

Fizyczna realizacja MS

Fizyczna implementacja serwera mediów może zostać zrealizowana na dwa sposoby:

- Programowo – w postaci aplikacji (software) umiejscowionej na serwerze sprzętowym (np. firmy Hewlett-Packard lub Sun w przypadku serwera mediów firmy Alcatel-Lucent).
- Sprzętowo – jako dedykowany moduł sprzętowy (np. MS firmy Radisys).

W ramach każdej z tych opcji istnieje szereg dalszych wariantów, np.

- implementacja programowa może być podzielona na osobne aplikacje odpowiadające MRFP i MRFC, zlokalizowane na oddzielnych serwerach
- MRFP i MRFC jako osobne aplikacje w ramach jednego serwera.
- MS jako aplikacja zlokalizowana na jednym serwerze. Kilka takich samych serwerów tworzy cluster w sieci celem zachowania redundancji^[1] i zwiększenia wydajności MS.
- implementacja sprzętowa może mieć postać kilku fizycznych maszyn odpowiadających np. jednemu MRFC i kilku MRFP.
- MRFC w postaci aplikacji może kontrolować kilka sprzętowych realizacji MRFP
- fizyczna realizacja rozproszona pomiędzy różne serwery (MS, serwer TTS, serwer z plikami audio, interpreter VXML)

Sposób rzeczywistej implementacji MS zależy zatem bardzo silnie od producenta i operatora sieci telekomunikacyjnej. Szczegółowa analiza wariantów architektury fizycznej nie jest celem tej pracy. Dalsze rozważania dotyczą serwera mediów jako

bloku funkcjonalnego w architekturze funkcjonalnej sieci.

MS a jakość usług (Quality of Service)

W sieciach typu V OIP wydajność pracy serwera mediów ma kluczowe znaczenie dla zapewnienia określonego poziomu jakości świadczonych usług^[2] (QoS). Organizacje standaryzacyjne nie definiują jednak wymagań wydajnościowych dla MS/MRF, określają natomiast kryteria dla konkretnych typów usług. W dokumencie [48] opisano kategorie usług sieciowych (obejmujące usługi związane w przetwarzaniem strumieni audio/video), wyróżnione ze względu na wymagany dla użytkownika końcowego sieci poziom QoS. Z kolei w dokumencie [49] zdefiniowane klasy QoS dla usług w sieciach IP oraz wytyczne dla różnych parametrów wydajnościowych sieci w ramach tych klas.

W związku z powyższym można jedynie rozpatrzyć kluczowe cechy fizycznej implementacji MS, które mają wpływ na jego wydajność i zapewniany przez sieć poziom QoS.

- Cechy implementacji MS wpływające na QoS

Do zbioru tych cech zaliczyć można:

- Aspekty logiczne MS

o Obsługa sygnalizacji – logiczna poprawność implementacji protokołów sygnalizacyjnych (wpływa na opóźnienia wynikające z przesyłania błędnych wiadomości i konieczności ponawiania ich wymiany) o Implementacja algorytmów przetwarzania strumieni a/v – rozpatrywana pod względem logicznej poprawności i optymalności (efektywności) – wpływa na przekłamanie danych i jakość sygnału oraz opóźnienia związane z przetwarzaniem sygnałów

- Aspekty fizyczne pracy MS

o Transport danych multimedialnych

- Szybkość transmisji strumieni a/v w ramach interfejsów serwera mediów do sieci transportowej IP
- Wydajność interfejsów transportowych MS – wpływa na ilość jednocześnie obsługiwanych sesji RTP

o Moc obliczeniowa maszyny na której pracuje MS

- Ilość pamięci
- Przestrzeń dyskowa
- Moc procesora o Stabilność systemu
- Odporność na zawieszanie się systemu
- Czas pracy bez konieczności ponownego uruchamiania MS

o Możliwość tworzenia clusterów – redundantnych struktur instancji serwera mediów – wpływa na zwiększenie mocy obliczeniowej MS oraz zabezpieczenie przed niedostępnością systemu

- Parametry opisujące wydajność pracy MS

Na podstawie wymienionych powyżej cech serwera mediów, które wpływają na poziom QoS jego pracy, można zdefiniować szereg parametrów, poprzez które opisuje się ten poziom.:

- Wydajność i pojemność MS

o Maksymalna ilość połączeń/sesji obsługiwanych jednocześnie przez MS

- Przez sesję RTP z terminalem końcowym użytkownika i wykonywanie wszelkich operacji na strumieniach a/v na potrzeby tego połączenia
- Można rozpatrywać sesje różnego typu (np. maksymalna ilość sesji z transkodowaniem pomiędzy kodekami określonych typów, bez transkodowania etc).

o Maksymalna ilość CAPS (Call Attempts Per Second) – maksymalna ilość obsługiwanych wywołań SIP przesyłanych do MS w ciągu 1 s.

o Maksymalna ilość obsługiwanych BHCA (Busy Hour Call Attempts) – wywołań SIP w godzinach największego ruchu w sieci o Maksymalny czas nawiązywania jednego połączenia (ustanawiania sesji SIP)

o Maksymalna ilość obsługiwanych jednocześnie sesji RTP o Maksymalna ilość obsługiwanych jednocześnie konferencji – tu konferencję należy rozumieć jako instancję miksera

o Maksymalna ilość użytkowników w ramach jednej konferencji – tu ilość użytkowników należy rozumieć jako ilość źródeł strumieni a/v przesyłanych do/z miksera (z każdego źródła mogą pochodzić maksymalnie 2 strumienie a/v)

o Opóźnienie (w ms) pomiędzy odbieraniem a przesyłaniem strumienia (np. przy transkodowaniu strumieni pomiędzy terminalami A i B – strumień odbierany od A jest przetwarzany i przesyłany do B) o Procent wykorzystania zasobów fizycznych maszyny (mocy procesora, pamięci operacyjnej przestrzeni dyskowej) podczas różnych operacji (np. obsługi połączeń konferencyjnych, transkodowanie etc). o Przepustowość (w Mbps) interfejsów sieciowych serwera mediów służących do przesyłania strumieni multimedialnych • Dostępność MS

o High Availability – parametr określający dostępność systemu, rozumianą jako dostępność dla użytkownika i/lub dla systemów z nim współpracujących. Serwer mediów jest zaliczany do systemów o wysokiej dostępności. Dla takich systemów wartość high availability wyraża się zazwyczaj w postaci liczby procentów opisywanej liczbą wchodzących w jego skład cyfr 9^[3]:

- 9% = niedostępność systemu przez 43.8 min/miesiąc lub 8.76 godzin/rok („trzy dziewiątki”)
- 99% = niedostępność systemu przez 4.38 min/miesiąc lub 52.6 min/rok („cztery dziewiątki ”)
- 999% = niedostępność systemu przez 0.44 min/miesiąc lub 5.26 min/rok („pięć dziewiątek ”)

o MTBF (Mean Time Between Failures) – maksymalny czas bezawaryjnej pracy systemu bez konieczności jego ponownego uruchamiania

o MTTR (Mean Time To Repair) – maksymalny czas potrzebny na naprawę systemu po wystąpieniu awarii

o Maksymalny czas trwania restartu systemu od wyłączenia go do chwili przywrócenia konfiguracji sprzed wyłączenia o Maksymalny czas na uaktualnienie (upgrade)

- Oprogramowania MS
- Konfiguracji sprzętowej MS

o Szacunkowa częstotliwość niezbędnych aktualizacji oprogramowania/ konfiguracji sprzętowej MS • Parametry sprzętowe MS

o Ilość pamięci operacyjnej, przestrzeni dyskowej, wydajność procesora o Ilość kart rozszerzeń

- Wspomagających przetwarzanie sygnałów
- Zawierających interfejsy sieciowe

o Miara wydajności jednostki centralnej (CPU) maszyny

- Liczba MIPS (Milion Instructions Per Seconds); określa liczbę milionów operacji stałoprzecinkowych wykonywanych przez CPU w ciągu 1 sekundy^[4]
- Liczba FLOPS (Floating Point Operations Per Second) – liczba operacji zmiennoprzecinkowych wykonywanych przez CPU w ciągu 1 sekundy.

Należy podkreślić, iż praktyka rynkowa pokazuje, że znacznie bardziej efektywna jest implementacja MS w postaci odrębnej maszyny niż w postaci aplikacji umieszczanej na różnego typu serwerach sprzętowych. Przetwarzanie sygnałów, a zwłaszcza transkodowanie, pochłania bowiem dużo mocy obliczeniowej. Operacje te są przeprowadzane bardziej wydajnie i zużywają mniej zasobów w przypadku sprzętowej implementacji serwera

mediów, zwykle przy wykorzystaniu kart sygnałowych z procesorami DSP. Wydajność spada dla programowych implementacji algorytmów przetwarzania sygnałów. Możliwe jest też podejście „mieszane” – tzn. programowa realizacja MS pracująca na serwerze i wykorzystująca zainstalowanie w nim karty z procesorami DSP.

[1] Tutaj: nadmiarowość w stosunku do tego co konieczne celem zmniejszenia prawdopodobieństwa awarii Media Servera w sieci. Serwery z aplikacją MS lub maszyny, które są realizacją MS, są duplikowane i w razie awarii któregoś z nich (lub silnego obciążenia) zadania przejmuje następny. Taka redundantna struktura (cluster) jest przez sieć „widziana” jako jeden Media Server. Algorytmy współdzielenia/przejmowania pracy przez instancje MS w ramach clustera zależą od implementacji i nie są przedmiotem standaryzacji.

[2] Poziom jakości usług świadczonych dla klienta jest określany zwykle w dokumencie SLA – Service Level Agreement

[3] Określanie wartości High Availability jest oparte o dokładne pomiary czasu dostępności danego systemu wykonane z wykorzystaniem niezawodnych narzędzi, jak i przy udziale osób oraz o szacunkowe obliczenia matematyczne. Ostateczna wartość parametru określa dostępność z punktu widzenia użytkownika (administratora systemu) jak i sieci (system może być bowiem dostępny dla administratora ale niedostępny w sieci i odwrotnie).

[4] Zamiennie wielkość ta określana jest terminem MOPS (Million Operations Per Second) – milion operacji na sekundę

Jeśli szukają Państwo pomocy w napisaniu własnej pracy - potrzebują Państwo fachowych konsultacji to polecamy stronę [pisanie prac](#) - profesjonalna pomoc w pisaniu prac w granicach prawa.