CITI Technical Report 94–3

Workload Characterization in a Large Distributed File System

Sarr Blumson
sarr@citi.umich.edu

ABSTRACT

Most prior studies of distributed file systems have focussed on relatively small communities of users, typically university computer science departments. At the Center for Information Technology Integration (CITI) we conjectured that these studies led to an excessively homogeneous view of the workload presented to servers. We therefore embarked on an effort to study the University of Michigan Institutional File System, a campus-wide system with a much more diverse user community. The results partially confirm our conjecture.

July 11, 1994

Center for Information Technology Integration
University of Michigan
519 West William Street
Ann Arbor, MI 48103-4943

Workload Characterization in a Large Distributed File System

Sarr Blumson

1. Introduction

Performance analysis of distributed file systems (and of systems in general) is always complicated by the question of workload definition: is the load being presented to the system being measured or modelled actually representative of any portion of the real world? The literature includes a variety of reports [1, 2] that describe the workload in particular systems, but for obvious reasons, these tend to be on small networks to which the researchers had convenient access, typically research networks operated by university computer science departments.

To characterize the workload of a more heterogeneous system, CITI instrumented and studied a very large distributed file system accessible (at least in principle) by every member of the University of Michigan community. As a result, we were able to record some 47,000,000 file system operations over a period of several weeks.

We begin by describing the environment in which our measurements were performed and the mechanisms we used to collect our data. We then describe the methodology used in our analysis and the results obtained. We then draw some conclusions from these results.

2. Measurement Environment

The Institutional File System (IFS) Project [3] is a joint effort of the University of Michigan and IBM to provide a campus-wide, integrated file service for the University of Michigan community. This community is both large and diverse; there are tens of thousands of potential client workstations of many different types with variations in storage and processing power of several orders of magnitude, from small personal computers to large multiuser systems.

The primary file storage mechanism for this service is AFS [4] running on portions of a pair of IBM ES/9000 mainframes. Our data collection mechanism was to implement tracing code that recorded information about every file operation performed on these servers [5]. This recording was turned on over a period of several weeks in April 1993; this was a period of high activity at the University, occurring just before and during final examinations for the winter term. While the usual sorts of accidents caused some gaps in the traces, some 47,000,000 operations were recorded.

While this approach allows us to observe a large user community without having to add instrumentation to thousands of individual workstations, it also gives us a very filtered view of what individual users are doing. AFS clients use comparatively large local caches (50 to 70 megabytes is not unusual) and even very small client machines (e.g. MacIntoshs) are served by intermediate translators that perform a caching function [6]. As a result, the read activity that we were able to record reflects only those requests that could not be satisfied from a cache.

3. Methodology

One problem that confronted us is that AFS is a connectionless system with no concept of a user being "signed on" or of a file being open. Therefore we are unable to distinguish between times when a particular user is working or home asleep. At the same time, simply averaging a user's activity over the full three-week period would give a distorted view of individual contributions to the total workload. Because of this, we have limited our examination to operations performed during weekdays between 9 AM and noon and between 1 PM and 4 PM, periods during which (to a first approximation) everyone is active.

A second issue is the impossibility of reliably recognizing connected operations. AFS breaks operations on large files into 64 kilobyte chunks with no definite way to determine that two successive operations are, from the user's point of view, parts of the same operation. As a result, we have chosen as our measures of user activity:

the total number of each server operation performed,

the total number of bytes read and written, and

the number of times file status is read or updated.

Reads are further separated into directories and data files.

These operations represent the vast majority of user operations and, since reads and writes are the only operations which transfer nontrivial amounts of data, they represent the vast majority of network data transferred as well. On four of the six servers we measured, these operations constituted well over 75% of the server operations (and over 60% on a fifth server), while the only other operation appearing over 3% was time synchronization.

From this data we began to classify users into subgroups based on their position within the university (faculty, staff, student) and job classification (for staff) using public university records. We then computed the mean and standard deviation for each of our metrics within each subgroup, and studied the results.

4. Results

As a sanity check on the data, we first examined the cross correlations among the variables we measured. We then attempted to identify significant subgroups of our users.

4.1 Correlations

One interesting discovery comes from looking at the cross correlations between the different metrics; the correlation coefficients are summarized in Table 1.

TABLE 1. Cross Correlations

	Reads	Bytes Read	Writes	Bytes Written	Dir Reads	Dir Bytes Read	Read Stats	Write Stats
Reads	1.00000	0.99649	0.06594	0.03394	0.99746	0.99056	0.99730	0.01639
Bytes Read	0.99649	1.00000	0.04020	0.06406	0.99261	0.98709	0.99694	0.00755
Writes	0.06594	0.04020	1.00000	0.61032	0.07738	0.11971	0.01937	0.32650
Bytes Written	0.03394	0.06406	0.61032	1.00000	0.03017	0.08933	0.01306	0.18882
Dir Reads	0.99746	0.99261	0.07738	0.03017	1.00000	0.99318	0.99459	0.02714
Dir Bytes Read	0.99056	0.98709	0.11971	0.08933	0.99318	1.00000	0.98639	0.03688
Read Stats	0.99730	0.99694	0.01937	0.01306	0.99459	0.98639	1.00000	0.01265
Write Stats	0.01639	0.00755	0.32650	0.18882	0.02714	0.03688	0.01265	1.00000

As one might expect, the various read measures are strongly correlated. The write measures, however, are not correlated either with reads or with each other. We suspect that this is because while all of our users do similar things—run large programs to process small files—the outputs of their work are more different; a writer produces small files while a software developer, at least in our environment, produces large executable files. Another possibility is that this difference is an artifact of the fact that our read data is skewed by caches while our write data is not.

4.2 User Classification

Our primary goal was to identify significant differences in the way different kinds of users make use of the file system. Since we knew who the users were, we broke them into groups based on their roles in the University and compared the means and standard deviations of our measured variables between the groups.

Our first attempt at this classification was a dismal failure; all of our subcategories had similar (small) means and enormous standard deviations. A closer examination of the individual user data revealed that our initial estimates of our user population were somewhat optimistic. Although at the time of our measurements some 3300 users had personal directories in IFS, only 925 identifiable individuals appear in our data, and many of these did very little. It is clear that for most of our theoretical user population IFS is not part of their normal day-to-day computing activity. In retrospect, this is not surprising; the university offers many other computing resources, including a large mainframe time sharing system which is still used by many. It was, however, something of a disappointment. It was equally clear that for many of the users who do appear in the data, their use of IFS is exploratory and tentative, while they continue to do their real work elsewhere; this impression was confirmed by personal contact with some of them.

Therefore, we created a new subcategory, “casual users” and transferred users whose usage seemed startlingly small compared to others in their initial subcategory to this new classification.

The appendix summarizes the results of this analysis.

5. Conclusions

Several observations suggest themselves:

- Many categories ended up too small to be of interest, some with only one member. In the future we still hope to study a truly large and diverse user community.
- Many categories are not significantly different from "casual" users. In particular our separation of students into undergraduates and graduate students and our breaking out of engineering students (who have an independent source of AFS access) turned out to be irrelevant; all student usage was casual. Again, this is due in part to the large variety of other resources available to them.
- Much of our usage continues to be users who are unauthenticated and so cannot be identified at all. This is yet another indication that much IFS usage is casual exploration, although some of this activity is users at other AFS sites accessing files in our system, since our AFS cell is exported to the world at large.

Despite these setbacks, there are clear differences among categories of users. Members of our manager category perform about the same number of reads as do clericals, but much larger ones, and they perform many more writes. Technical writers do much more of both. Our software developers perform about the same number of reads (of about the same size) as do writers, but they do fewer writes and the writes are larger.

It is interesting to speculate about some of these differences; one suspects for example that developers do larger writes because they are creating executable files, which tend to be large on modern computer systems. What we have demonstrated, however, is that there are differences among users, whatever their cause, that modelers of distributed systems need to take into account.

6. Acknowledgments

The original impetus for this work came from Peter Honeyman, who conceived the tracing collection model we used and wrote the original logging code. Sushila Subrmanian performed an analysis on an earlier collection of data, with similar goals but using a different methodology [7].

The IFS Deployment Group, led by Mark Giufridda, has operational responsibility for the servers we were instrumenting and was enormously helpful both in the mechanics of data collection and in ensuring that we had working logging code.

This work was supported by a research partnership with IBM.

7. References

1. J. Ousterhout, H.L. DaCosta, D. Harrison, J. Kunze, M. Kupfer, J. Thompson, "A Trace-Driven Analysis of the UNIX 4.2 BSD File System" Proceedings of the 10th ACM Symposium on Operating System Principles, Orcas Island (December 1985).
2. M.G. Baker, J.H. Hartman, M.D. Kupfer, K.W. Shirriff, and J.K. Ousterhout. "Measurements of a Distributed File System." Proceedings of the 13th Symposium on Operating System Principle ACM, (October 1991).

3. T. Hanss, “University of Michigan Institutional File System,” /AIXTRA: The AIX Technical Review, pp. 25-32, (January 1992).

4. J.H. Howard, M.L. Kazar, S.G. Menees, D.A. Nichols, M. Satyanarayanan, R.N. Sidebotham, and M.West, “Scale and Performance in Distributed File Systems,” ACM Transactions on Computer Systems, vol. 6, no. 1, pp. 51-81, (February, 1988).

5. S. Blumson, P. Honeyman, T. E. Ragland and M. T. Stolarchuk, “AFS Server Logging,” CITI Technical Report, 93-10, (December 10, 1993).

6. James Howe, “Intermediate File Servers in a Distributed File System Environment,” CITI Technical Report 92-4, (June 30, 1992).

7. S. Subramanian, “Workload Characterization of AFS File Servers,” CITI Technical Report 92-6, (November 23, 1992).

8. Appendix

The following table summarizes the statistics for each of our final user classifications. For each classification we report the number of members, and for each of our selected metrics the total for that metric, the mean per member, and the standard deviation over the members.

Note that the data for the “unauthenticated” classification is slightly misleading in that although these appear in the data as two different entities, each represents a (presumably large) number of individuals.

Class	Members	Operation	Mean	Std Deviation	Total
Casual	158	Reads	60.0000000	147.5455667	9480.00
Casual	158	Bytes Read	775594.15	3225972.84	122543876
Casual	158	Writes	31.2911392	101.4426702	4944.00
Casual	158	Bts Written	371099.83	1607553.69	58633772.79
Casual	158	Dir Reads	74.4240506	225.9997546	11759.00
Casual	158	Dir Bts Read	174507.75	488600.05	27572224.03
Casual	158	Stat Reads	513.8417722	3056.76	81187.00
Casual	158	Stat Writes	21.6582278	81.0914222	3422.00
Clerical	14	Reads	1484.50	5106.42	20783.00
Clerical	14	Bytes Read	2614659.92	9015772.85	36605238.82
Clerical	14	Writes	291.5714286	706.2305915	4082.00
Clerical	14	Bts Written	501548.43	1043418.91	7021678.09
Clerical	14	Dir Reads	1973.50	6523.48	27629.00
Clerical	14	Dir Bts Wtn	4187867.45	13579623.80	58630144.34
Clerical	14	Stat Reads	1302.57	3485.43	18236.00
Clerical	14	Stat Writes	95.9285714	244.5427332	1343.00
Grad Students	166	Reads	332.0722892	1037.49	55124.00
Grad Students	166	Bytes Read	7756129.12	47523962.49	1287517434
Grad Students	166	Writes	149.7951807	326.6674431	24866.00
Grad Students	166	Bts Written	3114634.31	15195425.57	517029296
Grad Students	166	Dir Reads	189.3614458	568.2153578	31434.00
Grad Students	166	Dir Bts Read	479700.82	1457796.36	79630335.95
Grad Students	166	Stat Reads	646.4457831	3097.91	107310.00
Grad Students	166	Stat Writes	257.7590361	2422.53	42788.00

Class	Members	Operation	Mean	Std Deviation	Total
Grad Students - Eng	39	Reads	219.7435897	569.0942350	8570.00
Grad Students - Eng	39	Bytes Read	3731117.35	12144688.12	145513577
Grad Students - Eng	39	Writes	191.3589744	544.1091457	7463.00
Grad Students - Eng	39	Bts Written	3170299.72	9687417.91	123641689
Grad Students - Eng	39	Dir Reads	147.5128205	430.4148471	5753.00
Grad Students - Eng	39	Dir Bts Wtn	359397.74	946581.74	14016512.02
Grad Students - Eng	39	Stat Reads	852.9743590	1906.18	33266.00
Grad Students - Eng	39	Stat Writes	133.0000000	471.2484539	5187.00
ITD Consultants	3	Reads	4950.33	4698.91	14851.00
ITD Consultants	3	Bytes Read	185764902	318274937	557294707
ITD Consultants	3	Writes	2374.00	3513.08	7122.00
ITD Consultants	3	Bts Written	118492661	204837483	355477982
ITD Consultants	3	Dir Reads	3431.33	3414.24	10294.00
ITD Consultants	3	Dir Bts Wtn	7541418.65	6580854.06	22624255.96
ITD Consultants	3	Stat Reads	6942.00	9432.91	20826.00
ITD Consultants	3	Stat Writes	445.0000000	633.1358464	1335.00
Managers	15	Reads	1466.33	1474.05	21995.00
Managers	15	Bytes Read	6032568.89	15343869.13	90488533.30
Managers	15	Writes	838.7333333	1086.12	12581.00
Managers	15	Bts Written	4099004.70	8898656.60	61485070.52
Managers	15	Dir Reads	2716.53	4621.98	40748.00
Managers	15	Dir Bts Wtn	6011153.06	9354385.93	90167295.84
Managers	15	Stat Reads	2405.67	3039.40	36085.00
Managers	15	Stat Writes	437.0666667	719.6451970	6556.00
Nurses	1	Reads	1277.00	.	1277.00
Nurses	1	Bytes Read	11035194.94	.	11035194.94
Nurses	1	Writes	0	.	0
Nurses	1	Bts Written	0	.	0
Nurses	1	Dir Reads	1438.00	.	1438.00
Nurses	1	Dir Bts Wtn	3364864.00	.	3364864.00
Nurses	1	Stat Reads	4847.00	.	4847.00
Nurses	1	Stat Writes	0	.	0
Researchers	20	Reads	801.2000000	831.5966634	16024.00
Researchers	20	Bytes Read	25612931.55	48217758.91	512258631
Researchers	20	Writes	552.8000000	909.3575291	11056.00
Researchers	20	Bts Written	9559567.77	16377892.45	191191355
Researchers	20	Dir Reads	425.4000000	502.5218927	8508.00
Researchers	20	Dir Bts Wtn	1150156.80	1118295.33	23003135.96
Researchers	20	Stat Reads	2039.80	2886.19	40796.00
Researchers	20	Stat Writes	190.8500000	326.2306699	3817.00
Software Developers	35	Reads	6352.69	10774.12	222344.00
Software Developers	35	Bytes Read	47681728.81	58798076.88	1668860508
Software Developers	35	Writes	2316.14	3496.93	81065.00
Software Developers	35	Bts Written	22044720.69	32768942.32	771565224
Software Developers	35	Dir Reads	4080.23	7194.61	142808.00
Software Developers	35	Dir Bts Wtn	15189372.47	27769842.44	531628036
Software Developers	35	Stat Reads	12668.77	18171.17	443407.00
Software Developers	35	Stat Writes	1350.91	2263.57	47282.00

Class	Members	Operation	Mean	Std Deviation	Total
Students	343	Reads	410.3673469	2040.13	140756.00
Students	343	Bytes Read	4002697.18	30689461.22	1372925133
Students	343	Writes	166.9679300	726.5652485	57270.00
Students	343	Bts Written	2753848.43	21106350.92	944570010
Students	343	Dir Reads	343.9941691	1150.21	117990.00
Students	343	Dir Bts Read	1335994.58	10098385.71	458246142
Students	343	Stat Reads	396.6151603	1178.63	136039.00
Students	343	Stat Writes	44.0670554	170.9022622	15115.00
Students - Eng	78	Reads	122.8333333	414.4818207	9581.00
Students - Eng	78	Bytes Read	1059595.68	3586942.94	82648463.03
Students - Eng	78	Writes	45.3205128	146.3595017	3535.00
Students - Eng	78	Bts Written	543196.59	1444452.21	42369334.26
Students - Eng	78	Dir Reads	134.5000000	524.9312324	10491.00
Students - Eng	78	Dir Bts Wtn	291078.56	1080197.81	22704128.00
Students - Eng	78	Stat Reads	164.9871795	277.4115745	12869.00
Students - Eng	78	Stat Writes	33.6153846	113.5709894	2622.00
System Administrators	14	Reads	2606.86	2701.08	36496.00
System Administrators	14	Bytes Read	23199845.57	21611290.81	324797838
System Administrators	14	Writes	1184.00	749.4336323	16576.00
System Administrators	14	Bts Written	9472930.07	12546544.72	132621021
System Administrators	14	Dir Reads	925.0714286	706.5371093	12951.00
System Administrators	14	Dir Bts Wtn	3118957.70	2145703.57	43665407.86
System Administrators	14	Stat Reads	10174.21	7381.88	142439.00
System Administrators	14	Stat Writes	1240.07	1465.22	17361.00
Technicians	3	Reads	1326.33	1089.58	3979.00
Technicians	3	Bytes Read	3097211.00	3670971.26	9291633.01
Technicians	3	Writes	743.6666667	958.2016141	2231.00
Technicians	3	Bts Written	9756682.28	15850251.84	29270046.84
Technicians	3	Dir Reads	2079.00	1519.12	6237.00
Technicians	3	Dir Bts Wtn	5182122.65	4492274.92	15546367.94
Technicians	3	Stat Reads	2250.33	1995.61	6751.00
Technicians	3	Stat Writes	294.0000000	444.2105357	882.0000000
UM Archive	8	Reads	161.1250000	184.0143143	1289.00
UM Archive	8	Bytes Read	3516023.75	3592732.70	28128190.00
UM Archive	8	Writes	37.2500000	61.0005855	298.0000000
UM Archive	8	Bts Written	1500698.76	2347575.37	12005590.07
UM Archive	8	Dir Reads	118.3750000	238.2531292	947.0000000
UM Archive	8	Dir Bts Wtn	349696.00	680379.77	2797567.98
UM Archive	8	Stat Reads	6540.75	14119.45	52326.00
UM Archive	8	Stat Writes	19.1250000	36.4787120	153.0000000
Unauthenticated	2	Reads	539953.00	763608.86	1079906.00
Unauthenticated	2	Bytes Read	7095306343	10034278459	14190612686
Unauthenticated	2	Writes	154.0000000	217.7888886	308.0000000
Unauthenticated	2	Bts Written	2370296.99	3352106.15	4740593.99
Unauthenticated	2	Dir Reads	286311.00	404904.90	572622.00
Unauthenticated	2	Dir Bts Wtn	776378366	1097964815	1552756732
Unauthenticated	2	Stat Reads	2035743.50	2878974.65	4071487.00
Unauthenticated	2	Stat Writes	35.0000000	49.4974747	70.0000000

Class	Members	Operation	Mean	Std Deviation	Total
Unclassified	20	Reads	165.4000000	594.5528528	3308.00
Unclassified	20	Bytes Read	9481083.25	38969116.24	189621665
Unclassified	20	Writes	34.0000000	75.4048721	680.0000000
Unclassified	20	Bts Written	703861.35	1784089.18	14077227.08
Unclassified	20	Dir Reads	15.8500000	25.6315492	317.0000000
Unclassified	20	Dir Bts Wtn	43417.60	71179.65	868352.00
Unclassified	20	Stat Reads	101.9500000	138.2672755	2039.00
Unclassified	20	Stat Writes	8.8000000	21.7052019	176.0000000
Vendors	1	Reads	431.0000000	.	431.0000000
Vendors	1	Bytes Read	1045323.96	.	1045323.96
Vendors	1	Writes	69.0000000	.	69.0000000
Vendors	1	Bts Written	179939.01	.	179939.01
Vendors	1	Dir Reads	687.0000000	.	687.0000000
Vendors	1	Dir Bts Wtn	1916928.02	.	1916928.02
Vendors	1	Stat Reads	311.0000000	.	311.0000000
Vendors	1	Stat Writes	1.0000000	.	1.0000000
Writers	7	Reads	7352.00	4886.81	51464.00
Writers	7	Bytes Read	45243831.43	35841350.25	316706820
Writers	7	Writes	4390.71	4163.57	30735.00
Writers	7	Bts Written	21683317.67	20322552.85	151783224
Writers	7	Dir Reads	5759.29	5260.26	40315.00
Writers	7	Dir Bts Wtn	20396909.68	18124032.99	142778368
Writers	7	Stat Reads	7994.43	5391.37	55961.00
Writers	7	Stat Writes	1333.57	1086.72	9335.00