

Procesamiento paralelo sobre arquitecturas serverless para tratamiento de datos provenientes del IoT

Nelson Rodríguez¹, María Murazzo¹, Diego Medel¹, Lorena Parra¹, Ana Laura Molina¹, Federico Sánchez¹, Adriana Martín¹, Fabiana Piccoli², Zamira Ale³, Lucas Vargas³

¹ Departamento e Instituto de Informática - F.C.E.F. y N. - U.N.S.J.

² Departamento de Informática - F.C.F.M. y N – UNSL

⁶ Alumno Avanzado Licenciatura en Sistemas de Información y Cs. de la Computación - F.C.E.F. y N. - U.N.S.J.

Complejo Islas Malvinas. Cereceto y Meglioli. 5400. Rivadavia. San Juan, 0264 4234129

nelson@iinfo.unsj.edu.ar, marite@unsj-cuim.edu.ar, vdiego.unsj@hotmail.com, lorenaparra152@yahoo.com.ar, almm95@gmail.com, fgsanchez@unsj-cuim.edu.ar, adrianamartin1@gmail.com, mpiccoli@unsl.edu.ar, zamiranahirale@gmail.com, lucasgvargas96@gmail.com

Resumen

Según un estudio realizado por la multinacional Ericsson, para 2022 existirán más de 18 mil millones de dispositivos IoT comunicados entre sí. Todos estos dispositivos generan una gran cantidad de datos que deberán ser procesados en el menor tiempo posible, con el objeto de poder realizar una toma de decisiones óptima. En este sentido, el cloud parece ser la solución obvia, pero hay casos en los que la latencia que viene con la transmisión de datos de un lado a otro podría ser una limitación. En determinadas situaciones, tampoco es factible esperar una conexión a Internet constante y confiable, ya sea porque podría no ser viable económicamente o porque puede que no sea posible desde el punto de vista de la infraestructura.

Con el objeto de solucionar estos problemas, se propone trabajar con el paradigma serverless, en articular con FaaS usando lenguajes de programación paralelos que permitan reducir el tiempo del procesamiento de los datos obtenidos desde los dispositivos de IoT.

Palabras Claves: IoT, Serverless Computing, FaaS, Paralelismo, Métricas.

Contexto

El presente trabajo se encuadra dentro del área de I/D Procesamiento Distribuido y Paralelo y es una de las líneas de investigación del proyecto: Computación Serverless para el tratamiento de datos provenientes de dispositivos de IoT, cuya propuesta ha sido aprobada y está en desarrollo para el período 2020-2021. Asimismo el grupo de investigadores viene trabajando en proyectos relacionados con la computación distribuida y de alta performance desde hace más de 20 años. Como continuación del proyecto anterior: Orquestación de Servicios para la Continuidad Edge al Cloud, se sigue trabajando con investigadores de otras universidades, lo cual favorece notablemente a todos las instituciones participantes.

Introducción

IoT [1] es una tecnología creada para optimizar procesos. Esta tecnología, que busca interconectar dispositivos a través de internet para facilitar y acelerar tareas, está avanzando a pasos agigantados. Según un estudio realizado por la multinacional Ericsson, para 2022 existirán más de 18 mil millones de dispositivos IoT comunicados entre sí. Pero este crecimiento de dispositivos inteligentes trae aparejado la generación de grandes cantidades de datos, los cuales no pueden ser tratados y procesados con los paradigmas de programación tradicionales [2]. Es por ello que es necesario aplicar algoritmos que permitan aprovechar la escalabilidad de recursos de cómputo y procesamiento de datos [3] [4].

En este sentido se plantea como solución al procesamiento de datos provenientes del IoT, técnicas de computación de alta prestaciones (HPC) con el fin de aumentar la velocidad de procesamiento. HPC es la evolución de los sistemas de cómputo convencional, los cuales permiten realizar operaciones de cómputo intensivo y mejorar la velocidad de procesamiento; involucrando diferentes tecnologías tal como los sistemas distribuidos y los sistemas paralelos.

Todos estos entornos son ideales para resolver aplicaciones científicas, computacionalmente costosas con manejo de grandes cantidades de datos, a fin de lograr resultados en menor tiempo. Dadas estas características, estas arquitecturas son candidatas ideales para procesar datos provenientes del IoT; puesto que estos son generados a una elevada tasa y requieren un tratamiento en tiempo real [5].

Si bien, las arquitecturas HPC han evolucionado en pos de obtener mejores tiempos de respuesta para las aplicaciones, presentan el inconveniente del escalado, tanto vertical como horizontal, de recursos de cómputo. Es por ello que una alternativa es el cloud, el cual permite contar con una cantidad de recursos computacionales virtualmente infinitos, administrados por terceros y accedidos bajo demanda pagando por el uso,

de esta manera es posible lograr calidad de servicio (QoS) garantizado por parte del proveedor de la infraestructura cloud [6] [7].

Cloud Computing se ha convertido en una tecnología centrada en ofrecer cómputo bajo demanda como cualquier otro servicio. Uno de los aspectos clave que afecta la performance de las aplicaciones al montarlas sobre arquitectura cloud, es la capacidad de virtualización de cualquier recurso (bases de datos, red, procesador, etc.) y ofrecerlo como un servicio (AaaS, Anything as a Service). Esto puede tener un impacto negativo en la ejecución de aplicaciones de cálculo intensivo, debido a que los ambientes cloud deben virtualizar la plataforma en la cual se ejecuta y esto conlleva un overhead de a contextualización de las máquinas virtuales [8] [9] [10].

El principal inconveniente que existe con migrar las aplicaciones al cloud, es que es responsabilidad de la organización mantener funcionando de forma correcta la infraestructura (IaaS) que se necesite para el despliegue de las aplicaciones. Esta responsabilidad lleva a cargar costos sobre el presupuesto para el mantenimiento y soporte de la infraestructura.

En este sentido, la aparición del Serverless Computing [11] [12] supone una revolución en el desarrollo de servicios web. Los desarrolladores ya no se tienen que preocupar por el aprovisionamiento y escalado de la infraestructura, por lo que se pueden centrar en la lógica de sus aplicaciones. Serverless Computing trata de abstraer de la gestión de servidores (aprovisionamiento, configuración, escalado, etc.) para que los usuarios, en este caso desarrolladores, puedan enfocarse en la lógica de sus aplicaciones. Este paradigma se relaciona normalmente con el modelo FaaS [13], que consiste en la definición de funciones en un determinado lenguaje de programación que son ejecutadas sobre un proveedor Cloud. Para gestionar este tipo de servicios se utilizan tecnologías de contenedores software, que son los entornos sobre los cuales se ejecutan las funciones, si bien su gestión suele quedar oculta a los desarrolladores.

FaaS Paralelas

Con FaaS se llega a la máxima expresión de abstracción pues no se despliegan microservicios si no funciones que se ejecutan bajo eventos. Es decir, se basa en una arquitectura bajo eventos donde se abstrae completamente de toda la infraestructura subyacente. Se pueden hacer configuraciones básicas, pero olvidándose de la gestión del escalado, disponibilidad, despliegue y configuraciones de recursos. Aquí es donde llega el concepto de arquitectura serverless a tener una capa de abstracción que donde no necesita saber las características de los recursos subyacentes. Esto permite que sea un servicio con auto-escalado completo y solo se pague por el consumo real, ya que no se reservan recursos de computación. Además, las funciones solo se ejecutan cuando ha habido un evento y, en caso contrario, se mantienen “apagadas” sin consumir recursos. Otra característica para tener en cuenta es que deben ser ejecuciones en tiempos mínimos, entre 5 y 10 minutos, ya que las funciones deberán consumir recursos el mínimo tiempo posible. En caso contrario es necesario dividir esa función en funciones más pequeñas para distribuir los procesos.

En función de estas características se puede ver que en FaaS la lógica es, desarrollada dentro de un flujo de trabajo clásico que automáticamente es desplegada en procesos sin estado que son capaces de responder a distintos eventos que se produzcan en la infraestructura, con tiempo de vida limitado y completamente gestionada por el proveedor.

Hay diversas razones para desplegar IoT serverless entre las que se incluyen [14]:

- Al construir un backend basado en Cloud es difícil de monitorear la escalabilidad y disponibilidad. Las redes IoT suelen estar compuestas de millones de nodos, a través de distintas zonas geográficas, utilizando una arquitectura serverless permite eliminar toda la gestión sobre los proveedores Cloud.
- Las redes IoT suelen tener variaciones de carga donde a ciertas horas del día o

debido a determinados eventos se puede desencadenar que se active un gran número de dispositivos. Debido a que los servicios IoT suelen ejecutar un gran número de dispositivos inactivos, estos están esperando a ser activados y, por lo tanto, tener todos los procesos con recursos reservados no es óptimo y causará sobrecostos por infrautilización. Construir estos servicios con una arquitectura serverless permite pagar por lo que realmente consumen.

- El software que ejecuta y habilita los servicios de IoT generalmente comienza de manera simple, pero se van agregando características continuamente. La adopción de metodologías serverless permite a las organizaciones centrarse en su valor comercial principal, agilizar los cambios y el time to market en vez de la orquestación y el mantenimiento.

En base a esto se puede observar como FaaS encaja perfectamente con una solución paralela para procesar datos del IoT [15].

Líneas de Investigación, Desarrollo e Innovación

Este trabajo presenta una línea de investigación dentro del proyecto que comienza a desarrollarse este año. Si bien se ha trabajado anteriormente con paradigmas de HPC para procesamiento de datos masivos provenientes de aplicaciones móviles y de dispositivos de IoT, esto se ha realizado sobre arquitecturas clustes y cloud.

Esta línea de investigación estará orientado a investigar el concepto de arquitectura serverless o FaaS (Functions as a Service) para soluciones IoT. Debido a la disrupción de las tecnologías y su gran velocidad de desarrollo aparece la necesidad de implementar soluciones escalables, flexibles, estables y a bajos costos. Por lo tanto, el concepto de arquitectura serverless encaja perfectamente con estas necesidades. Pudiendo conseguir desplegar soluciones paralelas para IoT sin la necesidad de tener una infraestructura, ni costos de mantenimiento y operación.

Con el objeto de trabajar con serverless computing, las investigaciones comenzarán con un análisis de las plataformas cloud más adecuadas para conducir las investigaciones. Posteriormente y una vez elegida la plataforma cloud se realizarán trabajos de procesamiento paralelo analizando las métricas y midiendo el grado de eficiencia y eficacia de las soluciones planteadas.

Resultados y Objetivos

Resultados Obtenidos

Durante los últimos doce años se trabajó en el área de Computación de Altas Prestaciones y distribuidas. Dicha experiencia impulsó el proyecto de investigación y la línea que en este trabajo se presenta.

Objetivos

El objetivo del grupo de investigación es trabajar con serverless computing sobre cloud públicos mediante FaaS que usen lenguajes de programación paralela. Una vez realizadas las ejecuciones, se realizará el análisis de las métricas con el objetivo de determinar cuál es el lenguaje que mejores resultados arroja en función de los datos procesados.

La finalidad última de esta línea de investigación es determinar el grado de satisfacción alcanzado con el procesamiento paralelo en serverless computing contra los paradigmas tradicionales de procesamiento.

Formación de Recursos Humanos

El equipo de trabajo de esta línea de investigación está compuesto de ocho investigadores que figuran en este trabajo de las universidades, Nacional de Salta y San Juan y dos alumnos de grado. Además, el proyecto marco donde se está desarrollando esta propuesta incluye a tres investigadores más de la Nacional de San Luis, de la Universidad Champagnat y de la Universidad Nacional de San Juan y tres alumnos de grado.

Se está desarrollando una tesis doctoral sobre paralelismo híbrido y Big Data, seis tesis de grado en el área de Serverless computing, Concurrencia y Computación distribuida y una tesis de maestría en áreas afines. Además se espera aumentar el número de publicaciones. Por otro lado también se prevé la divulgación de varios temas investigados por medio de cursos de postgrado y actualización o publicaciones de divulgación y asesoramiento a empresas y otros organismos del estado.

Bibliografía

- [1] Stoyanova, M., Nikoloudakis, Y., Panagiotakis, S., Pallis, E., & Markakis, E. K. (2020). A survey on the internet of things (IoT) forensics: challenges, approaches, and open issues. *IEEE Communications Surveys & Tutorials*, 22(2), 1191-1221.
- [2] Farhan, L., Kharel, R., Kaiwartya, O., Quiroz-Castellanos, M., Alissa, A., & Abdulsalam, M. (2018, July). A concise review on Internet of Things (IoT)-problems, challenges and opportunities. In *2018 11th International Symposium On Communication Systems, Networks & Digital Signal Processing (CSNDSP)* (pp. 1-6). IEEE.
- [3] Medel, D., Murazzo, M. A., Molina, A. L., Sánchez, F., Cornejo, M., Rodríguez, N. R., ... & Piccoli, M. F. (2019). La Computación de Alta Performance como soporte a los sistemas altamente distribuidos. In *XXI Workshop de Investigadores en Ciencias de la Computación (WICC 2019, Universidad Nacional de San Juan)*.
- [4] Barrionuevo, M., Escalante, J., Lopresti, M., Lucero, M., Miranda, N. C., Murazzo, M. A., & Piccoli, M. F. (2020). Solución de grandes problemas aplicando HPC multi-tecnología. In *XXII Workshop de Investigadores en Ciencias de la Computación (WICC 2020, El Calafate, Santa Cruz)*.
- [5] de Souza Cimino, L., de Resende, J. E. E., Silva, L. H. M., Rocha, S. Q. S., de Oliveira Correia, M., Monteiro, G. S., ... & de Castro Lima, J. (2017, November). IoT and HPC integration: revision and perspectives. In *2017 VII Brazilian Symposium on Computing Systems Engineering (SBESC)* (pp. 132-139). IEEE.

- [6] Mauch, V., Kunze, M., & Hillenbrand, M. (2013). High performance cloud computing. *Future Generation Computer Systems*, 29(6), 1408-1416.
- [7] Murazzo, M. A., Rodriguez, N. R., Guevara, M. J., & Tinetti, F. G. (2016). Identificación de algoritmos de cómputo Intensivo para big data y su implementación en clouds. In *Workshop de Investigadores en Ciencias de la Computación (WICC 2016, Entre Ríos, Argentina)* (Vol. 18).
- [8] Rodríguez, N. R., Murazzo, M. A., Medel, D., Chávez, S. B., Martín, A. E., & Valenzuela, A. (2015). Análisis de mejora en la escalabilidad de las infraestructuras de cloud computing. In *II Jornadas de Cloud Computing & Big Data (La Plata, 2015)*.
- [9] Murazzo, M. A., Castiglione, F., Medel, D., Tinetti, F. G., & Rodríguez, N. R. (2016). Análisis de la migración del cómputo intensivo a un cloud privado con OpenStack. In *XXII Congreso Argentino de Ciencias de la Computación (CACIC 2016)*.
- [10] Murazzo, M. A., & Rodríguez, N. R. (2016). Evaluación del Impacto de Migración al Cloud. *Revista Eletrônica Argentina-Brasil de Tecnologias da Informação e da Comunicação*, 1(4).
- [11] Baldini, I., Castro, P., Chang, K., Cheng, P., Fink, S., Ishakian, V., ... & Suter, P. (2017). Serverless computing: Current trends and open problems. In *Research Advances in Cloud Computing* (pp. 1-20). Springer, Singapore.
- [12] McGrath, G., & Brenner, P. R. (2017, June). Serverless computing: Design, implementation, and performance. In *2017 IEEE 37th International Conference on Distributed Computing Systems Workshops (ICDCSW)* (pp. 405-410). IEEE.
- [13] Fox, G. C., Ishakian, V., Muthusamy, V., & Slominski, A. (2017). Status of serverless computing and function-as-a-service (faas) in industry and research. *arXiv preprint arXiv:1708.08028*.
- [14] Ciuffoletti, A. (2020). Stateless IoT. *Information*, 11(2), 85.
- [15] Pinto, D., Dias, J. P., & Ferreira, H. S. (2018, October). Dynamic allocation of serverless functions in IoT environments. In *2018 IEEE 16th international conference on embedded and ubiquitous computing (EUC)* (pp. 1-8). IEEE.