



*The Decomposition of HPEC Applications Mapped to  
the Natural Decomposition of a Solution Architecture*

# Historical Solution Drivers for HPEC Applications

## ***Circa 1980 - 1990***

- ? Specialized Co-Processors Tightly Coupled to the Host CPU
- ? Dual-Ported Memories; albeit very Small Amounts
- ? Data Acquisition via Host CPU Bus

## ***Circa 1990 - 1995***

- ? First Wave of Small-Count Multi-Processor Embedded Applications
- ? Localized Non-Shared Memory
- ? Low Bandwidth, Low Functionality Bus-Based Interconnects
- ? Data Acquisition via Direct Parallel I/O Ports on Each Card

## ***Circa 1995 - 2003***

- ? First Wave of Larger-Count Multi-Processor Applications
- ? High Bandwidth, Crossbar Based Interconnects
- ? Lots of Processors and Lots of Small Distributed Memory Pieces
- ? Data Acquisition via Direct Fabric Based Ports on Each Card
- ? The era of the Compact Application Benchmark

# ***Historical Trends for Application Fitment to Hardware***

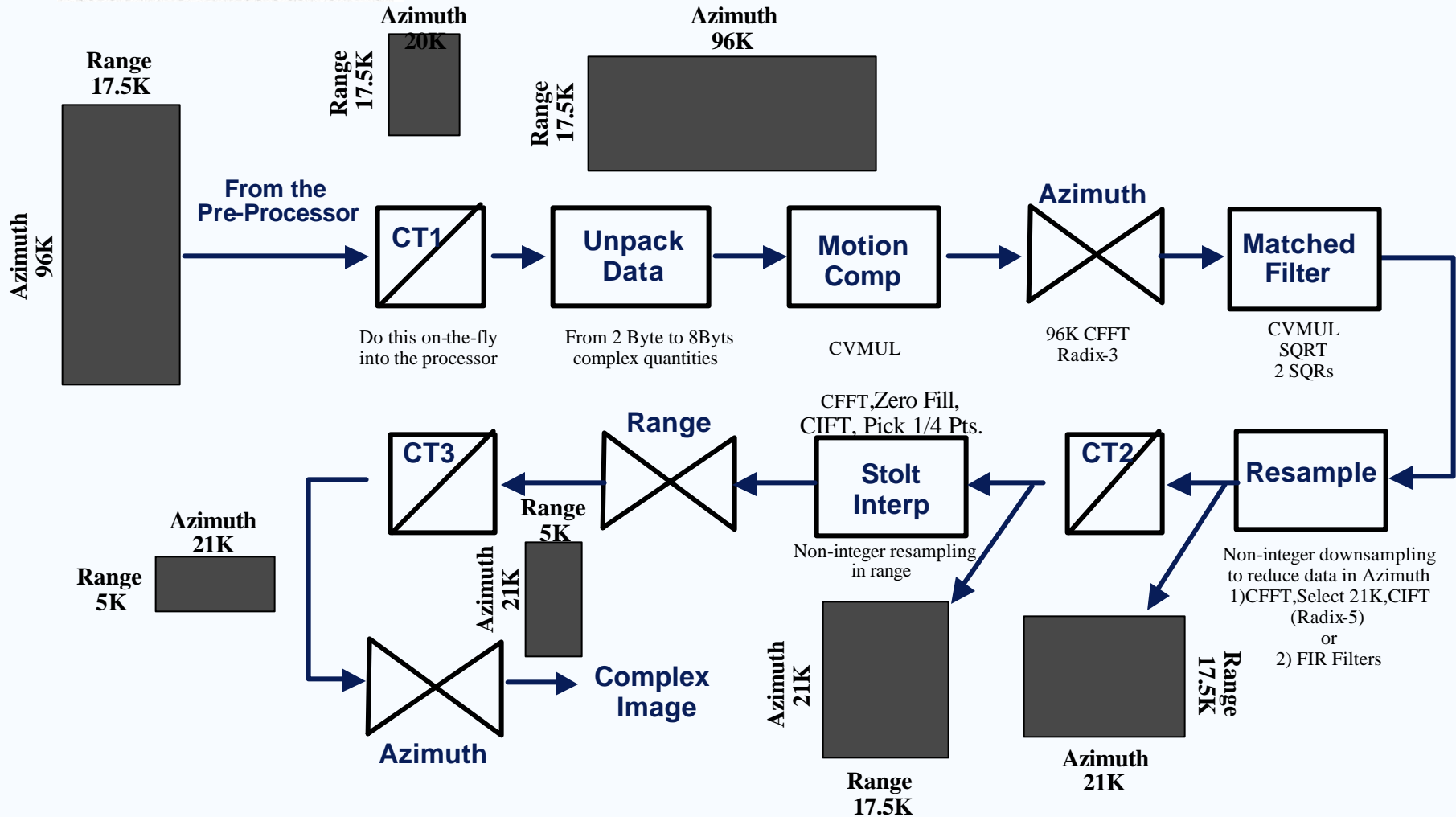
## **It was .....**

- ? Finesse the I/O *and* Computations to Fit the Computer's Architecture which included Integrated I/O and Computations
- ? Buffer, Rearrange, Move, and Process the Data *In-Fabric* using Widely-Distributed Small Buffers of Memory
- ? Tightly Couple the Application's Architecture to the Distributed Memory and Computer's Architecture

## **The Community had an Approach to Understand Behavior and Performance Estimates for Complex HPEC Systems**

- ? Standard Benchmarks: 2DFFT, Corner Turns, STAP, etc
- ? Help it with Middleware: MPI/RT, Data Re-Org, VSIPL
- ? But they did not really address the I/O and the impacts thereto to the Overall Processing and Processing Management of Data

# The FOPEN Example



**Note 1: To Preserve Memory pack and unpack the data before and after corner turns**  
**Note 2: To Preserve and Bandwidth Convert to/from float and Integer**

# The Architecture Problem

## ***Small Chunks of Distributed Memory can be Problematic***

- ? Lots of Small Memory Buckets = Lots of Data Movement
- ? Lots of Data Movement = Lots of Bandwidth Needed
- ? Hence the Problem: The Memory becomes *in-fabric*

## ***Small Memory Buckets are Challenging to Manage***

- ? Data Feeds are Scattered among the Fabric
- ? Partial Data Sets are Unnaturally Broken Up
- ? Many times, way too Much Scatter, Gather, and Re Organization

## ***HPEC Problems Naturally Decompose into Two Key Areas***

1. Data Acquisition, Buffering, and Re-distribution
2. High Speed and Highly Complex Computations on Well Bounded Data Sets utilizing Well Bounded Algorithms

**They have very different Architectural Needs if they are to be Optimally Served**



# What's Different in Today's Computer Architectures

- ✍ Processor-to-Memory Bandwidth is Huge*
  - ? 64-bit Wide DDR; > 3 GB/Sec memory access speed is possible
  - ? Memory is Cheap and Abundant
- ✍ I/O and Local System Bus Bandwidth is Very High*
  - ? Commodity Busses e.g. PCI-X > 1GB/Sec
  - ? Lots of Peripherals, Lots of Available Software (driver) Support
- ✍ Interconnect Fabrics are FAST and SMART*
  - ? 1GB/Sec per Port
  - ? Self-Discovery, Fault Detection, Recovery, Reliable
- ✍ Standard Processors are Readily Available; Specialized Devices are Becoming Easier to Use*
  - ? High Throughput SIMD's, DSP's, Fast Server-Centric Processors
  - ? FPGA's, ASIC's, and Other Custom Logic
  
- ✍ Today, it is Easier to Balance I/O and Computational Needs at the Computer Level rather than at the Application Level*

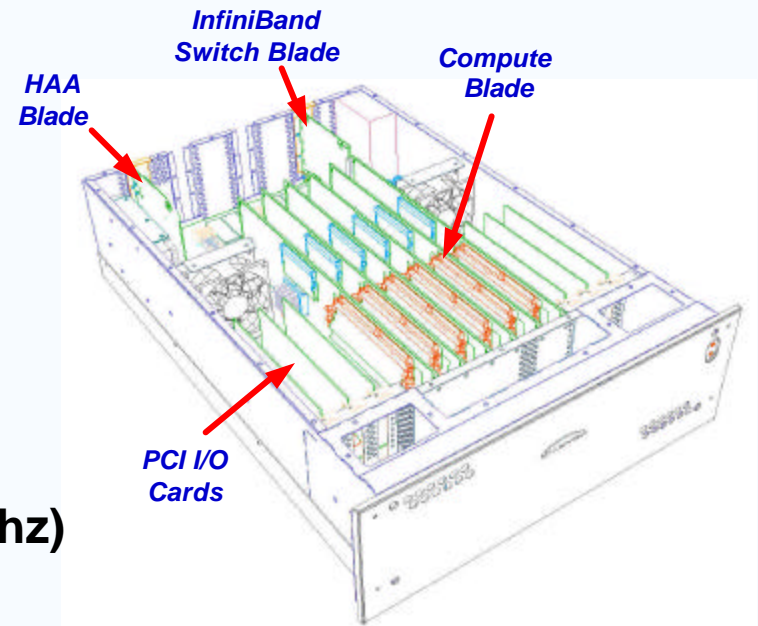
# Data Acquisition Server

## SMARTpac 600

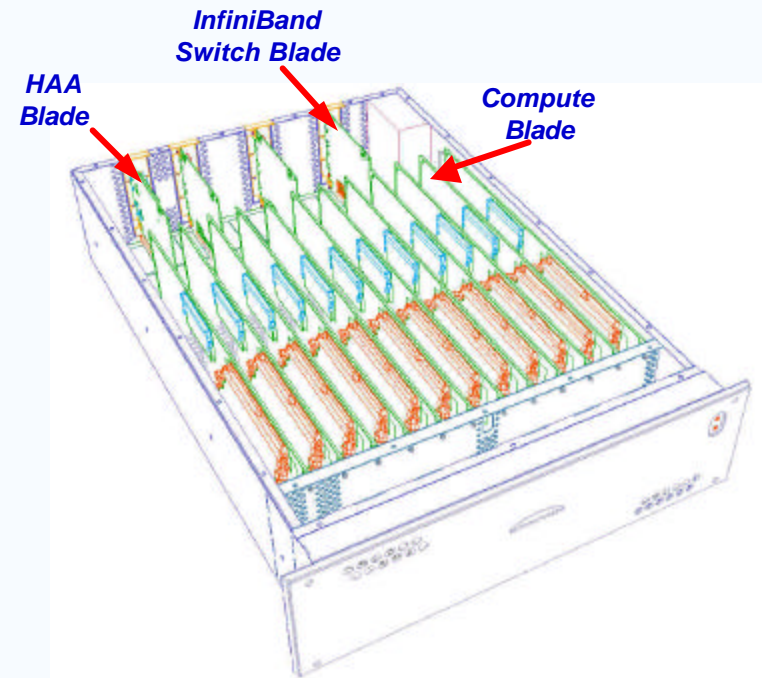


**Optimized for Data Acquisition and Data Management Services**

- ✎ Six PCI-X adapter slots (PCI 32-bit/33Mhz up to PCI-X 64 bit/100Mhz)
- ✎ Six compute blade slots for single or dual-processor compute blades
- ✎ Six InfiniBand connections at rear of chassis
- ✎ Two RJ45 Ethernet connections for High Application Availability Infrastructure



# Compute Server SMARTpac 1200

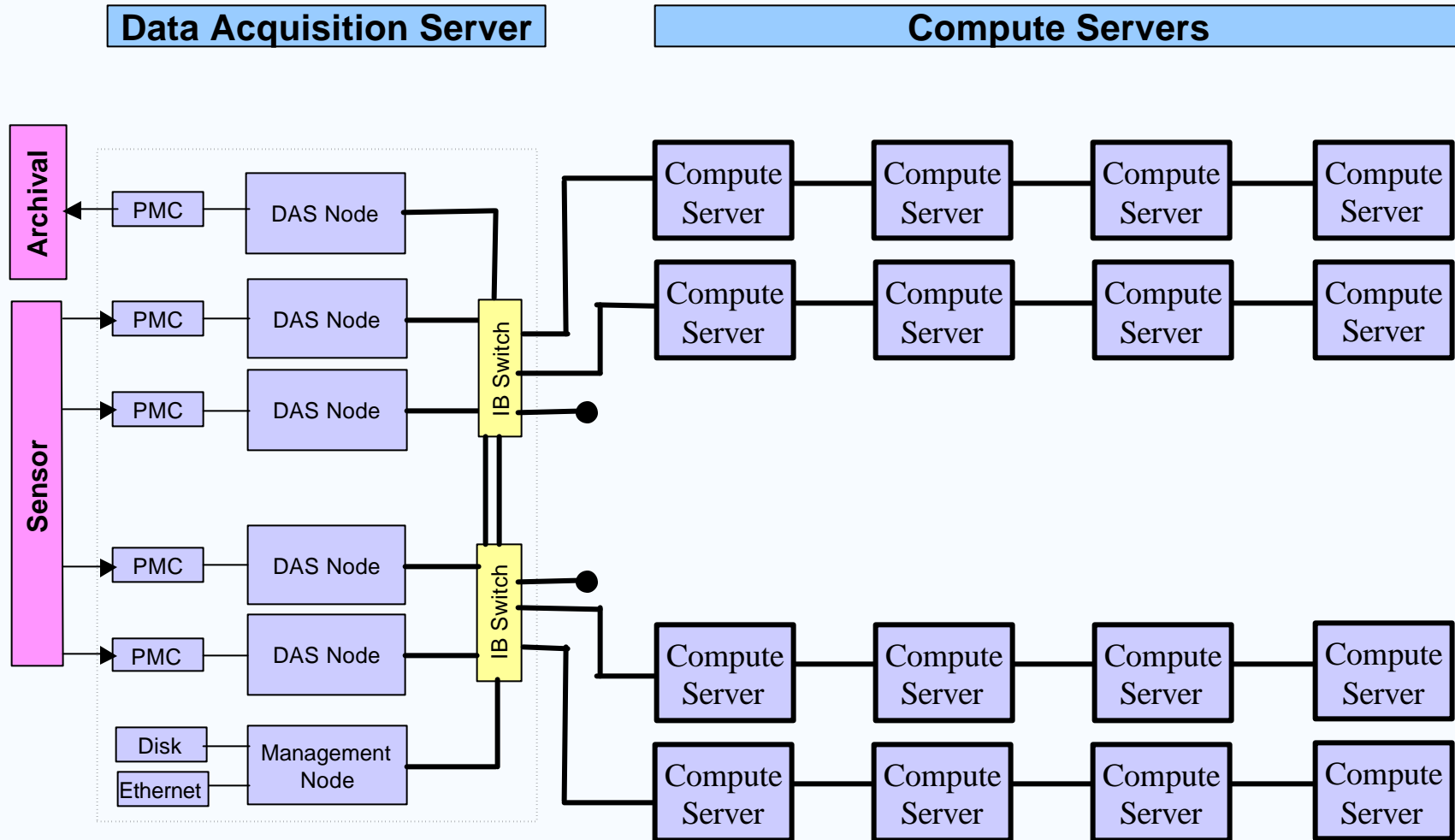


## Optimized for Computational Performance per \$/Watt/Area

- ✍ Twelve compute blade slots for single, dual-processor, or special function compute blades
- ✍ Six InfiniBand connections at rear of chassis
- ✍ Two RJ45 Ethernet connections for High Application Availability Infrastructure

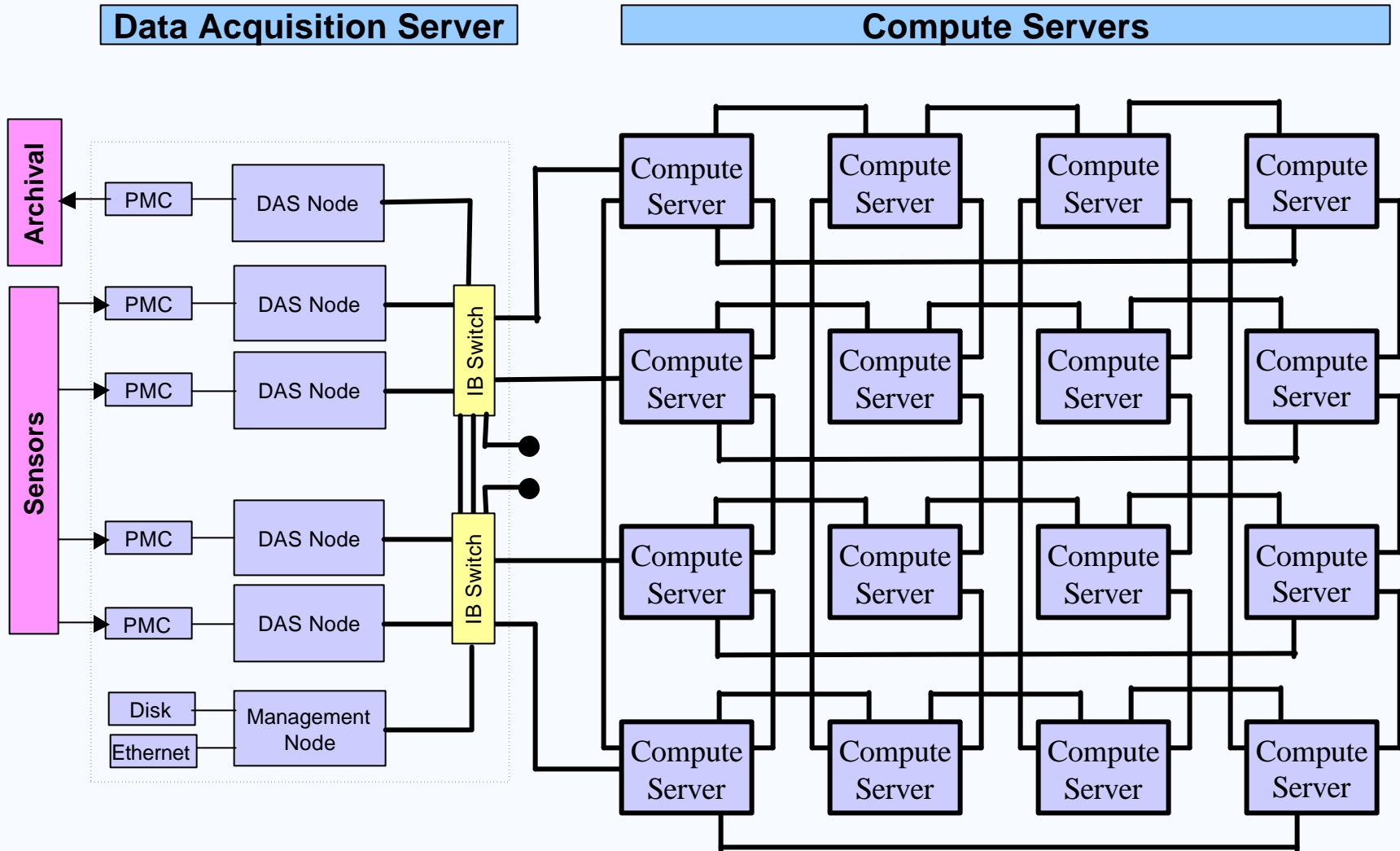
# Piped Connection

## Medium Bandwidth – Lower Complexity



# Fully Connection Mesh

## High Bandwidth – Higher Complexity



# *Decomposition Summary*

- ✍ **Most HPEC Problems Naturally Decompose into:**
  1. Data Acquisition and Management Services, and
  2. Computational Services
  
- ✍ **The Current HPEC Systems built around Raceway™ , SKYchannel™, and Myranet™ are representative of an “older school” approach, whereby I/O and Computes are tightly coupled, physically bound together, and use Small Buckets of Fabric-based Shared memory**
  
- ✍ **Today’s Technologies allow one to Think and Actually Implement Differently to Meet an Application’s Actual Decomposed I/O and Processing Needs**
  - ? **Data Acquisition Servers** optimized for I/O Services, Data Buffering, Data Management, and Data Distribution
  - ? **Compute Servers** optimized for Signal and Image Processing