

Abstract

Of master's certification on:

«Peculiarities of organization of high load asynchronous I/O applications running under Linux OS»

Artem Zosimchuk.

Relevance

With the rapid development of electronics and telecommunication volumes and data rates increase strides.

In this regard, becomes relevant to the establishment of software solutions based on effective core potential read / write / transfer data, as they are critical factors in ensuring the performance and scalability of software solutions rather broad class. Most servers and heavy-solutions-oriented systems running Linux. Review and selection of positive and negative aspects of different methods make it possible to create efficient I/O subsystem.

Objective

The aim of this work is to investigate and determine best practices for creating an effective I/O performance in applications with high load running Linux OS.

The solved problems

In achieving the goal in this paper the following tasks:

- identification of requirements and performance characteristics of I/O subsystem;
- study of existing high-performance solutions and the allocation of core architectures;
- consideration of the core and disk input/output scheduler operating system Linux;
- study the features of different methods of I/O;
- comparison of performance of the methods I/O to existing applications;
- writing test applications that demonstrate the influence of individual parameters on the rate of application.

Achieved

Having decided to put in work tasks, the author defends:

- a set of recommendations designed to create and optimization of input / output.

Scientific novelty

Scientific novelty of the work is to allocate according to the speed I/O on the implementation of various approaches, developed guidelines for creating high- performance I/O subsystem, the settings on Linux.

The practical value of

The paper provides an overview of existing innovative solutions (including web servers and databases), indicating the possibility of optimizing their performance. Principles of designing I/O subsystem showed, depending on the planned operating conditions, loads. The positive and negative aspects of different solutions consideration and their recommended application.

Findings

The paper discusses the main approaches to I/O architecture of high- performance applications: with a fixed number of processors and handlers queries created for each query. The paper provides an overview of possibilities of various methods of I/O, in terms of speed of work, scalability. As shown by comparison to most cases, one of the best architecture for I / O - is one in which multiple threads are treated with a common pool of network requests through the system call `epoll ()`, and for disk I/O library is used `aio` library. Also discussed solutions for certain operating conditions. Some practical advice on choosing the scheduler disk I/O OS Linux given - complete fair queuing is optimal for most purposes.

As a result, given an idea of how best to build a fast subsystem I/O application and determined the place where work can be customized and optimized.

The work contains 75 pp., 22 fig., 1 table, 22 sources.

Keywords: programming I/O IN HIGH LOAD APPLICATION, ASYNCHRONOUS I/O IN LINUX OS