Interface options for 100 GbE from a physical layer perspective

P. M. Krummrich, S. Pachnicke, M. Windmann
University of Dortmund, Friedrich-Wöhler-Weg 4, 44227 Dortmund, Germany

ABSTRACT

Due to the tenfold capacity increase of Ethernet from one generation to the next, 100 Gbps will be the straightforward next step after 10 Gbps. Demand can be predicted for the following years based on increasing internet data traffic and the lack of efficient packet based link aggregation mechanisms in the Ethernet protocol. Solutions have to be found for short haul intra-office and long haul inter-office interconnections of large routers. This contribution focuses on options for physical layer interfaces. Optical component aspects are discussed as well as transmission aspects. The selection of a suitable modulation format in combination with equalizer technologies opens a path towards robust transmission systems for this ultra-high datarate.

Keywords: 100 Gbps Ethernet, physical layer, transmission aspects, modulation formats, equalizer

1. INTRODUCTION

The foreseeable main drivers for deployment of 100 GbE are router interfaces. Traffic per port in some large internet or enterprise data network hubs is already approaching 40 Gbit/s today. Facing traffic growth rates of 50 … 100 % per year in these networks, demand for capacities up to 100 Gbit/s can be predicted for the coming years. Due to a lack of efficient packet based link aggregation mechanisms in the Ethernet protocol, adding another port does not provide a satisfactory option to increase capacity. Therefore, router manufacturers are already working on 100 Gbit/s interface cards to offer a solution for growing bandwidth demands. These cards represent a natural next step after 10 Gbit/s following the Ethernet tradition of a tenfold capacity increase for each new generation. Physical layer interface options have to be found to enable short range intra-office and long haul inter-office interconnection of routers with 100 GbE interface cards.

Standardization activities concerning 100 GbE are currently underway. The IEEE Higher Speed Study Group (HSSG) as well as the ITU-T Study Group 15 are working on solutions for a higher data-rate Ethernet. However, serious efforts are still required to evaluate the proposed alternatives from a scientific and technological point of view as well as from the system development and finally economical perspective. Especially the most appropriate data-rate for the transport layer currently contributes a heavily debated topic.

2. REALIZATION OPTIONS FROM A PHYSICAL LAYER PERSPECTIVE

The realization of 100 Gbit/s Ethernet interfaces comes with several challenges. Among them are packet buffering and processing for ultra high line rates as well as the physical layer interface. The following chapter focuses on options for the implementation of such interfaces. As optical fiber based transmission offers the only viable solution with sufficient reach for metro, regional and core networks with transmission lengths up to several hundred (long haul, LH) or even a few thousand (ultra long haul, ULH) kilometers, it was chosen as the technology being considered in this paper. Fig. 1 shows a selection of options for fiber based 100 Gbit/s interfaces in the form of a decision tree.
First deployment of 100 Gbit/s interfaces will most likely occur in large carrier offices or large enterprise data centers. The easiest way to implement these interfaces with currently available technology is to transmit multiple lower bitrate data streams over parallel fibers. For short range, i.e. intra-office links, fiber ribbon technology provides a feasible solution. The total data stream can be split by inverse multiplexing, launched into 10 fibers carrying 10 Gbit/s signals each and reassembled after detection. Transmitters and receivers for a data rate of 10 Gbit/s are mature and available from several sources. For use with fiber ribbon cables, ten of each have to be integrated in a single package, creating a significant challenge for current transponder form factors as well as thermal/power dissipation problems. Increasing the bitrate per fiber enables reduction of the transmitter, fiber and receiver count but requires development of new, higher bitrate components.

Despite the relatively straightforward implementation of this option, it is not very desirable from a logistical point of view. Some carriers prefer to avoid clogging of fiber trays by cutting off excess length of patch cables after installation and attaching new connectors on-site. The required procedures are well established for single fibers (one fiber per direction and per connector). Fiber ribbon cables would require more complex procedures and new equipment for the installation of multi-fiber connectors.

But even without the need for on-site installation of connectors, carriers prefer to avoid fiber ribbons, because they add another cable type to the already complex inventory. Managing patch cables is a challenging task in large offices due to different cable lengths. Planning the installation of patch panels would turn into a much more difficult process, if 100 Gbit/s ports were requiring fiber ribbon connections between them. Having the option to install a large number of single fiber connections and decide later, whether a given patch panel port will be used for a 10 Gbit/s or a 100 Gbit/s link, provides a much more flexible and hence desirable solution. In consequence, the single fiber option has to be considered as the preferable solution for short reach interfaces. It provides the only viable solution for longer reach links.

Another technology enabling to keep the speed requirements per channel low is wavelength division multiplexing (WDM). Again the total data stream is split into lower bitrate substreams by bitwise inverse multiplexing. In case of WDM, each one is transmitted using a separate wavelength in a common single fiber before it is detected and recombined with the other substreams in the receiver.
For short range connections, deploying another patch cable is no big issue. A dedicated patch cable can be reserved for each connection. As a result, a given fiber has to carry only one 100 Gbit/s channel. The WDM channels can be distributed over a broad wavelength range, enabling wide channel spacings. Coarse wavelength division multiplexing (CWDM) offers a cost efficient solution for this application, because filter and wavelength accuracy requirements can be relaxed.

In metro, regional and especially core networks, fibers are an expensive and usually scarce resource. Bandwidth efficiency has to be kept high to provide large transport capacity per fiber. Narrow WDM channel spacing helps to enable transmission of many channels in a given wavelength range, leading to dense wavelength division multiplexing (DWDM). Transporting a 100 Gbit/s service in a multiple channel group together with channels from other services in a single fiber is still an option, but not a desirable one. It provides less bandwidth efficiency than other options and it increases required port count in transparent optical networks.

A 100 Gbit/s service transported using N wavelengths needs the same number of available ports in an optical add drop multiplexer (OADM) or photonic cross-connect (PXC), whereas a service transported with a single wavelength needs only one. As the internal complexity of OADMs and PXC grows overproportionally with increasing port count, requiring multiple ports per service is not a desirable solution. In addition, guaranteeing that the channel group belonging to one service will follow the same path through the network increases the complexity of the network management.

Even if the group of channels carrying the substreams follows the same path through the network, there will be propagation delay differences between the channels, for example due to chromatic dispersion. Special measures have to be taken to keep these differences small. Additional measures are required in the receiver to resynchronize the bit streams. Resynchronization increases the complexity of the receiver, which has an impact on the cost. Resilience also contributes an issue, as any failing wavelength in the channel group potentially interrupts the whole connection, leading to an increased outage probability.

In consequence, transport of 100 Gbit/s services should not require more than one wavelength in metro, regional and core networks. For short range connections, CWDM can be considered offering a viable option in case it can be realized at sufficiently low cost.

The third decision deals with the number of bits per symbol. Transmission of one bit per symbol, especially in the form of amplitude shift keying (ASK) non-return to zero (NRZ) electrical time division multiplexing (ETDM), provides the most straightforward approach for 100 Gbit/s interfaces. It is also the most challenging one with respect to high speed electronics and component bandwidth requirements.

Optical time division multiplexing (OTDM) enables relaxing the speed requirements for electronic components by providing multiplexing and demultiplexing functions in the optical domain. Bitrates of 100 Gbit/s can be realized today using this technology. However, the need for multiple optical transmitters and receivers at the lower bitrate in combination with the optical components for multiplexing and demultiplexing does not lead to a very cost efficient solution.

Once realizable, ASK NRZ ETDM potentially provides a very low cost approach. For short reach interfaces, its cost advantages and low complexity (single fiber, single wavelength, small number of optical components with low complexity) create a lot of attractiveness. In case of LH and ULH applications, other criteria will dominate. One of them is bandwidth efficiency. Metro and regional networks installed today typically operate with channel spacings of 100 GHz, whereas spacings down to 50 GHz can be found in core networks. A transmission format compatible with these channel spacings would be highly desirable.

For channel spacings of 100 GHz, optical duobinary modulation may provide a solution due to its higher filtering tolerance compared to NRZ. Another advantage of this format are the lower bandwidth requirements for the optical modulator. Channel spacings of 50 GHz can potentially be realized with differential quaternary phase shift keying (DQPSK). This multi-level format features transmission of two bits per symbol. The resulting symbol rate of 50 Gsymbol/s reduces the speed requirements for the electronic components and leads to a higher tolerance towards linear signal distortions such as chromatic dispersion (CD) and polarization mode dispersion (PMD). Further reduction of the symbol rate can be realized by employing polarization multiplexing, i.e. transmission of an independent bitstream in each of the two orthogonal polarizations of a single mode fiber.

Tolerance towards linear and nonlinear signal distortions plays a major role in case of LH and ULH links. Orthogonal frequency division multiplexing (OFDM) potentially provides even more linear distortion tolerance. However, it may
suffer severely from nonlinear distortions. In addition, the implementation of this format imposes demanding requirements on high speed electronic as well as optical components.

3. IMPAIRMENTS AND SIGNAL DEGRADATION

Detrimental effects turning transmission at 40 Gbit/s into a challenging task are even more difficult to handle at 100 Gbit/s. The impact of some effects scales linearly with bitrate, while others scale quadratic or even with higher powers. Tolerance towards amplified spontaneous emission (ASE) is one example for a linear relation. Due to the bitrate increase of 2.5, 100 Gbit/s receivers need approx. 4 dB more optical signal to noise ratio (OSNR) than receivers for a given format at 40 Gbit/s to achieve the same bit error ratio (BER). Increasing the launch power into the transmission fiber usually does not provide an option to enhance OSNR due to limits imposed by nonlinear effects.

Forward error correction (FEC) helps to lower OSNR requirements. However, initial implementations for a higher bitrate typically achieve less coding gain than the more mature ones for lower bitrates. In the first generation of chipsets, less complex algorithms can be realized due to higher speed requirements for the electronic circuits. In consequence, the step from 40 Gbit/s to 100 Gbit/s may require an OSNR increase of 6 dB or even more.

Using more advanced modulation formats opens another path for the reduction of OSNR requirements. Currently deployed terrestrial 40 Gbit/s long haul transport systems operate with duobinary modulation for increased chromatic dispersion (CD) and optical filtering tolerance. The next generation will most likely use differential phase shift keying (DPSK) to achieve a reach of more than 1,000 km, comparable to link lengths of 10 Gbit/s systems. The reduction of required OSNR by more than 3 dB enabled by the new modulation format helps considerably in this direction. However, if 40 Gbit/s systems already use DPSK at the time when first 100 Gbit/s systems are scheduled to go into service, not many options remain to equip the higher bitrate system with a more ASE tolerant modulation format than the lower bitrate ones.

It may turn out that distributed Raman amplification evolves as an enabling technology for long haul 100 Gbit/s transmission. Counterdirectional pumping induces gain in the transmission fiber, raising the occurring minimum signal power levels and thereby improving the OSNR budget. Fig. 2 illustrates the concept.

Fig. 2. Distributed Raman amplification induced by counterdirectional pumping of the transmission fiber.

The shift of the minimum power can lead to an OSNR increase of up to approx. 6 dB. Stronger pumping results in more gain, i.e. further increases the power levels at the fiber output but has little effect on the minimum power levels. An additional noise contribution induced by double Rayleigh backscattering sets limits to the maximum useful gain.
Distributed Raman amplification was already proposed as an enabling technology for 40 Gbit/s transmission. However, the introduction of FEC provided a more cost efficient solution, reducing the need for Raman amplification. As additional OSNR improvement is required for 100 Gbit/s and FEC is already in use for the lower bitrate, the Raman option may help to achieve OSNR budget targets.

Chromatic dispersion provides an example for a linear signal distortion scaling quadratically with bitrate. Current 10 Gbit/s systems usually working with ASK NRZ modulation achieve a CD tolerance of approx. +/- 500 ps/nm with direct detection. Dispersion compensation modules typically come with a granularity of 170 ps/nm, corresponding to 10 km of standard single mode fiber (SSMF). A finer granularity would increase the number of module types, resulting in higher costs. Due to the bitrate increase by a factor of 10, the CD tolerance of a 100 Gbit/s system with the same modulation format is reduced by a factor of 100. Increasing the number of available dispersion compensation module types by the same factor, resulting in a granularity of 1.7 ps/nm, clearly does not provide a feasible approach. Instead, either modulation formats with more dispersion tolerance or adaptive solutions have to be applied.

In addition to the requirement for continuous or small step residual dispersion tuning, adaptive compensation for 100 Gbit/s systems with insufficient dispersion tolerance from the modulation format also has to be readjusted over time. Temperature variations in the transmission fiber result in small variations of the chromatic dispersion due to the temperature dependence of the refractive index of 1.07x10^{-5}/K at 1550 nm [1]. The refractive index variation with temperature leads to a chromatic dispersion temperature dependence of approx. -1.5x10^{-3} ps/nm/km/K in SSMF [2]. For a system with a length of 1,000 km and a temperature change of the transmission fibers of 20 K, the accumulated link CD changes by approx. 30 ps/nm. Dispersion compensating modules are usually installed in a temperature controlled environment, hence keeping their dispersion constant. In a 10 Gbit/s system, this change of the residual dispersion does not cause any problem due to the dispersion tolerance of +/-500 ps/nm. For a 100 Gbit/s NRZ based system with its tolerance of only +/-5 ps/nm, the CD variation cannot be neglected and has to be accounted for dynamically.

Temperature changes of 20 K are typical for cables buried in a depth of approx. 1 m. As daily temperature changes of the air don't reach that deep into the ground, dispersion changes in buried cables occur on timescales of weeks or months. Adaptive compensation for these changes can be readjusted rather slowly. Much faster and stronger dispersion changes have to be expected for aerial fibers.

Adaptive dispersion compensation can be implemented by tunable dispersion compensators such as fiber Bragg gratings (FBG), etalons or virtually imaged phase arrays (VIPA). Impairments from residual dispersion can also be mitigated using equalizer technologies such as feed forward equalization (FFE) / decision feedback equalization (DFE) or maximum likelihood sequence estimation (MLSE). The equalizer approach has the advantage of enabling the option to compensate for other signal distortions, such as PMD, for example. Due to much faster changes of polarization and also PMD over time down to ms timescales, adaptive compensation/mitigation covering also PMD needs much faster response times than pure CD compensation.

The need for PMD mitigation depends on the PMD tolerance of the modulation format and the characteristics of the transmission fiber. For non-multilevel modulation formats, mean DGD tolerances of approx. 0.9 ps have to be expected. State of the art fibers with typical PMD coefficients down to 0.02 ps/km^{1/2} would enable link lengths up to 2,000 km. However, most installed fibers feature higher PMD coefficients and contributions from optical components in the network elements also have to be included. In typical installations, mean DGD link values of several ps are not uncommon. Therefore, either higher level modulation formats or PMD compensation/mitigation have to be deployed.

Table 1 summarizes characteristics of modulation formats which are relevant for 100 Gbit/s transmission (adapted from [3] and our own results). NRZ would be the most cost efficient to implement due to the low complexity of the transmitter and receiver. Carrier suppressed return to zero (CS-RZ) typically provides better tolerance against intrachannel nonlinear effects. However, both formats will not be applicable to DWDM transmission due to optical bandwidth requirements. Duobinary features better optical filtering tolerance and also CD tolerance. It may provide an option for medium reach metro systems with 100 GHz channel spacing, which can compensate penalties due to narrowband optical filtering by sufficient OSNR margin.
Table 1: Comparison of modulation formats for a bitrate of 112 Gbit/s in the form of a decision matrix

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>NRZ</td>
<td>20.8</td>
<td>102</td>
<td>7</td>
<td>2.6</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>CS-RZ</td>
<td>19.8</td>
<td>131</td>
<td>6</td>
<td>3.0</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Duobinary</td>
<td>21.5</td>
<td>80</td>
<td>30</td>
<td>2.4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>DPSK</td>
<td>16.0</td>
<td>115</td>
<td>7</td>
<td>2.7</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DQPSK</td>
<td>17.1</td>
<td>65</td>
<td>23</td>
<td>5.2</td>
<td>-</td>
<td>--</td>
</tr>
<tr>
<td>PolMUX-DQPSK</td>
<td>20.0</td>
<td>33</td>
<td>75</td>
<td>10.3</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Real improvements with respect to optical filtering tolerance and CD as well as PMD tolerance can only be expected from multi-level formats such as DQPSK. Especially the combination with polarization multiplexing (PolMUX), i.e. transmission of an independent signal in each of the orthogonal polarizations of a single mode fiber, significantly helps. Besides the high complexity of the required transmitter and receiver structures translating into high cost, these formats suffer significantly stronger from nonlinear effects.

The detrimental effects of nonlinear phase noise, i.e. interactions between ASE and the signal radiation due to the Kerr effect and cross phase modulation (XPM) induced by adjacent signal channels, strongly decreases the possible launch power into the transmission fiber, especially in fibers with a mix of 100 Gbit/s and 10 G NRZ channels. Due to higher OSNR requirements of the higher bitrate signals, this translates into severely reduced reach. It is not clear yet, whether multilevel modulation formats with enhanced tolerance against linear signal distortions but smaller tolerance against nonlinear effects provide a better solution or binary formats with more tolerance against nonlinear effects in combination with options to increase the tolerance against linear signal distortions.

In principle, the insufficient tolerance of binary modulation formats against linear signal distortions can be enhanced easier by equalizer technologies than mitigation of detrimental nonlinear effects. Advanced MLSE, channel coding or coherent detection in combination with electrical digital signal processing provide powerful approaches to increase tolerance against linear signal distortions. More research has to be carried out in this area to identify the best option.

**AVAILABILITY OF COMPONENTS**

Good results have already been achieved in the area of components. Research groups were able to demonstrate devices for multiplexing of parallel data streams in the electrical domain to serial binary bitrates of 100 Gbit/s and above [4]. Some of these devices provide sufficient output voltage for a potential to drive the modulator directly. However, in many applications, driver amplifiers will be required. Realization of packaged modules with sufficient bandwidth and output voltage for 100 Gbit/s signals are a huge challenge and could not be realized so far.

A similar conclusion has to be drawn for modulators for binary signals with symbol rates beyond 100 Gsymbol/s. Mach-Zehnder-interferometer (MZI) type as well as electro-absorption modulator (EAM) based devices are difficult to realize with sufficient bandwidth and low drive voltage. Considerable research and development efforts have to be spent before devices suitable for product applications can be expected. Modulators for multi-level formats such as DQPSK are easier to implement due to their relaxed bandwidth requirements. The challenge for these devices lies in the more complex structure and stable operation over wide temperature ranges.

On the receiver side, good results could be achieved concerning photodiodes. Chips with bandwidths exceeding 100 GHz were demonstrated as well as packaged modules with a bandwidth of 100 GHz. Combining the photodiode with an electrical amplifier reduces the cut-off frequency of the module to 70 GHz, still leaving sufficient bandwidth for NRZ receivers. Together with already demonstrated integrated clock and data recovery chips for bitrates up to 107 Gbit/s and demultiplexers up to 110 Gbit/s, all major building blocks for ETDM receivers seem to be available with maturity levels close to product requirements.
Major remaining challenges on a component level are given by the packaging and integration of components with sufficient bandwidth. Promising results could be demonstrated for some specific components, but in many other cases, the maturity has to be improved significantly to meet product requirements.

On a subsystem level, integration for reduction of size, power consumption and cost is a task still in its infancy. Major efforts have to be spent in this area. Signal processing functions such as equalization and forward error correction already implemented at lower bitrates stretch the limits of currently available electronic processing capabilities. Finding a compromise between speed limits and complexity or the number of required gates will be a demanding endeavour.

**SUMMARY AND CONCLUSIONS**

Several options exist for realizing 100 Gbit/s Ethernet interfaces from a physical layer perspective. Serial transmission using a single wavelength probably is the most desirable one. For short range interfaces, CWDM may provide a faster realizable solution before more cost efficient serial implementations become available. In longer range transport applications, bandwidth efficiency dictates a restriction to single wavelength options.

It is not clear yet, which modulation format provides the best solution for long haul transport. Multilevel formats such as DQPSK, especially in combination with PolMUX, provide good tolerances against linear signal distortions. The drawback comes from reduced launch powers into the transmission fiber due to nonlinear effects. This power reduction in combination with increased OSNR requirements due to the higher bitrate turns achieving a similar reach as for lower bitrate systems into a difficult task. Maybe binary formats with their higher tolerance against nonlinear effects combined with powerful equalizer technologies offer a better solution.

**REFERENCES**

4. White paper on 100 GbE from the EIBONE working group on transmission technology, to be published under www.vde.com/Allgemein/Reports/positionspapiere/