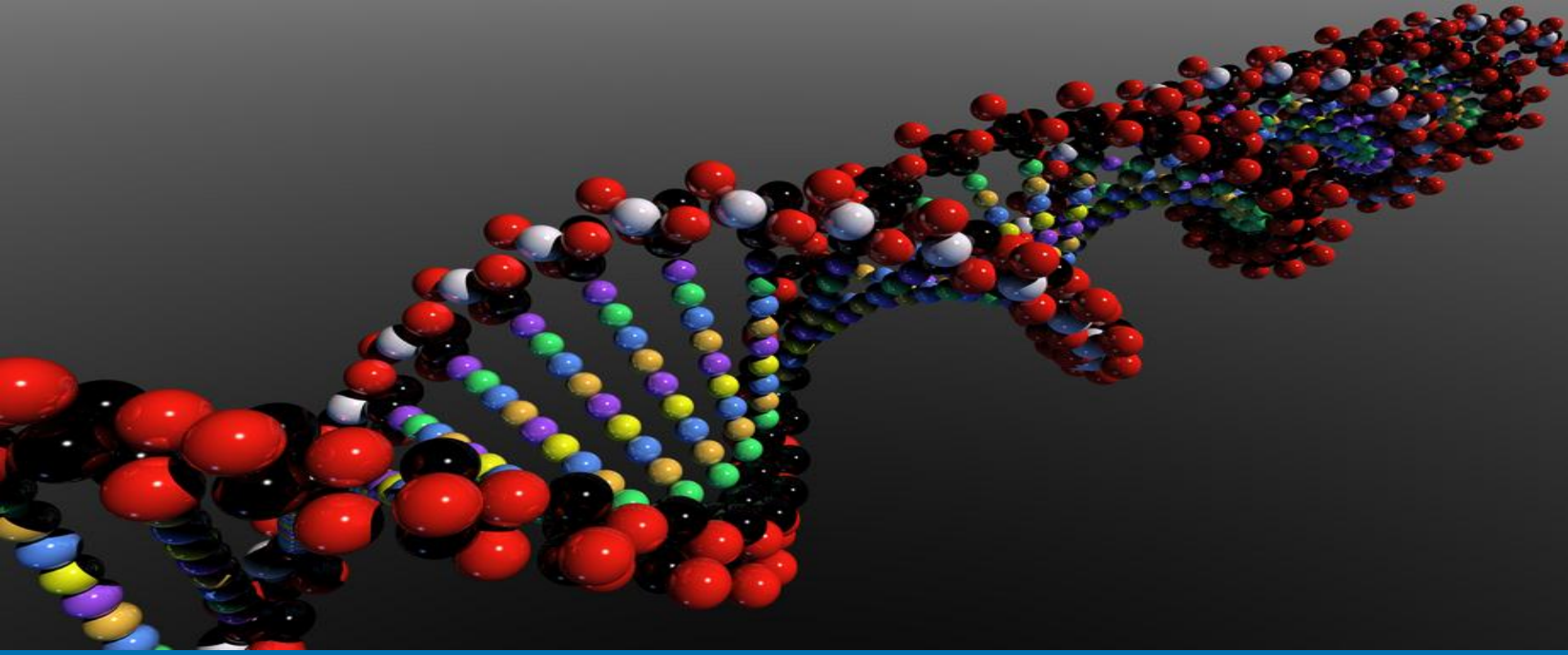
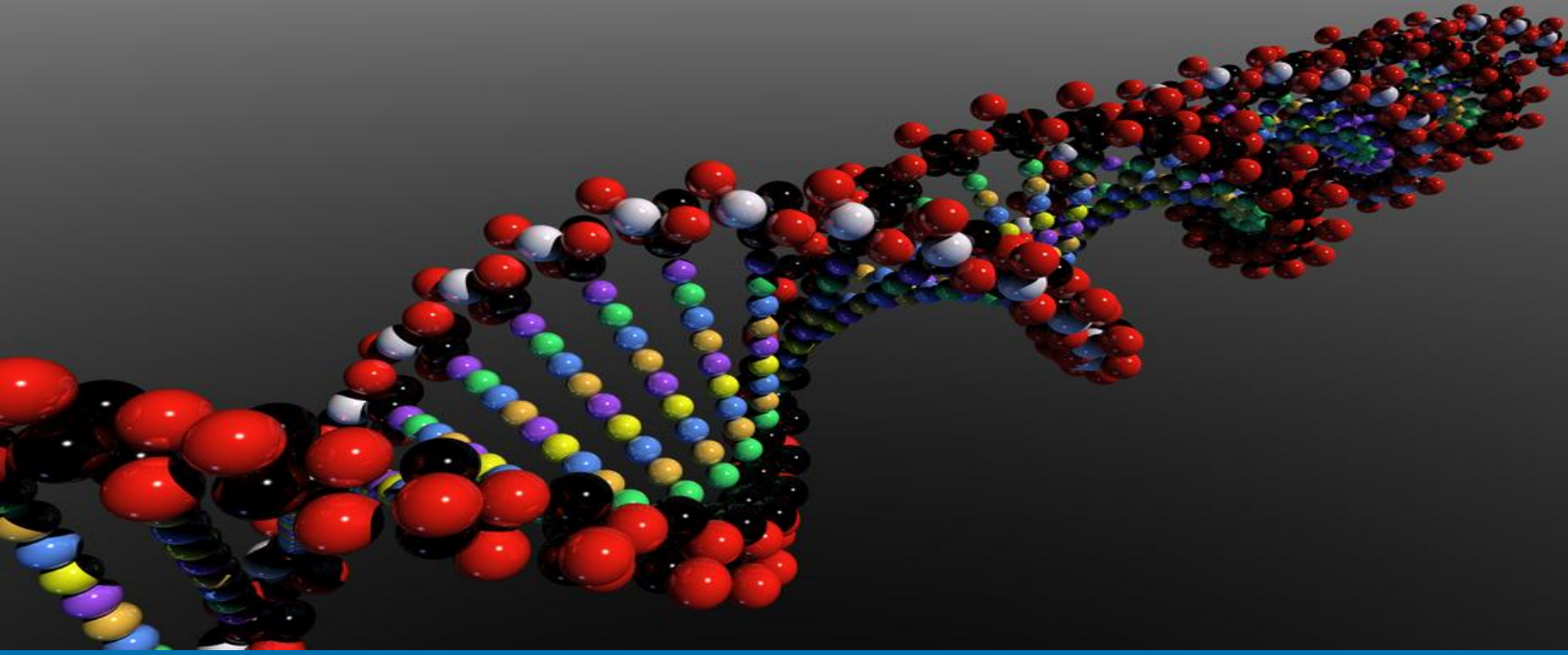


The Importance of Data Centricity in an Increasingly Parallel Computing World

Allan Cantle, President & Founder
a.cantle@nallatech.com

www.nallatech.com

Nallatech
Our Expertise. Your Success.

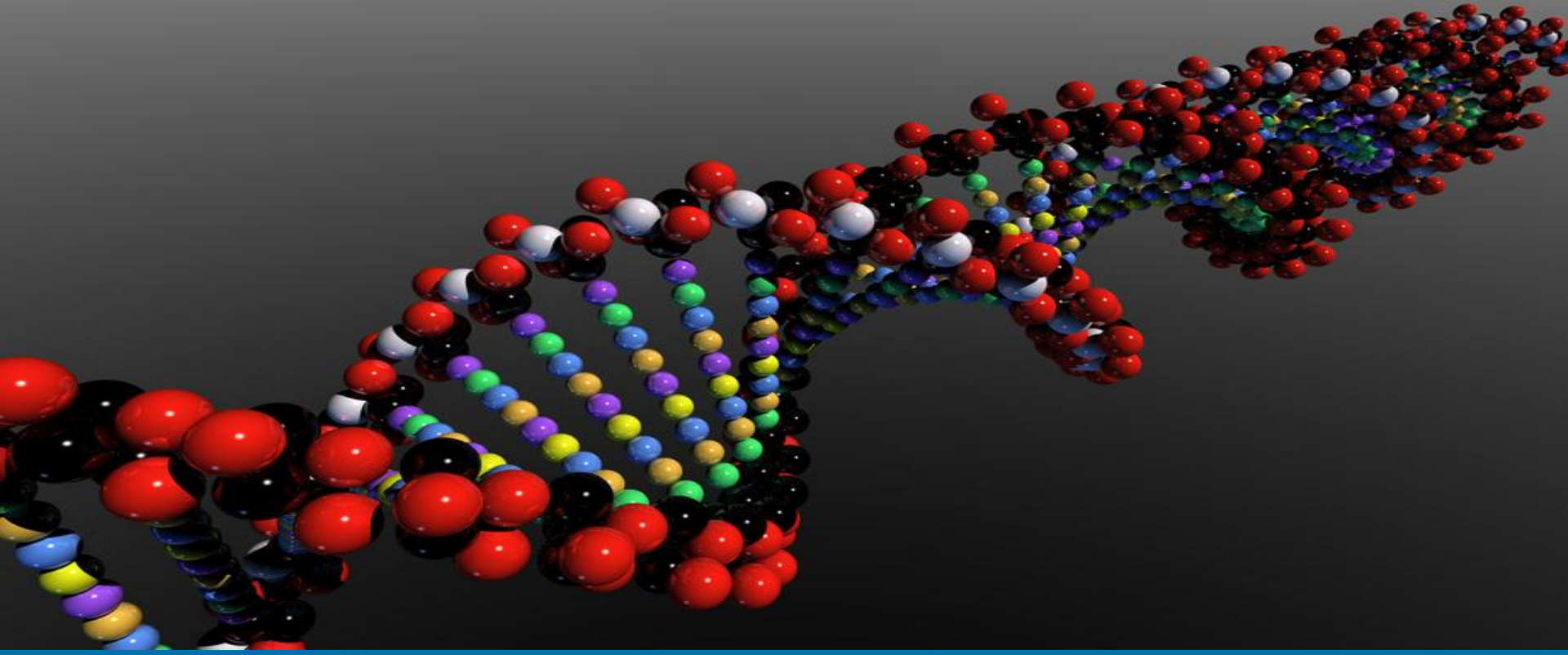


Stating the bleeding Obvious

Allan Cantle, President & Founder
a.cantle@nallatech.com

www.nallatech.com

Nallatech
Our Expertise. Your Success.

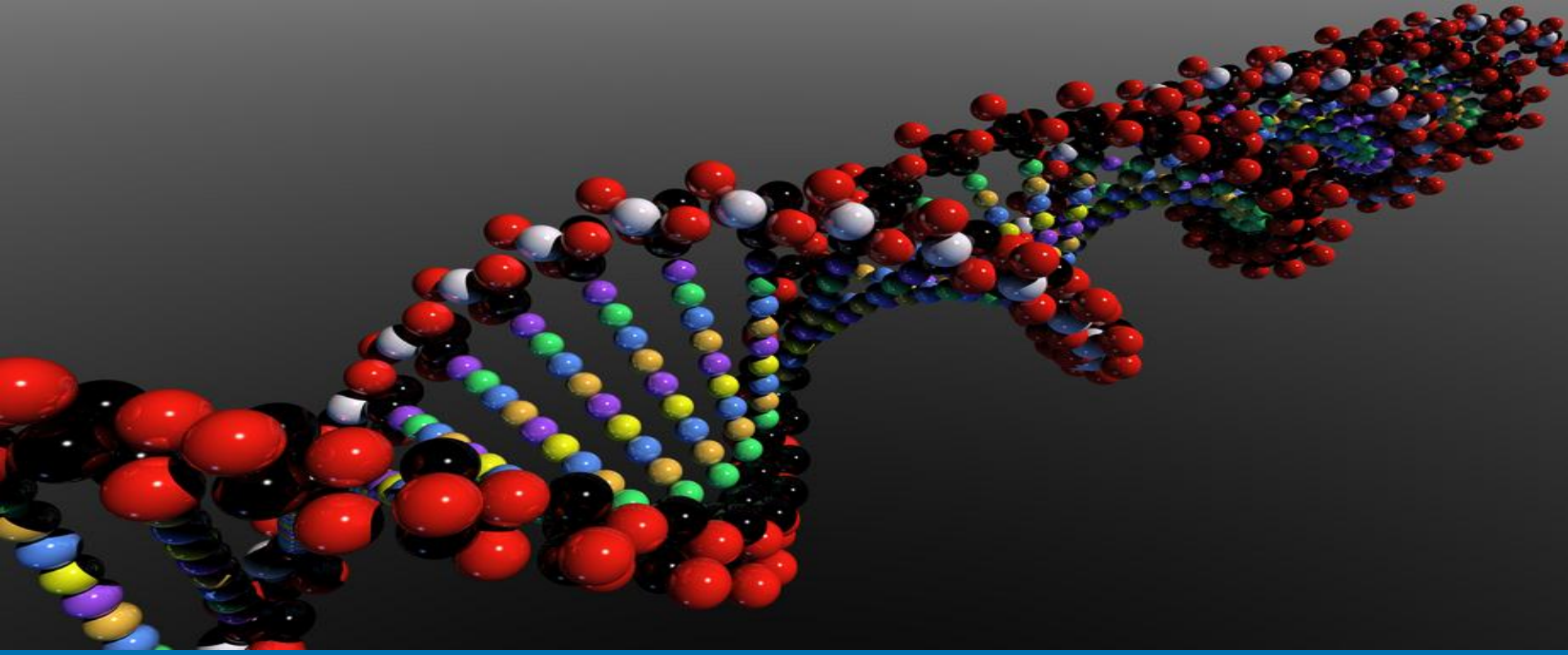


HPEC Fundamentals for HPC Engineers

Allan Cantle, President & Founder
a.cantle@nallatech.com

www.nallatech.com

Nallatech
Our Expertise. Your Success.



The Importance of Data Centricity in an Increasingly Parallel Computing World

Allan Cantle, President & Founder
a.cantle@nallatech.com

www.nallatech.com

Nallatech
Our Expertise. Your Success.

Agenda

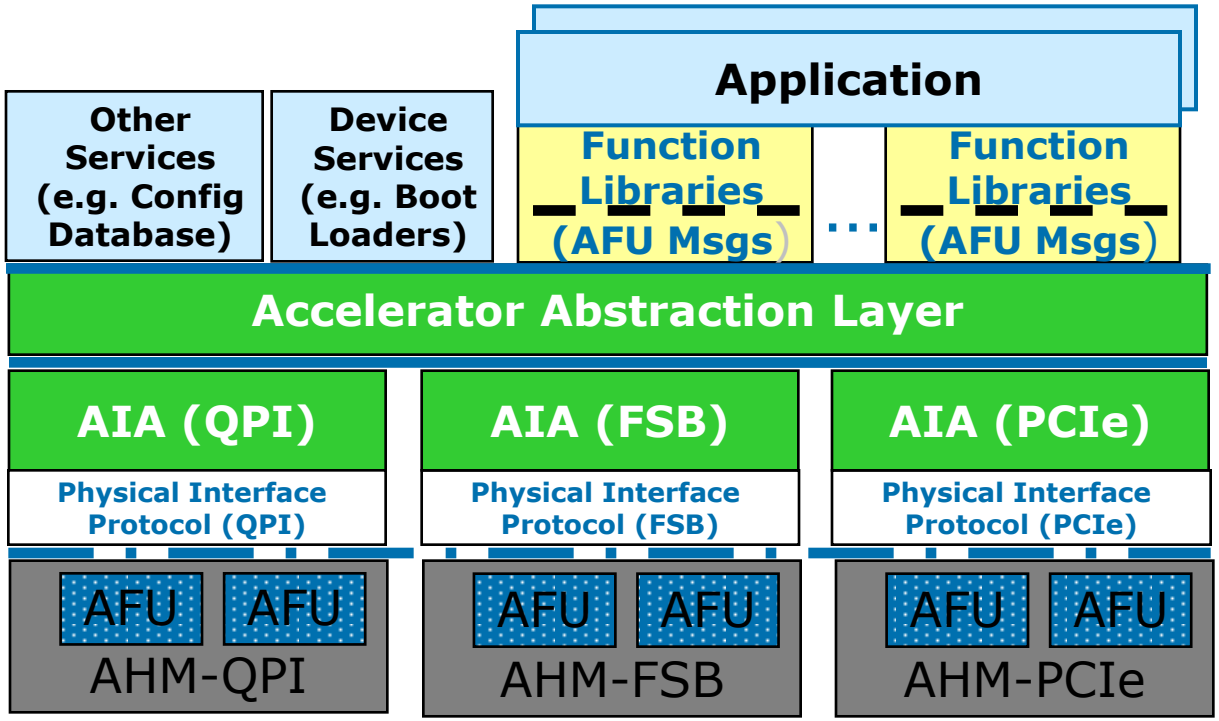
- Review – RSSI 2007
- Data Centric changes between FSB & QPI
- Peak versus Achievable bus bandwidths
- Quickassist Accelerator Model
- Bandwidth Issues when Accelerating Library Functions
- What Nallatech can do to help

Nallatech
Our Expertise. Your Success.

Review – RSSI 2007(Cantle)

- Two Fundamental Computing Models
  - Processor Centric
    - Bring all the data to the processor
    - Von-Neumann
    - X86 IA arguably the leading processor centric architecture today
  - Data Centric
    - Build the processor around the data
    - Data flow
    - FPGAs arguably the best Data Centric processors

Review – Intel's Quickassist Accelerator Model



Common Function APIs
e.g., MKL, IPP

Service APIs

Uniform Accelerator Interface

Accelerator Interface Adapters

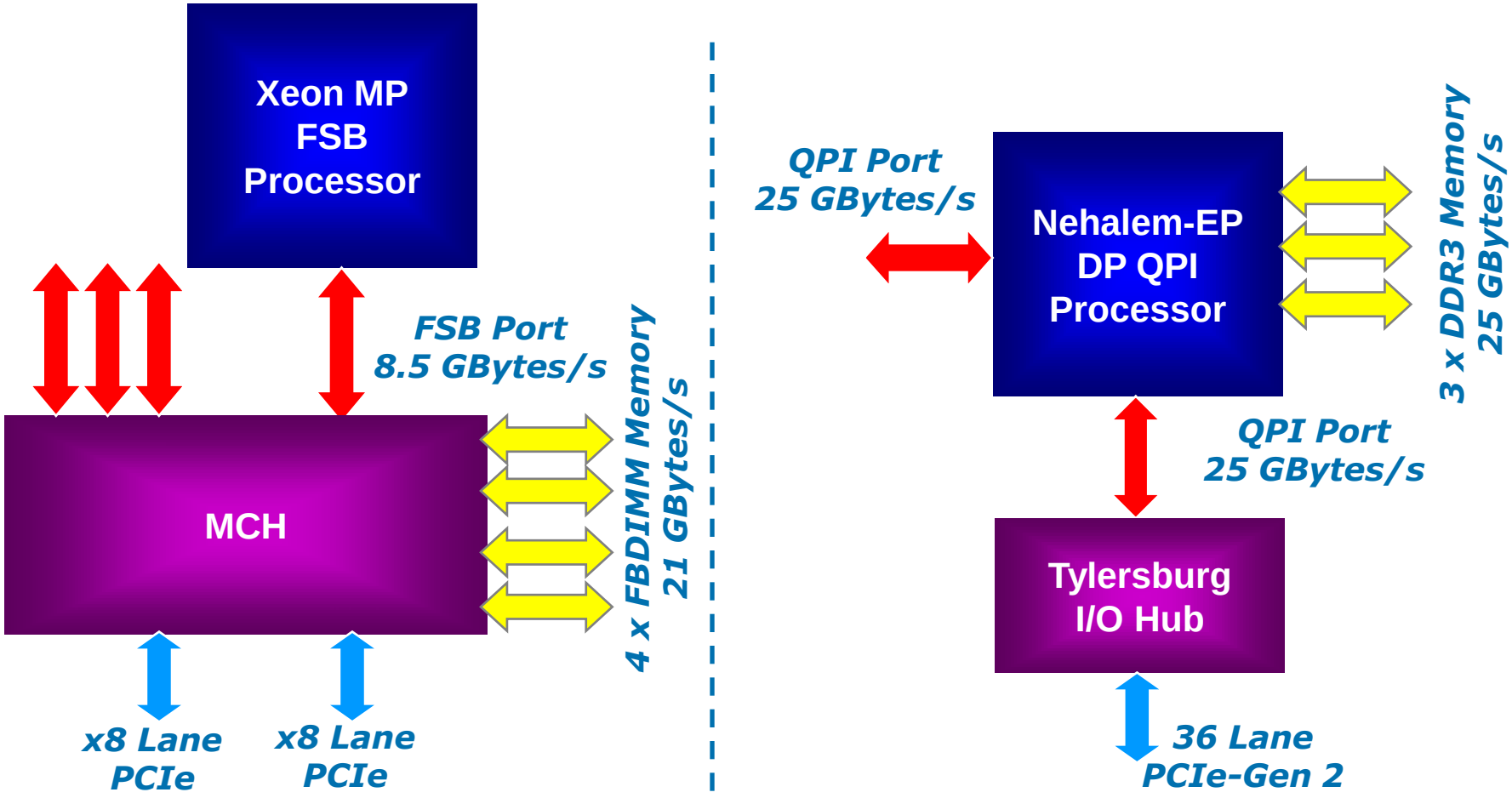
Physical Interconnect

Accelerators

Common function library over
FSB, QPI, PCIe or traditional IA based algorithms

Nallatech
Our Expertise. Your Success.

Review - FSB to QPI Architectural Changes

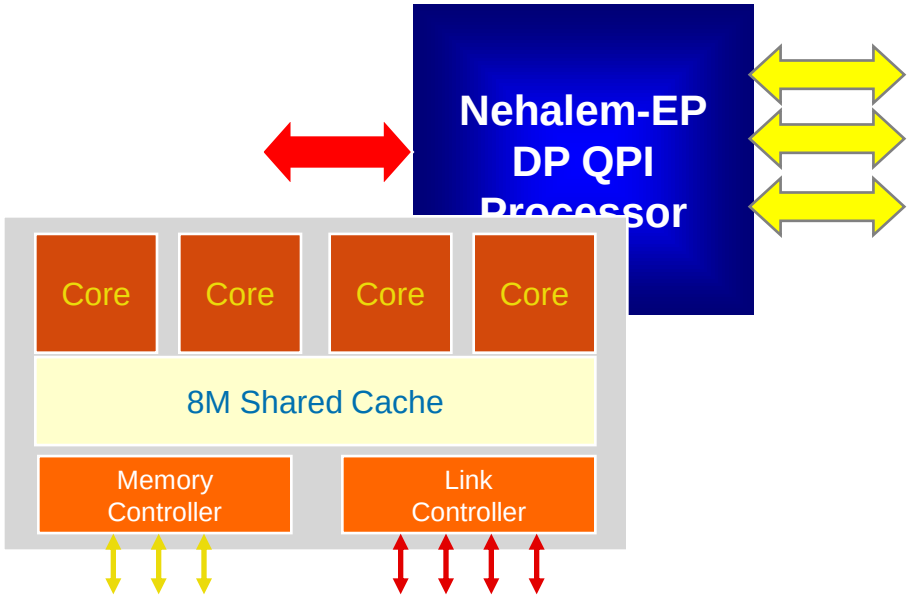


8.5 GBytes/s = Aggregate Theoretical Processor off chip Bandwidth = 75 GBytes/s

Nallatech
Our Expertise. Your Success.

QPI Architecture – Internal BW Considerations

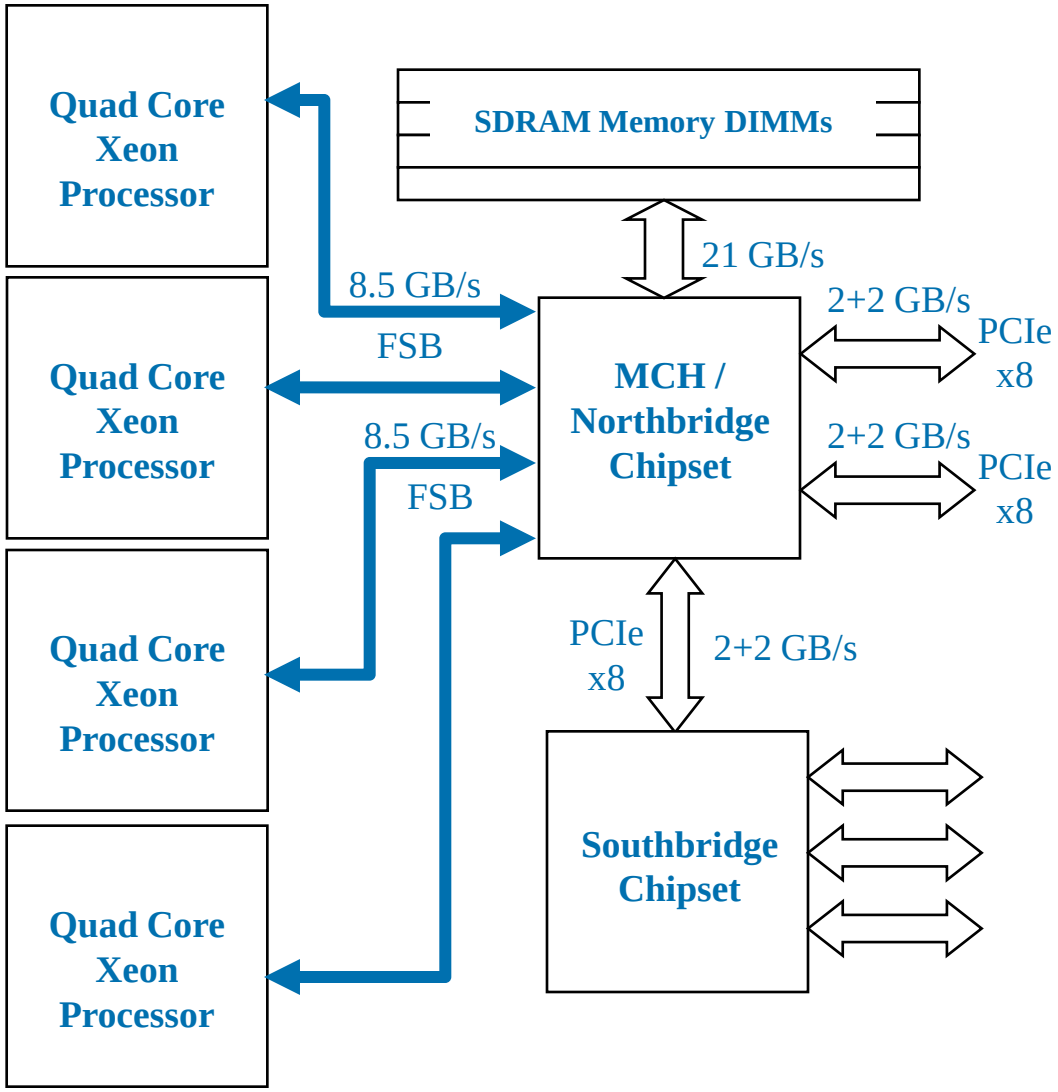
- Link Controller
  - Non Blocking?
  - Pass Thru?
  - BW Into Cache?
- Memory Controller
  - 25GB/s BW to Cache?
- 4 Cores
  - BW to Cache?
- All Roads Lead to Cache?
  - Cache Coherency
  - Potential bottleneck ?
  - I/O to Core BW?
  - Lots of Transistors!



Nallatech
Our Expertise. Your Success.

Achievable FSB Architecture Bus Bandwidths

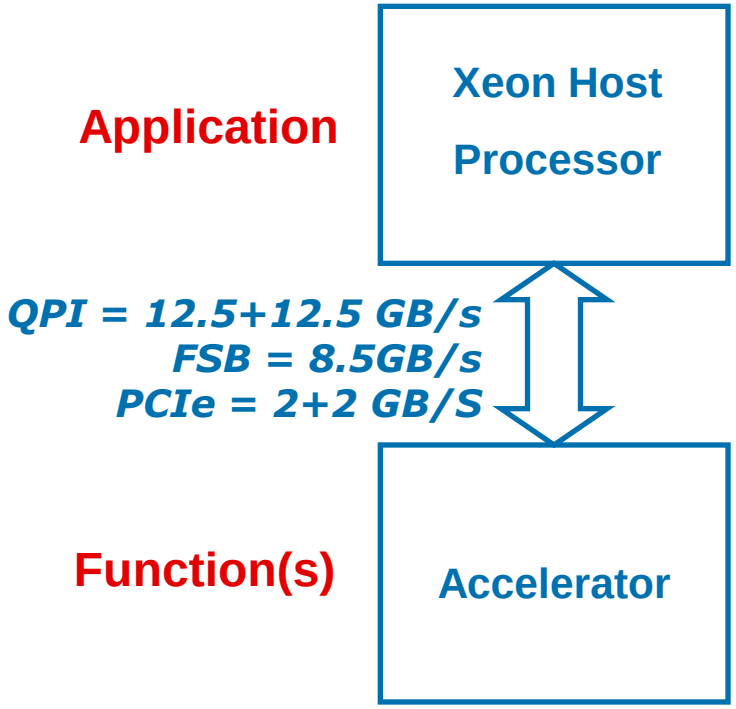
- Stream Benchmark
  - Processor to System Memory
  - Aggregate BW ~9GB/s
  - All 16 threads
  - On 16 cores
- PCIe x8 Gen 1
  - >1+1 GB/s sustained
  - No System Loading!



Nallatech
Our Expertise. Your Success.

Quickassist Accelerator Model

- Function Centric Acceleration
  - Common Function Libraries
  - FFT, RNG, Crypto etc
  - Host calls function on accelerator
  - Generic model for all accelerators
  - Ideally align libraries – e.g. MKL
- Accelerator may support multiple functions simultaneously
- Accelerator Refreshed with new functions
  - New bitstream loaded into FPGA



Nallatech
Our Expertise. Your Success.

Functional Centric Acceleration on FPGAs

FPGA Function I/O Bandwidth requirements & Performance	Single Function		Multiple Functions that will fit into a V5SX240	
	Required I/O Bandwidth GBytes/s	Performance G(Fl)ops	Required I/O Bandwidth GBytes/s	Performance G(Fl)ops
NTBody gravitational (Bandwidth ~ 0)	0	7.5	0	135
Random number (Sobol pipelined)	0.8	614.4	0.8	614.4
Convolution (floating point, Image processing, 0 Frame latency 11x11 kernel)	2.4	72.3	2.4	72.3
Complex FFT (pipelined, single precision, 1024 points)	4.8	30	19.2	120
Black Scholes (single precision)	6.4	14	19.2	42
Convolution (floating point, Image processing, 1 Frame latency 11x11 kernel)	2.4	6.3	28.8	75.6
Pre-Stack-Time-Migration	8	8	80	80
Encryption AES (pipelined)	22.4	NA	268.8	NA

Nallatech
Our Expertise. Your Success.

Functional Centric Acceleration on FPGAs

FPGA Function I/O Bandwidth requirements & Performance	Single Function		Multiple Functions that will fit into a V5SX240	
	Required I/O Bandwidth GBytes/s	Performance G(Fl)ops	Required I/O Bandwidth GBytes/s	Performance G(Fl)ops
NTBody gravitational (Bandwidth ~ 0)	0	7.5	0	135
Random number (Sobol pipelined)	0.8	614.4	0.8	614.4
Convolution (floating point, Image processing, 0 Frame latency 11x11 kernel)	2.4	72.3	2.4	72.3
Complex FFT (pipelined, single precision, 1024 points)	4.8	30	19.2	120
Black Scholes (single precision)	6.4	14	19.2	42
Convolution (floating point, Image processing, 1 Frame latency 11x11 kernel)	2.4	6.3	28.8	75.6
Pre-Stack-Time-Migration	8	8	80	80
Encryption AES (pipelined)	22.4	NA	268.8	NA

= I/O BW within PCIex8 & FSB Peak Bandwidth Capability

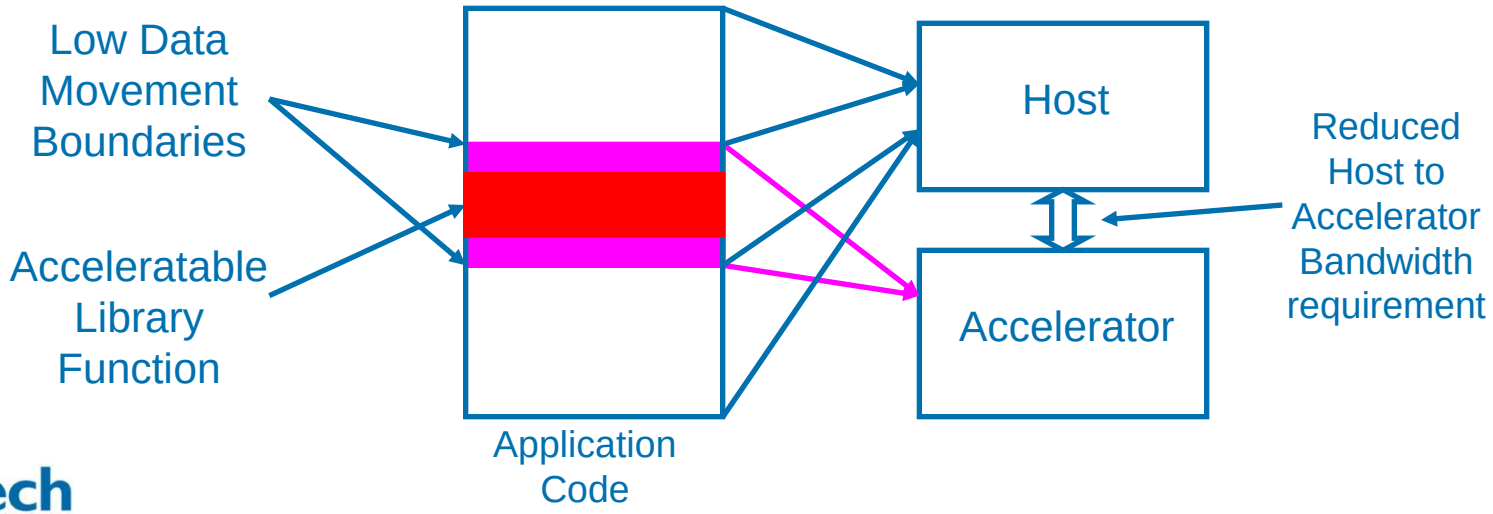
= I/O BW within QPI Peak Bandwidth Capability

= I/O BW not achievable (Acceleration is Data Bound)

Nallatech
Our Expertise. Your Success.

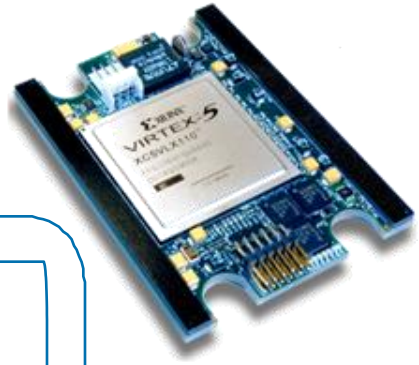
Functional Centric Acceleration on FPGAs

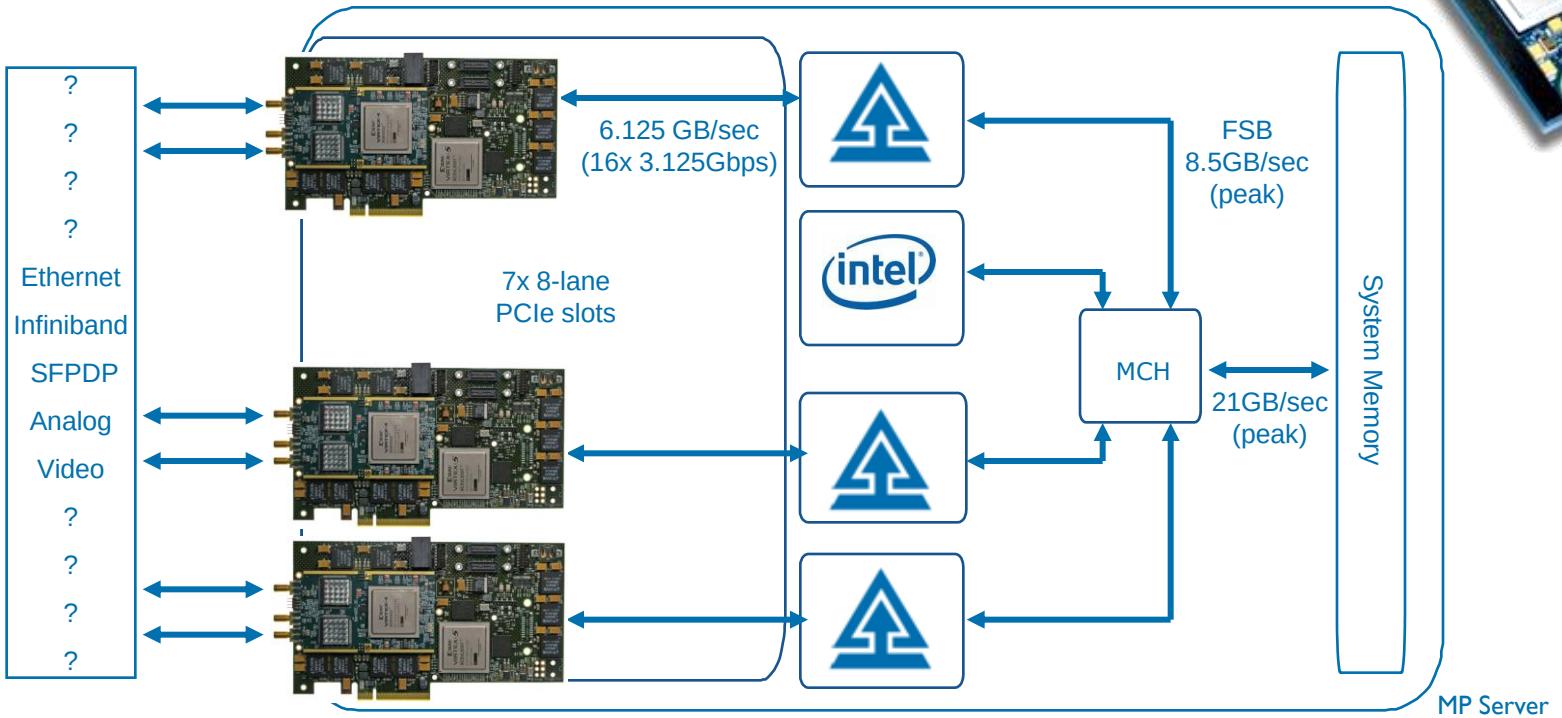
- Serious Issue with Accelerators being Data Bound
  - Even with QPI Bandwidths
  - Exacerbated by partitioning at the Library Function Interface
  - Actual Host to Accelerator Bandwidths will be less than peak
- Solution
  - Partition the accelerator at the point of least Data Movement in the application whilst encapsulating the Acceleratable Function



Nallatech
Our Expertise. Your Success.

Nallatech's Data Centric FPGA Computing Solutions





Nallatech
Our Expertise. Your Success.

Nallatech's Consultancy & Services Group

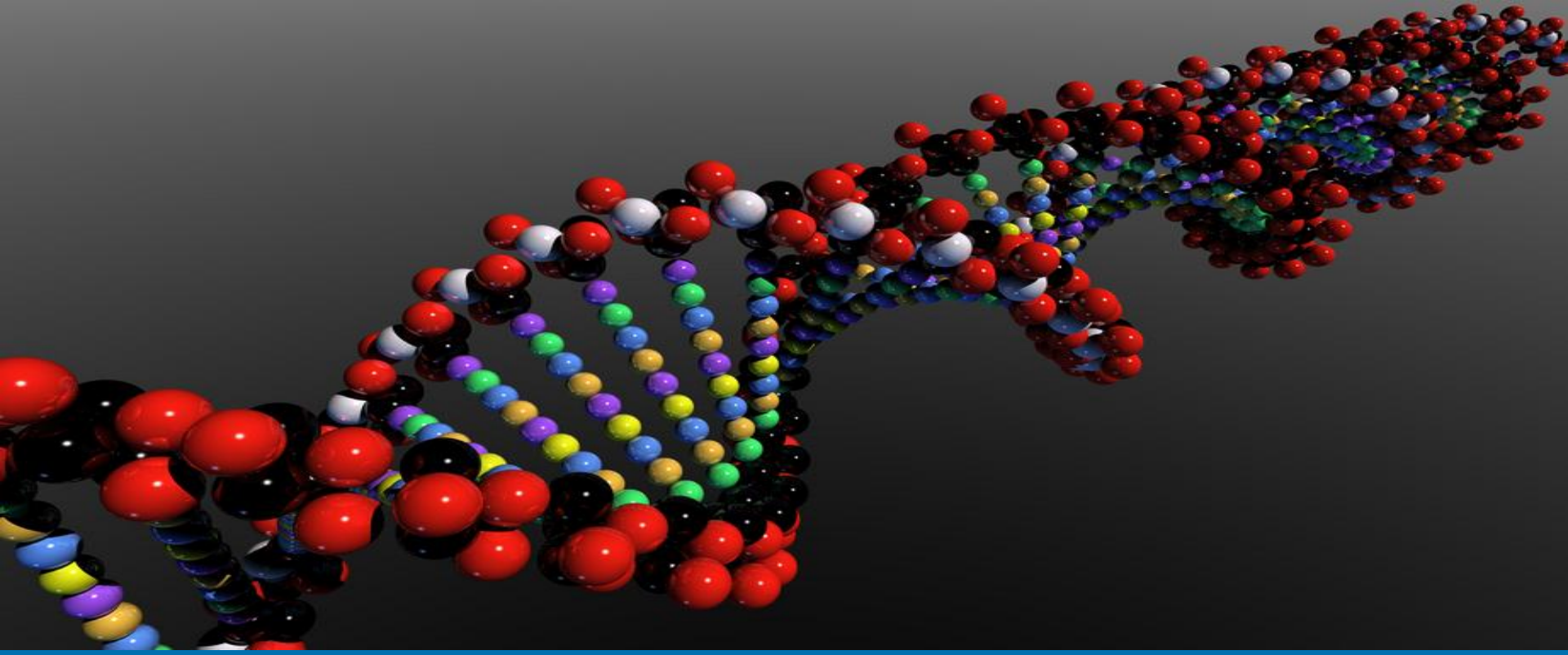
- Customers wishing to utilise accelerators often require support in getting up to speed with the challenges that they bring.
- Nallatech is actively developing this part of its business
  - Feasibility studies – Can your application benefit from accelerators?
  - Code Porting Services
  - Bespoke Training in Acceleration and Parallel concepts
- Additionally looking for Consulting Partners that have specific FPGA acceleration expertise and/or have vertical market expertise.

Nallatech
Our Expertise. Your Success.

Conclusions

- Careful thought should be put into assessing the viability of an accelerator for your application
  - Minimise data movement between processors and accelerators
- Fast Functions tend to need fast data feeds
- Nallatech's range of products cater for a wide range of Data Centric solutions that are compatible with Intel standard servers
  - PCIe, FSB, QPI, I/O
- Nallatech can help with your Accelerator success

Nallatech
Our Expertise. Your Success.



www.nallatech.com

Nallatech
Our Expertise. Your Success.