

БОЙКО Ю.В.,  
ГЛИБОВЕЦЬ М.М.,  
СРШОВ С.В.,  
КРИВИЙ С.Л.,  
ПОГОРІЛИЙ С.Д.,  
РОЛІК О.І.,  
ТЕЛЕНИК С.Ф.,  
ЯСОЧКА М.В.

## УПРАВЛІННЯ ВИСОКОПРОДУКТИВНИМИ ІТ-ІНФРАСТРУКТУРАМИ

Наведений аналітичний огляд сучасних моделей і методів управління ІТ-інфраструктурою. Моделі і методи згруповані у відповідності з визнаними ІТ-галуззю 15 групами функцій управління. Основна увага приділяється концептуальним основам управління такими складними ієрархічними розподіленими об'єктами управління, які сьогодні становлять глобальні, національні і корпоративні ІТ-інфраструктури. Найбільш широко представлені моделі розподілу ресурсів і навантаження та управління рівнем обслуговування користувачів.

An analytical survey of the modern IT-infrastructure management models and methods is presented. This models and methods are grouped in accordance with approved by IT-community 15 groups of management functions. The main attention is given to management basics of global, national and corporative IT-infrastructure — complex hierarchical distributed object of management. The models of resources and load allocation and users services level management are widely presented.

### Введення

Глобалізація світової економіки, формування соціально-орієнтованого середовища обумовлює новий рівень вимог до інформаційних технологій (ІТ). Бізнес, наукові дослідження і розроблення нових товарів і технологій, підтримка державного управління і комунального господарства вимагають вирішення все складніших проблем. Потрібні значні обчислювальні і телекомунікаційні ресурси, системи накопичення і зберігання великих обсягів даних [1, 2, 3, 4].

Вирішення проблем підтримки власних ІТ-інфраструктур відволікало підприємства, компанії, органи управління від основної діяльності. ІТ-інфраструктури створювалися в розрахунок на пікові навантаження і загалом використовувалися на рівні 15–20% від їх можливостей.

Реакція ІТ-галузі на цю ситуацію полягала у концентрації ресурсів у центрах оброблення даних (ЦОД) і їх наданні підприємствам, компаніям, органам управління у вигляді сервісу на потребу. Організаційно-технічно ці тенденції реалізовані за допомогою хмарних або Grid-технологій. Хмарні сервіси — Infrastructure as a Service (IaaS), Platform as a Service (PaaS) і Software as a Service (SaaS) — поступово змінюють уявлення про роль ІТ у бізнесі і власне ІТ як бізнес. Провайдери хмарних сервісів, насамперед Salesforce.com, Amazon, Microsoft,

Oracle, SAP, Softlayer, Terremark, Rackspace, NetSUIT, Rackspace постійно збільшують дохід від надання сервісів, знижуючи ціни.

Створення, підтримка і використання ІТ-інфраструктури для розв'язання наукових задач, базування автоматизованих інформаційних систем (АІС) породжує складні науково-практичні проблеми. Колектив авторів плідно працює над вирішенням зазначених проблем. У статті розглядаються проблеми управління високопродуктивною ІТ-інфраструктурою.

### Сучасні тенденції розвитку ІТ-галузі

Шукаючи шляхи вирішення проблем, провайдери інформаційно-комунікаційних сервісів (ІКС) удосконалюють діяльність з метою збереження достатнього рівня її рентабельності в умовах зростання конкуренції і жорстких вимог замовників [5]. До основних тенденцій розвитку ІТ-галузі, пов'язаних з темою статті, належать:

1. *Втілення ідеї глобалізації у розвитку корпоративних, національних і глобальних ІТ-інфраструктур* — організованих сукупностей взаємопов'язаних мереж, ІТ, інформаційних ресурсів і обладнання кінцевих користувачів та їх оточення, як організованої сукупності інформаційних застосувань, застосувань користувачів та інформаційних послуг, взаємодія яких

забезпечує підтримку процесів збору, оброблення і збереження інформації користувачів.

2. *Поява нових видів провайдерів* (Інтернет, контенту, сервісів за моделями IaaS, PaaS, SaaS) і формування класу конвергентних провайдерів.

3. *Зосередження провайдерів на проблемах користувачів і інтеграція ресурсів ЦОД* для надання їм сервісів для підтримки діяльності.

4. *Поява хмарних обчислень — Cloud Computing*, які забезпечують оптимізацію ресурсів ЦОД і ІТ-середовища в цілому.

5. *Віртуалізація платформ і ресурсів ІТ-інфраструктури*, як засіб для спрощення управління групами ресурсів і доступу до них.

6. *Поява мережоцентричних архітектур* як бази електронної комерції і поширення сервіс-орієнтованої і WEB-орієнтованої архітектур.

7. *Поширення компонентно-базованого підходу до створення АІС* шляхом комплексування готових компонентів єдиного ІТ-середовища.

8. *Розвиток розподілених систем збереження даних (СЗД)* об'єднанням сховищ різної архітектури з ІР-взаємодією з застосуваннями.

9. *Широке впровадження автоматизації і поява систем управління ІТ-інфраструктурою (СУІ)* надання сервісів внаслідок її ускладнення.

Конвергентними стають сервіси і провайдери. Мережі провайдерів, корпорацій, комунальні ІТ-системи інтегрувалися, утворивши єдине ІТ-середовище. ІТ-інфраструктура, як його основа, перетворилася у важливий об'єкт управління. Використання ІКС від провайдерів стимулюється економічними чинниками. Відносно користувача і провайдера регламентує SLA. Гарантований рівень обслуговування забезпечує провайдерів від збитків і сприяє залученню користувачів. Провайдерів ІКС приділяють велику увагу його забезпеченню, намагаючись скоротити витрати на надання послуг.

Системи управління провайдерів будуються на основі процесно-функціональної, структурно-організаційної і ресурсної концепції управління. В них виділяють business (BSS) і operations (OSS) support systems, які формують уніфіковані електронні образи бізнесу і ІТ-інфраструктури провайдера, забезпечуючи аналіз стану, генерування, оцінювання і вибір рішень. Їх будова і функціонування базуються на виділенні ІТ-інфраструктури як об'єкта управління, процесів, які задіюють її і підтримують діяльність як основу бізнесу, і організації, яка перетворює потенціал в реальні результати [6,7].

Розраховане на пікове навантаження традиційне горизонтальне структурування ІТ-се-

редовищ підприємств виявилось не гнучким і не адаптованим до змін у діяльності, мало вартість створення, підтримки і управління. Консолідована ІТ-інфраструктура об'єднує всі ресурси в єдиний пул для підтримки діяльності всіх бізнес-підрозділів [8]. Консолідація базується на технологіях віртуалізації платформ шляхом створення віртуальних машин (ВМ), ідеї якої переносять на мережі, СЗД та ін. [9].

Впровадження процесного підходу на основі стандартів передбачає реалізацію трьох груп функцій управління — підприємством, розвитком і операційною діяльністю. Перша група традиційна для всіх підприємств. Специфіку діяльності провайдерів ІКС визначають друга і третя групи, спрямовані на клієнтів, постачальників і партнерів. Процеси другої групи спрямовані на створення середовища функціонування, яке забезпечує підтримку і автоматизацію операційних процесів, визначаючи напрями розвитку ІТ-інфраструктури і продуктів. Їх підтримують підгрупи процесів розроблення стратегії розвитку провайдера; управління життєвим циклом (ЖЦ) інфраструктури і продукту. Операційні процеси спрямовані на роботу з клієнтами, їх потребами, вимогами, оцінками і історією співпраці, утворюючи вертикальні наскрізні угруповання процесів підтримки операцій і забезпечення готовності, виконання, управління якістю обслуговування і білінгу.

### Аналіз існуючих рішень

Основні постачальники ІТ вийшли на ринок з рішеннями, призначеними для управління ІТ-інфраструктурою. Цей сегмент ринку насичений продуктами, створеними на основі ІТ Infrastructure Library (ITIL) і концепції ІТ Service Management (ITSM). У процесній моделі ITIL виділяють 3 групи процесів управління: *управління інфраструктурою* (проектування і планування, розширення, супроводження і технічна підтримка); *підтримка обслуговування* (управління інцидентами і проблемами, конфігурацією і змінами, релізами); *управління наданням послуг* (рівнем обслуговування, фінансами, ІТ-послугами, готовністю, неперервністю обслуговування, потужностями).

Повнофункціональні рішення Microsoft, IBM, HP, EMC Symantec і Verland надають адміністраторам широкі можливості з управління ІТ-інфраструктурою ЦОД. Забезпечуючи створення єдиного інформаційного середовища, централізацію процесів розгортання, експлуатації і моніторингу ІТ-інфраструктури, вони

оцінюють вплив змін в застосуваннях, конфігураціях і топології на бізнес-процеси, підтримують роботи поширених платформ.

Продукти інших виробників реалізують базову функціональність. Крім того, деякі розробники систем ERP, наприклад, PeopleSoft і SAP включили до складу своїх основних продуктів компоненти управління ІТ-послугами.

Загалом рішення орієнтовані на підвищення гнучкості ІТ-інфраструктури і оптимізацію використання ресурсів. Але їх ефективність суттєво залежить від моделей розподілу ресурсів, особливо в умовах ресурсних обмежень. На жаль, у відкритих джерелах мало інформації про моделі і методи, покладені в основу реалізації наведених вище програмних продуктів. Окрім того, розглянуті рішення переважно орієнтуються на корпоративні комерційні структури і не враховують оперативні вимоги. Розглянуті компоненти не надають засобів оцінки показників оперативної діяльності та планування, оптимізації та управління АІС користувачів з урахуванням відповідних критеріїв і обмежень. Суттєвими недоліками згаданих продуктів є використання закритих фірмових технологій, моделей і алгоритмів управління. Висока вартість цих продуктів та орієнтованість на крупні підприємства з великим парком обчислювальної техніки не дозволяє розглядати їх якості СУІ для невеликих підприємств.

Отже, обґрунтоване вирішення всього комплексу задач управління ІТ-інфраструктурою відсутнє і виникає потреба у недорогій СУІ, яка відрізняється простотою адаптації та використанням ефективних моделей і методів.

### **Проблеми управління високопродуктивною ІТ-інфраструктурою**

Втілення потенційних переваг вимагає реалізації всіх груп функцій процесної моделі ІТІЛ. Це робить необхідним вирішення ряду проблем ефективного планування і управління ІТ-інфраструктурою і діяльністю провайдерів ІКС.

Загальні проблеми, викликані консолідацією: визначення напрямів розвитку ІТ-інфраструктури; забезпечення відмовостійкості її компонентів; побудова високопродуктивної ІТ-інфраструктури; безпека; взаємодія компонентів різних виробників; підтримка діяльності ІТ-служб.

Реалізація сервіс-орієнтованого підходу вимагає вирішення таких проблем: відстеження галузевих тенденцій; аналіз ринку; оцінювання перспективності і конкурентоздатності діяльності провайдера; вчасне і виважене введення но-

вих продуктів і сервісів; адресування серверів; взаємодія сервісів; маршрутизація доступу до сервісів; реєстрація сервісів та їх представлення користувачам; визначення раціональних параметрів SLA; реєстрація ресурсів і їх представлення користувачам; забезпечення визначеного QoS, безпеки і довіри; надання здатності до самоуправління і самоорганізації.

Орієнтація на проблеми користувачів вимагає вирішення проблем реалізації компонентно-базованого підходу до створення АІС компаній-користувачів на основі комплексування сервісів, управління ЖЦ сервісів, реалізації міжсервісної взаємодії з урахуванням SLA на основі політик, забезпечення рівня обслуговування, визначеного з потреб бізнесу, забезпечення необхідних для надання ІКС ресурсів і управління ними та ін. [10, 11].

Забезпечення надання сервісів незалежно від технологій транспортного і ресурсного рівнів вимагає вирішення таких проблем: управління ресурсами; оптимізація мережного трафіку; інформованість користувачів про ІКС і їх зв'язок з бізнес-процесами; вибір сервісів і їх комплексів для підтримки діяльності користувачами; управління безпекою на основі динамічно змінюваних політик; надання системам управління провайдерів і користувачів здатності до самоуправління і самоорганізації.

Розглянемо найважливіші проблеми управління ІТ-інфраструктурою, виявляючи чинники впливу, критерії, обмеження та інші аспекти.

*Проблема ресурсів.* Ресурси ІТ-інфраструктури вимагають обліку і аналізу відповідності вимогам. Гнучкі рішення на основі балансування навантаження і розподілу ресурсів, є важелем підтримки рівня надання ІКС.

*Проблема управління рівнем обслуговування користувачів.* Для її вирішення необхідно об'єктивно і своєчасно визначати причину погіршення параметрів сигналу, аналізувати зміни параметрів якості, передбачати завантаженість мережі, популярність того чи іншого контенту, і забезпечувати, за потреби, виділення додаткових ресурсів на надання сервісу. Визначальним для управління ІТ-інфраструктурою є забезпечення рівня обслуговування користувачів на основі раціонального співвідношення ціна/якість. Базою такого управління є моніторинг, оцінювання і контроль параметрів якості на всьому шляху контенту до користувача в розрізі мережі, медіапотоків, контенту та управління. Ядром такого управління є імпакт-аналіз.

*Проблема управління потоками даних.* Для управління мережними ресурсами ІТ-інфраструктури можна використовувати пріоритизацію пакетів, оптимізацію джитера і затримок, зменшення втрат пакетів і виділення смуги. Але управління потоками мережі з огляду на інтеграцію сервісного і процесного підходів та тенденції розвитку ІТ-галузі видається найбільш перспективним. Для забезпечення потрібної градації якості послуг проблема декомпозується на підпроблеми для окремих мереж, технологій і обладнання з врахуванням методів стиснення, роздільних здатностей та інших параметрів.

*Проблема опрацювання метрик.* Багаторівневість представлень ІТ-інфраструктури, наявність широкої номенклатури технологій, обладнання, ресурсів і компонентів обумовили наявність величезної кількості різноманітних метрик їх оцінювання. Тому виникає проблема впорядкування метрик, їх систематизації, зведення часткових метрик в інтегральні.

Вирішення цих проблем ускладнюють велика кількість користувачів, складність ІТ-інфраструктури, велика кількість типів і одиниць обладнання та інші чинники. Потрібні ефективні математичні моделі і методи.

### Концептуальні основи управління ІТ-інфраструктурою

Орієнтована на створення єдиного універсального середовища інтегрованого управління різноманітними ІТ, розподіленими застосуваннями, ресурсами і мережним обладнанням концепція управління ІТ-інфраструктурою враховує усі аспекти цього надскладного об'єкта управління. Її принципові положення [12]:

1. Включення контуру оперативного управління ІТ-інфраструктурою в зовнішній контур управління бізнесом і внутрішній контур управління операційною діяльністю.

2. Виділення в ІТ-інфраструктурі ієрархічних рівнів (сервісів, застосувань, обчислювальних ресурсів, мережної взаємодії) і розподіл функцій, команд і методів управління між ними.

3. Виділення категорій управління (наданням і продуктивністю сервісів, сервісами і процесами, підтримка сервісів і управління ресурсами) і загальних процесів (моніторинг, аналіз, управління, оптимізація, планування).

4. Використання декомпозиційно-компенсаційного способу організації управління рівнем послуг, заснованого на декомпозиції задач

управління і компенсації негативного впливу окремих чинників за рахунок виділення додаткових ресурсів критичним застосуванням [13].

5. Комплексне управління ІТ-інфраструктурою, інтеграція процесного і оперативного управління з організацією міжрівневої взаємодії на основі базової моделі дворівневої системи управління з координатором.

6. Врахування вимог бізнес-процесів.

7. Використання сервісно-ресурсних моделей (СРМ).

8. Впровадження процесно-оперативного підходу до управління рівнем ІТ-послуг.

9. Визначення універсального об'єкта моніторингу та управління (ОМУ) і робота СУІ не з реальними (РОУ), а логічними (ЛОУ) ОМУ.

10. Використання агентського підходу.

11. Застосування єдиного механізму взаємодії СУІ з агентами і ОМУ.

12. Використання універсальної схеми кодування стану ОМУ.

*Декомпозиційно-компенсаційний підхід до управління рівнем ІТ-послуг.* Розвиток бізнесу за рахунок множини  $S = \{s_i\}$ ,  $i = \overline{1, K}$  необхідних ІТ-послуг вимагає обґрунтованого співвідношення їх рівня якості  $Q$  і витрат  $C$ . Необхідність розкладання сукупної поведінкової моделі складної системи обумовлює декомпозицію управління [13]. Загальна схема підходу:

*Етап 1.* Комплексний аналіз співвідношення SLA/вартість (ціна рівня).

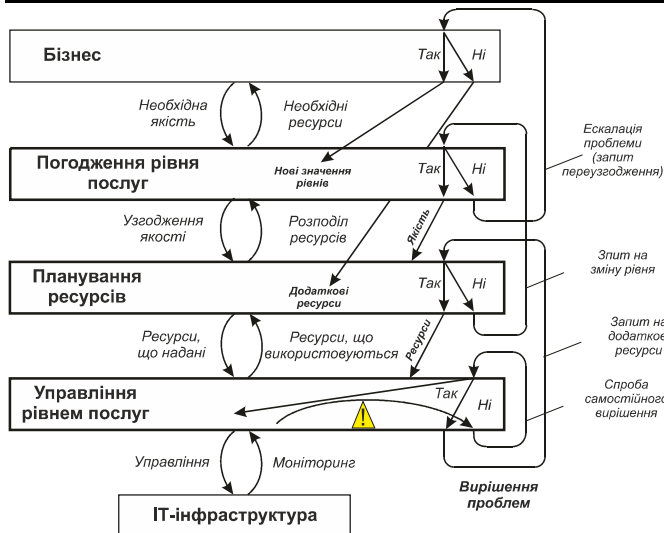
*Етап 2.* Визначення і обґрунтування інтегрального рівня SLA на основі можливих втрат провайдера внаслідок незадовільного рівня обслуговування і його витрат на надання сервісів з визначеним рівнем з урахуванням ризиків.

*Етап 3.* Аналіз залежності інтегрального показника SLA від базових.

*Етапи 4 — 7.* Декомпозиція проблеми визначення мінімальної вартості досягнення рівня обслуговування на підпроблеми компенсаційного визначення цієї вартості на підмережах, сервісах, ресурсах і технологіях.

*Етап 8.* Облік і контроль дотримання визначених параметрів.

Управління рівнем послуг вимагає інтегрованої взаємодії процесів погодження та управління рівнем послуг і планування ресурсів (рис.1).



**Рис. 1. Взаємодія процесів при управлінні рівнем послуг**

Процес погодження рівня послуг формує множину  $S$  і матрицю  $Q = \|q_{ki}\|$ , де  $q_{ki}$ ,  $k = \overline{1, M_i}$ ,  $i = \overline{1, K}$  — значення  $k$ -го показника якості  $i$ -ї послуги. Пара  $\langle Q, r \rangle$  визначається системою  $\langle \hat{D}, \hat{C}, \hat{L}, Q, M_0, N_0 \rangle$ , де  $\hat{D}$  — доходи бізнесу від ІТ,  $\hat{C}$  — витрати на ІТ-інфраструктуру,  $\hat{L}$  — втрати бізнесу від неякісних сервісів,  $Q = f_1(c)$  — витрати на підтримку якості сервісів,  $M_0$  — ризики і  $N_0$  — невизначеність. На основі пари  $\langle Q, r \rangle$ , де  $r$  — виділений послугам  $S$  ресурс, розв'язуються задачі на наступному рівні. Авторами запропоновані моделі, спрямовані на досягнення  $\max(\hat{D} - \hat{C})$ .

Процес планування формує матрицю  $P = \|\rho_{ij}\|$ , де  $\rho_{ij}$  — кількість виділеного послугі  $s_i$ ,  $i = \overline{1, K}$  ресурсу  $R_j$ ,  $j = 1, \dots, m$ . За дефіциту ресурсів підтримуються найважливіші застосування з врахуванням стану ІТ-інфраструктури і зміни інтенсивності запитів. Задача погодження рівня послуг зводиться до визначення матриці  $Q = F_1(S, r, W_p, Z_s)$ , де  $W_p = \{w_i | i = \overline{1, K}\}$ ;  $w_i$  — важливість послуги  $s_i$ ;  $Z_s = \{z_i | i = \overline{1, K}\}$ ;  $z_i$  — запланований ступінь підтримки  $i$ -ї послуги,  $r$  — сумарна кількість ресурсів.

За надлишку ресурсів забезпечують обумовлені значення показників якості  $q_{ki}$  застосувань із множини  $A = \{A_l\}$ ,  $l = \overline{1, L}$ , що підтримують послуги, за критерієм максимального вивільнення ресурсів з врахуванням їх надійності. Визначається якість  $Q = F_2(S, \hat{r}, r_r, C)$  при відомих ресурсах (задача першого роду), або ресурси  $r = F_3(S, Q, C)$ , потрібні для забезпечення

заданої якості (задача другого роду). Тут  $r_r$  — резервний обсяг ресурсів,  $\hat{r} = r + r_r$ . Авторами запропоновані відповідні моделі і методи.

Процес управління рівнем послуг підтримує відповідність фактичних значень  $q_{ki}^*$ ,  $k = \overline{1, M_i}$ ,  $i = \overline{1, K}$  показників якості послуг погодженим

$$q_{ki} - q_{ki}^* = 0, \quad \forall k = \overline{1, M_i}, i = \overline{1, K}. \quad (1)$$

Визначаються операції (зміна параметрів функціонування ІТ-інфраструктури, виділення додаткових чи вимикання незадіяних ресурсів, перерозподіл ресурсів між користувачами, ескалація проблеми на рівень планування), які забезпечують  $\min(q_{ki} - q_{ki}^*)$ ,  $k = \overline{1, M_i}$ ,  $i = \overline{1, K}$ , при  $q_{ki}^* < q_{ki}$ , або  $\min \hat{C}$ , при  $q_{ki}^* > q_{ki}$ ,  $k = \overline{1, M_i}$ ,  $i = \overline{1, K}$ . Відповідні моделі наведені нижче.

Нехай управління  $u^+ \in U$ , де  $U$  — множина управляючих дій, полягає у виділенні додаткових ресурсів застосуванню  $A_l^* \in A$ , для якого фактична якість  $q_{\phi}$  гірша цільової  $q_{\psi}$ ,  $q_{\phi} < q_{\psi}$ , а  $u^- \in U$  управління, яке вилучає ресурси у застосування  $A_l^* \in A$ , якщо  $q_{\phi} > q_{\psi}$ . Для монотонних функцій  $f_{qr}$ ,  $q = f_{qa}(\hat{a})$  і  $q = f_q(r, \hat{a})$ , враховуючи монотонний характер залежності  $q_{ki}$ ,  $\forall k = \overline{1, M_i}, i = \overline{1, K}$  від  $r$ , доведено існування бажаного управління [13].

**Теорема 1.** Для заданих значень  $q_{ki}$ ,  $\forall k = \overline{1, M_i}, i = \overline{1, K}$ , якщо  $q_{ki}^* < q_{ki}$ , існує управління  $u^+ \in U$ , яке дозволяє забезпечити  $q_{ki}^* = q_{ki}$  при  $\min r$  на підтримку рівня послуг.

Обґрунтовано, що управління  $u^+ \in U$  знаходиться ітеративно.

**Теорема 2.** Якщо  $\hat{r} = r + r_r$  і  $q_{ki}^* < q_{ki}$  і відомі значення  $\hat{a}$ , то управління  $u^+ \in U$ , яке дозволяє відновити умову  $q_{ki}^* = q_{ki}$ ,  $\forall k = \overline{1, M_i}, i = \overline{1, K}$  може бути знайдено без ітеративних процедур.

**Наслідок.** Якщо  $q_{ki}^* < q_{ki}$  і  $i$ -а послуга підтримується застосуванням  $A_l^* \in A$ , а за рахунок застосувань із  $A$ , для яких  $q_{\phi} > q_{\psi}$ , можна вивільнити обсяг ресурсів  $\Delta \hat{\rho}_e$ , причому  $\Delta \hat{\rho}_e \geq \Delta \hat{\rho}_{lj}$ , де  $\Delta \hat{\rho}_{lj}$  — додатковий обсяг ресурсів, який потрібно виділити  $l$ -му застосуванню,

щоб забезпечити виконання  $q_{ki}^* = q_{ki}$ , то управління  $u^+ \in U$  в умовах дефіциту ресурсів дозволяє відновити рівень  $i$ -ї послуги, що підтримується  $l$ -м застосуванням, за один прохід.

Процес погодження рівня послуг переформовує матрицю  $Q' = \|q'_{ki}\|$ . На її основі нижчий рівень вивільняє і виділяє ресурси послугам, для яких  $q_{ki}^* \neq q_{ki}$ ,  $k = \overline{1, M_i}$ ,  $i = \overline{1, K}$ . У випадку невдачі здійснюється ескалація проблеми на рівень бізнесу, який генерує матрицю  $Q' = \|q'_{ki}\|$ , або збільшує  $r_1, \dots, r_m$ .

Багаторівневі СУІ доцільно будувати з використанням універсальної компоненти — дворівневої системи управління з координатором.

**Базова дворівнева модель.** Багаторівневі ієрархічні СУІ будуються із дворівневих систем управління з координатором (рис. 2) [14].

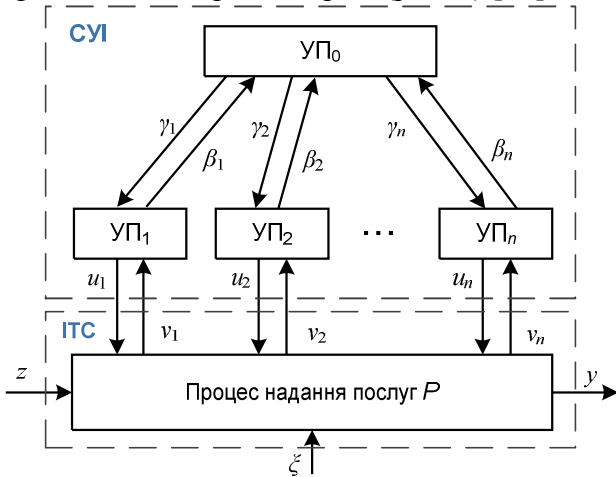


Рис. 2. Базова дворівнева модель представлення ієрархічних СУІ

Ці системи забезпечують не оптимальне функціонування ІТ-інфраструктури, а поліпшення якісних характеристик її використання. Координатор погоджує рішення підсистем з точки зору якості послуг. Процес  $P$  надання послуг задамо відображенням  $P: U \times Z \times \Xi \rightarrow Y$ , де  $U = U_1 \times U_2 \times \dots \times U_n$ . — множина управляючих сигналів (компоненту  $u_g \in U_g$  управляючого впливу  $u$  формує підсистема  $УП_g$ ,  $g = \overline{1, n}$ );  $Z$  — множина вхідних сигналів  $z$  (запитів користувачів);  $\Xi$  — множина збурень  $\xi$ ;  $Y$  — множина виходів (результатів виконання запитів). На вхід  $УП_g$ ,  $g = \overline{1, n}$  поступають сигнали: координуючий  $\gamma_g \in \Gamma$  — від управляючої підсистеми верхнього рівня  $УП_0$ ; інформаційний зворотного зв'язку  $v_g \in V_g$  — від системи моніторингу.

Підсистема  $УП_g$ ,  $g = \overline{1, n}$  реалізує відображення  $C_g: \Gamma \times V_g \rightarrow U_g$ . Сигнал  $v_g \in V_g$  функцією  $f_g: U \times Z \times \Xi \times Y \rightarrow V_g$  сигналів  $u$ , входів  $z$ , збурень  $\xi$  і виходів  $y$ . Управляюча підсистема  $УП_0$  як координатор виробляє сигнали  $\gamma_g \in \Gamma$ ,  $g = \overline{1, n}$ , реалізуючи відображення  $C_0: \mathbf{B} \rightarrow \Gamma$ , де  $\mathbf{B}$  — множина інформаційних сигналів  $\beta = (\beta_1, \dots, \beta_n)$  зворотного зв'язку. Сигнал  $\beta_g$ ,  $g = \overline{1, n}$  від  $УП_g$ , визначається відображенням  $f_0: \Gamma \times V \times U \rightarrow \mathbf{B}$ , де  $V = V_1 \times \dots \times V_n$ .

Координація погоджує дії підсистем для досягнення глобальної цілі системи — забезпечення максимальної якості  $Q$  ІТ-сервісів

$$\begin{aligned} \max Q &\Leftrightarrow \max Q_i, \forall i = \overline{1, K} \Leftrightarrow \\ &\Leftrightarrow \max q_{ki}, \forall i = \overline{1, K}, \forall k = \overline{1, M_i}, \end{aligned} \quad (2)$$

з мінімальною вартістю  $C$  (тут  $Q_i$ ,  $i = \overline{1, K}$  — якість  $i$ -ї послуги;  $q_{ki}$ ,  $k = \overline{1, M_i}$  — значення  $k$ -го показника якості  $i$ -ї послуги);

Ціллю координатора є підтримка якості  $Q$  послуг на погодженому рівні з мінімальними затратами  $C$  на задіяні ресурси  $Q = const|_{\min C}$ , локальні цілі підсистем полягають у підтримці заданих значень параметрів функціонування ІТ-інфраструктури з мінімальними витратами.

Координація на основі постулатів сумісності не можлива, оскільки точний прогноз значень параметрів процесу  $P$  і аналітичні вирази для вирішення задачі координації не відомі. Тому розроблена методика синтезу координатора на основі декомпозиції глобальної задачі.

Якщо  $Q_i$  і  $Q_i^*$ ,  $q_{ki}$  і  $q_{ki}^*$  — цільове і фактичне значення якості  $i$ -ї послуги і її  $k$ -го показника, ціль координатора набуде вигляду  $\min \Delta Q_i = \min(Q_i - Q_i^*)$ ,  $\forall i = \overline{1, K} \Leftrightarrow \min \Delta q_{ki} = \min(q_{ki} - q_{ki}^*)$ ,  $\forall k = \overline{1, M_i}, \forall i = \overline{1, K}$ ,

а задача СУІ зводиться до вибору управління, яке протидіє збуренню. Координатор виробляє координуючі сигнали, щоб мінімізувати відхилення, і природно використовувати принцип управління по відхиленню. Якщо можна виміряти збурення, вплив несправностей на якість сервісів, завантаженість каналів зв'язку, обгрунтовано перехід до комбінованого управління і виконана декомпозиція координатора.

Зміни глобальної цілі управління ІТ-інфраструктурою призводять до змін пріоритетів  $Pr$  застосувань  $\{A_i\}$  і цільових значень  $Q_i$ ,  $i = \overline{1, K}$

рівня послуг, а відображення  $C_0$  набуває вигляду  $C_0 : Pr \times Q_i \times \mathbf{B} \times V \times \mathbf{E} \rightarrow \Gamma, i = \overline{1, K}$ .

Вибір координуючого впливу визначається системою  $\langle Pr, \Delta\hat{q}, \mathbf{B}, V, \mathbf{E}, \mathcal{N} \rangle$ , де  $\Delta\hat{q}$  — вектор відхилень;  $\mathcal{N}$  — найбільш характерна для ІТ-інфраструктури ситуаційна невизначеність. Пріоритети  $Pr$  застосувань набувають значень із множини  $\{1, 2, \dots, Pr_m\}$ . Ступінь відхилення значень

елементів вектора  $\Delta\hat{q} = (\Delta q_{1,1}, \dots, \Delta q_{k,i}, \dots, \Delta q_{M_K, K})$  від нульових значень характеризує функція  $w(\Delta q_{kl}) = (q_{kl} - q_{kl}^*) / q_{кр_{kl}}, \forall k = \overline{1, M_i}, \forall i = \overline{1, K}$ ,

де  $q_{кр_{k,i}}$  — критичне значення показника якості  $i$ -ї послуги. Оскільки в аналітичному вигляді визначити відповідне відображення не видається можливим, розроблена ітеративна процедура координації [14], що передбачає участь усіх процесів, наведених на рис. 1. Рішення координатора впливають на вибір не управляючих, координуючих впливів [14], що визначають УП<sub>g</sub>,  $g = \overline{1, n}$ , якому віддається перевага при відновленні якості послуг і які методи доцільно використовувати. Авторами розроблений комплекс відповідних методів і продукційних систем для організації такої координації.

Отже, координація базується на ітеративних процедурах поліпшення координуючих сигналів на основі аналізу результатів координації або використання зворотного зв'язку для корегування координуючого сигналу [14]. В обох випадках для визначення сигналу похибки при оцінюванні впливу запропоновані методи зведення метрик, що вимірюються на рівні процесу  $P$ , до метрик, з якими оперує координатор.

**Реалізація декомпозиційно-компенсаційного підходу на основі СРМ.** Авторами розроблена методологія управління рівнем ІТ-послуг, яка визначає основи розроблення, впровадження і застосування СРМ [15]. Ці моделі зображуються ієрархічними графами, вершини яких відповідають ЛОУ у вигляді ресурсів, активів та інших сервісів, а ребра — зв'язкам між ними. Типи ЛОУ і зв'язки, що використовуються для створення СРМ, визначаються будовою сервісу і цілями створення моделі.

Методика створення СРМ розроблена на основі системної методології в рамках визначеного процесу управління ІТ-інфраструктурою. Вигляд моделі залежить від цілей процесу, для

якого створюється модель. Одна із основних задач СРМ полягає у організації моніторингу параметрів якості ІТ-послуг [15]. При цьому на практиці якість ІТ-послуг визначається комплексом вимірюваних характеристик, що відповідають вимогам користувачів.

*Методологія управління рівнем послуг на основі СРМ* надає розроблені авторами моделі, методи, методики і алгоритми для виконання таких комплексів робіт з управління рівнем послуг: формування каталогу ІТ-сервісів; визначення параметрів якості ІТ-послуг; проектування СРМ; визначення методів отримання значень показників якості сервісів; моніторинг і аналіз значень параметрів якості; управління ІТ-інфраструктурою з підтримки рівня послуг.

Реалізація декомпозиційно-компенсаційного підходу на основі СРМ передбачає комплексне управління рівнем ІТ-послуг, тісний взаємозв'язок процесного і оперативного управління з закріпленням за СУІ ролі зв'язуючої ланки для інтеграції управління [16–21]. Процесне управління корегує цільові значення  $q_{b,ki}, \forall b, k, i$  показників якості сервісів і стимулює вирішення проблем в ІТ-інфраструктурі. У контурі оперативного управління аналізуються ризики, планується зростання/скорочення кількості користувачів, обсягів даних збереження, компенсуються піки завантаженості з врахуванням змін умов надання послуг. Його Інструментом є перерозподіл ресурсів ІТ-інфраструктури з урахуванням пріоритетів бізнесу, критичності послуг і прогнозу вимог до їх доступності з забезпеченням прийнятого рівня якості.

Авторами розроблені математичні моделі, методи і відповідні алгоритми для реалізації усіх функцій управління ІТ-інфраструктурою. У статті розглядаються лише моделі і методи для: розподілу ресурсів і навантаження; імпаکت-аналізу; управління потоками в мережах; зведення метрик ІТ-інфраструктури і її складових.

### Математичні моделі і методи управління ІТ-інфраструктурою

*Управління ресурсами і навантаженням.* Класифікація моделей управління ресурсами і навантаженням ІТ-інфраструктури за ознаками рівень абстракції ресурсів (абстрактні, конкретні), етапом ЖЦ (узгодження, планування, впровадження і управління ІТ-сервісами), призна-

чення (управління власною ІТ-інфраструктурою, надання ІТ-послуг), технологія (виділені сервери, серверна віртуалізація, віртуальний хостинг, хмарні обчислення), забезпечення процесу ресурсами (цілком або ніяк, допускається часткове забезпечення) визначає потребу в  $n \cdot 100$  моделях управління ресурсами і навантаженням. Авторами розроблені моделі для всіх важливих класів. Розглянемо деякі з них.

**Розподіл абстрактних ресурсів.** Базова модель розподілу обмежених ресурсів враховує бізнес-процеси  $Z_1, \dots, Z_n$  ( $W = (w_1, \dots, w_n)$  – вектор коефіцієнтів їх важливості), що використовують ресурси  $R_1, \dots, R_m$  ІТ-інфраструктури. Нехай  $\|p_{ij}\|$  – матриця потреб бізнес-процесів у ресурсах, де  $p_{ij}$  відповідає приведеній кількості потрібного для процесу  $Z_i$  ресурсу  $R_j$ ,  $R = (r_1, \dots, r_m)$  – вектор обмежень на ресурси [22].

Якщо процес забезпечується ресурсами цілком або ніяк, вводиться булева змінна  $X = (x_1, \dots, x_n)$ , яка визначає підтримку бізнес-процесів. Тоді критерій набуває вигляду:

$$\max \sum_{i=1}^n x_i \cdot w_i, \quad (3)$$

і мають виконуватися ресурсні обмеження:

$$\sum_{i=1}^n x_i \cdot p_{ij} \leq r_j, j = 1, \dots, m. \quad (4)$$

Якщо допускається часткове забезпечення ресурсами, вводиться неперервна змінна  $x_i$ , яка визначає рівень підтримки бізнес-процесу  $Z_i$ . Тоді критерій (3) і ресурсні обмеження (4) залишаються і додається обмеження

$$0 \leq x_i \leq 1, i = 1, \dots, n, \quad (5)$$

Також розглянуто ряд інших корисних випадків, наприклад з: стохастичними оцінками наявних ресурсів, для якого визначено умови зведення до еквівалентної детермінованої задачі лінійного програмування; врахуванням надійності ресурсів, зведений до лінійної задача булевого програмування [23]: максимізувати (3) при виконанні обмежень (4) і вимоги до надійності [24]

$$\sum_{j=1}^m x_i \cdot d_{ij} \cdot \ln q_j \geq \ln s_i, \text{ для } i = 1, \dots, n, \quad (6)$$

де  $s_1, \dots, s_n$  – вимоги до показників надійності бізнес-процесів,  $i = 1, \dots, n$ ,  $q_1, \dots, q_m$  – показники надійності ресурсів,  $j = 1, \dots, m$ ;  $D = \|d_{ij}\|$  – булева матриця потреб бізнес-процесів у ресурсах.

**Розподіл конкретних ресурсів.** При плануванні і управлінні ресурсами і навантаженням

для підтримки бізнес-процесів реальними ресурсами з врахуванням технічних і географічних обмежень інтегрований ресурс  $R_j$  розглядається як група конкретних ресурсів одного типу але різного об'єму та надійності. У цьому класі запропоновано ряд лінійних і нелінійних булевих і неперервних задач розподілу ресурсів і навантаження з урахуванням ресурсних, надійнісних, технічних і географічних обмежень [25, 26].

**Розподіл ресурсів і навантаження ЦОД.** Розроблено підхід [27] до управління розподілом ресурсів ЦОД при віртуальному хостингу, серверній віртуалізації та хмарних обчисленнях. Розв'язано низку задач, постановки яких враховують моделі хостингу (віртуальний; віртуальні приватні сервери; виділені сервери; колокейшн), архітектуру побудови систем, технології побудови і організації доступу до прикладного програмного забезпечення, технології віртуалізації і кластеризації.

**Моделі і методи управління навантаженням і ресурсами для виділених серверів.** Сервери  $N_1, \dots, N_n$  обслуговують клієнтів  $K_1, \dots, K_m$ , надаючи їм доступ до застосувань. Перші характеризуються набором технічних параметрів, а другі – набором вимог. Сервер  $N_i, i = 1, \dots, n$ , має чотири параметри, які разом характеризують його потужність:  $\Omega_i$  – процесорна місткість; місткість  $\Gamma_i$  оперативної пам'яті і  $\Phi_i$  жорстких дисків;  $\Lambda_i$  – місткість каналів. Клієнт  $K_j, j = 1, \dots, m$ , має вимоги до процесорної місткості серверів ( $\omega_j$ ), оперативної пам'яті ( $\gamma_j$ ), пам'яті жорстких дисків ( $\phi_j$ ) і місткості каналів ( $\lambda_j$ ) серверів, на яких розташовані застосування, які обслуговують користувачів клієнта  $K_j$  [28].

Характеристики серверів і вимоги застосувань поділяються на залежні ( $\Omega_i, \omega_j, \Lambda_i, \lambda_j$ ) і незалежні від навантаження ( $\Gamma_i, \gamma_j, \Phi_i, \phi_j$ ). Вимоги до залежних від навантаження параметрів серверів визначаються сумарними вимогами користувачів клієнта, а вимоги до незалежних від навантаження параметрів серверів – сумарними вимогами усіх застосувань клієнта. Можливості сервера  $N_i$  розмістити застосування  $j$  визначає відповідний елемент булевої матриці  $R = \|R_{ji}\|_{m \times n}$ ,  $w_j$  – важливість застосування  $j$ . Запропоновано задачі планування розміщення застосувань на серверах і диспетчеризації навантаження.

Задача планування розподілу ресурсів полягає у визначенні матриці  $X = \|x_{ji}\|$  прив'язки клієнтів до серверів з врахуванням обмежень:



1) сумарні незалежні від навантаження вимоги застосувань сервера не перевищують його незалежні від навантаження параметри;

2) залежні від завантаження вимоги застосувань повинні бути забезпечені сумарними можливостями залежних від завантаження параметрів серверів, на яких вони розміщуються;

3) обмеження, визначені матрицею  $R$  і вектором технологічних вимог  $C$ .

Авторами розроблено декілька моделей із відмінностями у забезпеченості серверами за різними критеріями (максимальна сумарна важливість застосувань, що отримали ресурси; рівномірне завантаження серверів; максимальні сумарні експлуатаційні витрати на вивільнені сервери; мінімальна сумарна вартість реалізації операцій переходу від попереднього розміщення). Наприклад, задача вибору мінімальної за витратами на їх підтримку підмножини серверів, які зможуть задовольнити вимоги існуючих користувачів застосувань (тут  $s_i$  – експлуатаційні витрати на підтримку сервера  $N_i$ ):

$$\min \sum_{j=1}^m \sum_{i=1}^n s_i \cdot x_{ji} \quad (7)$$

за обмежень:

$$\omega_j x_{ji} \leq \Omega_i, \quad i = 1, \dots, n, j = 1, \dots, k; \quad \lambda_j x_{ji} \leq \Lambda_i, \\ i = 1, \dots, n, j = 1, \dots, k; \quad \gamma_j x_{ji} \leq \Gamma_i, \quad i = 1, \dots, n, \\ j = 1, \dots, k; \quad \phi_j x_{ji} \leq \Phi_i, \quad i = 1, \dots, n, j = 1, \dots, k;$$

$$\sum_{j=1}^m x_{ji} \leq 1, \quad i = 1, \dots, n; \quad \sum_{i=1}^n x_{ji} = 1, \quad j = 1, \dots, m; \quad (8)$$

$$\delta_j \leq \Delta_i, \quad i = 1, \dots, n; \quad x_{ji} \leq R_{ji}, \quad i = 1, \dots, n, j = 1, \dots, k.$$

Ці задачі є лінійними та нелінійними варіантами задач булевого програмування, що виправдовує використання евристик або методів «м'яких» обчислень, насамперед генетичних алгоритмів, модифікованих для отримання контрольованого наближення до результату.

Модель серверної віртуалізації. На фізичному сервері функціонує декілька віртуальних серверів  $S_i$ ,  $i=1, \dots, n$ , у кожному з яких встановлене одне чи декілька застосувань  $A_{ij}$ ,  $j=2, \dots, m_i$ ;  $m_i > 1$ . Нехай  $r_k$  – кількість ресурсів типу  $R_k$ , встановлена на фізичному сервері,  $r_{ki}$  – кількість ресурсів типу  $R_k$ , призначена віртуальному серверу  $S_i$ ,  $r_{kij}$  – кількість ресурсів типу  $R_k$ , що споживається застосуванням  $A_{ij}$ . Вочевидь, має виконуватись така умова:

$$\sum_{i=1}^n \sum_{j=1}^{m_i} r_{kij} \leq r_k, \quad k = 1, \dots, l. \quad (9)$$

Для множини  $K_0$  типів ресурсів, що не можуть динамічно перерозподілятися між віртуальними серверами, мають виконуватись умови:

$$\sum_{j=1}^{m_i} r_{kij} \leq r_{ki}, \quad i = 1, \dots, n, \quad k \notin K_0, \quad (10)$$

$$\sum_{i=1}^n r_{ki} \leq r_k, \quad k \notin K_0. \quad (11)$$

Якщо булева або неперервна змінна  $x_p$ , визначає ступінь підтримки, а  $w_p$  – важливість сервісу  $Z_p$ ,  $p=1, \dots, z$ , то критерій набуде вигляду

$$\max \sum_{p=1}^z x_p \cdot w_p. \quad (12)$$

Авторами розроблено ряд задач булевого і лінійного програмування для управління розподілом ресурсів та навантаження на етапах планування, управління і диспетчеризації [29–31].

Група задач управління для таких можливостей зупиняти застосування і віртуальні сервери для вивільнення ресурсів: не можна зупинити ні перші, ні другі; можна зупинити перші, але не другі; можна зупинити і перші, і другі; можна зупинити і перші, і другі (змінні неперервні).

Група задач планування: потрібно вибрати конфігурацію фізичного сервера і розподілити ресурси між віртуальними серверами та застосуваннями; при визначеній конфігурації сервера потрібно розподілити ресурси між віртуальними серверами та застосуваннями так, щоб найбільш ефективно підтримати сервіси у випадку дефіциту і надлишку ресурсів.

Задача диспетчеризації полягає у розподілі клієнтських запитів між віртуальними серверами пропорційно до значень змінних  $x_p$ , які були визначені на етапах планування та управління і розв'язується на основі системи правил.

Моделі і методи управління ресурсами і навантаженням для випадку віртуалізації серверів з можливістю міграції. Нехай  $r_{ki}$  – кількість ресурсу типу  $R_k$  сервера  $S_i$ ;  $p_{kj}$  – потреби ВМ  $V_j$  у ресурсах типу  $R_k$  для забезпечення заданого рівня SLA;  $x_{ij}$  – булева змінна, яка визначає, чи встановлена ВМ  $V_j$  на сервері  $S_i$ . Оскільки кожна ВМ одночасно розташована не більше, ніж на одному сервері, має виконуватись умова:

$$\sum_{i=1}^n x_{ij} \leq 1, \quad j = 1, \dots, m. \quad (13)$$

Для нормального функціонування ВМ повинні бути забезпечені достатнім обсягом ресурсів серверів, на яких вони розташовані:

$$\sum_{j=1}^m x_{ij} p_{kj} \leq r_{ki}, k = 1, \dots, l, i = 1, \dots, n. \quad (14)$$

Авторами розроблені моделі для розташування ВМ у випадку, якщо:

неможливо забезпечити усі ВМ ресурсами і визначається їх розташування, що підтримує сервіси з найбільшою сумарною важливістю:

$$\max \sum_{j=1}^m \sum_{i=1}^n x_{ij} w_j. \quad (15)$$

при обмеженнях (13), (14);

кількість наявних ресурсів суттєво перевищує потреби ВМ і визначається їх розташування ВМ, що забезпечує найменше енергоспоживання за рахунок вимкнення живлення серверів, які не використовуються:

$$\max \sum_{i=1}^n d_i e_i. \quad (16)$$

при обмеженнях (14), (17), (18):

$$\sum_{i=1}^n x_{ij} = 1, j = 1, \dots, m, \quad (17)$$

$$d_i = \prod_{j=1}^m \overline{x_{ij}}, j = 1, \dots, m. \quad (18)$$

Тут  $e_i$  позначає енергоспоживання сервера  $S_i$  коли на ньому не працює жодна ВМ.

Для цих задач застосовувалися методи булевого програмування – послідовного аналізу варіантів та ін. Врахування специфіки задачі дозволило суттєво зменшити перебір за рахунок симетричних рішень, однакових конфігурацій серверів тощо. Використання для планування наступних кроків наближених швидких алгоритмів теж дозволило зменшити перебір.

Моделі управління ресурсами і навантаженням для випадку надання сервісів IaaS провайдерами хмарних обчислень. Якщо сервіси IaaS планують у нодах, при укладанні договору фіксуються гарантовано (знаходиться у розпорядженні замовника в будь який момент часу) і додаткова (виділяється при потребі за наявності зайвих ресурсів) кількість нод [32, 33]. При цьому гарантовані ноди оплачуються за весь період, а додаткові – лише по факту їх використання. Оскільки додаткові ноди можуть виділятися замовнику з того ж фізичного сервера, де базується його ВМ, інструментом розподілу ресурсів стає міграція ВМ між серверами і виникає потреба у їх ущільненні. Нехай  $r_i$  – ресурси сервера  $S_i$ , у нодах,  $p_{oj}$  і  $p_j$  – кількість гарантованих і додаткових нод для ВМ  $V_j$ ,  $x_{ij}$  – булева

змінна, яка визначає, чи встановлена ВМ  $V_j$  на сервері  $S_i$ . Якщо можливості замовляти додаткові ресурси не надається, тобто  $p_j = p_{oj}$ , для  $j = 1, \dots, m$ , проблема зводиться до описаних вище задач. Також мають виконуватися такі умови:

$$\sum_{i=1}^n x_{ij} = 1, j = 1, \dots, m \quad (19)$$

$$\sum_{j=1}^m x_{ij} p_{oj} \leq r_i, i = 1, \dots, n. \quad (20)$$

За наявності вільних ресурсів провайдер:

– у випадку суттєвого надлишку ресурсів визначає найщільніший розподіл, щоб вивільнити частину серверів з метою економії електроенергії. При цьому мають бути задоволені усі додаткові потреби замовників у ресурсах

$$\sum_{j=1}^m x_{ij} p_j \leq r_i, i = 1, \dots, n. \quad (21)$$

Тоді задача формулюється: максимізувати (16) при обмеженнях (19), (20);

– у випадку несуттєвого перевищення наявними ресурсами вимог користувачів визначає найщільніший розподіл з менш суворими вимогами (20) замість (21). За критерій приймаємо

$$\min \sum_{i=1}^n \left| r_i - \sum_{j=1}^m x_{ij} p_j \right|, \text{ а обмеження – (19), (20).}$$

### Керовані і адаптивні генетичні алгоритми для задач управління ресурсами і навантаженням ІТ-інфраструктури

Використання традиційних методів розв'язання сформульованих задач розподілу ресурсів і навантаження виявилось неефективним через їх велику вимірність. Виправданою виявилась комбінована стратегія розв'язання зазначених задач – використання традиційних методів для задач невеликої вимірності і евристичних та генетичних алгоритмів для задач великої вимірності. Але експериментальні дослідження показали, що використання класичного генетичного алгоритму не забезпечує достатньої швидкості зміни значення функції пристосованості. Так, при розподілі обмежених ресурсів початкова популяція створювалась за допомогою рандомізованого емпіричного «жадібного» алгоритму, використовувався однокочковий кросинговер, імовірність мутації була фіксована і в різних експериментах набувала значення від 0,01 до 0,05. Значення функції пристосованості були досить далекими від оптимальних і

повільно зростали. Заміна одноточкового кросинговеру на двоточковий, підбір розміру популяції та інші технічні прийоми суттєво не покращили ситуацію. Позитивний вплив на якість знайденого рішення за рахунок наявності більшого об'єму генетичного матеріалу нівелювався збільшенням часу пошуку цього рішення, оскільки, по-перше, збільшується кількість епох, потрібна для випадкового використання в операторі кросинговеру підходящих одна до одної особин, а по-друге, час роботи оператора селекції, який використовується на початку кожної епохи, знаходиться у прямій залежності від розміру популяції. Було встановлено, генетичні алгоритми здатні конкурувати з іншими методами, забезпечуючи достатньо високу ефективність, однак їх ефективність суттєво залежить від кодування рішень, налаштування параметрів, реалізації операторів. Не існує строгих механізмів чіткого передбачення результатів для проблем, які розв'язуються за допомогою генетичних алгоритмів.

Тому виникла необхідність розроблення методології управління збіжністю генетичних алгоритмів для надання їм здатності ефективно розв'язувати різноманітні задачі. Авторами було розроблено керований і адаптивний варіанти генетичного алгоритму [34–36].

**Керований генетичний алгоритм.** Концептуально він відрізняється від класичного генетичного алгоритму механізмом, який, виходячи із розмірності задачі і заданої точності рішень, знаходить потрібний компроміс між кількістю кроків алгоритму і точністю рішень. Керований генетичний алгоритм враховує особливості задачі, характеристики попередньої популяції і послідовності найкращих рішень з метою визначення такої послідовності операторів кросинговеру, мутації та вибору, яка забезпечує найшвидшу збіжність. Основна ідея підходу полягає в застосуванні системи правил управління операторами вибору, кросинговеру і мутації в процесі розв'язання конкретної задачі на основі поточних результатів, історії і обчислення додаткових параметрів. Система правил підтримує баланс між розширенням дослідження простору пошуку і ефективним використанням усіх корисних для одержання найкращого рішення шим. Правила працюють таким чином, щоб наближення до найкращого рішення було контрольованим. Якщо кроки наближення досить великі, то необхідно прискорювати пошук, а якщо вони сповільнюються, то необхідно ро-

зширити використання корисних шим. Одночасно правила повинні утворити бар'єр на шляху збіжності до неоптимальних рішень.

У базовому варіанті генетичного алгоритму отримання популяції нової епохи відбувається шляхом послідовного виконання операторів кросинговеру (С), мутації (М) та селекції (S). Певного роду селекція відбувалась також перед кросинговером при відборі за певним принципом особин, що братимуть у кросинговері участь. Подібна організація процесу, хоча і давала збіжність пошуку, але не гарантувала оптимальності розв'язку. Це пов'язано із наявністю проблеми балансу «дослідження» і «використання», яка полягає у балансуванні між використанням вже існуючого матеріалу (закріплення оптимальних схем розв'язку за допомогою кросинговеру) та дослідженням нових оптимумів, що не входять до вивченого простору (отримання якісно нових схем за допомогою використання мутації). Велика кількість мутацій може призвести до руйнування схеми кращого представника поточної популяції, в той час як використання кросинговеру призводить до підвищення швидкості збіжності процесів пошуку до оптимуму (може бути як «локальним» так і «глобальним»).

На відміну від базового варіанту керований генетичний алгоритм забезпечує балансування між «дослідженням» і «використанням» за рахунок системи правил, яка визначає вибір наступного оператора. Кожна окрема епоха процесу пошуку присвячена, або «використанню» існуючого матеріалу (закріпленням оптимальних розв'язків), або «дослідженню» нових областей простору розв'язків. Селекція у кожному з цих випадків залишає у складі популяції нової епохи лише кращих представників із сукупності представників минулої епохи та отриманих на їх базі нащадків. У випадку, коли кількість отриманих нащадків буде занадто мала, сукупність, з якої робитиметься відбір, поповнюватиметься додатковим вливанням випадково згенерованих представників. Система правил формується на основі таких параметрів:

*коефіцієнт приросту популяції*  $k_{n,n} = \frac{l}{n}$ , що

характеризує її вироджуваність (тут  $l$  – кількість отриманих нащадків,  $n$  – розмір популяції);

перспективність популяції  $\rho = \frac{f_{\max}(x)}{f_{cp}(x)}$ , де

$f(x)$  – цільова функція або функція пристосованості, а  $f_{cp}(x)$  – її середнє значення для поточної популяції;

швидкість збіжності  $\Delta_{i-1,i} = \rho_{i-1} - \rho_i$ , що характеризує тенденції зміни збіжності при переході від епохи до епохи.

Авторами сформульований набір правил за якими обираються операції для отримання популяції наступної епохи:

1) оператор G для «вливання» нового генетичного матеріалу

IF ( $k_{n.n.} \leq k_{cp}$ ) AND ( $n \leq n_{cp}$ ) THEN G;

2) оператор кросинговеру (C) для формування популяції наступної епохи

IF ( $\Delta_{i-1,i} < 0$ ) THEN C;

або

IF (( $\rho_i \geq \rho_{cp}$ ) AND ( $\Delta_{i-1,i} \geq \Delta_{cp}$ )) THEN C;

3) оператор мутації для отримання популяції наступної епохи і переходу від стратегії «використання» до стратегії «дослідження»

IF (( $\rho_i < \rho_{cp}$ ) OR ( $\Delta_{i-1,i} < \Delta_{cp}$ )) THEN M.

Для організації управління пошуку розв'язку необхідно задатись граничними значеннями параметрів введених коефіцієнтів  $k_{cp}, n_{cp}, \Delta_{cp}, \rho_{cp}$ . Налаштуванням цих коефіцієнтів можна регулювати швидкість збіжності пошуку генетичного алгоритму та характер процесів що в ньому відбуваються на етапі модифікації системи правил, що дозволяє влаштувати свого роду зворотній зв'язок по оптимальності отриманого результату.

Авторами виконані об'ємні дослідження ефективності керованого генетичного алгоритму для усіх видів задач управління ресурсами і навантаження ІТ-інфраструктури. Результати досліджень продемонстрували переваги керованого ГА, як з точки зору якості рішення, так і часу, потрібного для його отримання, для всіх видів задач.

Наприклад, дослідження ефективності запропонованих алгоритмів для задачі виділення ресурсів для VM нодами показало, що керований генетичний алгоритм дає кращі результати. При цьому перевага його збільшується зі зростанням розмірів VM відносно до розміру фізичних серверів. Порівняльне дослідження запропонованого авторами керованого генетичного алгоритму і жадібного емпіричного алгоритму було виконане для віртуалізації серверів з мож-

ливістю міграції. Ефективність варіанту генетичного алгоритму суттєво зростає при збільшенні розміру кластеру, вимог VM до ресурсів, та розкиду параметрів VM. Ефект відчувається значно слабший для невеликих VM та зовсім відсутній для випадку використання VM з однаковими вимогами до ресурсів. Керований генетичний алгоритм забезпечував підвищення ефективності використання ІТ-інфраструктури в середньому від 2% до 7%.

### Моделі й методи імпакт-аналізу в управлінні ІТ-інфраструктурою

При визначенні ступеня впливу несправностей на рівень обслуговування, кожному пристрою ІТ-інфраструктури приписують рівень відповідності нормам функціонування — severity level. Враховуючи кількість і складний характер чинників впливу несправностей на рівень обслуговування, розміри ІТ-інфраструктури і кількість користувачів, зробити це в умовах жорстких часових обмежень дуже складно.

Традиційно алгоритм розрахунку рівня обслуговування для кожного абонента будують на основі шляху доступу для нього і severity level для кожного елемента ІТ-інфраструктури. Але змінність ІТ-інфраструктури, мінливість політики провайдерів щодо оцінювання впливу роблять не ефективним чисто алгоритмічний підхід. З іншого боку, використання продукційних систем також дозволяє на основі значення severity level для кожного елемента ІТ-інфраструктури, її структури, параметрів технологій, сервісів і обладнання, визначити рівень обслуговування для кожного абонента. Тут потреби у перепрограмуванні при змінах не виникає, але застосування побудованих на символічних перетвореннях продукційних систем не задовольняє жорстких часових обмежень.

Тому авторами запропонований оригінальний комбінований продукційно-алгоритмічний підхід. Його основна ідея полягає у використанні спрощеного представлення ІТ-інфраструктури, яке містить результати опрацювання правил продукційної системи у вигляді, готовому для алгоритмічних розрахунків. Тобто продукції застосовуються заздалегідь і на відповідному шляху у графі зберігаються потрібні для розрахунків значення коефіцієнтів поширення впливу несправностей. Коли ж виникають несправності, для визначення їх впливу потрібно лише обійти граф від вершин, де виникли несправності, до вершин-користувачів

і виконати арифметичні операції над попередньо визначеними коефіцієнтами.

Аналіз впливу несправностей на рівень обслуговування поділяється на дві фази. На першій правила прив'язуються до структури мережі у спрощеному представленні механізмом виведення. На другій фазі виконується перерахунок оцінок рівня обслуговування. Підготовлене заздалегідь спрощене представлення опрацьовується алгоритмічно після надходження повідомлення про несправність з використанням швидких алгоритмів обходу графів.

За параметр якості послуг вибрана швидкість доступу до мережі. Тоді імпакт-аналіз зводиться до визначення міри зниження швидкості доступу відносно її номінального значення, викликаного несправностями. Для визначення впливу несправностей на бізнес-процеси компаній і ефективного відновлення мережі необхідні її структура і схеми підтримки сервісів ресурсами і бізнес-процесів компаній-клієнтів сервісами провайдера. Спрощене представлення включає лише важливі для аналізу аспекти IT-інфраструктури:

1) топологічну схему мережі, задану графом, вершини якого відповідають вузлам мережі, а ребра — лініям зв'язку між вузлами;

2) модель вузла, яка описує проходження інформаційних потоків через нього і враховує реакцію обладнання і технологій на несправності;

3) систему правил, які описують особливості поширення впливу несправності, враховуючи особливості технологій, обладнання і функціонування мережевого вузла і його взаємодію з іншими вузлами.

Авторами розроблено базовий, швидкий і перспективний методи імпакт-аналізу на основі запропонованого продукційно-алгоритмічного підходу [37].

Основу **базового методу** складають: структурно-функціональна схема оцінювання впливу несправностей на якість сервісів і рівень обслуговування; моделі визначення послідовності дій і потрібних для їх виконання даних; механізм підключення моделей; моделі уточнення чинників, що визначають поширення впливу несправностей; система правил поширення впливу несправностей в залежності від її структури, сервісів, технологій, тощо [38].

Різноманіття сервісів і технологій вимагає побудови моделі для кожного сервісу чи їх групи. Для деталізації запропонованого підходу

були вибрані мережі доступу в IT-інфраструктурах провайдерів, високу трудомісткість і вимогливість до ресурсів задачі імпакт-аналізу в яких обумовлюють: велика кількість абонентів (до  $n \cdot 10^6$  юридичних і фізичних осіб); гоміодність мереж (до десяти і більше ієрархічних рівнів,  $n \cdot 10^4$  кількість типів і  $n \cdot 10^7$  одиниць обладнання і технологій); складний характер залежностей впливу несправностей на якість ІКС; складна інтерференційна картина впливу сукупності несправностей.

Для формального подання методу імпакт-аналізу введемо:

1) коефіцієнт зменшення швидкості доступу  $k_{i,j(n)}$ ,  $i = 1, \dots, I, j = 1, \dots, J, n = 1, \dots, N_{i,j}$ , де  $j$  — номер вузла на  $i$ -му рівні ієрархії мережі доступу,  $n$  — номер вихідної для  $i,j$ -го вузла  $U_{i,j}$  ( $j$ -й вузол мережі на  $i$ -му рівні її ієрархії) мережі лінії зв'язку,  $I, J$  і  $N_{i,j}$  — кількості рівнів ієрархії, вузлів на  $i$ -му рівні, ліній зв'язку, що виходять з вузла  $U_{i,j}$ . Коефіцієнт  $k_{i,j(n)}$  набуває значення в діапазоні  $[0,1]$  і визначає, у скільки разів зменшується швидкість доступу в  $n$ -й лінії зв'язку, що виходять з вузла  $U_{i,j}$ , відносно нормативного показника;

2) коефіцієнт  $K_{i,j}$ , який залежить від стану  $j$ -го вузла,  $j = 1, \dots, J, i$ -го,  $i = 1, \dots, I$ , рівня ієрархії мережі доступу  $U_{i,j}$ , обумовленого несправностями у мережі доступу, і визначає ступінь падіння швидкості передавання на виходах вузла  $U_{i,j}$  відносно нормативної ( $J_i$  — кількість вузлів  $U$  на рівні  $i$ ).

Якщо дію усіх несправностей в вузлах і лініях зв'язку можна визначити зміною коефіцієнтів  $k_{i,j(n)}$ , то фактична швидкість доступу в  $n$ -й лінії зв'язку, що виходять з  $i,j$ -го вузла,  $Vp_{i,j(n)} = k_{i,j(n)} \times Vn_{i,j(n)}$ , де  $Vn_{i,j(n)}$  — нормативна швидкість доступу в  $n$ -й лінії на виході  $i,j$ -го вузла. Результат порівняння фактичної швидкості  $Vp_m$  для абонента  $A_m$ ,  $m = 1, \dots, M$ , з визначеним в SLA пороговим значенням допустимого зменшення швидкості для  $m$ -го абонента разом зі значимістю абонента є основою прийняття рішення щодо дій провайдера.

Коефіцієнт  $K_{i,j}$  можна визначити на основі правил для типів вузлів, які враховують структурно-технологічні особливості мереж, сервісів, принципи їх функціонування, налаштування обладнання, резервування та інші чинники. Правила згруповані за наведеними на рис. 3 способами з'єднання вузлів [39].

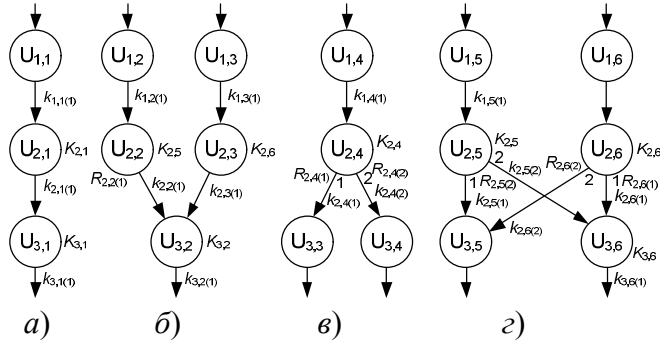


Рис. 3. Можливі способи з'єднання вузлів мереж доступу

Розглянемо, наприклад, спосіб а). При справному вузлі  $U_{2,1}$  швидкість доступу у вихідній лінії визначається вхідною швидкістю і обмеженнями на швидкість передавання і приймання, яка встановлена для портів вузла  $U_{2,1}$  при його конфігуруванні. Якщо вузол  $U_{2,1}$  не справний, то можливі такі випадки:

- 1) якщо  $K_{2,1} = 0, Vp_{2,1(1)} = 0$ , незалежно від швидкості на вході;
- 2) якщо  $K_{2,1} \neq 0$  і несправність у вузлі  $U_{2,1}$  призводить до падіння швидкості у вихідній лінії на коефіцієнт  $K_{2,1}$ , то коефіцієнт падіння швидкості відносно номінальної у вихідній лінії зв'язку вузла  $U_{2,1}$  буде становити  $k_{2,1(1)} = K_{2,1} \times k_{1,1(1)}$ . Падіння швидкості на виході вузла  $U_{3,1}$  буде визначатися за допомогою коефіцієнта  $k_{3,1(1)} = K_{3,1} \times k_{2,1(1)} = K_{3,1} \times K_{2,1} \times k_{1,1(1)}$ .

Авторами розглянуті всі способи з'єднань. На найнижньому рівні ієрархії мережі доступу коефіцієнт  $k_{i,j(n)}$  визначає падіння швидкості передачі на вході чи виході абонентського обладнання відносно швидкості, заданої в SLA. Тепер фактичну швидкість доступу абонента  $A_m, m = 1, \dots, M$ , можна визначити як

$$Vp_m = k_m \times V_{H_m}, \quad (22)$$

де  $k_m$  — коефіцієнт, відповідний  $k_{i,j(n)}$  на виході абонентського обладнання,  $V_{H_m}$  — нормативна швидкість доступу  $m$ -го абонента.

**Швидкий метод імпакт-аналізу** для підвищення ефективності механізму виведення використовує дерева виведення, а для обмеження пошукових просторів — методологію автоматичного доведення теорем. Ефективний перерахунок показників впливу несправностей елементів мережі доступу на інші елементи мережі забезпечується комбінованим застосуванням таких засобів: індексація елементів робочої області; структурування системи правил шляхом побудови дерев описання простих умов; прив'язування правил до вузлів спрощеного представлення мережі доступу; використання швидких алгоритмів обходу графів. Для

ефективної реалізації швидкого методу, побудованого на збереженні і опрацюванні спрощеного представлення у оперативній пам'яті [40], авторами розроблені ефективні способи врахування резервних вузлів і вирішена проблема розрахунку петель за рахунок використання еквівалентних перетворень. Результати експериментального дослідження швидкого методу для різних варіантів структур мереж, сервісів, кількостей користувачів, різних характеристик обладнання і технологій показали, що для мереж з 0,5 млн вузлів час створення спрощеного представлення склав трішечки більше хвилини, а час перерахунку коефіцієнтів поширення несправностей —  $n \cdot 100$  мс.

**Перспективний метод імпакт-аналізу** використовує комплекс прийомів пришвидчення розрахунків: структурування представлень мережі, стану і правил виконання розрахунків; обмеження сегменту мережі, на який вплинули несправності; оцінювання впливу несправностей на класи користувачів і перенесення його на користувачів за належністю до класу; оцінювання у порядку спадання важливості користувачів; використання попередньо підготовлених варіантів розрахунків, комбінуючи які можна охопити усі класи, на які вплинули несправності; використання представлення мережі, яке дозволяло б зберігати усі необхідні для проведення розрахунків дані у оперативній пам'яті і дозволяло б паралельне виконання розрахунків для визначення впливу несправностей.

ІТ-інфраструктура подається як система точок впливу (елементів розгалуження мережі, які впливають на функціонування пов'язаних з ними елементів, сегментів або сервісів) і відповідних їм сегментів впливу та коефіцієнтів розрахунку, що дозволяє задати напрям і об'єм необхідних обчислень для кожної точки впливу.

На основі наведених концепцій для розв'язання задачі імпакт-аналізу були розроблені алгоритми і відповідні реалізації типів А та В [39]. *Алгоритм типу А* комбінує «пошук в глибину» з «пошуком в ширину», що може дати складність  $O(an)$ , де  $n$  — кількість вершин,  $a > 1$ . Загальна складність алгоритму становить  $O(n+m)$ , де  $n$  — кількість вершин,  $m$  — кількість ребер, а на створення і сортування масивів потрібен час  $O(m \cdot \ln(m))$ . Оцінка пам'яті, необхідної для виконання становить  $O(d)$ , де  $d$  — обсяг буфера в оперативній пам'яті, необхідного для збереження чергової порції даних із бази даних. *Алгоритм типу В* працює з повним представленням, яке зберігається на дисках,

системою правил і спрощеним представленням, які зберігаються в оперативній пам'яті. Формування графа у найгіршому випадку виконується за час  $O(n+m)$ . Час виконання не перевищує  $O(n)$ . Оцінка пам'яті, необхідної для виконання:  $O(n^2)$ . Використовуючи ідею порівневого обходу дерева розповсюдження, можна пришвидшити час виконання. Для вирішення проблеми циклів запропонована спеціальна система позначок необхідності розрахунку чи використання обчислених раніше значення показника впливу для вузлів. Для швидкого встановлення у вузлах і їх прямих і непрямих спадкоємцях значень позначок встановлено корисну властивість цього процесу, яка дозволяє визначати значення позначок для великих підграфів. Дійсно, зміна значення показника впливу у вузлі  $u$  впливає лише на всіх прямих і непрямих спадкоємців і може вплинути на його батьківський вузол  $v$ , якщо  $u$  «живиться» тільки від резервних ліній і  $v$  є саме тим батьківським вузлом, від якого резервною лінією до  $u$  надходило живлення. Тоді час виконання можна оцінити складністю алгоритму позначення вершини  $u$  і її прямих і непрямих спадкоємців. Вона у середньому становить  $O(nu(t))$ , де  $nu(t)$  — кількість спадкоємців  $u$ , які позначені як  $t$ , а в найгіршому випадку —  $O(mu(t))$ , де  $mu(t)$  — кількість ребер, які виходять із спадкоємців  $u$ ). Оцінка необхідної пам'яті:  $O(n+m)$ .

Виконано експериментальне дослідження запропонованих алгоритмів типу  $B$ , типу  $A$ , базового методу за об'ємом використаної оперативної пам'яті та часом обчислень. Різноманітні випадки виходу з ладу (частково чи повністю) окремих елементів ІТ-інфраструктури генерувалися випадковим чином. Розрахунки виконувалися з різними варіантами структур мереж доступу, сервісів, кількостей клієнтів, характеристиками обладнання і технології. Вхідні дані генерувалися випадковим чином. Алгоритми типу  $B$  значно швидші ( $n \cdot 10$  мс на 0,5 млн вузлів), а їх потреби в оперативній пам'яті ( $n \cdot 10$  Мб на 0,5 млн вузлів) не надто великі з огляду на можливості сучасних серверів.

### Моделі і методи управління інформаційними потоками в ІТ-інфраструктурі

Управління ресурсами мережної складової ІТ-інфраструктури з метою зниження перевантажень, які призводять до погіршення якості послуг, доцільно зводити до управління потоками. За умови низької динаміки характеристик трафіка можна визначити дискретні часові ін-

тервали (інтервали управління  $T_y$ ), протягом яких параметри трафіка змінюються незначно, а стан мережі можна вважати стаціонарним. У такому випадку вирішити проблему перевантажень з періодом  $T_y$  можна як задачу математичного програмування, яка не містить час в явному вигляді [41].

Задача полягає в тому, щоб підтримувати значення фактичної завантаженості  $R'_C$  максимально близьким до максимально досяжної завантаженості  $R_C$ , не допускаючи перевантаження мережі, тобто управляти навантаженням так, щоб  $K_c \rightarrow 1$ . Якщо необхідно врахувати вимоги бізнес-процесів, критерієм ефективності може бути мінімальна затримка  $\tau_{cp}$  передавання даних. Тому реалізовані такі політики управління потоками: надання доступу до мережних ресурсів у повному обсязі; обмеження доступу до мережних ресурсів; відмова у доступі бізнес-процесу до мережних ресурсів.

Виникає потреба у методі управління інформаційними потоками в ІР-мережі корпоративної ІТ-інфраструктури, який би враховував вимоги бізнес-процесів, значимість інформації, що передається, для підтримки бізнес-процесів, і був здатен адаптуватися до мінливості умов передавання, пріоритетів бізнес-процесів і поточного стану мережі (завантаженості каналів зв'язку і відмов мережного обладнання).

Авторами розроблений такий метод для управління інформаційними потоками в мережах з комутацією пакетів. Концептуально управління здійснюється таким чином, щоб у випадку загрози виникнення мережних перевантажень обмежувати або забороняти передавання потоків, які генеруються низькопріоритетними застосуваннями.

Нехай кожний бізнес-процес  $b_k$  із їх множини  $B = \{b_k, k = \overline{1, K}\}$  породжує множину інформаційних потоків  $\lambda_k^P = \{\lambda_{kl}^P\}$ ,  $k = \overline{1, K}$ ,  $l = \overline{1, L_k}$ . Для кожного потоку  $\lambda_{kl}^P$  задається пріоритет  $Pr_{kl}$ ,  $k = \overline{1, K}$ ,  $l = \overline{1, L_k}$  із множини  $\Pi = \{1, \dots, Pr_{\max}\}$ . Легко визначити сумарний трафік  $\Lambda^P$  усіх потоків і  $\Lambda^U$  потоків, не пов'язаних безпосередньо з підтримкою бізнес-процесів. Трафік  $\Lambda^U$  обмежується або блокується СУІ у першу чергу, якщо він не обмежений або не заборонений корпоративною політикою використання мережних ресурсів. Кожному його потоку або типу можуть бути приписані

пріоритети із множини  $\Pi$ . Трафік  $\Lambda^T$ , що залишився, є технологічним. Його потокам або застосуванням, що їх генерують, можуть бути приписані пріоритети із множини. Сумарний трафік  $\Lambda^\Sigma$  корпоративної ІТ-інфраструктури становить  $\Lambda^\Sigma = \Lambda^P + \Lambda^U + \Lambda^T$ , причому він не має перевищувати сумарної пропускної здатності  $C^\Sigma$  мережі  $\Lambda^\Sigma \leq C^\Sigma$ . Оскільки виконання цієї умови не виключає перевантажень окремих каналів, доцільно управляти потоками так, щоб виконувалася умова  $\Lambda_i^\Sigma \leq C_i$  для всіх  $i = \overline{1, I}$ , де  $I$  — максимальна кількість каналів зв'язку,  $\Lambda_i^\Sigma$  — сумарний трафік в  $i$ -му каналі, а  $C_i$  — його пропускна здатність, причому  $C^\Sigma = \sum_{i=1}^I C_i$ .

Вхідними даними для управління інформаційними потоками є дані щодо завантаженості каналів зв'язку, вихідними — команди управління мережним обладнанням, що обмежують або забороняють проходження окремих потоків. Управління потоком здійснюється на вході кожного із вузлів мережі. При цьому СУІ намагається пропустити через кожний вузол мережі максимальну кількість потоків з обумовленими параметрами якості обслуговування. Задаються значення пріоритетів інформаційних потоків  $\Lambda^P$  і  $\Lambda^T$ , а також вимоги до якості обслуговування цих потоків. Обмеження потоків можливе як для окремого користувача, так і для окремого мережного сервісу.

Управління здійснюється протоколами динамічної маршрутизації і оптимізацією мережних потоків. Критерієм оптимального управління потоками є підтримка максимально можливої швидкості передавання пакетів при відсутності втрат внаслідок переповнення буферів:

$$S_q - S_q(t) \geq 0, \text{ при } \tau(t) + u(t) \rightarrow \min, \quad (23)$$

де  $S_q$  — обсяг вхідного буфера,  $S_q(t)$  — поточне значення черги,  $\tau(t)$  — міжкадровий інтервал,  $u(t)$  — змінна управління.

Основні функції управління інформаційними потоками:

1) аналіз стану завантаженості окремих каналів і мережі в цілому, полос пропускання усіх потоків кожного інтерфейсу вузлів комутації;

2) завантаженість каналу зв'язку більш ніж на 90% викликає попередження адміністратору щодо можливого перевантаження мережі чи окремих трактів;

3) завантаженість каналу зв'язку більш ніж на 95% викликає автоматичне обмеження вхід-

ного навантаження. Для цього із списку потоків, що проходять через інтерфейс, вибираються потоки, які попали в таблицю пріоритетності, і ті, які не попали. Полоса пропускання перших обмежується до вимог, указаних в таблиці, а другі автоматично послідовно забороняються. При цьому є можливість вводити обмеження для окремих користувачів, всього потоку чи групи потоків і користувачів. Відповідні управляючі впливи (у вигляді команд мережного обладнання) передаються на вузли комутації. Цей крок виконується доки завантаженість каналу не впаде нижче заданого порогового значення. Якщо цього недостатньо обмежуються найменш пріоритетні потоки  $\Lambda^P$  і забороняються потоки  $\Lambda^U$ .

Основною перевагою запропонованого авторами рішення є здатність побудованої на його основі системи обмежувати або блокувати тільки конкретні потоки окремих користувачів. Запропонований метод управління інформаційними потоками дозволяє: підтримувати погоджений рівень найважливіших послуг; підвищити ефективність використання мережних ресурсів; оптимізувати розподіл мережних ресурсів між бізнес-процесами, розподіленими застосуваннями і користувачами; балансувати вхідне навантаження і продуктивність мережі з метою попередження падіння ефективності мережі і можливості втрати працездатності в результаті перевантажень.

Для ефективної роботи підсистеми потрібно визначити величину и навантаження кожного потоку на мережу. Для цього було виконане моделювання різних підходів до управління потоками даних в корпоративній ІТ-інфраструктурі, що використовує ІР-протокол, порівняння їх ефективності у вирішенні проблеми уникнення перевантажень мережі.

Авторами побудовані моделі таких регуляторів для реалізації СУІ:

1. Апаратний регулятор виконаний по аналогії з методом управління трафіком в маршрутизаторах policing, який для обмеження навантаження на канали зв'язку затримує або відкидає пакети при перевищенні ним потоком заданого рівня. Цей регулятор відкидає трафік, який перевищує цільове значення. Його перевагами є простота реалізації та універсальність, а недоліком — здійснення управління за рахунок не регулювання, а обмеження у випадках перевантажень. Управляючі дії, які б синтезували сигнали для генераторів трафіка щодо необхіднос-



ті з метою запобігання перевантажень зменшення обсягів даних, що передаються, не виробляються.

2. Програмний регулятор у випадку необхідності обмеження потоку після порівняння цільової і фактичної його величин поступово зменшує коефіцієнт передавання, доки потік не буде достатньо обмеженим. Перевагою такого регулятора є універсальність, а недоліком — суттєвий вплив на якість управління затримок збору та оброблення інформації і поширення управляючих дій. Великі затримки викликають невідповідні управляючі дії. При цьому лінійне зменшення коефіцієнта може супроводжуватися стрибкоподібним збільшенням, коли потік уже достатньо обмежений, але через відсутність актуальної інформації продовжується його подальше обмеження.

3. Оптимальний за швидкодією регулятор, який на відміну від апаратного і програмного регуляторів, що функціонують у відповідності з алгоритмами, використовуваними мережним обладнанням, для управління інформаційними потоками використовує методи теорії автоматичного управління. Об'єктом управління є станція, підмережа або сукупність елементів, що генерують потоки інформації. Вихідною (регульованою) величиною цієї системи є завантаженість каналу зв'язку, вхідним впливом — пропускна здатність каналу зв'язку, помилкою — величина перевантаження мережі, а ціллю управління — мінімізація помилки за максимально короткий проміжок часу. Цю систему можна класифікувати як різновид дискретної системи стабілізації, оскільки при відсутності перевантажень управляти потоками не потрібно, а в супротивному випадку регулятор намагається підтримувати стабільне значення величини потоку не вище заданого вхідного впливу.

Авторами запропонована і обґрунтована методика синтезу регулятора, який забезпечує скінченій системі необхідні показники якості — відсутність статичної помилки і максимальну швидкодію — на основі передавальної функції і вибору бажаних характеристик замкненої системи.

Результати моделювання з регулятором, розрахованим в середовищі Matlab, підтвердили очікування. Порівняння перехідних характеристик систем управління потоками трафіка без регулятора і з оптимальним за швидкодією регулятором показало, що застосування запропонованих методів дозволяє отримати швидкий процес регулювання. Однак, використання цих

методів обмежене чітко описаними об'єктами управління, що складно зробити в реальних мережах, а також чітко заданим видом трафіка. У інших випадках можна використовувати розроблені авторами нечіткі моделі управління трафіком.

Для підвищення якості передавання даних високопріоритетних застосувань обґрунтована можливість ефективного управління потоками інформації в мережі з врахуванням важливості даних, що передається [40]. Авторами розроблений адаптивний регулятор, який дозволяє управляти потоками інформації для уникнення перевантаженості каналів зв'язку шляхом контролю швидкості передавання даних застосувань  $A_i$ ,  $i = \overline{1, m}$  з врахуванням важливості  $w_l$ ,  $l = \overline{1, m}$  даних, що передається. Для цього розроблені модель адаптивного регулятора, алгоритми управління, обґрунтована можливість застосування існуючих протоколів управління мережними елементами і розв'язано ряд інших задач. Напрацювання управляючих дій, спрямованих на підтримку потоків високопріоритетних застосувань за рахунок обмеження потоків даних, що генеруються низькопріоритетними застосуваннями, базується на інформації щодо завантаженості мережі з врахуванням затримок на передавання контрольної інформації і управляючих дій, а також затримок на визначення і реалізацію сигналів управління.

Об'єктом регулювання є елементи IT-інфраструктури, що генерують потоки даних з інтенсивністю  $\lambda(t)$ . Задаючий вплив або вхідна величина відповідає пропускній здатності каналу передавання даних і в реальних СУІ встановлюється на рівні 95—98% пропускної здатності. Вихідна величина — інтенсивність потоку  $\lambda_p(t)$  на виході об'єкта регулювання після реалізації управляючого впливу. Помилкою регулювання є величина, що відповідає перевантаженням у каналі зв'язку.

Авторами розроблена методика синтезу регулятора, який на основі аналізу вихідної величини визначає управляючу дію, що мінімізує помилку регулювання. При цьому управляюча дію не виробляється, коли генерований трафік  $\lambda(t)$  не перевищує задаючий вплив. У супротивному випадку СУІ виробляє управляючі дії як система стабілізації з принципом управління по відхиленню.

Стохастична природа потоків інформації в мережах IT-інфраструктур унеможливорює застосування систем диференціальних рівнянь. Од-

нак управління здійснюється на інтервалі  $T_y$ , значення якого вибирається із умови:

$$T_y \geq \sum (T_{изм} + T_{n1} + T_{yв} + T_{n2} + T_p), \quad (24)$$

де  $T_{изм}$  — час вимірювання завантаженості мережі;  $T_{n1}$  — час доставки результатів вимірювання в СУІ;  $T_{yв}$  — час визначення управляючої дії;  $T_{n2}$  — час передавання даних щодо характеру і величини управляючої дії в об'єкт регулювання;  $T_p$  — час реалізації управляючої дії в об'єкті регулювання. І при виконанні умови  $T_y \leq T_{вар}$ , де  $T_{вар}$  — інтервал між зафіксованими змінами характеристик трафіка, при визначенні передавальних функцій регулятора і управляючих дій можна виключити параметр часу.

Для об'єкта регулювання, що постійно змінюється, потрібен адаптивний регулятор. Модель трафіка можна розглядати як об'єкт регулювання зі змінними параметрами, коли у проміжках часу  $T_{вар}$ , що відповідають окремим стрибкам, величина трафіка не змінюється у часі. Передавальну функцію об'єкта регулювання  $W_{OP}(z)$  на цих проміжках можна представити пропорційною ланкою з коефіцієнтом  $K_{II}$ , рівним величині потоку, тобто:

$$W_{OP}(z) = K_{II}. \quad (25)$$

Зміна інтенсивності потоку викликає відповідну зміну значення параметра об'єкта регулювання. Для його передавальної функції змінюється не структура регулятора, а тільки значення параметрів. У цьому випадку регулятор можна розглядати як аналітична система з самоналаштуванням по характеристикам об'єкта управління. Такі зміни значень параметрів регулятора виконуються пристроєм налаштування параметрів регулятора. Авторами запропонована методика синтезу адаптивного регулятора, який працює тільки в той час, коли канал перевантажений. При допустимій завантаженості каналу зв'язку контур регулювання не задіяний, інформаційний потік  $\lambda(t)$  передається без обмежень. При появі перевантажень адаптивний регулятор переходить в активний стан і управляє потоком. Елементи затримки дозволяють врахувати потрібні інтервали часу.

Тестування роботи адаптивного регулятора виконувалося на трьох типах трафіка: стохастичний трафік із ступінчатими змінами інтенсивності потоку; трафік, що генерується IRC-клієнтами; реальний e-mail трафік, що генеру-

ється пакетом OPNET на основі визначених за результатами досліджень параметрів.

Результати показали, що синтезований адаптивний регулятор дозволяє управляти інтенсивністю вхідних інформаційних потоків, виключаючи перевантаження каналів зв'язку. Регулятор обмежує інформаційні потоки за відсутності апріорної інформації про об'єкт управління, при помилці регулювання, що не перевищує  $10^{-3}$ . Застосування запропонованих моделей і методів в СУІ дозволить управляти мережними ресурсами ІТ-інфраструктури в залежності від значимості застосувань, що генерують інформаційні потоки, обмежуючи при мережних перевантаженнях інтенсивність потоків малозначимих застосувань.

### Моделі й методи зведення метрик ІТ-інфраструктури і її складових

Для реалізації процесів збору, накопичення і оброблення інформації в СУІ на основі системи метрик розроблено засоби визначення і оперування метриками, що дозволяють швидко зберігати, вибирати потрібні метрики і агрегувати існуючі метрики з метою формування і аналізу нових корисних для різних служб СУІ метрик.

Для цього запропоновано використовувати опис всіх метрик у вигляді цілісної системи, в якій кожна метрика одержує системний код, складові якого визначають її сутність, зв'язки з іншими метриками системи і алгоритми оброблення [42]. Досвід розроблення СУ для провайдерів ІКС і аналіз відомих метрик вимірювання і оцінювання мереж дозволив визначити потрібний для цього набір системних параметрів описання метрик в рамках узагальненої структури:  $\langle$ значення $\rangle$   $\langle$ тип\_метрики $\rangle$   $\langle$ об'єкт\_рахунку $\rangle$   $\langle$ рівень\_ІТ-інфраструктури $\rangle$   $\langle$ компонент\_ІТ-інфраструктури $\rangle$   $\langle$ одиниця\_вимірювання $\rangle$ . Її складовими можуть бути: значеннями — безпосередні значення метрики; типами метрики — абсолютна кількісна, відносна часткова та ін.; об'єктами рахунку — реальні елементи мережі, події, час, стани та ін.; рівнями ІТ-інфраструктури — користувачі, технології, ресурси, сервіси, продукти; компонентами — підмережі та інші важливі складові; одиницями вимірювання: всі відомі одиниці.

Системний характер запропонованого підходу підкреслює комплекс взаємозв'язків значень параметрів, який включає ієрархічні, структурні та інші відношення.

Різноманіття суб'єктивних оцінок рівні обслуговування і потреба в моделях його

об'єктивного оцінювання за наявності багатьох чинників впливу в ІТ-інфраструктурі провайдерів визначило потребу у методі агрегування показників оцінювання якості, заданих у вигляді нечітких множин [42, 43]. Нехай  $n$  метрик ІТ-інфраструктури, мереж, сервісів, технологій  $K_1, K_2, \dots, K_n$  потрібно звести у одну метрику  $Q$ . Метрики, насамперед суб'єктивні, часто є не числами, а лінгвістичними змінними, які характеризують функціями належності (ФН). Кожному значенню лінгвістичної змінної відповідає певна нечітка множина зі своєю ФН. Наприклад, якість сервісу можна визначити за допомогою змінної «Якість», значення якої можуть належати множині  $S = \{\text{надзвичайно добра } (S_1); \text{ дуже добра } (S_2); \text{ добра } (S_3); \text{ середня } (S_4); \text{ нижча середньої } (S_5); \text{ погана } (S_6) \text{ і } \text{ дуже погана } (S_7)\}$ . У нечіткій логіці з  $n$  входами і одним виходом міркування здійснюються за нечіткими правилами  $R_j$  вигляду:

$$R_j: \text{Якщо } K_1 \in A_{1,j} \wedge \dots \wedge K_n \in A_{n,j}, \text{ то } Q \in S_j, \quad (26)$$

де  $K_i, i = 1, \dots, n, Q$  — вхідні і вихідна лінгвістичні змінні;  $S_j, A_{i,j}$  — нечіткі вихідна множина і множини вхідної лінгвістичної змінної  $K_i$ .

Оскільки в СУІ накопичується багато даних моніторингу ІТ-інфраструктури і її компонентів, будувати ФН вхідних і вихідних нечітких змінних та нечіткі правила зведення показників оцінювання якості сервісів доцільно на основі навчальних вибірок даних. Потрібен відповідний алгоритм навчання.

Авторами розроблено підхід до побудови ФН для нечіткої системи з  $n$  входами і одним виходом і правилами вигляду (26) на основі множини числових навчальних вибірок  $P = \{(k_{1,j}, \dots, k_{n,j}, q_j) \mid j = 1, \dots, m\}$  з  $n$  входами і одним виходом, де  $k_{i,j}$  — значення  $i$ -ї вхідної лінгвістичної змінної  $K_i$ , які складають множину  $(k_{1,j}, \dots, k_{n,j})$ ;  $q_j$  — відповідне значення вихідної лінгвістичної змінної  $Q, 1 \leq i \leq n$  і  $1 \leq j \leq m$  [42–45].

### Система управління ІТ-інфраструктурою

Розроблені авторами методологічні основи проектування, впровадження та експлуатації СУІ реалізовано в багатьох реальних об'єктах управління [4, 46–48]. Створення СУІ базується на ІТ управління ІТ-інфраструктурою. Архітектура інструментальних засобів ІТ побудована на принципах тривірневої клієнт-серверної технології, а реалізована вона на платформі швидкого розроблення і експлуатації застосувань SmartBase [49–53]. СУІ складається з трьох

програмних модулів — серверного, клієнтського та агентського. ІТ управління ІТ-інфраструктурою реалізує п'ять процесів — моніторингу, аналізу, управління оптимізації та планування. Кожний з процесів реалізується програмним модулем, що дозволяє нарощувати їх функціонал без внесення істотних змін в інші модулі, створювати версії СУІ зі скороченим функціоналом [4].

Модуль моніторингу визначає стан ІТ-інфраструктури та її елементів, використовуючи статистичний, реактивний і проактивний методи. Модуль аналізу визначає стан елементів ІТ-інфраструктури, забезпечує потрібною інформацією модулі управління, оптимізації і планування. Модуль управління вибирає управляючі дії на основі політик з урахуванням існуючих обмежень, реалізує функції координатора. Модуль оптимізації розподіляє та перерозподіляє ресурси ІТ-інфраструктури, а також вирішує завдання системної оптимізації у автоматизованому або автоматичному режимі роботи. Реалізація рішення виробляється зміною конфігурації ІТ-інфраструктури, в модуль управління надходить новий план розподілу ресурсів. Модуль планування вирішує завдання розвитку ІТ-інфраструктури. Кожен з модулів реалізує свої функції для всіх рівнів ієрархії ІТ-інфраструктури.

### Впровадження результатів

Тематика робіт, спрямованих на розроблення моделей, методів і технологій управління ІТ-інфраструктурою включена в науково-технічні плани Наукового парку «Київська політехніка», навчально-наукового центру «НТУУ «КПІ»–Неткрекер», КНУ імені Тараса Шевченка, НАУКМА, ІК ім. В.М.Глушкова. Робота виконувалась в рамках ряду держбюджетних та договірних науково-дослідних та дослідно-конструкторських робіт.

Результати роботи були використані при створенні ряду систем спеціального призначення. ІТ управління ІТ-інфраструктурою та інші результати використані при проектуванні та реалізації підсистеми управління інфраструктурою ТОВ «БМС Консалтинг», при розробленні системи управління ІТ-інфраструктурою ТОВ «ІНЛАЙН ГРУП ЗАХІД», в системі управління процесами діяльності ТОВ «САП Україна», при розробленні підсистеми управління експлуатацією системи управління інформаційно-телекомунікаційною мережею ТОВ «Система мобільних платежів України», при розробленні

системи управління функціонуванням інформаційно-телекомунікаційної інфраструктури ТОВ «Спеціалізовані інформаційні послуги». Практичні результати впровадження:

1. Результати досліджень можуть бути використані при розробленні як СУІ, так і окремих підсистем і модулів.

2. Використання положень і результатів роботи дозволило на 15–30% скоротити час на розробку на проектування і реалізацію СУІ.

3. Застосування декомпозиційно-компенсаційного підходу до управління рівнем послуг і методів імпаکت-аналізу дозволило скоротити час аналізу до 1 хв. і скарги користувачів на якість послуг на 30%, а витрати на експлуатацію серверів зменшити на 8–10%.

4. Впровадження моделей і методів управління ресурсами дозволило підвищити на 8–15% ефективність використання ресурсів, що виділяються для надання послуг користувачам, або обслуговувати на 15–20% більше користувачів наявними парком серверів із збереженням рівня обслуговування, скоротити витрати на експлуатацію серверного парку на 10–15%.

5. Використання методів, моделей та інструментів управління рівнем якості ІТ-послуг дозволяє збільшити на 12% дохід від надання послуг завдяки зменшенню відтоку клієнтів і витрат від штрафів за неякісне надання сервісів.

6. Методи аналізу дозволяють на 20% зменшити кількість тестових перевірок при виявленні несправних елементів ІТ-інфраструктури, на 10% знизити витрати на обслуговування, а також підвищити значення коефіцієнта готовності сервісу.

### Висновки

Розроблена і обґрунтована цілісна методологія управління складним об'єктом управління – сучасними глобальними, національними і корпоративними ІТ-інфраструктурами.

Запропоновано ряд нових методів імпакт-аналізу, управління потоками в мережах, нечіткого зведення метрик та ін.

Запропоновано керований генетичний алгоритм для розв'язання широкого класу задач управління ІТ-інфраструктурою.

На основі отриманих теоретичних результатів розроблено сучасну ІТ управління ІТ-інфраструктурами.

Впровадження розроблених на основі ІТ управління ІТ-інфраструктурами СУІ на багатьох об'єктах управління продемонструвало їх переваги як в процесі розроблення (скорочення витрат і термінів), так і в процесі експлуатації (підвищення якості послуг, зменшення витрат, збільшення доходів тощо).

### Список літератури

1. Крывый С.Л. Верификация программных систем. 1/ С.Л. Крывый, А.Н. Максимец // Кибернетика и системный анализ. – 2013. – № 6. – С. 3–12.
2. Крывый С.Л. Верификация программных систем. 2/ С.Л. Крывый, А.Н. Максимец // Кибернетика и системный анализ. – 2014. – № 1. – С. 14–21.
3. Ершов С.В. Мультиагентные модели на основе нечеткой логики высшего типа для высокопроизводительной среды / И.Н. Парасюк, С.В. Ершов // Проблемы програмування. – 2012. – № 2–3. – С. 260–269.
4. Ролик А.И. Тенденции и перспективы развития управления информационными технологиями / А.И. Ролик // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «ВЕК+», 2012. – № 55. – С. 81–109.
5. Боданюк М.Є. Управління системами збереження даних / М.Є. Боданюк, О.К. Карнаухов, О.І.Ролік, С.Ф.Теленик // Електроніка та зв'язок. – №5 (76). –2013. – С.81-90.
6. Павлов А.А. Информационные технологии и алгоритмизация в управлении / А.А. Павлов, С.Ф. Теленик // Киев. – Техника, 2002. – 344 с.
7. Теленик С.Ф. Военно-економічна безпека: управлінські та технічні аспекти / С.Ф. Теленик, Л.П. Артеменко, А.Є. Довгань, П.І. Терещенко// Київ: НТУУ «КПІ», 2008.– 476 с.
8. Погорілий С.Д. Методи кластерних обчислень // С.Д. Погорілий, Ю.В. Бойко, Р.І. Левченко, В.А. Мар'яновський/ За ред С. Д. Погорілого. – К: ВПЦ «Київський університет», 2013. – 416 с.
9. Погорілий С.Д. Технологія віртуалізації. Динамічна реконфігурація ресурсів обчислювального кластера / С.Д. Погорілий, І.В. Білоконь, Ю.В. Бойко// Математичні машини і системи. – №3. – 2012. – С. 3–18.
10. Крывый С.Л. Вступ до методів створення програмних продуктів/ С.Л. Крывый // Чернівці: Букрек. – 2012. – 423 с.
11. Ершов С.В. Трансформационный поход к разработке программных архитектур на основе нечетких графовых моделей / И.Н. Парасюк, С.В. Ершов // Кибернетика и системный анализ. – 2008. – № 5. – С. 139–150.

12. Ролик А.И. Концепция управления корпоративной ИТ-инфраструктурой / А.И. Ролик // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «ВЕК+», 2012. – № 56. – С. 31–55.
13. Ролик А.И. Декомпозиционно-компенсационный подход к управлению уровнем услуг в корпоративных ИТ-инфраструктурах / А.И. Ролик // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка: Зб. наук. пр. – К.: Век+, – 2013. – № 58. – С. 78–88.
14. Ролик А.И. Управление уровнем услуг корпоративной ИТ-инфраструктуры на основе координатора / Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: Век+. – 2013. – № 59. – С. 98–105.
15. Ролик А.И. Методика оперативного управління рівнем ІТ-послуг на основі сервісно-ресурсних моделей// О.І. Ролік // Вісник НТУУ «КПІ», Інформатика, управління та обчислювальна техніка». – №60. –2014. – С.119–129.
16. Теленик С.Ф. Логічний підхід до інтеграції програмних застосувань підтримки міждисциплінарних наукових досліджень / С.Ф.Теленик, О.А.Амонс, К.В.Єфремов, В.Т.Лиско // Наукові вісті НТУУ «КПІ». – №5 (91). – 2013. – С.53–72.
17. Теленик С.Ф. Створення гетеросистемного комплексу для поєднання різноманітних програмних систем на різних платформах / С.Ф. Теленик, Л.С. Глоба, О.В. Турчинов, Т.І. Кривець // Електроніка і зв'язок. – №20. – 2003.
18. Теленик С.Ф. Логика действий для интеллектуальных систем // Математические машины и системы. – 1999. – №2. – С.12–24.
19. Теленик С.Ф. Логика организации взаимодействия компонентов в адаптивной технологии // Системные технологии. Системы и процессы обработки информации и управления: Сб. науч. тр. – Вып.7.– Днепропетровск, 1999. – С.119–129.
20. Теленик С.Ф. Логика представления вычислительных процессов в интеллектуальной системе SmartBase // Системные технологии. Системное моделирование технологических процессов: Сб. науч. тр.– Вып.6.– Днепропетровск, 1999. – С.131–139.
21. Теленик С.Ф. Интеллектуальные системы и логики высших порядков // Радиоэлектроника, информатика, управление.– 1999.– №1. – С.96–105.
22. Теленик С.Ф. Технологія управління ІТ-інфраструктурою на основі ресурсного підходу/ С.Ф. Теленик, О.І. Ролік, М.М. Букасов // Вісник ЖДТУ. – 2008.– № 4 (47). – С. 180–189.
23. Telenyk S.F. Resource allocation and load management considering the assessed quality of the provided service and the use of agent technologies / S.F. Telenyk, A.I. Rolik, A.A. Pokotylo // 23rd International Crimean Conference “Microwave & Telecommunication Technology” (CriMiCo’2013). 9–13 September, Sevastopol, Crimea, Ukraine. – 2013. – p. 535–536.
24. Теленик С.Ф. Забезпечення процесів діяльності ресурсами з визначеним рівнем надійності в інформаційно-телекомунікаційній системі спеціального призначення / С.Ф. Теленик, О.І. Ролік, П.І. Терещенко, М.М. Букасов // Зб. наук. праць ВІПІ НТУУ «КПІ».–№ 3.–Київ:ВІПІ НТУУ «КПІ», 2006. – С. 134–138.
25. Теленик С.Ф. Управління доступом до обмежених ресурсів інформаційно-телекомунікаційної мережі АСУ військового призначення / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, П.І. Терещенко // Зб. наук. праць Центрального НДІ Збройних Сил України. – 2006.– №3 (37). – С. 33–43.
26. Теленик С.Ф. Моделі управління розподілом обмежених ресурсів в інформаційно-телекомунікаційній мережі АСУ/ С.Ф. Теленик, О.І. Ролік, М.М. Букасов. // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: Єкотех, 2006. – № 44.– С. 234–239.
27. Теленик С. Управління ресурсами центрів оброблення даних / С. Теленик, О. Ролік, М. Букасов, К. Крижова // Вісник Львів. ун-ту. Серія прик. матем. інформ. – Вип. 15. – 2009. – С. 325–340.
28. Теленик С.Ф. Управління навантаженням і ресурсами центрів оброблення даних при виділених серверах / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, Р.В. Римар, К.О. Ролік // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – 2009. – №2 (24). – С. 122–136.
29. Теленик С. Управління навантаженням і ресурсами центрів оброблення даних при віртуальному хостингу / С. Теленик, О. Ролік, М. Букасов, С. Андросов, Р. Римар // Вісн. Тернопільського держ. техн. ун-ту. – 2009. – Том 14. – № 4. – С. 198–210.
30. Теленик С.Ф. Моделі і методи розподілу ресурсів в системах з серверною віртуалізацією / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, О.А. Косован., О.І. Кобець // Зб. наук. праць ВІПІ НТУУ «КПІ». – Випуск № 3. – Київ: ВІПІ НТУУ «КПІ», 2009. – С. 100–109.
31. Теленик С.Ф. Моделі управління віртуальними машинами при серверній віртуалізації / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, А.Ю. Лабунський // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «ВЕК+», 2009. – № 51. – С. 147–152.
32. Telenyk S. Qualitative evaluation method of IT-infrastructure elements functionin / S. Telenyk, Y. Dorogiy, O. Rolick, D. Halushko, M. Bukasov, A. Pysarenko // Proc. of 2014 IEEE International Black Sea Conference on Communications and Networking (BlackSeaCom 2014), Chisinau, Moldova, May 27–30, 2014. – pp.165–169.

33. Telenyk S. Models and methods of resource management for VPS hosting / S. Telenyk, O. Rolik, M. Bukasov, D. Halushko//Technical transaction. Automatic control.–Politechnica Krakowska, 2013.–vol. 4-AC. – pp. 41–52.
34. Теленик С.Ф. Генетичні алгоритми вирішення задач управління ресурсами і навантаженням центрів оброблення даних / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, С.А. Андросов // Автоматика. Автоматизація. Електротехнічні комплекси та системи. – 2010. – №1 (25). – С. 106–120.
35. Теленик С. Ф. Управляемый генетический алгоритм в задачах распределения виртуальных машин в ЦОД / С. Ф. Теленик, О. І. Ролік, П. С. Савченко, М. Е. Боданюк // Вісник ЧДТУ. – 2011. – №2. – С.104–113.
36. Теленик С.Ф. Адаптивный генетический алгоритм для решения класса задач распределения ресурсов ЦОД / С.Ф. Теленик, А.И. Ролик, П.С. Савченко // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «БЕК+», 2011. – № 54. – С. 164–174.
37. Теленик С.Ф. Метод определения влияния неполадок на качество сервиса в IP-сетях с протоколами RIP и OSPF/ С.Ф. Теленик, А.И. Ролик, М.В. Ясочка, М.Е. Бутенко // Проблеми інформатизації та управління: Зб. наук. праць: Вип. 1 (37). – К.: НАУ, 2012. – С. 118–126.
38. Теленик С.Ф. Аналіз підходів до визначення впливу неполадок елементів мережі доступу на якість обслуговування клієнтів/ С.Ф. Теленик, О.І. Ролік, М.В. Ясочка, К.С. Голоднов // Вісник Вінницького політехнічного інституту. – 2011. – №3. – С. 69–83.
39. Теленик С.Ф. Определение распространения влияния неисправностей в сети доступа на качество предоставляемых сервисов / С.Ф. Теленик, А.И. Ролик, М.М. Букасов, М.В. Ясочка // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «БЕК+», 2009. – № 50. – С. 164–173.
40. Теленик С.Ф. Швидкий алгоритм визначення впливу неполадок елементів мережі доступу на якість обслуговування клієнтів / С.Ф. Теленик, О.І. Ролік, М.В. Ясочка, Д.В. Малюгін // Інформаційні технології та комп'ютерна інженерія. — 2010.— №2 (18).— С. 51–62.
41. Ролик А.И. Система управления информационными потоками в корпоративной IP-сети / А.И. Ролик, В.А. Иосифов // Адаптивні системи автоматичного управління. Міжвідомчий науково-технічний збірник. – Дніпропетровськ: Системні технології, 2009.– Вип. 14 (34).– С. 73–85.
42. Теленик С.Ф. Нечітке оцінювання в задачах управління рівнем обслуговування / С.Ф. Теленик, О.І. Ролік, М.В. Ясочка, О.М. Моргаль // Наукові записки УНДІЗ. – 2011. – № 2 (18). – С. 29–42.
43. Теленик С.Ф. Зведення метрик оцінювання рівня обслуговування користувачів на основі експертних оцінок / С.Ф. Теленик, О.І. Ролік, О.М. Моргаль, О.С. Квітко// Вісник Вінницького політехнічного інституту. – 2011. – №1. – С. 112–123.
44. Теленик С.Ф. Метод розподілу ресурсів між проектами / С.Ф. Теленик, П.І. Бідюк, О.А. Амонс, К.О. Крижова // І Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: «БЕК+» – 2008. – № 48. – С. 33–41.
45. Теленик С.Ф. Метод формування логічного висновку із залученням експертного комітету / С.Ф. Теленик, П.І. Бідюк, Л.О. Коршевнок, В.С. Хмелюк // Проблеми програмування. – 2008.– № 4. – С. 73 – 83.
46. Теленик С.Ф. Система управління інформаційно-телекомунікаційною системою корпоративної АСУ / С.Ф. Теленик, О.І. Ролік, М.М. Букасов, Р.Л. Соколовський // Вісник НТУУ «КПІ». Інформатика, управління та обчислювальна техніка. – К.: «БЕК+», – 2006. – № 45.– С. 112–126.
47. Теленик С.Ф. Модель управління розподілом ресурсів інформаційно-телекомунікаційної системи збройних сил України / С.Ф. Теленик, О.І. Ролік, П.І. Терещенко, О.В. Літвінцов // Наук.-техн. зб. ННДЦ оборонних технологій і воєнної безпеки України. – 2006. – №5 (34). – С. 117–124.
48. Ролик А.И. Система управления корпоративной информационно-телекоммуникационной инфраструктурой на основе агентского подхода / А.И. Ролик, А.В. Волошин, Д.А. Галушко, П.Ф. Можаровский, А.А. Покотило // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка. – К.: «БЕК+», 2010. – № 52. – С. 39–52.
49. Теленик С.Ф. Семантична інтеграція різнорідних інформаційних ресурсів / С.Ф.Теленик, О.А.Амонс, К.В.Ефремов, С.В.Жук // Вісник НТУУ «КПІ», Інформатика, управління та обчислювальна техніка». – №58. –2013. – С.29-45.
50. Теленик С.Ф. Проблеми розробки адаптивних технологій створення інформаційно-керуючих систем Науковій вісті НТУУ “КПІ”. – 2000. – № 1. – С.47–57.
51. Теленик С.Ф. Алгебро-логическая формализация в объектно-ориентированных технологиях/ С.Ф. Теленик, Г.Е. Цейтлин, А.А. Амонс//Проблемы программирования.–2002.–№1–2. Спец. Выпуск. – С. 136–146.
52. Матрична резолюція для клаузальних логік / С.Ф. Теленик, О.А. Амонс, Р.В. Смічик, В.С. Хмелюк // Вестник ХНАДУ. Сб. науч. тр. Вып.28, 2003. – С. 243–46.
53. Теленик С.Ф. Трансформационные преобразования и синтез классов алгоритмов / С.Ф. Теленик, Г.Е. Цейтлин, А.А. Амонс // Вісник НТУУ «КПІ»: Інформатика, управління та обчислювальна техніка, К., 2004, № 42.– С. 20–27.