

## SISTEMA DE MONITOREO Y CONTROL DEL MÓDULO DE ANÁLISIS DE SECUENCIAS DE LA PLATAFORMA DE SERVICIOS BIOINFORMÁTICOS.

Ing. Diana Navarro Rosquete<sup>1\*</sup>, Ing. Briseida Bussott Pérez<sup>2</sup>, Ing. Andry Daniel Díaz León<sup>3</sup>

1\* Departamento de Aplicaciones, Centro DATEC, Facultad 6, Universidad de las Ciencias Informáticas, dnavarro@uci.cu. Carretera a San Antonio de los Baños Km 2 ½, Boyeros, Ciudad de La Habana, Cuba.

2 Consejo de la Administración Municipal, Mariel, bry@ma.gobart.co.cu. Ave 67 Final, Artemisa, Cuba.

3 Centro de Estudio de Matemática y Computación, Centro DATEC, Facultad 6, Universidad de las Ciencias Informáticas, addiaz@uci.cu. Carretera a San Antonio de los Baños Km 2 ½, Boyeros, Ciudad de La Habana, Cuba.

---

### Resumen:

Con el decursar de los años, Cuba ha dado pasos agigantados en el desarrollo de las Tecnologías de la Información y las Comunicaciones, siendo pionera en este campo la Universidad de las Ciencias Informáticas. Dicha universidad posee una red de centros productivos, entre los que se destaca el Centro de Tecnologías de Gestión de Datos. Este centro cuenta con una Plataforma de Servicios Bioinformáticos, que posee una serie de módulos los cuales brindan varios servicios, específicamente el Módulo de análisis de secuencias que intenta proveer a los especialistas las funcionalidades comúnmente utilizadas en el área de la Bioinformática. Este sistema actualmente no está siendo gestionado por los administradores, por lo que el objetivo de la presente investigación es desarrollar un sistema de monitoreo y control de eventos para dicho módulo. Finalmente se obtuvo un componente de interfaz de usuario denominado Portlet, el cual visualiza diferentes indicadores del Módulo de análisis de secuencias, tales como: estado del disco duro, memoria RAM, procesos en ejecución, entre otros, a la vez que permite tomar acciones correctivas para garantizar el correcto funcionamiento de dicho módulo.

**Palabras clave:** Control; JMX; monitoreo; Módulo de análisis de secuencias; *portlet*.

### ABSTRACT:

*With the passage of the years, Cuba has made huge strides in the development of Information Technology and Communications, a pioneer in this field at the University of*

*Information Sciences. This university has a network of production centers, among which stands out the Center Data Management Technologies. This center has a research providing Bioinformatics Bioinformatics Services Platform, which has a series of modules which provide various services, specifically the sequence analysis module that attempts to provide the functionality commonly used specialists in the area of Bioinformatics. This system is not currently being managed by administrators, so that the objective of this research is to develop a monitoring and control of events for that module. Finally got a user interface component called Portlet, which displays different indicators of sequence analysis module such as hard disk status, RAM, running processes, among others, while allowing corrective action to ensure the correct operation of the module.*

**Keywords:** control; JMX; monitoring; Sequence analysis module; portlet.

---

## INTRODUCCIÓN

Desde el surgimiento de las primeras computadoras, el hombre se ha visto en la necesidad de expandir aún más sus conocimientos y romper todas las barreras impuestas por los nuevos retos y paradigmas de las ciencias de la computación. La gran cantidad de información manejada, genera un haz de conocimientos que aumenta de forma acelerada, lo cual trae consigo una mayor demanda de productos de software, soluciones informáticas y servicios que permitan gestionar la información eficientemente. Es por ello que las compañías, empresas y toda entidad laboral exigen soluciones rápidas y acordes a sus expectativas. Por estas razones fue obligatoria la interrelación de la informática con otras ciencias y aspectos de la vida social. Así surgen las Tecnologías de la Información y las Comunicaciones (TIC's), las cuales han llegado a ser uno de los pilares primordiales de la sociedad actual.

La aplicación de las Tecnologías de la Información y las Comunicaciones (TIC's) en la mayoría de los sectores de la sociedad ha reportado una mayor eficiencia en los procesos de gestión de la información, debido a la gran cantidad de conocimiento que esta engloba, esta era es catalogada como la era de la información y el conocimiento. La Bioinformática es una de las disciplinas emergentes que utiliza las bondades de las TIC's para manipular, almacenar y procesar la información biológica, este campo es crucial para países en vías de desarrollo como Cuba (1). Esta disciplina se interesa

por todos los aspectos relacionados con la adquisición, almacenamiento, procesamiento, distribución, análisis e interpretación de la información biológica, mediante la aplicación de técnicas y herramientas propias de las matemáticas, la química, la física, la biología y la informática, con el propósito de comprender el significado biológico de una gran variedad de datos (2).

Disímiles instituciones y centros cubanos hacen uso de aplicaciones bioinformáticas para dar solución a los problemas que presenta el país en el área de la salud, algunos ejemplos de estas entidades los constituyen el Centro de Ingeniería Genética y Biotecnología (CIGB), el Centro de Ingeniería Molecular (CIM) y el Instituto de Medicina Tropical Pedro Kourí (IPK).

Con el objetivo de ampliar los horizontes en esta ciencia y de incrementar el desarrollo tecnológico de la sociedad cubana, la Universidad de las Ciencias Informáticas posee una red de Centros Productivos en los que se desarrollan productos de *software*, tal es el caso del Centro de Gestión de Tecnologías de Datos (DATEC) que cuenta con una Plataforma de Servicios Bioinformáticos disponible a los especialistas del país a través de un portal web. Estos servicios permiten la manipulación, procesamiento y gestión de la información biológica, agrupados en diferentes módulos, siendo uno de los más importantes el de análisis de Secuencias.

Este módulo brinda una gama de servicios entre los que destacan el alineamiento de secuencias mediante varios programas, la traducción de secuencias de ácidos nucleicos a proteínas, visualización y edición de secuencias moleculares, entre otros. La cantidad de especialistas accediendo al sistema puede provocar una alta concurrencia de peticiones y traer consecuencias negativas en el correcto funcionamiento del mismo. El estado de los servidores, así como los recursos del sistema no están siendo actualmente controlados. Por tanto para consolidar y mantener el sistema en óptimas condiciones es necesario establecer indicadores de los datos procesados que a su vez permitan tomar acciones correctivas. Como necesidad de dar solución a la situación anteriormente expuesta, este trabajo tiene como objetivo desarrollar un sistema para el monitoreo y control del Módulo de análisis de secuencias de la Plataforma de Servicios Bioinformáticos.

## **DESARROLLO**

### **MATERIALES Y METODOS**

#### **Características de los sistemas de monitorización.**

Los sistemas de monitorización, teniendo en cuenta el lugar donde se ejecutan, pueden catalogarse como remotos o locales. El funcionamiento de los sistemas de monitorización remoto, exigen una alta velocidad de conexión para obtener un rendimiento óptimo cuando la distancia y la cantidad de equipos son considerables. Son muy usados por la comodidad con que reflejan sus datos y generalmente están orientados al control de redes. Se recomienda su uso cuando se cuenta con un servicio de red fiable, ya que dependen totalmente de esta, por lo que no garantizan el control total de los recursos del sistema permanentemente y su empleo puede ser muy costoso.

Por su parte, las aplicaciones locales son ligeras, escasamente requieren recursos del sistema operativo y aseguran el control del mismo mientras es usado. Son independientes de la conexión a una red y facilitan integrarse a esta. Se sugiere su empleo cuando es indispensable que un ordenador sea chequeado constantemente. El uso de estas herramientas pueden ser aplicadas con tecnología poco sofisticada (3).

El sistema a desarrollar posee ambas características, sistema de monitorización remota y agentes locales, ya que los servidores del módulo se encuentran distribuidos (Figura 1). El acceso se realizará a través de la red por medio de la Invocación Remota de Método (RMI, por sus siglas en inglés) y supervisará recursos del ordenador de aplicaciones, dígame conexión a la base de datos, procesos y parámetros físicos del hardware.

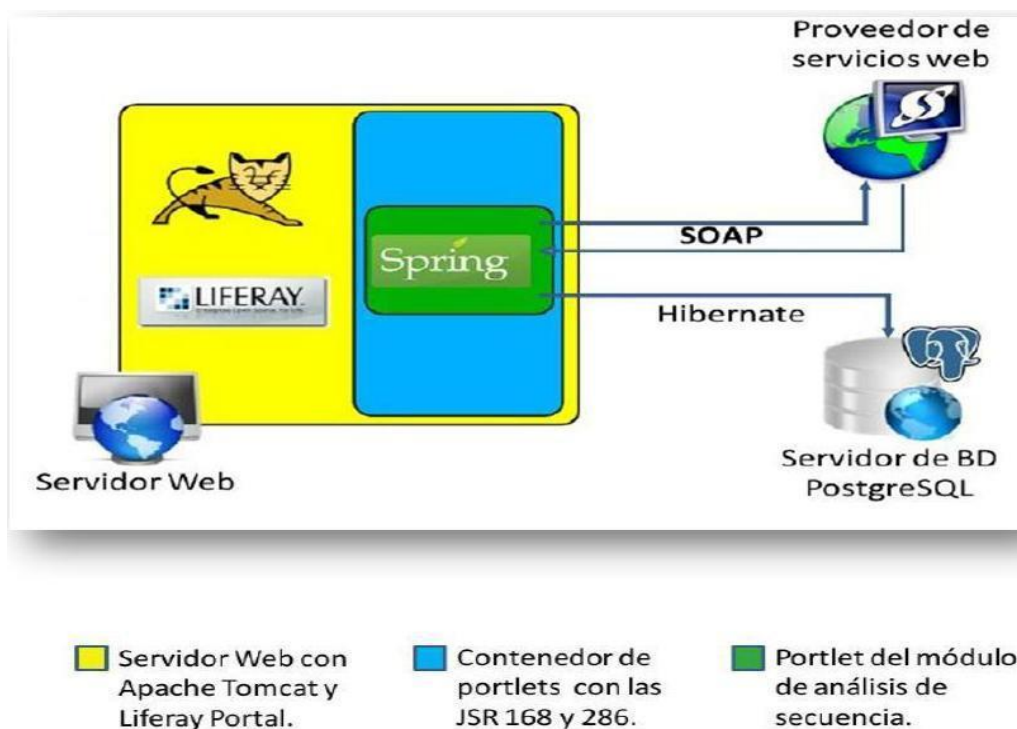


Figure 1: Componentes de la arquitectura del Módulo de análisis de secuencias.

## **Monitoreo de aplicaciones. Java Management Extensions**

Los sistemas de monitorización actuales se encuentran enmarcados en la detección de incidencias de un determinado recurso, dígame redes, bases de datos, servidor de aplicaciones o servicios web. Para el monitoreo de los componentes del Módulo de análisis de secuencias es necesario un sistema capaz de monitorizar servicios web, servidor web y servidor de bases de datos, es por ello que se determina desarrollar un sistema de monitorización y control utilizando Java Management Extensions (JMX, por sus siglas en inglés).

Utilizando JMX, se puede crear un entorno de gestión que es menos costoso y más flexible, en un corto período de tiempo. JMX es un marco añadido al lenguaje Java, que puede proporcionar una solución de gestión, permite encapsular todos sus recursos (hardware o software) con objetos Java y exponerlos en un entorno distribuido. Además, JMX proporciona un mecanismo para asignar fácilmente los protocolos de gestión existentes en sus propias estructuras de gestión. A continuación se mencionan algunas de las áreas de gestión definidas en las que JMX se desenvuelve:

- **Estado de la plataforma:** utilizando JMX, se pueden ajustar los recursos non-Java<sup>1</sup> y las interfaces de hardware con los objetos de Java para ajustarlos en un sistema de gestión JMX.
- **Rendimiento de aplicaciones y notificaciones:** controla fácilmente el sistema para los eventos críticos. Cuando un evento es notado, se emiten notificaciones a un proceso de escucha predefinido. Los oyentes pueden configurarse para enviar páginas, escribir mensajes de correo electrónico, o realizar cualquier otra operación.
- **Invocación de métodos remotos:** brinda soporte para establecer conexión mediante la invocación de métodos remotos (RMI, por sus siglas en inglés) desde un cliente para resolver problemas relacionados con la distribución de aplicaciones. RMI puede emplear para la comunicación la tecnología SSL (del inglés, Secure Sockets Layer), la cual asegura la privacidad y fiabilidad de las comunicaciones entre dos aplicaciones (4).

## **Términos esenciales de Java Management Extensions.**

---

<sup>1</sup> Dispositivos que pueden ser manipulados a través de interfaces Java, ejemplo: impresoras, módem, etc.

Los términos que se definen a continuación permitirán una mayor comprensión de la tecnología JMX. Además se explicará la arquitectura que presenta esta (5).

### **Recurso manejable**

Un recurso manejable es cualquier aplicación, dispositivo, o entidad existente que pueda ser accedida o está envuelta en Java. Es la entidad que estará expuesta para la gestión mediante JMX. Las aplicaciones pueden exponer componentes, APIs, o recursos adicionales para que JMX los pueda gestionar o monitorizar.

### **MBean**

Un Bean Gestionado o MBean (del inglés, Managed Bean) es una clase Java que cumple con las normas dictadas por la especificación JMX. Las aplicaciones crean instancias de clases Java que exponen interfaces de gestión para la manipulación y el acceso a los recursos manejables. Una interfaz MBean de gestión se compone de los atributos del MBean y las operaciones que se exponen para la gestión.

### **Servidor MBean**

Un servidor MBean es una clase Java que gestiona un grupo de MBeans. Este es el corazón de JMX, y actúa como un registro para buscar MBeans. El servidor de MBean expone la interfaz de gestión de cualquier MBean registrado, pero nunca expone el objeto de referencia.

### **Agente JMX**

Un agente JMX es un proceso que proporciona un conjunto de servicios para la gestión de MBeans. Los agentes JMX proveen servicios para la creación de relaciones MBean, cargar dinámicamente objetos de supervisión. Los agentes JMX pueden ser utilizados por una gran variedad de protocolos de gestión y herramientas diferentes.

### **Notificación**

Las notificaciones son objetos Java emitidos por MBeans y el servidor de MBean para encapsular los eventos, alertas o información. Otros MBeans u objetos de Java pueden inscribirse como oyentes para recibir notificaciones. El modelo de notificación de JMX es similar al modelo de eventos de Java.

### **Instrumentación**

La instrumentación es el proceso de exponer un recurso manejable utilizando un MBean (o conjunto de MBeans). La instrumentación de una aplicación puede llevarse a cabo junto con el desarrollo, o los desarrolladores pueden trabajar para crear MBeans que utilizan las API existentes de los sistemas activos (5).

### Arquitectura de Java Management Extensions.

La arquitectura JMX está diseñada para mejorar la flexibilidad y utilidad en un entorno de gestión, quedando dividida en tres niveles:

El nivel más bajo es la instrumentación de gestión. Cada recurso gestionable se describe en una interfaz que especifica los atributos que tiene, las operaciones a las que da soporte y las notificaciones que envía. Este recurso es un MBean.

El nivel medio es el agente de gestión. Cada proceso gestionado contiene un agente JMX que incluye un servidor de MBeans, este proporciona un registro y un punto de acceso para MBeans. Los clientes de gestión deben utilizar el servidor de MBeans para acceder a los objetos registrados.

El nivel más alto de la arquitectura es el nivel de servicios distribuidos y su rol es facilitar el acceso remoto a los agentes JMX. Esta tarea se consigue mediante conectores, que proporcionan una interfaz del lado del cliente, por ejemplo, un conector RMI o adaptadores de protocolo. A continuación se muestra la Figura 2 con las capas representadas.

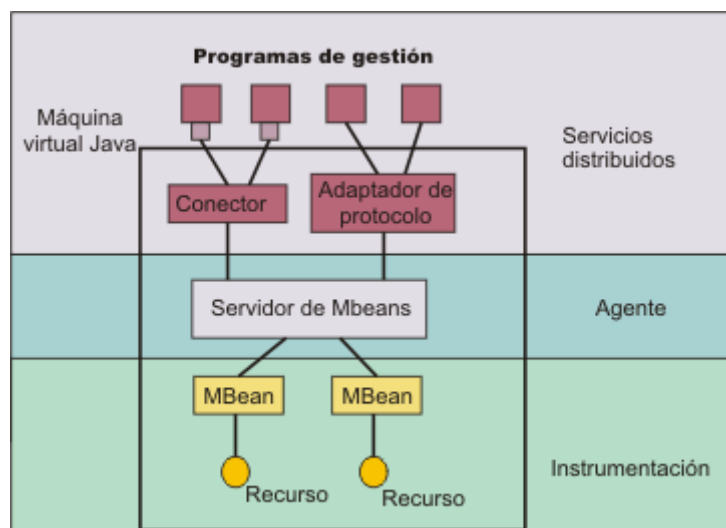


Figure 2: Arquitectura en tres capas de Java Management Extensions.



## RESULTADOS Y DISCUSIÓN

Los sistemas de monitorización se encuentran clasificados según el ámbito en que estos han sido aplicados, entre los que se encuentran, redes, bases de datos, servidores de aplicaciones y servicios web:

Uno de los sistemas ampliamente diseminados es Nagios, este es un sistema de monitorización de código abierto para redes, que vigila los equipos (hardware) y servicios (software) que se especifiquen, alertando cuando el comportamiento de los mismos no es el deseado. Entre sus características principales figuran el monitoreo de todos los componentes de la infraestructura de misión crítica – incluyendo aplicaciones, servicios, sistemas operativos, protocolos de red, métricas de sistemas e infraestructura de red (6). Nagios presenta algunas desventajas, ya que solo se encuentra disponible para sistemas operativos de la familia Unix, no existe actualmente una versión para Windows, por lo que no es multiplataforma. Consume recursos vinculados a la ejecución de chequeos predestinados a fallar y crea un gran número de notificaciones por fallos en el host y en cada uno de sus servicios, además la configuración tiene que ser a través de una terminal, esto provoca que las personas que no tengan un conocimiento en el Sistema Operativo Linux/Unix les resulte difícil la instalación.

Otro sistema de monitorización es Ipswitch WhatsUp, fabricado por la compañía Ipswitch, el cual permite monitorizar la red, incluyendo servidores, servicios, dispositivos, etc. Entre las versiones más importantes están WhatsUp Gold y WhatsUp® Professional 2006 Premium Edition. Monitorea servicios como: Hypertext Transfer Protocol (HTTP, por sus siglas en inglés), File Transfer Protocol (FTP, por sus siglas en inglés), Domain Name System (DNS, por sus siglas en inglés), entre otros. El funcionamiento de WhatsUp como servicio bajo Windows ofrece acceso vía web a su interfaz gráfica. Entre las principales características se encuentran la identificación y mapeo de todos los dispositivos de la red, envío de notificaciones, almacenamiento de información periódica sobre la red para generar reportes y monitoreo a cualquier hora y desde cualquier lugar (7).

Entre las herramientas encargadas de monitorizar bases de datos, se puede citar a Postgres Plus HQ Monitor. Su principal función es supervisar la disponibilidad, el rendimiento, la utilización y la infraestructura de las bases de datos. Monitoriza además el estado del servidor de bases de datos, los recursos que utiliza y las actividades que se realizan y las tablas por separadas. Presenta otras funcionalidades como: manejo de



recursos para supervisar servidores de aplicaciones y sistemas operativos, realiza monitoreo en tiempo real y envía alertas automáticas. Es una herramienta desarrollada por la EnterpriseDB, una compañía privada que proporciona soporte de clase empresarial para PostgreSQL (8).

Existen también herramientas de monitorización para propósitos generales como Hyperic HQ. Esta proporciona estadísticas de rendimiento sobre las aplicaciones y servidores de bases de datos, servidores web y servicios de red más comunes (9). Los elementos clave de la arquitectura son el servidor central, que proporciona una gestión centralizada y la persistencia de los datos, y el agente o demonio, que permite el seguimiento y control de cada componente distribuido. La definición de su arquitectura permite monitorizar aplicaciones mediante el uso de un modelo basado en agentes, ejecuta un agente (HQ Agent) en cada servidor que se desea administrar.

Luego de realizado un estudio de las características de los sistemas de monitorización mencionados anteriormente y teniendo en cuenta las características del módulo de Análisis de Secuencias, se hizo necesario la implementación de un sistema que cumpliera con los requisitos de dicho módulo y no conllevara a la realización de cambios en la arquitectura del mismo.

El sistema de monitorización y control del Módulo de Análisis de Secuencias brinda la posibilidad de monitorizar recursos de hardware tales como: servidores y conexiones de redes. De software: Base de datos del Módulo de análisis de secuencias, servicios web y servidor web, permitiendo mostrar el estado de estos en tiempo real y actualizados de forma automática cada cinco minutos. Brinda la opción de exportar esta información en Formato de Documento Portable (PDF, por sus siglas en inglés). Muestra las trazas de acciones de los usuarios donde el administrador puede filtrarlas por un criterio de búsqueda determinado. Posibilita la activación o desactivación de los servicios web brindados por dicho módulo, en caso de ocasionar congestión en las peticiones hechas al servidor. Ofrece las estadísticas de Hibernate en cuanto a: cantidad de conexiones a la base de datos, cantidad de entidades eliminadas, actualizadas e insertadas, así como el tiempo máximo de las consultas ejecutadas y la cantidad de sesiones abiertas y cerradas. Se utiliza un componente de interfaz de usuario, denominado *Portlet* para la interfaz del sistema, garantizando de esta manera, la independencia del sistema operativo.

## CONCLUSIONES

Una vez finalizada la investigación se puede concluir que: se obtuvo un sistema de monitoreo y control para el Módulo de análisis de secuencias, para ello:

- Se analizaron disímiles herramientas de monitoreo y se seleccionó JMX para implementar la solución propuesta, esto permitió crear un sistema de monitoreo y control propio, flexible y sin influir en la arquitectura del Módulo de análisis de secuencias. JMX proporcionó un entorno de gestión rápido y fácil, basado en el lenguaje Java y ofrece la posibilidad de cambiar el comportamiento de los componentes en tiempo de ejecución. El uso de componentes web de interfaz de usuarios, denominados *Portlets*, posibilitaron la portabilidad del sistema a través de diversos contenedores de *Portlets*.
- Fueron implementadas un conjunto de funcionalidades obteniendo un sistema capaz de monitorizar el estado de los componentes de la arquitectura del Módulo de análisis de secuencias. Brindando información acerca de los servidores, tales como: espacio en disco duro, memoria RAM, procesador, sistemas de archivos, entre otros. Deshabilita y habilita servicios web, gestiona las estadísticas del Hibernate, realiza salvadas de las bases de datos y brinda reportes de los datos monitorizados en formato PDF.

**REFERENCIAS**

1. *UCIBIOSOFT: A BIOINFORMATICS WEB PORTAL*. **Orlando Martínez Pérez, Alina Agramonte Delgado, Andry Daniel Díaz León**. La Habana : s.n., 2011.
2. *Primer Congreso Internacional de Bioinformática: reflexiones y perspectivas*. **Joan Valdivia A, Febles Rodríguez JP**. 2004.
3. *Selección e integración de una herramienta para monitorear la plataforma PlaTel*. **Miguel Makay Pérez, Raymon Tapia Castellanos**. La Habana : s.n., 2010.
4. **Benjamin G Sullins, Mark B Whipple**. *JMX in Action*. s.l. : Manning Publications Co., 2003. 1-930110-56-1.
5. Sullins, Benjamin G. y Whipple, Mark B. *JMX in Action*. s.l.: Manning Publications Co., 2003. 1-930110-56-1.
6. **Enterprises, Nagios. Nagios**. [En línea] 2009-20013. [Citado el: 20 de 11 de 2012.] <http://www.nagios.com/>.
7. Análisis del sistema para el monitoreo de la red UCI. Diseño de un prototipo no funcional. **Muro Fumero, Dorisbel y Verdecia López, Aimeé**. La Habana : s.n., 2007.
8. **García Mato, Rosa María**. *Diseño de bases de datos*. 1999. pág. 51.
9. **Guevara, P**. *Control de motores de corriente continua en Tiempo Real con capacidad de telecontrol y telemonitoreo. Tesis de Maestría en ciencias de la computación*. Instituto Politécnico Nacional, México : s.n., 1999.



[www.sociedadelainformacion.com](http://www.sociedadelainformacion.com)

Edita:



Director: José Ángel Ruiz Felipe  
Jefe de publicaciones: Antero Soria Luján  
D.L.: AB 293-2001  
ISSN: 1578-326x