Freenet-like GUIDs for Implementing Xanalogical Hypertext

Tuomas J. Lukka
Hyperstructure Group
Dept. of Mathematical Information Technology
University of Jyväskylä, PO. Box 35
FIN-40351 Jyväskylä
Finland
lukka@iki.fi

Benja Fallenstein
Oberstufen-Kolleg
University of Bielefeld, PO. Box 100131
D-33501 Bielefeld
Germany
b.fallenstein@gmx.de

ABSTRACT

We discuss the use of Freenet-like content hash GUIDs as a primitive for implementing the Xanadu model in a peer-to-peer framework. Our current prototype is able to display the implicit connection (transclusion) between two different references to the same permanent ID. We discuss the next layers required in the implementation of the Xanadu model on a world-wide peer-to-peer network.

Categories and Subject Descriptors

H.5.4 [Information Interfaces and Presentation]: Hypertext/Hypermedia—architectures; H.3.4 [Information Storage and Retrieval]: Systems and Software—distributed systems, information networks; C.4 [Performance of Systems]: [fault tolerance, reliability, availability, and serviceability]

General Terms

Design, Reliability, Security, Performance

Keywords

P2P, Xanadu, Permanence, Transclusion

1. INTRODUCTION

(Non-)Permanence of content is a growing problem on the Internet, and various attempts have been made to alleviate it[4, 2, 12].

The Xanadu hypermedia model[7, 8, 9] handles the problem by assigning a permanent ID to content when it first enters the system. The immutable, permanent ID relates to the physical act of entering the smallest units of data, as in “the character 'D' typed by Janne Kujala on 10/8/97 8:37:18". Mutable documents are represented as virtual files containing lists of permanent media IDs. Copying content from one document to another is done by referring to the same permanent IDs.

In the Xanadu model, all virtual files containing a given piece of content (transclusion) can be found and displayed to the user.

For example, an email quoting another email would be automatically and implicitly connected to the original via the transclusion. Bidirectional, non-breaking external links (content linking) can be resolved through the same mechanism. Nelson[9] argues that conventional software, unable to reflect such interconnectivity of documents, is unsuited to most human thinking and creative work.

In order to implement the Xanadu model, it must be possible to efficiently search for references to permanent IDs on a large scale. The original Xanadu design organized content IDs in a DNS-like hierarchical structure (tumblers), making content references arbitrary intervals (spans) in the hierarchy. Advanced tree-like data structures[6] were used to retrieve the content efficiently. Unfortunately, Project Xanadu’s implementation was never finished and it is unclear whether the tumbler model can be implemented securely to avoid e.g. spoofing attacks.

In this article, we discuss a less ambitious structure based on Freenet-like GUIDs. This allows distributed hashing[11, 14] to be used for looking up references to permanent IDs.

2. FREENET: GLOBALLY UNIQUE IDS

Freenet[1, 2] is a decentralized P2P (peer-to-peer) architecture for anonymous uncensorable publishing. Data is stored in immutable sequences of bytes which are identified by SHA-1 cryptographic hashes[10] of their contents. Since in practice SHA-1 hashes are unique, they can be assigned as Globally Unique IDs (GUIDs) without needing a central naming authority.

GUIDs are like Uniform Resource Names[13]: they do not specify a physical storage location but an identity. In such a system, the primitive operation is not “get me file X from location Y” but “get me the data X, wherever it may be”.

3. XANADU-MODEL HYPERTEXT ON TOP OF FRENET-LIKE GUIDS

While Freenet focuses on anonymity and uncensorability, the Xanadu model focuses on external linking, implicit connections via transclusions, and permanence. Despite the differences, Freenet-like GUIDs provide a useful lowest layer of abstraction for implementing the Xanadu model.

In this model, a block of media content (e.g. keystrokes) obtains a GUID when first saved to permanent storage. The byte sequence also contains metadata such as the author and the creation time. All documents transcluding the saved keystrokes use the GUID and an offset inside the block.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

HT’02, June 11-15, 2002, College Park, Maryland, USA.
Copyright 2002 ACM 1-58113-477-0/02/0006 ...$5.00.

1Metadata that is subject to change should naturally be stored outside the immutable block; in our model, it would be stored in other immutable blocks with a revocation or expiration mechanism.
of hypertext publications. The focus+context[3] view shows two things as the context of a transclusion from a PDF file: users’ annotations, which have explicitly been connected to it; and different transclusions of the same content, only implicitly connected to the focused transclusion through the Xanadu model.

This functionality is currently in a pre-alpha stage, but we expect to release a first version of it as a part of Gzz 0.8.0 (“chartreuse”) shortly (available at [5]).

5. CONCLUSIONS

We have shown how a subset of the Xanadu media model can be implemented using Freenet-like GUIDs. Our current prototype is still limited: to approach the full functionality of the Xanadu model, content links have to be implemented, and the system needs to be extended to 1) fetch data interactively through a P2P network, and 2) use distributed indices of references to permanent IDs.

The permanence of the Xanadu model affects performance in several ways: while checking hashes and assembling the content from different blocks may be slow, caching is on the other hand easier. More research on performance issues both on a single computer and on a network is necessary.

In addition to the implementation of the Xanadu model, research is also needed on visualizing and navigating the structures arising from it.

6. REFERENCES