

# Le doti nascoste di PICMG 2.16

Sebbene la versione 2.16 dell'architettura d'interconnessione PICMG2 sia stata propagandata essenzialmente per il suo nuovo approccio ai backplane, essa possiede anche altri pregi che si possono riassumere in termini di semplificazione dell'integrazione, di miglioramento della sicurezza, di maggiore disponibilità operativa, e d'ottimizzazione dell'uso delle risorse

**Robert Potter (PTI)**



Già dal suo lancio, avvenuto nel settembre del 2000, l'architettura PICMG 2.16 è stata oggetto di notevole interesse da parte della comunità dei progettisti di sistemi di telecomunicazione. Gli sviluppatori di schede e i costruttori di apparati si sono dimostrati rapidi nell'apprezzare questa tecnologia e nel cominciare a sviluppare soluzioni basate su di essa. La maggior parte degli osservatori che inizialmente si occuparono della versione 2.16 dello standard focalizzò la propria attenzione sui vantaggi offerti in termini di più alta affidabilità, migliore scalabilità, maggiore densità e più elevate prestazioni. Le promesse non sono andate disattese e questo backplane si sta dimostrando una soluzione avanzata rispetto alle attuali architetture PCI. Tuttavia, i miglioramenti a livello di backplane rappresentano solo il punto di partenza di una serie di benefici offerti dall'architettura PICMG 2.16. Questa nuova architettura offre agli sviluppatori di sistemi la possibilità di semplificare le problematiche d'integrazione, in modo da accelerare i tempi di sviluppo e di commercializzazione, migliorare la disponibilità operativa, migliorare le funzionalità di sicurezza, ed elevare il grado d'estensibilità. Vediamo il perché di tutto questo nel seguito.

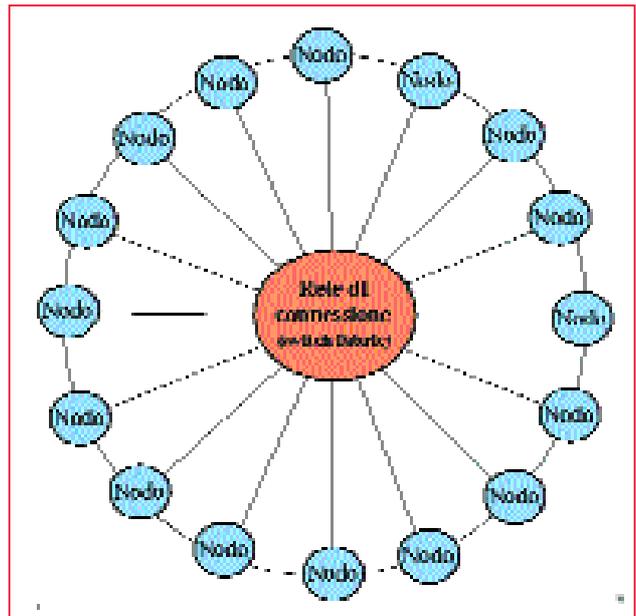


Fig.1 - In una rete con topologia a stella semplice, ciascuna scheda costituisce un nodo che si connette direttamente al centro di commutazione.

## Backplane a commutazione di pacchetto

È utile innanzitutto richiamare alcuni concetti base inerenti le reti di comunicazione Ethernet. In una rete con topologia a stella diversi nodi si scambiano informazioni comunicando esclusivamente attraverso un nodo centrale di commutazione che provvede a smistare i dati verso i vari punti della rete. Ne risulta che ci sono tanti collegamenti quanti sono i nodi utente, e questo rappresenta una drastica semplificazione rispetto alle tipologie di rete basate su connessioni dirette nodo-a-nodo, che sono caratterizzate da un'enorme quantità di collegamenti fisici. In una configurazione di rete a stella semplice, con un unico centro stella come illustrato nella figura 1, ciascun nodo si connette direttamente al centro di commutazione, che è costituito da una rete di connessione (switch fabric).

In una rete Ethernet, i dati sono distribuiti in pacchetti che vengono smistati e recapitati sulla base delle informazioni contenute nella loro inte-

# HARDWARE

## ARCHITETTURE D'INTERCONNESSIONE

stazione (header). In particolare l'header include due indirizzi: l'indirizzo di destinazione, che specifica dove i dati devono essere recapitati, e l'indirizzo del mittente, che indica la risorsa da cui i dati provengono.

Quando lo switch centrale riceve un pacchetto, ne esamina il campo header per leggerne l'indirizzo di destinazione (supponiamo DDD) che specifica dove spedire il pacchetto; in questo modo lo switch conclude che il pacchetto è destinato al nodo avente indirizzo uguale a DDD. Il centro di commutazione conosce già quale delle sue porte è collegata al nodo avente indirizzo DDD, perché, in base ad un meccanismo operativo di rete, ha ricevuto in precedenza un messaggio da quest'ultimo su una determinata porta (diciamo PPP), e tale messaggio gli consente di capire che il nodo avente indirizzo DDD è connesso alla porta PPP. Si tratta di una metodologia operativa molto potente, poiché la rete è in grado di verificare automaticamente i cambi di connessioni che si hanno quando un nodo viene disconnesso da una porta e riconnesso da qualche altra

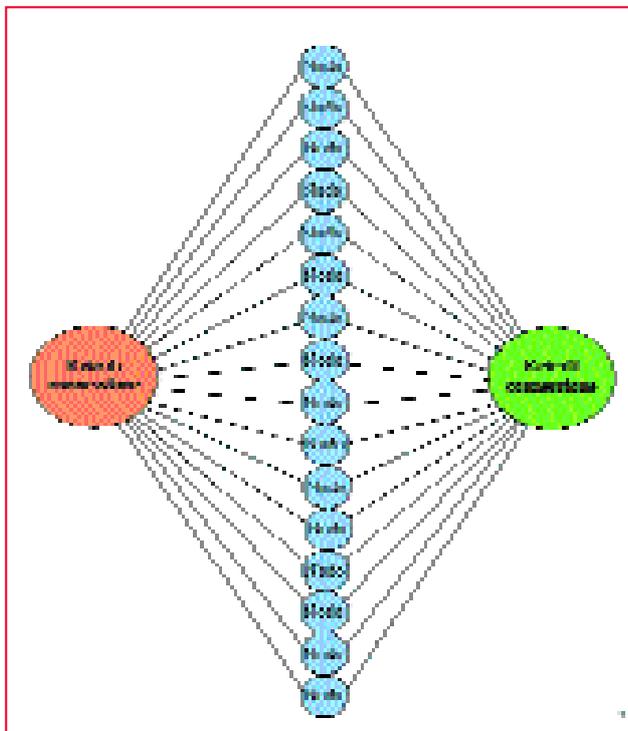


Fig. 2 - Nell'architettura a doppia stella dello standard PICMG 2.16, ciascun nodo possiede due porte che sono destinate a connettersi ai due centri di commutazione indipendenti della piattaforma (switch fabric), formando così percorsi di comunicazione ridondanti

parte, senza bisogno di interventi dell'utente. In questo modo l'integrazione di nuovi componenti risulta più semplice e più rapida rispetto ai sistemi d'indirizzamento tradizionali a bus.

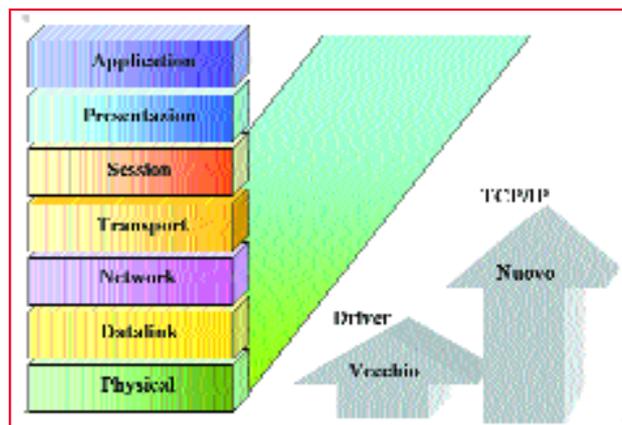
Verifichiamo ora cosa succede nel caso vi siano eventuali problemi nella comunicazione. In una rete a stella semplice ogni connessione non è supportata da percorsi alternativi ridondanti, vale a dire ciascun collegamento è unico, e in caso di guasto il relativo nodo rimane isolato. La situazione è diversa nell'architettura a doppia stella della versione 2.16 dello standard PICMG, dove ciascun nodo ha due porte collegate rispettivamente a due centri di commutazione indipendenti (Fig. 2).

Questa topologia a doppia stella può tollerare eventuali problemi operativi in molti punti differenti della rete, cosicché una comunicazione sostenuta in un certo periodo rimane effettiva per un'alta percentuale di tempo. Nella piattaforma PICMG 2.16 reale, tutte le connessioni tra nodi e centri di commutazione sono fisicamente realizzate sulla scheda madre di un cestello. Questo cestello ha 21 slot, di cui due sono riservati ai centri di commutazione e gli altri 19 sono disponibili per realizzare i nodi di sistema.

### Semplificazione dell'integrazione

Le problematiche d'integrazione sono un fattore fondamentale nello sviluppo dei sistemi. In particolare è di cruciale importanza l'integrabilità nel

Fig. 3 - Adottando l'architettura PICMG 2.16 i progettisti possono risolvere i problemi d'integrazione operando a livelli d'interfacciamento più alti del modello OSI



tempo di nuovi componenti in progetti datati e di vecchia generazione. Il processo d'integrazione deve essere anche semplice e poco dispendioso in termini di tempo, cosa che non si verifica nel caso delle architetture di sistema basate su bus condivisi. Infatti, in questi casi, l'integrazione di ciascuna scheda richiede lo sviluppo di un driver dedicato per ogni sistema operativo supportato. L'architettura PICMG 2.16 consente di ridurre i tempi e la complessità del processo di integrazione offrendo un approccio che prescinde dal tipo di sistema operativo usato e che opera ad alto livello, ovvero negli strati più alti del modello di interfaccia di comunicazione OSI (Fig. 3). Ne risulta una piattaforma adatta a facilitare i problemi d'integrazione anche nei sistemi più esigenti, e con tempi significativamente più brevi poiché ora l'integrazione dei componenti avviene ai livelli Rete/Trasporto del modello OSI. In sostanza, le varie schede individuali collocate nel chassis della piattaforma PICMG 2.16 sono entità indipendenti a basso livello, e possono avere un sistema operativo e una memoria propria, e la comunicazione tra loro avviene ad alto livello usando gli standard TCP, UDP, SCTP o ogni altro protocollo di trasporto.

In un sistema basato su bus, a ciascuna scheda è allocata una porzione dello spazio d'indirizzamento. Il trasferimento dei dati da e verso la scheda avviene sfruttando aree predefinite entro la porzione di spazio d'indirizzamento della scheda stessa. Questo spazio di indirizzamento è dedicato e viene appropriatamente definito al termine del processo di personalizzazione delle schede, dopo che sono state configurate per supportare le specifiche funzionalità richieste. Per comunicare tra loro, le varie schede devono essere capaci di interfacciarsi mutuamente. A livello di scheda questo viene assicurato tramite un driver, ovvero un'interfaccia software che fornisce il necessario livello di astrazione tra le direttive di comunicazione emesse in termini di concetti funzionali e le dettagliate procedure d'interfacciamento a livello di manipolazione dei bit richieste per implementarle. Sempre per architetture basate su bus condivisi, l'integrazione a livello di sistema è resa complicata dal fatto che tutte le schede si contendono un canale di comunicazione comune, e in ogni istante una sola risorsa può funzionare da trasmettitore, mentre un'altra agisce come ascoltatore. Integrare un sistema significa assicurare che tutte le schede siano in grado d'interfacciarsi corretta-

## A adattatore per telecomunicazioni

**Di recente introdotto da Performance Technologies, CPC388 è un adattatore per telecomunicazioni a otto porte T1/E1/J1 che risponde alle esigenze delle reti di telefonia IP e di telecomunicazione della futura generazione. Appartenente alla famiglia di prodotti IPnexus, CPC388 offre alte prestazioni per applicazioni voce/dati in formato 6U CompactPCI**

CPC388 si abbina perfettamente ai sistemi di telefonia sia di rete telefonica pubblica (PSTN) sia a protocollo IP, che gestiscono ampi volumi di circuiti vocali, per il trattamento dei protocolli o per il trasferimento tra interfacce differenti. Esempi di applicazione sono costituiti da elementi di reti SS7, apparecchiature di infrastrutture wireless, media e gateway di trasmissione e apparecchiature di commutazione e instradamento di telecomunicazioni.

CPC388 fornisce un supporto H.110 completo ed è in grado di commutare fino a 256 time slot in senso bidirezionale tra bus locale e H.110. Le interfacce 10/100 Ethernet duali consentono una connettività IP ridondante per applicazioni carrier-grade di alta affidabilità.

CPC388 utilizza la funzionalità avanzata del suo processore interno Motorola 8260 PowerQUICCII. Questo consente la gestione a un livello superiore di otto flussi T1/E1/J1 completamente canalizzati, incrementando il numero dei possibili link attivi dei protocolli.

L'ampio supporto software si basa su un sistema operativo Linux, che consente agli sviluppatori di sfruttare i protocolli e le tante applicazioni comunemente disponibili, riducendo così i tempi di commercializzazione. Con una ben definita API, la suite di protocolli integrata di Performance Technologies elimina i tempi superflui di sviluppo a livello di hardware/protocollo. I sistemi operativi supportati comprendono Solaris, Linux, Windows NT e VxWorks.

Per una manutenzione semplice e tempi di riparazione minimi, CPC388 è "hot swappable" secondo la più recente specifica PCI Industrial Computer Manufacturers Group (PICMG). L'hot swapping è reso più facile da un modulo di transizione I/O passivo, che elimina i punti di guasto e aumenta l'affidabilità.

La memoria DRAM del processore dedicato da 128 MB consente a CPC388 di gestire vaste esigenze interne di traffico e protocolli. CPC388 è dotato inoltre di un'area per PCI Telecom Mezzanine Card (PTMC) interna, fornendo così capacità di espansione per funzioni di potenziamento.

Come gli altri prodotti della famiglia IPnexus, CPC388 supporta lo standard proposto Compact Packet Switching Backplane (cPSB). Questo gli consente di operare in un cestello convenzionale CompactPCI o via Ethernet in una configurazione con backplane a commutazione di pacchetti. cPSB è una specifica PICMG di massima per la sovrapposizione di un'architettura a commutazione di pacchetti su un CompactPCI.

mente con le varie risorse con cui dovranno sostenere una comunicazione, e questo per un ammontare di tempo sufficiente a consentire alle risorse stesse di portare a termine la loro attività. Tutto ciò, in funzione della complessità del sistema, può richiedere mesi o anni. Inoltre, i sistemi che sono troppo grandi per essere contenuti in un unico chassis richiederebbero un tempo considerevolmente grande per essere integrati. Nei sistemi basati sui bus-base, quando una scheda deve essere estratta dal cestello per manutenzione, si avvia un meccanismo operativo graduale gestito centralmente che porta alla disattivazione del driver, nota come 'driver quiescing'. Nello stato di 'quiescing' il sistema smette di comunicare con la scheda, cosicché l'hardware può essere disconnesso dal bus senza ripercussioni sulla comunicazione in corso tra le rimanenti schede. Il processo d'inserzione presenta, analogamente, lo stesso grado di complessità.

### Integrazione con lo switch

In un sistema basato su una piattaforma di connessione Ethernet, le schede costituiscono i nodi della rete e usano il protocollo Internet (IP) per effettuare lo scambio dei dati.

In questo senso, a livello di scheda le esigenze di comunicazione necessarie a supportare le funzionalità proprie intrinseche della scheda vengono tramutate in direttive d'interfacciamento a pacchetti in accordo al protocollo IP. La scheda switch, che agisce come rete di connessione (switch fabric), permette di avere conversazioni multiple simultanee, così un sistema a N-nodi può supportare contemporaneamente N/2 conversazioni full-duplex. Ne risulta che i sistemi complessi che coinvolgono schede multi-nodo, nodi di commutazione e connessioni di rete esterne possono essere integrati nel giro di poche ore. Le schede di base della piattaforma, incorporate nello chassis, vengono collegate insieme con cablaggio Ethernet tradizionale per consentire facilmente di configurare architetture di sistema più grandi rispetto al singolo chassis.

Complessivamente, lo standard PICMG 2.16 è stato progettato per supportare fino a 24 slot per chassis. Se per esigenze di manutenzione una scheda nodo o uno switch dovessero essere estratti, basta semplicemente rimuoverli senza complicazione alcuna.

Non esiste nessuna necessità di disalimentare il sistema e non c'è il sovra flusso di traffico dati

associato alle procedure di 'quiescing' dei driver, e i rimanenti componenti del sistema sono immediatamente disponibili per continuare ad operare. Se viene estratta una scheda nodo, la scheda switch (centro di commutazione) rivela che tale nodo è stato rimosso, e i pacchetti vengono conseguentemente instradati verso una scheda nodo alternativa. Viceversa, se un nodo rileva che una scheda switch è stata rimossa, i pacchetti vengono instradati verso l'altro nodo di commutazione. Anche il processo d'inserzione risulta, analogamente, altrettanto semplificato.

### Riutilizzo del software esistente

Come precedentemente affermato, la maggior parte degli osservatori che si sono occupati dell'architettura PICMG 2.16 si è limitata solo a evidenziarne i miglioramenti offerti nei confronti dei sistemi basati su bus, e in particolare dei progetti Compact PCI esistenti.

Ora, un altro aspetto degno di essere rimarcato è che lo standard PICMG 2.16 consente ai progettisti di trarre vantaggio dall'abbondanza di software pre-esistente, sviluppato negli ambienti aziendali per le reti 10/100 Megabit Ethernet e Gigabit Ethernet. Tra i benefici che ne derivano si hanno l'alta disponibilità operativa, la migliore sicurezza e l'estensibilità dinamica.

Uno dei molti aspetti che i progettisti di sistemi basati sulle specifiche dello standard PICMG 2.16 possono vantaggiosamente risolvere sfruttando il software esistente è quello di riuscire a ottenere un'elevata disponibilità operativa della rete.

Specificamente, per incrementare il tempo operativo di rete e ottenere una più elevata disponibilità di sistema, si può usare l'architettura PICMG 2.16 in combinazione con i seguenti due protocolli: lo Spanning Tree Protocol (STP) e il Virtual Router Redundancy Protocol (VRRP).

Il protocollo STP è specificamente orientato alla gestione dei link ed è in grado di stabilire percorsi ridondanti, prevenendo nel contempo i ricircoli dei pacchetti dati su percorsi chiusi senza fine.

Il protocollo STP offre un tempo di convergenza relativamente breve (da 30 a 50 secondi), che risulta non rilevante per le applicazioni aziendali quali i sistemi di rete basati sugli standard HTTP o FTP.

Il protocollo Rapid Spanning Tree, un miglioramento dell'STP, è stato creato per fornire una convergenza più veloce (meno di un secondo) nei con-

## ARCHITETTURE D'INTERCONNESSIONE

fronti di guasti in un nodo di commutazione o in un collegamento, oppure nei riguardi di cambiamenti dei parametri STP.

Questa maggiore rapidità di convergenza tiene conto di potenziali applicazioni incorporate come servizi vocali o video.

Il protocollo VRRP offre il vantaggio di risolvere a livello di struttura interna della rete i problemi di definizione delle route di smistamento dei pacchetti dati, riuscendo ad offrire percorsi d'instradamento predefiniti a più elevata disponibilità operativa, senza la necessità di dover scaricare questi compiti su tutte le postazioni finali di rete, che altrimenti dovrebbero essere configurate con opportuni protocolli dedicati orientati alla ricerca dei percorsi d'instradamento (dynamic routing o router discovery protocol).

Il protocollo VRRP elimina il singolo punto di guasto rispetto alla configurazione d'instradamento predefinita statica.

Esso basa il suo funzionamento su un processo ad elezione di responsabilità che assegna dinamicamente il ruolo di router virtuale ad una risorsa di rete.

Se questa risorsa, funzionante nel ruolo di master, si guasta, il meccanismo ad elezione delle competenze del protocollo VRRP ovvia dinamicamente a questo inconveniente attuando un trasferimento di responsabilità ad un altro router; il tutto resta trasparente alle postazioni di rete che continuano a vedere lo stesso router virtuale anche se questo è fisicamente cambiato.

#### Miglioramento della sicurezza

In aggiunta all'elevata disponibilità operativa, i sistemi conformi allo standard PICMG 2.16 offrono anche i vantaggi di sicurezza delle reti a commutazione di pacchetto. In particolare, la nuova architettura PICMG consente di definire delle reti locali virtuali (VLAN) per migliorarne i livelli di sicurezza. Una VLAN è una rete locale definita logicamente, vale a dire identificata da un sotto insieme di dispositivi che, indipendenti dal punto di connessione e dalla topologia di rete, comunicano esclusivamente tra di loro come se facessero parte di una LAN fisica separata. I dispositivi di una VLAN possono comunicare solo con altre risorse appartenenti alla stessa VLAN. Le VLAN consentono di costruire domini di trasmissione broadcast senza bisogno di ricorrere a restrizioni sulle connessioni fisiche. Al fine di comprendere i potenziali benefici delle VLAN, conviene innanzi tutto passa-

re in rassegna i vari metodi d'indirizzamento dello standard Ethernet: unicast, broadcast e multicast. L'indirizzamento unicast consente ad una sorgente di indirizzare una singola destinazione, si tratta della modalità vista nell'esempio precedente relativo al recapito di un pacchetto dati ad un unico nodo. L'indirizzamento broadcast consiste nell'inviare i dati a tutte le destinazioni; infine la modalità multicast consente di indirizzare solo un sottoinsieme di destinatari. Vari protocolli usano l'indirizzamento broadcast e multicast quando l'indirizzamento unicast non è disponibile oppure il relativo indirizzamento non è pratico. L'indirizzamento broadcast e multicast non è giustificato nei casi in cui quello unicast basta a soddisfare gli scopi, tuttavia può accadere che questo non sia utilizzabile a causa della non disponibilità dell'indirizzo di destinazione. Consideriamo un nodo di commutazione che riceva un pacchetto con indirizzo di destinazione DDD. Per spedire questo pacchetto sulla porta corretta il nodo deve avere precedentemente appreso che il destinatario di indirizzo DDD è collegato alla porta PPP, secondo la modalità vista prima. Ora, qualora tale nodo non possieda ancora quest'informazione, il pacchetto viene trasmesso a tutte le porte. Così un intruso (snooper) sarà in grado di intercettare ed esaminare le trasmissioni broadcast e di catturarne informazioni di non propria pertinenza. Le reti virtuali VLAN offrono una soluzione a questo problema perché consentono di limitare la diffusione delle informazioni permettendo ai nodi di commutazione di trattare un pacchetto broadcast alla stregua di un pacchetto multicast.

Le reti VLAN funzionano sulla base di determinati criteri di filtraggio dei pacchetti dati; a seconda del tipo di criterio si hanno VLAN protocol-based, VLAN address-based, VLAN port-based, ecc. Una VLAN viene creata definendo un criterio di selezione in base al quale eliminare dal traffico dati i destinatari indesiderati. In sostanza, un nodo di commutazione che vuole eliminare un certo numero di porte dalla ricezione di dati inviati in modo broadcast, può introdurre un criterio di filtraggio VLAN per delimitare opportunamente il dominio di trasmissione. Questo sistema consente di eliminare dalla rete il traffico indesiderato e di incrementare l'efficienza operativa di rete, e permette infine di prevenire l'intercettazione delle informazioni trasmesse da parte di potenziali intrusi, visto che i dati non sono inviati indiscriminata-

mente a tutte le utenze.

### Scalabilità

I sistemi in architettura PICMG 2.16 possono essere impiegati congiuntamente a protocolli di aggregazione dei collegamenti, i quali permettono a due o più porte fisicamente indipendenti di essere aggregate per formare una singola connessione logica avente banda maggiore, vale a dire un gruppo di aggregazione di collegamenti (LAG). Questo schema di interfacciamento consente ai sistemi di avere bande di collegamento più elevate rispetto a quelle fornite individualmente da ciascuna connessione. I sistemi possono incrementalmente adattare la banda della porta di uplink proporzionalmente alle esigenze del sistema. Nel caso di un guasto ad un collegamento, il sistema automaticamente si riconfigurerà in modo da usare le porte fisiche rimanenti, secondo uno schema che incrementa l'affidabilità ottimizzando dinamicamente l'uso delle risorse disponibili (graceful degradation). I protocolli LAG consentono attualmente di avere sistemi che possono sfruttare vantaggiosamente bande di comunicazione maggiori di 1 Gb, e fino a 10 Gb con la maturazione tecnologica.

### Applicazioni PICMG 2.16

Lo standard PICMG 2.16 è idoneo per la sintesi di applicazioni sia nel settore dei sistemi cablati sia in quello senza fili. Nel settore wireless, lo standard PICMG 2.16 rappresenterà un riferimento per i progettisti di prodotti destinati alle infrastrutture a 2.5G/3G. Nei sistemi cablati invece, le potenziali applicazioni includono i sistemi di supporto per server autonomi (embedded server clustering), i controllori di gateway per dati/audio/video (media gateway controller), i gateway VoIP per dati vocali (VoIP media gateway), i gateway per dati di controllo, e altro ancora.

