

DOI: <https://doi.org/10.36910/6775-2524-0560-2024-56-04>

УДК 004.2

Пасічник Максим Юрійович, аспірант

<https://orcid.org/0009-0006-7885-3507>

Зайцев Володимир Григорович, д.т.н., професор

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», м. Київ, Україна

МЕТОДИ ДИСПЕТЧЕРИЗАЦІЇ ЗАВДАНЬ У СИСТЕМАХ РЕАЛЬНОГО ЧАСУ

Пасічник М.Ю., Зайцев В.Г. Методи диспетчеризації завдань в системах реального часу. Системи реального часу є невід'ємною частиною сучасних технологій, охоплюючи сфери автоматизації виробництва, медичні пристрої, транспортні системи та інші критично важливі області. Ці системи вимагають забезпечення своєчасного та надійного виконання завдань, що робить питання ефективної диспетчеризації особливо актуальним. У даній статті представлено огляд і аналіз сучасних методів диспетчеризації завдань у системах реального часу, зокрема статичних, динамічних і комбінованих підходів. У статті підкреслено важливість забезпечення балансу між передбачуваністю та гнучкістю методів диспетчеризації для підвищення ефективності та надійності систем. Особливу увагу приділено ключовим викликам, таким як збалансування навантаження між процесорами, забезпечення надійності та відмовостійкості, ефективне використання ресурсів, динамічна адаптація до змінних умов і врахування енергоспоживання. Розробка нових методів і вдосконалення існуючих підходів має враховувати ці виклики для досягнення оптимальної продуктивності. Особливу увагу приділено перспективам застосування машинного навчання для прогнозування навантаження та динамічної оптимізації алгоритмів диспетчеризації. Використання таких технологій дозволяє значно підвищити ефективність і надійність систем реального часу, забезпечуючи адаптивне управління завданнями в реальному часі. Математичне моделювання, проведене у статті, дозволяє оцінити ефективність різних методів диспетчеризації за ключовими критеріями, такими як час відгуку, пропускна здатність, здатність до масштабування, використання ресурсів, надійність і гнучкість. Результати показують, що жоден метод не є універсально найкращим, і вибір оптимального підходу залежить від специфічних вимог конкретної системи. У статті окреслено напрями для подальших досліджень, включаючи інтеграцію методів машинного навчання для підвищення адаптивності та продуктивності систем реального часу. Це є важливим кроком у подоланні викликів, пов'язаних з оптимізацією використання ресурсів і дотриманням дедлайнів у змінних умовах, що сприятиме подальшому прогресу в цій галузі.

Ключові слова: системи реального часу, диспетчеризація завдань, балансування навантаження, динамічна адаптація, пропускна здатність, час відгуку, здатність до масштабування, машинне навчання

Pasichnyk M., Zaitsev V. Task scheduling methods in real-time systems. Real-time systems are an integral part of modern technologies, covering the areas of production automation, medical devices, transport systems and other critical areas. These systems require ensuring timely and reliable execution of tasks, which makes the issue of effective dispatching particularly relevant. This article presents an overview and analysis of modern methods of dispatching tasks in real-time systems, including static, dynamic and combined approaches. The paper highlights the importance of balancing predictability and flexibility in dispatching methods to improve system efficiency and reliability. Special attention is paid to key challenges such as load balancing between processors, ensuring reliability and fault tolerance, efficient use of resources, dynamic adaptation to changing conditions and consideration of power consumption. The development of new methods and the improvement of existing approaches must take these challenges into account to achieve optimal performance. Special attention is paid to the prospects of using machine learning for load forecasting and dynamic optimization of dispatching algorithms. The use of such technologies allows to significantly increase the efficiency and reliability of real-time systems, providing adaptive management of tasks in real time. The mathematical modelling carried out in the article allows to evaluate the effectiveness of different dispatching methods according to key criteria such as response time, throughput, scalability, resource utilization, reliability and flexibility. The results show that no method is universally best, and the choice of the optimal approach depends on the specific requirements of a particular system. The paper outlines directions for further research, including the integration of machine learning techniques to improve the adaptability and performance of real-time systems. This is an important step in overcoming the challenges of optimizing the use of resources and meeting deadlines in a changing environment, which will contribute to further progress in this field.

Keywords: real-time systems, task scheduling, load balancing, dynamic adaptation, throughput, response time, scalability, machine learning

Постановка проблеми. Системи реального часу відіграють ключову роль у сучасних технологіях, знаходячи застосування у багатьох критично важливих галузях, таких як автоматизація виробництва, медичні пристрої, транспортні системи та інші. Ці системи потребують забезпечення своєчасного та надійного виконання завдань, оскільки навіть незначні затримки можуть призвести до серйозних наслідків. Наприклад, у медичних системах пропущений дедлайн може вплинути на життя пацієнта, а в системах управління транспортом – спричинити аварії.

Однією з основних проблем у системах реального часу є диспетчеризація завдань, яка визначає порядок їх виконання, розподіл ресурсів та забезпечення дотримання часових обмежень. Ефективні методи диспетчеризації повинні враховувати різні фактори, такі як пріоритети завдань, змінні навантаження та обмежені ресурси. Існуючі методи диспетчеризації, хоча й досягли значного

прогресу, все ще стикаються з численними викликами, зокрема зі збалансуванням навантаження, забезпеченням надійності та адаптивності систем до змінних умов.

Розробка нових і вдосконалення існуючих методів диспетчеризації є важливим науковим і практичним завданням. Це дозволить підвищити продуктивність і надійність систем реального часу, забезпечивши їхню здатність до роботи в різних умовах. Успішне вирішення цієї проблеми сприятиме покращенню якості та безпеки таких систем, що є критично важливим для багатьох сфер людської діяльності.

Аналіз останніх досліджень і публікацій. У роботі «Scheduling in Real-Time System» [1] автори досліджують різні алгоритми планування, що застосовуються в системах реального часу. Вони підкреслюють важливість своєчасного виконання завдань, критичного для систем, що вимагають високої надійності. У статті обговорюються планування з фіксованим пріоритетом (RMS) і планування з найпершим дедлайном (EDF), ілюструючи їх застосування та ефективність у різних сценаріях. Автори наголошують на важливості оптимальних стратегій планування для забезпечення стабільності та продуктивності системи в умовах жорстких часових обмежень.

У статті «Task Scheduling in Real-Time Systems with Energy Harvesting and Energy Minimization» [2] представлений інноваційний підхід, що інтегрує енергозбирання з плануванням завдань. Автори пропонують метод динамічного масштабування напруги (DVS) у поєднанні з плануванням EDF для підвищення енергоефективності. Їх алгоритм пріоритетизує завдання на основі доступної енергії та дедлайнів, значно знижуючи споживання енергії при збереженні продуктивності системи. Експериментальні результати демонструють ефективність алгоритму в балансуванні використання енергії та планування завдань, що робить його придатним для систем реального часу з обмеженою енергією.

У статті «A New FPGA-Based Task Scheduler for Real-Time Systems» [3] представлений планувальник завдань на базі FPGA, розроблений для покращення продуктивності систем реального часу. Дослідження детально описує архітектуру та впровадження планувальника FPGA, підкреслюючи його швидкість, гнучкість та надійність. Автори порівнюють продуктивність планувальника FPGA з традиційними програмними планувальниками, показуючи значні покращення в управлінні завданнями та часі виконання. У висновку зазначається, що планування на базі FPGA може суттєво покращити ефективність систем реального часу.

У статті «Aperiodic Task Scheduling for Hard-Real-Time Systems» [4] автори вирішують проблему інтеграції неперіодичних завдань у системи з жорсткими реальними часовими обмеженнями. Вони пропонують метод планування, що забезпечує своєчасне виконання як періодичних, так і неперіодичних завдань без шкоди для продуктивності системи. Підхід включає гібридний алгоритм планування, який динамічно коригує пріоритети на основі дедлайнів завдань та навантаження системи. Результати моделювання підтверджують здатність алгоритму задовольняти суворі часові вимоги, доводячи його придатність у різних реальних часових середовищах.

Постановка завдання. Реалізація ефективної диспетчеризації завдань в системах реального часу стикається з низкою викликів, які впливають на загальну продуктивність та надійність систем. Основні виклики включають:

- Збалансування навантаження: Забезпечення рівномірного розподілу навантаження між процесорами є критично важливим для уникнення перевантаження окремих компонентів системи. Нерівномірний розподіл може призвести до затримок у виконанні завдань та зниження загальної ефективності.

- Дотримання дедлайнів: У системах реального часу завдання повинні бути виконані у визначені часові рамки. Пропущення дедлайнів може призвести до серйозних наслідків, особливо у критичних додатках, таких як медичні системи або системи управління транспортом.

- Ефективне використання ресурсів: Максимальне використання доступних ресурсів, таких як процесорний час, пам'ять та енергетичні ресурси, є важливим для забезпечення ефективної роботи системи. Недостатнє використання ресурсів може призвести до марнування потужностей, а надмірне використання – до перегріву та зниження продуктивності.

- Динамічна адаптація: Системи реального часу повинні бути здатні адаптуватися до змінних умов, таких як зміна навантаження, виникнення нових завдань або зміна пріоритетів. Це вимагає використання адаптивних алгоритмів диспетчеризації, які можуть оперативно реагувати на такі зміни.

- Надійність та відмовостійкість: Забезпечення надійності та здатності системи до відновлення після збоїв є важливим аспектом диспетчеризації завдань. Системи повинні мати механізми для виявлення та корекції помилок, щоб запобігти втратам даних та забезпечити безперервну роботу.

- Врахування енергоспоживання: Зниження енергоспоживання стає все більш важливим, особливо для мобільних та вбудованих систем. Ефективні алгоритми диспетчеризації повинні враховувати енергетичні обмеження та оптимізувати використання енергії без шкоди для продуктивності.

- Складність реалізації: Впровадження складних алгоритмів диспетчеризації може вимагати значних витрат ресурсів на їх розробку та тестування. Необхідно знайти баланс між складністю алгоритму та його ефективністю, щоб забезпечити практичну реалізацію в реальних системах.

Подолання цих викликів є важливим для забезпечення ефективної та надійної роботи систем реального часу. Розробка нових методів диспетчеризації та вдосконалення існуючих підходів потребує врахування цих аспектів для досягнення оптимальної продуктивності.

Виклад основного матеріалу. Ефективність методів диспетчеризації завдань у системах реального часу можна оцінити за допомогою кількох ключових критеріїв:

- Час відгуку: важливий показник, що визначає час між надходженням запиту і початком його обробки. Методи, які забезпечують мінімальний час відгуку, є більш ефективними для систем реального часу.

- Пропускна здатність: Визначає кількість завдань, які система може обробити за одиницю часу. Висока пропускна здатність вказує на ефективне використання ресурсів.

- Здатність до масштабування: Важливо, щоб метод диспетчеризації міг ефективно працювати як при низькому, так і при високому навантаженні. Це показує гнучкість і адаптивність методу до змінних умов системи.

- Використання ресурсів: Оптимізація використання процесорного часу, пам'яті та енергетичних ресурсів є критичною для підтримки стабільної роботи системи. Методи, які максимально використовують доступні ресурси без перевантаження, є більш ефективними.

- Надійність: Здатність методу забезпечити стабільну роботу системи без збоїв або втрат даних. Надійні методи мають механізми для виявлення та корекції помилок.

- Гнучкість: Важливо, щоб метод диспетчеризації міг адаптуватися до змінних умов і нових вимог. Це включає можливість динамічної зміни пріоритетів завдань та адаптацію до нових типів завдань.

Для оцінки ефективності методів диспетчеризації завдань у системах реального часу розглянемо модель. Нехай є система з n завдань, де кожне завдання i характеризується:

- T_i — час виконання завдання i
- D_i — дедлайн завдання i
- P_i — період завдання i
- U_i — використання процесора завданням i
- R_i — час відгуку завдання i

Згідно теореми про верхню межу використання процесора:

$$U(n) = \sum_{i=1}^n \frac{T_i}{P_i} \leq n \left(2^{\frac{1}{n}} - 1 \right)$$

Верхня межа $U(n)$ наближається до 69% ($\ln 2$), коли число завдань наближається до нескінченності. Це оцінка для найгіршого випадку, але, як показано в роботі [5], для випадково обраної групи завдань ймовірна верхня межа дорівнює 88%.

Для кожного завдання i час відгуку можна оцінити з урахуванням інших завдань у системі. Припустимо, що завдання виконуються з пріоритетами. Нехай $HP(i)$ — множина завдань з вищим пріоритетом, ніж у завдання i .

Час відгуку R_i для завдання i можна знайти розв'язавши нерівність:

$$R_i \leq T_i + \sum_{j \in HP(i)} \left\lfloor \frac{R_i}{P_j} \right\rfloor T_j$$

Ця нерівність відображає час, який завдання і витратить на своє виконання та час очікування через завдання з вищими пріоритетами.

Затримка визначається як максимальний час від моменту надходження завдання до його завершення. Для системи з пріоритетами затримка для завдання і визначається як:

$$L_i = R_i + T_i$$

Розглянемо адаптивний метод, де пріоритети завдань можуть змінюватися залежно від стану системи. Нехай $\pi_i(t)$ — пріоритет завдання i в момент часу t . Тоді час відгуку для адаптивного методу можна визначити як:

$$R_i(t) = T_{start_i}(t) + T_i + \sum_{j \in HP_i(t)} \left\lfloor \frac{R_i}{P_j} \right\rfloor T_j$$

де $HP_i(t)$ — множина завдань з вищим пріоритетом у момент часу t .

Для оцінки ефективності введемо функцію ефективності E , яка враховує час відгуку і використання процесора:

$$E = \alpha \left(1 - \frac{1}{n} \sum_{i=1}^n \frac{R_i}{D_i} \right) + \beta (1 - U)$$

де α та β — вагові коефіцієнти, що визначають важливість відповідних критеріїв.

Ця модель дозволяє врахувати різні аспекти ефективності методів диспетчеризації та оптимізувати їх для забезпечення максимальної продуктивності та надійності систем реального часу.

Припустимо, що в системі є три завдання із наступними характеристиками (таблиця 1), які незалежні від стану системи і з'являються зі своїм періодом.

Завдання	T_i	D_i	P_i	R_i
1	2	5	10	3
2	3	3	6	2
3	3	8	12	5

Таблиця 1. Характеристики завдань

Загальне використання процесора U можна обчислити як:

$$U = \sum_{i=1}^n \frac{T_i}{P_i} = \frac{2}{10} + \frac{3}{6} + \frac{3}{12} = 0.95$$

Припустимо, що $\alpha = 0.5$ та $\beta = 0.5$. Тоді функція ефективності буде:

$$E = 0.5 \left(1 - \frac{1}{3} \left(\frac{3}{5} + \frac{2}{3} + \frac{5}{8} \right) \right) + 0.5 (1 - 0.95) \cong 0.2097$$

Отримане значення ефективності E вказує, що система працює на 20.97% від можливої максимальної ефективності за заданими критеріями. Таким чином, значення E надає загальну оцінку ефективності алгоритму диспетчеризації, що допомагає виявити можливості для оптимізації і покращення роботи системи реального часу. Подальші напрямки досліджень у сфері диспетчеризації завдань у системах реального часу можуть зосередитися на інтеграції методів машинного навчання для динамічного прогнозування навантаження та оптимізації алгоритмів у режимі реального часу. Використання моделей глибокого навчання для аналізу даних та прогнозування майбутніх станів системи дозволить більш точно і своєчасно адаптувати розподіл завдань, підвищуючи загальну ефективність і надійність. Крім того, розробка енергоефективних алгоритмів, які враховують змінні енергетичні потреби і можливості, стане ключовим аспектом для мобільних та вбудованих систем. Іншим перспективним напрямом є впровадження адаптивних методів, здатних автоматично налаштовувати параметри диспетчеризації у відповідь на зміни в робочому середовищі, що дозволить забезпечити оптимальну продуктивність навіть за умов непередбачуваних змін. Дослідження в цих напрямках сприятимуть створенню більш гнучких, стійких та енергоефективних систем реального часу, здатних ефективно виконувати критично важливі завдання у сучасних технологічних умовах.

Висновок. Методи диспетчеризації у системах реального часу є критично важливими для забезпечення своєчасного та ефективного виконання завдань. Статичні методи забезпечують передбачуваність і простоту впровадження, тоді як динамічні методи пропонують вищу гнучкість та адаптивність, але потребують більше ресурсів. Комбіновані підходи поєднують переваги обох методів, забезпечуючи баланс між передбачуваністю та гнучкістю. Основні виклики включають збалансування навантаження, дотримання дедлайнів, ефективне використання ресурсів, динамічну адаптацію, надійність, відмовостійкість та врахування енергоспоживання. Оцінка ефективності методів диспетчеризації може базуватися на таких критеріях, як час відгуку, пропускну здатність, здатність до масштабування, використання ресурсів, надійність та гнучкість. Подальші дослідження у цій галузі дозволять створити більш адаптивні, енергоефективні та стійкі до збоїв системи, що є критично важливим для сучасних технологій у різних сферах застосування.

Список бібліографічного опису:

1. Sirshar, M.; Rizwan, H.; Shabbir, M.; Nawaz, H. Scheduling in Real Time System. Preprints 2019, 2019120072. URL: <https://doi.org/10.20944/preprints201912.0072.v1>
2. Kumar, A. & Alam, B. (2018). Task Scheduling in Real Time Systems with Energy Harvesting and Energy Minimization. Journal of Computer Science, 14(8), 1126-1133. URL: <https://doi.org/10.3844/jcssp.2018.1126.1133>
3. Kohútka L. A New FPGA-Based Task Scheduler for Real-Time Systems. Electronics. 2023, 12(8), 1870. URL: <https://doi.org/10.3390/electronics12081870>
4. Sprunt, B., Sha, L. & Lehoczky, J. Aperiodic task scheduling for Hard-Real-Time systems. Real-Time Syst 1, 27–60 (1989). <https://doi.org/10.1007/BF02341920>
5. J. Lehoczky, L. Sha and Y. Ding, «The rate monotonic scheduling algorithm: exact characterization and average case behavior,» [1989] Proceedings. Real-Time Systems Symposium, Santa Monica, CA, USA, 1989, pp. 166-171. URL: <https://doi.org/10.1109/REAL.1989.63567>

References:

1. Sirshar, M.; Rizwan, H.; Shabbir, M.; Nawaz, H. Scheduling in Real Time System. Preprints 2019, 2019120072. URL: <https://doi.org/10.20944/preprints201912.0072.v1>
2. Kumar, A. & Alam, B. (2018). Task Scheduling in Real Time Systems with Energy Harvesting and Energy Minimization. Journal of Computer Science, 14(8), 1126-1133. URL: <https://doi.org/10.3844/jcssp.2018.1126.1133>
3. Kohútka L. A New FPGA-Based Task Scheduler for Real-Time Systems. Electronics. 2023, 12(8), 1870. URL: <https://doi.org/10.3390/electronics12081870>
4. Sprunt, B., Sha, L. & Lehoczky, J. Aperiodic task scheduling for Hard-Real-Time systems. Real-Time Syst 1, 27–60 (1989). <https://doi.org/10.1007/BF02341920>
5. J. Lehoczky, L. Sha and Y. Ding, «The rate monotonic scheduling algorithm: exact characterization and average case behavior,» [1989] Proceedings. Real-Time Systems Symposium, Santa Monica, CA, USA, 1989, pp. 166-171. URL: <https://doi.org/10.1109/REAL.1989.63567>