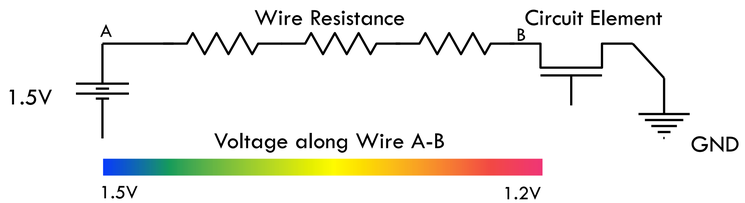
Quantifying the Relationship between the Power Delivery Network and Architectural Policies in a 3D-Stacked Memory Device

Manjunath Shevgoor, Jung-sik Kim*, Niladrish Chatterjee, Rajeev Balasubramonian, Al Davis, Aniruddha Udipi **
University of Utah, Samsung*, ARM**

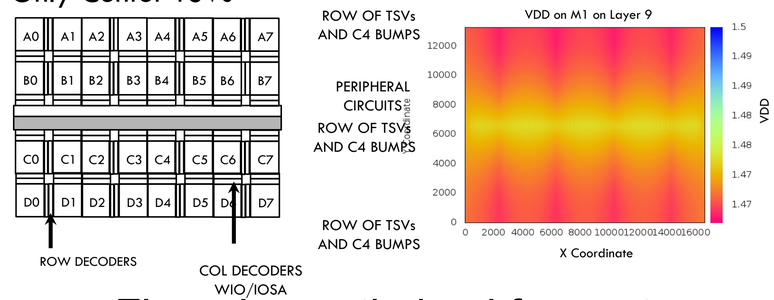
Wires are BAD!!!



- Only part of the Supply Voltage reaches the circuits.
- This loss of Voltage is called ***IR Drop***.

DRAM Layout

Only Center TSVs

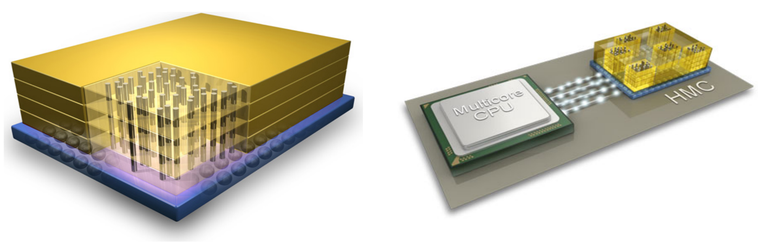


Floorplan optimized for cost

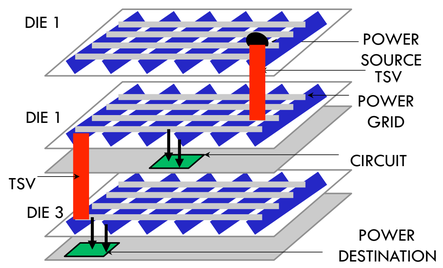
Architectural Solutions

- ***Region Based Constraints-*** Apply Constraints depending on location
- ***Reduce Controller Complexity-*** Define all constraints in terms of Read
- ***Request Scheduling-*** Avoid Starvation
- ***Data Mapping-*** Dynamically Map critical data to well supplied regions of DRAM

What is 3D DRAM?



- Regular DRAM stacked on top of each other for higher Capacity & Bandwidth



In 3D-chips, current has to travel longer distances , through TSVs.

Conventional Solutions

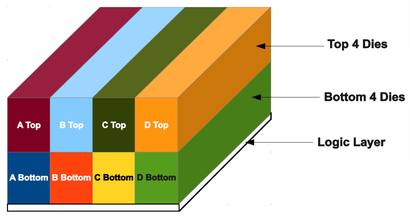
Reduce Resistance

- Make Wires Wider
- Add additional VDD/VSS bumps and TSVs
- ***Increases Cost***

Decrease Current

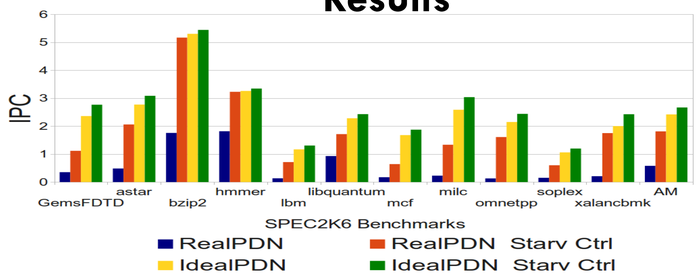
- Control Activity on Chip
- ***Decreases Performance***

Region Based Constraints



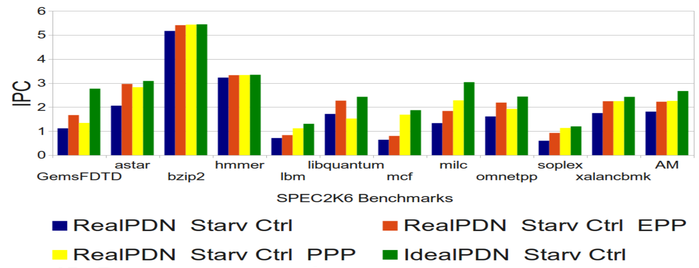
Single Region Constraints	1-4 Reads
Two Region Constraints	1-8 Reads
Four Region Constraints	1-16 Reads
Die Stack Wide Constraint	8-16 Reads

Results



Real PDN is 4.6x worse

Starvation Control Narrows this to 1.47x



IR Drop aware dynamic page placement narrows performance gap to 1.2x