utilizando los protocolos HTTP o FTP, o también mediante algún sistema de ficheros compartidos, como por ejemplo NFS. Los clientes pueden incluso almacenar localmente los paquetes a instalar, asegurándonos de esta manera la escalabilidad del sistema.

La lista de paquetes que han de ser instalados en cada equipo viene dada por su perfil de configuración. El componente SPM es el encargado de extraer esta lista de paquetes, con la ayuda de la librería NVA-API, y de crear un fichero de configuración local para la utilidad SPMA. A continuación, el componente SPM ejecuta la aplicación SPMA, que lee el fichero de configuración con los paquetes objetivo a instalar en el equipo, los compara con la lista de paquetes actualmente instalados, y calcula las operaciones de instalación, desinstalación o actualización necesarias para actualizar el equipo. Finalmente, SPMA ejecuta la herramienta avanzada de instalación de paquetes rpmt, que se encarga de llevar a cabo la transacción necesaria para hacer efectiva la nueva lista de paquetes.

La herramienta SPMA puede ser utilizada para instalar la totalidad de los paquetes de los equipos clientes, o sólo un subgrupo de ellos, dependiendo si se trata de equipos sobre los que tenemos un control total, o si lo que se quiere es garantizar que los equipos disponen de un grupo de aplicaciones determinado.

5.3. Instalación Automática de Equipos

El Subsistema de Instalación Automática de Equipos o All (del inglés Automated Installation Infrastructure) proporciona un conjunto de herramientas para la instalación inicial del sistema operativo en los equipos cliente. Esta primera instalación se realiza de manera automática y desatendida, a través de la propia red, y mediante el uso de servicios y herramientas estándares como son PXE, DHCP o TFTP.

Para poder llevar a cabo la instalación inicial de un equipo de manera remota, hay que realizar previamente los siguientes pasos:

- configurar los servicios de red necesarios, como añadir las correspondientes entradas en el servidor DNS y en el servidor DHCP para cada uno de los equipos a instalar,
- configurar el programa cargador; los equipos descargan desde la red un programa cargador o NBP (Network Bootstrap Program), que basado en un fichero de configuración, decidirá la manera de arrancar el equipo (desde disco local, o iniciar la instalación a través de la red), y
- configurar el programa de instalación del sistema operativo: si el programa NBP decide que hay que iniciar la instalación del sistema operativo, el programa instalador necesita de un fichero que contenga todos los datos básicos de configuración del sistema, tales como el esquema de particiones, el software inicial a instalar, los detalles del hardware, etc.

All proporciona las herramientas necesarias para simplificar las tareas descritas. Estas herramientas están totalmente integradas con el sistema CDB, desde donde recuperan todos los datos de configuración que son necesarios, aunque también podrían ser utilizadas de manera aislada.

El sistema All se compone de los siguientes módulos:

- aii-dhcp: que gestiona la parte de configuración del servidor DHCP para cada nodo (añadir, modificar o borrar una entrada correspondiente a un nodo),
- aii-nbp: encargado de gestionar la configuración de NBP para cada equipo (arrancar desde disco local, o iniciar un proceso de re-instalación remota desde red), y
- aii-osinstall: que gestiona los ficheros de configuración del instalador del sistema operativo.

Mientras que para redes del orden de varios centenares de equipos, los servicios involucrados (DHCP, TFTP, instalador del sistema operativo) pueden ejecutarse en una única máquina sin que ésta se sobrecargue, en redes de mayor tamaño es esencial que los equipos a instalar puedan ser particionados, debido a problemas de escalabilidad. Por esta razón, es posible instalar los tres módulos en máquinas diferentes. E incluso, para proporcionar redundancia al sistema, es también posible instalar el mismo módulo en más de una máquina (por ejemplo, tener dos servidores de DHCP, o tres de instalación del SO). La elección de qué servidores son utilizados por los clientes se realiza mediante los perfiles de configuración almacenados en CDB.

5.4. Administración de servicios

El subsistema de administración de servicios proporciona un conjunto de herramientas para la administración de los servicios ofrecidos por los nodos de la red. Desde la consola de administración se podrá monitorizar el estado de un servicio, arrancarlo, pararlo, etc. El subsistema estará dotado de una cierta inteligencia para resolver contingencias sencillas, como por ejemplo las dependencias entre servicios, de forma que será capaz de arrancarlos o pararlos en un orden concreto si fuera necesario.

5.5. Administración de hardware

El sistema de administración de hardware permite llevar un inventario actualizado del hardware que constituye el sistema, tanto a nivel de nodos como al de hardware incluido en esos nodos, así como de todos los elementos configurables susceptibles de ser incluidos en el ámbito de administración de TOMAS³, que será el de aquellos ordenadores donde corra el cliente de TOMAS³ y los dispositivos que ofrezcan una interfaz de administración WSDM.

El sistema de administración consiste en un repositorio de información sobre hardware y los servicios de almacenamiento y recuperación de información que actúan sobre ese repositorio. Además ofrecerá servicios de información sobre hardware de forma que será posible crear perfiles con unos determinados requisitos de hardware en función de su uso, y así consultar al sistema de administración de hardware si ciertas máquinas cumplen con esos perfiles.

5.6. Monitorización y alarmas

El sistema de monitorización permite saber en tiempo real el estado del sistema de forma visual. Permite agrupar la información de forma que se puede tener una visión global de un conjunto de nodos o una particular y específica de uno solo de ellos. Los grupos de nodos a monitorizar pueden ser configurados por el usuario según la arquitectura que desee, así como las métricas que desee obtener. También se guardará un histórico de los datos para poder realizar estudios estadísticos (datawarehouse).

Es posible determinar límites para las métricas monitorizadas de máquinas o grupos concretos y activar alarmas en base a éstos. Estas alarmas se visualizarán en la consola de administración y se notificará al administrador de diversas formas configurables.

5.7. Tolerancia a fallos

Utilizando el subsistema de monitorización y alarmas se podrán detectar problemas y se permitirá lanzar automáticamente y de forma desatendida procesos que solucionen los problemas detectados. Estas acciones podrán ser tan sencillas como reanunciar un servicio o tan complejas como balancear la carga de un servicio instalando desde cero una máquina, el software y la configuración necesaria y añadiéndolo al servicio. Estos procesos pueden ser determinados por el usuario.

5.8. Administración de usuarios

TOMAS³ proporcionará una interfaz para administrar de forma centralizada usuarios tanto para los sistemas administrados por TOMAS³ como para la administración de sus propias funcionalidades. La principal novedad aportada por el sistema de administración de usuarios de TOMAS³ es que se trata un punto único para múltiples clusters.

6. Conclusiones

La arquitectura de TOMAS³ definida por Ándago y el middleware resultante de implementarla permite incrementar significativamente la automatización de las tareas de instalación y mantenimiento de una red de ordenadores Linux, simplificandolas tareas de administración y por tanto, reduciendo los costes. Además, el sistema TOMAS³ permite mantener el control absoluto desde el punto de vista global del sistema y desde el particular de los nodos y la supervisión de las tareas resulta simple y rápida en las manos de un administrador de sistemas experto.

7. Referencias

- [1] Universidad autónoma de Madrid, <http://150.244.56.228>
- [2] Ministerio de turismo, industria y comercio, <http://www2.mityc.es>
- [3] Ministerios de educación y ciencia, <http://www.mec.es>
- [4] Programa de Fomento de la investigación Técnica (PROFIT) <http://www2.mityc.es/Profit/Profit/Guia/Index.htm>
- [5] Universidad politécnica de Madrid, <http://internetng.dit.upm.es>
- [6] Distributed Management Task Force, <http://www.dtmf.org>
- [7] Common Information Model, <http://www.dtmf.org/standards/cim>
- [8] Web Services Distributed Management, http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=wsdm
- [9] OASIS, <http://www.oasis-open.org>
- [10] Quattor, <http://www.quattor.org>
- [11] European Union DataGrid Project, <http://www.eu-datagrid.org>.
- [12] Foster, Ian, The Grid: Blueprint for a New Computing Infrastructure, Morgan-Kaufmann, Julio 1998.
- [13] CERN: European Organization for Nuclear Research, <http://www.cern.ch>.
- [14] EU DataGrid Fabric Management workpackage, <http://hep-proj-grid-fabric.web.cern.ch/hep-proj-grid-fabric>.
- [15] Barroso, Maite, "WP4 Report on Current Technology", DataGrid Technical Report DataGrid-O4-TED-0101-3_0, Apr. 2001 http://hep-proj-grid-fabric-config.web.cern.ch/hep-proj-grid-fabric-config/Tools/DataGrid-O4-TED-0101-3_0.pdf
- [16] García Leiva, Rafael y Peso, José del, "Open Source Solutions for Installation and Management of PC Clusters under Linux for ATLAS", ATLAS Notes ATL-SOFT-2003-001, Oct. 2002.

[17] Burgess, Mark, "Cfengine: a site configuration engine", USENIX Computing systems, Vol8, No. 3 1995.

[18] Anderson, Paul y Scobie, Alastair, "LCFG: The Next Generation", Proceedings of UKUUG Winter Conference, 2002.