

СПОСОБИ ВИКОРИСТАННЯ ВІРТУАЛІЗАЦІЇ У ГРІДІ

Розглянуто основні методи застосування віртуальних машин для роботи у середовищі ґрід. Надано опис декількох моделей та проведено аналіз доцільності їхнього використання. Наведено переваги та недоліки віртуалізації ресурсів для ґрід. Запропоновано альтернативний варіант запуску завдань через ґрід з використанням віртуальних машин.

Ключові слова: віртуалізація, ґрід, кластер, обчислювальний вузол, мережа, планувальник, задача, користувач, віртуальна машина.

Останнім часом застосування технологій віртуалізації набуло широкого розповсюдження, адже вони дозволяють відокремити робочі середовища користувачів, програм, операційних систем, зробити їх незалежними одне від одного, при цьому використано один фізичний комп'ютер. Це дає змогу підвищити рівень використання ресурсів та надати більше свободи кінцевим користувачам.

Віртуалізацією можна вважати прозору для користувача підміну певного апаратного чи програмного ресурсу. Віртуальні мережі дозволяють використовувати територіально чи адміністративно розподілені ресурси так, ніби вони безпосередньо під'єднані одне до одного, або, навпаки, виділити декілька мережевих інтерфейсів з одного фізичного. Віртуальні середовища виконання для програм забезпечують прозоре перенесення програм між різними операційними системами або їх версіями.

У середовищі високопродуктивних обчислень найбільш розповсюдженою є віртуалізація апаратного забезпечення для поділу його ресурсів або спільного їх використання. Її застосування має декілька важливих переваг. Незалежна конфігурація програмного забезпечення надає кожному користувачу можливість за бажанням налаштувати необхідну йому версію програми, встановлювати доповнення чи змінювати код. Завдяки вбудованим функціям віртуальних машин їх виконання може бути призупинене, що забезпечує легкість збереження проміжного стану при виконанні програм, які не мають підтримки контрольних точок [1]. Чітке розмежування між користувачами виділених ресурсів забезпечує неможливість використати більше, ніж виділено віртуальній машині.

Метою даної статті є огляд і аналіз існуючих способів використання віртуальних машин, визначення їх переваг та недоліків. Зокрема, застосування віртуальних машин для розгортання ґрід-сервісів та створення віртуальних кластерів на базі фізичних, а також використання клауд-комп'ютіну для ґрід. Також буде запропоновано альтернативний варіант застосування — запуск віртуальних машин як задач у ґріді — завданням якого є поєднання переваг попередників.

Віртуальні машини для розгортання сервісів. Мережа ґрід поєднує багато сервісів, які забезпечують реєстрацію та облік користувачів, розміщення та зберігання даних користувачів, планування та запуск завдань, надання інформації про ресурси тощо [2]. Кожен ґрід-сайт має встановити та налаштувати певну частину цих сервісів для забезпечення своєї інтеграції у ґрід. Зазвичай, сервіси мають бути встановлені на окремих серверах, що забезпечить їхнє незалежне функціонування. Важливою є також можливість окремо оновити або перезапустити сервіси. Таке використання вузлів може призвести до простою виділених під сервіси ресурсів, адже залежно від кількості користувачів їхня робота може потребувати лише незначних апаратних потужностей.

Для вирішення цих питань доцільним є розміщення ґрід-сервісів та служб на постійно працюючих віртуальних машинах. Декілька вузлів виділяється з кластера для розміщення ґрід-сервісів. На них встановлено програмне забезпечення для керування віртуальними машинами — гіпервізор [3] — та створено відповідні віртуальні машини. Кількість зайнятих фізичних вузлів при цьому, зазвичай, буде значно меншою за кількість віртуальних машин. На рис. 1 зображено схему роботи кластера

у цьому випадку та потік грід-задачі від користувача до обчислювального вузла. Як видно, задача та дані користувача через сервіси на віртуальних машинах, основними з яких є обчислювальний елемент (ОЕ) та зберігаючий елемент (ЗЕ), потрапляють до планувальника кластеру, а потім на фізичні обчислювальні вузли, де розпочинають виконання.

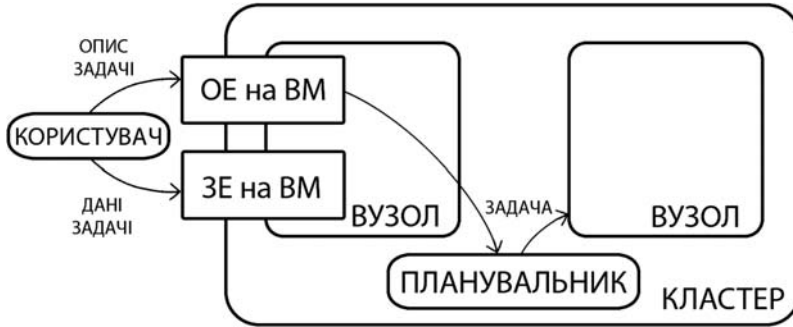


Рис. 1 — Рух задачі користувача через грід-сервіси на віртуальних машинах

Перевагами такої реалізації є економія ресурсів у разі незначного завантаження грід-сайту, а також незалежність версій програмних пакетів, необхідних для функціонування сервісу. До того ж, стандартні функції гіпервізорів дозволяють зберігати певний стан, наприклад, перед виконанням оновлення і відновлювати його, якщо виникла така потреба. При апаратному збої вузла є можливість перенести образ віртуальної машини на інший вузол для якнайшвидшого відновлення роботи.

Як недолік можна виділити те, що сервіси, які обслуговують великі гріди, потребують значних ресурсів, і це може означати, що один сервіс буде потребувати апаратних ресурсів цілого фізичного вузла. При цьому ще й будуть додаткові витрати на забезпечення віртуалізації.

Віртуальні кластери на базі фізичних. При виділенні ресурсів кластера для певного проекту може виникнути необхідність створення повної кластерної інфраструктури відокремленої та закритої від іншої частини кластера чи інших інфраструктур. При цьому на базі фізичних ресурсів створюється набір сервісів та робочих вузлів на віртуальних машинах [4]. Кожен проект отримує свою точку входу, планувальник тощо.

На рис. 2 зображено розміщення віртуальних вузлів на фізичних. Проект чи віртуальна організація отримує певну кількість фізичних вузлів кластера, які виокремлюються із загального пулу. На них розгортаються віртуальні машини з грід-сервісами, планувальником кластера та робочими вузлами. Користувачі взаємодіють як локально, так і через грід, з віртуальним інтерфейсним вузлом та виконують задачі на віртуальних робочих вузлах, проте для них це є цілком прозорим, тобто так, ніби вони працюють на фізичних ресурсах.

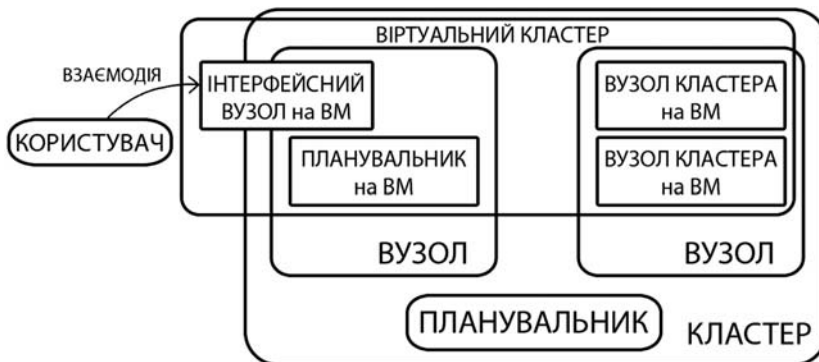


Рис. 2 — Розміщення віртуального кластера на фізичних вузлах

До переваг цієї реалізації можна віднести те, що проект отримує ресурси згідно з виділеною йому часткою у повному обсязі і вільно ними розпоряджається. Є можливість змінювати кількість та потужність віртуальних обчислювальних вузлів у межах виділених фізичних серверів, встановлювати параметри віртуальних машин згідно зі своїми потребами. Налаштування окремого планувальника для свого проекту дає можливість розподілити ресурси між членами віртуальної організації чи виконавцями проекту за власним бажанням незалежно від локальних політик кластера.

Як недолік слід зазначити необхідність витрати проектом певної частки отриманих ресурсів на розгортання сервісів, проте вона буде досить малою, адже ці сервіси обслуговуватимуть незначну кількість користувачів. Також є необхідність адміністрування з боку виконавців проекту, проте воно може бути за домовленістю виконано адміністраторами кластера.

Клауд-комп'ютинг через грід. У цьому випадку під клауд-комп'ютингом розуміється виділення користувачу ресурсів для запуску створених ним образів віртуальних машин та доступу до машин, які він запустив [5]. Таким чином, користувач надсилає на кластер не завдання, а лише запит на необхідні ресурси та вказує, який саме образ віртуальної машини на них розгорнути. Після виділення ресурсів користувач має доступ всередину віртуальної машини та може використовувати її на свій розсуд. Завершивши роботу, можна зберегти зміни в образі для подальшого використання. Початкові образи віртуальних машин загальнодоступні та, зазвичай, містять мінімальну установку певної операційної системи.

На рис. 3 зображено розміщення віртуальних машин (ВМ) та взаємодію користувача. Запит на ресурси надходить до планувальника, який виділяє необхідні ресурси, якщо є така можливість. Під час розгортання вказаного образу також налаштовується мережеве з'єднання, яке робить віртуальну машину доступною ззовні кластера, після чого користувач отримує параметри доступу до своєї віртуальної машини.

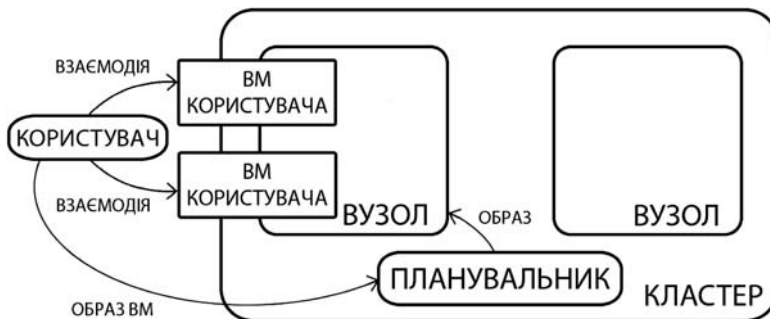


Рис. 3 – Виділення віртуальної машини користувачу

На рис. 4 показано послідовність команд при взаємодії з кластером через Nimbus. Будучи побудованим на базі GlobusToolkit [6], для аутентифікації на кластері він дозволяє використовувати грідівську систему проксі-сертифікатів. Як видно, користувач пересилає та запускає образ віртуальної машини, а потім використовує протокол віддаленого доступу для роботи. В кінці образ зберігається.

Основною перевагою цього підходу є повна свобода дій користувача всередині віртуальної машини, що дає йому можливість налаштувати робоче середовище оптимально для його задач. Однак значним недоліком є складність планування задач [7]. Оскільки взаємодія з користувачем має інтерактивний характер, то ресурси мають бути виділені або миттєво, або у зазначений користувачем період, тоді як кластерний планувальник зазвичай ставить задачу у чергу, в якій вона очікує в загальному випадку невизначений час на звільнення ресурсів.

Запуск віртуальних машин як задач у гріді. На відміну від клауду, грід не передбачає інтерактивної роботи, тому виділення користувачам ресурсів для віртуальних машин не обмежуватиметься зазначеними вище проблемами з плануванням. Саме тому пропонуємо реалізувати запуск задач через грід для їх подальшого виконання як віртуальних машин. При цьому основною метою є забезпечення роботи

проміжного програмного забезпечення грід та планувальника кластеру як з віртуальними машинами, так і зі звичайними задачами.

```
$ grid-proxy-init
$ cloud-client.sh --transfer --sourcefile /home/user1/my_vmimage
$ cloud-client.sh --run --name my_vmimage --hours 1
$ ssh root@hostname.cluster.edu
$ cloud-client.sh --save --handle vm-1 --newname my_vmimage_new
```

Рис. 4 — Команди для взаємодії з сервером Nimbus

На рис. 5 показано послідовність переміщення задачі при цій реалізації. Як і у випадку звичайної грід-задачі, користувач завантажує дані задачі на зберігаючий елемент (ЗЕ) гріду. Додатково має бути завантажено образ віртуальної машини (ВМ), якщо його до цього не було завантажено при виконанні попередніх задач. Опис задачі відправляють звичайним шляхом до планувальника гріду, який визначає підходящий обчислювальний елемент (ОЕ) — такий, що має доступ до ЗЕ з даними, після чого ОЕ відправляє задачу до черги локального планувальника кластеру. Після виділення ресурсів на вузлі запускається віртуальна машина користувача і до неї приєднується папка з даними задачі [8]. Наданий користувачем скрипт запуску виконується всередині віртуальної машини.

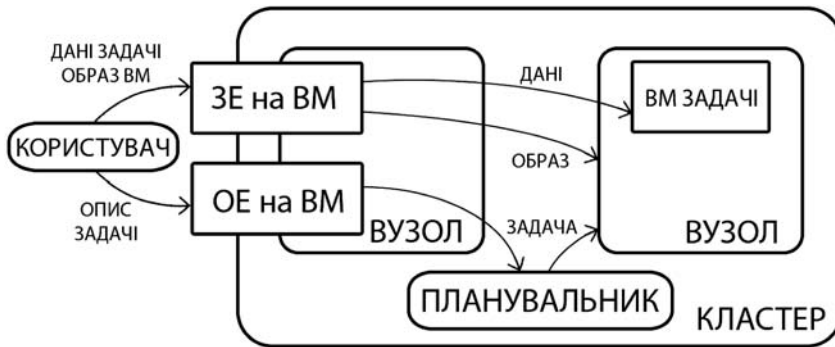


Рис. 5 — Рух задачі при використанні віртуальних машин як задач у гріді

Для реалізації цього підходу необхідно доповнити файл грід-завдання додаванням туди опцій вибору віртуальної машини та перевіркою наявності образу на ресурсах (рис. 6), модифікувати скрипти планувальника з додаванням можливості розподіляти віртуальні машини та створити скрипти для роботи з віртуальними машинами на робочих вузлах. Також потрібно використовувати окремий сервер для початкового налаштування віртуальних машин користувачами.

```
InputSandbox = {"myscript.sh"};
DataRequirements = {
  [ InputData = {"lfn:/grid/vo_name/my_vmimage"};
    DataCatalogType = "DLI";
    DataCatalog = "http://cluster.edu:8085"; ]
};
DataAccessProtocol = {"rfio", "gsiftp"};
```

Рис. 6 — Перевірка наявності образу на зберігаючому елементі

Перевагами цього рішення є те, що для запуску завдань та передачі файлів використано звичайне проміжне програмне забезпечення гріду, а планувальник працює з завданнями на запуск віртуальних машин, як із звичайними завданнями. При цьому залишились можливості збереження проміжного стану задачі, тимчасової зупинки виконання та переміщення задачі на інший вузол або кластер. Так само як і у випадку клауд-комп'ютіngu, користувач отримує свободу в налаштуванні робочого середови-

ща та незалежність від задач інших користувачів. Проте немає необхідності у налаштуванні зовнішнього мережевого інтерфейсу, що значно полегшує адміністрування.

До недоліків слід віднести непристосованість до запуску паралельних завдань на декілька вузлів. Такі завдання втрачатимуть на додаткових затримках, внесених віртуалізацією, при обміні даними. До того ж використання віртуалізації втрачає свої переваги через високу складність, а часто й неможливість, синхронної зупинки віртуальних машин на різних вузлах зі збереженням їхнього стану з урахуванням незавершених обмінів даними. Також будуть певні втрати при розгортанні та збереженні стану віртуальних машин, що при короткотривалих задачах суттєво зменшуватиме продуктивність.

Висновки. Використання віртуалізації дає змогу користувачу за допомогою адміністратора одного з кластерів один раз сформувати свій образ для віртуальної машини з власними налаштуваннями та використовувати його на будь-яких кластерах, які підтримують цю технологію, як локально, так і через грид. Також віртуальні машини дають можливість зупиняти своє виконання зі збереженням стану та відновлювати роботу як на тому самому, так і на іншому обчислювальному ресурсі.

При використанні віртуальних машин виникають і проблеми. Менеджери віртуальних машин певним чином змінюють налаштування операційної системи та її ядра. При цьому такі налаштування можуть бути несумісними з тим апаратним та програмним забезпеченням, що зараз використовується на кластерах.

Оскільки всередині віртуальної машини користувач має права адміністратора, то він може використовувати мережу для запуску певних небажаних програм. Потрібно визначити безпечний порядок взаємодії з мережевими ресурсами.

Віртуальні машини використовуватимуться дедалі ширше. Слід вибирати модель використання, виходячи з власних потреб. Модель використання клауду, зазвичай, передбачає стягнення певної винагороди за користування необмеженими ресурсами (on demand), тоді як модель використання гриду передбачає певний справедливий розподіл обмежених ресурсів між користувачами (fairshare). Запропонованим рішенням намагаємося поєднати переваги обох моделей при роботі у гріді, забезпечити зручність як користувачам, так і адміністраторам ресурсів.

Рассмотрены основные методы применения виртуальных машин для работы в среде грида. Предоставлено описание нескольких моделей и проведен анализ целесообразности их использования. Приведены преимущества и недостатки виртуализации ресурсов для гриду. Предложен альтернативный вариант запуска заданий через грид с использованием виртуальных машин.

Ключевые слова: виртуализация, грид, кластер, вычислительный узел, сеть, планировщик, задача, пользователь, виртуальная машина.

The basic methods of virtual machines to run on the grid. Given of several models and analyzes the feasibility of their use. The advantages and disadvantages of virtualization for Grid. An alternative version of the jobs to run through the grid using virtual machines.

Key words: virtualization, grid, cluster, the compute node, the network scheduler, task, user, the virtual machine.

Література

1. Jones W. M. Using checkpointing to recover from poor multi-site parallel job scheduling decisions / W. M. Jones // ACM/IFIP/USENIX 8th International Middleware Conference. — 2007.
2. Foster I., Kesselman C. The grid: blueprint for a future computing infrastructure / I. Foster, C. Kesselman. // Morgan Kaufmann Publishers. — 1998.
3. Hagen W. Professional Xen Virtualization / W. Hagen // Wiley Publishing. — 2008.
4. Keahey K., Freeman T. Contextualization: Providing One-Click Virtual Clusters / K. Keahey, T. Freeman // eScience. — December 2008.
5. Keahey K. Cloud Computing with Nimbus / K. Keahey // Proceedings of CCA. — October 2009.
6. Globus Toolkit — <http://www.globus.org/toolkit/>
7. Sotomayor B. Combining Batch Execution and Leasing Using Virtual Machines / B. Sotomayor, K. Keahey, and I. Foster // HPDC. — June 2008.
8. Surhone L. Openvz / L. Surhone, M. Timpledon, S. Marseken // Betascript Publishing. — July 2010.