UNIVERSITY of York

This is a repository copy of *Hex-Tile: Evaluation Prototype of a Novel Tile-based Processor Array:Preliminary Report*.

White Rose Research Online URL for this paper:
https://eprints.whiterose.ac.uk/190803/

Version: Published Version

Other:
Crispin-Bailey, Christopher orcid.org/0000-0003-0613-9698, Austin, Jim orcid.org/0000-0001-5762-8614, Thuphairo, Pakon et al. (1 more author) (2022) Hex-Tile: Evaluation Prototype of a Novel Tile-based Processor Array:Preliminary Report. UNSPECIFIED.

Reuse
Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown
If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

White Rose university consortium
Universities of Leeds, Sheffield & York

eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

Investigating Novel 3D Modular Schemes for Large Array Topologies: Power Modeling and Prototype Feasibility.

Pakon Thuphairo, Christopher Bailey, Anthony Moulds, Jim Austin

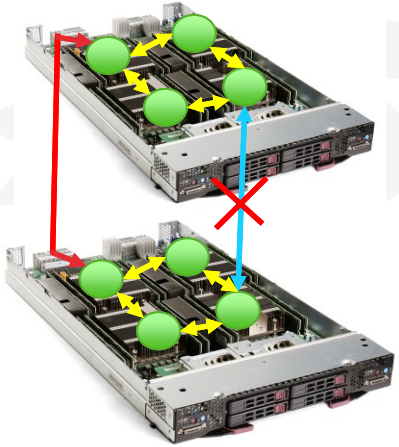
Department of Computer Science
University of York,
York, United Kingdom

Background and Motivation

Alternatives to Rack-Mount

Background - Structural comparison

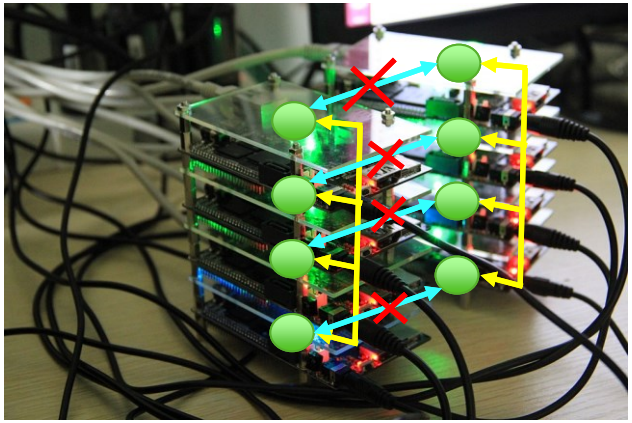
- Wiring effort (Power + data communication)
- Lengths of vertical and horizontal data channels
- Empty volume for cooling



Adapted from [1]

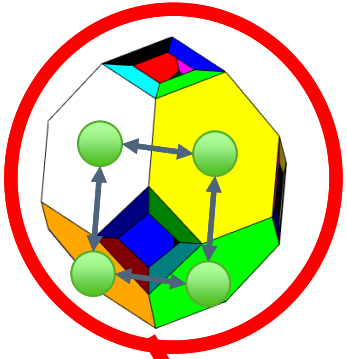
Blade server

(Vertical distance is for illustration purpose)

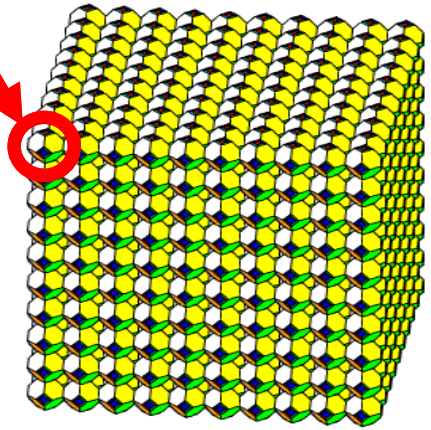


Adapted from [2]

Small single-board



External
DC Power supply



Our 'ball computer' packaging

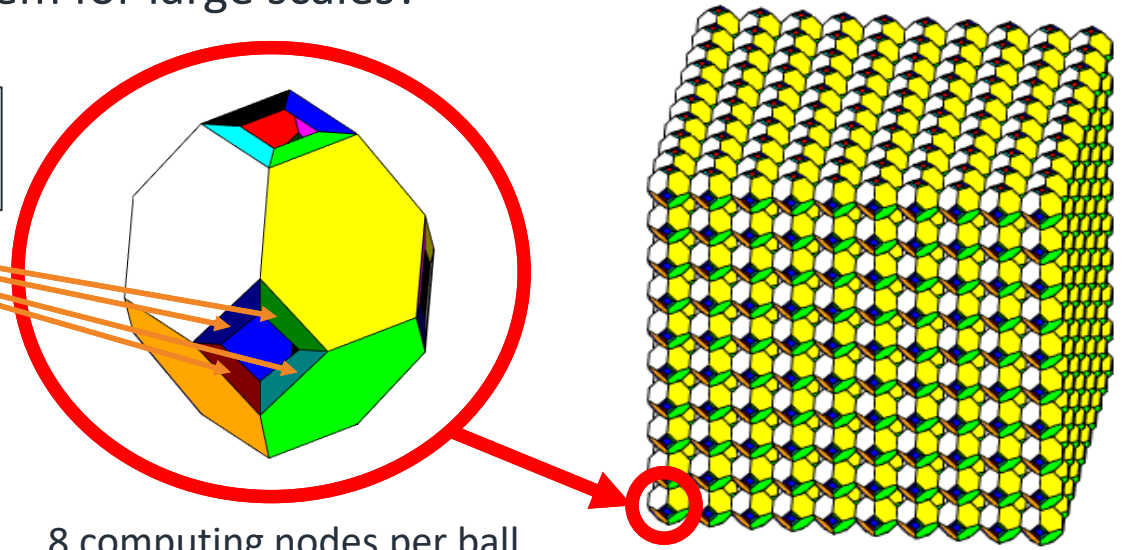
[1] https://upload.wikimedia.org/wikipedia/commons/thumb/d/d0/Supermicro_SBI-7228R-T2X_blade_server.jpg/1024px-Supermicro_SBI-7228R-T2X_blade_server.jpg

[2] https://upload.wikimedia.org/wikipedia/commons/thumb/2/27/Cubieboard_HADOOP_cluster.JPG/1024px-Cubieboard_HADOOP_cluster.JPG

Motivation – Power grid simulation

- This power grid system does not exist in conventional rack/cabinet systems.
  - Direct external power sources supplied to each blade/rack server
- In this work, in contrast, how does it impact on the scalability in the concept of hexagonal-tile system for large scales?

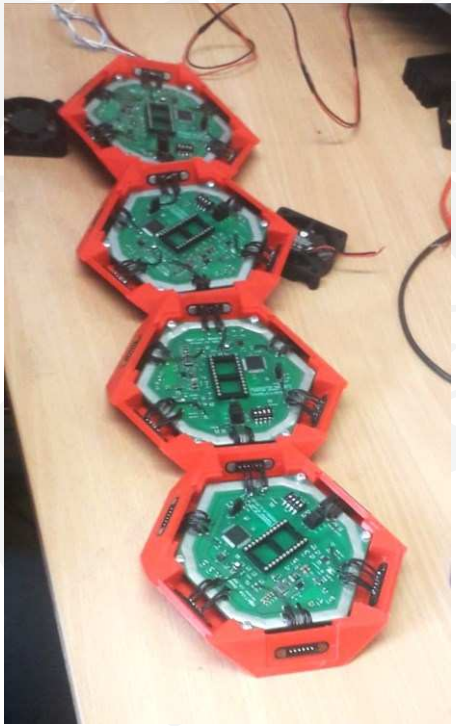
External power source connections (for any external trapezoidal faces)



8 computing nodes per ball

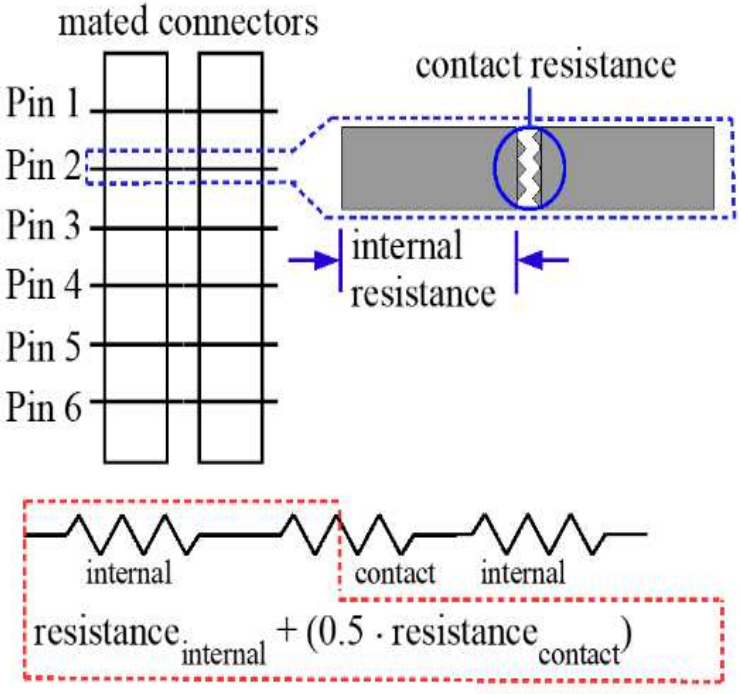
Introduction – from tile to ball

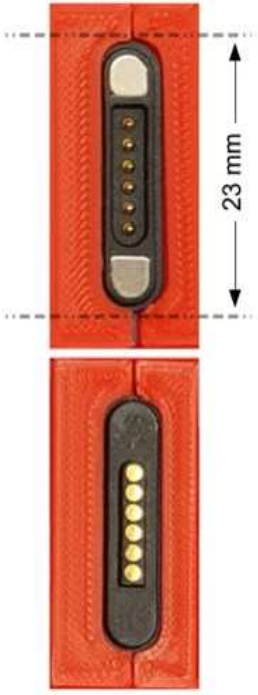
1D/2D/3D compossible configurations (prototype)



Simulation and Prototype Details

Model – Connector pin resistance

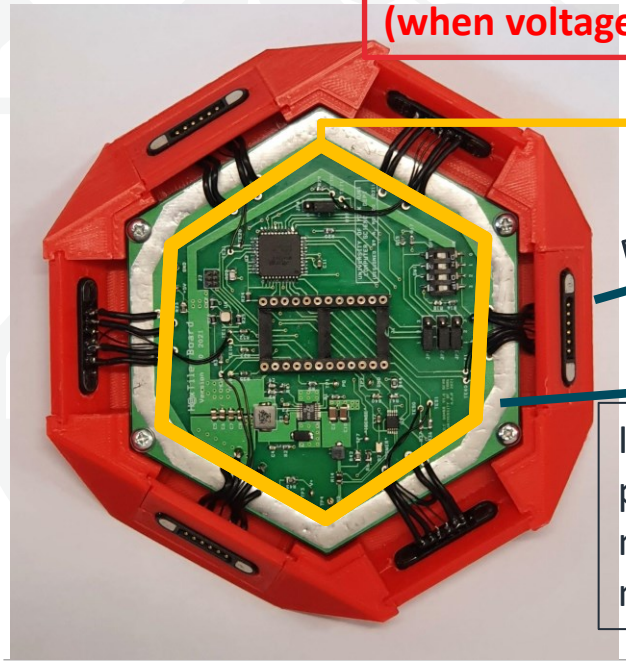




- ‘Off-the-shelf’ connectors in the current prototype
- Variants of (custom-made) more suitable connectors can be used for different power and data communication requirements.

Model – Simplified board-resistance model

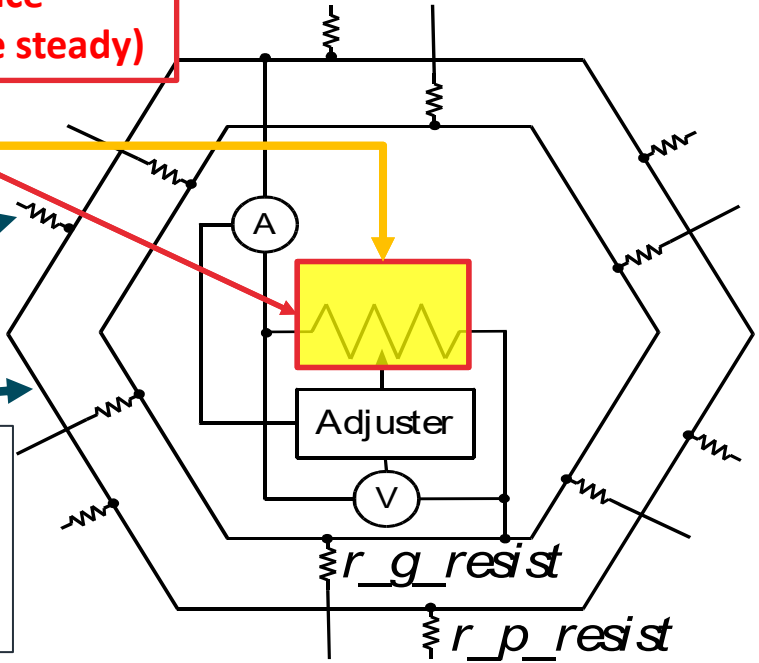
- Switching regulator models take long simulation times.
- A Simplified model has been created for our scalability simulations.



Entire board equivalent-resistance (when voltages and currents are steady)

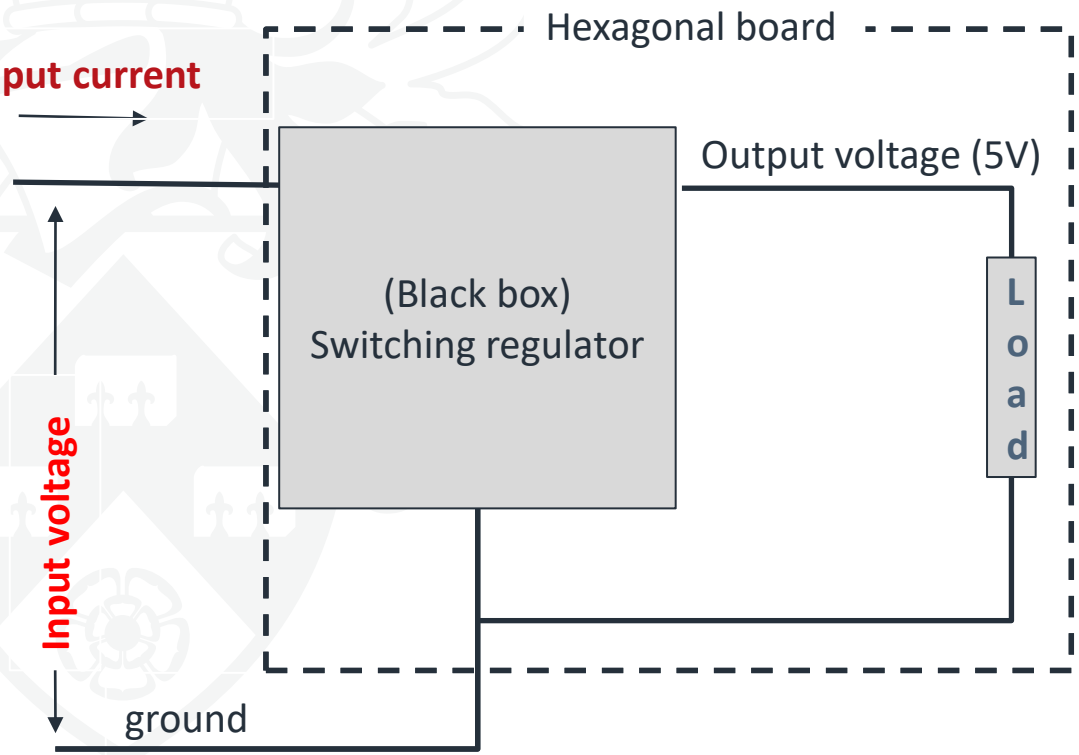
Pin(s) resistance

Intra-board power/gnd traces negligible in this model



Model – Simplified board-resistance model

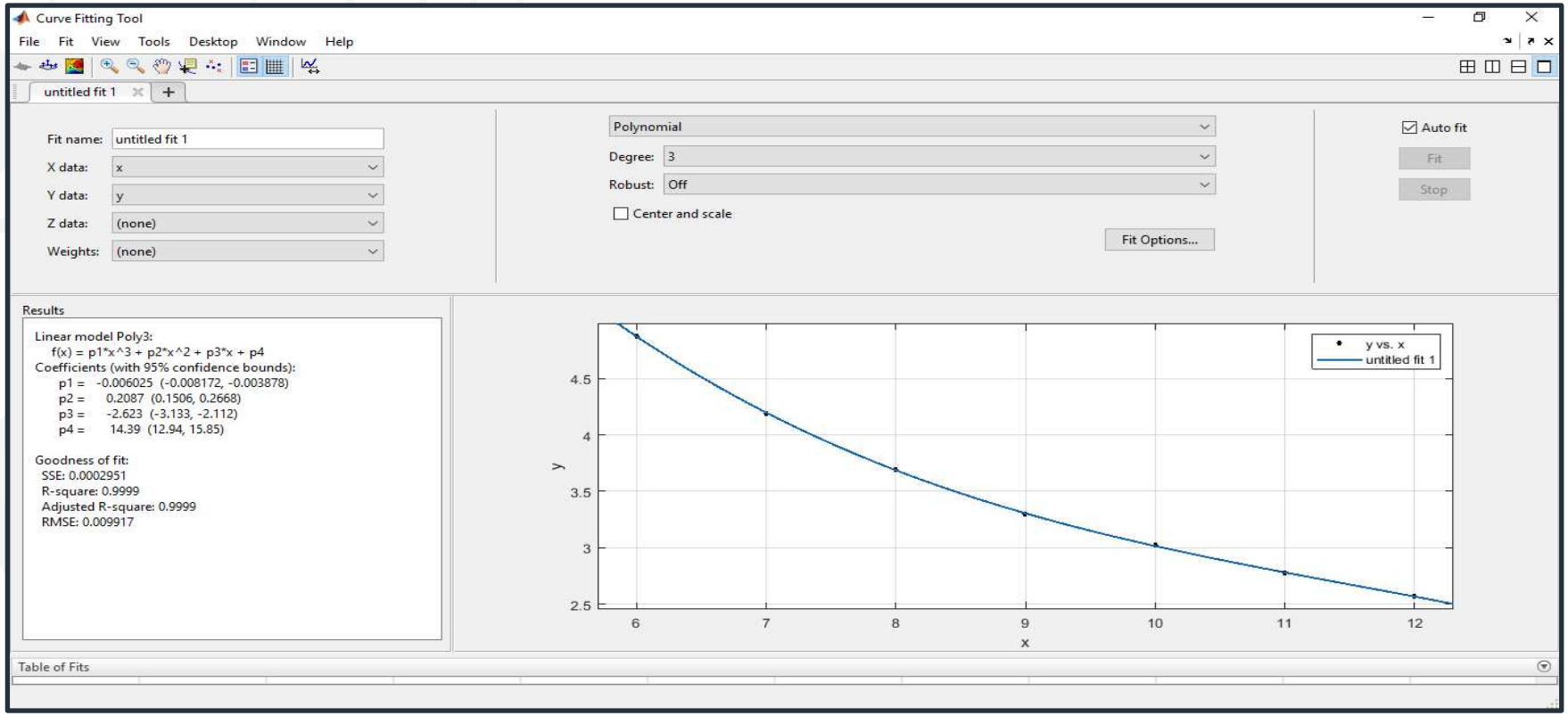
Curve fitting for the regulator and load



Input Voltage (V)	Load Resistance (Ω)	Input Current (A)
12	1	2.5706
11	1	2.775
10	1	3.0244
9	1	3.2982
8	1	3.695
7	1	4.1902
6	1	4.8733

Model – Simplified board-resistance model

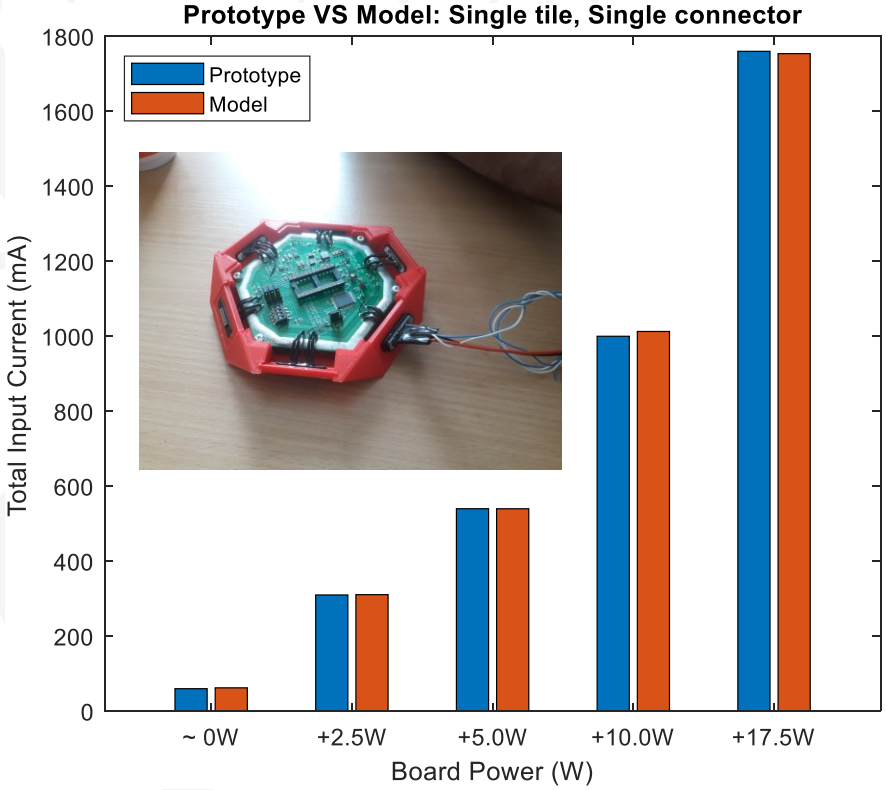
Curve fitting for the regulator and load (for a constant load-resistance)

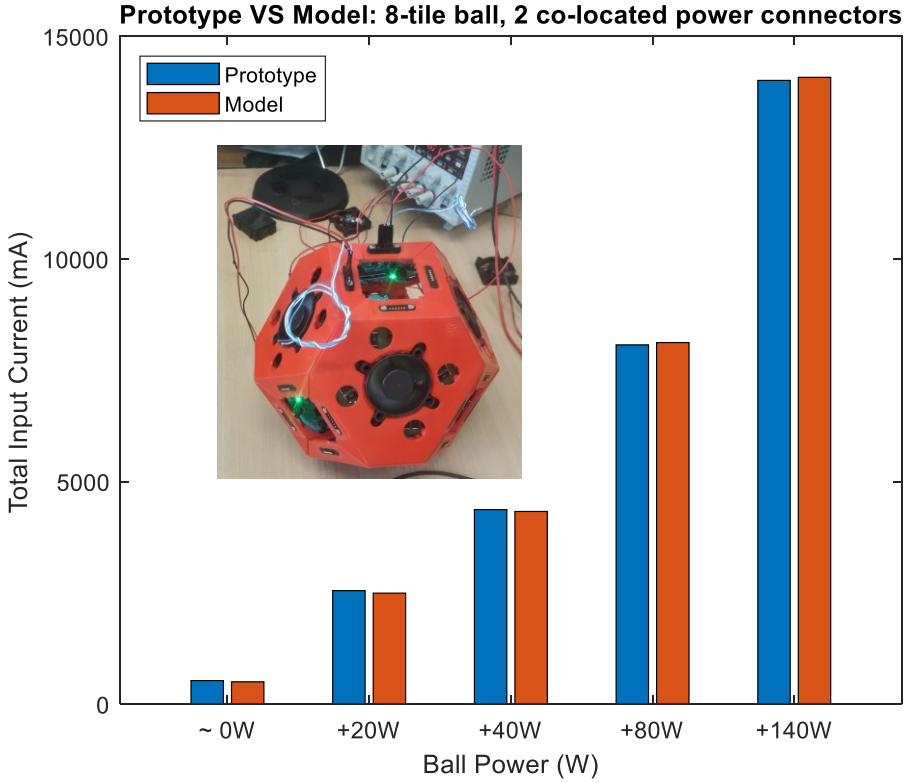


Validation
Simulator vs. Prototype

Model Validation – switching vs prototype

Input-current validation

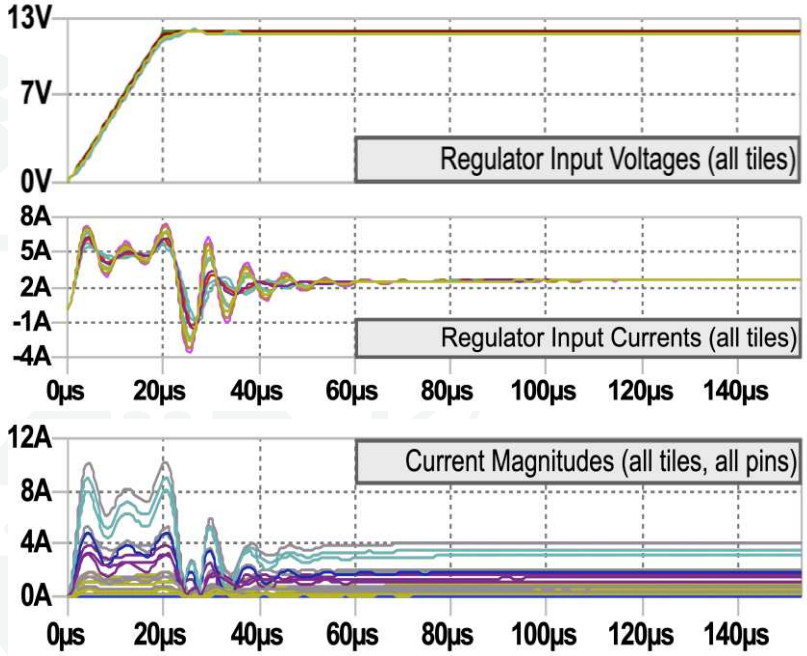




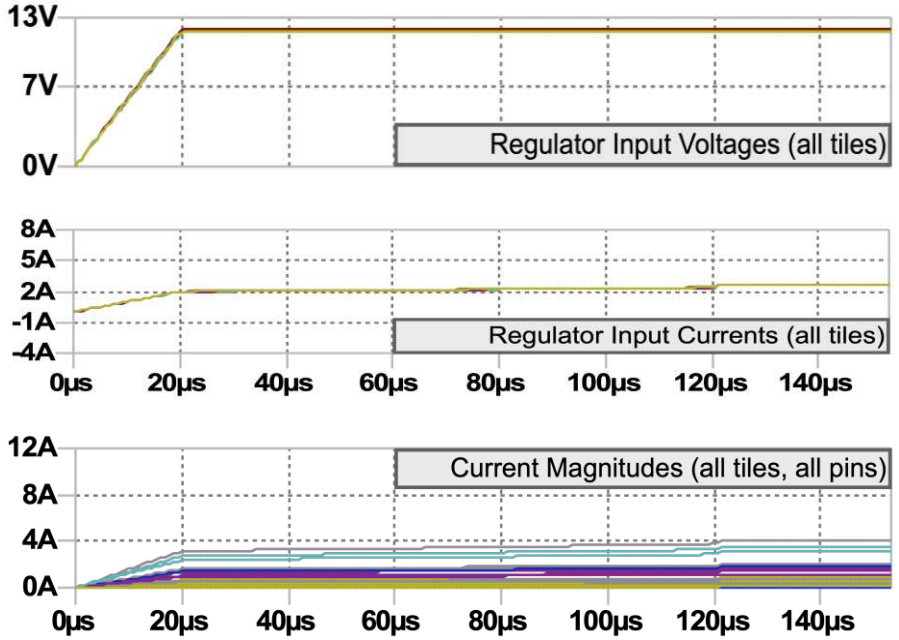
*LT3976 regulators, from Analog Devices, Inc., are used in our prototype.

Model validation

Switching VS Simplified model, 3x3x3-ball



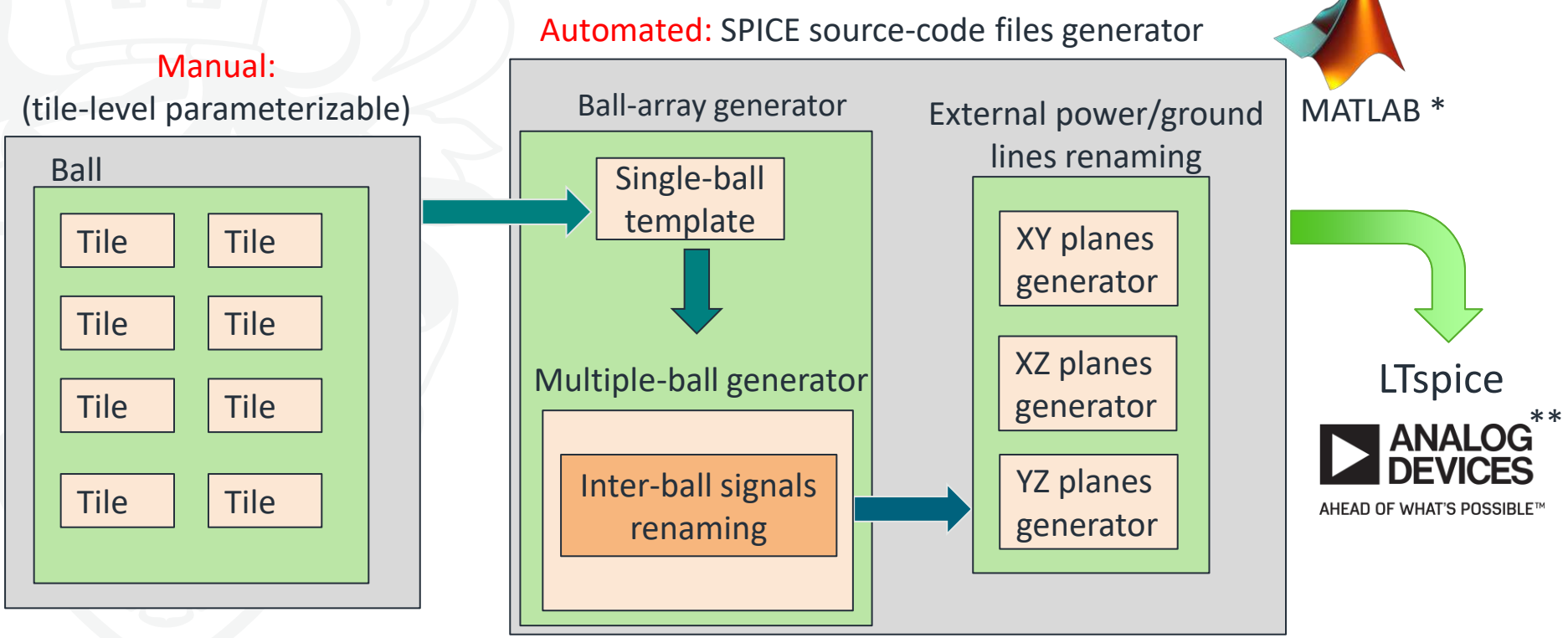
(a) Simulation based upon LT3976 regulator model



(b) Simulation using simplified (faster) model

* External power supplied to all surface power connectors

Simulation framework



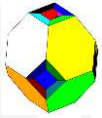
* https://upload.wikimedia.org/wikipedia/commons/thumb/2/21/Matlab_Logo.png/800px-Matlab_Logo.png

** https://upload.wikimedia.org/wikipedia/commons/thumb/8/86/Analog_Devices_Logo.svg/1920px-Analog_Devices_Logo.svg.png

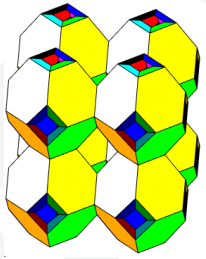
Scalability Evaluations

Scalability Results

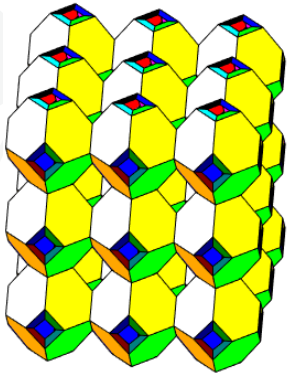
Experimental scenarios



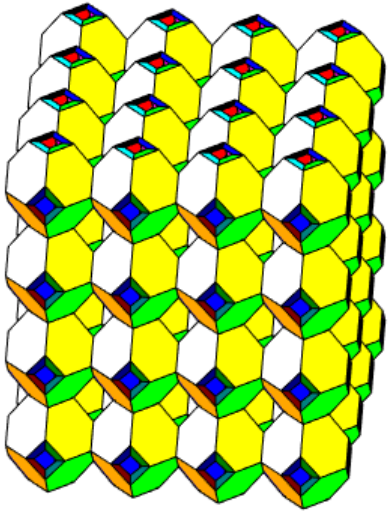
1 ball
8 tiles



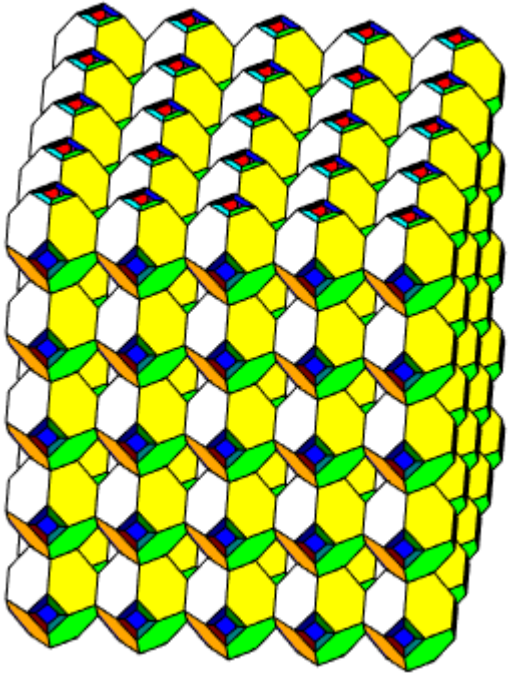
2x2x2 cube
64 tiles



3x3x3 cube
216 tiles



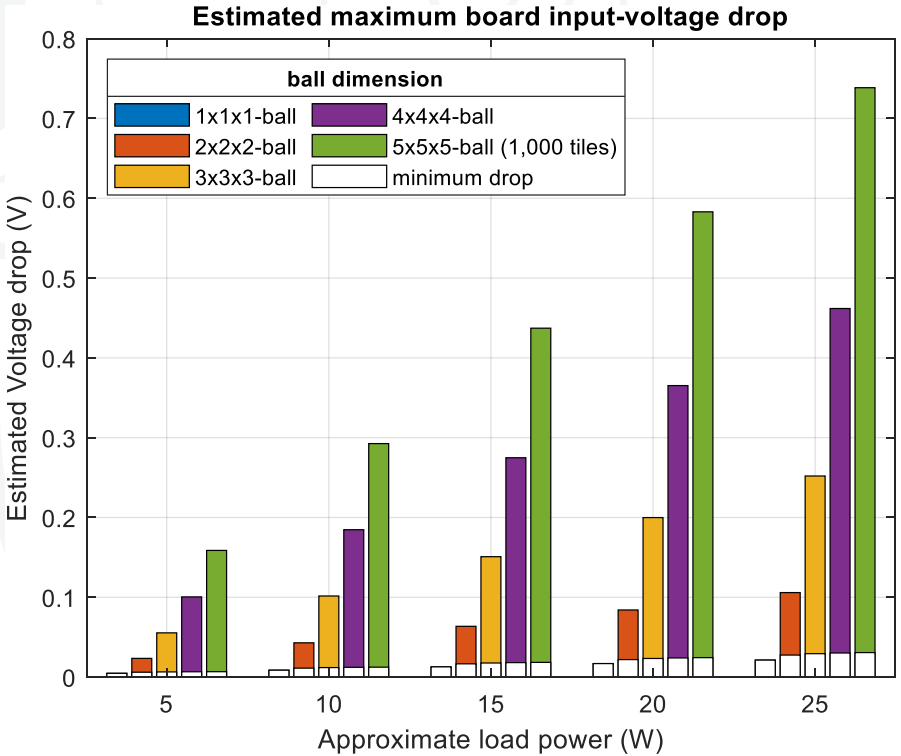
4x4x4 cube
512 tiles

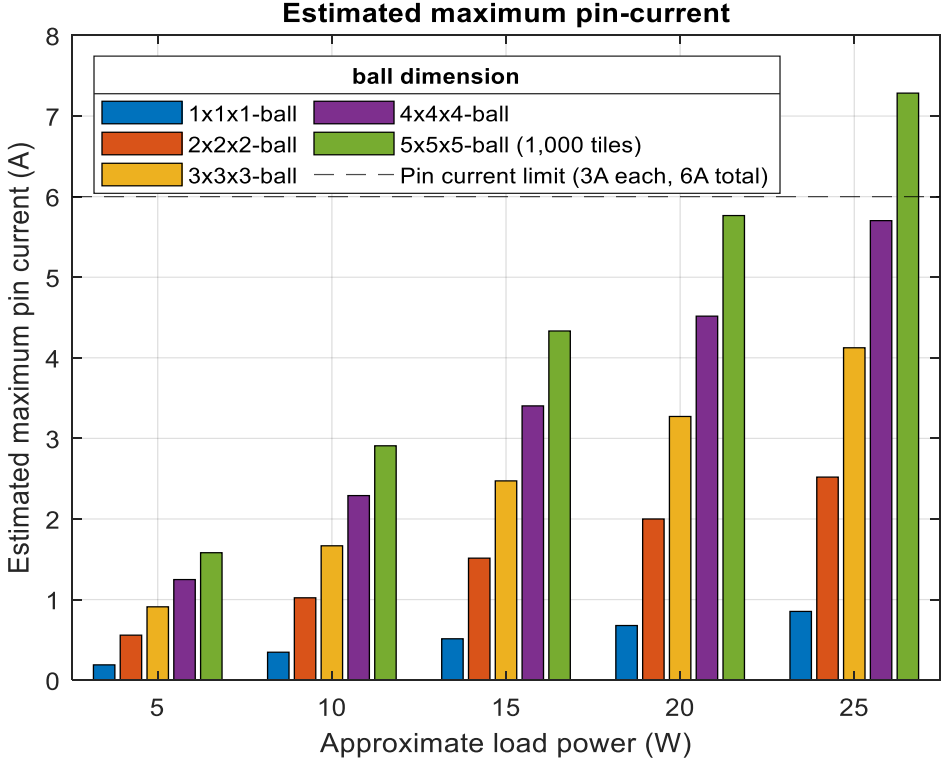


5x5x5 cube
1,000 tiles

Scalability Results

Uniform load-power per tile allocation





* Parameter: 50 mOhms mated pin-pair

Further Optimization

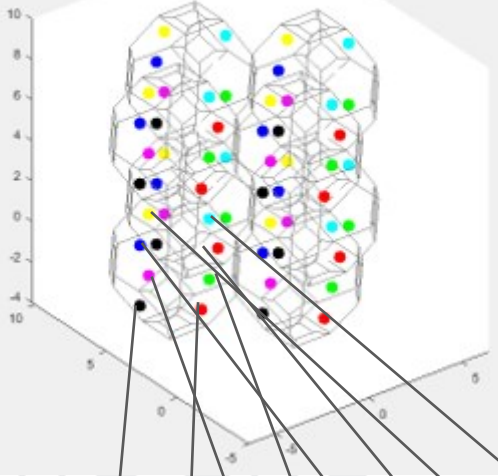
Brute Force ?

Genetic algorithms ?

GA load-power per tile optimization

UNIVERSITY of York

Non-uniform load-power per tile allocation

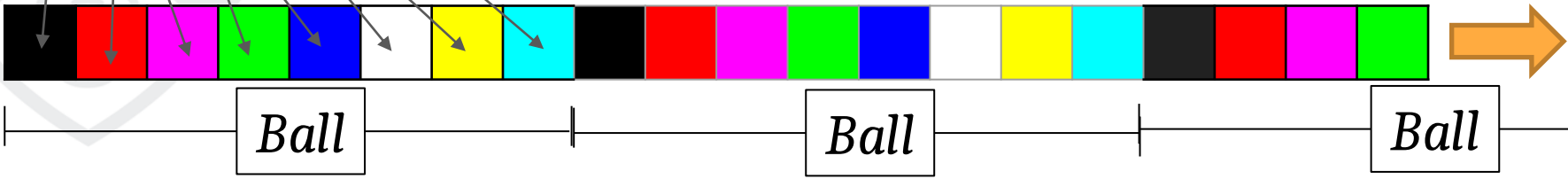


Method 1: Single-tile per gene

- Large search space (for a large system)

- Suitable for arbitrary…
  - non-symmetric external power connection
  - non-symmetric system shapes

Chromosome



GA load-power per tile optimization

Non-uniform load-power per tile allocation

Method 2: Center-distance allocation

Example: 2x2x2-ball system (64 tiles)

Single-tile per gene allocation:

- 5 power steps, 64 nodes

= 5^{64} cases

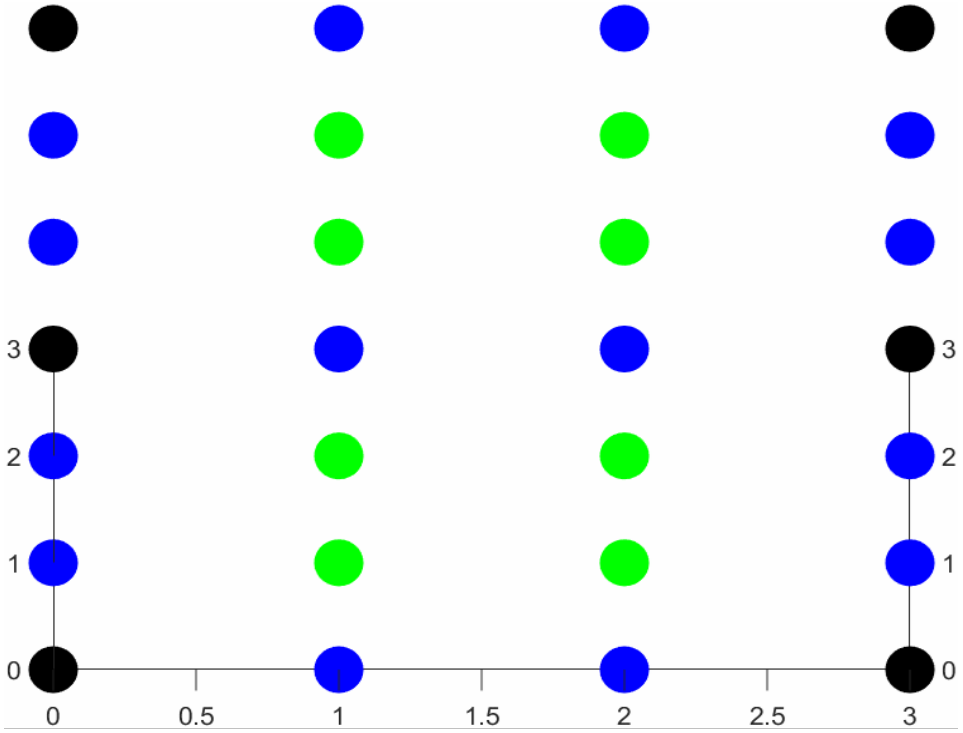
≈ 5.42 x 10^{44} cases!

Center-distance allocation:

- 5 power steps,

- 4 groups of nodes

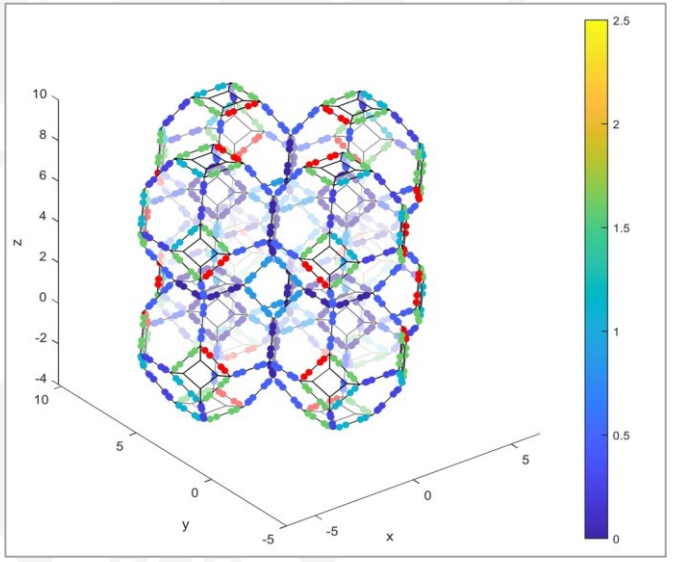
= 5^4 cases = 625 cases (Search space reduced)



GA load-power per tile optimization

Constraints: total 1000W-load per system, 3A connector pin

During optimisation

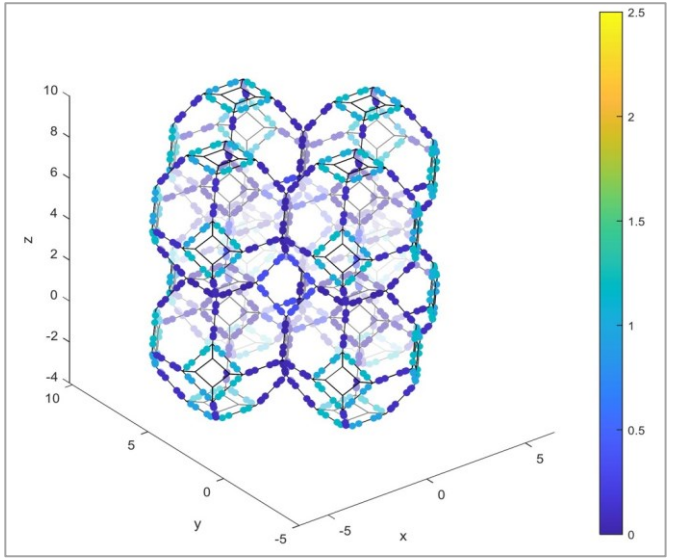


2.5 A

0 A

1000W, with pins overloaded

Stopping criteria reached



2.5 A

0 A

1000W, with pins under current limit

*Red dots = Overloaded pin currents (> 2.5A, for illustration purpose)

Outcomes and Implications

Outcomes

- **What we have done …**
  - **Hardware prototype system**
    - Testing the prototype
  - **Models and simulation framework**
    - Validating accuracy
      - Switching model vs hardware prototype
      - Switching model vs simplified model
    - Scalability projection
    - Power-grid optimization framework
      - Power pattern on a large scale
      - Visualization

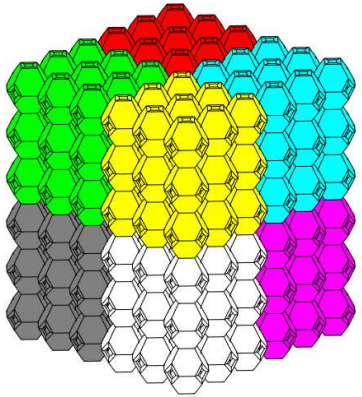
Implications

- **Existing prototype:** Allowing to achieve the system of the order of 1,000 processor tiles, even with a very basic prototype construction.
- With highly optimized fabrication, > 1000 tiles could be achievable.
- Reducing size = Higher density
- Current ball size: Many thousand processors in a server cabinet volume
- Cooling
  - More detailed investigation needed
  - But with the current tool capabilities, the power consumed at pins and tiles can be predictable.
    - Allowing a cooling model to be developed in the future

Possible future works

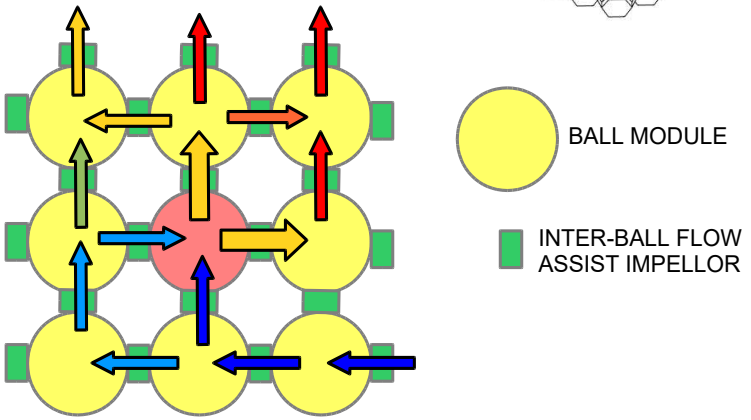
Simulation framework

- Model: Temperature/Manufacture-affected pin-resistance variability
- Opensource SPICE simulator (Ngspice) for simplified models. (In progress)
- Simulations on a computing cluster (In progress)
- Interfacing with an interconnection network simulator (BookSim2, In progress)
- Cooling design and simulation



Hardware developments

- **Reducing hops:** Localized shared physical wires?
  - Bus: Beneficial for broadcast-intensive workloads?
    Concern: Serialization, bandwidth issues?
- **Power reservoir:**
  - Intra/Inter-ball power storage?
  - Reducing voltage/current spike
- **In-System Cooling:**
  - Intra-ball fan/pump/impellor?



Q&A

Thank you for your attention