Experience with Grapevine: the Growth of a Distributed System (Summary)

Michael D. Schroeder, Andrew D. Birrell, and Roger M. Needham*
Xerox Palo Alto Research Center
3333 Coyote Hill Road
Palo Alto, CA 94304, USA

Grapevine is a distributed, replicated system that provides message delivery, naming, authentication, resource location, and access control services in an internet of computers. The system, described in a previous paper [1], was designed and implemented several years ago. We now have had operational experience with the system under substantial load. This experience has proved the original design sound in most aspects, but there also have been some surprises. In this paper we report what we have learned from using Grapevine. Our experience may offer some help to designers of new systems.

Grapevine is implemented as a program that is run on a set of dedicated server computers. Client programs of Grapevine run on various workstation and server computers attached to an internet. The services provided by Grapevine are divided into the message service and the registration service.

The message service accepts messages prepared by clients for delivery to individual recipients and distribution lists. Messages are buffered in inboxes on message servers until the recipient requests them. Any message server can accept any message for delivery, thus providing a replicated submission service. A computer system mail user has inboxes on at least two message servers, thus replicating the delivery path for the user.

The registration service provides naming, authentication, access control and resource location functions to clients. The registration service is based on a registration data base which maps names to information about the users, machines, services, distribution lists, and access control lists that those names signify. The registration data is distributed and replicated.

The primary clients of Grapevine are various computer mail system interface programs. Some other clients of Grapevine implement file server authentication and access controls, remote procedure call binding, and process controls for an integrated circuit facility.

When we reported on Grapevine in the Fall of 1981 there were 5 Grapevine. servers. The registration data base contained about 1500 individuals and 500 groups. The total number of messages handled by all servers (presented for delivery by clients plus forwarded from Grapevine servers) was about 2500 messages each working day. Since then the system has grown considerably. By the Spring of 1983 there were 13 Grapevine servers. The registration data base contained about 3700 individuals and 1300 groups. Over 8000 messages were handled in a work day. The growth and heavy use of the system have enabled us to see how well the design has met its goals in actual operation.

The body of the paper contains observations based on operational experience with the system. These observations are divided into six general topics that are relevant to the design of most distributed systems. The topics are:

Effects of Scale -- An important objective of Grapevine is the ability to increase system capacity over a large range by adding more servers of fixed power, rather than by using more powerful servers. Our goal was expansion to a maximum system size of 30 Grapevine servers and the total load generated by 10,000 users. With the problems discussed in the paper repaired as suggested, the system appears to meet this specification. Different approaches would be necessary in some cases to expand the system much beyond that specification.

Configuration Decisions -- Running Grapevine requires making configuration decisions about how many servers to have, where to place then, and how to distribute inboxes and registry replicas among them. It has been hard to develop guidelines for making such decisions. Factors affecting the configuration are discussed and the heuristics we have developed for adjusting the configuration are presented.

* Roger M. Needham's regular address is:
Computer Laboratory, Corn Exchange Street,
Cambridge, CB2 3QG, United Kingdom

© 1983 ACM 0-89791-115-6/83/010/0141 $00.75

Transparency of Distribution and Replication -- Most users treat Grapevine as if it were implemented on a single, large, reliable computer that contained all registration data and inboxes, to which all workstations were connected through high-speed links. Our experience is that most of the time the unitary model serves users well. We discuss the few ways in which Grapevine's distributed, replicated implementation shows through now and then to surprise users.

Adjusting to the Load -- When building a system like Grapevine many design decisions are made based on assumptions about the nature of the expected load. While in most instances the original design decisions have stood up well, in a few cases our prediction about the detailed nature of the load proved to be wrong, sometimes leading to bizarre performance problems. Several examples are discussed.

Operation of a Dispersed System -- Because Grapevine is geographically dispersed, it is important for smooth and efficient operation to make monitoring, control and repair functions accessible through the internet. The facilities built into Grapevine to permit smooth operation under these conditions are presented in the paper.

Reliability -- A design objective of Grapevine is high reliability. A primary technique for achieving high reliability is replication of function among several Grapevine servers. When one Grapevine server is unavailable, others can perform the same functions for clients. The goal is that failure of a single Grapevine server not make any service unavailable to any client. Our experience in running Grapevine shows that this approach has been extremely successful. Most users do not notice when a server fails. One cause of failure has been the gradual, undetected consumption of the spare capacity required for redundancy to produce reliablilty.

The paper concludes with a brief description of the changes made in adapting the prototype Grapevine system for use as the Xerox 8000 NS product message system and clearinghouse.

[1] Birrell, A.D., Levin, R., Needham, R.M., and Schroeder, M.D. Grapevine: an exercise in distributed computing. *Comm. ACM 25*, 4 (April 1982), 260-274.