

UNIVERSIDAD NACIONAL DE TRUJILLO
Facultad de Ciencias Físicas y Matemáticas
Escuela Profesional de Informática



**DISEÑO DE ALGORITMOS PARALELOS SOBRE LA
ARQUITECTURA FOG COMPUTING PARA REDUCIR LA
LATENCIA EN IoT**

Nombre de autor(es):

Panana Rosales Andy Martín
Valera Flores Cleiver

Nombre del Asesor:

Peralta Luján José Luis

Trujillo - La Libertad

2018

PROYECTO DE INVESTIGACIÓN PARA TRABAJO DE GRADUACIÓN

ESCUELA PROFESIONAL DE INFORMÁTICA

1. GENERALIDADES

El Internet de las cosas(IoT) ha tenido gran acogimiento a lo largo de los años, durante los cuales la cantidad de dispositivos con conexión a internet superó a la de las personas alrededor del mundo. Con el avance de la tecnología, grandes cantidades de objetos físicos han sido conectados a la red generando lo que se denomina como "big data", un término que expresa una gran complejidad con la cual el IoT viene lidiando en los últimos años.

Esta situación se ve reflejada según cifras establecidas por grandes compañías relacionadas a este campo, entre ellas Cisco, las cuales pronostican un mayor crecimiento en cuanto a objetos conectados a internet en un futuro no muy lejano.

El Internet de las Cosas es un campo del cual se viene hablando desde hace varios años atrás y se busca lograr establecer las bases para darle el mejor soporte en tecnologías de comunicación, pero también se mencionan soluciones que antes eran muy eficientes y que actualmente están siendo delegadas, como por ejemplo, las redes 4G, las cuales en los últimos años están perdiendo su poder en cuanto a la conectividad de dispositivos IoT, y están viviendo la gran necesidad de evolucionar para abordar los nuevos problemas que aquejan a esta tendencia; además, si se mencionan soluciones que antes eran muy buenas, no se puede dejar de lado a Cloud Computing, que es por excelencia una de las soluciones más usadas en temas de IoT debido a su alto poder de cómputo, pero como no todo es perfecto en este mundo, frente al campo de IoT con un alto nivel de tráfico, Cloud Computing deja entrever una de sus peores falencias, la centralización, la cual la convierte en una solución no muy sofisticada para este campo.

Con esta problemática que aqueja a todo IoT nace Fog Computing, solución que viene a resolver el problema antes mencionado; Fog Computing plantea una arquitectura más distribuida y local, manejando un paradigma donde hay que aprovechar al máximo los dispositivos finales.

Para comenzar a dilucidar esta interrogante es necesario ver primero el panorama actual del Internet de las Cosas, el cual ha tenido grandes avances en los últimos años gracias a la constante mejora en el desarrollo en RFID, sensores inteligentes, tecnologías de comunicación y protocolos de Internet (Al Fuqaha et al., 2015; Borgia, 2014). Además, las mejoras en hardware y software de cómputo, en diversos dispositivos como los de sistemas integrados, de red, de visualización, de control, etc., han respaldado enormemente a IoT para crecer de manera lenta pero a la vez constante, logrando superar grandes expectativas y posicionarse como una gran tecnología emergente (Vyas et al., 2017).

Con el cálculo(operaciones matemáticas), la conectividad y el almacenamiento de datos cada vez más avanzados y universales, ha habido una explosión de soluciones de aplicación basadas en IoT en dominios diversificados desde la atención médica hasta la seguridad pública,

desde la programación de la línea de ensamblaje hasta la fabricación y otros dominios tecnológicos como el industrial y ciudades inteligentes (Borgia, 2014; Vyas et al., 2017).

IoT es una tecnología que está constantemente evolucionando gracias a la convergencia de otras como: comunicación inalámbrica, MEMS (sistemas microelectromecánicos), red inalámbrica de sensores, comunicación móvil, etc. (Vyas et al., 2017). En particular sobre la comunicación inalámbrica se comienza a ver un nuevo horizonte para IoT, donde casi no existen problemas que no podían ser solucionados al emplear como comunicaciones de este tipo a las redes 4G y se empieza a ver con buenos ojos a la nueva y reciente mejora de dicha red, las redes 5G, las cuales se espera que aporten a la expansión del Internet de las Cosas, ayudando a impulsar el funcionamiento de las redes celulares, la seguridad de IoT, los desafíos de la red, y llevar el futuro de Internet al límite (Li et al., 2018).

Sin embargo, como toda tendencia, siempre se presentan nuevos retos que impiden un correcto desenvolvimiento. Algunos problemas que obstaculizan este paradigma son mencionados en (Borgia, 2014; Ibrahim Abaker et al., 2015; Mahmud et al., 2018), estos incluyen: un incremento masivo de los datos, el aumento en la latencia y la heterogeneidad de los dispositivos.

Cloud computing (Ibrahim Abaker et al., 2015) es considerada actualmente como “la piedra base” para el Internet de las cosas, en la cual muchas aplicaciones científicas han sido desarrolladas; pero esta posee una arquitectura centralizada, por ello en (Mahmud et al., 2018; Patel et al., 2018; Yi et al., 2015) se menciona a fog computing, que sería una arquitectura ideal para IoT. Particularmente, (Mahmud et al., 2018) menciona una interoperabilidad entre Cloud y Fog, y utilizando el simulador iFogSim se demuestran mejoras en cuanto al costo, latencia y el uso de energía, cubriendo las soluciones a ciertos problemas que presenta actualmente IoT y mejorando lo que se tiene actualmente.

Además de ello surge también la idea de Mist Computing (Uehara, 2017; Yogi et al., 2017) que permitiría construir sistemas de IoT a gran escala. Mist computing propone ser un centro de datos para cloudlets en Field Area Networks y alojarse entre pequeñas nubes (cloudlets) y la niebla (fog).

Actualmente se generan un promedio de 2,5 quintillones de bytes por día, de los cuales un gran porcentaje es gracias a las cosas que están conectadas a internet, donde dichos objetos son usados para capturar datos y producirlos para un futuro procesamiento. La cantidad de ellos que se necesita procesar está llegando a un nivel donde las técnicas actuales que son usadas para lidiar con estos problemas tienen demasiadas complicaciones, por lo cual IoT necesita buscar otros métodos de procesamiento de datos más eficientes y escalables; frente a esta problemática es por lo que se habla de una de las áreas que han surgido en las últimas décadas, la informática paralela y distribuida (Murazzo et al., 2017), donde sus técnicas se pueden aprovechar oportunamente para resolver problemas a gran escala y procesar los datos provenientes del paradigma de IoT (Piccialli et al., 2018).

En vista de lo mencionado anteriormente sale a relucir la siguiente interrogante ¿De qué manera se podría reducir la elevada latencia generada por las constantes peticiones desde los dispositivos IoT al servidor?

Ahora teniendo un conocimiento amplio de la actualidad de IoT y los problemas que tanto

lo aquejan, planteamos una solución utilizando Fog Computing como arquitectura distribuida para poder descentralizar las cargas que presentaba IoT con Cloud Computing e implementando algoritmos paralelos en cada punto donde se procesa la información, esto permitiría reducir la latencia de las peticiones al servidor.

1.1. Título

“Diseño de algoritmos paralelos sobre la arquitectura Fog Computing para reducir la latencia en IoT”

1.2. Autores

Tabla 1: Datos de los alumnos investigadores

| Código(s) | Nombres y Apellidos | Cargo en el proyecto | Email |
|-------------|----------------------------|----------------------|-----------------------|
| 10127003-14 | Andy Martín Panana Rosales | Estudiante invest. | apanana@unitru.edu.pe |
| 10527010-14 | Cleiver Valera Flores | Estudiante invest. | cvalera@unitru.edu.pe |

Fuente: Elaboración propia.

1.3. Tipo de investigación

1.3.1. De acuerdo al fin que se persigue:

Aplicada

1.3.2. De acuerdo al alcance de la investigación:

Explicativa o causal, porque se propone una solución concebida gracias a la combinación de la Arquitectura Fog Computing y la computación paralela con el objetivo de reducir la latencia que presentan las actuales propuestas en IoT, propuestas que en su gran mayoría funcionan bajo una arquitectura centralizada, Cloud Computing.

1.4. Área y línea de Investigación

1.4.1. Área de investigación :

Computación paralela y distribuida

1.4.2. Línea de Investigación:

Sistemas distribuidos

1.4.3. Tema de investigación :

Computación paralela y distribuida

1.5. Localidad e Institución donde se desarrollará el proyecto

1.5.1. Localidad:

Av. Juan Pablo II, Trujillo, Trujillo, La Libertad

1.5.2. Institución:

Universidad Nacional de Trujillo

1.6. Duración del trabajo de graduación (Plan TG y desarrollo del TG)

Del 20/08/2018 al 07/12/2018 (4 meses)

1.7. Cronograma del trabajo de graduación

Tabla 2: Etapas y actividades para el trabajo de graduación

| Etapas | Actividades/tareas | Fecha inicio | Fecha término | Hs. semanal |
|-------------------------|--|---------------------|----------------------|--------------------|
| Preparación del plan TG | Elaborar plan TG. | 20/08/2018 | 26/08/2018 | 12 - 16 |
| | Aprobación del plan TG por jurado. | 27/08/2018 | 07/09/2018 | 2 |
| Recolección de datos | Inv. bibliográfica. | 08/09/2018 | 03/11/2018 | 12 - 16 |
| | Instrumentos de medición. | 24/09/2018 | 30/09/2018 | 12 - 16 |
| Análisis de datos | Comparar datos recolectados. | 01/10/2018 | 14/10/2018 | 12 - 16 |
| Redacción del informe | Introducción | 15/10/2018 | 21/10/2018 | 12 - 16 |
| | Marco Teórico | 15/10/2018 | 21/10/2018 | 12 - 16 |
| | Desarrollo de la propuesta, diseño de algoritmos paralelos | 22/10/2018 | 03/11/2018 | 12 - 16 |
| | Informe I: Avance Tesis | 29/10/2018 | 03/11/2018 | 4 |
| | Avance de resultados | 04/11/2018 | 02/12/2018 | 8 |
| | Informe II: Avance Tesis. | 03/12/2018 | 07/12/2018 | 4 |
| | Desarrollo de los algoritmos paralelos | 08/12/2018 | 16/02/2019 | 12 - 16 |
| | Implementación y debugging de los algoritmos paralelos | 16/02/2019 | 18/05/2019 | 12 - 16 |
| | Resultados | 19/05/2019 | 10/06/2019 | 12 - 16 |
| | Conclusiones | 11/06/2019 | 30/06/2019 | 12 - 16 |
| | Trabajos futuros | 01/07/2019 | 13/07/2019 | 12 - 16 |
| Sustentación de tesis | Sustentación del trabajo de investigación. | 15/07/2019 | 26/07/2019 | 4 |

Fuente: Elaboración propia.

1.8. Recursos disponibles

1.8.1. Personal

Tabla 3: Recursos disponibles - Personal

| Código | Descripción | Nombre |
|---------------|--------------------|----------------------------|
| 2.3.2.7.2.2 | Asesor | José Luis Peralta Luján |
| 2.3.2.7.2.5 | Investigador | Andy Martín Panana Rosales |
| 2.3.2.7.2.5 | Investigador | Cleiver Valera Flores |

Fuente: Elaboración propia.

1.8.2. Materiales y equipos

1.8.2.1. Materiales de Consumo

Tabla 4: Recursos disponibles - Material de consumo

| Código | Cantidad | Descripción |
|---------------|-----------------|--------------------|
| 2.3.1.5.1.1 | 02 | Memoria USB 8 GB |
| 2.3.1.5.1.1 | 04 | Lapicero |
| 2.3.1.5.1.1 | 08 | Fólder Manila |
| 2.3.1.5.1.1 | 04 | Lápiz |
| 2.3.1.5.1.1 | 02 | Borrador |
| 2.3.1.5.1.1 | 1000 | Papel Bond A4 |

Fuente: Elaboración propia.

1.8.2.2. Hardware

Tabla 5: Recursos disponibles - Hardware

| Código | Cantidad | Descripción |
|---------------|-----------------|--------------------------------|
| 2.6.3.2.1.1 | 01 | Raspberry Pi 3 B |
| 2.6.3.2.1.1 | 01 | Laptop ASUS Core i3 |
| 2.6.3.2.1.1 | 01 | Laptop LENOVO Z50-45 AMD A8 |

Fuente: Elaboración propia.

1.8.2.3. Software

Tabla 6: Recursos disponibles - Software

| Código | Cantidad | Descripción |
|---------------|-----------------|--------------------|
| 2.6.6.1.3.2 | 02 | Microsoft Office |
| 2.6.6.1.3.2 | 01 | Zorin OS |
| 2.6.6.1.3.2 | 01 | Linux Mint |
| 2.6.6.1.3.2 | 02 | Windows 10 |

Fuente: Elaboración propia.

1.8.2.4. Servicios

Tabla 7: Recursos disponibles - Servicios

| Código | Descripción |
|---------------|--------------------|
| 2.3.2.2.1.1 | Energía Eléctrica |
| 2.3.2.1.2.1 | Transporte Público |
| 2.3.2.2.2.3 | Internet Movistar |

Fuente: Elaboración propia.

1.9. Presupuesto

1.9.1. Personal

Tabla 8: Presupuesto - Personal

| Código | Descripción | Costo estimado |
|---------------|--------------------|-----------------------|
| 2.3.2.7.2.2 | Asesor | S/. 0.00 |
| 2.3.2.7.2.5 | Investigador | S/. 0.00 |
| 2.3.2.7.2.5 | Investigador | S/. 0.00 |
| | Total | S/. 0.00 |

Fuente: Elaboración propia.

1.9.2. Materiales y equipos

1.9.2.1. Materiales de Consumo

Tabla 9: Presupuesto - Materiales de Consumo

| Código | Cantidad | Descripción | Costo |
|---------------|-----------------|--------------------|--------------|
| 2.3.1.5.1.1 | 02 | Memoria USB 8 GB | S/. 50.00 |
| 2.3.1.5.1.1 | 04 | Lapicero | S/. 8.00 |
| 2.3.1.5.1.1 | 08 | Fólder Manila | S/. 4.00 |
| 2.3.1.5.1.1 | 04 | Lápiz | S/. 3.00 |
| 2.3.1.5.1.1 | 02 | Borrador | S/. 1.00 |
| 2.3.1.5.1.1 | 1000 | Papel Bond A4 | S/. 22.00 |
| | | Total | S/. 88.00 |

Fuente: Elaboración propia.

1.9.2.2. Hardware

Tabla 10: Presupuesto - Hardware

| Código | Cantidad | Descripción | Vida útil | Costo |
|---------------|-----------------|-----------------------------|------------------|--------------|
| 2.6.3.2.1.1 | 01 | Raspberry Pi 3 B | 2 años | S/. 230.00 |
| 2.6.3.2.1.1 | 01 | Laptop ASUS Core i3 | 3 años | S/. 1500.00 |
| 2.6.3.2.1.1 | 01 | Laptop LENOVO Z50-70 AMD A8 | 3 años | S/. 1800.00 |
| | | Total | | S/. 3530.00 |

Fuente: Elaboración propia.

1.9.2.3. Software

Tabla 11: Presupuesto - Software

| Código | Cantidad | Descripción | Costo |
|---------------|-----------------|--------------------|--------------|
| 2.6.6.1.3.2 | 02 | Microsoft Office | S/. 365.00 |
| 2.6.6.1.3.2 | 01 | Zorin OS | S/. 0.00 |
| 2.6.6.1.3.2 | 01 | Linux Mint | S/. 0.00 |
| 2.6.6.1.3.2 | 02 | Windows 10 | S/. 555.00 |
| | | Total | S/. 920.00 |

Fuente: Elaboración propia.

1.9.2.4. Servicios

Tabla 12: Presupuesto- Servicios

| Código | Descripción | Costo |
|---------------|--------------------|-------------------|
| 2.3.2.2.1.1 | Energía Eléctrica | S/. 288.00 |
| 2.3.2.1.2.1 | Transporte Público | S/. 150.00 |
| 2.3.2.2.2.3 | Internet Movistar | S/. 480.00 |
| | Total | S/. 918.00 |

Fuente: Elaboración propia.

1.9.3. Costo del Proyecto

Tabla 13: Costo del Proyecto

| Descripción | Costo |
|-----------------------|--------------------|
| Personal | S/. 0.00 |
| Materiales de consumo | S/. 88.00 |
| Hardware | S/. 3530.00 |
| Software | S/. 920.00 |
| Servicios | S/. 918.00 |
| Total | S/. 5456.00 |

Fuente: Elaboración propia.

1.10. Financiamiento

1.10.1. Con recursos universitarios

Ninguno

1.10.2. Con recursos externos

Ninguno

1.10.3. Autofinanciación

Autofinanciado

2. PLAN DE INVESTIGACIÓN

2.1. Realidad problemática

Según cifras establecidas por grandes empresas del medio como Cisco, se puede predecir que en un futuro próximo tendremos cerca de 50 a 100 mil millones de objetos conectados electrónicamente a Internet, es un cifra bastante colosal que poco a poco genera problemas debido a las tecnologías que se usan actualmente para poder desplegar cualquier proyecto basado en Internet de las Cosas.

Desde el año 2000 se viene hablando sobre Internet de las Cosas y cómo cimentar las bases exactas para lograr darle el mejor soporte en tecnologías de comunicación, pero de acuerdo a ciertos autores, cada solución que se plantean para un problema es momentánea, es decir, funciona bajo el paradigma de una enfermedad terminal, dónde cura la herida por un cierto tiempo pero si no se plantean nuevas formas mas sofisticadas para solucionar el problema, volvemos a caer en el mismo hoyo. Tal es el caso de las redes 4G que han permitido que IoT no tenga problemas con la conectividad por mucho tiempo, pero la gran cantidad de dispositivos ha obligado a esta tecnología a evolucionar y lidiar con el problema que presenta IoT.

La tecnología más usada por parte de IoT, es Cloud Computing, dicha tecnología es buena solución ya que nos provee alto poder de cómputo, pero ya que nos enfrentamos a un contexto en el cual se tiene que lidiar con grandes cantidades de datos provenientes de diferentes puntos de información, esta deja de ser una buena solución, su arquitectura centralizada provoca grandes problemas, uno de ellos la latencia. Al enfrentarse a una masiva cantidad de datos, las 24 horas del dia, Cloud Computing no la pasa nada bien, su arquitectura centralizada hace que IoT no despliegue todo su potencial.

Partiendo de esta problemática nace Fog Computing, una solución que viene de una u otra manera a lidiar con la latencia que se presenta en la entrega de los datos, planteando un arquitectura distribuida y más local al lugar donde se generan los datos; además de eso Fog Computing lanza la idea de aprovechar la capacidad de cómputo de los dispositivos finales, pero es ahí donde tenemos un nuevo problema, los dispositivos finales cuentan con recursos limitados; recursos que son accesibles en cuestiones económicas debido al auge y el crecimiento de IoT pero con limitaciones en su nivel de procesamiento y almacenamiento.

2.2. Antecedentes

Al Fuqaha et al. (2015) expresan que IoT está habilitado por los últimos desarrollos en RFID, sensores inteligentes, tecnologías de comunicación y protocolos de Internet. Donde la premisa básica es que los sensores inteligentes colaboren directamente sin participación humana para ofrecer una nueva clase de aplicaciones. También enfatizan en que la revolución actual en las tecnologías de Internet, dispositivos móviles y máquina a máquina (M2M) se puede ver como la primera fase del IoT. En los próximos años, esperan que IoT conecte diversas tecnologías para permitir nuevas aplicaciones conectando objetos físicos en apoyo de una toma de decisiones inteligente. Este trabajo nos muestra el panorama actual de IoT entorno a los ultimos avances tecnologicos de algunos campos y culmina enfatizando la necesidad de muchas aplicaciones con respecto a la evolucion de IoT.

Borgia (2014) explica IoT como un nuevo paradigma que combina aspectos y tecnologías

provenientes de diferentes enfoques donde la computación ubicua, el protocolo de Internet, las tecnologías de detección, las tecnologías de comunicación y los dispositivos integrados se fusionan para formar un sistema en el que el mundo real y digital se encuentran y están continuamente en interacción simbiótica. Además mencionan que la gran cantidad esperada de dispositivos interconectados y la gran cantidad de datos disponibles abren nuevas oportunidades para crear servicios que traerán beneficios tangibles a la sociedad, el medio ambiente, la economía y los ciudadanos individuales. Esta investigación nos aporta un enfoque donde la cantidad de dispositivos que serán conectados a internet en los próximos años, abre un sinnúmero de posibilidades entorno a la creación de servicios en diferentes campos como la economía, el medio ambiente, etc.

Mahmud et al. (2018), esta investigación trata sobre IoT en el cuidado de la salud, y los problemas que existen relacionados a la latencia, carga de datos desigual y la heterogeneidad de las aplicaciones. Muchos servicios de la computación en la nube han sido utilizados en este campo, sin embargo, la centralización geográfica de los centros de datos de la nube no ha podido solucionar del todo los problemas antes mencionados, por ello se habla de una interoperabilidad entre las arquitecturas Cloud y Fog Computing lo cual implica una compleja coordinación entre aplicaciones y servicios. Este trabajo nos brinda resultados esperados en cuanto a la solución de dichos problemas, simulando el uso de la arquitectura Fog Computing con el software iFogSim.

Verma et al. (2015) mencionan los problemas que presenta Cloud Computing como por ejemplo: limitación con respecto al equilibrio de la carga de datos y el alojamiento de centros de datos, los cuales crean un nivel de latencia de red grande e impredecible. Además, presentan como solución para dichos problemas a un nuevo modelo de computación llamado Fog Computing, el cual es análoga a la nube, con la única diferencia es que se encuentra en el borde de la red; otra ventaja es que las aplicaciones que requieren dependencia de ubicación pueden ser viables a través de ella, proponen una nueva arquitectura basada en el algoritmo de equilibrio de carga en el entorno de Fog Computing de manera menos compleja y efectiva, debido a la gran cantidad de usuarios finales los cuales aumentan cada día a un ritmo muy rápido. Dicha investigación nos brindará información acerca del gran poder que tiene Fog Computing frente a temas de latencia y su manera eficiente de como solucionarlo, además nos muestran un panorama donde la combinación entre Fog Computing y el algoritmo de equilibrio de carga logran ser un dúo perfecto frente a estos problemas.

Con el transcurso de los años los dispositivos IoT requieren mayor espacio para poder almacenar datos, los cuales aumentan masivamente (Murazzo et al., 2017), señala la necesidad de mejorar los tiempos de respuesta y la escalabilidad de sistemas; menciona ejemplos de aplicaciones web con datos masivos que podrían manejarse con técnicas de computación de alto rendimiento permitiendo que la velocidad de procesamiento mejore. Este trabajo indica los campos necesarios a investigar para lidiar con grandes volúmenes de datos, tales como: arquitecturas híbridas, problemas de datos masivos y las herramientas existentes para resolverlos, entre otros.

Peng et al. (2015) presentan una red de acceso de radio basado en Fog Computing (F-RAN) como iniciativa para lograr alcanzar los objetivos propuestos para el sistema de comunicación inalámbrica de quinta generación (5G). Además, resaltan que en las actuales redes propuestas los operadores necesitan desplegar una gran cantidad de recursos, lo que los convierte en un serio desperdicio cuando el volumen de tráfico de entrega no es lo suficientemente grande. Por ende, proponen a Fog Computing como una buena alternativa frente a Cloud Computing, cuya

propuesta viene a colocar una cantidad sustancial de almacenamiento, comunicación, control, configuración, medición y administración en el borde de una red. Este trabajo nos aporta una aplicación importante de Fog Computing en un campo como las redes 5G, donde se busca lograr el aprovechamiento de los dispositivos finales (RRH) frente a las limitaciones que presenta la propuesta planteada con el fin de lograr los objetivos de 5G.

Piccialli et al. (2018) nos explica sobre la cantidad de almacenamiento que se va generando por el incremento de dispositivos conectados a internet, en su mayoría debido a dispositivos IoT, y la estimada para el 2020; a su vez indica que el procesamiento actual de los datos no es ideal pero mejoraría con la inclusión de técnicas de computación paralela y distribuida. Como aporte, esta investigación reafirma nuestra idea de que al utilizar algoritmos paralelos podemos lograr una mejora notable en cuanto a la velocidad del procesamiento de los datos y por ende disminuir la latencia.

T. Shanmugapriya et al. (2018) muestra un panorama donde una variante de un algoritmo de agrupamiento más reciente basado en el hallazgo rápido y la búsqueda de picos de densidad (CFS), tiene la capacidad de afrontar la agrupación de grandes cantidades de datos dinámicos en el Internet de las cosas industriales, donde muestra que dicha variación puede encontrar grupos de forma arbitraria y determinar la cantidad de clústeres automáticamente, solucionado de cierta manera las limitaciones que presentaba el algoritmo inicial frente al agrupamiento de grandes volúmenes de datos. Como aporte, esta investigación nos permite ver el potencial oculto que tiene este algoritmo CFS propuesto por Alessandro Laio, Alex Rodriguez y Maria d'Errico frente al agrupamiento de datos, que si bien es cierto tiene una limitación bien marcada frente al volumen de estos, podemos mejorarlo de acuerdo a nuestras necesidades o al problema que queremos resolver y obtener mejores resultados.

2.3. Justificación

La necesidad de hacer frente a la problemática que vive actualmente la tecnología IoT y la que vivirá en los próximos años, induce a pensar rápido en la manera de cómo mitigarla y lograr solucionar los posible cuellos de botella que sufre y sufrirá esta tecnología. Por ende, la propuesta planteada en este trabajo requiere de un involucramiento a fondo en el paradigma donde cada objeto puede estar conectado a internet o de forma resumida, el Internet de las Cosas, también no llama a sumergirse en las diferentes arquitecturas enfocadas a dar soporte a esta tecnología y a lograr sacar a relucir todo el potencial de la computación paralela por medio de la creación de algoritmos de este tipo, que para procesar grandes cantidades de datos son mucho más eficaces que los comúnmente usados (algoritmos secuenciales), de tal manera que la complementación de estos permite plantear una nueva forma o enfoque de cómo solucionar la latencia que presenta IoT frente a un contexto donde se generan una gran cantidad de información de manera distribuida y la enorme necesidad de analizar toda esta data. De igual modo frente al gran crecimiento de la cantidad de dispositivos conectados a internet, el tema de la entrega de datos puede sufrir grandes complicaciones debido a factores como: la centralización que presenta la arquitectura sobre la cual funciona la mayoría de proyectos IoT y la gigantesca información proveniente de dispositivos finales que se necesitan procesar; dado a conocer este complicado panorama es que se necesita tener una solución frente al tema de la latencia existente en la transmisión de datos, por lo cual se propone desarrollar un mecanismo concebido gracias a la complementación entre la arquitectura Fog Computing y algoritmos paralelos,

permitiendo lograr reducir los tiempos de transmisión de datos entre los dispositivos donde se originan de los datos y los servidores alojados en la nube o centros de datos. De esta manera, se busca lograr que cada idea o proyecto sobre IoT, principalmente aquellos que manejan grandes cantidades de información de diferentes puntos de origen, no tengan complicaciones con respecto al tiempo de la transmisión de datos y el tema de análisis de la información se desarrolle la mejor manera posible.

2.4. Problema

¿De qué manera se podría reducir la elevada latencia generada por las constantes peticiones desde los dispositivos IoT al servidor?

2.5. Hipótesis

El uso de algoritmos paralelos sobre la arquitectura Fog Computing reducirá la latencia en el Internet de las cosas.

2.6. Variables

2.6.1. Variable Independiente

Tabla 14: Variable Independiente

| Variable | Definición Conceptual | Dimensiones | Indicadores |
|--|--|---|--|
| Algoritmos paralelos sobre la arquitectura Fog Computing | Algoritmos enfocados al uso de varios procesadores trabajando juntos para resolver una tarea común | Rendimiento de los Algoritmos Paralelos | Tiempo de ejecución. Aceleración (speedup) ^a . Eficiencia (efficiency) ^b . |

Fuente: Elaboración propia.

^a El cálculo de la aceleración (SpeedUp) esta expresado por la siguiente fórmula:

$$SU = \frac{T_s}{T_p}$$

Figura 1: Fórmula para el cálculo del SpeedUp.

Fuente: Rasúa (2009).

Donde:

Ts = Tiempo en ejecución de un programa secuencial.

Tp = Tiempo de ejecución de la versión paralela de dicho programa.

^b El cálculo de la eficiencia (efficiency) esta expresado por la siguiente fórmula:

$$E(n) = \frac{S(n)}{n}$$

Figura 2: Fórmula para el cálculo de la Eficiencia.

Fuente: Rasúa (2009).

Donde:

S(n) = SpeedUp.

n = Cantidad de procesadores.

2.6.2. Variable Dependiente

Tabla 15: Variable Dependiente

| Variable | Definición Conceptual | Dimensiones | Indicadores |
|--------------------------------------|--|-------------|--|
| Latencia en el Internet de las Cosas | Retardo existente en la entrega de datos dentro la red | Latencia | Tiempo existente entre la transmisión y la recepción de los datos. |

Fuente: Elaboración propia.

2.7. Objetivos

2.7.1. Objetivo general

Disminuir el tiempo de latencia en los dispositivos de internet de las cosas, utilizando algoritmos paralelos sobre la arquitectura Fog Computing.

2.7.2. Objetivos específicos

- Determinar el nivel de latencia que presenta actualmente la arquitecturas IoT bajo Cloud Computing.
- Comparar el performance que existe entre Cloud Computing y Fog Computing como arquitectura para IoT.
- Diseñar algoritmos paralelos con la menor complejidad posible.
- Comprobar que los algoritmos paralelos permiten disminuir la latencia.
- Analizar los resultados obtenidos de la investigación.

2.8. Método de trabajo

2.8.1. Tipo investigación

Estudio Explicativo

2.8.2. Diseño de la investigación

a) Tipo de Diseño

Diseño experimental - Experimento puro

a) Explicación del tipo

Los experimentos puros son aquellos que reúnen los dos requisitos para lograr el control y la validez interna: grupos de comparación, y equivalencia de grupos. Asimismo, pueden utilizar prepuebas y pospruebas para analizar la evolución de los grupos antes y después del tratamiento experimental Sampieri et al. (1997).

a) **Diseño con posprueba únicamente y grupo de control**

Este diseño incluye dos grupos, uno recibe el tratamiento experimental y el otro no (grupo de control).

Esquema del diseño:

RG1: - O1

RG2: X O2

Donde:

X: variable independiente (Algoritmos Paralelos sobre Fog Computing).

O1: medición previa (antes de la aplicación de Algoritmos Paralelos sobre Fog Computing) de la variable dependiente.

O2: medición posterior (después de la aplicación de Algoritmos Paralelos sobre Fog Computing) de la variable dependiente.

2.8.3. Población y muestra

Universo: Transmisiones de datos en Internet de las Cosas.

Población: Transmisiones de datos desde dispositivos IoT hasta un servidor.

Muestra: Transmisiones de datos desde dispositivos IoT hasta un servidor.

Tipo de muestra: Muestra Aleatoria.

Unidad de estudio: Transmisión de los datos(paquetes) de IoT donde se manejan grandes cantidades de información.

Cálculo de la muestra:

Para calcular el tamaño de la muestra suele utilizarse la siguiente fórmula:

$$n = \frac{Z^2 S^2}{E^2} = 153,664$$

Figura 3: Fórmula para el cálculo de la Muestra.

Fuente: Metodología de la Investigación Sampieri et al. (1997).

Donde:

n = el tamaño de la muestra.

S = Desviación estándar de la población.

Z = Valor obtenido mediante niveles de confianza. Es un valor constante que, si no se tiene su valor, se le toma en relación al 95 % de confianza que equivale 1.96 (como más usual) o en relación al 99 % de confianza que equivale 2.58, valor que queda a criterio del investigador.

E = Límite aceptable de error muestral que, generalmente cuando no se tiene su valor, suele utilizarse un valor que varía entre 1 % (0.01) y 9 % (0.09), valor queda a criterio del encuestador Sampieri et al. (1997).

En nuestro caso:

Tabla 16: Cálculo de la Muestra.

| | | Elevado al cuadrado |
|---------------------------------|---------|----------------------------|
| E | 0.5ms | 0.0025ms |
| S | 1ms | 1ms |
| Z | 1.96 | 3.8416 |
| | | Redondeando |
| n (tamaño de la muestra) | 153.664 | 153 |

Fuente: Elaboración propia.

2.8.4. Técnicas e Instrumentos de recolección de datos

Para la recolección de datos, se usaran las siguientes técnicas según Sampieri et al. (1997):

Investigación documental: Mediante la recolección de información en diferentes documentos como tesis, artículos y libros, se lograra tener una referencia formal y valida para el diseño de los Algoritmos Paralelos, asi como la manera de adaptarla con la arquitectura Fog Computing.

Observación: Se observará como trabajan actualmente los dispositivos IoT y como se llevan a cabo sus transmisiones.

Tabla 17: Técnicas e Instrumentos de recolección de datos.

| Variable | Técnica | Instrumento |
|---------------------------------------|---------------------------|--------------------------------|
| Latencia en el Internet de las Cosas. | Investigación documental. | Tesis, Artículos, Libros, etc. |
| | Observación. | Guías de observación. |

Fuente: Elaboración propia.

2.8.5. Etapas

Para llegar a los objetivos propuestos, el desarrollo de la investigación comprendió las siguientes etapas de trabajo:

- Análisis de los principales problemas con respecto a la transmisión de datos a los que se enfrenta actualmente IoT y a los que va a enfrentarse en un futuro cercano.
- Formulación del problema principal de la investigación, justificando su importancia.
- Levantamiento bibliográfico de los diferentes temas necesarios para la elaboración de la investigación, tales como IoT, Cloud Computing, Fog Computing, Algoritmos Paralelos, entre otros.
- Estudio y análisis del panorama actual y remoto de IoT, con la finalidad de aportar al estado del arte del problema formulado.
- Estudio y análisis de los diferentes métodos de solución evaluados en la revisión bibliográfica, así como los que serán propuestos en este estudio.
- Diseño de los algoritmos paralelos.
- Implementación y debugging de los algoritmos paralelos previamente diseñados sobre la arquitectura Fog Computing.
- Selección de muestras y realización de pruebas.
- Análisis de resultados obtenidos con la herramienta “Excel”.

Referencias

- Al Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., and Ayyash, M. (2015). Internet of Things: A survey on enabling technologies, protocols and applications. *IEEE*, 17(4):2347–2376.
- Borgia, E. (2014). The Internet of Things vision: Key features, applications and open issues. *Computer Communications*, 54:1–31.
- Ibrahim Abaker, T. H., Ibrar, Y., Nor Badrul, A., Salimah, M., Abdullah, G., and Samee Ullah, K. (2015). The rise of “big data” on cloud computing: Review and open research issues. *Information Systems*, 47:98–115.
- Li, S., Zhao, S., and Xu, L. D. (2018). 5G Internet of Things: A survey. *Industrial Information Integration*.
- Mahmud, R., Luiz Koch, F., and Buyya, R. (2018). Cloud-Fog Interoperability in IoT-enabled Healthcare Solutions. *ACM New York*.
- Murazzo, M. A., Piccoli, M. F., Rodríguez, N. R., Medel, D., Mercado, J., Sánchez, F., and Tello, M. (2017). Paralelismo Híbrido aplicado a soluciones de problemas de datos masivos. *Buenos Aires: XIX Workshop de Investigadores en Ciencias de la Computación*.
- Patel, H. M., Chaudhari, R. R., Prajapati, K. R., and Patel, A. A. (2018). The Interdependent Part of Cloud Computing: Dew Computing. *Intelligent Communication and Computational Technologies*.
- Peng, M., Yan, S., Zhang, K., and Wang, C. (2015). Fog Computing based Radio Access Networks: Issues and Challenges. *National Natural Science Foundation of China*.
- Piccialli, F., Cuomo, S., and Jeon, G. (2018). Parallel Approaches for Data Mining in the Internet of Things Realm. *International Journal of Parallel Programming*.
- Rasúa, R. A. T. (2009). *Algoritmos paralelos para la solución de problemas de optimización discretos aplicados a la decodificación de señales*. PhD thesis, Universidad de València.
- Sampieri, R. H., Fernández-Collado, C., and Lucio, P. B. (1997). *Metodología de la Investigación*. McGraw-Hill.
- T. Shanmugapriya, M., Anbuselvi, A., and Thendral, A. (2018). An Incremental CFS Algorithm for Clustering Large Data in Industrial Internet of Things. *CSE ULTRA College of Engineering & Technology for Womena*.
- Uehara, M. (2017). Mist Computing: Linking Cloudlet to Fogs. *Computational Science/Intelligence and Applied Informatics*, 726:201–213.
- Verma, M., Bhardawaj, N., and Yadav, A. K. (2015). An architecture for Load Balancing Techniques for Fog Computing Environment. *UPTU University Agra*, 6:269–274.
- Vyas, D. A., Bhatt, D., and Jha, D. (2017). IoT: Trends, Challenges and Future Scope. *Computer Science & Electronics Journals*, 7:186–197.

Yi, S., Li, Q., and Li, C. (2015). A Survey of Fog Computing: Concepts, applications and issues. *Williamsburg: College of William and Mary*.

Yogi, M. K., Chandrasekhar, K., and Kumar, G. V. (2017). Mist Computing: Principles, trends and future direction. *SSRG International Journal of Computer Science and Engineering*, 4(7):19–21.

Andy Panana Rosales

Cleiver Valera Flores

Jose L. Peralta Lujan
Asesor