УДК 681.3.

Федюков Д. А., Рябчун С. Г.

МАТЕМАТИЧНА МОДЕЛЬ ДЛЯ ПОБУДОВИ КОМПЛЕКСУ ВИЯВЛЕННЯ НЕСПРАВНОСТЕЙ ВИСОКОПРОДУКТИВНОЇ ОБЧИСЛЮВАЛЬНОЇ СИСТЕМИ

У роботі показано математичну модель на основі мереж Петрі для вирішення завдань виявлення, обробки та реакції на несправності кластера. На відміну від відомих моделей і підходів до моделювання, викладена концепція моделі дозволяє моделювати одиночні несправності та несправності будь-якої кратності для високопродуктивної обчислювальної системи. Модель дозволяє проаналізувати можливі наслідки роботи єдиного комплексу виявлення, обробки та реакції при появі несправності для виявлення слабких місць системи.

Ключові слова: мережі Петрі, виявлення несправностей кластера, суперкомп'ютер.

Вступ

Нині суперкомп'ютери поширені не лише у окремих галузях, що отримують добру фінансову підтримку (військові об'єкти, ядерні центри, спеціалізовані лабораторії) вони стали звичним явищем у комерції, промисловості та інших побутових сферах діяльності. Змінився і якісний склад суперкомп'ютерів. Аналіз, проведений в [1], показує, що в складі рейтингу Top500 найбільш продуктивних суперкомп'ютерів у період після 2000 року більше 70% складають кластери типу Beowulf, що працюють під керуванням ОС Linux. Для позначення кластерних систем нині вживають термін **високопродуктивні обчислювальні системи** (High Performance Computing, HPC) [1-3].

Після розробки й введення в експлуатацію високопродуктивної обчислювальної системи (ВПОС) з'являється проблема її обслуговування, супроводу програмного забезпечення, ремонту та вирішення інших поточних проблем. Серед них виділимо контроль технічного стану, оповіщення системного адміністратора в разі ознак відмов або несправностей, забезпечення систем захисту, керування та контролю в цілодобовому режимі. Зазвичай такі завдання вирішує єдиний комплекс виявлення, обробки та реакції на несправності кластера (ЄКВОР). Його програмною частиною є система аварійного оповіщення і відключення устаткування кластера (САОВ), яка забезпечує працездатність обладнання ВПОС й збереження результатів обчислень у разі аварійних ситуацій, цілодобово інформує системного адміністратора про виникнення цих ситуацій. Важливість ЄКВОР зумовлена тим, що за перший рік експлуатації обчислювальних центрів виходить із ладу до 10 % апаратної частини ВПОС.

Упродовж роботи датчики ЄКВОР приймають інформацію від апаратної частини ВПОС, програмна частина ЄКВОР обробляє цю інформацію. Результати обробки є основою для запуску системи захисту та інформування системного адміністратора. Найчастіше в якості ЄКВОР використовують вільно поширюване програмне забезпечення [3]. Наприклад, моніторинг ВПОС вирішують пакети ECMT, Big Sister (Big Brother), autocheck, Ganglia тощо. Однак вони мають певні недоліки: всі вони орієнтовані на вузьке коло завдань, вимагають наявності висококваліфікованого персоналу [3-6].

На прикладі проаналізуємо структуру ВПОС зі списку TOP50.RU [7, 8]. Результати аналізу наведено в таб. 1. Тут більшість ВПОС використовують вільно поширюване програмне забезпечення.

Досвід конструювання й експлуатації ВПОС показав, що у розробці ЄКВОР виникають труднощі, пов'язані з особливістю конструкції ЄКВОР. Серед них виділимо кілька причин, що ускладнюють процес створення і налагодження ЄКВОР і САОВ.

Першою причиною є велика кількість типів можливих відмов. На мал. 1 проілюстровано їхню малу частину, реальне число відмов ВПОС у кілька десятків разів більше. До другої причини належать високі вимоги, які висуваються до ЄКВОР і САОВ. Наприклад, передбачається в режимі реального часу за секунду знімати від 300 до 500 діагностичних параметрів з обчислювального вузла, кількість обчислювальних вузлів мо-

Таблиця 1. **Системи життєзабезпечення ВПВС**

Приналежність, назва	**Час розробки**	**Архітектура, продуктивність**	**Система моніторингу**	**Система безперебійного електроживлення, потужність, особливості**	**Система життє-забезпечення**
Московський Державний Університет шляхів сполучення МИИТ, Т-4700	2003	1. 64 двопроцесорних вузлів Discus з процесорами AMD Opteron™ (Barcelona), 2. пікова продуктивність 4,7 трильйонів операцій у секунду (Тфлопс). 3. Реальна продуктивність Linpack 3,89 Тфлопс (83% від пікової).	Вільно поширювані засоби керування та моніторингу	потужністю 60 Квт, рівень резервування N+1, автономна робота до 10 хвилин	15–20 хвилин після відмикання електро-живлення
Томський держ-університет, СКІФ CYBERIA	Лютий 2007 р.	1. 566 двоядерних процесорів Intel® Xeon® серії 5150, 2. пікова продуктивність 12 трильйонів операцій на секунду; 3. Реальна продуктивність на тесті Linpack 9.019 Тфлопс (75% від пікової)	Те саме	потужністю 160 квт, автономна робота до 10 хвилин	15–20 хвилин після відмикання електро-живлення
Науково-дослідний обчислю-вальний центр Московського держуніверситету, СКІФ МГУ	19 березня 2008 року	1. 1250 чотириядерних процесорів Intel® Xeon® E5472, 2. пікова продуктивність 60 трильйонів операцій у секунду (TFlops). 3. Реальна продуктивність системи на тесті Linpack - 47,17TFlops, (78,6% від пікової)	Програмно-апаратна система моніторингу розробки ІПС РАН.	п'ять модульних джерел безперебійного живлення Symmetra PX потужністю 80 квт із резервуванням N+1.	10 хвилин після відмикання електро-живлення

же досягати сотень тисяч. Інформацію слід обробляти в реальному часі. Третьою причиною є конструктивна складність ЄКВОР. Система – це розподілений програмно-апаратний комплекс, показники надійності якого аналогічні показникам ВПОС. До четвертої причини слід віднести складність обробки результатів вимірювань датчиків і прийняття взаємовиключних рішень, показаних на мал. 1 («Повідомлення ЄКВОР»). При цьому внаслідок відмови ЄКВОР рішення можуть бути як правильними, так і неправильними. Як видно з мал. 1, несанкціоноване спрацьовування несправної ЄКВОР призведе до фінансових втрат або непоправного пошкодження обладнання. П'ятою причиною є неприпустимість появи прихованих відмов у ЄКВОР і ВПОС. Практика показала, що одночасний вихід з ладу системи охолодження та своєчасне виявлення цієї відмови системою ЄКВОР (протягом 10 хвилин) призводить до виходу з ладу апаратної частини ВПОС.

Отже, розробку ефективної системи моніторингу ЄКВОР не можна здійснити інженерно-інтуїтивними методами. Потрібна наочна й ефективна математична модель ЄКВОР, що дозволить проаналізувати роботу ЄКВОР в більшості можливих технічних станах за різних видів відмов, показаних на мал. 1.

Математичні моделі побудови систем технічного діагностування і функціонального контролю складних систем

Класичні підходи [9-13] до моделювання ЄКВОР у вигляді таблиць функцій несправностей або моделювання одиночної несправності в дискретній системі не можна використати, оскільки неможлива декомпозиція ЄКВОР на окремі фізичні елементи із двома видами технічних станів (справне, несправне). Як показав досвід розробки й експлуатації в ЄКВОР і ВПОС, більшість фізичних елементів систем і підсистем може мати кілька видів несправностей. Наприклад, датчик температури в ЄКВОР може перебувати в справному стані та мати мінімум два види несправностей.

У [14] розроблено математичні моделі, що розглядають два види несправностей: коротке замикання й обрив для елемента підсистеми. Такі моделі також не підходять, оскільки ЄКВОР є програмно-апаратним комплексом, більшу його частину складає програмне забезпечення. У цьому випадку модель несправності має враховувати роботу програмної й апаратної частин.

У [15] розглянуто математичні моделі, які припускають наявність несправностей кількох видів і логічні взаємозв'язки між ними. Але використання цих моделей також неприйнятне,

Джерела відмов • • №
п/п • Підсистема • Конкурентні ресурси • Ознаки відмови • Тип відмови • • 1 • Система автоматичного газового пожежогасіння (АГПТ) • SNMP протокол • Сигнал EMS • Спрацювання АГПТ • • • • SNMP протокол • Сигнал EMS • Несправність АГПТ • • • • SNMP протокол • Сигнал EMS • Ручний режим • • 2 • Система контролю клімату (КК) • SNMP протокол • 5 датчиків t > 30 C^0, або 1 датчик t > 40 C^0 • Перевищення температури • • • • SNMP протокол • 3 датчика t > 40 C^0 • Перевищення температури • • • • SNMP протокол • 1 датчик t > 40 C^0 тривалий час • Передвідмовний стан датчика • • 3 • Система охолодження (ОВ) • SNMP протокол • SNMP-Trap • Несправність датчиків • • 4 • Підсистема електроживлення (ЭП) • SNMP протокол • SNMP-Trap • Зміна конфігурації ИБП • • • • SNMP протокол • SNMP-Trap • Відмова модуля ИБП • • • • SNMP протокол • SNMP-Trap • ИБП в обхідному режимі • • • • SNMP протокол • • ИБП робота від батареї • • 5 • САОО • SNMP протокол • Негативний результат перевірки • Загублення зв'язку КК • • • • SNMP протокол • Негативний результат перевірки • Загублення зв'язку з одним кондиціонером • • • • SNMP протокол • Негативний результат перевірки • Загублення зв'язку з двома кондиціонерами • SNMP протокол • Негативний результат перевірки • Загублення зв'язку з ИБП • •

1.2 Підсумки роботи САОО • • Стан ЄКВОР • Стан ВПОС • Повідомлення САОВ • Результаты
експлуатації, дії персоналу • • 1 • 2 • 3 • 4 • • Працездатна • Працездатна • Коректна робота • Нормальна експлуатація • • • • Інформаційна подія • Нормальна експлуатація • • • • Важлива подія • Оповіщення персоналу • • • Несправна • Небезпечна подія • Оповіщення персоналу, штатне відключення обладнання • • • • Критична подія • Оповіщення персоналу, аварійне відключення обладнання • • Несправна • Працездатна • Небезпечна подія • Відключення обладнання, штрафні санкції • • • • Критична подія • • • • Несправна • Коректна робота • Пошкодження обладнання • • • • Інформаційна подія • • • • • Важлива подія • • •

Мал. 1. Результати роботи системи ЄКВОР

оскільки в [15] подано коректну постановку завдання й недостатньо розкрито шляхи її рішення. Це не дозволяє враховувати використання загальних комунікаційних ресурсів, особливо конфлікт доступу до обмеженого ресурсу або моделювання процесу очікування одним компонентом результатів роботи іншого. Для рішення цих завдань доцільніші моделі на основі мереж Петрі [16], але у відомих роботах цього напрямку система контролю та захисту є ідеалізованою, безвідмовною. Як показав досвід експлуатації й розробки, надійна система ЄКВОР може відмовляти (мал. 1, стовпець «Підсистеми»). Отже, необхідно розробити нову модель, яка б враховувала відмови ЄКВОР і вирішувала вищевикладені завдання.

Прототип моделі ЄКВОР

Результати аналізу діагностичних моделей дають змогу подати прототип моделі ЄКВОР у вигляді мережі Петрі, як це відбувається при вирішенні завдань контролю технічного стану складних об'єктів. Визначимо нову концепцію моделювання. Мережа Петрі є впорядкованою множиною, що включає чотири елементи:

$$C=(P, T, I, O), \quad (1)$$

де

$$P=\{p_1, p_2, p_3, p_4, \ldots, p_n\} \quad (2)$$

є кінцевою множиною позицій $n \geq 0$, а

$$T=\{t_1, t_2, t_3, t_4, \ldots, t_m\} \quad (3)$$

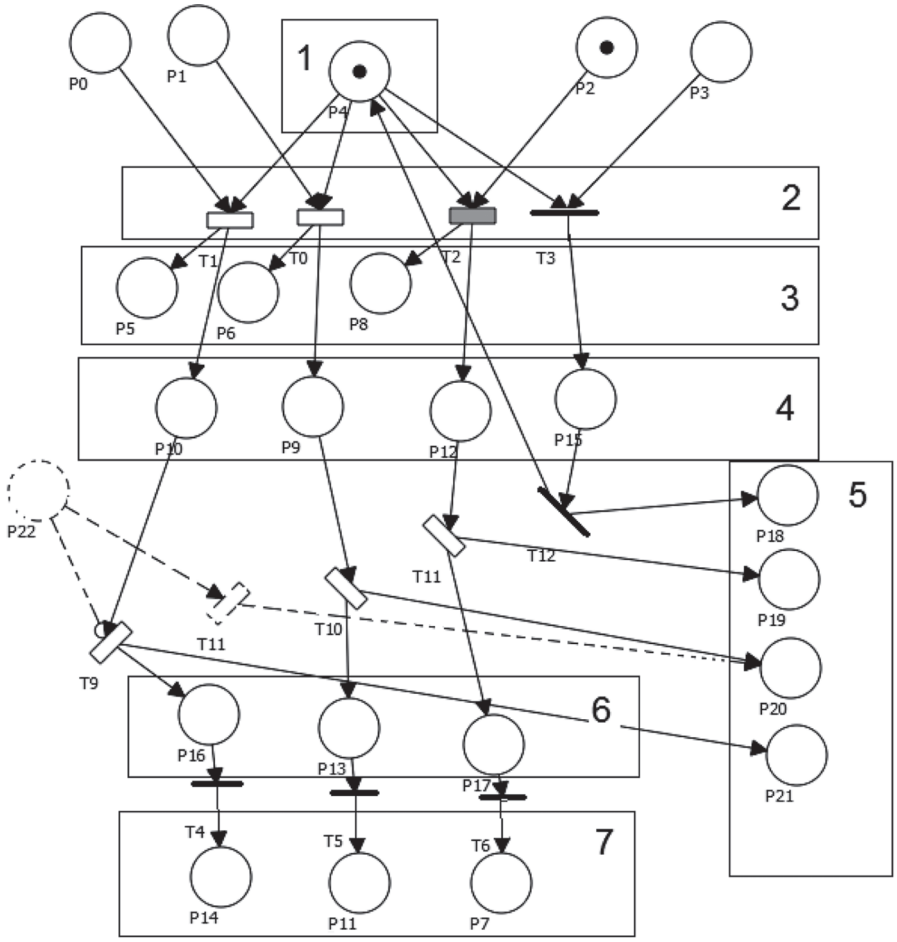
є множиною переходів $m \geq 0$. Причому множини позицій P і T не перетинаються. Вхідна функція $I:T \rightarrow P^\infty$ є вхідною функцією – відображенням переходів у комплекти позицій. Вихідна функція $C:T \rightarrow P^\infty$ є відображенням переходів у комплекти позицій [17].

З використанням математичної моделі (1) потрібно вирішити наступні завдання:

– формалізувати технічні стани ЄКВОР (мал. 1 стовпчик 1) і ВПОС (мал. 1 стовпчик 2) так, щоб дати можливість показати перехід ВПОС у непрацездатні стани;
– показати, як ЄКВОР відображає у своїх повідомленнях стан ВПОС і активізує систему захисту (мал. 1 стовпчик 3);
– проаналізувати можливі наслідки відпрацьовування ЄКВОР або ВПОС (мал. 1, стовпчик 4) у разі приходу керівного сигнал у , щоб не допустити небажані наслідки експлуатації.

З огляду на досвід експлуатації та розробки ЄКВОР, на вищевикладені завдання для реалізації моделлю, розкриємо основні елементи моделі та її концепцію:

– відображати технічні стани ВПОС і ЄКВОР у вигляді елементів множини (2) і взаємозв'язок між ними;



Мал. 2. Математична модель елемента підсистеми

– показати процес формування сигналів, що інформують, сигналів, що управляють у вигляді елементів множини (2);
– забезпечити адекватне моделювання типових і потенційних відмов у ВПОС і ЄКВОР у вигляді тривалих переходів [17, 18] (елементів множини (3)) або додаткових місць (елементів множини (2));
– розкрити процес формування інформаційних повідомлень ЄКВОР і керівних сигналів ВПОС у випадку виникнення несправності;
– формалізувати результати реагування ВПОС і ЄКВОР у випадку надходження інформаційного сигналу у вигляді елементів множини (2);
– показати потенційні результати відмов і реагування на них ВПОС і ЄКВОР у вигляді елементів множини (2).

Використовуючи концепцію моделі, відобразимо елементи ВПОС і ЄКВОР, що беруть участь у процесі контролю технічного стану і захисту у вигляді орієнтованого мультиграфа мережі Петрі G (мал. 2),

$$G = (V, A), \qquad (4)$$

де $V = \{u_1, u_2, u_3, \ldots, u_s\}$ – множина вершин, які визначають стан ВПОС, ЄКВОР. Множина $A = \{a_1, a_2, a_3, \ldots, a_t\}$ є комплектом спрямованих дуг, які визначають взаємозв'язки між станами ВПОС і ЄКВОР, $a_i = (v_j, v_k)$, де $v_j, v_k \in V$ [17].

Множину V розбито на дві неперетинні підмножини, де P визначається множиною (2) і T визначається множиною (3). Для будь-якої спрямованої дуги $a_i \in A$, якщо $a_i = (v_j, v_k)$, тоді або $v_j \in P$ *і* $v_k \in T$, або $v_j \in T$, $v_k \in P$ (мал. 2).

Щодо аналізу розробки й практичної експлуатації ВПОС повідомлення ЄКВОР доцільно розділити на 4 основних типи, відображені на мал. 1 (стовпець 3). У математичній моделі (4) ці повідомлення відображено в блоці 5 на мал. 2. Позиція *P18* позначає відображення інформаційної події, позиція *P19* позначає відображення важливої події, позиція *P20* – небезпечної події, позиція *P21* – критичної події.

Блок 1 мал. 2 показує початковий стан системи. Зважаючи на те, що у вихідному стані система справна та працездатна, місце *P4* має одну мітку.

Блок 2 мал. 2 призначено для моделювання відмов у системі ВПОС і містить у собі прості та тривалі переходи. Тривалі переходи *T0-T2* дозволяють моделювати стани ВПОС перед відмовою, відмови та несправні стани ВПОС.

Прості переходи внаслідок миттєвого спрацьовування *T3*, *T12* дозволяють моделювати інформаційну подію.

Блок 3 мал. 2 містить місця *P5-P8*, що моделюють несправні стани ВПОС (критичного, небезпечного й важливого станів).

Блок 4 мал. 2 містить місця *P10, P9, P12, P15*, що моделюють справні стани ЄКВОР, які виявляють критичні, небезпечні, важливі й інформаційні події відповідно. Тривалі переходи *T9-T12* відображають коректну роботу ЄКВОР. Переходом *T11* (пунктир) показаний спосіб моделювання несправності в системі ЄКВОР.

Блок 6 мал. 2 містить місця, що моделюють результати роботи системи захисту *P16, P13, P17*, аварійне відключення ВПОС, штатне відключення ВПОС, вимоги до виконання регламентних робіт відповідно.

Для адекватного відображення процесу функціонування ЄКВОР й ВПОС і адекватного моделювання наслідків спрацьовування системи захисту необхідно зробити розмітку мережі. У цьому випадку маркуванням μ мережі (1) будемо вважати функцію, що відображає множину позицій (2) у множину натуральних чисел *N* [17, 18]

$$\mu: P \rightarrow N$$

або у векторному вигляді

$$\mu = (\mu_1, \mu_2, \mu_3, \ldots, \mu_n).$$

З огляду на маркування подамо нашу модель як марковану мережу

$$M = (3, \mu). \qquad (5)$$

Введемо функцію $\delta(\mu, t_i)$, названу функцією наступного стану мережі (5), і початкове маркування μ^0. Початкове маркування, виходячи з концепції моделі, визначає наявність розмітки в місці *P4*, формалізуючи вихідний, працездатний стан ВПОС. Крім того, для аналізу моделі з метою вирішення завдань, викладених вище, необхідно змоделювати несправність, увівши нову вихідну розмітку з несправністю. За функцією $\delta(\mu, t_i)$ одержимо послідовність розміток або спрацьовувань переходів для одержання результатів моделювання.

Наприклад, для однократної несправності ВПОС застосовують одну мітку на одне з місць *P2* (мал. 2), дозволяючи у ході виконання мережі (5) змоделювати несправність у ВПОС. Для моделювання несправності в ЄКВОР у модель додано місце, що має мітку, як показано на мал. 2, *P22* (штрих). Вищесказане дозволяє сформувати множину вихідних розміток для моделювання несправностей, яку позначимо

$$\mu^{0d} = (\mu^{0d}{}_1, \mu^{0d}{}_2, \mu^{0d}{}_3, \ldots, \mu^{0d}{}_k), \qquad (6)$$

де елементом множини є початкова розмітка мережі з несправністю. Для кожного елемента множини (6) можна провести дослідження з метою одержання всіх можливих послідовностей маркувань і послідовностей переходів, які можуть бути запущені. Наприклад, для маркування $\mu^{0d}{}_1$ можна одержати послідовність переходів і послідовність маркувань, що позначимо у вигляді множини

$$\mu/\mu^{0d}{}_1 = (\mu_0, \mu_1, \mu_2, \ldots). \qquad (7)$$

Множина (7) дозволяє побудувати дерево досяжності для рішення завдань аналізу й комп'ютерного моделювання ЄКВОР і ВПОС.

Висновки

У роботі побудовано математичну модель, що дозволяє формалізувати технічні стани ЄКВОР і ВПОС для аналізу можливих наслідків роботи системи захисту з метою запобігання небажаних наслідків експлуатації і для побудови системи аварійного оповіщення та відключення обчислювального кластера.

На відміну від відомих моделей і підходів до моделювання, викладена концепція моделі дозволяє моделювати одиночні несправності та несправності будь-якої кратності як самого об'єкта діагностування (ВПОС), так і системи для контролю технічного стану (ЄКВОР). Модель дозволяє формалізувати відмови в програмній та апаратній частинах об'єкта діагностування. Модель легко адаптується для автоматизації та розробки програмного забезпечення.

Незважаючи на всі ці переваги, для ефективного використання моделі необхідні подальші дослідження й розвиток у наступних напрямках:

– розробити моделі процесу формування сигналів ЄКВОР у вигляді мереж Петрі;
– зібрати статистичну інформацію й описати найчастіші типові відмови ВПОС і ЄКВОР, виявити найнебезпечніші відмови у ВПОС і ЄКВОР;
– описати вищезгадані відмови в теорії мереж Петрі для формування моделей тривалих переходів (елементів множини (3)) або додаткових місць (елементів множини (2));
– описати процеси формування інформаційних повідомлень ЄКВОР у вигляді мереж Петрі для відповідного програмного й апаратного забезпечення;
– описати процеси реагування ВПОС і ЄКВОР у випадку приходу інформаційного сигналу у вигляді мереж Петрі.

1. Черняк Л. Суперкомпьютинг вглубь и вширь [Электронный ресурс] / Леонид Черняк // Открытые системы. – 2007. – № 9. – Режим доступа : http://www.osp.ru/os/2007/09.htm. – Название с экрана.
2. Андреев А. А. Кластеры и суперкомпьютеры – близнецы или братья? [Электронный ресурс] / Андреев А. А., Воеводин Вл. В., Жуматий С. А. // Открытые системы. – 2000. – № 5–6. – Режим доступа: http://www.osp.ru/os/2000/05-06/009.htm. – Название с экрана.
3. Жуматий С. А. Комплекс мониторинга распределённых информационно-вычислительных систем / Жуматий С. А., Кальянов А. А. // Научный сервис в сети Интернет. Труды всероссийской научной конференции. – М. : Изд-во МГУ, 2002. – С. 47.
4. Воеводин В. В. Параллельные вычисления / В. В. Воеводин, Вл. В. Воеводин. – СПб. : БХВ-Петербург, 2002. – 608 с. – ISBN 5-94157-160-7.
5. Николаев А. Автоматизация процессов эксплуатации ИТ / Николаев А., Кузубов С., Тукмаков Р., Ланцова Л. // Jet info. Информационный бюллетень. – 2008. – № 10 (185).
6. Карасев В. Там, где живут серверы / Вячеслав Карасев // Jet info. Информационный бюллетень. – 2008. – № 4 (179).
7. Аветисян А. И. Архитектура и системное программное обеспечение вычислительных кластерных систем / Аветисян А. И., Самоваров О. И., Грушин Д. А. // Высокопроизводительные параллельные вычисления на кластерных системах. Материалы пятого Международного научно-практического семинара / Под ред. проф. Р. Г. Стронгина. – Нижний Новгород : Изд-во Нижегородского госуниверситета, 2005. – 253 с.
8. Пархоменко П. П. Основы технической диагностики. Оптимизация алгоритмов диагностирования, аппаратные средства / Пархоменко П. П., Сагомонян Е. С. – М. : Энергоиздат, 1981. – 320 с.
9. Бережной В. П. Выявление причин отказов РЭА / В. П. Бережной, Л. Г. Дубицкий – М. : Радио и связь , 1983. – 232 с.
10. Ярмольник В. Н. Контроль и диагностика цифровых узлов ЭВМ. – М. : Наука и техника, 1988. – 240 с.
11. Гольдаман Р. С.Техническая диагностика цифровых устройств / Гольдаман Р. С., Чипулис В. П. – М. : Энергия, 1976. – 224 с.
12. Сагунов В. И. Контролепригодность структурно-связанных систем / Сагунов В. И., Ломакина Л. С. – М. : Энергаатомиздат, 1990. – 112 с.
13. Ксёнз С. П. Диагностика и ремонтопригодность радиоэлектронных средств / Ксёнз С. П. – М. : Радио и связь, 1989. – 248 с.
14. Катін П. Ю. Функціонально-статистична модель аналогового пристрою // Зб. наук. пр. ВІТІ НТУУ «КПІ». – № 1. – К. : ВІТІ НТУУ «КПІ», 2004. – С. 49–62.
15. Murata T. Petri Nets: Properties, Analysis, and Applications / T. Murata. – Proceedings of the IEEE. – 1989. –Vol. 77. – No. 4. – PP. 541–580.
16. Питерсон Дж. Теория сетей Петри и моделирование систем / Дж. Питерсон. – М. : Мир, 1984. – 264 с.
17. Котов В. Е. Сети Петри / Котов В. Е. – М. : Наука, 1984. – 160 с.

D. Fedyukov, S. Riabchun

THE HELPER MATHEMATICAL MODEL TO BUILD HARDWARE FAILURE DETECTOR FORHIGH PERFORMANCE COMPUTING SYSTEMS

The article describes the Petri Nets-based mathematical model that used as theoretical basis to build the uniform solution for problem detection, issues processing and system response (PD/IP/SS). In article it is shown that having such model is critical to built working solution, since it cannot be carried out by engineering-intuitive methods. Unlike known models and approaches to the modelling, the stated concept of model introduces modelling of faults of any multiplicity (including single faults) for both HPC system and PD/IP/SS itself. The model allows to provide the analysis of possible consequences of PD/IP/SS operation for revealing system's weaknesses.

Keywords: supercomputer, cluster system, Petri Nets, mathematical model, failure, denial, alarm message, monitoring.