

Plataforma de monitorización avanzada basada en tecnologías web

Antonio Cuadra¹, Francisco Garcés¹, Marcos Reyes¹, José M. Cantera¹, Damaris Fuentes², Luís Sánchez³

¹ Telefónica I+D, Emilio Vargas 6, 28043 Madrid, España

² IMDEA Networks, Av. del Mar Mediterráneo 22, 28918 Madrid, España

³ Universidad Carlos III, Av. de la Universidad 30, 28911 Madrid, España
{cuadras, fgarcés, mru, jmcf}@tid.es, damaris.fuentes@imdea.org, luiss@it.uc3m.es

Abstract — En la actualidad, tanto los operadores Telco como los proveedores de servicios disponen de sistemas de información que aglutinan contenidos muy valiosos, procedentes de plataformas de soporte al negocio y a operaciones que se encuentran desplegadas a lo largo de las diferentes redes subyacentes. Esta información se utiliza para propósitos muy diversos, en donde en ocasiones la gestión de tal cantidad de datos no está razonablemente optimizada. Sin embargo, la aplicación de nuevas tecnologías web –incluyendo los conceptos de web semántica– permitirán el manejo de la información existente de una forma eficiente. En este artículo se muestran las actividades de innovación realizadas en el ámbito de la supervisión de redes y servicios, en donde se han aplicado tecnologías web avanzadas sobre las herramientas tradicionales de monitorización mediante sondas no intrusivas de los servicios ofrecidos a clientes. Esta propuesta se ha materializado en un prototipo de supervisión de plataformas de IPTV (televisión sobre IP).

I. INTRODUCCIÓN

Esta propuesta describe los resultados de aplicar a los sistemas de información de un operador de telecomunicaciones una arquitectura software basada en tecnologías web avanzadas. De esta forma, la información orientada a red se transforma en recursos web, de manera que los datos puedan ser accesibles mediante servidores estándar a través de lenguajes habituales (XML, HTML, etcétera).

Dentro de las distintas áreas en las que intervienen los sistemas de información, esta aportación se ha centrado en el campo de la supervisión de redes y servicios, en donde se ha realizado una prueba de concepto que abarca tres actividades. Por un lado se ha elegido un sistema de supervisión de la calidad de servicio (QoS) existente dentro de la operadora. Parte de una arquitectura distribuida en donde el tráfico de la red se extrae a partir de una serie de sondas pasivas desplegadas a lo largo de las diferentes redes monitorizadas [1], como son la red GSM, GPRS, UMTS, IMS, IPTV, etcétera. Por otro lado, se ha realizado una conversión de la información de la red en recursos web convenientemente estructurados, de forma que los datos puedan ser accesibles mediante navegadores web estándar. Además, se ha habilitado un buscador web para explotar la información en bruto aplicando diferentes técnicas de indexación de recursos web registrados en un servidor de aplicaciones.

Estas actividades se han desarrollado dentro de un proyecto de innovación semillero denominado SEMNET (*SEMantic Web-based NETwork & Service Monitoring*) en el que están participando diferentes empresas de I+D y universidades. Los trabajos se están enfocando en la monitorización de los diferentes servicios existentes en las plataformas de IPTV (televisión sobre IP a través de un acceso de banda ancha) [2]. En concreto se ha desarrollado un prototipo basado en una serie de escenarios de IPTV y, posteriormente, se extenderá a nuevos servicios, como voz sobre IP (VoIP) [3] y a sistemas de supervisión de redes ya existentes en la red telefónica básica y en las redes móviles GSM y GPRS.

II. DISEÑO Y DESARROLLO

A. Diseño de la arquitectura

La implementación del prototipo tiene como pilar fundamental un conjunto de principios de diseño de arquitecturas software para sistemas hipermedia distribuidos, denominados en su totalidad como REST (*Representational State Transfer*) [4]. En REST todo recurso es una entidad capaz de ser identificada mediante una URL (*Uniform Resource Locator*) única con la que se puede interactuar a través del protocolo HTTP (*HyperText Transfer Protocol*), con los métodos estándar que éste ofrece (GET, POST, PUT y DELETE). Sus características aplicadas a la web surgen del enfoque más purista de la misma, destacando así una serie de diseños fundamentales clave:

- Protocolo cliente/servidor sin estado.
- Operaciones bien definidas y estándar que se aplican a todos los recursos de información.
- Sintaxis universal para definir los recursos de información.

Las mejoras conceptuales que se han conseguido con la aplicación de REST al prototipo de SEMNET son las siguientes:

- URLs limpias: Las URLs representan recursos y no acciones. De esta manera, se consiguen dos aspectos básicos:
 - Direccionamiento: La información de qué recurso está siendo manejado se encuentra en la URL.
 - Interfaz uniforme: La operación viene dada por el método HTTP, y son los mismos para cualquier recurso.

- Descubrimiento: Los recursos, al tener su propia URL, pueden ser fácilmente cacheados, guardados como marcador o indexados en el motor de búsqueda.
- Formatos de respuesta variados: Una misma acción puede devolver HTML, XML o cualquier otro formato de datos de manera simple y sencilla, según los requisitos de la interfaz cliente.
- Interoperabilidad más fácil: Para manipular un recurso de otra aplicación sólo se precisa su URL.
- Escalabilidad: Es fácil distribuir una aplicación REST, pues no tiene estado.

Estas ventajas han situado a REST como la infraestructura más adecuada para la consecución de los objetivos globales de SEMNET, permitiendo la definición de un conjunto de recursos y servicios capaces de proporcionar importantes ventajas frente a las proporcionadas por otras implementaciones de las arquitecturas SOA.

Para el diseño del prototipo se ha escogido una arquitectura REST concreta. Esta arquitectura, considerada un subconjunto de la arquitectura SOA, es la denominada ROA (*Resource-Oriented Architecture*) [5], que se fusiona perfectamente tanto con el patrón de diseño conocido como MVC (*Model View Controller*) [6], utilizado aquí como patrón global para la arquitectura, como con la arquitectura basada en capas descrita por Craig Larman [7].

B. Entorno de implementación

La implementación de arquitecturas REST presenta un alto grado de independencia frente a las tecnologías de desarrollo. Existen múltiples opciones, capaces de coexistir y cooperar entre sí, que permiten la correcta implementación de recursos REST (cualquier tecnología o tecnologías capaces de implementar y desplegar un contenido dinámico accesible vía HTTP constituyen una opción a tener en cuenta para el desarrollo de recursos).

En nuestro caso se ha construido la aplicación sobre un entorno Rails [8], un *framework* para el desarrollo de aplicaciones web, situándose encima de un servidor de páginas web. Dos de las ventajas más importantes y que han sido decisivas a la hora de seleccionar Rails como entorno para el desarrollo del prototipo es que implementa de manera automática el patrón MVC y que incluye toda una estructura para generar aplicaciones REST de una manera semi-automática. El lenguaje de programación elegido para la implementación ha sido Ruby, que además de ser potente y simple, es portable, capaz de ser ejecutado de manera indistinta en cualquier plataforma, con lo que la elección de ésta no supone ningún problema o restricción.

III. DESCRIPCIÓN FUNCIONAL DEL PROTOTIPO

A. Escenarios y fuentes de información para el prototipo

Como sistema de información de la plataforma de IPTV se ha seleccionado un sistema ya desplegado que tiene la capacidad de supervisar extremo a extremo los servicios de la red. Se basa en la elaboración de Registros Detallados que contienen la información más importante intercambiada por el usuario y los servidores de aplicaciones (vídeo, compras, *webmail*, etc.), desde la solicitud del servicio hasta la finalización del mismo.

De esta forma se reconstruyen los IPDRs (*IP Detailed Records*) de cada servicio de Imagenio. Este concepto se deriva de los registros reconstruidos en otros protocolos de comunicaciones, como es el caso de CDRs para el tráfico asociado a llamadas de voz (*Call Detailed Record*). Para ello es necesario correlacionar todas las tramas asociadas a una única sesión. Cada uno de estos registros contiene la información más relevante del diálogo completo para un servicio único solicitado entre un origen y un destino. Estos procedimientos se implementan para los niveles de aplicación considerado (RTSP, IGMP, TFTP, HTTP, etc). En este caso concreto de IPDRs, se almacenan aquellos parámetros más significativos de los mensajes asociados a una sesión de tráfico IP, y se particularizan para cada sesión bajo la nomenclatura SDR (*Session Detailed Record*). Actualmente el sistema contempla los escenarios de servicios siguientes, soportados sobre HTML:

- Servicio de Videoclub (VoD)
- Servicio de Pay per view (PPV)
- Servicio de Webmail
- Arranque de descodificadores.

La figura 1 muestra el escenario del servicio de vídeo bajo demanda (VoD), en el cual se reconstruye un IPDR para cada subservicio y sesión. En este escenario se desencadenan tres sesiones secuenciales que permiten llevar a cabo el servicio de vídeo bajo demanda (VoD). En primera instancia se verifica que el contenido no haya sido tarificado con anterioridad, dentro del plazo fijado para la visualización (normalmente 24 horas). En caso negativo se procede a facturar la película, y posteriormente se solicita la descarga de contenidos sobre el descodificador. Cada sesión generará el IPDR correspondiente.

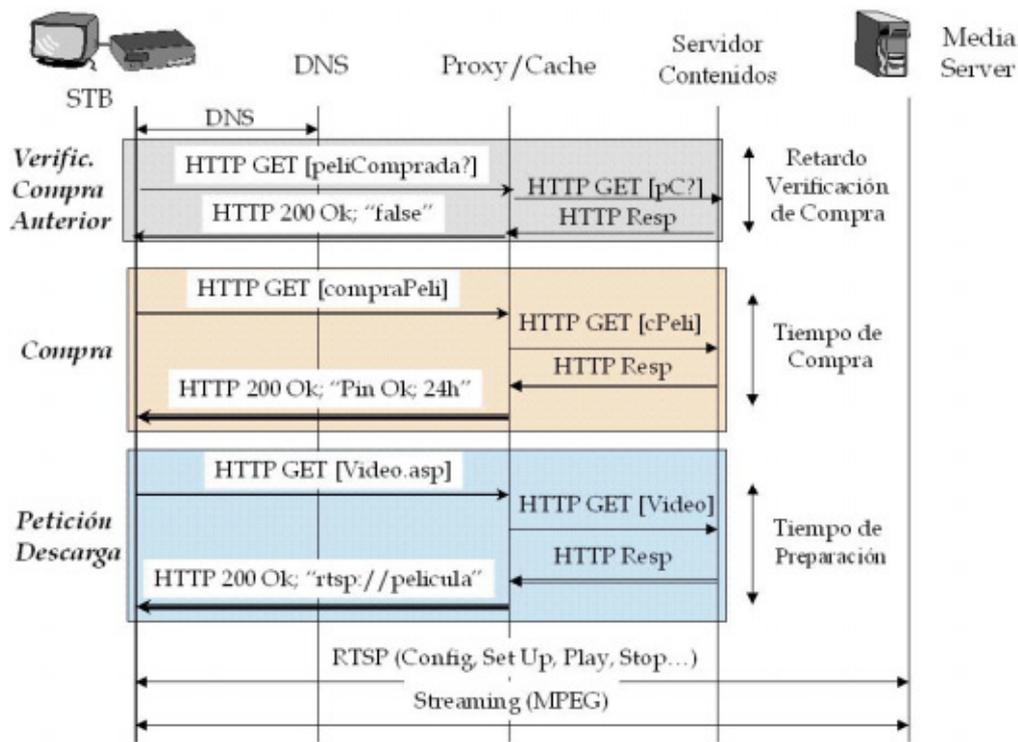


Fig. 1. Reconstrucción de IPDRs para el escenario de VoD en IPTV.

B. Portal de acceso y navegación

El acceso al portal puede realizarse con cualquier interfaz cliente HTTP, como puede ser un navegador web. El tipo de formato en el que pueden realizarse las peticiones (y recibir sus correspondientes respuestas) puede ser tanto HTML como XML, aunque pueden añadirse más formatos de una manera sencilla con la arquitectura MVC de la aplicación.

Dentro del portal, los usuarios pueden realizar las consultas pertinentes a la información almacenada. Esta información almacenada es básicamente la siguiente:

- Transacciones y registros detallados de servicio (SDRs).
- Servicios solicitados en el escenario IPTV.
- Clientes, representando entidades que inician un servicio y se corresponden con una persona física o un servidor.

El usuario podrá navegar directamente a cada uno de estos conjuntos de información desde la entrada del portal, o bien desde cada uno de esos mismos puntos de información, debido a que los distintos datos están relacionados entre sí. La navegación dispone también de mecanismos para consultas complejas, donde podrán realizarse búsquedas de determinada información en base a una y/o más propiedades de esa información. Por ejemplo, se permite la búsqueda de SDRs cuya fecha de inicio de transacción sea una determinada, el origen esté localizado en una comunidad concreta, y así con el resto de propiedades. Esta navegación se realiza mediante hipervínculos dentro de las páginas, sin necesidad de rellenar ningún formulario a la hora de seleccionar los parámetros de búsqueda o consulta, pues estos también tienen una URL propia con la que poder ser identificados y almacenados para posteriores consultas. En la figura 2 se muestran varias capturas de pantalla que representan la página principal, la colección de recursos, y el gráfico estadístico del uso de un determinado servicio.

La información almacenada se encuentra declarada en forma de recursos REST. Cada recurso cuenta con su propia dirección URL dentro del espacio de direcciones y el modelo de información definidos en el prototipo. Este modo de representar los datos permite emplear mecanismos genéricos, uniformes y bien conocidos para la manipulación, consulta y explotación tanto de la información como de los servicios implementados.

Para el prototipo se han diseñado e implementado los siguientes recursos:

- Contenidos simples: Formados por objetos únicos de información (un cliente, un servicio, un SDR).
- Feeds o contenidos complejos: Formados por colecciones de elementos (el conjunto de clientes, de servicios, etc.).
- Servicios: Formados por algoritmos aplicados a una colección de elementos (por ejemplo, el conjunto de SDRs que cumplen una determinada propiedad o propiedades).

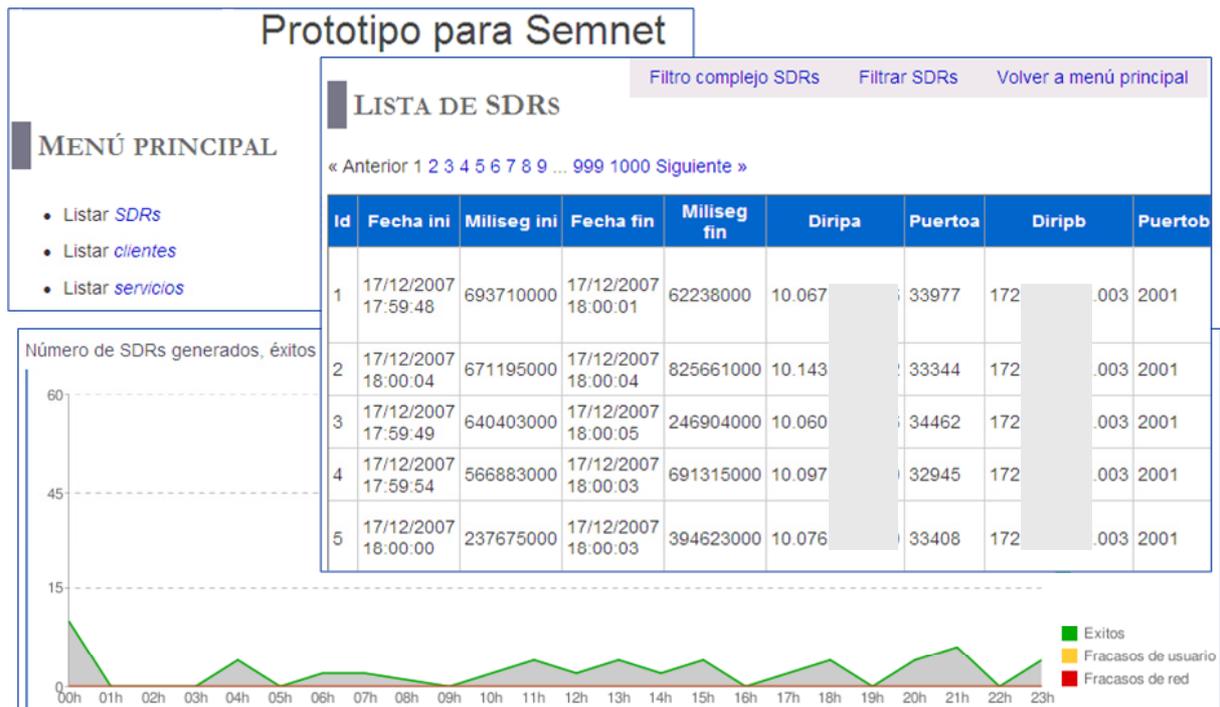


Fig. 2. Pantallas del prototipo: menú principal (izquierda), registros (derecha) y estadísticas de calidad de servicio (abajo).

Toda la información es así modelada en forma de recursos web con identidad propia fácilmente accesibles.

C. Buscador

El usuario dispone de la posibilidad de acceso a la información del sistema mediante un sistema de búsqueda web tradicional que incorpora los siguientes componentes:

- Módulo de *crawling* de páginas basado en la herramienta Nutch.
- Módulo de indexado de páginas basado en Apache Lucene.
- Módulo de consulta basado también en Apache Lucene.
- Interfaz de consulta tradicional mediante caja de texto libre.

La disponibilidad de la información del sistema como recursos Web con URIs estables y conocidas ha facilitado la incorporación del sistema de búsqueda permitiendo al usuario consultar la información de forma sencilla, versátil y familiar.

III. CONCLUSIONES Y LÍNEAS FUTURAS

En esta propuesta se han aplicado una serie de técnicas avanzadas de conversión reglada de la información mediante la adopción de una arquitectura REST, con el fin de aprovechar las facilidades web de navegación y búsqueda de datos aplicados a un sistema de supervisión de redes y servicios. El modelo adoptado en SEMNET facilita las tareas de recuperación, agrupación y explotación de datos, simplificando la integración y difusión de información en otros sistemas, y permitiendo el uso de funcionalidades colaborativas (etiquetado, posibilidad de compartir marcadores, informes, etc.) a partir de los recursos web. Además, este modelo permite utilizar buscadores estándar para localizar información relevante de forma rápida, adaptable y sencilla: tal y como se hace en Internet.

Por otro lado, se han identificado algunas limitaciones en la generalización de esta propuesta a los sistemas de información, que sin embargo permitirán abrir nuevas oportunidades dentro de esta misma línea de innovación:

- Nuevas técnicas de consolidación de los datos alternativas a BD, de cara a una ampliación de las fuentes de datos.
- Los buscadores actuales y las técnicas de *crawling* no están pensados para la indexación de información en cuasi tiempo real ("*right time*"), por lo que sería necesario adaptarlos y evolucionarlos a las necesidades de los sistemas.
- Se hace necesario explorar técnicas automáticas o asistidas para la creación de las capas de recursos de acceso web a fin de evitar la necesidad de realizar desarrollo ad-hoc.
- No se ha abordado la gestión del conocimiento, aunque el prototipo sienta las bases para favorecer el flujo del conocimiento, como utilizar la experiencia de los usuarios mediante análisis de búsquedas, o facilitar las tareas de análisis de información compartida.

Con todo, los resultados de este proyecto son aplicables a diferentes entornos nativos de supervisión: seguimiento de la calidad de servicio (QoS), *troubleshooting* y análisis detallado de la dimensión del usuario (relaciones con el cliente en diferido) y otras actividades de gestión como el dimensionado y planificación de la red.

Los resultados de esta propuesta plantean una serie de beneficios palpables para los usuarios de las herramientas de gestión derivados del acceso web a la información, como son:

- Acceso mediante web a la información del sistema de gestión.
- Homogeneización de las interfaces de los sistemas de información.
- Reducción en la formación derivada de la familiaridad con el entorno de navegación y buscador estándar.
- Ampliación de los potenciales grupos usuarios de la herramienta.
- Integración en interfaces personalizadas (*mashup*).

Por otro lado, se están realizando nuevas actividades enmarcadas en la evolución de esta iniciativa, aportando principalmente conocimiento a los recursos generados, mediante técnicas de etiquetado colaborativo de recursos, y explorando mejoras en el descubrimiento de la información por parte de los usuarios. Además, se está explorando la posibilidad de que la información de negocio generada pueda utilizarse para la composición y explotación de recursos de interacción humana (*gadgets*) compatibles con plataformas de *mashup*. Finalmente, otro de los retos futuros que se pretenden adoptar es la implantación de tecnologías de web semántica y ontologías aplicadas como solución para la interoperabilidad futura entre sistemas heterogéneos con diferente estructura de información.

REFERENCIAS

- [1] A. Cuadra, F. Garcés et al., "Tecnologías para auditorías de calidad de servicios triple play en redes Internet de nueva generación", primer premio "Nuevas Aplicaciones para Internet" de la Cátedra Telefónica en la UPM, Madrid, diciembre de 2007.
- [2] A. Cuadra, "Supervisión de la calidad percibida en plataformas de IPTV", XVII Jornadas Telecom I+D. Valencia, octubre de 2007.
- [3] A. Cuadra, "Supervisión de redes y servicios en NGN", ponencia en XIV Telecom I+D. Madrid, noviembre de 2004.
- [4] R. Fielding y R. Taylor, "Principled design of the modern web architecture", artículo en ACM Trans. Internet Technol, vol.2, 2002.
- [5] L. Richardson y Sam Ruby, "RESTful Web Services", O'Reilly, 2007.
- [6] T. Reenskaug, "The original MVC reports", Oslo, 1979.
- [7] C. Larman, "Applying UML and patterns: An introduction to object-oriented analysis and design and iterative development" New York: Prentice Hall, 2005
- [8] D. Thomas, D. Heinemeier Hansson et al., "Agile Web Development with Rails: A Pragmatic Guide." The Pragmatic Bookshelf. Raleigh, 2005.