

# Information Centric Networking

André Diegues - 201206858

Fábio Teixeira - 201305725

Tópicos Avançados em Redes - CC4037

Departamento de Ciencia de Computadores

Faculdade de Ciencias da Universidade do Porto

**Abstract**—Hoje em dia, o enorme aumento de tráfego de conteúdo e dados entre utilizadores na *Internet* motivou o desenvolvimento de novas ideias de arquitetura de *Internet* que resolvam eficazmente este problema. Uma delas é a que vamos abordar neste artigo, a *Information Centric Networking* (ICN), que através de uma abordagem de pesquisa de informação nas redes permite fornecer à rede um serviço mais resiliente a falhas que cumpre as exigências atuais de distribuição de conteúdo [1]. Vamos abordar o seu funcionamento, as arquiteturas que nasceram desta ideia e estudar se a mudança de arquitetura de *Internet* é ou não viável.

## I. INTRODUÇÃO

O paradigma ICN foi baseado numa primeira abordagem de arquitetura *Internet* denominada de TRIAD [2], cujo principal objectivo seria facilitar e aliviar cerca de 80% do tráfego de *Internet* que servia apenas para entrega de conteúdo através de uma arquitetura *data-centric*, isto é, um utilizador pede o conteúdo ao servidor em vez de pedir ao *host* que detém esse mesmo conteúdo [1]. A TRIAD define uma nova camada de conteúdo que está implementada por *content routers* que encaminham os pedidos aos *content servers* que, de seguida, fornecem o conteúdo .

O ICN procura substituir a arquitetura atual, que é um modelo de comunicação *host-to-host*, por uma arquitetura baseada num modelo *data-centric*, tratando o conteúdo como entidade principal na arquitetura das redes. Uma rede com este tipo de arquitetura ganha inúmeras vantagens em relação ao modelo *host-to-host*, nomeadamente, na distribuição de conteúdo, segurança e desenvolvimento de aplicações [3].

Existem várias arquiteturas baseadas neste paradigma. Neste artigo vamos abordar algumas destas arquiteturas, nomeadamente as que ganharam mais apoio ao longo do tempo:

- *Data-Object Network Architecture* (DONA)
- *Publish-Subscribe Internet Technology* (PURSUIT)
- *Named Data Networking* (NDN) baseada em *Content-Centric Networking* (CCN)

## II. COMO FUNCIONA O ICN?

O ICN utiliza o modelo de comunicação *Publish/Subscribe*. Neste modelo, os emissores, chamados *Publishers*, em vez de enviarem as mensagens diretamente aos recetores específicos, caracterizam as mensagens publicadas em classes sem

conhecimento de quem vão ser os subscritores. Por outro lado, os subscritores, chamados de *Subscribers*, demonstram interesse numa ou mais classes e apenas recebem mensagens da(s) classe(s) que subscreveram, também sem conhecimento de quem as enviou.

Neste novo paradigma, os dados tornam-se independentes de localização, memória e meio de transporte, permitindo o armazenamento e replicação em *cache*. Algo que impulsiona uma melhoria na eficiência, na escalabilidade e na robustez de comunicação em cenários difíceis. Existem várias abordagens do ICN, todas focadas num novo desenho da arquitetura da *Internet* atual, com o objetivo de substituir o atual modelo centrado nos *hosts*, para implementar um modelo mais orientado a dados e centrado nos conteúdos [4]. A saber:

### A. DONA

A resolução de nomes no DONA [5] é da responsabilidade de uns servidores especializados chamados *Resolution Handlers* (RHs). Existe pelo menos um RH em cada sistema autónomo. Estes *Handlers* estão interconectados, formando um serviço hierárquico de resolução de nomes por cima das relações interdomínio de *routing* existentes, para que seja possível conciliar resolução de nomes com *routing* de informação no mesmo sistema. O *Publisher* envia uma mensagem *REGISTER* com o nome do objeto para o seu RH local, que guarda um apontador para o *Publisher*. O RH depois propaga o *REGISTER* para os RHs que tem ligação, seguindo as rotas estabelecidas, guardando estes RHs o mapeamento entre o nome do objeto e o endereço do RH que encaminhou o registo.

Com isto, os *REGISTERs* são replicados nos RHs até aos *tier-1*, e já que estes estão conectados com os outros todos, o RH que está localizado nesse *tier* está ciente do que se passa em toda a rede. Para localizar um item, o *Subscriber* envia uma mensagem *FIND* para o RH local, que também propaga esta mensagem aos RHs parentes, até que é feito um *match* com o *tier-1*. Depois seguem-se os apontadores para encontrar o *Publisher*, já que o *tier-1* é conhecedor de toda a estrutura da rede.

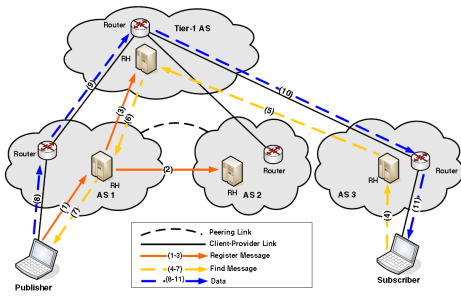


Fig. 1. DONA implementation.

## B. PURSUIT

Esta arquitetura substitui completamente a *stack* do protocolo IP por uma *stack* do protocolo *publish-subscribe*. Consiste em três funções distintas: *rendezvous*, *topology management* e *forwarding*. Quando o *rendezvous* encontra a subscrição de uma publicação, direciona a função de gestão da topologia para criar uma rota entre o *Publisher* e o *Subscriber*. Esta rota é usada pela função de encaminhamento para realizar a transferência de dados.

A resolução de nomes no *PURSUIT* é tratada pela função *rendezvous*, a qual é implementada por uma coleção de *Rendezvous Nodes* (RNs), pertencentes à *Rendezvous Network* (RENE). Quando um *Publisher* quer anunciar um objeto de informação, emite uma mensagem *PUBLISH* para o RN local, o qual é encaminhado por tabelas de *hash* distribuídas para o RN atribuído com o ID do *scope* correspondente. Quando o *Subscriber* emite um *SUBSCRIBE* para o mesmo objeto de informação ao seu RN local, é encaminhado pela tabela de *hash* para o mesmo RN. Depois, o RN informa um nó da *Topology Management* (TM) para criar uma rota ligando o *Publisher* ao *Subscriber* para entrega de dados. A gestão da topologia envia a rota ao *Publisher* numa mensagem *START PUBLISH*, que já utiliza a rota para enviar o objeto de informação via *Forwarding Nodes* (FNs).

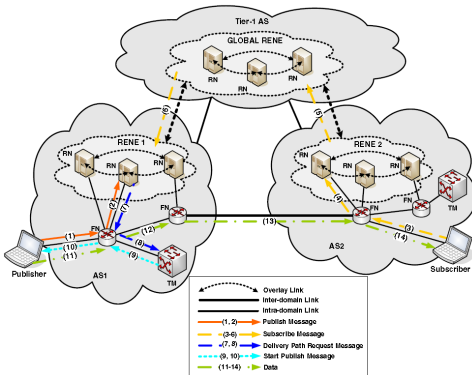


Fig. 2. PURSUIT implementation.

## C. NDN

Os *Subscribers* emitem mensagens *INTEREST* para pedir informação sobre objetos que chegam na forma de dados. As mensagens são encaminhadas por *Content Routers* (CRs), e cada um dos CRs mantém três estruturas de dados: *Forwarding Information Base* (FIB), *Pending Interest Table* (PIT) e *Content Store* (CS).

O FIB mapeia informação para as interfaces de saída que devem ser usadas para encaminhar as mensagens *INTEREST*. O PIT segue as interfaces de entrada nas quais as mensagens de *INTEREST* chegou. Já o CS funciona como *cache* local para objetos de informação que passaram pelo CR. Quando um *INTEREST* chega, o CR extrai a informação do nome e procurar por um objeto nesse CS cujo nome coincida com o prefixo pretendido. Se for bem-sucedido, é imediatamente enviado através da interface de entrada numa mensagem *DATA* e o *INTEREST* é descartado. Senão, o *router* executa uma procura do prefixo mais longo no seu FIB, para decidir a direção do encaminhamento. Se a entrada for encontrada no FIB, o *router* regista a interface de entrada do *INTEREST* no PIT e empurra o *INTEREST* para o CR indicado pelo FIB.

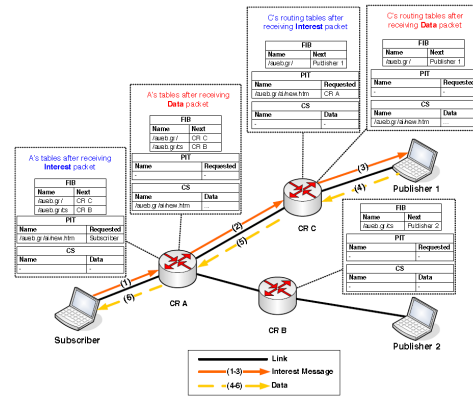


Fig. 3. NDN implementation.

## III. QUAIS OS OBSTÁCULOS E CUSTOS DE IMPLEMENTAÇÃO O ICN?

### A. Requisitos

O ICN ainda está longe de estar pronto a ser implementado. Os *researchers* têm lidado com vários desafios para tornarem este novo modelo disponível para poder ser colocado ao serviço dos utilizadores. Entre os desafios que têm encontrado destacam-se:

#### 1) Naming dos dados

Dar nomes aos dados é tão importante para o ICN como dar nomes aos *hosts* é para a Internet de hoje. Sendo assim, o ICN requer nomes únicos para os *Named Data Objects* (NDOs), já que os nomes são

utilizados para identificar objetos independentemente da sua localização ou conteúdo. Usar tabelas de *hash* é uma forma possível de resolver este imbróglio, já que permite depois que se possa comparar isso com o próprio nome do componente.

## 2) Proteção da privacidade

Como a rede pode ver quem faz o pedido pela informação e já que a tendência do ICN é guardar a história dos utilizadores, torna-se um problema para o utilizador não ter garantias de privacidade, já que como os nomes devem ter um longo tempo de uso, seria uma limitação desperdiçar nomes.

## 3) Atualizações

Se um NDO pode ser replicado e guardado na rede para futuros tratamentos, os nomes têm de ter um prazo de validade longa e o conteúdo do nome não deve ser alterado, o que impossibilita a atualização de objetos.

## 4) Integridade dos dados

A verificação da integridade dos dados é um passo importante para a consolidação do ICN. O facto de os NDOs não só serem recuperados a partir da copia original como também a partir de qualquer ponto da rede em que estejam guardados em cache e de poderem ser modificados faz com que não se possa confiar a 100% na integridade dos dados. Utilizar uma *hash* como parte do nome de objeto é também uma possível solução deste problema, embora a utilização de chaves criptográficas seja melhor aplicada nestes casos.

## 5) Cifragem

possível cifrar NDOs no ICN e apenas os consumidores que tiverem as chaves podem aceder a esse conteúdo privado. No entanto distribuir e gerir estas chaves, tal como fornecer as interfaces para os utilizadores é ainda uma matéria de estudo.

## 6) Agregação e filtragem de tráfego

Uma mensagem de um pedido *request* para receber um objeto de dados pode agregar vários pedidos de vários consumidores. Esta agregação reduz o tráfego na rede, mas torna a filtragem mais difícil. O desafio neste caso é fornecer um mecanismo que pretenda agregação, mas ao mesmo tempo uma pré-filtragem dos *request* dos utilizadores. Uma possível solução é indicar o conjunto de utilizadores que fizeram o *request* nesse *request* agregado, permitindo assim gerar numa resposta apenas o subconjunto dos utilizadores que fizeram *request* e têm acesso aos dados. No entanto, esta solução requer

a utilização de outros nós na rede e não permite fazer *caching*.

## 7) Roteamento pelo nome

Uma vez que o número de objetos de dados tem tendência a aumentar, o tamanho das tabelas de encaminhamento é um problema a pensar, pois pode ser proporcional ao número de objetos de dados, a menos que seja introduzido um mecanismo de agregação. Por outro lado, o *Route-By-Name Routing* (RBNR) reduz a latência e simplifica o processo de roteamento devido à omissão do processo de resolução.

### B. Vantagens vs Custo de Implementação

### C. Testes de trabalhos anteriores

### IV. ICN: O FUTURO DA Internet

Os utilizadores estão cada vez mais interessados em receber conteúdos, seja qual for a sua origem, do que ter de aceder a um servidor para receber essa informação. E o facto de a Internet ainda ser centrada nos *hosts* implica que o utilizador tenha de especificar em cada pedido não só a informação que deseja receber, como também especificar o servidor do qual a informação pode ser retirada. Com o ICN isso já não acontece [4].

O pressuposto básico por trás do ICN é que a informação é nomeada, endereçada e encontra o conteúdo independentemente da sua localização. Uma implicação indireta da implementação do ICN é que a informação se torna orientada para o recetor, em contraste com a atual realidade da Internet em que os emissores têm controlo total sobre os dados trocados [6]. No ICN só são recebidos os dados que o recetor tenha pedido. Depois de ser enviado o pedido, a rede é responsável por localizar a melhor origem para fornecer a informação desejada ao recetor.

Quando um elemento da rede receber um pedido por conteúdo, pode ter duas ações: se estiver em *cache*, responde imediatamente com o conteúdo; se não estiver, faz um pedido aos elementos com os quais tem ligação e depois guardar em *cache* o conteúdo quando for encontrado [4].

Nesse sentido, e já que os conteúdos chegam de elementos da rede ao invés da origem, o desenho do ICN tem de garantir a segurança dos conteúdos, contrariamente à estrutura atual da Internet que se foca no caminho. Para isso, quem fornece os dados assina um modelo de segurança para que os elementos da rede e os consumidores apenas tenham de verificar essa assinatura para garantir a sua fiabilidade [7].

### V. CONCLUSÃO

The conclusion goes here.

## REFERENCES

- [1] B. Ahlgren, C. Dannewitz, C. Imbrenda, D. Kutscher, and B. Ohlman, "A survey of information-centric networking," *IEEE Communications Magazine*, vol. 50, no. 7, pp. 26–36, July 2012.
- [2] D. Cheriton and M. Gritter, "Triad: A new next-generation internet architecture," 2000.
- [3] "Guest editorial [information centric networking]," *China Communications*, vol. 12, no. 7, pp. iii–iv, July 2015.
- [4] G. Xylomenos, C. N. Ververidis, V. A. Siris, N. Fotiou, C. Tsilopoulos, X. Vasilakos, K. V. Katsaros, and G. C. Polyzos, "A survey of information-centric networking research," *IEEE Communications Surveys Tutorials*, vol. 16, no. 2, pp. 1024–1049, Second 2014.
- [5] *A Data-Oriented (and Beyond) Network Architecture*, vol. 37. Kyoto, Japan: ACM, 08/2007 2007.
- [6] S. Arianfar, P. Nikander, and J. Ott, "On content-centric router design and implications," in *Proceedings of the Re-Architecting the Internet Workshop*. ACM, 2010, p. 5.
- [7] A. Ghodsi, S. Shenker, T. Koponen, A. Singla, B. Raghavan, and J. Wilcox, "Information-centric networking: Seeing the forest for the trees," in *Proceedings of the 10th ACM Workshop on Hot Topics in Networks*, ser. HotNets-X. New York, NY, USA: ACM, 2011, pp. 1:1–1:6. [Online]. Available: <http://doi.acm.org/10.1145/2070562.2070563>