

Analiza porównawcza rozwiązań wysokiej dostępności

Michał Sylwester Borsewicz, Daniel Bieniek

Politechnika Lubelska, Instytut Informatyki, Nadbystrzycka 36B, 20-618 Lublin, Polska

Streszczenie. Celem niniejszej pracy jest przedstawienie analizy porównawczej wybranych rozwiązań wysokiej dostępności. Problematyka wyboru odpowiednich rozwiązań HA jest nieodzowną częścią każdej firmy korzystającej z infrastruktury IT. Jako, że komercyjnie najczęściej stosowane są rozwiązania wirtualizacyjne firm Microsoft: Hyper-V oraz VMWare: vSphere, to zostały one poddane dogłębniejszej analizie. W tym celu wykonano testy wydajnościowe systemów gości na powyższych hypervisorach oraz przetestowano działanie klastra w sytuacji awarii jednego z hostów.

Słowa kluczowe: klastrer; wysoka dostępność; Hyper-V; vSphere

*Autor do korespondencji.

Adres e-mail: borsewicz.michal@gmail.com

Comparative Analysis of High Availability Solutions

Michał Sylwester Borsewicz, Daniel Bieniek

Institute of Computer Science, Lublin University of Technology, Nadbystrzycka 36B, 20-618 Lublin, Poland

Abstract. The aim of the thesis was to perform a comparative analysis of selected high availability solutions. The issue of choosing the right HA solutions is an indispensable part of any company that uses IT infrastructure. As the most commonly used solutions are the virtualization solutions of Microsoft: Hyper-V and VMWare: vSphere, they have been subjected to in-depth analysis. For this purpose, we have performed performance tests of guest systems on the above hypervisors and we tested the behavior of the cluster in the event of failure of one of the hosts.

Keywords: cluster; high availability; Hyper-V; vSphere

*Corresponding author.

E-mail address: borsewicz.michal@gmail.com

1. Wstęp

Obecnie jednym z najważniejszych problemów, z którym muszą się zmierzać wszystkie firmy wykorzystujące rozwiązania IT, bez względu na ich wielkość, czy typy wykorzystywanych aplikacji jest zapewnienie nieprzerwanej dostępności do infrastruktury IT. Każdy przestój w działaniu aplikacji ukierunkowanych na obsługę klientów stanowi poważne straty finansowe, a zarazem może być przyczyną utraty zaufania przez klientów. Aby zapobiegać takim sytuacjom, gdy np. awaria serwera może spowodować przestój, stosuje się rozwiązania wysokiej dostępności tzw. High Availability (HA).

Dostępność to stopień, w jakim aplikacje, usługi lub funkcje są dostępne dla użytkowników. Stopień dostępności mierzony jest poprzez odczucia użytkownika aplikacji. Użytkownicy odczuwają frustrację, gdy ich dane są niedostępne lub system komputerowy nie działa zgodnie z oczekiwaniami. Zwykły, szary użytkownik komputera nie rozumie zależności pomiędzy komponentami składającymi się na działanie aplikacji i zazwyczaj wcale go to nie interesuje. Najważniejsze dla niego jest to, aby aplikacja na której musi pracować po prostu działała. W sytuacji, gdy użytkownik nie

może uzyskać dostępu do systemu, mówi się, że system jest niedostępny.

Celem niniejszego artykułu jest przedstawienie analizy porównawczej wybranych rozwiązań wysokiej dostępności czyli produktów Hyper-V oraz vSphere.

2. Analiza obecnego stanu wiedzy

Od początku lat 90 tematy wysokiej dostępności były często poruszane na łamach czasopism naukowych. J. Gray i D. Siewiorek w artykule [1], przedstawiają ówczesne techniki stosowane do budowy wysoce dostępnych systemów komputerowych począwszy od początku lat 70. Wtedy to elementy sprzętowe były głównym źródłem usterek i przestojów. Obecnie usterki sprzętu stanowią niewielki procent wszystkich awarii systemu. Teraz najczęstszymi przyczynami są błędy w oprogramowaniu i błędy środowiska. Dzisiaj na tę pozycję można patrzeć jedynie pod kątem historycznym.

Inną pozycją, która dobrze opisuje działanie i wykorzystanie klastrów, jest publikacja [2] autorstwa K. Kopper. Pomimo, że obecne rozwiązania są inne, to zasady działania pozostały podobne. Publikacja opisuje w głównej mierze działanie klastrów. Ukazuje darmowe rozwiązania

dotyczące wysokiej dostępności. Książka zawiera informacje o tym, jak połączyć parę serwerów aby uzyskać wysoką dostępność za pomocą pakietu Heartbeat, jak korzystać z oprogramowania do równoważenia obciążenia Linux Virtual Server, jak skonfigurować niezawodny system drukowania w środowisku klastra Linux, jak zbudować harmonogram zadań w systemie Linux.

W artykule autorstwa Prashanta K. D. [3], porównane są dwa opensourcowe rozwiązania wirtualizacyjne: hyperwizory Xen, KVM oraz komercyjne VMware vSphere (ESXi), Microsoft Hyper-V. Autor opisuje ich architekturę, sposoby implementacji.

W następnym artykule R. Alrobaea, A. Ghiduk, M. Ikbal [4] przeprowadzono analizę porównawczą wymienionych w tytule hyperwizorów. Porównywano w głównej mierze wpływ hyperwizorów na wydajność. Z przeprowadzonych eksperymentów wynika, że najlepsze wyniki wydajnościowe osiągnął KVM. Jedynie w teście wydajności systemu plików lepszy okazał się XEN. Autorzy stwierdzili, że KVM najlepiej nadaje się do aplikacji wymagających dużej mocy obliczeniowej i pamięci np. aplikacje internetowe. XEN natomiast lepiej nadaje się do aplikacji zależnych od systemu plików np. aplikacje bazodanowe będą lepiej pracować jeśli zostaną wdrożone na hyperwizorze XEN.

Granszewski i Arciszewski w artykule [5] porównali Hyper-V, ESXi, OVM, VirtualBox i Xen. W swoim artykule sprawdzili wpływ powyższych hyperwizorów na CPU, NIC oraz pamięć RAM. Ich wyniki ukazały, że najlepszym hyperwizorem pod względem wydajności jest ESXi.

W kolejnym ważnym artykule [6] autorstwa A. Bhatia, G. Bhattal przeprowadzono testy wydajnościowe dla hyperwizorów: VMware ESXi, Xen, Hyper-V i KVM. Autorzy wykorzystali w tym celu narzędzie SIGA oraz Passmark Software. W swoim artykule autorzy wykazali, że najlepszym rozwiązaniem pod względem wydajności jest ESXi.

W artykule [7] autorzy H. Fayyad-Kazan, L. Perneel, M. Timmerman przetestowali wydajność wymienionych w tytule hyperwizorów. Z przeprowadzonych badań autorom wyszło, że najlepszym z hyperwizorów jest XEN, gdy ostatnie miejsce zajął ESXi od VMWare.

Jak widać na powyższych przykładach, różni autorzy doszli do różnych konkluzji, jeżeli chodzi o testy wydajnościowe najpopularniejszych hyperwizorów.

Problematyka wyboru odpowiednich rozwiązań HA jest nieodzowną częścią każdej firmy korzystającej z infrastruktury IT. Analiza wybranych rozwiązań może okazać się pomocna dla administratora systemów czy baz danych. Przeprowadzenie analizy porównawczej dwóch najczęściej stosowanych rozwiązań wirtualizacyjnych, czyli VMware ESX 6.0.0 U3 oraz Microsoft Hyper-V 2019 Version 1809 pozwoli na pewno na lepsze zrozumienie ich działania. Poza tym będzie można odpowiedzieć na pytanie, które rozwiązanie wypadło lepiej w testach wydajności oraz czasu

reakcji potrzebnego na ponowny dostęp do danych w momencie awarii jednego z hostów.

3. Wstęp

Hyper-V to oprogramowanie stosowane do wirtualizacji fizycznych maszyn, którego autorem jest firma Microsoft. Występuje w dwóch wariantach: jako samodzielny produkt o nazwie Hyper-V Server lub jako doinstalowywany zewnętrzny komponent do systemu Windows Server. MS Hyper-V wymaga procesora ze sprzętową obsługą wirtualizacji. Wymagane jest, aby procesor był zgodny z technologią Intel VT lub AMD-V. Powinien również wspierać technologię SLAT (Second Level Address Translation) oraz DEP (Data Execution Prevention).

Hyper-V obsługuje izolację maszyn wirtualnych wykorzystując partycje. Partycja to logiczna jednostka izolacji obsługiwana przez hyperwizor, w którym uruchamiany jest każdy system operacyjny gościa. Jest to zasadniczo po prostu maszyna wirtualna. Oprogramowanie do wirtualizacji Hyper-V działa w partycji nadrzędnej i jest to jedyna maszyna wirtualna która posiada bezpośredni dostęp do zasobów sprzętowych. Partycja nadrzędna tworzy partycje podrzędne, które obsługują systemy gości. Wszystkie partycje gości, aby uzyskać dostęp do urządzenia muszą przejść przez partycję nadrzędną. Partycja podrzędna nie ma dostępu do procesora. Hyperwizor obsługuje przerwania do procesora i przekierowuje je do odpowiedniej partycji za pomocą SynIC. Hyper-V może przyspieszyć sprzętową translację adresów przestrzeni adresów wirtualnych gościa za pomocą translacji adresów drugiego poziomu, nazywanej EPT dla Intel i RVI (dawniej NPT) dla AMD [8]. Partycje podrzędne nie mają dostępu do zasobów sprzętowych, ale mają dostęp do zasobów urządzeń wirtualnych. Każde żądanie wysłane do urządzeń wirtualnych jest przekierowywane przez VMBus do urządzeń w partycji nadrzędnej, która zarządza żądaniami. VMBus umożliwia komunikację między partycjami. Odpowiedź jest również zarządzana przez VMBus. Jeśli urządzenia w partycji nadrzędnej są również urządzeniami wirtualnymi, zostaną przekierowane dalej, aż dotrą do partycji nadrzędnej, gdzie uzyskają dostęp do urządzeń fizycznych. Partycje nadrzędne uruchamiają VSP, który łączy się z VMBus i obsługuje żądania dostępu do urządzeń z partycji podrzędnych [9].

vSphere (ESXi) to hyperwizor typu 1 (tzw. bare metal), czyli jest instalowany bezpośrednio na sprzęcie systemowym. Został opracowany przez VMWare. Nazwa ESX powstała jako skrót od Elastic Sky X. ESXi zawiera i integruje ważne komponenty systemu operacyjnego, takie jak jądro VMKernel. Należy zaznaczyć, że Jądro Linuksa jest uruchamiane jako pierwsze, a następnie jest używane do ładowania innych składników wirtualizacji w tym m.in. VMKernel [10].

Inne najważniejsze produkty VMWare działające z ESXi to:

- vCenter Server - umożliwia monitorowanie i zarządzanie wieloma hostami ESX, ESXi i GSX. Ogólnie rzecz ujmując, Center Server umożliwia zarządzać platformą vSphere.

- vSphere Client - umożliwia monitorowanie i zarządzanie pojedynczą instancją serwera ESX lub ESXi.

4. Skalowalność Microsoft Hyper-V i VMWare vSphere

W tabeli 1 zostały porównane najważniejsze z zasobów dotyczących obydwu rozwiązań: Hyper-V oraz vSphere. Jak widać w pewnych aspektach lepiej wypada produkt VMWare – np. w maksymalnej liczbie procesorów logicznych oraz w liczbie wirtualnych CPU przypadających na hosta, jak i odwrotnie, w innych lepiej wypada produkt Microsoftu – np. w obsłudze maksymalnej pamięci fizycznej hosta, czy w liczbie wirtualnych CPU przypadających na maszynę wirtualną.

Oczywiście podczas wyboru konkretnego hyperwizora nie należy polegać wyłącznie na skalowalności. Ważne są również kwestie dostępnych funkcjonalności, kompatybilności z systemem operacyjnym, czy chociażby dostęp i profesjonalizm pomocy technicznej.

Tabela 1. Skalowalność zasobów Hyper-V 2019 i vSphere 6.7

| System | Resources | Microsoft Hyper-V 2019 | VMware vSphere 6.7 | | |
|---------|-----------------------|---|--------------------|----------------|-----------------|
| | | | Free Hypervisor | Essential Plus | Enterprise Plus |
| Host | Logical Processors | 512 | 768 | 768 | 768 |
| | Physical Memory | 24 TB | 4 TB | 4 TB | 16TB |
| | Virtual CPUs per Host | 2048 | 4096 | 4096 | 4096 |
| | VM per Host | 1024 | 1024 | 1024 | 1024 |
| | Nested Hypervisor | Yes | Yes | Yes | Yes |
| VM | Virtual CPUs per VM | 240 for Generation2 64 for Generation1 | 8 | 128 | 128 |
| | Memory per VM | 12 TB for Generation2 1 TB for Generation1 | 6128 GB | 6128 GB | 6128 GB |
| | Maximum Virtual Disk | 64 TB for VHDX format 2040 GB for VHD format | 62 TB | 62 TB | 62 TB |
| | Number of disks | 256 (SCSI) | 256 (SCSI) | 256 (SCSI) | 256 (SCSI) |
| Cluster | Maximum Nodes | 64 | 64 | 64 | 64 |
| | Maximum VMs | 8000 | 8000 | 8000 | 8000 |

5. Metodyka badań

W celu wykonania dokładnego porównania obu hyperwizorów użyto dwóch serwerów:

HPE MicroServer Gen8 o następujących parametrach:

- Procesor: Intel® Xeon® Processor E3-1265L v2 – 4 rdzenie i 8 wątków o taktowaniu 2,50 GHz i 3,50 GHz w trybie turbo, 8MB pamięci cache L3;
- Pamięć RAM: 12 GB DDR3 1Rx8 PC3L 12800E-11-13-D1; Unbuffered ECC (8GB + 4GB);
- Karty sieciowe: 2 x Ethernet 1 Gbit + 1x port iLO;
- Dysk: 240GB SSD MTL;

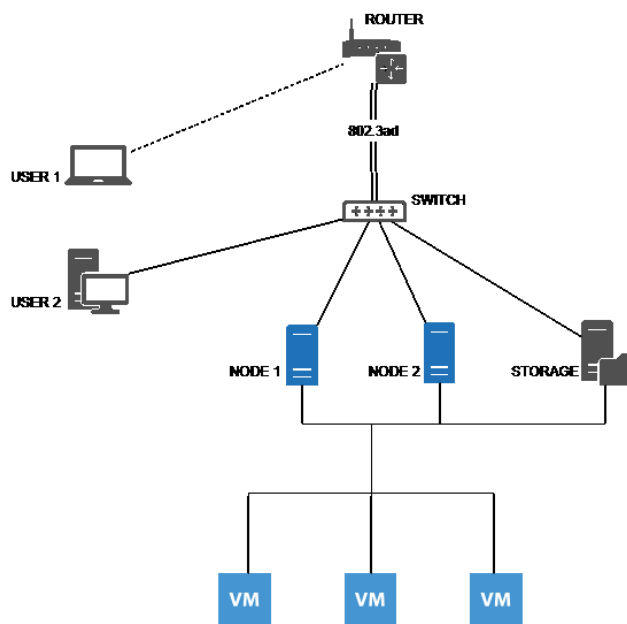
Dysk został wykorzystany jako storage do instalacji poszczególnych nadzorców.

Drugim serwerem został Fujitsu PRIMERGY TX120 S3P, który służył jako iSCSI Target Server:

- Procesor: Intel Celeron G1610T 2.3 GHz- 2 rdzenie i 2 wątki o taktowaniu 2,30 GHz, 2MB pamięci cache L3
- Pamięć RAM: 8 GB DDR3 1Rx8 PC3L 12800E-11-13-D1; Unbuffered ECC (2 x 4GB);
- Karty sieciowe: 2 x Ethernet 1 Gbit + 1x port iRMC;
- Karta RAID:LSI SAS 1064E.

Na dysku SSD została zainstalowana dystrybucja Linuxa odpowiedzialna za udostępnianie dysków po protokole ISCSI. Dyski SAS 2 x 300GB zostały skonfigurowane z RAID0 i przeznaczone dla maszyn wirtualnych, natomiast dysk HDD 500GB pełnił rolę magazynu na obrazy ISO czy pliki instalacyjne programów.

Wszystkie serwery zostały podłączone do przełącznika Ubiquiti ES-24-LITE za pomocą skrętki komputerowej Cat.6A S/FTP 2m. Każdy z serwerów został podłączony 3 takimi przewodami. Jeden służy do zarządzania serwerem a dwie kolejne do wykorzystania przez system zainstalowany na serwerach. Wszystkie elementy platformy testowej działają z przepustowością 1Gbps. Na potrzeby testów została stworzona sieć 10.200.200.0/24 z dostępem do Internetu. Na rysunku 1 przedstawiono schemat środowiska testowego.



Rys. 1. Schemat środowiska testowego

Testowano VMWare ESX 6.0.0 U3 oraz Microsoft Hyper-V 2019 Version 1809. Jako gośzczonych systemów operacyjnych, pracujących na hypervisorach w czasie testów, używano kombinacji instancji systemów operacyjnych Windows 10 Version 1809 oraz dystrybucji systemu Linux CentOS 7.6.1810. Zastosowano edycje 64-bitowe dla Windows 10 i CentOS 7 oraz VMWare i Hyper-V. Windows i Linux instalowano, używając tej samej pamięci i ustawień vCPU dla obu hyperwizorów.

Dla systemu Windows było to:

- Procesor: 2 rdzenie i 2 wątki
- Pamięć RAM: 2 GB
- Karty sieciowe: 1 x 1 Gbit
- Dysk: 50 GB

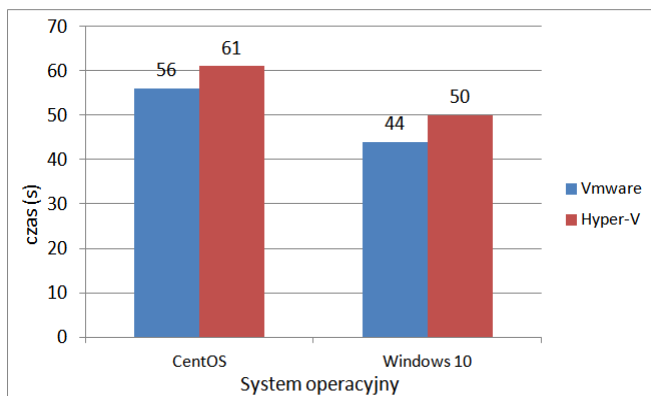
Dla systemu Linux było to:

- Procesor: 1 rdzeń i 1 wątek
- Pamięć RAM: 1 GB
- Karty sieciowe: 1 x 1 Gbit
- Dysk: 16 GB

6. Wyniki badań

6.1. Awaria hosta

Pierwszym etapem badań było zachowanie klastra w sytuacji awarii jednego z węzłów. Symulacja awarii polegała na wyłączeniu zasilania hosta, na którym znajdowały się maszyny wirtualne. W czasie awarii sprawdzany był czas jaki jest potrzebny aby wszystkie maszyny wirtualne były ponownie dostępne. Za pomocą polecenia PING sprawdzano utratę pakietów. Wykonano trzy próby a wyniki uśredniono. Średnia liczba utraconych pakietów dla VMware wyniosła 14 pakietów, a dla Hyper-V 16 pakietów.



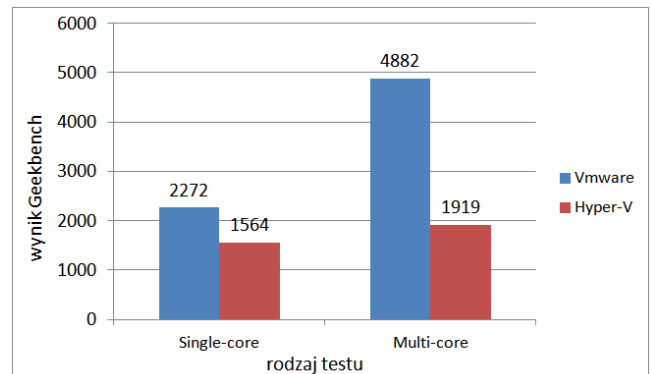
Rys. 2. Średnie czasy po jakich systemy były gotowe do działania w zależności od systemu operacyjnego

Jak obrazuje wykres (Rys. 2), dla obydwu systemów operacyjnych, zarówno na Windows 10, jak i na CentOS 7, lepsze czasy uzyskał hyperwizor VMware. Na systemie Windows hosty gotowe były do ponownej pracy po 44 sekundach, a na systemie CentOS po 56 sekundach. Czasy wymagane do ponownej pracy przy użyciu rozwiązania wirtualizacyjnego firmy Microsoft to 50 sekund dla systemu Windows oraz 61 sekund dla systemu CentOS. Wyraźnie widać, że po awarii hosta na ponowną pracę szybciej pozwala hyperwizor VMware. Różnica względem Hyper-V wynosiła 5 oraz 6 sekund.

6.2. Wydajność

W testach wydajności przebadano prędkość procesora oraz średnie czasy operacji dyskowych oferowanych przez badane rozwiązania. Do testowania pomiarów prędkości procesora na systemie operacyjnym Windows 10

wykorzystano aplikację Geekbench. Aplikacja ta testuje m.in. szybkość zapisu i odczytu danych oraz obsługę strumieni. Sprawdza jak PC radzi sobie podczas szyfrowania, kompresji tekstu i obrazów. Wbudowane w nim benchmarki pozwalają na testowanie procesorów jedno- lub wielordzeniowych w aspektach wielowątkowości i wielordzeniowości. Testy procesora Single-Core oraz Multi-Core szacują, jak szybko procesor jest w stanie wykonać wiele różnych obliczeń. Im szybciej procesor jest w stanie wykonać obliczenia, tym wyższy wynik testu porównawczego. Wyniki testów przedstawiono na rysunku 3.

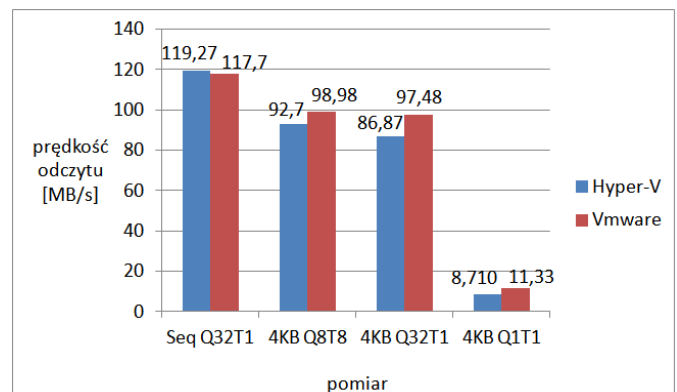


Rys. 3. Wyniki pomiaru prędkości CPU w programie Geekbench dla systemu Windows

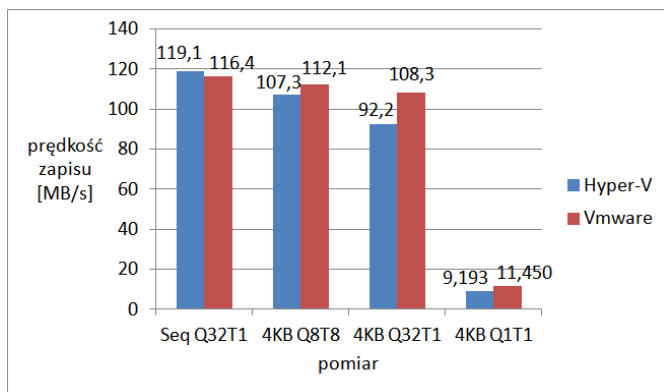
W teście wydajności operacji wejścia/wyjścia dysku wykorzystano aplikację CrystalDiskMark. Jest to niewielki benchmark służący do testowania wydajności dysków twardych, który na bazie wykonanych testów losowych oraz sekwencyjnych operacji zapisu/odczytu danych, potrafi dokonać pomiaru prędkości dysku. Prędkość wyznaczana jest dla 4 pomiarów:

- Seq Q32T1: testuje w trybie wielowątkowości. Plik jest zapisany w segmentach o wielkości 128 KB.
- 4K Q8T8: wielkość segmentów to 4KB, 8 kolejek, wątków 8.
- 4K Q32T1: wielkość segmentów to 4KB, 32 kolejki, wątków 1.
- 4K Q1T1: gdzie, wielkość segmentów to 4KB, 1 kolejka, wątków 1.

Uzyskane wyniki zostały przedstawione na rysunku 4 oraz 5.



Rys. 4. Średnie prędkości odczytu danych z dysku dla obydwu hyperwizorów dla systemu Windows

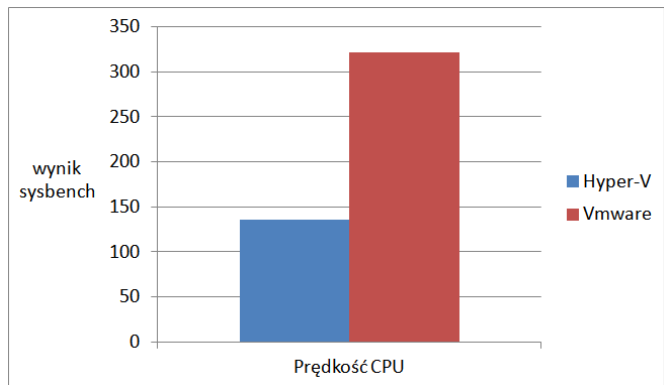


Rys. 5. Średnie prędkości zapisu danych z dysku dla obydwu hyperwizorów dla systemu Windows

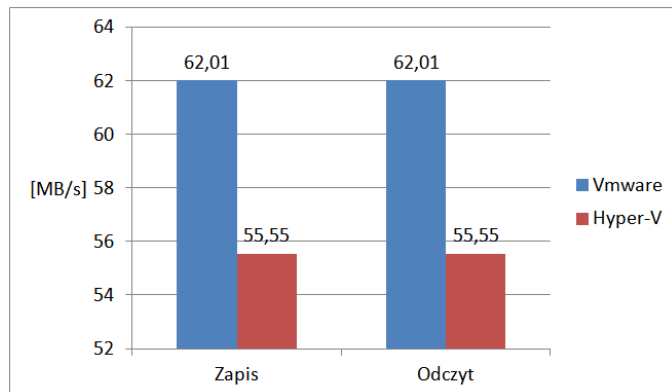
Dla systemu CentOS oba benchmarki zostały wykonane za pomocą programu SysBench. Jest to prosta w obsłudze aplikacja, za pomocą której szybko przetestujemy podstawowe parametry komputera/serwera:

- Procesor (CPU)
- Pamięć RAM
- System plików (transfer I/O)
- Bazę danych (OLTP)

Uzyskane wyniki zostały przedstawione na rysunku 5 i 6.



Rys. 6. Wynik pomiaru prędkości CPU w programie sysbench dla systemu CentOS



Rys. 7. Średnie prędkości odczytu i zapisu danych z dysku dla systemu CentOS wykonane przy użyciu programu sysbench

7. Wnioski

Przeprowadzone badania wykazały, że produkt VMWare – vSphere był lepszy w teście dotyczącym awarii jednego z hostów. Maszyny wirtualne przy użyciu rozwiązania vSphere były szybciej dostępne do ponownej pracy. Różnica względem Hyper-V wynosiła nieco ponad 5 sekund na korzyść vSphere. Różnica niby niewielka, ale w środowisku produkcyjnym, gdzie liczy się każda sekunda niedostępności danego sprzętu, systemu, czy usługi lepszym wyborem będzie VMWare.

W testach wydajności goszczonych systemów również lepszym okazał się produkt VMWare.

W przeprowadzonych testach dotyczących goszczonego systemu operacyjnego Windows 10 w teście CPU zdecydowanie lepiej wypadł produkt Vmware gdzie różnica w pomiarze pojedynczego rdzenia wynosiła ponad 30% a w przypadku wielu rdzeni już prawie 40%. Natomiast w teście prędkości zapisu/odczytu plików, Hyper-V poradził sobie jedynie lepiej w pracy z większymi plikami. System Windows goszczony na VMWare dużo szybciej pracował z mniejszymi plikami. Z testów wynika, że hyperwizora VMWare lepiej używać np. do utrzymania baz danych, gdy Hyper-V powinien lepiej się sprawować w aplikacjach internetowych.

W testach dotyczących goszczonego systemu CentOS 7, w teście prędkości CPU, prawie dwukrotnie lepszy wynik, wynoszący ponad 40%, otrzymał hyperwizor vSphere. Natomiast w teście prędkości zapisu/odczytu, zarówno zapis jak i odczyt odbywa się szybciej przy wykorzystaniu rozwiązania VMWare. Średnio te operacje wykonują się o 10% szybciej przy użyciu hyperwizora VMWare, niż w przypadku Hyper-V.

Literatura

- [1] Gray J., Siewiorek D.: High-Availability Computer Systems. Carnegie Mellon University, 1991.
- [2] Kopper K.: The Linux Enterprise Cluster - Build a Highly Available Cluster with Commodity Hardware and Free Software. No Starch Press, 2005.
- [3] Prashanta K. D.: Comparative Study on XEN, KVM, VSphere, and Hyper-V. [W]: Emerging Research Surrounding Power Consumption and Performance Issues in Utility Computing, 2016.
- [4] Alroobaea R., Ghiduk A., Ikbali M.: Performance Evaluation of Xen, KVM, and Proxmox Hypervisors. [W]: International Journal of Open Source Software and Processes. 9(2), 2018.
- [5] Graniszewski W., Arciszewski A.: Performance analysis of selected hypervisors (Virtual Machine Monitors – VMMs). [W]: International Journal of Electronics and Telecommunications, 62(3), 2016.
- [6] Bhatia A., Bhattal G.: A comparative study of Various Hypervisors Performance. [W]: International Journal of Scientific & Engineering Research. Tom 7, zeszyt 12, 2016.
- [7] Fayyad-Kazan H., Perneel L., Timmerman M.: Benchmark the Performance of Microsoft Hyper-V server, VMWare ESXi and Xen Hypervisors. [W]: Journal of Emerging Trends in Computing and Information Sciences, 2013.

- [8] <https://www.microsoft.com/en-us/download/details.aspx?id=56495> [01.06.2019].
- [9] <https://docs.microsoft.com/en-us/virtualization/hyper-v-on-windows/reference/hyper-v-architecture> [01.06.2019].
- [10] <https://blogs.vmware.com/vsphere/2018/04/introducing-vmware-vsphere-6-7.html> [01.06.2019].