Paper abstract: SPRAY protocol

Jérémy Bardon Anthony Pena

***Abstract*—Under the capstone project in master ALMA, we present you an abstract of one article related to our work. SPRAY [3] is an adaptive random peer sampling protocol used in CRATE, the million user live collaborative editor we contributed to.**

I. Introduction

One of the big concerns today is to connect everyone together. WebRTC helps to handle browser-to-browser communication, however it raises inefficient and unreliable peer-sampling mechanism.SCAMP [1] and CYCLON [2] are two solutions similar to SPRAY because of their concepts based on neighbor-to-neighbor interactions to solve this problem.

Browsers can be used with small devices and these solutions are not compatible with a large networks. Indeed, the number of connections will increase quickly with the number of peers.SPRAY is designed to keep a number of connection as low as possible and handle failures easily.

This paper is organized in 4 sections: WebRTC protocol description, presentation of the scientific problem, SPRAY features explanations and experimentation's conclusions.

II. Context

WebRTC is the technology used to make real-time pear-to-pear communication possible on large networks. To connect browsers together, WebRTC is based on hand-checking mechanics with a third party service.

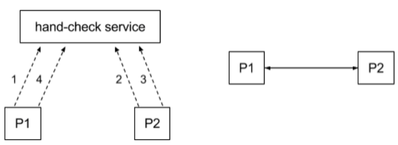

Fig. 1. WebRTC hand-check with two peers

This protocol does not manage routing so each pear have to know all other pears in the network to communicate. It means if a third peer will have to connect itself with P1 and P2.

Random peer sampling protocols as CYCLON and SPRAY helps to keep the number of connection as low as possible in order to reduce the traffic and limit resources consumption. These solutions involved that each peer have a partial view of the entire network.

III. Motivation

With the complexity of protocols using adaptive random peer sampling, the following scientific problem emerged:

Problem Statement 1. *Let t be an arbitrary time frame, let $\mathcal{N}^t$ be the network membership at that given time t and let $\mathcal{P}^t_x$ be the partial view of peer $p_x \in \mathcal{N}^t$. A cost-efficient random peer sampling should provide the following best-case properties:*

1. *Partial view size:* $\forall p_x \in \mathcal{N}^t,\ |\mathcal{P}^t_x| = \Theta(\ln |\mathcal{N}^t|)$
2. *Connection establishment:* $O(1)$

The first condition defines that the partial view of each peer will grow logarithmically with the network size. In addition, the second condition ensures a constant complexity for each connection establishment no matter the number of peers.

IV. Spray

Inspired by SCAMP and CYCLON , SPRAY is an adaptive random peer sampling protocol that manages connections between peers.

Joining is the first step in SPRAY . When a new peer comes in the network, this step will make sure the number of arcs will be fine according to the network size. After, a shuffling periodic protocol is run in any peer to uniform the peers partial view knowledge and balancing. Finally, SPRAY ensures that peers leaving and crashing are handled correctly through the network.

A. Joining

Each peer in the network, have a partial view of all other peer (their neighbors). On a new connection, the receiver will add the new peer to his partial and forward to his neighbors. Thanks to the gossip-concept, every peer in the network will know the new peer so the second condition of the scientific problem (Statement III) is verified.

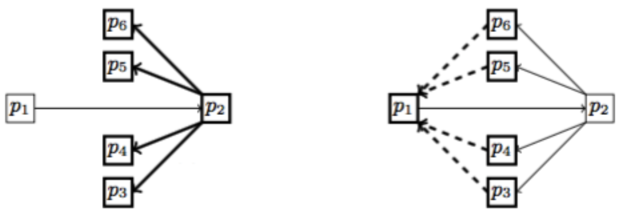

Fig. 2. Example of P1 joining a network *(extracted from* SPRAY *paper)*

In the example below, we can see that P1 joined the network via P2 peer. After the joining step, all peers in the networks have a direct connection with P1.

Partial views grows during this step, they handle duplicates and for each peer an *age* is associated to accelerate the shuffling step.

The peer is added in other network peers' partial view. However the view of the new peer is unbalanced and breaks the first condition of the scientific problem (Statement III).

B. Shuffling

This periodic shuffling protocol aims to balance peers partial view size and randomly mix their neighborhood.

Consider a peer *p* which choose its oldest know peer *q* and then exchange half of its partial view with. The list sent to *q* is randomly filled with *p* partial view peer. In addition, one occurrence of *q* is removed and one occurrence of *p* is added before sending.

Once the list received, all occurrences of *q* are replaced with *p* to avoid self-loop issues. Finally, both of them remove the sample they sent and add the received one.

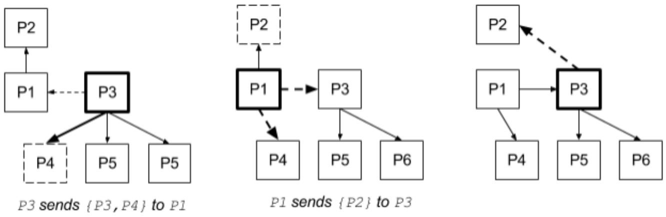

Fig. 3. Example of P3 shuffling with P1

After the shuffling protocol execution on *P3* and *P1*, both partial views are balanced. Indeed, *P1* view increase from 1 to 2 peers and *P3* view decrease from 4 to 3 peers.

With this algorithm, SPRAY can verify the first condition of the scientific problem (Statement III). As a result, it provides adaptiveness of partial views for peers in the network.

C. Leaving and crashing

Once SPRAY is able to deals with joining and peer balancing, it is necessary to focus on leaving and crashing peer issues.

Imagine a peer *P1* with 4 peers in his partial view. If *P1* leaves, it will remove all arcs in *P1* partial view plus arcs pointed to *P1* in peers it was connected to. Now, look at the numbers of arcs in the network. It decreases too much considering the added arcs when *P1* joining.

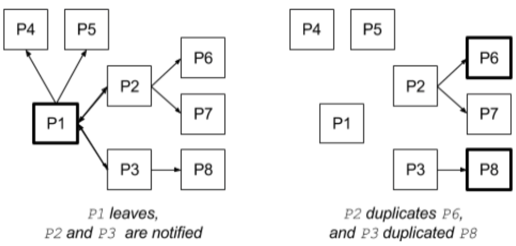

Fig. 4. Example of P1 leaving with P2 and P3 notified

When the peer *P1* leaves, connections to *P2-5* are removed. *P3* and *P4* are notified because of their dual link with *P1*. To keep partial views balanced, SPRAY replaces occurrences to the lost peer randomly with a know peers.

V. EXPERIMENTATION

In order to demonstrate capabilities of SPRAY protocol, the original paper contains 7 experimentations. For most of them, SPRAY results are compared with similar existing protocols SCAMP and CYCLON .

A. Clustering coefficient

Observing SPRAY adaptiveness with 0.1k, 1k, 10k and 100k peers shows SPRAY convergence time of clustering coefficient is lower than CYCLON a the beginning. However, after an average of 7 cycles SPRAY clustering already converged when CYCLON takes more depending on peers number.

B. Average shortest path length

With measurements performed after network convergence for 0.1K to 100K peer it demonstrates that SPRAY and CYCLON are equivalent with best CYCLON configuration for each test. The strength of SPRAY is its partial view which gives better performances on large networks than on small networks.

C. In-degree distribution

To show network robustness, it is possible to observer the number of references to a single peer among other peers partial view. Independently of the network size, CYCLON gives a total of 7 but SPRAY can switch from 2 to 12 depending on its the network size.

D. Dynamic network

This experiment is focused on number of connections and partial views size variation while peers are added and removed randomly. For both of these measurement, SPRAY have better results: 1k to 2.5k fewer connections and partial view size is up to 2 peers lower than CYCLON results.

E. Massive failure

Robustness can be demonstrated by removing 5% of peers in cycles and measure the number of weak and strong components in the network. Due to SPRAY randomness and duplicate partial view, results are a bit lower than CYCLON but still really close to a degrade at 70% of removed peers.

F. Duplicates in partial views

The birthday paradox concern is to estimated the number of people which have the birthday in a sample of *n* choosen persons. Compared with the amount of duplicates in peers partial views in SPRAY , the result is there are less duplicates while network size grows.

G. Failures in connection establishment

VI. CONCLUSION

Compared to similar adaptive-by-design random peer sampling protocol as CYCLON , experiments shows SPRAY is much better and adaptive to network variations. The principal strengths of SPRAY are on the average shortest path and the in-degree which scales better, evolves with network and converges faster.

REFERENCES

[1] A. Ganesh, A.-M. Kermarrec, ad L. Massoulié. Peer-to-peer membership management for gossip-based protocols. Computers, IEEE Transactions on, 52(2):139–149, Feb 2003.
[2] S. Voulgaris, D. Gavidia, and M. van Steen. Cyclon: Inexpensive membership management for unstructured p2p overays. Journal of Network and Systems Management, 13(2):197–217, 2005.
[3] B. Nédelec, J. Tanke, D. Frey, P. Molli and A. Mostefaoui. SPRAY: an Adaptive Random Peer Sampling Protocol, 2015.