MoSAIC Meeting at Eurecom, February 15th, 2005

MoSAIC

Introduction

This document summarizes the MoSAIC meeting which took place at Eurecom, Sophia-Antipolis, France, on February 15th, 2005. The attendees were:

- Yves Roudier, Refik Molva, and Slim Trabelsi, PhD student at Eurecom;
- Michel Banâtre and Paul Couderc from IRISA;
- Ludovic Courtès, Yves Deswarte, Marc-Olivier Killijian and David Powell from LAAS-CNRS.

M-O. Killijian and L. Courtès stayed at Eurecom for a further day and a half after the meeting. This was an opportunity to discuss topics that had not been discussed during the meeting, to further detail some of the ideas expressed earlier, and to carry out brainstorming sessions with several people from Eurecom.

The first section below summarizes the discussions held on February the 15th. The next section lists the discussions held during the next two days.

1. Meeting Contents

This section summarizes the main discussions held on February the 15th.

1.1. Internet-based Cooperative Backup Systems

The meeting started at 10am with a talk of Ludovic Courtès presenting works in the area of collaborative backup over the Internet, as well as the result of several brainstorming sessions at LAAS regarding MoSAIC's desirable features and overall design. The slides of the talk are available online at *http://www.laas.fr/mosaic/papers.html*. The talk basically summarizes the contributions of various Internet-based cooperative backup systems such as *Pastiche*, and *PeerStore* [2,3,5,7].

The talk also sketched the "mailbox" or "repository" abstraction which has been thought of in previous meetings at LAAS. The *mailbox* or *repository* is basically a mechanism allowing for asynchronous data retrieval. A data saver ultimately sends backed up data to their owner's repository. The owner may eventually access its repository in order to restore his/her data. This abstraction fits both the *push* and *pull* models presented in the MPAC paper [6], depending on whether the repository is close to the owner or close to the data saver.

1.2. Cooperating Incentives and Trust Establishment

Yves Roudier then had a talk describing his plans with respect to the implementation of both *cooperation incentives* and *trust establishment* in MoSAIC, following the ideas expressed in the project proposal (the slides are also available from the website). His proposal for cooperation incentives is to have data owners 'pay' data savers for whenever they want to store data for a certain time. Transactions could be made safe thanks to tamper-proof hardware and in particular smartcards that should be used on both the owner-side and the saver-side. Additionally, Yves proposes that a trusted third-party (TTP) implements a *reputation service* online. The goal of this service would be to allow any participant to know whether a given participant is dishonest. Therefore, if one cheats (e.g. by unexpectedly aborting a transaction), others can eventually learn about it when they connect to the TTP.

Yves went on describing the first implementation of these mechanisms students have been working on. It uses JavaCards and a USB smartcard reader device.

1.3. On the Attractiveness of the Cooperative Approach

In the afternoon, the attendees discussed the choice of a development platform. On one hand, several people emphasized the fact that the use of PDAs was seemingly on the decline while *smartphones* are becoming increasingly widespread.

This yielded the questioning of the point in having such fully connected devices perform ad hoc cooperative backup when they could simply use the infrastructure. Several arguments were given in favor of the cooperative approach:

1. access to the infrastructure of a smartphone (GPRS, UMTS, etc.) may be costly, while proximity communications (using 802.11, Bluetooth, etc.) are not;
2. similarly, short-distance communications such as 802.11 may require much less energy;
3. short-distance communications may as well offer a much higher throughput (e.g. GPRS allows up to 9600bps while 802.11g allows up to 54Mbps);
4. cooperative backup is inherently an interesting option because
   - it leverages excess resources (disk storage) which makes it cheap;
   - it can be easily set up since it doesn't require any infrastructure and is self-administered;
   - it can benefit from a high diversity of nodes with independent failure modes.

1.4. Development Platform

On the other hand, people at LAAS explained that GNU/Linux was considered a good choice as a first implementation platform for MoSAIC during previous meetings, notably because it is portable and already available on a wide range of hardware architectures (15 as of today). However, while there are GNU/Linux PDAs (such as Sharp's Zaurus) there are still few GNU/Linux-based smart-phones. However, several vendors announced the availability of such smart-phones shortly. Therefore, the choice of GNU/Linux as the main implementation platform will likely not prevent eventual experiments or demos on smart-phones.

1.5. MoSAIC Use Cases

Several relevant use cases of the cooperative backup system we envision were discussed.

Intermittent Internet access

> This scenario looks like the most common use case. In such a scenario, a mobile device usually backs up its data using opportunistic ad hoc connections. Data may then be restored *online*, using the Internet.

No Internet access

> In this scenario, both the back up and restoration processes are initiated using the ad hoc network. Ad hoc routing may optionally be used in order to allow for multi-hop communications. David Powell suggested that in such scenario, data owners may want to emphasize *data locality* by having back-up copies of their data kept by data savers available in their vicinity; data savers leaving an owner's vicinity may then be able to discard its data. Alternatively, there may be a *chain of contributors* (data savers) that contributed to the transport of a node's data all the way to the Internet.

2. Further Discussions

This section summarizes the various discussions and brainstorming sessions that took place on Feb. the 16th and 17th with Y. Roudier, M-O. Killijian, L. Courtès and other people.

2.1. Cooperation Enforcement

This section summarizes a brainstorming session held on Feb. 16th with Pietro Michiardi. The discussions tried to address the cooperation enforcement techniques that may be considered in MoSAIC. Pietro noted that, in order to protect from Sybil attacks (where the attacker benefits from the community-supported services by constantly renewing its identity), there must be no interest for an attacker in changing its identity. In other words, the reputation of a newcomer must be the lowest possible reputation so that a wrongdoer cannot wash its history by simply changing identities.

Cooperation may be enforced using some sort of "reward". This raised the question of *when* this rewarding should take place. Marc-Olivier and Pietro noted that there are four places where a contributor (a data saver) can be rewarded for its service:

1. a data owner can reward a saver *in advance*, when they meet for the first time;

2. a data owner can reward a saver *once they have agreed* on the terms of the storage service (amount of allocated storage, amount of time during which the saver guarantees to hold the data, etc.);

3. a data saver can be rewarded *once it has delivered the data to the owner's online repository*;

4. finally, a data owner can reward its contributor only once it has been able to effectively restore its data.

2.2. Peer-to-peer Backup

Ernst Biersack, from Eurecom, is currently working on a Microsoft-supported cooperative backup system for local area networks. He noted that Microsoft is already working on such a system, called BitVault. Ernst mentioned the contribution of works such as Venti [8] in the area of fragmentation, dissemination, and data sharing with secrecy, all of which are relevant to cooperative backup.

2.3. On Trust, Reputation, and Collaboration Incentives

In an attempt to try and define a generic programming interface between the backup subsystem and the « security » subsystems, Y. Roudier, M-O. Killijian and L. Courtès discussed the various mechanisms that may be used to solve MoSAIC's security concerns.

There are basically two security-related problems that need to be solved in systems that leverage cooperation among nodes with no prior trust relationship:

- *cooperation enforcement* in order to actually make participating nodes collaborate;

- *trust establishment* is needed to allow nodes to quantify the risks taken in collaborating with another given node.

In order to implement cooperation enforcement, Y. Roudier had proposed the use of a *credit mechanism* where contributors are "paid" for the service they give using "virtual coins". "Tamper-proof" hardware (smart cards) ensures that one cannot spend money one does not have. The trust issue needs to be solved in another way since there may be, for instance, wrongdoers that get paid and then do not honor the agreement made with a data owner, or there may be owners who decide not to pay for the service they were given, depending on when the transaction occurs (see section 2.1).

For this reason, Yves Roudier proposed the use of a *reputation mechanism* to help participating nodes know which nodes are trustworthy and which are not. Reputation information for each node would be updated by each node dealing with it, and then maintained and broadcasted by an online trusted third party (TTP).

In this scheme, both mechanisms are needed to fulfill the cooperation and trust requirements. However, M-O. Killijian and L. Courtès questioned the usefulness of a full-blown credit mechanism when a reputation mechanism is still needed. The idea is that good trust relationships may yield a good level of cooperation. For example, one may never refuse to cooperate with a good friend; however, collaborating with a stranger or a famous wrongdoer may obviously be much less systematic. Also, the more one interacts or cooperates with someone, the more one gets to trust or distrust him. Finally, the decision of whether to cooperate with someone is a function of both one's trust in the other and one's current situation: when one has lots of resources available, one does not mind cooperating with strangers since the risk is limited anyway; however, in times or resource scarcity, one may refuse to spend resources on behalf of a stranger.

Several papers discuss the use of a trust relationship or reputation model as a cooperation incentive [1,4]. These options look interesting, especially in the absence of a reachable central authority (TTP) as is the case in pure ad hoc mode. We agreed on the need to study related works more thoroughly.

References

[1] C. Grothoff. An Excess-Based Economic Model for Resource Allocation in Peer-to-Peer Networks. *Wirtschaftsinformatik*, 3-2003, June, 2003.

[2] E. Sit, J. Cates, R. Cox. A DHT-based Backup System. MIT Laboratory for Computer Science, August, 2003.

[3] J. Cooley, C. Taylor, A. Peacock. ABS: The Apportioned Backup System. MIT Laboratory for Computer Science, 2004.

[4] K. LAI, M. FELDMAN, J. CHUANG, I. STOICA. Incentives for Cooperation in Peer-to-Peer Networks. *Workshop on Economics of Peer-to-Peer Systems.*

[5] L. P. COX, B. D. NOBLE. Pastiche: Making backup cheap and easy. *Fifth USENIX Symposium on Operating Systems Design and Implementation*, 2002.

[6] M-O. KILLIJIAN, D. POWELL, M. BANÂTRE, P. COUDERC, Y. ROUDIER. Collaborative Backup for Dependable Mobile Applications. *Proceedings of 2nd International Workshop on Middleware for Pervasive and Ad-Hoc Computing (Middleware 2004)*, 2004.

[7] M. LANDERS, H. ZHANG, K-L. TAN. PeerStore: Better Performance by Relaxing in Peer-to-Peer Backup. *Proceedings of the Fourth International Conference on Peer-to-Peer Computing*, 2004.

[8] S. QUINLAN, S. DORWARD. Venti: A new approach to archival storage. *First USENIX conference on File and Storage Technologies*, 2002.