



FACULTAD DE INGENIERÍA
INGENIERÍA CIVIL INFORMÁTICA

Una arquitectura caché para redes ICN basada en comportamiento de usuario

Tesis para optar al grado de ingeniero civil en informática.

Autor:

Mathias Nicolas Velilla Brandau.

Profesores guía:

Carlos Gomez-Pantoja
Miguel Guitierrez

Santiago de Chile, Chile.

Abril, 2017

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Desafíos	3
1.3. Contribución de la tesis	3
1.4. Estructura de la tesis	4
2. Descripción del problema	5
2.1. Contextualización	5
2.2. El problema	5
2.2.1. Declaración del problema	5
2.2.2. Diagrama de ishikawa	5
2.3. Objetivos	5
2.3.1. Objetivo general	5
2.3.2. Objetivos específicos	5
2.3.3. Alcance de los objetivos	6
3. Marco teórico	7
3.1. Information Centric Network	7
3.2. Caché	10
3.3. Comportamiento de usuario	10
4. Estado del arte	11

Índice de figuras

3.1. Nodo CCN	8
3.2. Paquetes CCN	9
3.3. Nodo CCN	10

Capítulo 1

Introducción

1.1. Motivación

El nacimiento del Internet en el año 1958 en EEUU a través de ARPA (*Advanced Research Projects Agency*) tuvo como objetivo la comunicación entre diversas universidades y las entidades militares con propósitos investigativos, la invención del telégrafo, el teléfono, la radio y el ordenador sentaron las bases para el desarrollo de esta nueva tecnología, la cual hoy en día ha revolucionado la informática y las comunicaciones como ninguna otra cosa, convirtiéndose en una herramienta de índole mundial, un mecanismo el cual nos permite diseminar información de manera inmediata, generando un medio de colaboración e interacción entre las personas y los ordenadores, desconociendo su ubicación física.

Durante los años 2000 - 2008 el uso diario del Internet en las personas americanas se desplazaba entre un 60 y 70 por ciento [Digital CITIZE], este aumento en el uso de la Internet, conlleva también a un aumento de la información presente en la web, dando como resultado el desarrollo de aplicaciones las cuales deben interactuar con un gran número de usuarios y analizando una sobresaliente cantidad de información, motivo por el cual Internet impulsado por las demandas de las aplicaciones cada vez más emergentes y las capacidades de las nuevas redes de comunicación, se ha convertido en un mosaico arquitectónico que resulta en una creciente complejidad y vulnerabilidades imprescindibles, entregando como resultado violaciones de capas, proliferación de subcapas, y la erosión del modelo de extremo a extremo, es bajo esta eventualidad que todos los cambios efectuados concluyen en un aumento de la complejidad, lo cual se traduce en una Internet osificada, es por esta razón que los problemas anteriormente señalados no se encuentran directamente relacionados con los protocolos o mecanismo específicos del Internet actual, mas bien son causados esencialmente por

la incapacidad de integrar nuevos mecanismos, lo que quiere decir que los problemas son causados por la arquitectura de Internet y podrían ser resueltos con un nuevo diseño de arquitectura de Internet[1].

La creciente demanda de una distribución altamente escalable y eficiente del contenido ha motivado el desarrollo de futuras arquitecturas de Internet basadas en objetos de datos nombrados (NDO, por sus siglas en inglés), por ejemplo, páginas web, videos, documentos u otras informaciones. El enfoque de estas arquitecturas se conoce comúnmente como redes centradas en información (ICN). Por el contrario, las redes actuales se centran en el host, donde la comunicación se basa en hosts nombrados, por ejemplo, servidores web, PC, portátiles, teléfonos móviles y otros dispositivos. Las arquitecturas ICN aprovechan el almacenamiento en red para el almacenamiento en caché, la comunicación multilateral a través de la replicación y los modelos de interacción que desacoplan a los remitentes y receptores. El objetivo común es lograr una distribución eficiente y confiable de los contenidos en donde cada uno de estos es nombrado única e independientemente desde la ubicación del productor(servidor), facilitando el almacenamiento en cache en los nodos intermediarios. Así, por ejemplo, los consumidores solicitaran contenidos enviando el denominado paquete de interés, el cual lleva el nombre del contenido, mientras que el productor o cualquier nodo que mantenga una copia del contenido puede responder a la petición realizada.[4]

A su vez, los usuarios ya antes mencionados como consumidores, realizan peticiones en la web, las cuales siguen una conducta dinámica caracterizándose por un elevado sesgo entre los diferentes conjuntos de peticiones, en otras palabras, dentro del universo de peticiones generadas existen conjuntos de peticiones que son regularmente solicitadas por los consumidores en intervalos de tiempos distintos, por lo contrario, otras peticiones escasamente son solicitadas. Dicho lo anterior, también existen situaciones dentro de un intervalo acotado de tiempo, donde surgen explosivamente peticiones las que se caracterizan por poseer un contenido en común, las cuales nacen por el desarrollo de un evento de interés popular teniendo como resultado un aumento sustancial en la demanda generada a los nodos, teniendo como consecuencia latencia y congestión en las redes.

1.2. Desafíos

Los desafíos que se afrontaran para la realización de proyecto de titulo I, son inicialmente el diseño de una nueva arquitectura de memoria caché para los nodos de las redes *ICN* (*Information Centric Network*), considerando el comportamiento de los usuarios por medio del tráfico de red. En segundo lugar, se debe incorporar una estrategia de caché(políticas de admisión, desalojo y reemplazo) y que será implementada en un simulador denominado *ndnSIM* (*Named Data Networking Simulator*), todo esto con el fin de mejorar resultados(Queresultados) en comparación con otras arquitecturas caché que se encuentran por defecto dentro del simulador(LFU, FIFO).

1.3. Contribución de la tesis

El aporte entregado por el proyecto de titulo, se puede identificar inicialmente por la creación diseño de una nueva arquitectura de memoria caché para los nodos de las redes *ICN* (*Information Centric Network*), considerando el comportamiento de los usuarios por medio del tráfico de red. En segundo lugar, se debe incorporar una estrategia de caché(políticas de admisión, desalojo y reemplazo) y que será implementada en un simulador denominado *ndnSIM* (*Named Data Networking Simulator*), todo esto con el fin de mejorar resultados de eficiencia de la memoria caché (Hit's, Miss) en comparación con otras arquitecturas caché que se encuentran por defecto dentro del simulador(LFU, FIFO). No obstante, a continuación se detallaran las contribuciones efectuadas por este trabajo:

- Diseño de una nueva arquitectura cache bajo el paradigma de las redes ICN, que contenga una subdivisión de tres grupos capaces de retener diferentes tipos de peticiones(intereses) de modo que la eficiencia del nodo se vea afectada positivamente. En cuanto a la división de la memoria caché, el primer segmento se encarga del almacenamiento de peticiones del tipo ráfaga, la segunda de guardar las de tipo permanente y para la ultima sección, la recaudación de peticiones de tipo variables.
- Creación de una topologia de redes ICN, utilizando el simulador ndnSIM, la cual contenga dentro de sus nodos la arquitectura caché anteriormente señalada

con la finalidad de inyectar tráfico de datos en base al comportamiento usuario para la obtención de resultados empíricos respecto a la nueva arquitectura caché diseñada.

1.4. Estructura de la tesis

A continuación se realizara una breve reseña de como se estructura el siguiente trabajo:

El capitulo 2 abarca la descripción del problema, en el se define el objetivo general, específicos, hipótesis y los alcances que contiene el proyecto de titulo. Incluyendo también la descripción de la metodología de trabajo utilizada para el progreso del trabajo.

El capitulo 3 posee el marco teórico, en el cual se definen diferentes conceptos esenciales para el entendimiento del proyecto del titulo. Dicho lo anterior se encuentran los siguiente conceptos: Redes centradas en la información (ICN), ndnSIM, aplicaciones web de gran escala, el cache, estructura y políticas de remplazo del caché existentes y finalizando con el comportamiento del usuario de manera donde se explica la ley de Zipf y se detallan los tipos de peticiones de usuario.

En el capítulo 4 se presenta una revisión bibliográfica de los trabajos realizados respecto a las arquitecturas cache y las políticas de cache implementadas en redes ICN, ademas del revisión de trabajos dentro del simulador ndnSIM, todo esto con el fin de diseñar una arquitectura de cache que no exista en la literatura.

Capítulo 2

Descripción del problema

2.1. Contextualización

-¿Que es el cache y para que sirve?

2.2. El problema

2.2.1. Declaración del problema

2.2.2. Diagrama de ishikawa

2.3. Objetivos

2.3.1. Objetivo general

Optimizar la eficiencia del cache en redes *ICN*, por medio del diseño de una arquitectura caché basada en el comportamiento de los usuarios, dentro del simulador *ndnSIM*.

2.3.2. Objetivos específicos

- Determinar los componentes estructurales de las redes *ICN*.
- Proponer una arquitectura caché para redes *ICN* que considere el comportamiento usuario.
- Incorporar una arquitectura cache en el simulador *ndnSIM*
- Comparar el rendimiento de la estrategia caché propuesta con estrategias de la literatura

2.3.3. Alcance de los objetivos

Capítulo 3

Marco teórico

3.1. Information Centric Network

Como una futura propuesta de arquitectura de Internet, ICN (Information-Centric Networking) pretende motivar la transición arquitectónica de la actual arquitectura centrada en el host a centrada en la información para difundir de manera eficiente y flexible la enorme información generada por una variedad de aplicaciones. En esta dirección de investigación, se han propuesto muchos enfoques durante los últimos años (Ver Figura 3.1)

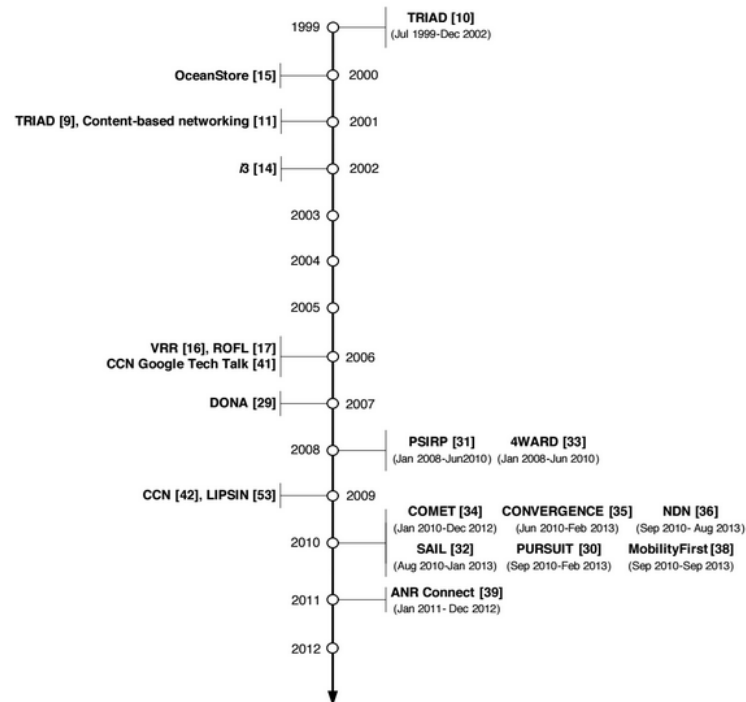


Figura 3.1: Nodo CCN

Dicho lo anterior, para el desarrollo de este trabajo, nos centraremos en la investigación NDN la cual pretende mover el Internet hacia el futuro con un diseño centrado en el contenido el cual es capaz de realizar una distribución de contenido eficiente y soporte de movilidad sin problemas.[11]

Así mismo la principal abstracción de ICN es el NDO (Named Data Objects), la que se define como la unidad de datos direccionables que se encuentra dentro de una red centrada en la información, siendo capaces de representar una colección de bytes o una información en específica. Así mismo cada uno de los objetos de datos existentes posee un nombre vinculado a el, pudiendo verse como fragmentos de datos nombrados sin semántica alguna caracterizándose por su granularidad la cual varia desde el tamaño del paquete hasta los objetos completos. [1][4]

Los objetos de datos con nombre se desligan de la ubicación, la aplicación, el almacenamiento y los medios de los transporte, dando paso al almacenamiento en cache y un bajo costo y ubicua replicación en la red, esperando como resultado una mayor eficiencia y seguridad, una mayor escalabilidad respecto a la demanda de información/ancho de banda y una superior robustez en escenarios de comunicación desafiantes[1].

Volviendo al tema que nos ocupa, las redes ICN se componen de dos tipos de paquetes, los paquetes de interés y los paquetes de datos (Figura 3.1), los cuales son enviados por dos tipos de usuarios el primero denominado consumidor, entidad encargada de enviar una solicitud de un NDO a la red (Interest) y por otra parte un productor, que es la entidad encargada publicar NDO's en la red, los cuales serán solicitados por un consumidor y enviados (Data)

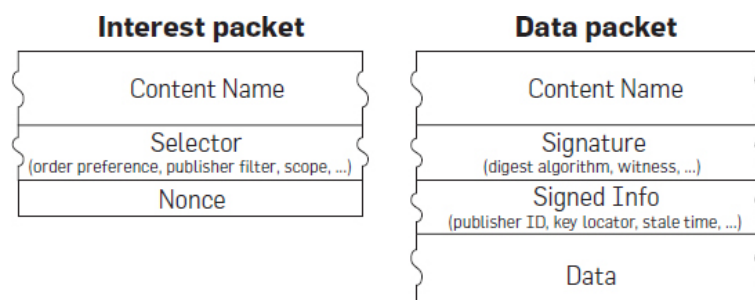


Figura 3.2: Paquetes CCN

Ademas este tipo de redes también cuentan con nodos intermediarios entre los consumidores y los productores, lo que le entrega a la red un aumento en la replicación de información y consigo una mayor resolución de consultas. Los nodos contienen tres estructuras de datos principales(Figura 3.2):

- **Forwarding Interest Table (FIB):** El FIB se utiliza para reenviar paquetes de interés hacia fuentes potenciales de datos coincidentes. Es casi idéntico a un IP FIB, excepto que permite una lista de caras salientes en lugar de una sola. Esto refleja el hecho de que CCN no está restringido a reenvío en un árbol de expansión. Permite múltiples fuentes de datos y puede consultarlas todas en paralelo.[3]
- **Almacen de contenido (Content Store (CS):** El almacenamiento para almacenar en caché NDO es una parte integral del servicio ICN. Todos los nodos potencialmente tienen cachés, incluyendo nodos en redes de infraestructura de ejecución de operadores y redes domésticas dirigidas por usuarios, así como terminales móviles. Las solicitudes de NDO pueden ser satisfechas por cualquier nodo que contenga una copia en su caché. ICN combina así el almacenamiento en caché en el borde de la red, como en P2P y otras redes de superposición, con almacenamiento en caché en la red (por ejemplo, cachés de web transparentes). El almacenamiento en caché es genérico, es decir, es independiente de la aplicación y se aplica a todos los proveedores de contenido, incluido el contenido generado por el usuario.[3]
- **Pending Interest Table (PIT):** El PIT realiza un seguimiento de los Intereses enviados hacia arriba hacia la(s) fuente(s) de contenido de forma que los datos devueltos puedan ser enviados aguas abajo a su(s) solicitante(s).[3]

3.2. Caché

3.3. Comportamiento de usuario

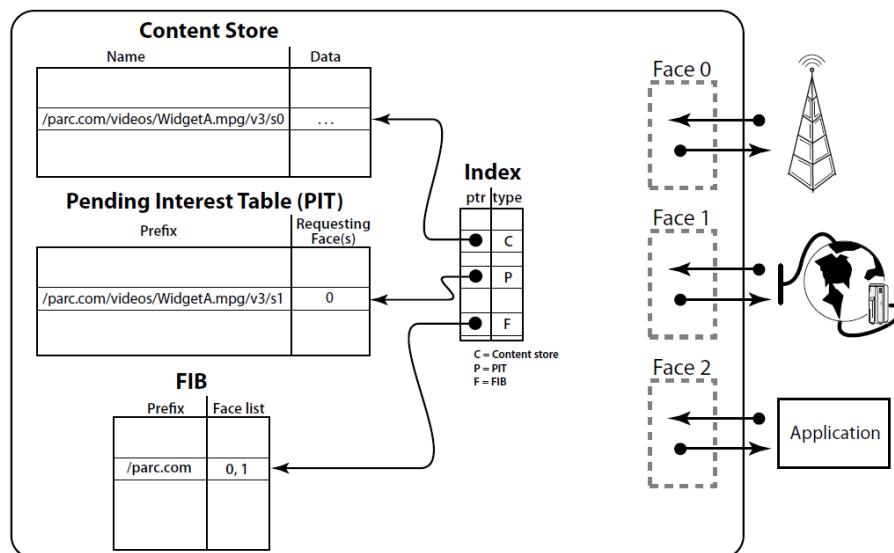


Figura 3.3: Nodo CCN

Capítulo 4

Estado del arte

Bibliografía

- [1] KUTSCHER, D., EUM, S., PENTIKOUSIS, K., PSARAS, I., CORUJO, D., SAUCEZ, D., ... & WAEHLISCH, M., *Information-centric networking (ICN) research challenges*. 2016
- [2] P. MÜLLER Y B. REUTHER, *Future Internet Architecture—A Service Oriented Approach*, vol. 50, no 6, pp. 383-389, 2009.
- [3] JACOBSON, V., SMETTERS, D. K., THORNTON, J. D., PLASS, M. F., BRIGGS, N. H., & BRAYNARD, R. L., "*Networking named content.*", Proceedings of the 5th international conference on Emerging networking experiments and technologies. ACM, 2009.
- [4] AHLGREN, BENGT, ET AL., "*A survey of information-centric networking.*", IEEE Communications Magazine 50.7 (2012).
- [5] JOKELA, P., ZAHEMSZKY, A., ESTEVE ROTHENBERG, C., ARIANFAR, S., & NIKANDER, P., *LIPSIN: line speed publish/subscribe inter-networking.*, ACM SIGCOMM Computer Communication Review (2009).
- [6] B. AHLGREN, M. D'AMBROSIO, C. DANNEWITZ, A. ERIKSSON, ET AL., *Second netinf architecture description*, 4WARD EU FP7 Project, Deliverable D-6.2 v2. 0. (2010).
- [7] AHLGREN, B., D'AMBROSIO, M., MARCHISIO, M., MARSH, I., DANNEWITZ, C., OHLMAN, B., ... & VERCELLONE, V. , *Design considerations for a network of information*, In Proceedings of the 2008 ACM CoNEXT Conference (p. 66). ACM. (2008, December).
- [8] PURSUIT, [HTTP://WWW.FP7-PURSUIT.EU/](http://www.fp7-pursuit.eu/)
- [9] KOPONEN, T., CHAWLA, M., CHUN, B. G., ERMOLINSKIY, A., KIM, K. H., SHENKER, S., & STOICA, I. , *A data-oriented (and beyond) network architecture.*,

In ACM SIGCOMM Computer Communication Review (Vol. 37, No. 4, pp. 181-192). ACM.(2007, August).

[10] NDN, [HTTP://WWW.NAMED-DATA.NET/.](http://www.named-data.net/),

[11] RAN, J., LV, N., ZHANG, D., MA, Y., & XIE, Z. , *On performance of cache policies in named data networking.*, In International Conference on Advanced Computer Science and Electronics Information (pp. 668-671). (2013, July).