

A model for integrating cloud and fog computing for IoT applications

Osvaldo Marianetti¹, Ernesto Chediak¹, Joaquín Zorrero Maio¹, Enzo Manzur¹,
Nicolás Da Via¹

¹ Universidad de Mendoza. Facultad de Ingeniería. Mendoza. Argentina.
olmarianetti@gmail.com

Abstract. Although Fog computing solves some of the challenges faced by implementing some IoT applications or cyber-physical systems (CPS), Fog nodes may not be able to meet the performance, energy, and latency constraints of future IoT or CPS applications unless the Fog computing architecture is tailored to meet these application requirements. This adaptation is required at both the system and node levels for Fog computing.

This paper describes an integrated cloud and fog architecture with IoT that enables increased performance, energy efficiency, reduced latency, and scalability for future IoT applications. Fog nodes (e.g., edge servers, smart routers, base stations) receive computing process requests for sensed data from various IoT devices. To improve application performance, energy efficiency, and real-time responsiveness, the design proposes a reconfigurable Fog node architecture that analyzes application characteristics and reconfigures architectural resources to better meet peak load demands. In particular, the focus is on a reconfigurable hardware layer, in which FPGA devices are used to provide the hardware resilience features required by the applications of these systems.

Keywords: Fog Nodes, Reconfigurable architecture, IoT applications.

Un modelo de integración de computación en la nube y computación en la niebla para aplicaciones de IoT

Resumen. Aunque la computación en niebla resuelve algunos de los problemas que enfrenta la realización de algunas aplicaciones IoT o los sistemas ciberfísicos (CPS), los nodos de niebla pueden no ser capaces de cumplir con las restricciones de rendimiento, energía y latencia de las futuras aplicaciones IoT o CPS, a menos que la arquitectura de computación en niebla se adapte para satisfacer estos requisitos de aplicación. Esta adaptación es necesaria tanto a nivel del sistema como a nivel del nodo para la computación en niebla.

En este trabajo se describe una arquitectura integrada de nube y niebla con IoT que permite un aumento en el rendimiento, eficiencia energética, reducción de latencia y escalabilidad para futuras aplicaciones IoT. Los nodos de niebla (por

ejemplo, servidores de borde, routers inteligentes, estaciones base) reciben solicitudes de proceso de cómputo de datos sensados de varios dispositivos IoT. Para mejorar el rendimiento, la eficiencia energética y la capacidad de respuesta en tiempo real de las aplicaciones, el diseño plantea una arquitectura de nodo de niebla reconfigurable que analiza las características de las aplicaciones y reconfigura los recursos arquitectónicos para satisfacer mejor las demandas de carga máxima. En particular se enfoca una capa reconfigurable de hardware, en la cual se utilizan dispositivos FPGAs para aportar características de resiliencia al hardware que requieren las aplicaciones de estos sistemas.

Palabras clave: Nodos en la niebla, Arquitectura reconfigurable, Aplicaciones IoT.

1 La heterogeneidad de tecnologías

La oferta de plataformas y kits de desarrollo para aplicaciones IoT es muy amplia, ya que algunos proveedores y/o fabricantes ofrecen dispositivos específicos como simples sensores con las especificaciones de algún protocolo, módulos con sensores, actuadores, microcontroladores e interfaces de comunicaciones. Otros presentan soluciones integrales con todos los componentes que requiere la implementación de la arquitectura de un sistema de IoT. Un ejemplo de una solución del tipo integral que ofrecen distintos proveedores se presenta en la Figura 1, donde se abarca una amplia gama de posibilidades de implementación.

En ella se pueden observar los distintos componentes del sistema en cada nivel del mismo. En el nivel de Internet, esta presenta la utilización de tecnologías de computación en la nube (cloud). Esta nube puede ser un servicio prestado por proveedor o se puede acceder a otra nube pública o privada a la cual se accede mediante servicios REST. Actualmente la mayoría de las empresas como Google, Microsoft, Dropbox, LinkedIn, Twitter, etc. permiten a los programadores conectarse a sus servicios mediante servicios REST (REpresentational State Transfer). Estos servicios reciben las peticiones por el protocolo HTTP, el cual se usa para ver cualquier página en un navegador web. Los servicios REST permiten acceder y/o modificar la información mediante métodos HTTP, por lo cual se puede acceder a ellos mediante URLs. Por lo general regresan la información en formato JSON, aunque también pueden regresar archivos XML o csv. Al desarrollar sistemas móviles es recomendable utilizar servicios REST, los cuales manejan toda la lógica de una aplicación y puede ser consumida por aplicaciones de escritorio o sistemas Web, ya que los teléfonos celulares no tienen tanto poder de procesamiento como lo tiene un servidor.

En el nivel de Gateway, componente que permite el acceso a los recursos físicos de los distintos dispositivos sensores y actuadores (cada uno con su protocolo particular), también pueden encontrarse componentes Gateway implementados por terceras partes o por el mismo cliente. Unos dispositivos muy utilizados son microcomputadoras con los recursos de procesamiento y de módulos de entradas y salidas con las interfaces de buses y protocolos presentes en las aplicaciones de sistemas de IoT.

En este tipo de implementación genérica se observa la heterogeneidad de tecnologías presentes, lo cual hace necesario tener en cuenta la compatibilidad y la flexibilidad requerida, dados los mecanismos de resiliencia y seguridad implementados por distintos fabricantes. Figura 1.

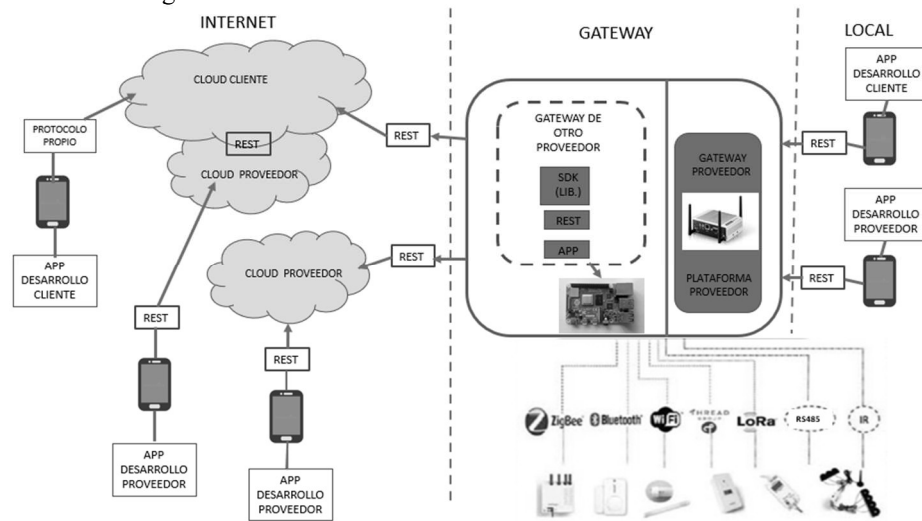


Fig. 1. Solución del tipo integral que ofrecen distintos proveedores.

Integración de sistemas de IoT y de computación en la nube.

En los últimos años, el despliegue generalizado de la computación en la nube impactó tremendamente a la industria de Tecnologías de la Información (TI), debido a su potencia de cálculo, almacenamiento y procesamiento. Sin embargo, la integración de Internet of the Things (IoT) y Cloud Computing ha sido un tema vigente para muchos investigadores (también conocido como CloudIoT), en relación con su complementariedad.

Un gran número de investigadores han seguido y desarrollado esta idea, plasmada en una amplia variedad de artículos científicos, libros y congresos. En todos ellos persiste la visión de integrar la inteligencia en los objetos que nos rodean. Esto se logra proporcionando a los objetos capacidades de detección, acción, almacenamiento y procesamiento, pero sobre todo proporcionándoles interconectividad a través de Internet, todo esto con el fin de brindar servicios a diferentes usuarios. Es precisamente en los usuarios, donde radica la principal diferencia del IoT con respecto a Internet tradicional, ya que mientras que en los usuarios tradicionales son personas conectadas a Internet a través de un PC, portátil o algún dispositivo móvil (smartphone o Tablet), en el IoT los usuarios pueden ser otros objetos o "cosas inteligentes" que requieran algún servicio o interacción. Esta interacción puede tener lugar incluso con fuentes de información adicionales, como las redes sociales. Además, el concepto de IoT implica algunas actividades muy importantes que con frecuencia se olvidan: el análisis, el almacenamiento y la utilización de los datos recopilados por diferentes dispositivos/objetos. En este contexto, IoT representa la evolución importante respecto de la Internet tradicional hacia una red interconectada de "objetos inteligentes" que no sólo recopilan datos

del entorno (detección) e interactúan con el mundo físico (actuación, comando y control), sino que también utilizan Internet para proporcionar servicios de transferencia de información, análisis, aplicaciones y comunicaciones.

En resumen, en IoT, los dispositivos crean datos que se envían a la aplicación principal para ser distribuidos, procesados y utilizados. Dependiendo del dispositivo, la red y las restricciones de consumo de energía, los datos pueden enviarse en tiempo real o en lotes en cualquier momento. Debido a la heterogeneidad de las fuentes de datos, la naturaleza de los datos generados y la heterogeneidad de los servidores de procesamiento de datos, un sistema de IoT también puede verse como una red computacional que consta de una gran cantidad de dispositivos que generan datos para su procesamiento y una gran cantidad de recursos capaces de procesar dichos datos. Para los dispositivos y sensores inteligentes, cada evento que perciben o que se registra puede y creará datos. Estos datos luego se pueden enviar a través de la red a la aplicación central. En este punto debe recordarse que se requiere un estándar para aplicar a los datos y cómo estos se enviarán a través de la red. Ya se han citado los protocolos más utilizados como el transporte de telemetría de cola de mensajes (MQTT), el protocolo de transferencia de hipertexto (HTTP) y el protocolo de aplicación restringida (CoAP) que son los protocolos estándares más comunes utilizados.

Por otra parte, durante años, la computación en la nube fue la tecnología base más utilizada para muchas arquitecturas empresariales y las empresas decidieron trasladar todos sus datos, cálculos, procesamiento, etc., desde la infraestructura "local" a la propia nube. La nube parece ofrecer espacio de almacenamiento infinito y escalabilidad para la computación sin ninguna preocupación desde el punto de vista de la empresa, que puede configurar todas las funciones para que cambien automáticamente. El resultado podría ser menos tiempo para gestionar infraestructuras a demanda y menos dinero para invertir. La computación en la nube permite un acceso conveniente y bajo demanda a un conjunto compartido de recursos informáticos configurables, como redes, servidores, almacenamiento, aplicaciones y servicios, que pueden aprovisionarse y liberarse rápidamente con un mínimo esfuerzo de gestión o interacción con el proveedor de servicios. Sin embargo, para muchas operaciones, un modelo basado únicamente en la nube no es necesariamente la opción perfecta. El costo del almacenamiento en la nube está cayendo, pero transmitir y almacenar cantidades masivas de datos en la nube para su análisis rápidamente se vuelve prohibitivamente costoso. Por lo general, todos los datos se transmiten con su fidelidad actual, sin posibilidad de seleccionar qué datos son de mayor valor comercial y qué datos pueden ignorarse. Las plataformas de IoT basadas en la nube también requieren una conexión de red constante, lo que las hace menos ideales para empresas con operaciones remotas, que no pueden permitirse el lujo de detener sus operaciones cuando su conexión se corta. Como consecuencia, se ha propuesto un nuevo conjunto de paradigmas para el manejo de datos de IoT.

En concordancia con lo citado anteriormente en el modelo de tres capas para la arquitectura de los sistemas de IoT, se presentan tres enfoques importantes diferentes: la computación en la nube, la computación en niebla y la computación en el borde, con sus principales características, ventajas, desventajas y diferencias entre sí. Los tres enfoques se basan en la necesidad de contar con recursos de computación distribuidos.

2 Arquitectura de nodo de computación en la niebla reconfigurable

Aunque la computación en niebla alivia algunos de los problemas que enfrenta la realización de algunas aplicaciones IoT o los sistemas ciberfísicos (CPS), los nodos de niebla pueden no ser capaces de cumplir con las restricciones de rendimiento, energía y latencia de las futuras aplicaciones IoT o CPS, a menos que la arquitectura de computación en niebla se adapte para satisfacer estos requisitos de aplicación. Esta adaptación es necesaria tanto a nivel del sistema como a nivel del nodo para la computación en niebla.

En esta sección se presenta una arquitectura integrada de nube y niebla para aplicaciones Internet de las Cosas. En él, sus autores, Munir A, Kansakar P y Khan S, en “Integrated Fog Cloud IoT Architectural Paradigm for Future Internet of Things”, proponen un nuevo paradigma arquitectónico integrado de nube y niebla con IoT que permite un aumento en el rendimiento, eficiencia energética, reducción de latencia, tiempo de respuesta más rápido, escalabilidad y mejor precisión localizada para futuras aplicaciones IoT. Los nodos de niebla (por ejemplo, servidores de borde, routers inteligentes, estaciones base) reciben solicitudes de descarga de computación y datos sensados de varios dispositivos IoT. Para mejorar el rendimiento, la eficiencia energética y la capacidad de respuesta en tiempo real de las aplicaciones, el diseño plantea una arquitectura de nodo de niebla reconfigurable en capas, que analiza las características de las aplicaciones y reconfigura los recursos arquitectónicos para satisfacer mejor las demandas de carga máxima. Las capas de la arquitectura propuesta del nodo de niebla incluyen capa de aplicación, capa de análisis, capa de virtualización, capa de reconfiguración y capa de hardware. La arquitectura en capas facilita la abstracción e implementación para el paradigma de computación en niebla que es distribuido por naturaleza y donde están involucrados múltiples proveedores (por ejemplo, aplicaciones, servicios, proveedores de datos y contenido).

El análisis de la carga de trabajo en un servidor/nube revela que diferentes aplicaciones tienen diferentes horas pico de carga en diferentes momentos. Esta variación temporal de las cargas de trabajo máximas de las aplicaciones permite desarrollar una arquitectura multinúcleo adaptable y reconfigurable para el servidor de borde que puede adaptarse de acuerdo con la carga de la aplicación que se ejecuta en un momento dado para mantener mejor la velocidad de datos proyectada, el volumen de datos y el tiempo real. requisitos de las aplicaciones IoT o CPS. La arquitectura de servidor de borde propuesta consta de varias capas: capa de aplicación, capa de análisis, capa de virtualización, capa de reconfiguración y capa de hardware, como se muestra en la Figura 2. Capa de aplicación: esta es la capa superior de la arquitectura del servidor de aplicaciones de borde. Esta capa consta de servicios de plataforma de aplicaciones que el servidor puede proporcionar a varias aplicaciones alojadas en el servidor perimetral. Los servicios de plataforma de aplicaciones proporcionados por el servidor incluyen servicios de descarga de cálculos, agregación de contenido, bases de datos y respaldo, información de red, etc. Cuando un dispositivo IoT conectado al servidor de borde solicita que se ejecute una aplicación en particular, se crea un entorno de VM. creado para la aplicación. Esto significa que cada aplicación tiene su propia instancia del entorno

de VM que se ejecuta en la capa de aplicación. La capa de aplicación para el servidor perimetral actúa como un proveedor de PaaS, es decir, la capa de aplicación abstrae toda la arquitectura del servidor para proporcionar una plataforma estándar para los desarrolladores de aplicaciones de IoT.

Capa de análisis: La capa de análisis consta de tres módulos: el módulo de estadísticas de uso de servicios de plataforma, el módulo de aprendizaje automático y el módulo de administrador de energía. El módulo de estadísticas de uso de servicios de plataforma analiza el uso de los servicios de aplicación proporcionados en la capa de aplicación. El módulo de aprendizaje automático toma como entrada la información del tipo y volumen de las solicitudes de servicio y analiza esta información para predecir los requisitos de recursos de hardware, que pueden ser aprovechados por la capa reconfigurable. El módulo de administración de energía analiza el tipo y la información del volumen de las solicitudes de servicio para determinar utilización del servidor. Basado en la utilización servidor, el módulo de administración de energía ajusta dinámicamente el voltaje operativo y la frecuencia de los componentes de hardware del servidor.

Capa de virtualización: la capa de virtualización abstrae los recursos de hardware subyacentes (los recursos de hardware pueden ser de diferentes proveedores) para proporcionar una interfaz común para los servicios de aplicaciones. La capa de virtualización actúa como un proveedor de IaaS, es decir, la capa de virtualización abstrae los recursos de hardware del servidor de la capa de servicios de aplicaciones y, por tanto, de las aplicaciones que se ejecutan en el servidor de borde.

Capa Reconfigurable: La capa reconfigurable consta de un administrador de reconfiguración y un conjunto de módulos reconfigurables. El administrador de reconfiguración en la capa reconfigurable recibe información del módulo de aprendizaje automático en la capa de análisis y reconfigura los recursos arquitectónicos para cumplir mejor con los requisitos de la aplicación de carga de trabajo máxima en un momento determinado. La capacidad de adaptar el hardware del servidor a los requisitos cambiantes de la carga de trabajo es un concepto incipiente.

Capa de hardware: la capa de hardware consta de un administrador de escalado dinámico de voltaje y frecuencia (DVFS), controladores de almacenamiento y recursos de red. El módulo DVFS adquiere información del módulo de administrador de energía en la capa de análisis. El módulo DVFS ajusta el voltaje operativo y la frecuencia de varios componentes de hardware del servidor perimetral (por ejemplo, núcleo del procesador, memoria y periféricos) según las demandas de la carga de trabajo. Los controladores de almacenamiento y las unidades de almacenamiento se utilizan para servicios de bases de datos y servicios de respaldo. El módulo de red gestiona la conectividad entre el servidor y los dispositivos IoT, y entre el servidor de borde y la nube.

De la descripción de este modelo de infiere que la computación en la niebla proporciona varias ventajas sobre la computación en la nube para aplicaciones que requieren un procesamiento más rápido con latencia y fluctuaciones de retardo reducidas, capacidad de respuesta en tiempo real, soporte de movilidad y personalización basada en la ubicación. Sin embargo, la computación en la niebla no reemplaza a la computación en la nube, ya que la computación en la nube seguirá siendo viable para trabajos de procesamiento por lotes de alto nivel que son muy frecuentes en los mundos empresarial y científico. La sinergia de la computación en la nube y la niebla ayudará a realizar

futuras aplicaciones de IoT y CPS. Y también, como ya se manifestó, la importancia de la posibilidad de la reconfiguración de los nodos fogs, lo que permite incorporar el concepto de resiliencia a esta implementación.

En los párrafos anteriores se hace referencia a la reconfiguración del hardware del servidor de borde, en particular las instancias de virtualización de la infraestructura requerida para distintas aplicaciones y el uso de aceleradores de hardware. En particular un administrador de escalado dinámico de voltaje y frecuencia (DVFS) en nodos de red para soportar la computación en la niebla.

A continuación se complementa este modelo, ya que se avanza en las posibilidades de reconfiguración de mayores especificaciones y requerimientos de los nodos Fogs, incluso también es aplicable a un nodo de borde. Esta solución se propone a partir de las posibilidades que caracterizan a los dispositivos lógicos programables como las FPGAs.

La tecnología de lógica reconfigurable es un prestador de soluciones eficientes, escalables y sostenibles para estos escenarios, donde el creciente número de sensores que se implementan en hogares inteligentes, fábricas y ciudades inteligentes plantea a los desarrolladores un abordaje acorde a estos sistemas. Por un lado, las nuevas aplicaciones de IoT requieren una conectividad mejorada para comunicarse de manera eficiente a través de una amplia gama de interfaces y protocolos nuevos y heredados. Por otro lado, a medida que los usuarios exijan mayores niveles de inteligencia, los diseñadores necesitarán más recursos computacionales para el procesamiento en tiempo real de los datos cercanos a los sensores. En un número cada vez mayor de aplicaciones, se hace necesario una solución programable que combine E/S flexibles, una arquitectura configurable, operación de energía baja en un factor de forma pequeño y precio para producción de alto volumen. Figura 2.

Esta solución, puede considerarse como la de un componente que ofrezca los recursos lógicos de un coprocesador de alto rendimiento capaz de computación altamente paralela mientras agrega simultáneamente altos niveles de conectividad y soporte para una amplia gama de protocolos y estándares de E/S. Finalmente, se requiere una solución programable que esté lista para producción para los sistemas actuales y que sea compatible con módulos IP (propiedades intelectuales), diseños de referencia, placas de desarrollo y herramientas de software necesarias para llegar al mercado más rápido.

Los fabricantes de FPGAs ofrecen chips con funciones adecuadas para abordar estas necesidades. Están diseñados específicamente para permitir a los desarrolladores crear rápidamente soluciones energéticamente eficientes que satisfagan las demandas de procesamiento de los productos industriales y de Internet de las cosas.

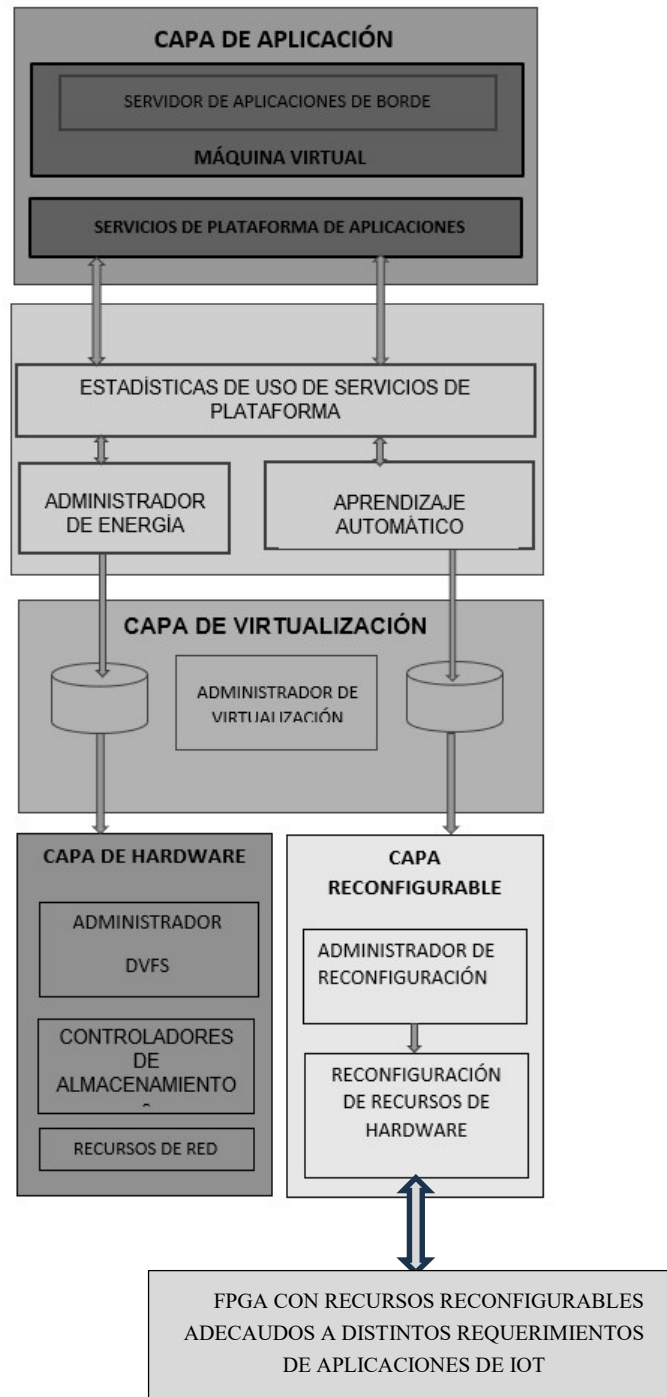


Fig. 2. Modelo con hardware reconfigurable con tecnologías de FPGAs.

3 Desarrollo de un prototipo de hardware de nodo de computación en la niebla reconfigurable

Se desarrolló un sistema embebido basado en soft_processors. Se diseñó una arquitectura de un sistema embebido basada en este tipo de procesadores. Esta arquitectura presenta la misma funcionalidad de los sistemas comerciales con componentes de seguridad en su diseño, pero incorpora las características de la naturaleza del hardware reconfigurable, para que el sistema sea más resiliente. En la arquitectura basada en FPGA sus unidades funcionales (memorias, puertos, controladores, temporizadores, etc.) son reconfigurables y adaptables a nuevos requerimientos, incluso en forma remota. Incorpora un módulo de seguridad (encriptación). El único componente externo al sistema programa en el chip, es el módulo de comunicaciones (protocolo WI FI). Figura 3.

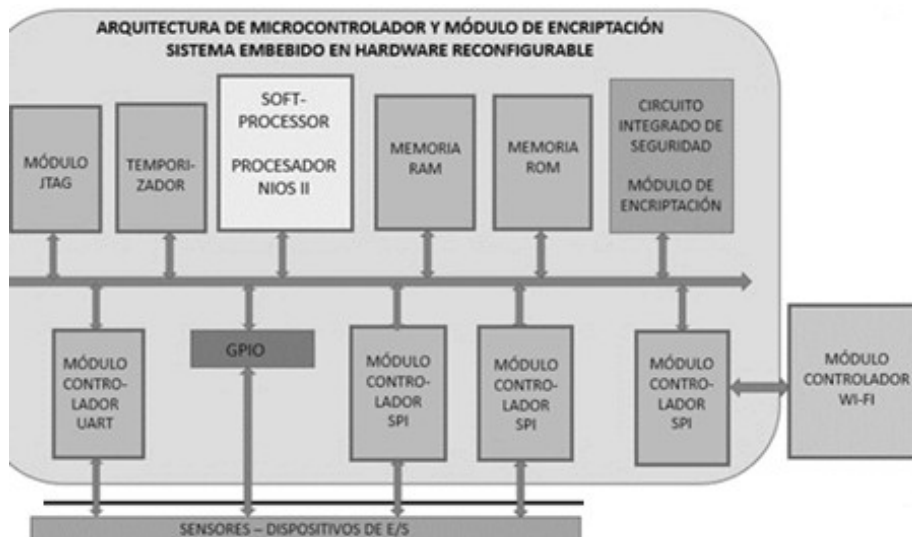


Fig. 3. Implementación de un hardware con las unidades funcionales que requiere un nodo de una capa de borde en arquitectura de computación en la niebla. Arquitectura basada en FPGA.

Respecto de las herramientas de software para el desarrollo con FPGA, se utilizaron: Quartus II (versiones 13.1 web edition y Quartus Prime Lite Edition 17.0) La herramienta QSYS de estos entornos para la generación del SOPC (sistema programable en el chip) y el entorno NIOS II software build tool for Eclipse para la programación del soft_processor NIOS II/e.

Se han utilizado todos dispositivos y herramientas disponibles en el medio y el contexto en el cual se desarrolló el prototipo.

4 Conclusión

Este prototipo corresponde a un modelo genérico de sistema embebido, en base a un dispositivo de tecnología Field Programmable Logic Array (FPGA), cuya arquitectura esté optimizada para resolver alguno de los distintos requerimientos de resiliencia y seguridad que presentan los dispositivos de IoT que permite avanzar en el desarrollo de arquitecturas de sistemas programables en un chip (SOPC) con dispositivos reconfigurables, FPGAs, optimizadas para operar como nodo de una WSN y también en aplicaciones de sistemas de IoT que pueda actuar como concentradores o nodos de la capa fog y/o de borde, utilizando tecnologías y herramientas de disponibles y accesibles en el contexto local.

Este modelo se adapta a requerimientos de una mejora en la implementación segura de sistemas de IoT e IoE, teniendo en cuenta las vulnerabilidades que se presentan en las distintas capas de estos sistemas y en la reducción de la superficie de ataque, permite cumplir determinadas especificaciones de seguridad desde la fase de diseño, como así también, condiciones de flexibilidad para adaptarse a distintas situaciones de la integridad y la disponibilidad de los datos del Sistema.

Referencias

1. Asemani M, Abdollahei F, Jabbari F (2019). "Understanding IoT platforms: towards a comprehensive definition and main characteristic description". In: 2019 5th international conference on web research (ICWR) IEEE. pp. 172–177.
2. Bonomi F, Milito R, Natarajan P, Zhu J (2014). "Fog computing: a platform for internet of things and analytics. Big data and Internet of Things: a roadmap for smart environments". Springer, Berlin, pp 169–186.
3. Mahmud R, Kotagiri R, Buyya R (2018). "Fog computing: a taxonomy, survey and future directions. Internet of everything". Springer, Singapore, pp 103–130.
4. Chiang M, Zhang T (2016). "Fog and IoT: an overview of research opportunities". IEEE Internet Things J 3(6):854–864.
5. Albouq SS, Abi Sen AA, Namoun A, Bahboub NM, Alkhodre AB, Alshanqiti A (2020). "A double obfuscation approach for protecting the privacy of IoT location based applications". IEEE Access 8:129415–129431. author, F. (2016). Article title. Journal 2(5), 99–110.
6. A Lattice Semiconductor White Paper IoT Sensor Connectivity and Processing with Ultra-Low Power, Small Form-Factor FPGAs. 2018.
7. Argyrios Sideris, Theodora Sanida, Minas Dasygenis. "Hardware Acceleration of the AES Algorithm using Nios-II Processor". Panhellenic Conference on Electronics & Telecommunications (PACET). 2019.
8. <https://www.intel.com/content/www/us/en/products/docs/processors/xeon/platform-firmware-resilience.html>.
9. Lattice to Speak on Cyber Resiliency and Supply Chain Security at 2021 FPGA. Conference June 28, 2021/<https://ir.latticesemi.com/node/20301/pdf>.