Klaster Windows dla własnych usług: rola, zasób, zależności, kworum i testy przełączenia

Poziom: średnio

Dotyczy: Windows Server 2022, Failover Clustering (moduł FailoverClusters), PowerShell 5.1, Zabbix agent 2 (7.0 LTS)

Czas: ~90 min

Klaster pracy awaryjnej w Windows Server potrafi pilnować nie tylko maszyn wirtualnych, ale też zwykłej usługi albo programu: gdy węzeł z usługą przestaje działać, usługa startuje na drugim, a użytkownicy łączą się dalej pod tą samą nazwą. Na tej stronie opisujemy pojęcia (rola, zasób, zależności, punkt dostępu klienta, kworum), walidację klastra, dodanie usługi przez Generic Service i Generic Script oraz testy przełączenia i monitoring. Wracamy tu, zanim wstawimy do klastra Perforce albo Lore, o których piszemy na następnej stronie.

> Własnych ról nie dodajemy na węzłach `hv01` i `hv02`. Microsoft pisze wprost, że Storage Spaces Direct nie wspiera serwera plików ogólnego użytku i że takie usługi mają działać w maszynach wirtualnych [1]. Przykłady z tej strony stoją więc na klastrze gościnnym z dwóch maszyn wirtualnych, a wynik sprawdzenia i jego skutki opisujemy niżej, w sekcji „Gdzie wolno uruchamiać własne role".

> Wyniki w przykładach pokazują kształt odpowiedzi i nie pochodzą z uruchomienia na prawdziwych hostach Zielonej. Klaster gościnny `ftgc01` (10.20.10.76) z węzłami `p4n1` (.62) i `p4n2` (.63) to wariant (b) ze strony o Perforce i Lore, wybierany wtedy, gdy zmierzona przerwa po awarii przekracza RTO.

> **ZAŁOŻENIA SCENARIUSZA**
>
> klaster własnych usług składa się z maszyn wirtualnych na `hv01` i `hv02`, a nie z samych hostów, w jednej serwerowni i w VLAN 10. Klaster hostów jest na początku bez świadka;

świadkiem może być udział na fizycznym `bkp01` albo Cloud Witness, który wymaga wyjścia na port 443 przez pojedynczy `fw01`.

Skróty użyte na tej stronie (19)

Skrót	Co znaczy
AD	Active Directory: katalog Microsoftu z kontami użytkowników, komputerów i grup w firmie
API	Application Programming Interface: sposób, w jaki programy rozmawiają ze sobą
CNO	Cluster Name Object: konto komputera w Active Directory, które reprezentuje sam klaster i z którego klaster tworzy konta punktów dostępu klienta
CSV	w Windowsie: Cluster Shared Volume, dysk widoczny naraz dla wszystkich węzłów klastra; w danych: Comma-Separated Values, plik z wartościami oddzielonymi przecinkami
DFS	Distributed File System: usługa łącząca udziały z wielu serwerów pod jedną wspólną ścieżką sieciową
DNS	Domain Name System: książka telefoniczna internetu, zamienia nazwy na adresy IP
gMSA	group Managed Service Account: konto usługi, któremu Windows sam zmienia hasło
IP	Internet Protocol: adres komputera w sieci; w elektryce: stopień ochrony obudowy (Ingress Protection)
LTS	Long Term Support: wersja z długim wsparciem poprawkami
NAS	Network Attached Storage: dysk sieciowy
NTFS	New Technology File System: system plików Windows z wbudowanymi uprawnieniami ACL na poziomie pliku i folderu
RTO	Recovery Time Objective: jak długo wolno nie działać przy awarii
S2D	Storage Spaces Direct: programowa macierz Windows Server z dysków wielu serwerów
SMB	Server Message Block: protokół udostępniania plików i drukarek w sieci Windows, na nim stoi zwykły udział sieciowy
SQL	Structured Query Language: język zapytań do relacyjnych baz danych
VBScript	Visual Basic Scripting Edition: język skryptów Windows, w którym piszemy skrypty roli Generic Script w klastrze; Microsoft zapowiada jego wycofanie w przyszłych wydaniach Windows
VCO	Virtual Computer Object: konto komputera w Active Directory dla punktu dostępu klienta roli klastra, czyli nazwy sieciowej, pod którą klienci łączą się z rolą
VHD	Virtual Hard Disk: plik, który Hyper-V pokazuje maszynie wirtualnej jako dysk; wspólny dysk dwóch maszyn ma format VHD Set (.vhds)
VLAN	Virtual LAN: wydzielona logicznie sieć na tym samym sprzęcie

Fiszka

Polecenia wykonujemy w PowerShell uruchomionym jako administrator, na węźle klastra gościnnego (maszyna wirtualna z rolą Failover Clustering), chyba że przy poleceniu napisano inaczej. Zamiast `NazwaKlastra`, `NazwaWezla`, `NazwaRoli`, `NazwaUslugi` i `AdresIP` wpisujemy własne wartości.

TYLKO ODCZYT

Sprawdzenie ról klastra i ich właściciela: Get-ClusterGroup

```
Get-ClusterGroup
```

Pokazuje każdą rolę, jej stan i węzeł, który jest jej właścicielem. Rola w stanie `Online` ma właściciela, a w stanie `Failed` albo `Partial Online` zaglądamy do zasobów.

Przykład:

```
Name           OwnerNode  State
----           ---------  -----
Cluster Group  p4n1       Online
lore           p4n1       Online
perforce       p4n2       Online
```

TYLKO ODCZYT

Sprawdzenie zasobów roli: Get-ClusterResource

```
Get-ClusterResource
```

Wypisuje wszystkie zasoby klastra z rodzajem (`ResourceType`) i rolą, do której należą. Dla roli z usługą spodziewamy się adresu IP, nazwy sieciowej, dysku i samej usługi.

Przykład:

```
Name               State   OwnerGroup  ResourceType
----               -----   ----------  ------------
Cluster Disk 1     Online  perforce    Physical Disk
perforce           Online  perforce    Network Name
```

```
IP Address 10.x    Online  perforce    IP Address
Perforce           Online  perforce    Generic Service
```

IP = Internet Protocol: adres komputera w sieci; w elektryce: stopień ochrony obudowy (Ingress Protection)

TYLKO ODCZYT

Sprawdzenie zależności zasobu: Get-ClusterResourceDependency

```
Get-ClusterResource -Name "NazwaZasobu" | Get-ClusterResourceDependency
```

Pokazuje, od czego zasób zależy. Zasób startuje dopiero po swoich zależnościach i zatrzymuje się przed nimi [2].

TYLKO ODCZYT

Sprawdzenie świadka i głosów: Get-ClusterQuorum

```
Get-ClusterQuorum | Format-List Cluster, QuorumResource
Get-ClusterNode | Format-Table Name, State, NodeWeight, DynamicWeight
```

Dwa polecenia tylko odczytują stan: pierwsze pokazuje zasób świadka, drugie głosy węzłów. Pusty `QuorumResource` oznacza klaster bez świadka, a to dla dwóch węzłów jest luka (sekcja „Kworum i świadek").

TYLKO ODCZYT

Walidacja węzłów: Test-Cluster

```
Test-Cluster -Node "NazwaWezla1","NazwaWezla2" -ReportName "C:\Temp\walidacja-
NazwaKlastra"
```

Uruchamia wszystkie testy konfiguracji i zapisuje raport. Testy dysków nie sprawdzają dysków używanych przez role, chyba że wcześniej zatrzymamy rolę albo użyjemy `-Force`, który wyłącza dyski na czas testu [3]. Listę testów pokazuje `Test-Cluster -List`.

ZMIENIA SYSTEM

Utworzenie klastra: New-Cluster

```
New-Cluster -Name "NazwaKlastra" -Node "NazwaWezla1","NazwaWezla2" -StaticAddress "AdresIP"
```

Zakłada klaster, jego obiekt w Active Directory (CNO) i nazwę w DNS. Konto, którym uruchamiamy polecenie, musi mieć prawo tworzenia obiektów komputerów w kontenerze, w którym powstanie CNO.

CNO = Cluster Name Object: konto komputera w Active Directory, które reprezentuje sam klaster i z którego klaster tworzy konta punktów dostępu klienta · **DNS** = Domain Name System: książka telefoniczna internetu, zamienia nazwy na adresy IP

ZMIENIA SYSTEM

Dodanie dysku współdzielonego do klastra: Add-ClusterDisk

```
Get-ClusterAvailableDisk | Add-ClusterDisk
```

Dodaje do klastra dyski, które widzą oba węzły, a których jeszcze nie ma w klastrze. W klastrze gościnnym to wspólny dysk maszyn wirtualnych (VHD Set), którego tworzenie pokazujemy na stronie Pacemaker na Linuksie: własna usługa z adresem wirtualnym (te same polecenia Hyper-V działają dla gości Windows).

VHD = Virtual Hard Disk: plik, który Hyper-V pokazuje maszynie wirtualnej jako dysk; wspólny dysk dwóch maszyn ma format VHD Set (.vhds)

ZMIENIA SYSTEM

Dodanie usługi jako roli: Add-ClusterGenericServiceRole

```
Add-ClusterGenericServiceRole -ServiceName "NazwaUslugi" -Name "NazwaRoli" -StaticAddress "AdresIP" -Storage "Cluster Disk 1"
```

Tworzy rolę z usługą Windows, punktem dostępu klienta (nazwa i adres) oraz dyskiem [4]. `-ServiceName` to nazwa usługi z `Get-Service`, `-Name` to nazwa punktu dostępu, pod którą łączą się użytkownicy.

ZMIENIA SYSTEM

Dodanie programu jako roli: Add-ClusterGenericApplicationRole

```
Add-ClusterGenericApplicationRole -CommandLine "SciezkaDoProgramu" -Parameters "ParametryProgramu" -Name "NazwaRoli" -StaticAddress "AdresIP" -Storage "Cluster Disk 2"
```

Dla programu, który nie jest usługą Windows. Klaster uruchamia wskazany plik, a jego zakończenie uznaje za awarię zasobu [5].

□□ ZMIENIA SYSTEM

Dodanie skryptu jako roli: Add-ClusterGenericScriptRole

```
Add-ClusterGenericScriptRole -ScriptFilePath "SciezkaDoSkryptu" -Name "NazwaRoli" -StaticAddress "AdresIP" -Storage "Cluster Disk 3"
```

Skrypt musi działać w Windows Script Host (VBScript albo JScript) i sam odpowiada za stan aplikacji [6]. Szkic skryptu pokazujemy niżej.

□□ **VBScript** = Visual Basic Scripting Edition: język skryptów Windows, w którym piszemy skrypty roli Generic Script w klastrze; Microsoft zapowiada jego wycofanie w przyszłych wydaniach Windows

□□ ZMIENIA SYSTEM

Ustawienie preferowanego właściciela roli: Set-ClusterOwnerNode

```
Set-ClusterOwnerNode -Group "NazwaRoli" -Owners "NazwaWezla1","NazwaWezla2"
```

Dla roli lista `-Owners` oznacza preferowanych właścicieli w kolejności. Dla pojedynczego zasobu ta sama komenda ustawia możliwych właścicieli, czyli węzły, na których zasób wolno uruchomić [7].

□□ ZMIENIA SYSTEM

Przeniesienie roli na drugi węzeł: Move-ClusterGroup

```
Move-ClusterGroup -Name "NazwaRoli" -Node "NazwaWezla2"
```

Zatrzymuje rolę na bieżącym węźle i uruchamia na wskazanym. Bez `-Node` klaster wybiera sam, a `-Wait 0` oddaje wiersz poleceń od razu [8].

□□ ZMIENIA SYSTEM

Wstrzymanie węzła z opróżnieniem: Suspend-ClusterNode

```
Suspend-ClusterNode -Name "NazwaWezla1" -Drain -Wait
```

Przenosi role z węzła i blokuje na nim nowe. Dla maszyn wirtualnych oznacza to migrację na żywo, a dla ról z usługą zwykłe przeniesienie z krótką przerwą [9].

◫◫ ZMIENIA SYSTEM

Wznowienie węzła po pracach: Resume-ClusterNode

```
Resume-ClusterNode -Name "NazwaWezla1" -Failback NoFailback
```

Węzeł znów przyjmuje role. `-Failback` ma wartości `NoFailback`, `Immediate` i `Policy` i decyduje, czy role wracają same [10].

◫◫ TYLKO ODCZYT

Wyświetlenie zasad restartu zasobu: Get-ClusterResource

```
Get-ClusterResource -Name "NazwaZasobu" | Format-List Name, RestartAction,
RestartThreshold, RestartPeriod, PendingTimeout
```

Pokazuje, ile razy klaster próbuje zrestartować zasób na tym samym węźle w danym okresie, zanim uzna rolę za wymagającą przeniesienia [11]. Zmieniamy je przez przypisanie właściwości, na przykład `$zasob.RestartThreshold = 3`.

◫◫ TYLKO ODCZYT

Zebranie dziennika klastra: Get-ClusterLog

```
Get-ClusterLog -UseLocalTime -TimeSpan 15 -Destination "C:\Temp"
```

Zbiera z każdego węzła dziennik z ostatnich 15 minut do wskazanego katalogu. Sięgamy po niego po nieudanym przełączeniu, żeby zobaczyć, który zasób zawiódł i na którym węźle.

Z życia wzięte

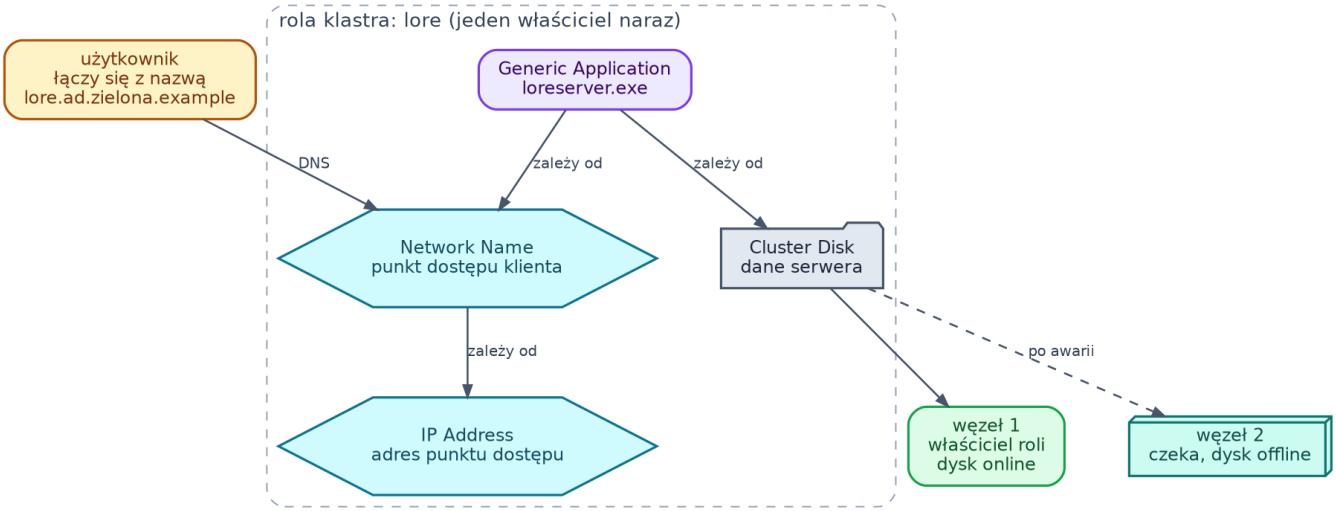
Z ŻYCIA WZIĘTE

Środa, 10:40. Tomek przygotowuje na `hv01` rolę Generic Application dla serwera Lore i już ma wpisane `Add-ClusterGenericApplicationRole`, gdy Ola pokazuje mu przypis w dokumentacji Microsoftu: ogólne usługi na Storage Spaces Direct mają działać w maszynach wirtualnych, a nie na węzłach z magazynem. Tomek kasuje polecenie i zakłada na kartce dwie maszyny wirtualne, które staną na różnych hostach. Rola Lore ma się przenosić między nimi, a `hv01` i `hv02` zostają od tego z daleka.

Jak to działa

Rozwiń: jak to działa pod spodem

Klaster pracy awaryjnej to grupa serwerów (węzłów), które razem pilnują ról. W danej chwili rolę prowadzi jeden węzeł, a pozostałe obserwują go przez sygnały kontrolne, zwane heartbeat. Gdy węzeł przestaje odpowiadać albo zasób roli zawodzi, klaster uruchamia rolę na innym węźle [12]. Poniżej pojęcia, z których składa się każda rola.



Rysunek 1. Rola klastra z serwerem Lore. Użytkownik łączy się z nazwą, nazwa zależy od adresu, a usługa zależy od nazwy i od dysku. Dysk jest online tylko na węźle, który prowadzi rolę, a po awarii przechodzi na drugi. (kliknięcie otwiera pełny rozmiar)

Pojęcie: rola klastra (grupa zasobów). Zestaw zasobów, które klaster przenosi zawsze razem, na przykład adres, nazwa, dysk i usługa. Rola ma jednego właściciela naraz i stan (`Online`, `Offline`, `Failed`).

Pojęcie: zasób. Pojedynczy element roli, który klaster umie włączyć, wyłączyć i sprawdzić: adres IP, nazwa sieciowa, dysk, usługa Windows, program. Każdy zasób ma rodzaj (`ResourceType`), na przykład `IP Address`, `Network Name`, `Physical Disk`, `Generic Service`.

Pojęcie: zależność. Reguła „zasób A startuje dopiero po zasobie B". Usługa zależy od dysku i od nazwy, nazwa od adresu, więc przy starcie roli klaster podnosi najpierw adres, potem nazwę i dysk, a na końcu usługę, a przy zatrzymaniu robi to w odwrotnej kolejności [2].

Pojęcie: punkt dostępu klienta (CNO i VCO). Nazwa i adres, pod którymi użytkownicy łączą się z rolą, niezależnie od tego, który węzeł ją prowadzi. CNO (Cluster Name Object) to konto komputera klastra w Active Directory, a VCO (Virtual Computer Object) to konto komputera nazwy roli. CNO zakłada VCO, więc musi mieć prawo tworzenia obiektów komputerów w odpowiednim kontenerze albo VCO przygotowujemy z wyprzedzeniem.

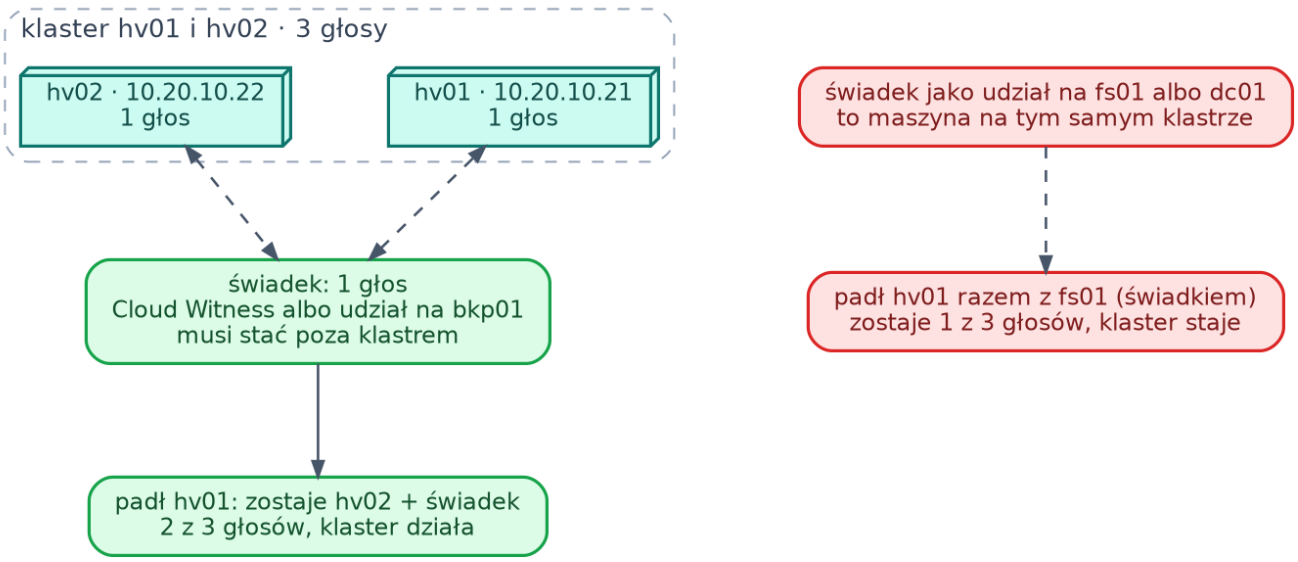
Pojęcie: dysk klastrowy i CSV. Dysk klastrowy (Physical Disk) w danej chwili należy do jednego węzła: tylko on widzi system plików, a po przeniesieniu roli dysk przechodzi razem z nią. CSV (Cluster Shared Volume) to wolumen widoczny pod `C:\ClusterStorage` na wszystkich węzłach naraz, stworzony dla maszyn wirtualnych i serwerów plików z dostępem aktywny-aktywny [13]. Usłudze z jednym właścicielem wystarcza zwykły dysk klastrowy w NTFS, a CSV nie jest potrzebny.

Pojęcie: preferowany właściciel i failback. Preferowany właściciel to węzeł, który rola woli prowadzić. Failback to powrót roli na niego po jego naprawie. Powrót włączamy świadomie, bo każdy powrót to kolejna przerwa. Ustawia go właściwość `AutoFailbackType` roli (0 wyłącza powrót, 1 go włącza), najlepiej razem z oknem czasowym (`FailbackWindowStart` i `FailbackWindowEnd`) [14].

> **Pojęcie: zasady restartu i przełączenia.** Zasób, który zawiedzie, klaster najpierw restartuje na tym samym węźle (`RestartThreshold` razy w okresie `RestartPeriod`). Gdy to nie pomaga, a `RestartAction` pozwala, przenosi całą rolę na inny węzeł. Rola może się przenieść tylko tyle razy, ile pozwalają `FailoverThreshold` w okresie `FailoverPeriod` (godziny), a po przekroczeniu rola zostaje w stanie `Failed` [14][11].

Kworum i świadek

Węzły klastra głosują, kto ma rację, gdy zniknie łączność. Klaster działa, jeśli widzi więcej niż połowę głosów, a świadek (witness) to dodatkowy głos spoza węzłów. Microsoft zaleca świadka zawsze, także dla klastra dwuwęzłowego [15]. Bez niego utrata jednego z dwóch węzłów zostawia drugiemu połowę głosów, a to za mało.



Rysunek 2. Świadek klastra hv01 i hv02. Świadek poza klastrem daje trzeci głos, więc po awarii jednego hosta zostają dwa z trzech. Świadek na maszynie, która sama stoi na tym klastrze, znika razem z hostem i klaster staje. (kliknięcie otwiera pełny rozmiar)

Świadek ma trzy odmiany: dysk, udział plikowy (SMB 2 lub nowszy) i Cloud Witness w koncie Azure Storage [15]. Udział plikowy ma stać fizycznie osobno od węzłów, służyć tylko jednemu klastrowi i nie być DFS ani replikowanym magazynem. Dla klastra z Windows Server 2019 lub nowszym może leżeć także na urządzeniu spoza domeny, na przykład na NAS [15]. Cloud Witness wymaga konta Azure Storage ogólnego przeznaczenia v2 i wyjścia na port 443 z każdego węzła [16].

W Zielonej wszystkie serwery Windows są maszynami wirtualnymi na `hv01` i `hv02`, więc udział na `fs01` albo `dc01` odpada: znika razem z hostem. Zostają Cloud Witness albo udział na `bkp01`, który jest fizycznym serwerem poza klastrem. Dziś klaster `hv01` i `hv02` świadka nie ma (strony o S2D o nim milczą), więc to luka do zamknięcia przed jakąkolwiek rozmową o „niewidocznym" przełączeniu.

Gdzie wolno uruchamiać własne role

Przy każdym pomyśle klastra własnych usług pada to samo pytanie: czy Microsoft wspiera własne role na węzłach hiperkonwergentnego S2D, czy zaleca tam tylko Hyper-V i Scale-Out File Server. Dokumentacja Microsoftu mówi dwie rzeczy. Pierwsza: w układzie hiperkonwergentnym na węzłach z magazynem działają maszyny wirtualne Hyper-V albo bazy SQL Server [17]. Druga, w uwagach na stronie o planowaniu woluminów: „Storage Spaces Direct does not support a file server for general use. If you need to run the file server or other generic services on Storage Space Direct, configure it on the virtual machines" [1].

Wynik dla Zielonej jest więc taki. Generic Service, Generic Application i Generic Script na samych `hv01` i `hv02` nie są drogą wspieraną, a Microsoft kieruje takie usługi do maszyn wirtualnych. Mamy dwie wspierane drogi: sama maszyna wirtualna jako rola klastra Hyper-V (migracja na żywo przy pracach planowych, restart na drugim węźle przy awarii) albo klaster gościnny z dwóch maszyn wirtualnych, w którym działają role z tej strony. Którą wybieramy dla Perforce i Lore, rozstrzygamy na następnej stronie.

Klaster gościnny potrzebuje wspólnego magazynu widocznego dla obu maszyn. Na S2D robimy to współdzielonym dyskiem wirtualnym w formacie VHD Set (`.vhds`), który Windows Server 2016 i nowszy wspiera w klastrach gościnnych [18]. Ograniczenie, które znamy z dokumentacji: kopia maszyn z takim dyskiem na poziomie hosta sprawia kłopoty (błędy 32770 i 32775), więc kopię robimy wewnątrz gościa [19].

Krok po kroku: usługa Windows jako rola

Przykład dotyczy usługi `Perforce` z klastra gościnnego `ftgc01`. Zakładamy, że klaster z dwoma węzłami już istnieje (`New-Cluster`), ma świadka, a wspólny dysk jest dodany jako `Cluster Disk 1` i sformatowany jako NTFS pod literą `D:`.

Przygotowanie usługi na obu węzłach

Usługa musi istnieć na obu węzłach w tej samej wersji i z tą samą konfiguracją, a jej dane leżeć na dysku klastrowym, pod tą samą literą. Na obu węzłach ustawiamy ręczny typ uruchamiania, bo to klaster ma ją włączać, a nie system przy starcie:

```
Invoke-Command -ComputerName p4n1,p4n2 -ScriptBlock {
    Set-Service -Name "NazwaUslugi" -StartupType Manual
    Stop-Service -Name "NazwaUslugi" -ErrorAction SilentlyContinue
    Get-Service -Name "NazwaUslugi" | Select-Object Name, Status, StartType
}
```

Polecenie wykonujemy z `-ComputerName` wskazującym oba węzły klastra, a `NazwaUslugi` to nazwa z `Get-Service`.

Utworzenie roli

```
Add-ClusterGenericServiceRole -ServiceName "NazwaUslugi" -Name "NazwaRoli" -StaticAddress
"AdresIP" -Storage "Cluster Disk 1"
```

W AD pojawia się VCO o nazwie `NazwaRoli`, a w DNS rekord A tej nazwy wskazujący `AdresIP`. Jeżeli polecenie kończy się błędem o tworzeniu konta komputera, przyczyną są zwykle uprawnienia CNO (tabela „Częste błędy").

Sprawdzenie zależności

```
Get-ClusterResource -Name "NazwaUslugi" | Get-ClusterResourceDependency
```

Spodziewamy się, że usługa zależy od dysku i od nazwy sieciowej. Brakującą zależność dodajemy poleceniem `Add-ClusterResourceDependency -Resource "NazwaUslugi" -Provider "NazwaZasobuZaleznosci"` [2].

Ułożenie roli ręcznie

Gdy kreator nie wystarcza (na przykład usługa ma zależeć od innej usługi), składamy rolę z zasobów. Poniższy blok można wykonać w całości, po podmianie wartości:

```
Add-ClusterGroup -Name "NazwaRoli"
```

```
Add-ClusterResource -Name "NazwaRoli-IP" -ResourceType "IP Address" -Group "NazwaRoli" |
    Set-ClusterParameter -Multiple @{ Address = "AdresIP"; SubnetMask = "255.255.255.0";
Network = "NazwaSieciKlastra" }

Add-ClusterResource -Name "NazwaRoli-Nazwa" -ResourceType "Network Name" -Group "NazwaRoli" |
    Set-ClusterParameter -Multiple @{ Name = "NazwaRoli"; DnsName = "NazwaRoli" }
Add-ClusterResourceDependency -Resource "NazwaRoli-Nazwa" -Provider "NazwaRoli-IP"

Move-ClusterResource -Name "Cluster Disk 1" -Group "NazwaRoli"

Add-ClusterResource -Name "NazwaRoli-Usluga" -ResourceType "Generic Service" -Group
"NazwaRoli" |
    Set-ClusterParameter -Name ServiceName -Value "NazwaUslugi"
Add-ClusterResourceDependency -Resource "NazwaRoli-Usluga" -Provider "NazwaRoli-Nazwa"
Add-ClusterResourceDependency -Resource "NazwaRoli-Usluga" -Provider "Cluster Disk 1"

Start-ClusterGroup -Name "NazwaRoli"
```

Rodzaj `Generic Service` ma parametr prywatny `ServiceName`, a opcjonalnie `StartupParameters` i `UseNetworkName` [4]. Dokumentacja starszego API ostrzega, że zasób generyczny daje tylko podstawową obsługę: klaster uznaje go za działający, dopóki Service Control Manager widzi usługę jako uruchomioną [20]. Usługa, która wisi, ale nie odpowiada, nadal jest dla klastra „online", więc dochodzi monitoring z sekcji o Zabbixie.

Preferowany właściciel i zasady przełączenia

```
Set-ClusterOwnerNode -Group "NazwaRoli" -Owners "p4n1","p4n2"
$grupa = Get-ClusterGroup -Name "NazwaRoli"
$grupa.AutoFailbackType = 0
$grupa.FailoverThreshold = 2
$grupa.FailoverPeriod = 6
```

Ustawiamy kolejność właścicieli, wyłączamy powrót automatyczny i pozwalamy na dwa przełączenia w ciągu sześciu godzin. Po przekroczeniu limitu rola zostaje w stanie `Failed`, żeby awaria powtarzająca się co kilka minut nie przenosiła usługi bez końca, a człowiek wziął ją pod uwagę. Wartości dobieramy do usługi i sprawdzamy w teście.

Krok po kroku: program bez usługi (Generic Application)

Lore ma serwer w pliku `loreserver.exe`, uruchamiany z pliku konfiguracyjnym, a dokumentacja nie opisuje instalacji jako usługi Windows [21]. Dla takich programów służy Generic Application, w której klaster sam uruchamia plik i uznaje jego zakończenie za awarię zasobu:

```
Add-ClusterGenericApplicationRole -CommandLine "C:\loreserver\loreserver.exe" -Parameters "--config D:\loreserver\config" -Name "NazwaRoli" -StaticAddress "AdresIP" -Storage "Cluster Disk 2"
```

Dwa ograniczenia znamy wprost. Pierwsze: program uruchamia host zasobów klastra, więc nie wybieramy dla niego konta gMSA, a konto, na którym działa, sprawdzamy w teście. Jeśli konto ma znaczenie dla dostępu do plików, program opakowujemy w usługę (narzędzie firmy trzeciej, jego licencję sprawdzamy) i dajemy Generic Service albo uruchamiamy go z zadania harmonogramu poza klastrem. Drugie: `InteractWithDesktop` nie jest w Generic Application obsługiwane [5].

Krok po kroku: skrypt jako rola (Generic Script)

Generic Script wybieramy, gdy potrzebujemy własnej logiki: na przykład prawdziwego sprawdzenia, czy usługa odpowiada, a nie tylko, czy proces żyje. Skrypt działa w Windows Script Host, więc to VBScript albo JScript. Microsoft zapowiada, że VBScript będzie najpierw funkcją na żądanie, a potem zostanie wycofany w przyszłych wydaniach Windows, bez podanych dat [22]. Dla nowych wdrożeń wolimy Generic Service z monitoringiem z Zabbixa i traktujemy Generic Script jako rozwiązanie przejściowe.

Skrypt musi mieć funkcje `LooksAlive` i `IsAlive`. Funkcje `Open`, `Online`, `Offline`, `Close` i `Terminate` są opcjonalne. Zwracana wartość `True` oznacza powodzenie [6]. Do dziennika piszemy przez obiekt `Resource`, a nie przez API klastra, bo wołanie klastra ze skryptu grozi zakleszczeniem [6]. Poniżej szkic dla `Perforce`, którego składnię i zachowanie sprawdzamy na węźle testowym, bo w bazie go nie uruchamialiśmy:

```
' ft-p4-rola.vbs: szkic roli Generic Script dla serwera P4
Function Open()
    Resource.LogInformation "ft-p4: otwarcie zasobu"
    Open = True
```

```
End Function

Function Online()
    Dim powloka
    Set powloka = CreateObject("WScript.Shell")
    powloka.Run "net start Perforce", 0, True
    Online = True
End Function

Function Offline()
    Dim powloka
    Set powloka = CreateObject("WScript.Shell")
    powloka.Run "net stop Perforce", 0, True
    Offline = True
End Function

Function LooksAlive()
    LooksAlive = True
End Function

Function IsAlive()
    Dim powloka, kod
    Set powloka = CreateObject("WScript.Shell")
    kod = powloka.Run("""C:\Program Files\Perforce\p4.exe"" -p localhost:1666 info", 0, True)
    IsAlive = (kod = 0)
End Function

Function Terminate()
    Terminate = True
End Function
```

Dodajemy rolę poleceniem z Fiszki (`Add-ClusterGenericScriptRole`). Uwaga na `IsAlive`: gdy `p4.exe` zawiśnie, `Run` z oczekiwaniem zawiśnie razem z nim, więc w prawdziwym wdrożeniu dodajemy limit czasu albo sprawdzamy sam port.

Wymagania wobec usługi

Nie każda usługa nadaje się do klastra. Zanim ją dodamy, sprawdzamy sześć rzeczy.

Wymaganie	Dlaczego	Jak sprawdzamy
Wszystkie dane (baza, pliki, dziennik) na dysku klastrowym	po przeniesieniu roli drugi węzeł musi zobaczyć te same dane	ścieżki w konfiguracji usługi wskazują literę dysku klastrowego
Ta sama wersja i konfiguracja na obu węzłach	po przeniesieniu usługa startuje z plików drugiego węzła	`Get-Package` albo wersja pliku wykonywalnego na obu węzłach
Typ uruchamiania Manual	automatyczny start usługi zderza się z klastrem, który włącza ją sam	`Get-Service NazwaUslugi \`
Konto usługi dostępne na obu węzłach	inaczej usługa nie wstanie po przeniesieniu	gMSA z obiema maszynami w grupie, patrz Konta usługowe MSA i gMSA, albo konto lokalne systemu
Łączenie po nazwie punktu dostępu, nie po nazwie węzła	klient musi trafić na aktualnego właściciela	konfiguracja klientów, DNS i licencje oparte o nazwę lub adres
Czysty start po nagłym zatrzymaniu	po awarii węzła usługa startuje tak, jakby ktoś wyrwał wtyczkę	test: wyłączenie węzła i start na drugim (sekcja niżej)

Ostatni punkt jest najważniejszy, bo klaster nie przenosi pamięci procesu. Użytkownik z otwartym połączeniem je traci, a operacja w toku przerywa się i wymaga powtórzenia. Konto gMSA wymaga jeszcze, żeby oba węzły były w grupie z prawem pobierania hasła (`PrincipalsAllowedToRetrieveManagedPassword`).

Testy przełączenia

Test robimy poza oknami zamknięcia miesiąca i poza zakazem zmian od 15 listopada do 6 stycznia. Do każdego testu mierzymy przerwę z komputera, który nie jest węzłem klastra: pętlą, która co sekundę sprawdza port punktu dostępu i zapisuje wynik z godziną.

```
$cel = "NazwaPunktuDostepu.ad.zielona.example"
$port = NumerPortu
$koniec = (Get-Date).AddMinutes(10)
while ((Get-Date) -lt $koniec) {
```

```
    $stan = Test-NetConnection -ComputerName $cel -Port $port -InformationLevel Quiet -
WarningAction SilentlyContinue
    "{0};{1}" -f (Get-Date -Format "HH:mm:ss"), $stan | Add-Content -Path
"C:\Temp\przerwa.csv"
    Start-Sleep -Seconds 1
}
```

Przerwę odczytujemy jako czas od pierwszego wiersza z `False` do pierwszego wiersza z `True` po nim. `Test-NetConnection` na niedostępnym porcie potrafi czekać kilka sekund, więc wynik traktujemy jako pomiar z dokładnością do kilku sekund.

Trzy testy, od najłagodniejszego:

1. Przeniesienie roli: `Move-ClusterGroup -Name NazwaRoli -Node NazwaWezla2`. Sprawdzamy `Get-ClusterGroup`, a w pomiarze szukamy krótkiej luki. Usługa jest zatrzymana na jednym węźle i uruchomiona na drugim, więc połączenia się zrywają.
2. Prace planowe: `Suspend-ClusterNode -Name NazwaWezla1 -Drain -Wait`, potem restart węzła i `Resume-ClusterNode`. To scenariusz aktualizacji Windows. Dla roli z usługą też jest to przeniesienie z przerwą, a nie migracja na żywo.
3. Awaria: wyłączamy maszynę wirtualną węzła z jej hosta, bez zamykania systemu: `Stop-VM -Name p4n1 -TurnOff`. Klaster wykrywa brak sygnałów i uruchamia rolę na drugim węźle, a przerwa jest najdłuższa ze wszystkich trzech testów.

Po każdym teście zbieramy `Get-ClusterLog` i zapisujemy czas przerwy obok daty, wersji usługi i liczby użytkowników, którzy zauważyli problem. Z trzech pomiarów robimy tabelę na stronie usługi, a deklarowany czas przełączenia w umowie z właścicielem usługi opieramy na najgorszym z nich, a nie na średniej.

Monitoring w Zabbixie

Klaster sam wie, co jest `Online`, ale ten stan nie mówi, czy usługa odpowiada. Dlatego patrzymy z trzech stron: z zewnątrz na punkt dostępu, od środka na rolę i w dzienniku zdarzeń.

Pierwsza strona to sprawdzenie portu z `mon01` przez nazwę punktu dostępu, które widzi to samo, co użytkownik. Na hoście w Zabbixie dodajemy element danych typu prosta kontrola: `net.tcp.service[tcp,NazwaPunktuDostepu.ad.zielona.example,NumerPortu]`. Wynik 0 przez dwie minuty to alarm. Klucz i wartości sprawdzamy w dokumentacji Zabbixa dla używanej wersji [23].

Druga strona to stan i właściciel roli. Na każdym węźle w pliku `zabbix_agent2.d\ft-klaster.conf` dopisujemy parametry użytkownika:

```
UserParameter=ft.klaster.stan[*],powershell.exe -NoProfile -Command "(Get-ClusterGroup -Name
'$1').State"
```

```
UserParameter=ft.klaster.wlasciciel[*],powershell.exe -NoProfile -Command "(Get-ClusterGroup -Name '$1').OwnerNode.Name"
```

Po restarcie agenta element `ft.klaster.stan[NazwaRoli]` zwraca `Online`, a zmiana wartości `ft.klaster.wlasciciel[NazwaRoli]` oznacza, że rola się przeniosła. Alarm na zmianę właściciela jest informacją dla nas: po cichym przełączeniu nikt nie dzwoni, a my chcemy wiedzieć, że klaster został użyty.

Trzecia strona to zdarzenia. Element typu dziennik zdarzeń Windows `eventlog[System,,,"Microsoft-Windows-FailoverClustering","^(1069|1135|1177)$",,skip]` łapie awarię zasobu, usunięcie węzła z klastra i utratę kworum. Numery zdarzeń sprawdzamy w dzienniku `System` na węźle testowym, zanim zbudujemy na nich wyzwalacze. Stan usługi na biernym węźle jest zatrzymany z założenia, więc nie ustawiamy alarmu „usługa nie działa" na każdym węźle z osobna.

Zabbix pilnuje też świadka: osobny parametr użytkownika z `(Get-ClusterQuorum).QuorumResource.State` pokazuje, czy zasób świadka jest `Online`. Świadek, który cicho przestał działać, zamienia klaster dwuwęzłowy w klaster bez zapasu, a tego nikt nie widzi, dopóki nie padnie węzeł.

Częste błędy

Objaw	Przyczyna	Co robimy
Rola nie powstaje, błąd o tworzeniu konta komputera	CNO nie ma prawa tworzenia obiektów komputerów w kontenerze albo konto VCO już istnieje	nadajemy CNO prawo „Create Computer objects" w OU albo przygotowujemy VCO z wyprzedzeniem i dajemy CNO pełną kontrolę nad nim
Usługa wstaje na obu węzłach po restarcie	typ uruchamiania Automatic zamiast Manual	`Set-Service -StartupType Manual` na obu węzłach
Po przeniesieniu usługa startuje z pustymi danymi	dane leżą na dysku lokalnym węzła, nie na klastrowym	przenosimy dane na dysk klastrowy i poprawiamy ścieżki w konfiguracji
Zasób usługi w stanie `Failed` po kilku przełączeniach	przekroczony `FailoverThreshold` w `FailoverPeriod`	ustalamy przyczynę w `Get-ClusterLog`, potem `Start-ClusterGroup`, a progi zmieniamy świadomie
Rola `Online`, ale użytkownicy nie mogą się połączyć	klaster widzi proces, usługa nie odpowiada	monitoring portu z zewnątrz, a w roli Generic Script prawdziwe sprawdzenie w `IsAlive`
Klaster staje po awarii jednego hosta	świadek był maszyną na tym samym klastrze albo go nie ma	Cloud Witness albo udział na serwerze poza klastrem
`Test-Cluster` ostrzega o dyskach w użyciu	storage testy nie ruszają dysków używanych przez role	zatrzymujemy rolę na czas testu albo `-Force`, w oknie serwisowym

Objaw	Przyczyna	Co robimy
Klienci wciąż łączą się z nazwą węzła	konfiguracja klienta wskazuje węzeł, nie punkt dostępu	zmieniamy konfigurację klientów na nazwę roli

Ćwiczenia

1. Dlaczego usługi z klastra nie ustawiamy na start automatyczny?

Odpowiedź

Bo to klaster decyduje, na którym węźle usługa ma działać. Przy starcie automatycznym system uruchomiłby ją na każdym węźle po restarcie, jeszcze zanim klaster przydzieli dysk i adres, a to mogłoby prowadzić do dwóch kopii usługi nad tymi samymi danymi albo do usługi bez dostępu do dysku. Typ Manual zostawia włączanie klastrowi.

1. Co robi `Suspend-ClusterNode -Drain` dla roli z usługą Generic Service, a co dla maszyny wirtualnej jako roli?

Odpowiedź

Dla roli z usługą oznacza przeniesienie z krótką przerwą: usługa zatrzymuje się na jednym węźle i startuje na drugim, a połączenia się zrywają. Dla maszyny wirtualnej jako roli klaster wykonuje migrację na żywo, więc maszyna działa przez cały czas, a użytkownik zwykle widzi tylko krótką pauzę, którą mierzymy w teście. To różnica, na której opieramy porównanie dróg na stronie o Perforce i Lore.

1. Klaster dwuwęzłowy ma świadka w postaci udziału na `fs01`. Co się stanie, gdy padnie `hv01`, na którym działa `fs01`?

Odpowiedź

Znikają jednocześnie węzeł i świadek, więc zostaje jeden głos z trzech. To za mało na większość i klaster przestaje obsługiwać role, chociaż `hv02` żyje. Świadek musi stać poza klastrem, który obsługuje: w chmurze albo na fizycznym serwerze, takim jak `bkp01`.

1. Usługa ma `RestartThreshold` równe 3 i `FailoverThreshold` równe 2 w sześciu godzinach. Co się dzieje, gdy zawodzi co kilka minut?

Odpowiedź

Klaster najpierw próbuje restartować zasób na tym samym węźle, do trzech razy w okresie restartu. Gdy to nie pomaga, przenosi rolę na drugi węzeł, ale tylko dwa razy w ciągu sześciu godzin. Przy kolejnym zawodzeniu rola zostaje w stanie `Failed` i czeka na człowieka, żeby awaria powtarzająca się nie przenosiła usługi bez końca.

Powiązane: Wysoka dostępność: pojęcia, liczenie dostępności i poziomy rezerwy · Replikacja danych między hostami: DRBD, GlusterFS, Ceph, ZFS i replikacja w aplikacji · Hyper-V na co dzień: przełącznik, maszyny, punkty kontrolne i migracja · Wolumeny współdzielone klastra (CSV) · S2D: wymiana dysku bez przestoju, gdy pula jest prawie pełna · Konta usługowe MSA i gMSA · Perforce i Lore w klastrze (strona następna w tej książce, plik `wd2-02-perforce-i-lore-w-klastrze.md`)

Źródła: (odczyt 3 października 2026)

1. learn.microsoft.com: Plan volumes, Storage Spaces Direct, uwaga na początku strony
2. learn.microsoft.com: Add-ClusterResourceDependency
3. learn.microsoft.com: Test-Cluster
4. learn.microsoft.com: Add-ClusterGenericServiceRole oraz Generic Service
5. learn.microsoft.com: Add-ClusterGenericApplicationRole i Generic Application Private Properties
6. learn.microsoft.com: Add-ClusterGenericScriptRole i Scripting Entry Points
7. learn.microsoft.com: Set-ClusterOwnerNode
8. learn.microsoft.com: Move-ClusterGroup
9. learn.microsoft.com: Suspend-ClusterNode
10. learn.microsoft.com: Resume-ClusterNode
11. learn.microsoft.com: Resource Failure
12. learn.microsoft.com: Failover clustering overview
13. Ta baza wiedzy: Wolumeny współdzielone klastra (CSV); learn.microsoft.com: Failover clustering overview, część o Cluster Shared Volumes
14. learn.microsoft.com: Group Common Properties
15. learn.microsoft.com: Manage the cluster quorum i quorum witness i Set-ClusterQuorum
16. learn.microsoft.com: Deploy a Cloud Witness
17. learn.microsoft.com: Storage Spaces Direct overview, sekcja o wdrożeniach
18. learn.microsoft.com: Create Hyper-V VHD Set files
19. learn.microsoft.com: Errors when you back up VMs that belong to a guest cluster
20. learn.microsoft.com: Generic Service, sekcja o ograniczeniach

21. epicgames.github.io: Deploy a local Lore Server
22. learn.microsoft.com: Deprecated features resources
23. zabbix.com: Zabbix agent items i Windows Zabbix agent, klucze `net.tcp.service` i `eventlog` do sprawdzenia w wersji agenta

---

Revision #18
Created 2026-10-03 13:44:01 UTC by Tomasz Konieczny
Updated 2026-10-26 14:21:26 UTC by Tomasz Konieczny