



ДЕРЖАВНА СЛУЖБА
ІНТЕЛЕКТУАЛЬНОЇ
ВЛАСНОСТІ
УКРАЇНИ

УКРАЇНА

(19) **UA** (11) **97237** (13) **U**
(51) МПК (2015.01)
G06F 7/00

(12) ОПИС ДО ПАТЕНТУ НА КОРИСНУ МОДЕЛЬ

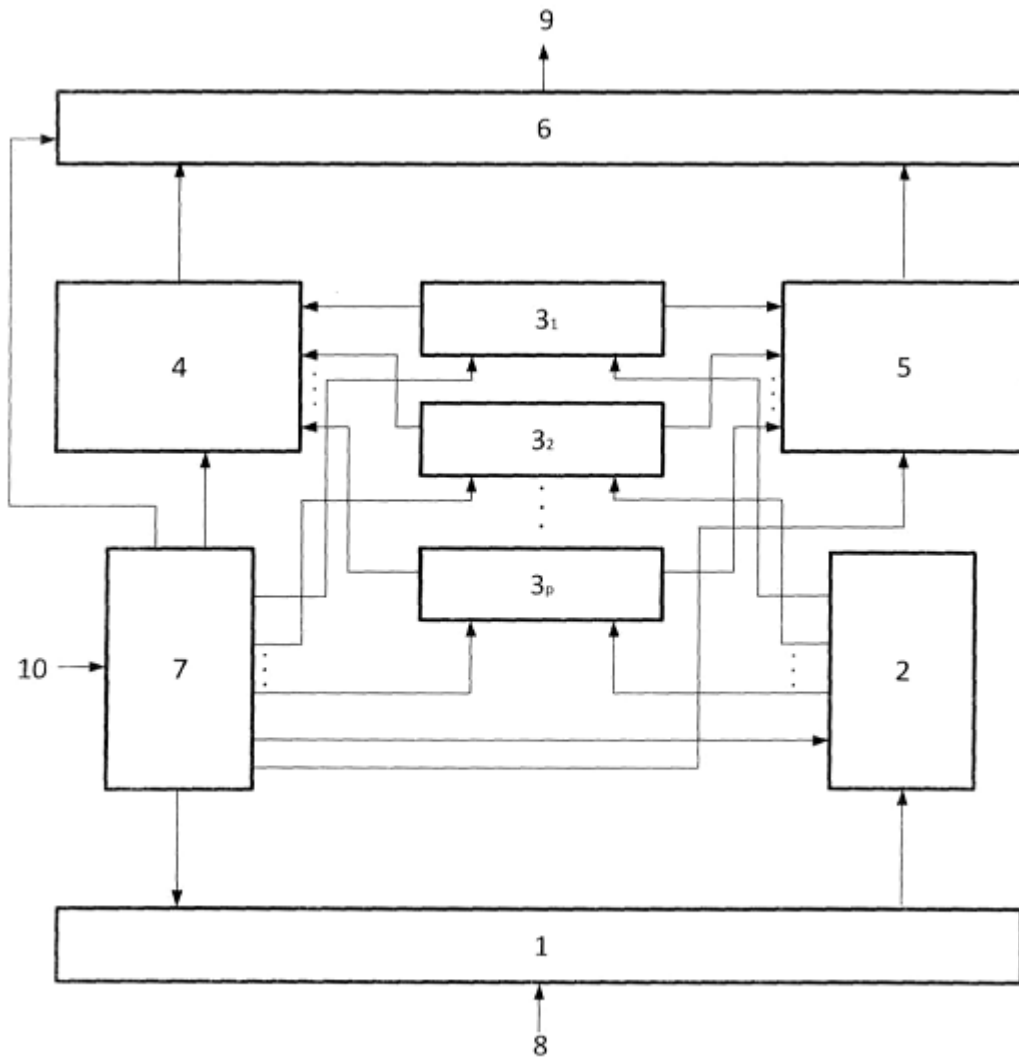
<p>(21) Номер заявки: u 2014 07879</p> <p>(22) Дата подання заявки: 14.07.2014</p> <p>(24) Дата, з якої є чинними права на корисну модель: 10.03.2015</p> <p>(46) Публікація відомостей про видачу патенту: 10.03.2015, Бюл.№ 5</p>	<p>(72) Винахідник(и): Семеренко Василь Петрович (UA), Матрос Дмитро Олександрович (UA)</p> <p>(73) Власник(и): ВІННИЦЬКИЙ НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ, Хмельницьке шосе, 95, м. Вінниця, 21021 (UA)</p>
---	---

(54) ПРИСТРІЙ ДЛЯ ПАРАЛЕЛЬНОГО ЗОВНІШНЬОГО СОРТУВАННЯ ДАНИХ

(57) Реферат:

Пристрій для паралельного зовнішнього сортування даних складається із р блоків часткового сортування, блока розподілу даних. Додатково в пристрій введені блок пам'яті вхідних даних, блок пошуку мінімального елемента, блок пошуку максимального елемента, блок пам'яті вихідних даних, блок керування.

UA 97237 U



Фиг. 1

Корисна модель належить до обчислювальної техніки і може застосовуватись в системах управління базами даних та інших комп'ютерних системах, що потребують сортування великих обсягів даних.

Відомий пристрій та спосіб для паралельного сортування злиттям [патент США US 2005/0144167 A1 МПК G06F7/00, Jun. 30, 2005], який містить блок для злиття символічних рядків, блок для поділу даних на два рядки, блок сортування, блок адміністрування системою, блок зберігання даних і блок керування, вихід якого зв'язаний з входом блока адміністрування системою, який за допомогою шини керування зв'язаний з блоком для злиття символічних рядків, блоком для поділу даних на два рядки і блоком сортування, причому блок зберігання даних за допомогою шини даних зв'язаний з блоком сортування, блоком для поділу даних на два рядки і блоком для злиття символічних рядків.

Недоліком цього пристрою є великі апаратні витрати і великі витрати часу, обумовлені можливістю лише попарного злиття відсортованих символічних рядків на останньому етапі сортування.

Найбільш близьким по технічній суті до запропонованого є пристрій для сортування [Патент США US 5727200, МПК G06F17/30, Mar. 10, 1998], що містить генератор початкової множини даних, блока розподілу даних, набір p блоків часткового сортування і блок порівняння та вибору входів, p входів якого з'єднані відповідно з виходами p блоків сортування, входи яких з'єднані відповідно з p виходами блока розподілу даних, вхід якого з'єднаний з виходом генератора початкової множини даних.

У відомому пристрої сортування даних складається з трьох етапів: етапу розподілу даних, етапу часткового сортування (сортування векторів) і етапу повного сортування (злиття даних).

У відомому пристрої суміщаються перші два етапи, за винятком сортування останнього вектора. Етап повного сортування розпочинається лише після повного завершення сортування першого вектора. Основним недоліком такого пристрою є необхідність зберігання відсортованих векторів даних, що вимагає додаткових витрат обладнання та додаткового часу.

В основу корисної моделі поставлена задача створення пристрою для паралельного зовнішнього сортування даних, в якому за рахунок введення нових блоків та зв'язків, паралельного виконання етапів часткового та повного сортування, а також відсутності в потребі зберігання відсортованих векторів даних, досягається можливість прискорення операції зовнішнього сортування та розширення функціональних можливостей пристрою.

Поставлена задача вирішується за рахунок того, що у пристрій для паралельного зовнішнього сортування даних, який складається із p блоків часткового сортування, блока розподілу даних, p виходів якого з'єднані з входами p блоків часткового сортування, введені блок пам'яті вхідних даних, блок пошуку мінімального елемента, блок пошуку максимального елемента, блок пам'яті вихідних даних, блок керування, виходи якого з'єднані з входами керування блока розподілу даних, p блоків початкового сортування, блока пам'яті вихідних даних і блока пам'яті вхідних даних, інформаційний вхід якого з'єднаний з інформаційним входом пристрою, а вихід з'єднаний з інформаційним входом блока розподілу даних, p виходів якого з'єднані, відповідно, з інформаційними входами p блоків сортування, виходи мінімуму і виходи максимуму яких з'єднані відповідно, з входами блока пошуку мінімального елемента і блока пошуку максимального елемента, виходи яких з'єднані відповідно з першим інформаційним входом і другим інформаційним входом блока пам'яті вихідних даних, інформаційний вихід якого під'єднаний до інформаційного виходу пристрою, вхід керування якого з'єднаний з входом початкового установа блоку керування.

На фіг. 1 представлена функціональна схема пристрою; на фіг. 2 - функціональна схема i -го блока часткового сортування; на фіг. 3 - функціональна схема блока пошуку мінімального елемента; на фіг. 4 - функціональна схема блока пошуку максимального елемента.

Пристрій для паралельного зовнішнього сортування даних (фіг. 1) містить блок 1 пам'яті вхідних даних, блок 2 розподілу даних, p блоків $3_1, \dots, 3_p$ часткового сортування, блок 4 пошуку мінімального елемента, блок 5 пошуку максимального елемента, блок 6 пам'яті вихідних даних і блок 7 керування, виходи якого з'єднані з входами керування блока 2 розподілу даних, p блоків $3_1, \dots, 3_p$ часткового сортування, блока 6 пам'яті вихідних даних і блока 1 пам'яті вхідних даних, інформаційний вхід якого з'єднаний з входом 8 пристрою, а вихід з'єднаний з інформаційним входом блока 2 розподілу даних, p виходів якого з'єднані, відповідно, з інформаційними входами p блоків $3_1, \dots, 3_p$ початкового сортування, виходи мінімуму і виходи максимуму яких з'єднані відповідно, з входами блока 4 пошуку мінімального елемента і входами блока 5 пошуку

максимального елемента, виходи яких з'єднані відповідно з першим інформаційним входом і другим інформаційним входом блока пам'яті вихідних даних, інформаційний вихід якого під'єднаний до інформаційного виходу 9 пристрою, вхід 10 керування якого з'єднаний з входом початкового установалення блока 7 керування.

5 Блок 3_i ($i = 1 \div p$) часткового сортування даних (фіг. 2) містить інформаційний вхід 11_i , вхід 12_i керування, вихід 13_i мінімуму і вихід 14_i максимуму.

Блок 4 пошуку мінімального елемента (фіг. 3) містить $(2^{\log_2 p} - 1)$ вузлів 15 вибору мінімального з двох елементів, які утворюють $(\log_2 p)$ -рівневе бінарне дерево, у якому вузол 15, що відповідає кінцевій вершині дерева, має входи 16_{j-1} і 16_j ($j = 2 \div p$), а вихід вузла 15, що
10 відповідає кореневій вершині дерева, з'єднаний з виходом 17 блока.

Блок 5 пошуку максимального елемента (фіг. 3) містить $(2^{\log_2 p} - 1)$ вузлів 18 визначення максимального з двох елементів, які утворюють $(\log_2 p)$ -рівневе бінарне дерево, у якому вузол 18, що відповідає кінцевій вершині дерева, має входи 19_{j-1} і 19_j ($j = 2 \div p$), а вихід вузла 18, що
15 відповідає кореневій вершині дерева, з'єднаний з виходом 20 блока.

15 Пристрій працює таким чином.

В блок 1 пам'яті вхідних даних по входу 8 надходить невідсортована послідовність S із n чисел.

У пристрої реалізовано спосіб для зовнішнього сортування даних, який складається з трьох етапів:

- 20
- етапу розподілу даних,
 - етапу часткового сортування (сортування векторів),
 - етапу повного сортування (злиття даних).

На етапі розподілу даних початкова послідовність S за допомогою блока 2 розподілу даних поділяється на p векторів із m елементів ($n = p \times m$). На наступному етапі p векторів
25 одночасно надходять на p блоків $3_1, \dots, 3_p$ часткового сортування. В блоці 3_i часткового сортування протягом $m/2$ ітерацій відбувається сортування вектора v_i методом простого вибору ($i = 1 \div p$). В кінці першої ітерації визначається мінімальний елемент e'_{\min} вектора v_i , який поміщається на початок вектора v_i і визначається максимальний елемент e'_{\max} вектора v_i , який поміщається в кінець вектора v_i . Мінімальні елементи векторів v_1, \dots, v_p утворюють
30 вектор мінімальних елементів v_{\min} , а максимальні елементи v_1, \dots, v_p утворюють вектор максимальних елементів v_{\max} .

На наступних ітераціях визначають чергові мінімальні і максимальні елементи кожного вектора серед невідсортованих елементів цих векторів. Після завершення етапу часткового сортування в початкових позиціях блоків $3_1, \dots, 3_p$ знаходяться мінімальні елементи, а в кінцевих
35 позиціях цих блоків - максимальні елементи відповідних векторів.

Етап повного сортування полягає у формуванні відсортованої послідовності S через злиття відсортованих векторів v_i . Для цього використовується метод вибору мінімальних і максимальних елементів за допомогою бінарних дерев.

Спочатку перший вектор v_{\min} мінімальних елементів з перших позицій блоків $3_1, \dots, 3_p$
40 надходить на p входів блока 4 пошуку мінімального елемента. Одночасно перший вектор v_{\max} максимальних елементів з останніх позицій блоків $3_1, \dots, 3_p$ надходить на p входів блока 5 пошуку максимального елемента.

Протягом однієї ітерації в блоці 4 пошуку мінімального елемента визначається мінімальний елемент першого вектора v_{\min} , тобто мінімальний елемент початкової послідовності S , і
45 записується за першою адресою блока 6 пам'яті вихідних даних. Протягом цієї ітерації в блоці 5 пошуку максимального елемента визначається максимальний елемент першого вектора v_{\max} , тобто максимальний елемент початкової послідовності S , і записується за останньою адресою блока 6 пам'яті вихідних даних. Таким чином, в блоці 6 пам'яті вихідних даних починає формуватися відсортована послідовність S_{sort} .

На j -й ітерації етапу повного сортування із j -го вектора v_{\min} мінімальних елементів за допомогою блока 4 пошуку мінімального елемента визначається мінімальний елемент, який записується за j -ю адресою блока 6 пам'яті вихідних даних, тобто переміщується в j -у позицію послідовності S_{sort} . Одночасно із j -го вектора v_{\max} максимальних елементів за допомогою блока 5 пошуку максимального елемента визначається максимальний елемент, який записується за $(n-j+1)$ -ю адресою блока 6 пам'яті вихідних даних, тобто переміщується в $(n-j+1)$ -у позицію послідовності S_{sort} $\left(j = 1 \div \frac{n}{2} \right)$.

Елементи, які поміщаються у відсортовану послідовність S_{sort} , вилучаються із блоків $3_1, \dots, 3_p$, а на їх місце надходять із відповідних векторів v_i наступні по зростанню та спаданню елементи.

Через $\frac{n}{2}$ ітерацій буде завершено етап повного сортування і в блоці 6 пам'яті вихідних даних буде сформована повністю відсортована послідовність S_{sort} .

Якщо тривалість часу ітерацій на етапах часткового і повного сортування буде однакою, тоді ці етапи можна частково сумістити і виконувати паралельно. Можливість суміщення в часі двох етапів сортування основана на тому, що після кожної ітерації етапу часткового сортування векторів підготовляються початкові дані для проведення однієї ітерації етапу повного сортування. Відразу після закінчення першої ітерації етапу часткового сортування будуть визначені перші вектори v_{\min} і v_{\max} , тому можна розпочинати етап повного сортування. Тоді загальна кількість ітерацій для сформування відсортованої послідовності S_{sort} становитиме $\left(\frac{n}{2} + 1 \right)$.

Блок 4 пошуку мінімального елемента працює таким чином.

На p входів $16_1, \dots, 16_p$ блока 4 надходить вектор v_{\min} мінімальних елементів. Для визначення найменшого серед елементів вектора v_{\min} використовується спосіб $(\log_2 p)$ -рівневого бінарного дерева. Бінарне дерево реалізоване за допомогою $(2^{\log_2 p} - 1)$ вузлів 15 визначення мінімального з двох елементів. На виході вузлів 15 першого рівня формується мінімальні елементи серед пар сусідніх вхідних елементів, які надходять з входів $16_1, \dots, 16_p$ блока. На виході вузлів 15 $(j+1)$ -го рівня формується мінімальні елементи серед пар сусідніх вхідних елементів, які отримані з вузлів 15 j -го рівня $(j = 1 \div \log_2 p)$. На виході 17 блока 4 буде сформований мінімальний елемент серед всіх елементів вектора v_{\min} .

Блок 5 пошуку максимального елемента працює таким чином.

На p входів $19_1, \dots, 19_p$ блока 5 надходить вектор v_{\max} максимальних елементів. Для визначення найбільшого серед елементів вектора v_{\max} використовується спосіб $(\log_2 p)$ -рівневого бінарного дерева. Бінарне дерево реалізоване за допомогою $(2^{\log_2 p} - 1)$ вузлів 18 визначення максимального з двох елементів. На виході вузлів 18 першого рівня формується максимальні елементи серед пар сусідніх вхідних елементів, які надходять з входів $19_1, \dots, 19_p$ блока. На виході вузлів 18 $(j+1)$ -го рівня формується максимальні елементи серед пар сусідніх вхідних елементів, які отримані з вузлів 18 j -го рівня $(j = 1 \div \log_2 p)$. На виході 20 блока 5 буде сформований максимальний елемент серед всіх елементів вектора v_{\max} .

Блоки 4 і 5 працюють одночасно.

Таким чином, у відомому пристрої відбувається суміщення у часі етапів часткового та повного сортування, за винятком першої ітерації.

Загальний час T_1 сортування n елементів даних у запропонованому пристрої складає:

$$T_1 = p t_{\text{роз}} + t_{\text{ітер}} + t_{\text{пов}}^1,$$

де $t_{\text{роз}}$ - тривалість розподілу n елементів даних по p векторах;

$t_{\text{ітер}}$ - тривалість однієї ітерації сортування;

$t_{\text{пов}}^1$ - тривалість етапу повного сортування.

Загальний час T_2 сортування n елементів даних у відомому пристрої складає:

$$T_2 = pt_{\text{роз}} + t_{\text{сорт}} + t_{\text{пов}}^2,$$

де $t_{\text{сорт}}$ - тривалість етапу сортування останнього вектора даних;

$t_{\text{пов}}^2$ - тривалість етапу повного сортування.

Тривалість однієї ітерації сортування менша за час сортування одного вектора даних:

$$5 \quad t_{\text{ітер}} < t_{\text{сорт}}.$$

Завдяки одночасному пошуку мінімальних та максимальних елементів вдвічі зменшується тривалість етапу повного сортування:

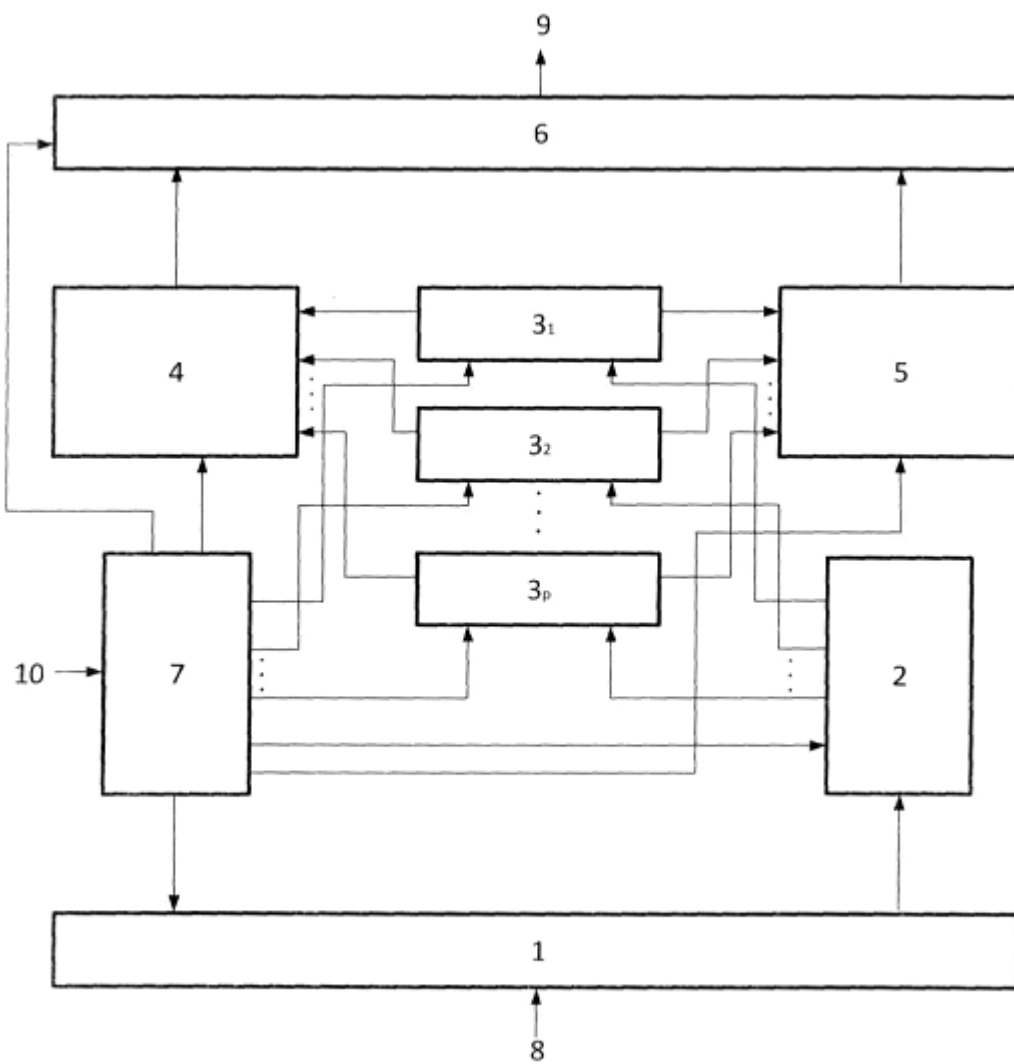
$$t_{\text{пов}}^1 = \frac{1}{2} t_{\text{пов}}^2.$$

10 Крім того, не потрібно зберігати відсортовані вектори даних, що дає додаткову економію часу та обладнання. Таким чином:

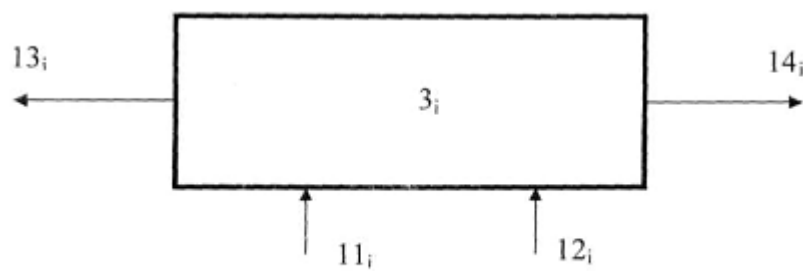
$$T_1 < T_2.$$

ФОРМУЛА КОРИСНОЇ МОДЕЛІ

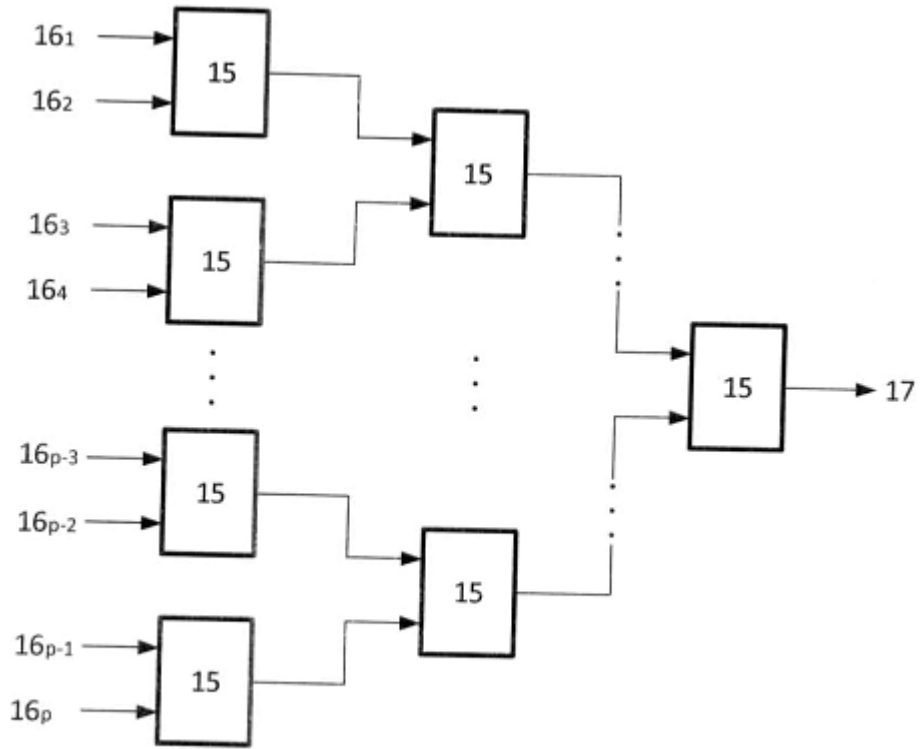
- 15 1. Пристрій для паралельного зовнішнього сортування даних, який складається із p блоків часткового сортування, блока розподілу даних, p виходів якого з'єднані з входами p блоків часткового сортування, який **відрізняється** тим, що в нього введені блок пам'яті вхідних даних, блок пошуку мінімального елемента, блок пошуку максимального елемента, блок пам'яті вихідних даних, блок керування, виходи якого з'єднані з входами керування блока розподілу
- 20 даних, p блоків початкового сортування, блока пам'яті вихідних даних і блока пам'яті вхідних даних, інформаційний вхід якого з'єднаний з інформаційним входом пристрою, а вихід з'єднаний з інформаційним входом блока розподілу даних, p виходів якого з'єднані, відповідно, з інформаційними входами p блоків сортування, виходи мінімуму і виходи максимуму яких з'єднані, відповідно, з входами блока пошуку мінімального елемента і блока пошуку
- 25 максимального елемента, виходи яких з'єднані, відповідно, з першим інформаційним входом і другим інформаційним входом блока пам'яті вихідних даних, інформаційний вихід якого під'єднаний до інформаційного виходу пристрою, вхід керування якого з'єднаний з входом початкового устанавлення блока керування.
2. Пристрій за п. 1, який **відрізняється** тим, що блок пошуку мінімального елемента містить
- 30 $(2^{\log_2 p} - 1)$ вузлів вибору мінімального з двох елементів, які утворюють $(\log_2 p)$ -рівневе бінарне дерево, у якому вузол, що відповідає кінцевій вершині дерева, має входи $(j-1)$ і j -й ($j = 2 \div p$), а вихід вузла, що відповідає кореневій вершині дерева, з'єднаний з виходом блока.
3. Пристрій за п. 1, який **відрізняється** тим, що блок пошуку максимального елемента містить
- 35 $(2^{\log_2 p} - 1)$ вузлів вибору максимального з двох елементів, які утворюють $(\log_2 p)$ -рівневе бінарне дерево, у якому вузол, що відповідає кінцевій вершині дерева, має входи $(j-1)$ і j -й ($j = 2 \div p$), а вихід вузла, що відповідає кореневій вершині дерева, з'єднаний з виходом блока.



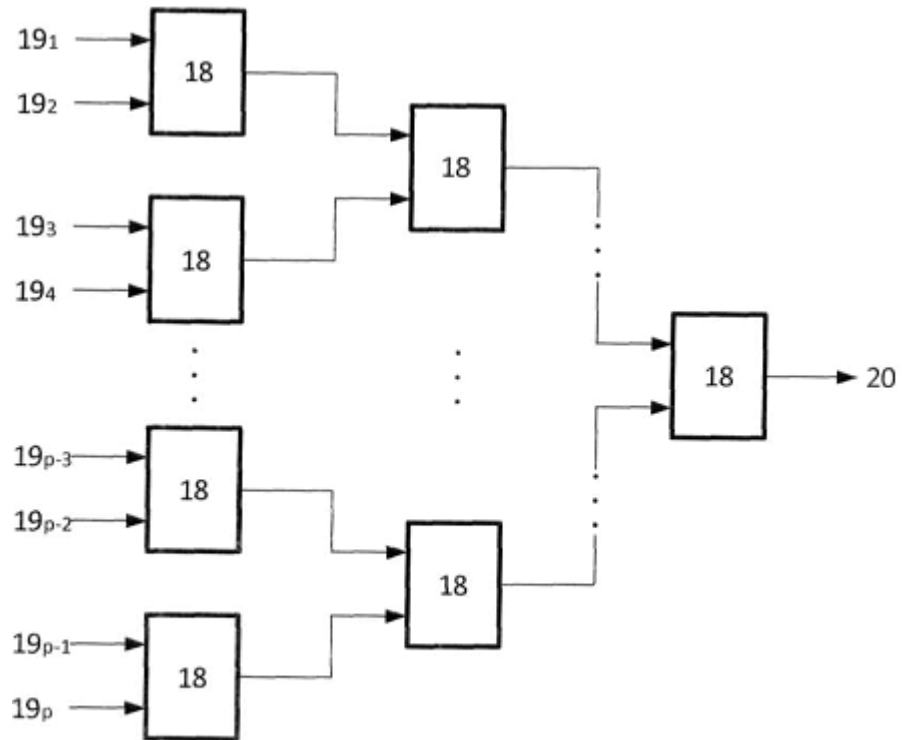
Фиг. 1



Фиг. 2



Фиг. 3



Фиг. 4

Комп'ютерна верстка А. Крулевський

Державна служба інтелектуальної власності України, вул. Урицького, 45, м. Київ, МСП, 03680, Україна

ДП "Український інститут промислової власності", вул. Глазунова, 1, м. Київ – 42, 01601