

Piotr KASPRZYK

Politechnika Śląska, Instytut Informatyki

WIDEOKONFERENCJA OD WEWNĄTRZ

Streszczenie. Systemy wideokonferencyjne składają się z elementów sprzętowych i programowych. W artykule zostały przedstawione funkcje elementów sprzętowych takich systemów, doświadczenia zdobyte podczas tworzenia aplikacji wideokonferencyjnej w systemie Linux oraz wnioski z eksploatacji gotowego systemu wideokonferencyjnego dla Windows 95.

INSIDE THE VIDEOCONFERENCE

Summary. The videoconferencing systems are a mix of hardware and software elements. The functions performed by hardware items of videoconferencing system are presented. The experiences acquired while developing a videoconferencing application for Linux operating system are described. Results of running a commercial videoconferencing system in Windows 95 environment are also presented.

1. Wstęp

W ostatnich latach można zaobserwować rosnące zainteresowanie wykorzystaniem sieci komputerowych. Stosują je nie tylko uczelnie i duże firmy, coraz częściej można je spotkać także w małych firmach, gdzie pozwalają na pracę kilku osób nad wspólnym projektem lub umożliwiają wymianę informacji między oddległymi oddziałami firmy. Wraz ze zwiększeniem dostępu do różnego rodzaju sieci komputerowych można zaobserwować różnicowanie się jej zastosowań. Pierwotnie głównym zadaniem sieci było umożliwienie transmisji plików, zdalna praca i racjonalne wykorzystanie drogich zasobów (np. drukarki), w chwili obecnej rozpowszechnia się zastosowanie sieci komputerowych do celów rozrywkowych, co pozwala na prowadzenie wspólnej gry przez kilku uczestników. Atrakcyjność takiej gry zachęca do

instalowania sieci komputerowych w celu połączenia prywatnych mieszkań. Większość komputerów przeznaczonych m.in. do zabawy jest wyposażona w karty dźwiękowe, co pozwala na wprowadzanie (za pomocą mikrofonu) oraz nadawanie (za pomocą słuchawek i głośników) dźwięku. Takie wyposażenie pozwala na prowadzenie rozmów z wykorzystaniem sieci komputerowej.

Użytkownik ma do wyboru szereg aplikacji wideokonferencyjnych (patrz [3]), począwszy od NV/VAT, przez CU-SeeMe aż do Netscape CoolTalk i Microsoft NetMeeting.

Kolejnym sposobem uatrakcyjnienia procesu komunikacji jest przesył ruchomego obrazu. Komputer musi być wtedy wyposażony w kamerę, którą można podłączyć do portu Centronics, USB lub do specjalizowanej karty wyposażonej w złącze ISA lub PCI.

Poniżej zostaną przedstawione doświadczenia zdobyte podczas eksploatacji półprofesjonalnego systemu Intel Business Video Conferencing (IBVC).

2. Anatomia sprzętowa systemu

System IBVC (patrz [7]) jest rozwiązaniem sprzętowo-programowym firmy Intel. Składa się z dwóch kart dla komputera klasy PC: karty ISA dla sieci ISDN i wizyjno-dźwiękowej karty PCI. Na karcie PCI są umieszczone specjalizowane układy scalone znanych firm, co pozwala zorientować się w możliwościach karty na podstawie dostępnych opisów tych układów.

W skład systemu wideokonferencyjnego wchodzi karta ISA dla obsługi sieci ISDN. Jest to zwykła karta 2B+D o nazwie ix-micro 3.0 firmy ITK. Pozwala ona na prowadzenie wideokonferencji przez stosunkowo popularną sieć ISDN z szybkością dochodzącą do 128 kb/s.

Na karcie dźwiękowo-wizyjnej znajduje się układ AD1847 firmy Analog Devices (patrz [4]). Pozwala on na obsługę głośników, słuchawek i dwóch mikrofonów (jeden w głośniku, drugi w słuchawkach). Od strony parametrów technicznych AD1847 pozwala na wprowadzanie dźwięku z czterech stereofonicznych źródeł, konwersję analogowo-cyfrową, przetworzenie i przesłanie zarówno do komputera, jak i na linię wyjściową, do przetwornika cyfrowo-analogowego. Transmisja danych dźwiękowych do komputera odbywa się za pomocą szeregowej magistrali cyfrowej, która może wykorzystywać jeden wspólny przewód (w trybie transmisji półdupleksowej) lub dwa przewody, gdzie jednym przewodem komputer przysyła dane dźwiękowe do układu, a drugim otrzymuje dane od układu AD1847. Transmisja danych odbywa się w jednym z czterech zaimplementowanych formatów (jeden 16-bitowy i trzy 8-bitowe), ale wewnętrznie układ wykorzystuje zawsze format 16-bitowy, aby nie tracić precyzji podczas przetwarzania. Protokół transmisji danych posługuje się pojęciem szczelin

czasowych, które są zwykle łączone po 3 w celu przesłania instrukcji sterujących i danych dla obu torów dźwięku stereofonicznego. Do tej magistrali może zostać podłączony cyfrowy procesor sygnałowy, który pozwala na realizację dodatkowych efektów, takich jak filtrowanie szumu. Tę funkcję może wykonywać układ TMS320C52 (opisany poniżej). W chwili obecnej mikrofony zestawu są monofoniczne, a cały system wideokonferencji także wykorzystuje dźwięk monofoniczny, ale można myśleć o jego rozbudowaniu w przyszłości o dźwięk stereofoniczny. Układ AD1847 posiada także inne cechy, które na karcie nie zostały wykorzystane, takie jak możliwość potokowego łączenia układów tej rodziny.

Karta wizyjno-dźwiękowa stosuje układ SAA7196 firmy Philips do wprowadzania obrazu (patrz [5]). W zestawie jest kolorowa kamera CCD o rozdzielczości 510 punktów w poziomie i 492 punktów w pionie, która generuje obraz w systemie NTSC. Za pomocą przełączników umieszczonych na kamerze użytkownik może regulować stopień nasycenia kolorów, ostrość i jasność obrazu. Obraz z kamery jest wprowadzany za pomocą dwóch 8-bitowych przetworników analogowo-cyfrowych (dla sygnału chrominancji i luminancji) do układu SAA7196. Układ ten jest specjalizowanym cyfrowym dekodery sygnału wizyjnego, przeznaczonym dla domowych urządzeń wizyjnych. Konfiguracja układu odbywa się przez odczyt i zapis 64 rejestrów dostępnych przez magistralę I²C. Do układu musi być podłączony kwarc o częstotliwości 26.8 MHz, dzięki czemu odbierane są obrazy o maksymalnej rozdzielczości 768 punktów w linii. Układ SAA7196 akceptuje sygnały standardów PAL, SECAM i NTSC, pozwala też na analizowanie obrazów z przeplotem - przetwarzanie wszystkich lub tylko niektórych (np. parzystych lub nieparzystych) klatek obrazu. Rejestry sterujące układem pozwalają na zmianę jasności, kontrastu i stopnia nasycenia barw. Układ wysyła obraz za pomocą 32-bitowej magistrali w jednym z kilku formatów, np. w zapisie RGB o szerokości 15 lub 24 bitów, albo jako obraz monochromatyczny w zapisie 8-bitowym. Układ SAA7196 jest wyposażony dodatkowo w dwukierunkowy port wizyjny o szerokości 16 bitów, do którego można podłączyć dodatkowe źródło obrazu.

Na karcie wizyjno-dźwiękowej jest zainstalowany cyfrowy procesor sygnałowy TMS320C52, taktowany zegarem 100 MHz (patrz [6]). Jest to 16-bitowy procesor stałoprzecinkowy, który może być wykorzystywany dla efektów dźwiękowych, ale prawdopodobnie głównym jego zadaniem będzie pobieranie informacji wizyjnej od układu SAA7196 i przekazywanie jej do pamięci RAM komputera.

Ostatnim z interesujących układów karty wizyjno-dźwiękowej jest CLT 87022 firmy GE Plessey Semiconductors. Jest to układ ASIC zrobiony na zamówienie firmy Intel. Współpracuje on z magistralą PCI i łączy wszystkie elementy karty w całość. Prawdopodobnie obsługuje cyfrową magistralę szeregową dla układu AD1847 oraz magistralę I²C dla układu SAA7196.

Oprogramowanie IBVC jest elementem integrującym dwie karty w jeden system. Zapewnia jednolity sposób nawiązywania połączeń zarówno przez sieć lokalną, jak i przez ISDN lub tradycyjną sieć telefoniczną.

System IBVC nie zawiera karty sieciowej Ethernet, gdyż dla wideokonferencji z użyciem ISDN karta sieciowa nie jest potrzebna, a użytkownik chce ją realizować w sieci lokalnej, to taki komputer już zwykle jest wyposażony w kartę sieciową. W przypadku eksploatacji sieci lokalnej podczas prowadzenia wideokonferencji warto zwrócić uwagę na konstrukcję i możliwości zastosowanej karty sieciowej. Praktycznie każda dostępna karta sieciowa zapewnia odpowiednią szybkość przesyłania informacji multimedialnej, ale warto spojrzeć na sposób wymiany informacji między kartą a komputerem. Wymianę tę w prostych kartach, zgodnych z Novell NE-2000, realizuje się za pomocą sekwencyjnego odczytywania, bajt po bajcie (w nowszych konstrukcjach można czytać po 2 lub 4 bajty), całej treści ramki. Ten sposób angażuje główny procesor, co zmniejsza jego moc obliczeniową, potrzebną do kompresji danych o dźwięku i obrazie. Innym problemem może być zbyt mała wielkość buforów odbiorczych karty, co powoduje zgubienie ramek przy dużym obciążeniu sieci, jakie może generować wideokonferencja, co obniża jakość odbioru informacji przez człowieka, gdyż dochodzi do opóźnień w przesyłaniu poszczególnych ramek.

Przykładem karty odpowiedniej dla wideokonferencji jest 3c90x firmy 3Com (patrz [2]). Wprowadzono w niej szereg rozwiązań, które zmniejszają stopień wykorzystania głównego procesora podczas transmisji treści ramek. Program sterownika tej karty ma za zadanie utworzyć w pamięci dwie listy buforów, z których każdy służy do przechowywania treści jednej ramki. Jedna lista zawiera ramki odebrane od karty sieciowej (składa się na początku z pustych buforów), a druga jest listą ramek, które należy przesłać do karty sieciowej (na początku ta lista jest pusta). Po przesłaniu adresów początków tych list do karty sieciowej za pomocą zapisu odpowiednich rejestrów, karta jest gotowa do pracy. W momencie odbioru ramki z sieci Ethernet karta zaczyna przysyłać jej treść do pierwszego pustego bufora na liście odbiorczej, przy czym transmisja odbywa się za pomocą mechanizmu "Bus Mastering" - karta sieciowa przejmuje sterowanie magistralą PCI i dokonuje odpowiedniego zapisu do pamięci RAM, bez komunikowania się z głównym procesorem. Karta może wygenerować odpowiednie przerwanie po odebraniu kilku lub kilkunastu ramek, co zmniejsza średni narzut czasowy, jaki wprowadza wywołanie procedury obsługi przerwania. Ta procedura nie musi już sprowadzać treści ramki do pamięci RAM, tylko wprowadza treść odebranych ramek do systemu operacyjnego. Wykorzystanie listy buforów odbiorczych pozwala także na ominięcie konsekwencji zastosowania zbyt małego (8 KB) bufora w obrębie samej karty sieciowej, gdyż wystarczy zdefiniować odpowiednio długą listę buforów odbiorczych. Z testów przy maksymalnym obciążeniu sieci Ethernet wynika, że przy 32 buforach odbiorczych ramki były gubione raz na kilka minut. Analogiczne mechanizmy

działają przy wysyłaniu ramek, które są zapisywane na liście nadawczej, a karta zgłasza przerwanie tylko po wysłaniu treści wszystkich ramek z listy lub w razie wystąpienia błędu. Dodatkowym, dosyć ciekawym sposobem przeniesienia części zadań z głównego procesora na kartę sieciową jest obliczanie (przy wysyłaniu) lub sprawdzanie (przy odbieraniu ramek) sum kontrolnych dla najczęściej stosowanych w przypadku wideokonferencji protokołów IP, TCP i UDP. Jedyną operacją realizowaną przez kod sterownika jest odczyt określonych bitów w początkowej części elementu listy odbiorczej (o znaczeniu IP_OK, TCP_OK, UDP_OK) lub ustawieniu bitów (o znaczeniu Generuj_IP, Generuj_TCP i Generuj_UDP) w tym samym miejscu elementu listy. Samo obliczanie sum kontrolnych jest prostsze niż w przypadku kodów CRC stosowanych przez sieć Ethernet i polega na obliczeniu sumy 4-bajtowych słów tworzących daną ramkę, ale warto je stosować, aby zmniejszyć obciążenie głównego procesora.

3. Rozwój oprogramowania dla wideokonferencji

Aktualnie w Instytucie Informatyki są prowadzone prace nad stworzeniem stanowiska do prowadzenia wideokonferencji pod systemem Linux. Opisane poniżej doświadczenia są rezultatem tych prac.

Na świecie trwają prace nad stworzeniem darmowej, dostępnej w postaci kodu źródłowego, implementacji protokołu H.323 (patrz [1]). Taka wersja pozwoli na jej wykorzystanie na komputerach, gdzie nie można uruchomić darmowego produktu firmy Microsoft.

System Linux zdefiniował sposób wprowadzania obrazu w postaci specyfikacji Video4Linux. W celu zapoznania się z tym standardem była testowana specjalizowana karta Intel Smart Video Recorder III (wyposażona w układ Bt848A firmy BrookTree) firmy Intel, wprowadzająca obraz z kamery lub magnetowidu do komputera. Może ona przekazywać obraz bezpośrednio z wejścia wizyjnego do pamięci karty graficznej. Jest to możliwe poprzez wykorzystanie mechanizmu "Bus Mastering" na magistrali PCI - karta jest w stanie z własnej inicjatywy (bez ingerencji głównego procesora) dokonać operacji odczytu lub zapisu do pamięci RAM. Ten sam mechanizm jest stosowany w kartach sieciowych 3c90x firmy 3Com (patrz [2]).

Podczas rozwoju i testowania aplikacji z wykorzystaniem wideokonferencji istnieje możliwość diagnostyki na każdym poziomie, zarówno końcowej aplikacji, jak i pojedynczych sterowników wykorzystywanych urządzeń. Mechanizm dynamicznie dołączanych i usuwanych sterowników w jądrze 2.2.x (patrz [8]) pozwala na szybkie (nie wymaga ponownego

uruchomienia komputera) i mało kłopotliwe (wymianę kodu sterownika można zrealizować za pomocą prostego skryptu) dokonywanie poprawek i sprawdzanie ich skuteczności.

Linux (a w ogólności UNIX) ma budowę modułową, co pozwala na uruchomienie tylko tych procesów, które są w danej chwili potrzebne.

W czasie testowania kodu nowego sterownika zdarza się, że system operacyjny przestaje pracować. Jeśli w takim momencie byłyby prowadzone operacje zapisu na dysku, system plików mógłby zostać uszkodzony. Aby tego uniknąć, został opracowany sposób bezpiecznego testowania zawodnych aplikacji. Zakłada on wykorzystanie dwóch komputerów, połączonych siecią lokalną. Oznaczmy te komputery literami L i T. Na pierwszym z nich została zainstalowana pełna wersja systemu Linux, a na drugim zostanie uruchomiony Linux nie korzystający z twardego dysku komputera. Dysk twardy komputera T posiada partycję FAT i 64 MB pamięci RAM. Na partycji FAT znajdują się pliki potrzebne do uruchomienia systemu Linux: komenda "loadlin.exe", jądro systemu Linux i skompresowana zawartość systemu plików. Za pomocą zlecenia "loadlin" jest wczytywane jądro, jest tworzony system plików w pamięci RAM i Linux zaczyna działać. Podstawowe pliki zajmują 10 MB. W ten sposób Linux działa całkowicie w pamięci RAM; po wstępnym wczytaniu potrzebnych danych nie następują już żadne odwołania do twardego dysku komputera T. Wstępnie przygotowany system plików zawiera jedynie wstępne informacje, dlatego istnieje konieczność modyfikacji i dodania nowych plików, które będą dobrane stosownie do aktualnie testowanej aplikacji. Istotną rolę odgrywa tutaj demon sshd (secure shell daemon). Przy jego wykorzystaniu istnieje możliwość kopiowania plików między komputerami L i T, a także wykonywania komend, które są wydawane na komputerze L, a ich realizacja odbywa się na komputerze T. Taki mechanizm pozwala na przygotowanie i uruchomienie skryptów adaptujących system ogólnego zastosowania (wstępnie istniejący na komputerze T) w specjalizowane środowisko do testowania konkretnej aplikacji. Wnoszenie poprawek i kompilacja kodu sterownika odbywają się na komputerze L, a jego testowanie na komputerze T. W ten sposób unika się ryzyka uszkodzenia systemu plików podczas uruchamiania oprogramowania.

Skrypt dostosowujący komputer T do testowania sterownika karty wprowadzającej obraz z kamery kopiuje aktualne wersje modułów jądra, tworzy dodatkowy system plików o pojemności 10 MB, do którego są następnie kopiowane aplikacje pozwalające na uruchomienie okrojonego środowiska XWindow (serwer X dla karty ATI Mach64, plik konfiguracyjny XF86Config, prosty zarządca okienek mwm, aplikacja konsoli tekstowej xterm oraz testowany program - w tym przypadku xtvscreen). Tworzenie dodatkowego systemu plików w pamięci RAM jest konieczne, gdyż Linux nie obsługuje systemu plików o zmiennej wielkości, a dodatkowe pliki zajmują więcej miejsca, niż pozostało na głównym systemie plików. Innym rozwiązaniem byłoby utworzenie większego głównego systemu

plików (np. 20 MB), ale taka zmiana zmniejszyłaby liczbę komputerów, na których można uruchomić w ten sposób system Linux, gdyż nie wszystkie dysponują pamięcią operacyjną o wielkości 64 MB. Dodatkowo kopiowane są dynamicznie dołączane biblioteki, konieczne dla uruchomienia aplikacji systemu XWindow, a następnie aktualizowany jest ich spis za pomocą polecenia "ldconfig". Użytkownik następnie uruchamia system XWindow na komputerze T, tworzy za pomocą skryptu "MAKEDEV" potrzebne urządzenia, a za pomocą skryptu "update" usuwa poprzednie wersje modułów jądra (jeśli były załadowane) i dołącza aktualne wersje modułów.

Cykl testowania poprawek w kodzie sterownika ogranicza się do wniesienia poprawek do kodu źródłowego sterownika na komputerze L, skompilowania za pomocą polecenia "make", przesłania kodu wykonywalnego do komputera T za pomocą zlecenia "scp", a następnie do aktualizacji kodu jądra za pomocą skryptu "update".

Polecenie "scp" (secure copy) służy do bezpiecznego (w przeciwieństwie do zlecenia "ftp", które przesyła jawnie hasło użytkownika) przesyłania plików między dwoma komputerami. Zwykle wymaga ono podania hasła użytkownika na odległym komputerze, ale można zautomatyzować tę operację przez nadanie użytkownikowi pustego hasła (to rozwiązanie jest mało bezpieczne) lub przez dopisanie odpowiednich informacji w pliku "authorized_keys" na zdalnym komputerze, dzięki czemu użytkownik będzie miał dostęp do niego bez konieczności podawania hasła.

4. Eksploatacja systemu IBVC w środowisku Windows 95

W czasie pracy czasami zdarza się błąd w aplikacji, który po wyświetleniu komunikatu na niebieskim tle uniemożliwia dalszą pracę w systemie.

Dostępny jest dodatkowo pakiet Program Developers Kit dla IBVC, dzięki czemu można łatwo integrować usługę wideokonferencji z własnymi aplikacjami w języku Microsoft Visual C i Visual Basic, a nawet z przeglądarkami stron WWW.

Jednym ze standardów opisujących przesyłanie dźwięku i obrazu przez sieć komputerową jest H.323. Protokół ten jest wykorzystywany przez Microsoft NetMeeting i przez to jest dostępny w IBVC. Pozwala on na współpracę z wieloma innymi systemami, np. Trinicom firmy Sony i VCS3000 firmy 3M.

W wideokonferencji można dostrzec ciekawy aspekt - nie korzysta ona w czasie właściwej pracy z pamięci masowych, a jej jakość zależy przede wszystkim od parametrów kanału transmisji oraz urządzeń wizyjnych.

W czasie połączenia z Los Angeles we wczesnych godzinach rannych (po godzinie piątej) obraz był wyraźny tylko co pewien czas, po wykonaniu jakiegoś ruchu ulegał zamazaniu. Można wyciągnąć wniosek, że zmiany w kolejnych klatkach są przekazywane, ale nie są prawidłowo odwzorowywane.

Wchodzący w skład IBVC program NetMeeting firmy Microsoft pozwala na współdzielenie aplikacji. Pozwala to na wspólną pracę w dowolnej aplikacji, która działa pod kontrolą systemu operacyjnego Windows 95. Praca jest naprzemienna, a przełączanie odbywa się przez naciśnięcie dowolnego klawisza klawiatury lub lewego klawisza myszy.

Przesyłanie dźwięku odbywa się z wykorzystaniem algorytmu, który po wykryciu ciszy zawiesza transmisję informacji o dźwięku. Przy odległych połączeniach dźwięk dochodził w postaci oderwanych fragmentów, które były trudne do zrozumienia. Użytkownik powinien mieć możliwość ustawienia trybu automatycznej adaptacji parametrów przesyłanego dźwięku do zmian warunków panujących w sieci. Inne ustawienie powinno umożliwić transmisję metodą przesyłania plików dźwiękowych, aby była pewność, że informacja dotarła w sposób zrozumiały. Użytkownik powinien mieć możliwość wyboru algorytmu kodującego dźwięk, aby uzyskać jak najlepszą jego jakość.

5. Wnioski

W chwili obecnej są dostępne gotowe systemy wideokonferencyjne, które można stosować bez wnikania w szczegóły ich działania. Można także się włączyć w rozwój aplikacji tego typu, co pozwoli na lepsze zrozumienie sposobu ich działania i implementację tych cech, które są potrzebne w niekonwencjonalnych zastosowaniach.

Z opisu powyższych elementów składowych systemów wideokonferencyjnych można wysunąć wniosek, że na rynku występują elementy, z których można złożyć (po dodaniu jednego specjalizowanego układu ASIC i oprogramowania) system wideokonferencyjny. Znając dokładnie elementy składowe takiego systemu, można oszacować maksymalne parametry jakościowe i ilościowe wideokonferencji prowadzonej w takim systemie.

LITERATURA

1. Darmowa implementacja protokołu H.323, "<http://www.h323.org>".
2. Linux i karta 3c90x, "<http://cesdis.gsfc.nasa.gov/linux/drivers/vortex.html>".
3. Filipiak J. (red.): Sieci dostępne dla usług szerokopasmowych, tom I, Wydawnictwo Fundacji Postępu Telekomunikacji, Kraków 1997.

4. Układ AD1847, "<http://products.analog.com/products/info.asp?product=AD1847>".
5. Układ SAA 7196, "<http://www-us.semiconductors.philips.com/pip/SAA7196U>".
6. Układ TMS 320C52, "<http://www.ti.com/sc/docs/products/dsp/tms320c52-100.html>".
7. Systemy wideokonferencyjne firmy Intel, "<http://www.intel.com/proshare/conferencing/>".
8. Jądro systemu Linux, "<http://www.linuxhq.com>".

Recenzent: Dr inż. Ryszard Winiarczyk

Wpłynęło do Redakcji 12 kwietnia 1999 r.

Abstract

The computer networks become recently more and more popular, new network applications, such as videoconferencing, are introduced. This article describes internals of a commercial videoconferencing system. The functions performed by audio and video chips are presented. The essential requirements for network interface card for videoconferencing system are analyzed. The process of videoconferencing application development in Linux operating system is described. The features of bundled software are discussed.