

Collaborative Authoring of Documents with Eventual Consistency

Name(s) Withheld for Review

November 2014

Abstract

We describe a system for authoring and editing documents by many participants in a distributed fashion. Traditional document authoring tools permit only a single user to change modify a document at a time, inhibiting collaboration, especially among large teams. The move to the cloud has brought TODO: talk about cloud computing several commercial systems for collaborative document editing, but these systems struggle to handle massive numbers of concurrent . To address these shortcomings, we describe a system that allows an arbitrary number of users to edit a document. Each user can make arbitrary edits and see the results immediately, and the system guarantees eventual consistency of all edits made by users of the system. We provide implementation details We outline the design of our system, compare it to existing alternatives we should say how we're going to evaluate this in the literature and in industry, and evaluate different use cases the performance of the system under

1 Introduction

Take abstract draft and make it longer

2 Related Work

* cite Dynamo, of course * talk about Google Docs * mention paxos, but we're targeting web-scale so this is too expensive, etc. * we can nail this down later

3 Document Editing Model

Each edit consists of the insertion or removal of a contiguous sequence of characters at some location in the document. This model suffices to express more complicated operations; for example, moving a section of text corresponds to a

deletion of the section of text followed by an insertion of the same text at the new location.

This is made complicated by the fact Consequently, an absolute offset to the location of an edit in a document as found at one replica may correspond to a different section of text in the version at another replica, even if the intended location of the edit in this version would be readily apparent to a human reader. We would like to thank our shepherd

6 Results

We now describe several experiments evaluating the performance of our system relative to under a synthetic workload delivered by 100 users, 30 percent of whom edit individual sentences, 50 percent of whom delete or rearrange words or phrases within a sentence, and 20 percent of whom replace or move entire paragraphs. of such artifacts is future work.

4 Distributed Editing and Consistency

Pending edits to a document are propagated to the other replicas, and are evaluated relative to the local state of the document at the current replica. If any edits with an earlier timestamp subsequently arrive from another replica, any later-timestamped edits are rolled back, the newly received edit is applied, must be taken to ensure that they are applied in the correct location.

5 roughly 15000 lines of C++ for the document servers, in addition to 4000 lines of Javascript for the client-side editor interface

7 Conclusion

8 References

8 References