WHITE PAPER


# HIGH DATA RATES OVER THE VPX INFRASTRUCTURE

Revision 2

# HIGH DATA RATES OVER THE VPX INFRASTRUCTURE
## INTRODUCTION


The VPX VITA 46 has been one of the first modular computer open standards to define a connector and backplane Infrastructure allowing data transfers at rates in excess of one gigabit per physical channel. This allowed embedded computers of such form factors to implement the same state of the art interconnect technologies that currently exist in the servers and consumer PC area. Thus, users of the VPX could benefit from the same level of performance for input/output, graphics, and more generally computer data exchanges, at least up to an acceptable wattage for the targeted environment. And the ratio of I/O throughput to CPU processing speed ,with their associated software architectures, can remain coherent.

This is why it is important for the VPX technology to keep up with the latest data rates available in the computer industry. After having successfully addressed the 5 Gbits/s level for the PCI Express® gen2, VPX is now ready to adopt the higher data rates required by the latest version of two fundamental protocols: PCIe gen3 at 8 Gbits/s and Ethernet at 10 Gbits/s.

The first part of this article concentrates on the possible technical issues and workarounds that could be encountered when implementing 10 Gbits/s rates over a VPX copper backplane. The second part details Kontron's achievements in these areas in terms of qualification of the technology and available architectures and products. Readers only interested by the results and the new possibilities of transferring data at 10 Gbits/s on VPX could jump directly to the second part.

# CONTENT

# HIGH DATA RATES OVER THE VPX INFRASTRUCTURE

## TRANSFERRING 10 GBITS/S OVER VPX BACKPLANES

### Normative context

Apart from the VPX VITA 46 standard defining the mechanical and the utility structures, plus its sub-standards focusing on each individual physical protocol, OpenVPX VITA 65, built on top of the VPX norm, is the subsystem level standard specifying how the modules and the backplanes interoperate, in which protocols and with what pin assignments. The current ANSI ratified OpenVPX version is "ANSI/VITA 65-2010 (R2012) OpenVPX™ System Specification", approved in 2012.

This R2012 version of VITA 65 defines a nomenclature for daughter modules and backplanes up to a rate of 6.25 Gbauds which is sufficient to run protocols such as Serial RapidIO and PCIe gen2. The next revision currently under consideration includes the PCIe gen3 and the 10G Ethernet up to a rate of 10 Gbits/s. It still relies mainly on the Tyco-Electronics connectors specified from the beginning in VPX VITA 46. Other connector manufacturers are offering interesting alternate solutions which are compatible at the PCB footprint level on the module and backplane side, but are not compatible with the original connectors: this means that you cannot plug a module with one brand of connector in a backplane with another brand. One exception to this are the backplane connectors from EPT which Kontron has determined to be compatible and equivalent in term of signal integrity with Tyco connectors.

None of the VITA 46 or VITA 65 standards specify the requirement for high speed interoperability of modules and backplanes: timing, eye openness, jitter budget, and more general electrical channel characteristics. Rather they point to the electrical requirements inside the existing base protocol standards such as PCI Express® or Ethernet 802.3, and also to the VITA 68 standard, currently a draft. The VITA 68 standard attempts to fill the gap between signal integrity requirements defined at the chip level or at dedicated form factors, and the necessity to standardize the interoperability requirements at the card level for the VPX form factors. The VPX Compliance Channel VITA 68 standard mainly takes as a reference, the Ethernet 802.3 10Gbase KR standard

### The IEEE802.3 KR Channel

This IEEE standard defines 10 Gbits/s Ethernet transfers over a backplane at the raw speed of 10.3125 Gbits/s (the physical data encoding is of type 64b/66b). In part 3, annex 69B of IEEE 802.3 specifies the channel characteristics for a single differential copper pair, one per direction, impedance of 100 Ohms +/- 10%. The said channel is defined for up to 1 meter of trace length, from the pin of the transmitter circuit to the receiver device, excluding the receiver blocking capacitor. One should remark here that, contrary to PCI Express®, the blocking capacitor is located on the receiver side of the Ethernet channel, which is more adequate from the perspective of signal integrity: the signal arrives at the receiver already attenuated due the connectors having been crossed and the PCB losses, so the discontinuities generated by the capacitor are less critical.

A compliant channel is required to be better than the specified limits in terms of:

- Insertion loss, a curve measuring sinusoidal signal losses as a function of its frequency. This is the magnitude in dB of the S21 S parameter, the smaller is the better
- Fitted attenuation, the least mean square line of the previous curve within 2 defined frequencies
- Insertion loss deviation, the allowed deviation, as a function of the frequency, of the actual insertion loss from the fitted attenuation
- Return loss, a curve measuring the sinusoidal signal losses being echoed back to the transmitter as a function of its frequencies. This is the magnitude, in dB, of the S11 S parameter, the higher absolute value is the better.

- Insertion loss to crosstalk ratio as a function of frequency

On figure 1, extracted from annex 69B of the IEEE 802.3 standard, the insertion loss and return loss parameter limits for 10GBase-KR Ethernet signals are shown.

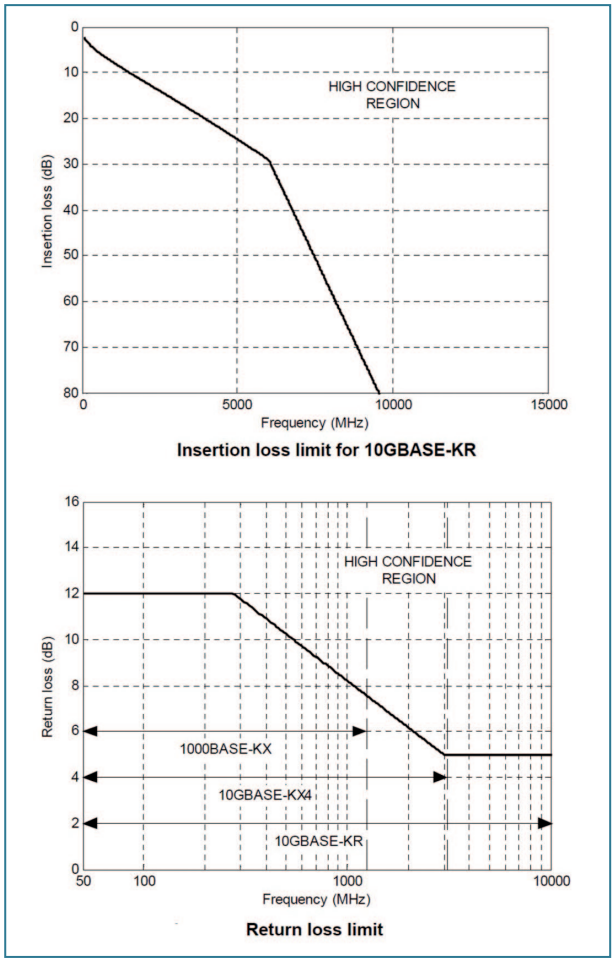


Figure 1: IEEE 802.3 10GBase-KR Insertion loss and return loss limits

## Stubs are the first limiting factor

Among the several factors limiting the channel performance, stubs, these short undesired conductive tracks not belonging to the main signal path, have a direct negative impact. At 10Gbits/s, with signal traveling in copper at close to half the speed of the light, each bit transiting over the backplane is separated from the next bit by less than 2 cm. For example, short stubs of 5 mm would typically refl ect part of the signal back to main path and to the transmitter, thereby affecting half of the length of a following transitioning bit, in addition to the losses resulting in the main direction forward.

The connector contact itself, depending on its technology, could be a source of stubs. A rear VPX connector with populated contacts, even without an RTM attached, would also represent an unacceptable stub at multiple Gigabit/s. In fact, even a metalized via drilled in the PCB for going from one layer to another could form a significant stub: the backplane itself, typically 5 mm thick when it is designed to host front and rear VPX connectors, will exhibit stubs at the VPX connector via unless they are removed by using techniques such as blind via or back-drilling. The stub formed by the backplane via, with the back drilled and blind via workarounds are represented on figure 2.

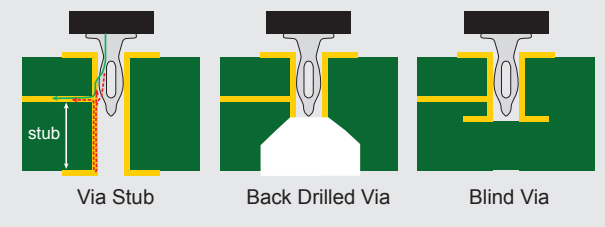


Figure 2: Via stub elimination

## Impedance mismatch inside vias

The differential impedance of the pairs carrying the data shall be 100 Ohms +/- 10% according the IEEE 802.3. Having impedance discontinuities along the path would create partial signal reflections and ringing, resulting in degradation of attenuation and return loss parameters. However, inside the via of the VPX connectors, both at the module connector and the backplane connector, the differential impedance is generally much lower, in the range of 80 Ohms and below. This can be easily explained since the impedance per unit length is given by the formula L/C.

The inductance, L, is decreased inside the via because the perimeter of the section of the via is much larger than the copper track section connected to the via. There is not much that can be done to mitigate this since the diameter of the connector via is determined from the press fit pin dimension of the connector. The capacitor, C, is generally increased due to the many ground and power planes facing the via cylinder on the different layers of the PCB. This factor can be optimized to a certain level by increasing the air gap between the via and the ground or power planes, by suppressing some ground or power planes facing the via, and also by grouping the plus and minus plane voids of the via into a single larger void.

As a summary, inside the via, the inductance is lower, the capacitance is generally higher, and the outcome is a decrease of the impedance, L/C.

Figure 3 represents the typical impact of the VPX connector via on the impedance of the differential pairs.

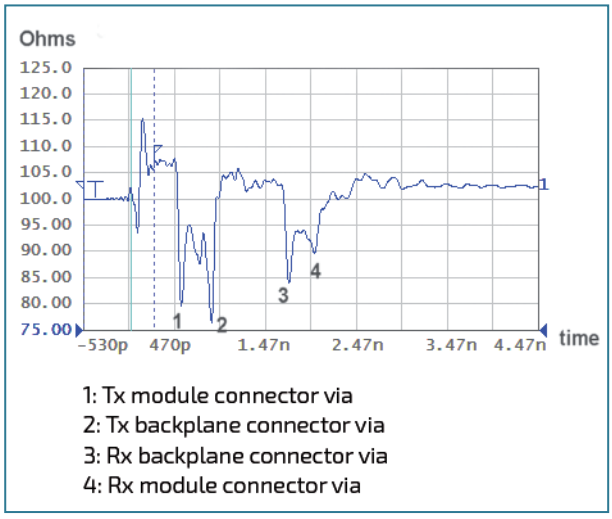


Figure 3: VPX connector via impedance mismatch

## Other undesirable effects

There are many possible other unwanted effects, but fortunately, these can be avoided by good design practices. In particular, the following potential effects require attention:

- Crosstalk: proper gap with neighbors must be applied, in addition to providing the required reference planes, especially when the attenuation is significant.
- PCB losses: selecting the appropriate PCB materials will help to avoid high attenuation that would lead to high Inter Symbol Interference. This ISI phenomenon is described as the distortions due to the different attenuation of bit sequences depending on the bit patterns: a bit pattern of 01010101 is more attenuated than 00001111.
- Length mismatch: most protocols are specified not to require a precise differential pair length adjustment, however, the plus and minus components of the pair should be the same. Any deviation is directly impacting the eye width, thus the bit error rate. Attention should be paid in particular to compensate for any length mismatch that could occur inside the connector itself.
- Return current: adequate ground reference planes shall be provided. No split planes or unwanted plane voids should be crossed that would impact the return current path. Adding stitching via may also help.
- Fiber weave effect: this effect characterizes the risk of having one wire of the pair traveling faster than the other due to an impedance variation caused by the parallelism of one wire with a fiberglass of the PCB. This effect could be cancelled either by selecting a sufficiently fine pitch weave for the dielectric, or by avoiding routing the differential pairs in parallel to the weave: this is possible on the module by applying some small angle on the copper tracks. On the backplane side, applying such an angle might be difficult, and one solution might be to apply a small rotation to the whole PCB itself.

## Common transfer clock

Some protocols like the PCI Express® were designed with the option of sharing a common reference clock between the transmitter and the receiver. The common clock simplifies the circuit design at the receiver which may only implement a DLL (Delay Locked Loop) instead of a PLL (Phase Locked Loop). Synchronization in the elastic buffer of the receiver is also simpler and faster. And from a user perspective, it allows for an option to modulate a little bit the base frequency so as to spread the radiated spectrum, thereby easing EMI compliance and also lowering the internal noise level that could impact other analog sections inside the box.

For the PCI Express®, the motherboard shall provide this common clock, and the PCIe adaptor card decides to use it or not. The PCI Express® substandard of the VPX, the VITA 46.4, mandates that all compliant modules shall have the capability to operate without a common reference clock. Depending on the PCIe silicon used, this sometimes could require the use of an intermediate PCIe switch or retimer that has this capability to operate without a common clock on the VPX side. Unfortunately, during the early days of the VPX PCI Express® standard committees, very few companies like Kontron requested that a 100 MHz common clock differential pair be provisioned in the connector pin assignment; after a vote, it was decided not to specify a common clock for PCIe over VPX, and as a result, there was no standard way to implement it. Nowadays, recent versions of the standard allows to use the REF_CLK pin at 100 MHz instead of 25 MHz to distribute PCIe clock with a radial topology. Kontron PCIe VPX modules always feature such a provision to synchronize the PCIe interface through a standard 100 MHz common reference clock.

## DELIVERING THE 10GBITS/S TECHNOLOGY OVER THE VPX

### Actual channel characterization

Using actual good quality PCIe backplanes, compliant with OpenVPX profiles like BKP3-CEN08-15.2.16-n for example, we characterized all PCIe and 10GBase-KR Ethernet channels. Figure 4 is showing the network analyzer setup and test module cards used.

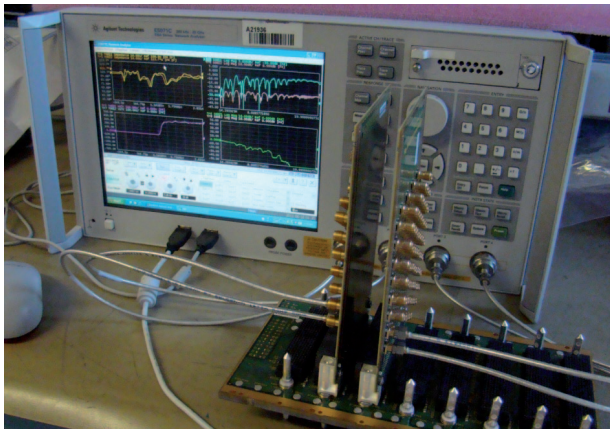

Figure 4: VPX channel test setup

Every channel proved to be compliant with the IEEE 802.3 10GBase-KR Ethernet channel definition, having a base baud rate of 10.3125 Gbits/s. The test module card allowed to evaluate the impact of different trace lengths on the backplane side, as well as on the module side.

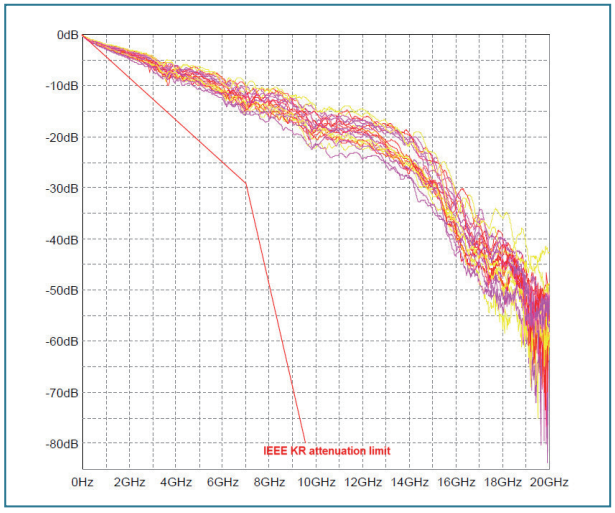


Figure 5: VPX attenuation for all channels

The attenuation criteria of the KR channel definition is probably the most easy to meet for those backplane lengths up to 8 slots of one inch. A larger backplane would use center slots for the switches, meaning that the validation is also applicable at least for an implementation of 16 slots of one inch. In figure 5, all channels attenuation were gathered on single graph.

The other compliance criteria were also met with adequate margins, some of the most sensitive having been the insertion loss deviation and the return loss for the small channel lengths.

Figure 6 shows the return loss diagram for all the channels.

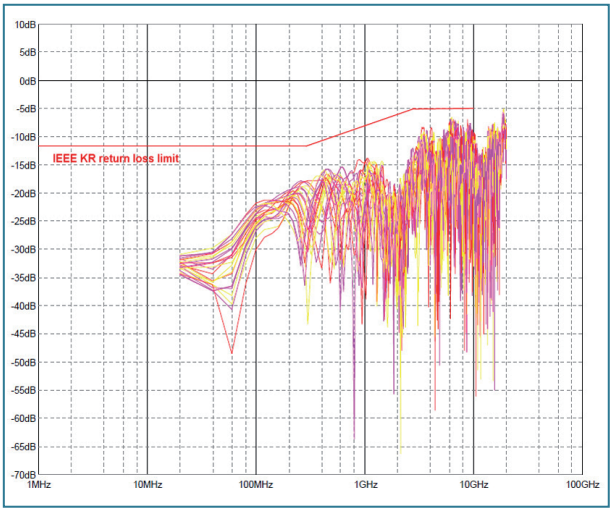


Figure 6 : VPX return loss for all channels

### Product level implementations

Kontron has developed a set of VPX modules which takes advantage of these high data rates over the VPX backplane. The gen3 PCIe protocol operating at 8 Gbits/s and the Ethernet 10GBase-KR links were successfully qualified at extended temperature environments. The measured bit error rates are order of magnitudes better than the standardized 10E-9 figure.

The 3U VPX VX3052/VX3058 CPU products are based on Intel® Xeon D-1500 technology in dual and eight core version.

Figure 7: VX3058 single board computer with 8 and 10 Gbits/s port

These models feature a VPX PCIe gen3 x8 interface configurable as one x8 port, two x4 ports or four x2 ports. A dual Ethernet 10GBase-KR interface is also offered in the same pin assignment and with backward compatibility with the traditional 1000BaseKX 1Gbits/s backplane interface. The 10GBase-KR links available on the VPX connectors can also be configured as SFP+/SFI interfaces for external fiber or copper cables. An SFP+ cage is available on the PB-VX3-400 RTM for that purpose. The VX6090 is the 6U version of the card, also hosting PCIe gen3 and 10Gbits/s Ethernet; it comes with two Intel® Xeon D-1548 CPUs and four 10G Ethernet links to the backplane. Figure 7 represents the VX3058 single board computer which is available in air cooled and conduction cooled builds.

A 24 ports 10Gbits Ethernet switch card, managed at L3 level, is also available, based on the 10GBase-KR protocol: the VX3920model. Featuring 2 SFP+ cages and one 1000BaseT RJ45 on the front, this switch can interconnect up to 24 CPU cards at 10 Gbits/s with a flat and symmetric architecture. Alternatively, it can operate in 40G Ethernet mode by grouping some or all the 24 lanes, four by four. An RTM is available to output from the rear up to one 40G Ethernet QSFP+, two SFP+ and one RJ45 1000BaseT. Extended temperature, air cooling and conduction cooling options are supported.

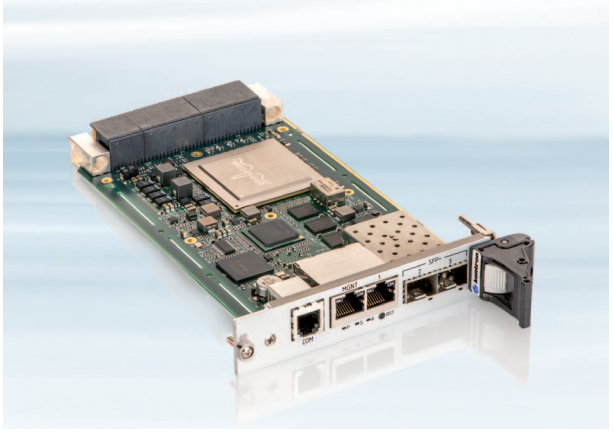

Figure 8: VX3920 10G Ethernet switch, 24 ports

## Box level architecture

The availability inside a VPX chassis of a high number of switched 10 Gbits/s Ethernet links is changing the way parallel processing can be organized. A flat CPU board interconnect could be easily setup with standard software architecture. More than 20 CPU boards could compute in parallel and exchange data at Gigabyte speed without having any bandwidth limitation due to the target board position.

A reference system has been built, called the StarVX, featuring 6 Xeon D1548 eight core CPU VX3058 boards communicating through the VX3920 10 Gbits Ethernet switch. The input data stream from the sensors is coming through a QSFP+ 40 Gbits Ethernet link, and is distributed to each processing board by the Ethernet switch. The resulting system based on a standard 3U OpenVPX backplane, BKP3-CEN08-15.2.16-n, is shown on figure 9a and figure 9b. Similar chassis and interconnect can also be assembled with 6U dual eight cores Xeon D-1548 boards like the VX6090.

Figure 9a: StarVX parallel VPX computer, front view

Figure 9b: StarVX parallel VPX computer engine, rear view

When higher speed data transfers are needed in complement, such as CPU to CPU communication or GPGPU accelerator attachment, the PCIe gen3 interfaces could offer an additional supplementary data path. The PCIe x8 gen3 interfaces can provide more than 5 Gbytes/s of usable bandwidth. The Kontron VxFabric data exchange middleware offers those transfer rates under Linux, either in TCP/IP over PCIe mode where the standard TCP socket API can be used, or in raw DMA mode. More details on VxFabric middleware can be found on a dedicated document, the VxFabric White Paper.

## CONCLUSION

10 Gbits/s baud rates over the VPX infrastructure are deploying. They allow standard protocols such as Ethernet and PCIe to run at scalable speeds defined by the number of lanes forming a port. These rates are backward compatible with previous generation protocol speeds, and can thereby be activated only where necessary in the critical portion of the data path.

Having correctly dimensioned pipes without unnecessary topology restrictions is a key factor for using standard software architecture. The CPU to I/O ratio is kept in the same range as on non embedded computer equipment. The Ethernet over PCI Express® VxFabric middleware, from Kontron, allows great flexibility in the tuning of the streams while preserving the standard Ethernet socket API.

Now what will be the next step forward, after the 10 Gbits/s baud rate range, and when ? Optical interconnect may look easier than copper to transport the biggest data streams, however radical changes of the infrastructure would probably be required for optical modular computing. We know the PCIe gen4 at 16 Gbits/s is scheduled for 2018. Will the copper technology allow to go through 2 backplane connectors at that speed ? Personally, I would bet YES.

**Serge Tissot**
Technical Strategy Manager
at Kontron Modular
Computers, France

## About Kontron

Kontron, a global leader in embedded computing technology and trusted advisor in IoT, works closely with its customers, allowing them to focus on their core competencies by offering a complete and integrated portfolio of hardware, software and services designed to help them make the most of their applications.

With a significant percentage of employees in research and development, Kontron creates many of the standards that drive the world's embedded computing platforms; bringing to life numerous technologies and applications that touch millions of lives. The result is an accelerated time-to-market, reduced total-cost-of-ownership, product longevity and the best possible overall application with leading-edge, highest reliability embedded technology.

Kontron is a listed company. Its shares are traded in the Prime Standard segment of the Frankfurt Stock Exchange and on other exchanges under the symbol "KBC".

For more information, please visit: **www.kontron.com**


### CORPORATE OFFICES

**FRANCE**
150 rue Marcelin Berthelot
ZI de Toulon-Est - BP 244
83078 TOULON CEDEX 9
Tel.: + 33 4 98 16 34 00
Fax: + 33 4 98 16 34 01
sales.KFR@kontron.com

**HEAD OFFICE**
Lise-Meitner-Str. 3-5
86156 Augsburg
Germany
Tel.: + 49 821 4086 0
Fax: + 49 821 4086 111
info@kontron.com

**NORTH AMERICA**
14118 Stowe Drive
Poway, CA 92064-7147
USA
Tel.: + 1 888 294 4558
Fax: + 1 858 677 0898
info@us.kontron.com

**ASIA PACIFIC**
1~2F, 10 Building, No. 8 Liangshuihe 2nd Street,
Economical & Technological Development Zone,
Beijing, 100176, P.R.China
Tel.: +86 10 63751188
Fax: +86 10 83682438
info@kontron.cn