

Лабораторна робота №5 (РГР)

Дослідження роботи планувальників роботи систем реального часу

Мета роботи — змодельовати роботу планувальника задач у системі реального часу

Основні теоретичні відомості

Планування виконання завдань (англ. Scheduling) є однією з ключових концепцій в багатозадачності і багатопроцесорних систем, як в операційних системах загального призначення, так і в операційних системах реального часу. Планування полягає в призначенні пріоритетів процесам в черзі з пріоритетами.

Найважливішою метою планування завдань є якнайповніше завантаження доступних ресурсів. Для забезпечення загальної продуктивності системи планувальник має опиратися на:

- Використання процесора(-ів) — дати завдання процесору, якщо це можливо.
- Пропускна здатність — кількість процесів, що виконуються за одиницю часу.
- Час на завдання — кількість часу, для повного виконання певного процесу.
- Очікування — кількість часу, який процес очікує в черзі готових.
- Час відповіді — час, який проходить від подання запиту до першої відповіді на запит.
- Справедливість — Рівність процесорного часу для кожної ниті

У середовищах обчислень реального часу, наприклад, на пристроях, призначених для автоматичного управління в промисловості (наприклад, робототехніка), планувальник завдань повинен забезпечити виконання

процесів в перебігу заданих часових проміжків (час відгуку); це критично для підтримки коректної роботи системи реального часу.

Система масового обслуговування (СМО) — система, яка виконує обслуговування вимог (заявок), що надходять до неї. Обслуговування вимог у СМО проводиться обслуговуючими приладами. Класична СМО містить від одного до нескінченного числа приладів. В залежності від наявності можливості очікування вхідними вимогами початку обслуговування СМО (наявності черг) поділяються на:

- 1) системи з втратами, в яких вимоги, що не знайшли в момент надходження жодного вільного приладу, втрачаються;
- 2) системи з очікуванням, в яких є накопичувач нескінченної ємності для буферизації надійшли вимог, при цьому очікують вимоги утворюють чергу;
- 3) системи з накопичувачем кінцевої ємності (чеканням і обмеженнями), в яких довжина черги не може перевищувати ємності накопичувача; при цьому вимога, що надходить в переповнену СМО (відсутні вільні місця для очікування), втрачається.

Основні поняття СМО:

- *Вимога* (заявка) — запит на обслуговування.
- *Вхідний потік вимог* — сукупність вимог, що надходять у СМО.
- *Час обслуговування* - період часу, протягом якого обслуговується вимогу.

Вимоги до системи

Вхідні задачі

Вхідними заявками є обчислення, які проводилися в лабораторних роботах 1-3, а саме обчислення математичного очікування, дисперсії, автокореляції, перетворення Фур'є.

Вхідні заявки характеризуються наступними параметрами:

- 1) час приходу в систему – T_r – потік заявок є потоком Пуассона або потоком Ерланга k -го порядку (інтенсивність потоків та їх порядок задаються варіантом);
- 2) час виконання (обробки) – T_o ; математичним очікуванням часу виконання є середнє значення часу виконання відповідних обчислень в попередніх лабораторних роботах;
- 3) крайній строк завершення (дедлайн) – T_d – задається (випадково?); якщо заявка залишається необробленою в момент часу $t = T_d$, то її обробка припиняється і вона покидає систему.

Потік вхідних задач

Потоком Пуассона є послідовність випадкових подій, середнє значення інтервалів між настанням яких є сталою величиною, що дорівнює $1/\lambda$, де λ – інтенсивність потоку.

Потоком Ерланга k -го порядку називається потік, який отримується з потоку Пуассона шляхом збереження кожної $(k+1)$ -ї події (решта відкидаються). Наприклад, якщо зобразити на часовій осі потік Пуассона, поставивши у відповідність кожній події деяку точку, і відкинути з потоку кожну другу подію (точку на осі), то отримаємо потік Ерланга 2-го порядку. Залишивши лише кожну третю точку і відкинувши дві проміжні, отримаємо потік Ерланга 3-го порядку і т.д. Очевидно, що потоком Ерланга 0-го порядку є потік Пуассона.

Пристрій обслуговування

Пристрій обслуговування складається з P незалежних рівноправних обслуговуючих приладів - обчислювальних ресурсів (процесорів). Кожен ресурс обробляє заявки, які йому надає планувальник та може перебувати у двох станах – вільний та зайнятий. Обробка заявок може виконуватися повністю (заявка перебуває на обчислювальному ресурсі доти, доки не обробиться повністю) або поквантово (ресурс обробляє заявку лише

протягом певного часу – кванту обробки – і переходить до обробки наступної заявки).

Пріоритети заявок

Заявки можуть мати пріоритети – явно задані, або обчислені системою (в залежності від алгоритму обслуговування або реалізації це може бути час обслуговування (обчислення), час до дедлайну і т.д.). Заявки в чергах сортуються за пріоритетом. Є два види обробки пріоритетів заявок:

- 1) без витіснення – якщо в чергу до ресурсу потрапляє заявка з більшим пріоритетом, ніж та, що в даний момент часу обробляється ним, то вона чекає завершення обробки ресурсом його задачі.
- 2) з витісненням – якщо в чергу до ресурсу потрапляє заявка з більшим пріоритетом, ніж та, що в даний момент часу обробляється ним, то вона витісняє її з обробки; витіснена задача стає в чергу.

Дисципліна обслуговування

Вибір заявки з черги на обслуговування здійснюється за допомогою так званої дисципліни обслуговування. Їх прикладами є FIFO (прийшов першим - обслуговується першим), LIFO (прийшов останнім - обслуговується першим), RANDOM (випадковий вибір). У системах з очікуванням накопичувач в загальному випадку може мати складну структуру.

Дисципліна RR

Алгоритм Round-Robin (від англ. round-robin — циклічний) – алгоритм розподілу навантаження на розподілену (або паралельну) обчислювальну систему методом перебору і впорядкування її заявок по круговому циклу. Даний алгоритм не враховує пріоритети вхідних заявок.

Нехай є P ресурсів (з порядковими номерами p) та X заявок (з порядковими номерами x), які необхідно виконати. Тоді перша заявка ($x = 1$) назначається для виконання на першому ресурсі ($p = 1$), друга ($x = 2$) –

другому і т.д., до досягнення зайнятості останнього ресурсу ($p = P, x = P$) або до вичерпування необроблюваних заявок ($x = X$). Усі наступні заявки будуть розподілені по ресурсах аналогічно до попередніх, починаючи з першого ресурсу ($x = P + 1 \rightarrow p = 1, x = P + 2 \rightarrow p = 2$ і т.д.). Іншими словами відбувається перебір ресурсів по циклу (по колу – round).

Обчислення задач розділене на кванти часу, причому по закінченню кванту завершені та прострочені задачі виходять з системи, незавершені – здвигаются по колу на 1 ресурс (тобто задача першого об'єкта передається другому, другого – третьому і т.д., останнього – першому).

Дисципліна EDF

Алгоритм планування Earliest Deadline First (по найближчому строку завершення) використовується для встановлення черги заявок в операційних системах реального часу.

При настанні події планування (завершився квант часу, прибула нова заявка, завершилася обробка заявки, заявка прострочена) відбувається пошук найближчої до крайнього часу виконання (дедлайну) заявки і призначення її виконання на перший вільний ресурс або на той, який звільниться найшвидше.

Дисципліна FB

Алгоритм планування Foreground-Background є алгоритмом багаторівневого циклічного планування, який забезпечує швидшу реакцію системи на заявки з малим часом виконання.

В системі є Q FIFO-черг заявок (з порядковими номерами q). Вхідні заявки поступають в чергу $q = 1$. Пріоритет обслуговування черг відповідає порядковому номеру q черги, черга з порядковим номером $q = 1$ має найвищий пріоритет, тобто якщо в черзі з порядковим номером q є хоча б одна незавершена заявка, то черга $q+1$ не обслуговується.

При обслуговуванні черги q заявка назначається на виконання на перший вільний ресурс або на той, який звільниться найшвидше і обробляється на ньому протягом кванту часу; якщо за цей час заявка була

оброблена або прострочена, то вона покидає систему. В іншому випадку заявка переходить у чергу $q + 1$, $\max(q+1) = Q$.



Рисунок 1. Схема черг дисципліни FB

Завдання на лабораторну роботу

1. Змодельовати планувальник роботи системи реального часу. Дві дисципліни планування: перша – RR, друга задається викладачем або обирається самостійно.
2. Знайти наступні значення:
 - 1) середній розмір вхідної черги заявок, та додаткових черг (за їх наявності);
 - 2) середній час очікування заявки в черзі;
 - 3) кількість прострочених заявок та її відношення до загальної кількості заявок
3. Побудувати наступні графіки:
 - 1) Графік залежності кількості заявок від часу очікування при фіксованій інтенсивності вхідного потоку заявок.
 - 2) Графік залежності середнього часу очікування від інтенсивності вхідного потоку заявок.
 - 3) Графік залежності проценту простою ресурсу від інтенсивності вхідного потоку заявок.

Приклад графіків

Алгоритм – випадковий вибір

Кількість заявок = 1000

Мінімальна вага заявки = 1.0

Максимальна вага заявки = 10.0

Інтенсивність = 1.0

Мінімальна інтенсивність = 0.05, максимальна = 10.0, крок = 0.05



