

УДК 004.713

Б.Дунець

Національний університет “Львівська політехніка”

АРХІТЕКТУРА ПРИСТРОЮ ПЛАНУВАННЯ КОМУТАЦІЄЮ

В статті запропоновано архітектуру пристрою планування комутацією, в якій реалізовано швидкий паралельний метод планування комутацією. Ця архітектура дозволяє побудувати комутатори пакетів матричного типу з вхідною буферизацією, які при меншій часовій складності, ніж у відомих, забезпечують максимальну пропускну здатність та мінімальний середній час затримки комутації пакетів при опрацюванні пакетів на швидкості поступлення вхідних даних.

Вступ

Активний розвиток комунікаційних систем та мереж (ATM, BISDN, Ethernet) вимагає наявності високопродуктивних засобів передачі інформації. Одним з компонентів таких системи є комутатори [1]. Сучасні комутатори є складними пристроями, на які покладено задачі встановлення, підтримки та знищення зв'язку, а також керування потоком даних. Незважаючи на те, що в сучасних стандартах передачі інформації закладено підтримку швидкостей передачі даних в каналі OC-192(10Gbps) та навіть OC-768 (40Gbps), відомі комутатори підтримують швидкості передачі не більше 3,125Gpps в каналі [2]. Питання побудови високопродуктивного комутатора є важливою задачею, розв'язання якої вимагає розробки нових методів комутації та високопродуктивних архітектур комутаторів.

Основним елементом комутатора, від продуктивності якого залежить продуктивність комутатора в цілому, є комутуюче середовище (КМС) [2]. Оскільки пакети даних, що поступають в КМС не є зплановані, то виникає ситуація, при якій декілька пакетів вхідних даних одночасно мають бути скомутовані на один вихідний порт. Така ситуація називається вихідним змаганням (конфліктом). Подолання вихідного конфлікту та наявного в деяких типах комутаторів внутрішнього блокування є основною задачею, що розв'язується різними методами побудови комутаторів із застосуванням ефективних алгоритмів розв'язання конфліктів [3]. Від ефективності вибраного методу побудови комутаторів та алгоритму розв'язання конфліктів залежать характеристики КМС та комутатора загалом.

В статті розглядається архітектура КМС матричного типу з вхідною буферизацією, яка є перспективним напрямком його розвитку, оскільки забезпечує максимальну пропускну здатність при опрацюванні пакетів на швидкості поступлення вхідних даних [4]. Побудова комутаторів та КМС з вхідною буферизацією наштовхується на трудність подолання обмеження пропускну здатності при використанні вхідної буферизації типу FIFO. В роботі [5] показано, що при використанні вхідної буферизації типу FIFO пропускну здатність КМС є обмеженою на рівні 0,586, оскільки виникає блокування черги пакетом в заголовку (HOL Blocking). Усунення цього недоліку та забезпечення максимальної пропускну здатності КМС досягається застосуванням вхідної буферизації, відмінної від FIFO, та ефективних алгоритмів планування комутацією.

Відомі два методи побудови КМС з вхідною буферизацією, які дозволяють позбутися блокування вхідної черги пакетом в заголовку, – це віртуальна вихідна буферизація (VOQ) [3] на основі FIFO та буферизація, відмінна від FIFO [5]. В першому випадку використовується організація буферної пам'яті, при якій вхідні дані, що поступають по вхідному каналу, розміщуються у відповідному буфері FIFO, номер якого відповідає номеру вихідного каналу. Планування комутацією полягає у оптимальному встановленні зв'язку за допомогою комутуючої мережі між вибраними вхідними буферами та вихідними каналами. Наявна низка алгоритмів планування комутацією для КМС з віртуальною вихідною буферизацією, основними з яких є iSLIP, SLIP[6], LQF, LPF, iLQF, iLPF[7], MSM[8], MWM[9], MUCFA[10], LAURA[11]. Проте,

алгоритмічна складність кращих алгоритмів є доволі високою а саме, $O(N^2)$, що є перепоною на шляху створення високопродуктивного пристрою планування комутацією та комутатора загалом. В другому випадку вхідна буферизація організовується таким чином, що КМС може одночасно опрацьовувати W пакетів кожного вхідного буферу. Величину W називають вікном обслуговування запитів. Відомо [5], що при зростанні вікна обслуговування W зростає відповідно і пропускна здатність КМС. Питання планування комутацією в даному випадку широко не досліджувалось, а в роботі [12] запропоновано лише використовувати нейронні мережі для планування комутацією, що суттєво ускладнює пристрій планування комутацією.

В статті розглядаються комутатори з вхідною буферизацією, відмінною від FIFO, з аналізом та вибором пакетів на основі вікна обслуговування. Запропоновано архітектуру пристрою планування комутацією, в якій реалізовано швидкий паралельний метод такого планування.

1. Швидкий паралельний метод планування комутацією

З метою подолання блокування вхідної черги пакетом в заголовку автором запропоновано структуру КМС (рис.1) та ефективний для даної структури швидкий паралельний алгоритм планування комутацією – FPFA. При такій організації вхідна буферизація є відмінною від FIFO, а блок планування комутацією може одночасно опрацьовувати W пакетів кожного вхідного буферу. Розглядається КМС розміром $M \times N$, де M – кількість входів, N – виходів, в якому за кожен часовий проміжок t поступає тільки один пакет в кожен з вхідних портів.

Комутуюче середовище (рис.1) складається з вхідного буферу Q , що представляється масивом буферів пакетів $Q=[q_{i,k}]$, $i=1,2,\dots,N$, $k=1,2,\dots,W$, пристрою планування комутацією (ППК) та комутуючої мережі. При поступленні в момент часу t множини вхідних пакетів $A(t)$, дані пакетів зберігаються в масиві $Q(t)$, а відповідні ідентифікатори $H(t)$ поступають в ППК. Для моменту часу t (часового проміжку) ППК на основі отриманих вхідних ідентифікаторів визначає вектор комутації $S(t)=[s_j]$, $j=1,2,\dots,N$, який містить слово комутації s_j для кожного j -го вихідного порта. Слово комутації s_j задає номер пакету масиву Q , який комутується за допомогою комутуючої мережі на вихід j .

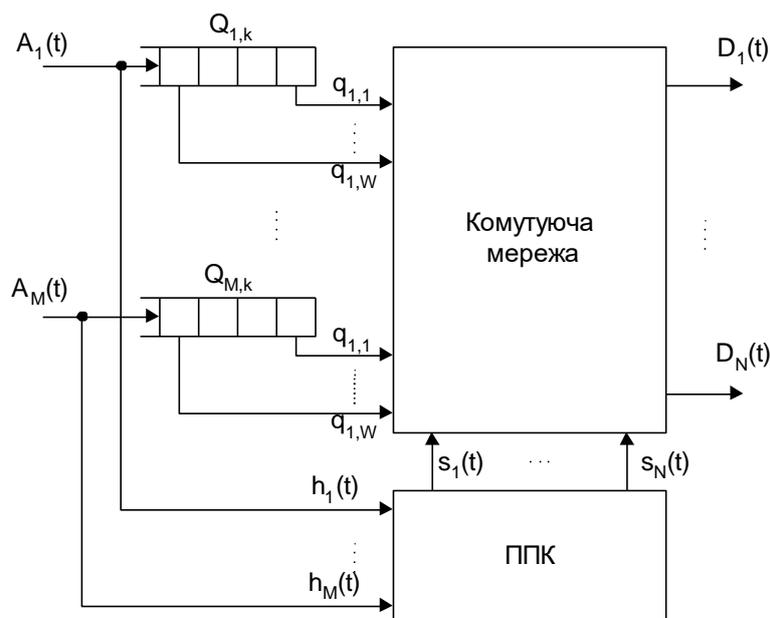


Рис.1 Структура КМС з вхідною буферизацією, відмінної від FIFO

Швидкий паралельний алгоритм планування комутацією в кожному часовому проміжку t виконує планування комутацією згідно з такою послідовністю дій:

Крок 1. Вертикальний пошук

Обчислення за допомогою функції вертикального вибору значень матриці вибору $g_{i,k}=V_SEL(V(t), j,k)$, для кожної пари j та k , де $j=1,2, \dots, N$ та $k=1,2, \dots, W$.

Крок 2. Горизонтальний пошук

Обчислення за допомогою функції горизонтального вибору значень вектора комутації $s_j=H_SEL(G(t), j)$, для кожного $j=1,2, \dots, N$.

Крок 3. Підтвердження та комутація

Для кожного вихідного порта $j=1,2, \dots, N$ виконується комутація, згідно з вектором комутації, та встановлення в 0 значення відповідних комірок матриці комутації, що відповідають зкомутованим пакетам на даному кроці:

$$D_j(t)=v_{i,k}, v_{i,k}=0, \text{ де } \{i,k\}=s_j$$

Функція вертикального вибору $V_SEL(V(t),j,k)$ виконує пошук елемента, значення якого дорівнює j серед елементів k -того стовпця матриці $V(t)$. Функція повертає номер рядка елемента, чи 0 при відсутності в рядку елемента, значення якого дорівнює j . За наявності декількох елементів в k -му стовпці матриці $V(t)$ функція повертає найбільший номер рядка.

Функція горизонтального вибору $H_SEL(G(t), j)$ виконує пошук серед елементів j -того рядка матриці $G(t)$ значення ненульового елемента. Якщо такий елемент $g_{i,k}$ знайдено, то функція повертає пару чисел $\{i,k\}$, де $i=g_{j,k}$ та k – номер стовпця знайденого елемента. При наявності декількох ненульових елементів в j -му рядку матриці $G(t)$, функцією вибирається елемент з найбільшим номером стовпця. За відсутності в j -му рядку ненульових елементів функція повертає $\{0,0\}$.

Асимптотична складність запропонованого алгоритму цьому дорівнює $O(\log_2 N)$ та в порівнянні з часовими складностями наведених в [3-11] кращих алгоритмів планування комутацією є найменшою, що є основою створення високопродуктивних пристроїв планування комутацією. Автором показано, що характеристики КМС залежать від розміру вікна обслуговування W , проте застосування гнучкої буферизації при переповненні ППК дозволяє забезпечити необхідні характеристики при фіксованих розмірах вікна обслуговування.

3. Архітектура пристрою планування комутацією

Основним елементом КМС є ППК, який виконує аналіз та планування вхідних запитів згідно з запропонованим швидким паралельним алгоритмом планування. Відповідно до алгоритму ППК складається з (рис.2):

- W блоків вертикального пошуку (БВП) (рис.3),

- блоку горизонтального пошуку (БГП) (рис.4) та блоку обчислення кількості втрачених пакетів (БОВП). БВП виконують паралельне обчислення елементів матриці вибору. Зокрема, БВП з номером k виконує обчислення елементів $g_{1,k}, g_{2,k}, \dots, g_{N,k}$ матриці вибору серед елементів k -того стовпця матриці ідентифікаторів $V(t)$ шляхом паралельного виконання N функцій вертикального вибору: $V_SEL(V(t),1,k), V_SEL(V(t),2,k), \dots, V_SEL(V(t),N,k)$. БВП містить пам'ять для збереження значень стовпця матриці ідентифікаторів.

БГП виконує обчислення значень вектора комутації s_1, s_2, \dots, s_N на основі матриці вибору $G(t)$, елементи якої обчислені блоками БВП. Обчислення виконується шляхом паралельного виконання N функцій горизонтального вибору: $H_SEL(G(t),1), H_SEL(G(t),2), \dots, H_SEL(G(t), N)$. Для забезпечення контролю комутації додатково в ППК можна використати БОВП, який виконує лік кількості втрачених пакетів.

ППК є синхронним пристроєм, робота якого синхронізується часовими проміжками t поступлення вхідних запитів, що дозволяє виконувати планування

комутацією на швидкості поступлення вхідних запитів. На початку часового проміжку t , множина вхідних ідентифікаторів переписується з БВП W в БВП $W-1$ аналогічно до реєстрова зсуву. В БВП з номером W записується множина вхідних ідентифікаторів $H(t)$. На основі накопиченої та отриманої інформації БВП виконують пошук елементів матриці вибору $G(t)$. Після обчислення матриці вибору БВП виконує пошук та формування вектора комутації $S(t)$. Інформація про вибрані пакети поступає в БВП, ідентифікатори ж вибраних елементів при поступленні наступного часового проміжку будуть вилучені.

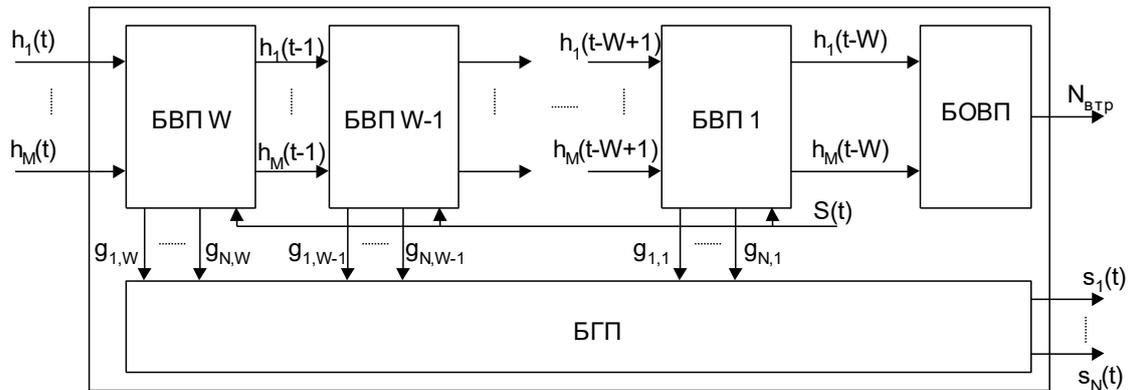


Рис.2. Структура ППК

БВП складається з M реєстрів $Рг1, Рг2, \dots, РгM$ ідентифікаторів, M блоків вилучення запиту (БВЗ) БВЗ1, БВЗ2, ..., БВЗМ та блоку реалізації функції вертикального вибору (БФВВ).

Ідентифікатори $h(t)$ в момент часу t поступають з попереднього БВП та записуються у реєстри $Рг1-РгM$. Таким чином, в реєстрах містяться номери вихідних портів, на які мають бути зкомутовані пакети, та 0, якщо пакет відсутній. БФВВ виконує пошук N кандидатів на комутацію на основі аналізу значень $РГ1-РГM$ згідно з алгоритмом планування комутацією. БВЗ призначені для вилучення ідентифікатора пакету, що підлягає комутації в момент часу t . Вилучення виконується шляхом скиду в нульове значення ідентифікатора, при наявності у векторі комутації номера ідентифікатора даного пакету.

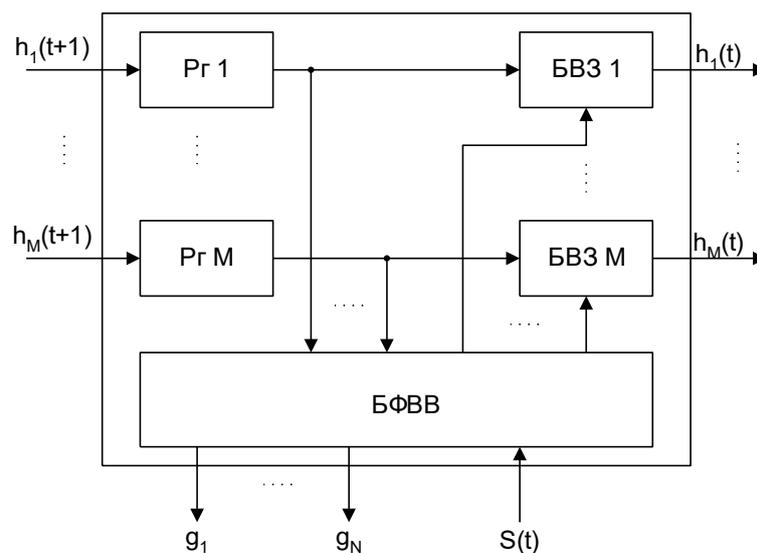


Рис.3. Структура БВП

Завдяки алгоритму планування комутацією пошук N елементів матриці вибору g_1, g_2, \dots, g_N виконується паралельно. БФВВ в такому випадку повинен містити N

незалежних блоків реалізації вертикального вибору (БФВВ) БФВВ1, БФВВ2,..., БФВВN, кожен з яких виконує пошук елемента, згідно з алгоритмом роботи функції вертикального вибору V_SEL швидкого паралельного алгоритму. Зокрема, БФВВ з номером j виконує пошук елемента серед вхідних ідентифікаторів, значення якого дорівнює j , та подає на свій вихід g_j номер рядка елемента, чи 0 у випадку відсутності елемента, значення якого дорівнює j .

БГП складається з N блоків реалізації функції горизонтального вибору (БФГВ) БФГВ1, БФГВ2,...,БФГВ N, кожен з яких виконує пошук пакету, що підлягає комутації у відповідності до алгоритму роботи функції горизонтального вибору H_SEL швидкого паралельного алгоритму планування комутацією.

Зокрема, БФГВ з номером j виконує пошук серед вхідних значень j -того рядка матриці вибору $G(t)=\{g_{j,w}, g_{j,w-1}, \dots, g_{j,1}\}$ значення ненульового елемента. Якщо такий елемент знайдено, то на вихід БФГВ поступає значення елемента вектора комутації $s_j=\{g_{j,k};j\}$, де $k=1,2,\dots,W$ – номер стовпця знайденого елемента. В протилежному випадку на вихід пристрою подається $s_j=\{0;0\}$.

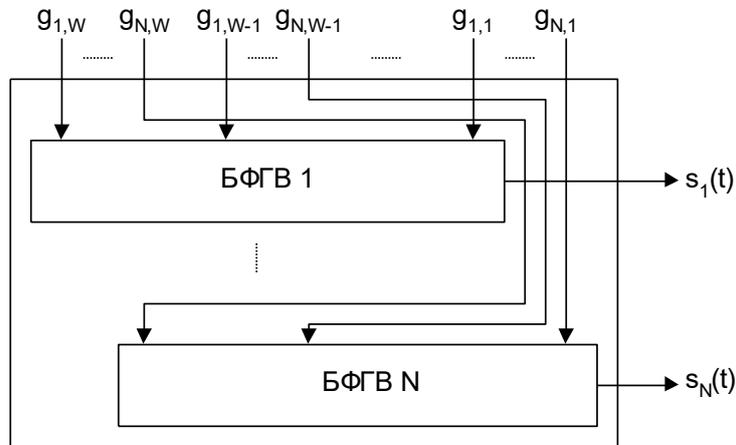


Рис.4. Структура БГП

Основним елементом БГП є БФГВ. Згідно з швидким паралельним алгоритмом планування комутацією БФГВ повинен виконати пошук після виконання процедури вертикального пошуку (Крок 1 алгоритму). Проте, можна здійснити пошук пакету, що підлягає комутації, паралельно до процедури вертикального пошуку.

3. Характеристики пристрою планування комутацією

З метою порівняння запропонованого методу планування комутацією, основною складовою якого є швидкий паралельний алгоритм такого планування, скористаємося загальноприйнятими критеріями оцінки КМС, а саме: часовою та асимптотичною часовою складнощами алгоритмів планування комутацією, залежностями пропускної здатності від розміру вікна обслуговування та середнього часу затримки від інтенсивності поступлення вхідних пакетів.

На рис. 5 наведено графіки залежності пропускної здатності (Throughput) КМС, в яких використано відомі алгоритми планування комутацією, що використовують вікно обслуговування (криві McCulloch-Pitts, Hopfield Memory, Simple Heuristic), та КМС з використанням запропонованого підходу в залежності від розміру вікна обслуговування W ($N=M=16$) (крива FPSA). Результати симуляції отримано для характеру поступлення вхідних пакетів, що описуються процесами Бернуллі з рівномірним розподілом вхідних запитів. Аналіз цих результатів показує, що запропонований алгоритм в порівнянні з іншими відомими алгоритмами планування комутації дозволяє досягти більшої пропускної здатності.

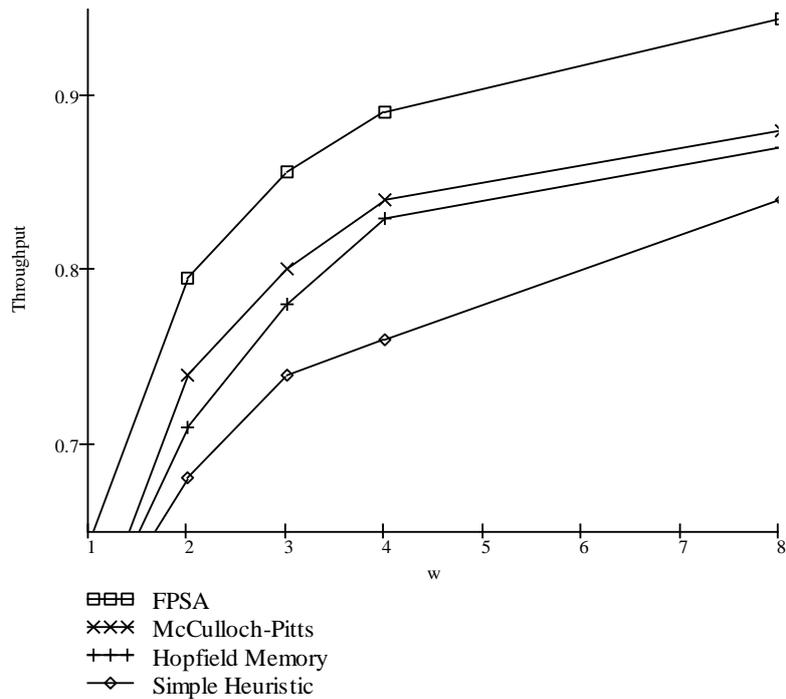


Рис.5 Порівняння пропускної здатності відомих та запропонованого КМС

Одним з важливих критеріїв порівняння КМС є середній час затримки пакетів при комутації. Дана характеристика дозволяє оцінити середню затримку передачі даних в інформаційній мережі, де використовується КМС. На рис.6 наведено графіки залежності середнього часу затримки (Mean delay time) від інтенсивності поступлення вхідних пакетів (Load). Як бачимо з результатів симуляції, середній час затримки КМС, в якому використано запропонований метод (крива FPSA), є ідентичний до середнього часу затримки КМС з вихідною буферизацією (крива Output Queue).

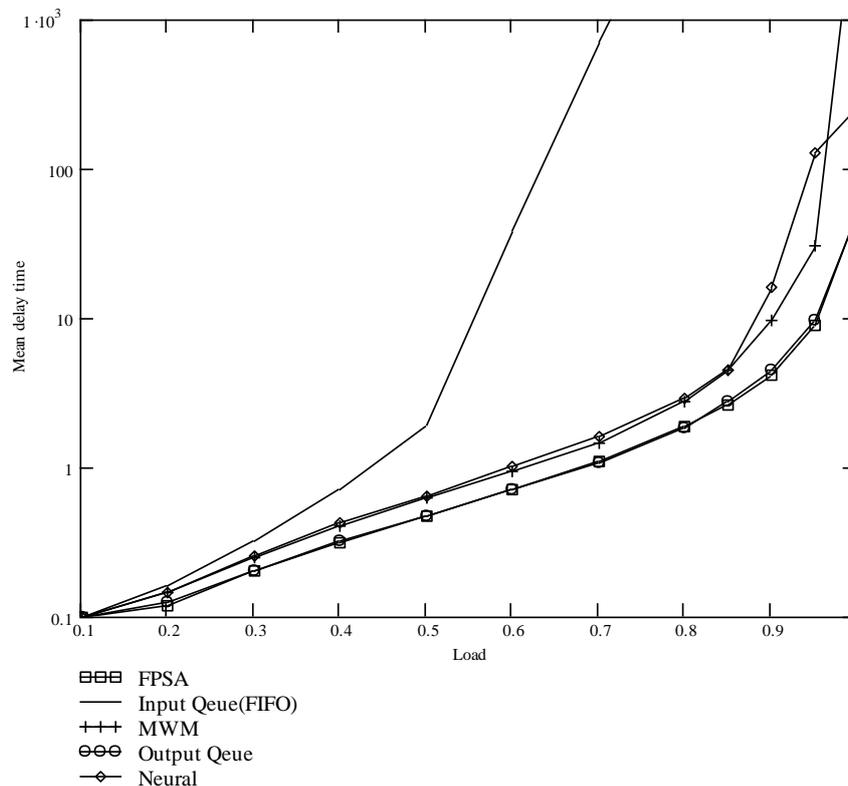


Рис.6 Порівняння характеристик КМС, в яких використані різні алгоритми планування комутацією

Висновки

1. Запропоновано нову архітектуру пристрою планування комутацією, що ефективно реалізує швидкий паралельний метод планування комутацією для комутуючих середовищ з вхідною буферизацією.
2. У порівнянні з відомими підходами до побудови КМС з плануванням комутацією на основі вікна обслуговування, пристрій планування комутацією забезпечує більшу пропускну здатність КМС.
3. Середній час затримки передачі пакетів КМС при використанні запропонованого пристрою планування комутацією є на рівні оптимального.
4. Пристрій планування комутацією працює на швидкості поступлення даних.

In the article the architecture of device for commutation planning, in which the rapid parallel planning method of commutation is realized, is proposed. This architecture allows to build the switchboards of packages of matrix type with the entrance spooling, which at less sentinel complication, than known switchboards have, provide a maximal carrying capacity and minimum mean time of delay of commutation of packages at working of packages on speed of entrance data .

Література

1. Кульги М. Технология корпоративных сетей: Энциклопедия. – СПб.: Издательство “Питер”, 2000. – 704с.
2. www.agere.com
3. Chao H. J., Lam Ch. H., Oki E. Broadband Packet Switching Technologies: A Practical Guide to ATM Switches and IP Routers. – New York.: John Wiley & Sons, 2001. – 459p.
4. Marsan M. A., Bianco A., Filippi E., Giaccone P., Leonardi E., Neri F. A Comparison of Input Queuing Cell Switch Architectures // IEEE BSS'99, 3rd International Workshop on Broadband Switching Systems, Kingston, Canada, June 1999.
5. Karol M.; Hluchyj M. Queueing in high-performance packet-switching // IEEE J. Selected Area Communications, Dec. 1988, vol.6, P.1587-1597.
6. McKeown N., Izzard M., Mekkittikul A., Ellersick B., Horowitz M. The Tiny Tera: A Packet Switch Core // IEEE Micro Jan/Feb.- 1997.- P. 26-33.
7. Mekkittikul A., McKeown N. A Practical Scheduling Algorithm to Achieve 100% Throughput in Input-Queued Switches // IEEE Infocom 98, Vol 2, April 1998, San Francisco.- P. 792-799.
8. Iyer S., McKeown N. Maximum Size Matching and Input Queued Switches // Proceedings of the 40th Annual Allerton Conference on Communication, Control and Computing. Vol 1, San Francisco, March 2002. - P. 296-302
9. Keslassy I., McKeown N. Analysis of Scheduling Algorithms That Provide 100% Throughput in Input-Queued Switches, Proceedings of the 39th Annual Allerton Conference on Communication, Control, and Computing. Monticello, Illinois, October 2001. - P. 375-379.
10. Prabhakar B., McKeown N. On the Speedup Required for Combined Input and Output Queued Switching // Computer Systems Technical Report CSL-TR-97-738. November 1997.
11. Shah D., Giaccone P., Prabhakar B. An Efficient Randomized Algorithm for Input-Queued Switch Scheduling // IEEE Hot Interconnects 9, Stanford, CA, USA, August 2001. - P.154-158.
12. W.M.Moh, Y.-F. Chung. Design and evaluation od cell scheduling algorithms for ATM switches // Networks: The Next Millenium, Proc. of IEEE Singapore Int. Conf. on Networks, World Scientific, 1997 - P. 355-369.

Одержано 27.09.2003 р.