Feasibility of a 100GE MAC

Med Belhadj
Alex Nicolescu
David Walsh
Victor Uras
Summary

- A 10 lane MAC/PMD interface for 100 GE (CTBI)
- Impact of CTBI on the PCS Feasibility
- Feasibility of the Major MAC functions
- Conclusion
A 10 lanes 100G Interface

CTBI = 10 lanes at 10.3 Gbps (10Gbps 64/66 coded)
Why is 10 Lanes “a sweet spot”

- 10.3Gbps is feasible in standard CMOS
- Connectors running at 10.3Gbps are reasonable
- We can make a low power, low complexity interface between the MAC/PCS chip and PMA/PMD with all lanes frequency locked (similar to OIF SFI5.2)
- Simple Data Muxing and Demuxing
- Simplify the MAC/PCS digital design
Interface Chip complexity

- PMD choice will drive interface chip complexity

- Digital Complexity:
  - Number of “lanes”:
    - If 10x or 5x then the design is very simple, if 8x or 4x require a Gearbox
    - Other width possible but involve additional gear-boxing
  - Muxing level:
    - Bit or Word (66bit)?
  - Scrambling & Alignment

- Analog complexity:
  - 20.6 Gbps or 25.7 Gbps Serdes in CMOS (e.g. 65nm)?

- PMD will likely have a CMOS and/or Exotic depending on the rate supported and the functionality
CTBI Interface

1. Each Lane is 64/66 Encoded
   - One scrambler per lane (assumption made for feasibility study)

2. Data striping mechanism (how MAC data map into multiple lanes)
   - 64-bit Word Based
   - 32-bit Column Based

3. Multi-lane Align & Clock Compensation Mechanisms
   - Similar to Align/Skip characters in XAUI
Looking at CGMII (the 100G equivalent of XGMII)
The following bus width can be used

<table>
<thead>
<tr>
<th>CGMII as 10x</th>
<th>CGMII as power of 2 (not a natural width if using CTBI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>640bit @ 156.25 Mhz (80B)</td>
<td>512bit @ 195.3125 Mhz (64B)</td>
</tr>
<tr>
<td>320bit @ 312.5 Mhz (40B)</td>
<td>256bit @ 390.625 Mhz (32B)</td>
</tr>
<tr>
<td>160bit @ 625 Mhz (20B)</td>
<td>128bit @ 781.25 Mhz (16B)</td>
</tr>
<tr>
<td>80bit @ 1250 Mhz (10B)</td>
<td>64bit @ 1562.5 Mhz (8B)</td>
</tr>
</tbody>
</table>

Speed = Conservative
Width = Too wide

Speed = Sweet spot
Width = wide

Speed = Very aggressive
Width = narrow

Speed = Not in SC ASIC
Width = very narrow

=> We will assume CGMII 40 Bytes for this presentation
CGMII Example (40 Bytes)

Keep Frames the Same as 10GE:

- Start ||S|| can occur in any 32-bit
  - 10 different possibilities
- End /T/ can occur in any Byte
  - 40 different possibilities
- Same IFG rules
64-bit Word Striping

CGMII = 40 Bytes Words (+40 bit control)

10 columns (of 4 Bytes+4bit ~ XGMII like)

Ethernet stream In 4 Bytes “Columns”

IEEE Meeting Nov 2006
### 32-bit Column Striping

#### CGMII = 40 Bytes Words

<table>
<thead>
<tr>
<th>Lane#0</th>
<th>0</th>
<th>40</th>
<th>80</th>
<th>120</th>
<th>160</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>41</td>
<td>81</td>
<td>121</td>
<td>161</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>42</td>
<td>82</td>
<td>122</td>
<td>162</td>
<td>202</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>43</td>
<td>83</td>
<td>123</td>
<td>163</td>
<td>203</td>
</tr>
<tr>
<td>Lane#1</td>
<td>1</td>
<td>44</td>
<td>84</td>
<td>124</td>
<td>164</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>45</td>
<td>85</td>
<td>125</td>
<td>165</td>
<td>205</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>46</td>
<td>86</td>
<td>126</td>
<td>166</td>
<td>206</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>47</td>
<td>87</td>
<td>127</td>
<td>167</td>
<td>207</td>
</tr>
<tr>
<td>Lane#2</td>
<td>1</td>
<td>48</td>
<td>88</td>
<td>128</td>
<td>168</td>
<td>208</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>49</td>
<td>89</td>
<td>129</td>
<td>169</td>
<td>209</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>50</td>
<td>90</td>
<td>130</td>
<td>170</td>
<td>210</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>51</td>
<td>91</td>
<td>131</td>
<td>171</td>
<td>211</td>
</tr>
<tr>
<td>Lane#3</td>
<td>1</td>
<td>52</td>
<td>92</td>
<td>132</td>
<td>172</td>
<td>212</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>53</td>
<td>93</td>
<td>133</td>
<td>173</td>
<td>213</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>54</td>
<td>94</td>
<td>134</td>
<td>174</td>
<td>214</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>55</td>
<td>95</td>
<td>135</td>
<td>175</td>
<td>215</td>
</tr>
<tr>
<td>Lane#4</td>
<td>1</td>
<td>56</td>
<td>96</td>
<td>136</td>
<td>176</td>
<td>216</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>57</td>
<td>97</td>
<td>137</td>
<td>177</td>
<td>217</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>58</td>
<td>98</td>
<td>138</td>
<td>178</td>
<td>218</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>59</td>
<td>99</td>
<td>139</td>
<td>179</td>
<td>219</td>
</tr>
</tbody>
</table>

#### Frames in 4 Bytes “Columns”

- Lane#0: 00, 01, 02, 03, 04
- Lane#1: 10, 11, 12, 13, 14
- Lane#2: 20, 21, 22, 23, 24
- Lane#3: 30, 31, 32, 33, 34
- Lane#4: 40, 41, 42, 43, 44
- Lane#5: 50, 51, 52, 53, 54
- Lane#6: 60, 61, 62, 63, 64
- Lane#7: 70, 71, 72, 73, 74
- Lane#8: 80, 81, 82, 83, 84
- Lane#9: 90, 91, 92, 93, 94

---

IEEE Meeting Nov 2006
Data Striping

Data striping across 10 lanes can be done using:

- **64 bits Striping:**
  - Same coding as 10GBase-R
  - Align or Skip has to be 80Bytes
  - Simple Muxing/De-muxing

- **32 bits Striping**
  - Require new Control Types
  - Align and Skips can be 40 Bytes

- **Note:** Even Byte striping is possible across the 10 Lanes, but requires a different 64/66 coding
Codes required for Word & Column Striping

Start column

End Column(s)

Ordered Set Column

Control Column

Data Column

Valid 64-bit in 10GBase-R (15 Block Types)

Needed 64-bit Combinations to allow for 32-bit Striping
(15 + 48 = 63 Block Types)

IEEE Meeting Nov 2006

Nov 13-15
MAC Digital Requirement

100 Gbps MAC Digital core Implementations options:

- **Bus width:**
  - 320 bit Bus (not a power of two)
  - 256 bits (require a 4:5 Gear box to connect to CTBI interface)

- **Speed Up**
  - The implementation can be simplified by:
    - Making the Start Byte always at Byte 0 of the bus, while keeping the same IFG as in 10GE require a speed Up of:
      - **256bit Bus** @ 390.625 Mhz with a Speed Up of 1.2x for ~ 470 Mhz
      - **320bit Bus** @ 312.5 Mhz with a Speed Up of 1.36x for ~ 425 Mhz
Example of a PCS/MAC Receive

PCS64/66 Serdes

Gearbox

Multi-lane Align
De-skew

CGMII

MAC Functions
CRC32 etc

Option# A

256bit @ 391 Mhz or
256bit @ 470 Mhz

320:256 Gear-Box

1.2x Speed-Up
(Start Align)

MAC Functions
CRC32 etc

Option# B

256bit @ 391 Mhz or
320@ 321.8 or 312.5 Mhz

1.36x Speed-Up
(Start Align)

MAC Functions
CRC32 etc

Option# C

320@ 425 Mhz

MAC Functions
CRC32 etc

Option# D

320@ 321.875
CRC32 checker with 320 bit data path

CRC32 checker
(Size is large because last Byte can be in 40 different positions)

320bit@ 425Mhz
Utilization 88%
Process 90nm CMOS
Gate Equivalent ~ 200Kgate
Size 1.026x1.026 mm²

As a reference:
CRC32 checker 64bit@350Mhz ~ 10K
CRC32 checker 128bit@350Mhz ~ 60K
## Major Blocks (90nm CMOS)

<table>
<thead>
<tr>
<th>Blocks</th>
<th>Size (in Kgate)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRC32 Check and Generator</td>
<td>320 @ 425 Mhz</td>
<td>400 Post Place &amp; Route result ~200Kgate per 1mm²</td>
</tr>
<tr>
<td></td>
<td>256 @ 470 Mhz</td>
<td></td>
</tr>
<tr>
<td>Gear-Box</td>
<td>(320:256)@ 390 Mhz</td>
<td>10 Rough estimate using Synthesis only with no optimization (No P&amp;R)</td>
</tr>
<tr>
<td>Striping across 10 Lanes</td>
<td>Columns (32 bits)</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>Words (64 bits)</td>
<td>10</td>
</tr>
<tr>
<td>De-scrambler (self-sync 10GE)</td>
<td>32 bits x (10 instances)</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>320 bits</td>
<td></td>
</tr>
<tr>
<td>Scrambler</td>
<td>32bits x (10 instances)</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>320bit @425 Mhz</td>
<td>35</td>
</tr>
<tr>
<td>Full PCS64/66 (10 instances)</td>
<td>32bits @ 321.8 Mhz (10 instances)</td>
<td>440 Post Place &amp; Route result 10GE PCS with 32:33 Gearbox</td>
</tr>
</tbody>
</table>

These sizes are used as a reference point, actual implementation might vary (based on tools, process, additional features such as diagnostic, etc)
Conclusion

- A Single rate 100G MAC is feasible and of a reasonable complexity

- Multiple implementation are possible and feasible in Standard CMOS

- A PMD interface of 10x 10.3 will leverage design and specifications for SERDES, connectors, etc
To do

- Narrow down choices for:
  - PCS/PMD interface (e.g. CTBI 10 lanes)
  - CGMII interface (e.g. 40 Bytes)

- Add estimate for other blocks related to:
  - Align/Sync & de-skew buffer size, when a mechanism is defined

- Determine the feasibility of of PMD interface chip:
  - 10x 10.3 to 5x 20.6
  - 10x 10.3 to 4x 25.7