

Autotuning OpenCL Workgroups Sizes

Collaborative autotuning for Stencils outperforms human experts

Chris Cummins

c.cummins@ed.ac.uk

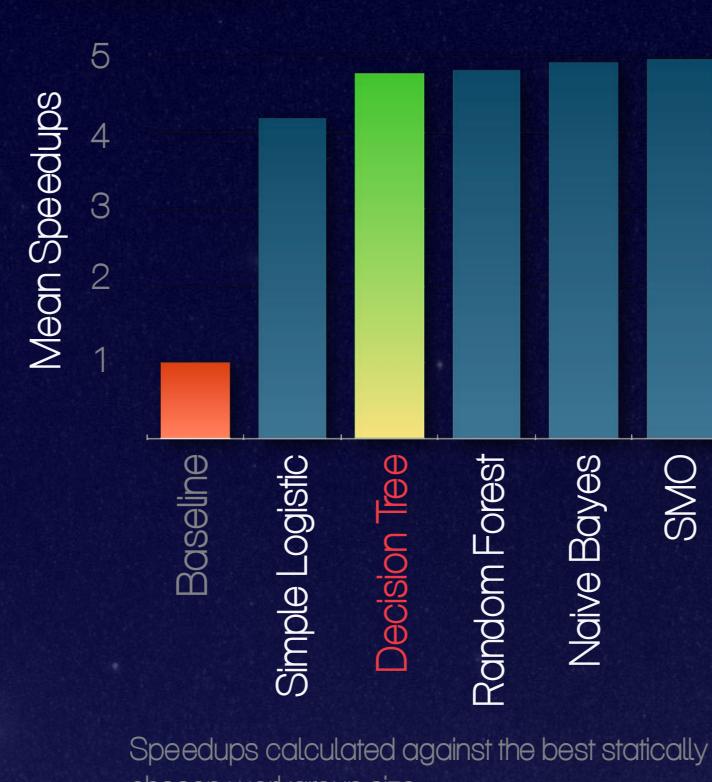
Pavlos Petoumenos

Michel Steuwer

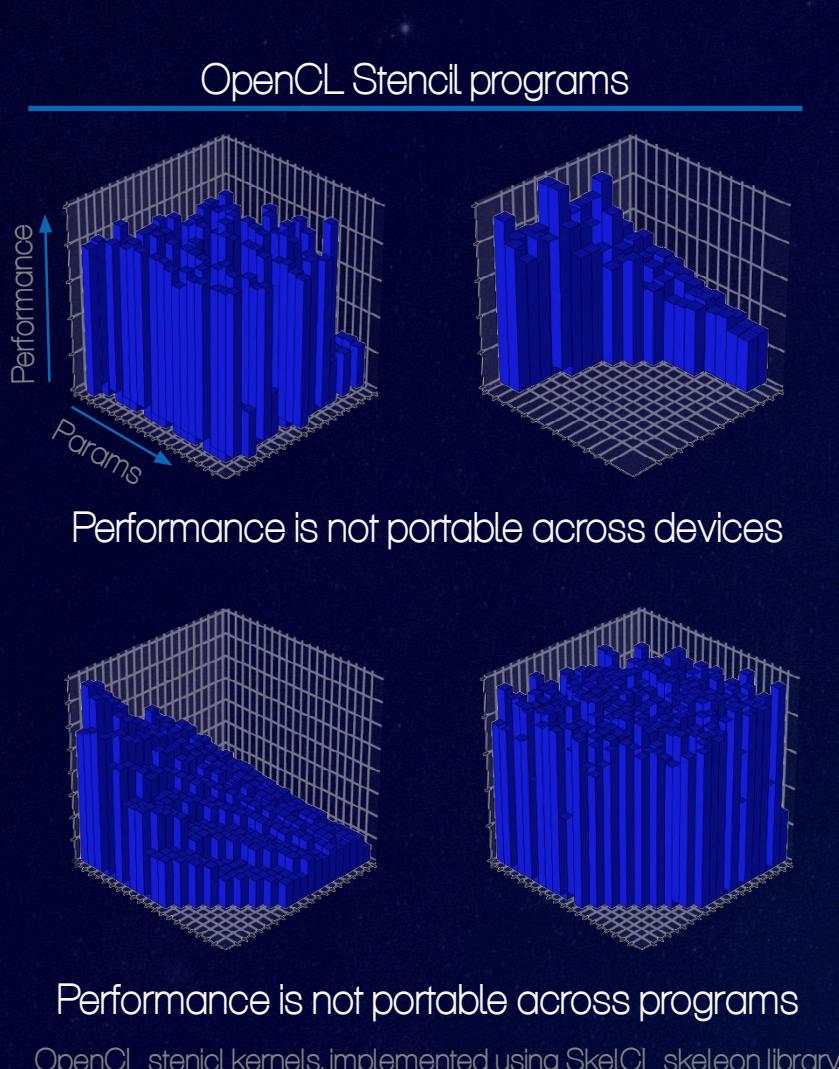
Hugh Leather

5.0X
speedup!

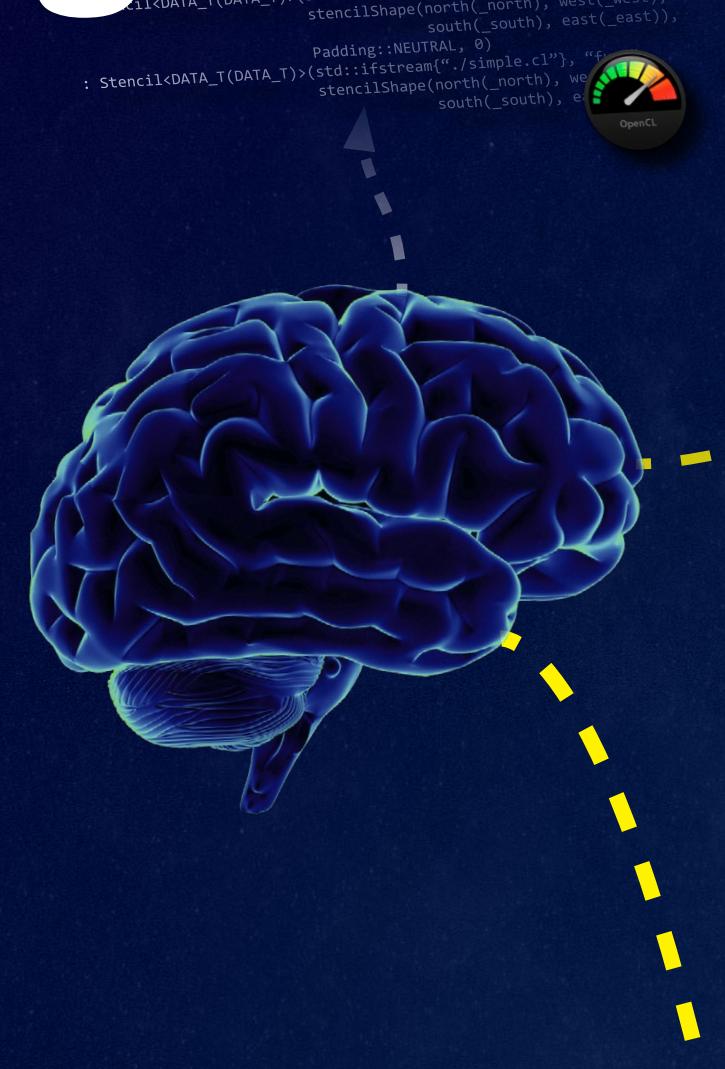
Predicting OpenCL workgroup sizes
of 429 stencil programs, execution
devices, and datasets.



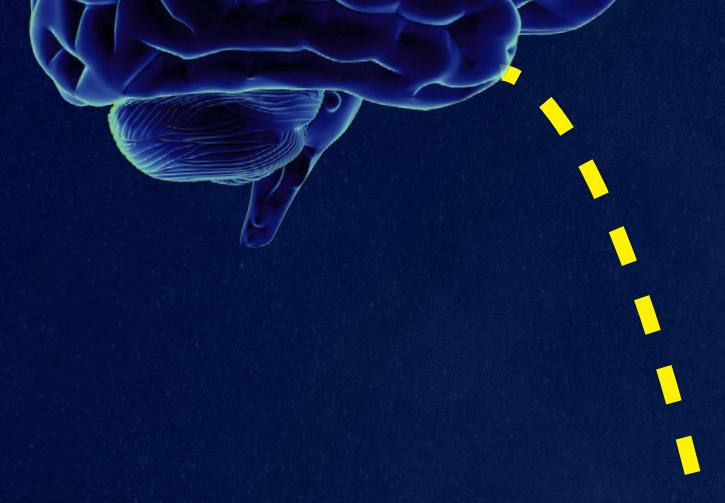
**Hand tuning programs
is expensive and time
consuming**



OpenCL stencil kernels, implemented using SkelCL skeleton library

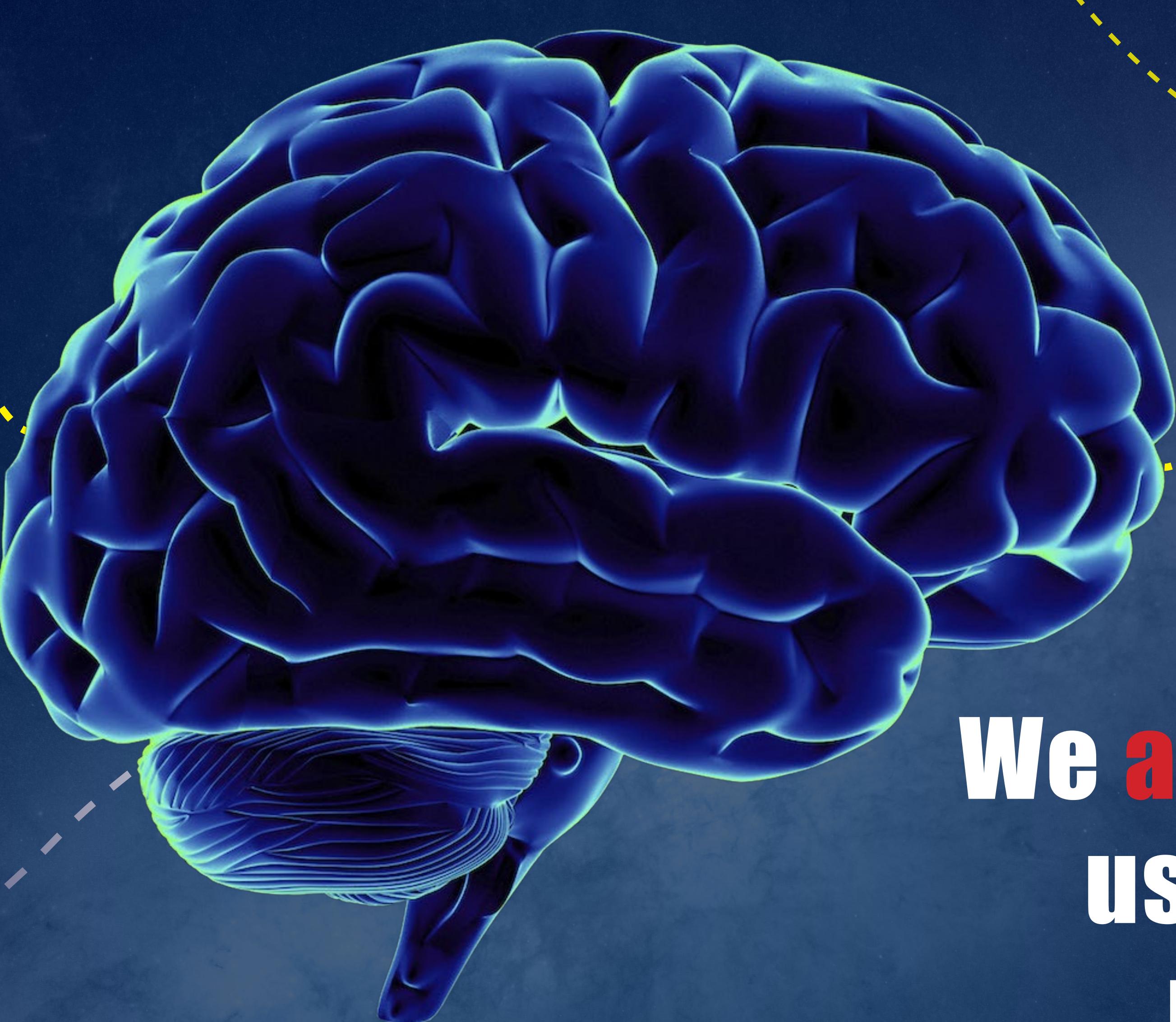
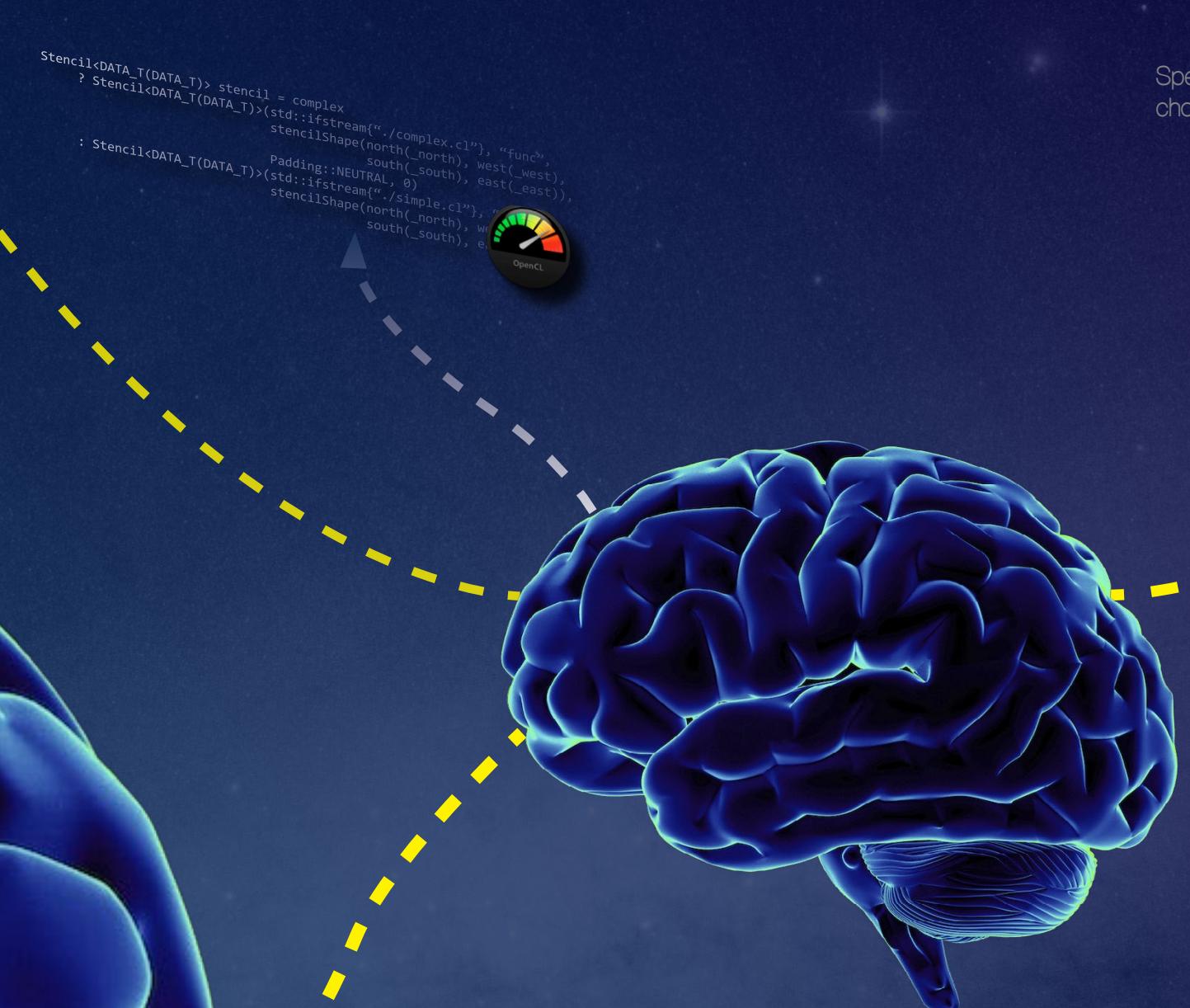


Performance is not portable across devices



Performance is not portable across programs

OpenCL stencil programs



We **automate** this task
using collaborative
machine learning

Introducing OmniTune ...



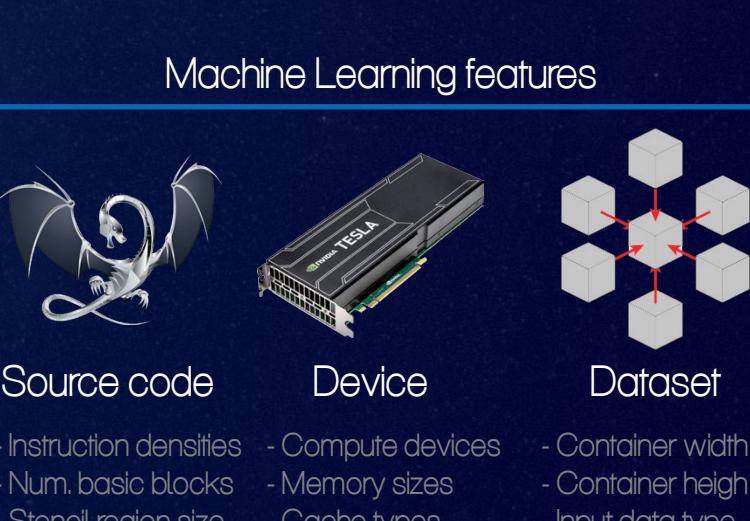
OmniTune generates synthetic benchmark
programs to use for empirical testing



OmniTune collaboratively gathers performance
data by testing different parameter values



OmniTune uses machine learning to predict
parameters for unseen programs at runtime



Machine Learning features

Source code

- Instruction densities

- Num. basic blocks

- Stencil region size

- etc.

Device

- Compute devices

- Memory sizes

- Cache types

- etc.

Dataset

- Container with

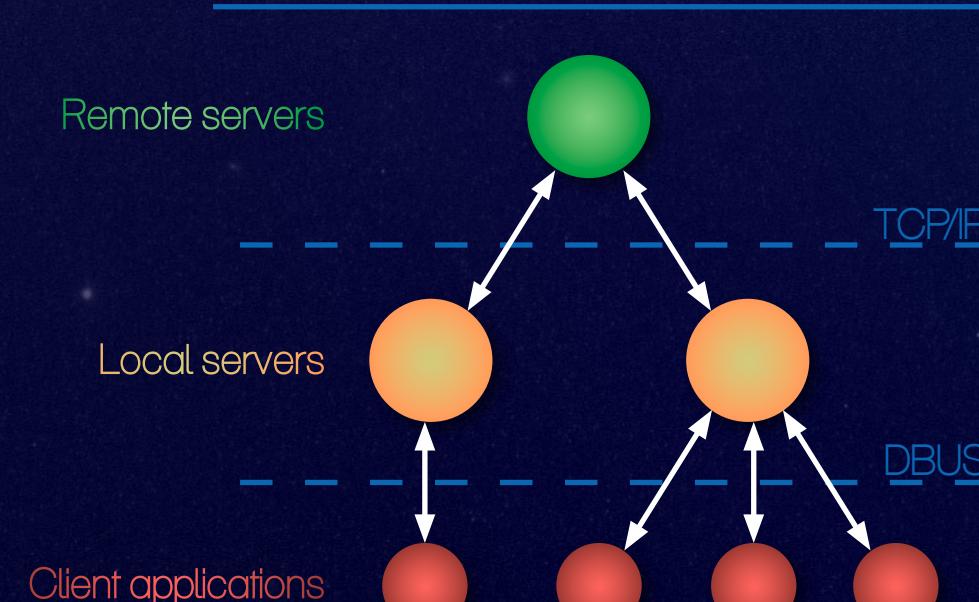
- Input data type

Publications

Chris Cummins, Pavlos Petoumenos, Michel Steuwer and Hugh Leather. "Autotuning OpenCL Workgroup Size for Stencil Patterns". ADAPT 2016.

Chris Cummins, Pavlos Petoumenos, Michel Steuwer and Hugh Leather. "Towards Collaborative Performance Tuning of Algorithmic Skeletons". HPGPU 2016.

Read more ...
<http://chriscummins.cc>



OmniTune architecture



THE UNIVERSITY OF EDINBURGH
informatics

EPSRC Centre for Doctoral Training in
Pervasive Parallelism

EPSRC
Engineering and Physical Sciences
Research Council

Image credits: 'Brain' Department of Molecular & Cell Biology, Berkeley, M82 NASA.