

Arquitectura Software y Almacén de Datos altamente escalable para guardar información agronómica en la Nube de forma eficiente.

N. Pavón-Pulido¹, J.A. López-Riquelme², P. Soto², R. Morais³, A. Iborra²

¹ Departamento de Tecnología de la Información. Universidad de Huelva, Ctra. Huelva-Palos de la Fra. s/n, 21819, Palos de la Fra., Huelva, e-mail: npavon@dti.uhu.es

² DSIE - División de Sistemas e Ingeniería Electrónica. Universidad Politécnica de Cartagena, Campus Muralla del Mar s/n, 30202, Cartagena, e-mail: {jantonio.lopez,roque.torres,pencho.soto,andres.iborra}@upct.es

³ UTAD - Universidade de Trás-os-Montes e Alto Douro, Quinta de Prados, 5001-801 Vila Real, Portugal, e-mail: rmorais@utad.pt

Resumen

Este trabajo describe una arquitectura en la Nube que permite el despliegue de un completo sistema de gestión de cultivos en un contexto de Agricultura de Precisión, validado en un entorno real, que consta de una red inalámbrica de nodos sensores de coste económico reducido. Los módulos software se han desarrollado usando la Plataforma como Servicio de Google Cloud. El almacén de datos agronómicos se ha implementado utilizando el Datastore de Google, lo que asegura una alta escalabilidad y constituye un aspecto considerablemente novedoso, ya que la mayoría de las aplicaciones existentes usan bases de datos relacionales que dificultan la escalabilidad. Un análisis de las ventajas del sistema propuesto se presenta como conclusión del trabajo.

Palabras clave: Cloud Computing, Agricultura de Precisión, Redes Inalámbricas de Sensores

A Software architecture and a highly scalable Data Store for efficiently storing agronomic information in the Cloud.

Abstract

This work describes an architecture in the Cloud that allows a complete crop management system to be deployed and validated in a context of Precision Agriculture. The system comprises a wireless sensor network that uses economically inexpensive nodes. The software components have been developed by using the framework provided by the Google's Cloud Platform as a Service. The agronomic data store has been implemented by using the Google's Datastore; this ensures a high scalability degree and it could be considered as a significantly novel feature, since most existing applications use relational databases that make scalability difficult. The advantages of the proposed system are analyzed as a conclusion of this work.

Keywords: Cloud Computing, Precision Agriculture, Wireless Sensor Networks

1) Introducción y/o Justificación

El concepto de Agricultura de Precisión (AP), está fuertemente ligado a las Tecnologías de la Información y las Comunicaciones (TICs), ya que son éstas las que lo hacen viable,

posibilitando la adquisición de datos que proporcionan conocimiento acerca del estado del suelo, de la planta y del ambiente (Zhang et al., 2002). Los interesados usan este conocimiento para realizar predicciones y tomar decisiones sobre los cultivos. En la última década los investigadores han llevado a cabo interesantes avances en el contexto de las Redes Inalámbricas de Sensores (WSNs), aplicadas a la AP, permitiendo el despliegue de dispositivos de forma eficiente y compatible con las actividades agrícolas (Akyildiz et al., 2002); prestando especial atención a cuestiones esenciales: robustez, mejora de las comunicaciones, reducción del consumo energético y del coste económico. Sin embargo, el intercambio de información entre la WSN e Internet sigue llevándose a cabo mediante un modelo Cliente/Servidor tradicional, basado en el uso de bases de datos con esquema relacional alojadas en un servidor o servidores donde los módulos software realizan diversas tareas (almacenamiento, recuperación y agregación de datos, etc.). Además, se incluye el desarrollo de clientes que extraen información útil y la presentan a los interesados. No obstante, este modelo tradicional no está exento de inconvenientes: (i) El mantenimiento de los equipos y del software es difícil y necesita de personal especializado, lo que puede suponer un obstáculo que genere resistencia en los usuarios a la hora de incluir sistemas de AP en sus modelos de negocio. (ii) La inclusión de nuevos componentes software capaces de aplicar técnicas novedosas de análisis de datos (mediante técnicas estadísticas o de Inteligencia Artificial avanzadas), no es sencilla. (iii) El aumento del volumen de la información conlleva la necesidad de incluir técnicas apropiadas de escalado de las aplicaciones y de los almacenes de datos. (iv) La seguridad de la información es un aspecto complejo y crucial, ya que las decisiones deben estar basadas en información fiable. Una posible solución a los inconvenientes mencionados puede hallarse en el concepto de Cloud Computing o Computación en la Nube (Marotta et al., 2015). En este trabajo se muestran las ventajas de usar la Nube en el contexto de la AP, mediante el desarrollo de una nueva arquitectura hardware/software implementada usando las herramientas de la Plataforma como Servicio (PaaS) de Google Cloud, entre las que destaca el Datastore (un almacén de datos sin esquema altamente escalable). Su uso constituye un aspecto novedoso ya que la mayoría de las aplicaciones de AP usan bases de datos relacionales poco escalables. El artículo se estructura en varias secciones. En la sección II se describen los componentes hardware del sistema y las herramientas de Google Cloud usadas. En la sección III se detalla el funcionamiento de los diferentes módulos que componen la arquitectura, se describe el modelo de datos diseñado que sirve como base para desplegar las distintas entidades en el Datastore de Google y se presenta un análisis que compara la arquitectura propuesta con otros sistemas de AP tradicionales, resaltando las ventajas del uso de la Nube. Por último, en la sección IV se valora el trabajo realizado y se menciona la línea futura a seguir.

II) Material y Métodos

El sistema de AP propuesto incluye una WSN compuesta de nodos sensores (NS), que incluyen un Waspmote (López et al., 2015) y un módulo de radio XBee, y un nodo enrutador (NE) basado en el SBC (Single Board Computer), Raspberry Pi 2, con acceso a la Nube. Los NS permiten la conexión de los siguientes sensores: Estación meteorológica WS300 (con anemómetro, veleta y pluviómetro), conjunto de sensores de suelo (watermarks analógicos y digitales y sonda multi-paramétrica hydra probe II). La autonomía de los nodos sensores se garantiza mediante paneles solares de 3 W y baterías de polímero de litio de 3,7 V y 6600 mAh. El NE está conectado a la red eléctrica y posee

conexión a Internet. Dependiendo de la localización del código, los módulos de la arquitectura software se han implementado usando lenguajes de programación y herramientas de desarrollo distinta: (i) En los NS (Waspnotes) se ha usado Processing. (ii) El NE usa PHP, Apache y un Sistema Gestor de Base de Datos Relacional. (iii) Los módulos alojados en la Nube se han desarrollado usando la PaaS Google App Engine (GAE) y componen una aplicación que se ejecuta en la infraestructura de Google. (iv) El modelo de datos para almacenar la información agronómica en la Nube se ha implementado usando Java Data Objects (JDO 3.0). El acceso al Datastore de Google se realiza a través del “plugin” que permite guardar objetos de datos POJO (Plain Old Java Objects) en el Datastore. (v) La compartición de datos entre el NE, la Nube y las aplicaciones cliente que permiten a los usuarios acceder a los datos es posible gracias a una arquitectura basada en Endpoints de Google, que garantiza un intercambio de información y un método de autenticación seguros, ya que se basan en la utilización de los protocolos “https” y OAuth 2.0.

III) Resultados y Discusión

La Figura 1 (a) muestra la arquitectura hardware/software del sistema diseñado. Los NS desplegados intercambian información con el NE usando el protocolo IEEE 802.15.4. El NE almacena la información en una base de datos local. Cada cierto tiempo (configurable por el administrador del sistema), dicho nodo envía los nuevos datos desde el almacén local a la Nube utilizando un “Endpoint” a través de una conexión GPRS. Una vez que el sistema se asegura de que la nueva información se ha almacenado correctamente en el Datastore de Google, la base de datos local es liberada. A su vez, los usuarios del sistema pueden interactuar con el sistema de AP mediante clientes que también acceden a GAE a través de “Endpoints”. Además, es posible enviar información para reconfigurar ciertos parámetros de los nodos de la WSN, tales como, el tiempo entre envíos de información y la frecuencia de volcado de la base de datos local. Por otra parte, cuando un usuario realiza una petición de información, la aplicación GAE analiza si los datos se encuentran disponibles en el Datastore, si no es así (aún no se ha producido el último volcado de la base de datos local), realiza una petición de actualización de los datos al NE, que a su vez puede requerir información de los NS. Este comportamiento es configurable, es decir, se puede especificar si la aplicación GAE realiza la petición de los últimos datos o si obtiene, simplemente los últimos datos guardados en el Datastore. El comportamiento del sistema depende del rol del usuario. Algunos sólo están autorizados para descargar cierta información del cultivo y no pueden cambiar la configuración. El rol de administrador permite ejecutar un abanico mayor de acciones. Para identificar a los usuarios y el rol de cada uno, se usa un esquema de autenticación basado en el protocolo OAuth 2.0. Los usuarios son autenticados a través de su cuenta de correo electrónico de Google, lo que facilita enormemente el control de acceso.

IV) Conclusiones

En este trabajo se presenta un sistema de AP, para monitorización de parámetros relevantes de cultivos, basado en una arquitectura software que incluye módulos implementados usando GAE. Los datos agronómicos se guardan en el Datastore, así, tanto la aplicación como el sistema de almacenamiento son altamente escalables de manera automática. Este paradigma (esencial para construir el IoT o Internet de las

Cosas) está teniendo un fuerte crecimiento. De hecho, la fusión de la AP y el IoT puede constituir una nueva revolución en el ámbito de la Agricultura y de la Ganadería altamente positiva. La combinación AP-IoT va a permitir que los dispositivos desplegados en el campo pueden convertirse en objetos inteligentes integrados en la Nube de forma transparente, accesibles de un modo sencillo y seguro, en cualquier momento y desde cualquier lugar. Como la información reside en la Nube, el tamaño de los datos, así como su escalado (redistribución remota en los servidores de la Nube), se convierte en un problema menor y, además, surge la posibilidad de emplear nuevos métodos para obtener conocimiento mediante técnicas de Minería de Datos o Clustering aplicadas a Big Data, lo que permite ampliar fácilmente las aplicaciones del sistema propuesto (no sólo para monitorización, sino también para control automático de tareas agrícolas, entre otras).

Agradecimientos

Los autores quieren agradecer a los proyectos cDrone (ref. TIN2013-45920-R) y RIDEFRUT (ref. AGL2013-49047-C2-1-R) su apoyo para realizar este trabajo.

Bibliografía

Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E. 2002. Wireless sensor networks: a survey. *Computer Networks* 38, 393-422.

López, J.A., Navarro, H., Soto, F., Pavón, N., Suardíaz, J., Torres, R. 2015. GAIA2: A multifunctional wireless device for enhancing crop management, *Agricultural Water Management*, 151, 75-86.

Marotta, M.A., Roveda, L., Klafke, M.A., Zambenedetti, L., Rochol, J., Bonato, C. 2015. Managing mobile cloud computing considering objective and subjective perspectives, *Computer Networks*, 93, 531-542.

Zhang, N., Wang, M., Wang N. 2002. Precision agriculture-a worldwide overview. *Computers and Electronics in Agriculture* 36, 113-132.

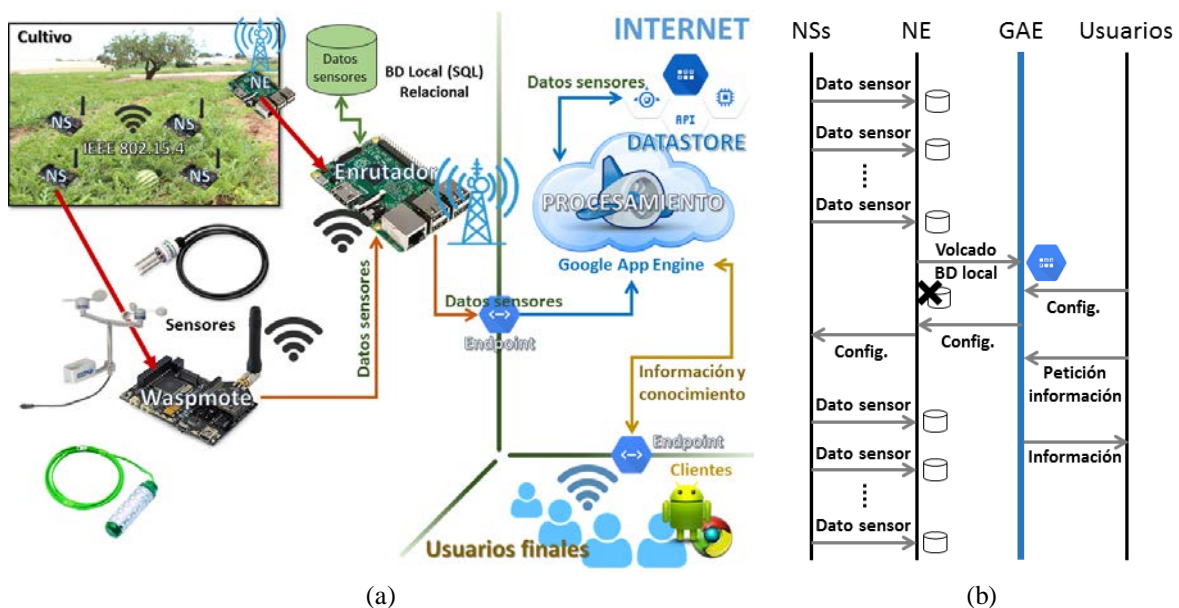


Figura 1. Arquitectura y funcionamiento del sistema de AP propuesto.