Topics for T-110.5190 Seminar on Internetworking
Tancred Lindholm

Note: the references given are merely initial pointers; you are free to focus on systems/papers not mentioned here (as long as you stay on topic).

The formation of ad-hoc data storage spaces in groupwork We look at a scenario where a group of people, for the most part previously unknown to each other, meet. The activity of the group calls for some sort of storage space, where the group may gather relevant data, such as hyperlinks, images, video clips, textual documents etc. By what means should the group set up such a group storage? What will be the underlying technologies for implementing the group space? How should we name data in such spaces, so that we may easily make references to it? Note that the group members, at least initially, do not have access to a common server infrastructure. References Bluetooth (`http://www.bluetooth.com`), Jini (`http://www.jini.org/`), Apple's Bonjour (`http://www.apple.com/macosx/features/bonjour/`), Service Location Protocol (`http://www.openslp.org/`) (service discovery), [Sat96, E⁺97, K⁺00] (shared storage), [BLFM98] (naming).

Data share models in environment with intermittent disconnections Continuing the groupwork scenario from the previous item, we ask what models for sharing data make sense in an environment where network connectivity cannot be assumed at all times (due to technological or monetary reasons[1]). The topic may be approached from several angles, such as: should users explicitly control when data is synchronized, what data should be cached locally (if any), when should data be synchronized, what should be done in cases of concurrent and conflicting updates? References: [Sat96, E⁺97, K⁺00] (shared storage systems), [Syn02, BP98, SS02, Lin03] (synchronization mechanisms, reconciliation).

Managing changing data in a decentralized/P2P architecture The most prevalent systems for groupwork use a centralized server that keeps the current, up-to-date state of the data. However, a centralized system may not always be possible, in which case it becomes much harder to maintain a consistent state for the data. In this topic,

---

[1]Some people may not want to pay for continuous network access!

we explore models for managing changing data in a decentralized architecture. References: [E$^+$97, K$^+$00, SS02, KWK03, P$^+$83].

Storage as a utility: scalability, security and archival It has been visioned that we will someday rent or buy the storage capacity we need from an operator in the "Internet cloud", and that local storage devices will merely act as local caches. What would such a system look like? Specifically, what are the issues of scalability, privacy, performance, durability and availability that need to be addressed, and how have these been addressed in past and current research systems? References: [E$^+$97, K$^+$00, Sat96, SS02]

References

[BLFM98] Tim Berners-Lee, Roy Fielding, and Larry Masinter. *RFC 2396: Uniform Resource Identifiers (URI): Generic Syntax*. Internet Engineering Task Force, August 1998.

[BP98] Sundar Balasubramaniam and Benjamin C. Pierce. What is a file synchronizer? In *Proceedings of the Fourth Annual ACM/IEEE International Conference on Mobile Computing and Networking*, pages 98–108, October 1998. Full version available as Indiana University CSCI technical report #507, April 1998.

[E$^+$97] W. Keith Edwards et al. Designing and implementing asynchronous collaborative applications with Bayou. In *Proceedings of 10th annual ACM Symposium on User Interface Software and Technology*, pages 119–128, October 1997.

[K$^+$00] John Kubiatowicz et al. Oceanstore: An architecture for global-scale persistent storage. In *Proceedings of the Ninth international Conference on Architectural Support for Programming Languages and Operating Systems*, November 2000.

[KWK03] Brent ByungHoon Kang, Robert Wilensky, and John Kubiatowicz. The hash history approach for reconciling mutual inconsistency. In *ICDCS '03: Proceedings of the 23rd International Conference on Distributed Computing Systems*, pages 670–670, Washington, DC, USA, 2003. IEEE Computer Society.

[Lin03] Tancred Lindholm. XML three-way merge as a reconciliation engine for mobile data. In *Third ACM International Workshop*

on Data Engineering for Wireless and Mobile Access, pages 93–97, September 2003.

[P⁺83]   Douglas Stott Parker, Jr. et al. Detection of mutual inconsistency in distributed systems. *IEEE Transactions on Software Engineering*, 9(3):240–247, May 1983.

[Sat96]   Mahadev Satyanarayanan. Mobile information access. *IEEE Personal Communications*, 3(1):26–33, February 1996.

[SS02]   Yasushi Saito and Marc Shapiro. Replication: Optimistic approaches. Technical Report HPL-2002-33, Hewlett Packard Laboratories, February 2002.

[Syn02]   SyncML Initiative. *SyncML Sync Protocol, version 1.1*, February 2002.