



БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ
ИНСТИТУТ ПО ИНФОРМАЦИОННИ И КОМУНИКАЦИОННИ
ТЕХНОЛОГИИ



Генка Иванова Славова-Торрес

МЕТОДИ И СРЕДСТВА ЗА ОПТИМИЗИРАНЕ НА ЗАПИСИ В МРЕЖОВА СРЕДА ЗА СЪХРАНЕНИЕ

АВТОРЕФЕРАТ

на дисертация
за придобиване на образователна и научна степен „доктор”

по докторска програма **Компютърни системи, комплекси и мрежи**
професионално направление **5.3 Комуникационна и компютърна техника**

научен ръководител: доц. д-р Владимир Иванов

София 2026

Дисертационният труд е обсъден и насрочен за защита на заседание на Научния съвет на Институт по информационни и комуникационни технологии при Българска академия на науките (ИИКТ–БАН), проведено на

Публичната защита на дисертационния труд ще се състои на от часа в ИИКТ–БАН.

Дисертационният труд е структуриран в Увод, 6 глави, Заключение, Научни приноси, Списък с Публикации, Декларация за оригиналност на резултатите и Библиография. Трудът е с общ обем от 147 страници и съдържа 24 фигури и 16 таблици.

Защитата на дисертационния труд ще се състои на2026 г. от часа в зала на блок 2 на ИИКТ-БАН на открито заседание на научно жури в състав:

Научно жури

1.
2.
3.
4.
5.

Рецензиите и становищата на членовете на научното жури, както и авторефератът, ще бъдат публикувани на интернет страницата на ИИКТ-БАН.

Материалите по защитата са на разположение на интересуващите се в стая 315 на ИИКТ-БАН, ул. „Акад. Г. Бончев“, бл. 2.

Автор: **Генка Иванова Славова-Торрес**

Заглавие: **Методи и средства за оптимизиране на записи в мрежова среда за съхранение**

I. ОБЩА ХАРАКТЕРИСТИКА НА ДИСЕРТАЦИЯТА

Актуалност на проблема

Съвременните мрежови среди за съхранение на данни (Storage Area Networks – SAN) са основен компонент на информационните инфраструктури, обслужващи виртуализирани среди, облачни платформи, системи за управление на бази данни и други приложения с високи изисквания към производителността и надеждността. Развитието на технологиите за съхранение и увеличаването на степента на паралелизъм съществено повишават пропускателната способност на тези системи, но едновременно с това поставят нови предизвикателства, свързани с поддържането на устойчиво и предсказуемо латентностно поведение при конкурентен достъп до споделени ресурси.

Традиционната оценка на SAN системите е насочена предимно към показатели като максимална пропускателна способност, средна латентност и брой входно-изходни операции за единица време. Практиката показва, че тези агрегирани показатели не са достатъчни за характеризиране на поведението на системата при високо натоварване, тъй като не отразяват влиянието на опашковите процеси, координационните взаимодействия и крайните латентностни отклонения, които определят качеството на обслужване в реални експлоатационни условия.

Особено значение придобиват операциите по запис, тъй като тяхното изпълнение включва взаимодействие между множество механизми за управление на заявки, буфериране, кеширане и синхронизация, което ги прави по-чувствителни към координационни ефекти и натрупване на заявки в сравнение с операциите по четене. Формирането на входно-изходни опашки, конкуренцията за споделени ресурси и координацията между отделните елементи по пътя на записите могат да доведат до нелинейно нарастване на латентността, което не може да бъде обяснено единствено чрез характеристиките на отделните компоненти на системата.

В този контекст възниква необходимост от системен подход, който разглежда латентността като резултат от взаимодействието между натоварването, степента на паралелизъм и процесите на обслужване по целия път на записите. Подобен подход създава предпоставки не само за по-задълбочено разбиране на поведението на SAN системите, но и за формулиране на принципи и практически насоки за оптимизация на операциите по запис чрез управление на паралелизма, входно-изходните опашки и координационните механизми.

Актуалността на настоящото изследване се определя от необходимостта за разработване на аналитичен модел и експериментално обоснована методология, които позволяват идентифициране на устойчивите режими на работа, ранно откриване на координационен стрес и избор на оптимизационни действия, насочени към поддържане на предсказуемо латентностно поведение при конкурентно натоварване.

Цел на изследването

Целта на настоящия дисертационен труд е да се изследва поведението на операциите по запис в мрежови среди за съхранение на данни (Storage Area Networks - SAN) чрез разработване на системен модел на пътя на записите (write path), експериментално изследване на влиянието на натоварването, степента на паралелизъм и управлението на входно-изходните опашки върху латентността, както и чрез разработване на аналитична рамка за анализ и оптимизация на операциите по запис.

За постигането на тази цел се анализират зависимостите между характеристиките на генерираното натоварване, механизмите за управление на входно-изходните опашки, координацията между отделните елементи по пътя на записите и поведението на латентностните разпределения. Особено внимание е отделено на анализа на крайните персентили на латентността (p95 и p99), които служат като индикатори за проявление на опашкови ефекти и координационен стрес при конкурентно натоварване.

Изследването има за цел да идентифицира устойчивите режими на работа на SAN системите и да формулира принципи за оптимизация, основани на управление на паралелизма, входно-изходните опашки и конкурентния достъп до споделени ресурси. В този контекст оптимизацията се разглежда не като максимизиране на пропускателната способност, а като поддържане на устойчиво, предсказуемо и контролируемо латентностно поведение при запазване на необходимата производителност.

Чрез съчетаването на разработения системен модел, аналитичната интерпретация, експерименталната методология и рамката за оптимизация настоящото изследване предлага цялостен подход за анализ, оценка и оптимизация на операциите по запис в SAN среда който може да бъде използван при проектирането, конфигурирането и експлоатацията на съвременни инфраструктури за съхранение на данни.

Задачи на изследването

За постигане на поставената цел в дисертационния труд са формулирани следните основни изследователски задачи:

1. Да се анализират архитектурните особености на съвременните мрежови среди за съхранение на данни (SAN) и факторите, влияещи върху поведението на операциите по запис.
2. Да се разработи системен модел на пътя на записите (write path), който разглежда процеса на запис като взаимодействие между хост системата, транспортния слой, входно-изходните опашки и системата за съхранение.
3. Да се разработи аналитичен модел за интерпретация на латентностното поведение, основан на взаимодействието между натоварването, паралелизма и процесите на обслужване по пътя на записите.
4. Да се дефинират показатели и методи за анализ на латентността, основани на емпирични разпределителни функции (ECDF), крайните персентили (p95 и p99) и принципите на теорията на масовото обслужване.
5. Да се разработи експериментална методология и да се проведат контролирани експерименти за изследване влиянието на параметри като степен на паралелизъм, дълбочина на входно-изходните опашки, форма на натоварването и конкуренцията между хостове върху поведението на операциите по запис.
6. Да се извърши аналитичен синтез и интерпретация на експерименталните резултати с цел идентифициране на устойчивите режими на работа, проявлението на координационен стрес и факторите, определящи латентностното поведение на SAN системите.
7. Да се разработи рамка за оптимизация на операциите по запис в SAN среда, базирана на анализа на латентностните разпределения, управлението на паралелизма и входно-изходните опашки, която да подпомага избора на устойчив режим на работа.

Обект и предмет на изследването

Обект на изследването в настоящия дисертационен труд са мрежовите среди за съхранение на данни (Storage Area Networks – SAN), използвани за осигуряване на блоков достъп до споделени ресурси в съвременни информационни инфраструктури.

Предмет на изследването е поведението на операциите по запис в SAN среда, разглеждано чрез взаимодействието между натоварването, степента на паралелизъм, входно-изходните опашки и архитектурните компоненти по пътя на записите, както и тяхното влияние върху латентността и нейното разпределение.

Изследването е фокусирано върху анализа на латентностните разпределения, включително крайните персентили (p95 и p99), като показатели за оценката на устойчивостта и предсказуемостта на поведението на системата при конкурентно натоварване. Особено внимание е отделено на ролята на входно-изходните опашки и механизмите за координация между отделните слоеве на SAN архитектурата като фактори, определящи проявлението на опашкови ефекти и координационен стрес.

В този контекст операциите по запис се разглеждат като системен процес, при който латентността е резултат от взаимодействието между процесите на обслужване и изчакване по целия път на записите. Това позволява анализът да бъде използван като основа за разработване на аналитичен модел и рамка за оптимизация на операциите по запис в SAN среда.

Изследователска теза

Основната изследователска теза на настоящия дисертационен труд е, че **поведението и оптимизацията на операциите по запис в SAN среда се определят от съгласуването между характеристиките на натоварването, степента на паралелизъм и управлението на входно-изходните опашки по пътя на записите.**

За разлика от традиционния подход, при който оптимизацията се свързва предимно с повишаване на производителността на отделните хардуерни компоненти, настоящото изследване приема, че при висока степен на конкурентност определящо значение имат взаимодействията между входящия поток от заявки, процесите на обслужване и формирането на входно-изходните опашки. В тези условия устойчивостта и предсказуемостта на латентностното поведение се определят не само от характеристиките на системата за съхранение, а и от начина, по който се управляват паралелизмът и конкуренцията по целия път на записите.

Въз основа на тази теза се приема, че анализът на латентностните разпределения, крайните персентили (p95 и p99) и опашковите ефекти позволява идентифициране на

устойчивите режими на работа и създава основа за разработване на аналитична рамка за оптимизация на операциите по запис в SAN среда.

Методология на изследването

Методологията на настоящото изследване се основава на системен експериментално-аналитичен подход, предназначен за изследване и оптимизация на операциите по запис в мрежови среди за съхранение на данни (SAN). Тя обединява концептуален анализ, разработване на системен и аналитичен модел, експериментално изследване, статистически анализ на резултатите и формулиране на оптимизационна рамка.

В началния етап е извършен анализ на архитектурните особености на SAN средите и е разработен системен модел на пътя на записите (write path), който разглежда операцията по запис като последователност от взаимодействащи компоненти – хост система, транспортна инфраструктура, входно-изходни опашки и система за съхранение. На тази основа е изграден аналитичен модел за интерпретация на латентността поведение, отчитащо взаимодействието между натоварването, степента на паралелизъм и процесите на обслужване.

За проверка на изследователската теза е разработена експериментална методология, реализирана в контролирана тестова среда. Експериментите са проведени при систематично изменение на отделни параметри на натоварването, включително степен на паралелизъм, дълбочина на входно-изходните опашки, форма на натоварването и конкурентен достъп, при запазване на останалите параметри непроменени. Този подход позволява надеждно идентифициране на влиянието на всеки фактор върху латентността поведение на системата.

Интерпретацията на експерименталните резултати се основава на използването както на традиционни показатели за производителност, така и на анализ на емпиричните разпределителни функции (ECDF), медианата, средната стойност и крайните персентили на латентността (p95 и p99). За интерпретация на наблюдаваните зависимости са използвани принципите на теорията на масовото обслужване, които позволяват разграничаване на приноса на процесите на обслужване и на времената за изчакване при формиране на крайната латентност.

В заключителния етап е извършен аналитичен синтез на експерименталните резултати, въз основа на който са идентифицирани устойчивите режими на работа,

проявленията на координационен стрес и основните фактори, определящи латентностното поведение на SAN системите. Получените резултати са използвани за разработване на аналитична рамка за оптимизация на операциите по запис, която интегрира системния модел, експерименталните наблюдения и аналитичната интерпретация в последователен процес за анализ, оценка и управление на латентността.

Ограничения на изследването

Настоящото изследване е фокусирано върху анализа и оптимизацията на операциите по запис в мрежови среди за съхранение на данни (Storage Area Networks – SAN) при ясно дефинирана архитектурна и експериментална рамка. Ограниченията на изследването са съзнателно приети с цел осигуряване на контролируеми условия и надеждна интерпретация на получените резултати.

Изследването разглежда единствено операциите по запис, тъй като те включват по-сложни механизми на координация, управление на входно-изходните опашки и взаимодействие между компонентите на системата. Операциите по четене не са предмет на настоящото изследване.

Експериментите са проведени в контролирана среда с фиксирана инфраструктурна конфигурация, като последователно се изменят единствено параметрите на натоварването. Този подход позволява да бъдат идентифицирани причинно-следствените зависимости между характеристиките на натоварването, входно-изходните опашки и латентността, без влиянието на допълнителни променливи, свързани с промяна на хардуерната или софтуерната конфигурация.

Изследването не разглежда детайлно вътрешните механизми на конкретни търговски системи за съхранение, като алгоритми за кеширане, вътрешно планиране на заявки, контролерни политики и други проприетарни механизми. Тези компоненти са третираны като част от обобщения системен модел, като акцентът е поставен върху крайното латентностно поведение, а не върху специфичните особености на отделни реализации.

Също така не се анализират в детайли специфичните характеристики на транспортната инфраструктура, включително особеностите на SAN комутаторите, механизмите за управление на трафика и конкретните реализации на транспортните протоколи. Тези фактори могат да окажат влияние върху количествените стойности на измерваните показатели, но не променят основните закономерности, установени при

анализа на взаимодействието между натоварването, паралелизма и входно-изходните опашки.

Въпреки посочените ограничения разработените системен модел, експериментална методология и аналитична рамка осигуряват надеждна основа за анализ и оптимизация на операциите по запис в SAN среда.

Практическа приложимост на резултатите

Получените резултати имат практическа приложимост при проектирането, конфигурирането и експлоатацията на мрежови среди за съхранение на данни (Storage Area Networks – SAN), в които е необходимо поддържане на устойчиво и предсказуемо латентностно поведение при конкурентно натоварване.

Разработените системен модел, аналитична интерпретация и рамка за оптимизация предоставят методологична основа за анализ на операциите по запис и за идентифициране на факторите, които водят до формиране на опашкови ефекти и координационен стрес. Това позволява по-информиран избор на параметри, свързани със степента на паралелизъм, управлението на входно-изходните опашки и режима на работа на системата.

Предложената експериментална методология и използваните показатели за анализ, включително емпиричните разпределителни функции (ECDF) и крайните персентили на латентността (p95 и p99), могат да бъдат използвани при оценка, диагностика и мониторинг на SAN среди, където традиционните показатели за производителност не предоставят достатъчно информация за устойчивостта на системното поведение.

Практическата приложимост на предложената рамка се изразява във възможността за идентифициране на устойчивите режими на работа, ранно откриване на координационен стрес и избор на подходящи оптимизационни действия, насочени към управление на паралелизма, входно-изходните опашки и конкурентния достъп до споделени ресурси.

Получените резултати са приложими при проектиране и експлоатация на SAN инфраструктури, в които операциите по запис са част от критични за производителността процеси - например транзакционни бази данни, виртуализирани среди, мултитенант платформи и системи с интензивен конкурентен достъп до споделени ресурси. В такива среди устойчивото латентностно поведение и контролът върху крайните стойности на

латентността са съществени условия за поддържане на предсказуемо качество на обслужване.

II. СТРУКТУРА НА ДИСЕРТАЦИОННИЯ ТРУД

Дисертационният труд е структуриран в шест глави, заключение, списък на научните приноси, списък на използваната литература и приложения.

Първа глава представя предметната област и формулира научния проблем. Разгледана е актуалността на изследването, дефинирани са целта, задачите, обектът, предметът и изследователската теза на дисертационния труд. Представени са използваната методология, ограниченията на изследването и практическата приложимост на получените резултати.

Втора глава е посветена на разработването на системен модел на операциите по запис в SAN среда. Анализирани са архитектурните особености на SAN инфраструктурата, пътят на записите (write path) и ролята на взаимодействието между хост системата, транспортната инфраструктура, входно-изходните опашки и системата за съхранение при формиране на латентността. Представен е аналитичен модел за интерпретация на латентностното поведение, основан на принципите на теорията на масовото обслужване.

Трета глава представя разработената експериментална методология. Описани са експерименталната среда, тестовите сценарии, използваните натоварвания и показателите за анализ. Формулирана е логиката на експерименталния дизайн, която позволява систематично изследване на влиянието на различни фактори върху поведението на операциите по запис.

Четвърта глава съдържа резултатите от проведените експериментални изследвания. Анализирано е влиянието на степента на паралелизъм, дълбочината на входно-изходните опашки, формата на натоварването и конкурентния достъп върху латентността на операциите по запис. Представени са анализи на латентностните разпределения, емпиричните разпределителни функции (ECDF) и крайните персентили на латентността.

Пета глава представя аналитичния синтез на експерименталните резултати. Извършена е интерпретация на наблюдаваните зависимости чрез принципите на теорията на масовото обслужване, анализирани са опашковите ефекти и режимите на

работа на системата, формулирани са принципи за оптимизация на операциите по запис и е разработена аналитична рамка за тяхната оптимизация. Представени са също практическите приложения на получените резултати и възможностите за използването им при проектиране, конфигуриране и експлоатация на SAN среди.

Шеста глава съдържа заключението на дисертационния труд. Обобщени са основните научни резултати, формулирани са научните приноси, представени са ограниченията на изследването и са очертани основните направления за бъдеща научноизследователска работа.

III. СЪДЪРЖАНИЕ НА ДИСЕРТАЦИЯТА

ГЛАВА 1: ПРОБЛЕМ, ЦЕЛИ И ЗАДАЧИ

В първа глава е обоснована актуалността на изследването и е формулиран научният проблем, свързан с анализа и оптимизацията на операциите по запис в съвременни мрежови среди за съхранение на данни (Storage Area Networks – SAN). Показано е, че традиционните показатели за производителност, основани предимно на средна латентност и пропускателна способност, не са достатъчни за оценка на поведението на SAN системите при конкурентно натоварване. Аргументирана е необходимостта от системен подход, който отчита влиянието на натоварването, степента на паралелизъм и входно-изходните опашки върху устойчивостта и предсказуемостта на латентностното поведение.

Показано е, че традиционните показатели за производителност, основани предимно на средна латентност и пропускателна способност, не са достатъчни за оценка на поведението на SAN системите при конкурентно натоварване. Аргументирана е необходимостта от анализ на латентностните разпределения и крайните персенти като основа за по-точна оценка на устойчивостта и предсказуемостта на операциите по запис.

В главата са формулирани целта, задачите, обектът, предметът и изследователската теза на дисертационния труд. Основната теза е, че **поведението и оптимизацията на операциите по запис в SAN среда се определят от съгласуването между характеристиките на натоварването, степента на паралелизъм и управлението на входно-изходните опашки по пътя на записите**, което налага използването на системен подход при анализа и оптимизацията на латентностното поведение.

Представени са използваната методология, ограниченията на изследването и практическата приложимост на очакваните резултати, както и общата логика на дисертационния труд.

ГЛАВА 2: СИСТЕМЕН МОДЕЛ НА ОПЕРАЦИИТЕ ПО ЗАПИС В МРЕЖОВИ СРЕДИ ЗА СЪХРАНЕНИЕ

Във втора глава е разработен системен модел на операциите по запис в SAN среда, който служи като концептуална основа за последващия експериментален анализ и за формулирането на предложената оптимизационна рамка. Моделът разглежда операцията по запис като процес, реализиран чрез взаимодействието между хост системата, транспортната инфраструктура и системата за съхранение, като отчита влиянието на механизмите за управление на паралелизма и входно-изходните опашки върху латентността.

Анализирана е еволюцията на факторите, определящи производителността на системите за съхранение. Показано е, че **при съвременните SSD и NVMe архитектури ограниченията все по-рядко се определят от физическите характеристики на устройствата и все по-често произтичат от взаимодействията между конкурентните заявки, механизмите за синхронизация и управлението на входно-изходните опашки.**

На тази основа е въведена концепцията за **пътя на записите (write path)** като последователност от взаимодействащи архитектурни слоеве, включващи хост системата, SAN транспортната инфраструктура и сторидж подсистемата. Разработеният модел представя всеки от тези слоеве като система от обслужващи звена с входно-изходни опашки, чието съвместно поведение определя крайната латентност на операциите по запис.

Концептуалният модел на пътя на записите е представен на **Фигура 1**, която илюстрира разпределението на входно-изходните опашки и механизмите за управление по различните нива на системата.



Фигура 1. Концептуален модел на пътя на записите в мрежова среда за съхранение, илюстриращ разпределението на опашките и регулаторите по различните нива на системата

Особено внимание е отделено на ролята на теорията на масовото обслужване при интерпретацията на експерименталните резултати. Общата латентност се разглежда като следствие от взаимодействието между процесите на обслужване и натрупването на заявки във входно-изходните опашки, което позволява тя да бъде интерпретирана като системна характеристика на SAN средата.

Показано е, че при увеличаване на степента на натоварване натрупването на заявки във входно-изходните опашки води до нелинейно нарастване на времената за изчакване и до формиране на опашкови ефекти. В тази връзка са анализирани различни показатели за оценка на латентността, като е аргументирано използването на емпиричните разпределителни функции (ECDF) и високите персентили (p95 и p99) като по-информативни показатели за оценка на устойчивостта и предсказуемостта на системното поведение.

В резултат е разработен системен модел, който описва зависимостите между натоварването, паралелизма, входно-изходните опашки и латентността и служи като аналитична основа за експерименталното изследване, аналитичния синтез и предложената рамка за оптимизация на операциите по запис в SAN среда.

ГЛАВА 3: ЕКСПЕРИМЕНТАЛНА МЕТОДОЛОГИЯ

Трета глава е посветена на разработването на експерименталната методология, използвана за изследване на поведението на операциите по запис в SAN среда. Основната цел е изграждането на възпроизводима експериментална рамка, която позволява систематично изследване на влиянието на натоварването, степента на паралелизъм и

управлението на входно-изходните опашки върху латентностното поведение на системата.

Методологията е изградена върху принципа за контролирано изменение на експерименталните параметри, при който във всеки експеримент се изменя само един фактор, а останалите условия се запазват непроменени. Този подход позволява еднозначно определяне на влиянието на отделните параметри върху наблюдаваното поведение и осигурява висока възпроизводимост на резултатите.

В главата е представена експерименталната среда, включваща хост система, SAN транспортна инфраструктура и сторидж подсистема. Дефинирани са използваните конфигурационни параметри, характеристиките на натоварването и условията за провеждане на експериментите.

Таблица 1 представя фиксираните параметри на експерименталния дизайн.

категория	параметър	стойност
<i>Хост (инициатор)</i>	Брой хостове	1
	Тип достъп	Блоков достъп до SAN (без файлови системи)
	Операционна система	Linux-базирана
	Генериране на натоварване	fio
	Multipath	Активен (2 независими I/O пътеки)
<i>SAN транспорт</i>	Технология	Fibre Channel
	Свързаност	2 × 16 Gbit/s
	Топология	Редундантна, без single-path зависимости
<i>Система за съхранение</i>	Архитектура	Двуконтролерна SAN система
	Write cache	Активен
	Back-end	Flash-базирана среда
<i>Работен набор</i>	Логически капацитет	1 TB
	Работен набор	256 GB
<i>Натоварване (baseline)</i>	Тип операции	100% write
	Размер на блока	4 KB random write
<i>Измерване</i>	Продължителност на тест	300 s
	Warm-up период	30 s
	Повторяемост	3 повторения

Таблица 1. Фиксирани параметри на експерименталния дизайн

За анализ на различните аспекти на поведението на системата са разработени няколко експериментални сценария.

Първият сценарий представлява базов експериментален режим при равномерно натоварване и служи като референтна конфигурация за сравнение с останалите експерименти.

Вторият сценарий изследва влиянието на дълбочината на входно-изходната опашка (Queue Depth – QD) върху латентността на операциите по запис и производителността на системата. Целта е да се установи как увеличаването на броя едновременно поддържани входно-изходни заявки влияе върху средната латентност, високите персентили и пропускателната способност, както и дали нарастването на паралелизма води до устойчиво подобрене или до натрупване на заявки в опашките.

Третият сценарий анализира влиянието на профила на натоварването върху латентността на операциите по запис чрез сравнение между равномерно и пулсиращо генериране на входно-изходни заявки. Целта е да се оцени как временните колебания в интензивността на натоварването влияят върху натрупването на заявки и латентностното поведение.

Четвъртият сценарий изследва влиянието на степента на координация между конкурентните операции по запис чрез сравнение между координирано и некоординирано подаване на заявки. Целта е да се оцени как синхронизацията на заявките влияе върху взаимодействието между входно-изходните опашки и върху формирането на крайната латентност.

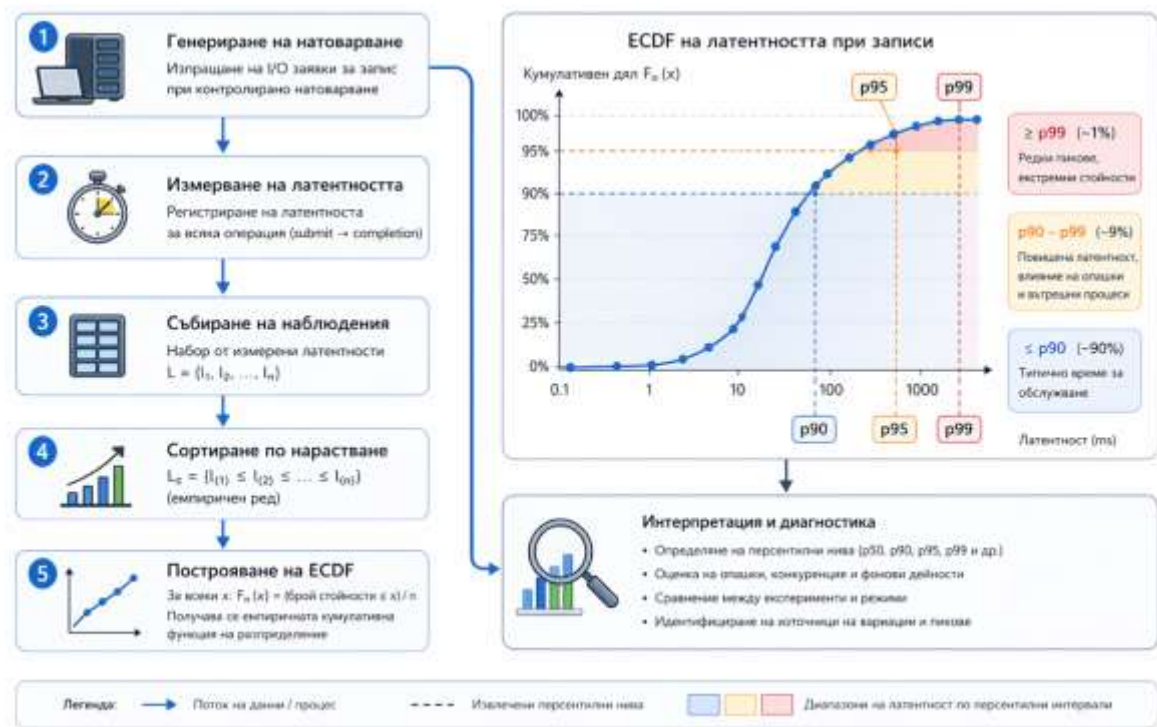
Петият сценарий изследва влиянието на конкурентния достъп на независим хост (Noisy Neighbor) върху латентността на операциите по запис при използване на споделен том. Целта е да се анализира как конкуренцията за общите ресурси на системата за съхранение променя разпределението на латентността, води до нарастване на крайните персентили и влияе върху устойчивостта на латентностното поведение.

Петте експериментални сценария са разработени така, че последователно да изследват различни аспекти на поведението на SAN системата - от базовото латентностно поведение до влиянието на паралелизма, профила на натоварването, координацията между заявките и конкуренцията за споделени ресурси.

Особено внимание е отделено на избора на показателите за анализ. Освен традиционните показатели за производителност като пропускателна способност и средна латентност, методологията използва анализ на емпиричните разпределителни функции (ECDF), както и на високите латентностни персентили (p95 и p99). Тези показатели

позволяват идентифициране на опашковите ефекти и ранните признаци на координационен стрес, които не могат да бъдат установени чрез анализ, основан единствено на усреднени стойности.

Потокът на измерване и анализ на латентността е представен на **Фигура 2**.



Фигура 2. Поток на измерване и интерпретация на латентността чрез емпирична кумулативна функция на разпределение (ECDF)

Методологията включва също последователна обработка и интерпретация на експерименталните резултати чрез статистически анализ и съпоставка между различните експериментални сценарии. Анализът е насочен към идентифициране на зависимостите между натоварването, входно-изходните опашки и латентностното поведение, както и към определяне на условията, при които системата преминава от устойчив режим към координационен стрес.

Разработената експериментална методология създава необходимата основа за аналитичния синтез и за формулирането на предложената рамка за оптимизация на операциите по запис в SAN среда, представени в следващите глави на дисертационния труд.

ГЛАВА 4: ЕКСПЕРИМЕНТАЛНИ РЕЗУЛТАТИ

Четвърта глава представя резултатите от проведените експериментални изследвания на поведението на операциите по запис в SAN среда. Основната цел е експериментално да се изследва влиянието на различни фактори, включително степента на паралелизъм, дълбочината на входно-изходните опашки, характеристиките на натоварването, координацията между заявките и конкурентния достъп до споделените ресурси, върху латентността на операциите по запис.

Анализът започва с базов експериментален сценарий, който служи като референтна конфигурация за сравнение с останалите експерименти. При умерено натоварване системата демонстрира компактно латентностно разпределение и ограничена вариабилност, характерни за устойчив режим на работа. Получените резултати служат като отправна точка за оценка на промените в латентностното поведение при систематично изменение на отделните параметри на натоварването.

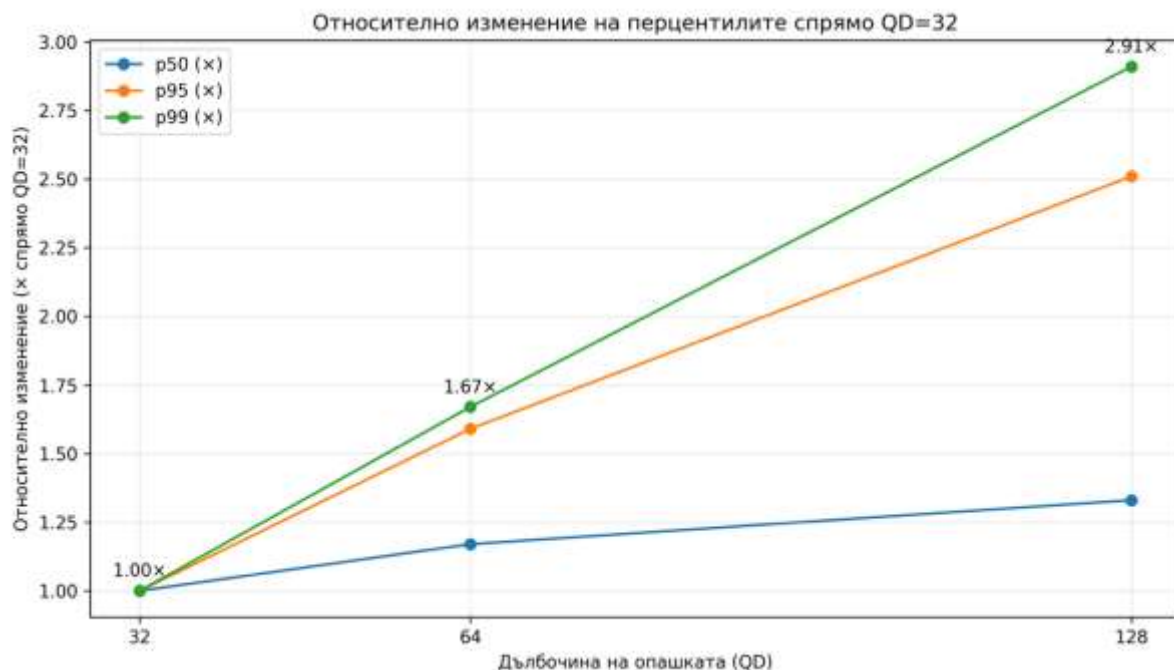
Експеримент 1: Влияние на дълбочината на входно-изходната опашка (Queue Depth)

Първият експеримент изследва влиянието на дълбочината на входно-изходната опашка (Queue Depth – QD) върху латентността на операциите по запис и производителността на системата. Чрез систематично увеличаване на стойността на Queue Depth се анализира как повишаването на степента на паралелизъм влияе върху различните латентностни персентили.

Получените резултати показват, че увеличаването на QD води до едновременно повишаване на производителността и влошаване на латентностното поведение. При увеличаване на QD от 32 на 128 производителността нараства от 42 000 на 52 000 IOPS (с приблизително 24%), докато средната латентност се увеличава умерено – от 0.8 ms на 1.1 ms, а медианната латентност (p50) – от 0.6 ms на 0.8 ms.

Значително по-силно изменение се наблюдава при високите латентностни персентили. Стойността на p95 нараства от 3.9 ms на 9.8 ms, а p99 – от 11.7 ms на 34.0 ms, което представлява почти трикратно увеличение спрямо референтния сценарий. Получените резултати показват, че повишаването на степента на паралелизъм засяга предимно крайната част на латентностното разпределение, докато типичното време за обслужване на по-голямата част от операциите остава сравнително стабилно.

Относителното изменение на латентностните персенти при различни стойности на Queue Depth е представено на Фигура 3.



Фигура 3. Относително изменение на латентностните персенти (p50, p95 и p99) спрямо дълбочината на входно-изходната опашка

Анализът показва, че увеличаването на паралелизма подобрява използването на наличните ресурси, но след определено ниво води до непропорционално нарастване на крайните латентностни стойности. Това показва, че при високи стойности на QD деградацията на латентността се проявява главно чрез усилване на редките, но критични случаи на забавяне, а не чрез равномерно влошаване на латентността за всички операции.

Експеримент 2: Влияние на формата на натоварването върху латентността на операциите по запис

Вторият експеримент изследва влиянието на различните форми на натоварване върху латентността на операциите по запис. За целта е извършено сравнение между равномерно (steady) и пулсиращо (bursty) генериране на входно-изходни заявки при фиксиран входен паралелизъм и непроменена системна конфигурация.

Получените резултати показват, че промяната във времевата структура на натоварването оказва съществено влияние върху латентностното поведение, въпреки че средната производителност остава практически непроменена. Производителността

намалява с едва 0.6% (от 48 200 на 47 900 IOPS), а медианната латентност (p50) се увеличава умерено – от 1.4 ms на 1.6 ms.

Съществените различия се проявяват при високите латентностни персенти. Стойността на p95 нараства от 6.9 ms на 13.7 ms, а p99 – от 11.8 ms на 26.4 ms, което представлява увеличение с приблизително 124%. Това показва, че пулсиращото натоварване води до временни натрупвания на заявки във входно-изходните опашки и значително увеличава вероятността от възникване на операции с висока латентност.

Получените резултати показват, че при еднакво средно ниво на натоварване времевата структура на подаване на заявките представлява самостоятелен фактор, определящ поведението на латентностното разпределение. Следователно оценката на производителността не може да се основава единствено на средните показатели, тъй като те не отразяват същественото влошаване на крайните латентностни стойности при пулсиращо натоварване.

Експеримент 3: Координация срещу сурова скорост

Третият експеримент изследва влиянието на координацията между операциите по запис върху латентностното поведение на SAN системата. Сравнени са два режима на работа – некоординиран режим (R1), ориентиран към максимална скорост на подаване на заявките, и координиран режим (R2), при който подаването на операциите по запис е съобразено с текущото състояние на системата.

Получените резултати са представени в Таблица 2.

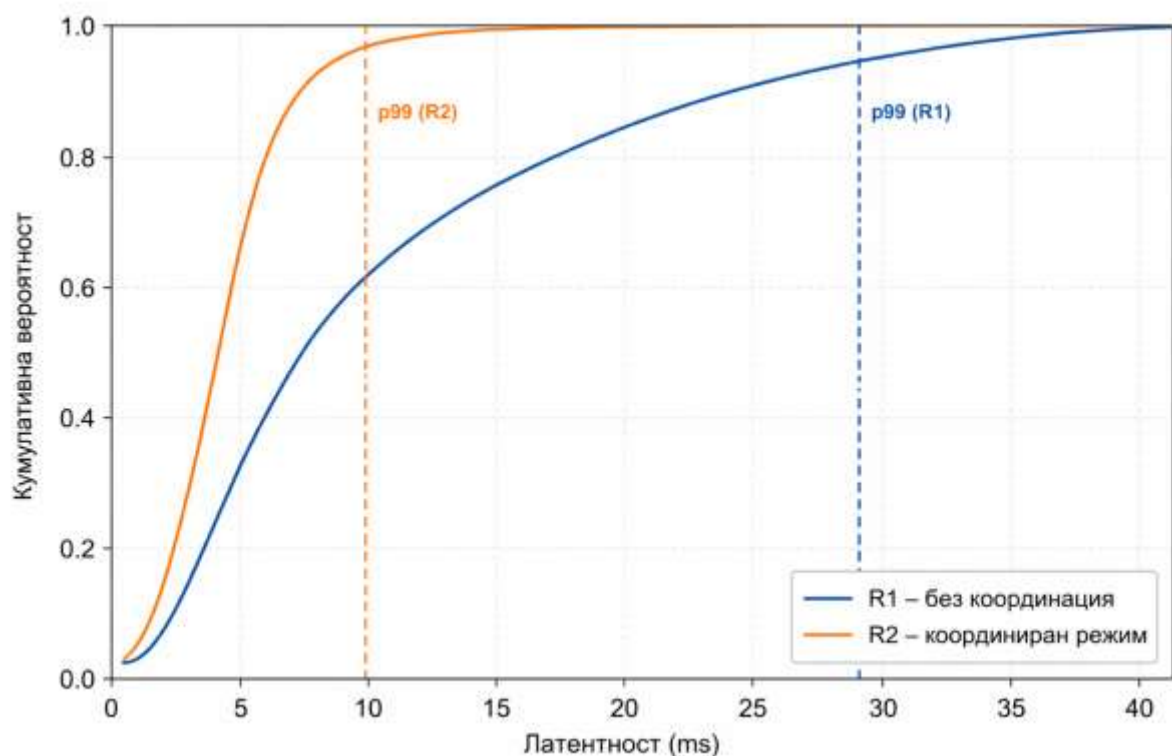
Режим	Производителност (IOPS)	p50 (ms)	p95 (ms)	p99 (ms)
R1 – без координация (raw speed)	43 000	1.8	14.5	28.0
R2 – координиран режим	41 000	1.9	7.5	12.0

Таблица 2. Агрегирани показатели за латентност и производителност при некоординиран и координиран режим

На пръв поглед двата режима се различават незначително по производителност и медианна латентност. Производителността намалява от 43 000 IOPS при некоординирания режим до 41 000 IOPS при координирания режим, докато медианната латентност (p50) остава практически непроменена (1.8 ms срещу 1.9 ms). Ако анализът се ограничи единствено до тези показатели, може да се направи извод, че поведението на системата е почти идентично.

Съществените различия се проявяват при високите латентностни персенти. Стойността на p95 намалява от 14.5 ms до 7.5 ms, а p99 – от 28.0 ms до 12.0 ms. Това показва, че координацията между заявките оказва ограничено влияние върху типичното време за изпълнение на операциите по запис, но значително намалява вероятността от възникване на редки операции с висока латентност.

По-задълбочен анализ е извършен чрез емпиричните кумулативни функции на разпределение (ECDF), представени на Фигура 4.



Фигура 4. ECDF на латентността на операциите по запис при некоординиран режим (R1) и координиран режим (R2) при фиксиран входен паралелизъм

ECDF анализът потвърждава количествените резултати, представени в Таблица 2. При координирания режим производителността намалява с приблизително 4.7% (от 43

000 на 41 000 IOPS), докато стойността на p99 се понижава с приблизително 57% (от 28.0 ms на 12.0 ms). В същото време медианната латентност (p50) остава практически непроменена. Това показва, че сравнително малко намаление на производителността води до съществено подобрене на латентностното поведение и значително ограничава появата на операции с екстремно висока латентност.

ECDF анализът показва още, че при некоординирания режим латентностното разпределение е значително по-разтеглено и се характеризира с по-дълга латентностна опашка. Следователно използването единствено на производителността или на средните показатели за латентност не е достатъчно за оценка на поведението на SAN системата, тъй като те не отразяват редките, но критични случаи на силно забавени операции по запис.

Получените резултати показват, че надеждната оценка на поведението на SAN среда изисква едновременно използване на анализ на латентностните разпределения чрез ECDF и на високите латентностни персенти (p95 и p99). Именно този подход позволява идентифициране на координационните ефекти и ранните признаци на нестабилност, които остават скрити при използване единствено на усреднени показатели.

Експеримент 4: Влияние на конкурентен хост (Noisy Neighbor)

Четвъртият експеримент изследва влиянието на конкурентния достъп до споделен ресурс върху латентността на операциите по запис. Анализирано е поведението на системата при едновременно генериране на входно-изходни заявки от два независими хоста, използващи общ том за съхранение.

Получените резултати показват, че добавянето на конкурентен хост оказва съществено влияние върху латентностното поведение, въпреки че производителността на наблюдавания хост остава практически непроменена. Производителността намалява с едва около 1.2% (от 42 000 на 41 500 IOPS), докато медианната латентност (p50) се увеличава незначително – от 1.6 ms на 1.7 ms.

Съществените изменения се наблюдават при високите латентностни персенти. Стойността на p95 нараства от 6.8 ms на 15.2 ms (увеличение с приблизително 124%), а p99 – от 11.5 ms на 31.0 ms (увеличение с приблизително 170%). Това показва, че ефектът Noisy Neighbor се проявява предимно чрез значително увеличаване на вероятността от

възникване на редки операции с висока латентност, без съществено влияние върху типичното време за обслужване.

Получените резултати потвърждават, че при SAN среди със споделени ресурси поведението на дадено натоварване зависи не само от неговите собствени параметри, но и от активността на останалите конкурентни натоварвания. Следователно конкуренцията за общите ресурси следва да се разглежда като самостоятелен фактор за възникване на латентностна нестабилност и координационен стрес.

Независимо дали причината е увеличаване на Queue Depth, промяна във формата на натоварването, липса на координация между заявките или конкуренция за споделени ресурси, всички експериментални сценарии показват една и съща закономерност – деградацията на поведението на SAN системата се проявява най-напред в крайната част на латентностното разпределение. Това потвърждава необходимостта от използване на разпределителен анализ и високи латентностни персентили при оценката на поведението на операциите по запис и създава експерименталната основа за аналитичния синтез, представен в следващата глава.

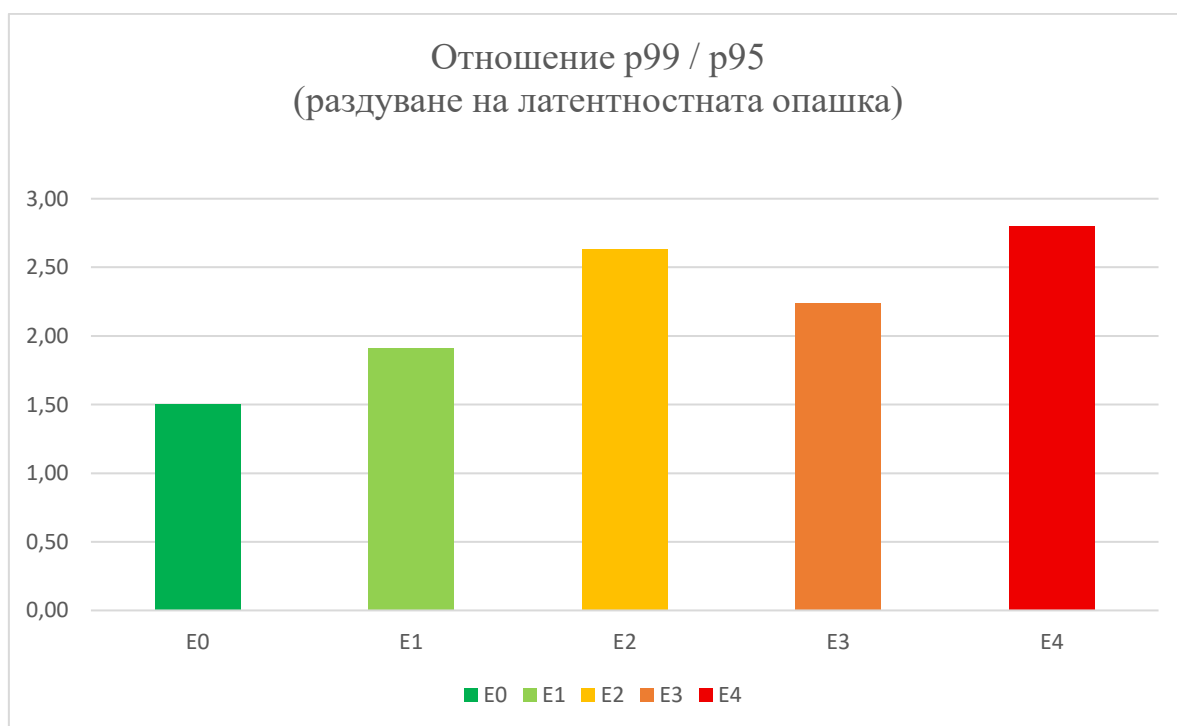
ГЛАВА 5: АНАЛИТИЧЕН СИНТЕЗ И ИНТЕРПРЕТАЦИЯ

Пета глава представя аналитичния синтез на проведените експериментални изследвания и извежда основните закономерности в поведението на операциите по запис в SAN среда. За разлика от предходната глава, в която резултатите са представени по отделни експериментални сценарии, тук те са интерпретирани като единна система от взаимосвързани зависимости между натоварването, управлението на входно-изходните опашки, степента на паралелизъм и латентностното поведение.

Анализът показва, че поведението на SAN системите не може да бъде описано чрез единични показатели за производителност. Получените резултати демонстрират наличието на ясно изразени режими на работа, които се различават както по начина на използване на ресурсите, така и по характера на латентностното разпределение. Установено е, че с увеличаване на конкурентното натоварване системата преминава последователно през устойчив режим, преходен режим и режим на координационен стрес, като преходът между тях се характеризира с нелинейно изменение на крайните латентностни стойности.

Особено внимание е отделено на анализа на латентностните разпределения. Показано е, че средната латентност не отразява в достатъчна степен поведението на системата при високо натоварване. Докато основната част от операциите продължава да се изпълнява в относително тесен интервал от стойности, малък процент от заявките формира удължена латентностна опашка, която определя устойчивостта на системата.

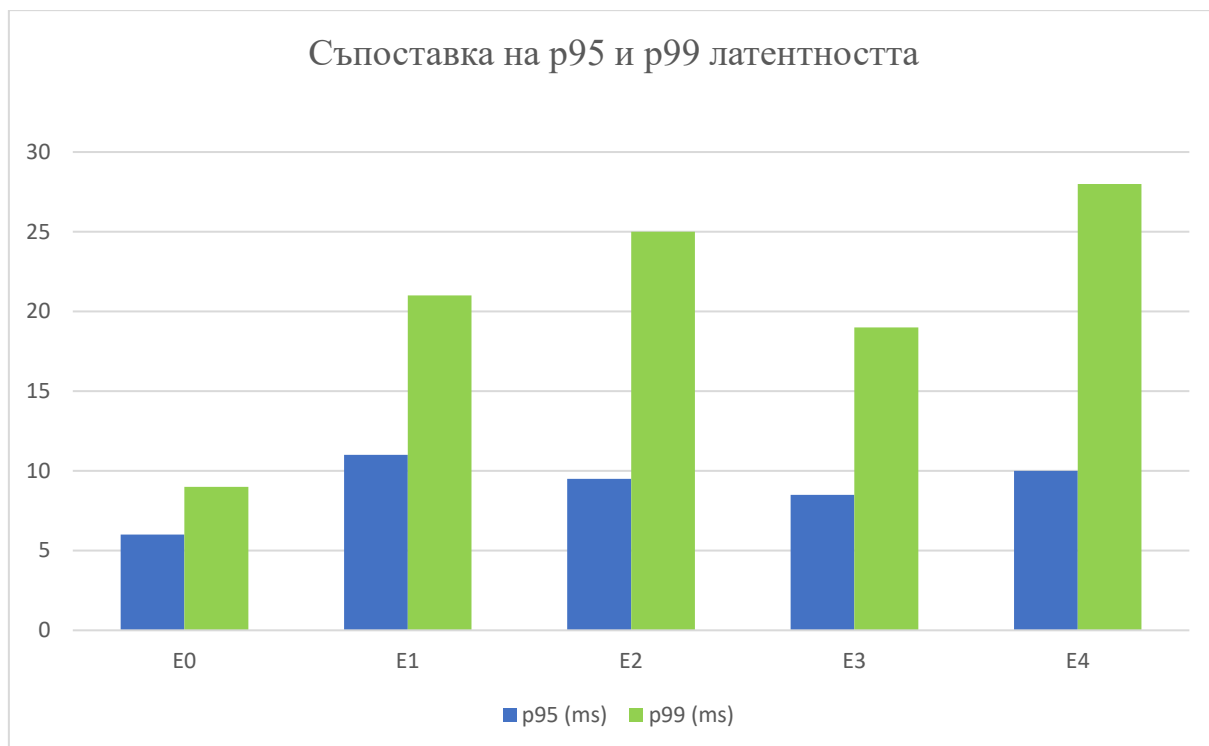
Раздуването на латентностната опашка е анализирано чрез отношението $p99/p95$, използвано като безразмерен показател за промяната във формата на разпределението.



Фигура 6. Раздуване на латентностната опашка ($p99/p95$) при различни експериментални сценарии

Получените резултати показват, че при ниско натоварване отношението $p99/p95$ остава близко до единица, което е характерно за компактно латентностно разпределение. При увеличаване на конкуренцията се наблюдава ускорено нарастване на $p99$, което показва разширяване на латентностната опашка и увеличаване на вероятността от редки, но значително забавени операции. Тази зависимост показва, че отношението $p99/p95$ може да се използва като количествен индикатор за началото на деградацията на латентностното поведение.

За по-подробно изследване на този ефект е извършена съпоставка между високите латентностни персентили.



Фигура 7. Съпоставка между p95 и p99 латентност при различни експериментални сценарии

Съпоставката показва, че при увеличаване на натоварването p95 остава сравнително стабилен, докато p99 нараства значително по-бързо. Това потвърждава, че деградацията на латентното поведение се проявява преди всичко в крайните случаи и не може да бъде установена чрез анализ на средните стойности. Получените резултати показват, че високите персентици представляват чувствителен индикатор за настъпващи опашкови ефекти и могат да бъдат използвани като критерий за ранно откриване на координационен стрес.

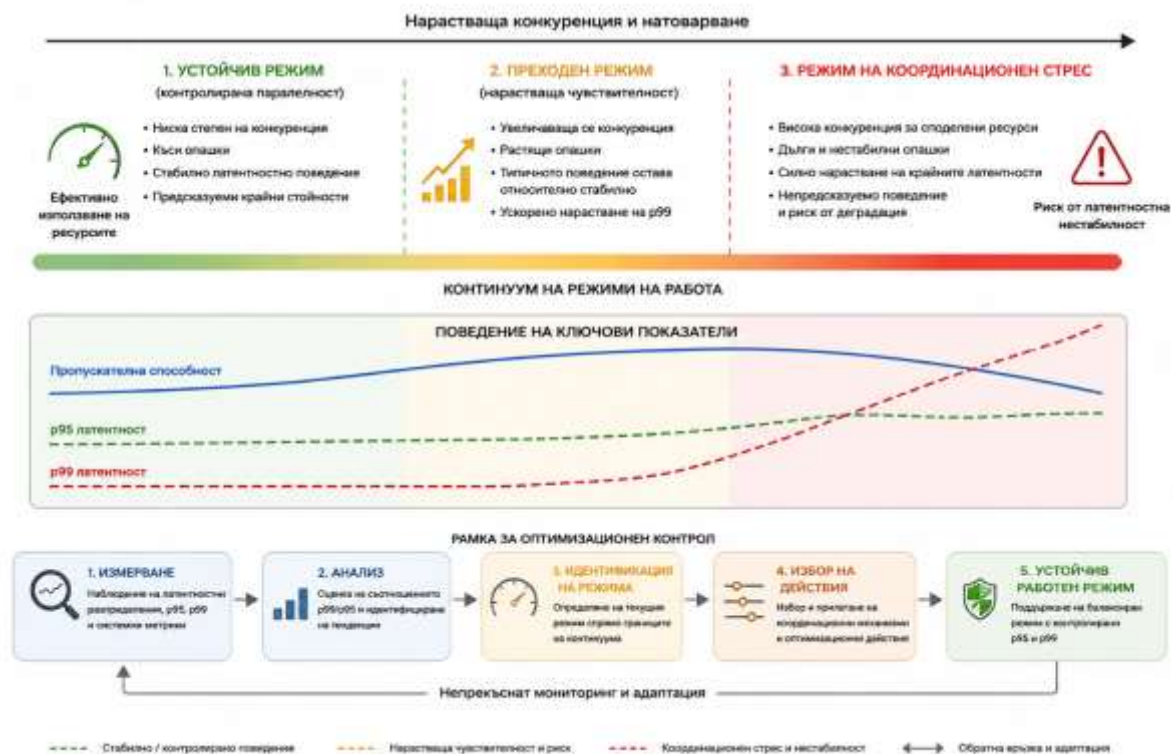
Наблюдаваното поведение е интерпретирано чрез принципите на теорията на масовото обслужване. При увеличаване на интензивността на входящия поток спрямо капацитета на обслужване времето за изчакване във входно-изходните опашки започва да доминира над времето за обслужване, което води до ускорено нарастване на крайната латентност.

Концептуално това поведение може да бъде представено чрез зависимостта

$$T_{latency} = T_{service} + T_{queue},$$

където крайната латентност се формира като сума от времето за обслужване и времето за изчакване във входно-изходните опашки. Анализът показва, че при ниско натоварване доминира компонентът $T_{service}$, докато при висока степен на конкуренция основен принос към крайната латентност започва да има T_{queue} .

Следващата част на аналитичния синтез разглежда компромиса между производителност, координация и устойчивост на латентностното поведение. Установено е, че увеличаването на степента на паралелизъм първоначално води до по-добро използване на ресурсите, но след достигане на определен праг започва да усилва координационните взаимодействия между заявките и да увеличава латентностната вариабилност.



Фигура 8. Режим на работа по пътя на записа при нарастваща конкуренция – от стабилно поведение към координационен стрес

Получените резултати показват, че оптималният режим на работа не съответства на максималната пропускателна способност, а на диапазона, в който производителността се увеличава без съществено разширяване на латентностното разпределение. Това позволява оптимизацията да бъде формулирана като задача за поддържане на системата в устойчив работен диапазон. Този диапазон съответства на областта, в която увеличаването на натоварването все още не води до ускорено нарастване на крайните латентностни персентили.

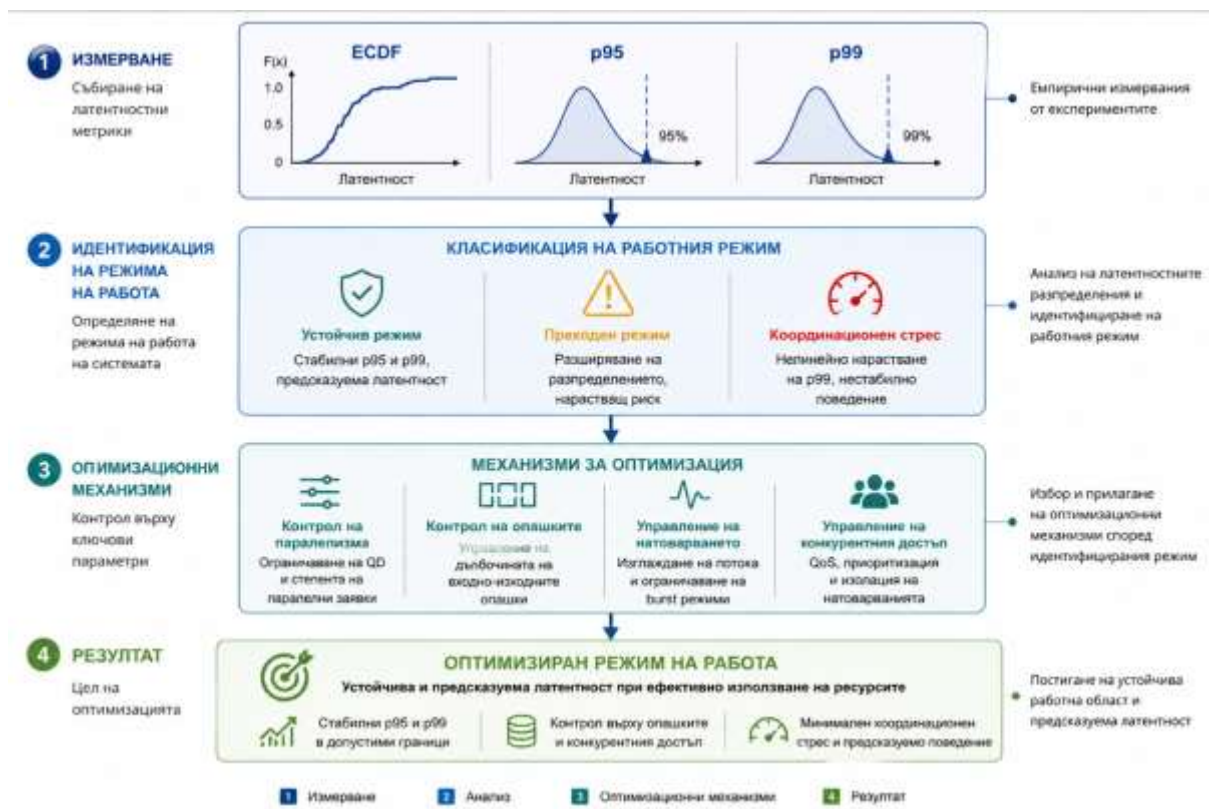
На тази основа е разработена оптимизационна интерпретация на експерименталните резултати. Предложени са принципи за управление на степента на паралелизъм, дълбочината на входно-изходните опашки, конкурентното натоварване и

анализа на латентностните разпределения като основа за избор на оптимални режими на работа.

Наблюдаван ефект	Вероятна причина	Основен показател	Оптимизационно действие	Очакван резултат
Нарастване на p99	Натрупване на заявки във входно-изходните опашки	p99, p99/p95	Намаляване на дълбочината на входно-изходната опашка (QD)	Намаляване на крайната латентност
Разширяване на латентностното разпределение	Координационен стрес	ECDF, p99/p95	Ограничаване на степента на паралелизъм	По-стабилно латентностно поведение
Латентностни пикове	Пулсиращо (burst) натоварване	ECDF, p99	Регулиране на интензивността на входно-изходните заявки	Намаляване на вариационността на латентността
Нестабилност при конкурентен достъп	Конкуренция за споделени ресурси	p99, вариационност на латентността	Прилагане на QoS, изолация или приоритизация на натоварванията	Ограничаване на взаимното влияние между натоварванията
Висока пропускателна способност при нарастващ p99	Работа в близост до капацитета на системата	Пропускателна способност, p99	Работа в устойчив работен диапазон	Баланс между производителност и предсказуема латентност

Таблица 3. Оптимизационни насоки, изведени от експерименталния анализ на операциите по запис в SAN среда

Разработена е и рамка за оптимизация на операциите по запис, реализирана като затворен цикъл, включващ измерване на латентностното поведение, идентификация на режима на работа, избор на оптимизационни действия и последваща оценка на тяхната ефективност чрез механизъм за обратна връзка.



Фигура 9. Рамка за оптимизация на операциите по запис в SAN среда

Предложената рамка трансформира експерименталните резултати в практически приложима методология за управление на SAN среди и позволява идентифициране на устойчивите режими на работа, ранно откриване на координационен стрес и избор на подходящи оптимизационни действия.

Наблюдаван ефект	Основен индикатор	Причина	Оптимизационно действие	Цел
Постепенно нарастване на p95/p99	p95, p99	Увеличаване на паралелизма	Ограничаване на паралелизма	Запазване на устойчив режим
Рязко нарастване на p99	p99/p95	Натрупване на заявки	Намаляване на дълбочината на входно-изходните опашки	Ограничаване на времената за изчакване
Разширяване на ECDF	ECDF	Координационен стрес	Настройка на натоварването	Стабилизиране на разпределението
Латентностни пикове	p99	Пулсиращо натоварване	Ограничаване на пулсиращото натоварване чрез регулиране на интензивността на заявките	Намаляване на вариативността
Конкуренция	p99	Споделени ресурси	Механизми за качество на обслужването (QoS); Изолация и приоритизиране на конкурентните натоварвания	Изолация

Наблюдаван ефект	Основен индикатор	Причина	Оптимизационно действие	Цел
Висока пропускателна способност	Пропускателна способност и p99	Работа близо до капацитета	Работа в устойчив диапазон	Баланс

Таблица 4. Връзка между наблюдаваните латентностни ефекти и препоръчителните оптимизационни действия

В заключение аналитичният синтез потвърждава основната изследователска теза на дисертационния труд. Показано е, че устойчивото латентностно поведение не се определя единствено от производителността на отделните компоненти, а от взаимодействието между натоварването, входно-изходните опашки и механизмите за координация по целия път на записите. На тази основа е разработена аналитична рамка за оптимизация, която интегрира системния модел, експерименталните резултати и тяхната интерпретация в единен подход за анализ, оценка и оптимизация на операциите по запис в SAN среда. По този начин експерименталните резултати се трансформират в практически приложима методология за избор и поддържане на устойчиви режими на работа в SAN среди.

На основата на проведения аналитичен синтез могат да бъдат формулирани следните основни изводи:

1. Латентностното поведение има системен характер. Латентността на операциите по запис в SAN среда се определя от взаимодействието между натоварването, степента на паралелизъм, входно-изходните опашки и механизмите за координация по целия път на записите, а не единствено от производителността на отделни компоненти.
2. Съществуват ясно разграничими режими на работа. При нарастващо натоварване системата преминава през устойчив режим, преходен режим и режим на координационен стрес. Преходът между тях се характеризира с нелинейно нарастване на крайната латентност вследствие на усилване на опашковите ефекти.
3. Анализът на латентностните разпределения предоставя по-надеждна оценка на поведението на системата. Показателите p95 и p99, както и анализът на емпиричните кумулативни функции на разпределение (ECDF), позволяват ранно откриване на деградация на

производителността и координационен стрес, които остават скрити при използване единствено на средни стойности.

4. Оптимизацията следва да бъде насочена към управление на системното поведение. Получените резултати показват, че устойчивото латентностно поведение се постига чрез управление на паралелизма, входно-изходните опашки и конкурентното натоварване, а не чрез максимизиране на пропускателната способност или увеличаване на хардуерните ресурси.
5. Разработената аналитична рамка осигурява основа за практическа оптимизация на SAN среди. Предложената рамка интегрира системния модел, аналитичната интерпретация и експерименталните резултати в последователен процес за анализ, идентифициране на работния режим, избор на оптимизационни действия и последваща оценка на тяхната ефективност.

В заключение на главата е показано, че разработеният системен модел, аналитичната интерпретация и проведените експерименти формират единна методологична основа за анализ и оптимизация на операциите по запис в SAN среда. Получените резултати демонстрират, че оптимизацията следва да се разглежда като процес на поддържане на устойчиво и предсказуемо латентностно поведение чрез управление на взаимодействията между натоварването, паралелизма и входно-изходните опашки. Предложената рамка трансформира експерименталните резултати в практически приложим подход за избор на устойчиви режими на работа, ранно откриване на координационен стрес и формулиране на оптимизационни действия при проектиране и експлоатация на SAN системи.

ГЛАВА 6: ЗАКЛЮЧЕНИЕ И БЪДЕЩА РАБОТА

В шеста глава на дисертационния труд са обобщени основните резултати от проведеното изследване, формулирани са ограниченията на изследването и са очертани перспективите за бъдеща научна работа.

Извършеният анализ потвърждава основната изследователска теза, че поведението на операциите по запис в SAN среда се определя не само от производителността на отделните компоненти, а от взаимодействието между

натоварването, степента на паралелизъм, входно-изходните опашки и механизмите за координация по целия път на записите. Показано е, че при нарастващо натоварване системата преминава през различни режими на работа - устойчив режим, преходен режим и режим на координационен стрес, като преходите между тях се характеризират с нелинейно нарастване на крайната латентност.

Експерименталните резултати показват, че анализът на латентностните разпределения, емпиричните кумулативни функции на разпределение (ECDF) и високите персентири (p95 и p99) позволява откриване на опашкови ефекти и ранни признаци на деградация, които не могат да бъдат установени чрез използване единствено на средна латентност или пропускателна способност. На тази основа е разработен аналитичен модел за интерпретация на поведението на SAN системите, който свързва експерименталните наблюдения с принципите на теорията на масовото обслужване и обяснява влиянието на входно-изходните опашки върху латентностното поведение.

Съществен резултат на изследването е формулирането на оптимизационна интерпретация на получените резултати. Показано е, че ефективната оптимизация на операциите по запис не се свежда до максимизиране на пропускателната способност, а представлява процес на поддържане на устойчиво и предсказуемо латентностно поведение чрез управление на паралелизма, входно-изходните опашки и конкуренцията между заявките. В тази връзка е разработена аналитична рамка за оптимизация, която интегрира измерването на латентностното поведение, идентифицирането на режима на работа, избора на оптимизационни действия и последващата оценка на тяхната ефективност.

Направеният аналитичен синтез показва, че разработеният системен модел, експерименталната методология и предложената оптимизационна рамка формират единна методологична основа за анализ и оптимизация на операциите по запис в SAN среда. Получените резултати могат да подпомогнат проектирането, конфигурирането и експлоатацията на SAN системи, при които устойчивостта и предсказуемостта на латентността са критични изисквания.

Като основни направления за бъдещи изследвания могат да бъдат посочени разработването на адаптивни механизми за управление на натоварването и входно-изходните опашки, използването на ентропийно базирани показатели за ранно откриване на деградация, прилагането на методи от машинното обучение и изкуствения интелект (методи за машинно обучение, включително Random Forest и Case-Based Reasoning) за автоматична класификация на режимите на работа и препоръчване на оптимизационни

действия, както и разширяване на анализа към взаимодействието между SAN слоя и останалите компоненти на системната архитектура.

В обобщение, дисертационният труд показва, че устойчивото латентностно поведение в SAN среда може да бъде постигнато чрез системен подход, основан на анализ на латентностните разпределения, управление на паралелизма и контрол върху входно-изходните опашки. Разработеният модел, аналитична интерпретация и рамка за оптимизация превръщат експерименталните резултати в практически приложима методология за анализ и оптимизация на операциите по запис в SAN системи.

IV. ОСНОВНИ РЕЗУЛТАТИ ОТ ДИСЕРТАЦИЯТА

В резултат на проведеното изследване в дисертационния труд са получени следните основни резултати:

1. Разработен е системен модел на пътя на записите (write path) в SAN среда, който представя операциите по запис като последователност от взаимодействащи механизми за управление на заявки, входно-изходни опашки и координация между архитектурните слоеве.
2. Разработена е експериментална методология за контролирано изследване на поведението на операциите по запис при систематично изменение на параметрите на натоварването, степента на паралелизъм и конкурентния достъп.
3. Проведен е експериментален анализ, който показва нелинейната зависимост между натоварването, паралелизма и латентността и позволява разграничаване на устойчив режим, преходен режим и режим на координационен стрес.
4. Показано е, че анализът на емпиричните разпределителни функции (ECDF) и крайните персентили (p95 и p99) позволява ранно откриване на опашкови ефекти и деградация на производителността, които не могат да бъдат установени чрез използване само на средни показатели.
5. Разработена е аналитична интерпретация на експерименталните резултати, основана на принципите на теорията на масовото обслужване, която обяснява влиянието на входно-изходните опашки върху латентностното поведение.
6. Формулирани са оптимизационни принципи и е разработена рамка за оптимизация на операциите по запис в SAN среда, която свързва анализа на

латентността с избора на практически оптимизационни действия за поддържане на устойчив режим на работа.

V. ПРИНОСИ

1. **Разработено е системно представяне на операциите по запис в SAN среда, при което латентността се разглежда като резултат от взаимодействието между множество слоеве и опашкови механизми по пътя на записите.**
2. **Разработен е аналитичен модел за интерпретация на поведението на SAN системите при нарастващо натоварване, позволяващ разграничаване на режими на работа - контролирана паралелност, преходна зона и координационен стрес.**
3. **Разработена е експериментална методология за изследване на операциите по запис в SAN среда, основана на систематично изменение на параметрите на натоварването при контролирани експериментални условия.**
4. **Разработен е подход за анализ на латентностното поведение, базиран на емпирични разпределителни функции и крайните персентири (p95 и p99), който позволява идентифициране на опашкови ефекти и ранни признаци на деградация, недостъпни при анализ, основан единствено на усреднени показатели.**
5. **Установено е експериментално, че увеличаването на степента на паралелизъм не води еднозначно до подобряване на производителността, а след определен праг усилва вариационността на латентността и води до непропорционално нарастване на крайните персентири.**
6. **Установено е, че опашковите ефекти представляват надежден индикатор за възникване на координационен стрес и осигуряват по-висока чувствителност при оценката на поведението на SAN системите в сравнение с традиционните агрегатни показатели.**
7. **Разработена е рамка за оптимизация на операциите по запис в SAN среда, която интегрира разработения системен модел, аналитичната интерпретация и експерименталните резултати в последователен процес за наблюдение, анализ, идентифициране на режима на работа и избор на оптимизационни действия.**
8. **Формулирани са практически препоръки за избор на оптимизационни действия в зависимост от наблюдавания режим на работа, основани на управление на паралелизма, входно-изходните опашки и крайните латентностни персентири.**

VI. ПУБЛИКАЦИИ ПО ДИСЕРТАЦИЯТА

- [1] Torres G.. *Entropy-Aware AI-Powered Ransomware Detection Using IBM FlashCore Module 4 in SAN Environments*. International Conference Automatics and Informatics (ICAI), 2026, ISBN:979-8-3315-8611-9, DOI:10.1109/ICAI67591.2025.11325084, 100-106; 2025
- [2] Torres G.. *Optimization of Write Operations in SAN Environments Through Intelligent Compression, Deduplication and Thin Provisioning*. TELECOM 2025, 33, 33rd National Conference with International Participation (TELECOM), 2026, ISBN:978-1-6654-6564-9, ISSN:2837-5246, DOI:10.1109/TELECOM66943.2025.11304087, 1-4; 2025
- [3] Torres G.. *Storage Virtualization: Optimizing Data Center Efficiency and Performance*. 2/2024, 2024, ISSN:2535-0668, 65-66; 2024
- [4] Torres G., Ivanov Vl.. *A Survey on Storage Area Network*. Robotics, Automation and Mechatronics '23, Prof. Marin Drinov Publishing House of Bulgarian Academy of Science, 2023, ISSN:1314-4634, 22-26; 2023
- [5] Торрес Г., Иванов Вл.. *Приложения на Fiber Channel, iSCSI, Infiniband и FCOE в Storage Area Network*. Robotics, Automation and Mechatronics '23, RAM 2023, Prof. Marin Drinov Publishing House of Bulgarian Academy of Sciences, 2023, ISSN:1314-4634, 30-32; 2022
- [6] Торрес Г.. *Анализ на методи за оптимизация на записи в SAN среда*. Prof. Marin Drinov Publishing House of Bulgarian Academy of Sciences, 2022, ISSN:1314-4634, 69-72; 2022

ИЗПОЛЗВАНА ЛИТЕРАТУРА:

- L. Kleinrock, *Queueing Systems, Volume I: Theory*. New York, NY, USA: Wiley-Interscience, 1975.
- E. D. Lazowska, J. Zahorjan, G. S. Graham, and K. C. Sevcik, *Quantitative System Performance: Computer System Analysis Using Queueing Network Models*. Englewood Cliffs, NJ, USA: Prentice Hall, 1984.
- R. Jain, *The Art of Computer Systems Performance Analysis*. New York, NY, USA: Wiley, 1991.
- M. Harchol-Balter, *Performance Modeling and Design of Computer Systems: Queueing Theory in Action*. Cambridge, U.K.: Cambridge University Press, 2013.
- A. B. Downey, *Think Stats: Probability and Statistics for Programmers*, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2014.
- S. Asmussen, *Applied Probability and Queues*, 2nd ed. New York, NY, USA: Springer, 2003.
- A. S. Tanenbaum and H. Bos, *Modern Operating Systems*, 4th ed. Boston, MA, USA: Pearson, 2015.
- J. H. Saltzer and M. F. Kaashoek, *Principles of Computer System Design*. Burlington, MA, USA: Morgan Kaufmann, 2009.
- C. Min et al., "Understanding the characteristics of flash storage," in *Proc. IEEE MSST*, 2014.
- J. Dean and L. A. Barroso, "The tail at scale," *Communications of the ACM*, vol. 56, no. 2, pp. 74–80, 2013.
- A. Fox and E. Brewer, "Harvest, yield, and scalable tolerant systems," in *Proc. HotOS*, 1999.
- S. Kanev et al., "Profiling a warehouse-scale computer," in *Proc. ISCA*, 2015.
- A. Klimovic et al., "Understanding ephemeral storage performance in the cloud," in *Proc. USENIX ATC*, 2018.
- J. Leverich and C. Kozyrakis, "Reconciling high server utilization and sub-millisecond QoS," in *Proc. EuroSys*, 2014.
- R. Lu et al., "NVMe SSD failures in the field: The fail-stop and the fail-slow," in *Proc. USENIX ATC*, 2022.
- I. Jo, "Toward ultra-low latency SSDs: Analyzing the impact on data-intensive workloads," *Electronics*, vol. 12, no. 1, 2023.
- Y. Zhang et al., "Characterizing tail latency in cloud storage systems," in *Proc. IEEE IC2E*, 2022.
- Z. Huang et al., "FairNVMe: State-driven fairness control for efficient I/O queue scheduling," *Future Generation Computer Systems*, vol. 155, 2025.

- A. Yang et al., “Set-associative mapping to improve data throughput and reduce long tail latency in SSDs,” *Journal of Systems Architecture*, vol. 155, 2025.
- S. Li et al., “Understanding performance variability in modern NVMe SSDs,” in *Proc. USENIX FAST*, 2023.
- M. Bjørling et al., “The Linux block layer and NVMe: Performance implications for modern storage,” *ACM Transactions on Storage*, vol. 19, no. 2, 2023.