

УДК 004.042

А. П. Гантимуров, А. В. Босов

Московский государственный технический университет им. Н. Э. Баумана

Анализ возможностей оптимизации структуры распределенной системы хранения данных при известном профиле нагрузки

Приведен анализ возможностей для определения оптимальных структурных характеристик распределенной системы хранения данных, исходя из базовых характеристик однородного массива жестких дисков, входящих в систему, и измеренных показателей обмена данными по сети на базе математической модели и известного профиля нагрузки. Прежде всего определена постановка задачи на основе многолетней практики проектирования, разработки, внедрения и мониторинга распределенных систем хранения данных. Указаны ключевые факторы, влияющие на показатели времени ожидания между запросом на операцию и стартом передачи данных и суммарная производительность информационного обмена данными. Выполнены математическое моделирование влияющих параметров и характеристик, а также их анализ с учетом данных, получаемых при мониторинге работающих распределенных систем хранения данных различного уровня сложности и архитектуры. Исследования, описанные в статье, будут полезны специалистам в области проектирования, разработки и организации распределенных систем хранения данных различной сложности.

Распределенная система хранения данных, профиль нагрузки, информационный обмен данными, сеть передачи данных, производительность передачи данных

Ключевыми показателями эффективности информационного обмена между бизнес-процессами в высоконагруженной сети передачи данных являются показатели времени ожидания между запросом на операцию и стартом передачи данных и суммарная производительность информационного обмена данными [1]–[3]. Это подтверждается многочисленными наблюдениями и многолетним опытом в области разработки систем хранения данных [4]. Указанные величины являются базовыми, на которые ориентируются при проектировании как конкретных бизнес-процессов, так и всей сети передачи данных. Соответственно, сбои по указанным показателям являются крайне нежелательными, поскольку могут привести к серьезной разбалансированности всего информационного обмена.

При формировании профиля нагрузки учитываются, например, следующие параметры: процентное соотношение операций чтения и записи; режим процессов чтения/записи (последовательный или случайный); размер информационного блока данных; длительность теста. Также может указываться ряд дополнительных параметров, например количество хостов, используется ли

дедупликация и компрессия данных и пр. На сформированном профиле нагрузки определяются требования по значениям среднего и максимального времени ожидания, а также минимальной производительности информационного обмена.

Постановка задачи. Традиционно, наиболее узким местом в информационном обмене высоконагруженной сети передачи данных является обмен данными с массивом жестких дисков системы хранения данных. Даже если при построении системы хранения данных использовать скоростные жесткие ssd-диски, то и в этом случае, в зависимости от режима обращения к диску, реальная производительность может отличаться от номинальной и разброс значений может варьироваться в широком диапазоне значений (рис. 1). Помимо этого скорость чтения/записи на ssd-диск также зависит от его заполненности, что проиллюстрировано графиком тестирования ADATA DashDrive Durable HD710 (рис. 2: кривая 1 – USB 3.0.чтение; кривая 2 – USB 3.0.запись; кривая 3 – USB 2.0.чтение; кривая 4 – USB 2.0.запись). Указанное подтверждается многолетним мониторингом за состоянием систем хранения данных различного назначения и разной сложности [4].

Кроме указанных трудностей, при проектировании распределенной системы хранения данных с заданными показателями производительности информационного обмена данными определенную сложность представляет и учет физических возможностей сети передачи данных, связанных с варьированием скорости передачи пакетов данных. Другими словами, при проектировании распределенной системы хранения данных, используемой в информационном обмене данными, с заданным профилем нагрузки и требуемыми показателями эффективности необходимо учитывать нестабильность показателей скорости обмена данными между внутренними подсистемами системы хранения данных и предусматривать механизмы, компенсирующие эту нестабильность.

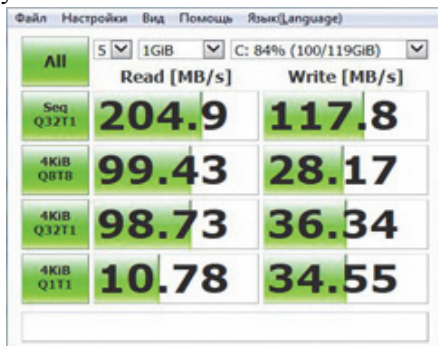


Рис. 1

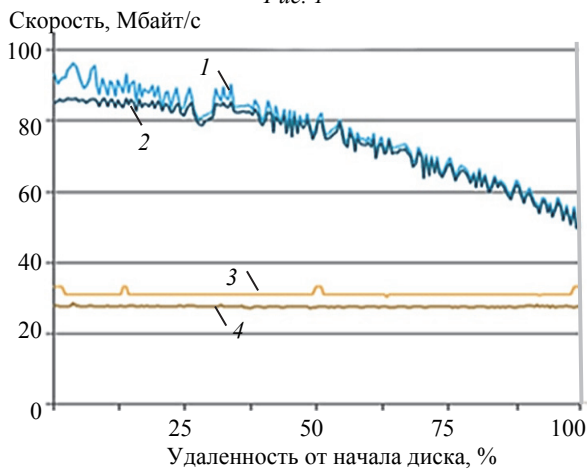


Рис. 2

Как правило, оптимизация работы и производительности распределенных систем хранения данных выполняется по результатам мониторинга, включая нагрузочное тестирование, уже после ее разработки и внедрения с целью поиска способов и методов рационализации структуры в целом, в том числе и в замене используемого оборудования [5]–[7].

Целью исследований в области проектирования, построения, внедрения и дальнейшего обслуживания распределенных систем хранения данных является анализ эффективных способов разработки распределенных систем хранения данных различного назначения.

Один из результатов данной работы – анализ возможностей оптимизации структуры распределенной системы хранения данных при известном профиле нагрузки, выполненный на основе данных как математического моделирования, так и экспериментального наблюдения.

Анализ причин влияния. Опишем механизм возникновения времени ожидания между запросом на операцию и стартом передачи данных в распределенной системе хранения данных, приведенной в [8], [9].

На операции записи время ожидания складывается из времени формирования и посылки запроса на установление соединения и времени установления соединения. В подавляющем большинстве случаев длительность данных операций не превышает десятков микросекунд и существенно не влияет на время ожидания.

При выполнении операции чтения из распределенной системы хранения данных, согласно [8]–[10], необходима процедура сборки матрицы кодирования/декодирования из блоков данных, хранимых на жестких дисках системы. Следовательно, на время ожидания в первую очередь влияет время, затрачиваемое на чтение с жестких дисков системы требуемого количества блоков для составления матрицы кодирования/декодирования. Далее, считая, что чтение со всех дисков системы осуществляется в параллельном режиме, получаем, что время ожидания напрямую зависит от количества жестких дисков в системе. Характер этой зависимости, рассчитанный на базе математической модели распределенной системы хранения данных вида [8]–[10], приведен на рис. 3.



Рис. 3

Помимо времени, требуемого на чтение с жестких дисков блоков данных матрицы кодирования/декодирования, на время ожидания также влияет запуск процедуры восстановления потерянных данных в матрице кодирования/декодирования. Запуск данной процедуры приводит к спонтанному росту времени ожидания. Как правило, этот рост учитывается в профиле как максимально допустимое время ожидания и накладывает ограничение на показатели производительности процедуры восстановления потерянных данных, учитываемые в процессе разработки системы.

При анализе графика на рис. 3 можно выделить минимально необходимую среднюю скорость обмена данными с массивом жестких дисков, при которой достигается требуемый по профилю норматив времени ожидания для фиксированного количества дисков в массиве. Соответственно, на базе математической модели можно построить кривую зависимости минимально необходимого количества жестких дисков, требуемых для обеспечения среднего времени ожидания по профилю, от средней скорости чтения блоков данных с жесткого диска (рис. 4).

Требуемое минимальное количество дисков для средней задержки менее 5 мс

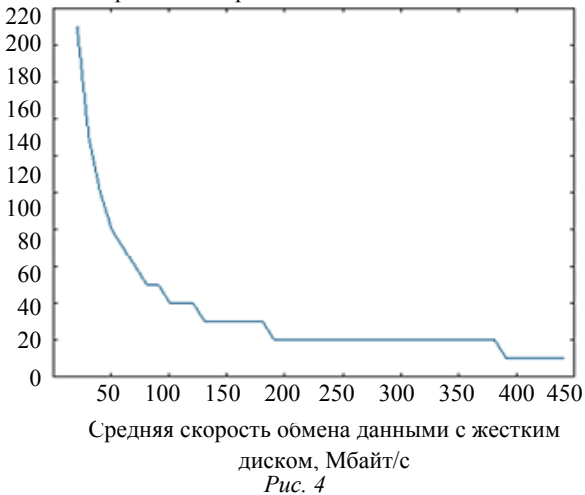


Рис. 4

Аналогичные допущения можно применить и к параметру общей производительности передачи данных между клиентом и распределенной системой хранения данных с учетом физических возможностей сети передачи данных. Производительность процесса обмена данными зависит от скорости прохождения пакета данных, которая может варьироваться в некотором диапазоне. График зависимости производительности процесса обмена данными от среднего времени про-

хождения пакета при заданном количестве жестких дисков (100 шт.) и для обеспечения требуемой производительности (300 kiops, iops – количество операций ввода-вывода в секунду), по данным математического моделирования распределенной системы хранения данных вида [8]–[10], приведен на рис. 5.

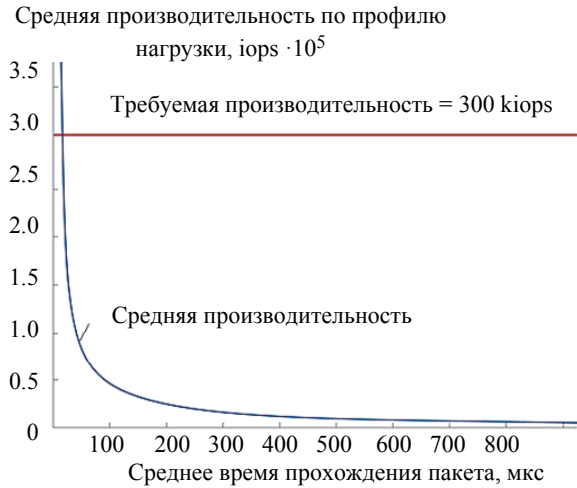


Рис. 5

Рост показателя производительности обмена данными при фиксированных показателях времени прохождения возможен при увеличении числа независимых потоков передачи данных в пределах пропускной способности канала связи. Соответственно, на базе математической модели распределенной системы хранения данных можно построить кривую зависимости минимально необходимого количества независимых потоков передачи данных, требуемого для обеспечения общей производительности обмена данными между клиентом и системой по профилю, от среднего времени прохождения пакета данных, что проиллюстрировано графиком зависимости количества потоков от среднего времени прохождения пакета при заданном количестве жестких дисков (100 шт.) для обеспечения требуемой производительности (300 kiops) (рис. 6).

На основе представленных графиков, их аппроксимации и анализа можно сделать вывод, что при априори известном профиле нагрузки с применением процедур сборки и восстановления потерянных данных в матрице кодирования/декодирования из блоков данных в параллельном режиме рекомендуется использовать минимально необходимое количество независимых потоков передачи данных и жестких дисков для обеспечения среднего времени ожидания и общей

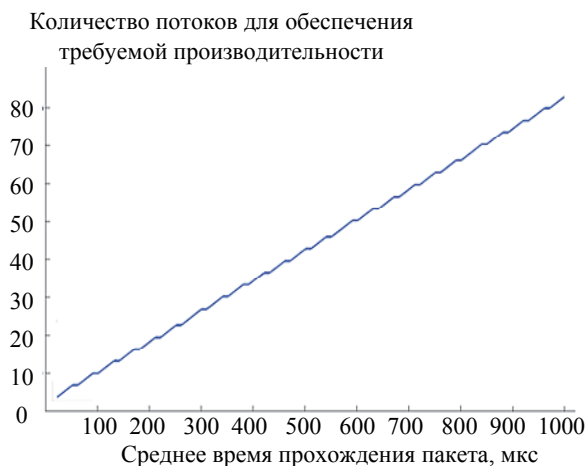


Рис. 6

производительности обмена данными между клиентом и системой по профилю.

Обсуждение результатов. Полученные и аппроксимированные результаты позволяют сделать вывод о правильном подходе к проводимым исследованиям [8]–[10], а выполненный анализ

находит подтверждение при разработке распределенных систем хранения данных [4].

В результате моделирования на базе полученных зависимостей (рис. 4 и 6) для известного профиля нагрузки появилась возможность проектирования минимально необходимой конфигурации распределенной системы хранения данных типа [8]–[10] с однородным массивом жестких дисков фиксированного типа в сети передачи данных с реально измеренными параметрами передачи пакета данных для обеспечения требуемых показателей производительности обмена данными и среднего времени ожидания.

Статья подготовлена в рамках выполнения работ по соглашению от 03 июля 2018 г. № 074-11-2018-012 с Министерством науки и высшего образования Российской Федерации (Минобрнауки России).

СПИСОК ЛИТЕРАТУРЫ

1. Разработка архитектуры распределенного интерконнекта / Г. Масич, А. Вотинова, А. Созыкин, А. Масич, В. Щапов, А. Игумнов, Бобров, А. Латыпов, С. Береснев, А. Куклин // Вестн. Пермского федерального исследовательского центра. 2016. № 4. С. 51–56.
2. Baranov A. V., Leshchev S. A. Methods and Means of Distributed Storage Systems Implementation // Software J.: Theory and Applications. 2019. № 3. P. 14–24.
3. Исследование и анализ производительности распределенного интерконнекта вычислительной среды в УрО РАН / А. С. Игумнов, А. Ю. Берсенев, А. Г. Масич, Г. Ф. Масич, В. А. Щапов // Суперкомпьютерные дни в России. 2016. URL: <http://2016.russianscdays.org/files/pdf16/199.pdf> (дата обращения 09.10.2020)
4. Сайт ООО «НПО Баум». URL: <http://npobaum.ru/> (дата обращения 09.10.2020)
5. Силин А. Производительность СХД: как определить, какая система подходит бизнесу? // Журн. сетевых решений. LAN. 2017. № 6. URL: <https://www.osp.ru/lan/2017/06/13052244> (дата обращения 26.10.2020)

6. Сайт Издания itWeek. URL: <https://www.itweek.ru/ecm/article/detail.php?ID=209396> (дата обращения 26.10.2020)

7. Трофимова П. СХД для потоковых данных: проблемы и решения // Журн. «Storage News». 2011. № 2 (46). URL: www.storagenews.ru (дата обращения 26.10.2020)

8. Выбор архитектурных решений построения системы хранения данных в технологии Wes-Scale IT / И. П. Иванов, А. П. Гантимуров, А. Д. Виниченко, А. В. Босов // Перспективы науки. 2019. № 8 (119). С. 6–70.

9. Методы оптимизации информационных потоков в бизнес-процессах с распределенной системой хранения данных / И. П. Иванов, А. П. Гантимуров, А. Д. Виниченко, А. В. Босов // Перспективы науки. 2019. № 9 (120). С. 35–39.

10. Повышение общей производительности информационной системы путем решения оптимизационной задачи подбора параметров технологии QoS / И. П. Иванов, А. П. Гантимуров, А. Д. Виниченко, А. В. Босов // Перспективы науки. 2019. № 12 (123). С. 78–82.

A. P. Gantimurov, A. V. Bosov
Bauman Moscow State Technical University Saint

ANALYSIS OF OPPORTUNITIES TO OPTIMIZE DISTRIBUTED STORAGE STRUCTURE UNDER KNOWN LOAD PROFILE

An analysis of the possibilities for determining the optimal structural characteristics of a distributed storage system is given, based on the basic characteristics of a homogeneous array of hard drives included in the system and measured indicators of data exchange over the network based on a mathematical model and a known load profile. First of all, the task is defined based on the multi-year practice of designing, developing, implementing and monitoring distributed storage systems. The key factors affecting the waiting times between the operation request and the start of data transmission and the overall performance of the data exchange. Mathematical modeling of influencing parameters and characteristics is performed, as well as their analysis taking into account data obtained during monitoring of operating distributed storage systems of various levels of complexity and architecture. The research in the paper will be useful to specialists in the field of design, development and organization of distributed storage systems of various complexity.

Distributed storage, load profile, data communications, data network, data performance