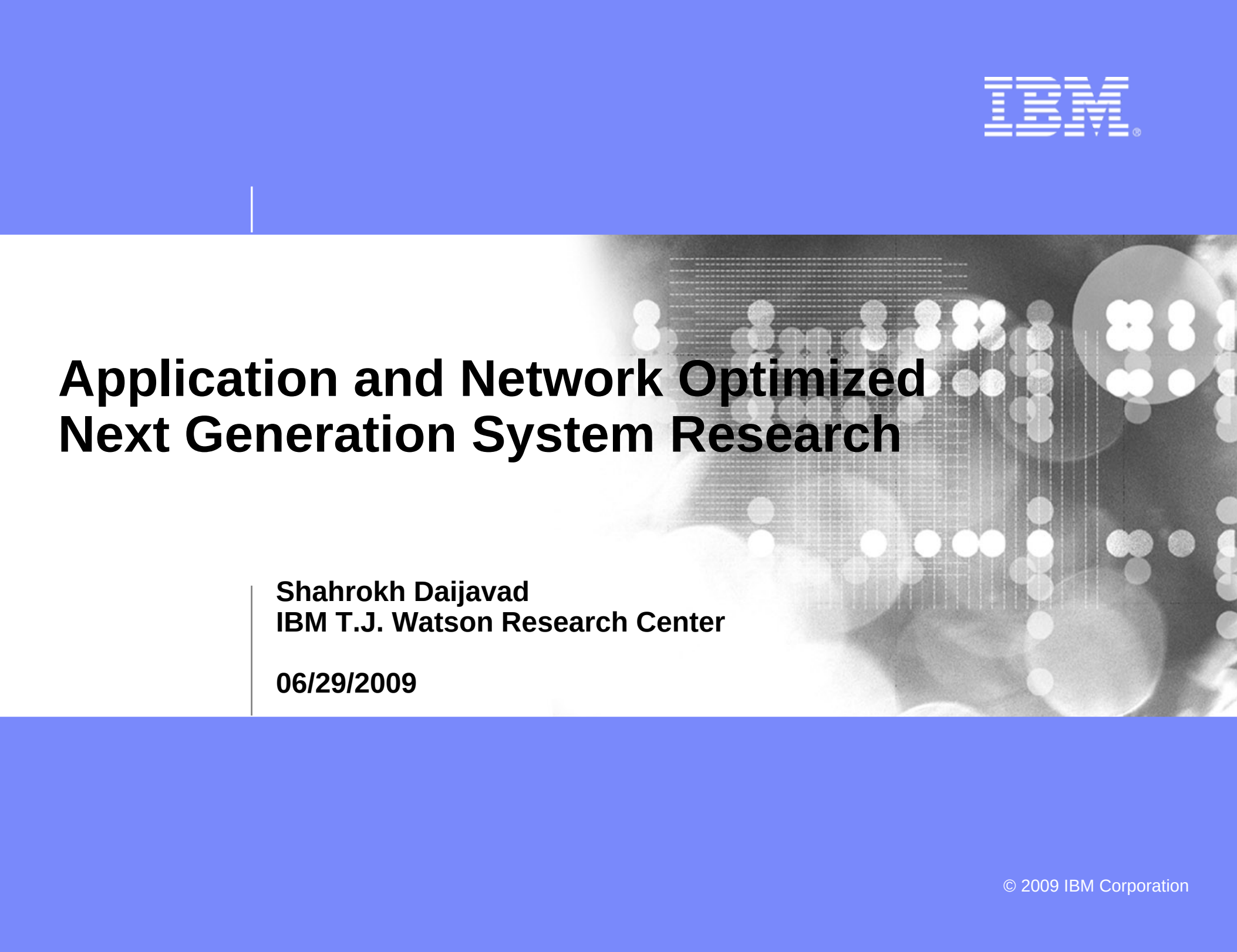
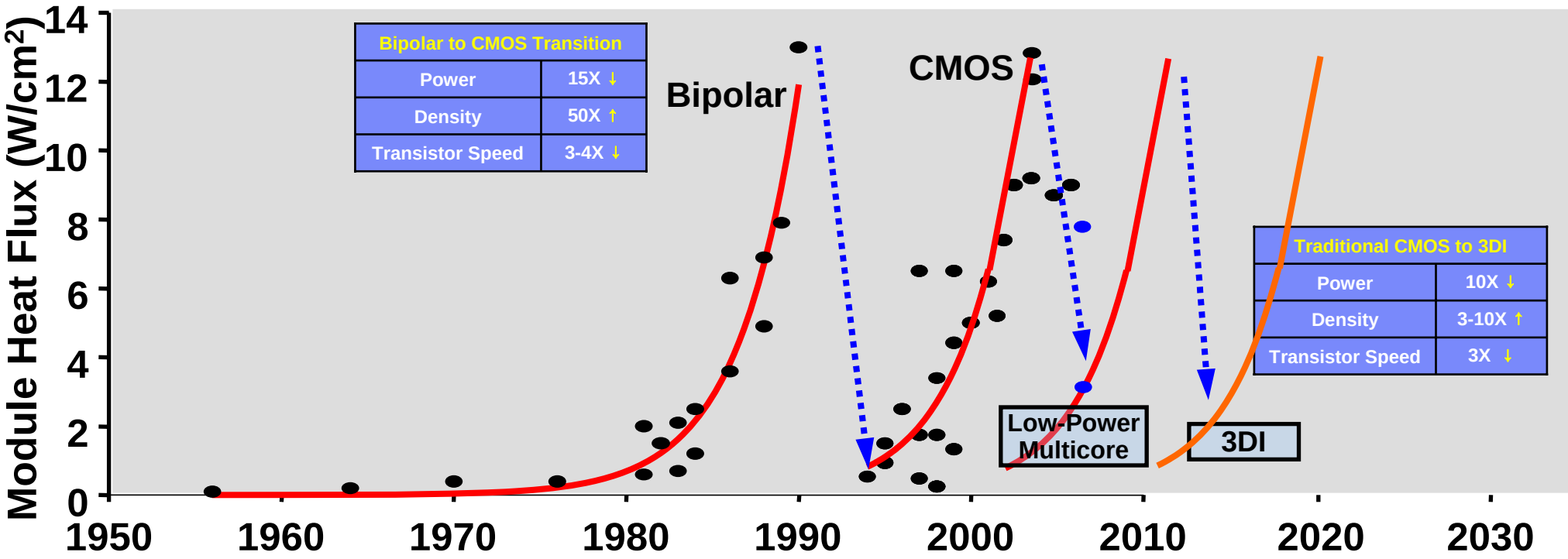


Application and Network Optimized Next Generation System Research

Shahrokh Daijavad
IBM T.J. Watson Research Center

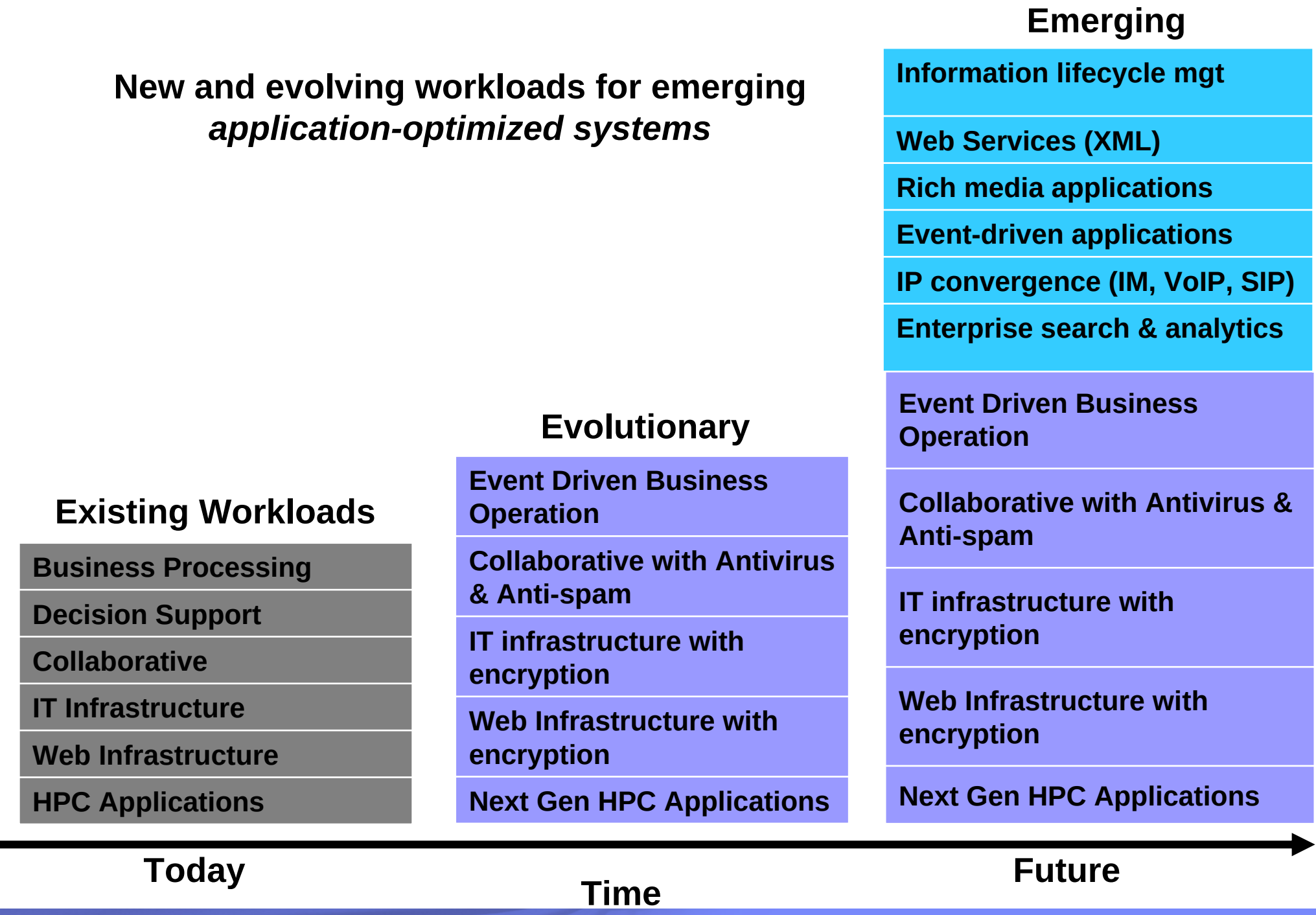
06/29/2009

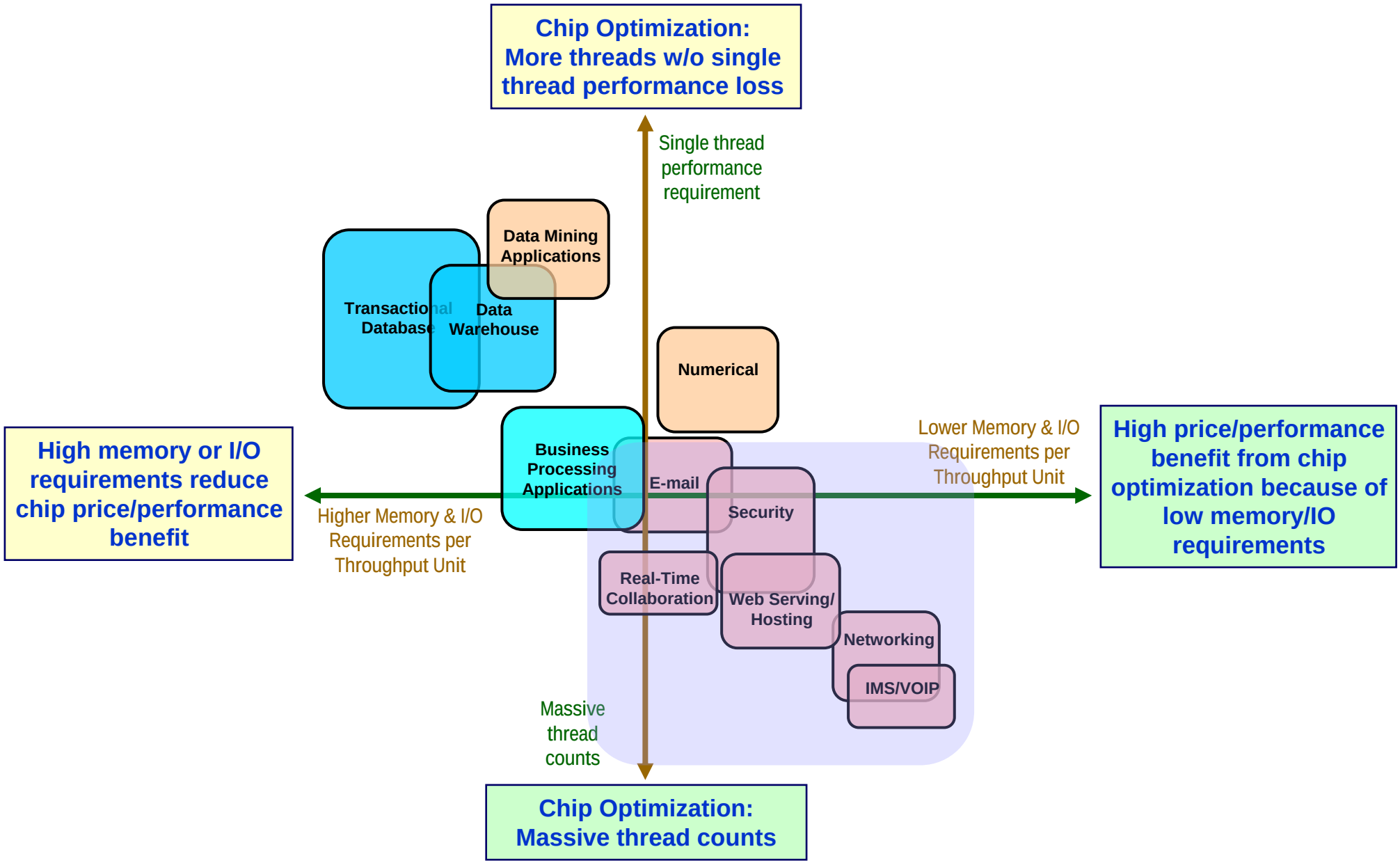


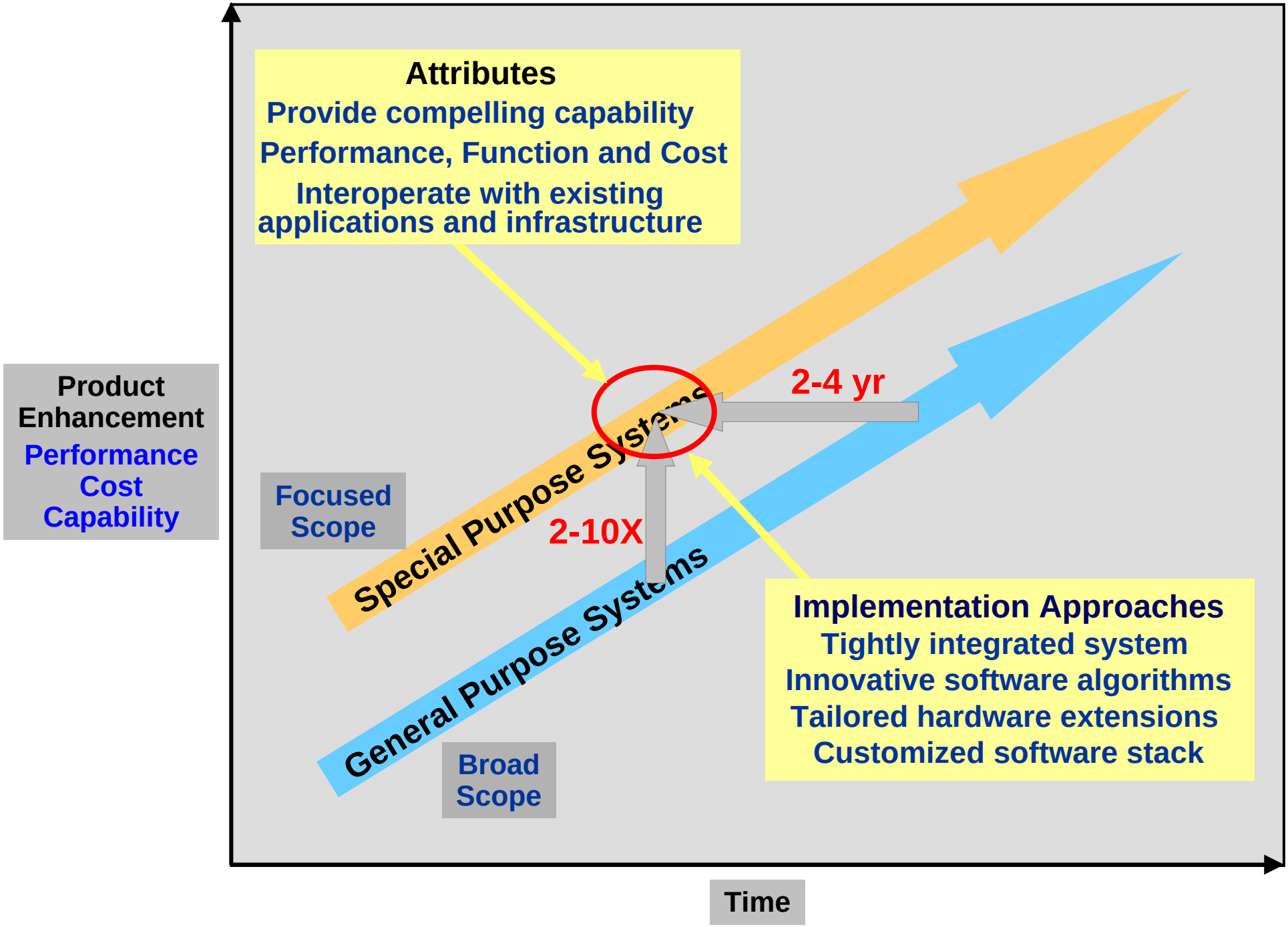
Trends:

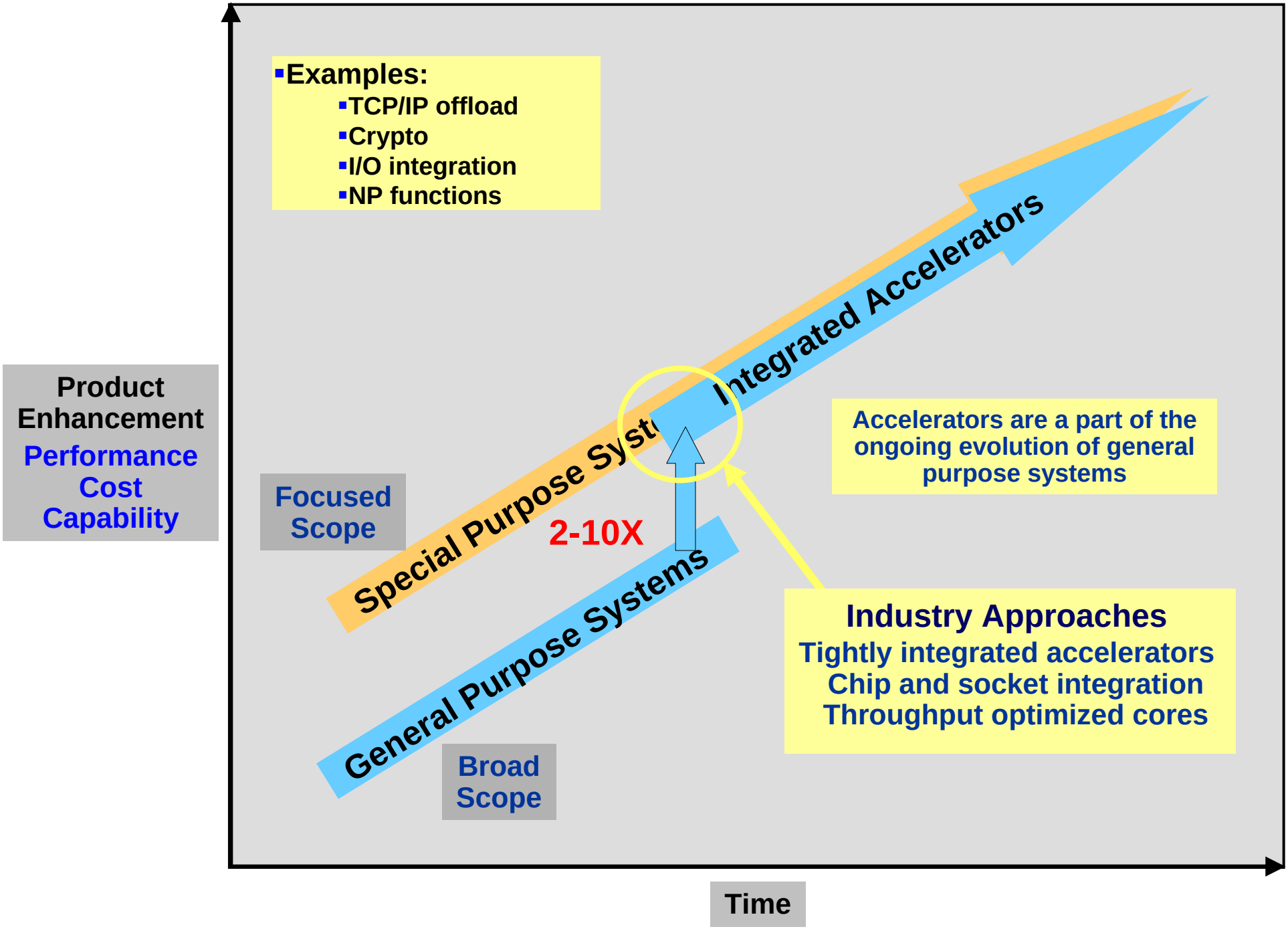
- **Highly parallel homogeneous / heterogeneous systems built with multiple small processors**
  - Weaker single-thread performance, good chip-level throughput performance, and excellent power-performance
- **Increasing cores / die, threads / core, transistors / chip, and virtualization containers will saturate on-chip cache capacity and off-chip bandwidth**
- **Increased computational density drives the need for greater I/O bandwidth and more efficient I/O processing**
- **Application optimized systems and system-level accelerators grow in importance**
- **Parallelism is exploited at all levels of the software stack**

New and evolving workloads for emerging *application-optimized systems*

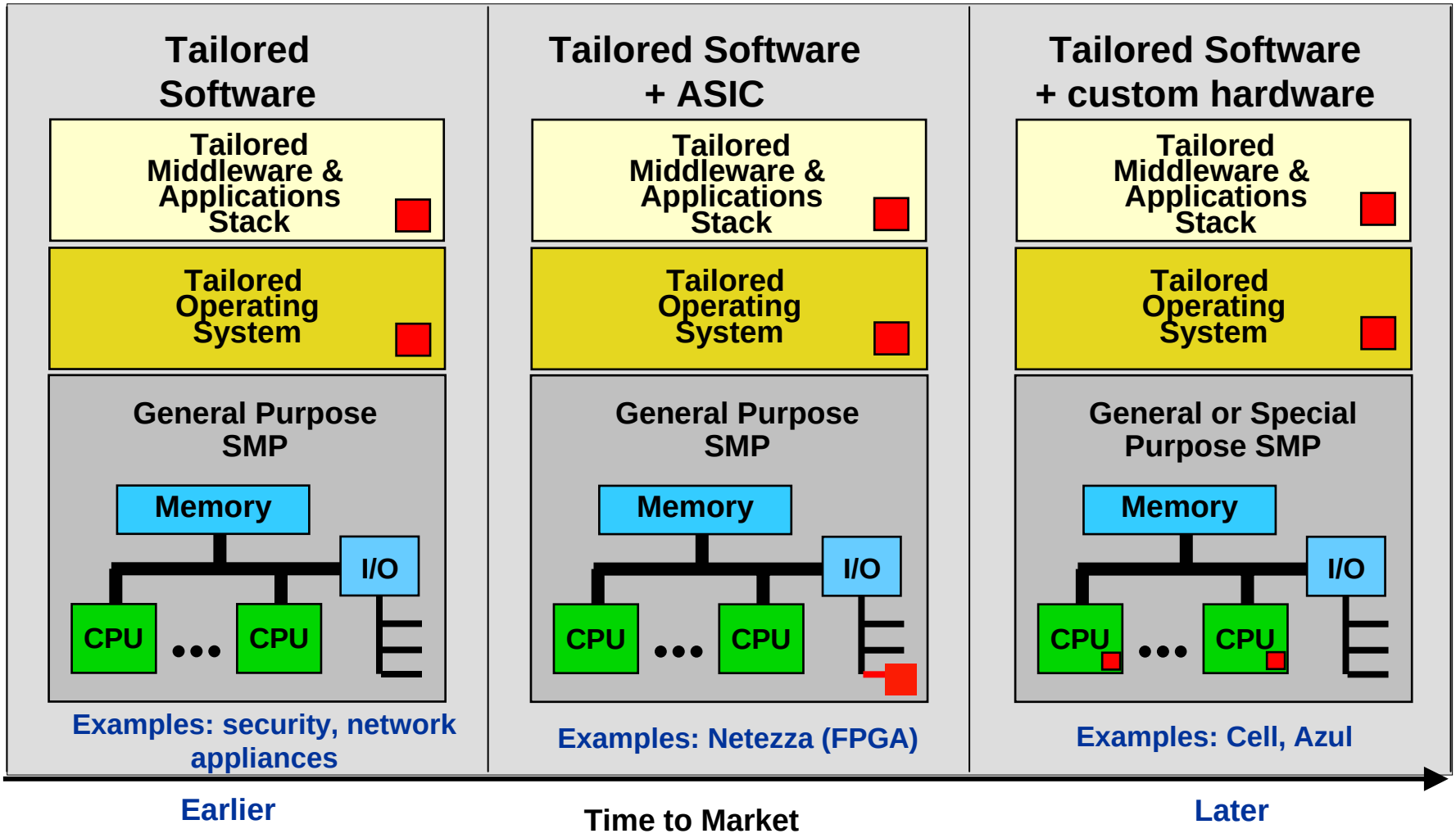



Chip Optimization:
More threads w/o single thread performance loss
Single thread performance requirement
Data Mining Applications
Transactional Database
Data Warehouse
Numerical
Business Processing Applications
E-mail
Security
Real-Time Collaboration
Web Serving/ Hosting
Networking
IMS/VOIP
High memory or I/O requirements reduce chip price/performance benefit
Higher Memory & I/O Requirements per Throughput Unit
Lower Memory & I/O Requirements per Throughput Unit
High price/performance benefit from chip optimization because of low memory/IO requirements
Massive thread counts
Chip Optimization:
Massive thread counts


Attributes
Provide compelling capability
Performance, Function and Cost
Interoperate with existing applications and infrastructure
Product Enhancement
Performance
Cost
Capability
Special Purpose Systems
General Purpose Systems
2-4 yr
2-10X
Focused Scope
Broad Scope
Implementation Approaches
Tightly integrated system
Innovative software algorithms
Tailored hardware extensions
Customized software stack
Time

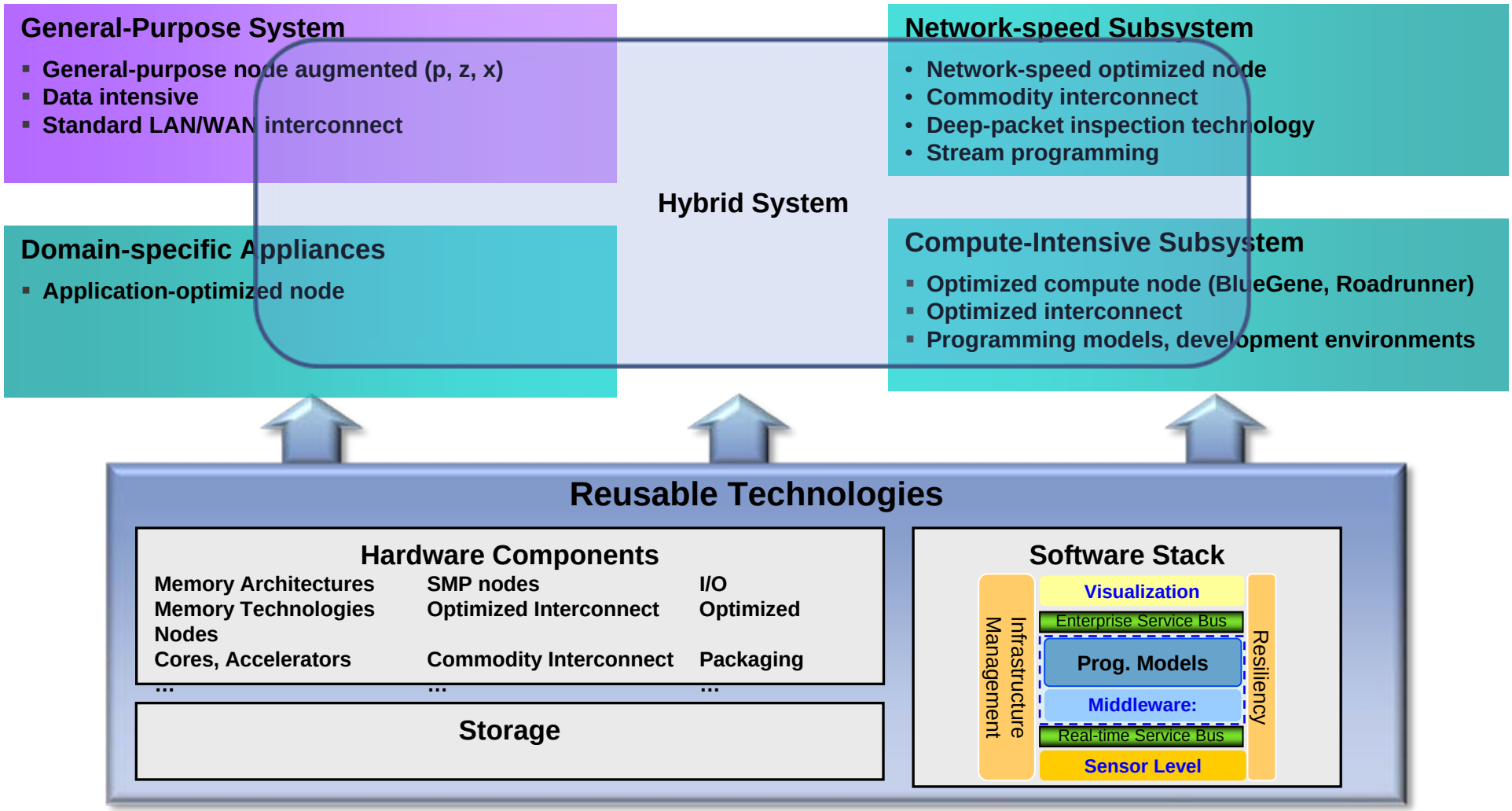

Examples:
TCP/IP offload
Crypto
I/O integration
NP functions
Integrated Accelerators
Special Purpose Syst
General Purpose Systems
2-10X
Accelerators are a part of the ongoing evolution of general purpose systems
Industry Approaches
Tightly integrated accelerators
Chip and socket integration
Throughput optimized cores
Product Enhancement
Performance
Cost
Capability
Focused Scope
Broad Scope
Time

Application Optimization Approaches

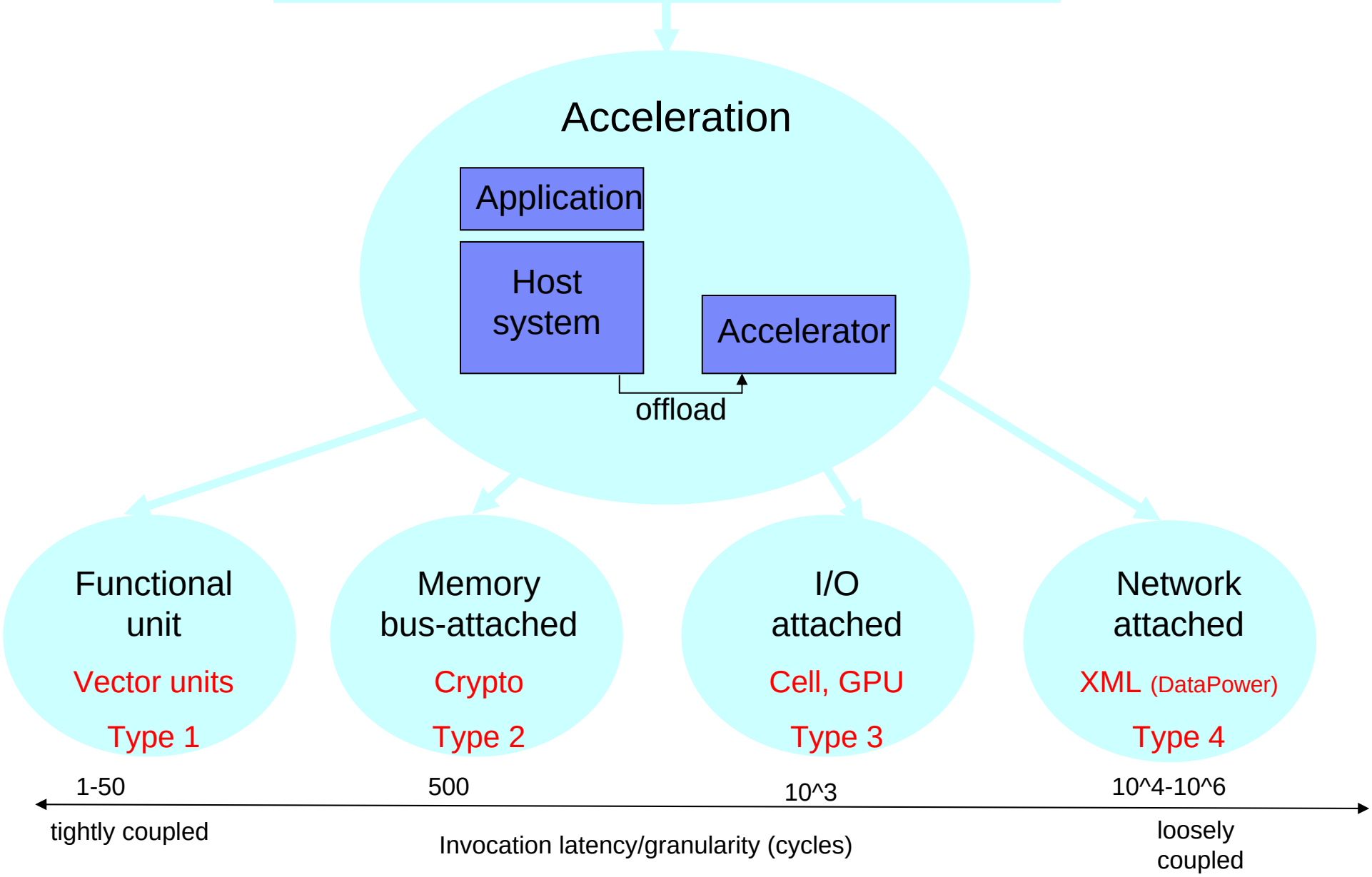


- **Loosely coupled: Traditional Heterogeneous Systems Approach**
  - Each step in a computational work stream is a separate application which runs on a subset of the systems
- **Tightly coupled: Processor with Attached Coprocessor Approach**
  - Master application on the base system spawns work threads to the accelerator system as needed

Compelling differentiation and accelerated system improvement can only be achieved through a multilevel Hybrid System architecture that integrates complementary scalable subsystems optimized throughout the stack. These subsystems will evolve into a few main forms of scalable parallel architectures, with significant reuse of common technology components across all of them.



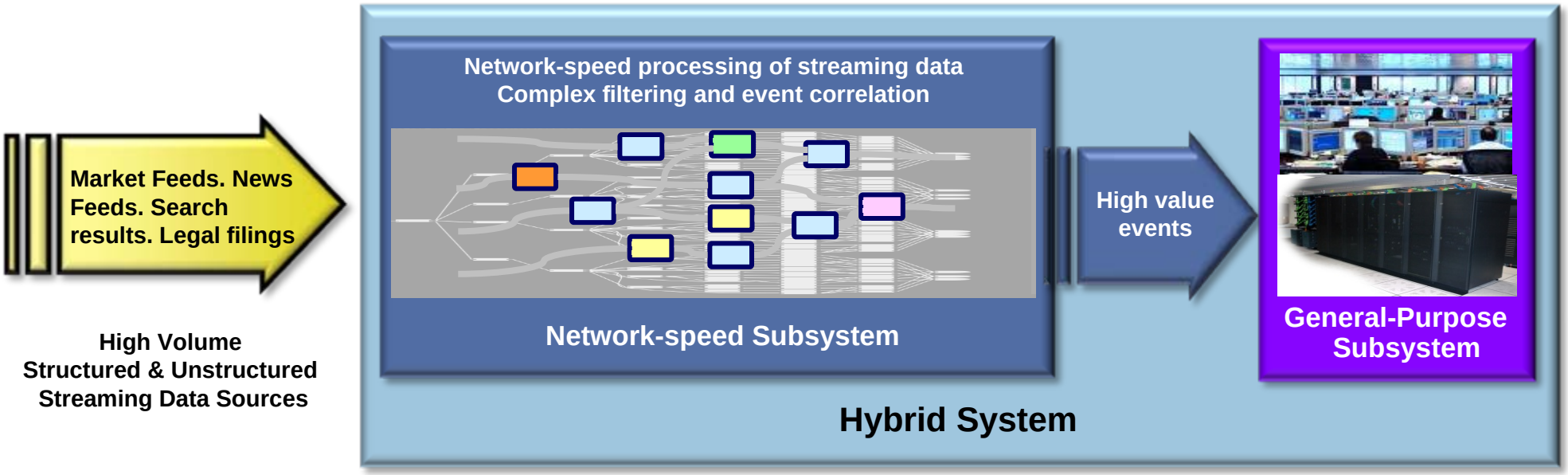
Different Flavors of Acceleration for Hybrid Systems



Network Optimized Hybrid Systems

In a Smarter Planet, vast amounts of data will be created and disseminated at networking speeds. Systems will need to process data in near real-time for information efficiency, and scale with the growth in the number of data sources and information types at networking data rates.

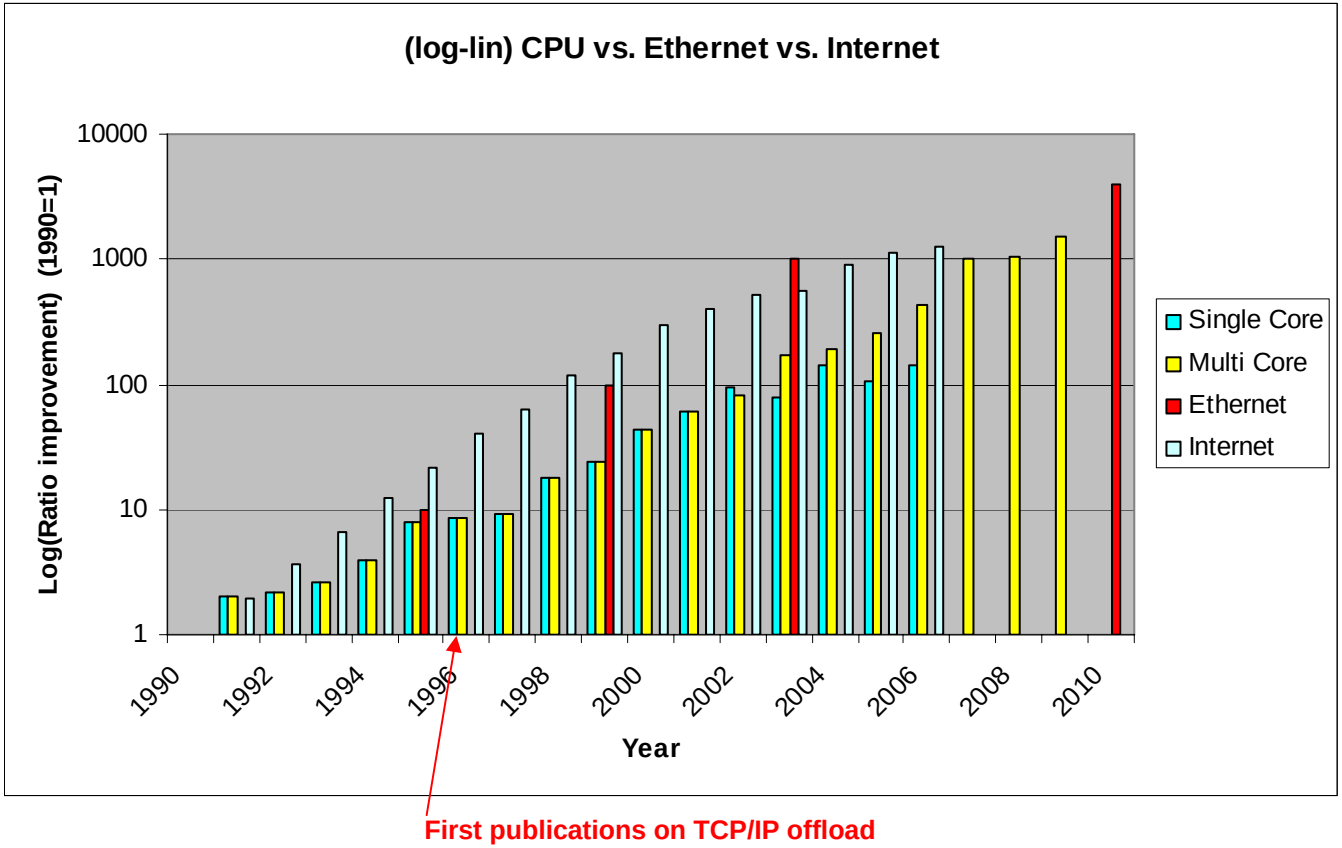
Example: Network-speed arrival of financial information from disparate sources (news feeds, tickers, subscriptions, exchanges, clearing houses) pre-processed for risk analytics and automated trading. This will allow **real-time response** to complex market and credit conditions and greater visibility into real economic conditions



Opportunities:

- Financial, Security, Streaming
- Health and Diagnostics, Data Analysis
- Transformations and Correlations
- Climate, Population
- Global-threat containment, Power grid
- ….

IO Performance vs. CPU Performance



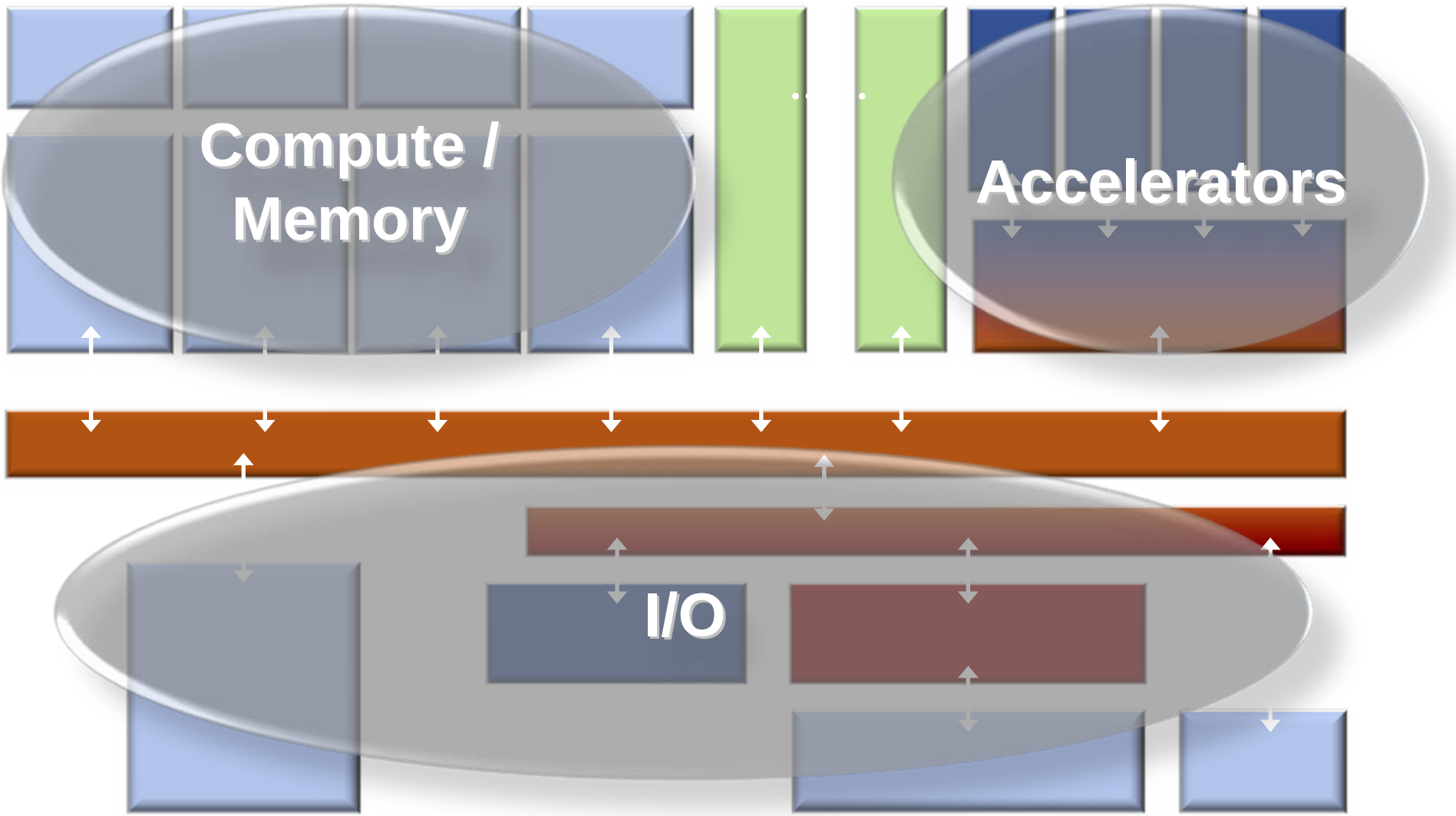
Notes:

1. **Assumes 40G Ethernet in 2010**
2. **Multi-core performance of TCP/IP, or multiple cores doing IO on one CPU die is not fully considered.**

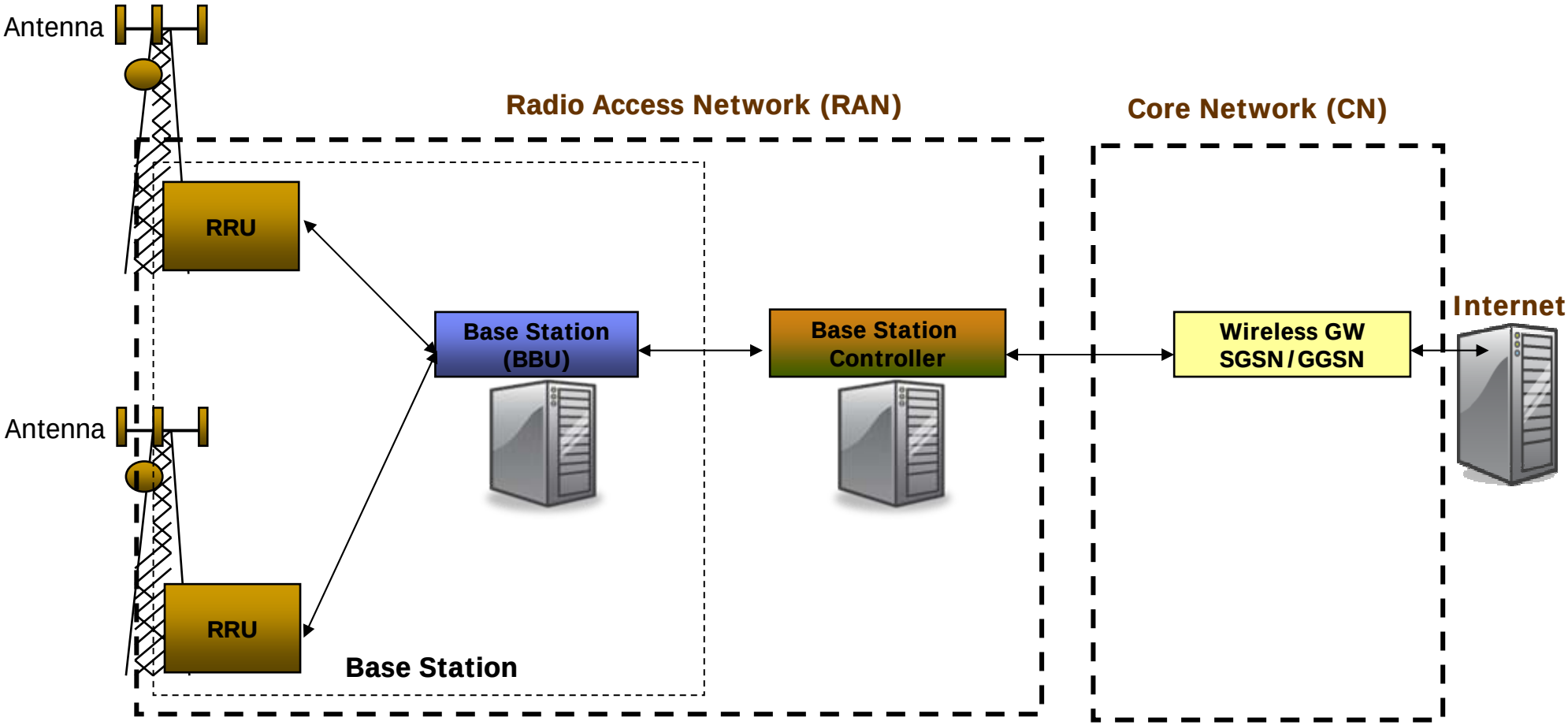
- **Since mid 1990's IO speed grows faster then CPU performance: delta ~ 0.5 – 1 Order of M.**
- **Multi-core processor chips *seem* to allow catch up, though:**
  - **Only if off-chip bus speed scales linear with number of cores**
  - **Only if IO tasks can be well balanced over all cores**
- **In addition, scale-out will put more demands on I/O – more synchronization / messaging between cores, not necessarily limited to one multi-core processor chip**

A Throughput-Optimized MMT System

- ***Network Processing + General Purpose Computing + Targeted Accelerators***



Example: Wireless Network Infrastructure Base Station



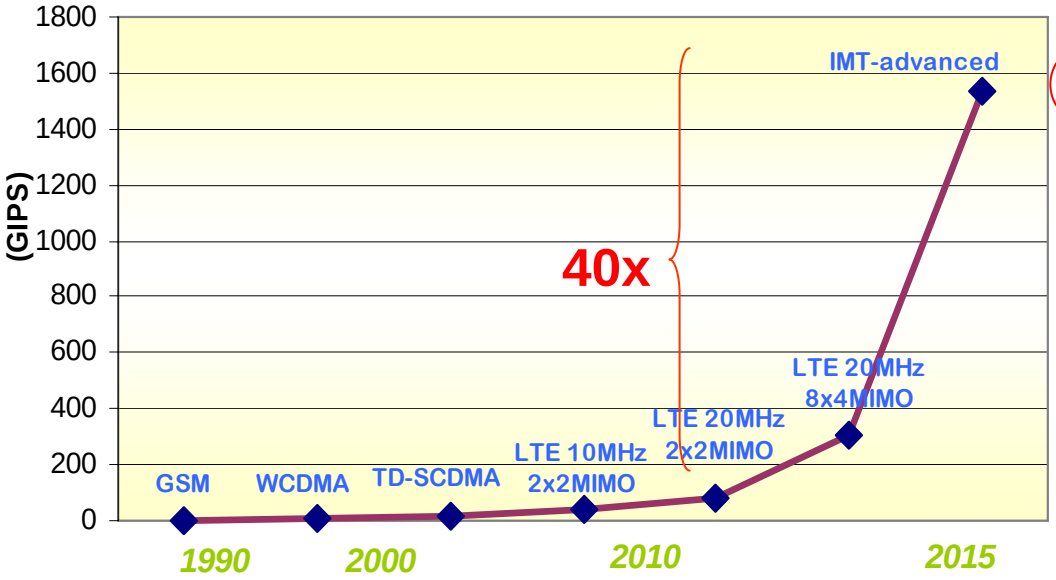
RRU: Remote Radio Unit

BBU: Base Band Unit

SDR: Software Defined Radio

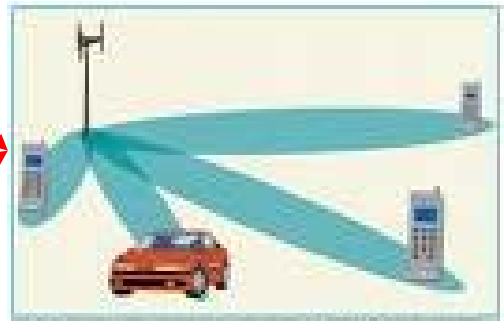
GGSN/SGSN: Gateway/Serving GPRS Support Node

Computation requirement per wireless spectrum carrier

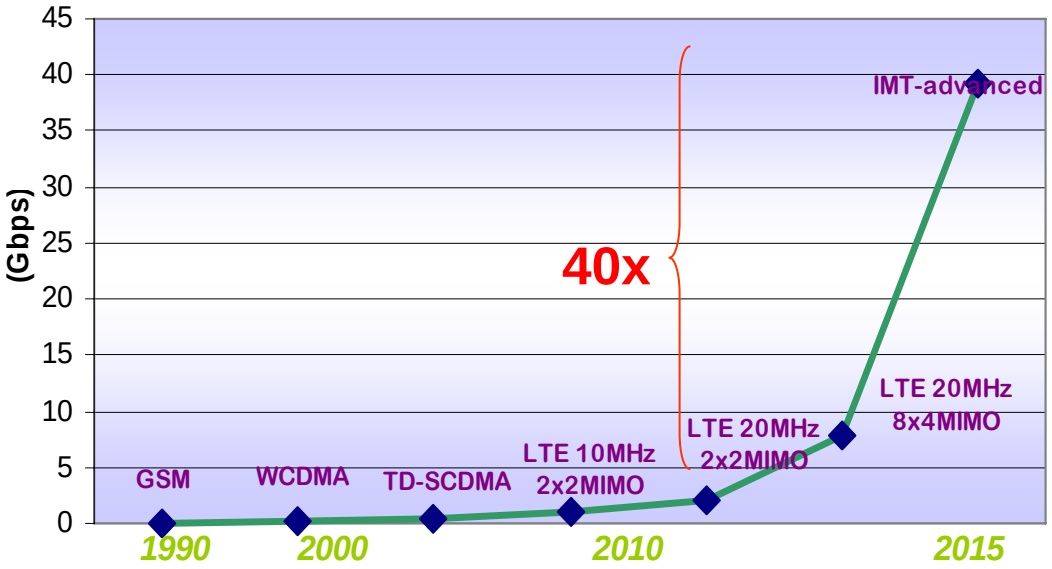


*

In 2015, a base station with 3 sectors requires > 4500GIPS computation capability, and 120Gbps I/O capability



I/O throughput requirement per wireless spectrum carrier



*

40x computation and I/O requirement

2G -> 3G -> 4G

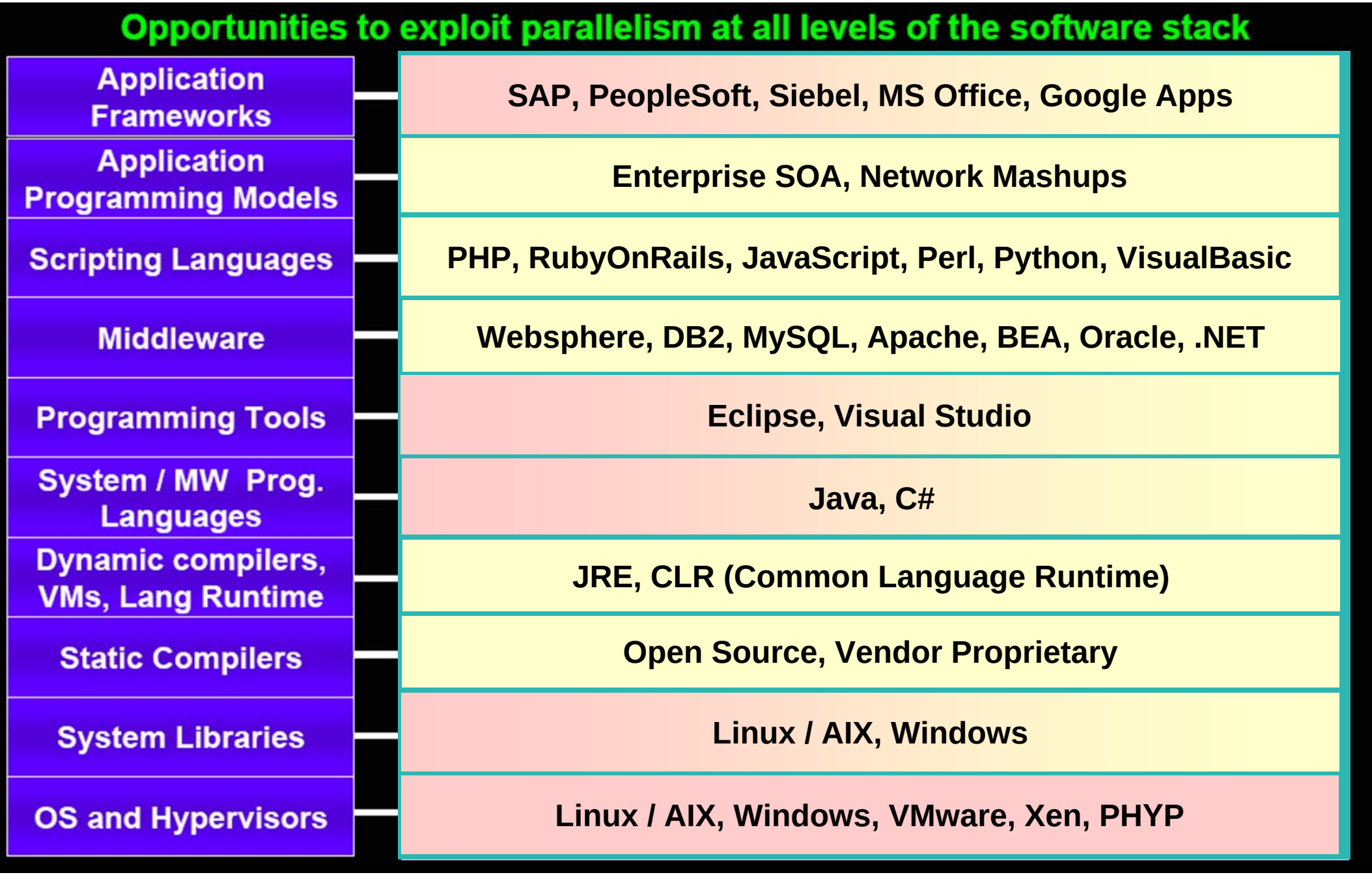
❑Multicore

- ➢ Applications must contain inherent parallelism (Amdahl's law)
  - ▪ Options
    - – explicitly expressed in programming model
    - – implicitly implemented in middleware (JEE)
    - – automatically detected by compiler (long history of limited success)
  - ▪ w/ appropriate tooling (development, performance)
- ➢ SW stack (middleware, OS, etc.) must preserve and map parallelism to HW
- ➢ Has been a very challenging problem for several decades

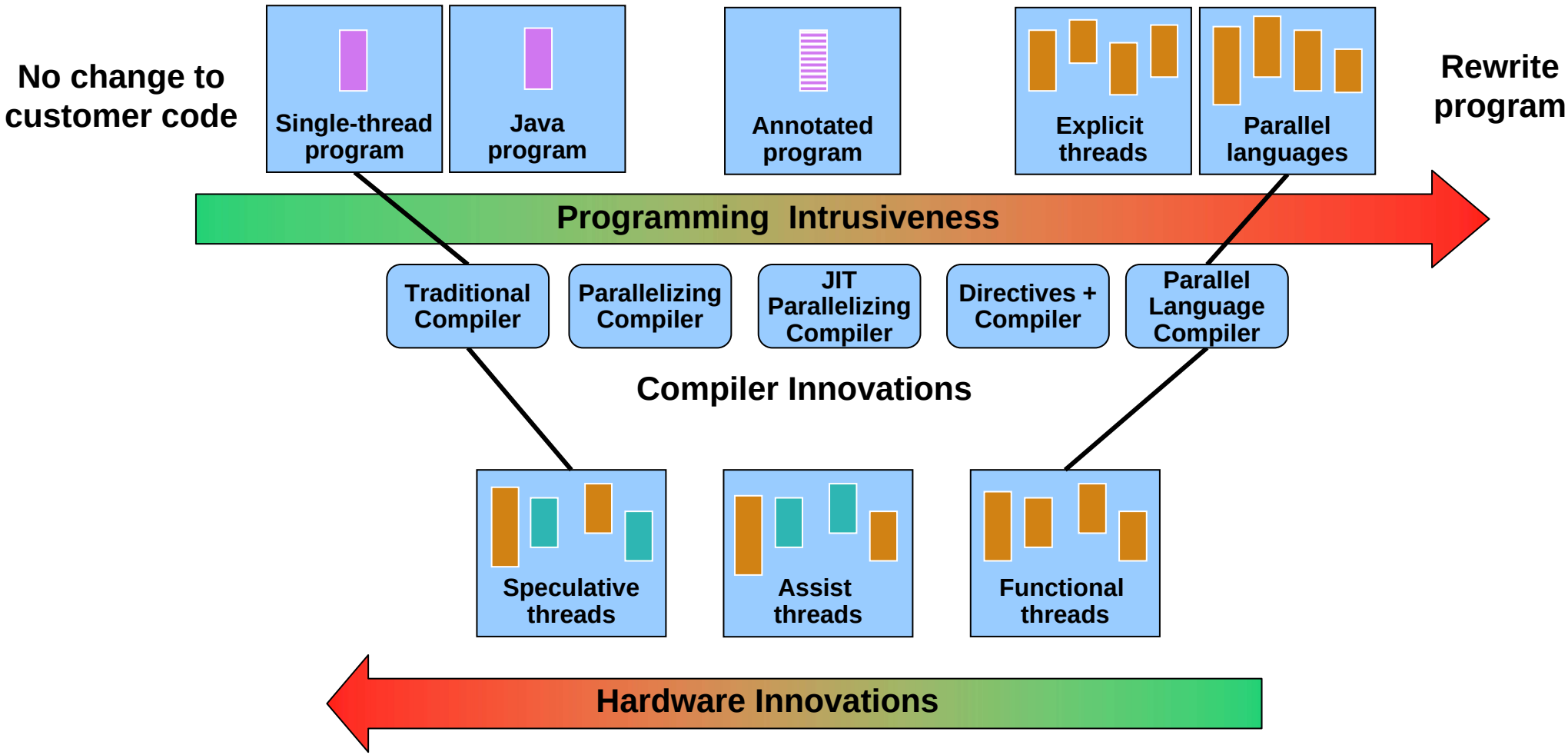
❑Hybrid

- ➢ Different communication, coordination, programming model assumptions
- ➢ Applications must have components that can be accelerated
  - ▪ Must be explicitly expressed in programming model or
  - ▪ Common function that maps directly to semantics of an accelerator
- ➢ SW stack must preserve and map effectively to HW
- ➢ All the problems of multicore, plus …

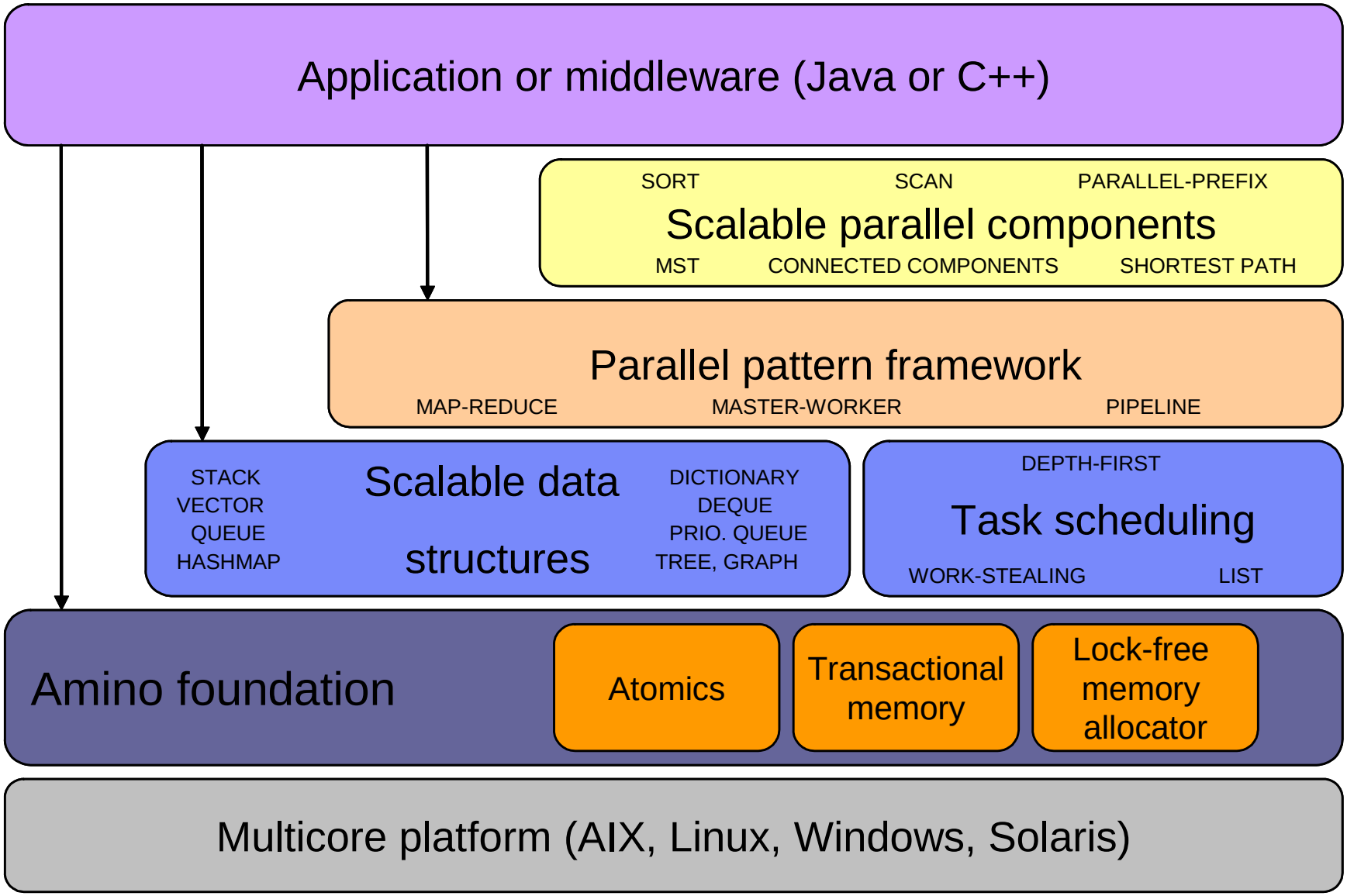
Multi-threading Enablement will Occur at Multiple Levels of the Stack



Systems built around multi-core processor chips are driving the development of new techniques for automatic exploitation by applications

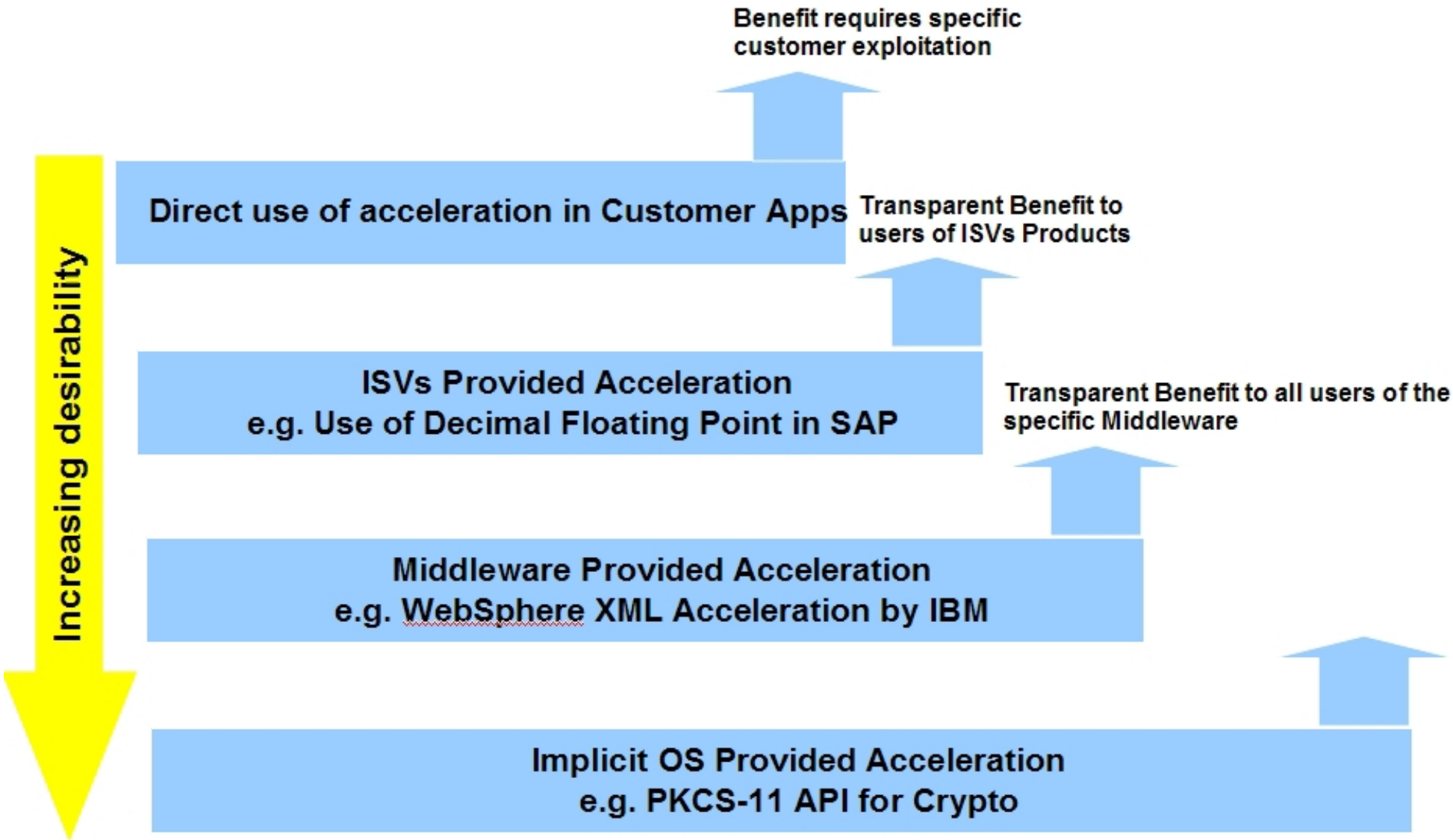


- **Open Source Software for dealing with Multicores**
  - Sourceforge: http://sourceforge.net/projects/amino-cbbs/



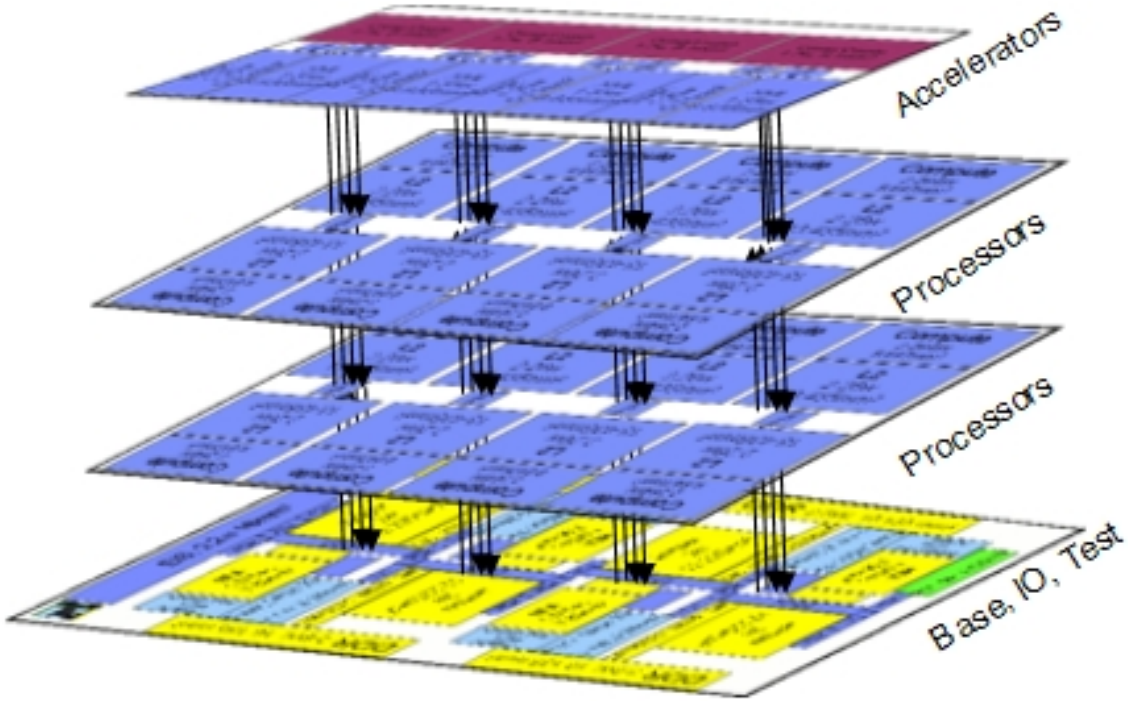
Exploitation of Accelerators

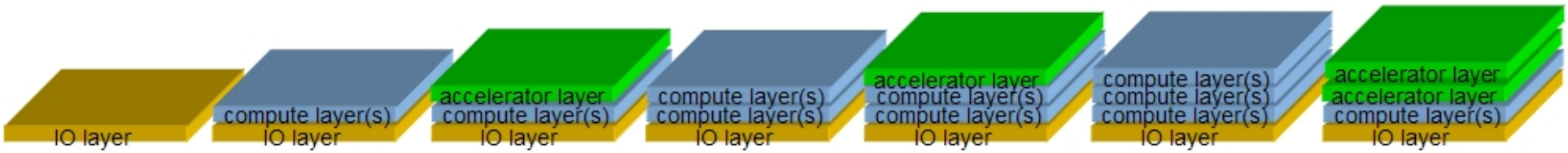
- Exploitation of accelerators at various layers of the software stack
- The “lower” the layer of the software stack, the more “desirable” from the point of view of Independent Software Vendors (ISVs)



PKCS-11: Public Key Cryptography Standard #11 is the Cryptographic Token Interface Standard

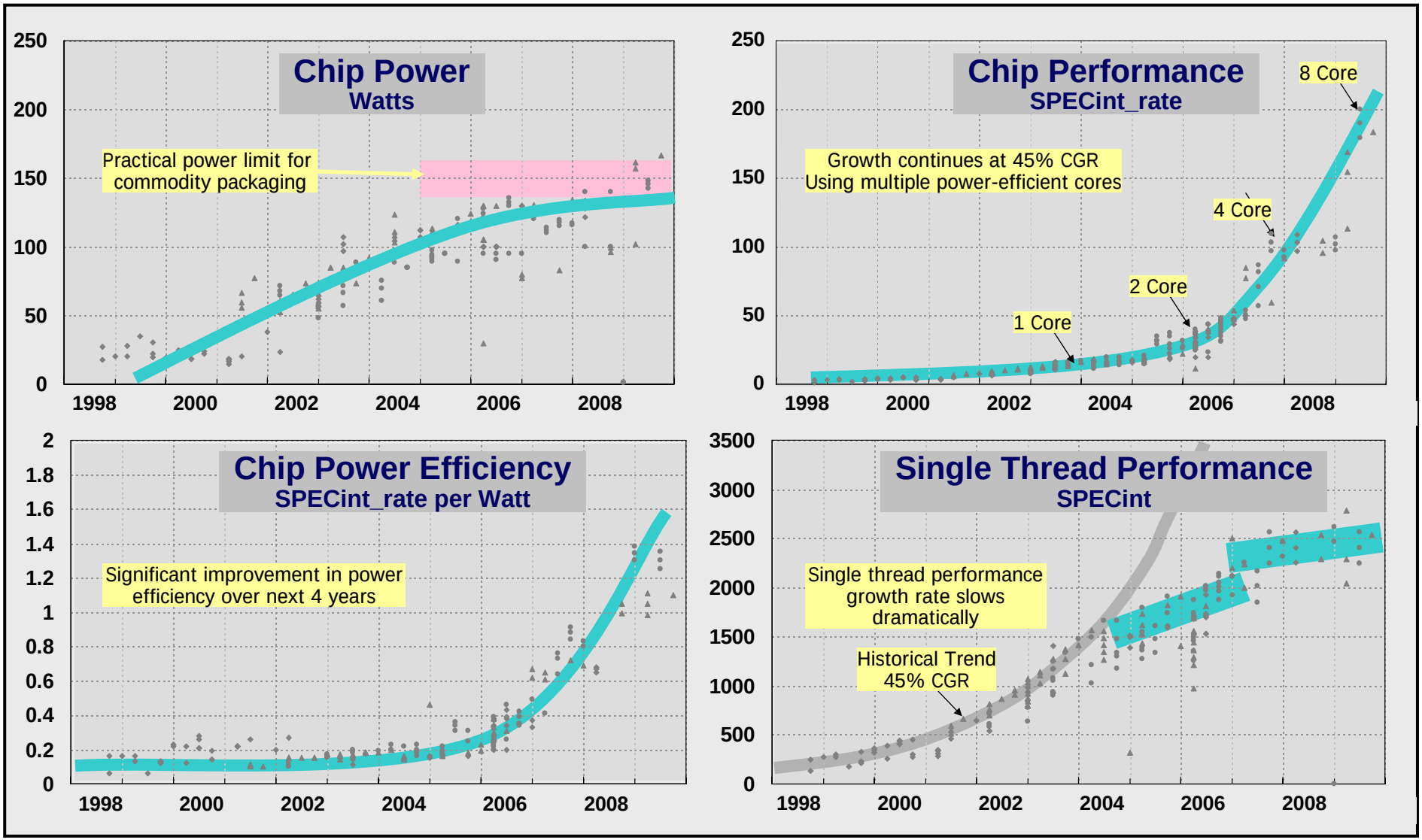
- 3D technology allows for the equivalent of 2x to 4x density improvements beyond normal semiconductor density scaling
- The modular layer approach to 3D allows for a multitude of application scenarios





Backup

Processor Technology Trends



The industry is adopting multi-core chips and energy efficient cores for aggressive chip and system-level performance growth.