

УДК 004.728.4

В. Ф. ЗАЙКА, доктор техн. наук, доцент;

М. П. ТРЕМБОВЕЦЬКИЙ, доктор техн. наук, ст. наук. співробітник;

В. В. ЖЕБКА, канд. техн. наук,

Державний університет телекомунікацій, Київ

Реалізація вузла мережної зовнішньої пам'яті з передаванням файлів і блоків

Для розвитку функціональних можливостей систем і мереж зовнішнього зберігання і обробки даних доцільно використовувати ідеї віртуалізації, не обмежуючись виключно віртуалізацією пам'яті, а розширюючи це поняття і на обчислювальні потужності сховищ даних. Запропоновано як основний метод віртуалізації даних систем і мереж у розширеному сенсі використовувати принципи проектування на основі імітаційного поведінкового моделювання, «керованого» знаннями і правилами.

Щоб знизити рівень складності структурних вирішень, запропоновано як точки віртуалізації вибирати вузли мережі зберігання, які здатні поєднувати функції віртуалізації з функціями обробки структурованих даних.

Ключові слова: зовнішнє зберігання даних; багатофункціональна мережа; віртуальна машина; обробка даних; агентно-орієнтована технологія.

Вступ

Наприкінці ХХ століття, у деяких ранніх роботах як вітчизняних, так і зарубіжних учених [1–3] було побудовано низку структур мережних вузлів зовнішньої пам'яті на основі багатофункціональних спеціалізованих процесорів зовнішніх запам'ятовувальних пристроїв (ЗЗП). Такі публікації, по суті, сприяли появі систем зберігання даних, які за сучасною зарубіжною класифікацією належать до мереж зберігання, або мережним програмам типу NAS.

Основна частина

У мережному вузлі зовнішньої пам'яті, спрощену структуру якого зображено на рис. 1, використовуються спеціальний процесор ЗЗП, або каналний процесор із розширеними функціями, пристрій управління і нагромаджувач на магнітних дисках.

Метою побудови моделі є аналіз черги *quel* на ініціалізацію каналних програм з боку зв'язувального процесора, а також черги *main* сегментів даних в оперативній пам'яті, які очікують на закінчення операції обміну. Проста імітаційна модель (*модель 1*) зв'язувального процесора відбиває лише функцію буферизації вхідних повідомлень, які надходять на його вхід з лінії зв'язку та видаються на лінії зв'язку за запитами. Канальні програми (повідомлення в моделі) очікують на початок обробки каналом *chan* у черзі *quel*. Вхідні повідомлення змінної довжини з лінії зв'язку, керовані зв'язувальним процесором, надходять до оперативної пам'яті. Після закінчення запису зв'язувальний процесор видає запит на ініціалізацію каналної програми, що реалізує запис повідомлення на доріжку диска. Отримуючи запит на видачу повідомлення, в оперативній пам'яті резервується певна область для запису в неї повідомлення, а зв'язувальний процесор видає запит на ініціалізацію каналної програми, під управлінням якої відбудеться зчитування повідомлення з доріжки диска.

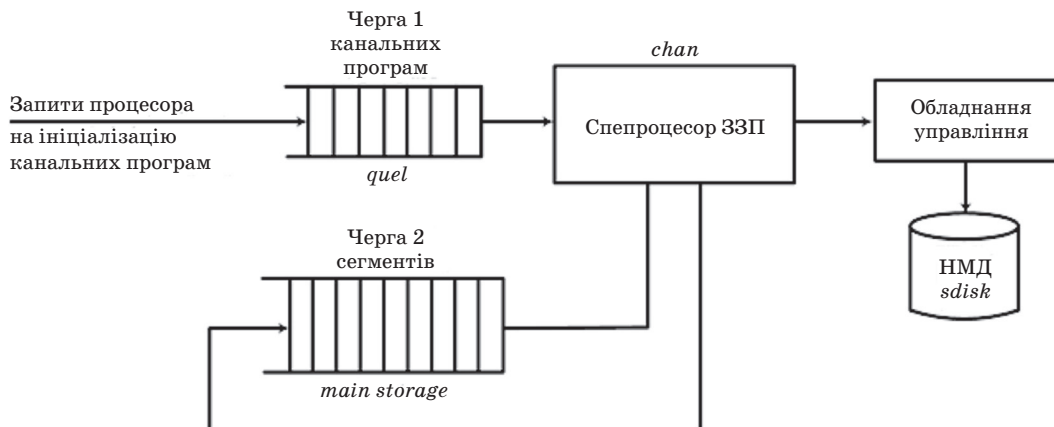


Рис. 1. Спрощена структура мережного вузла зовнішньої пам'яті

Характеристики системи, поданої на рис. 1, можна значно поліпшити, якщо розширити можливості спецпроцесора ЗЗП. *Модель 2* розроблено для системи, структуру якої ілюструє рис. 2. На відміну від моделі 1 тут береться до уваги наявність засобу *rps* (прискореного визначення положення записів при обертанні). Як і у попередній моделі, ініціалізовані зв'язувальним процесором каналні програми (повідомлення в моделі) очікують на початок обробки каналом *chan* у черзі *quel* (черга 1 на рис. 2). На основі експериментів з моделями встановлено, що продуктивність системи, зображеної на рис. 2, у 3–5 разів вища (залежно від вхідних навантажень), ніж продуктивність системи, спрощену структуру якої унаочнює рис. 1.

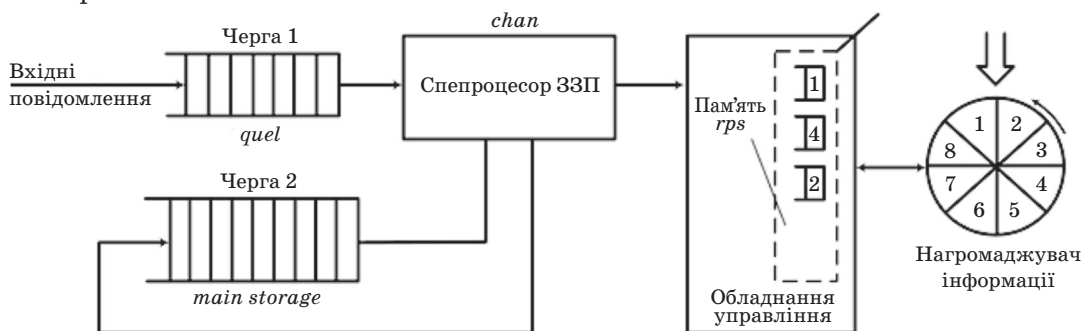


Рис. 2. Мережний вузол зовнішньої пам'яті з прискореним пошуком записів

Проаналізуємо далі мережний вузол зовнішньої пам'яті з розвиненими функціональними можливостями. Раніше було розглянуто моделі виконання операцій буферизації повідомлень під управлінням периферійного або зв'язувального процесора в системі з непрямым зв'язком між зв'язувальним і центральним процесором. *Модель 3* було покладено в основу побудови алгоритму функціонування центру концентрації повідомлень обчислювальної мережі, що приймає повідомлення і видає їх за запитами центрального процесора. Структуру тривірневої системи пам'яті концентратора повідомлень у складі мережного вузла зовнішньої пам'яті з розвиненими функціональними можливостями наведено на рис. 3.

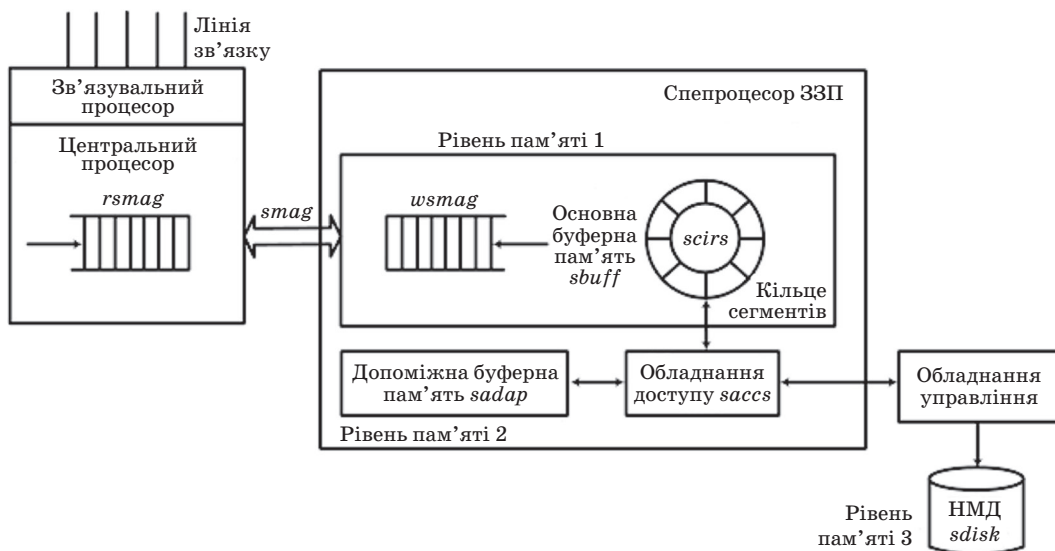


Рис. 3. Структура мережного вузла зовнішньої пам'яті з розвиненими функціональними можливостями

Повідомлення довільної довжини при занесенні у концентратор розбиваються на безліч незалежних сегментів (сторінок) однакової довжини, причому для кожного сегмента з огляду на інформацію, що міститься в ключі, визначається своя фізична адреса хешуванням ключа. Пам'ять концентратора поділено за рівнями так. *Перший рівень* — це основна буферна пам'ять *sbuff* (служить для попереднього розміщення вхідних і вихідних повідомлень); *другий рівень пам'яті sadap* має обсяг, достатній для розміщення одного або кількох сегментів (відіграє роль допоміжного буфера між основною буферною пам'яттю і ЗЗП на дисках); *третій рівень пам'яті* є основним і формується на базі ЗЗП *sdisk* із прямим доступом.

Уведення другого рівня пам'яті між першим і третім дозволяє спростити організацію структури пам'яті першого рівня, а також узгодити швидкість передавання даних із ЗЗП на дисках через канал із розподілом циклів основної пам'яті.

Аналіз функціонування зазначеної мережної підсистеми виконується за допомогою імітаційної системи GPSS. Імітаційна модель, що записана в термінах блоків GPSS, передбачає виконання таких дій:

- 1) створення чергового повідомлення;
- 2) генерування значення довжини повідомлення в сегментах;
- 3) генерування номера необхідного сектора у нагромаджувачі на дисках;
- 4) встановлення необхідного режиму запису або читання для повідомлення;
- 5) перевірка заданого у повідомленні режиму (у режимі запису здійснюється перехід до п. 6, у режимі читання — до п. 8);
- 6) заняття магістралі *smag* і передавання повідомлення сегмент за сегментом до основної буферної пам'яті;
- 7) формування виконавчих адрес для кожного сегмента;
- 8) поділ повідомлення на сегменти і спрямування сегментів у програмне кільце *scire* покажчиків звернень до пам'яті третього рівня;
- 9) виконання дій, запропонованих блоками GPSS-моделі щодо занесення або зчитування сегментів;
- 10) формування вихідного повідомлення з сегментів;
- 11) перевірка заданого у повідомленні режиму (у режимі читання здійснюється перехід до п. 12, у режимі запису — до п. 13);
- 12) заняття магістралі *smag* і передавання повідомлення послідовно сегмент за сегментом з основної буферної пам'яті джерела запиту;
- 13) збір статистичних даних про час відповіді; при цьому для подальшого графічного виведення формуються три гістограми: перша для запису повідомлень, друга — для читання, третя — загальна;
- 14) знищення повідомлення.

Повний текст моделі містить майже 900 блоків GPSS.

Внаслідок виконаних статистичних експериментів було отримано залежності для часу відповіді та встановлено вплив «синонімів» при багаторазовому оновленні даних у ЗЗП. Також було здобуто статистичні дані стосовно завантаження компонент трирівневої системи пам'яті і щодо черги. Деякі основні результати подано на рис. 4. При моделюванні взято, що на одному тракті нагромаджувача може бути розміщено 128 сегментів, пам'ять другого рівня має обсяг, достатній для розміщення одного сегмента. Передбачалося, що довжина повідомлення розподілена відповідно до функції:

$$\text{length function rnl, c4}$$

$$0,2 / .5,3 / .9,5 / 1,9,$$

де у першому рядку зазначено ім'я функції *length*, ключове слово *function*, ім'я потоку псевдовипадкових величин *rnl*, що використовуються при генерації значень довжини повідомлень і число координат *c4* функції при її кусково-лінійній апроксимації, а у другому рядку — відповідні координати функції.

Залежність середнього часу $t_{\text{оп}}$ обробки повідомлень, що вимагають запису (крива 1), читання (крива 2) і повідомлень без розрізнення за режимами (крива 3) від інтенсивності λ вхідного потоку зображено на рис. 4, а. Імовірність α появи «синоніму» (накладення) при читанні сегмента беруть такою, що дорівнює 0,3, а це відповідає заповненню третього рівня пам'яті на 30%.

Аналізуючи ці залежності, доходимо висновку, що поява «синонімів» призводить до суттєвих відмінностей у характеристиках часу відповіді для повідомлень зазначених типів. Залежність 2 показує, що при читанні повідомлень, розбитих на сегменти, повернення запитів у кільце покажчиків спричинює додаткову затримку, яка позначається на загальному часі складання повідомлення з сегментів. Під час виконання операцій запису розбитих на сегменти даних накладення сегментів вдається уникати завдяки відображенню стану пам'яті третього рівня у пам'яті першого рівня мережного вузла зовнішньої пам'яті і попередній перевірці зайнятості відповідної області пам'яті сегментом, тому час запису менший, ніж час виконання операції читання.

Залежності середнього часу запису (крива 1) і читання повідомлення (крива 2), а також загального часу відповіді незалежно від типів повідомлень (крива 3) від імовірності α «синоніму» при читанні сегмента для двох різних інтенсивностей λ подано на рис. 4, б. Поводження цих залежностей свідчить про те, що величина α суттєво впливає не тільки на збільшення часу обробки повідомлень на читання, а й побічно на збільшення часу обробки повідомлень на запис завдяки зросту внаслідок появи «синонімів» загальної кількості сегментів у трирівневій системі пам'яті.

Залежності кількості сегментів L у динамічному кільці від λ при $\alpha = 0,3$ (рис. 4, в) і від α при $\lambda = 400$ 1/с (рис. 4, г) унаочнює рис. 4. Тут кривими 4 позначено залежності для середньої кількості, а кривими 5 — залежності для максимальної кількості сегментів у кільці при ймовірності переповнення буфера не більш ніж 0,001.

Ці характеристики дають корисну інформацію стосовно довжини кільця сегментів (i , відповідно, про необхідний обсяг буферної пам'яті) та дозволяють оцінити обсяг апаратних засобів для його реалізації.

Як засіб імітаційного моделювання функціонування вузлів мережної пам'яті при побудові останніх трьох моделей було використано відому імітаційну систему GPSS [4]. Імітаційні моделі, записані на мові GPSS, відносно добре передають поведінковий аспект, що моделюється. Кожна подія у моделі при цьому моделюється окремим блоком. Широке поширення системи GPSS зумовлено не тільки наявністю для неї добре апробованих трансляторів, а й внутрішніми особливостями симулятора.

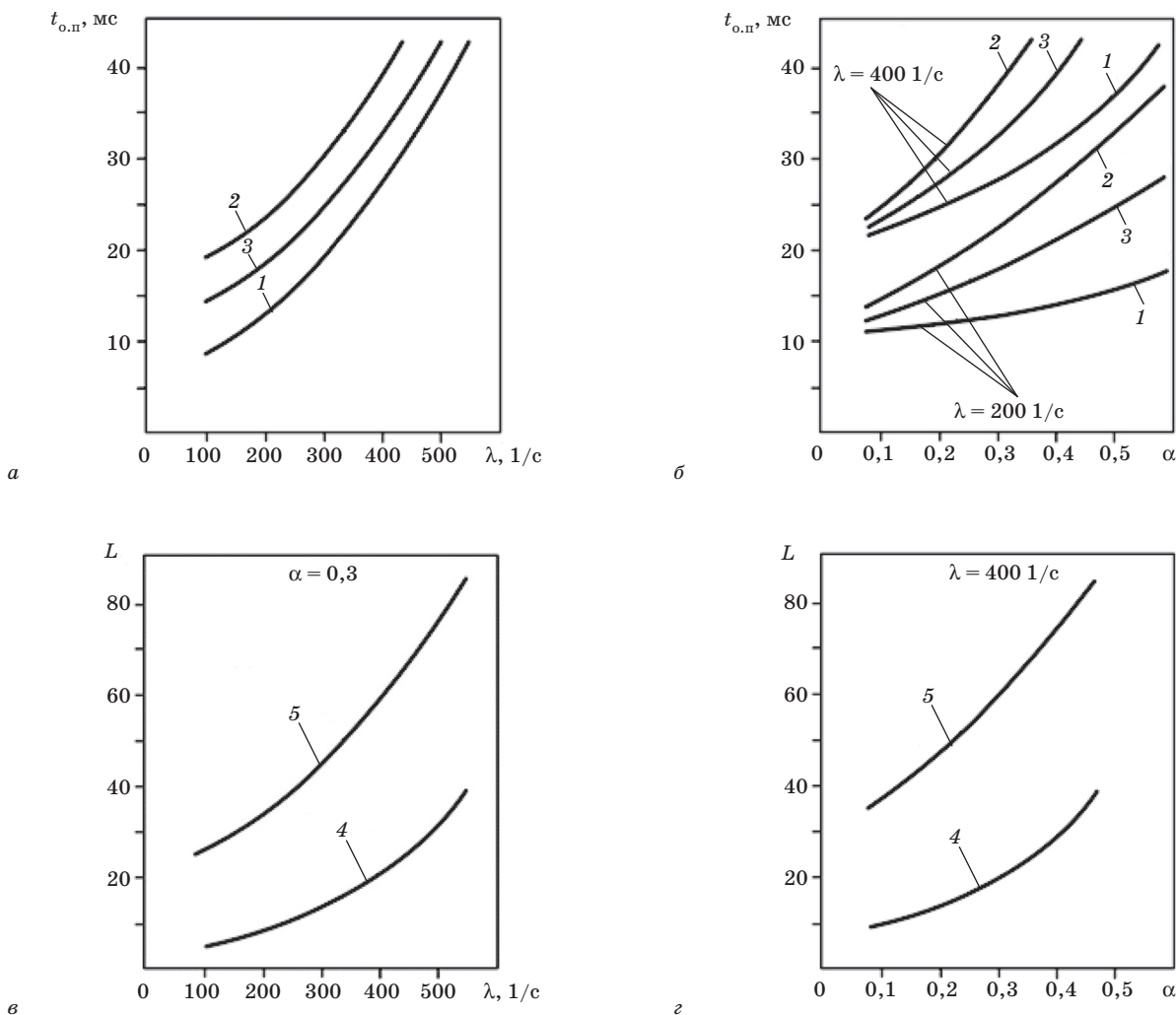


Рис. 4. Основні результати моделювання мережного вузла зовнішньої пам'яті з розвиненими функціональними можливостями (пояснення позначень — у тексті)

Мова GPSS також відносно добре передає структурний аспект проектованої системи і дозволяє архітекторам EOM приділяти більшу увагу логічній організації, що розробляється, а не складанню моделі. Складання моделі на GPSS полегшується завдяки тому, що у ній задається просторове переміщення повідомлень (іноді так званих транзактів) від одного блока до іншого. Остання версія системи GPSS має вільний формат, більшу кількість параметрів і блоків; користувачу тут надана можливість застосувати непряму адресацію об'єктів не тільки через параметри повідомлень, а й через будь-який стандартний числовий атрибут.

Імітаційна система GPSS внаслідок своєї неформалізованості не дозволяє безпосередньо використовувати текст моделі при проектуванні та апробації алгоритмів функціонування обчислювальних систем. Вона також не підтримує модульне ієрархічне проектування, що істотно ускладнює її використання для архітектурного і функціонально-структурного моделювання та проектування.

Висновок

Досвід функціонально-структурної організації, концептуального і архітектурного моделювання показує, що кошти моделювання мають підтримувати необхідний рівень абстрактності при описі поводження системи, що проектується залежно від критеріїв і цілей проектування. Формальний опис може використовуватися для розробки і деталізації специфікацій алгоритмів функціонування моделюваної системи. Оскільки на різних етапах моделювання систем та мереж зберігання і обробки даних потрібні специфікації, що значно різняться за рівнем деталізації і показовості, доцільно використовувати уніфіковане подання імітаційних моделей, наприклад, що засноване на логіко-алгебраїчних і сценарних моделях.

Список використаної літератури

1. **Вашкевич, Н. П.** Вычислительные сети коллективного пользования: учеб. пособие / Н. П. Вашкевич.— П.: Изд-во Пенз. политехн. ин-та.— 1979.— 72 с.
2. **Зинкин, С. А.** Управление страничным обменом в трехуровневой системе памяти / С. А. Зинкин, С. И. Каток // Автоматизация проектирования систем автоматического и автоматизированного управления: тр. Всесоюз. науч.-техн. совещ.— Челябинск: Изд-во Челяб. политехн. ин-та.— 1979.— С. 49–55.
3. **Раков, Б. М.** Оценка характеристик структур ЭВМ с использованием GPSS/360 и GPSS V / Б. М. Раков, В. Г. Мельник // Вопросы радиоэлектроники.— 1978.— Вып. 13.— С. 63–71 (Серия ЭВТ).
4. **Кудрявцев, Е. М.** GPSS World. Основы имитационного моделирования различных систем / Е. М. Кудрявцев.— М.: ДМК Пресс, 2004.— 320 с.
5. **Дубинин, В. Н.** Принципы построения и исследование сети внешних запоминающих устройств с моноканальной архитектурой / В. Н. Дубинин, С. П. Чеботарев // Моделирование, проектирование и производство систем ВЗУ ЭВМ: тр. Всесоюз. конф.— П., 1990.— С. 69–71.
6. **Вашкевич, Н. П.** Проектирование систем внешних ЗУ на магнитных дисках для супер-ЭВМ / Н. П. Вашкевич, П. Д. Морозов // Периферийные устройства ЭВМ: материалы семинара.— М.: МДНТП.— 1991.— С. 22–26.
7. **Дубинин, В. Н.** Формализованное описание и верификация коммуникационных структур сетей внешних ЗУ / В. Н. Дубинин, С. В. Демидов // Периферийные устройства ЭВМ: материалы семинара.— М.: МДНТП.— 1991.— С. 27–32.

Рецензент: доктор техн. наук, ст. наук. співробітник **М. М. Степанов**, Державний університет телекомунікацій, Київ.

V. F. Zaika, M. P. Trembovetskiy, V. V. Zhebka

РЕАЛИЗАЦИЯ УЗЛА СЕТЕВОЙ ВНЕШНЕЙ ПАМЯТИ С ПЕРЕДАЧЕЙ ФАЙЛОВ И БЛОКОВ

Для развития функциональных возможностей систем и сетей внешнего хранения и обработки данных целесообразно использовать идеи виртуализации, не ограничиваясь исключительно виртуализацией памяти, а расширяя это понятие и на вычислительные мощности хранилищ данных. Предложено в качестве основного метода виртуализации данных систем и сетей в расширенном смысле использовать принципы проектирования на основе имитационного поведенческого моделирования, «управляемого» знаниями и правилами.

Для снижения сложности структурных решений предлагается в качестве точек виртуализации выбирать узлы сети хранения, которые способны сочетать функции виртуализации с функциями обработки структурированных данных.

Ключевые слова: внешнее хранение данных; многофункциональная сеть; виртуальная машина; обработка данных; агентно-ориентированная технология.

V. F. Zaika, M. P. Trembovetskiy, V. V. Zhebka

IMPLEMENTATION OF NODE NETWORK EXTERNAL MEMORY WITH THE TRANSFER OF FILE AND BLOCK

For the development of the functionality of systems and networks of external data storage and processing is advisable to use the idea of virtualization, virtualization is not limited only memory, and extending this concept to the computing power of data warehousing. Serves as the primary method virtualization systems and data networks in the broader sense to use the design principles based on behavioral modeling simulation, «managed» knowledge and rules.

To reduce the complexity of the proposed structural solutions as the virtualization storage nodes points to choose the network that are able to combine with virtualization function structured data handling functions.

Keywords: network external storage; multifunctional system; virtual machine; data processing; agent-oriented technology.