Process Cruise Control

Event-Driven Clock Scaling for Dynamic Power Management

Frank Bellosa

Andreas Weißel

Department of Computer Science 4 (Operating Systems)

University of Erlangen, Germany

[bellosa,weissel]@cs.fau.de

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Outline

- The Playground for Power Management
  - Mobile Devices
  - Workstations
  - Servers
- Power/performance trade–offs for variable speed systems
- Power/performance characterization
- Event-driven energy characterization
  - System level
  - Task level
- Event-driven frequency scaling (task level)
- Evaluation
  (Intel IQ80310 with XScale 80200 processor @333-733MHz)

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Playground #1: Mobile Devices

- Goal of power management
  - ◆ Increasing the lifetime of batteries
  - ➜ Improvement of the energy efficiency
    - ➜ Throttling the speed of execution (Berkeley, HP Labs, Delft)
    - ➜ Avoiding unnecessary device activation (Stanford, Bologna, Duke)
    - ➜ Cooperation between applications
    - ➜ Page compression and paging in multibank systems
  - ◆ Guarantee of a pre-defined stand-by and active time
    - ➜ Scheduling with respect to an energy budget (Duke)
- Question:
  When and at which speed is a certain activity allowed to take place?

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Playground #2: Workstations

- Goals of power management
  - ◆ Avoiding the pollution of the working environment
  - ➜Reducing the noise
  - ➜Thermal management (HP Labs)
  - ◆ Cutting the operating costs (power/cooling)
  - ➜Reducing the energy consumption
    - ➜Throttling the speed of execution
    - ➜Halting unused devices
- Questions:
  What is the environmental status?
  When and at which speed is a certain activity allowed to take place?

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Playground #3: Servers

- Goal of power management
  - ◆ Reducing the operating costs (power, A/C, floor space)
  - ➜Reducing the energy consumption
    - ➜Throttling the speed of execution
    - ➜Batching of requests (IBM ARL)
    - ➜Balancing the load among servers (IBM Raleigh & Duke)
  - ◆ Billing of operating costs to customers
  - ◆ Operation with emergency power supply
  - ➜Short term throttling of average power
- Question:
  Where, at which speed and at which price
  is an activity to be allowed to take place?

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

The Need for Energy Characterization

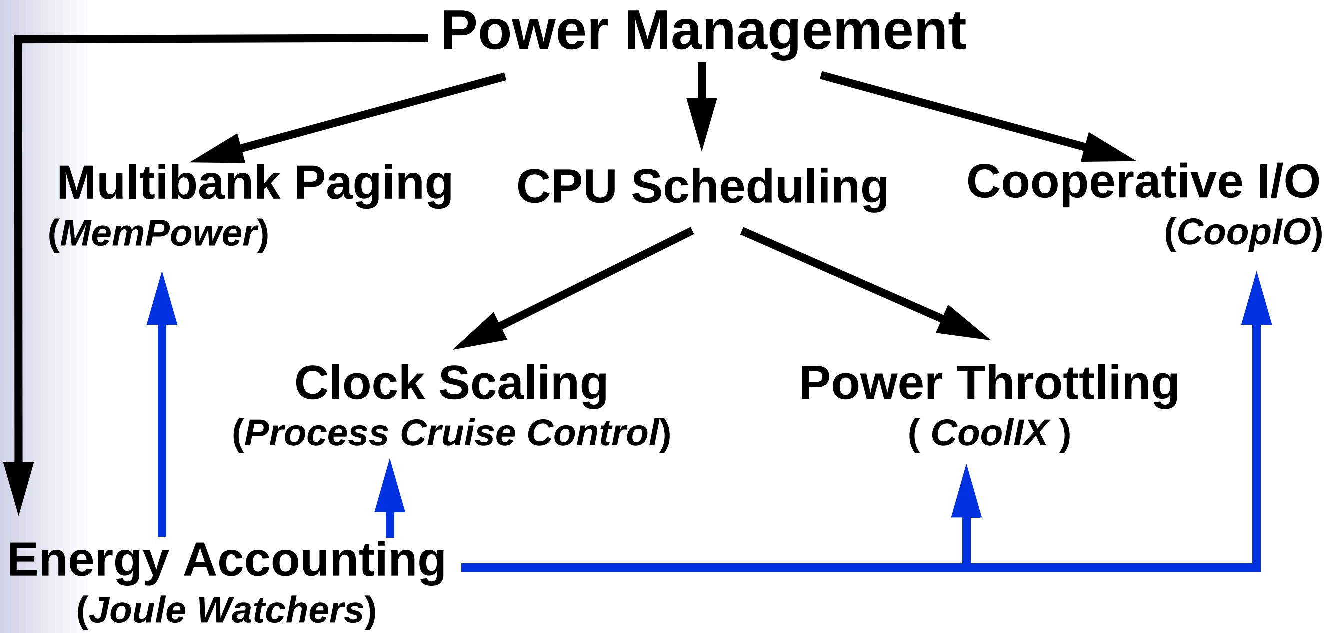
- Power Mangement has to decide
  - ◆ When
  - ◆ Where
  - ◆ How fast
  - ◆ Under which environmental conditions
  - ◆ At which price for power and cooling

  any activitiy in the system is allowed to take place.
- The energy-related behavior of the system is the basis of all power-mangement decisions.
- Because we cannot predict the behavior of the system off-line, we have to observe the system!

  ➜**The Case for On-Line Energy Characterization**

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Power Management Projects at the U of Erlangen



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Power/Performance Trade-Offs

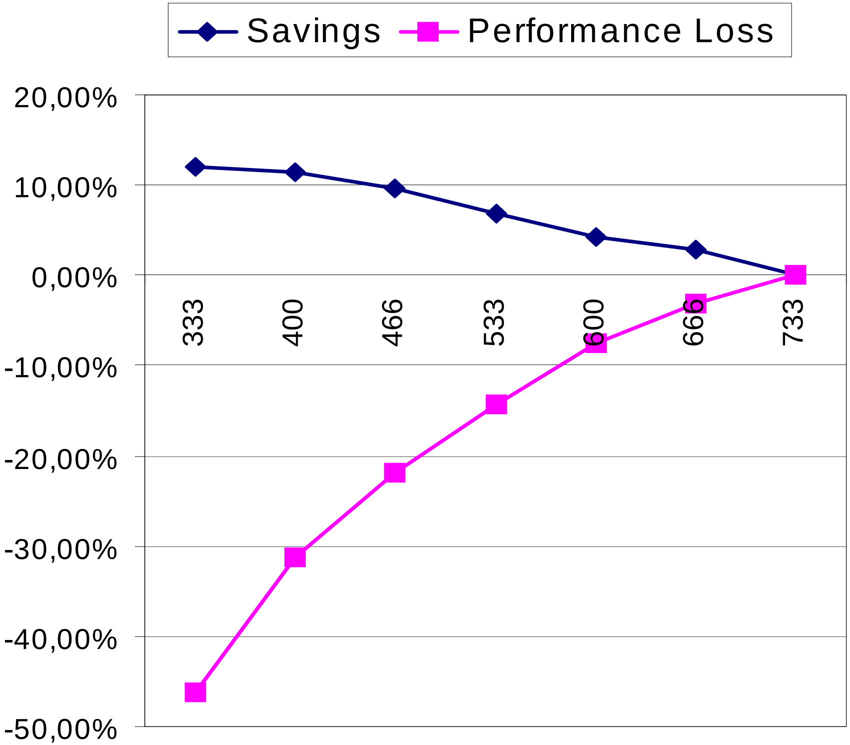
- Minimizing the energy consumption
- Reducing energy consumption for given performance requirements
  - Maximal 10 % loss in computational performance
  - Maximal 30 % increase in response time
- Task-specific power/performance trade-offs
  - Background tasks (e.g., daemons, low-priority system tasks)
  - High-priority system tasks
  - User tasks with individual performance requirements

➜ Power/performance trade-offs require the knowledge about the power/performance characteristics of the system and the tasks!

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Clock Scaling: Energy vs. Performance Profile I

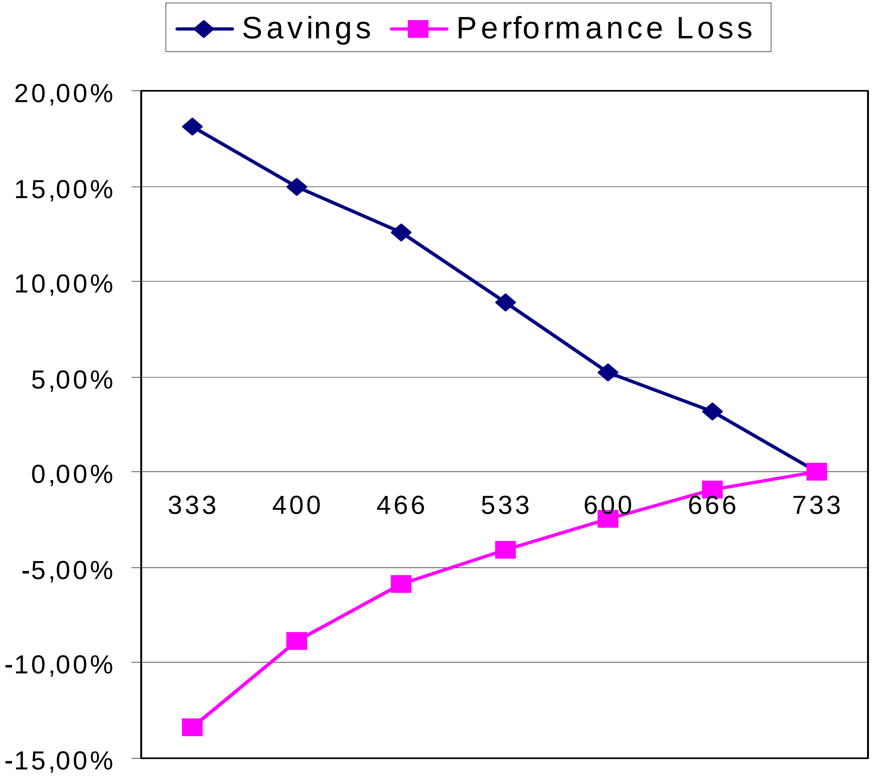
- Energy savings vs. performance loss for "`factor`"



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Clock Scaling: Energy vs. Performance Profile II

- Energy savings vs. performance loss for "`grep`"



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

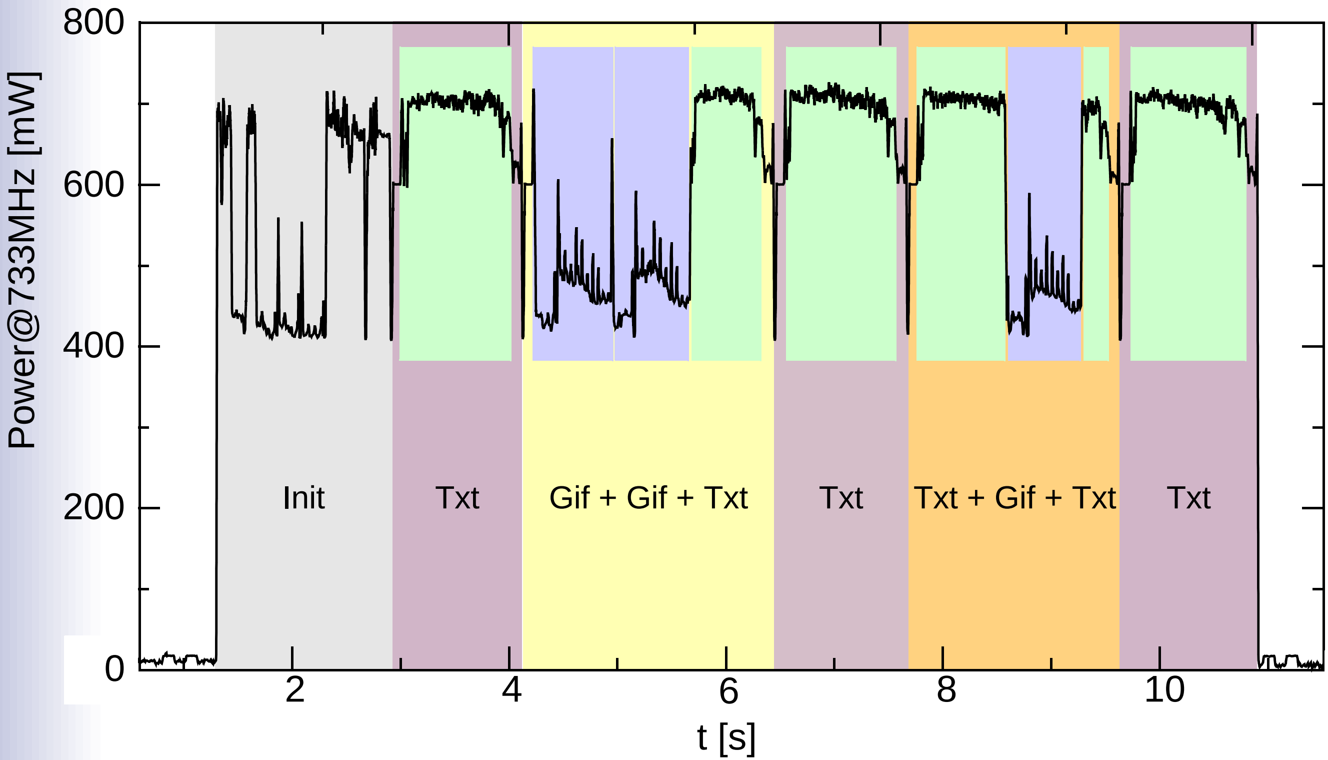
Profiling of Complex Applications

- Complex applications show unpredictable characteristics due to
  - Multiple blocks of computations with unknown execution order
  - Unknown input data
  - Unknown working set size
  - Examples: interpreters, decoders, virtual machines

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Power Profile of an Interpreter

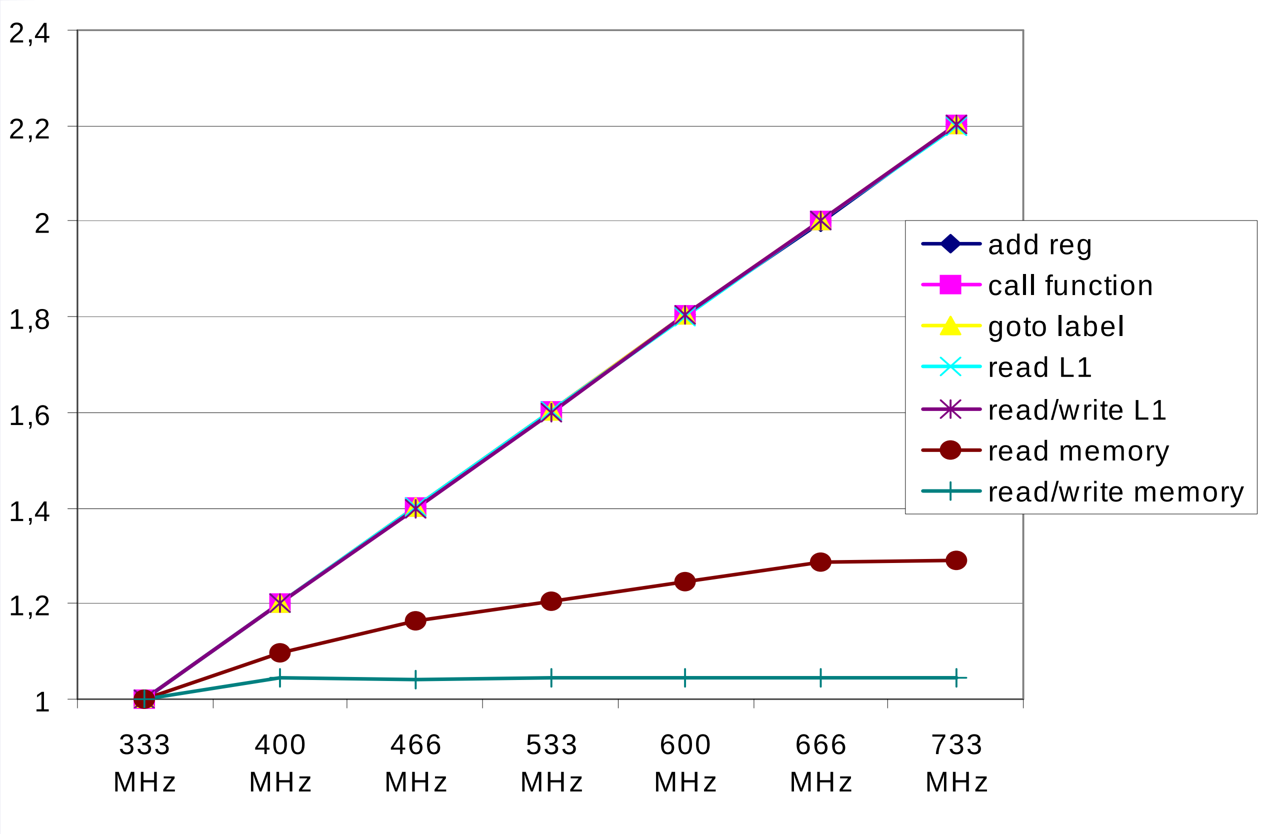
- Example: ghostscript interpreter gs



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Clock Scaling: Performance Characterization

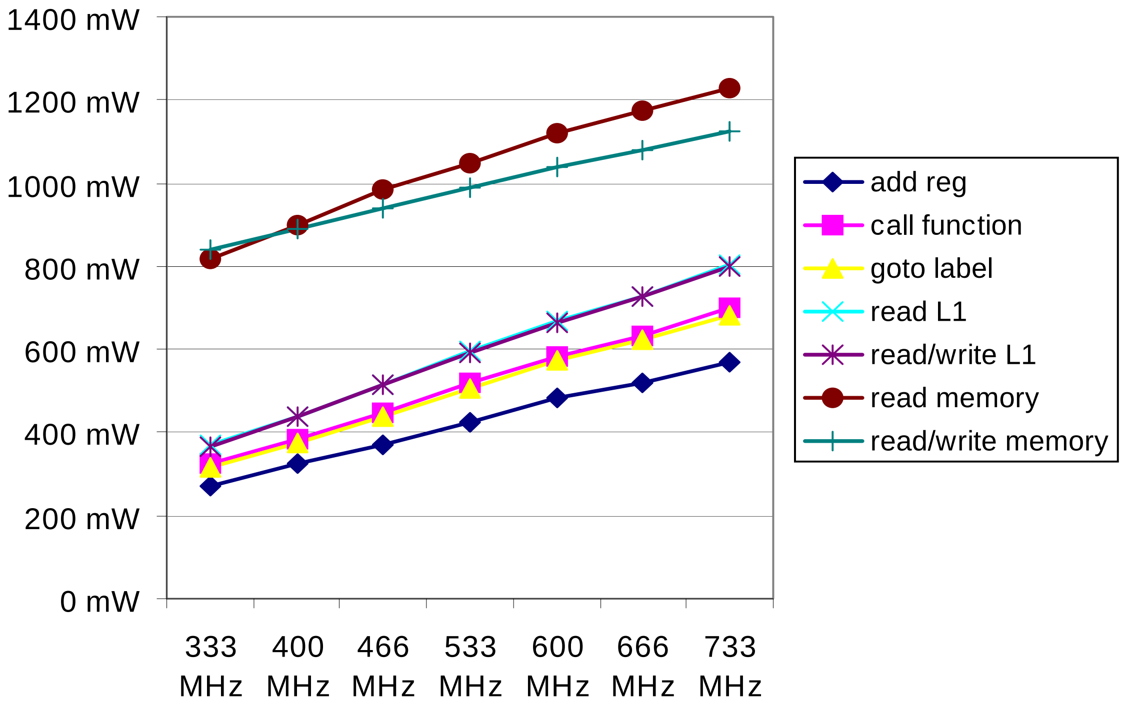
- Correlation between performance and clock frequency



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Clock Scaling: Power Characterization

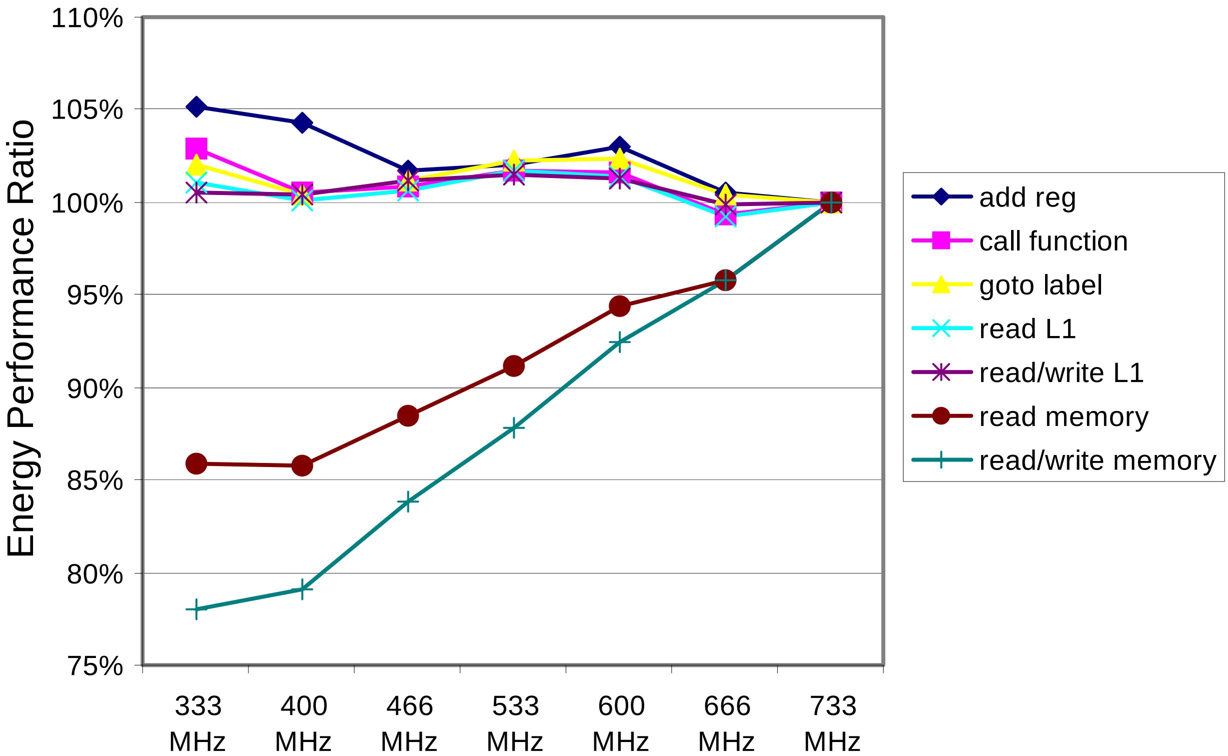
- Correlation between power consumption and clock frequency:



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

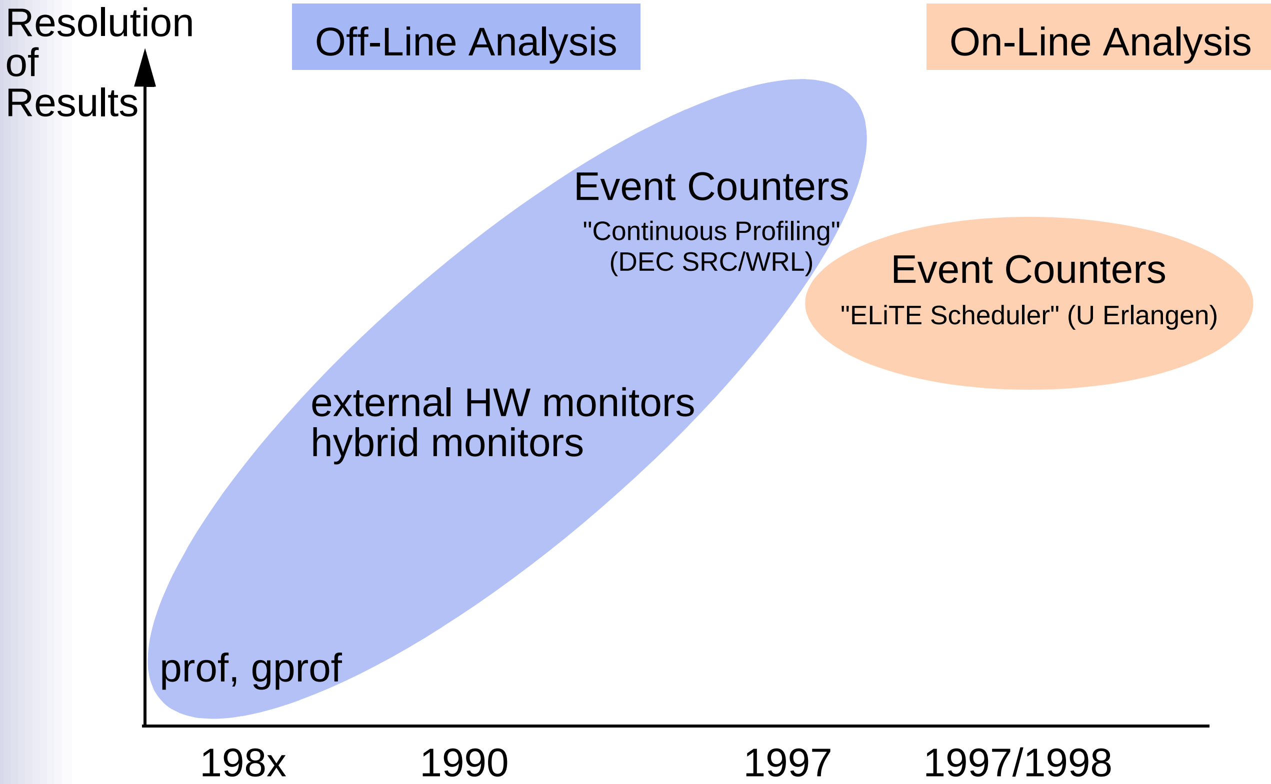
Energy Efficiency is Unpredictable

- Correlation between energy efficiency and clock frequency



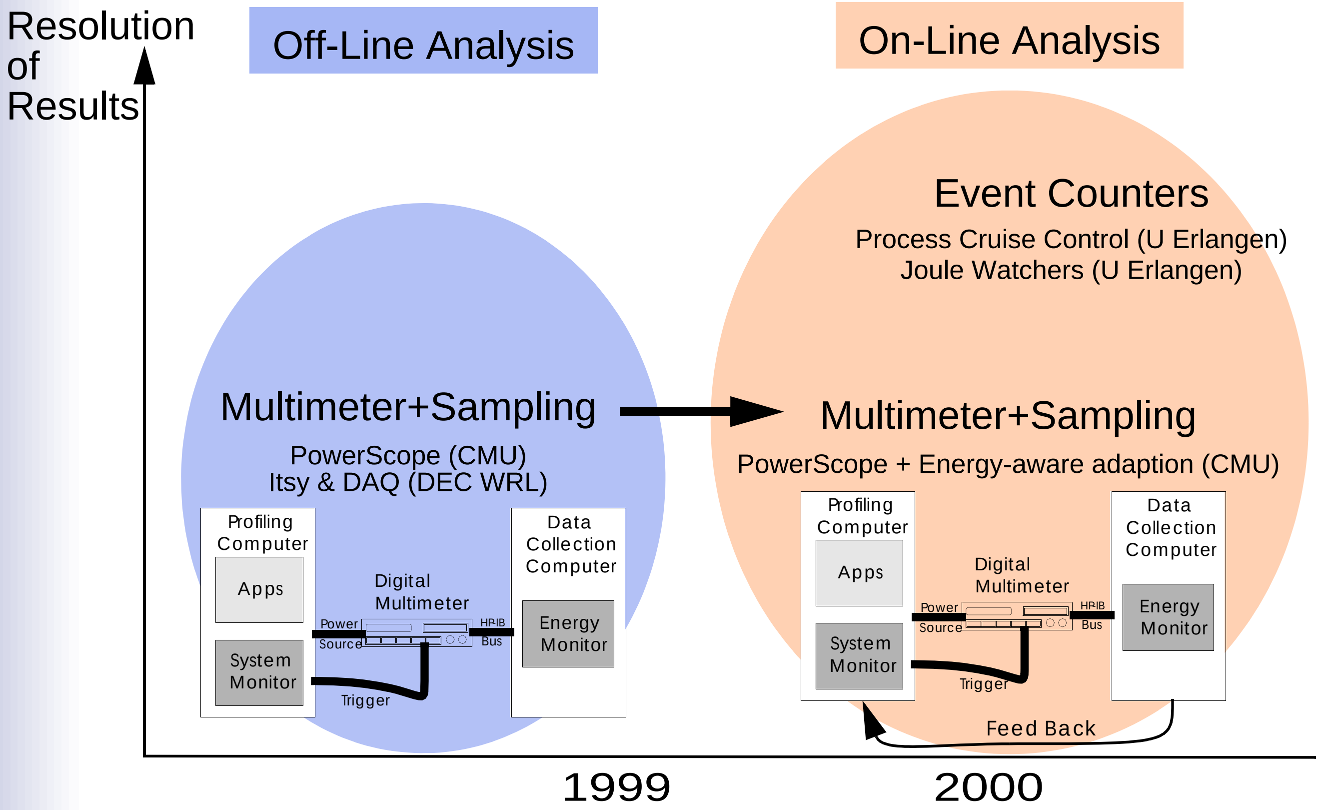
Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Evolution of Performance-Profiling



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

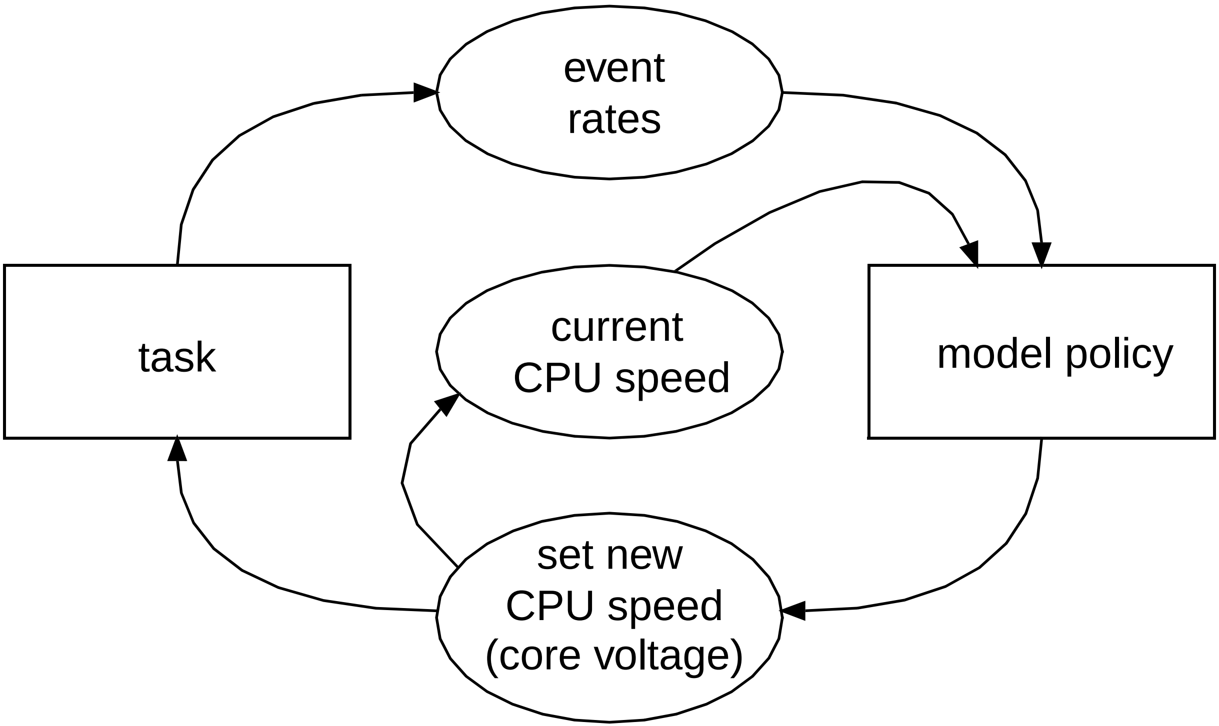
Evolution of Energy-Profiling



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Process Cruise Control: Principles of Operation

- The rates at which some events happen correspond to a specific performance/energy-efficiency profile.



© Frank Bellosa • Universität Erlangen-Nürnberg • Computer Science 4 (Operating Systems), 2002

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

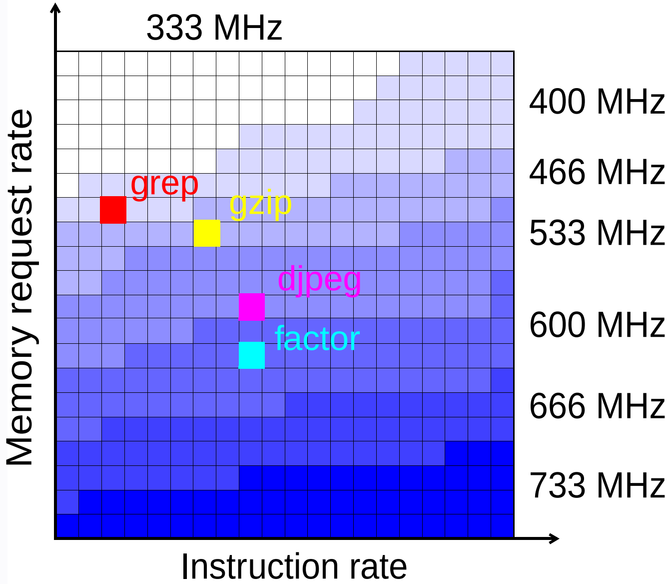
Event-Based Characterization

- Challenge:
  - Selection of the appropriate events
    - Clock cycles, retired instructions, L1 cache & memory/bus access...
  - Finding the correlation between event rates and clock speed for predefined power/performance demands.
- Methodology of event-based power/performance characterization:
  - Countable events are triggered by synthetic training applications.
  - Energy at various clock speeds is measured by a DAQ system.
  - The rate of various events is determined.
  - The performance of the training application is evaluated.
  - For each training application the optimal clock speed is determined for a predefined performance penalty.

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Frequency Domains for IQ80310

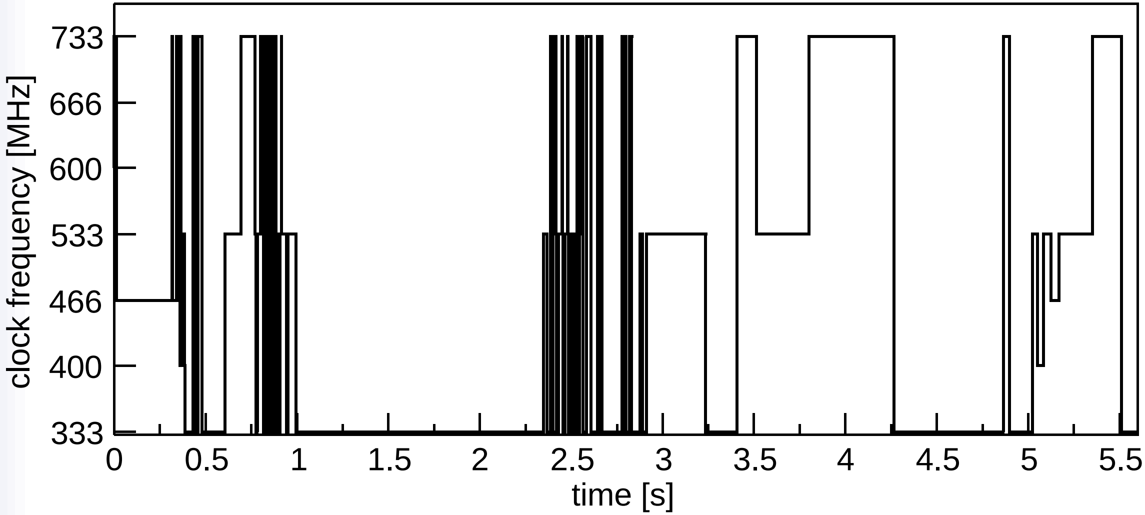
- Selection of events:
  - Instruction rate is an indicator for performance loss.
  - Memory request rate is an indicator for energy efficiency gains.
- Example: Partitioning of the event space into frequency domains for a 10 % loss in computational performance



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Measurements

Application	optimal speed	Process Cruise Control: clock scaling	Process Cruise Control Energy Savings
grep	400 MHz	400 MHz	15%
gzip	466 MHz	466 MHz	10%
djpeg	600 MHz	533 MHz	8%
factor	600 MHz	600 MHz	4%
ghostscript		dynamic	< 5%



Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Advantages

- Power/Performance characteristics become a first class element of the process context.
- Low memory overhead:
  Just one entry in the task structure per event
- Low algorithmic overhead**:**
  Counters are evaluated
  - in the timer interrupt handler
  - in the scheduler
- High temporal resolution:
  Characteristics can be measured for arbitrarily short periods.
- Fast response:
  Changes in the energy-related characteristics of a process can be registered promptly.

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Limitations

- The characterization of Cruise Control is
  - sufficient for best-effort improvement of the energy efficiency.
  - inadequate for quantitative guarantees.
- Just 2 counters are available in most architectures.

  ➜Let's embed more counters!
- HW events were selected with focus on performance not on energy profiling.

  ➜Let's find events that refer to a specific energy consumption!

  ➜**E**nergy **M**onitoring **C**ounters (EMCs)

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.

Conclusion

- Energy aware OS strategies require information that is
  - precise
  - immediately available

- Performance penalties are the result of HW latencies.

  ➜Let's count performance critical HW activations!

- Energy consumption is the result of HW activity.

  ➜Let's count energy intensive HW activations!

➜ Event-driven power/performance characterization is the basis of energy-aware systems.

© Frank Bellosa • Universität Erlangen-Nürnberg • Computer Science 4 (Operating Systems), 2002

Reproduktion jeder Art oder Verwendung dieser Unterlage, außer zu Lehrzwecken an der Universität Erlangen-Nürnberg, bedarf der Zustimmung des Autors.