

R&T PCle400 : Problématique refroidissement



14 avril 2022



Julien Langouët,
Kévin Arnaud, Paul Bibron, Jean-Pierre Cachemiche,
Renaud Le Gac, CPPM

Sommaire

Spécification thermique PCIe400

- Estimation de puissance FPGA
- Modèle thermique FPGA
- Bilan de puissance dissipée

Spécification thermique serveur HPC

- Flux d'air serveur
- Ventilateurs et flux d'air

Spécification radiateur PCIe400

- Inspiration GPU
- Dimensions radiateur
- Types de radiateur
- Comparaison simple type de radiateurs
- Exemple évaluation d'un modèle de radiateur

Sous-traitance

Conclusion

Estimations de Puissance

Modèles de consommation Agilex

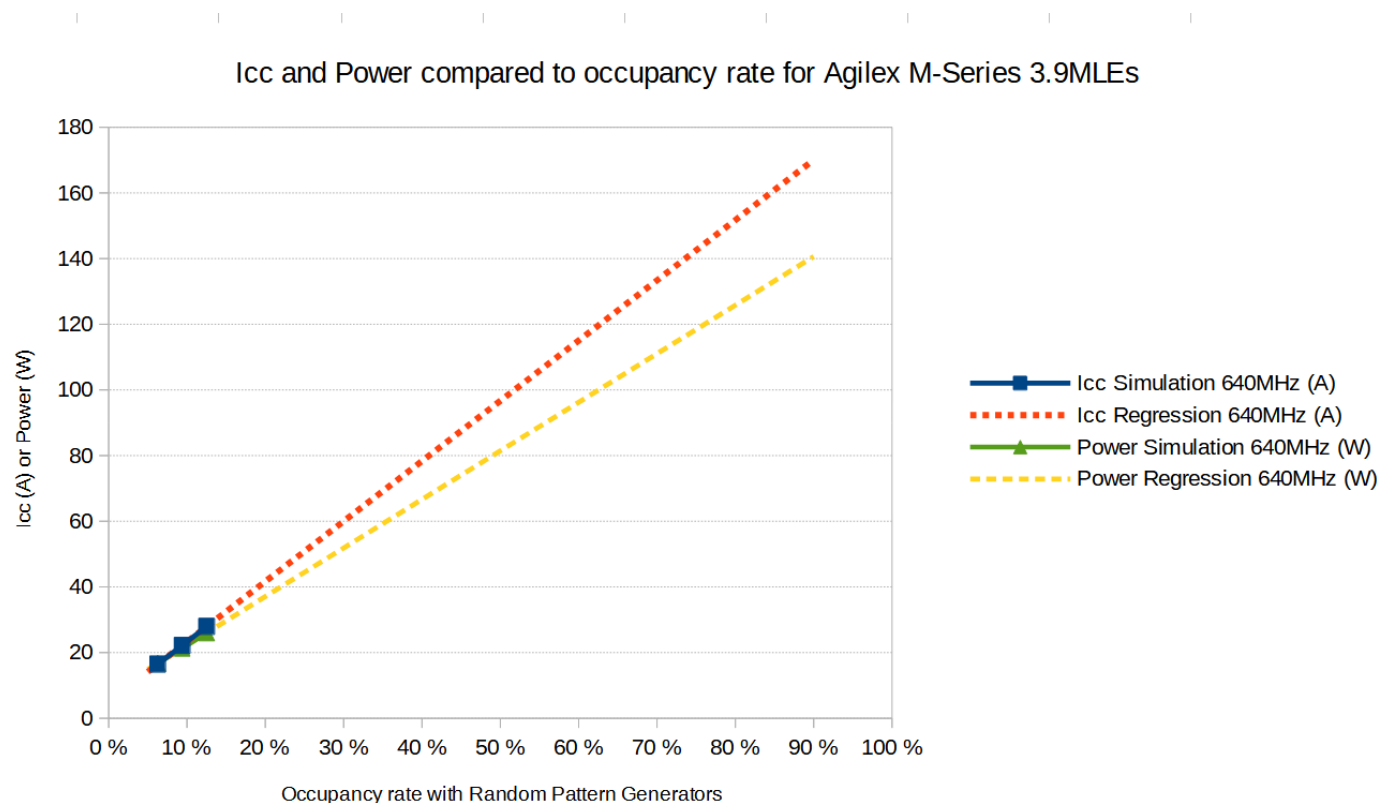
- Le AGMF039R47AxE2Vxx non disponible à ce jour (56x66mm) 3xF-tile + 1xR-tile + 2xHBM2 16GB
- Le AGIB027R31BxxxxR0 disponible sur PTC >21.4pro (56x45mm) 4xF-tile
 - I-series utilise même cœur que M-series

Simulation Post Fit + Power Analyzer : précis mais lent

- Random Pattern/640MHz/90 % Occupation : 170A ou 140W sur logique du cœur

PTC :

- Estimation Transceivers
 - R-Tile PCIe Gen5
 - F-Tile 48x 32G
- **35W**

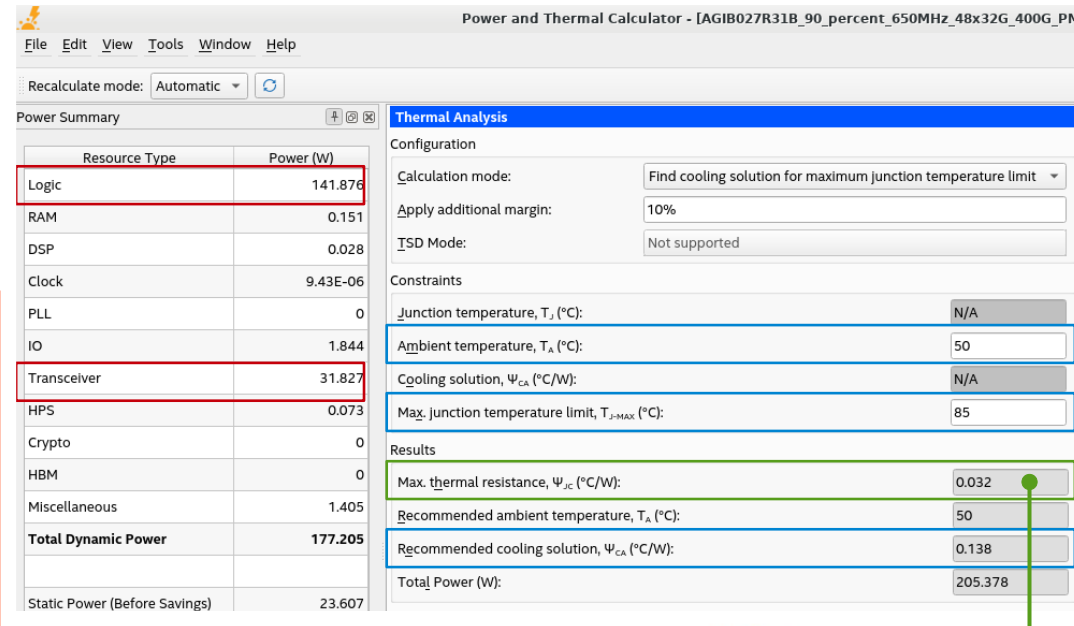
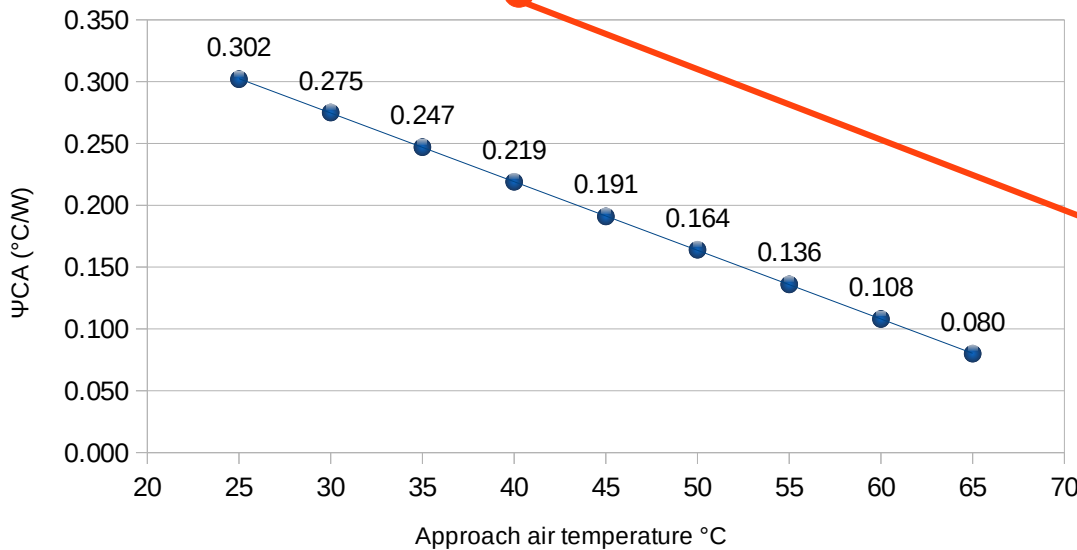


Solution refroidissement FPGA

Utilisation du PTC

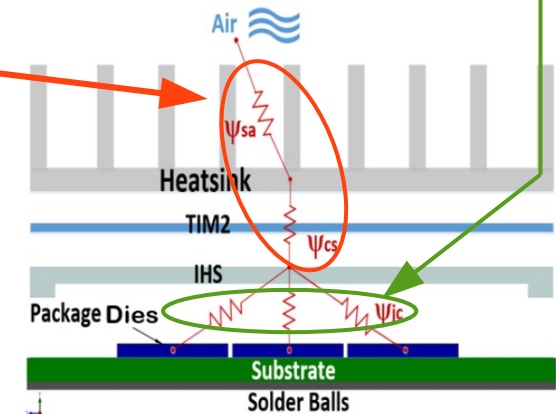
- Modification d'un fichier PTC ciblant AGIB027R31BxxxxR0 pour imiter la consommation extrapolée du AGMF039R46B
- Utilisation du mode *Find cooling solution for maximum junction temperature*

Heat sink thermal resistance Ψ_{CA} ($^{\circ}\text{C}/\text{W}$) vs ambient temperature



- Thermal Interface Material (TIM)

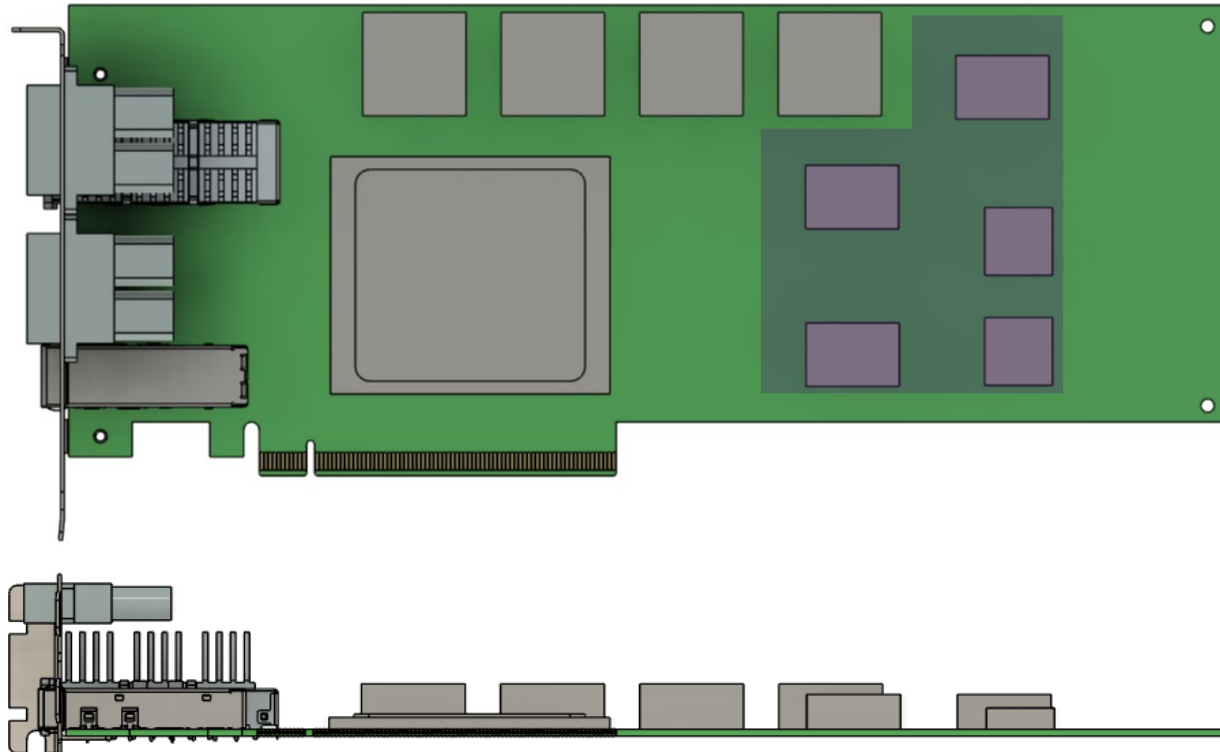
○ Ex $k=5\text{ W}/\text{m}\cdot\text{K} \Rightarrow \Psi_{cs} = \frac{e}{A \cdot k} = \frac{0.25\text{ mm}}{52 \times 50\text{ mm} \cdot 5\text{ W}/\text{m}\cdot\text{K}} = 0.019^{\circ}\text{C}/\text{W}$



Bilan puissance dissipée

PCIe AIC double slot full height 3/4 length

- **FPGA** max TDP 180W (90 % fill factor), $T_{j\ min} = 0^{\circ}\text{C}$, $T_{j\ max} = 85^{\circ}\text{C}$, contact pad $\sim 25\text{cm}^2$ (49x52mm)
- **4 BOA** : TPD 4x 7.5W, $T_{j\ max} = 70^{\circ}\text{C}$,
- **1 QSFP112** : TPD $\sim 12\text{W}$ (?), $T_{j\ max} = 70^{\circ}\text{C}$,
- **1 SFP** : TDP $< 1\text{W}$, $T_{j\ max} = 85^{\circ}\text{C}$
- **Alimentation** : 2 LTM4681 VCC Core + 1 LTM4681 + 2 LTM4677 + ...
 - Contrainte sur l'encombrement du radiateur en hauteur ! (LTM4681 8.17mm)
 - Pas de radiateur requis, 5 à 7W par module largement dissipé dans le PCB

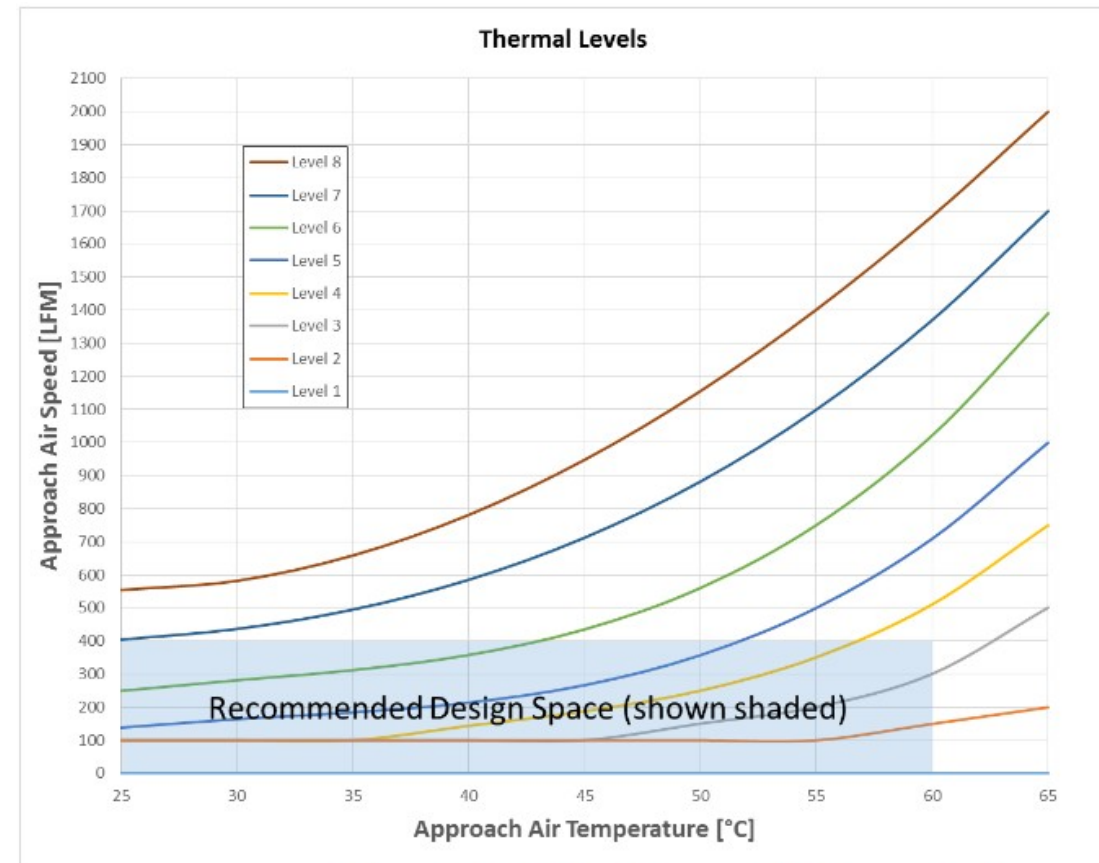
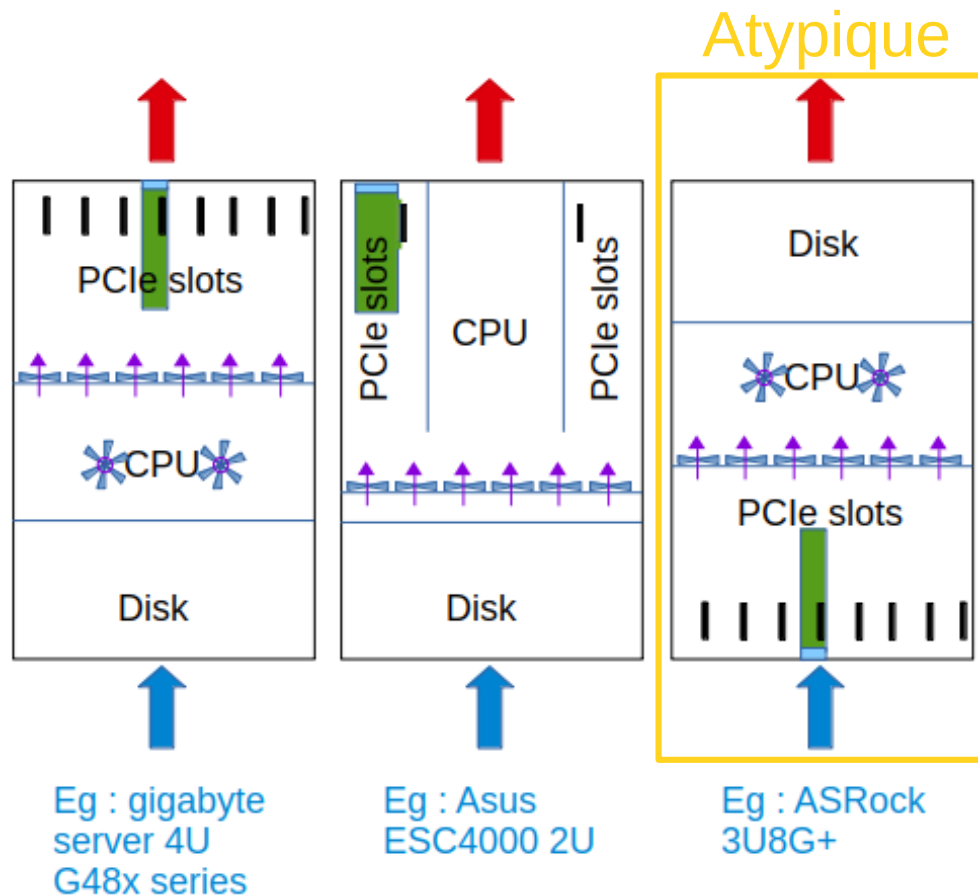


NB : illustration AIC de 274mm
AIC 3/4 = 254mm

Flux d'air serveur

19'-rack serveur

- Form factor 2U à 4U (~90 à ~175mm)
- Direction du flux d'air : **pas de standard**, architecture **front to back** est largement **répandue**
- Spécification PCI SIG CEM rev 5 : fonctionnement au TDP avec air ambiant entre 25 et 60°C et débit entre 100 et 400LFM (0.5m/s à 2m/s)



Ventilateurs et flux d'air

Capacité de flux d'air dans un serveur

- Caractéristique P-Q (pression vs Débit)
 - Varie en fonction des RPM
 - Nombre de ventilateurs en série
 - Nombre de ventilateurs en parallèle

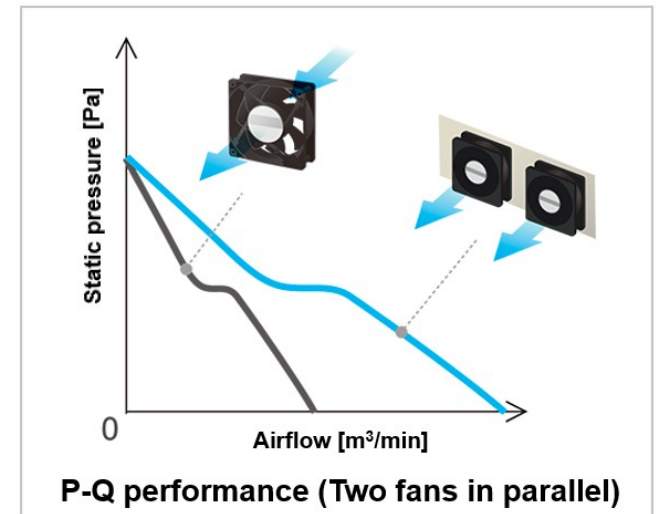
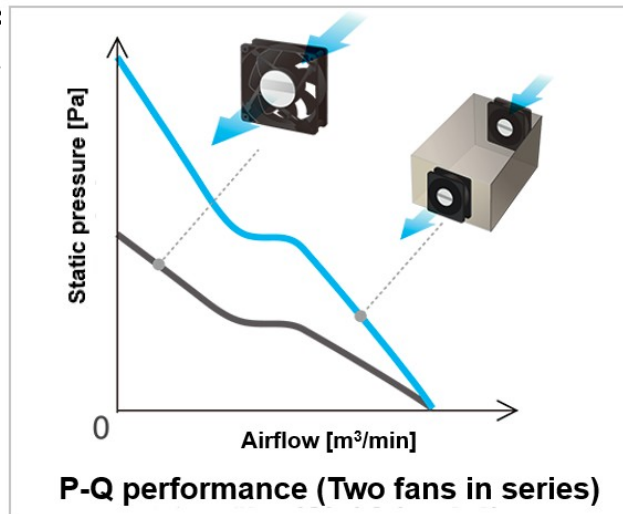
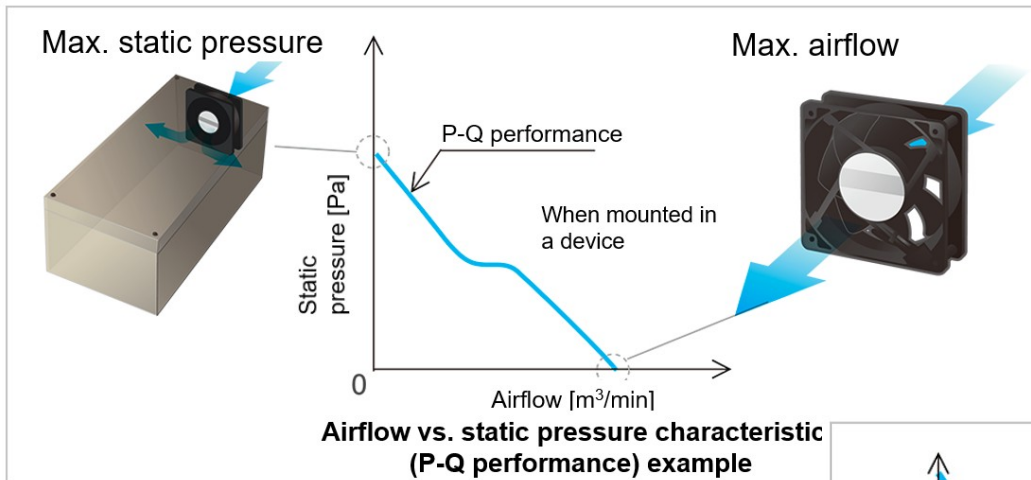
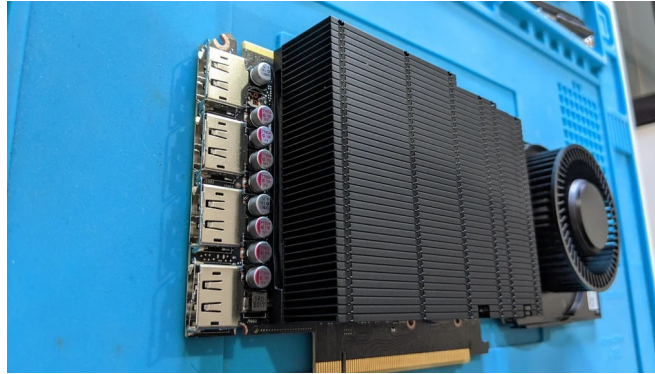
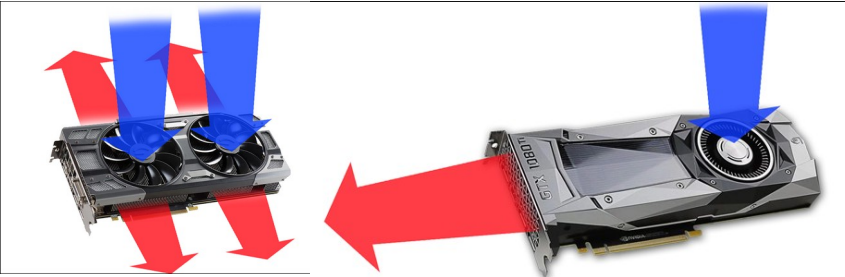


Illustration from [Sanyo Denki cooling training fan basic](#)

Inspiration cartes GPU

Cartes « gaming » pour PC tour (PC ATX)

- Architecture active « open air » ou « blower ». TDP jusqu'à **300W**

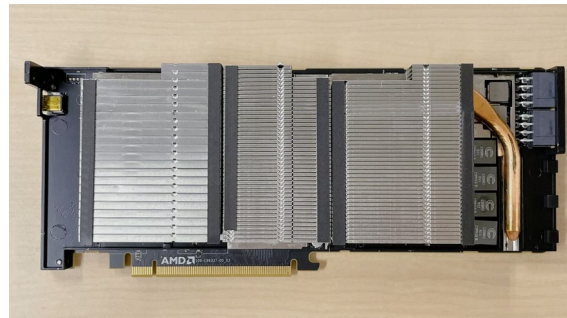


Nvidia A5000



Cartes « serveur »

- Architecture passive. TDP jusqu'à **350W** accepte un flux d'air dans les deux sens.



AMD Firepro S9300



Nvidia A100

Choix d'une architecture passive

- Flux d'air contrôlé par le serveur (cf thermal reporting PCI SIG CEM rev 5)
- Complexité réduite sur la PCIe400

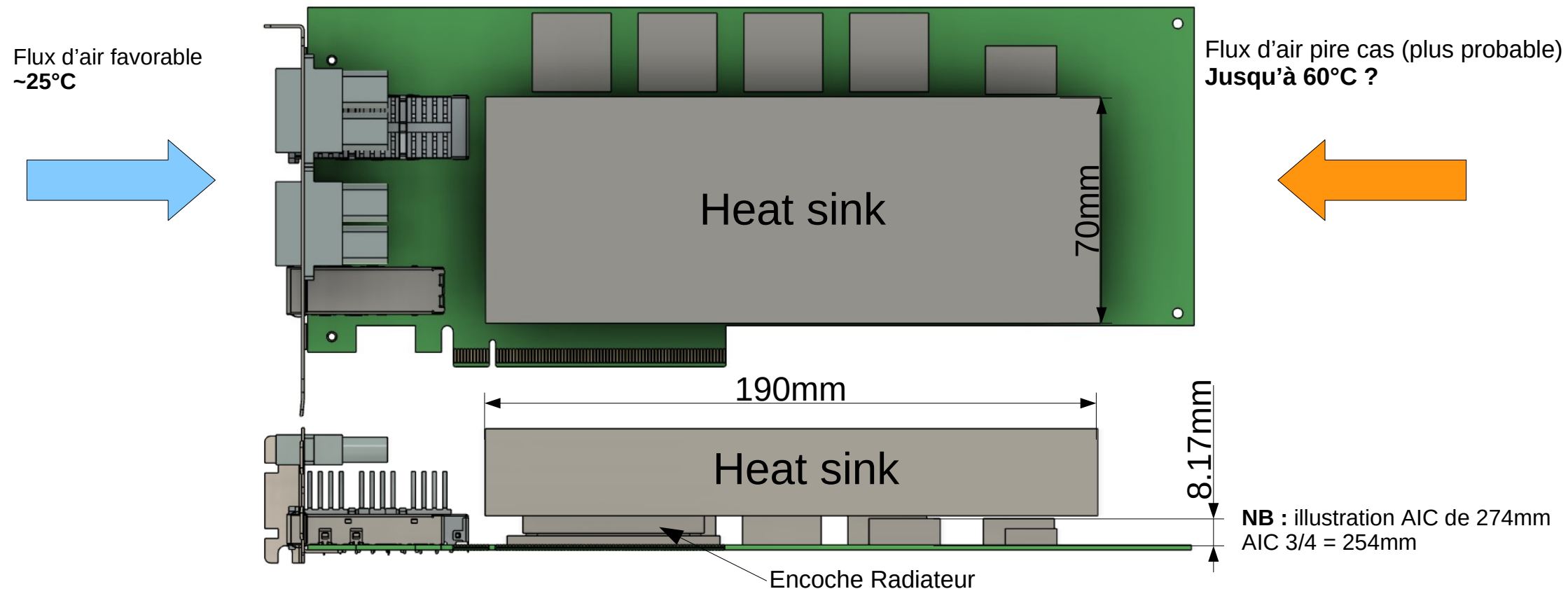
Dimension radiateur

Dimensions maximum

- Largeur ~70mm
- Hauteur ~27mm + encoche pour contact avec FPGA (~5.4mm)
- Longueur jusqu'à ~190mm (si format 3/4)
- Flux d'air : Deux cas à considérer

Carénage

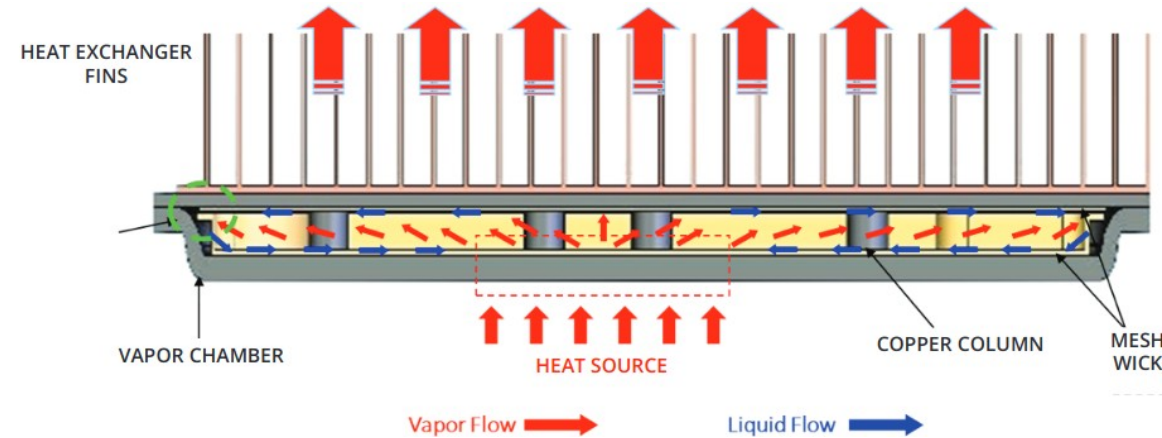
- Ajout possible si amélioration de l'efficacité



Type de radiateur

Base

- Aluminium conductivité $\sim 200 \text{ W/m.K}$
- Cuivre conductivité $\sim 380 \text{ W/m.K}$
- Caleoduc (vapor chamber / heat pipe) conductivité $\times 5 - 10$ (dépend de la surface de la source / surface du radiateur)

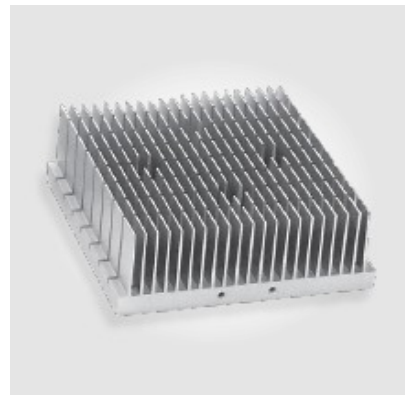


Ailette

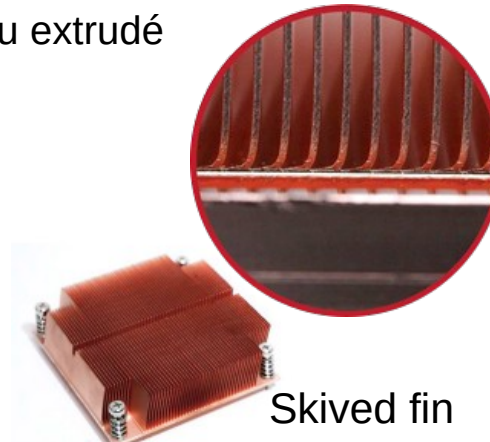
- Pin fin
- Cross cut fin : ailettes taillées dans la masse ou extrudé
- Skived fin
- Zipper fin



Pin fin



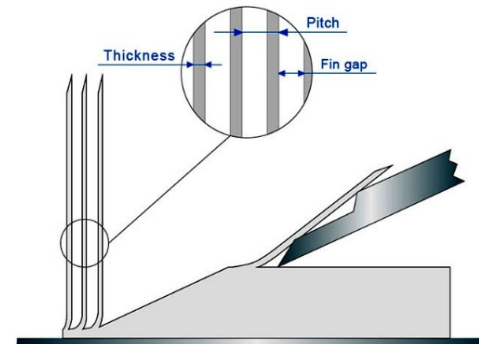
Cross cut



Skived fin

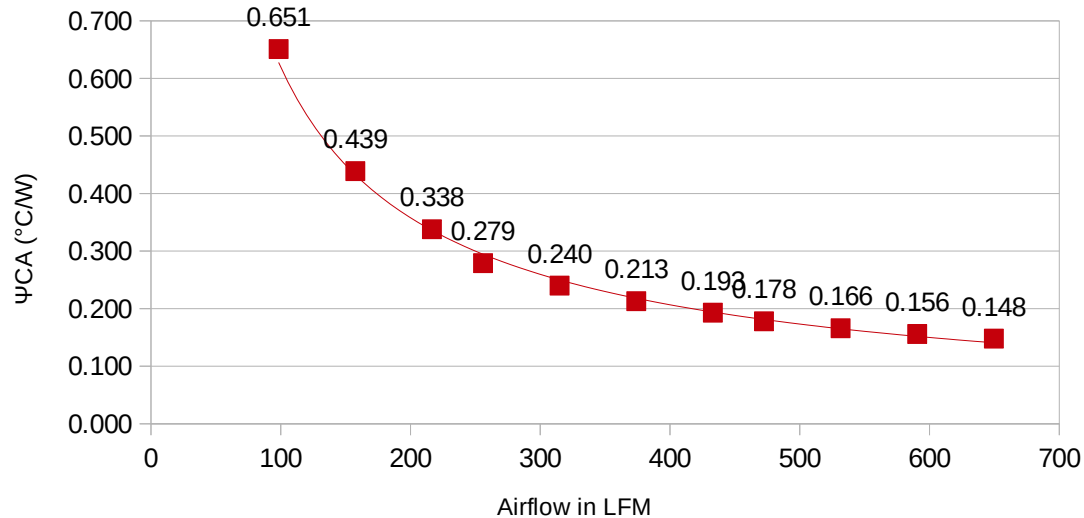


Zipper fin



Exemple évaluation d'un modèle de radiateur

Heat sink thermal resistance vs airflow



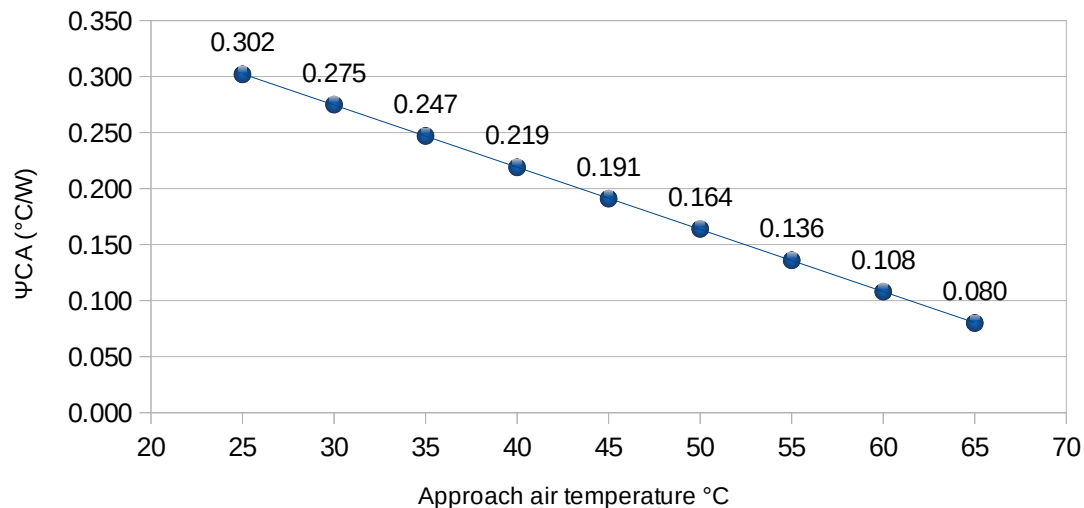
Caractéristique radiateur

- Vapor chamber + copper fins
- Heat sink WxLxH 70x190x27mm
- Base 2mm
- Fins 47x0.15mm
- Airduct WxH 32x105mm

Domaine de fonctionnement pour $T_j < 85^\circ\text{C}$

- $T_A = 25^\circ\text{C} \rightarrow \sim 250\text{LFM}$ (1.3m/s)
- $T_A = 52^\circ\text{C} \rightarrow \sim 650\text{LFM}$ (3.3m/s)

Heat sink thermal resistance ΨCA (°C/W) vs ambient temperature



Les ventilateurs du serveur sont-ils capables de fournir une pression statique suffisante ?

- Vérifier la caractéristique P-Q du ventilateur
- Vérifier la caractéristique P-Q du radiateur
- Trouver le point de fonctionnement

Liste sous-traitants conception radiateur

- Pour pallier à la difficulté sur les ressources au sein du projet, possibilité de faire appel à un sous-traitant
- Recherche et contact de plusieurs sous-traitants

COMPANY	Origin	Services	Technology	Contact
ALPHA NOVATECH	JPN	CFD, COTS customizable	cross cut machined / extruded	FAE
RADIAN HEATSINKS	USA	CFD, PCIe, Prototyping, industrialization, Production, test bench	VC/HP, cross cut,	FAE
CECLA METAL PROCESS	FR	CFD, PCIe, Prototyping, industrialization, Production, test bench	HP, machined, skived,	Commercial
BOYD CORP	USA+FR	CFD, PCIe, Prototyping, Production	VC/HP, skived, zipper, bonded, extrusion,	Commercial
CELSIA INC	USA	CFD, prototyping ?, production ?	VC/HP, zipper, skived, extruded, machined	CEO
ATS INC	USA	CFD, prototyping, production ?, airflow instrumentation, liquid cooling	VC/HP, zipper, skived, extruded, machined, liquid cooling	FAE
PIONEER THERMAL HEATSINK TECHNOLOGY	CHN	COTS VC HS, CFD, prototyping, industrialization, production	VC/HP, zipper, skived, extruded, machined	-
MYHEATSINKS	USA+CHN	custom HS, COTS HS, test bench	VC/HP, zipper, skived, extruded, machined	-
HEATSCAPE	USA	custom HS, COTS HS, large clients	VC/HP, zipper, skived, extruded, machined	-
FUJIKURA THERMAL	JPN	large clients	VC/HP	-
DYNATRON	USA	COTS VC	VC	-

Conclusion

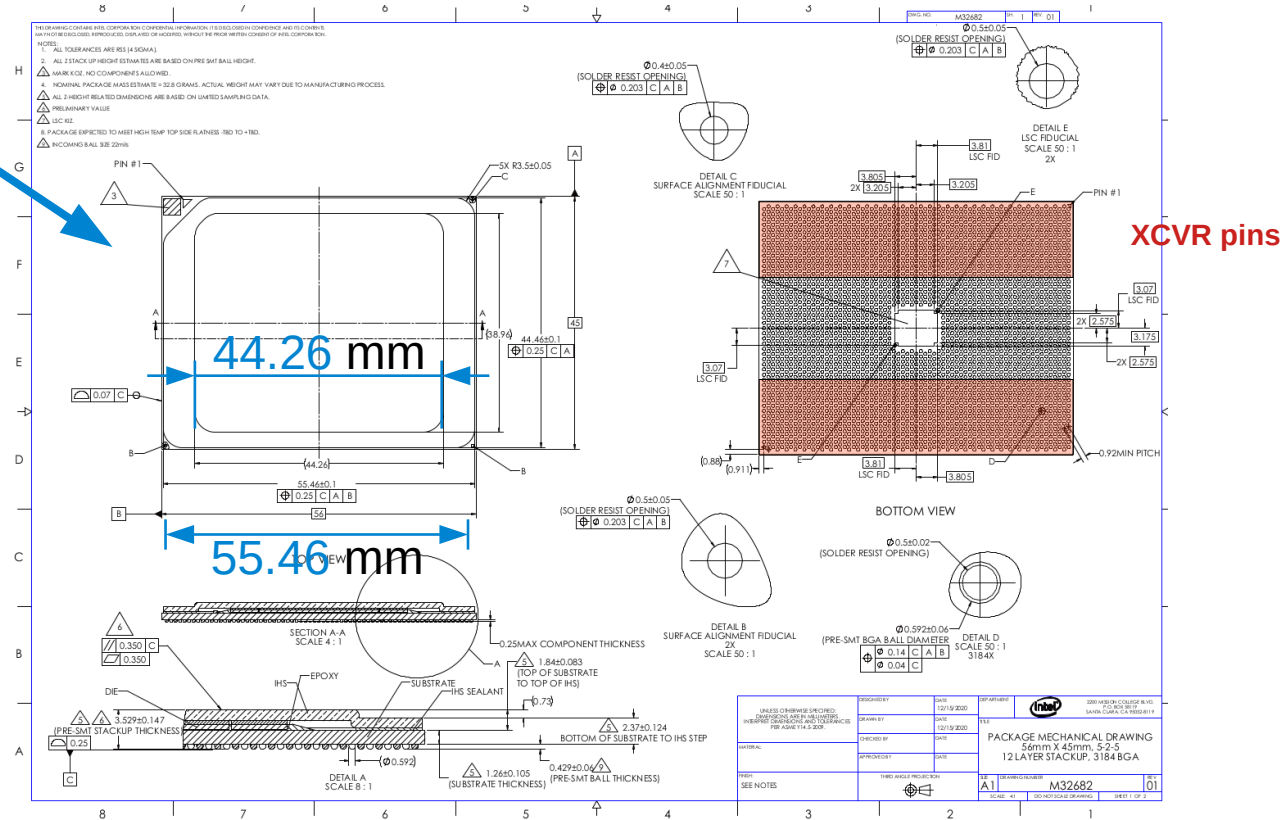
Faisabilité d'une solution de refroidissement à air

- Une solution passive permet une versatilité du flux d'air des serveurs cibles
- Les contraintes d'encombrement mécanique exigent une haute performance du radiateur
 - Utilisation de caleoducs
- Solution semble faisable dans une gamme de température de **25 à 50°C (pas 60°C !)** avec un débit max de 650LFM (3.3m/s), proche des conditions réelles ?
 - ➔ Capacité du refroidissement à eau largement supérieur ?

Next step

- Sous traitance de la conception de prototype de radiateur
 - Evaluation des sous traitants, savoir faire/coûts
- Maîtrise des outils de CFD en interne
 - Confirmation du modèle PCIe40 avec mesures in situ
 - ➔ Confirmer les conditions réelles de fonctionnement PCIe40
 - T° ambiante
 - Flux d'air en LFM
 - T_j constatée
- Choix planning pour prototypage radiateur
 - Mise au point rapide prototype radiateur + test avec carte « fantôme » émulant le TDP du FPGA ?
 - Retarder la conception du radiateur pour le prototype PCIe400 ?

Back up : Extrapolation dimension FPGA



- Déduction de la dimension du IHS (contact pad) par extrapolation des dimensions externes du boîtier
 - Ex sur boîtier Stratix 10 GX et SX (1SG210HUF50 et 1SX210BHUF53)

➔ Erreur relative < 1 % à 3 %

- Agilix AGMF039R46B
 - ➔ Boîtier WxL 56x66mm
 - ➔ IHS WxL 49x52mm

- Déduction de l'épaisseur du boîtier

- Ex sur boîtier Stratix 10 GX et SX (1SG210HUF50 et 1SX210BHUF53)

➔ Même épaisseur

- Agilix AGMF039R46B

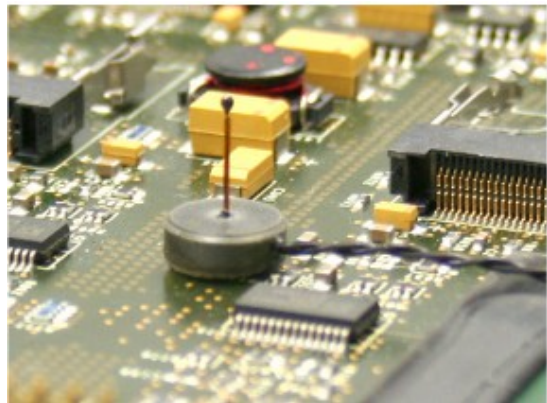
➔ Épaisseur globale (avec BGA avant reflow) 3.529mm

➔ BGA → -33 % sur 0.429mm

Back up : Mesure du flux d'air in situ

Anémométrie

- Plusieurs techniques :
 - **Hot-wire / Hot-film** : mesure par transfert de chaleur calibré
 - Pitot tube : Equation de bernouilli pour déduire la vitesse par mesure de prèssion statique et dynamique, adapté pour les hautes vitesse de flux d'air
 - Laser doppler velocimetry : nécessite l'ajout de particules dans l'air et de laser/photodiode pour la mesure de diffusion
- Principe du **Hot wire anemometry (HWA)** (plus d'info [ici](#) et [ici](#))
 - Maintien du fil chaud à température constante ($\sim 150-250^{\circ}\text{C}$), **mesure de la puissance délivrée pour compenser la dissipation** due au flux d'air, requiert une **mesure de température ambiante** pour évaluer le **coefficient de convection du fil chaud vers l'air**



Solution
d'instrumentation
ATS



Solution
d'instrumentation
DEGREE CONTROLS
INC.

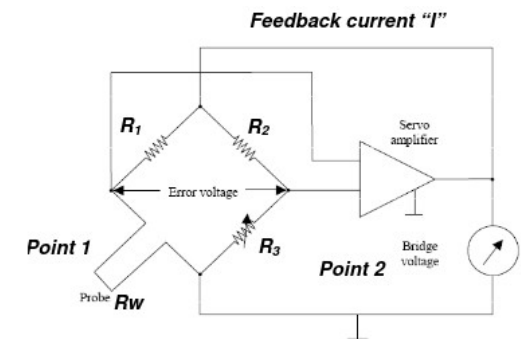
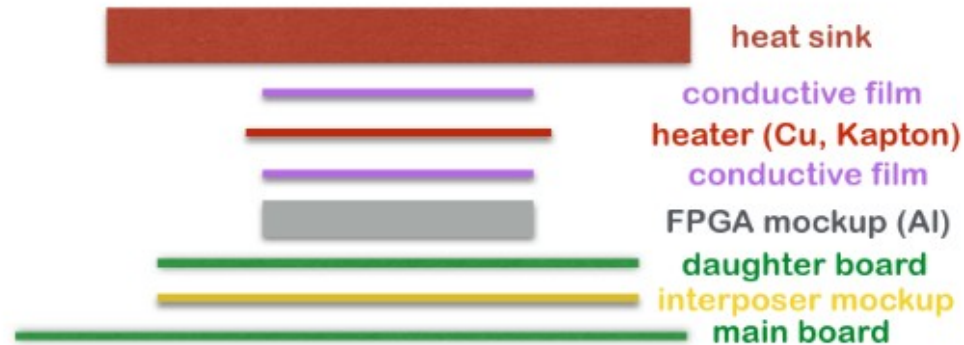
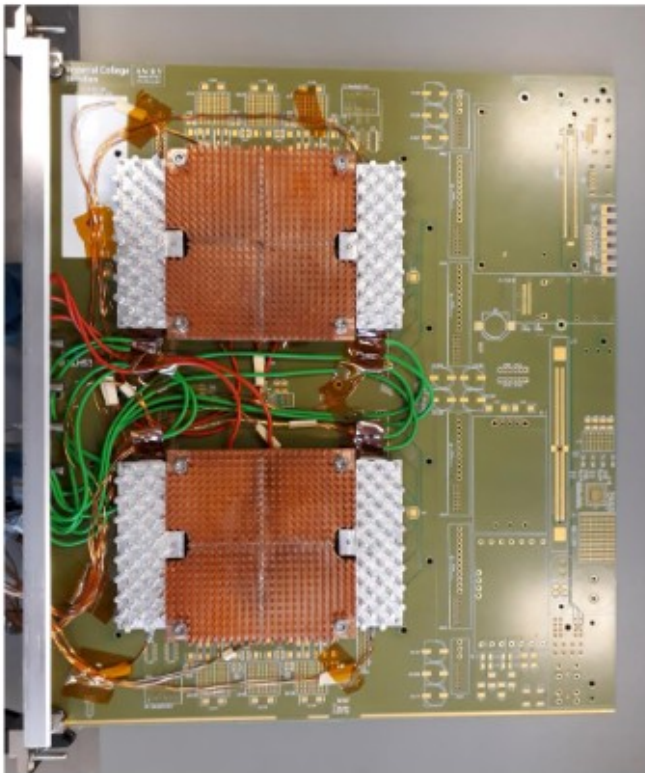
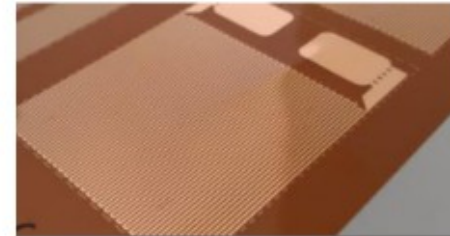


Schéma de principe fil chaud

Back up : carte fantôme

CARD MOCK-UP

Kapton heaters made at CERN



6

Giacomo FEDI (INFN), Gregory Iles (ICL) @TWEPP 2019

Maquette équipée avec radiateur prototype

- Émulation FPGA via bloc aluminium
- Émulation puissance dissipée via résistance chauffante en Kapton

Vérification simulation CFD