



**HAL**  
open science

# An Optimal Broadcast Algorithm for Content-Addressable Networks – Extended Version

Ludovic Henrio, Fabrice Huet, Justine Rochas

► **To cite this version:**

Ludovic Henrio, Fabrice Huet, Justine Rochas. An Optimal Broadcast Algorithm for Content-Addressable Networks – Extended Version. [Research Report] 2013. hal-00866228v1

**HAL Id: hal-00866228**

**<https://inria.hal.science/hal-00866228v1>**

Submitted on 26 Sep 2013 (v1), last revised 8 Oct 2013 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# An Optimal Broadcast Algorithm for Content-Addressable Networks – Extended Version

Ludovic Henrio, Fabrice Huet, Justine Rochas

**RESEARCH  
REPORT**

**N° 8375**

September 2013

Project-Team Oasis





# An Optimal Broadcast Algorithm for Content-Addressable Networks – Extended Version

Ludovic Henrio, Fabrice Huet, Justine Rochas

Project-Team Oasis

Research Report n° 8375 — September 2013 — 16 pages

**Abstract:** Structured peer-to-peer networks are powerful underlying structures for communication and storage systems in large-scale setting. In the context of the Content-Addressable Network (CAN), this paper addresses the following challenge: how to perform an efficient broadcast while the local view of the network is restricted to a set of neighbours? In existing approaches, either the broadcast is inefficient (there are duplicated messages) or it requires to maintain a particular structure among neighbours, e.g. a spanning tree. We define a new broadcast primitive for CAN that sends a minimum number of messages while covering the whole network, without any global knowledge. Currently, no other algorithm achieves those two goals in the context of CAN. In this sense, the contribution we propose in this paper is threefold. First, we provide an algorithm that sends exactly one message per recipient without building a global view of the network. Second, we prove the absence of duplicated messages and the coverage of the whole network when using this algorithm. Finally, we show the practical benefits of the algorithm throughout experiments.

**Key-words:** Broadcast, Peer-to-Peer, Content-Addressable Network

**RESEARCH CENTRE  
SOPHIA ANTIPOLIS – MÉDITERRANÉE**

2004 route des Lucioles - BP 93  
06902 Sophia Antipolis Cedex

## Un algorithme de broadcast optimal pour réseaux pair-à-pair de type CAN

**Résumé :** Ce document présente un nouvel algorithme de broadcast pour réseaux pair-à-pair de type CAN. Cet algorithme de broadcast est optimal dans le sens où tous les pairs ne reçoivent le message qu'une seule fois, sans connaissances globales. Après avoir introduit les étapes fondamentales de cet algorithme, ses principales propriétés sont exposées et prouvées. Des expériences à taille réelle montrent la validité de cet algorithme et ses bénéfices.

**Mots-clés :** Broadcast, Pair-à-Pair, Content-Addressable Network

## 1 Introduction

In this work, we are interested in *Structured Overlay Networks* (SONs) where peers are organised in a well-defined topology and resources are stored at a deterministic location. The underlying geometric topology is used by communication primitives and ensures their efficiency. We are interested in CAN (Content-Addressable Network) [1] P2P networks, where peers are organised according to a multi-dimensional cartesian space. This space is organised in a geometrical way; the geometrical organisation dictates the dependencies between peers, as we will see in Section 2.

This paper presents a broadcast algorithm for the CAN overlay network that prevents a peer from receiving the same message more than once. We call such a broadcast algorithm *efficient*, in the sense that it minimises the number of exchanged messages between peers. Of course, a broadcast algorithm also has to be *correct*, and reach every peer of the network.

In previous works, Bongiovanni and Henrio proved, using the Isabelle/HOL theorem prover, that an efficient broadcast protocol for CAN *existed* [2]. However, the algorithm that was exhibited to prove the *existence* of an optimal solution was naive and had a very high latency, making it unusable in practice. In this work, we are interested in the design and implementation of an effective broadcast protocol that, in practice, also has an acceptable latency.

The contribution of this paper is as follows:

- Firstly, we propose a new broadcast algorithm that greatly improves the state of the art.
- Secondly, we prove that this algorithm is both correct (it covers the whole network) and optimal in terms of exchanged messages.
- Thirdly, we set an experimental comparison of the algorithm with others in a realistic distributed environment and show its efficiency in practice.

This paper is organised as follows. First, Section 2 will show that several broadcast algorithms exist for CAN but none of them was able to completely remove duplicated messages purposes. Section 3 will present our broadcast algorithm, together with its proof of efficiency and correctness. Section 4 will present the evaluation of our algorithm over a distributed peer-to-peer network. Finally Section 5 concludes this paper.

## 2 Related Works and Objectives

### 2.1 Context and Motivation

A *CAN* [3] is a structured P2P network based on a  $d$ -dimensional Cartesian coordinate space labeled  $\mathcal{D}$ . This space is dynamically partitioned among all peers in the system such that each node is responsible for storing data, in the form of  $(key, value)$  pairs, in a sub-zone of  $\mathcal{D}$ . To store a  $(k, v)$  pair, the key  $k$  is deterministically mapped onto a point in  $\mathcal{D}$  and the value  $v$  is stored by the node responsible for the zone comprising this point. The search for the value corresponding to a key  $k$  is achieved by applying the same deterministic function on  $k$  to find the node responsible for storing the corresponding value. These two mechanisms are performed by an iterative routing process starting at the query initiator and which traverses its adjacent neighbours (a peer only knows its neighbours), and so on and so forth until it reaches the zone responsible for the key to store/retrieve. One can find several definitions for a valid CAN, i.e. which shape can the zone of each peer have and how peers can be organised (see [2]). Here we rely on a very generic and simple definition: each zone is an hyperrectangle, and the only structure is the neighbouring relation: each peer only knows the peers whose zones are adjacent to its own zone. Additionally, a CAN is a torus, and the peers on the left border know the ones