ENERGY EFFICIENT SOCS
A REINFORCEMENT LEARNING APPROACH

Anca Molnos, CEA LETI, Grenoble, France

16th MPSoC forum – Nara, Japan, July 11-15, 2016

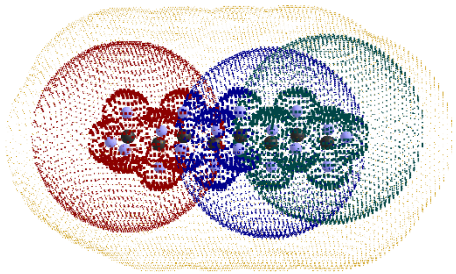
OVERVIEW

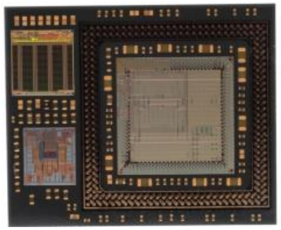
1. State-of-the-art and challenges
2. Reinforcement learning formulation of the energy minimisation problem
3. Some experimental results
4. Conclusions and perspectives
    - FDSOI: DFVS vs DBB

FROM RESEARCH TO INDUSTRY
cea tech

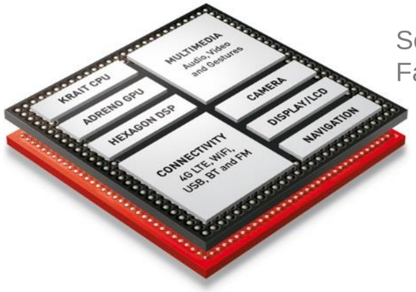
SOC ENERGY MANAGEMENT CONTEXT

Source: Urban Farming News





Source: NXP



Source: Fairphone

Source: HPE

- **Problem: design a *manager* (power governor, run-time,…) that:**
  - Minimise energy under performance constraints
  - Optimize energy efficiency under power and performance constraints
  - Mitigate heating
  - Maximise performance under power capping
  - …
- **Actuators**
  - Hardware units operating point (mode idle, mode sleep, DVFS, …)
  - Load balancing of the application software on the hardware units

STATE-OF-THE-ART: 2 MAIN VIEWS

Many-core/dark silicon	Heterogeneous processors (CPU/GPU)
Assumptions	
• One can compute the power • The exactly execution profile is known (start/end/execution/idle times, …)	• One can measure the power • Only the values of performance counters are known

$$E_i(V_i, V_{ti}) = R_i C_i V_i^2 + T_i k_i V_i e^{\left(-\frac{V_t}{S_t}\right)}$$

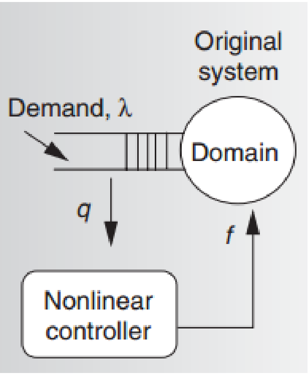
$$\overline{q}'_k = \bar{q}_{k-1} + \frac{T}{2}\left(\bar{\lambda}_{k-1} - \bar{\mu}_{k-1}\right)$$

$$\bar{q}_k = \bar{q}_{k-1} + \left(\bar{\lambda}_{k-1} - \bar{\mu}_{k-1}\right)T$$

$$e_k = \overline{q}'_k - q_{\mathrm{ref}}$$

$$\bar{\mu}_k = \bar{\mu}_{k-1} + K_I e_k + K_P\left(e_k - e_{k-1}\right)$$

$$f_k = \frac{\bar{C}_2 \bar{\mu}_k}{1 - \bar{t}_1 \bar{\mu}_k}$$



Counters	
CSBusy	HSALUTexRatio
CSTime	HSTexBusy
DepthStencilTestBusy	HSTexInstCount
DSBusy	HSPatches
DSTime	HSSALUBusy
GPUTime	HSSALUInstCount
GPUBusy	HSVALUBusy
GSBusy	HSVALUInstCount
GSTime	GSALUBusy
HSBusy	GSALUEfficiency
HSTime	GSALUInstCount
InterpBusy	GSALUTexRatio
PrimitiveAssemblyBusy	GSExportPct
PSBusy	GSPrimsIn
PSTime	GSSALUBusy
ShaderBusy	GSSALUInstCount
ShaderBusyCS	GSTexBusy
ShaderBusyDS	GSTexInstCount
ShaderBusyGS	GSVALUBusy
ShaderBusyHS	GSVALUInstCount
ShaderBusyPS	GSVerticesOut
ShaderBusyVS	ClippedPrims
TessellatorBusy	CulledPrims
TexUnitBusy	PAPixelsPerTriangle
VSBusy	PAStalledOnRasterizer
VSTime	PrimitivesIn

…

STATE-OF-THE-ART: 2 MAIN VIEWS

Many-core/dark silicon	**Heterogeneous processors (CPU/GPU)**
Assumptions	
• One can compute the power • The exactly execution profile is known (start/end/frame times, …)	• One can measure the power • Only the values of performance counters are known
Main challenges addressed	
• Models to fit application dynamics and methods to solve for, e.g., frequency	• Load balancing and metrics for application performance power ↔ consumption
Approaches	
• Control theory: simple/complex controllers [Wu04,Garg10,David12,Bartolini12, Bogdan13, Rahmani15 ..] • Learning [Tan09,Liu10,Jung10,Dhiman11,Wang11,Ye14, Das14,Triki14,Khan14,Chen15] • Theoretical proofs	• Power-performance prediction models [Dhiman09,Paul13,VRodriguez13,Pathania15] • Control theory • Simple controllers [Wang14,Dietrich14, Pathania14] • Experimental validation (large traces)
Evaluation	
• Simulation (no thorough cost investigation)	• Execution and measurements
• **Over-simplified view on software** • **Assume a too high observability of application and hardware**	• **Too low observability of application in relation to the hardware and power consumption**

CHALLENGES

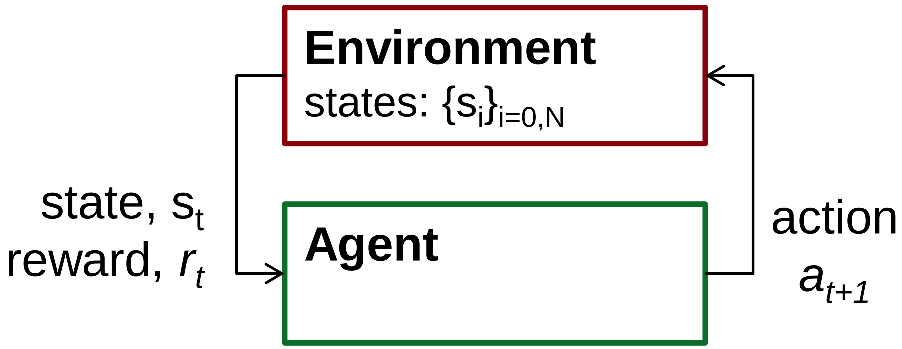
- **Deal with "non-ideal" cases**
  - Application are too dynamic, exact models are hard to obtain
    - Dependent on the content of the input data
  - Application performance is not proportional to the operating point (e.g., core clock frequency)
    - Contention at shared resources, memory accesses
    - System software
      - Task scheduling, timers, interrupts
- **(Relatively) low computation cost of energy optimization**
    - No large matrix inversion
    - Small state space
    - No/limited costly functions (exponential, trigonometric)
    - No/limited least mean squares
    - …
- **Define hardware ↔ manager ↔ application interfaces for good application and power consumption visibility**
  - Power consumption should probably be the concern of the application and the operating system
  - In which parts of the application is the power consumed

CHALLENGES

- **Deal with "non-ideal" cases**
  - Application are too dynamic, exact models are hard to obtain
    - Dependent on the content of the input data
  - Application performance is not proportional to the operating point (e.g., core clock frequency)
    - Contention at shared resources, memory accesses
    - System software
      - Task scheduling, timers, interrupts
- **(Relatively) low computation cost of energy optimization**
    - No large matrix inversion
    - Small state space
    - No/limited costly functions (exponential, trigonometric)
    - No/limited least mean squares
    - …
- **Define hardware ↔ manager ↔ application interfaces for good application and power consumption visibility**
  - Power consumption should probably be the concern of the application and the operating system
  - In which parts of the application is the power consumed

REINFORCEMENT LEARNING (RL) APPROACH

- **RL algorithm**
  - an *agent*, which aims to learn from the interaction with an *environment* (trial and error) to achieve a *goal (maximize reward)* [Sutton1998]
- **Adaptive (indirect) controller**

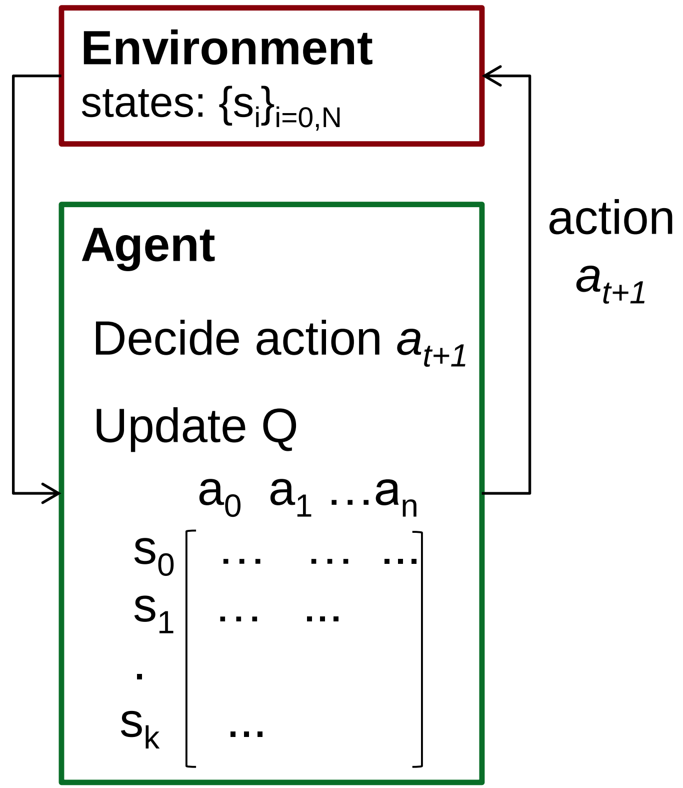


- **Advantages**
  - no model of the dynamics of the system
  - the learning is not (really) supervised
    - training happens in the same time with the optimization process (exploration/exploitation)
    - no need to know relevant use-cases
- **Disadvantages**
  - may take poor decisions during exploration
  - fully-fledged versions may have a large overhead
  - theoretical proofs (convergence speed, optimality) are valid only for discrete&finite state space, stationary, markovian systems (with some extension to semi-markovian cases)
    - however known to work ok in non-stationary, non-markovian, systems.

LOW COST Q-LEARNING

- Q-learning: the learned knowledge of the agent → a value table Q(states,actions)
  - Decided action, given s, Q → Ɛ-greedy
    - Exploitation: argmax(Q(s))
    - Exploration: some random value
  - Construct the Q table
    - Initial value of the Q table
    - Update formula:

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha \left[r_{t+1} + \gamma \max_{a'} Q_t(s',a') - Q_t(s,a) \right]$$



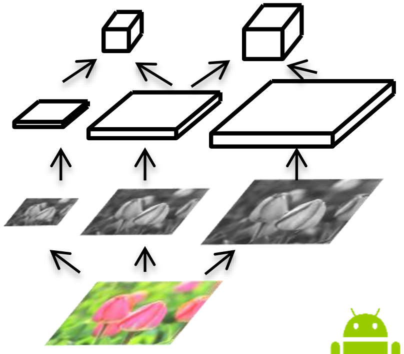
- Potential to learn complex behavior
  - No free lunch:
    - define the states
      - queues filling indicates application progress
    - set its parameters
      - experimental investigation
    - define a reward function
    - …

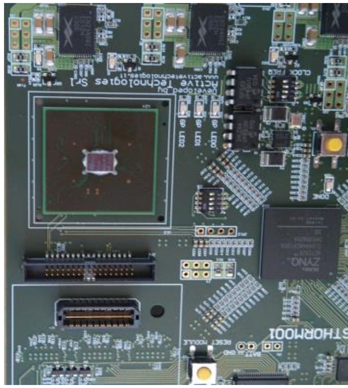
A.Molnos, S.Lesecq, J.Mottin, D.Puschini, "Investigation of Q-learning applied to DVFS management of a System-on-Chip". In 4th IFAC International Conference on Intelligent Control and Automation Sciences, ICONS (2016).

FROM RESEARCH TO INDUSTRY cea tech

RESULTS

- Test board
  - ARM host processor and an SoC with 16 processors elements.
  - Android OS + in-house run-time
  - Application: a part of a HMAX object recognition
  - 15%-44% energy reduction wrt. state-of-the-art
  - similar number of throughput violations
  - lightweight manager: 0.7% of application time, 1KB footprint

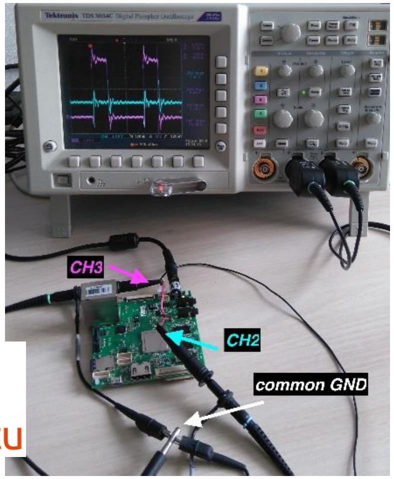




android

- Commercial board (on going work)
  - IMX6: dual-core ARM + GPU
  - Ubuntu OS, OpenCL
  - Application: obstacle detection

ubuntu



CONCLUSIONS

- **Challenges in managing energy consumption in SoCs**
  - Deal with “non-ideal” cases
  - (Relatively) low computation cost of energy optimization
  - Define hardware ↔ manager ↔ application interfaces
- **An adaptive Q-learning-based approach**
  - Model-free, potential to learn complex behaviour
- **Experiments with promising results**
  - Higher energy reductions than state-of-the-art methods addressing the same application domain.

PERSPECTIVES

- **More advanced RL methods that lead to quicker convergence and higher adaptability**
  - e.g., actor-critic, hierarchical, TD(λ), best-match → their applicability and costs should be evaluated.
  - multi-criteria optimization: temperature, aging
- **Better methods to deal with non-Markovian, non-stationary systems**
- **Deal with complex energy consumption models in future technologies**

FDSOI28 PARAMETRIC EXPLORATION

- **How to choose the best power/energy optimization strategy (DBB vs DVFS) to implement?**
- **Do we really need Dynamic Body Bias (DBB) mechanism in our future design?**

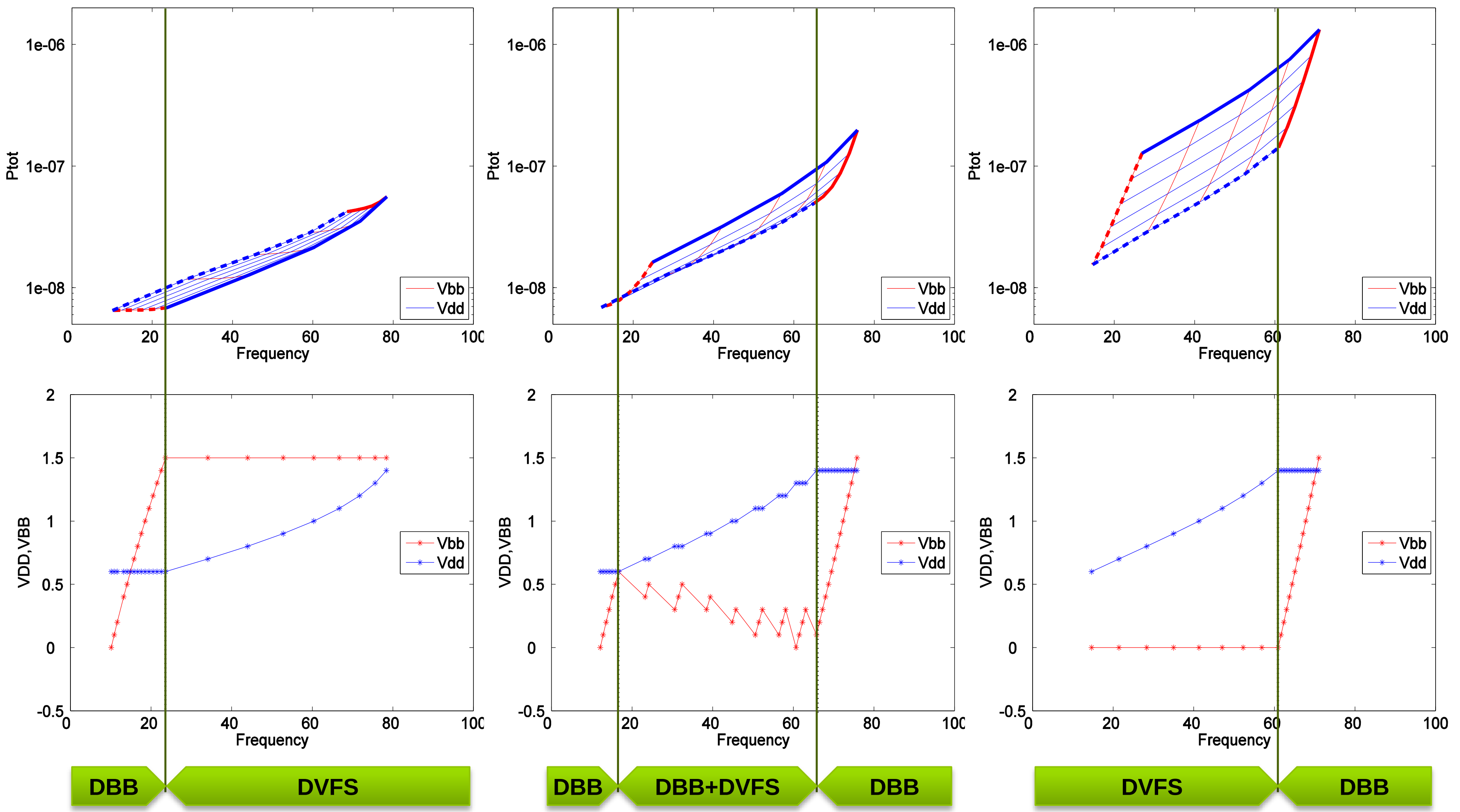
Inputs	Static Parameters		Substrate	RVT, LVT
			Poly Bias	0nm, 4nm, 10nm, 16nm
			Process Corner	SS, TT, FF
	Dynamic Parameters	User controlled	Vdd	[0.6V, 1.4V]
			Vbb in LVT	[0V, 1.5V]
			Vbb in RVT	[−1.5V, 0V]
		Operating Conditions	Temp	[−40°C, 125°C]
			Toggle Rate	[0.1%, 50%]
Measurements			Fmax	
			Iswitch	
			Ileak	

Electrical simulation of a ring oscillator

"Body Bias usage in UTBB FDSOI designs: A parametric exploration approach", D. Puschini, J. Rodas, E. Beigne, M. Altieri, S. Lesecq. Special Issue: Planar Fully-Depleted SOI technology, Solid-State Electronics Elsevier Journal, Volume 117, March 2016, Pages 138-145

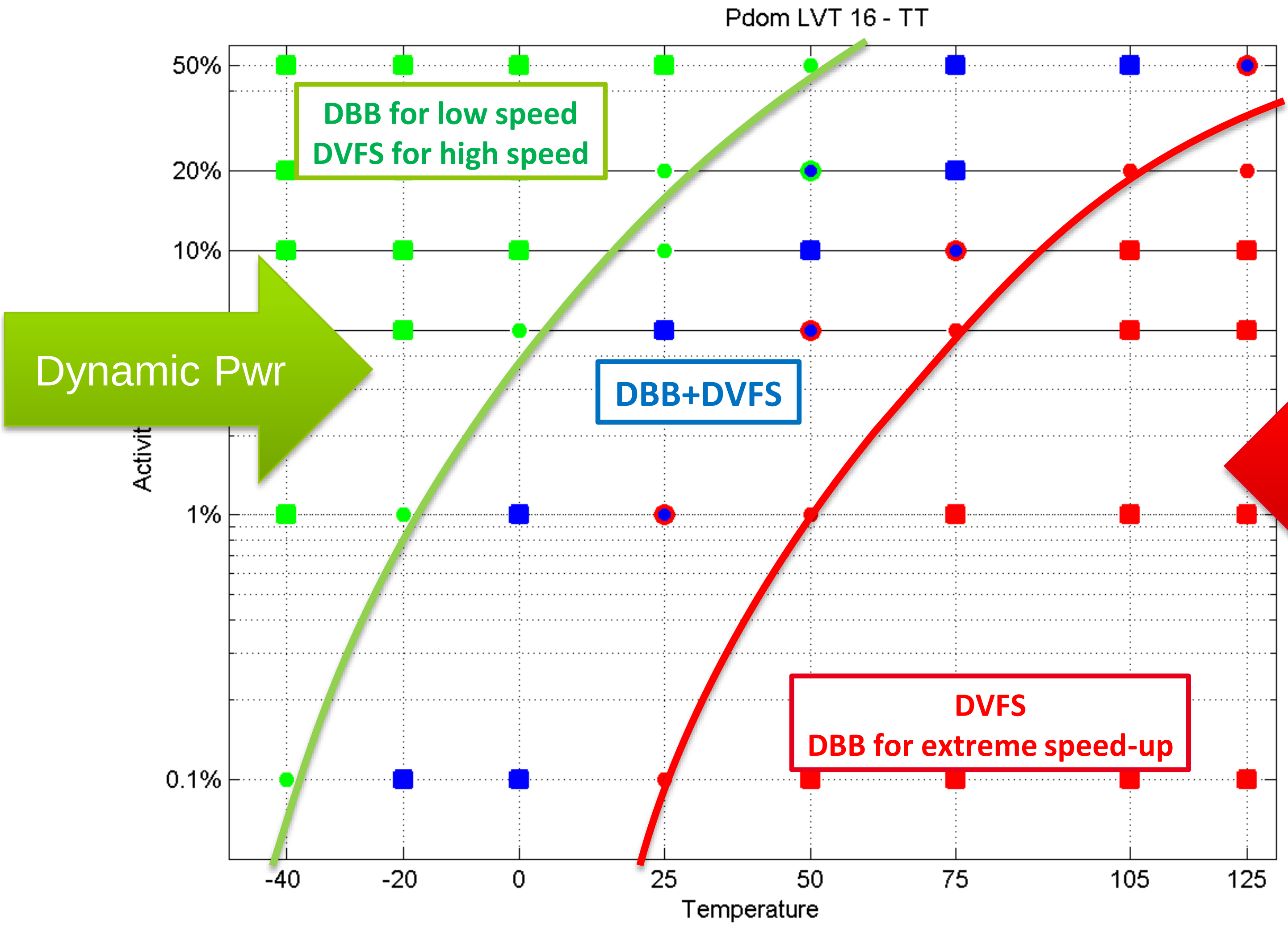
STATIC AND DYNAMIC POWER CONSUMPTION

Dynamic Power Dominates | Dynamic ≅ Static | Static Power Dominates



T= -20° | T= 50° | T= 125°

DBB VS. DVFS



Process TT
Vbb = -1.5V to 0V
Vdd = 0.6V to 1.4V

THANK YOU!

REFERENCES

[Sutton98]. Sutton, R.S. and Barto, A.G. Reinforcement learning: An introduction, volume 1. MIT press Cambridge.
[Chen15] Chen, Z. and Marculescu, D. Distributed Reinforcement Learning for Power Limited Many-core System Performance Optimization. DATE 1521-1526.
[Das14] Das, A. et al. Reinforcement Learning-Based Inter and Intra-Application Thermal Optimization for Lifetime Improvement of Multicore Systems. In Proceedings of the The 51st Annual Design Automation Conference on Design Automation Conference, 2014
[Khan14] Khan, U. and Rinner, B.. Online learning of timeout policies for dynamic power management. ACM TECS, 13(4).
[Tan09] Tan, Y.T.Y., Liu, W.L.W., and Qiu, Q.Q.Q. (2009). Adaptive power management using reinforcement learning. IEEE/ACM ICCAD.
[Triki15] Triki, M., Wang, Y., Ammari, A.C., and Pedram, M. . Hierarchical power management of a system with autonomously power-managed components using reinforcement learning. Integration, 48, 10-20.
[Alimonda09] Alimonda, A., Carta, S., Acquaviva, A., Pisano, A., and Benini, L. "A feedback-based approach to DVFS in data-flow applications". IEEE Trans. Comput. Aided Des. Integr. Circuits Sys. 28, 11, 1691–1704.
[Ogras09] U.Y. Ogras, Umit Y. Ogras, R. Marculescu, D. Marculescu, and E. Gu Jung, "Design and Management of Voltage-Frequency Island Partitioned Networks-on-Chip," IEEE Trans. on Very Large Scale Integration Systems, vol. 1 7, no. 3, pp. 330-341 , 2009.
[Garg10] Garg, S., Marculescu, D., and Marculescu, R. "Custom feedback control: Enabling truly scalable on-chip power management for MPSoCs". In Proceedings of the 16th ACM/IEEE International Symposium on Low Power Electronics and Design (ISLPED'10). IEEE, 425–430.
[David12] R. David, P. Bogdan, R.Marculescu, "Dynamic power management for multicores: Case study using the Iintel SCC". VLSI-SoC 2012: 147-152
[Wu04] Q. Wu, P. Juang, M. Martonosi, M., and Clark, D. W. "Formal online methods for voltage/frequency control in multiple clock domain microprocessors". In Proceedings of the 11th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'04). ACM, 248–259,2004
[Carta07] Salvatore Carta, Andrea Alimonda, Alessandro Pisano, Andrea Acquaviva, and Luca Benini. "A control theoretic approach to energy-efficient pipelined computation in MPSoCs". ACM Trans. Embed. Comput. Syst. 6, 4, Article 27, September 2007.
[Bogdan13] Paul Bogdan, Radu Marculescu, and Siddharth Jain. "Dynamic power management for multi-domain system-on-chip platforms: An optimal control approach". ACM Trans. Des. Autom. Electron. Syst. 18, 4, Article 46, October 2013.
[Zanini12] F. Zanini, D. Atienza, C. N. Jones, "Online Thermal Control Methods for Multiprocessor Systems, ACM Transactions on Design Automation of Electronic Systems", Vol. 18, No. 1, Article 6, December 2012.
[Bartolini12] A. Bartolini, M. Cacciari, A. Tilli, and L. Benini, "Thermal and energy management of high-performance multicores: Distributed and self-calibrating model-predictive controller," IEEE Trans. Parallel Distrib. Syst., vol. 24, no. 1, pp. 170–183, 2012
[Almeida11] G. Almeida, R. Busseuil, L. Ost, F. Bruguier, G. Sassatelli, P. Benoit, L. Torres, and M. Robert, "PI and PID regulation approaches for performance-constrained adaptive multiprocessor system-on-chip," Embedded Systems Letters, IEEE, vol. 3, no. 3, pp. 77–80, Sept 2011.
[Paul13] I. Paul, V. Ravi, S. Manne, M. Arora, and S. Yalamanchili, "Coordinated energy management in heterogeneous processors," in Proceedings of SC13: International Conference for High Performance Computing, Networking, Storage and Analysis, ser. SC '13. ACM, 2013.
[Pathania15] Anuj Pathania, Alexandru Eugen Irimiea, Alok Prakash, and Tulika Mitra. 2015. Power-Performance Modelling of Mobile Gaming Workloads on Heterogeneous MPSoCs. In Proceedings DAC 2015.
[Rahmani15] A.-M. Rahmani et al. Dynamic Power Management for Many-Core Platforms in the Dark Silicon Era: A Multi-Objective Control Approach. In Proc. Int. Symp. on Low Power Electronics and Design, ISLPED, 2015
[Dhiman09] Gaurav Dhiman, Giacomo Marchetti, and Tajana Rosing. 2009. vGreen: a system for energy efficient computing in virtualized environments. In Proceedings ISLPED 2009
[VRodriguez13] N. Vallina-Rodriguez and J. Crowcroft, "Energy management techniques in modern mobile handsets," IEEE Commun. Surveys Tuts., vol. 15, no. 1, pp. 179–198, 2013.
[Wang14] Yue Wang and Nagarajan Ranganathan. 2014. A Feedback, Runtime Technique for Scaling the Frequency in GPU Architectures. In Proceedings ISVLSI 2014
[Dietrich14] B. Dietrich and S. Chakraborty, "Forget the battery, let's play games!" in ESTIMedia, 2014.
[Pathania14] A. Pathania, Q. Jiao, A. Prakash, and T. Mitra, "Integrated CPU-GPU power management for 3D mobile games," in DAC, 2014.