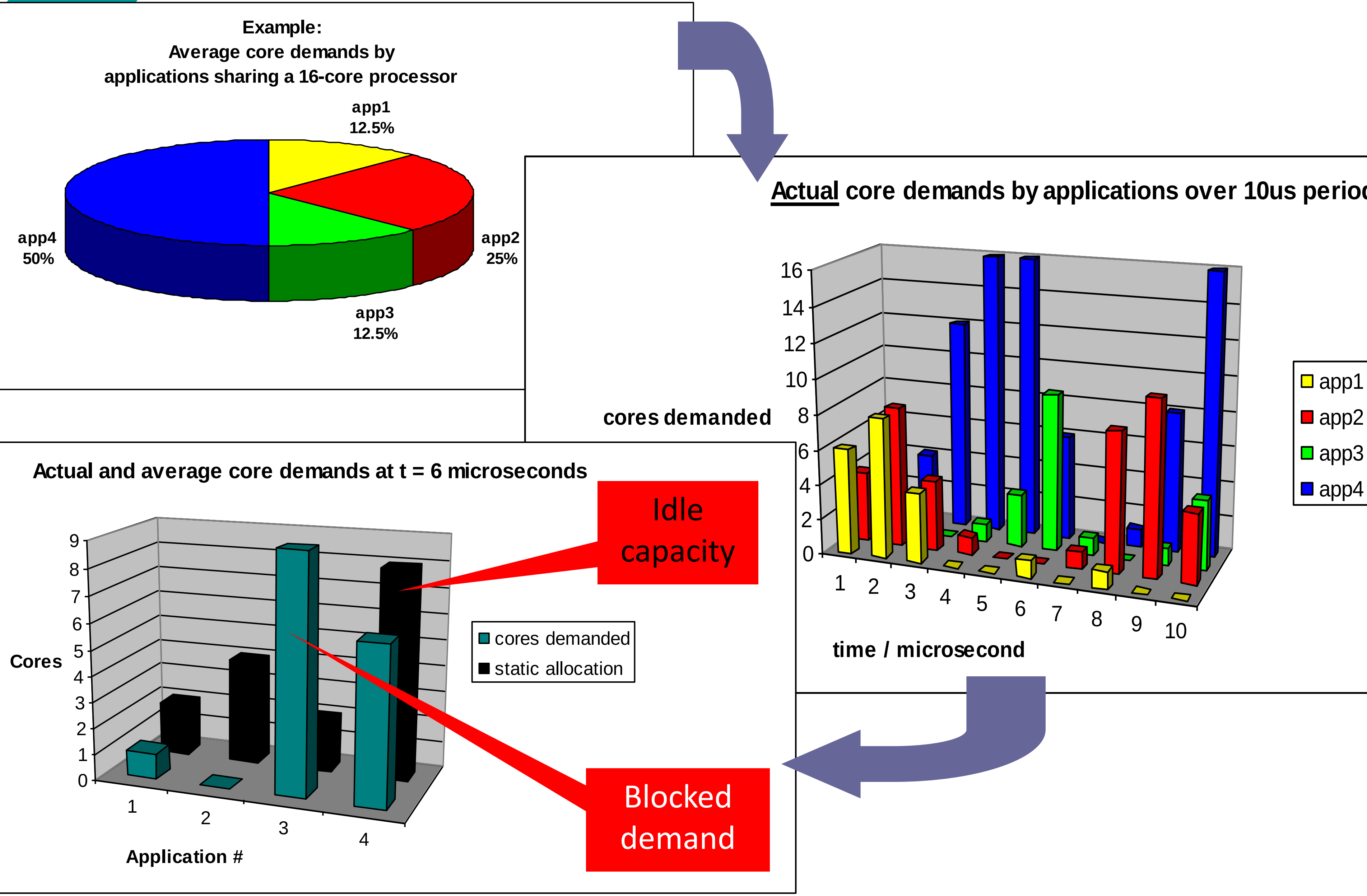
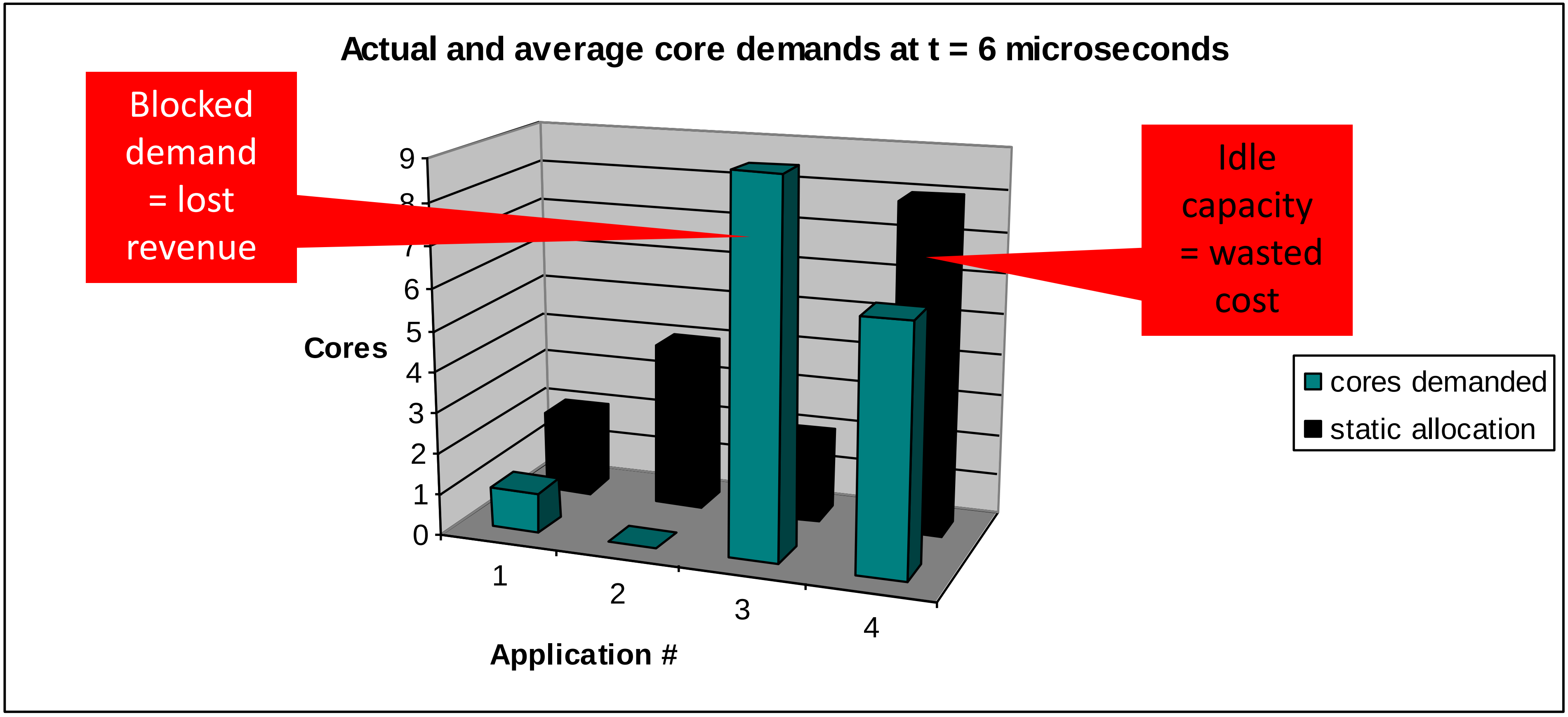
THROUGHPUTER

Parallel Program
Development and Execution
Platform as a Service

Manycore Cloud Computing Challenge - Technical



Manycore Cloud Computing Challenge - Economical

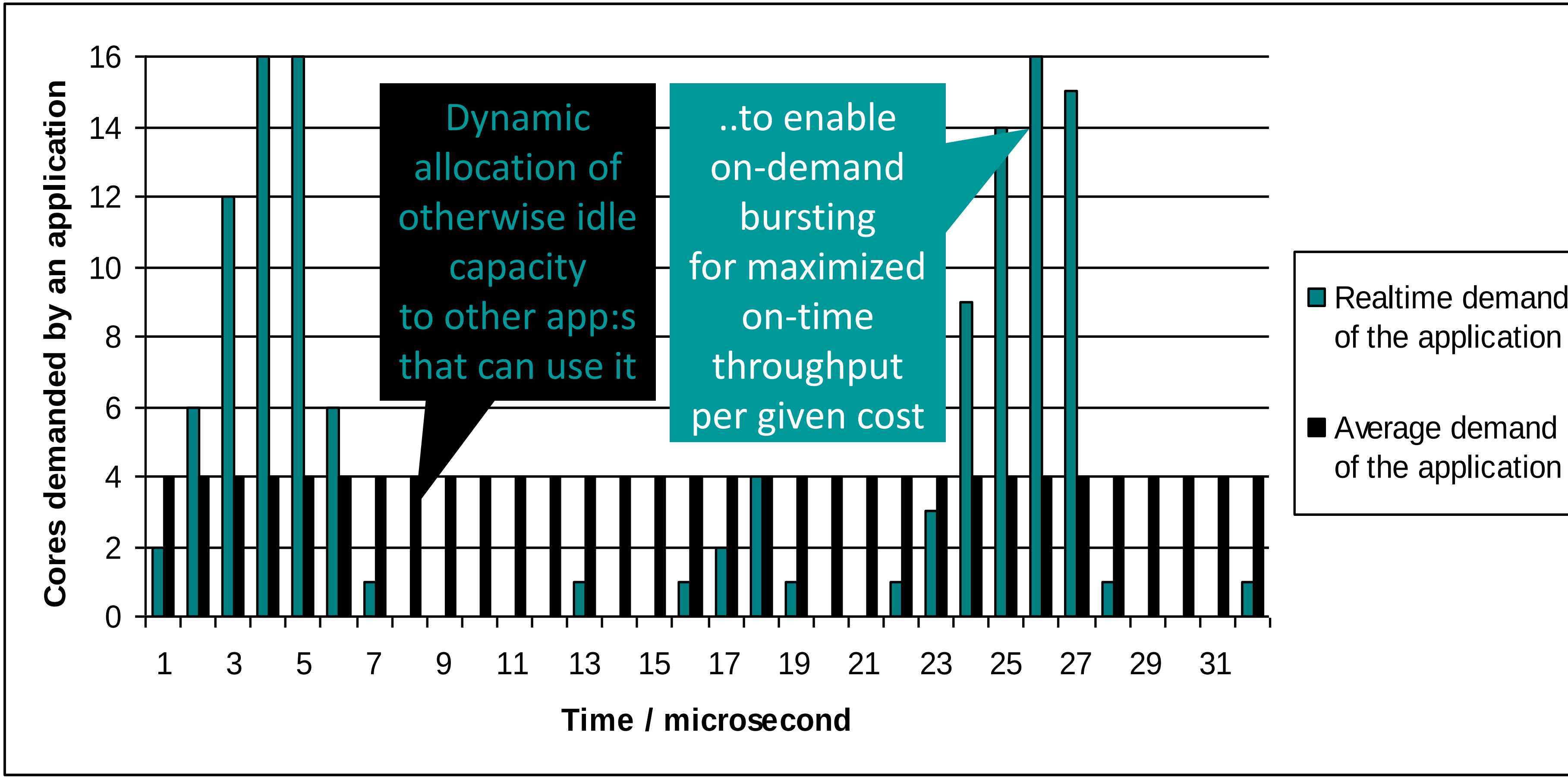


- The actual, momentary processing capacity demand by any given individual application program hardly ever equals its ‘average’ demand

→ 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

Manycore Cloud Challenge – ThroughPuter's Solution

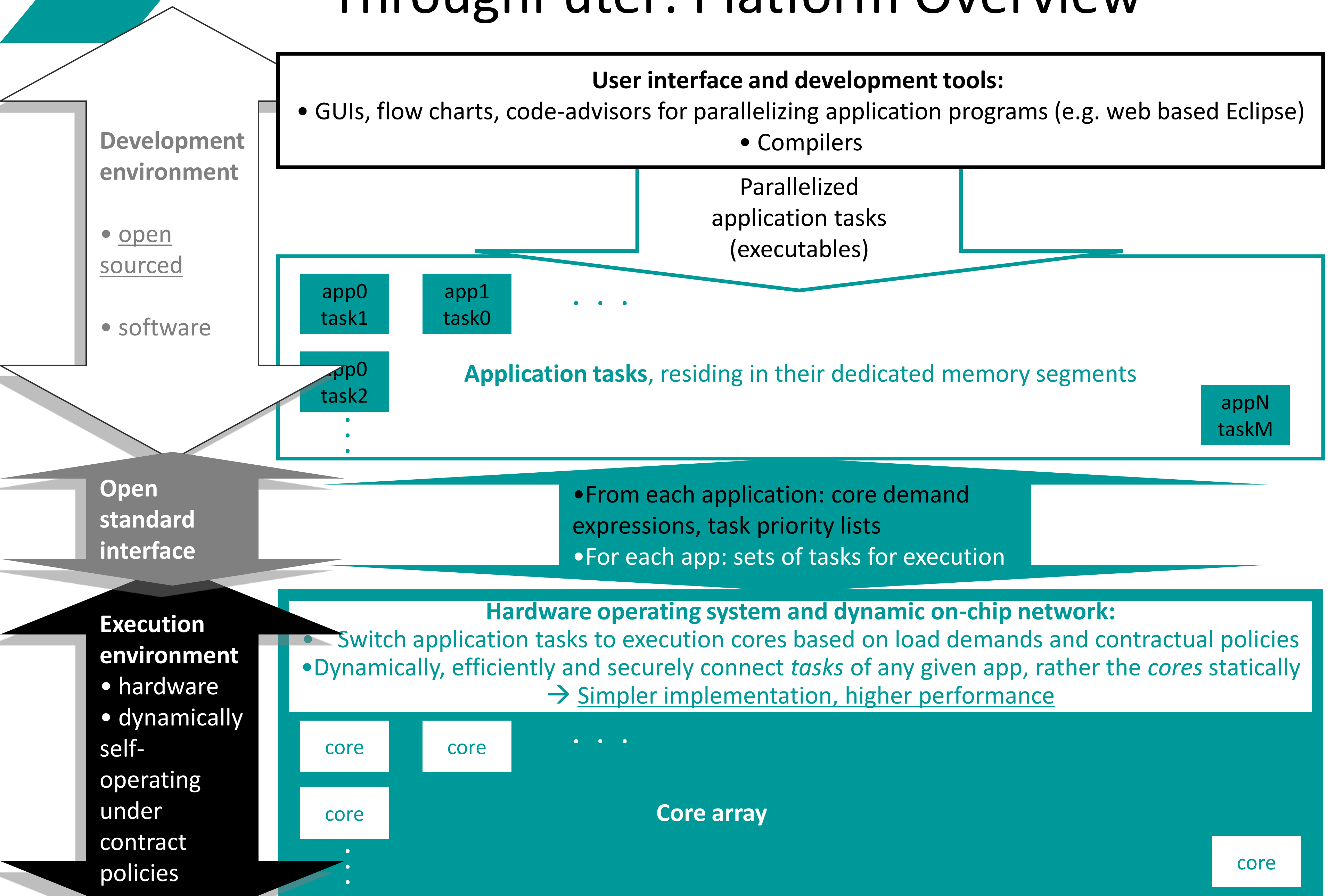


- ThroughPuter enables application load adaptive, dynamic parallel cloud computing

What ThroughPuter Brings To Party

- New manycore processor architecture *designed from ground up* for secure, dynamic parallel cloud computing
  - More about this under *Key Advantages*…
- User-friendly access to high-performance of hardware through Platform-as-a-Service (PaaS) business model
  - New hardware architecture to eliminate, not cause, parallel processing software headaches
- Open-source collaboration for comprehensive parallel computing PaaS:
  - Entire PaaS platform software, as well as the processor core hardware model (Verilog), open-sourced
  - Parallel processing an industry-wide challenge
    - → Industry-wide collaborative solution and holistic platform approach needed
  - *See architecture overview & motivation at next 2 slides*

ThroughPuter: Platform Overview



Patents issued and pending.

Open Parallel Computing PaaS

Reasons for software ecosystem for parallel computing platform to be based on the dynamic parallel execution environment interface *standard*:

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**: unauthorized interactions between different applications *simply not enabled* in the hardware

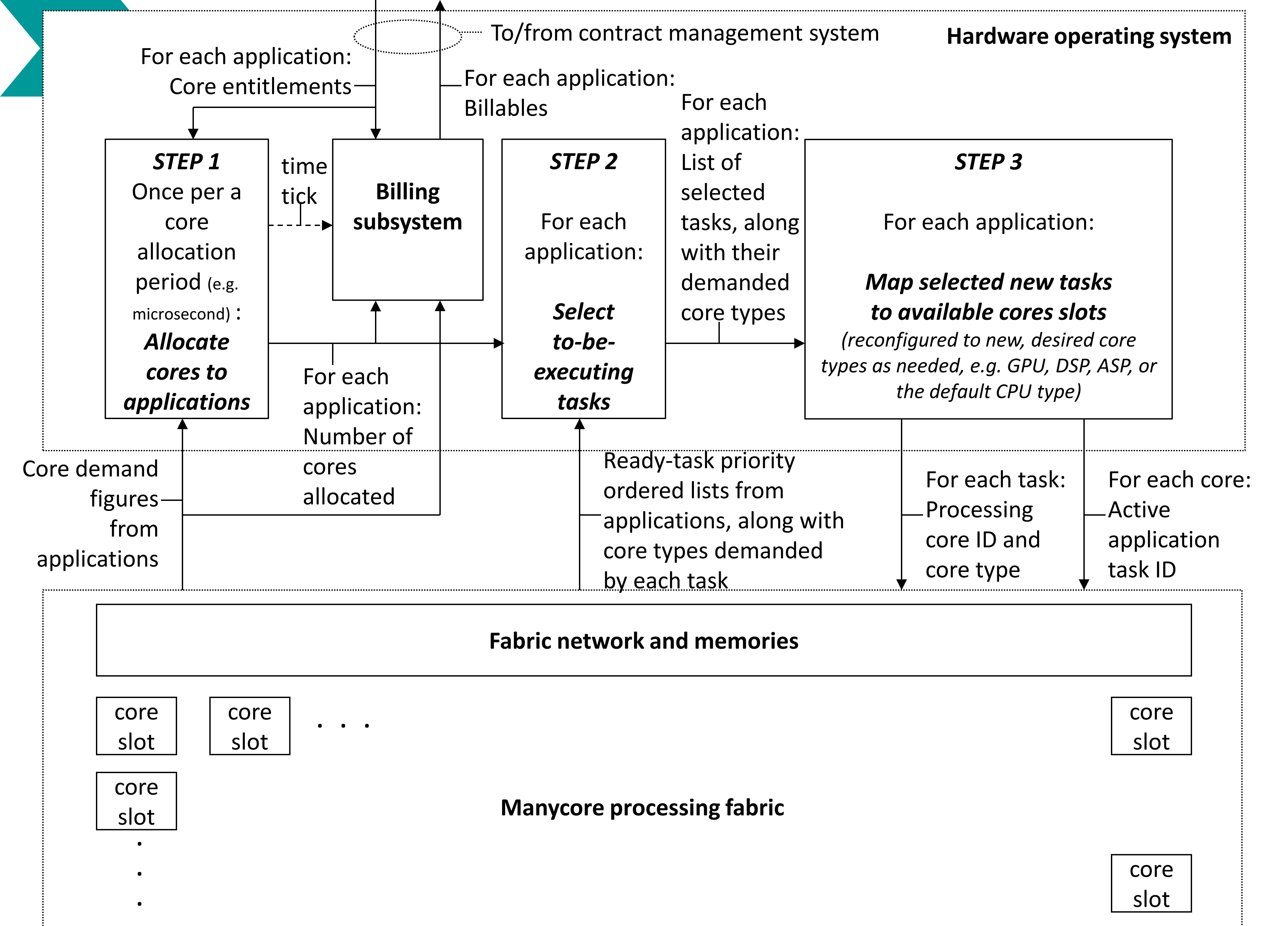
4) **Open standard interface** between development and execution environment

Key Advantages #1 – Dynamic Parallel Processing

- Hardware operating system designed for parallel processing on multi-client shared manycore processors
  - Dynamic parallel computing with strict minimum capacity access guarantees:
    - Each client program gets the maximized number of cores that it can use at any given time so long as all clients get their materialized demand for cores met at least up to their entitled and fair share
  - Automated, minimum-overhead interaction with applications:
    - PaaS development tools configure client programs so that they express to the HW OS their parallel execution ready tasks in priority order
    - HW OS selects and maps tasks for execution at their assigned cores
    - HW OS and on-chip network take care of inter-task communication, without requiring the tasks to know whether or where any given task is running at any given time

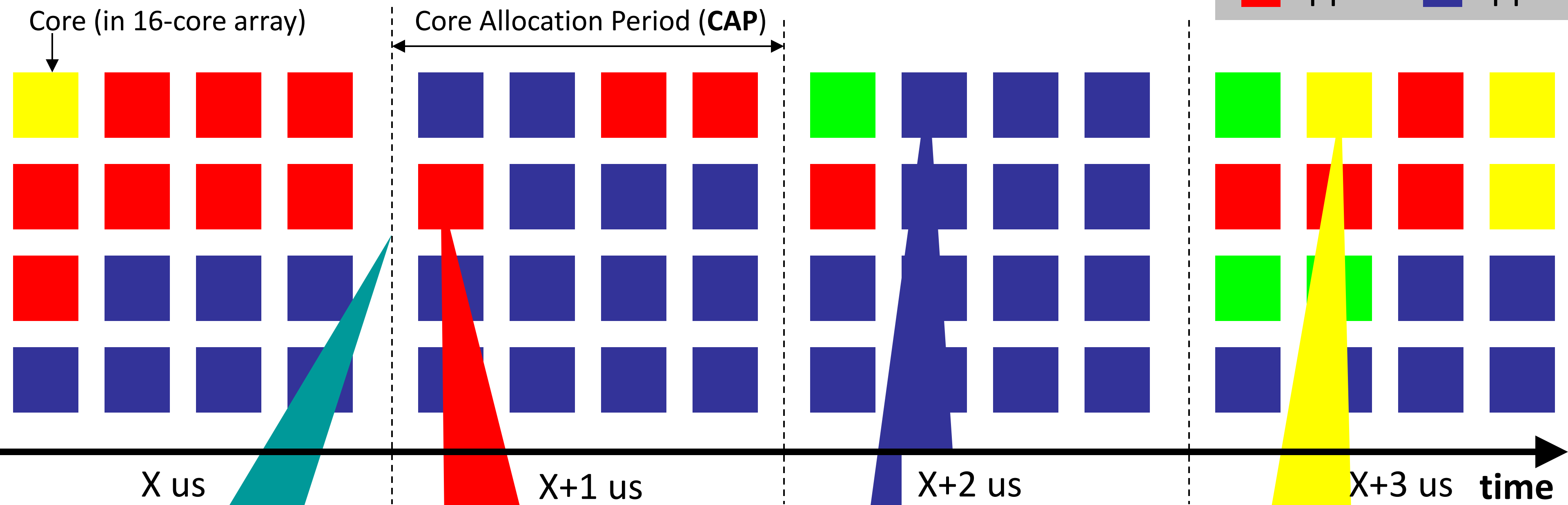
See system and operating diagrams on next slides →

→ Maximized application processing throughput per unit cost



Patents issued and pending.

Dynamic Core Allocation Example



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 uninterruptedly 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 the application without having to know whether or at which core any given task is running

 Patents issued and pending.

Key Advantages #2 – Inbuilt Security = Productivity

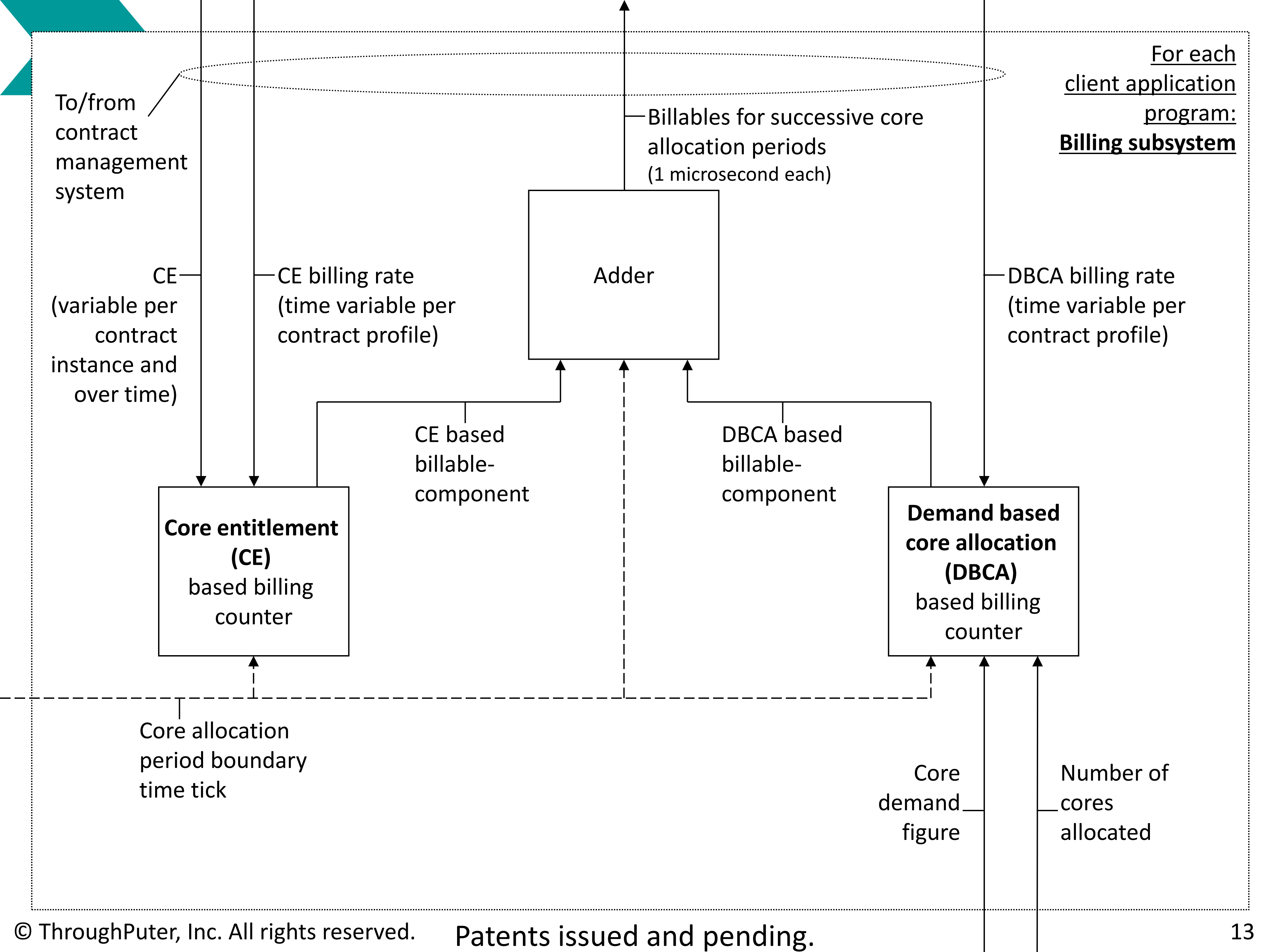
- Dynamic cloud computing with built-in security:
  - Unauthorized interactions and undesired interferences between client applications or the system not even enabled in the hardware logic
  - Each client program completely resides in its dedicated memory segments
  - At any given core allocation period, tasks of only one chosen application to run on any given core
  - Clients isolated from each others already at hardware resource access level
- Productivity:
  - Each application, while able to access the full manycore array processing capacity, can be developed, tested and run completely independent of other applications on the system

Security risks prevented at hardware level, not pushed to software

Key Advantages #3 – Contracts Optimized for Cloud

- Straightforward contracts:
  1) Each app assured to get at least up to its contractually entitled share of core capacity (Core Entitlement, CE) whenever its actual demand so warrants
  2) All cores assigned among the apps for each core allocation period (CAP), using demand-driven allocation after condition 1) met
     - → Each client app allowed to burst to up to the full capacity of the shared core pool, so long as condition 1) above met for all apps sharing given pool of cores
- Incentives for system-wide optimized resource usage:
  - Contracts available with differing CE-time profiles (e.g. targeted for apps with business hours, evening-hours and overnight peak demands)
    - Contracts of minimally overlapping CE time profile peaks to share pools of cores
    - → Common costs shared among greater # clients w/o performance drawback
  - Apps billed at CAP (1 microsecond) time granularity based on:
    - CE based on the app's contract, and
    - Numbers of cores allocated to meet the app's demand
    - → Apps incentivized to not demand more cores than what they are able to effectively utilize at any given CAP
  - → <u>Maximum amount of cores available for demand-driven allocation</u> to meet demand peaks of applications able to most effectively use the extra cores

See diagram of billing subsystem on next slide →



 Patents issued and pending.

Key Advantages #4 – Maximized Cloud Cost-Efficiency

- Efficiency:
  - Processing capacity of manycore array used for client programs, not OS tasks
    - Client applications can have their independent software OSs if/as desired
- Ultra-fast responsiveness:
  - Cores of manycore array reassigned among application tasks at microsecond intervals (based on applications' core demands and entitlements), practically without overhead
- Productivity:
  - Parallelization of the programs in development and execution automated through the PaaS integrated development environment and hardware OS
- Flexibility and scalability:
  - Programmable logic implementation enables customizing instances for desired # of cores and # of applications/tasks supported
- Billing techniques to incentivize optimized resource usage

→ Architecturally maximized cost-efficiency for parallel computing era

ThroughPuter – 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 among client applications dynamically sharing a pool of cores from hardware level up
- PRODUCTIVITY:
  - PaaS integrated development environment automate parallel program development and deployment
- OPEN SYSTEM:
  - PaaS software and execution core hardware to be open-sourced
  - Host anywhere; ThroughPuter commercial hosting an option

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 & billing
- Collaboration opportunities:
  - Development environment and tools to enable client programs to most effectively utilize the new hardware features for secure, high-performance and cost-efficient parallel processing in the cloud

Collaborators in parallel processing PaaS domain - please contact:

info@throughputer.com

www.throughputer.com