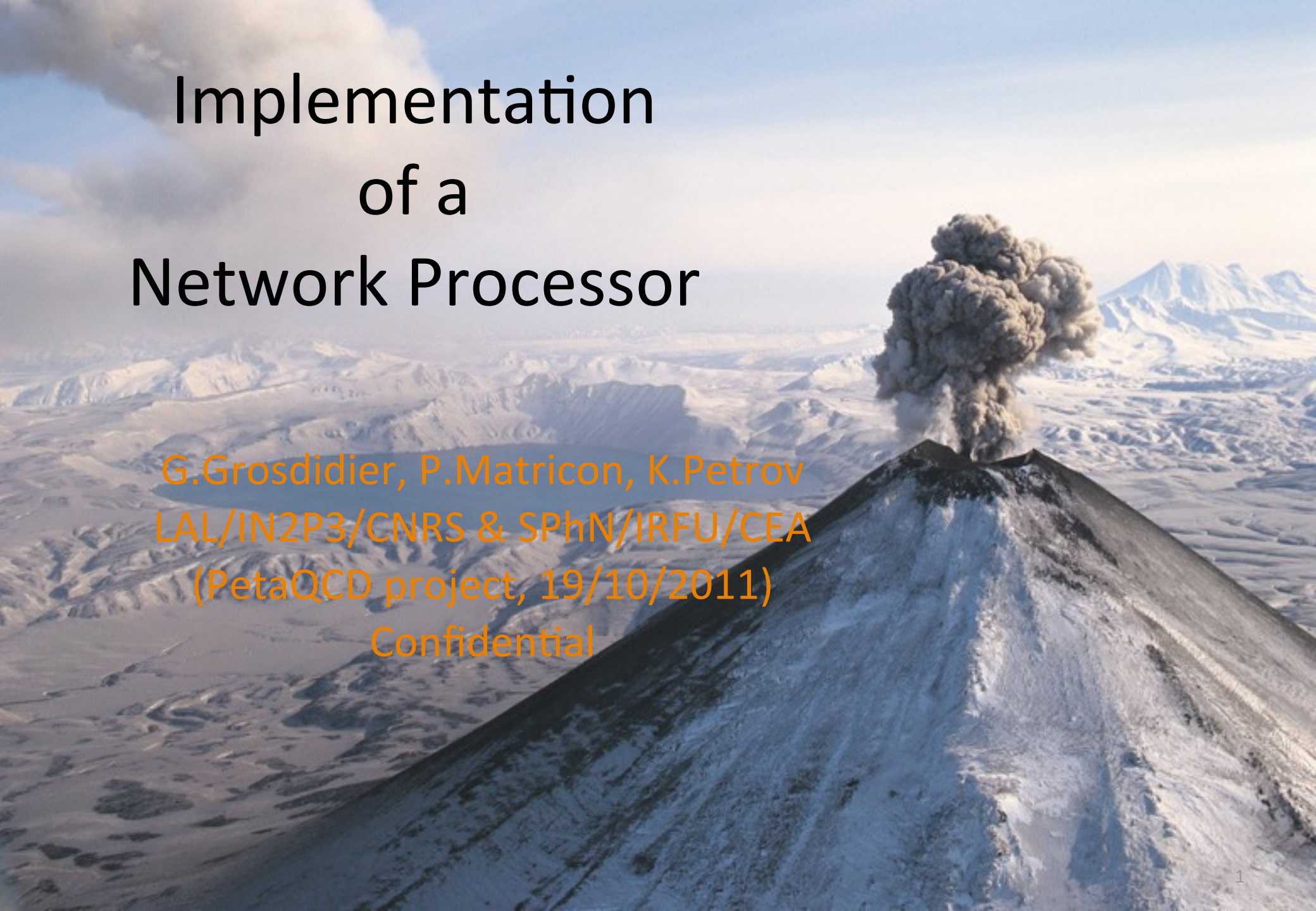
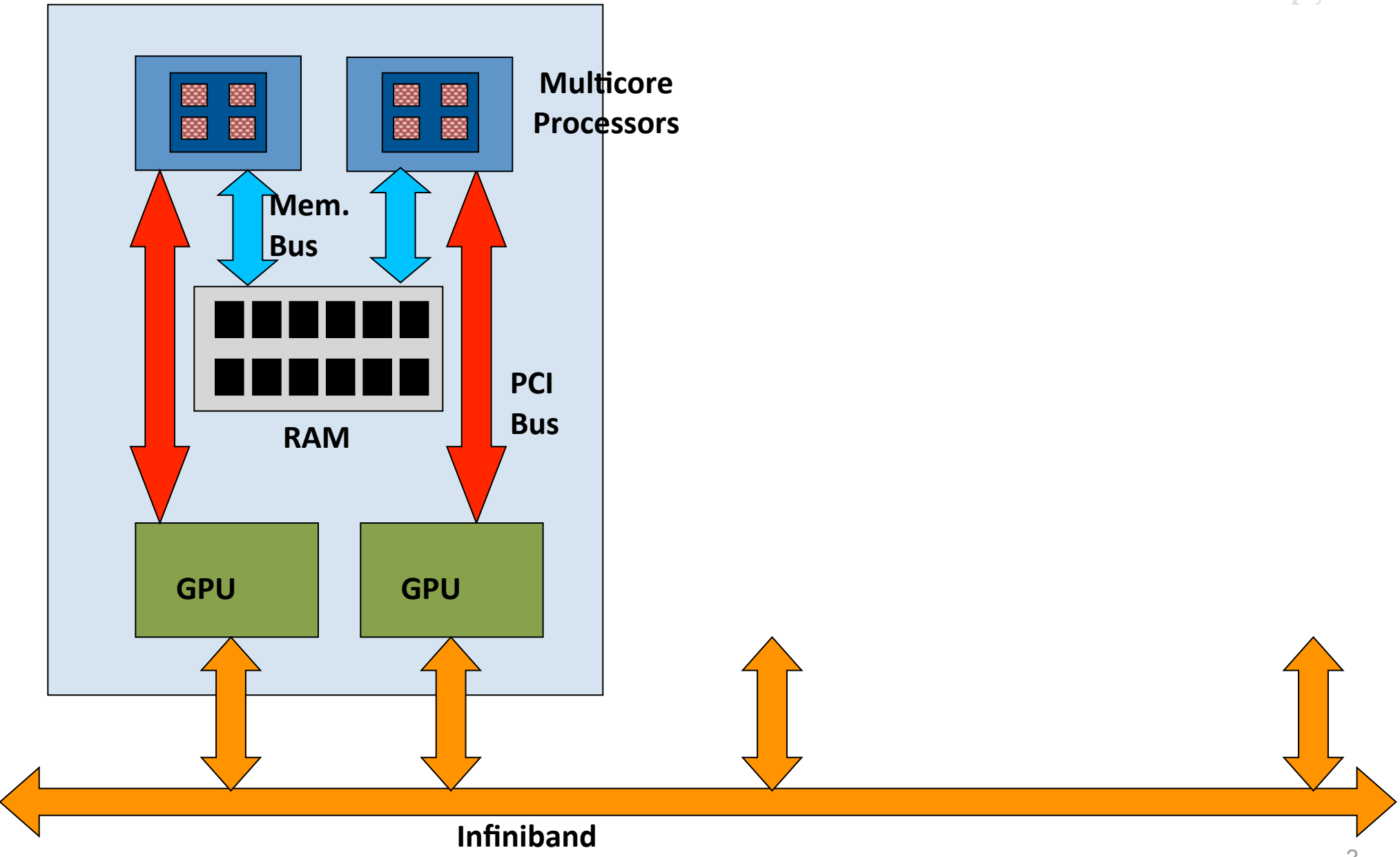
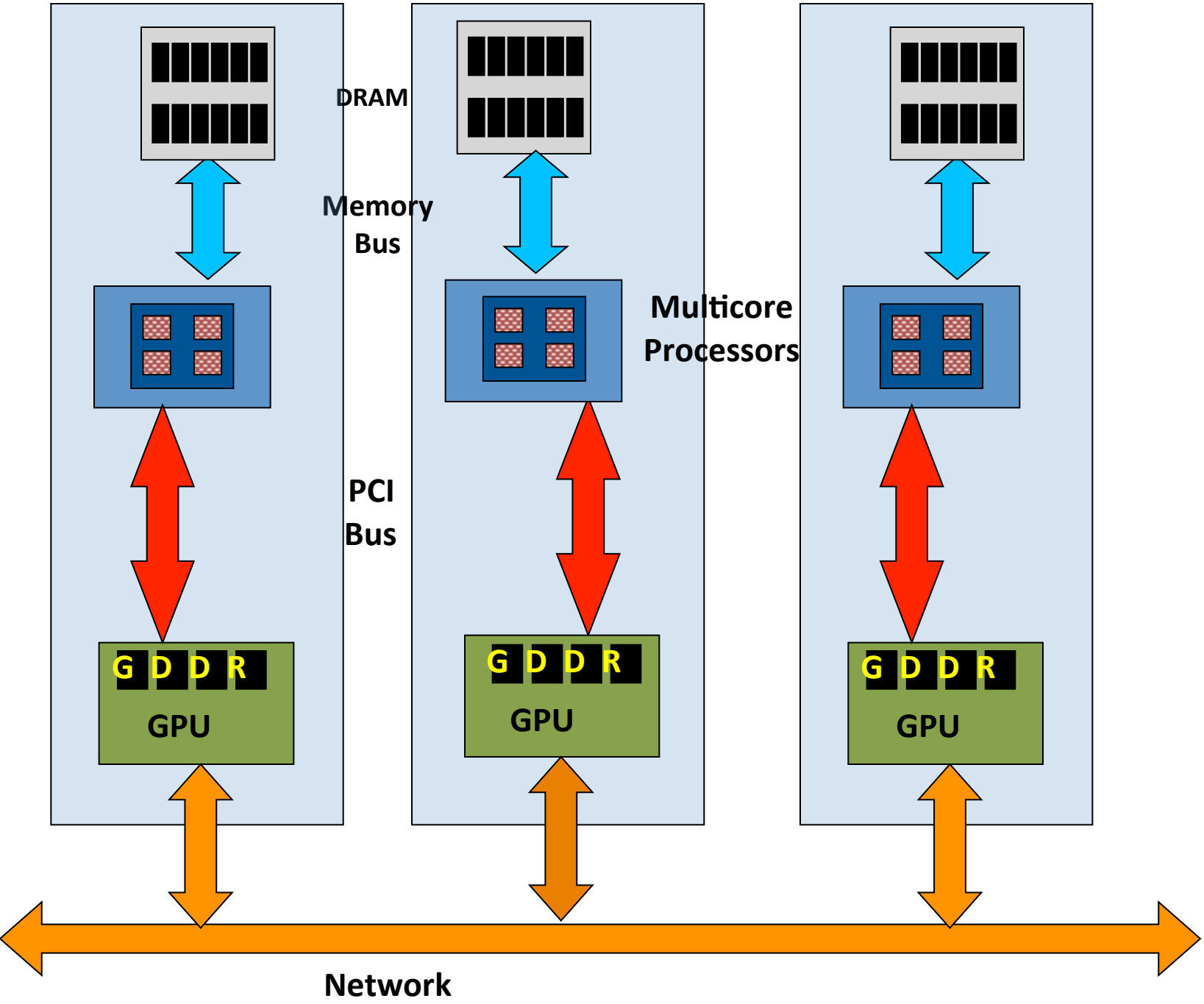


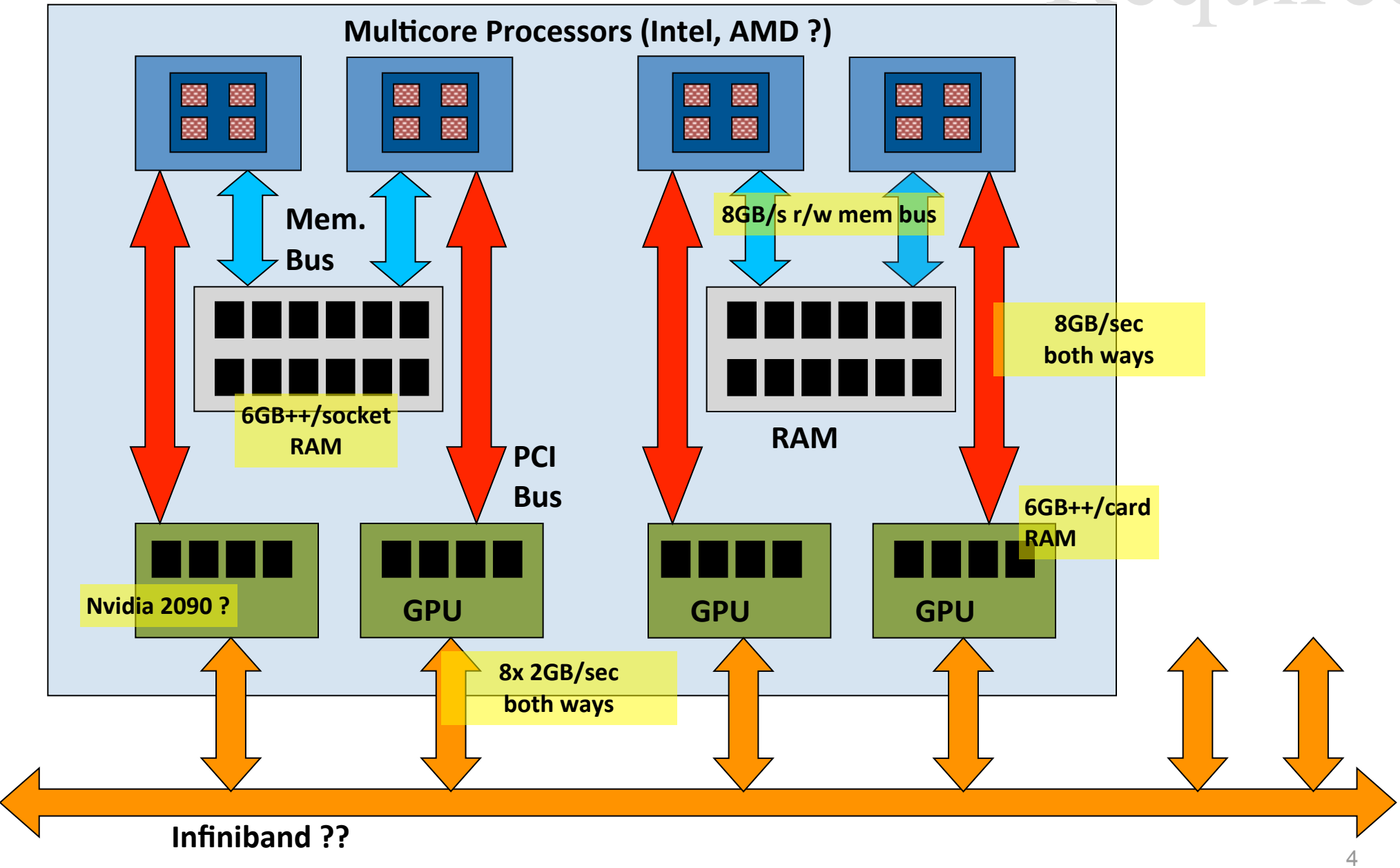
Implementation of a Network Processor

G.Grosdidier, P.Matricon, K.Petrov
LAL/IN2P3/CNRS & SPhN/IRFU/CEA
(PetaQCD project, 19/10/2011)
Confidential

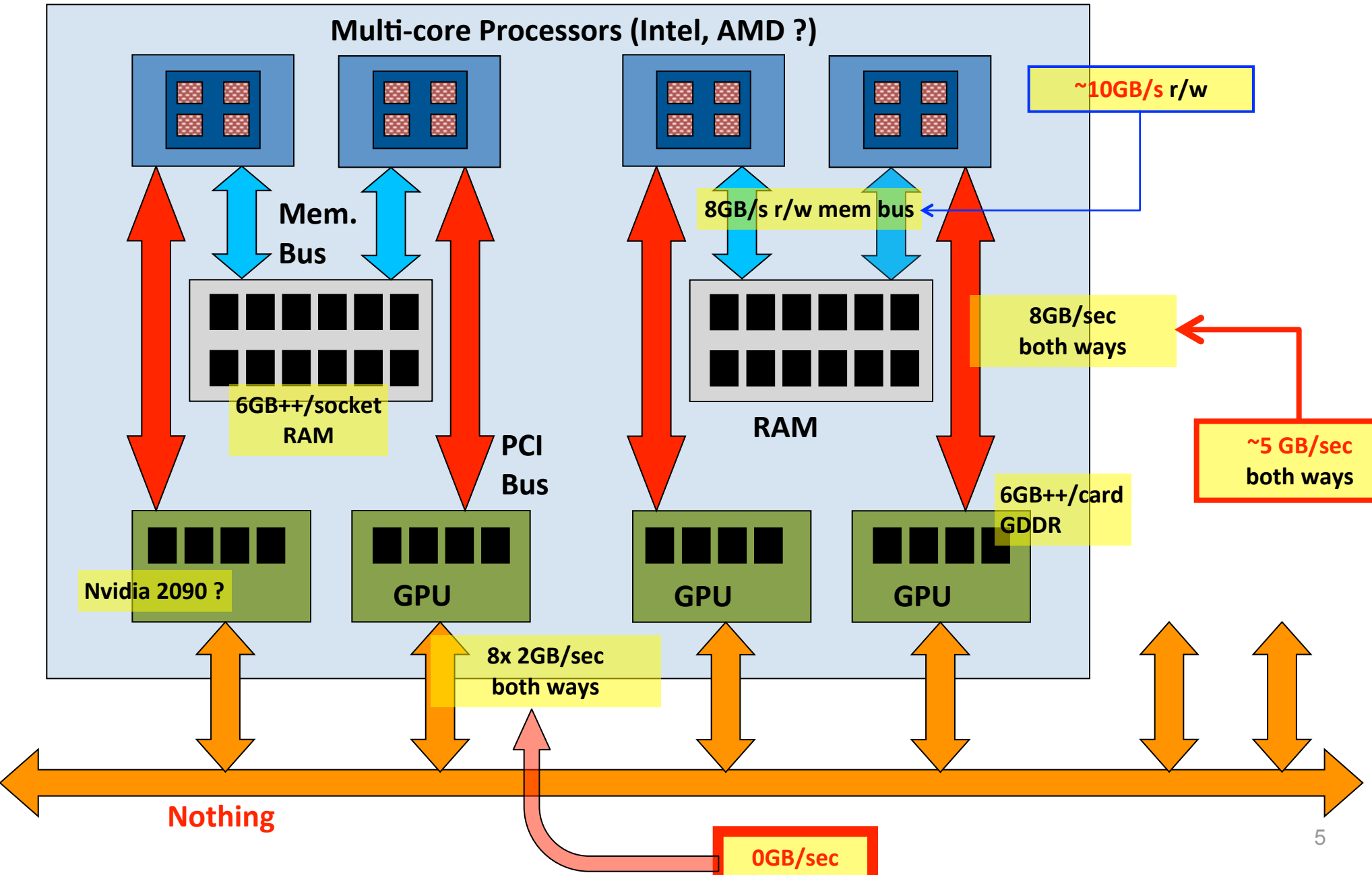
Block Design




DRAM
Memory
Bus
Multicore
Processors
PCI
Bus
G D D R
GPU
G D D R
GPU
G D D R
GPU
Network

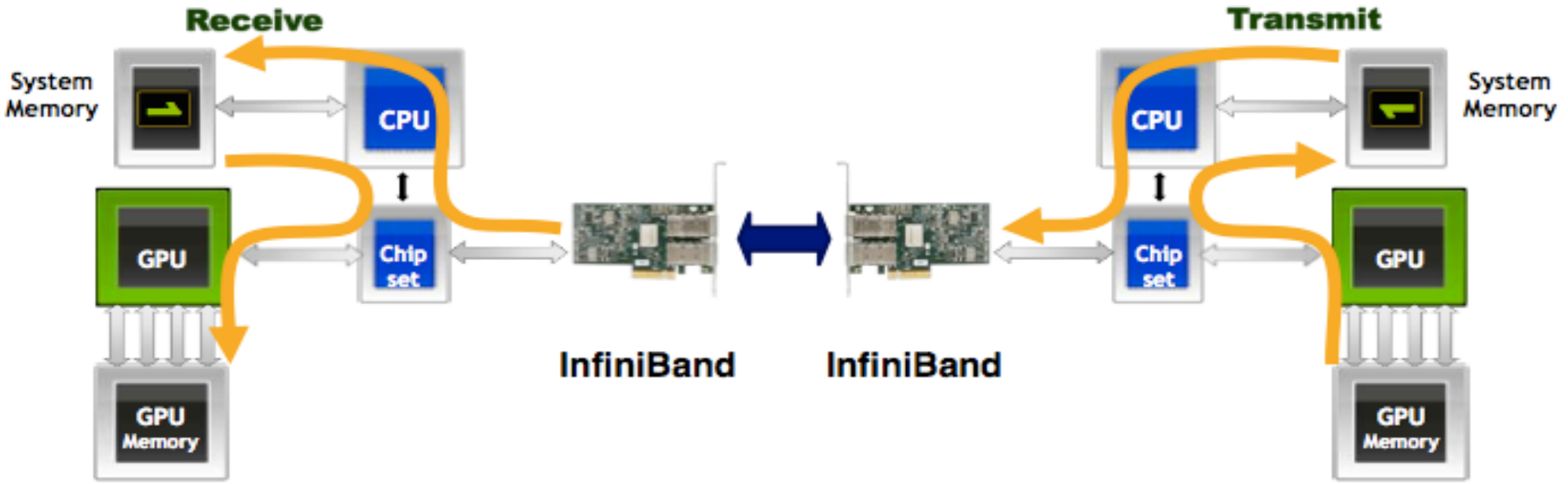

Multicore Processors (Intel, AMD ?)
Mem.
Bus
8GB/s r/w mem bus
8GB/sec
both ways
6GB++/socket
RAM
RAM
PCI
Bus
6GB++/card
RAM
Nvidia 2090 ?
GPU
GPU
GPU
8x 2GB/sec
both ways
Infiniband ??

Current


Multi-core Processors (Intel, AMD ?)
~10GB/s r/w
Mem.
Bus
8GB/s r/w mem bus
8GB/sec
both ways
6GB++/socket
RAM
RAM
~5 GB/sec
both ways
PCI
Bus
6GB++/card
GDDR
Nvidia 2090 ?
GPU
GPU
GPU
8x 2GB/sec
both ways
Nothing
0GB/sec

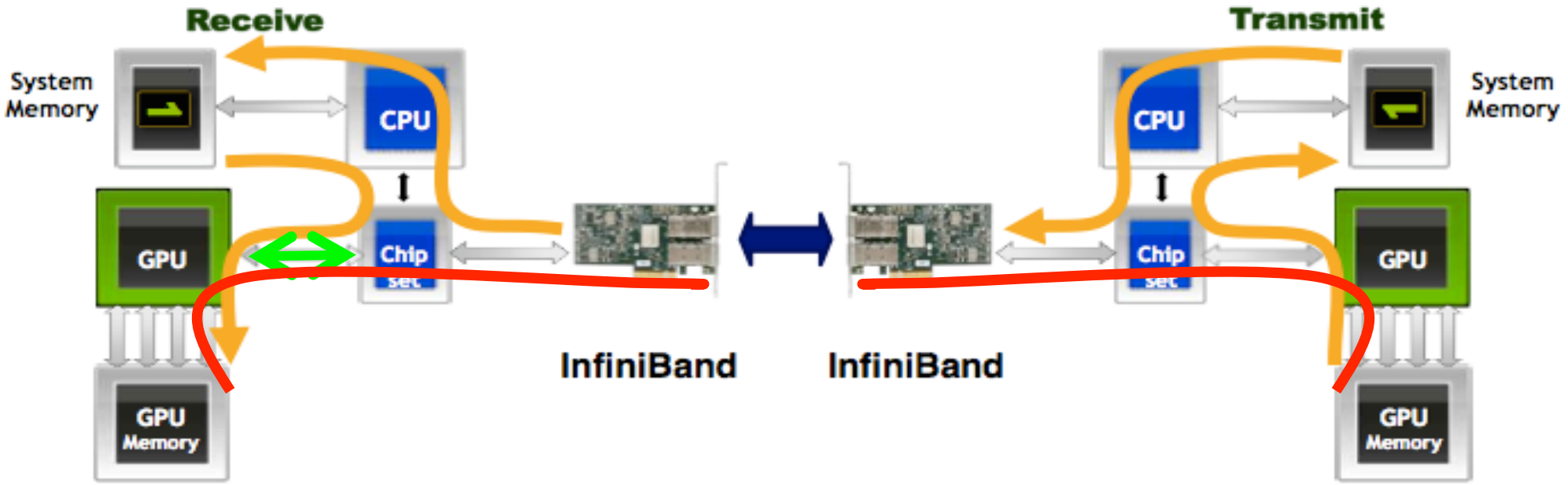
Current GPU-Direct

- The performances have been measured in INRIA-Bordeaux
  - This is obviously preliminary
  - They got 50 µsec latency ‘one-way’
  - This is by far too large for our requirements
  - What else ?



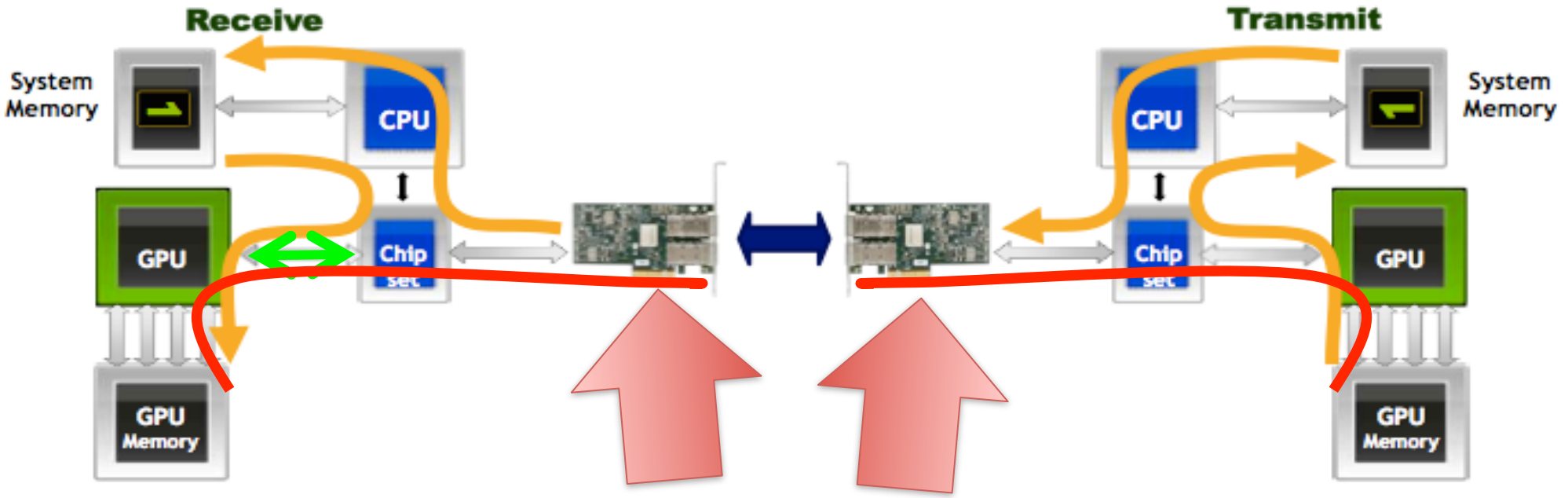
Next step ?

- What can be done to enhance current setup ?
  - There seems to be a link already between the GPU chip and the HCA one
    - Green arrow below
  - Why not to use a straight route thru this link towards the other GPU ?
    - Red path

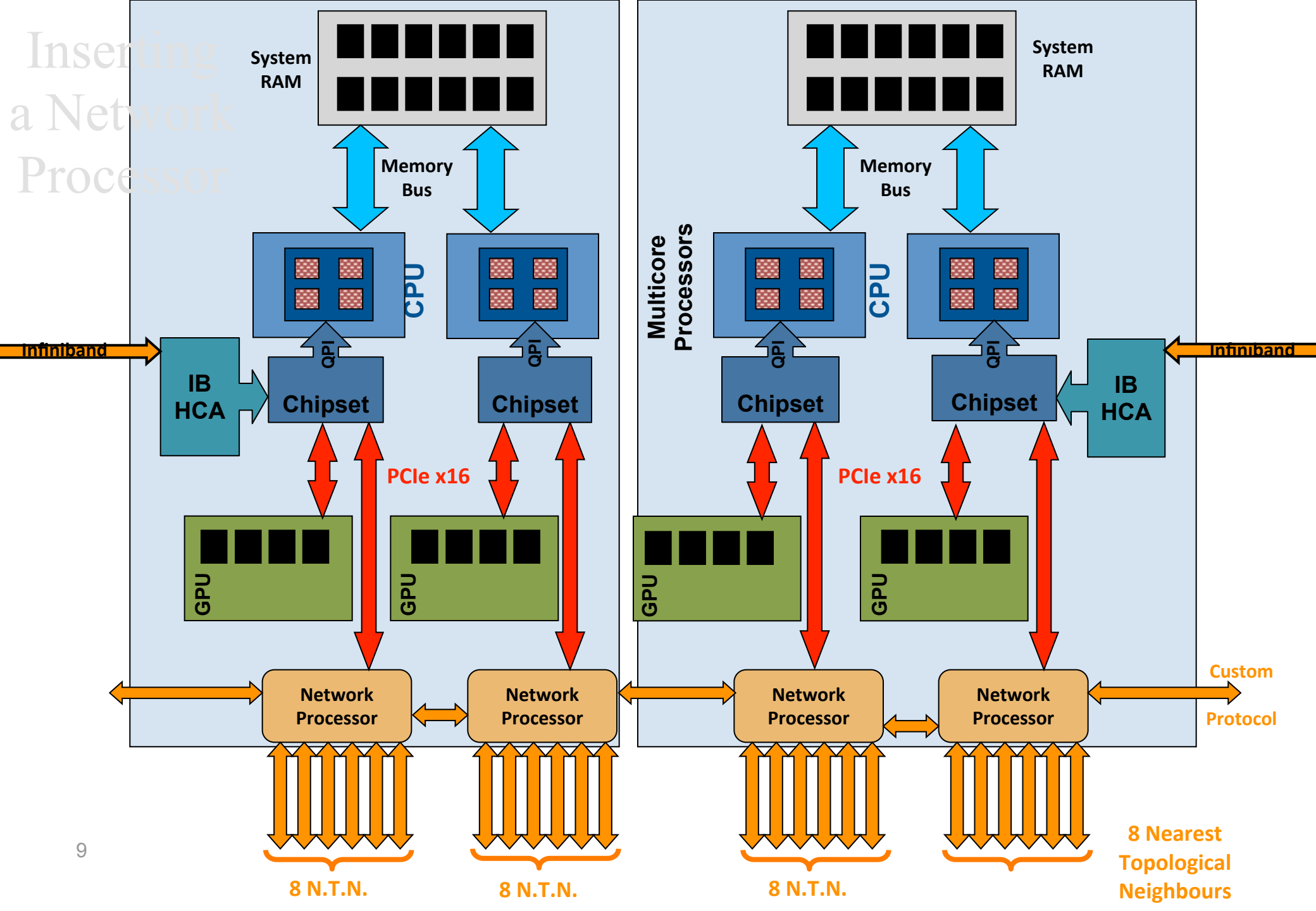


In the end

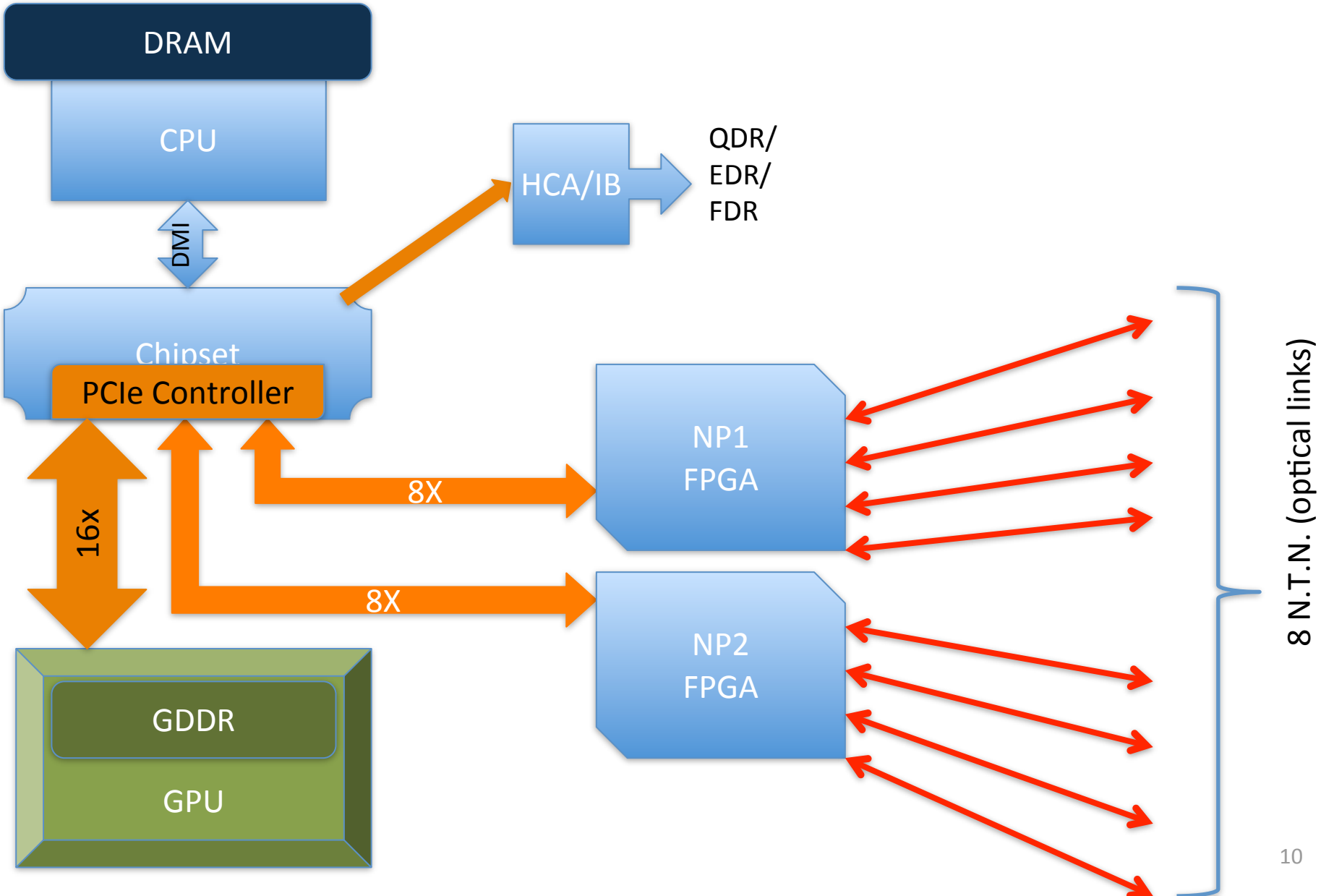
- Replace the Infiniband HCA with Custom Network Processors
  - Better suited to our 8 NTN specific connectivity toplogy
  - Get rid of any network switches
  - Does not avoid transit through PCIe though (latencies …)
  - Increases the PCIe bottleneck on GPU exit



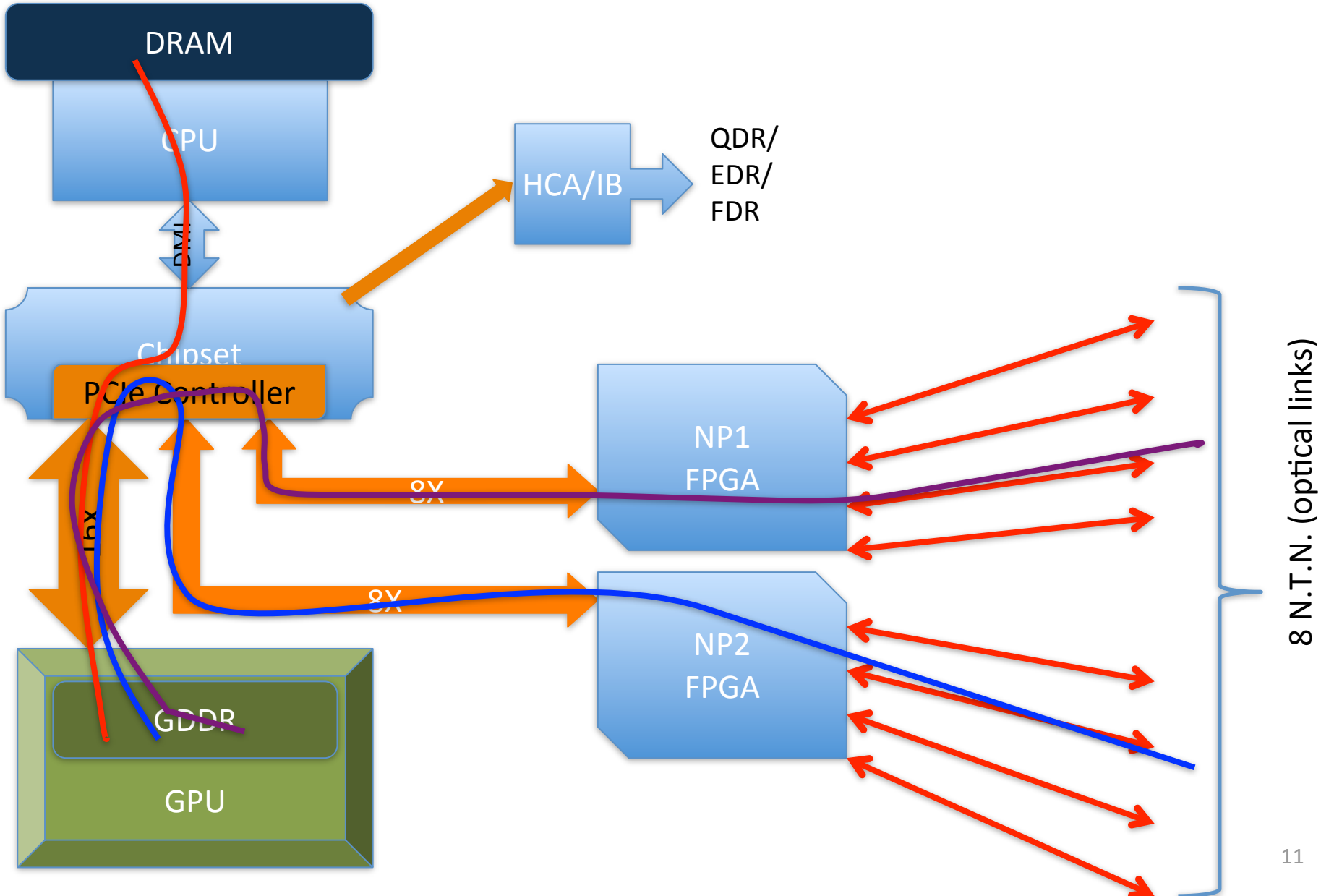
Custom Network Processors


System RAM
Memory Bus
CPU
QPI
Chipset
IB HCA
Infiniband
PCIe x16
GPU
Network Processor
8 N.T.N.
Multicore Processors
Custom Protocol
8 Nearest Topological Neighbours

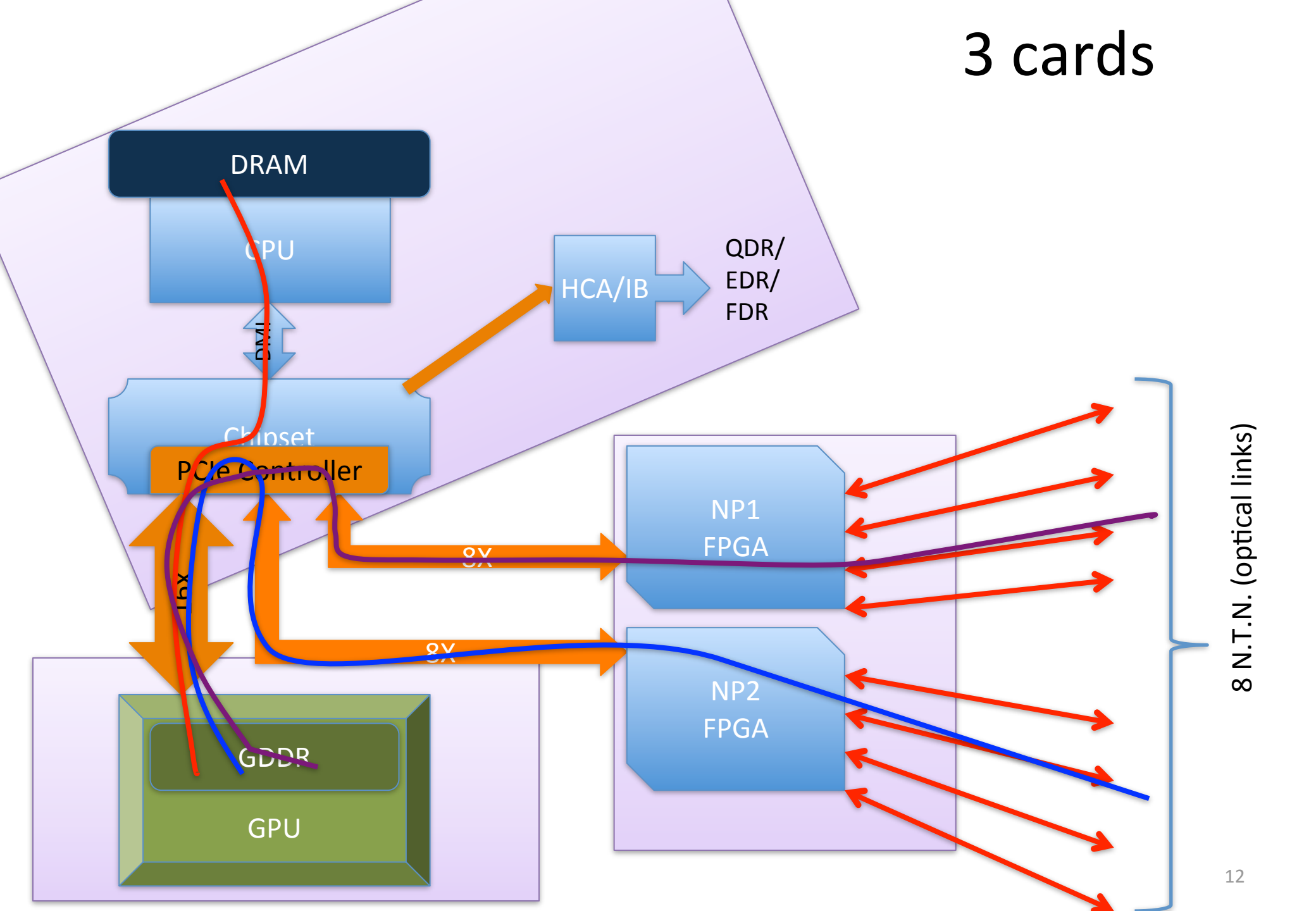
Simplest case : 1 socket and 1 Chipset



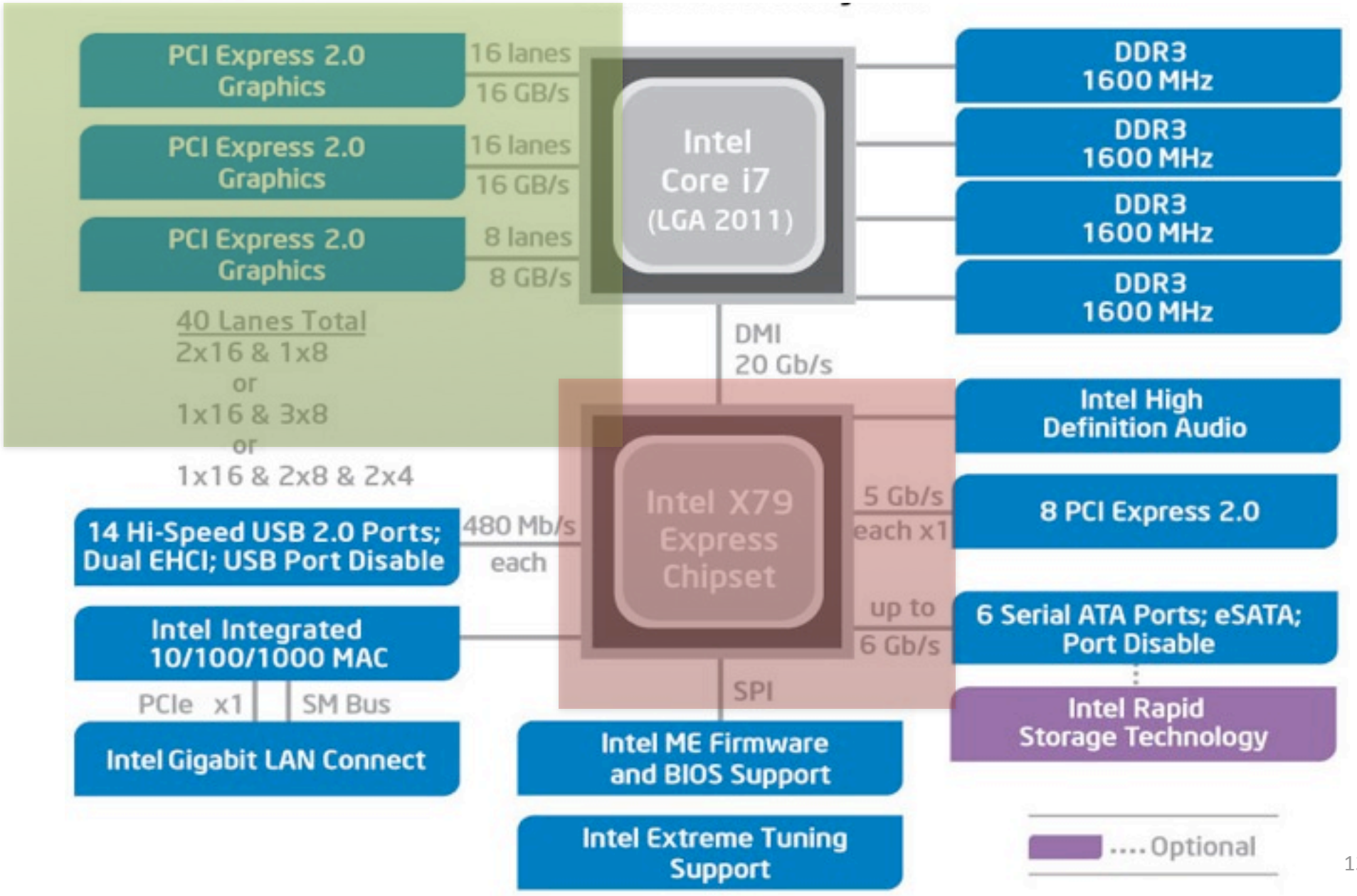
Data routes


DRAM
CPU
DMI
HCA/IB
QDR/
EDR/
FDR
Chipset
PCIe Controller
16x
8x
8x
NP1
FPGA
NP2
FPGA
GDDR
GPU
8 N.T.N. (optical links)

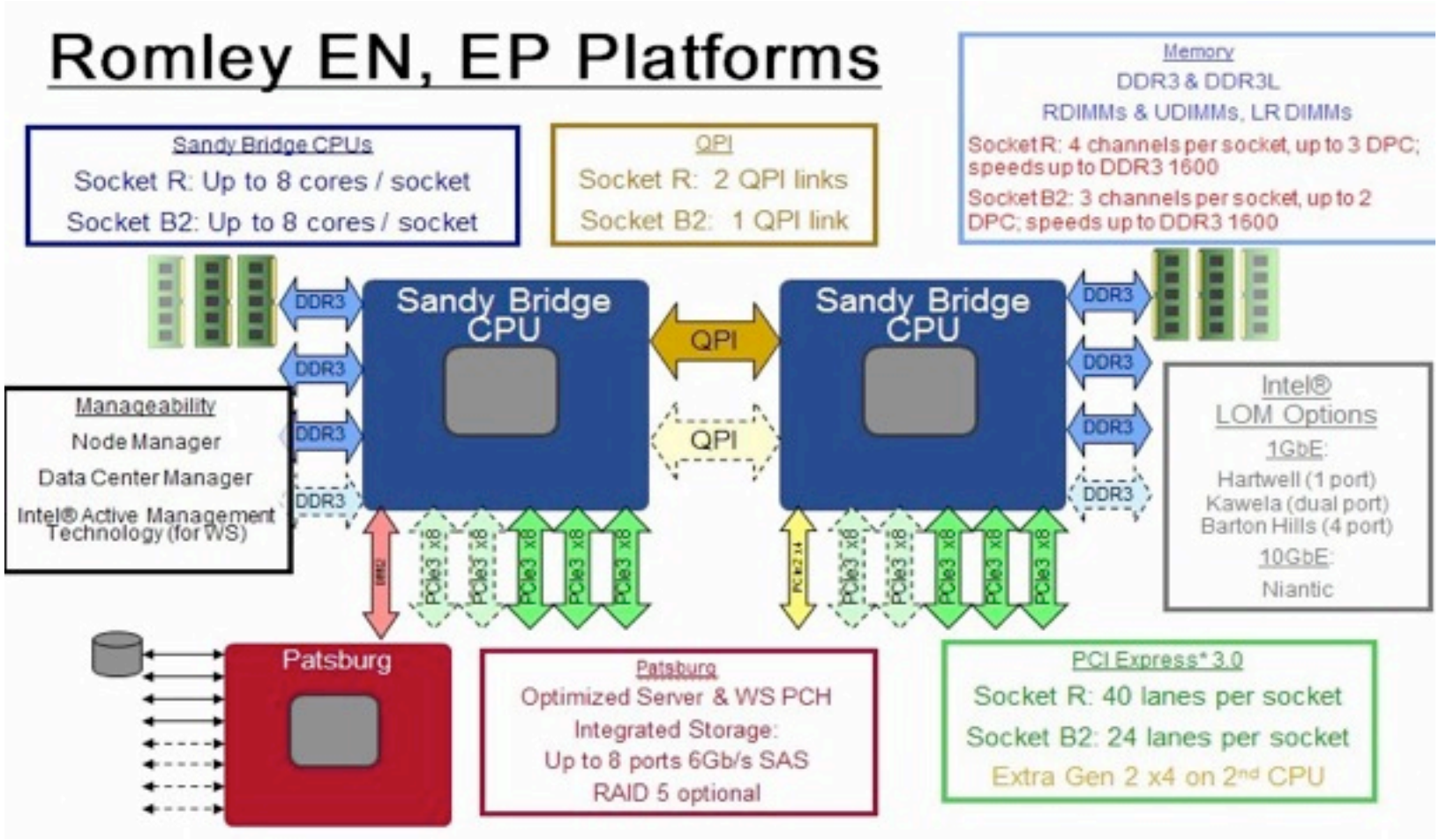
3 cards


DRAM
CPU
HCA/IB
QDR/
EDR/
FDR
Chipset
PCIe Controller
16x
8X
8X
NP1
FPGA
NP2
FPGA
GDDR
GPU
8 N.T.N. (optical links)

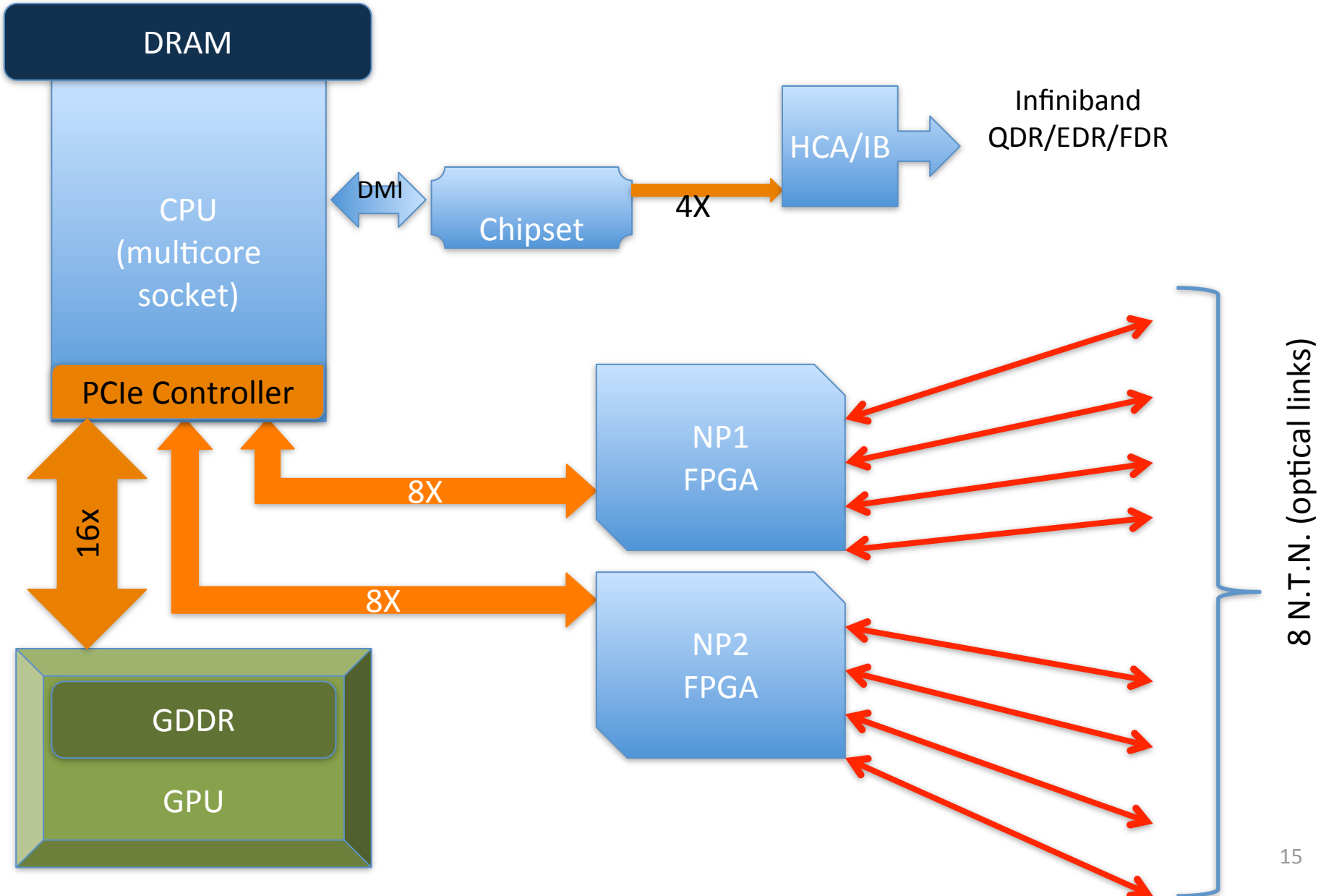
Sandy Bridge (expected)

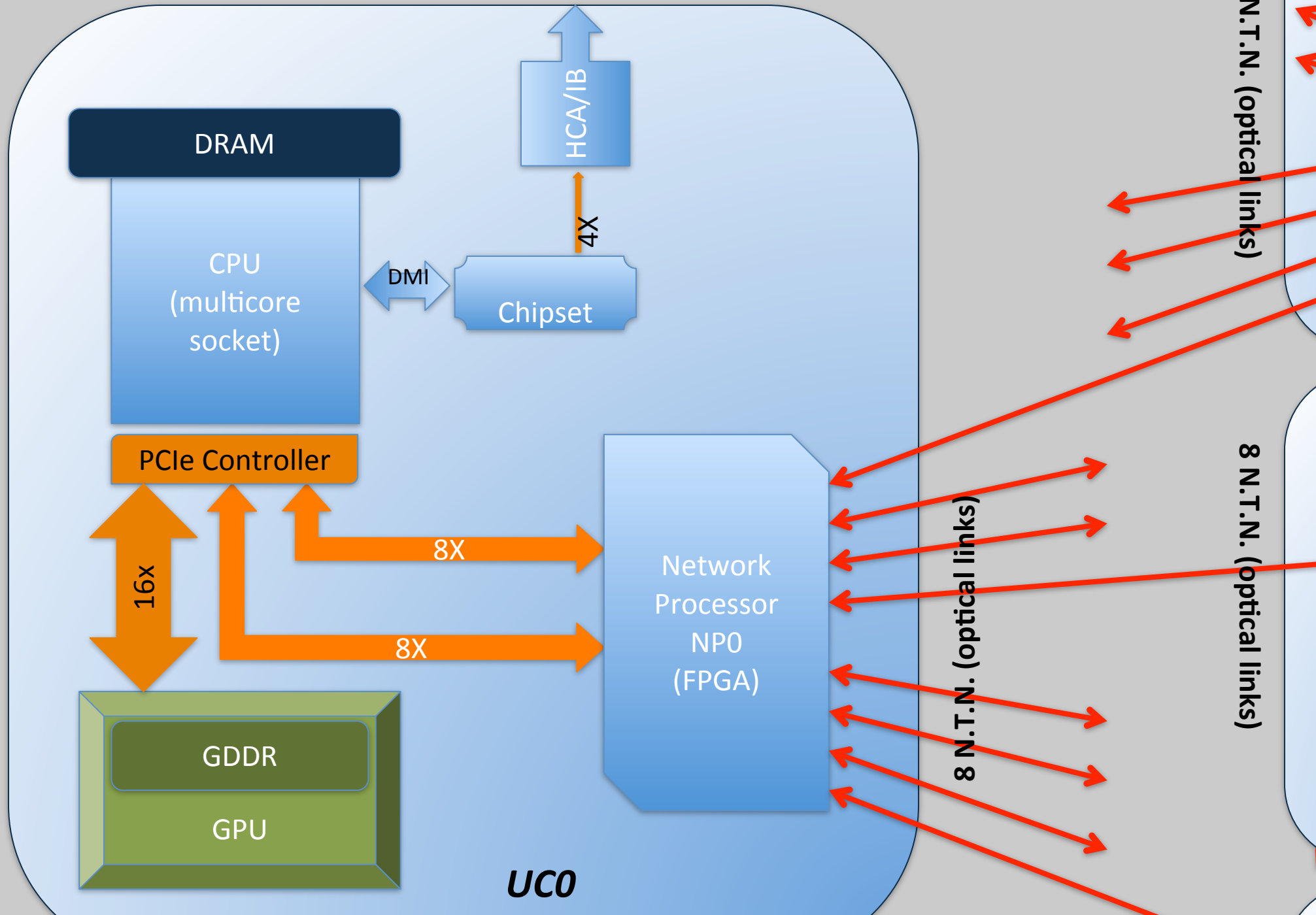


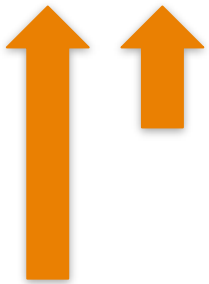
Sandy Bridge setup, 2 sockets

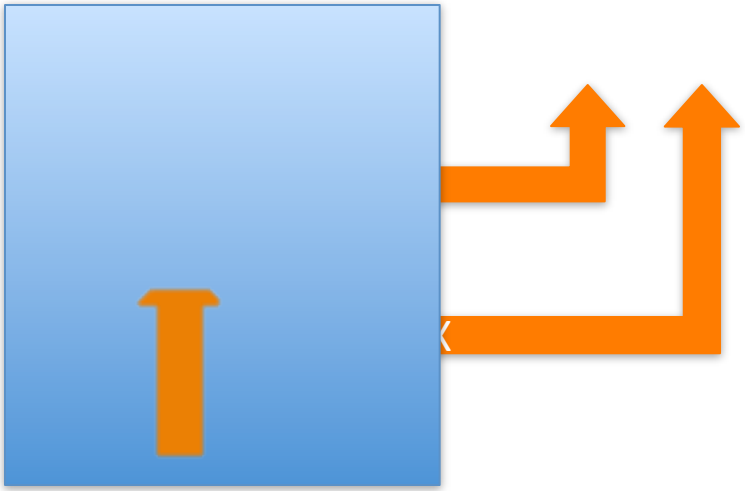


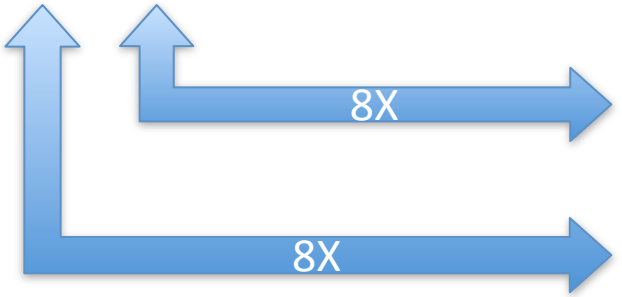
Sandy Bridge case(?)

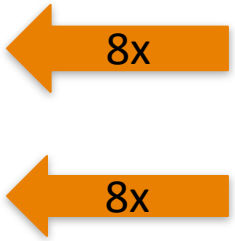



QDR/EDR/FDR
HCA/IB
DRAM
4X
CPU
(multicore
socket)
DMI
Chipset
PCIe Controller
16x
8X
8X
Network
Processor
NP0
(FPGA)
GDDR
GPU
UC0
8 N.T.N. (optical links)
8 N.T.N. (optical links)
8 N.T.N. (optical links)






8X
8X


8x
8x