ISSN 2328-1391 (print), ISSN 2227-930X (online)
International Journal of Advanced Studies (iJAS)
2015, Vol. 5, No 2, pp. 12-17
DOI: 10.12731/2227-930X-2015-2-2
Received January 20, 2015; revised February 21, 2015; accepted March 6, 2015

Publishing House
Science and Innovation
Center, Ltd.

RESEARCH OF PERFORMANCE LINUX KERNEL FILE SYSTEMS

Andrey Vladimirovich Ostroukh
full member RAE (Russian Federation) and EAEN (Germany), Dr. of Tech. Sci., Professor,
Department of «Automated Control Systems»
State Technical University – MADI
Leningradsky prospekt, 64, Moscow, 125319, Russian Federation
http://www.madi.ru
ostroukh@mail.ru

Alexander Gennadievich Salniy
Undergraduate Student, Department of «Automated Control Systems»
State Technical University – MADI
Leningradsky prospekt, 64, Moscow, 125319, Russian Federation
http://www.madi.ru
alexandr.salny@yandex.ru

Abstract. *The article describes the most common Linux Kernel File Systems. The research was carried out on a personal computer, the characteristics of which are written in the article. The study was performed on a typical workstation running GNU/Linux with below characteristics. On a personal computer for measuring the file performance, has been installed the necessary software. Based on the results, conclusions and proposed recommendations for use of file systems. Identified and recommended by the best ways to store data.*

Keywords: *File System (FS); Operating System (OS); file; performance; Linux; experimental data; personal computer (PC); software.*

1. INTRODUCTION

File System (FS) – order that defines a way to organize, store, and naming data storage media for computers and other electronic equipment such as digital cameras, mobile phones and so on. [1] The file system determines the format of the content and the method of the physical storage of information, which is usually grouped in the form of files. Specific file system determines the size of the file names and (catalogs), the maximum possible size of the files and set file attributes. Some file systems provide service opportunities, such as access control and encryption of files.

The file system is strictly monitors the organization, storage and naming data. For example, when a program accesses a file, it does not know anything about where the information about the file, on what physical medium it is recorded. The only thing that he knows the program is the name of the file to which it refers, its size, and other parameters that the target file is different from many other files (the date and time the file was created, the name of the file owner, the

rights and the method of access to the file, etc.). All these data it receives from the file system driver. This file system determines where and how to file will be stored on physical media (such as hard disk or removable device). All in one place catalogs (in a different folder), sub-directories as well as files and have the data of the file system.

2. OVERVIEW OF LINUX FILE SYSTEMS

Currently, the most common the following file systems: **ext2, ext3, ext4, ReiserFS, btrfs, jfs** [1 – 21].

Second Extended File System, abbreviated as **ext2**– filesystem kernel Linux. Remy has been developed to replace the existing cardio then ext. The speed and productivity, it can serve as a benchmark in performance tests filesystems.

The main drawback ext2 (and one of the reasons why such high performance demonstration) is that it is not a journaling file system. He was eliminated in the file system ext3 – the next version of Extended File System, is completely compatible with ext2. But Solid State Drives (SSD) is rather a plus, extends the life of the drive. This is the main reason why ext2 is still supported in Anaconda and Ubiquity.

The ext2 file system is still used on a flash card and SSD, as the lack of logging is an advantage when working with hard drives, which have a limit on the number of write cycles.

Third Extended File System (third version of the extended file system), **ext3**, or abbreviated **ext3fs** – journaling file system used in the operating system kernel on Linux, it is the default file system in many Linux distributions. Based on ext2 file system, which laid the beginning of the development of Stephen Tweedie.

The main difference from ext2 is that ext3 journaling that is, it provides a record of some data to restore the file system in case of failures in the computer.

The standard provides three modes of logging:

writeback:only logged in the file system metadata, i.e. information about its change. It can not guarantee the integrity of the data, but noticeably reduces scan time compared to ext2;

ordered: same as the writeback, but writing data to the file produced is guaranteed to record information about changing this file. Slightly degrades performance, can not guarantee the integrity of the data (although increases the likelihood of their safety when appending to an existing file);

journal: full journaling file system metadata and user data. The slowest but safest mode; can ensure the integrity of data stored on a separate section of the journal (and better - on a separate hard drive).

The ext3 file system can support files up to 1 TB. With the Linux-2.4 kernel file system volume limited to a maximum size of a block device that is 2 terabytes. In Linux 2.6 (32-bit processors), the maximum size of the block device is 16 TB, but only supports ext3 up to 4TB

The maximum number of blocks for ext3 equals 232. The block size can vary, which affects the maximum number of files and the maximum file size in the file system.

Table 1.

Restrictions on the size of the ext3 file system

Blocksize	MaximumFileSize	Maximum Size of the File System
1 KB	16 GB	<2 TB
2 KB	256 GB	< 8 TB
4 KB	2 TB	< 16 TB
8 KB	2 TB	< 32 TB

Ext4 – is the result of evolution ext3, the most popular file system in Linux. In many ways, Ext4 is a big step forward compared to ext3, ext3 than was in relation to the ext2 [2]. The most significant improvement compared to ext3 ext2 was logging, while ext4 involves important changes in data structures, such as for storing data files.

One improvement is to increase the maximum size of a single file for the FS. Today, the maximum size ext3 file system is 16 terabytes, and the file size is limited to 2 terabytes. In Ext4 adds 48-bit block addressing, which means that the maximum size of this file system is equal to one exabyte and files can be up to 16 Terabytes. 1 EB (exabytes) = 1,048,576 TB (terabytes), 1 EB = 1024 PB (petabytes), 1 PB = 1024 TB, 1 TB = 1024 GB. 48-bit addressing of blocks used

for a number of limitations that need to be removed, to make Ext4 fully 64-bit, and this problem was not put before the Ext4. The data structures in the Ext4 designed taking into account the changes required, so one day in the future, support for 64 bits per Ext4 will appear. In the meantime, we have to settle for one exabytes.

As one of the important features of the file system is online defrag, which reduces file fragmentation. For example: You create three files in the same directory, and they are located on disk for each other. Then, once you decide to update the second file, and this file is somewhat more – so that the space is not enough for him. At the same time there are no other solutions than to separate the fragment can not contain the file and put it on a disk or another place select the file area of the disk larger elsewhere, away from the first two files that will lead to disk head movements, if an application need to be considered all the files in the directory (for example, the file manager will create thumbnails for image files). In addition, the file system can take care of only certain types of fragmentation and it can not know, for example, that it should store all the files required at boot time, next to each other, because it does not know which of them are required for downloading.

XFS – a high-performance journaling file system created by Silicon Graphics for its own operating system IRIX. May 1, 2001 Silicon Graphics has released XFS GNU (General Public License). XFS differs from other file systems in that it was originally designed for use on discs 2 terabytes. XFS support was included in the Linux kernel version 2.4 (starting from 2.4.25 when Marcelo Tozatti found it stable enough) and 2.6, and thus, it has become quite universal for Linux-based systems. Installers distributions openSUSE, Gentoo, Mandriva, Slackware, Ubuntu, Fedora and Debian offer XFS file system as an option for installation. FreeBSD now supports XFS in read mode in December 2005.

However, the file system has a very critical to the operation of the server flaws.

It is impossible to reduce the size of the existing file system.

Recovery of deleted files in the XFS – a very complex process, so at this time (2014) for this purpose there are only a few software products, for example, «Raise Data Recovery for XFS» for OS Windows.

The possibility of data loss during the recording if the power fails, because a large number of data buffers stored in memory while the metadata recorded in the log (to disk) efficiently. This is true for other file systems metadata journaling.

Btrfs (B-tree FS, «Better FS» or «Butter FS») – file system for Linux, based on the structure of B-trees and working on the principle of "copy on write» (copy-on-write). Published by Oracle Corporation in 2007, is licensed under GNU General Public License (GPL). One of the original purposes of developing this file system was to ensure a decent competition popular ZFS. Btrfs will be relieved of many of the disadvantages of other modern file systems for Linux.

Btrfs is considered to be stable, but as of 2010 does not create a tool to check the file system, and bug fixes. Btrfs v0.19 version released in June 2009.

It was originally planned to release the Btrfs v1.0 (and record to disk storage format) at the end of 2008, but the format was recorded only 12 June 2010.

ReiserFS file system a large number of small files leads to the greatest "lost" disk space. [3] But with this method of storage move / create / copy a file takes much less time.

The advantage of the file system is the use of ReFS recovery features built into the file system that allow you to immediately correct mistakes without starting a long global test. This is especially true for large volumes of data. ReFS can handle up to 1YB (Yottabyte).

To store the information about the free sites do not use ReiserFS simple lists, but somewhat more complex data structures. The system ReiserFS used for this so-called "balanced trees" or "B + Trees", the search for which is not proportional to the number of objects (files in the directory, or the number of blocks on the disk), and the logarithm of the number. In all branches of the tree balanced (the path from the root to the "sheet") have the same (or approximately the same) length. ReiserFS uses a balanced trees to store all file system objects: files, directories, information

about free blocks, and so on. D. This can significantly improve the performance of disk accesses.

ReiserFS also has a number of features aimed specifically to improve the performance with small files. ReiserFS is not related restriction in the allocation of memory for the whole file including 1-2-4 KB blocks. If necessary, the file can allocate the exact size. ReiserFS also includes some special kinds of optimization file "tail" end portions for storing files that are smaller than a logical block in the file system. To increase the speed, ReiserFS capable of storing content files directly within the tree b * tree, rather than a pointer to a disk block (in ext2 is the concept fastlink, when the contents of the "soft" links 60 bytes stored in iNode).

This achieves two things. First, productivity is greatly increased, as the data and metadata (stat_data, in other words, inode) information can be stored there and read a disk I / O. Second, ReiserFS is able to pack the tails (tail) file, saving disk space. In fact, the resolution of ReiserFS tail packing to perform (the default setting) will save approximately six percent of disk space (in comparison to ext2).

It should be borne in mind that the tail packing requires more work, as you resize the files necessary to "re-packaging". For this reason, ReiserFS tail packing can be disengaged, allowing the administrator to choose between speed and efficient use of disk space.

Journaled File System (JFS) – 64-bit journaling file system created by IBM, available under the GNU GPL [1].

The AIX operating system, there are two generations of JFS called JFS (JFS1) and JFS2 respectively. In other operating systems such as OS/2 and Linux, there is only the second generation, which is simply called JFS. Also called the JFS filesystem VxFScompanyVeritas Software, OS used in HP-UX.

Initially, JFS was developed by IBM for the operating system AIX. JFS second generation was developed by IBM for OS Warp Server for e-Business. Later, she was transferred to IBMAIX and Linux. The goal of the developers was to ensure high performance, reliability and scalability for multiprocessor computers.

Unlike ext3, which has been added to support journaling, JFS was originally journaling. JFS keeps a log only metadata supporting structure of the file system integrity but not necessarily the data. Power failure or system crash can lead to the preservation of outdated copies of the files, but the files themselves remain usable. Journaling JFS like journaling XFS, which only logs a part iNode.

To manage disk partitions formatted JFS was released a set of tools called JFSutils.

3. METHODS

When testing the speed of file system was used:

- based server Core i7-4960 16Gb DDR3;
- hard disk (HDD) WD black 7000rpm;
- hypervisor VmWare ESXI 5.0;
- virtual machine with two cores and 8 Gb of RAM;
- CENTOS 6 with the latest updates.

For comparison, the velocity used the following Bash script:

```
cmd1="cp -r /media/media4/video/best $dest"
cmd2="rsync -rlhtgopu /media/media4/backup
$dest"
cmd3="greplinux -sir $dest/backup/wine-src/"
cmd4="find $dest -type f -delete"
```

Test conditions:

- virtual machine with 2 cores and 4 GB of RAM;
- measurements were performed using c / usr / bin / time;
- between tests 10 minute break to uptime;
- partition with a PS is chosen so that the data to fill it 2/3;
- file sizes used in the test.
- smaller files – 1,7G / media / media1 – 40,285 files.
- large files – 17,4G / media / media2 – 4 files.

4. RESULTS

Comparative characteristics of the file systems listed in Table 2.

Performance comparison of file systems are shown in Table 3 and the charts of Figure 1 [15].

Table 2.

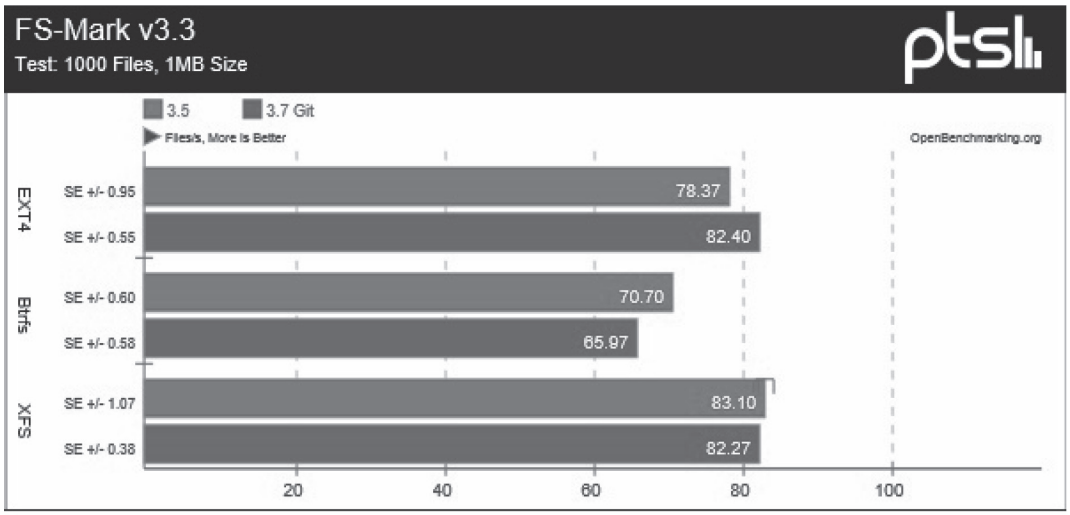
Comparative characteristics of the File Systems

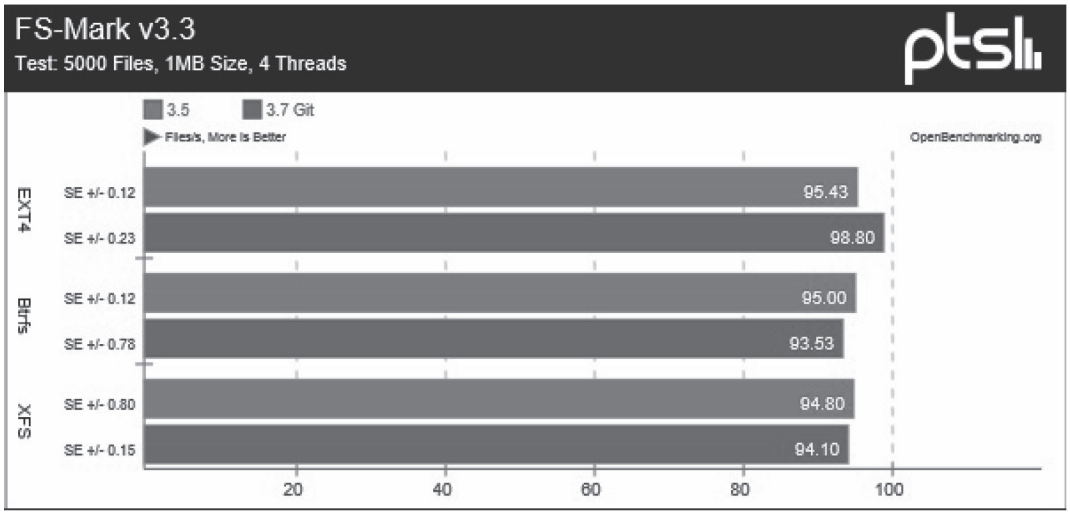
	ext2	ext3	xfs	btrfs	reiserfs	jfs
Folder Contents	Table	Table	B+ Trees	B+ Trees	B+ Trees	B+ Trees
File Table	Bit Maps	Bit Maps	B+ Trees	Extend	Bit Maps	Bit Maps
Maximum File Size	from 16 GigaByte to 2 TeraByte	Depends on the Block Size	8 ExaByte	8 ExaByte	1 ExaByte	4 PetaByte
Maximum Length of the File Name	255 байт	Depends on the Block Size	256 Byte	255 Byte	4032 Byte	255 Byte
Maximum Volume Size	from 2 to 32 TeraByte	Depends on the Block Size	16 ExaByte	16 ExaByte	16 TeraByte	32 PetaByte

Table 3.

Comparative performance of the File Systems

	ext2	ext3	ext4	xfs	reiserfs	btrfs	jfs
Copy Large Files	116.03	122.69	116.45	137.47	138.67	130.25	130.98
ZIP Small Files	115.33	124.25	99.61	220.50	119.25	98.44	172.21
Search Small Files	66.71	63.69	68.76	47.02	66.45	77.18	107.21
Search again Small Files	100.47	97.27	102.36	80.70	96.48	101.27	135.29
Search and Delete Files	8,09	7,51	6,40	82,59	10,22	13,53	15.67
Average Load on the File System	1.85, 1.37,	1.95, 1.39,	1.99, 1.26	2.02, 1.64	2.00, 1.47	2.09, 1.37	2.55, 1.99





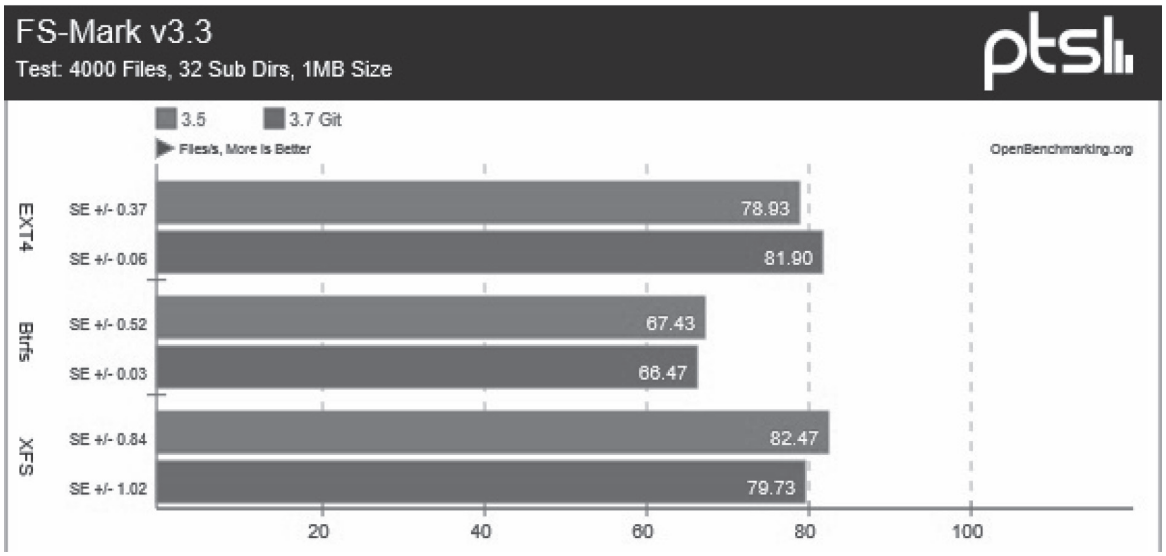


Fig. 1. Comparing the performance of File Systems Linux Kernel

Based on data from Table 3, the optimal performance file systems are ext2 and ext, but must take into account the fact that ext4 is a journaling file system, and more resistant to crashes.

5. CONCLUSIONS

Increased loads in local area networks connected in most cases with an increase in file circle network that raises the bar for file systems on servers. In this article we looked at several file systems the kernel Linux, widely used in the present time. Specifications file systems were given to a structured form and speed measurements were carried out at various operations with files. Optimum performance has ext4 file system used by default in many distributions Linux.

REFERENCES

[1] https://ru.wikipedia.org/.
[2] https://ru.wikipedia.org/wiki/Ext4
[3] https://ru.wikipedia.org/wiki/ReiserFS
[4] https://ru.wikipedia.org/wiki/ExFAT
[5] http://habrahabr.ru/post/179821/
[6] http://citforum.ru/operating_systems/linux/robbins/fs02.shtml
[7] http://xgu.ru/wiki/ext4
[8] http://www.linuxcenter.ru/lib/books/kostromin/gl_16_08.phtml
[9] http://www.opennet.ru/docs/RUS/reiserfs_ondisk_layout/
[10] http://www.opennet.ru/opennews/art.shtml?num=21343
[11] http://habrahabr.ru/post/191136/
[12] http://habrahabr.ru/post/179821/
[13] http://habrahabr.ru/post/54043/
[14] http://habrahabr.ru/post/45873/
[15] http://www.linuxcenter.ru/lib/books/kostromin/gl_16_01.phtml
[16] http://www.phoronix.com/scan.php?page=article&item=linux_37_fsthree&num=2
[17] Ostroukh A.V. Input and processing of digital information. Moscow: Publishing House “Academy”. 2012. 288 p. ISBN 978-5-7695-9457-1.
[18] Krasnyanskiy M.N., Karpushkin S.V., Obukhov A.D., Ostroukh A.V. Automated control system for university research projects // International Journal of Advanced Studies (iJAS). 2014. Vol. 4, Issue 1, pp. 22-26. DOI: 10.12731/2227-930X-2014-1-4.
[19] Ostroukh A.V., Polgun M.B. Automation of processes supervisory control urban passenger transport // International Journal of Advanced Studies (iJAS). 2013. Vol. 3. No 3. pp. 3-9. DOI: 10.12731/2227-930X-2013-3-1.
[20] Ostroukh A.V., Polgun M.B. New approaches to development of automated supervisory systems of industrial enterprises transport // International Journal of Advanced Studies (iJAS). 2013. Vol. 3. No 4. pp. 3-9. DOI: 10.12731/2227-930X-2013-4-1.
[21] Kuftinova N.Ya.G.E., Ostroukh A.V. Development of an automated system of survey passenger traffics // International Journal of Advanced Studies (iJAS). 2014. Vol. 4. No 4. pp. 3-9. DOI: 10.12731/2227-930X-2014-4-2.