STUDIA INFORMATICA 2002
Volume 23 Number 3 (50)

Andrzej KWIECIEŃ, Rafał CUPEK, Marcin FOJCIK
Politechnika Śląska, Instytut Informatyki

WYBÓR METODY INTEGRACJI SIECI PRZEMYSŁOWYCH NAJNIŻSZEGO POZIOMU Z SIECIAMI OPARTYMI NA RODZINIE PROTOKÓŁÓW TCP/IP

Streszczenie. W artykule na tle podziału sieci przemysłowych najniższego poziomu opisana jest metoda ich integracji z zastosowaniem protokołów TCP/IP. Zwrócono również uwagę na środowisko, jakim jest system czasu rzeczywistego i jego wymagania czasowe. Opisano dokonaną próbę integracji sieci najniższego poziomu z użyciem protokołu TCP/IP i przedstawiono ocenę takiego połączenia pod kątem gwarantowanego czasu przesłania informacji.

THE CHOICE OF INTEGRATION METHODS BETWEEN THE LOWEST LEVEL INDUSTRIAL NETWORK AND NETWORKS BASED ON TCP PROTOCOLS

Summary. This article describe division of the lowest level industrial networks and one of the integration method with using TCP/IP protocols. Also is described the method of integration this network according of kind of used real time system and it's time parameters. The authors suggest using TCP protocol to integration of industrial networks and try to check possibility of fulfill guaranteed time requirement in this type of integration.

1. Wstęp

Wzrost popularności protokołu TCP/IP związany jest z rozbudową globalnej sieci Internet. Zdobywanie przez protokół TCP/IP nowych obszarów zastosowań powodowane jest również powszechną dostępnością tego standardu Nie bez znaczenia są także czynniki ekonomiczne, do których należy zaliczyć brak opłat licencyjnych związanych z jego wykorzystywaniem oraz powszechna dostępność rozwiązań sprzętowo – programowych.

Znaczący wzrost zainteresowania możliwościami protokołu TCP/IP nie ominął także dziedziny informatyki przemysłowej. Obserwuje się tu coraz częstsze próby jego wykorzystywania, co nie oznacza, że zawsze i wszędzie będzie on powszechnie zalecany i stosowany. Wymienione na wstępie czynniki, powodujące stały wzrost zainteresowania protokołami TCP/IP, nie umożliwiają jednak ich bezkrytycznego wykorzystania w każdej instalacji automatyki przemysłowej. Dotyczy to przede wszystkim tych zastosowań, w których determinizm czasowy i silne uwarunkowania czasowe są absolutnie niezbędne i wymagają bezwzględnego respektowania.

Z drugiej zaś strony warto jednak zastanowić się nad pewnymi grupami zastosowań, w których w pełni da się wykorzystać zalety protokołów TCP/IP, bez rezygnacji z walorów systemów czasu rzeczywistego. Jedną z takich grup zastosowań wydaje się być integracja sieci najniższego poziomu, jak również integracja tych systemów z sieciami rozległymi i siecią Internet.

2. Sieci najniższego poziomu

Systemy komunikacyjne spotykane w przemyśle można podzielić generalnie na dwie kategorie:

- systemy komunikacyjne o zdeterminowanym w czasie dostępie do medium,
- systemy komunikacyjne o swobodnym, nie gwarantowanym w czasie dostępie do medium.

Pierwsza z wymienionych kategorii dotyczy zazwyczaj systemów czasu rzeczywistego, druga zaś to albo systemy czasu rzeczywistego o słabych wymaganiach czasowych, gdzie czas jako parametr globalny nie jest czynnikiem krytycznym, albo nie są to po prostu systemy czasu rzeczywistego.

Ponadto istnieje cały szereg innych podziałów czy też typów systemów, w zależności od przyjętych kryteriów. Mogą nimi być szybkość wymiany informacji (w tym sprawność i przepustowość użyteczna), przeznaczenie, protokół dostępu do łącza, topologia, rozległość sieci, poziom pracy sieci. Ze względu na poziom pracy sieci możemy podzielić na: sieci rozległe, lokalne i sieci przemysłowe. W sieciach rozległych i lokalnych zachodzi wymiana danych na poziomie aplikacji i zwykle nie dochowuje się wymagań dotyczących żądanego czasu wymiany informacji. W sieciach przemysłowych, które ściśle są związane z obiektem przemysłowym, przesyłane są informacje dotyczące stanu urządzeń, sygnały wymagane do sterowania, parametryzacji, a także alarmy. Sieci, w których są przesyłane sygnały dotyczące stanu obiektu i niezbędne dla jego prawidłowej pracy, a pochodzące od abonentów bezpośrednio z nim sprzęgniętych (na przykład pochodzące ze sterowników swobodnie

programowalnych, będących abonentami sieci i urządzeniami sterującymi jednocześnie), będziemy nazywać **sieciami najniższego poziomu.** W sieciach tych podstawowym parametrem branym pod uwagę jest czas przesyłania informacji o zdarzeniach zachodzących na obiekcie. Czas ten jest wypadkową wielu czynników typowych dla wymian „obiekt-abonent" (patrz: sterownik obiektowy), „abonent-sieć", co ma wpływ na sprawność komunikacji i przepustowość użyteczną. Szybkość wymiany informacji powinna być jak największa w odniesieniu do wymagań technologicznych obiektu, a dodatkowo powinno się zagwarantować jej minimalną wartość. Gwarantowany, zdeterminowany w czasie, dostęp do medium, w połączeniu z gwarantowanym czasem reakcji pojedynczego elementu systemu automatyki przemysłowej, daje pewność prawidłowej pracy całego systemu.

2.1. Protokoły sieci przemysłowych

Podstawową cechą sieci przemysłowych jest zdeterminowany czas dostępu do danych. Właściwość ta umożliwia określenie maksymalnego czasu, po jakim wytworzona informacja dotrze do adresata.

Parametry czasowe sieci zależą od dwóch czynników: parametrów elektrycznych łącza i zastosowanego protokołu komunikacyjnego. Cechy fizyczne medium transmisyjnego, jego rozległość, parametry elektryczne oraz podatność na zakłócenia decydują o fizycznej prędkości transmisji informacji możliwej do uzyskania za pomocą określonego medium komunikacyjnego. Kolejnym czynnikiem wpływającym na parametry czasowe jest sposób realizacji algorytmów obsługi protokołów sieci przemysłowej, a także dostępu do zasobów sprzętowych abonenta sieci. Znając parametry czasowe warstwy fizycznej i realizowany protokół, można dokonać pełnej analizy czasowej transmisji danych dla konkretnej sieci [3, 9].

Klasyfikując sieci przemysłowe ze względu na stosowane protokoły komunikacyjne, zawierające gwarancje czasowe, można wyodrębnić trzy główne grupy: sieci z protokołem „Master-Slave", sieci z protokołem „krążący żeton" (*ang „Token-Bus"*) oraz Producent-Dystrybutor - Konsument (PDK) (*ang. PDC - "Producer – Distributor – Consumer"*).

2.1.1. Sieci MASTER - SLAVE

Architektura systemów Master-Slave charakteryzuje się wyróżnieniem jednego z abonentów i przypisaniem mu właściwości urządzenia nadrzędnego, które nosi miano „master". Pozostali abonenci sieci pełnią rolę urządzeń podrzędnych, noszących nazwę „slave". O czasie i rodzaju przeprowadzanych wymian decyduje wyłącznie abonent nadrzędny, zaś rola abonentów podrzędnych ogranicza się jedynie do udzielania odpowiedzi na transmisje inicjowane przez stację „master". Sieć oparta na architekturze Master-Slave nadaje się szczególnie dla systemów, w których dane produkowane przez urządzenia

peryferyjne przesyłane są do centrum. Na podstawie zebranych informacji centrum podejmuje decyzje i rozsyła polecenia sterujące do poszczególnych elementów wykonawczych.

Ze względu na rozpowszechnienie tego standardu, a także niewielkie wymagania sprzętowe, rozwiązania oparte na protokole Master-Slave w prosty sposób można wykorzystać do łączenia urządzeń pochodzących od różnych producentów. Dobrym przykładem jest tu protokół komunikacyjny MODBUS, który z większymi lub mniejszymi zmianami implementacyjnymi doczekał się adaptacji przez większość producentów sterowników swobodnie programowalnych, regulatorów prędkości obrotowej czy też zdalnych modułów pomiarowych.

Sieć typu „Master-Slave" nadaje się dobrze do wizualizacji systemów pomiarowych i telemetrycznych, o niewygórowanych wymaganiach czasowych. Stacja kontrolno-nadzorcza pełni zazwyczaj rolę stacji „master" sieci, decydując o tym, jakie informacje są w danym momencie pobierane lub wysyłane do poszczególnych abonentów. Ze względu na słabe parametry czasowe opisywanego protokołu oraz konieczność każdorazowego udziału stacji nadrzędnej w transmisji, wykorzystanie sieci typu „Master-Slave" ogranicza się do systemów o średnich wymaganiach czasowych.

2.1.2. Sieci TOKEN BUS

W zastosowaniach przemysłowych sieci lokalne typu „krążący żeton" są dedykowane do komunikacji pomiędzy wieloma abonentami, dając gwarancję dużej niezawodności transmisji. Sieci tego typy charakteryzują się zdecentralizowaną architekturą, łatwą rozbudową w systemy rozległe, gwarantowanym czasem wymiany informacji, a wreszcie możliwością pracy sieci nawet w przypadku awarii dowolnego abonenta.

W sieciach przemysłowych wykorzystujących protokół „krążący żeton" nie wyróżnia się w sposób szczególny żadnego z abonentów. Funkcje związane z zarządzaniem pracą sieci realizowane są kolejno przez wszystkich jej abonentów. Transmitowany cyklicznie żeton jest uprawnieniem dla abonenta, który z chwilą jego przejęcia jest uprawniony do nadawania. Po zrealizowaniu własnych żądań, dotyczących transmisji, abonent przekazuje żeton następcy, który według ustalonego protokołu zaczyna pełnić funkcję urządzenia nadrzędnego. Realizację postulatu zdeterminowanego czasu dostępu do łącza gwarantuje ograniczenie maksymalnego czasu przetrzymywania żetonu przez poszczególnych abonentów sieci, co jest ściśle związane z kontrolą czasu nadawania. Indywidualnie ustalany maksymalny czas dostępu do łącza oraz stosowanie różnych priorytetów transmitowanych danych stwarza możliwości konfiguracji sieci w sposób stosowny do wymagań czasowych procesu. Protokół typu „krążący żeton" doczekał się szeregu implementacji firmowych, różniących się

pomiędzy sobą dość znacznie, co utrudnia integrację urządzeń pochodzących od różnych producentów za pomocą tego typu sieci.

2.1.3. Sieci typu Producent-Dystrybutor-Konsument

Głównym założeniem dla sieci opartych na protokole PDK jest rozproszone przetwarzanie informacji znajdującej się w rozproszonej bazie danych. W sieci tej konwencjonalne połączenia typu „punkt-punkt" pomiędzy zadajnikami sygnału i urządzeniami wykonawczymi a jednostką centralną, są zamienione na połączenia typu „multipoint". Mówiąc inaczej, każdą ramkę z informacją wysyłaną przez dowolnego abonenta, mogą jednocześnie pobierać wszyscy abonenci potrzebujący tych informacji. Znaczącą zaletą tej sieci jest ogromna łatwość w jej modyfikacji, rozbudowie i obsłudze, przy zachowaniu zalet związanych z takimi parametrami, jak czas odpowiedzi lub bezpieczeństwo poprawnej transmisji. Sieć ta jest rozwiązaniem dającym standardowy dostęp wszystkim producentom i wewnętrznym użytkownikom do rozległego systemu informatycznego. W sieci PDK wyróżnia się trzy rodzaje abonentów: „Producenta", „Dystrybutora" i „Konsumenta".

„Producent" jednostki danych jest odpowiedzialny za ich produkcję. „Dystrybutor" jest natomiast odpowiedzialny za transfer danych od „Producenta" do wszystkich „Konsumentów". „Konsumenci" danych są aplikacjami, które żądają danych. Model PDK stawia, w odróżnieniu od modelu „klient-serwer", na uprzywilejowanej pozycji „Konsumentów". Poza tym funkcja dystrybucji, statycznie skonfigurowana i zaprojektowana, może być modyfikowana, spełniając czasowe wymagania „konsumentów".

2.2. Klasy sieci przemysłowych

Ze względu na wykorzystywany protokół sieci przemysłowe najniższego poziomu, możemy podzielić na 3 klasy:

- Sieć klasy I – MASTER-SLAVE

Rys. 1. Sieć klasy I MASTER-SLAVE
Fig. 1. Network Class I MASTER-SLAVE

- Sieć klasy II – TOKEN BUS

Rys. 2. Sieć klasy II TOKEN-BUS
Fig. 2. Network Class II TOKEN-BUS

- Sieć klasy III – PDK

Rys. 3. Sieć klasy III PDK
Fig. 3. Network Class III PDK

Każda klasa zawiera jednorodną sieć przemysłową o zagwarantowanym czasie wymiany informacji. Typowy system informatyczny, zastosowany na obiekcie przemysłowym, może składać się z szeregu jednostek jednorodnych sieci najniższego poziomu, będących reprezentantami różnych klas. Wielokrotnie zdarza się tak, że istnieje konieczność ich integracji albo w rezultacie wymaganej całkowitej integracji systemu, albo w efekcie zaistniałej konieczności rozbudowy systemu już istniejącego. Rodzi się zatem pytanie o zasadność integracji owych systemów, przez powiązanie ich sieciami opartymi na protokóle TCP/IP, w celu konstruowania systemów o większych możliwościach i lepszej kontroli nad sterowanym obiektem.

3. Integracja sieci najniższego poziomu

Integracja sieci polega na połączeniu dwóch lub więcej sieci przemysłowych, należących do tych samych lub różnych klas, w sposób sprzętowy lub programowy. Dzięki temu powstaje możliwość gromadzenia danych z różnych sieci celem uzyskania informacji do prawidłowego sterowania i monitorowania procesów przemysłowych. Konieczność integracji

sieci przemysłowych może być spowodowana, na przykład wymaganiem przesyłania informacji pomiędzy segmentami sieci, wydawaniem poleceń do abonenta innego segmentu, gromadzeniem danych na potrzeby wizualizacji lub zapamiętywaniem zmian stanu obiektu w czasie. Integracja wymaga zachowania specyficznych cech łączonych sieci, niezbędnych do prawidłowego sterowania. Szczególne znaczenie mają wymagania systemu komunikacyjnego, a zwłaszcza wielokrotnie już podkreślana cecha, dotycząca determinizmu czasowego.

Integracja sieci przemysłowych może odbywać się na trzech poziomach:

- na poziomie sieci przemysłowych,
- na poziomie stacji nadrzędnych,
- na poziomie sieci lokalnej i rozległej.

Na najniższym poziomie mamy do czynienia z wymaganiami dotyczącymi bieżącej pracy obiektu, co związane jest z zapewnieniem gwarantowanych parametrów czasowych.
Na poziomie stacji nadrzędnych wymiana informacji może wymagać determinizmu czasowego lub nie, a na poziomie sieci rozległej przesył informacji z zachowaniem determinizmu czasowego w ogóle nie będzie możliwy.

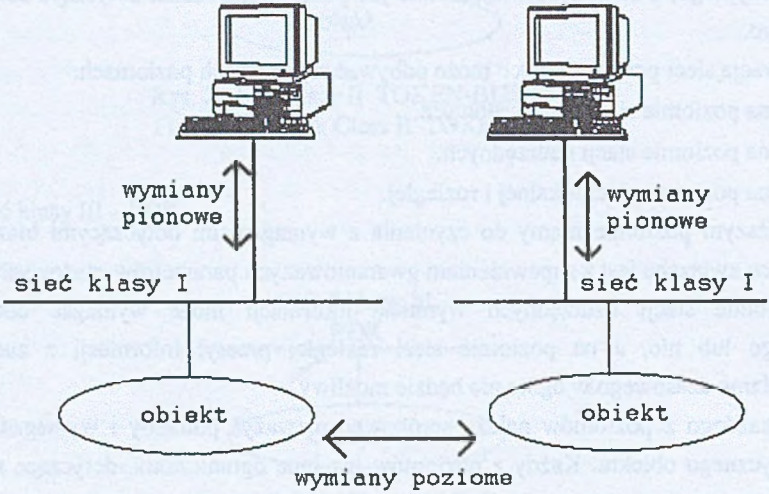
Dla każdego z poziomów należy spróbować rozważyć potrzeby i wymagania systemu informatycznego obiektu. Każdy z poziomów ma inne ograniczenia, dotyczące możliwości jego zastosowania – sterowanie i wizualizacja, lub tylko wizualizacja i monitorowanie. Problem zakresu stosowalności integracji sieci jest związany z gwarantowanym czasem wymiany informacji, tak istotnym z punktu widzenia bezpieczeństwa sterowania procesami. Integracja sieci prowadzi prawie zawsze do pogorszenia parametrów dotyczących czasu wymiany informacji w całym systemie. Ale ważne jest również zachowanie możliwości zmiany parametrów systemu, jego łatwa rozbudowa i zapewnienie otwartości, a także niezawodności, oraz redundancja i mechanizmy kontroli komunikacji [10].

3.1. Cele integracji – rodzaje wymian

Patrząc na funkcje realizowane przez sieci przemysłowe najniższego poziomu, przez pryzmat hierarchicznej struktury systemu automatyki, można wyróżnić dwa rodzaje wymian realizowanych w tych sieciach. Nazwijmy je wymianami „poziomymi" i „pionowymi". Podział ten związany jest z funkcjami wymian i wymaganiami co do gwarancji czasowych ich realizacji.

Wymiany „poziome" dotyczą najniższego poziomu sieci i polegają na monitorowaniu, sterowaniu i kontroli działania urządzeń. Są to wymiany krytyczne dla pracy systemu z punktu widzenia czasu ich realizacji. Musi być dla nich zagwarantowany minimalny czas wymiany informacji.

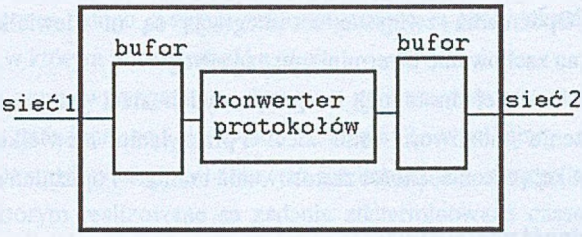
Drugim typem wymian są wymiany „pionowe". Polegają one na monitorowaniu, zbieraniu informacji, gromadzeniu raportów i tworzeniu archiwów oraz baz danych na temat pracy systemu. Są to zazwyczaj dane nie mające bezpośredniego wpływu na jego poprawną pracę. Informacje tu zgromadzone mogą być udostępniane w sieciach lokalnych. Wymienione tutaj wymiany nie muszą spełniać wymogów determinizmu czasowego.



Rys. 4. Wymiany pionowe i poziome w sieci klasy I
Fig. 4. The „vertical" and "horizontal" exchanges in Class I networks

3.2. Urządzenia pośredniczące

Często, aby umożliwić połączenie dwóch lub więcej sieci przemysłowych, niezbędne jest dodatkowe urządzenie pełniące funkcję układu pośredniczącego (*ang. „Interconnection Device"*). Urządzenie takie stosuje się najczęściej wtedy, kiedy integrowane sieci mają różne parametry komunikacyjne, np. protokoły, prędkości transmisji i kiedy bezpośrednie fizyczne połączenie, na przykład magistral, nie przyniesie żadnego efektu.



Rys. 5. Urządzenie pośredniczące
Fig. 5. Interconnection Device

Patrząc na schemat budowy urządzenia pośredniczącego widać, że ważnym czynnikiem jest czas obsługi programowej konwersji informacji z jednej obsługiwanej sieci do drugiej. Czas obsługi jest tym większy, im więcej różnych typów sieci przemysłowych umie ona obsługiwać, im pełniejszy (dokładniejszy) jest proces konwersji protokołów. Można powiedzieć, że im większe różnice w łączonych sieciach, tym czas obsługi ich łączenia jest większy. W granicznych przypadkach może dojść do niemożności dotrzymania gwarancji czasowych.

3.3. Podstawowe problemy integracji poziomej

W niniejszym opracowaniu główny nacisk został położony na problematykę związaną z wymianami poziomymi w zintegrowanych sieciach przemysłowych i na możliwości wykorzystania do celów integracji protokołów TCP/IP.

Decydując się na jakiś sposób integracji, musimy pamiętać, że jeżeli integracja na najniższym poziomie miałaby spowodować pogorszenie się parametrów systemu komunikacyjnego i zakłócić wymagany determinizm, to musimy z niej zrezygnować i stworzyć nowy niezależny segment sieci przemysłowej. Jeżeli wymagania czasowe nie są krytyczne dla pracy sieci lub istnieje przewidziana rezerwa czasowa, należy zastosować drugie kryterium, którym są koszty realizacji.

3.3.1. Jednorodność sieci

Integracja homogenicznej sieci stanowi przypadek najprostszy. Mamy tu dwie lub więcej sieci tej samej klasy. Integracja będzie się więc odbywała przy zachowaniu tego samego, jednorodnego protokołu. Niektóre protokoły przemysłowe oferują mechanizmy do prostego łączenia sieci w segmenty dzięki odpowiedniemu adresowaniu zmiennych. W takim przypadku wystarczy fizyczne połączenie obu sieci i odpowiednia adresacja w postaci na przykład <adres zmiennej : adres sieci>. Mechanizm ten jest dostępny między innymi w sieciach PDK. W przypadku sieci nie obsługujących adresowania segmentu należy zastosować urządzenie rozpoznające i zamieniające adresy jednej z podsieci na drugą. Często

w takich przypadkach do integracji wystarczy repeater, gdy na przykład potrzebne jest tylko wydłużenie sieci. Opóźnienia związane z integracją są tu niewielkie i wpływają w niewielkim stopniu na zachowanie determinizmu czasowego.

W przypadku sieci niejednorodnej integracja takich sieci jest sensowna tylko przy znacznym ograniczeniu możliwości obu sieci i przesyłaniu niewielkiej ilości danych. Ograniczenia te wynikają z konieczności zastosowania bramy – i opóźnień czasowych [7].

3.3.2. Jednorodność parametrów

Dla przypadku łączenia sieci o różnych czasach wymiany informacji (a może to mieć miejsce nawet dla sieci o tych samych protokołach), integracja musi zostać przeprowadzona przez dodanie urządzenia pośredniczącego. Różnice w czasach pracy obu sieci powodują, że integracja dla wymian poziomych w połączonych sieciach będzie powodować ujednolicenie czasu pracy wszystkich sieci do poziomu najwolniejszej. W niektórych przypadkach sytuacja taka może być wystarczająca, w innych musimy zastosować integracje dla wymian pionowych.

Oznacza to, że gdy na przykład dla sieci typu „Master-Slave" (z implementacją firmową MODBUS) czas cyklu wynosi około 500 ms., a dla sieci typu „Token-Bus" (z implementacją firmową N10) około 50 ms, to łączenie tych sieci wymianami poziomymi nie ma sensu, gdyż spowoduje kilkakrotne opóźnienie szybszej sieci, a tym samym najprawdopodobniej uniemożliwi prawidłową pracę systemu. Wykorzystanie natomiast wymian poziomych spowoduje stworzenie hierarchii sieci. Sieć o mniejszej prędkości będzie zbierała informacje i przekazywała je do sieci szybszej. Wymiany w drugim kierunku nie będą możliwe.

4. Systemy czasu rzeczywistego

Analiza rozwiązań sieci najniższego poziomu powinna obejmować zagadnienia przepustowości i analizy czasowej prowadzonej ze względu na kryterium determinizmu czasowego, a także wymogów technologicznych stawianych przez konkretny obiekt. W różnych obiektach mogą występować odmienne wymagania zarówno co do ilości, jak i jakości usług sieciowych. Jednak zawsze najważniejszym parametrem będzie czas. Wszystkie systemy czasu rzeczywistego możemy podzielić na 2 grupy:

- systemy, w których czas wymiany informacji jest parametrem krytycznym,
- systemy, w których dozwolona jest pewna tolerancja dla czasu wymiany.

Oba systemy muszą jednak gwarantować nieprzekraczalny czas reakcji T_G, który jest globalnym parametrem całego systemu informatycznego, a który określa nieprzekraczalny interwał czasowy, w którym musi nastąpić reakcja.

Dla pierwszej grupy wymagania, związane z gwarantowanym czasem realizacji usług komunikacyjnych, stawiane są w sposób „sztywny", wynikający tak z właściwości technologicznych kontrolowanego procesu, jak i z konstrukcji samego systemu automatyki. Taki system, w którym realizowane są zadania zdeterminowane czasowo, nazywany jest systemem typu *„hard real-time system"*. Znaczna część zadań tego typu systemów, związana jest z poziomymi wymianami informacji, jak np.: akwizycja informacji o stanie procesu, synchronizacja pracy kilku współdziałających ze sobą urządzeń. W przypadku wymian pionowych wymagania mogą być związane np.: z rejestracją zdarzeń bądź też przekazywaniem zdeterminowanej czasowo informacji w złożonych hierarchicznych systemach sterowania.

Do drugiej grupy należy zaliczyć usługi komunikacyjne nie obwarowane tak ściśle przez wymogi determinizmu czasowego. Szereg funkcji systemu związanych np. z realizacją niektórych działań związanych z wizualizacją, czy udostępnianiem danych do wyższych warstw infrastruktury informatycznej, może być z powodzeniem realizowana za pomocą systemów, choć zdeterminowanych czasowo, to jednak z pewną tolerancją wartości bezwzględnej parametru, jakim znów jest czas. Systemy te noszą miano *„soft real-time systems"*, w których czas nie jest parametrem krytycznym. Zakres usług komunikacyjnych tego typu koncentruje się wokół wymian pionowych. Przykładem może być system bankomatów, w których czas graniczny T_G nie może być przekroczony, ale który liczy się w dziesiątkach sekund, a nie w mikro czy milisekundach i na dodatek może być przekroczony o definiowalną wartość T_{TOL}.

W systemach czasu rzeczywistego mogą też występować pewne typy sieci, dla których nie występuje w ogóle gwarancja czasowa. Może to być prezentacja podstawowych parametrów produkcji na stronach WWW. W takim przypadku można nawet wyobrazić sobie sytuację, w której osoba zainteresowana, mająca dostęp do informacji za pomocą sieci rozległej, nie otrzyma w ogóle potrzebnej informacji. (np. po opublikowaniu w Internecie zdjęć z Marsa, strona ta została „zatkana" przez tysiące chętnych do ich zobaczenia).

5. Protokół TCP/IP w sieciach przemysłowych

Głównym celem integracji jest uzyskanie możliwości gromadzenia danych z różnych sieci celem uzyskania informacji do prawidłowego sterowania i monitorowania procesów przemysłowych przy zachowaniu dotychczasowych parametrów sieci.

Zastosowanie protokołu TCP/IP w sieciach przemysłowych najniższego poziomu sprawiło, iż jednym z głównych wyzwań stawianych przed systemami czasu rzeczywistego stała się potrzeba wypracowania mechanizmów pozwalających na połączenie różnych typów usług. Istotnym zagadnieniem jest zapewnienie determinizmu czasowego. Protokół TCP, zastosowany w sieciach lokalnych, opiera się na nie dającym żadnych gwarancji co do czasu realizacji usługi, rywalizacyjnym algorytmie CSMA/CD, będącym podstawą funkcjonowania sieci Ethernet. Taka realizacja warstwy aplikacji protokołu TCP nie nadaje się więc do szeregu zastosowań typu *„hard real-time system"*.

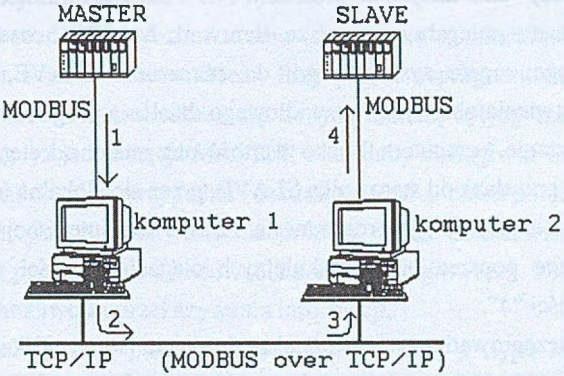
Sieci przemysłowe wykorzystujące protokół TCP/IP nie mogą być jednak rozpatrywane w oderwaniu od architektury rozpatrywanego rozwiązania. Można protokół TCP/IP połączyć z protokołami przemysłowymi, np. z protokołem MASTER-SLAVE. Dominuje tu technika enkapsulacji ramek sieci przemysłowej w segmentach TCP.

Analizując powyższe rozwiązanie integracji sieci, można zastosować kryteria pozwalające na określenie w sposób deterministyczny parametrów danej sieci. Zastosowanie protokołu MODBUS opartego na modelu „Master-Slave" gwarantuje dla wydzielonego segmentu sieci brak występowania kolizji w warstwie fizycznej sieci Ethernet, a stosując mechanizmy analizy czasowej sieci przemysłowych [3], można pokusić się o pełną analizę takiego rozwiązania. Wymaganym warunkiem poprawnej pracy i gwarancji czasowych jest stworzenie wydzielonego segmentu sieci, w którym zostanie zastosowany protokół TCP, a abonenci będą pracować zgodnie z zasadami MASTER-SLAVE. Tylko wtedy będziemy mogli zagwarantować brak kolizji, tylko stacja MASTER będzie mogła rozpoczynać przesył informacji, a stacje SLAVE będą tylko odpowiadały.

Jednym z rozwiązań jest zaproponowane przez firmę Modicon rozwiązanie MODBUS/TCP [15]. Zakłada ona kapsułkowanie ramek zgodnych ze standardem MODBUS/RTU w segmencie TCP. Na poziomie warstwy aplikacji oprogramowanie odpowiedzialne za obsługę protokołu MODBUS generuje ramki zgodne ze standardem MODBUS/RTU. Po usunięciu z ramki sumy kontrolnej CRC pozostała część kapsułkowana jest w segmencie TCP. Segmenty TCP, zgodnie z warstwową budową stosu protokołów TCP/IP, kapsułkowane są na poziomie Intersieci w datagramy IP. Datagramy IP na poziomie interfejsu sieciowego kapsułkowane są w ramki sieci Ethernet. Warstwa sprzętu zapewnia dostarczenie ramki do adresata. Oprogramowanie komunikacyjne po stronie odbiorcy realizuje odwrotną procedurę. Ramka sieci Ethernet poprzez datagram IP i segment TCP doprowadzana jest do postaci zgodnej ze standardem MODBUS/RTU.

6. Stanowisko laboratoryjne

W ramach badań nad integracją sieci przygotowano stanowisko laboratoryjne. Stanowisko to miało posłużyć do badań nad metodami integracji sieci najniższego poziomu, za pomocą sieci opartych na protokóle TCP/IP, a także praktycznej weryfikacji przyjętych założeń. Stanowisko składało się z 2 sterowników wyposażonych w moduły sieci typu MASTER- SLAVE, 2 komputerów klasy IBM PC (procesor P166, 64MB RAM, WINDOWS NT) i segmentu sieci lokalnej wykorzystującej protokół Ethernet o prędkości 10Mb/s.



Rys. 6. Schemat stanowiska badawczego
Fig. 6. The laboratory stand

Komputery były połączone między sobą siecią lokalną, a każdy z nich dodatkowo był połączony ze sterownikiem swobodnie programowalnym, a dokładnie z modułem obsługującym sieć przemysłową MODBUS. Jeden ze sterowników pełnił rolę stacji MASTER, a drugi stacji SLAVE. Obieg informacji był następujący. Sterownik „MASTER" wysyłał ramkę do komputera 1, program w komputerze 1 „kapsułkował" ją do protokołu TCP i przesyłał siecią lokalną do komputera 2. Tam następował proces „enkapsulacji", a następnie wysłanie ramki siecią MODBUS do sterownika „SLAVE". Sterownik wykonywał polecenie i tą samą drogą odsyłał odpowiedź.

Dla porównania zmierzono czas wymiany informacji bezpośrednio pomiędzy sterownikami. Wynosił on około 100-200 ms. Obsługa sieci MODBUS była dokonywana programowo z uwzględnieniem informacji potrzebnych do kontroli działania programu. W warunkach przemysłowych, przy wymianie obsługiwanej sprzętowo, czas ten wynosi około 50-100 ms dla interfejsu RS232 i 20-50 ms dla interfejsu RS485.

Pierwsza wersja oprogramowania laboratoryjnego na komputery powstała podczas laboratorium z przedmiotu "komputerowe systemy rozproszone". Przeprowadzone testy wykazały, że program napisany pod kątem ogólnej idei integracji sieci nie jest optymalny.

Komunikacja odbywała się w większości przypadków prawidłowo, jednak czasami występowały przerwy w transmisjach informacji lub przesyłana informacja nie była poprawna. Aby tę sytuację poprawić, dokonano modyfikacji programu zarówno w sterownikach, jak i w komputerach. Zmiany te polegały na optymalizacji wymian (cykl pracy sieci a cykl pracy jednostki centralnej sterownika [2,8]) i synchronizacji komunikacji w sieciach lokalnej (TCP) i przemysłowej (MODBUS). Jednym z napotkanych problemów były zdecydowanie różne prędkości przesyłania informacji w obu sieciach. Opóźnienie w przekazywaniu informacji było najbardziej widoczne dla obsługi sieci MODBUS w komputerze, przekazywania danych do protokołu TCP i kapsułkowania.

Główna idea testu polegała na tym, że sterownik MASTER wysyłał wartość sekund pobraną ze swojego zegara systemowego, do sterownika SLAVE, a ten odsyłał ją z powrotem. Dla ułatwienia obserwacji prawidłowego działania programu wartość sekund była wyświetlana na ekranie komputera 1 jako wartość otrzymana od sterownika MASTER, a także jako wartość przesłana od sterownika SLAVE, przez sieć lokalną (w obie strony).

Dla nieoptymalnie napisanego programu na ekranie komputera pojawiały się przerwy w transmisji, widoczne poprzez braki w kolejnych ciągach wartości sekund – np. liczby 1,2,3,5,6 bez wartości "4".

Na podstawie przeprowadzonych badań okazało się, że istnieje kilka elementów znacząco wpływających na szybkość transmisji w takim połączeniu. Są to w kolejności od najbardziej znaczących:

- Sposób obsługi łącza szeregowego,
- buforowanie informacji przez program w komputerze,
- parametry sieci MODBUS, a konkretnie prędkość transmisji,
- sposób napisania programu w sterowniku (szybkość działania, zajętość pamięci),
- sposób obsługi wyświetlania w programie komputerowym.

Znaleziono też elementy mające znacząco mniejszy wpływ. Są to:

- Rodzaj protokołu TCP czy UDP,
- obciążenie sieci lokalnej (Ethernet),
- ilość przesyłanej informacji i różnice w prędkościach sieci MODBUS (tu: 19kb/s) a TCP (tu: 10000kb/s).

Uwagi te są słuszne tylko wtedy, gdy protokół TCP służy do integracji, gdyż wszystkie ramki powstają tylko w urządzeniach obsługujących sieć MODBUS z szybkością dostępną dla sieci MODBUS. Z racji bardzo znaczących różnic prędkości obu sieci (rzędu około 500 razy) można wysnuć wniosek, że bez względu na sposób obsługi protokołu TCP (optymalny czy nie, z obciążeniem czy bez) sieć jest w stanie przetransmitować wszystkie dane bez zauważalnych opóźnień.

Można wyliczyć rozmiary ramek i czasy ich przesyłów.

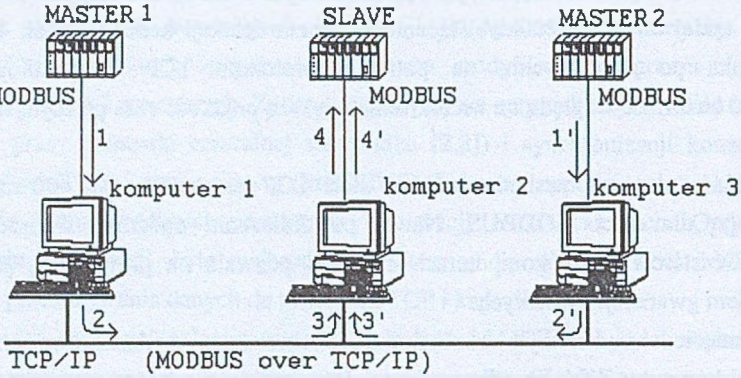
Dla sieci MODBUS ramka zawierając rozkaz odczytu 1 słowa ma długość 81 bitów. Czas przesłania takiej ramki (19200kb/s) łącznie z czasem detekcji końca wynosi: 4,394 ms. Ta sama ramka po dopasowaniu na potrzeby protokołu TCP („kapsułkowaniu") ma długość150 bitów. Ze względu na wielokrotnie większą prędkość czas przesyłu ramki wynosi ok. 0,01 ms.

Jak widać, czas na przesłanie ramki w sieci TCP jest krótszy ok. 400 razy od czasu wymaganego dla sieci MODBUS. Nawet po doliczeniu opóźnień wnoszonych przez programy w sterowniku i komputerach czas ten pozwala na prawidłową pracę sieci z zachowaniem gwarancji czasowych.

Rozwinięciem badań było rozbudowanie stanowiska o dodatkowy trzeci zestaw stanowiący komputer wraz ze sterownikiem. Przy próbie zaprogramowania go jako stacji SLAVE, sytuacja w sieci nie zmieniłaby się. Jedyną zmianą byłyby rzadsze zapytania do każdego sterownika, gdyż jedna stacja MASTER musiałaby obsługiwać dwie stacje SLAVE. Sytuacja taka może zaistnieć w przypadku rozbudowy sieci o nowe pomiary. Spowoduje to wydłużenie się czasu cyklu wymiany informacji, który można dokładnie wyliczyć dla samej sieci MODBUS, a w przybliżeniu oszacować dla sieci zintegrowanej. Nie uzyskamy jednak tu żadnych nowych możliwości przekazywania informacji.

Aby zasymulować sytuację połączenia dwóch sieci klasy I (typu MASTER-SLAVE), dołożona została druga stacja MASTER. Obie stacje MASTER po połączeniu powinny wykonywać poprzednie zadania, a także wykonywać komunikację poziomą pomiędzy zintegrowanymi sieciami.

Aby sprawdzić, czy takie połączenie będzie możliwe do wykonania i będą spełnione wymagania czasowe, trzecie stanowisko zaprogramowano jako MASTER. Zmodyfikowany układ składał się więc z dwóch stacji MASTER i jednej stacji SLAVE. W normalnej sieci MODBUS jest to niemożliwe, bowiem stacja MASTER może być tylko jedna ze względu na adresowanie tylko modułów SLAVE. W sytuacji kapsułkowania (tunelowania) adresuje się komputer (adresem w sieci TCP/IP), który z kolei jednoznacznie określa sterownik.



Rys. 7. Rozbudowane stanowisko badawcze
Fig. 7. The extended laboratory stand

Podczas kolejnych testów okazało się, że przy dwóch stacjach MASTER duże znaczenie ma właściwa obsługa programowa w komputerach. Pojawiają się tu takie problemy, jak:

- rozpoznawanie i pamiętanie do której stacji MASTER należy wysłać informację,
- szybkość odpowiedzi, bowiem ramki mogą nadchodzić dwa razy częściej niż mogą być obsługiwane.

Na podstawie przeprowadzonych badań można powiedzieć, że w przypadku integracji dwóch sieci klasy MODBUS za pomocą protokołu TCP system przy właściwie zaprojektowanym i wykonanym oprogramowaniu zachowuje właściwości determinizmu czasowego. Można powiedzieć, że prędkość w tak zintegrowanej sieci na pewno nie będzie większa niż w pojedynczych sieciach, ale nie będzie też znacząco mniejsza. W wielu wypadkach będzie to całkowicie wystarczające.

Inną zaletą takiej integracji może okazać się większa wydajność przekazywania informacji. W standardowej sieci MODBUS abonent SLAVE odpowiada tylko na ramki od „swojego" MASTERA. W przypadku 10 stacji SLAVE można stwierdzić, że średnio każdy moduł obsługuje co 10 ramkę, a więc w ciągu 90% czasu nie przekazuje żadnych informacji.

Można zastanowić się, czy w przypadku podłączenia dwóch stacji MASTER, jak w opisywanym wyżej stanowisku laboratoryjno-badawczym, nie uda się zwiększyć wydajności przez odpytywanie stacji SLAVE przez dwie stacje MASTER. Porównanie czasów przesyłania dla sieci MODBUS i TCP pokazuje, że w czasie przesłania jednej ramki w sieci MODBUS, możemy przesłać co najmniej 400 ramek dla sieci TCP. Przy odpowiednio przygotowanym algorytmie, zawierającym priorytety dla stacji MASTER, można podczas jednego cyklu wymiany informacji (np. dla sterowania) przez pierwszą stację MASTER, dokonać zbierania przynajmniej części danych dla drugiej stacji MASTER. Taki układ może

zostać wykorzystany do przekazywania informacji na poziom wyższy, przy jednoczesnej normalnej pracy sieci MODBUS.

7. Obserwacje

Z przedstawionych częściowych badań wypływa wniosek o możliwości stosowania połączenia protokołów MODBUS i TCP. Sieć oparta na takim protokole charakteryzuje się gwarancjami czasowymi, nie wymaga modyfikacji istniejących urządzeń przemysłowych, jest łatwa w rozbudowie. Bardzo dużą rolę odgrywa jednak sposób wykonania dopasowania protokołu MODBUS do protokołu TCP. W przypadku stanowiska badawczego dopasowanie to odbywało się w sposób programowy. Od jakości realizacji tego programu będzie zależała poprawność działania sieci. Podczas eksperymentów zauważono bardzo duże różnice w czasie przesyłania informacji, wynikające z niewłaściwego dopasowania buforowania i braku optymalizacji.

Dodatkowo dokonano wielu testów przy pracy dwóch abonentów „MASTER". W takiej sytuacji zdarzało się, że do stacji „SLAVE" przychodziły dwie ramki, każda od innej stacji „MASTER" i z innej sieci. Sytuacja taka nie występuje dla sieci MODBUS. Stacja „SLAVE" nie ma możliwości adresowania stacji „MASTER". Program integrujący musi zostać wyposażony w mechanizm zapamiętywania, która stacja „MASTER" wysyłała obsługiwaną ramkę i gdzie wysyłać odpowiedź.

Powyższy przykład integracji stanowi początek badań nad możliwościami zastosowań protokołu w sieciach przemysłowych do wykonywania zarówno wymian poziomych, jak i pionowych dla systemów czasu rzeczywistego.

8. Wnioski

Projektowanie rozwiązań komunikacyjnych dla sieci przemysłowych najniższego poziomu wymaga realizacji różnych wymagań. Część funkcji wymagać będzie zapewnienia usług należących do klasy *„hard real-time system"*, inne zadowolą się usługami typu *„soft real-time system"*. Wykorzystanie protokołu TCP/IP stwarza nowe możliwości, dotyczące integracji różnego typu systemów sieciowych, i to zarówno w zakresie sieci przemysłowych najniższego poziomu, jak i w zakresie pozostałych warstw hierarchicznego systemu sterowania i wizualizacji. Z opisanych eksperymentów można wyciągnąć wniosek, iż zastosowanie protokołu TCP/IP w sieciach przemysłowych najniższego poziomu jest możliwe. Wykorzystanie tego protokołu nie może być jednak automatyczne. Protokół sam z

siebie nie gwarantuje wymagań czasowych, a integracja TCP/IP z protokołami przemysłowymi wymaga zastosowania wydzielonego segmentu sieci, zastosowania odpowiednio zaprojektowanego oprogramowania (driverów) kart sieci TCP i analizy czasowej, aby wyliczyć, czy czas graniczny T_G po integracji będzie odpowiadał wymaganiom systemu.

LITERATURA

1. Comer D.: Sieci komputerowe TCP/IP (tom 1, 2, 3). WNT, Warszawa 2001.
2. Cupek R., Fojcik M.: Budowa modułów komunikacyjnych stacji nadzorczej z sieciami przemysłowymi. ZN Pol. Śl. s. Informatyka z. 32, Gliwice 1997.
3. Kwiecień A., Cupek R.: Ocena przydatności protokołu TCP/IP dla sieci przemysłowych najniższego poziomu. Materiały konferencyjne SCR'01, AGH, Kraków 2001.
4. Cupek R.: Protokół TCP/IP w systemach wizualizacji procesów przemysłowych. ZN Pol. Śl. s. Studia Informatica Vol. 22 Number 3, Gliwice 2001.
5. Cupek R.: Metody hierarchizacji rozproszonych procesów przemysłowych. ZN Pol. Śl. s. Informatyka z. 28, Gliwice 1995.
6. FIP NETWORK General Introduction, DPS 50249 aA, Clamart 1990.
7. Fojcik M.: Ograniczenia w wymianie informacji w zintegrowanych sieciach przemysłowych. ZN Pol. Śl. s. Informatyka z. 28, Gliwice 1995.
8. Fojcik M.: Problemy integracji sieci przemysłowych. Zeszyty Naukowe Pol. Śl., z.36, Gliwice 1999.
9. Kwiecień A., Gaj P.: Kryteria doboru protokołów komunikacyjnych w sieciach przemysłowych. ZN Pol.Śl. s.Informatyka z.45, Gliwice 2001.
10. Kwiecień A., Fojcik M.: Kryteria integracji sieci przemysłowych najniższego poziomu. ZN Pol. Śl. s. Studia Informatica Vol. 22 Number 3, Gliwice 2001.
11. Kwiecień A.: Sieć rozległa FIP. ZN Pol. Śl. s. Informatyka z. 24, Gliwice 1993.
12. Halang W. A., Real-Time Systems: Another Perspective, Real Time Systems: Abstractions, Languages and Desing Methodologies, K. M. Kavi (ed.) IEEE Computer Society Press, 1992.
13. Kwiecień A.: Analiza przepływu informacji w komputerowych sieciach przemysłowych. Wydawnictwo Jacka Skalmierskiego, Gliwice 1999.
14. Praca zbiorowa: Rozproszone systemy komputerowe. Pronet, Gliwice 1994.
15. Materiały firmowe: OPEN MODBUS/TCP SPECIFICATION. Schneider Electric, marzec 1999.

Recenzent: Dr inż. Włodzimierz Boroń

Wpłynęło do Redakcji 18 kwietnia 2002 r.

Abstract

This paper describes industrial network protocols divided into 3 classes. All of this classes are suitable for fulfill time requirements and can be used in real time systems. Article describes division of real time system into 2 types: "soft" and "hard". The division base on the time reason.

Authors shows purposes and method of integration, and 2 types of data exchanges in integrated network – "vertical" and "horizontal".

In spite of popularity of TCP/IP protocol in computer networks an article shows possibilities of using this protocol in industry solutions. There are made some experiments with integration TCP/IP and MODBUS protocols. And first results shows that these kind of integration can be used with guaranteed time of exchange information.