Para FortiGate w klastrze: FGCP aktywny-pasywny

Poziom: średnio

Dotyczy: dwa FortiGate 100F (fw01-a i fw01-b), FortiOS 7.6 i 8.0, FGCP, rdzeń sw01 i sw02, VLAN 98

Czas: ~80 min

Jedna zapora to dziś pojedynczy punkt awarii całej Zielonej: internet, VPN handlowców, poczta i DHCP dla skanerów w Swarzędzu zależą od jednego pudełka. Ta strona opisuje, jak dwa FortiGate pracują jako jedna zapora w protokole FGCP: kto jest główny, jak urządzenia sprawdzają się nawzajem, co się dzieje z sesjami, jak okablować parę, jak ją zbudować, zaktualizować i przetestować. Wracamy tu przy zakupie drugiej sztuki, przy każdej aktualizacji pary i po każdym przełączeniu.

ZAŁOŻENIA SCENARIUSZA

zapora to para `fw01-a` i `fw01-b` w FGCP aktywny-pasywny, z rdzeniem `sw01` i `sw02` i adresem pary 10.20.98.1 w VLAN 98. Zielona ma dwa łącza: światłowód (`wan1`) i Starlink (`wan2`) z publicznym, ale zmiennym adresem, przy czym `wan2` zostaje poza monitorowaniem pary. Magazyn jest w sieci centrali (VLAN 30 na parze). Pozostałe strony zwykle przyjmują pojedynczy `fw01`: FortiGate w zarządzaniu siecią: interfejsy, VLAN-y, polityki i NAT.

Skróty użyte na tej stronie (27)

Skrót	Co znaczy
BMC	Baseboard Management Controller: osobny malutki komputer na płycie serwera, który działa nawet przy wyłączonym serwerze i odczytuje m.in. temperatury czujników
CLI	Command Line Interface: praca przez wpisywanie komend w terminalu
CRC	Cyclic Redundancy Check: suma kontrolna dopisana do odpowiedzi czujnika, dzięki której wykrywamy przekłamany odczyt
DHCP	Dynamic Host Configuration Protocol: usługa, która sama rozdaje komputerom adresy IP
FGCP	FortiGate Clustering Protocol: mechanizm Fortinet, w którym dwie zapory FortiGate działają w parze i druga przejmuje ruch, gdy pierwsza przestaje odpowiadać

Skrót	Co znaczy
FSSO	Fortinet Single Sign-On: mechanizm FortiGate, który podgląda logowania w Active Directory i sam wie, z jakiej grupy korzysta dana osoba, bez proszenia jej o dodatkowe hasło
HA	High Availability: wysoka dostępność, czyli działanie mimo awarii jednego elementu
HTTPS	HTTP zabezpieczony szyfrowaniem TLS
ICMP	Internet Control Message Protocol: protokół diagnostyczny sieci, którym działa między innymi ping
IKEv2	Internet Key Exchange w wersji 2: nowsza odmiana IKE, wbudowana w klienta VPN Windows, szybciej wraca do życia po zerwaniu łącza niż starsze protokoły VPN
IP	Internet Protocol: adres komputera w sieci; w elektryce: stopień ochrony obudowy (Ingress Protection)
IPsec	Internet Protocol Security: zestaw protokołów szyfrujących i uwierzytelniających ruch sieciowy, na nim stoi tunel VPN typu IKEv2
LACP	Link Aggregation Control Protocol: protokół z normy IEEE 802.1AX, w którym dwa urządzenia uzgadniają łączenie kilku kabli w jedno łącze logiczne; jedna sesja zostaje zawsze na jednym kablu
MAC	Media Access Control: unikalny adres fizyczny karty sieciowej, przypisany fabrycznie
MC-LAG	Multi-Chassis Link Aggregation Group: agregacja łączy do dwóch osobnych przełączników, która działa jak jeden przełącznik dla serwera
RADIUS	Remote Authentication Dial-In User Service: standardowy protokół pytania serwera, czy podane konto i hasło są poprawne
SD-WAN	Software-Defined WAN: funkcja zapory FortiGate, która łączy kilka łączy WAN w jedną strefę, mierzy ich jakość sondą i wybiera łącze według reguł
SMS	Short Message Service: krótka wiadomość tekstowa na telefon, która dociera także bez internetu
SPOF	Single Point of Failure: element, którego awaria zatrzymuje całą usługę, bo nie ma dla niego zamiennika
SSH	Secure Shell: szyfrowane zdalne logowanie do serwera
SSL	Secure Sockets Layer: dawna nazwa szyfrowania, dziś formalnie TLS
TCP	Transmission Control Protocol: protokół sieciowy z potwierdzaniem dostarczenia
UDP	User Datagram Protocol: szybki protokół sieciowy bez potwierdzeń
VLAN	Virtual LAN: wydzielona logicznie sieć na tym samym sprzęcie
VPN	Virtual Private Network: szyfrowany tunel do sieci firmowej
WAN	Wide Area Network: sieć rozległa, czyli strona zapory zwrócona do internetu, na przykład interfejsy wan1 i wan2 zapory fw01
WMS	Warehouse Management System: system zarządzania magazynem

Fiszka

Polecenia wykonujemy w CLI FortiOS po zalogowaniu przez SSH na adres zarządzania wybranego urządzenia (10.20.98.30 dla `fw01-a`, 10.20.98.31 dla `fw01-b`) albo na adres pary 10.20.98.1, który odpowiada na urządzeniu głównym. Składnię opatrzoną źródłem sprawdziliśmy w artykułach Fortinet; strony podręcznika `docs.fortinet.com` o konfiguracji klastra otwierały nam się w tym środowisku tylko jako spis treści [11], więc resztę sprawdzamy w dokumentacji naszej wersji FortiOS przed pierwszym użyciem.

TYLKO ODCZYT

Stan klastra: get system ha status

```
get system ha status
```

Pokazuje tryb pracy, listę członków z numerami seryjnymi, który z nich jest główny i czy konfiguracje są zgodne. Pierwsze polecenie po każdym przełączeniu, przed aktualizacją pary i po niej [3].

Przykład (poglądowy, układ pól sprawdzamy w naszej wersji):

```
HA Health Status: OK
Mode: HA A-P
Primary: fw01-a, FG100FTK20000001, HA cluster index = 0
Secondary: fw01-b, FG100FTK20000002, HA cluster index = 1
```

get = FortiOS: polecenia odczytu stanu urządzenia, na przykład get system status pokazuje wersję i numer seryjny

TYLKO ODCZYT

Porównanie konfiguracji obu urządzeń: diagnose sys ha checksum cluster

```
diagnose sys ha checksum cluster
```

Wypisuje sumy kontrolne konfiguracji każdego członka: globalną i dla domeny `root`. Różnica oznacza, że urządzenia nie są zsynchronizowane. Rozbieżną część znajdujemy poleceniem `diagnose sys ha checksum show root NazwaTabeli`, a po naprawie wymuszamy ponowne policzenie przez `diagnose sys ha checksum recalculate` [7].

diagnose = FortiOS: polecenia diagnostyczne zapory i zarządzanych przełączników, na przykład diagnose sys sdwan member pokazuje członków SD-WAN

TYLKO ODCZYT

Pełna konfiguracja HA wraz z wartościami domyślnymi: show full-configuration system ha

```
show full-configuration system ha
```

Pokazuje także ustawienia, których nie zmienialiśmy, na przykład interwał heartbeatu i margines czasu pracy. Sprawdzamy tym polecenie na obu urządzeniach, bo część ustawień HA nie synchronizuje się między nimi [3][5].

HA = High Availability: wysoka dostępność, czyli działanie mimo awarii jednego elementu · **show** = FortiOS: wypisuje konfigurację, w podstawowej postaci tylko ustawienia różniące się od domyślnych, a z wariantem full-configuration także wartości domyślne

TYLKO ODCZYT

Wejście na urządzenie podrzędne: execute ha manage

```
execute ha manage NumerUrzadzenia NazwaAdministratora
```

> (gdzie `NumerUrzadzenia` to numer członka klastra z listy podpowiedzi polecenia, a jego znaczenie sprawdzamy w dokumentacji naszej wersji; konta z RADIUS potrzebują zakresu 169.254.0.0/24 wśród `trusthost`, o czym piszemy w FortiGate w zarządzaniu siecią [12])

Otwiera sesję CLI na drugim urządzeniu z poziomu pierwszego, co przydaje się, gdy adres zarządzania podrzędnego jeszcze nie działa.

RADIUS = Remote Authentication Dial-In User Service: standardowy protokół pytania serwera, czy podane konto i hasło są poprawne · **CLI** = Command Line Interface: praca przez wpisywanie komend w terminalu · **execute** = FortiOS: polecenia, które coś wykonują na urządzeniu, na przykład execute backup config wysyła kopię konfiguracji, a execute ha manage otwiera sesję na drugiej zaporze w parze

OSTROŻNIE · NIEODWRACALNE

Ustawienie HA na jednym urządzeniu: config system ha

```
config system ha
    set group-id 1
    set group-name "fw01-klaster"
    set mode a-p
    set password HasloKlastra
    set hbdev "ha1" 100 "ha2" 50
    set session-pickup enable
    set override disable
    set priority 200
    set monitor "wan1" "LAG_Rdzen"
end
```

To ustawienia `fw01-a` (priorytet 200). Na `fw01-b` identyczne są `group-id`, `group-name`, `mode`, `password` i `hbdev`, a `priority` wynosi 100. Włączenie HA zmienia adresy MAC interfejsów na wirtualne (sprawdzamy w dokumentacji naszej wersji), więc na chwilę przerywa ruch, a urządzenie o niższej randze dostaje konfigurację od drugiego i traci własną. Dlatego robimy to w oknie zmian i w kolejności z sekcji o budowie pary, nie na działającej zaporze w środku dnia. Składnię `hbdev` z priorytetami i `monitor` potwierdzają artykuły o heartbeatach i wyborze urządzenia głównego [4][5], a nazwę interfejsu `LAG_Rdzen` w `monitor` oraz zachowanie agregacji sprawdzamy w dokumentacji naszej wersji.

HA = High Availability: wysoka dostępność, czyli działanie mimo awarii jednego elementu · **MAC** = Media Access Control: unikalny adres fizyczny karty sieciowej, przypisany fabrycznie

ZMIENIA SYSTEM

Zarezerwowany interfejs zarządzania: ha-mgmt-status

```
config system ha
    set ha-mgmt-status enable
    config ha-mgmt-interface
        edit 1
            set interface "mgmt"
            set gateway 10.20.98.1
        next
    end
end
config system interface
```

```
    edit "mgmt"
        set ip 10.20.98.30 255.255.255.0
    next
end
```

Pierwszy blok jest wspólny dla pary. Drugi, z adresem interfejsu `mgmt`, ustawiamy osobno na każdym urządzeniu: 10.20.98.30 na `fw01-a`, 10.20.98.31 na `fw01-b` [1]. Jak zrobić to tak, żeby synchronizacja nie nadpisała adresu drugiej sztuki, sprawdzamy w dokumentacji naszej wersji. Uwaga na podsieć: Fortinet zaleca dla tego interfejsu podsieć inną niż ta, w której pracuje para [1], a nasze adresy leżą w tej samej podsieci 10.20.98.0/24 co interfejs `VLAN98_Zarzadzanie`. Opisujemy to w sekcji o nazwach i adresach.

□□ ZMIENIA SYSTEM

Kontrolowane przełączenie: diagnose sys ha reset-uptime

```
diagnose sys ha reset-uptime
```

Wykonane na urządzeniu głównym zeruje jego licznik czasu pracy w klastrze, więc przy wyłączonym `override` rolę głównego przejmuje drugie urządzenie [6]. W FortiOS od 7.4.0 jest też wariant `diagnose sys ha reset-uptime-primary-only`, który działa tylko na urządzeniu głównym, więc nie pomylimy sztuk [6]. Sesje bez session pickup zostaną zerwane, więc robimy to w oknie zmian.

□□ **diagnose** = FortiOS: polecenia diagnostyczne zapory i zarządzanych przełączników, na przykład diagnose sys sdwan member pokazuje członków SD-WAN

□□ ZMIENIA SYSTEM

Aktualizacja pary: upgrade-mode

```
config system ha
    set upgrade-mode uninterruptible
end
```

Od FortiOS 7.4.1 tryb aktualizacji pary ustawia `upgrade-mode` z wartościami `simultaneous`, `uninterruptible`, `local-only` i `secondary-only`; wcześniej był parametr `uninterruptible-upgrade` [2]. Wartości domyślnej nie potwierdziliśmy, więc ustawiamy ją jawnie i sprawdzamy `show full-configuration system ha`.

Z życia wzięte

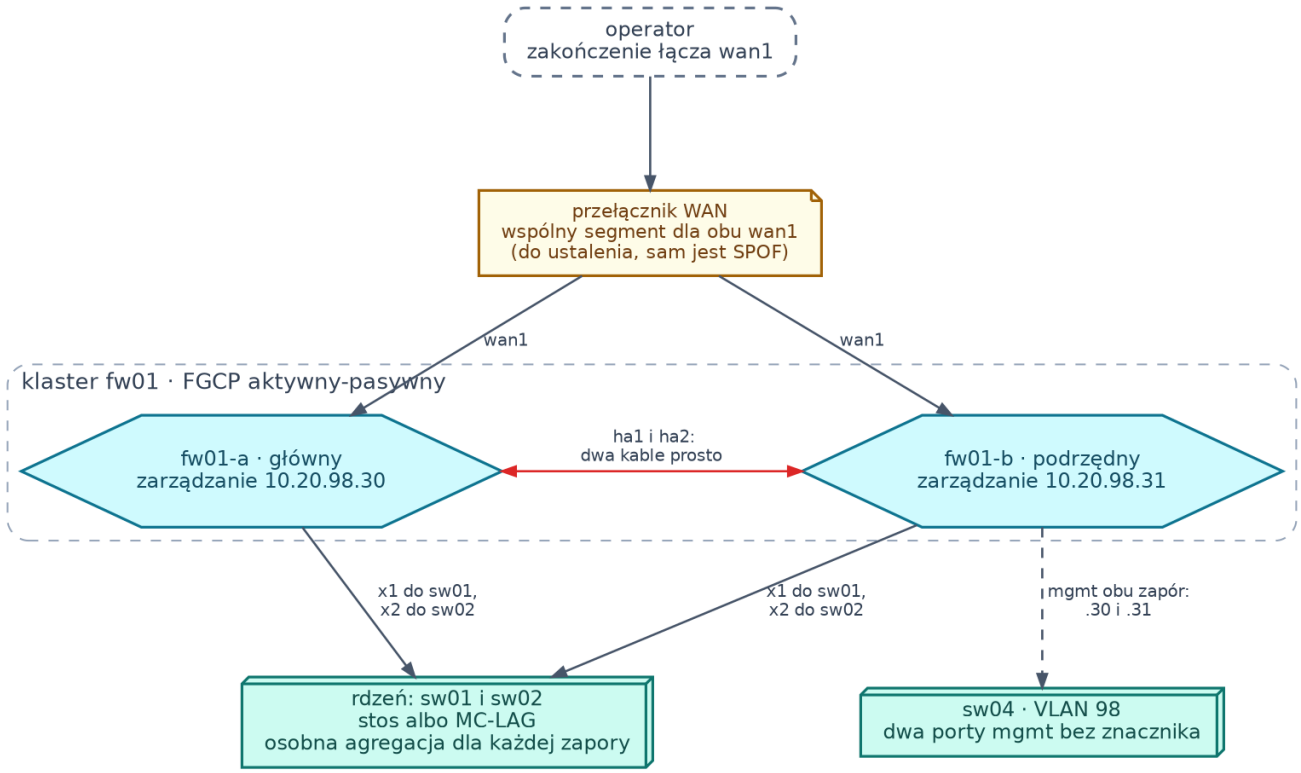
Z ŻYCIA WZIĘTE

Wtorek, 2:40. Zabbix wysyła SMS: `fw01-a` nie odpowiada na ping. Nikt w firmie nie budzi się od tego, bo VPN handlowców i poczta działają dalej: `fw01-b` przejął rolę głównego i odebrał ruch. Rano Tomek wchodzi na adres 10.20.98.1, w `get system ha status` widzi `fw01-b` jako głównego i tylko jednego członka na liście, a po zalogowaniu na konsolę uszkodzonej sztuki znajduje padnięty zasilacz. Zgłasza naprawę, wpisuje wydarzenie do dziennika zmian i nie zmienia ról w ciągu dnia: powrót `fw01-a` na stanowisko głównego zaplanuje na wieczorne okno.

Jak to działa

Rozwiń: jak to działa pod spodem

FGCP (FortiGate Clustering Protocol) to protokół, którym dwa FortiGate rozmawiają ze sobą kablami heartbeat, żeby pracować jak jedno urządzenie: mają wspólne adresy IP i wirtualne adresy MAC na interfejsach, jedno z nich obsługuje ruch, a drugie kopiuje jego konfigurację i czeka. W Zielonej para zastępuje dzisiejsze `fw01`: nazwa `fw01` zostaje nazwą pary, a pojedyncze sztuki nazywamy `fw01-a` i `fw01-b`.



Rysunek 1. Para fw01-a i fw01-b: dwa kable heartbeat prosto między sztukami, osobne łącza X1 i X2 każdej sztuki do sw01 i sw02, porty zarządzania obu sztuk w sw04. Warto zauważyć, że przełącznik WAN po stronie operatora sam jest pojedynczym punktem awarii. (kliknięcie otwiera pełny rozmiar)

Pojęcie: FGCP. Protokół klastrowy FortiGate (FortiGate Clustering Protocol). Dwa urządzenia tego samego modelu z tą samą wersją FortiOS udają jedno, a administrator widzi jedną konfigurację.

Pojęcie: heartbeat. Pakiety, którymi urządzenia klastra co chwilę sprawdzają, czy drugie żyje, i wymieniają stan: dane o konfiguracji, sesjach i interfejsach. Idą osobnymi kablami między dedykowanymi portami HA, nie przez sieć użytkowników.

Pojęcie: urządzenie główne i podrzędne (primary i secondary). Główne obsługuje ruch i jest źródłem konfiguracji. Podrzędne dostaje od niego kopię konfiguracji, a ruch przejmuje, gdy główne zawiedzie. Starsze materiały Fortinet używają słów „master” i „slave”.

Tryb: aktywny-pasywny czy aktywny-aktywny

Cecha	Aktywny-pasywny (a-p)	Aktywny-aktywny (a-a)
Kto obsługuje ruch	tylko główne	oba, ale z zasady rozkładane są tylko sesje wymagające inspekcji treści (UTM) [9]
Co dodatkowo rozkładamy	nic	opcją `load-balance-all` także sesje TCP, opcją `load-balance-udp` także UDP; ICMP, multicast i broadcast nie są rozkładane nigdy [9]
Przepustowość	jak jedna sztuka	zysk tylko tam, gdzie liczy się inspekcja; dla zwykłego TCP i UDP rozkładanie kosztuje narzut [9]
Złożoność	mała	większa: dwa urządzenia naraz przetwarzają ruch, trudniej szukać błędu
Dla Zielonej	proponujemy	nie, dopóki jedna sztuka wystarcza z inspekcją

Zielona ma około 180 osób i jeden operator łącza, więc para ma dać ciągłość, a nie dwukrotną moc. Proponujemy tryb aktywny-pasywny, zgodny z Macierz wysokiej dostępności usług Zielonej. To decyzja autora bazy.

Heartbeat: jak urządzenia sprawdzają się nawzajem

Minimum to jeden interfejs heartbeat, maksimum osiem, a domyślnie większość modeli ma dwa [4]. Interfejs heartbeat nie może być podinterfejsem VLAN, tunelem IPsec, interfejsem redundantnym ani agregacją 802.3ad [4], więc `ha1` i `ha2` w 100F to zwykłe porty fizyczne, a nie część `LAG_Rdzen`. Priorytet każdego heartbeatu (`hbdev`) mieści się w przedziale od 0 do 512, wyższy znaczy preferowany, a domyślnie jest 50 [4]. Domyślnie `hb-interval` wynosi 2 jednostki po 100 ms, a `hb-lost-threshold` 6 pakietów, więc zniknięcie partnera uznajemy po około 1,2 sekundy [4].

Kable heartbeat prowadzimy bezpośrednio między sztukami, bez przełącznika: żaden przełącznik nie psuje wtedy ramek heartbeatu. Fortinet ostrzega, że sprzęt pośredni musi przepuszczać ramki warstwy 2 z typami Ethernetu 0x8890, 0x8891 i 0x8893, a niektóre

przełączniki (jako przykład podano Cisco Nexus) zgłaszają przy nich błędy CRC [4]. Bezpośredni kabel nas z tego zwalnia.

> ⚠ Nie wyjmujemy obu kabli heartbeat naraz w pracującej sieci. Każda sztuka uzna wtedy, że druga zniknęła, i obie spróbują być główne. Taki stan nazywamy podziałem klastra (split brain). Dlatego w teście przełączenia odłączamy tylko jeden kabel.

Które urządzenie jest główne

Wybór zależy od ustawienia `override`. Przy wyłączonym `override` kolejność kryteriów to: liczba działających monitorowanych interfejsów, czas pracy w klastrze, priorytet, numer seryjny. Przy włączonym `override` priorytet wchodzi przed czas pracy: monitorowane interfejsy, priorytet, czas pracy, numer seryjny [5]. Jeśli różnica czasu pracy jest mniejsza niż `ha-uptime-diff-margin` (domyślnie 300 sekund), czas pracy jest pomijany i rozstrzyga priorytet [5].

Każdy restart urządzenia i każda awaria monitorowanego interfejsu zeruje jego „wiek” w klastrze [5], dlatego po restarcie głównego rolę zwykle przejmuje drugie urządzenie i je zostawiamy w tej roli.

Proponujemy `override disable`. Przy włączonym `override` naprawione `fw01-a` wracałoby na głównego samo, w środku dnia, a to drugie przerwanie sesji bez naszej decyzji. Przy wyłączonym wybieramy moment sami poleceniem `diagnose sys ha reset-uptime` [6]. Cena jest taka, że po awarii role zostają odwrócone, aż je zamienimy, co dla sieci nie ma znaczenia, a dla porządku w dokumentacji ma, więc zapisujemy to w dzienniku zmian. Priorytety 200 i 100 zostają jako wskazówka przy remisie i przy równym czasie pracy. `override` i priorytet nie synchronizują się między sztukami, więc ustawiamy je ręcznie na każdej [5].

Monitorowane interfejsy

Interfejs monitorowany to taki, którego zniknięcie zmienia wybór głównego: sztuka z większą liczbą działających monitorowanych interfejsów wygrywa [5]. Dla `fw01` proponujemy monitorować `wan1` i `LAG_Rdzen`. `wan2` zostawiamy poza monitorowaniem, bo drugie łącze obsługuje SD-WAN ze strony Łącza operatora, a utrata łącza zapasowego nie powinna zmieniać roli sztuk.

Monitorowanie pomaga tylko wtedy, gdy awaria jest lokalna: wyrwany kabel, martwy port, awaria przełącznika. Gdy pada łącze u operatora, obie sztuki tracą `wan1` tak samo i przełączenie niczego nie naprawi. Agregacja `LAG_Rdzen` wypada z gry dopiero wtedy, gdy nie

działa żaden jej port (minimalną liczbę łączy sprawdzamy w dokumentacji naszej wersji), więc wyrwanie jednego kabla do rdzenia nie przełącza pary.

Sesje i session pickup

Bez session pickup przełączenie zrywa sesje: użytkownicy przeglądarek odświeżają stronę, ale przesyłanie plików i niektóre protokoły trzeba zacząć od nowa [8]. Z włączonym session pickup urządzenie podrzędne dostaje kopię tablicy sesji, więc część sesji przeżywa przełączenie. Dotyczy to sesji, które nie przechodzą przez inspekcję treści [8]. Sesje UDP i ICMP wymagają osobnej opcji (`session-pickup-connectionless`, składnię sprawdzamy w dokumentacji naszej wersji).

Cena to większy ruch na kablach heartbeat. Fortinet radzi wtedy włączyć `session-pickup-delay`, który ogranicza synchronizację do sesji trwających dłużej niż chwilę, oraz wydzielić interfejs do synchronizacji sesji (`session-sync-dev`) zamiast dzielić go z heartbeatem [8]. W Zielonej w pierwszym kroku zostajemy przy dwóch kablach heartbeat, bo ruch jest umiarkowany, a ewentualne wydzielenie interfejsu planujemy po obserwacji.

Co to znaczy dla Zielonej: sesje skanerów do WMS to zwykły TCP bez inspekcji treści, więc powinny przeżyć przełączenie. Sesje przeglądarek w polityce z inspekcją SSL z FortiGate w zarządzaniu siecią prawdopodobnie nie. Czy tunele IKEv2 handlowców przeżywają przełączenie bez ponownego połączenia, sprawdzamy w teście, a nie zakładamy.

Synchronizacja konfiguracji

Konfigurację zmieniamy tylko na urządzeniu głównym, a pary dbają o resztę. Nie synchronizują się między innymi nazwa urządzenia, priorytet i `override` [5][8], a adres zarezerwowanego interfejsu zarządzania ustawiamy osobno na każdej sztuce [1]. Fortinet zaleca też, żeby sztuki miały różne nazwy hosta przed utworzeniem klastra [8], stąd `fw01-a` i `fw01-b`. Jeśli reguły Grayloga albo Zabbixa szukają `devname=fw01`, poprawiamy je tak, żeby obejmowały obie nazwy.

Zgodność konfiguracji sprawdzamy poleceniem `diagnose sys ha checksum cluster` [7].

> □□ Para nie jest kopią zapasową. Błędna reguła dodana na głównym trafia na podrzędne w tej samej chwili, więc zmiany nadal robimy według procedury z FortiGate w zarządzaniu siecią: kopia pliku przed zmianą, wpis w dzienniku zmian, test po zmianie.

Wymagania i licencje

Obie sztuki muszą być tym samym modelem (100F) i pracować na tej samej wersji i kompilacji FortiOS, a przed dołączeniem drugiej sztuki aktualizujemy ją do wersji pierwszej. Każda sztuka ma własną rejestrację i własne licencje, według Macierz wysokiej dostępności usług Zielonej potrzebne na obie. Tu zapisujemy tylko, że te licencje istnieją i są osobne dla obu urządzeń; warianty i ceny wybiera i płaci pan Zbigniew, a dokładny zakres wymagań sprawdzamy w dokumentacji naszej wersji.

Nazwy i adresy w Zielonej

Element	Wartość	Uwagi
Nazwa pary	`fw01`	tak jak w całej bazie
Urządzenie główne	`fw01-a`, priorytet 200	dzisiejsza zapora produkcyjna
Urządzenie podrzędne	`fw01-b`, priorytet 100	nowa sztuka
`group-name` i `group-id`	`fw01-klaster`, 1	`group-id` zmienia wirtualne adresy MAC, więc dwa klastry w jednym segmencie L2 muszą mieć różne (sprawdzamy w dokumentacji naszej wersji)
Adres pary w VLAN 98	10.20.98.1	odpowiada urządzenie główne
Zarządzanie `fw01-a`	10.20.98.30	port `mgmt` do sw04
Zarządzanie `fw01-b`	10.20.98.31	port `mgmt` do sw04
Adres publiczny	203.0.113.10 na `wan1`	wspólny dla pary, zmienia właściciela przy przełączeniu

Sprawdziliśmy po całej bazie, że adresy 10.20.98.30 i 10.20.98.31 nie występują nigdzie poza tą stroną. Są wolne, ale zwracamy uwagę na regułę z Zdalne zarządzanie serwerem: BMC: karta BMC ma w VLAN 98 ten sam ostatni oktet co serwer w VLAN 10. Serwer `sql01` ma adres 10.20.10.31, ale jest maszyną wirtualną, więc BMC nie ma. Gdyby jednak kiedyś pojawił się fizyczny serwer pod adresem 10.20.10.30 albo 10.20.10.31, jego BMC trafiłoby pod adres zarządzania jednej z zapór. Proponujemy zapisać w NetBoksie oba adresy jako zarezerwowane dla `fw01-a` i `fw01-b`.

> ⚠ **Decyzja do rozstrzygnięcia na początku.** Adresy .30 i .31 leżą w tej samej podsieci 10.20.98.0/24, w której działa interfejs `VLAN98_Zarzadzanie`

> pary (10.20.98.1). Fortinet zaleca dla zarezerwowanego interfejsu zarządzania podsieć inną niż podsieć dostępu do klastra [1]. FortiOS zwykle nie pozwala mieć dwóch interfejsów w tej samej domenie z nakładającymi się podsieciami. Czy zarezerwowany interfejs zarządzania, który pracuje we własnej domenie `vsys_hamgmt`, nakładanie dopuści, sprawdzamy w dokumentacji naszej wersji i w teście na sprzęcie, zanim ułożymy kable. Jeśli nie, mamy dwie drogi: osobna podsieć zarządzania urządzeń (propozycja do decyzji autora bazy, która zmienia adresy .30 i .31) albo rezygnacja z zarezerwowanego interfejsu i wchodzenie na podrzędne poleceniem `execute ha manage`. Pozostaje też ryzyko, że ta sama domena rozgłoszeniowa dochodzi do `fw01` dwiema ścieżkami: przez `LAG_Rdzen` i przez port `mgmt` w sw04.

Na politykę dostępu z VPN administratorów do VLAN 98 (opisaną w Zdalne zarządzanie serwerem: BMC) dopisujemy adresy .30 i .31 jako cel SSH i HTTPS, bo ruch z puli 10.20.200.0/24 do zarządzania urządzeniem podrzędnym przechodzi przez urządzenie główne.

Okablowanie

Skład szafy R1 opisuje Przełączniki: L2 i L3, tablica MAC, VLAN, LACP i STP: `fw01` łączy się z rdzeniem agregacją dwóch łączy 10G, a sw04 zbiera zarządzanie. Przy parze dochodzą kable drugiej sztuki.

Połączenie	Kabel	Uwagi
`fw01-a` `ha1` do `fw01-b` `ha1`	bezpośredni, krótki	pierwszy heartbeat
`fw01-a` `ha2` do `fw01-b` `ha2`	bezpośredni, krótki	drugi heartbeat
`fw01-a` `X1` do sw01, `X2` do sw02	światłowód lub DAC 10G	osobna agregacja LACP tej sztuki
`fw01-b` `X1` do sw01, `X2` do sw02	jak wyżej	druga, niezależna agregacja LACP
`mgmt` obu sztuk do sw04	miedź, port dostępowy VLAN 98 bez znacznika	zarezerwowany interfejs zarządzania
`wan1` obu sztuk	do wspólnego segmentu warstwy 2 z łączem operatora	wymaga małego przełącznika albo konwertera z dwoma portami po stronie zapór
`wan2` obu sztuk	jak wyżej, dla drugiego łącza	według Łącza operatora

Fortinet zaleca, żeby przy agregacji do dwóch przełączników każda sztuka klastra aktywno-pasywnego miała po stronie przełącznika własną grupę LACP [8][10]. Czy dwie osobne grupy

działają przy stosie albo MC-LAG, które wybierzemy w rdzeniu, sprawdzamy w dokumentacji przełączników i w teście, patrząc w tablicę MAC, czy wirtualny adres MAC zapory nie „skacze” między grupami. Dla pary potrzebujemy w rdzeniu czterech portów 10G dla zapór: dwóch w sw01 i dwóch w sw02.

Dwie rzeczy z tabeli są nowe w stosunku do opisu jednej zapory. Po pierwsze, `wan1` obu sztuk musi trafić do tej samej sieci warstwy 2 u operatora, więc potrzebny jest mały przełącznik albo urządzenie z dwoma portami po stronie zapór, a on sam jest nowym pojedynczym punktem awarii. Po drugie, wiele łączy operatorskich wiąże adres MAC urządzenia po stronie klienta, więc przed uruchomieniem pary pytamy operatora, czy wirtualny MAC zapory mu nie przeszkodzi.

Budowa pary krok po kroku

Pierwszy raz robimy to w oknie zmian z kalendarza z FortiGate w zarządzaniu siecią, czyli w środę albo czwartek po 19:00, nie w ostatnich trzech dniach roboczych miesiąca, nie w piątek po 15:00 i nie między 15 listopada a 6 stycznia.

1. Robimy kopię konfiguracji `fw01` ręcznie, na bkp01, i zapisujemy jej nazwę w dzienniku zmian.
2. Wpisujemy w NetBoksie `fw01-a` i `fw01-b`, adresy .30 i .31 i kable z tabeli powyżej.
3. Aktualizujemy `fw01-b` do tej samej wersji i kompilacji co `fw01-a` (`get system status` na obu) i nadajemy mu nazwę `fw01-b`. Nowa sztuka na tym etapie nie ma jeszcze kabli do rdzenia ani do operatora.
4. Na `fw01-b` ustawiamy HA z priorytetem 100, a na `fw01-a` HA z priorytetem 200, według bloku `config system ha` z Fiszki. Hasło klastra zapisujemy w Vaultwardenie i w zalakowanej kopercie obok hasła konta awaryjnego opisanego w FortiGate w zarządzaniu siecią, bo przy awarii `fw01` Vaultwarden bywa nieosiągalny.
5. Łączymy kable `ha1` i `ha2` oraz porty `mgmt`. Sprawdzamy `get system ha status` na obu: `fw01-a` ma być głównym, a `diagnose sys ha checksum cluster` ma pokazywać zgodne sumy.
6. Ustawiamy zarezerwowany interfejs zarządzania i adresy .30 oraz .31 po rozstrzygnięciu kwestii podsieci z sekcji o adresach. Sprawdzamy, że oba adresy odpowiadają z VPN administratorów.
7. Dopiero teraz podłączamy kable `X1`, `X2` i `wan1` drugiej sztuki. Przy rozłączonych portach danych nowa sztuka ma mniej działających monitorowanych interfejsów, więc nie wygra wyboru głównego [5].
8. Wykonujemy test przełączenia z następnej sekcji i dopiero po nim kończymy okno.
9. Aktualizujemy dokumentację: wiersz w macierzy HA, NetBox, reguły Grayloga i Zabbixa dla dwóch nazw.

Kolejność kroków 4 i 7 ma jeden powód. Sztuka, która przegra wybór głównego, dostaje konfigurację od wygranej i własną traci. Gdyby czysta fabrycznie `fw01-b` została głównym, na `fw01-a` zostałaby domyślna konfiguracja zamiast produkcyjnej. Przed krokiem 5 nic nie robimy, dopóki nie mamy świeżej kopii z kroku 1.

Aktualizacja pary bez przerwy

Dla pary ustawiamy `upgrade-mode uninterruptible`, co od FortiOS 7.4.1 zastąpiło parametr `uninterruptible-upgrade` [2]. Przy takim trybie kolejność jest następująca [3]:

1. Wgrywamy obraz FortiOS na urządzenie główne.
2. Główne wypycha obraz na podrzędne.
3. Podrzędne aktualizuje się pierwsze i restartuje, potem wysyła potwierdzenie do głównego.
4. Główne aktualizuje się i restartuje, a w tym czasie podrzędne przejmuje rolę głównego i ruch.
5. Pierwotne główne wraca do klastra już na nowej wersji. Czy znów będzie głównym, zależy między innymi od czasu pracy w klastrze i od `override` [3].

Nazwa „bez przerwy” opisuje zamiar, a nie gwarancję: w chwili przełączenia sesje bez session pickup są zrywane, więc okno zmian zostaje takie jak w tabeli z FortiGate w zarządzaniu siecią, a ścieżkę aktualizacji sprawdzamy w Upgrade Path Tool. Czasu przerwy nie szacujemy, tylko mierzymy w pierwszym oknie (sekcja o teście). Postęp aktualizacji śledzimy poleceniem `get system ha status`, a diagnostykę zaawansowaną opisuje artykuł [3]. Gdy chcemy najpierw przetestować nową wersję na jednej sztuce, tryby `secondary-only` i `local-only` aktualizują tylko wybrane urządzenie, a przełączamy się ręcznie [2].

Po aktualizacji sprawdzamy to samo co przy jednej zaporze: `get system status`, tunel IKEv2 z laptopem testowym, kolektor FSSO, przekierowania poczty i logi w Graylogu.

Test przełączenia

Test robimy po zbudowaniu pary i po każdej większej zmianie, w oknie zmian, z laptopem w VLAN 20 i działającym `ping -t` do adresu 8.8.8.8 i do `ora01`, żeby widzieć, ile odpowiedzi zginęło. Liczby zapisujemy w dzienniku zmian; nie podajemy ich tu z góry, bo zależą od naszego sprzętu.

Test	Jak	Czego oczekujemy	Co zapisujemy
Kontrolowane przełączenie	`diagnose sys ha reset-uptime` na głównym	`fw01-b` zostaje głównym, internet działa	liczba utraconych odpowiedzi ping, czas do powrotu pracy

Test	Jak	Czego oczekujemy	Co zapisujemy
Awaria zasilania głównego	odłączamy zasilanie głównego (po uprzedzeniu Marka)	jak wyżej	jak wyżej; czy VPN laptopu testowego wrócił sam
Wyrwanie kabla `wan1` głównego	odłączamy kabel `wan1`	monitorowany interfejs zmienia wybór głównego	czas do powrotu ruchu
Wyrwanie jednego kabla `X1`	odłączamy jeden kabel do rdzenia	agregacja działa dalej, para się nie przełącza	czy cokolwiek zniknęło z ruchu
Wyrwanie jednego kabla heartbeat	odłączamy `ha1`	klaster zostaje zdrowy dzięki `ha2`	stan w `get system ha status`
Sesje	przed testem otwieramy SSH na `ora01` i transfer pliku, po teście sprawdzamy, czy trwają	SSH i TCP bez inspekcji powinny przeżyć, sesje z inspekcją nie	które sesje przeżyły

Czego nie robimy: nie wyjmujemy naraz obu kabli heartbeat (podział klastra), nie testujemy w piątek po 15:00 ani w zakazie zmian od 15 listopada do 6 stycznia, i nie testujemy o 6:00, kiedy 40 skanerów w Swarzędzu loguje się do WMS.

Inne rozwiązania: FGSP i VRRP

Obok FGCP są dwa rozwiązania, które znamy tylko z opisu i dla Zielonej nie proponujemy ich teraz.

- **FGSP (FortiGate Session Life Support Protocol).** Dwa niezależne FortiGate, z których każdy obsługuje ruch, wymieniają tylko tablicę sesji. Konfiguracji same nie synchronizują, a rozkład ruchu między nimi musi zapewnić coś poza zaporą, na przykład routing. Pasuje do sieci z kilkoma bramami, nie do jednego łącza z jednym adresem publicznym.
- **VRRP.** Protokół, w którym urządzenia dzielą wirtualny adres bramy i jedno z nich go obsługuje. Zakres przełączania obejmuje sam adres, a nie sesje ani konfigurację, więc każdą zmianę musielibyśmy robić na obu sztukach ręcznie. Wygrywa tylko tam, gdzie para ma być zbudowana z różnych urządzeń.

Dla jednej firmy z jednym łączem FGCP daje to samo, a konfigurację prowadzimy raz. Szczegóły obu rozwiązań sprawdzamy w dokumentacji naszej wersji, bo ich tu nie potwierdziliśmy.

Czego para nie chroni

Para usuwa awarię jednej zapory, ale nie usuwa pozostałych punktów z listy w Macierz wysokiej dostępności usług Zielonej.

Dalej pojedyncze	Dlaczego
Szafa R1	obie sztuki stoją w tej samej szafie, więc zasilanie szafy, zalanie czy pożar zabierają obie
UPS	według macierzy HA jest jeden UPS z torami A i B z tego samego urządzenia
Łącze operatora	jedno światłowodowe łącze i jeden adres publiczny; drugie łącze to temat SD-WAN w Łącza operatora
Przełącznik WAN	wspólny segment `wan1` obu sztuk to nowy punkt awarii
Konfiguracja	błąd w regule trafia na obie sztuki naraz, więc kopia konfiguracji jest nadal konieczna
Usługi obok zapory	kolektor FSSO i RADIUS stoją na `dc01` i `dc02`, a nie w parze

Wpis do macierzy HA

Macierz w Macierz wysokiej dostępności usług Zielonej ma dziś wiersz „Zapora, VPN i brama internetowa” z docelowym mechanizmem „druga sztuka w parze HA, tryb aktywny-pasywny (FGCP)”. Wprowadzenia zmian w tej stronie nie robimy; proponujemy autorowi bazy tę treść do wiersza i do SPOF 9.

Pole	Propozycja
Mechanizm docelowy	para `fw01-a` i `fw01-b` w FGCP aktywny-pasywny, `override disable`, session pickup, osobne licencje na obie sztuki (etap 3)
RTO / RPO	RTO 15 min zostaje górną granicą, a wartość rzeczywistą wpisujemy po pierwszym teście przełączenia; RPO 0
Strona	Para FortiGate w klastrze: FGCP aktywny-pasywny
SPOF 9	po parze zostają: jedno łącze, przełącznik WAN, szafa R1 i UPS

Częste błędy

Objaw	Przyczyna	Co robimy
Nowa sztuka przejęła ruch i zniknęła konfiguracja	wygrała wybór głównego, więc podrzędna dostała jej pustą konfigurację	przed łączeniem kabli danych sprawdzamy `get system ha status`, a przywracamy konfigurację z kopii z kroku 1
Klaster się nie tworzy	inna wersja FortiOS, inne `group-name`, `password` albo `group-id`	porównujemy `show full-configuration system ha` i `get system status` na obu sztukach
`diagnose sys ha checksum cluster` pokazuje różnice	zmiana wykonana poza urządzeniem głównym albo przerwana synchronizacja	szukamy tabeli z różnicą według [7], po naprawie wymuszamy `recalculate`
Po restarcie `fw01-a` rolę przejęło `fw01-b`	restart zeruje wiek w klastrze, a `override` jest wyłączone	zostawiamy role, a zamianę robimy poleceniem `diagnose sys ha reset-uptime` w oknie
Po awarii kabla `wan1` nic się nie przełączyło	awaria po stronie operatora, obie sztuki straciły łącze tak samo	to nie jest awaria zapory, sprawdzamy łącze i SD-WAN
Zerwane sesje po przełączeniu	brak session pickup albo sesja z inspekcją treści	sprawdzamy `session-pickup` w `show full-configuration system ha`, a aplikacje z inspekcją testujemy osobno
Przełącznik rdzeniowy widzi „skaczący” adres MAC zapory	jedna agregacja na obie sztuki albo wspólna grupa LACP	każda sztuka ma własną grupę LACP po stronie przełącznika [8][10]
Zalogowany administrator z RADIUS nie wejdzie na podrzędną	brak 169.254.0.0/24 wśród `trusthost`	dopisujemy zakres, jak w FortiGate w zarządzaniu siecią [12]
Adres .30 albo .31 odpowiada, ale nie z VPN	polityka z VPN administratorów do VLAN 98 nie obejmuje nowych adresów	dopisujemy .30 i .31 do obiektu adresowego VLAN 98
Nowa para gubi pakiety do operatora	operator wiąże adres MAC urządzenia po stronie klienta	pytamy operatora o MAC wirtualny pary przed oknem

□□ Ćwiczenia

1. Dlaczego kabli heartbeat nie prowadzimy przez przełącznik rdzeniowy, tylko bezpośrednio między `fw01-a` i `fw01-b`?

Odpowiedź

Heartbeat to ramki warstwy 2 ze specjalnymi typami Ethernetu, które muszą przejść nieruszone. Niektóre przełączniki traktują je jako błędne ramki [4]. Bezpośredni kabel usuwa to ryzyko, a awaria przełącznika nie zabiera nam jednocześnie heartbeatu i ruchu danych.

1. `fw01-a` było głównym, po restarcie jest podrzędnym, a `override` jest wyłączone. Czy to usterka?

Odpowiedź

Nie. Restart zerował wiek urządzenia w klastrze, więc przy wyłączonym `override` czas pracy wyższy ma `fw01-b`, a różnica większa niż 300 sekund rozstrzyga o wyborze [5]. Jeśli chcemy, żeby `fw01-a` wróciło na głównego, robimy to w oknie poleceniem `diagnose sys ha reset-uptime` na `fw01-b` [6].

1. Dlaczego przy dołączaniu nowej sztuki kable `X1`, `X2` i `wan1` podłączamy dopiero po utworzeniu klastra?

Odpowiedź

Przy rozłączonych portach danych nowa sztuka ma mniej działających monitorowanych interfejsów, więc nie wygra wyboru głównego [5]. Gdyby wygrała, jej pusta konfiguracja zastąpiłaby produkcyjną na drugim urządzeniu.

1. Pan Zbigniew pyta, czy dzięki parze FortiGate Zielona przeżyje awarię łącza operatora. Co odpowiadamy?

Odpowiedź

Nie. Obie sztuki dzielą to samo łącze i ten sam adres publiczny, więc awaria po stronie operatora zatrzymuje obie. Para chroni przed awarią zapory, zasilacza, portu albo kabla w rdzeniu. Na awarię łącza potrzebne jest drugie łącze z SD-WAN, o którym piszemy w Łącza operatora.

Powiązane: FortiGate w zarządzaniu siecią: interfejsy, VLAN-y, polityki i NAT · VPN: tunel przez cudzy internet · Przełączniki: L2 i L3, tablica MAC, VLAN, LACP i STP · Łącza operatora · Zdalne zarządzanie serwerem: BMC · Macierz wysokiej dostępności usług Zielonej · Graylog: codzienna praca

Źródła:

1. Technical Tip: HA Reserved Management Interface, Fortinet Community: składnia `ha-mgmt-status` i `ha-mgmt-interface`, osobny adres na każdej sztuce, zalecenie innej podsieci, brak routingu i reguł zapory

2. Technical Tip: New FGCP Cluster Upgrade Mode, Fortinet Community: `upgrade-mode` od FortiOS 7.4.1, cztery wartości
3. Technical Tip: FortiGate HA upgrade procedure and the status during the upgrade, Fortinet Community: kolejność aktualizacji pary, `get system ha status`, `show full system ha`
4. HA heartbeat and communication between cluster units, FortiOS 5.6 Handbook: liczba i rodzaje interfejsów heartbeat, priorytet `hbdev`, typy Ethernetu, `hb-interval` i `hb-lost-threshold` (starsza wersja dokumentacji, wartości domyślne sprawdzamy w naszej wersji)
5. Technical Tip: FortiGate HA primary unit selection process when override is disabled vs enabled, Fortinet Community: kolejność kryteriów, `ha-uptime-diff-margin`, reset wieku, brak synchronizacji `override`
6. Technical Tip: Restoring HA primary role after a failover using diagnose sys ha reset-uptime, Fortinet Community: `diagnose sys ha reset-uptime` i `reset-uptime-primary-only` od 7.4.0
7. Technical Tip: Troubleshooting a checksum mismatch in a FortiGate HA cluster, Fortinet Community: `diagnose sys ha checksum cluster`, `show` i `recalculate`
8. FGCP High Availability Best Practices, FortiOS 6.4 Best Practices: różne nazwy hosta, session pickup, `session-pickup-delay`, `session-sync-dev`, osobne interfejsy przełącznika na każdą sztukę (starsza wersja dokumentacji)
9. Load balancing overview, FortiOS 5.2 Help i Active-passive HA, FortiOS 5.2 Help: co rozkłada tryb aktywny-aktywny, `load-balance-all` i `load-balance-udp` (starsza wersja dokumentacji)
10. FortiGate HA with dual home L3 switches, Fortinet Community: osobna grupa LACP dla każdej sztuki (wątek forum, nie dokumentacja)
11. FortiGate / FortiOS 7.6, Fortinet Document Library i FortiGate / FortiOS 8.0: wersje 7.6.7 i 8.0.1 w dokumentacji; strony o konfiguracji klastra otwierały się tu tylko jako spis treści
12. Technical Tip: How to allow remote administrators to login to the slave unit, Fortinet Community: `trusthost` 169.254.0.0/24 i `execute ha manage`
13. Ta baza wiedzy: Macierz wysokiej dostępności usług Zielonej, Przełączniki: L2 i L3, tablica MAC, VLAN, LACP i STP, Zdalne zarządzanie serwerem: BMC

Revision #11
Created 2026-10-03 19:10:53 UTC by Tomasz Konieczny
Updated 2026-10-26 14:18:17 UTC by Tomasz Konieczny