Architectural Design Study for a 10Gb/s Ethernet Switch

referat de doctorat

Doctorand: ing. Alexandra Dana OLTEAN
Conducator stiintic: prof. dr. ing. Vasile BUZULOIU

August 2004
# Content

1  Introduction........................................................................................................................................... 3  
   1.1  Context........................................................................................................................................... 3  
   1.2  Approach....................................................................................................................................... 6  
2  Connectors ............................................................................................................................................. 8  
3  Backplane technology.......................................................................................................................... 14  
   3.1  Terminology.................................................................................................................................. 14  
   3.2  Routing Constraints ..................................................................................................................... 17  
   3.3  Impedance controlled traces ........................................................................................................ 19  
   3.4  Trace Constraints ....................................................................................................................... 21  
4  Connector Pin-field Simulation ............................................................................................................. 27  
   4.1  The problem................................................................................................................................. 27  
   4.2  The tools....................................................................................................................................... 28  
   4.3  The single via............................................................................................................................... 29  
   4.4  Coupled VIAs ............................................................................................................................ 30  
   4.5  Via trace crosstalk....................................................................................................................... 32  
5  Advanced TCA ..................................................................................................................................... 41  
   5.1  Conclusion ................................................................................................................................... 43  
6  Conclusions........................................................................................................................................... 45  
7  References............................................................................................................................................ 47
1 Introduction

1.1 Context

The demand for 10Gb/s switches at this early stage in the market is primarily for modular solutions that can grow as do the bandwidth requirements. This indicates a chassis based solution where individual line cards can be added to a chassis infrastructure. The way that data is switched between different cards depends on the chipset(s) available. There are two main classes of switch depending on the choice of switch chipset.

In the first case the switch functionality is split between what is port specific and that which is system specific. This leads to a star architecture as shown in Fig. 1.1., where part of the switching function is done on the line card and the rest in the central or core switch which may be made up of one or several chips.

Fig. 1.1. Star topology for a 10Gb/s switch

One major drawback of this architecture is that the core switch card constitutes a single point of failure and so the preferred solution is the redundant star topology, as shown in figure 1.2. If one core switch fails then the switch can quickly be reconfigured to use the other one.
In the second case all of the switching functionality can be accommodated on the line card. Now, each line card needs an interconnect to each and every other line card to build up the complete switch. This is a *mesh architecture* and is shown in figure 1.3.

In either case there is a growing demand for moving data across the copper backplane at 10Gb/s as a goal. However, it is not practical to transmit data between cards at 10Gb/s over one or two conductors, because the losses involved in traditional printed circuit board technology are too high. Therefore the data transmission technologies of choice run
typically between 2.5 and 3.5 Gb/s per conductor pair. Each 10Gb/s link therefore needs at least four conductor pairs to operate a 10Gb/s link.

This architecture for a 10Gb/s switch is not easily scalable since the available space on the backplane for routing all these connections is finite and the space that is available has to be carefully managed to avoid crosstalk between adjacent connecting pairs. Our task is to see to what extent we can have confidence that the interconnect will not be a source of system failure.

The weakest point for this type of design is traditionally the connector. It is possible to keep printed circuit traces well separated just by making larger the cards area, despite the cost. However the design is severely constrained since the pins of the connector passing through the printed circuit limit the available space for the routing channels.

![Fig. 1.4. A typical star fabric backplane](image)

The figure 1.4 above shows a typical backplane which underlines the point to which such boards have more connectors to traverse than open backplane. When routing signal across such a backplane, the portion most geometrically constrained in terms of space availability and ability to maintain the target impedance, is the area where the signals must be routed through connector pin-fields. For instance, a too large via anti-pad defined by the manufacturer, can cause an important deviation from theoretical design
impedances and strongly affect the return current distribution. On the other hand, a too
constraint manufacturer definition of the routing channels may provide less than adequate
separation space between trace and via anti-pad. In addition, the connector manufacturers
characterise their products in isolation from the PCB in which they are mounted. The role
of this task is to consider the system aspects of the interconnect and not just any one
component in isolation.

1.2 Approach

At the time this project was being put together, papers in DesignCon2001 were warning
of the difficulties of achieving 1Gb/s backplanes with standard technologies. By 2002 the
same conference was talking about terabit backplanes and 6Gbps channels using standard
printed circuit material, FR4. What made this difference were advances on two fronts.
From the chip manufacturers came different solutions in the form of adaptive equalisation
on receivers or pre-emphasis and multilevel encoding on transmitters. From the connector
manufacturers came a second generation of differential connectors with a strong
emphasis on pair shielding that allowed them to characterise differential transport up to
5Gbps.

The present technology study is trying to match the chip technologies and packaging
limitations into a 10Gb/s backplane system, corresponding to the market demands. For
the most part this approach was successful and we came to a good understanding of how
to package a backplane that would sustain the design criteria. In the event, the
improvements in transmitter technology have probably bought enough headroom that the
packaging and interconnecting become second order effects, but it was preferable to be
sure.

In order that a 10Gb/s backplane system to be successful and result in a practical and
economical solution, all aspects of the mechanical and electrical interconnect
infrastructure must be considered. This paper will examine the influences of PCB routing
and those of the traces constraints. Another study will be provided for the influence of the connector choices on the design of the 10Gb/s backplane and successful design rules will be provided.

There are still a number of unresolved issues concerning signal integrity that occur in the transition zone between the connectors and their interconnecting traces and as well in the area where traces are passing across the connector pin-fields on the backplane. Furthermore, it will be examined the quality of a transmitted digital bit sequence along differential traces when they are routed across the connector aggressing pin-field. It will be demonstrated that the chosen routing across the connector pin-field can provide reasonable eye opening at the receiver.

Late in 2002 over one hundred companies forming the PICMG consortium addressed the needs of future switch fabric infrastructure with a mesh or star fabric backplane standard. The result of this was the specification of Advanced TCA (Advanced Telecom Computing Architecture) proprietary standard, which defines rack and chassis form factors, core backplane fabric connectivity, power, cooling, management interfaces and the electromechanical specifications of the boards for scalable solution of next generation telecommunication equipment. The design of the Advance TCA passive backplane turned out to be so close to our own conclusions that it was adopted for housing the 10Gb/s Ethernet switch.

This paper covers the main aspects and the factors involved in selecting the appropriate interconnect components and packaging techniques required by the engineering aspects of a 10Gb/s copper backplane and finally validates the prototype design.
2 Connectors

The 10 Gigabit Ethernet standardization efforts specified the XAUI protocol for 10Gb/s data transmission using four parallel channels of 3.125Gbit/sec over copper. This was at the limit of what was possible with the pre-existing standards of 2mm ‘hard metric’ (HM) connector technology by the end of 2001 and the beginning of 2002, especially as far as allowable crosstalk was concerned. More work needed to be done on pair to pair screening and several vendors brought out new products to meet the standard.

An article in Electronics Weekly -- Electronic News, 4/19/2002 provided a useful overview of the available choices.

“Teradyne is proposing its GbX platform connector as a speed upgrade to its existing 2mm connectors with less than 2 percent crosstalk. The GbX connector is available as a 4-pair platform, which offers 55 differential pairs per linear inch, while the 5-pair configuration offers 69 differential pairs per linear inch.

Molex Inc. supplies a 70-pin PCB mounted edge card connector aimed specifically at the emerging 10Gbit/sec (IEEE 802.3ae) family of Z-axis pluggable transceivers.

FCI can also support 10Gbit/sec connections to opto-transceivers with its 300-pin MEG-Array for BGA package types. The key standards for these optical transceiver mounting connectors, as well as all high speed XAUI compatible signaling connectors, are the Telcordia GR-1217-CORE reliability standards.

Tyco Electronics Corp. is working with at least one chip supplier, Gennum Corp., to demonstrate that its Z-PACK HM-Zd connector system using FR-4 board backplanes will run serial data at 10Gbit/sec in a 100ohm differential signaling environment. The Z-PACK HM-Zd product, first introduced in 2001, extends Tyco’s existing IEC 61076-4-101 hard metric connector family with differential signaling that supports 3.5Gbit/sec data rates.
**FCI Electronics** is also offering a 100ohm differential version of its 2mm HM family, **Metral 4000.** Its design uses a stripline structure in both the receptacle and header and modified contact geometry better impedance match and lower insertion loss. The aim says FCI is to allow designers to mix different data rates on the same daughtercard connection.

**Harting Group** has a multi-source agreement with **3M** for its high-speed hard metric (HSHM) family of connectors. While **Molex** and **Teradyne** offer compatible high-speed differential versions of the VHDM 2mm connector system. Called **VHDM-HSD, it is designed for data rates up to 5Gbit/sec in commonly used daughter card-to-backplane applications.**

If we assume all the above vendors claims of achievable bandwidth and cross-talk as credible, then making a choice is divided between technical and commercial considerations. On the technical side we weigh the pin density versus the difficulties of routing the signals on the same backplane between the connector pins. From the commercial point of view, our considerations include cost and multiple source availability.

The figure 2.1. below shows the footprints of a backplane connector (left) and its matching daughter card connector (right) and the main parameters of interest, which are presented in Table 2.1. for each of the competing products.

Fig. 2.1. Critical connector footprint dimensions for motherboard and daughter cards
<table>
<thead>
<tr>
<th>Manufacturers</th>
<th>Teradyne &amp; Molex</th>
<th>Teradyne</th>
<th>Tyco &amp; ERNI</th>
<th>FCI &amp; ITT-Cannon</th>
<th>3M &amp; Harting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Connector type Parameter</td>
<td>HSD</td>
<td>GbX</td>
<td>HM-Zd &amp; ErmetZd</td>
<td>Metral 4000</td>
<td>MetPak HSHM</td>
</tr>
<tr>
<td>Intra-pair spacing on backplane, Dp</td>
<td>2.25 mm</td>
<td>1.85 mm</td>
<td>1.5 mm</td>
<td>2.0 mm</td>
<td>2.0 mm</td>
</tr>
<tr>
<td>Intra-pair spacing on linecard, Dp</td>
<td>2.0 mm</td>
<td>1.5 mm</td>
<td>2.5 mm</td>
<td>2.0 mm</td>
<td>2.0 mm</td>
</tr>
<tr>
<td>Pair pitch in column direction, Cp</td>
<td>2.0 mm</td>
<td>1.85 mm</td>
<td>2.5 mm</td>
<td>2.0 mm</td>
<td>4.0 mm</td>
</tr>
<tr>
<td>Recommended through hole drill size</td>
<td>0.66 mm</td>
<td>0.57 mm</td>
<td>0.7 mm</td>
<td>0.6 mm</td>
<td>0.6 mm</td>
</tr>
<tr>
<td>Recommended through hole pad size</td>
<td>1.0 mm</td>
<td>0.9 mm</td>
<td>1.02 mm</td>
<td>0.9 mm</td>
<td>0.86 mm</td>
</tr>
<tr>
<td>BP Routing channel</td>
<td>1.34 mm</td>
<td>1.28 mm</td>
<td>1.8 mm</td>
<td>1.4 mm</td>
<td>1.4 mm</td>
</tr>
<tr>
<td>LC routing Channel</td>
<td>1.34 mm</td>
<td>0.93 mm</td>
<td>1.8 mm</td>
<td>1.4 mm</td>
<td>1.4 mm</td>
</tr>
<tr>
<td>Pairs per inch for 4 pair columns</td>
<td>38</td>
<td>55</td>
<td>40</td>
<td>38</td>
<td>38</td>
</tr>
<tr>
<td>Worst case differential crosstalk</td>
<td>3.2%</td>
<td>&lt;5%</td>
<td>3.1%</td>
<td>&lt;5%</td>
<td>6%</td>
</tr>
</tbody>
</table>

Table 2.1. Comparative figures for commercial available connector footprints

The two ‘hard metric’ connectors from 3M and FCI have apparently included extra shielding onto their pre-existing 2mm Compact PCI products. This shows up in their comparatively worse crosstalk figures.

If we are really forced into a pin count problem because of the architectural choices then the GBX is a good candidate. On the other hand the narrowness of the routing channel on the backplane for the GBX connector is a cause for concern.

The HSD offers nearly as many pairs per inch as the HM-Zd but has nearly half a millimeter less in its routing channel because it adheres to the 2mm rectangular pitch.
From a purely electro-mechanical point of view then we would rate the **HM-Zd connector as a first choice** and the **GBX as a second choice**, if space becomes a real problem.

Each manufacturer provides, either publicly, or under NDA, SPICE models of their connectors behaviour. The crosstalk figures obtained for the available ‘hard metric’ 2mm connectors are impressive and show great progress over the previous generation. However the models given are only of the connector itself, in isolation from any backplane or daughter-board they are connected to. This is understandable, as the vendors have no control over how their connectors are used. It is important therefore very important to understand the implications and the factors which affect a connector electrical characteristics when it is placed in a ‘real’ PCB stack-up environment. In this direction AMP did some typical measurements of a ‘real’ environment and reported their results in the paper: [www.amp.com/products/simulation/files/papers/21GC009.pdf]. The following extract is highly significant:

“**At 3.125 Gb/s the connector has a significant impact on system performance. Typically when reviewing the performance of a connector, little consideration is given to the footprint of the connector in the PWB. [...] Using the conditions specified by the XAUI Channel Working Group the worst case simulated noise of the Z–PACK HM-ZD was 3.1%. This meets the [...] proposed connector noise budget of 4%.**

*The inclusion of the footprint significantly changed the overall shape and magnitude of the near end and far end noise pulses. For all three pin-outs considered the amount of near-end noise increased by at least double. The impact on far-end noise was pin-out dependent, ranging from a slight decrease (Pin-out 2) to a significant increase (10.5 times for Pin-out 1 and 2 times for Pin-out 3”*

The essential conclusion of this study is that we can almost discount the admitted effect of the connector since the act of plugging it into a backplane can increase the crosstalk
noise by anything up to a factor of ten or not at all. This mainly depends on how carefully
the connector footprint is designed. We investigate this further in the chapter IV.

Both our connector choices - the HM-Zd and GBX connectors - are carefully thought out
and well engineered parts. The figures 2.2 and 2.3 below show the main features of the
GBX part and of the HM-Zd connector.

Fig. 2.2.a.) GBX header and receptacle
Fig. 2.2.b.) GBX mated cross section

The header has columns of contact pairs separated by ground shields. This can be seen in
the left hand picture and show up as the vertical strips in the right hand picture. Shielding
between the rows is done with the backplate to the receptacle wafer that holds the spring
contacts that mate with the header’s pins. This shows as the horizontal strips on the right
hand picture. The pin pair thus propagates through an impedance controlled waveguide
fully shielded from its neighbours. This is true inside the body of the connector, however
the pin leads that enter the printed circuit stack-up no longer have ground shielding and
that is where the main crosstalk effects occur.
In this case the header has a wrap round L-shaped section whose matching other half is in the receptacle. When mated, the receptacle has a spring leaf that makes contact with the headers screen to ensure ground continuity through the connector.

In the next figure 2.3 we see a typical test installation as built by every vendor working at higher speed and used to characterize their products. It is interesting to note that very few channels are exercised concurrently. Although at least one set of sockets runs traces through successive connector pin-fields these connectors are not actually mounted neither are they fed with aggressor signals that could be the source of further crosstalk.
3 Backplane technology

3.1 Terminology

Before discussing fine backplane terminology details consider the figure 3.1. of a hypothetical plated through hole, or ‘via’. A via is typically used to receive connector pins and route signals between different layers in the stack-up of a printed circuit board (PCB). The stack-up is the cross-section of a PCB, defining the number of layers and the thickness and the dielectric properties of each layer used in its fabrication. The figure below is necessary to explain some of the terms used later in the discussion.

![Fig. 3.1. View of signal via through a PCB](image)

The central barrel of the via is usually as narrow as can be, when used to transport signals from one layer to another. As the thickness of the board stack-up increases due to many layers being used, however so must the barrel diameter. That is because there is a practical limit to the aspect ratio (via height divided by via diameter) that can be
successfully plated. Conventionally this limit is about 5. For backplane applications, however, the barrel diameter limit is set by the requirements of the connector pin that is inserted into it and it is much larger than is needed for a signal via. For example, signals can be routed through 6 mil diameter vias whereas a connector can require from 20 to 40 mils in diameter. This increase in dimension has repercussions on the capacitance of the interconnect and hence on the signal integrity.

At either end of the barrel is a round **pad**, which brings the conducting barrel of the via onto the top and bottom layers of the PCB. In addition, similar pads are used to make the connection to traces on internal PCB layers. Typically the pad diameter will be between 10 and 18mils wider than the barrel diameter, to ensure a certain tolerance in the take-up of copper (the drilling process). More exactly the manufacturers want to prevent that even if the drill is off-centre, there will still be sufficient copper to ensure the connection as shown in the next figure 3.2. Case A is a well-centered barrel, B is off centre but within tolerance and C is an out of tolerance drilling.

![Fig. 3.2. Drilling tolerance requirements](image)

When a via passes through a power or ground plane there must be sufficient clearance around the barrel and its pad so that it cannot make accidental contact with the plane it passes through. This clearance is called the **anti-pad**. Typically an anti-pad will offer 10mils of clearance around the drill hole. Now this causes problems for the traces routed near the via barrel, since the anti-pad eats into the area of copper that these traces are referenced to for their controlled impedance. For the exemplification, let us consider the case in the next figure 3.3., where two traces - A and B - are asymmetrically disposed next to a via.
Fig. 3.3. Reduction in return path due to anti-pad

The trace A is sufficiently far from the anti-pad gap, having sufficiently copper in the ground plane. The return current through it is then symmetrically distributed along the axis of the trace and it is spreading over about three times the trace width in the reference ground plane. In the case of trace B, it is as close as it can be located to the via without overlapping the anti-pad. The return current now has some unknown distribution, which by definition will have some possibly negative effects on signal integrity.

A secondary issue affecting the signal integrity is the fringe capacitance. Let us consider two situations, as presented in figure 3.4., in order to demonstrate the anti-pad effects. For the case on the left, we opt for a minimal mechanical tolerance, which will add a significant capacitance between the ground plane and the via. This will affect the signal integrity! For the case on the right, expanding the anti-pad size, the capacitance is reduced. However having the ground plane now further away from the via, there is less continuity for the return signal path (yellow arrows).

Fig. 3.4. Anti-pad effects
To reach the stack-up definition for a PCB, we have to consider the routing constraints, the number of traces that need to be laid, their constraints and their impedance control requirements.

### 3.2 Routing Constraints

For the routing constraints exercise we study the most demanding case which was the first architecture that had been proposed for the 10Gb/s Ethernet switch design. We were considering this case even after that proposal was suspended, on the grounds that it push the limit further as possible, setting an upper bound for what ultimately may be required.

The original topology is shown in the figure 3.5. Each 10Gigabit Ethernet port can be accessed by 8 links from each of the two switch fabric boards. Each link consists of four traces, which carry two differential pairs running at 3.125Gb/s in each direction. Each switch fabric has 16 such ports.

![Switch fabric based on IBM fabric](image)
So one switch plane carries $16 \times 8 \times 4 = 512$ traces and with two planes that makes a total of 1024 traces on the backplane. This is not counting at least one and possibly two PCI busses for control as well as other interconnects for power monitoring.

One way to layout this out on the backplane is shown in the figure 3.6. We have placed the two switch planes in the centre of the board with eight ports on each side of them. Each port is equipped with three connectors. The top and bottom ones carry interconnects and the middle one, power and extra ground. No other logic excepting traces is routed on the backplane. First we need to see if they will fit at all before considering the ancillary logic.

The red lines in the figure bottom indicates where groups of eight links (times four traces) need to be routed between the ports and the secondary switch plane. This is the most complicated route since it has to pass through the primary plane before reaching the secondary one. It is clear even at this level that there is unlikely to be any routing path that avoids other connectors. At least one and probably more connectors will have to be crossed by traces that don’t make connections there.

![Fig. 3.6. Backplane of candidate architecture](image)
The cross-over area between the two switch planes is the obvious bottleneck, so we consider how many routing planes will be needed for this.

![Diagram of switch fabrics](image)

**Fig. 3.7. Routing through the connectors**

The figure above shows a horizontal slice through the switch fabric connectors. We assume four pairs per row and the ground pins are not shown for clarity. We choose to only route one pair between adjacent pin rows so it is obvious that four layers will be needed to route the escapes from each row, or five if we had chosen a five pair connector.

Our previously choices –HM-Zd and GBX– are four pairs connectors. The first choice, the HM-Zd connector, has a lower pair/inch ratio than GBX and will supply 40pairs/inch. Thus 512 traces is 256 pairs or 6.4inches of connector to carry the interconnect. A 6U board has just less than 9 inches of connector space so there is no major problem with providing enough pins for the above mentioned pairs of traces. However, we did not yet take into account the other services (such as control and power monitoring), which have to find room as well. Adding all their corresponding traces, we might finally run out of available space. The solution for this is either make the board a bit taller or use a higher density connector, such as our second choice GBX.

### 3.3 Impedance controlled traces

Reaching an optimum dimension for impedance controlled traces is not a trivial exercise since it involves making continual tradeoffs between what is ‘desirable’ and what is ‘necessary’. We will use the results of a stack-up based calculator, recognizing that while
this is not accurate enough for the final design, it is good enough to show the best way to go at this moment.

The first assumption is that we do not need to resort to exotic dielectrics to achieve the desired results. This has been studied for long in the industry and the consensus is that it is necessary to use special dielectric materials only for rates of 5Gb/s and upward. This limit has been upwardly revised over the past two or three years essentially as a result of the more advanced techniques of pre-emphasis and equalization used in modern backplane transceivers and receivers.

As a consequence of the above assumptions, we consider a poor definition of the permittivity. For our first pass impedance calculations we assume a dielectric constant of 4.3, but what is it really? For FR4 the dielectric constant varies as function of temperature, frequency and resin content. Typical values can range from 4.8 to 4.1. Tests for different frequencies on one particular sample are shown in the following figure.

Fig. 3.8. Variations of permittivity with temperature and frequency
The graph in the next figure shows the effect of such a range of values on the differential impedance of a stripline built in FR4.

![Graph showing the effect of permittivity on differential impedance.]

**Fig 3.9. Variation of Z\textsubscript{diff} with permittivity**

So if we base our calculations on an arbitrary permittivity value of 4.3 our actual manufactured differential impedance could be anywhere within +2.5% to -4.5% of nominal, depending on the dielectric stock used and the frequency of operation of the electrical signals.

This does not mean that we should abandon any hope of accurately predicting the impedance of our transmission lines. The theoretical analysis of the final product must be done in conjunction with the board manufacturer who knows the materials he is working with and can adapt the nominal dimensions we supply to achieve the desired result.

### 3.4 Trace Constraints

A typical routing through a slice of a generic connector footprint is shown in the figure 3.10.
The design width of the traces has some conflicting constraints, as follows:

- The two traces shall fit between the connector pins without overlapping the anti-pad footprints.
- The traces shall be as far from adjacent vias as possible to reduce cross talk effects.
- The traces shall be as wide as possible to reduce skin effect losses over long distances.
- The traces shall be as narrow as possible to minimize the distance between the reference planes and thus the overall board thickness.
- The traces shall be as close together as possible to ensure equal traces lengths over the whole routed path.
- The traces shall be as far apart as possible to maintain a differential impedance twice that of the single ended impedance.
However, in practice the degrees of freedom are quite limited for any given connector. Let us consider in the following table the critical dimensions for the HM-Zd connector (our first choice as discussed in chapter II).

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Pin Pitch</td>
<td>98.425 Mils</td>
</tr>
<tr>
<td>Via Hole Diameter</td>
<td>27.6 Mil</td>
</tr>
<tr>
<td>Anti-Pad Diameter</td>
<td>50 Mils</td>
</tr>
<tr>
<td>Max Routing Channel</td>
<td>48.425 Mils</td>
</tr>
</tbody>
</table>

The smallest trace width for volume production is 5 Mils and the minimum trace to trace separation is also 5 Mils. We would also prefer a trace to anti-pad spacing of 5 Mils. This then leaves us options of 23.425mils (from trace extremity edge to the anti-pad) up to 33.425mils (from trace extremity edge to the via hole).

Now we have to consider the design impedance. The accepted standard for this technology requires a Zo (characteristic) impedance of 50 Ohms and a Zdiff (differential) impedance of 100 Ohms. It is possible to achieve this impedance by building a stack to give 50 Ohms Zo and then running the two traces as far apart as possible. The main problem with that is the difficulty of maintaining equal length traces and managing the changes in impedance occurring as these traces are separated away after the transmitting device and then brought closer together at the connector or the receiving device. Consider now what happens as we bring the traces closer together. They will have very little effect on each other until they get within a distance of about three line widths of each other. At that point the differential impedance starts to drop because now the differential coupling is not just between the traces and their reference plane, but also between adjacent traces. On the other side, running traces close together will minimize EMI effects since their radiated fields will cancel out. This is the case if the traces have to pass any distance on an external layer of a PCB, between, say, the chip on the daughter card and the connector to the backplane. It is preferable then to define and maintain from source to destination a fixed spacing, having the traces close together to fit in the routing channel and minimize the EMI.
As the differential impedance drops by having the traces close one of each other, we could increase Zo by altering the geometry and this would restore 100 ohms Zdiff. However, the cost to pay for this is a much thicker backplane and an increasing mismatch for the single ended impedance. Alternatively we can use very narrow traces. The spacing between them is less than for wider traces for the same Zdiff and the final thickness of the board is much less. Figure 3.11. shows the options.

![Graph showing spacing needed for design impedance](image)

**Fig. 3.11.** Variation of Zdiff with trace width and trace separation

The figure above shows that even for the narrowest trace it is difficult to reach 100 Ohms within the practical limits imposed by routing through a connector. As it can be seen the spacing between traces is going from a minimum 10mil to a maximum 42mil. Some of these values are not at all ‘realistic’, since the whole routing channel is 48.425mil for HM-Zd connector. By processing the graph 3.11. and taking into account only the ‘realistic’ values for the HM-Zd connector, we obtain the next graph.
The dotted line in the graph above delimits the routing space available for HM-Zd. The figure 3.12. shows the remaining options when we take the available routing space into account. Corresponding to it, true 100 Ohm Zdiff is only possible with 5 or 6mil traces. We can be within 1% for traces up to 9mils wide, 2% for 10, 3% for 11 and 4% for 12.

The following question comes then naturally: ‘How can we choose the optimum value for the trace width, that it will determine a close as possible 100 Ohms Zdiff?’ At this point in the procedure it often comes down to a question of personal preference. The designers with great faith in their analysis tools will favor the narrower traces to achieve the design target of 100 ohms, the ones with production cost experience will favor the wider traces that are cheaper and can more reliably be reproduced. The manufacturer studies show that the narrower the traces are, they are much more exposed to tolerance errors (only to mention the variation of the dielectric thickness or those of the trace width) than the wider traces.
In addition, most of the current in a transmission line at high frequencies is carried in the outer skin of the conductor, which is a layer typically only a few microns thick. The effect of this on the eye opening is shown in the following figure 3.13.

![Figure 3.13. Signal loss for different trace widths](image)

For a 25 inch trace at 5 Gb/s there is 10% improvement in eye opening for a 10mil trace compared with a 5mil one. For our application, at only 3.125 Gb/s this will not be so serious, nevertheless the combination of other effects as well points clearly to the wider trace being preferred.

The overall magnitude of all manufacturing errors, such as dielectric thickness or trace width variation, is alarmingly high. However, in practice the two effects are cancelled out in the production stage where the manufacturer is given the target controlled impedance to work with plus a test coupon on the board to measure the finished result. This process is of course made easier (read cheaper) if the wider traces were chosen from the beginning. It does however put discussions about 100 Ohm versus 99Ohm impedances into perspective and we should also consider the effect these choices have on the final thickness of the backplane.
4 Connector Pin-field Simulation

4.1 The problem

Given the findings of the AMP tests on connector crosstalk (see chapter II), and concerns about the routing constraints through connector pin-fields, it was felt that simulation of the problem was a prudent precaution. Consider the subset of a connector pin-field shown in figure 4.1. below. The two traces of a differential pair are passing between vias that carry other, equally differential signals. The traces will have an influence on the vias and vice versa. Our concern is not particularly for the victim via since the zone of interaction is so short and noise coupled to one via is very close to the noise coupled to the other via of the pair. Hence most of the coupling is in common mode and the receiver is largely insensitive to it. However noise coupled from via to the nearest trace of the differential pair is more than the noise coupled to the farther trace. Thus differential noise is induced into the trace pair. In a complex backplane the traces pass between many such vias, eight pairs per connector in our design case (using HM-Zd connectors), and the suspicion is that the noise will accumulate.

Fig. 4.1. Differential traces through a connector pin-field
4.2 The tools

We investigated several possibilities to use for system simulation. *Specctraquest* is the CADENCE signal integrity tool including a 2D solver with advertised ‘capabilities’ for including the effects of vias. It is well integrated with the CADENCE design (Concept HDL) and layout toolset (Allegro). We calibrated our results against published TDR results for different trace widths and printed circuit stack-ups with excellent correlation for both far and near end crosstalk (FEXT and NEXT). When we included vias in proximity to the test traces we could detect no sign of coupling, even when we violated the design rules and moved them very close together.

We concluded that we could use Specctraquest to extract SPICE models for the ‘standard’ PCB structures, such as long transmission lines. However, for the smaller three dimensional structures we would have to use a more complex field solver. We could then join the two models together in a SPICE simulation deck. We evaluated several 3D field solvers, such as Microwave Studio from CST and Maxwell Q3D Extractor from ANSOFT, but none of these two gave sufficient accurate results for our requirements. Finally we settled on ANSOFT 3D full-wave field solver – HFSS. It has a very steep learning curve but this was accepted as the cost of having trustworthy results for the high speed connector regions on the backplane.

With HFSS we quickly ran into serious limitations. We started with a small section of a connector with two traces running between two rows of pins but the analysis ran for more than 24 hours and still failed to converge on a stable solution. We then spent a long time iterating downwards until we could identify a tractable level of complexity. Along the way we discarded design file import in favour of drawing the structures with two dimensional conducting planes rather than three dimensional solid conductors. We also ran into problems with the definition of boundary conditions and maximum volume of the structure under investigation. This gave rise to a set of inconsistent results so we were forced back to a 'bottom up' approach where we would model the simplest structures and verify their results against the literature before moving on to more complex structures. Even so, by the time we reached a structure like the one in Fig. 4.1., we were again
experiencing compute times in excess of 36 hours with only limited signs of convergence. Nonetheless we have extracted some useful results which are presented below.

4.3 The single via

The figure below shows the geometry of a simple via hole used to check the full-wave field solver HFSS, to see if we could obtain the expected modelling results. This was principally an exercise in choosing the correct boundaries for the bounded volume and the optimal dimensions so that the 3D passive geometry could be computed in a reasonable time and converges quickly to a stable solution.

Fig. 4.2. Simulation of a single via through a ground plane

It was difficult to find in the literature a reference characteristic for the PCB via. We had to accept that a short signal via between one or two ground planes has very little effect on transmitted signals. The real problems don’t appear until there are backplanes of 250 mils thickness or more. We therefore accepted that a ‘generic’ reflective curve (the red curve in the left hand picture), rising up to about –5dB reflection for frequencies around 15 to 25 GHz, was sufficient proof of concept. The following figure 4.3.a shows our simulated result and in figure 4.3.b is shown a reference design from Ansoft, who worked on a design for Finisar [9]. This appears to indicate that our results are going in the right direction.
4.4 Coupled VIAs

The next figure shows the HFSS geometry of a pair of vias connecting traces on the top and bottom layers of a PCB. Each trace is referenced to a ground plane. The two vias shown to the left and right of the signal pair are ground vias interconnecting the ground pairs. This is the subset of a typical high speed connector footprint.
This geometry was simulated in HFSS and the results shown below in figure 4.5.

![Fig. 4.5. HFSS S-parameter results](image)

As in the single via case (fig.4.3.), we note a reflection free increasing loss up to –5dB above 15GHz. A similar geometry has been considered by Ansoft for the Finisar design. They are calling this geometry GSSG (‘Ground-Signal-Signal-Ground’), due to the corresponding vias functionalities. The S-parameter results are taken as a reference and they are presented in the figure 4.6.
Fig.4.6. Reference GSSG geometry S-parameters

We don’t have to forget that our geometry is not identical with the one used by Ansoft and the smallest modification in the geometry is reflected in the S-parameter values. We have been using the HM-Zd connector with the dimensions given by the manufacturer, while Ansoft made an optimization study for the via parameters (anti-pad, pad, via hole etc.) and presented the reflection and transmission results for the optimal dimensions. However, our results and Finisar results are similar in a way, which indicates that our results are going in the right direction.

### 4.5 Via trace crosstalk

In this test we are now close to the desired structure as shown in figure 4.7. The geometry is a region of a HM-Zd connector with 8 Mil differential traces located on the top and on the bottom layer of the PCB and a pair of traces passing between the connector pins. The figure below shows one frame of a current density vector animation from the differential stimulation of the two vias and the currents induced in the adjacent traces due to the coupling effects.
Fig. 4.7. Differentially driven vias adjacent to differential tracks

Figure 4.8. below shows three signals. The uppermost one in dark blue shows the far-end cross-talk induced in the farthermost trace. The light blue signal shows it for the closest trace. The difference between the two is shown in the black signal.

Fig. 4.8. Generation of differential cross-talk due to via proximity
The amplitude of the signal in black effectively confirms our concern that differential noise is induced by this routing topology. In order to highlight these effects, we plotted below as an ‘eye’ diagram the far-end cross-talk magnitude against one bit time of the aggressor signal.

![Fig. 4.9.a) Differential FEXT](image1) ![Fig. 4.9.b) Differential NEXT](image2)

The magnitude of the far-end cross-talk effect is of the order of 0.002% of the aggressor signal, which is much lower than expected. So low in fact that there was concern that the field-solver or the simulator was not being used properly and that an incorrect evaluation would result. In order to exaggerate the coupling effect between the vias and the traces, we moved the differential stripline pair from its default position which was 19.6 mils from the via. It was moved to the closest distance that design rules allow, 5 mils separation between via and the trace edge. In the figures 4.10.a) and 4.10.b) are presented the cross-talk from the via to both the nearest and furthest trace. In the figure 4.10.a) we show the results from the default position of 19.6 mils from the via. At the frequency of interest for our signals (10Ghz) there is only about 40dB of cross-talk even for the closest trace (the signal in red), so the differential noise is extremely low, as expected. In the figure 4.10.b) the results for the closest separation (5mil) are shown. There is the same difference between the two traces, but the overall signal has risen by an order of magnitude from -40 to -20dB.
This demonstrates that the field-solver and the simulator are ‘seeing’ the correct effect and confirms that it is indeed a very small one. However, the full signal path consists of the source circuit on a daughter-board driving a short trace segment between it and the backplane connector. Noise can be picked up from adjacent vias in this connector. Then there are the traverses through one or more successive connectors on the backplane and finally, the signal is routed off the backplane and through a short trace it arrives to the receiving chip on another daughter-board. Figure 4.11. below shows the physical representation of the daughter-board to backplane connector (fig. 4.11.a) and the traversed backplane connector in the figure 4.11.b).
The above geometries have been electromagnetically modeled in HFSS and their equivalent full-wave models have been further used in an HSPICE simulation. The HSPICE analysis considers a ‘realistic’ situation when the signals from the transmitter board pass via the backplane to the receiver daughter-board. Figure 4.12. shows such a channel circuit.

**Fig. 4.12.** PSPICE circuit of connection between successive connector pin-fields

Starting at the transmitter (TX_model) on the upper left side, the signal traverses two inches of FR4 PCB trace before arriving at the connector that leads it to the backplane. This connector is described as a four port device. Two ports handle the ingress and egress of the signal and the two other ports carry signals that pass through adjacent vias. These vias are stimulated by other transmitters just as would be the case in a fully populated connector. The exiting signal traverses 10 inches of FR4 and then passes through the pins of a connector it is not itself connected to. This model is a six port device: the ingress and egress ports for the signal plus four other ports representing stimulated vias. After another 10 inches of FR4 the signal exits to the receiver’s line card through the final daughter-board connector. This is also a source of noise to the signal. The eye-diagram of the resulting signal is computed at the input to the receiver (RX_model).
For the driving and the receiving circuits, we used proprietary models received under NDA from a well-known manufacturer. Measurements of signals at this level are done with ‘eye-diagrams’ than one-off pulse analysis. There are several reasons for that, as follows. Firstly, real data transmission does not include the equivalent of an isolated single pulse. Transitions from zero to one and one to zero are bounded by the encoding algorithm which limits the number of bit periods a signal can remain at a one or a zero. Secondly, the encoding sequence can generate successive bit patterns which are more or less susceptible to interference. Thirdly, chip manufacturers define their tolerance for correctly discriminating a ‘one’ or a ‘zero’ in terms of the opening of the eye diagram of the received signal.

It is current practice to define a pseudo-random bit pattern (PRBS) and use that for characterizing a transmission channel. We use such a PRBS bit sequence at the design bit rate of 3.125 Gbps to stimulate the transmitter circuit. We first measure the received signal after the simplest case of sending from transmitter to receiver with just the daughter-board connectors and 14 inches of backplane between them. This is then compared with the same length signal but this time traversing three other backplane pin-fields. The pin-field models used were set to the limiting case of 5mils distance from trace to vias. The results are shown in the figure 4.13.
In blue is shown the simpler case and in red the more complex one including the extra pin-field regions on the backplane. There is just a small difference in the maximum eye opening from 870mV without extra pin-fields down to 810mV with the pin-field models added on the backplane, a loss of 6%.

Considering now the crosstalk measurement we ran the simulation for the two cases, one with close proximity to the aggressing vias (5mils) and one at the design distance (19.6 mils). We used the same stimulus, a PRBS stream at 3.125Gbps.
Fig. 4.14. Differential noise after traversing three backplane pinfields.

In red is the differential cross-talk induced in the traces by the aggressing vias located at only 5mil distance, while for the black signal the vias are 19.6 mil away from the traces. The level of crosstalk appears to be less than 1% (0.6%) when the stripline is located at 19.6 mil from the vias and increases to about 3% for the limiting case of 5mils distance. What this means in terms of the eye diagram is shown in the following figure 4.15.
The eye opening difference between the 'close' (5mil) and the 'normal' (19.6mil) position of the trace is 30mV, which corresponds to an eye closing of about 3.3%.

This confirms that while there is certainly a measured effect that causes the injection of differential noise the magnitude of this effect is so small as to be negligible. The remaining simulated eye-opening has more than enough margin to sustain error free operation at the design speed of 3.125Gbps. A stated eye opening of 400mV is called for in the specifications which are themselves highly conservative and this eye opening exceeds that value for over 50% of the bit period.
5 Advanced TCA

As our design for the 10Gb/s copper backplane was going on, the risks of developing a proprietary backplane became more evident for more than one manufacturer. In particular the PICMG (PCI Industrial Computer Manufacturer Group) group of over 100 participating companies had to address the fact that the Compact PCI market had fallen between two targets. It had been too late to capture a significant portion of the embedded VME market and it was too slow, from a technical point of view, to capture the future telecoms infrastructure market. They faced this squarely by dropping the traditional bus-based architectures and embracing a point-to-point switching architecture that would have the possibility of scaling into the next generation of telecom equipment.

The result was the specification of Advanced TCA (Advanced Telecom Computing Architecture) proprietary standard. The base specification for the new Advanced TCA family (Revision 1 of PICMG 3.0) was adopted in January 2003. It defines the physical (rack and chassis form factors, electromechanical specifications of the board) and electrical characteristics (backplane connectivity, power, cooling, management interfaces) of a switch based platform delivering a scalable solution for next generation telecommunications equipment.

We studied the nominal specification of the new standard and compared them with the guidelines we had already established in our studies, previously presented.

Fig. 5.1. Recommended stripline dimensions for 100 Ohm Z\text{diff}
This stack is defined for FR4 having a relative permittivity of $\varepsilon_r = 4.0$, which we consider to be optimistic. It is left up to the manufacturers to produce the required 100 Ohms using this as a starting point. The trace specification in Advanced TCA standard is looser than the solution we had recommended, but compensates by defining an overall tolerance that is quite realistic for FR4 technology. Comparing this with our own estimations, and using the same stack-up calculator as before, then we obtain:

- at $\varepsilon_r=4.0$ $Z_{\text{diff}} = 94.5$ Ohms and $Z_0 = 51.6$ Ohms

- at $\varepsilon_r=4.3$ $Z_{\text{diff}} = 91.1$ Ohms and $Z_0 = 49.8$ Ohms

In practice either of these choices still meets the specification which states that “the differential impedance of the backplane and board serial links for the Base and Fabric Interfaces shall be 100 Ohms ± 10%.”

The connector of choice is the Tyco/Erni HM-ZD 2.0mm high speed connector in a 4x10 differential pair configuration, which was also our choice.

One area of concern is the new anti-pad specification. This is more generously dimensioned than the connector specs originally called for (50Mil diameter) and will have the effect of reducing the routing channel from 48.425 Mils in our design to 34.425 Mils in theirs. This is one of those compromises that we mentioned earlier. By increasing the anti-pad clearance there is a corresponding reduction of via to ground or power plane capacitance and therefore, less signal loss is injected into via. However, the traces routed through the connector pins don’t have anymore an infinitely wide GND or power reference plane, as it is assumed by all the formulas, and that will cause an un-known distribution of the return current. The next figure shows the resulting difference in stripline configuration between the ATCA specification (A) and the model (B) we had developed.
Fig. 5.2. Comparison of stripline geometry’s

The drawing is to scale and shows that the increasing in via to plane proximity in Advanced TCA standard has come at the cost of reduced ground plane coverage for the return current. This will affect as well the impedance of the traces as they pass through the via pin-field. It must be assumed that the Advanced TCA simulations took this into account and that the trade-off was considered as acceptable.

5.1 Conclusion

The differences between our model and that adopted by the Advanced TCA group are so small as to make building a proprietary backplane a poor economic choice. Our backplane design therefore will be based on the Advanced TCA specifications.

We have confidence in the design simulation work done by the Advance TCA consortium, which is obviously extensive and professional. We have less confidence in our 10Gb/s prototype, as experience shows that it can take a couple of iterations with any given design to obtain the design spec performance. Problems can be cause by the PCB manufacture, but as well by the first generation of switch fabric chips or some
combination of out of spec transceivers and out of spec backplanes. On the other hand, the same problems of fabrication control that we have ourselves are encountered by the Advanced TCA, since they are prototype backplanes with no past history of application successes. In order to be able to subtract at least the backplane from the set of things to worry about, we will be studying in the near future a backplane ‘exerciser’ (the future work of the ‘Backplane Tester’ system) that will fully load a number of adjacent slots and provide ‘real’ conditions traffic across the backplane. With the new ‘Backplane Tester’ we hope to achieve a degree of confidence to distinguish between the two kinds of problem.
In the present paper we have addressed the concerns of building a copper backplane to carry switched Ethernet traffic at 10Gb/s and identified the main engineering points of best practice.

We have examined different connector options from several manufacturers and we have selected the most appropriate ones. Our connector choices were done in favor of Tyco HM-Zd and Teradyne GBX.

We have investigated the routing constraints for the most demanding proposed architecture and found solutions. The influences of the PCB materials and of the routing on the trace controlled impedance have been also considered. We have finally derived an optimum printed circuit board design taking into account the high frequency effects and the high routing density. The tradeoffs between conflicting criteria have been described and evaluated.

We have worked systematically through the development of an analytical model of the connector pin-field regions on the backplane, which includes the crosstalk effects from adjacent structures. We found the effect to be less than expected and conclude that the backplane design will meet its requirements.

We have investigated the proposed Advanced TCA infrastructure and found that it meets the same design criteria that we had established. The differences between the design we proposed and those adopted by the Advanced TCA group were considered as small as to make building a proprietary backplane a poor economic choice. Our copper 10Gb/s Ethernet backplane design therefore will be based on the solution proposed by the Advanced TCA standard.
The experience shows that it can take a couple of iterations to any given design to obtain the spec performance. In this context, our 10Gb/s Ethernet switch is a prototype and the Advanced TCA standard does not have such a long past history either to have been proven in the field application successes. The errors could be anywhere in the silicon or in the backplane. Therefore, in order to subtract at least the backplane from the set of things to worry about, we will be studying in the near future an Advanced TCA backplane ‘exerciser’ that will fully load a number of adjacent slots and provide 10Gb/s traffic in ‘real’ conditions across the copper backplane.
7 References

“Creating High-Performance, Cost Effective PCBs”, Ansoft Corporation presentation

“Strategies for High Density and High Speed Packaging”, Ansoft Corporation presentation

“ERmet ZD High Speed Differential Hard Metric Connector System”


