# The Ficus Scalable File System

Ficus Project
Computer Science Department
University of California Los Angeles

## Personnel

| | |
|---|---|
| Principal investigators | Gerald Popek |
| | Thomas Page |
| Research staff | Richard Guy |
| | Dieter Rothmeier |
| | Wai Mak |
| Students | John Heidemann |
| | Yuguang Wu |

## Contacts

Dr. Thomas Page
Computer Science Department
UCLA
Los Angeles, CA
90024-1596
page@cs.ucla.edu

Dr. Richard Guy
Computer Science Department
UCLA
Los Angeles, CA
90024-1596
guy@cs.ucla.edu

## Project Description

Ficus is a scalable distributed file system in development at UCLA. The design uses a "stackable layers" architectural paradigm and an optimistic approach to file replication as a basis for providing high quality filing services despite the unusually rich failure modes of wide area, large scale systems.

The stackable layers methodology allows custom filing services to be constructed in a dynamic fashion from building block "layers," each of which provides a relatively thin (narrow) set of services when compared to traditional filing mechanisms. Because stackable layers each support an identical calling interface, the order in a layer stack and stack composition are not determined *a priori* when a layer is implemented; instead, order and composition are functions of layers' semantics.

A number of layers have been designed and implemented. These include encryption, compression, file versions, measurement, token-based consistency for NFS, general optimistic replication (a pair of layers), and a transport layer to map a stack across address space boundaries (cross machine, user-to-kernel, kernel-to-user). The Ficus replication service currently leverages existing UFS and NFS layers to provide persistent storage and remote file access.

High availability for data access (both read and update) is provided by a pair of stackable layers that support optimistic data replication. Optimistic replication incorporates a *one copy availability* policy for general file access, while guaranteeing *no lost updates* semantics. Independent, uncoordinated, concurrent update to distinct replicas is reliably detected; such conflicting updates to directory files are automatically reconciled using a new *two-phase algorithm* family.

Ficus supports a large scale hierarchical name space by clustering sub-trees of files into AFS-reminiscent volumes. Ficus volumes are a primary unit of replication: a logical volume is represented by multiple physical replicas, each of which may contain a subset of the logical volume. Volumes are dynamically located and *grafted* using a novel *graft point* mechanism. Graft point consistency is automatically maintained by the normal Ficus reconciliation mechanism.

Ficus is currently operational at UCLA, and is the native environment for its developers. Testing outside the lab is expected later this year. Ficus is structured as a set of modules which can be linked to a standard SunOS 4.1 kernel. The Ficus project has been supported under DARPA contract number F29601-87-C-0072.

## Report Availability

Ficus technical reports and articles may be obtained by anonymous FTP from ftp.cs.ucla.edu [131.179.128.34] in the directory /pub/ficus. All files are compressed PostScript suitable for direct printing by a LaserWriter+ or compatible printer. Reports may also be acquired by mailing a request to ficus@cs.ucla.edu (please specify e-mail or hardcopy response) or writing to the contacts above.

## References

[1] R. G. Guy, "A replicated filesystem design for a distributed UNIX system," Master's thesis, University of California, Los Angeles, 1987.

[2] R. G. Guy and G. J. Popek, "Reconciling partially replicated name spaces," Tech. Rep. CSD-900010, University of California, Los Angeles, Apr. 1990.

[3] R. G. Guy, J. S. Heidemann, W. Mak, T. W. Page, Jr., G. J. Popek, and D. Rothmeier, "Implementation of the Ficus replicated file system," in *USENIX Conference Proceedings*, pp. 63–71, USENIX, June 1990.

[4] R. G. Guy, T. W. Page, Jr., J. S. Heidemann, and G. J. Popek, "Name transparency in very large scale distributed file systems," in *Second IEEE Workshop on Experimental Distributed Systems*, Oct. 1990.

[5] R. G. Guy and G. J. Popek, "Algorithms for consistency in optimistically replicated file systems," Tech. Rep. CSD-910006, University of California, Los Angeles, Mar. 1991.

[6] R. G. Guy, *Ficus: A Very Large Scale Reliable Distributed File System.* Ph.D. dissertation, University of California, Los Angeles, June 1991. Also published as CSD-910018.

[7] J. S. Heidemann and G. J. Popek, "An extensible, stackable method of file system development," Tech. Rep. CSD-900044, University of California, Los Angeles, Dec. 1990.

[8] J. S. Heidemann, "Stackable layers: an architecture for file system development," Master's thesis, University of California, Los Angeles, 1991. In progress.

[9] J. S. Heidemann and G. J. Popek, "A layered approach to file system development," Tech. Rep. CSD-910007, University of California, Los Angeles, Mar. 1991.

[10] T. W. Page, Jr., G. J. Popek, R. G. Guy, and J. S. Heidemann, "The Ficus distributed file system: Replication via stackable layers," Tech. Rep. CSD-900009, University of California, Los Angeles, Apr. 1990.

[11] T. W. Page, Jr., G. J. Popek, and R. G. Guy, "Stackable layers: An object-oriented approach to distributed file system architecture," in *Proceedings of the Workshop on Object Orientation in Operating Systems*, Oct. 1990.

[12] T. W. Page, Jr., R. G. Guy, G. J. Popek, and J. S. Heidemann, "Architecture of the Ficus scalable replicated file system," Tech. Rep. CSD-910005, University of California, Los Angeles, Mar. 1991.

[13] T. W. Page, Jr., R. G. Guy, G. J. Popek, J. S. Heidemann, W. Mak, and D. Rothmeier, "Management of replicated volume location data in the Ficus replicated file system," in *USENIX Conference Proceedings*, USENIX, June 1991.

[14] G. J. Popek, R. G. Guy, T. W. Page, Jr., and J. S. Heidemann, "Replication in Ficus distributed file systems," in *Proceedings of the Workshop on Management of Replicated Data*, pp. 20–25, IEEE, Nov. 1990.