Georgia Tech

Georgia Tech PACKAGING RESEARCH CENTER

Thermal Analysis of Integrated Voltage Regulators

Students: Venkatesh Avula, Lydia Mele and Claudio Alvarez
Faculties: Dr. Vanessa Smet, Prof. Madhavan Swaminathan and Prof. Yogendra Joshi

Acknowledgements*

SRC Liaisons

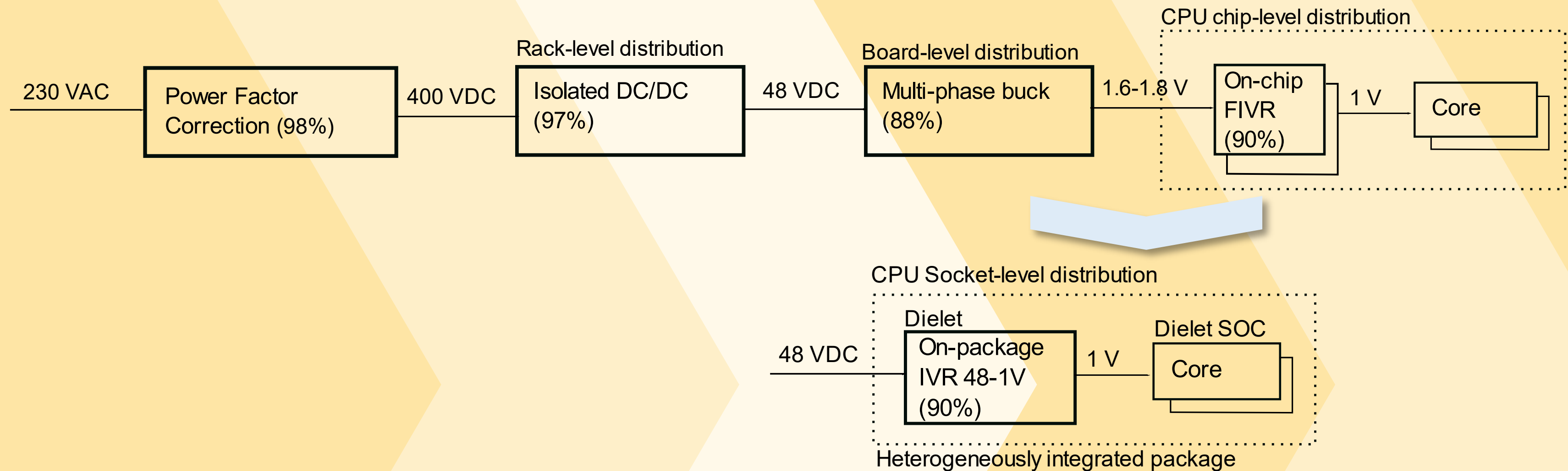
- Ashish Agrawal (Intel)
- Henning Braunisch (Intel)
- Tsu-Hsi Chang (DARPA)
- Brian Doyle (Intel)
- Tanay Gosavi (Intel)
- Augusto L. Gutierrez (Northrop Grumman)
- Nazila Haratipour (Intel)
- Jack T. Kavalieros (Intel)
- Harish K. Krishnamurthy (Intel)
- Elizabeth Kunkee (Northrop Grumman)
- Ravi Pillarisetty (Intel)
- Willy Rachmady (Intel)
- Abhishek A. Sharma (Intel)
- Shriram Shivaraman (Intel)
- Seung Hoon Sung (Intel)
- Han Wui Then (Intel)
- Xin Zhang (IBM)

* This work was supported in part by ASCENT, one of six centers in JUMP, a Semiconductor Research Corporation (SRC) program sponsored by DARPA.

1. Research Objectives

Thermal management of HPC/Server

- For high performance servers in data centers, a typical power delivery sequence[1,2] shown below has around 75% power efficiency.



- If a 48-1V integrated voltage regulator is heterogeneously integrated with CPU SOC at socket-level, as shown below, the overall efficiency can be increased to 85%.

→ Such functional integration comes with an increase in thermal density, in systems that are **already thermally limited**. Thermal management, therefore, is a key challenge that must be addressed to enable such integrated and efficient power conversion.

→ **Research objective**: model, design and demonstrate substrate-integrated cooling solutions for PRC's 48V-1V IVR modules with embedded inductors.

[1] Corey Gough, Ian Steiner and Winston Saunders, Energy Efficient Servers, Apress, 2015
[2] Philip T. Krein, Data Center Challenges and Their Power Electronics, CPSS TRANSACTIONS ON POWER ELECTRONICS AND APPLICATIONS, VOL. 2, NO. 1, MARCH 2017

2. Server power delivery challenges

2.1 Thermal impact on voltage regulator performance

- Power loss, dissipated as heat, in a voltage regulator depends on the regulator's efficiency (η) and power output.

$$P_{loss} = P_{out}\left(\frac{1}{\eta} - 1\right)$$

- Major components of the power loss are switching and conduction losses. Switching loss includes power MOSFET and ac magnetic core loss, while conduction loss is due to DC resistance of power MOSFET and inductor.

$$P_{loss} = P_{sw} + P_{cond} = P_{sw} + I^2 R_{DSon_mosfet} + I^2 R_{inductor}$$

- Conduction loss depends on temperature, as resistance varies with temperature.

$$R_{Tf} = Ro\left(1 + \alpha(T_f - To)\right)$$

Where T_o, Tf are reference and current temperatures respectively.
α is temperature coefficient. R_{Tf} and R_o are resistance at temperatures T_o and Tf respectively.

$$\alpha_{Cu} = 0.00386 \text{ per degree Celsius}$$
$$\alpha_{GAN_MOSFET} = 1.5 \text{ to } 2.0 \text{ per degree Celsius } [1]\ [2]$$

- Above positive temperature coefficients cause positive feedback loop between temperature and heat generating losses, ultimately resulting in thermal runaway.

→ Thus thermal management is critical for integrated voltage regulator (IVR) design.

[1] Eric Persson, CoolGaN application note, AN_201702_PL52_010, Infineon, 2018
[2] Stephen L. Colino and Robert A. Beach, Fundamentals of Gallium Nitride Power Transistors, APPLICATION NOTE: AN002, EPC, 2011

2. Server power delivery challenges

2.2 Chip-level thermal limitations

- The current on-chip integrated voltage regulators have power dissipation heat flux densities in excess of 2 W/mm^2.

Power density per on-chip IVR (W/mm^2)	**Intel FIVR 1.8:1V [1]** 140MHz 30A/16A from 16 phases 22nm process 90% peak efficiency- Buck	**IBM 1.8:0.9V [2]** 100MHz 19.1mA 32nm SOI process 86% peak efficiency – Switched Cap	**IBM Power8 1.1:1V [3]** 32nm SOI process 90.5% peak efficiency – LDO
Power regulation	33.6	4.6	34.5
Power dissipation as heat flux[4]	3.73	2.14	3.62

- But conventional air-based cooling solutions are limited to 1.5 W/mm^2 , motivating us to look for inexpensive, high performance liquid-based alternatives.

Heat rejection type	Heat flux limit (W/mm^2)
Air [5]	1.5
Microchannel liquid [6]	7.9

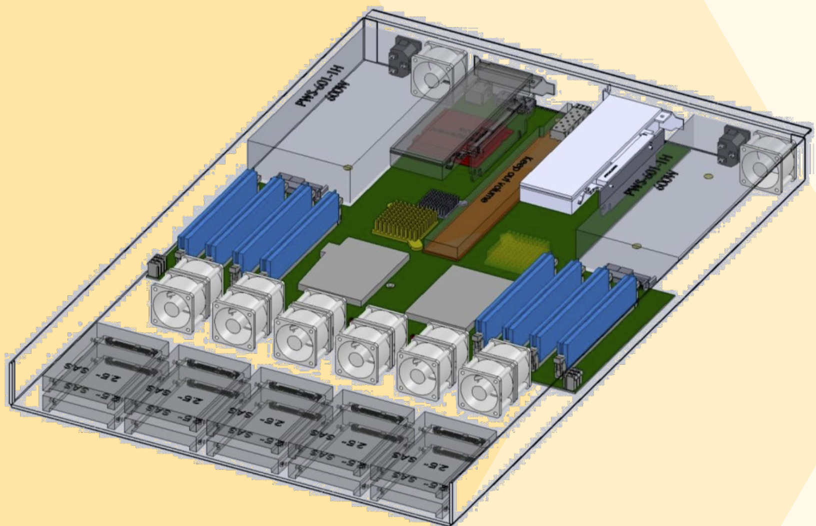
- If the heat is not properly rejected, die-level local thermal hotspots form, negatively affecting reliability and performance of CPU.
- In such cases, thermal gating or selective shutdown of IVR phases is possible, but it slows down the IVR's response to the power demands of CPU.

[1] FIVR — Fully integrated voltage regulators on 4th generation Intel® Core™ SoCs, Edward A. Burton et al, IEEE APEC 2014
[2] A 4.6 W/mm2 Power Density 86% Efficiency On-Chip Switched Capacitor DC-DC Converter in 32 nm SOI CMOS, T M. Andersen et al
[3] Distributed system of digitally controlled microregulators enabling per-core DVFS for the POWER8TM microprocessor, Zeynep et al, 2014
[4] Karen Khatamifard et al, ThermoGater: Thermally-Aware On-Chip Voltage Regulation, ISCA '17, June 24-28, 2017, Toronto, ON, Canada
[5] Interaction of Scaling Trends in Processor Architecture and Cooling, W Huag et al, 2010
[6] High-performance heat sinking for VLSI, Tuckerman and Pease, 1981

3. State-of-the-art of thermal management for IVRs

Board-level VRM (industry standard)

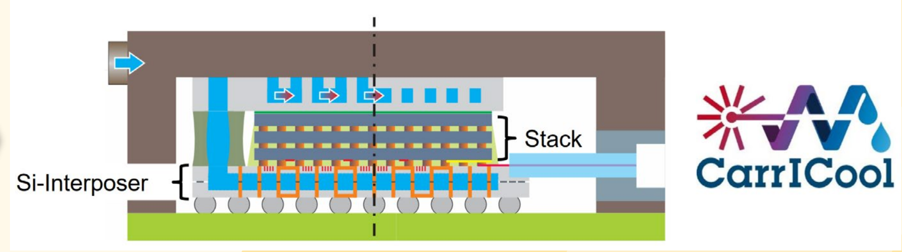
- **Forced air cooling in a rack**



Direct liquid-based solutions are costly and hard to implement at large scale.
→ **How can we create hybrid liquid/air-based cooling solutions that are cost competitive and deliver superior thermal performance?**

IVR (R&D)

- **Indirect microfluidic cooling [1]**
  - Si interposer with integrated microfluidic cooling
    - Expensive process, but costs are coming down



- **Direct liquid immersion cooling [2] [3]**
  - Cooled with a 3M Fluorinert dielectric fluid[3]

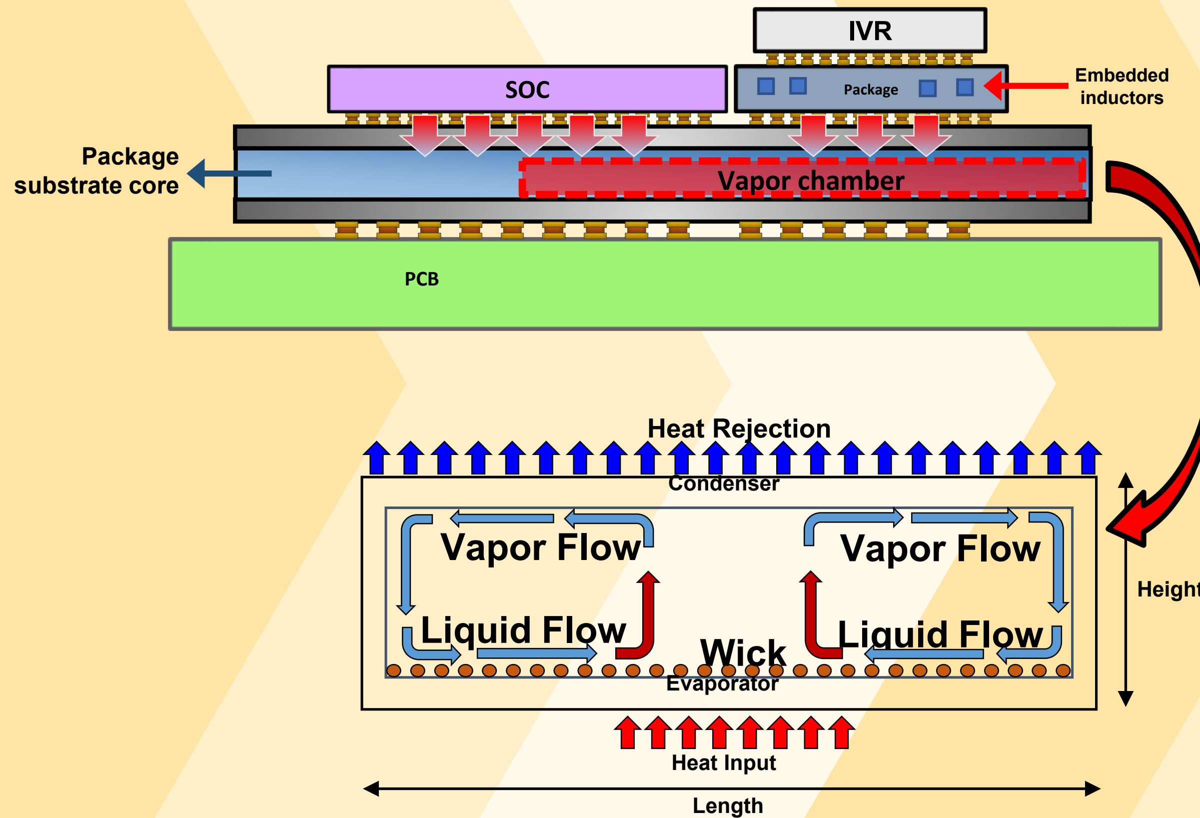
- Possible candidate for supercomputers
- Cost high for most commercial products
- Material compatibility issues

[1] Arvind Sridhar, IBM, Switzerland , Advanced Power and Thermal Packaging Enabling Dense Integration of Compute Nodes, PwrSoC 2018
[2] http://www.vicorpower.com/press-room/unveiling-the-first-super-computer
[3] https://news.3m.com/press-release/company/3m-fluorinert-helps-pezyexascaler-rise-top-green500-list

4. Technical approach

4.1 Thermal management with integrated vapor chambers

- Board- or package-integrated **vapor chamber** is proposed as a thermal solution for the IVR, in combination with forced air cooling.



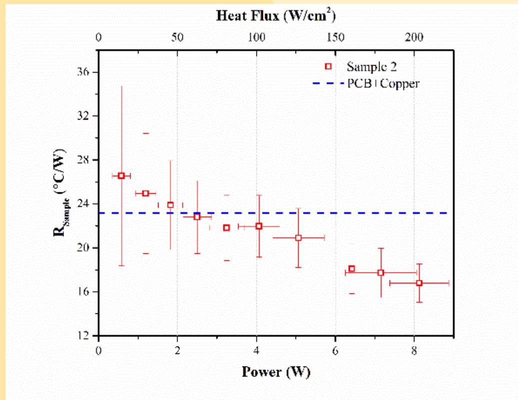
- A vapor chamber [1] is a two-phase heat spreader, utilizing latent heat of vaporization. It is also passive, not requiring additional power input.

[1] Sangbeom Cho, Thermal performance enhancement of packaging substrates with integrated vapor chamber, PhD thesis, 2018

4. Technical approach

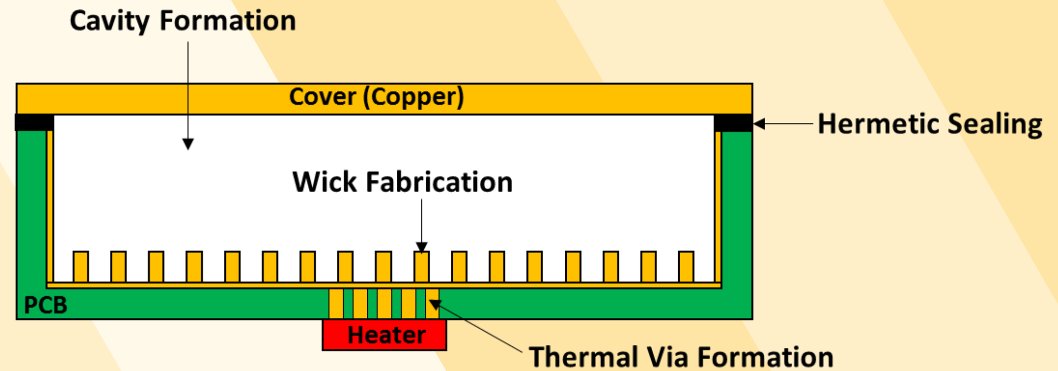
4.2 Prior Work on Vapor Chambers *(courtesy of Dr Sangbeom Cho, Prof. Joshi)*

- 25% improvement in heat spreading compared to Cu RDL

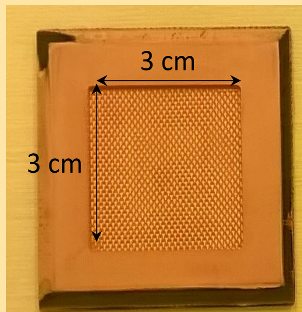


(a) thermal resistance of PCB integrated with vapor chamber vs. Cu plated PCB (total thickness: ~ 1 mm)

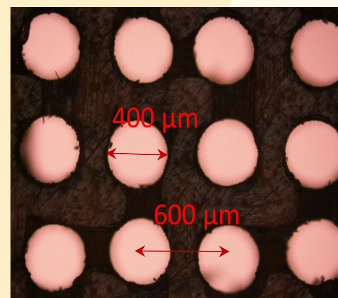
- Established fabrication processes leveraging standard SAP processes

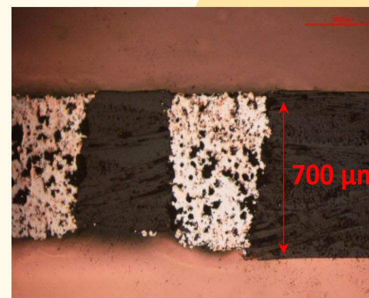


- **Cavity formation**

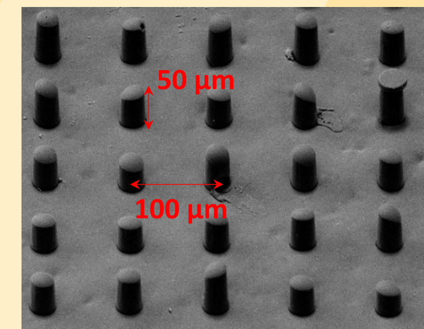


- **Thermal via formation**





- **Wick structure fabrication**

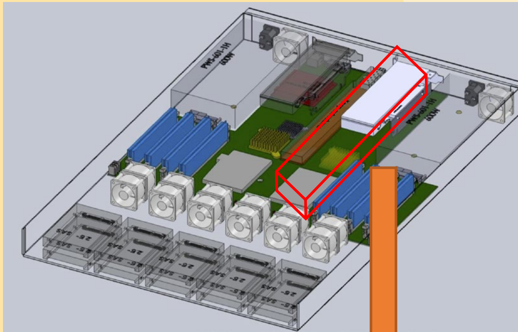


5. Thermal analysis

5.1 Thermal model of VRM vs. IVR in conventional Server rack

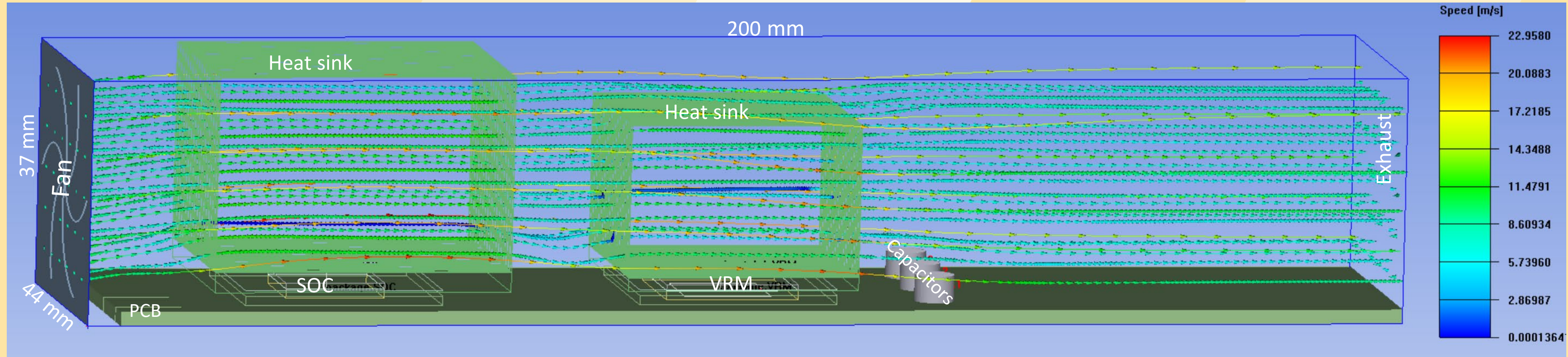
- It is essential to analyze thermal limitations at system and chip-level before developing the necessary thermal solutions.
- For the thermal analysis, the following enclosure consisting of a CPU and VRM on a typical Server motherboard is considered.

(a) 1U HPC System[1]



Computational fluid flow analysis
Software package: Ansys Fluent
Type of analysis: Steady state
Type of flow regime: Turbulent
Initial Conditions: ambient temperature of 42°C
Boundary Conditions of cabinet walls: impermeable adiabatic boundary.
Left axial surface boundary: intake fan with fixed volumetric flow of 30 cfm
Right axial surface boundary: exhaust
Governing equations: conduction and convection
Variables solved: temperature and flow

(b) Thermal model enclosure [2]



[1] Bob Ogrey, Eugene Chung, Otto Joe, AMD Motherboard Hardware v3.0, Open Compute project, 2012
[2] Miguel Rodriguez, Stephen Kosonocky and Ali Merrikh, Server Power Delivery Challenges and Opportunities, PowerSOC Conference 2014, Boston, MA, USA

VRM-Voltage regulator module on board
SOC-System on chip, CPU
HPC-High performance computing

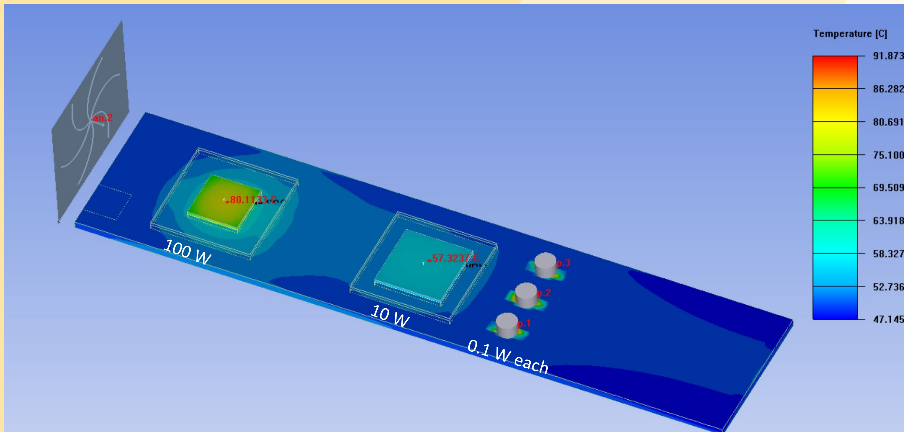
5. Thermal analysis

5.2 System-level impact of IVR

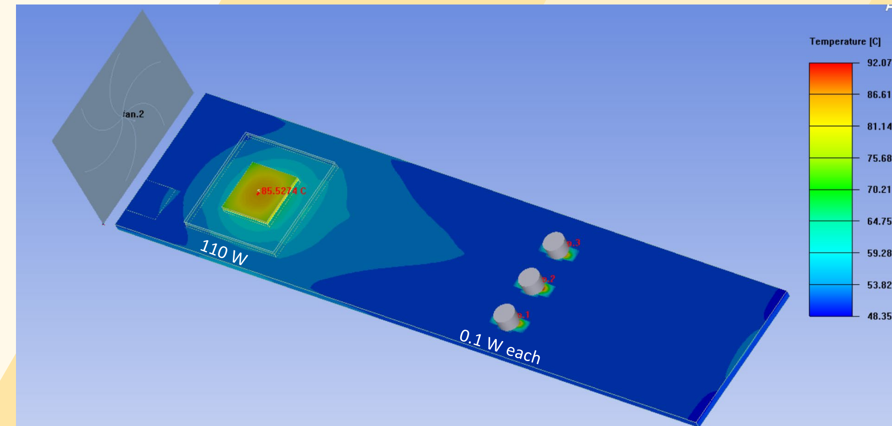
- To ascertain the severity of thermal challenges, we compared the traditional regulator on-board configuration with that of an on-chip IVR.
- An increase of 5.5°C on CPU junction temperature is observed, when the voltage regulator is moved from on-board to on-chip.

Configuration	Fan flow (CFM)	$T_{ambient}$ (°C)	T_{pcb} (°C)	T_{vr} (°C)	$T_{cpu_junction}$ (°C)
VRM	30	42	91	57	80
IVR	30	42	82	85.5	85.5

(a) VRM on-board configuration: PCB surface temperature profile



(b) IVR on-chip configuration : PCB surface temperature profile

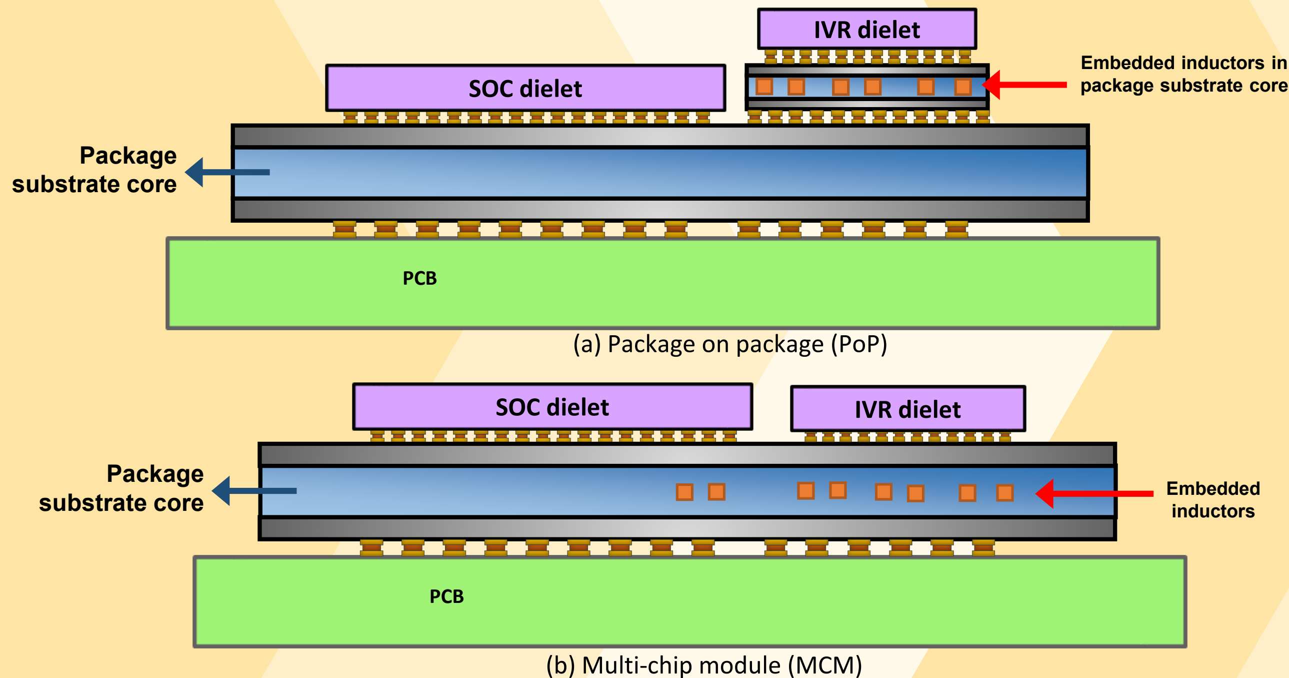


T_{VR}-temperature of voltage regulator

5. Thermal analysis

5.3 Heterogenous integration architectures

- For the IVR integration, two heterogenous architectures are being explored.
- Inductors are embedded either in the separate IVR package or in the substrate of the multi-chip package, as shown below.



(a) Package on package (PoP)

(b) Multi-chip module (MCM)

- For the thermal analysis, TDPs assigned are:

Dimensions in mm	**SOC** (12x12x.0.75)	**IVR** (10x10x0.75)	**Embedded inductor** (2x1x0.35)
TDP (W)	100	10	0.15

*TDP-Thermal design power

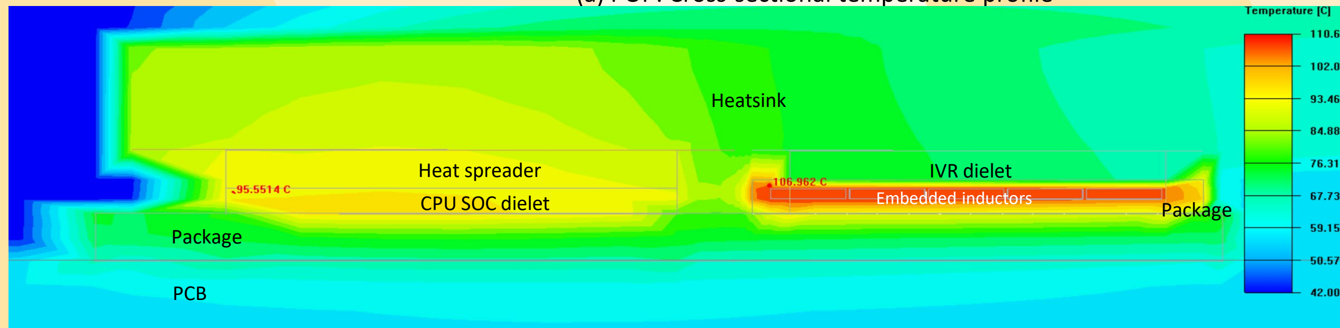
5. Thermal analysis

5.4 Heterogenous integration architectures: temperature profiles

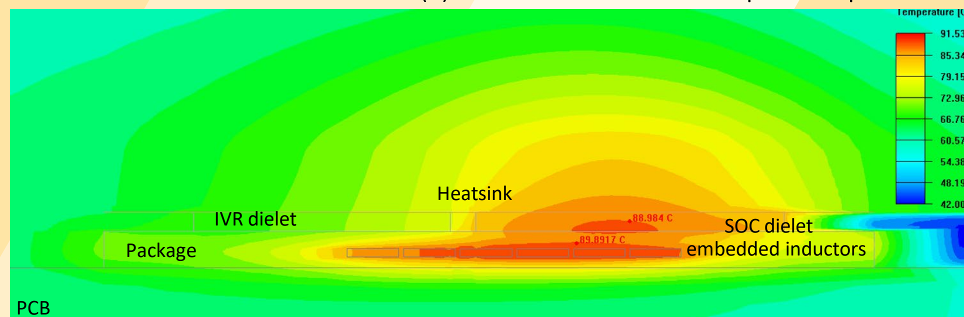
- The embedded inductors, as shown below, reach high temperatures in the range of 89-107°C.

Configuration	Fan flow (CFM)	$T_{ambient}$ (°C)	$T_{inductor}$ (°C)	$T_{cpu_junction}$ (°C)
POP	30	42	107	93.5
MCM	30	42	89.89	88.98

(a) POP: Cross-sectional temperature profile



(b) MCM : Cross-sectional temperature profile

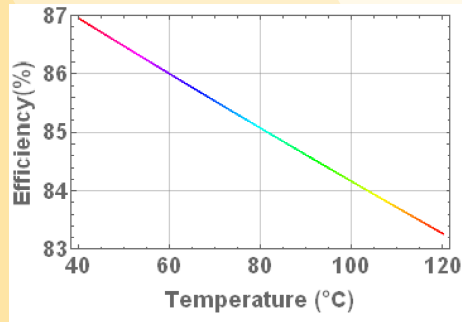


- Joule heating is neglected in this analysis. With joule heating, the temperatures are going to be even higher.

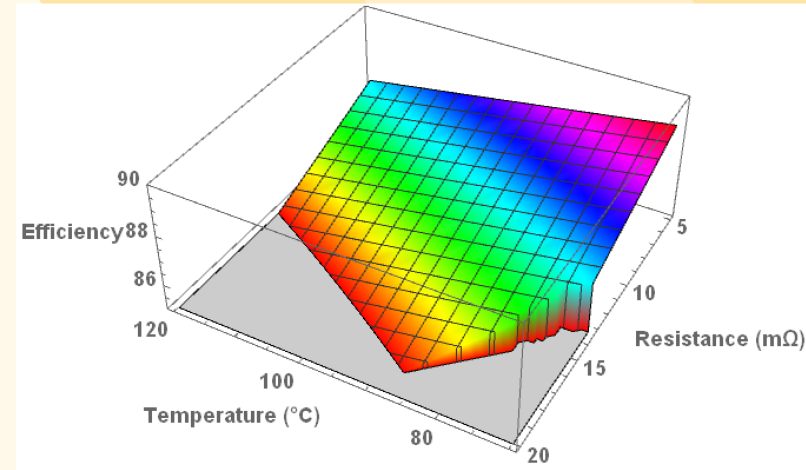
5. Thermal analysis

5.5 Thermal impact on IVR's overall efficiency

- At such high temperatures of both inductor and IVR chip, the IVR power efficiency drops, as shown below.



(a) Temperature impact on efficiency

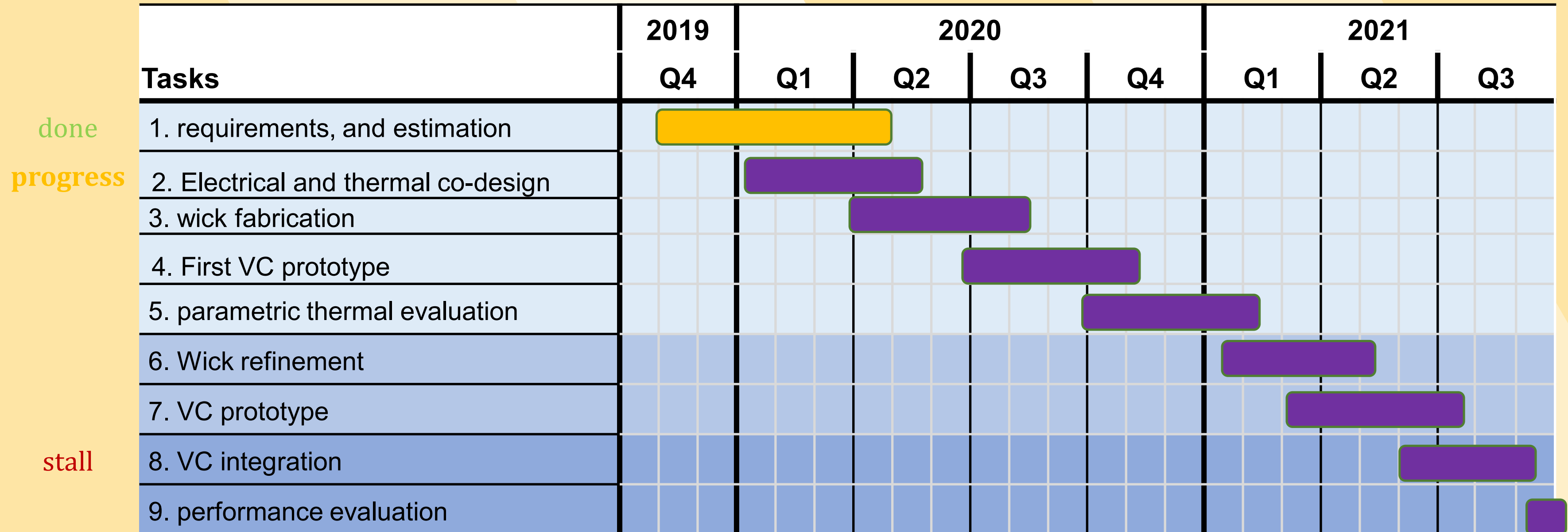


(b) Design space response

→ We need more efficient thermal management to keep the IVR cool and operate at its peak efficiency operating point.

- For low resistance inductors, of around 5mΩ, the design space is open and acceptable efficiency figures of 85–90% can be achieved.
- But for high resistance inductors, of more than 15mΩ DC resistance, the design space has shrunk, and an acceptable design point cannot be reached unless cooling solutions and thermal isolation techniques are employed.

6. Schedule



Task 1: Thermal characterization of magnetic core, determination of required heat removal requirements, and estimation of VC dimensions through simulations.
Task2: Co-design of package architecture: electrical thermal multiphysics package system architecture design (number of levels, manufacturing, thermal, reliability), electro-thermal modeling (analysis of effect of thermal solution on electrical performance), Joule heating – electrical-thermal coupling
Task3: Wick materials and working fluid(s) selection, and wick fabrication demonstration
Task 4: First VC prototype fabrication, assembly, and charging
Task5: First VC prototype parametric thermal evaluation
Task 6: Wick design/manufacturing refinement through detailed VC modeling
Task 7: Second generation VC prototype fabrication, assembly, and testing
Task 8: Integration of VC with test bed
Task 9: Thermal performance evaluation with test bed