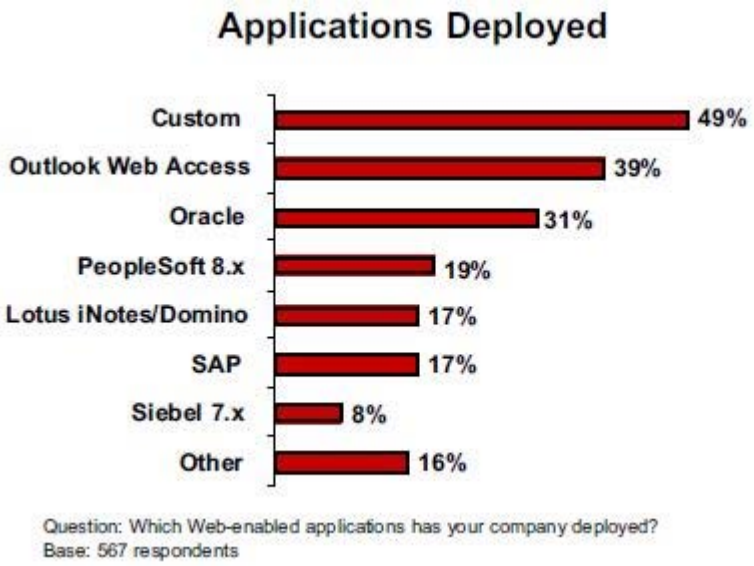
Akceleracja datacenter

– techniczne zasady działania i konfiguracji na przykładzie urządzeń Juniper Networks serii DX

Ewolucja wymagań w datacenter:

Migracja aplikacji korporacyjnych z tradycyjnej architektury klient–serwer do architektury opartej o protokoły HTTP i HTTPS stawia przed centrami przetwarzania danych nowe wyzwania. W wypowiedziach niezależnych ekspertów, między innymi z grupy Gartnera, dominuje opinia, że obecnie trudno sobie wyobrazić aplikację, która nie będzie migrować do postaci web-enabled. Dotyczy to zarówno kosztownych, profesjonalnych narzędzi tworzonych przez duże firmy (np. SAP, Oracle, PeopleSoft), jak i własnego oprogramowania wewnętrznego opartego o technologie PHP, ASP, lub Java.



Rys. 1 - Aplikacje web-enabled wdrażane w firmach (źródło - "Computerworld")

Zmiany te mają na celu obniżenie kosztów rozbudowy tych aplikacji, wystarczy bowiem wprowadzić zmiany tylko po stronie serwera, a nie serwera i klientów, jak musi się dziać w przypadku architektury tradycyjnej. Z drugiej strony proces ten zwiększa wymagania stojące przed projektantami centrów przetwarzania danych. Szacuje się, że komunikacja między przeglądarką WWW, pełniącą rolę klienta w architekturze web-enabled a serwerem aplikacji pochłania więcej pasma niż wymiana danych między klientem natywnym a serwerem. Przykładowo, w przypadku klienta WWW aplikacji SAP zaobserwowano 10-krotny wzrost zużycia pasma.

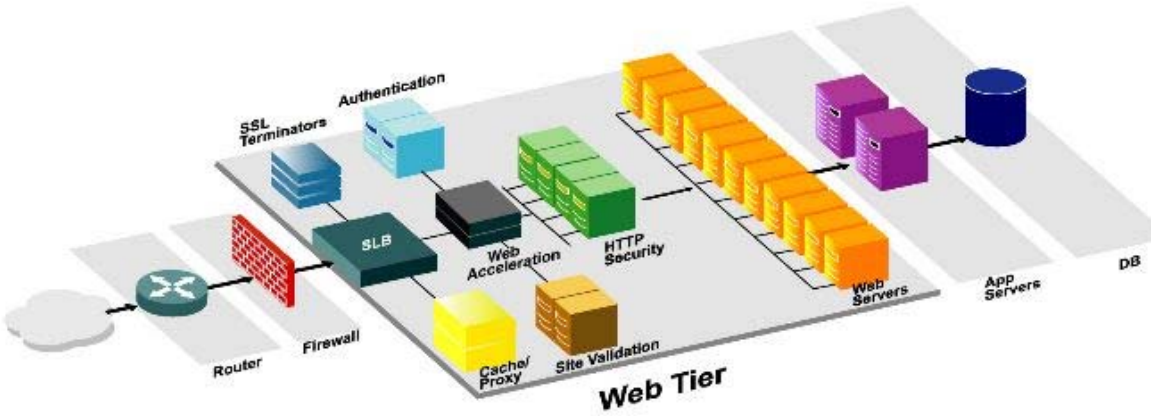
Zadaniem projektantów datacenter w nowych warunkach wymuszonych przez zmianą architektury działania aplikacji sieciowych jest:

- zwiększenie wydajności,
- zwiększenie dostępności aplikacji,
- uproszczenie architektury,
- zabezpieczenie infrastruktury,
- sprawne zarządzanie i monitorowanie całości.

Zwiększenie wydajności stałego rozbudowywania infrastruktury serwerów WWW w datacenter oraz serwerów back-end'owych, wspierających funkcjonowanie udostępnianych usług (serwerów aplikacyjnych i bazodanowych). Z drugiej strony wzrastają trudności związane z utrzymaniem rosnącej ilości serwerów oraz synchronizowaniem ich zawartości. Zamiast ciągłego zwiększania liczby serwerów efektywniejsze jest odciążanie istniejących serwerów od realizacji zadań, które nie są związane z udostępnianiem oferowanych treści. Do takich najważniejszych wymagających obliczeniowo zadań należy zarządzanie tysiącami połączeń TCP.

Równie ważny jest problem dostępności zasobów. Należy zapewnić taki czas odpowiedzi aplikacji, aby przeglądarki WWW nie generowały komunikatów o przekroczeniu czasu ściągania strony (timeout). Aplikacje powinny również odpowiadać na tyle szybko, by nie wywoływać zdenerwowania i frustracji użytkownika końcowego, a co za tym idzie i działu helpdesk. Dodatkowo należy wyeliminować ilość błędnych odpowiedzi widzianych przez użytkowników aplikacji lub oferowanego serwisu.

Nowoczesne datacenter składa się obecnie z wielu elementów budujących warstwę WWW “web tier”, których zadaniem jest optymalizacja efektów działania serwerów WWW. Bardzo często są to urządzenia dedykowane do realizacji zaawansowanych, specyficznych zadań, takich jak Server Load Balancing, akceleracja, caching, terminowanie sesji SSL, autentykacja, czy inspekcja zawartości transakcji HTTP. Zarządzanie taką ilością urządzeń, często pochodzących od różnych producentów staje się coraz bardziej kłopotliwe. Często również wszystkie współpracują ze sobą w topologii gwiazdy, z load balancerem umieszczonym w centralnym miejscu, który dokonuje odpowiedniego przekierowania. Taki system wnosi zbyt duże opóźnienia, a jego stabilność jest trudna do przewidzenia. Im więcej funkcji będzie realizowanych przez serwery WWW, tym więcej ściśle wyspecjalizowanych urządzeń będzie pojawiało się na rynku, a ich wdrożenie będzie coraz bardziej skomplikowane.



Rys. 2 - Klasyczna architektura datacenter

Kolejnym istotnym zagadnieniem jest zabezpieczenie udostępnianych serwisów przed atakami na warstwie aplikacyjnej oraz atakami denial-of-service. Ataki na warstwie aplikacyjnej mogą doprowadzić do penetracji chronionych zasobów przez intruza, kradzieży danych lub zablokowania usługi poprzez zakłócenia pracy serwerów back-end'owych generowaniem błędnych zapytań. Zadaniem ataków denial-of-service jest zablokowanie lub zakłócenie ciągłości działania udostępnianej zawartości lub aplikacji.

Całość struktury datacenter powinna być stale monitorowana, podobnie jak każdy najdrobniejszy aspekt jego funkcjonowania. Platforma zarządzająca powinna umożliwiać sprawne dokonywanie zmian konfiguracyjnych niezależnie od rodzaju urządzenia, nawet w przypadku środowiska heterogenicznego, czyli składającego się z elementów pochodzących od różnych producentów. System generowania raportów ma obejmować zarówno krótkie okresy czasu, jak i długie terminy, umożliwiające wykrycie pojawiających się tendencji kształtowania się ruchu oraz analizę trendów.

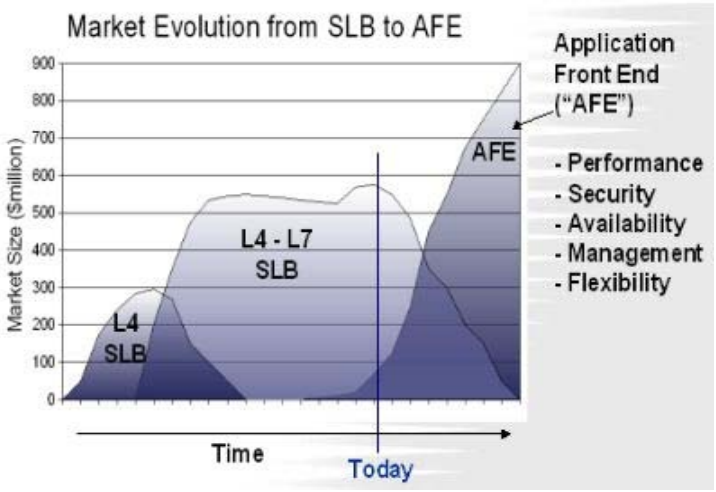
Front-end aplikacyjny:

Biorąc pod uwagę powyższe czynniki należy dążyć do uproszczenia struktury datacenter, aby zastąpić złożoną strukturę pojedynczym rozwiązaniem, zawierającym w sobie wszystkie funkcjonalności niezbędne do obsługi warstwy WWW. Nowo powstała kategoria urządzeń nazywana jest front-endem aplikacyjnym (AFE – application front end). AFE jest w stanie sprostać nowym wymaganiom dzięki realizowaniu na tej samej platformie następujących mechanizmów:

- akceleracji,
- kompresji,
- cachingu,
- server load balancingu,
- terminowaniu sesji SSL,
- high availability,
- systemowi zarządzania, monitorowania i raportowania.

Rys. 3 - Architektura datacenter z wykorzystaniem platformy AFE

Instytut Gartnera przewiduje nadejście zmierzchu zaawansowanych urządzeń do load balancingu warstw 4-7 i zastąpienie ich przez wielofunkcyjne front-endy aplikacyjne, które są lepiej skalowalne oraz umożliwiają wydajne wdrażanie nowych funkcjonalności. Rezultatem ich zastosowania jest przyspieszenie działania firmowych aplikacji sieciowych, przy zmniejszeniu kosztów oraz uproszczeniu architektury sieciowej, jak i samych aplikacji.



Rys. 4 - Ewolucja rynku Application Delivery (źródło: Gartner Institute)

Juniper Networks DX-series:

Urządzenia Juniper Networks DX to wiodący wielofunkcyjny front-end aplikacyjny, znany wcześniej na rynku jako rozwiązania firmy Redline Networks. W ramach jednego rozwiązania integruje mechanizmy obsługiwane przez wąsko wyspecjalizowane urządzenia innych producentów. Jego zastosowanie upraszcza i optymalizuje architekturę datacenter. Podstawowe zadanie tych urządzeń polega na odciążaniu serwerów WWW i serwerów aplikacyjnych od obsługi połączeń sieciowych oraz funkcji wejścia-wyjścia (I/O). Zapewnia wzrost wydajności oraz zwiększenie dostępności aplikacji. Jednocześnie posiada zestaw mechanizmów zabezpieczeń, pozwalając na ochronę krytycznych zasobów. Platforma akceleracji DX jest skalowalna, prosta w zarządzaniu i stanowi solidny fundament pod budowę nowoczesnego centrum przetwarzania danych.

Rys. 5 - Urządzenia Juniper Networks serii DX

W przeciwieństwie do tradycyjnych rozwiązań akceleracji, architektura DX jest oparta na transakcjach HTTP. Oznacza to, że urządzenie w pełni rozumie kontekst transakcji HTTP i HTTPS i może w dowolny sposób manipulować przetwarzanymi danymi z utrzymaniem wysokiej wydajności. Taki stopień funkcjonalności i wydajności nie może być zapewniony przez load balancery wykorzystujące techniki oparte o pakiety, sesje TCP, czy mechanizmy inspekcji warstwy aplikacyjnej. DX pracuje jako pełne proxy HTTP 1.1, dzięki czemu rozumie komunikację i wszelkie zależności aplikacji sieciowych obsługiwanych przez serwery WWW.

Rezultatem zastosowania akceleratorów DX są następujące korzyści:

- wydajność – szybszy dostęp użytkowników końcowych oraz większa efektywna pojemność serwerów – urządzenie tak optymalizuje i kompresuje przepływające dane, że zwiększa się obserwowana pojemność łącza, więc czas ładowania się strony w przeglądarce internetowej użytkownika jest krótszy niezależnie od lokalizacji, z której się łączy. Pojemność aplikacji uzyskiwana jest dzięki temu, że akcelerator pośredniczy w nawiązywaniu sesji między użytkownikami i serwerami, co zwalnia zasoby serwera z konieczności obsługi tych zadań i pozwala mu na realizację tego, do czego jest przeznaczony – udostępnianiem zawartości
- bezpieczeństwo strefy transakcyjnej – odizolowanie warstwy serwerów od niebezpiecznego środowiska sieci zewnętrznych oraz zabezpieczenie przesyłanych danych są zagadnieniami kluczowymi. Platforma DX w pełni wspiera protokół SSL, przy czym dostarcza klientom zaszyfrowaną zawartość szybciej, niż mogłyby to zrobić serwery tzw. czystym tekstem. Ochrona przed atakami DDoS i SYN flood zapewnia utrzymanie dostępności zasobów w przypadku anormalnych warunków występujących w trakcie trwania ataków sieciowych. Pełna znajomość kontekstu transakcji HTTP oraz możliwość modyfikacji danych w locie udostępniają elastyczne i wydajne mechanizmy zabezpieczeń warstwy aplikacyjnej, które można dowolnie dostosować do specyfiki udostępnianych aplikacji;
- dostępność – platforma DX posiada w pełną funkcjonalność load balancingu warstw 4-7. Pozwala na uruchamianie różnych aplikacji obsługiwanych przez różne farmy serwerów za pojedynczym urządzeniem DX. Stan poszczególnych sieci, serwerów i aplikacji jest monitorowany przez dowolnie kształtowany mechanizm health checkingu. Skalowalność i redundancja zapewniana jest przez unikalny mechanizm Active-N. Nawet do 64 urządzeń może pracować równolegle, widocznych jako jeden lub kilka wspólnych wirtualnych adresów IP. Aby zwiększyć możliwości aktualnej struktury akceleratorów, wystarczy dodać tylko jedno urządzenie, w przeciwieństwie do konieczności dodawania urządzeń parami i czasochłonnych zmian w strukturze sieci, jak to się dzieje w przypadku architektury active – standby lub active – active;

- zarządzalność i monitorowanie – ponad 200 statystyk udostępnianych w czasie rzeczywistym zapewnia szczegółowy wgląd w ruch wchodzący i wychodzący. Bogate kryteria definiowania statystyk danych historycznych ułatwiają proces analizy i rozwiązywania problemów. Dzięki temu, że DX jest front-endem dla wszystkich serwerów aplikacyjnych w datacenter umożliwia konsolidację logów oraz analizę dostępności zasobów serwerów. System zarządzania umożliwia zróżnicowanie przywilejów poszczególnych użytkowników, a wszystkie czynności administratorskie rejestrowane są w logach audytowych;
- elastyczność – środowisko programistyczne OverDrive udostępnia API (application programming interface) pozwalające na modyfikację przesyłanych danych w locie, czyli w efekcie zmianę zachowania aplikacji bez konieczności zmiany ich kodu źródłowego. Zestaw intuicyjnych reguł umożliwia definiowanie czynności, które mają zostać podjęte przez urządzenie w zależności od dowolnie definiowanych kryteriów występujących w zapytaniach klientów i odpowiedziach serwerów. W ten sposób zaawansowane reguły bezpieczeństwa mogą zostać stworzone i zaimplementowane bez wniesienia zauważalnych opóźnień. Podobnie stare aplikacje, w których zmiana kodu źródłowego jest niemożliwa lub zbyt kosztowna, mogą zostać dostosowane do nowych wymagań. Dostępność aplikacji może zostać zmaksymalizowana dzięki sprawdzaniu poprawności odpowiedzi serwerów.

Najmniejszym obecnie dostępnym urządzeniem z serii DX jest Juniper DX 3200. Posiada 2 interfejsy Fast Ethernet (100TX), można na nim uruchomić do 64 klastrów aplikacyjnych, czyli aplikacji obsługiwanych niezależnie, w każdym z nich może pracować do 32 serwerów. Nominalna przepustowość urządzenia to 100 Mb/s.

Kolejny model, Juniper DX 3250 posiada te same właściwości co 3200, a dodatkowo wzbogacony jest w sprzętowy akcelerator kryptograficzny, czyli posiada większą wydajność obsługi protokołu SSL.

Następne urządzenie, Juniper DX 3600 można zamówić w dwóch wariantach, jeśli chodzi o konfigurację interfejsów. Pierwszy wyposażony jest w 4 interfejsy Gigabit Ethernet 10/100/1000 TX, drugi posiada 2 interfejsy Gigabit Ethernet miedziane (10/100/1000 TX) oraz 2 Gigabit Ethernet optyczne. Jest to najniższy model, w którym znajduje się redundantne źródło zasilania. Obsługuje do 128 klastrów aplikacyjnych, w których w każdym może działać do 64 serwerów. Pojedyncze urządzenie obsługuje posiada przepustowość do 550 Mb/s.

Model Juniper DX 3650 jest odpowiednikiem wersji 3600, rozbudowanym o sprzętowe wspomaganie szyfrowania SSL.

Dla użytkowników, którzy muszą spełniać szczególne wymagania w zakresie bezpieczeństwa przeznaczone jest urządzenie DX 3650-FIPS. Model ten jest zgodny ze standardem FIPS 140-2, na poziomie 3. Posiada te same możliwości, co DX 3600, ale wyposażony jest jedynie w 2 interfejsy Gigabit Ethernet 10/100/1000 TX.

Juniper DX 3670 jest największym urządzeniem całej serii. Posiada 2 interfejsy Gigabit Ethernet 10/100/1000 TX. Charakteryzuje się wydajniejszą obsługą sesji szyfrowanych SSL, parametry dla ruchu nieszyfrowanego podobne są do osiągów wersji DX 3600.

	DX 3200	DX 3250	DX 3600	DX 3650	DX 3650-FIPS	DX 3670
Application Clusters Supported	64	64	128	128	128	128
Servers per Application Cluster	32	32	64	64	64	64
SSL Transactions per Second	3,000	5,000	5,000	9,000	5,000	18,000
New SSL Connections per Second	250	800	500	1,600	500 (FIPS card specification)	11,000
Simultaneous Connections	50,000	50,000	500,000	500,000	500,000	500,000
FIPS 140-2 Level 3 Certified					Yes	
Network Interface Options	100TX (x2)	100TX (x2)	a) GbE-TX (x4) b) GbE-TX (2x) and GigE Fiber (x2)	a) GbE-TX (x4) b) GbE-TX (2x) and GigE Fiber (x2)	GbE-TX (2x)	1GbE-TX (x2)
Form Factor	1U	1U	2U	2U	2U	2U

Rys. 6 - Zestawienie parametrów urządzeń serii Juniper DX

Akceleratory Juniper DX są urządzeniami przeznaczonymi dla dużych firm, utrzymujących infrastrukturę datacenter, wykorzystujących aplikacje web-based lub oferujących usługi WWW dla klientów zewnętrznych.

DX Framework:

W rozwiązaniach serii DX Juniper Networks zastosowano zestaw mechanizmów DX Framework, aby sprostać wymaganiom stawianym przed platformami AFE. DX Framework składa się z mechanizmów:

- akceleracji,
- dostępności,
- zabezpieczeń,
- kontroli.

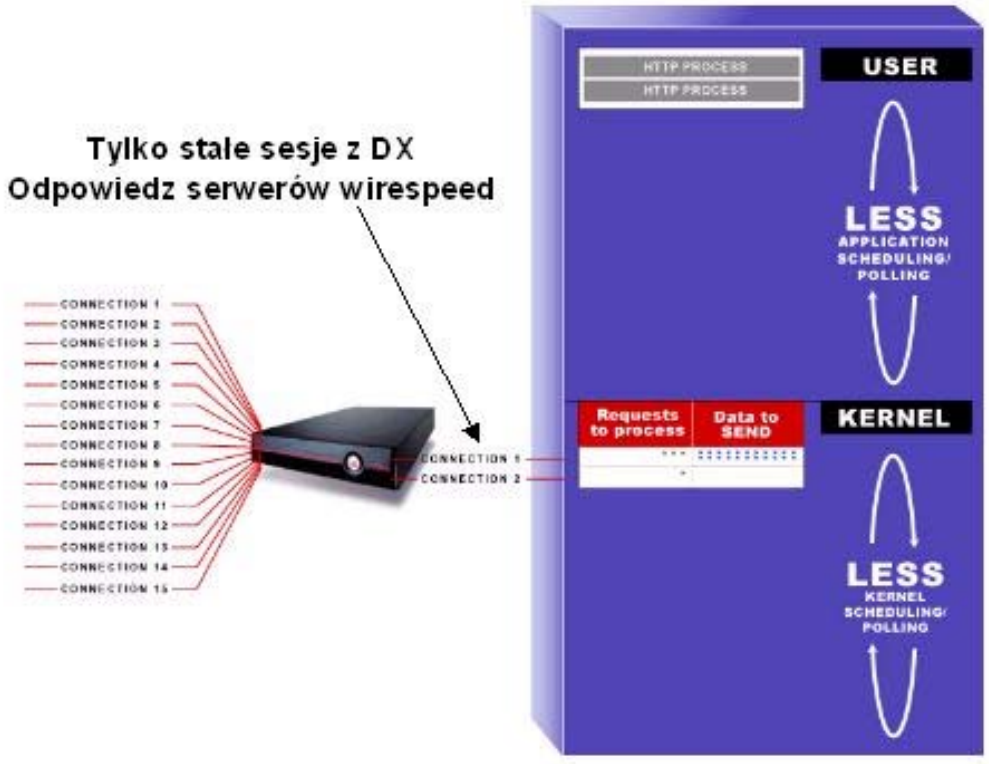
Rys. 7 - DX Framework

Bezpośrednią korzyścią wynikającą z zastosowania mechanizmów akceleracji jest przyspieszenie działania aplikacji oraz skrócenia czasu ściągania udostępnionych stron WWW. Funkcje dostępności zwiększają liczbę transakcji zakończonych powodzeniem oraz zapewniają nieprzerwaną dostępność zasobów. Dodatkowo w dalszej perspektywie przyczyniają się do redukcji wydatków na infrastrukturę datacenter – ta sama liczba urządzeń i serwerów jest w stanie obsłużyć większy ruch. Mechanizmy zabezpieczeń chronią zarówno same serwery, jak i poszczególne transakcje HTTP. Funkcje kontroli i widoczności pozwalają na monitorowanie udostępnianych aplikacji end – to end. Ponadto oferują możliwość modyfikacji danych w locie, aby poprawić wydajność, stabilność i bezpieczeństwo aplikacji.

DX Framework – akceleracja:

Zadaniem mechanizmów akceleracji jest przyspieszenie wydajności aplikacji web-based poprzez odciążenie serwerów oraz kompresję i caching.

Akcelerator DX pracuje w trybie reverse-proxy. Oznacza to, że wszystkie zapytania przychodzące od klientów terminowane są na urządzeniu, a następnie multipleksowane i przesyłane do serwerów docelowych. W ten sposób zwalniane są zasoby serwerów – jądro systemu i procesy odpowiedzialne za działanie oprogramowania serwera WWW, które nie muszą zajmować się zarządzaniem tysiącami TCP nawiązywanymi przez poszczególnych użytkowników. Serwery mogą więc zająć się wyłącznie tym, do czego są przeznaczone – udostępnianiem zawartości. Do komunikacji z akceleratorem wykorzystywanych jest jedynie 6 sesji, które cały czas pozostają aktywne. Szybkość odpowiedzi serwerów jest ograniczona jedynie przepustowością łącza, oczywiście po odliczeniu overhead'u TCP i HTTP. Średnie efektywne zwiększenie pojemności serwerów wynosi od 3 do 4 razy. Średni współczynnik multipleksowania sesji TCP to 1000:1.



Rys. 8 - Architektura reverse proxy

Urządzenie DX może automatycznie kompresować odpowiedzi wysyłane do przeglądarek WWW użytkowników końcowych. Pozwala to zmniejszyć utylizację pasma poszczególnych klientów. Kiedy nagłówek żądania HTTP zawiera parametr “Content-Encoding” oznacza to, że przeglądarka jest w stanie zaakceptować zawartość skompresowaną metodami gzip lub i deflate. Aby zmniejszyć wykorzystanie łącza pomiędzy akceleratorem a serwerami, można również kompresować dane przesyłane pomiędzy nimi. W ten sposób DX może rozkompresowywać dane przychodzące od serwerów, dokonywać ich analizy oraz ewentualnej modyfikacji, z powrotem je kompresować i przesyłać do klientów. Podobnie dzieje się z komunikacją w drugą stronę. Modyfikacja skompresowanych danych odbywa się z wykorzystaniem reguł AppRules Page Translator Content (PTC), na poziomie nagłówków lub treści. Stosując reguły PTC można również kontrolować, które obiekty mają zostać poddane kompresji, w zależności np. od typu MIME. Dotyczy to również tak często spotykanych w sieciach firmowych niestandardowych typów MIME, jak “application/pdf”, “application/msword”, “application/vnd.ms-excel”, czy “application/vnd.ms-powerpoint”.

Rys. 9 - Przykładowy nagłówek HTTP z parametrem "Content-Encoding"

Do przyspieszania transakcji oraz działania aplikacji platforma DX może wykorzystywać mechanizm cachingu. Te elementy stron WWW, które pojawiają się najczęściej w zapytaniach użytkowników, przechowywane są bezpośrednio na urządzeniu. Jeżeli do urządzenia przychodzi zapytanie użytkownika, sprawdzane jest, czy obiekt z wywoływanej strony przechowywany jest cache'u, oraz czy jest aktualny. Jeżeli tak, to przesyłany jest do przeglądarki z pominięciem serwerów, a odbywa się to w sposób całkowicie przezroczysty dla użytkownika. Juniper DX wykorzystuje caching trzeciej generacji (3GCaching), co oznacza, że przy pomocy reguł AppRules w dowolny sposób można określać, które obiekty mają się znaleźć w cache'u oraz na jak długo. Obiekty przechowywane są bezpośrednio w pamięci urządzenia, a nie na twardym dysku, co zwiększa szybkość transmisji. W cache'u może znajdować się ten sam obiekt w formie rozkompresowanej, jak i skompresowanej, a to, kiedy i do kogo dana postać ma być wysłana określa się w AppRules.

```
PTH: http_reply_code equals "200" and
                reply_header "Content-Type" contains "image" then
                 cache "86400"

PTH: http_reply_code equals "200" and
                reply_header "Content-Type" contains "javascript" then
                 cache "604800"

PTH: http_reply_code equals "200" and
                reply_header "Content-Type" contains "css" then
                 cache "604800"
```

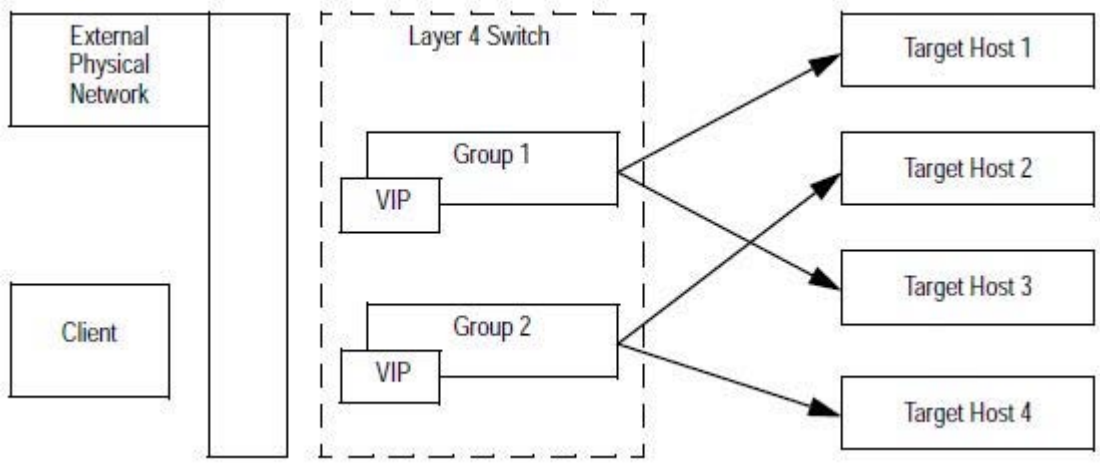
Rys. 10 - Przykładowe zastosowanie AppRules w 3GCaching - przechowywanie elementów graficznych przez jeden tydzień oraz Cascading Style Sheets i JavaScript przez jeden dzień

3GCaching jest funkcjonalnością dodatkowo licencjonowaną. Ponadto, aby korzystać z możliwości udostępnianych przez AppRules, potrzebna jest również licencja na OverDrive.

DX Framework – dostępność

DX Framework zawiera 2 mechanizmy dostępności – Server Load Balancing oraz Active-N, których zadaniem jest zminimalizowanie widocznych przerw w funkcjonowaniu datacenter oraz ułatwienie skalowania całej infrastruktury w miarę rosnących potrzeb.

Obecnie większość aplikacji i serwisów WWW udostępniana jest z kilku lub więcej serwerów, aby dostarczyć odpowiednią moc obliczeniową. Server Load Balancing pozwala na zarządzanie przepływem ruchu do poszczególnych serwerów docelowych lub ich grup. Zwiększenie wydajności odbywa się w sposób niewidoczny dla użytkownika, bez przeryw w działaniu serwisu, poprzez dodanie kolejnej maszyny do istniejącej grupy. Platforma DX rozkłada ruch protokołów HTTP, HTTPS, FTP oraz pozostały ruch TCP i UDP. Użytkownicy odwołują się do wykreowanych na urządzeniu wirtualnych adresów IP (VIP), które reprezentują grupy serwerów (application clusters).



Rys. 11 - Schemat funkcjonowania Server Load Balancingu

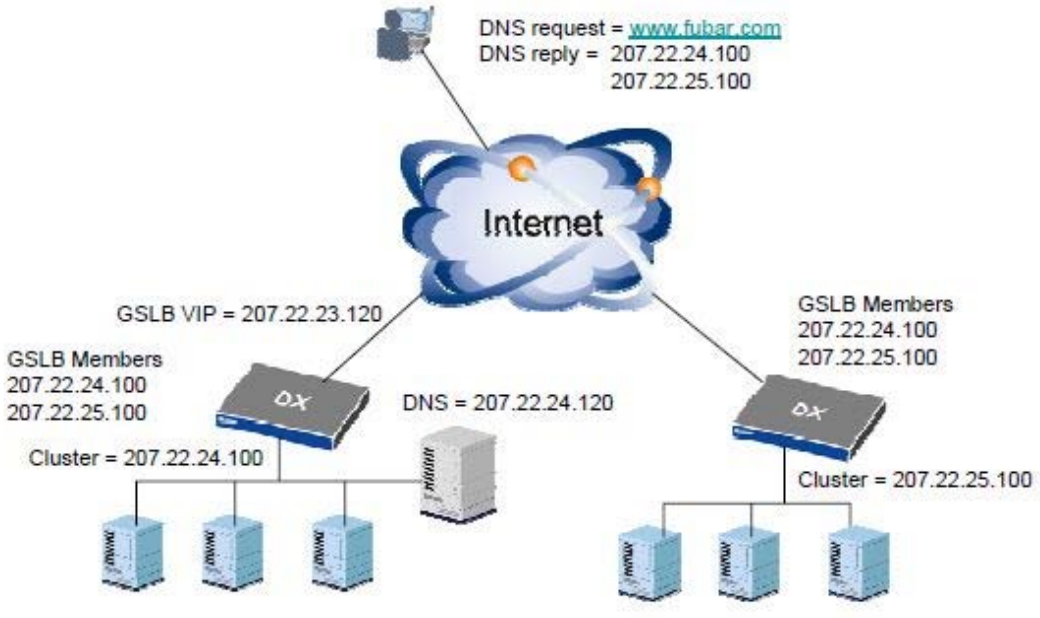
W celu zapewnienia wsparcia dla balansowania protokołów dynamicznych i statycznych urządzenie funkcjonuje jako przełącznik warstwy 4 OSI, w trybie half-NAT lub full-NAT. W half-NAT klient odwołuje się do adresu VIP, który zostaje zamieniony na adres hosta docelowego – adres IP klienta pozostaje nie zmieniony. W trybie full-NAT dodatkowo adres klienta zostaje zastąpiony przez adres z puli NAT akceleratora.

Ruch TCP i UDP może być rozdzielany przy użyciu następujących metod:

- round robin – każde nowe zapytanie jest kierowane sekwencyjnie do następnego serwera w grupie; jeżeli odpytany zostanie ostatni serwer, kolejne zapytanie jest wysyłane do pierwszego;
- weighted round robin – zapytania przydzielane są do serwerów w zależności od przypisanych im wartościom wag; im większa waga, tym większe prawdopodobieństwo, że serwer otrzyma zapytanie; gdy wagi poszczególnych serwerów mają równe wartości, zapytania są przekierowywane sekwencyjnie, czyli tak samo jak w algorytmie round robin;
- least connection – serwer docelowy jest wybierany na podstawie liczby połączeń utrzymywanych przez aktywne serwery; każde nowe zapytanie jest kierowane do serwera, który w danym momencie utrzymuje najmniejszą ilość aktywnych połączeń;
- weighted least connections – jest rozwinięciem metody least connection, które zostało wzbogacone o możliwość definiowania wartości wag przypisywanym poszczególnym serwerom;
- maximum connections – dla każdego serwera definiuje się limit “maxconn”, które określa, jaką maksymalną liczbę połączeń serwer może obsłużyć; następnie wszystkie nadchodzące zapytania kierowane są do pierwszego serwera w grupie, aż do osiągnięcia wartości “maxconn”, potem do kolejnego, itd.;
- backup chaining – wszystkie nowe połączenia kierowane są do pierwszego aktywnego serwera w grupie; ruch do kolejnego urządzenia w grupie kierowany jest jedynie w wypadku awarii pierwszego serwera.

Do optymalnego zarządzania ruchem HTTP i HTTP Juniper DX wykorzystuje opatentowany algorytm Fewest Outstanding Requests. Od tradycyjnych algorytmów load balancingu odróżnia się tym, że podejmuje decyzje na podstawie kontekstu transakcji HTTP, a nie kryteriów z niższych warstw modelu OSI. W przypadku, gdy w danym momencie serwery nie obsługują żadnego zapytania, ruch kierowany jest do pierwszego serwera w grupie. Następne transakcje od tego samego użytkownika kierowane są do tego samego serwera, pod warunkiem, że nie jest zajęty obsługą innych zapytań – pozwala to lepiej wykorzystać lokalny cache serwera i dodatkowo przyspieszyć czas odpowiedzi. Podobnie czas alokacji pamięci na serwerze, algorytmy kolejkowania systemu operacyjnego oraz wiele innych czynników wpływają na to, że najszybciej odpowiada ten serwer, który właśnie ukończył obsługę poprzedniej transakcji. Jednak gdy serwer zajęty jest obsługą ruchu od innego użytkownika, lepiej jest przekierować zapytanie do innego urządzenia w grupie, niż je kolejkować na akceleratorze i czekać na zwolnienie zasobów zajętego serwera. DX analizuje obciążenie każdego serwera w grupie i kieruje połączenia do najmniej obciążonego. Algorytm bierze pod uwagę nie tylko liczbę połączeń, ale również różnice w prędkościach CPU, ilości pamięci, a także specyfiki wywoływanych URL-i – przesyłanie plików graficznych zwykle trwa szybciej, niż stron będących rezultatem działania skryptów (np. cgi-bin). Wynikiem działania metody Fewest Outstanding Requests jest szybsze przetwarzania nadchodzących zapytań, niż w przypadku klasycznych algorytmów balansowania.

Global Server Load Balancing (GSLB) umożliwia dynamiczne rozkładanie ruchu pomiędzy lokalizacjami rozproszonymi geograficznie, poprzez przekierowywanie zapytań do lokalizacji najszybszych lokalizacji z punktu widzenia poszczególnych użytkowników. Gdy dostęp do jednego z oddziałów zostanie odcięty, automatycznie jest on usuwany z grupy do chwili, kiedy znów stanie się dostępny. GSLB polega na manipulowaniu rekordami DNS. Kiedy klient pyta o nazwę hosta, otrzymuje od serwera DNS odpowiedź zawierającą adresy IP, z którymi może się skontaktować, a następnie próbuje skontaktować się z pierwszym adresem z listy, a potem z następnym, gdy kontakt z pierwszym nie jest możliwy, itd. Poprzez manipulację rekordami w odpowiedzi serwera DNS można osiągnąć funkcjonalność load balancingu.



Rys. 12 - Zasada działania Global Server Load Balancingu

W celu sprawdzenia, czy serwery w grupie wykonują poprawnie swoje zadania wykorzystywane są mechanizmy “health checking”. Serwery odpytywane są cyklicznie co określony czas, w przypadku braku odpowiedzi na kilka kolejnych próbek uznawane są za nieaktywne i ruch do nich nie jest przekierowywany. Na platformie DX dostępne są następujące rodzaje health checkingu:

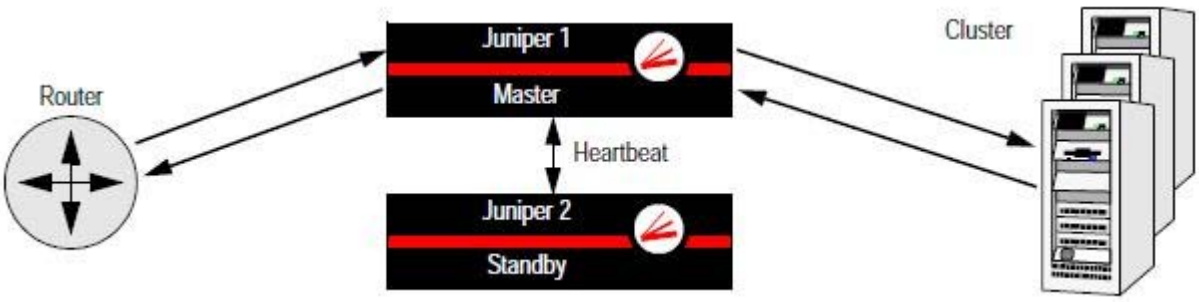
- ICMP (ping),
- TCP SYN,
- UDP,
- Layer 7,
- Scriptable Layer 7.

Health checking w warstwie 7 polega na wysyłaniu do serwerów WWW zapytań o określony URL i sprawdzaniu statusu odpowiedzi. Może być również zastosowany do monitorowania serwerów SMTP – nawiązywane jest wtedy połączenie na port 25 i wydawana jest komenda HELO, po czym sprawdzany jest kod odpowiedzi. Scriptable Layer 7 health checking pozwala dostosować mechanizm stanu sprawdzania serwerów do specyfiki udostępnianych stron lub aplikacji. Za pomocą skryptów Expect/Tcl można definiować np. wywołania HTTP-Get do specyficznych stron i monitorować poprawność odpowiedzi. Skrypty uruchamiane są w specjalnie wydzielonym obszarze pamięci (sandbox), uniemożliwiającym ich ewentualne nieuprawnione lub szkodliwe działanie. Wykorzystanie ich wiąże się z koniecznością wykupienia dodatkowej licencji na OverDrive.

Drugim komponentem mechanizmów dostępność DX Framework jest high availability. HA na platformie DX można zapewnić w następujący sposób:

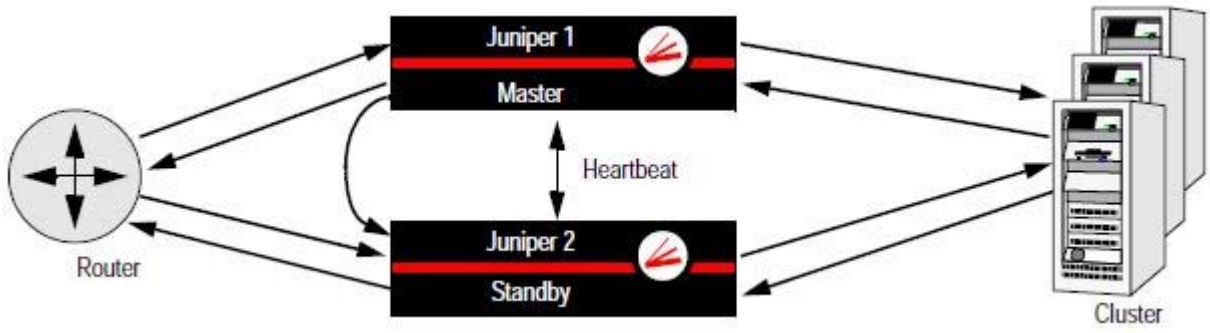
- active - standby,
- active – active,
- activeN.

W konfiguracji active–standby pracują dwa urządzenia, z których jedno aktywnie przetwarza ruch pomiędzy klientami i serwerami, a drugie pasywnie oczekuje na awarię pierwszego. Stale nasłuchuje, czy urządzenie aktywne wysyła sygnały heartbeat. Jeżeli kilka kolejnych heartbeat nie dotrze do pasywnego DX, wysyła pakiet Gratuitous ARP, ogłaszając przejęcie adresów VIP i VMAC. Od tej chwili jest on rozpoznawany jako urządzenie aktywne i cały ruch przechodzi przez niego.



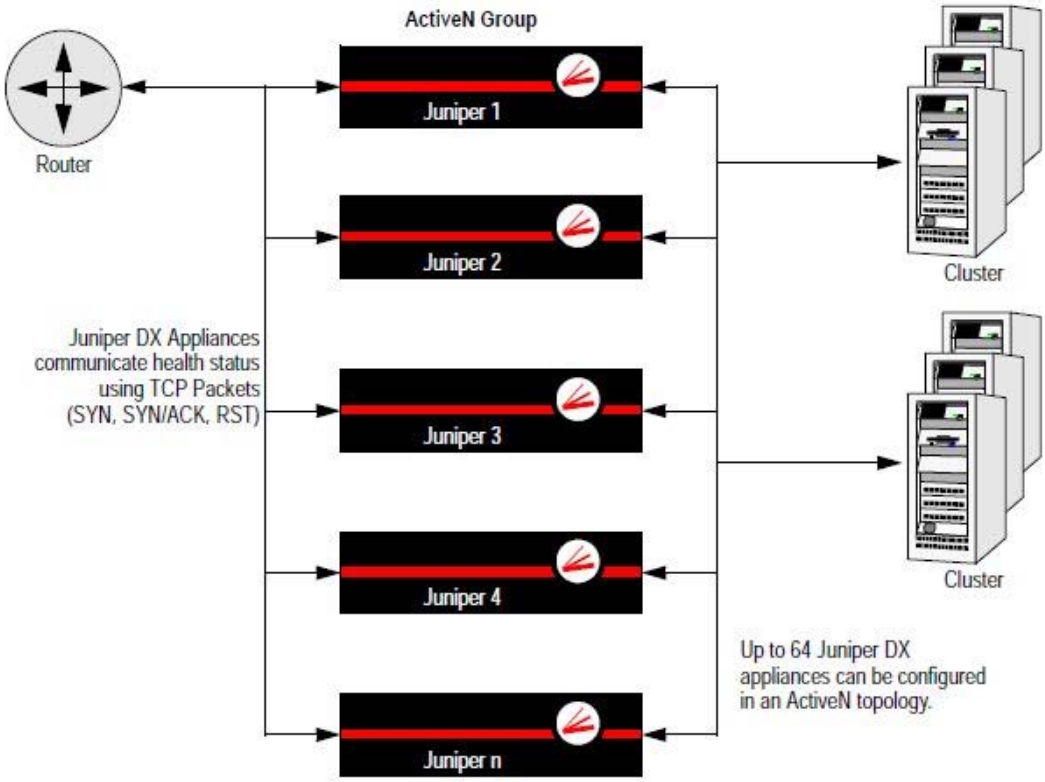
Rys. 13 - Dwa urządzenia DX w konfiguracji active-standby

Konfiguracja active-active to zestaw dwóch urządzeń, które aktywnie przetwarzają transakcje. Jeden DX funkcjonuje jako master, decydując o tym, który ruch ma zostać obsłużony przez niego samego lub przez drugie urządzenie – standby, które w tym przypadku również zajmuje się obsługą ruchu. Jeżeli master ulegnie awarii, jego rola jest przejmowana przez standby.



Rys. 14 - Dwa urządzenia DX w konfiguracji active - active

ActiveN jest unikalnym rozszerzeniem konfiguracji active–active, które pozwala na proste i efektywne skalowanie sieci i jednocześnie wzmacnia odporność całej struktury na awarie. ActiveN umożliwia jednoczesną aktywną pracę do 64 urządzeń DX, które mogą być widziane nawet jako jeden adres VIP, bez konieczności użycia zewnętrznego load balancera do podejmowania decyzji, które urządzenie powinno zająć się konkretną transakcją. Jedno urządzenie pracuje jako master, dystrybuując ruch do pozostałych członków klastra oraz sprawdzając ich dostępność. W przypadku zakłóceń w pracy mastera, pozostałe urządzenia decydują o tym, które z nich zastąpi go w tej roli, bez wpływu na ciągłość przesyłanego ruchu. Dodawania kolejnych urządzeń do klastra nie wymaga zmian ani w strukturze datacenter, ani w konfiguracji pozostałych urządzeń sieciowych, co jest konieczne w przypadku tradycyjnych konfiguracji active-standby i active-active. Jeżeli wolumen ruchu obsługiwanego przez klaster DX zbliża się do granicy jego wydajności, wystarczy dołożenie kolejnego akceleratora – nie ma potrzeby wymiany całej infrastruktury AFE na nowsze modele. Wydajność całego klastra skaluje się liniowo, zwiększają się o możliwości każdego następnego urządzenia.



Rys. 15 - Akceleratory Juniper DX w klastrze ActiveN

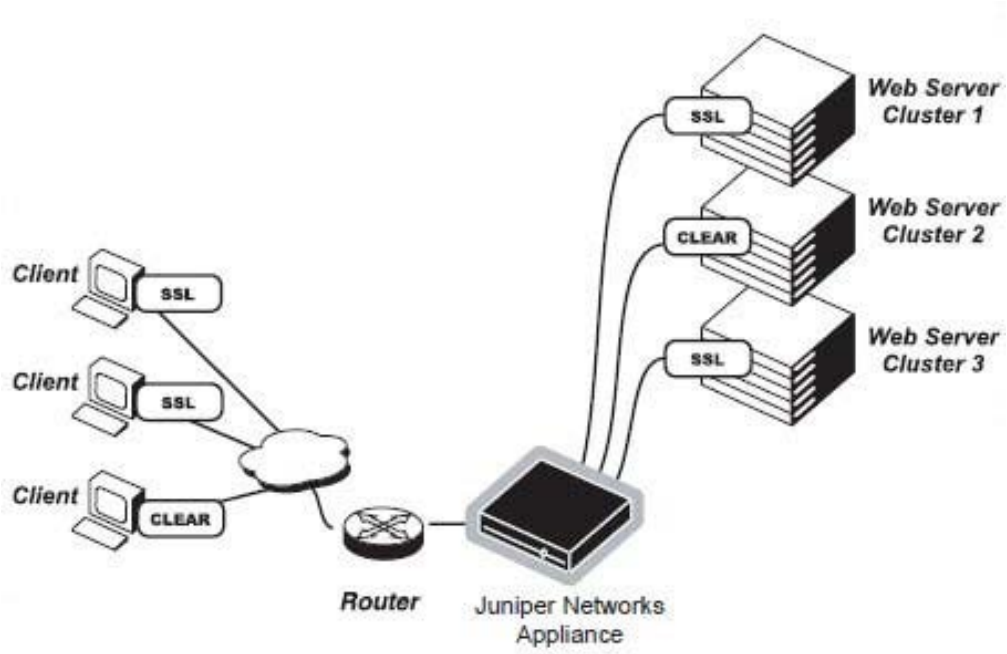
Ze względu na swoje zalety, konfiguracja ActiveN jest rekomendowana przez producenta.

DX Framework – mechanizmy zabezpieczeń:

Zestaw DX Framework udostępnia również mechanizmy pozwalające zabezpieczyć nie tylko ruch klientów, ale i samą infrastrukturę – serwery i akcelerator.

Najpopularniejszym sposobem zabezpieczenia komunikacji aplikacji web-based jest ich szyfrowanie za pomocą protokołu SSL/TLS. Nawiązywanie sesji SSL, zarządzenie nimi, weryfikacja kluczy i certyfikatów, wykonywanie operacji szyfrowania i deszyfrowania to zadania w znacznej mierze pochłaniające zasoby serwerów WWW. Akcelerator DX to urządzenie, które ma pełną funkcjonalność terminatora SSL, może więc koncentrować sesje szyfrowane od klientów i przesyłać do serwerów czysty ruch HTTP, odciążając je w ten sposób od realizacji zadań nie związanych z udostępnianiem zawartości.

Urządzenie DX może pracować jako “forwarder”, pełniąc rolę koncentratora dla sesji SSL, po czym może ponownie zaszyfrować ruch i przesłać do serwerów – jest to SSL end-to-end. To samo urządzenie może wysyłać ruch nieszyfrowany i szyfrowany, ale do różnych klastrów aplikacyjnych, czyli grup serwerów.



Rys. 16 - DX w trybie SSL forwarder

Fakt terminowania sesji SSL na urządzeniu daje możliwość zastosowania AppRules do analizy lub modyfikacji przetwarzanych danych. Transakcje SSL przesyłane dalej do serwerów nie stanowią dla nich znaczącego obciążenia – są optymalizowane podobnie, jak zwykłe transakcje HTTP. Pomiędzy akceleratorem a serwerami zestawionych jest jedynie kilka stałych sesji.

Do wzajemnego uwierzytelniania klientów i serwerów można wykorzystać certyfikaty wystawione przez główne Centra Certyfikacji (Trusted Root CA) lub certyfikaty wystawione przez samego właściciela serwerów (tzw. self-signed). W trybie SSL forwarder podczas realizacji transakcji SSL end-to-end, aby użytkownicy mogli mieć pewność, że dostają się na właściwe strony WWW, trzeba przechowywać certyfikaty na samych serwerach oraz akceleratorze, co nie jest najlepszym rozwiązaniem z punktu widzenia bezpieczeństwa. Lepszym sposobem jest wykorzystanie tzw. certyfikatu pośredniego (Intermediate Certificate), za pomocą którego można wygenerować certyfikaty poszczególnych serwerów, które również będą uznawane za ważne przez przeglądarkę użytkowników.

Juniper DX wspiera wszystkie najpopularniejsze algorytmy szyfrowania i uwierzytelniania: DES, 3DES, AES, RC2, RC4, IDEA, MD5, SHA1.

Oprócz certyfikatów, DX obsługuje również uwierzytelnianie AAA z wykorzystaniem serwerów RADIUS, LDAP oraz RSA SecurID. W razie potrzeby połączenia użytkowników, którzy poprawnie przeszli proces weryfikacji, mogą być przechowywane w cache'u urządzenia, więc osoby te nie muszą się ponownie uwierzytelniać – oczywiście w odpowiednio ograniczonym czasie.

Dodatkowym mechanizmem jest Auto-SSL, pozwalający na przekierowywanie przeglądarki do nowej lokalizacji lub innego protokołu (HTTP lub HTTPS). Można go zastosować, np. w razie szybkiej konieczności zabezpieczenia serwisu lub potrzeby uniknięcia wprowadzania zmian w kodzie stron. Nadsyłane połączenia mogą być przekierowywane na tę samą stronę po HTTPS, nową stronę po HTTP lub nową stronę po HTTPS. DX wysyła odpowiedzi o kodzie HTTP 302 “Temporarily Moved”, a także odpowiednio zmienia sekcję “HTTP Location” w nagłówku HTTP. Dotyczy to również linków wkodowanych “na twardo” w strukturę stron WWW. W ten sposób jedynym nieszyfrowanym komponentem stron przesyłanych do użytkowników są komunikaty HTTP 302.

Ochrona przed atakami denial-of-service, distributed denial-of-service oraz SYN flood jest włączona automatycznie. DX pracuje jako proxy, więc terminuje połączenia do serwerów i w sposób ciągły monitoruje, czy ilość połączeń do danego hosta nie przekracza zdefiniowanych wartości progowych. W niektórych konfiguracjach (np. ActiveN), w których włączony jest tryb DSR (Direct Server Return), należy dodatkowo wyspecyfikować maksymalny czas aktywności sesji TCP, zakończenia sesji przez klienta oraz trwania handshake'u TCP. W trybie DSR odpowiedzi do użytkowników przesyłane są inną drogą – albo bezpośrednia, albo przez inny akcelerator znajdujący się w klastrze.

Niektóre błędne zapytania HTTP mogą doprowadzić do niestabilności lub zawieszenia się oprogramowania serwerów WWW. Co więcej, nawet zapytania zgodne ze specyfikacją protokołu HTTP mogą doprowadzić do zakłóceń w pracy serwerów back-endowych, np. baz danych lub starych aplikacji, których z różnych przyczyn nie można poprawić. Platforma DX udostępnia narzędzia do monitorowania transakcji i wykonywania dowolnych akcji na podstawie kryteriów z nagłówków lub payload'u HTTP. Zapytania użytkowników niezgodne z regułami bezpieczeństwa mogą być blokowane lub przekierowywane na inne strony, czy serwery.

```
RS: url contains ".exe" then close_conn FIN and log
#Deep packet inspection blocks URLs

RS: url contains "%255" then route_request "10.0.0.5"
#Deep packet inspection sends suspicious URLS to honeypot server

RS: request_header "Host" not_contains "mysite.com" then reply 302 "http://www.mysite.com" "/"
#Hosts supplied by clients should always match the name of the site. Virues often use IPs instead

PTH: url starts_with "/" then
                update_reply_header "Server" "Apache 2.0.47 (Amiga)" "Netscape-Enterprise/4.1" "GWS/2.1"
#Server Cloaking functionality
```

Rys. 17 - Przykładowe reguły bezpieczeństwa na DX

Do definiowania reguł bezpieczeństwa wykorzystuje się AppRules.

DX Framework: monitorowanie, zarządzanie, kontrola.

Czwarty, ostatni komponent DX Framework, to mechanizmy zarządzania, monitorowania i kontroli. Ich zadaniem jest zapewnienie administratorom pełnego wglądu w stan datacenter, oraz dostarczenie narzędzi do kontrolowania każdego aspektu komunikacji pomiędzy klientami i serwerami.

Do zarządzania urządzeniami Juniper DX wykorzystuje się DXSHELL, czyli interfejs linii komend CLI (Command Line Interface) lub graficzny interfejs WWW. CLI jest dostępny przez port konsoli, Telnet oraz SSH, do interfejsu graficznego można się dostać po protokołach HTTP lub HTTPS. Dostępne jest również monitorowanie SNMP w wersji 2c.

Baza danych użytkowników akceleratora może być zdefiniowana lokalnie, lub na zewnętrznych serwerach (RADIUS, LDAP). Można określić 5 poziomów uprawnień użytkowników:

- administrator – posiada pełny dostęp do wszystkich poleceń DXSHELL, może również dodawać i usuwać użytkowników oraz zmieniać ich uprawnienia,
- network administrator – może wykonywać wszystkie komendy DXSHELL z wyjątkiem tych, które odnoszą się do SSL,
- network operator – posiada dostęp do tych poleceń DXSHELL, które nie zmieniają ustawień urządzenia, z wyjątkiem komend związanych z SSL, ponadto może włączać lub wyłączać poszczególne serwery w klastrach aplikacyjnych, oraz zmieniać ustawienia dotyczące protokołów zarządzania urządzeniem,
- security administrator – ma pełny dostęp do tych poleceń DXSHELL, które służą do konfigurowania SSL,
- security operator – może oglądać konfigurację i statystyki SSL, bez prawa dokonywania zmian,
- user – może oglądać wszystkie konfiguracji i statystyki, poza SSL, bez prawa dokonywania zmian,
- target operator – może oglądać to, co user, ponadto ma prawo włączać lub wyłączać serwery w klastrach aplikacyjnych.

Dashboard, dostępny przez CLI i interfejs webowy zbiera w jednym miejscu najistotniejsze statystyki dotyczące pracy urządzenia i serwerów: obciążenie pamięci i procesora DX, zajętość łączy, status VIP-a i poszczególnych serwerów – członków klastra aplikacyjnego, ilość bajtów danych przesłanych do klientów, liczba zaakceptowanych i odrzuconych połączeń, obsłużone zapytania, a także liczbę zaoszczędzonych danych.

Można również uzyskać szczegółowe statystyki w zależności od metody zapytania HTTP, rodzaju przeglądarki użytkownika, kodu odpowiedzi serwera, rodzaju zawartości przesyłanej transakcji (z rozróżnieniem na typy MIME), transakcji i rodzaju protokołu SSL.

Część statystyk można oglądać w czasie rzeczywistym. Są to informacje o czasie pracy urządzenia, połączeniach, transakcjach HTTP oraz ilościach danych przychodzących i wychodzących w bajtach.

System monitorowania oferuje również narzędzie przydatne przy podejmowaniu decyzji w procesie capacity planning. Jeżeli któraś z wartości obciążenia urządzenia lub łączy przekroczy wcześniej zdefiniowane wartości progowe, administrator jest informowany o tym fakcie. Zbierane są również wszelkiego rodzaju dane historyczne dotyczące statystyk obciążenia lub zużycia łączy, które mogą być przydatne przy analizie bieżących i przyszłych potrzeb. Wszystkie dane historyczne mogę zostać wyeksportowane do zewnętrznych aplikacji w formacie CSV (Comma Separated Value), a potem mogą zostać poddane dowolnej obróbce w innej aplikacji.

Unikalną cechą platformy DX Juniper Networks jest możliwość ingerencji w dane przesyłanych przez urządzenie aplikacji opartych o protokół HTTP. W tym celu producent udostępnił dedykowane środowisko programistyczne OverDrive, którego interfejs API (application programming interface) przy pomocy intuicyjnego zestawu reguł “if-then” umożliwia zmianę kodu HTML. Zbiór tych reguł nosi nazwę Application Rules Translator, w skrócie AppRules. Już kilkakrotnie w tym dokumencie pojawiły się odniesienia do tego mechanizmu, szczególnie w kontekście wyboru obiektów, które mają być przechowywane w cache'u oraz zabezpieczenie aplikacji. Możliwości zastosowania AppRules jest wiele i praktycznie są one ograniczone jedynie wyobraźnią administratora. Niektóre z nich to:

- przepisywanie odnośników zawartych w stronach z HTTP na HTTPS,
- dodawanie własnych sekcji do nagłówka HTTP,
- przekierowanie niektórych rodzajów zapytań na dedykowane do ich obsługi serwery,
- blokowanie niewłaściwych zapytań HTTP lub usuwanie z nich sekwencji znaków,
- wyszczególnianie typów danych do przechowywaniu w cache,
- modyfikacja zawartości przesyłanych stron, np. dodawanie bannerów,
- ponowne wysłanie zapytania do serwerów przypadku zauważenia błędnej odpowiedzi.

Podsumowując, platforma akceleracji aplikacji Juniper Networks DX reprezentuje nowe podejście do optymalizacji działania farm serwerów WWW i środowisk Datacenter. Minimalizuje niekorzystny wpływ czynników związanych z niespójną architekturą sieci i protokołów na wydajną pracę aplikacji opartych o HTTP/S. Efektem jej wykorzystania jest uproszczenie struktury sieci, lepsze działanie aplikacji oraz ograniczenie kosztów operacyjnych. Dla użytkowników końcowych najważniejszymi korzyściami będą krótszy czas ściągania stron i odpowiedzi aplikacji oraz minimalna ilość komunikatów o błędach.

Sławomir Karaś, JNCIA-J

