

УДК 004.94

Одарущенко О.Б. доцент, к.т.н.

Радченко І.А. магістрант

Полтавський національний технічний університет

імені Юрія Кондратюка

МОДЕЛЮВАННЯ БРОКЕРА РЕСУРСІВ В ГАРАНТОЗДАТНІЙ ГРІД З ВИКОРИСТАННЯМ ТЕОРІЇ МАСОВОГО ОБСЛУГОВУВАННЯ

Розглядається моделювання брокера ресурсів у гарантоздатній грід-системі з використанням теорії масового обслуговування. Побудовано модель грід-системи та досліджено її характеристики і потік заявок у систему, що дозволяє оцінити ефективність самої даної системи масового обслуговування.

Ключові слова: система масового обслуговування, грід-система, часові ряди, розподілена система, модель ARIMA, граничні ймовірності.

Вступ

Технологія GRID (англ. grid - решітка, мережа) використовується для створення географічно розподіленої обчислювальної інфраструктури, що об'єднує ресурси різних типів з колективним доступом до цих ресурсів у рамках віртуальних організацій, що складаються з підприємств та спеціалістів, що спільно використовують ці ресурси [1].

Існує ряд важливих наукових задач, без розв'язання яких неможливе повномасштабне використання можливостей GRID-технологій в науці та промисловості. Однією з таких актуальних задач є ефективне керування ресурсами в розподіленому середовищі. Відсутність хорошого планувальника, що забезпечує керування потоком задач, не тільки значно знижує ефективність використання всієї GRID-інфраструктури, але може зробити необґрунтованим її створення.

Огляд [2-4] свідчить про різноманітність підходів до розв'язання проблеми, а також про необхідність розробки ефективних способів моделювання різних стратегій управління ресурсами розподілених систем.

Одним з розділів моделювання GRID-систем є теорія масового обслуговування [5], що є теоретичною основою ефективного конструювання і експлуатації систем масового обслуговування (СМО). СМО зустрічаються в багатьох областях і призначені для багатократного використання при виконанні однотипних завдань.

Дослідження даної теми є актуальним, оскільки велика кількість питань у сфері моделювання комп'ютерних GRID-систем на даному етапі є невирішеними.

Мета роботи

Моделювання брокера ресурсів грід-системи та дослідження потоку задач у системі з використанням системи масового обслуговування.

Завдання

Дослідження потоку задач у грід-системі та моделювання системи з подальшим аналізом її характеристик.

Аналіз та прогнозування задач і моделювання GRID з використанням СМО

Основною задачею ГРІД є інтеграція великої кількості розподілених ресурсів, що можуть взаємодіяти одне з одним для досягнення деякої цілі.

Брокер ресурсів у GRID-системі розподіляє задачі між сайтами. Одним із варіантів його моделювання є представлення його як елемента у системі масового обслуговування. Для кращого розподілу задач можна зробити аналіз і прогнозування вхідного потоку задач та аналіз завантаженості обчислювальних ресурсів.

Прогнозування характеру потоку задач може здійснюватися за допомогою різних способів, проте як показала практика найбільш ефективним є застосування аналізу часових рядів [6]. Існує дві основні цілі аналізу часових рядів: визначення природи ряду та прогнозування.

Аналіз часового ряду починається з побудови графіка, адже з його вигляду можна відразу зробити припущення про наявність чи відсутність тренду, циклічності та сезонної компонент (та її характер – адитивна чи мультиплікативна). На рисунку 1 видно, що часовий ряд не має ні періодичних складових, ні тренду, а самі значення варіюються у околі фіксованого рівня, що приблизно рівний 2700, а отже модель поведінки є горизонтальною.

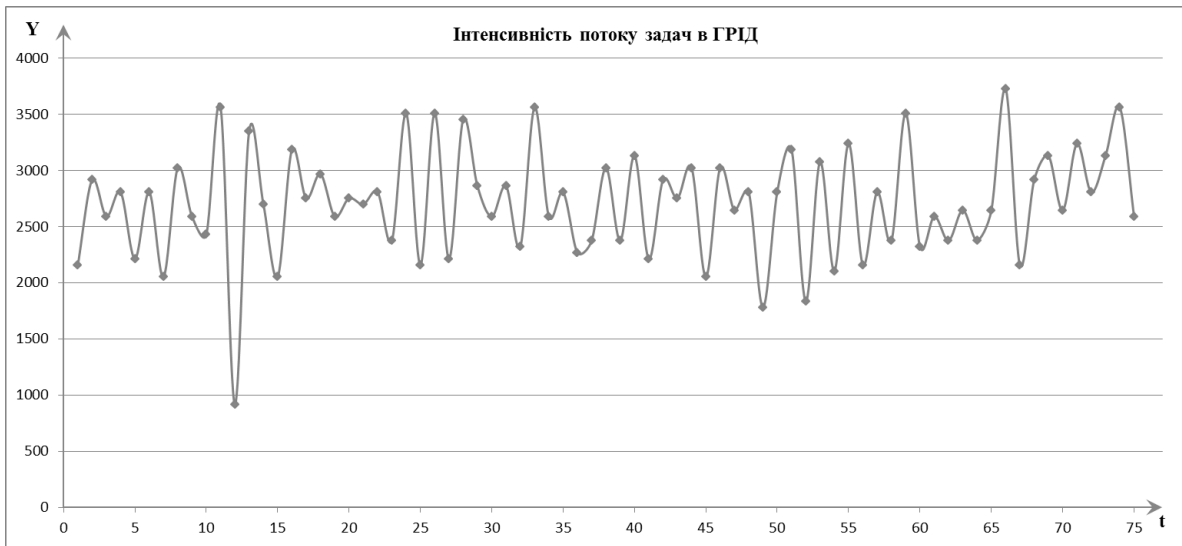


Рис.1. Графік часового ряду

Далі знаходяться числові характеристики часового ряду. Середній рівень ряду за формулою (1):

$$\bar{y} = \frac{1}{n} \sum_{t=1}^n y_t, \quad (1)$$

і допоміжні величини $y_t - \bar{y}$ для $t = 1, 2, 3, \dots, n$, та їх квадрати – $(y_t - \bar{y})^2$.

Знаючи всі необхідні значення знаходяться дисперсія часового ряду:

$$D = \frac{1}{n} \cdot \sum_{t=1}^n (y_t - \bar{y})^2$$

та середньоквадратичне відхилення, що дорівнює квадратному кореню з дисперсії ($\sigma = \sqrt{D}$). У роботі значення дорівнюють $D = 2441169,51$ і $\sigma = 494,14$.

Видно, що середньоквадратичне відхилення та дисперсія є досить великими, а це говорить про значне розсіяння досліджуваної величини.

Далі, щоб перевірити або спростувати припущення про структуру ряду проводиться автокореляційний аналіз. Для цього спочатку обчислюються коефіцієнти автокореляційної функції за формулою (2):

$$r_k = \frac{\sum_{t=1}^{n-k} (y_t - \bar{y}) \cdot (y_{t+k} - \bar{y})}{n \cdot \sigma^2}. \quad (2)$$

Слід також врахувати, що величина максимального лагу не повинна перевищувати $n/4$. У роботі це $75/4 = 18,75$ тобто $k = 1, 2, \dots, 18$.

Рівняння (3) є простою випадковою моделлю, що часто називають моделлю білого шуму. Спостереження $y_t, t = 1, 2, \dots, n$ складається з двох частин: c – загальний рівень і ε_t – компонента, що представляє собою випадкову помилку. Важливо відмітити, що компонента ε_t не корельована від періоду до періоду.

$$y_t = c + \varepsilon_t \quad (3)$$

Необхідно перевірити чи співпадають початкові дані з цією моделлю.

Існує визначений критерій, розроблений для перевірки гіпотези про те, чи є коефіцієнт кореляції суттєво відмінним від нуля. Далі приведено нульова та альтернативна гіпотеза перевірки значущості коефіцієнтів автокореляційної функції із запізненням на k періодів:

$$H_0: r'_k = 0, \quad H_1: r'_k \neq 0.$$

Для перевірки нульової гіпотези може бути використана t -статистика, визначена рівнянням (4):

$$t_k = \frac{r_k - r'_k}{SE(r_k)} \quad (4)$$

Перш ніж перевірити дану гіпотезу знаходяться стандартні похибки для автокореляцій за формулою (5):

$$SE(r_k) = \sqrt{\frac{1 + 2 \cdot \sum_{i=1}^{k-1} r_i^2}{n}}, \quad (5)$$

де $SE(r_k)$ – стандартна похибка автокореляції із запізненням k , r – автокореляція із запізненням i , k – час запізнення, n – кількість спостережень у часовому ряді. Для автокореляції, що відповідає запізненню в один період, використовується стандартна похибка $1/\sqrt{n}$, тобто $SE(r_1) = 1/\sqrt{n}$. Далі на основі отриманих значень знаходяться t -статистики за формулою (6):

$$t_k = \frac{r_k - r'_k}{SE(r_k)}. \quad (6)$$

Протестувати коефіцієнти автокореляції можна використавши модифіковану Q -статистику запропоновану Льюїнгом та Боксом. Вона має розподіл Пірсона з m степенями свободи, де m – кількість запізнень, для яких робиться перевірка:

$$Q = n \cdot (n + 2) \cdot \sum_{k=1}^m \frac{r_k^2}{n - k}.$$

У роботі для рівня значущості $\alpha = 0,05$ і $m = 18$ степенів свободи χ^2 -значення дорівнює 28,8693, а $Q = 57,758 > 28,8693$, це говорить про те, що часовий ряд не є випадковим.

На основі отриманих даних будується корелограма (рисунок 2).

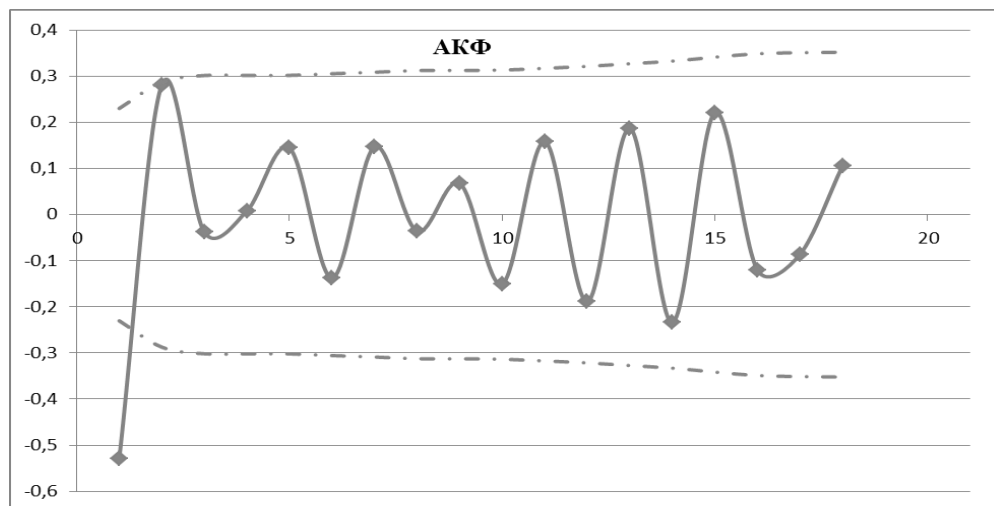


Рис.1. Автокореляційна функція часового ряду

Звідси можна зробити висновок, що дані не є випадковими. Окрім того не спостерігається викидів автокореляційної функції, що свідчить про відсутність

сезонної компоненти. Загалом же коефіцієнти не повільно спадають до нуля, а вже на третьому лазі майже рівні нулю. Таким чином, можна стверджувати, що заданий часовий ряд не має тренду і є стаціонарним.

Не завжди можна чітко визначити наявність або відсутність тренду. Для перевірки гіпотези існування тренду в рамках роботи розглядається критерій серій, а точніше його модифікація – критерій «висхідних та низхідних» серій.

На першому кроці даного методу будується послідовність із плюсів та мінусів за правилом:

$$d_i = \begin{cases} +, y_{t+1} - y_t > 0, t = 1, 2, \dots, n - 1, \\ -, y_{t+1} - y_t < 0, t = 1, 2, \dots, n - 1. \end{cases}$$

Потім перевіряються наступні умови:

$$v(n) > \left[\frac{1}{3} \cdot (2n - 1) - 1,96 \cdot \sqrt{\frac{16n - 29}{30}} \right],$$

$$\tau_{max}(n) < \tau_0(n)$$

де $v(n)$ – число серій, що у даній роботі рівне 66, $\tau_{max}(n)$ – довжина найдовшої серії, що у даній роботі рівна 2, значення $\tau_0(n) = 6$ для $n = 75$.

Умови виконуються:

$$v(n) = 66 > 42, \quad \tau_{max}(n) = 2 < \tau_0(n) = 6,$$

тому гіпотеза про відсутність тренду приймається.

Перед переходом до прогнозування часового ряду, будується часна автокореляційна функція. Це потрібно для того, щоб ідентифікувати модель ARIMA, обрану в даній роботі для знаходження майбутніх значень часового ряду.

На лазі один, часна автокореляційна функція рівна звичайній автокореляції – $p_1 = r_1$, решта обчислюється за наступною формулою:

$$p_k = \frac{r_k - \sum_{j=1}^{k-1} p_{k-1,j} \cdot r_{k-j}}{1 - \sum_{j=1}^{k-1} p_{k-1,j} \cdot r_j}, p_{k,j} = p_{k-1,j} - p_k \cdot p_{k-1,k-j},$$

де $p_{k,k} = p_k$.

Далі знаходяться межі інтервалів надійності:

$$\text{Верхня межа} = t_{0,975} \cdot \left(\frac{1}{\sqrt{n}}\right) = 1,9927 \cdot 0,1155 = 0,2301,$$

$$\text{Нижня межа} = t_{0,975} \cdot \left(\frac{1}{\sqrt{n}}\right) = -1,9927 \cdot 0,1155 = -0,2301.$$

На основі отриманих розрахунків будується графік (рисунок 3).

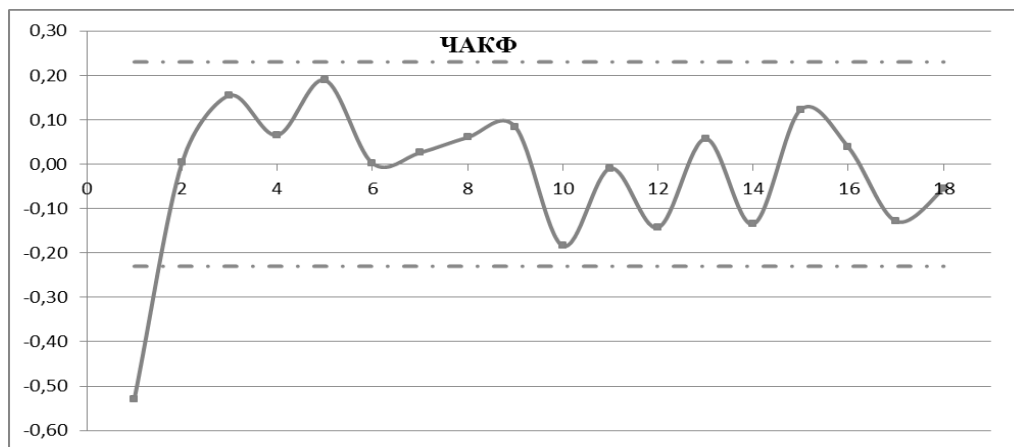


Рис.3. Часна автокореляційна функція

Після аналізу автокореляційної та часної автокореляційної функції для часового ряду отримаємо дві моделі ARIMA: AR(1) та MA(2), параметри яких треба оцінити:

$$y_t = \phi_0 + \phi_1 \cdot y_{t-1} + \varepsilon_t, \quad (7)$$

$$y_t = \mu + \varepsilon_t - \omega_1 \cdot \varepsilon_{t-1} - \omega_2 \cdot \varepsilon_{t-2}. \quad (8)$$

Для моделі (7) невідомих параметрів два: ϕ_0 та ϕ_1 . Для моделі (8) їх три: μ , ω_1 та ω_2 . Для знаходження оцінок використовується додаток Minitab. Значення невідомих параметрів для (7) та (8) дорівнюють $\phi_0 = 4170,33$, $\phi_1 = -0,5379$ та $\mu = 2714,77$, $\omega_1 = 0,5667$, $\omega_2 = -0,356$ відповідно. Отже:

$$\hat{y}_t = \phi_0 + \phi_1 \cdot y_{t-1} = 4170,33 - 0,5379 \cdot y_{t-1}, \quad (9)$$

$$\hat{y}_t = \mu - \omega_1 \cdot \varepsilon_{t-1} - \omega_2 \cdot \varepsilon_{t-2} = 2714,77 - 0,5667 \cdot \varepsilon_{t-1} + 0,356 \cdot \varepsilon_{t-2}. \quad (10)$$

Далі перевіряється адекватність моделей. Спочатку знаходимо залишки для обох випадків, побудувавши прогнози для відомих значень.

Для першого рівня попереднє значення невідоме, отже і скласти прогноз у випадку моделі (7) для нього неможливо, для решти рівнів прогноз знаходиться за формулою (9). На основі цих даних знаходяться залишки для першої моделі за формулою:

$$e_i = y_i - \hat{y}_i.$$

Далі будується залишкова автокореляційна функція для першої моделі:

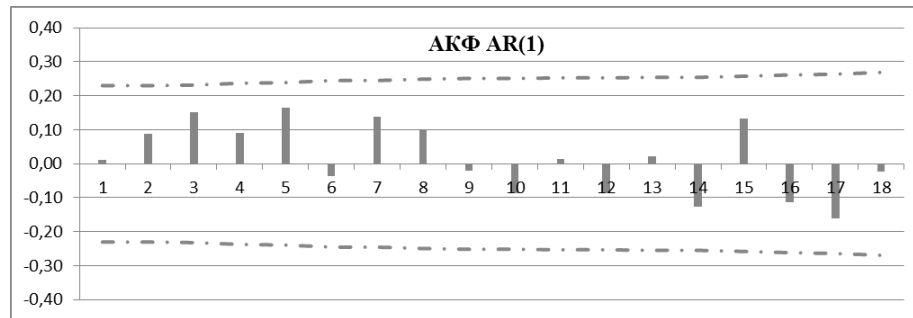


Рис.4. Залишкова автокореляційна функція для моделі AR(1)

Усі залишкові коефіцієнти досить малі і знаходяться в межах надійного інтервалу, отже можна стверджувати, що помилки є випадковими. Перевіряємо це твердження розрахувавши для залишків статистику Льюїнга-Бокса. Значення χ^2 -розподілу рівне 26,2962.

$$Q = n \cdot (n + 2) \cdot \sum_{k=1}^m \frac{r_k^2}{n - k} = 16,362.$$

Оскільки $16,362 < 26,2962$ – похибки не суттєві.

Розраховується також залишкова середньоквадратичну похибка:

$$S^2 = \frac{\sum_{t=1}^n e_t^2}{n - r} = \frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n - r} = 178692.$$

Аналогічно і для другої моделі:

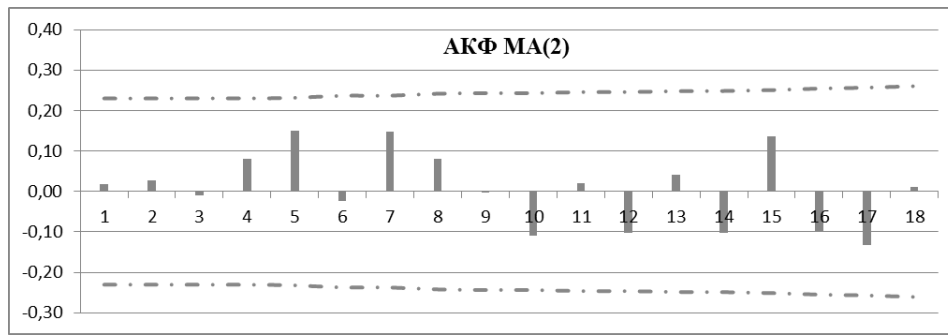


Рис.5. Залишкова автокореляційна функція для моделі MA(2)

Так як усі залишкові коефіцієнти автокореляції є близькими до нуля та знаходяться в межах допустимого інтервалу, похибки випадкові. Додатково розрахуємо Q-статистику та порівняємо її із табличним значенням χ^2 -розподілу для рівня значущості $\alpha = 0,05$ із степенями свободи $m - 3 = 15$, тобто з 24,9958. Отримаємо $12,7106 < 24,9958$, що свідчить про те, що похибки несуттєві. Залишкова середньоквадратична похибка (S^2) для моделі (8) дорівнює 175049.

Обидві моделі показали гарну відповідність даним. Моделі є адекватними і похибки є випадковими. А залишкові середньоквадратичні похибки майже однакові. Тому для прогнозування підходить як модель AR(1), так і MA(2).

Для таких випадків, було розроблено декілька підходів вибору моделі, що враховують як якість підгонки моделі, так і кількість її параметрів. Інформаційний критерій Акайке (Akaike), або AIC, дозволяє вибрати найкращу модель із групи моделей претендентів. Відповідно до цього критерію, обирається модель, що мінімізує вираз:

$$AIC = \ln \hat{\sigma}^2 + \frac{2}{n} \cdot r \quad (11)$$

де $\hat{\sigma}^2$ – залишкова сума квадратів, поділена на кількість спостережень, n – кількість спостережень, r – загальна кількість доданків (включаючи постійний доданок) в моделі ARIMA. Відповідно до Бессовського інформаційного критерію, що розроблений Шварцем (Schwarz), або BIC, обирається та модель, котра мінімізує наступний вираз:

$$BIC = \ln \hat{\sigma}^2 + \frac{\ln n}{n} \cdot r \quad (12)$$

Другий доданок у формулах (11-12) – це «штрафний фактор», що враховує включення в модель додаткових параметрів. Критерій BIC накладає більше обмеження на кількість параметрів в порівнянні з критерієм AIC. Тому мінімізація критерію BIC при виборі моделі завжди дасть кількість параметрів, що не перевищує кількості, встановленої відповідно критерію AIC. Часто обидва критерії дають один і той же результат. В даному випадку відразу дві моделі ARIMA дають цілком адекватний опис даних. Перша – це модель AR(1) з $r = 2$ параметрами (включаючи постійний член) та $\hat{\sigma}^2 = 173926,797$. Другою була модель MA(2) з кількістю параметрів $r = 3$ та $\hat{\sigma}^2 = 168047,466$. Розрахуємо за формулами (11-12) для них критерії AIC та BIC. Отримаємо:

$$\begin{aligned} AIC_{AR(1)} &= 12,12, & AIC_{MA(2)} &= 12,112, \\ BIC_{AR(1)} &= 12,182, & BIC_{MA(2)} &= 12,204. \end{aligned}$$

Критерії AIC та BIC показали результати, що мають протиріччя. Критерій AIC менший для моделі MA(2), в той час як критерій BIC менший для моделі AR(1). Хоча в другому випадку різниця більш суттєва ніж в першому, тому враховуючи також принцип економії в якості моделі прогнозування краще взяти AR(1).

Отже розрахуємо значення майбутніх на даних на три періоди вперед за допомогою моделі AR(1).

Рівняння прогнозу AR(1) має вигляд:

$$\hat{y}_t = \phi_0 + \phi_1 \cdot y_{t-1} = 4170,33 - 0,5379 \cdot y_{t-1}.$$

Отже значення для наступних періодів:

$$\hat{y}_{76} = 2579,16; \quad \hat{y}_{77} = 2705,43; \quad \hat{y}_{78} = 2713,31;$$

Тепер розглянемо 5-канальну ($n = 5$) модель СМО з відмовами, та знайдемо її характеристики. Хай маємо глобальний потік задач. Час надходження між заявками розподілений за експоненціальним законом з параметром λ (хай буде близьке періоду \hat{y}_{76} , а саме 2579). Час обслуговування глобального потоку задач на всіх

пристроях розподілених за експоненціальним законом з параметром $\mu = 864$. У якості показників ефективності багатоканальної СМО з відмовами будемо розглядати:

A – абсолютна пропускна спроможність СМО;

Q – відносна пропускна спроможність;

$P_{\text{відм}}$ – ймовірність відмови;

\bar{k} – середнє число зайнятих каналів,

які обчислюються за формулами:

$$P_{\text{відм}} = \frac{\rho^n}{n!} p_0, \quad (13)$$

$$Q = 1 - P_{\text{відм}} = 1 - \frac{\rho^n}{n!} p_0, \quad (14)$$

$$A = \lambda Q = \lambda \left(1 - \frac{\rho^n}{n!} p_0 \right), \quad (15)$$

$$\bar{k}_{\text{зайн}} = \sum_{k=0}^n k p_k = \frac{A}{\mu} = \rho \left(1 - \frac{\rho^n}{n!} p_0 \right). \quad (16)$$

Розглянемо граф станів системи (рисунок 6).

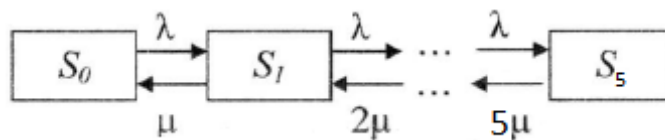


Рис.6. Граф станів системи

S_0 – всі канали вільні;

S_1 – один канал зайнятий;

...

S_5 – всі канали зайняті;

Ймовірність того, що всі канали вільні знаходиться за формулою (17):

$$p_0 = \left(1 + \frac{\rho}{1!} + \frac{\rho^2}{2!} + \dots + \frac{\rho^n}{n!}\right)^{-1} \quad (17)$$

де $\rho = \frac{\lambda}{\mu}$ – інтенсивність навантаження каналу.

Ймовірність завантаженості n каналів системи обчислюється за формулою (18):

$$p_n = \frac{\rho^n}{n!} p_0 \quad (18)$$

Отже граничні ймовірності станів системи: $p_0 = 0,055$, $p_1 = 0,1644$, $p_2 = 0,2454$, $p_3 = 0,2442$, $p_4 = 0,1822$, $p_5 = 0,1088$.

Таким чином 5,5% часу система не працює через відсутність заявок, 16,44% працює лише один пристрій. 24,54%, 24,42%, 18,22% та 10,88% відповідно завантажені 2, 3, 4 та 5 пристроїв.

Ймовірність відмови $P_{\text{відм}} = p_5 = 10,88\%$.

Відносна пропускна спроможність системи: $Q = 0,8912$, тобтодесь 89,12% заявок будуть опрацьовані.

Абсолютна пропускна спроможність системи: $A = 2,66$ (середнє число заявок, що обслуговуються у одиницю часу).

Середнє число зайнятих каналів: $\bar{k}_{\text{зайн}} = 0,003$.

Таким чином розраховано характеристики та ймовірності для 5-канальної СМО з відмовами для періоду \hat{y}_{76} .

Отже якщо взяти значення періодів як інтенсивність потоку та розрахувати характеристики і граничні ймовірності (рисунок 7), можна отримати уявлення про завантаженість системи у майбутньому для прийняття правильного рішення відносно профілактичних робіт, не погіршуючи продуктивність системи, чи вдосконалення алгоритму розподілу задач з певними параметрами.

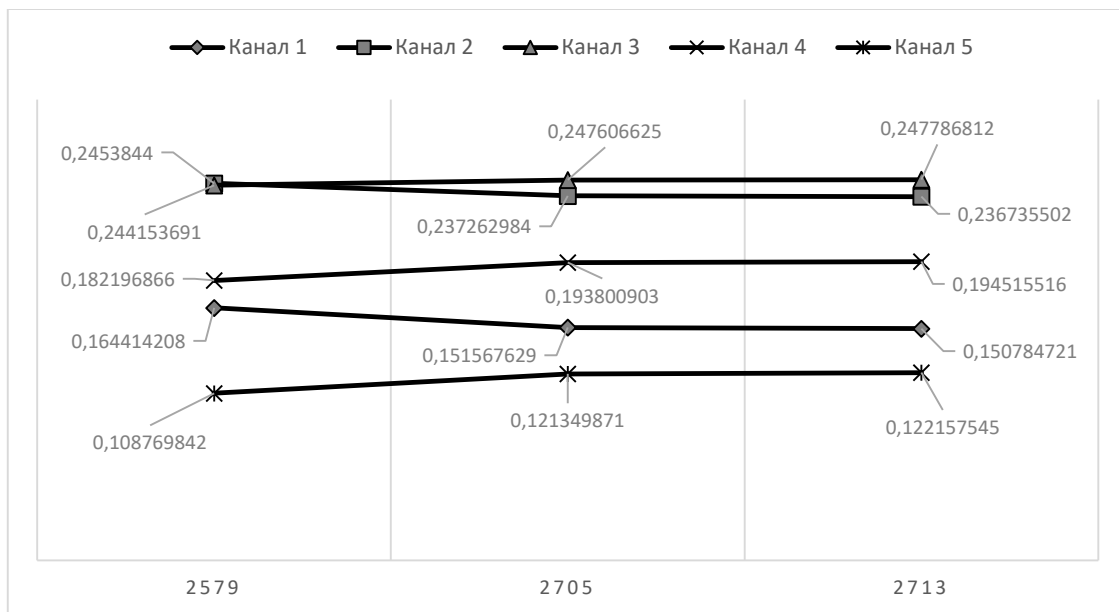


Рис.7. Граничні ймовірності системи для \hat{y}_{76} , \hat{y}_{77} та \hat{y}_{78} періодів

Висновки

У роботі було розглянуто базові поняття GRID-систем та один із напрямків досягнення їх відмовостійкості, а саме прогнозування характеристик потоку задач в GRID, таких як інтенсивність потоку або пропускна здатність системи. Також було розглянуто GRID-систему як систему масового обслуговування та розраховано її характеристики. Для автоматизації процесу прогнозування характеру потоку задач та розрахування характеристик системи було створено відповідний додаток.

Список використаної літератури

1. Кирьянов А.К., Рябов Ю.Ф., Введение в технологию Грид. Петербургский институт ядерной физики им. Б.П. Константинова РАН. – Гатчина, 2006 – 39 с.
2. Харченко В.С., Боярчук А.В., Куланов С.О., Локазюк В.М., Одарущенко О.М., Поморова О.В., Фурманов О.А. Моделирование гарантоздатных систем и сетей. Лекційний матеріал. - Міністерство освіти і науки України, Національний аерокосмічний університет ім. М.Є. Жуковського «ХАІ», 2008. - 336 с.
3. Коваленко В.Н. Организация ресурсов грид [Електронний ресурс] / В.Н. Коваленко, Д.А. Корягин // Preprint, Inst. Appl. Math., the Russian Academy of Science. – Москва, 2004. – Режим доступу: http://www.keldysh.ru/papers/2004/prep63/prep2004_63.html#_Toc85204258.

4. Литвинов В.В., Стеценко І.В. Управління розподіленими ресурсами ґрид-системи. *Математичні машини і системи*, 2012, №2, 3-12с.
5. Кошуняева Н.В., Патронова Н.Н., Теория массового обслуживания (практикум по решению задач) / САФУ имени М.В. Ломоносова. - Архангельск; САФУ, 2013 - 107 с.
6. Бокс Дж., Дженкінс Г., Аналіз часових рядів. Прогноз та керування: Переклад з англійської під редакцією Писаренко В.Ф. – М.: Мир, 1974. – 406 с.

Одарущенко Е.Б. доцент, к.т.н.

Радченко И.А. магистрант

Полтавский национальный технический университет

имени Юрия Кондратюка

МОДЕЛИРОВАНИЕ БРОКЕРА РЕСУРСОВ В ГАРАНТОСПОСОБНОЙ ГРИД С ИСПОЛЬЗОВАНИЕМ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ

Рассматривается моделирование брокера ресурсов в гарантоспособной ґрид-системе с использованием теории массового обслуживания. Построена модель ґрид-системы и исследованы ее характеристики и поток заявок в систему, что позволяет оценить эффективность самой системы массового обслуживания.

Ключевые слова: *система массового обслуживания, ґрид-система, временные ряды, распределенная система, модель ARIMA, предельные вероятности.*

O.B. Odaruschenko, PhD

I.A. Radchenko, student

Poltava National Technical Yuri Kondratyuk University

MODELING RESOURCES BROKER OF DEPENDABLE GRID USING QUEUING THEORY

Modeling resource broker of dependable GRID system using queuing theory is regarded. We constructed a model of GRID system and explored its characteristics and flow applications in the system, that allows asses the effectiveness of the system of queuing.

Keywords: *queuing system, GRID-system, time series, distributed system, ARIMA model, limit propability.*