Vincent Cate, Thomas Gross, Garth Gibson, HT Kung

Wide Area Network Filesystems

OBJECTIVES

Transparent access to FTP files around the world

Good performance

Handle heterogeneous machine types

Easy to use

APPROACH

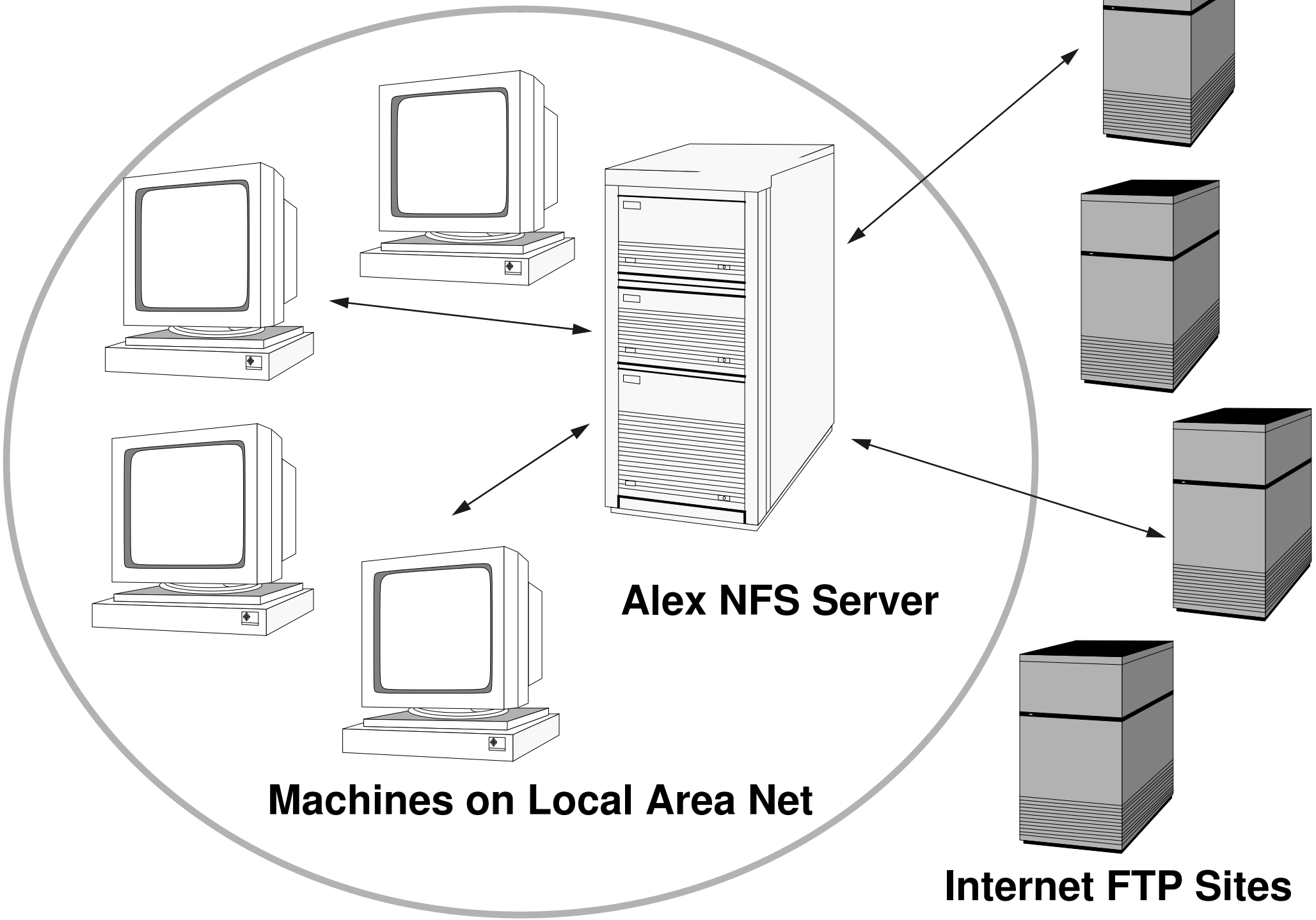
Make hostname into part of pathname:

FTP Host	Remote Path	Alex Path
ftp.cse.ucsc.edu	/pub/tr	/alex/edu/ucsc/cse/ftp/pub/tr

Extensive Long-Term Caching

Relaxed Cache Consistency Specification

APPROACH / IMPLEMENTATION



BACKGROUND

Stateless - NO callbacks

Read/Write ratio (Read-Only filesystem)

Heterogeneous - (commands used and dir returned)

Thousands of servers and millions of files already on Internet (and users who want to access these)

Lots of servers down at any one time

Whole file transfer

Bulk stats (whole directory at once) - very important

FTP, NFS, and UFS are interfaces Alex deals with

Files stored in FTP sites don’t change very often

RESULTS / DISCUSSION

About 50 users and growing

Can transparently access about:

- **3,000,000 files**
- **1 Tera-Byte**
- **3,000+ machines**
- **2,000+ CS Tech-Reports**
- **25,000+ README files**

Files stay in cache for about 6 weeks after last access

Estimated hit ratio above 75%

Different users accessing same files

CONCLUSIONS

A leap forward in terms of accessing remote data

Can easily access files all around the world

Combines well with other information technologies

Unique cache consistency constraints merit further study