Performance-Influence Models

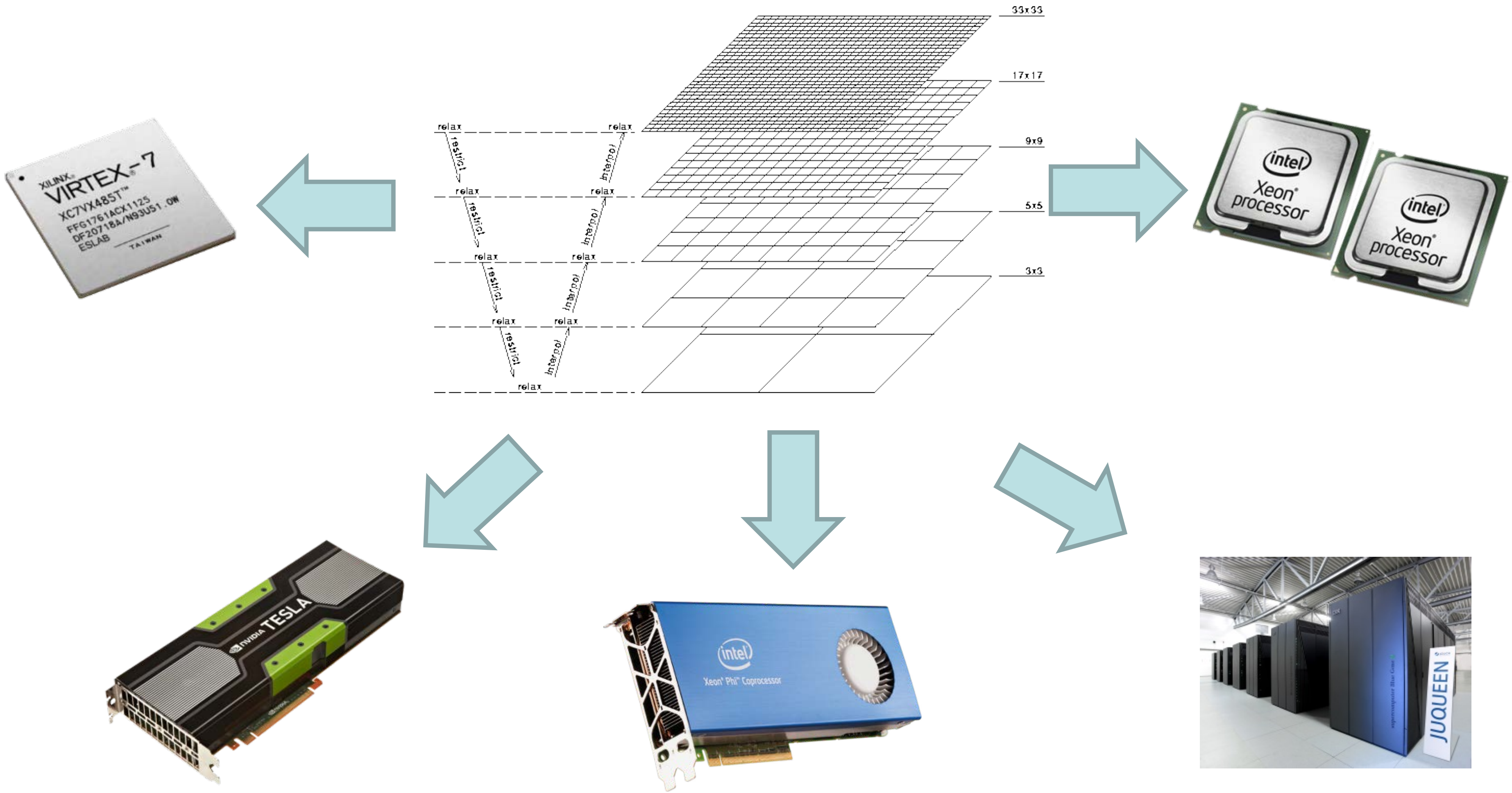
Alexander Grebhahn*, *Norbert Siegmund, Sven Apel

University of Passau

Performance Modeling: Methods & Applications

July 2015

Which is the Optimal Configuration for a given Hardware Platform?



320 optinal binary options lead to more configurations than the expected number of atoms in the universe

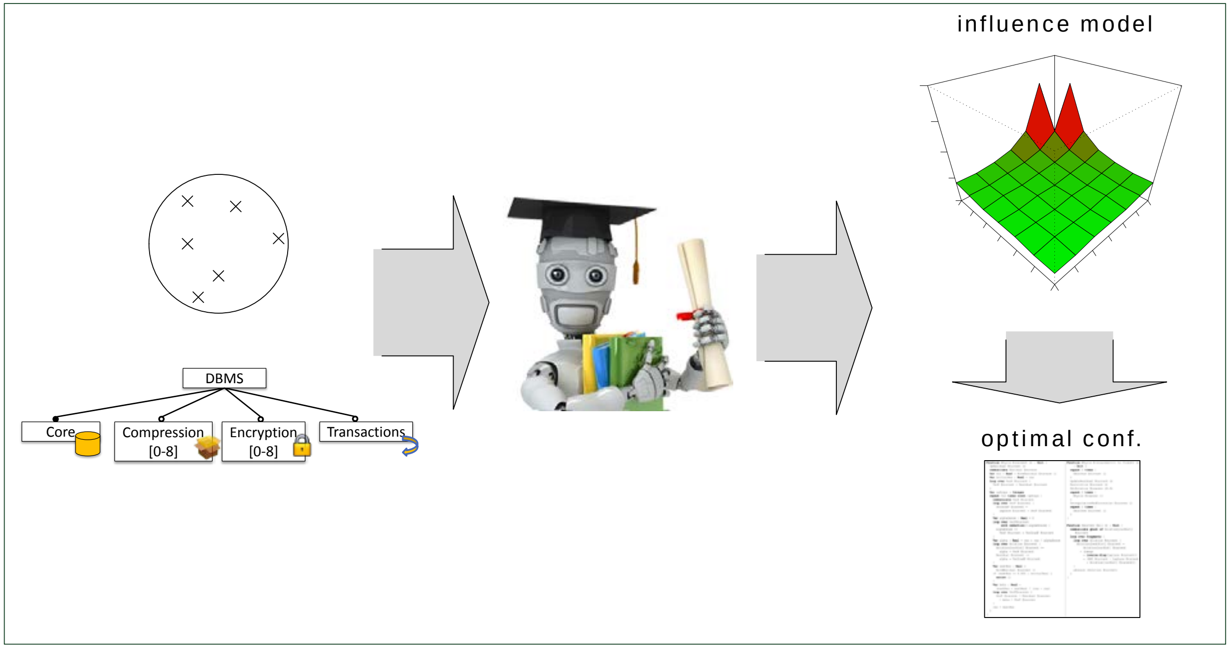
Numeric options make things much worse

Optimal configuration for

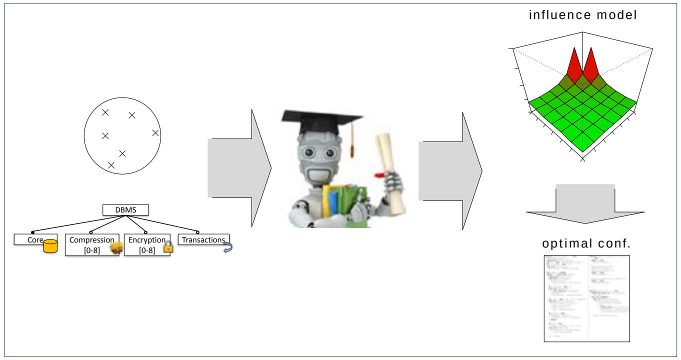
Optimal configuration for

- Use Machine Learning
  - Pros:
    - Automated
    - Many tools, much research
  - Cons:
    - Overfitting, underfitting
    - Not tailored to the application domain

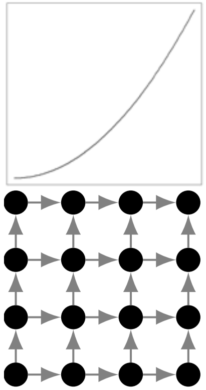
- Use Machine Learning
  - Pros:
    - Automated
    - Many tools, much research



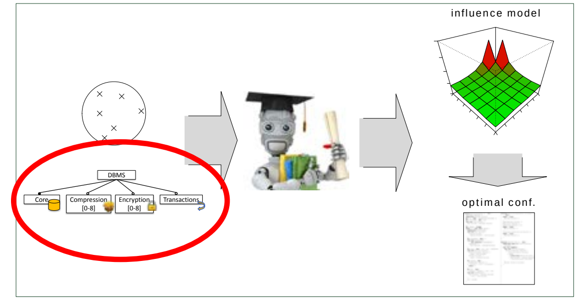
- Use Machine Learning
  - Pros:
    - Automated
    - Many tools, much research
  - Cons:
    - Overfitting, underfitting
    - Not tailored to the application domain

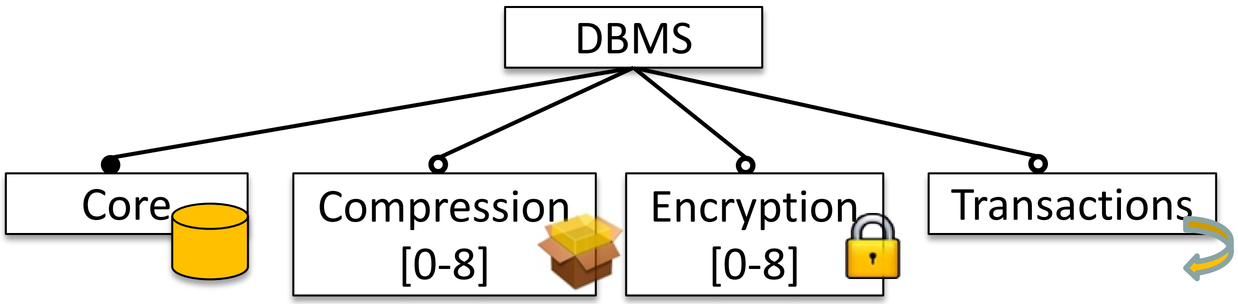


- Use Domain Knowledge
  - Pros:
    - Knowledge about asymptotic behavior
    - No measurement overhead
  - Cons:
    - Expensive, hard to incorperate
    - Sometimes misleading

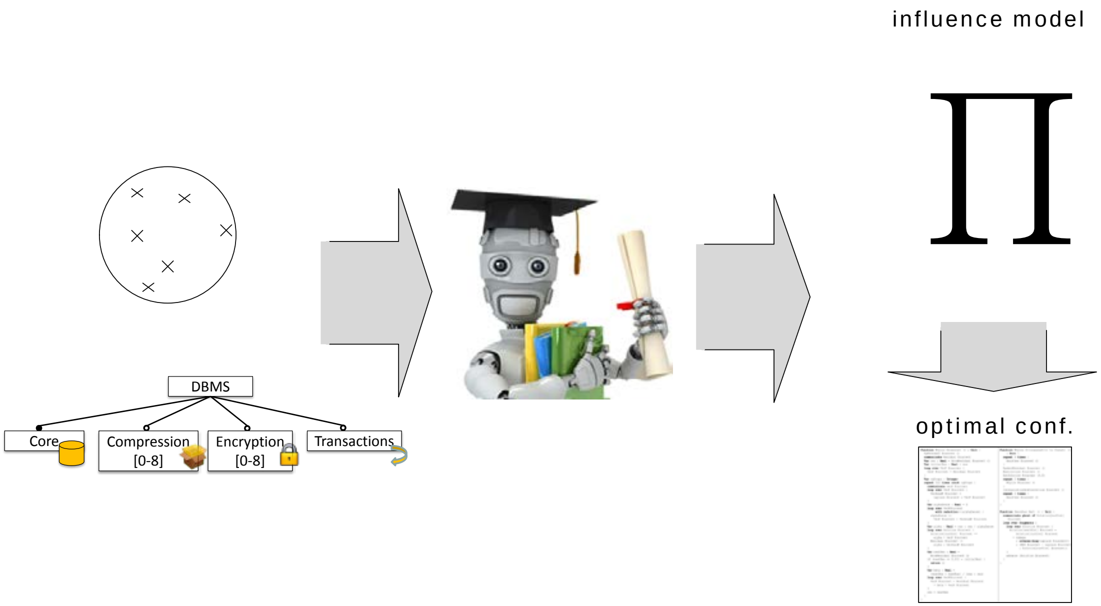


Variability Model

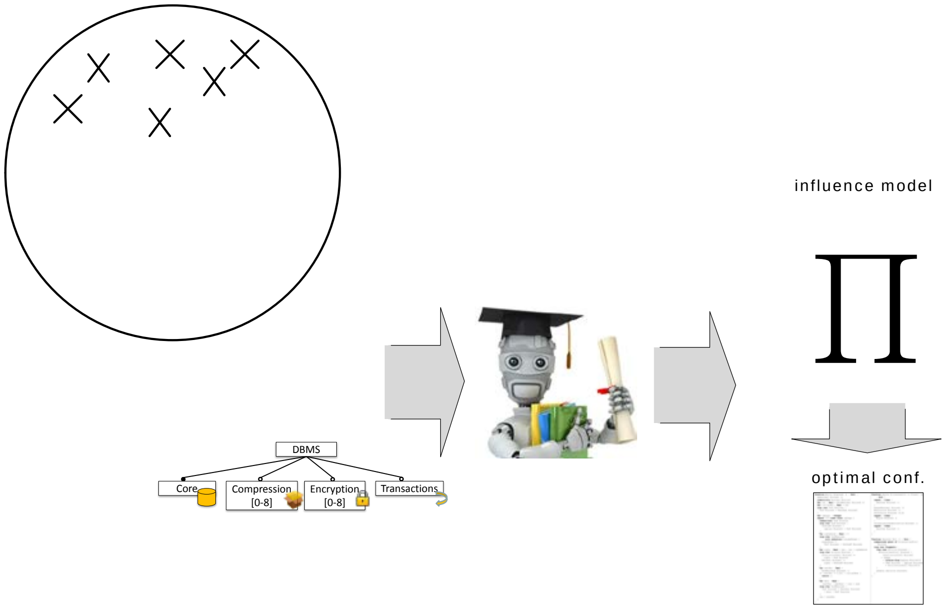




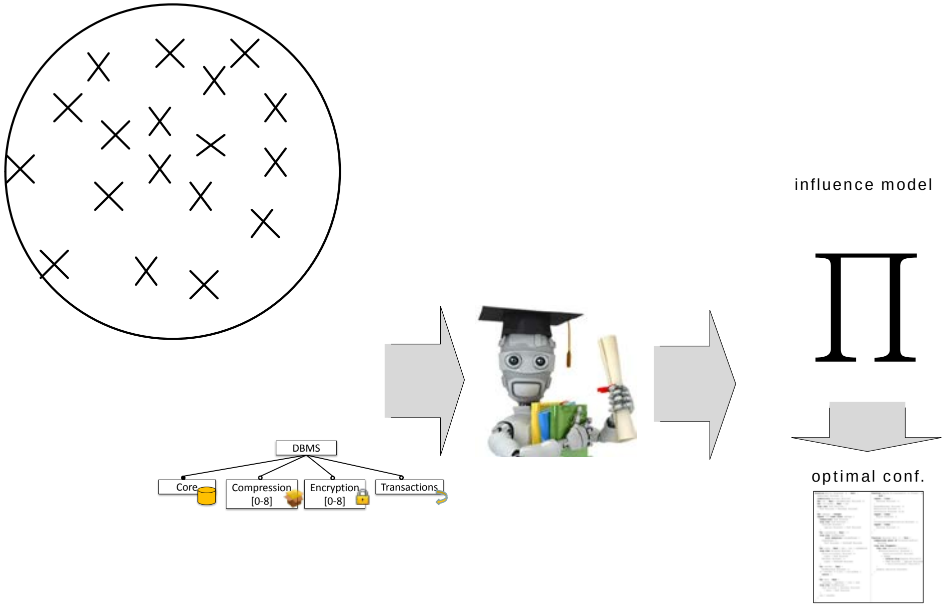
Sampling


influence model
Π
optimal conf.
DBMS
Core
Compression
[0-8]
Encryption
[0-8]
Transactions

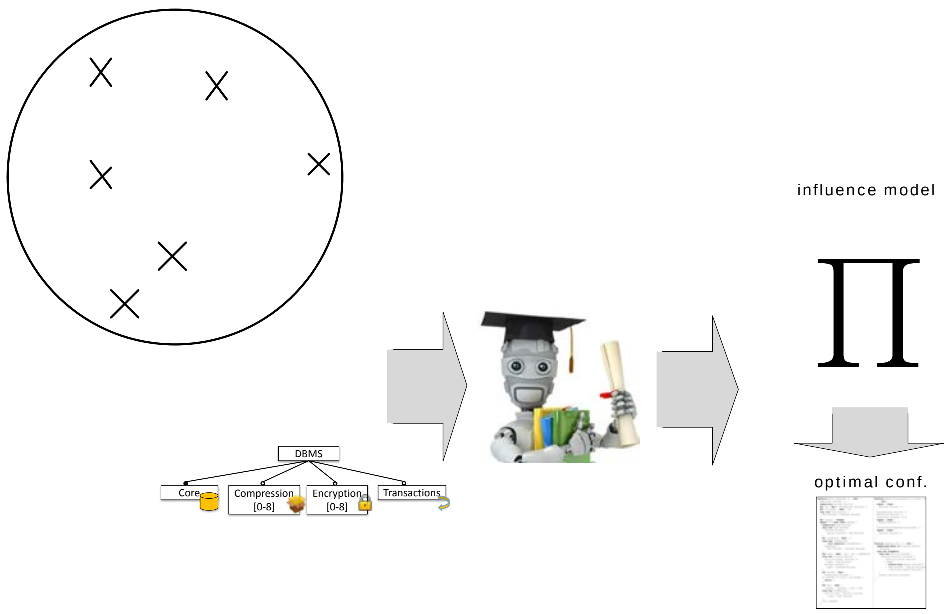
Sampling

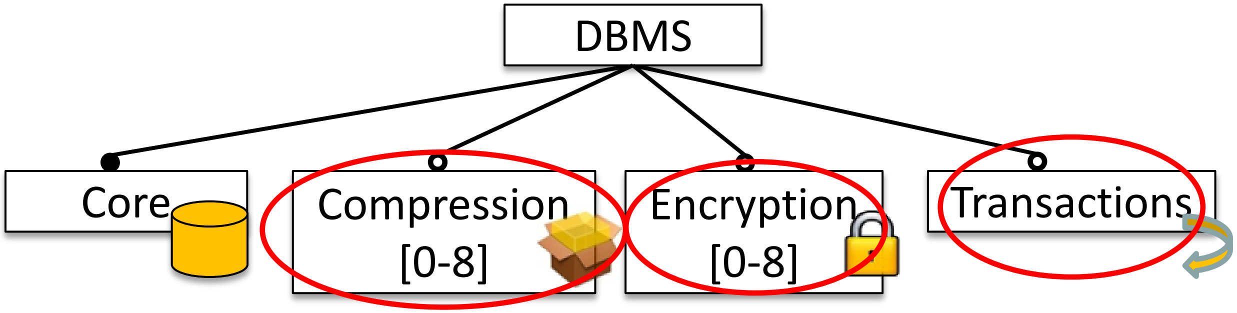

influence model
Π
optimal conf.
DBMS
Core
Compression
[0-8]
Encryption
[0-8]
Transactions

Sampling

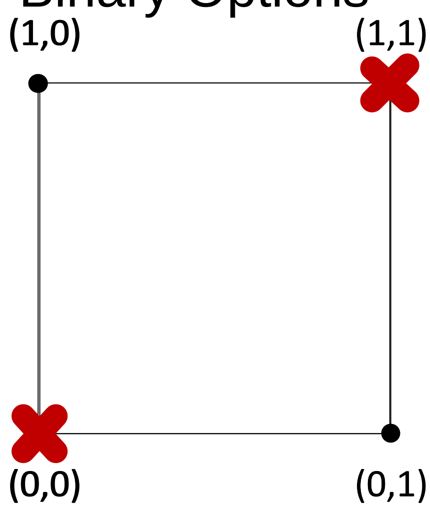

influence model
Π
optimal conf.
DBMS
Core
Compression
[0-8]
Encryption
[0-8]
Transactions

Sampling

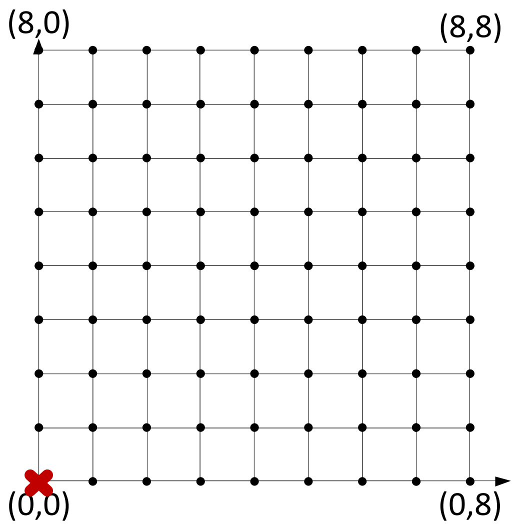

influence model
Π
optimal conf.
DBMS
Core
Compression
[0-8]
Encryption
[0-8]
Transactions



Binary Options

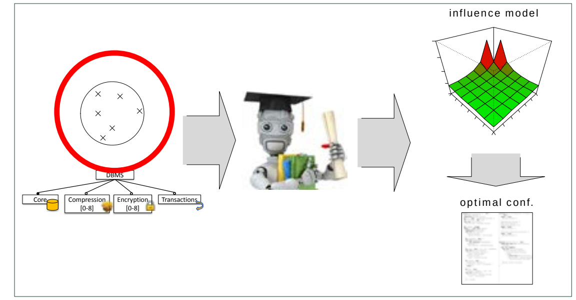


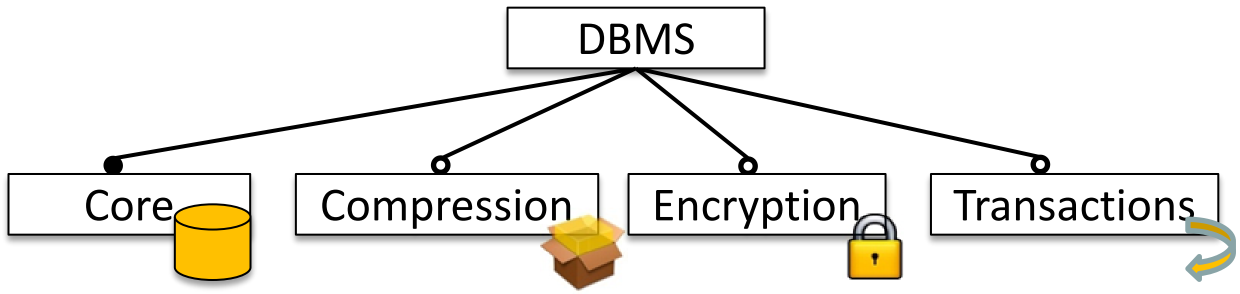
Numeric Options



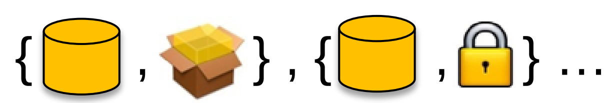
Structured sampling approaches for the different kinds of options

- Random?
  - Unlikely to select a valid configuration
  - Only locally clustered solutions using SAT

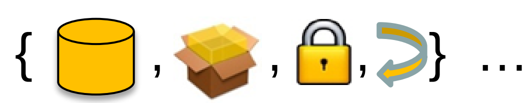




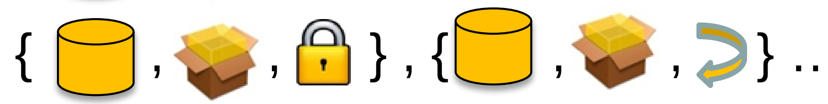
- Heuristics
  - Option-Wise (OW)

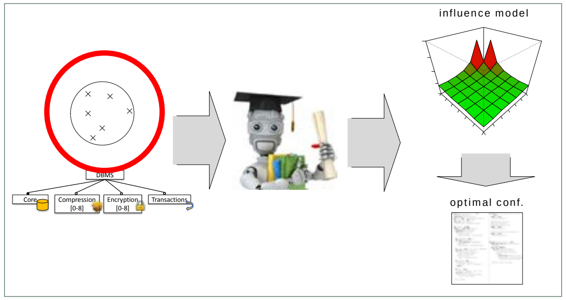


  - Negative Option-Wise (nOW)

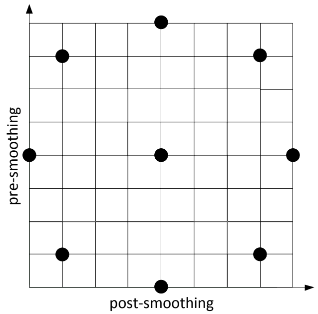


  - Pair-Wise (PW)

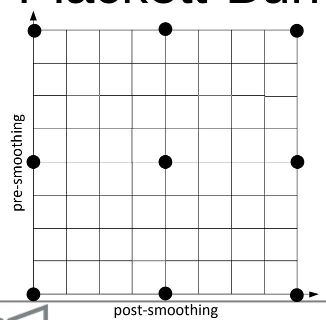




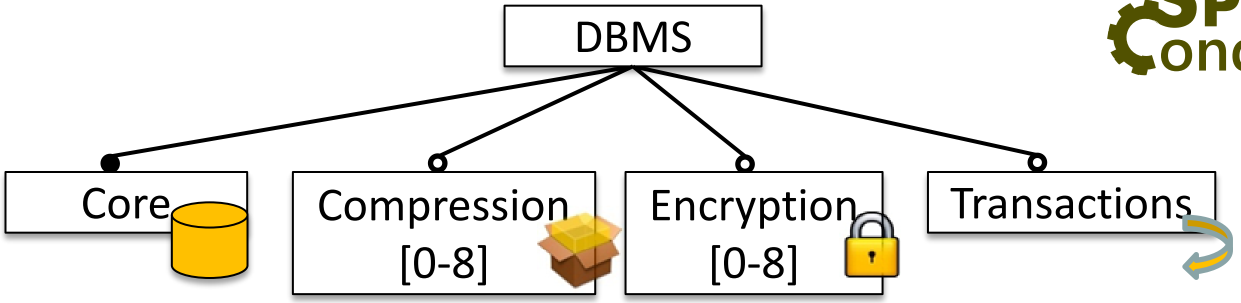
- Response surface models
  - Identify the influence of independent variables on a parameter
  - Scale to multiple numeric options
- Central Composite Design (CCD)



- Plackett-Burman Design (PBD)



(Siegmund et al., FSE'15)

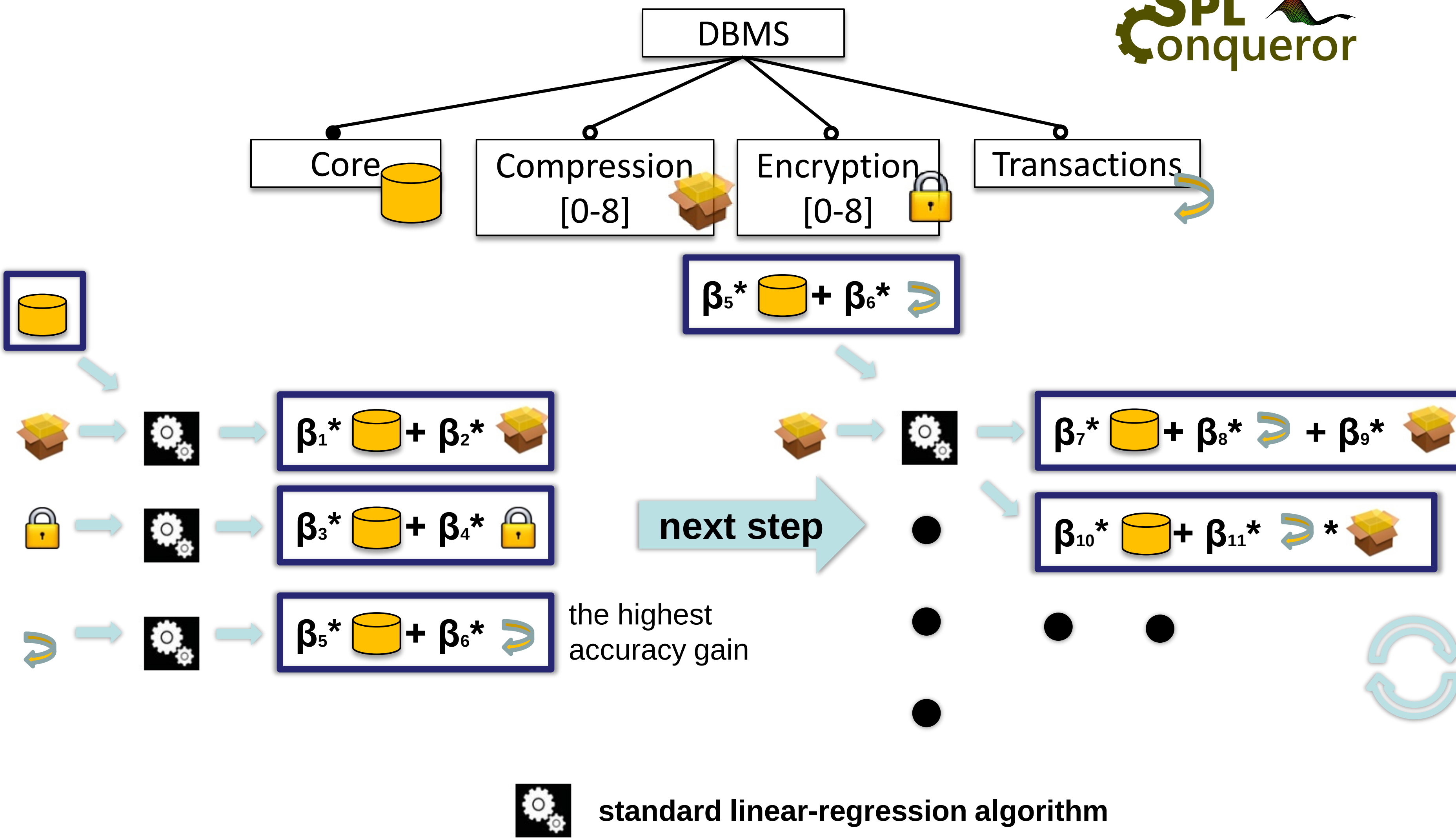


Performance-Influence Model (Π)

$$\Pi : \mathbb{C} \rightarrow \mathbb{R}$$

Performance-Influence Model

(Siegmund et al., FSE’15)



Experiments: Subject Systems

DUNE MGS

clasp

264

ORACLE
BERKELEY DB

APACHE
HTTP SERVER

AJStats

SAC
www.sac-home.org

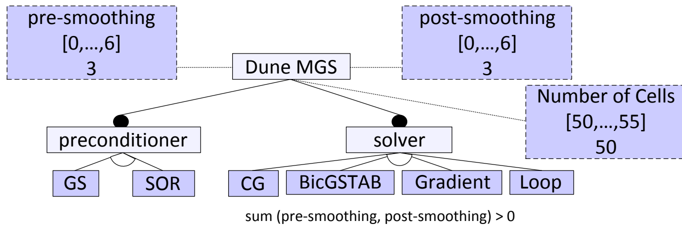
Java

lrzip

HIPAcc

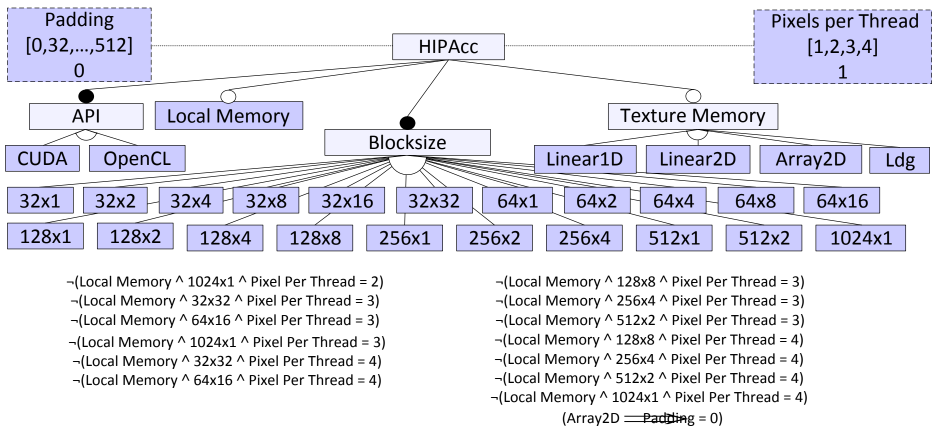
HSMGP

- DUNE MGS



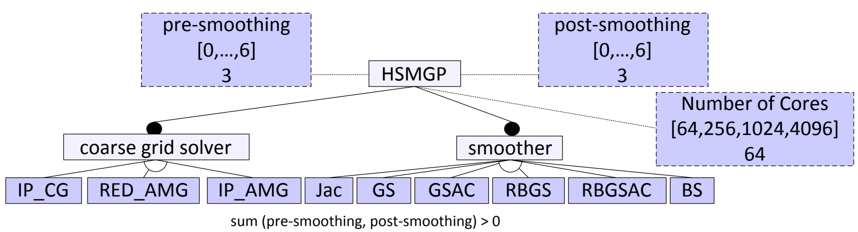
- 2 304 configurations
- Intel i5-4570 Quad Code and 32 GB RAM

- HIPAcc



- 13 485 configurations
- nVidia Tesla K20 with 5GB RAM and 2495 cores

- HSMGP



- 3 456 configurations
- JuQueen at Jülich

	Option-Wise $\bar{e}/\|C\|$	Pair-Wise $\bar{e}/\|C\|$	Negative Option-Wise $\bar{e}/\|C\|$
Dune MGS			
PBD(9,3)	14.1%/45	14.9%/72	15.8%/45
PBD(49,7)	11.4%/245	11.9%/392	11.6%/245
CCD	11.1%/75	11.9%/120	10.8%/75
HIPAcc			
PBD(9,3)	14.7%/240	13.8%/1221	49.3%/85
PBD(49,7)	13.9%/736	11.1%/3645	41.4%/161
CCD	14.2%/242	10.5%/1247	48.2%/102
HSMGP			
PBD(9,3)	2%/72	2.4%/162	3.3%/72
PBD(49,7)	2.1%/392	1.5%/882	2.4%/392
CCD	3.2%/120	2.7%/270	3.7%/120

$\bar{e}$: average prediction error, $|C|$: number of measurements
PBD: Plackett-Burman Design, CCD: Central Composite Design

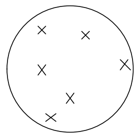
- *Option-Wise* is the best trade-off between prediction accuracy and measurement overhead
- *Option-Wise* combined with *PBD(49,7)* has best accurcy (~avg. error of 9.1%) compared to measurement overhead

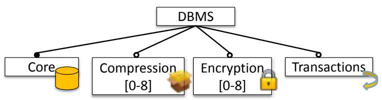
What about Domain Knowledge?

Tailor numeric option sampling to known shape of function

Tailor binary option sampling to known interactions

Tailor numeric option sampling to known absence of interactions





Learn specific functions (do not probe for any function)

Learn separate models for independent configuration spaces

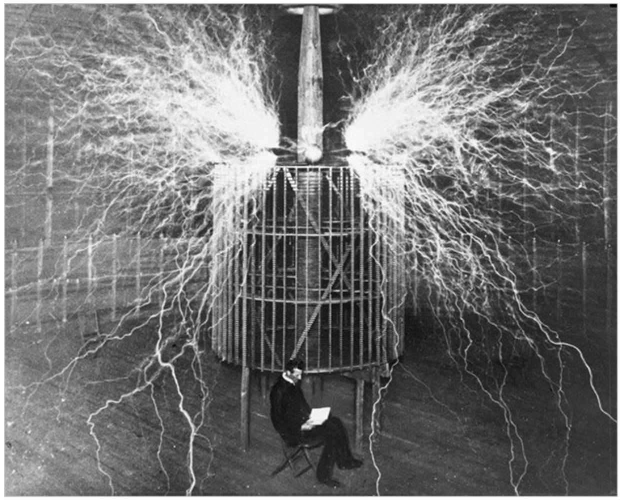
influence model

$$\Pi_1 + \Pi_2 + \Pi_3$$

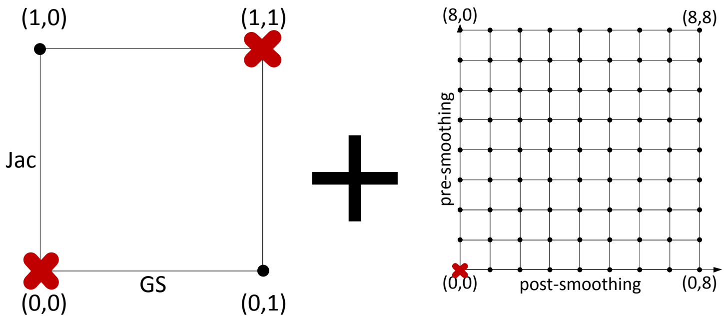
optimal conf.

Outlook

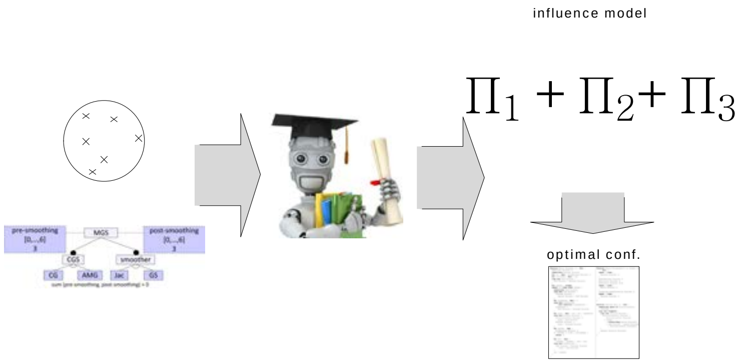
Energy efficiency



Combined sampling of binary and numeric options



Domain-knowledge integration and validation



- Norbert Siegmund, Alexander Grebhahn, Sven Apel, Christian Kästner. **Performance-Influence Models for Highly Configurable Systems**. In *ESEC/FSE*. ACM Press, August 2015; to appear.
- Alexander Grebhahn, Sebastian Kuckuk, Christian Schmitt, Harald Köstler, Norbert Siegmund, Sven Apel, Frank Hannig, and Jürgen Teich. Experiments on Optimizing the Performance of Stencil Codes with SPL Conqueror. *Parallel Processing Letters*, 24(3):Article 1441001, September 2014.
- Alexander Grebhahn, Norbert Siegmund, Sven Apel, Sebastian Kuckuk, Christian Schmitt, and Harald Köstler. Optimizing Performance of Stencil Code with SPL Conqueror. In *HiStencils*, January 2014.

SPL Conqueror