





THROUGHPUTER



Application load and type
adaptive parallel computing





OVERVIEW

1) Fundamental transformation in computing:

- Parallel computing on dynamically shared (cloud) infrastructure

2) Industry-wide, must-solve challenge:

- Calls for an open-source platform approach

3) Innovation to enable performance-critical cloud computing:

- Dynamic parallel execution environment

4) Call for collaboration



1) Fundamental transformation in computing

- A. What higher clock rates *were* needed for (=everything), parallel processing *will be* needed for:
- For decades, application program speed-up was automatic through higher processor clock rates
 - Processor clock rate increases no longer feasible/economical
 - Going forward performance gain demands parallel processing
 - Including at intra-application level
 - Very complex and distracting!



1) Fundamental transformation in computing

B. Computing is increasingly taking place on dynamically shared cloud infrastructure

- A. (parallel processing) and B. (cloud computing) together:

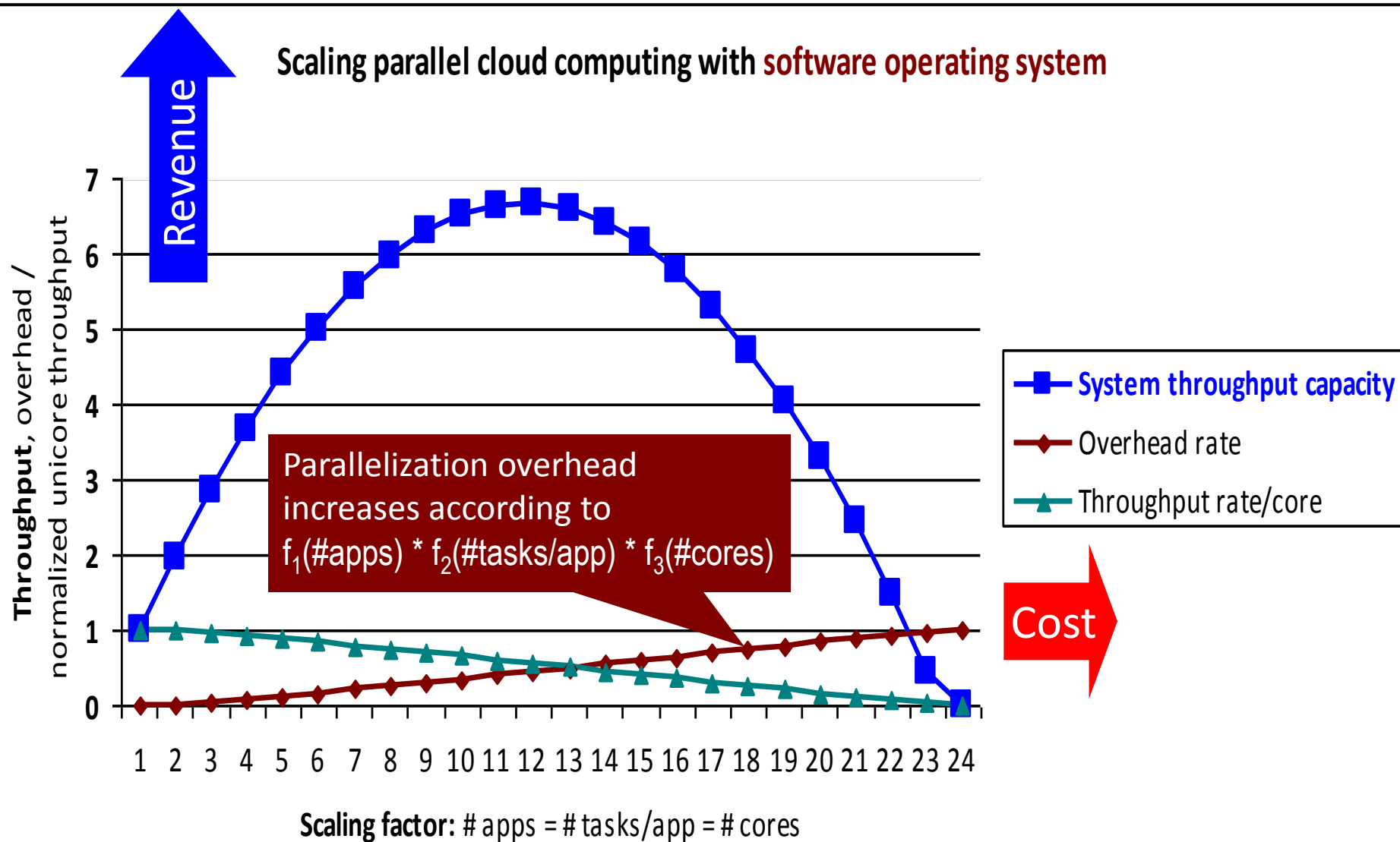
→ ***Need for dynamic parallel computing***

- Must-solve challenge for application developers, who however are ill-equipped to address it

→ **Fundamental challenge for software industry calling for a solution**

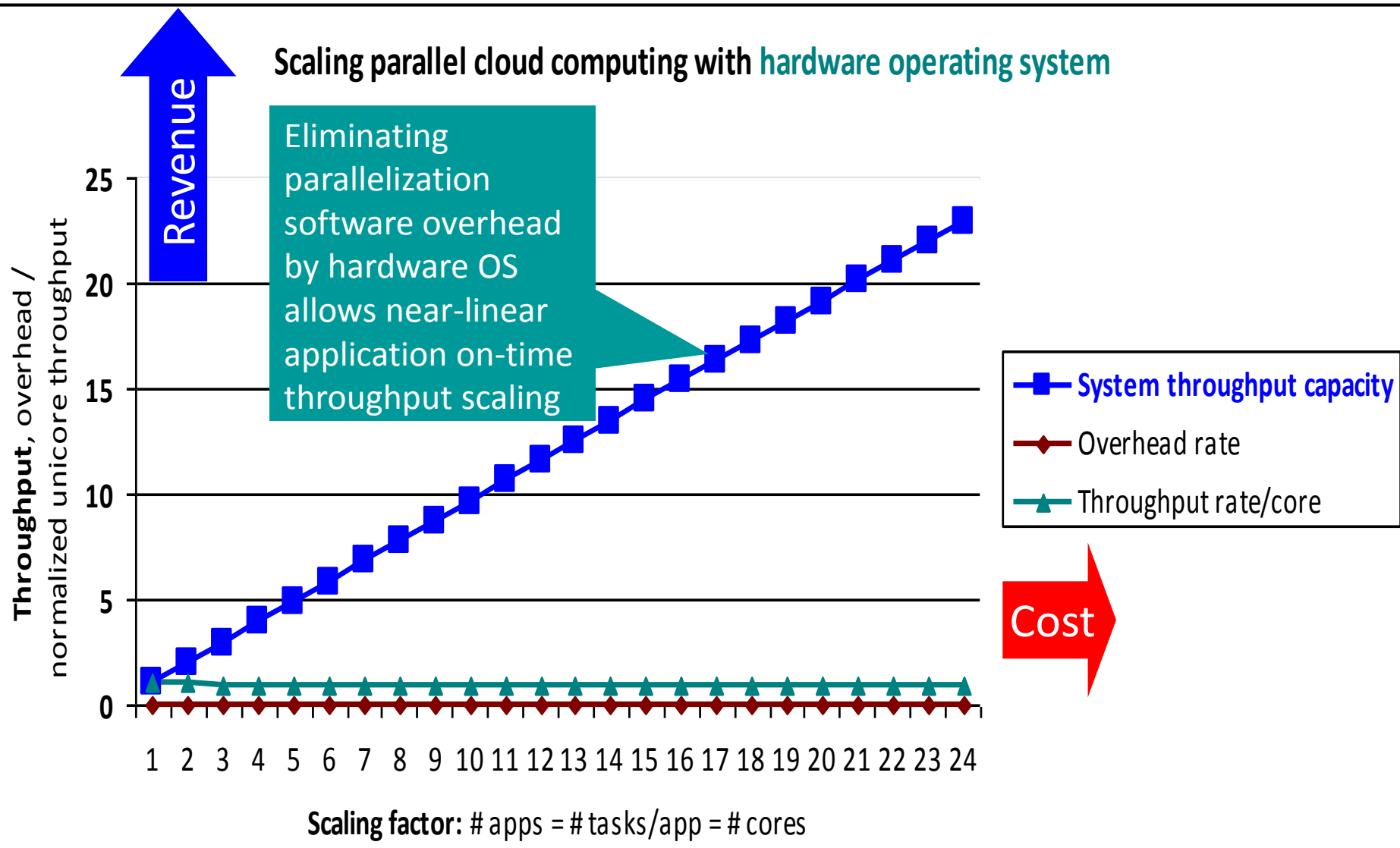
1) Scalability **Problem** in Parallel Cloud Computing = **SYSTEM SOFTWARE OVERHEAD**

Scaling parallel cloud computing with **software operating system**



Scaling factor: # apps = # tasks/app = # cores

1) Scalability **Solution** for Parallel Cloud Computing = AUTOMATE SYSTEM FUNCTIONS IN **HARDWARE**





2) Industry-wide, must-solve challenge

- Existing computing tech suppliers offer point-products or half-solutions:
 - parallel programming language (extensions), frameworks, tools, middleware, manycore processors, etc.
 - Burden of cost-efficient parallel cloud computing left for individual application developers (SaaS vendors)
- Status-quo not sustainable

→ A holistic platform solution needed for parallel cloud computing



3) Parallel cloud computing platform

- Existing parallel computing tools mainly limited to parallel *programming* aspect of parallel computing challenge
- Any parallel *execution* tools etc. designed for acceleration on dedicated machines, not on the cloud
- Parallel execution in cloud i.e. *dynamic parallel execution*, though critical for cloud computing cost-efficiency, left unaddressed by legacy vendors

→ **Dynamic parallel execution critical piece of performance-critical cloud computing**

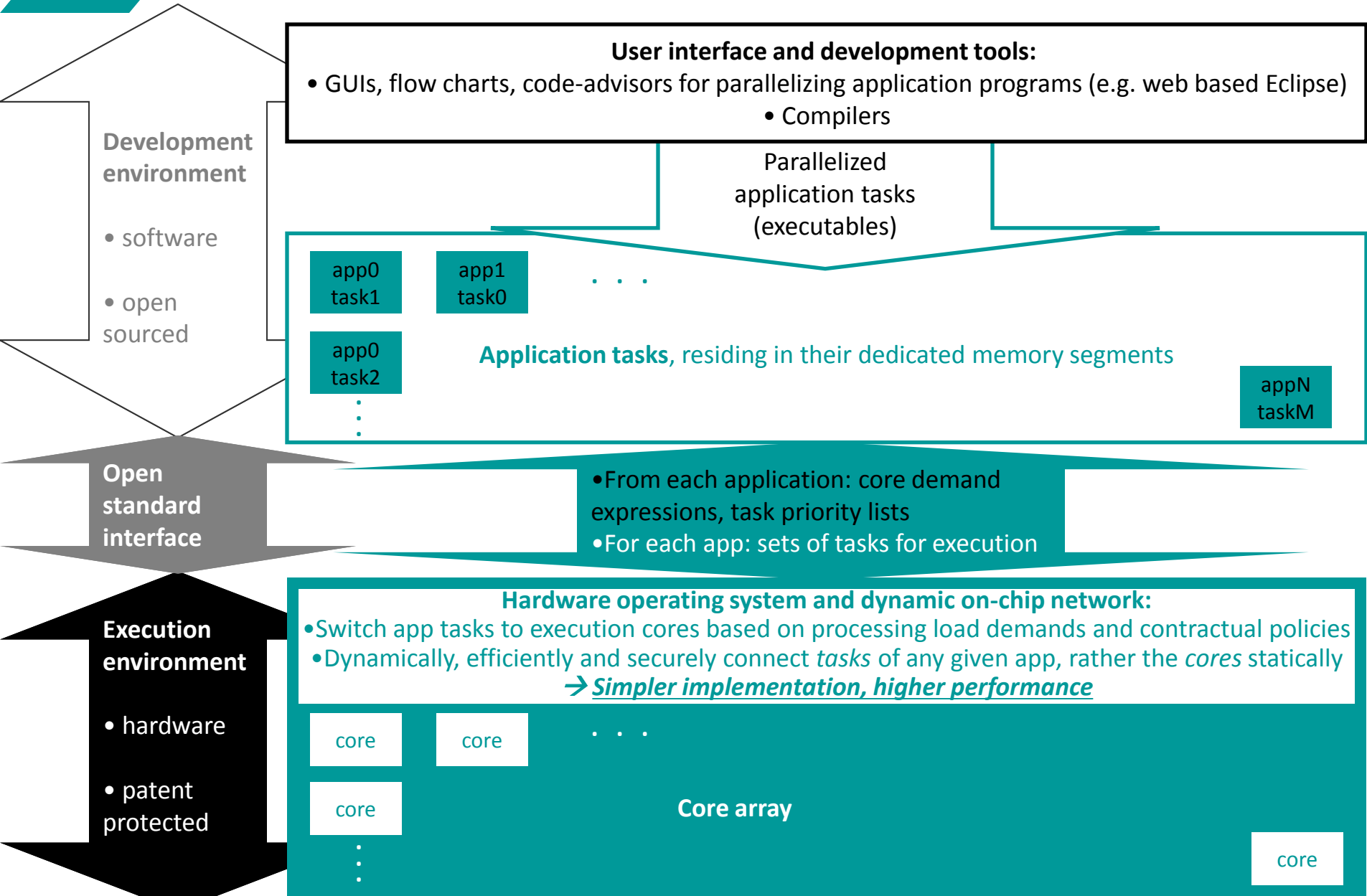


3) ThroughPuter: parallel computing PaaS

- ThroughPuter has developed a critical technology enabling effective parallel cloud computing platform:
 - Dynamic parallel execution architecture
 - Implemented in HDL code
 - User-friendly PaaS business model
 - Incentives to maximize computing on-time throughput performance *and* resource efficiency as well as user's productivity
 - Intellectual property rights for dynamic parallel execution
 - 45 patents issued and pending worldwide, incl. 19 granted US/UK patents

Performance and efficiency of dynamic parallel execution with high development productivity and deployment flexibility

3) ThroughPuter: Platform Overview



3) ThroughPuter: open parallel computing PaaS

Reasons for open-source collaboration for parallel computing platform to be based on ThroughPuter execution environment:

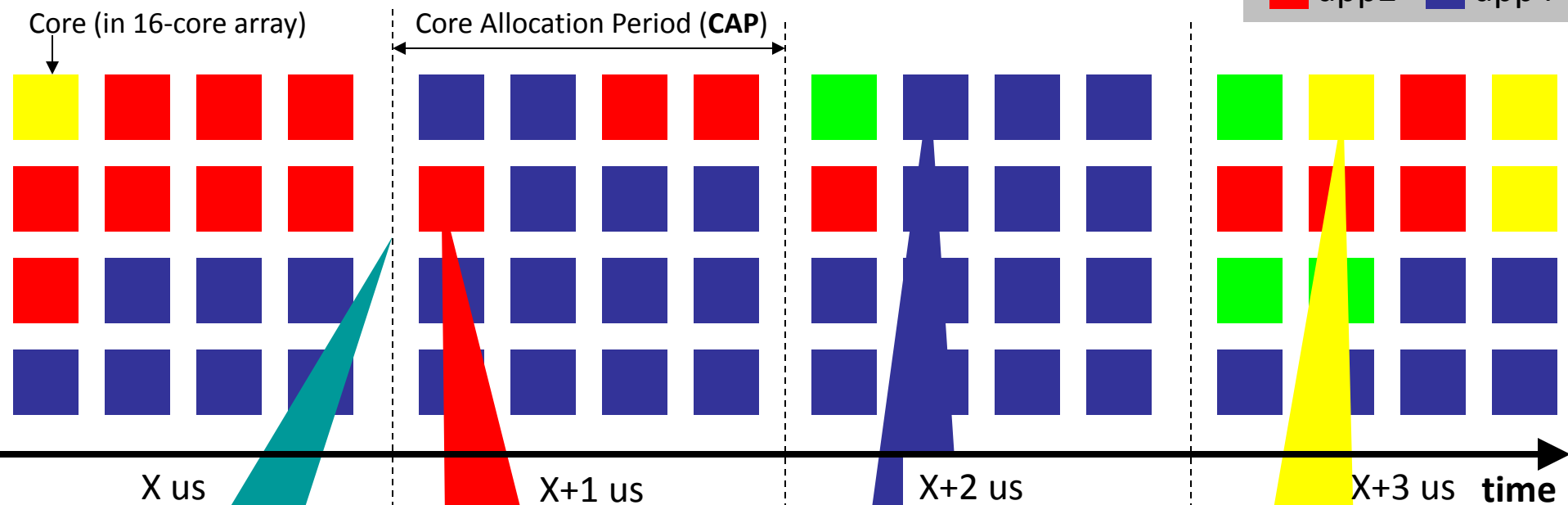
- 1) **Less low-level work:** ThroughPuter's execution environment automates parallel execution routines in (programmable) hardware, providing higher level interface (API) for the development environment software
- 2) **Higher performance** due to minimum-overhead hardware automation of system tasks such as optimally allocating processing capacity, scheduling and placing application tasks for execution, inter-task communications, billing etc.
- 3) **Built-in cloud computing security:** mechanisms for unauthorized interactions between different applications simply non-existent in the hardware
- 4) **Open standard interface** between development and execution environment

Dynamic parallel capacity allocation example and associated functional processor architecture block diagram on next slides →

Open parallel computing PaaS needed -- ThroughPuter execution environment goes a long way toward reaching that goal

3) Dynamic Core Allocation Example

app1 app3
app2 app4



Allocation of cores re-optimized among applications for each new CAP based on core demands and entitlements of the applications

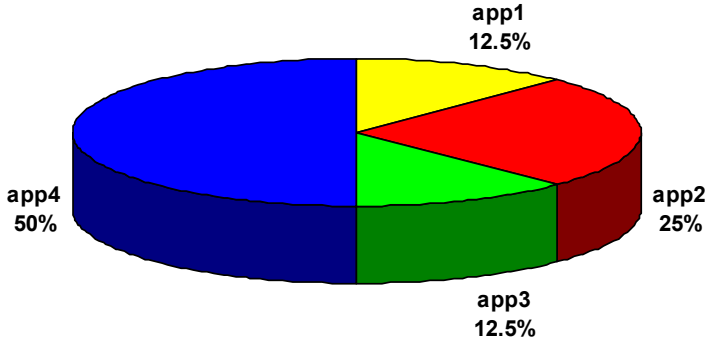
All tasks continuing on consecutive CAPs stay on their existing core, and continue processing uninterrupted through CAP boundaries

Any application can burst even up to full capacity of the shared core pool, as long as actually materialized core demands by other apps are met up to their entitlements

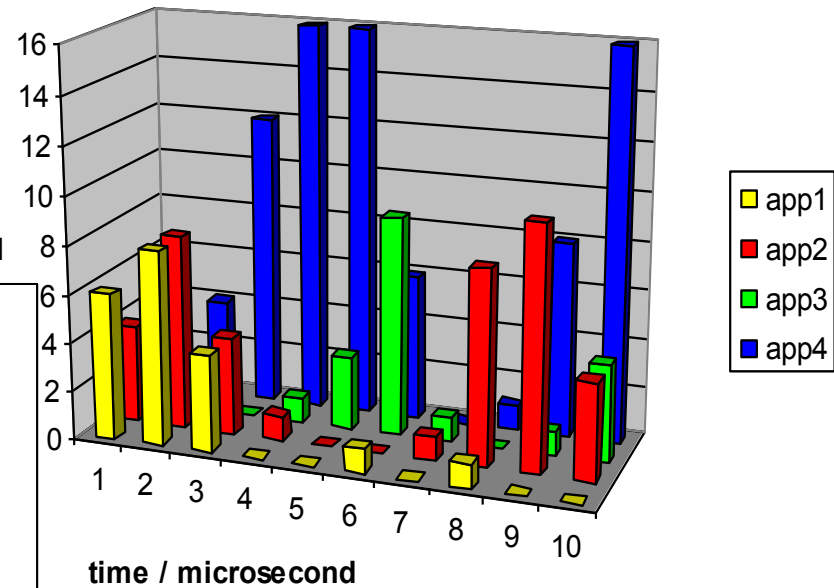
Any task can communicate with all other tasks of its app instance without knowledge of whether or at which core any given task is running

3) Manycore Cloud Computing Challenge – Technical

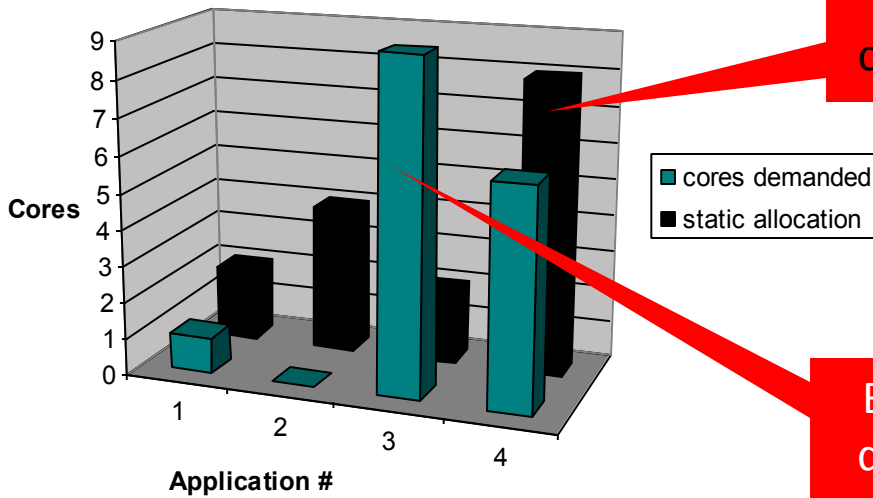
Example:
Average core demands by applications sharing a 16-core processor



Actual core demands by applications over 10us period



Actual and average core demands at t = 6 microseconds

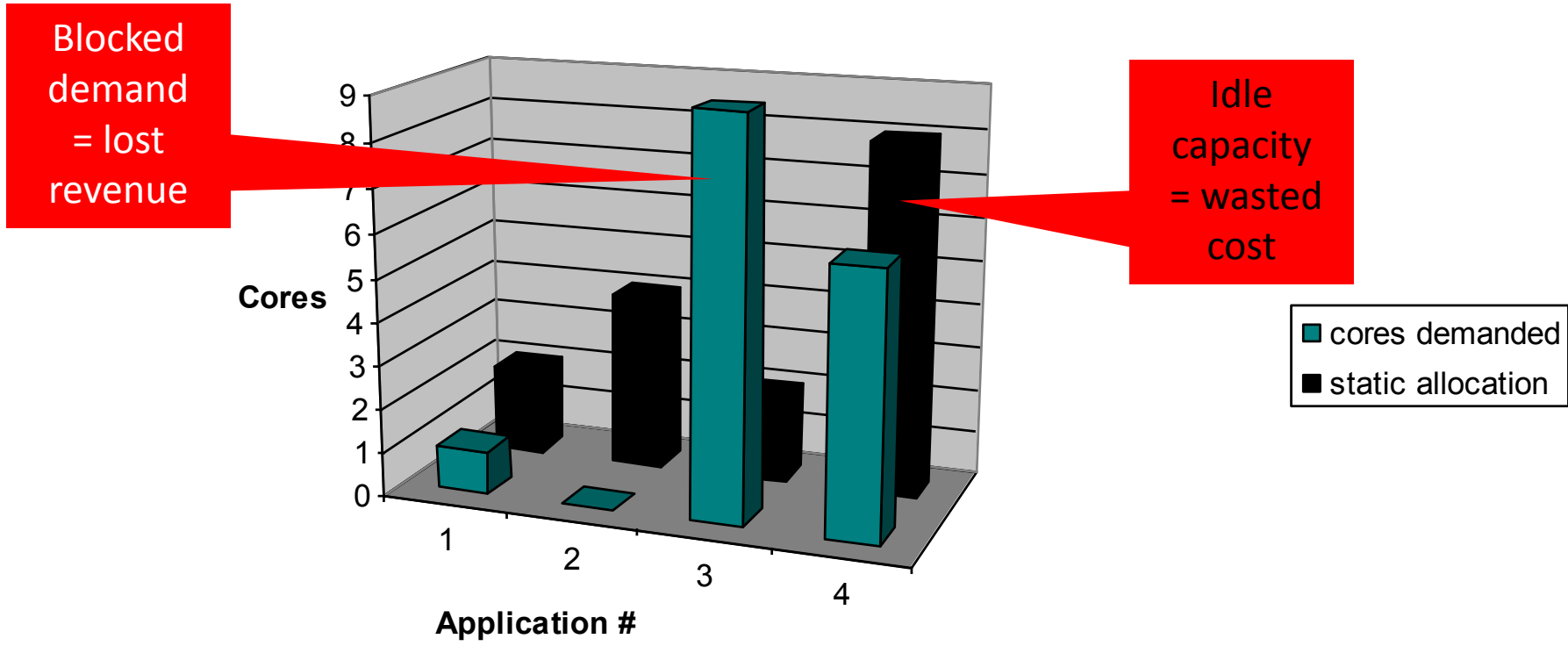


Idle capacity

Blocked demand

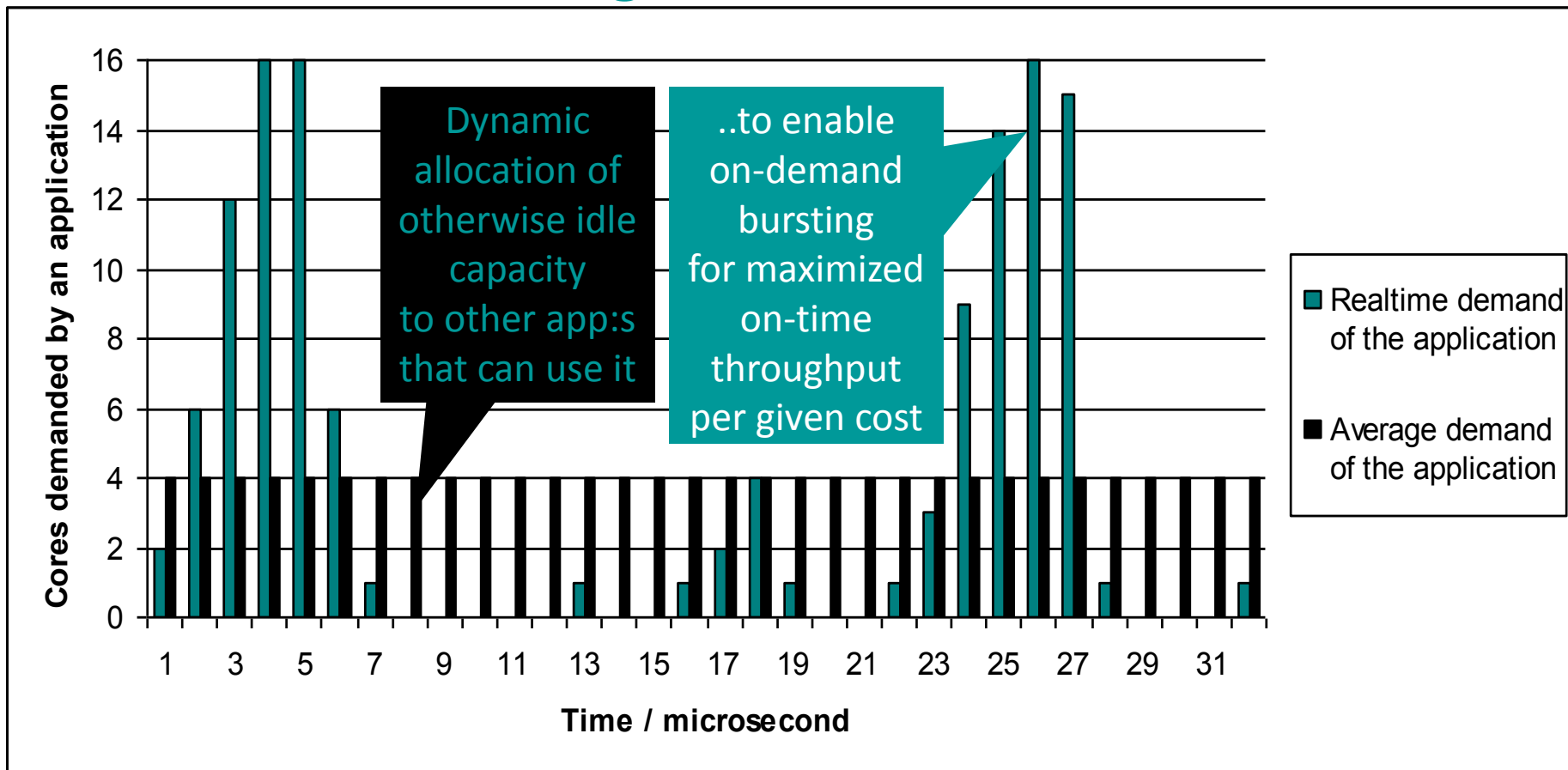
3) Manycore Cloud Computing Challenge - Economical

Actual and average core demands at t = 6 microseconds

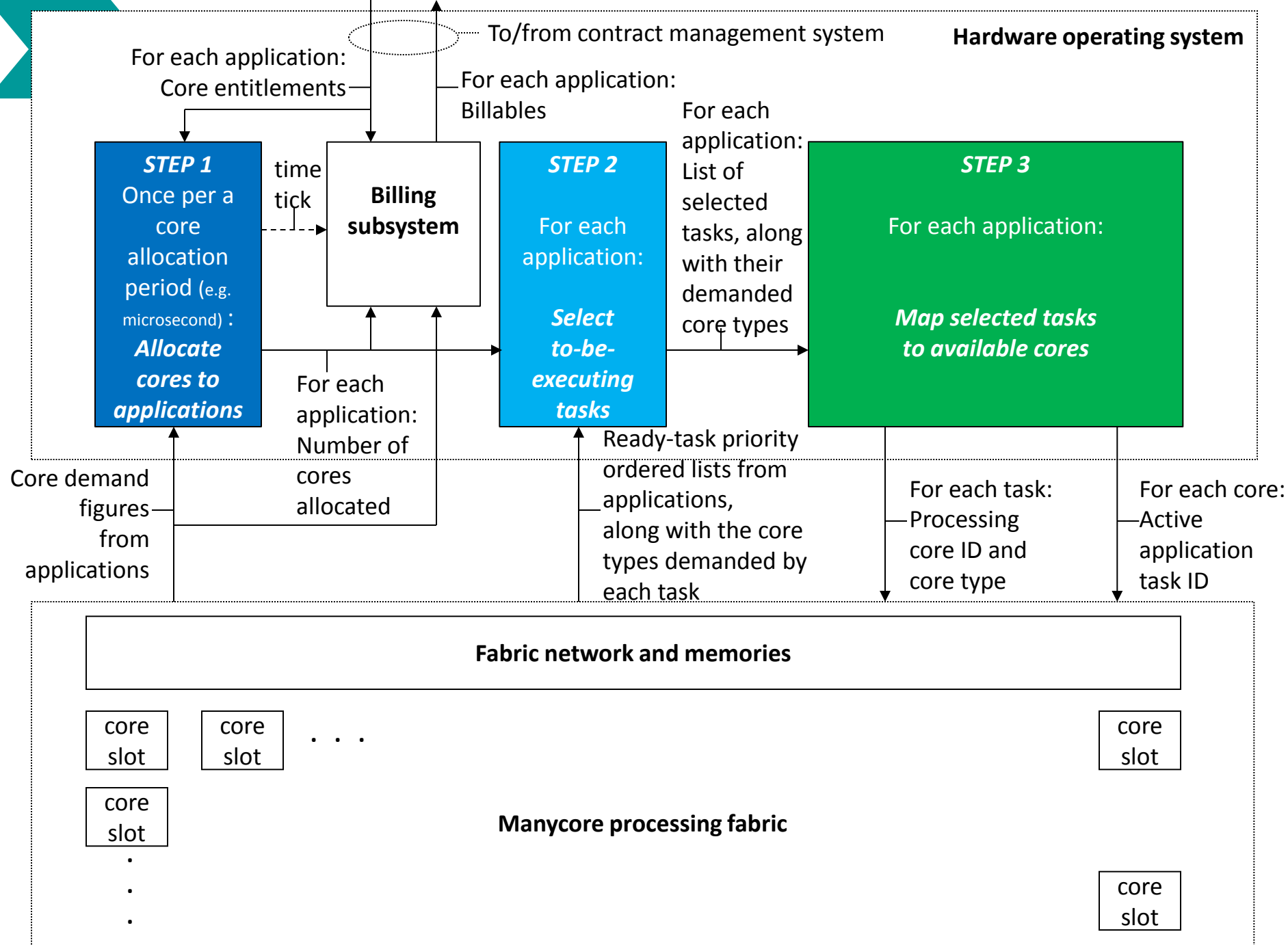


- The actual, momentary processing capacity demand by any given individual application program hardly ever equals its 'average' demand
- Non-adaptive capacity partitioning leads to wasting of resources and blocking on-time throughput
- Capacity being held statically in reserve for idling applications should have been allocated to other applications on the manycore processor that at that time would have been able to use it

3) Manycore Cloud Challenge – ThroughPuter’s Solution



- ThroughPuter enables application load adaptive, dynamic parallel cloud computing





3) Summary of Advantages

- **PERFORMANCE and COST-EFFICIENCY:**
 - Architecturally maximized application processing on-time throughput per unit cost
 - Hardware operating system and on-chip network optimized for dynamic parallel processing on multi-client shared manycore processors
- **SECURITY:**
 - Full isolation, right from hardware level up, among client applications dynamically sharing a pool of cores
- **PRODUCTIVITY:**
 - Integrated development environment of PaaS automate parallel program development and deployment
- **OPEN SYSTEM:**
 - PaaS software and processor-core hardware to be open-sourced
 - Host anywhere; ThroughPuter commercial hosting an option



4) Call for Collaboration

- The need for parallel processing an emerging, MAJOR industry and profession wide challenge
 - Open-source collaboration a natural approach
- Need for architectural optimization *across* traditional application, system and hardware layer boundaries
- SOLUTION: **Open-source PaaS reaching all the way to parallel cloud computing optimized hardware**
 - ThroughPuter’s contribution: Hardware architecture designed for dynamically shared multi-user parallel cloud computing
 - Secure hardware OS for manycore fabric with on-chip network, taking care of dynamic capacity allocation, parallel program execution mgmt
 - Collaboration opportunities:
 - Development environment and tools
 - Extensions of the PaaS for specific user domains: channel partnerships
 - Processor-core IP for the cloud processors;
 - Physical hardware supply, physical hosting (IaaS) etc.



tech@throughputer.com

www.throughputer.com