

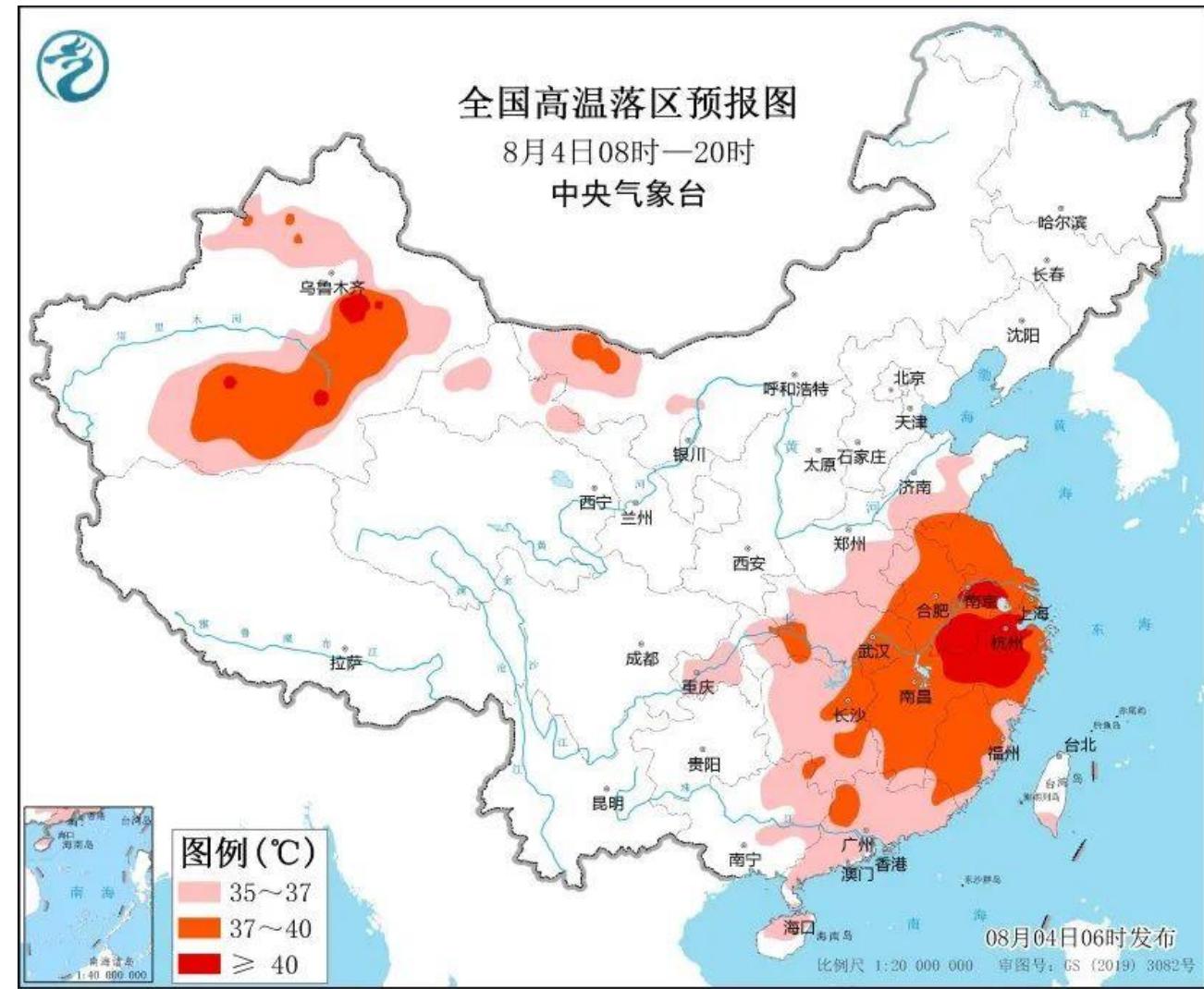
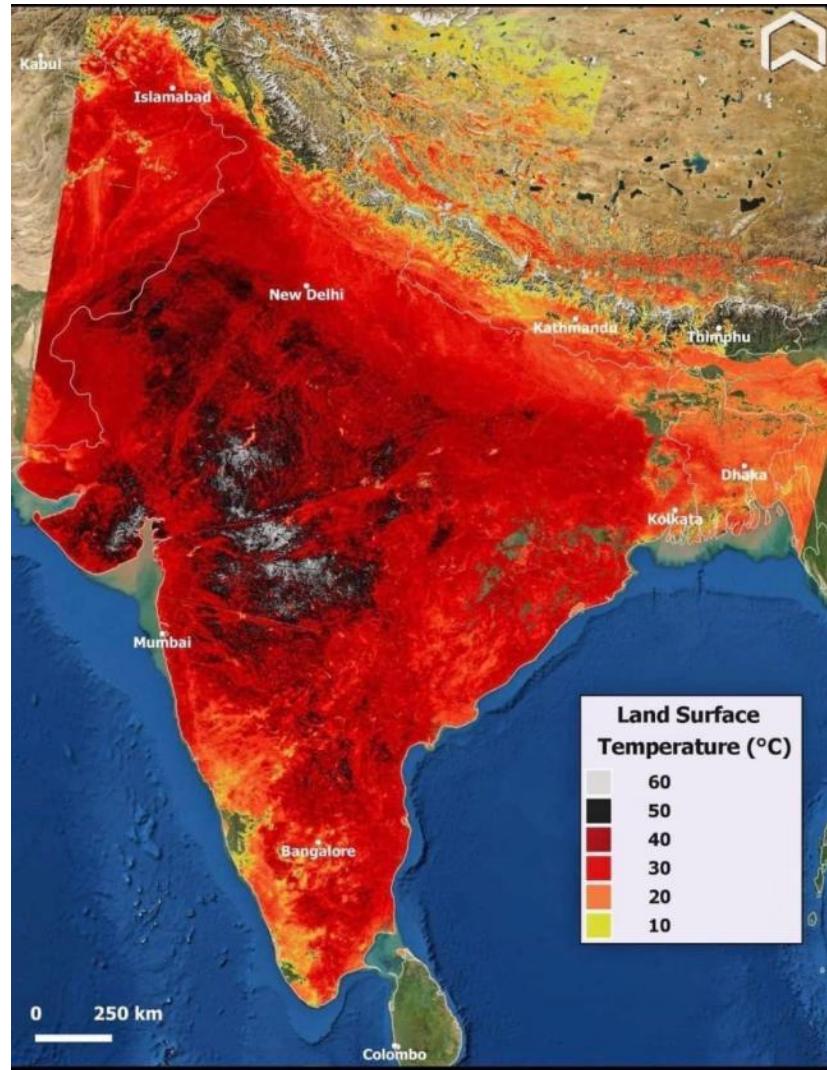


Question

1. 亲耳听过大规模风冷服务器的噪音吗？
2. 你知道风冷中的空调等设备占了多少电力吗？
3. 为啥风冷仍然是当前 DC 的主流？
4. 关于 L0 层投资电力还是液冷？



感受到印度和杭州人们的热情!



Question

- 太热啦受不了，那么怎么办？

1. 服务器丢到大山沟里面！
2. 服务器丢到北极上面！
3. 服务器丢到大海里面！



服务器丢到大山沟里面

- 蘋果中國（貴安）數據中心是蘋果公司在美國和歐洲之後設立的第三個數據中心
- 以“算”提“质” 中国移动内蒙古公司加速培育发展以算力为代表的新质生产力



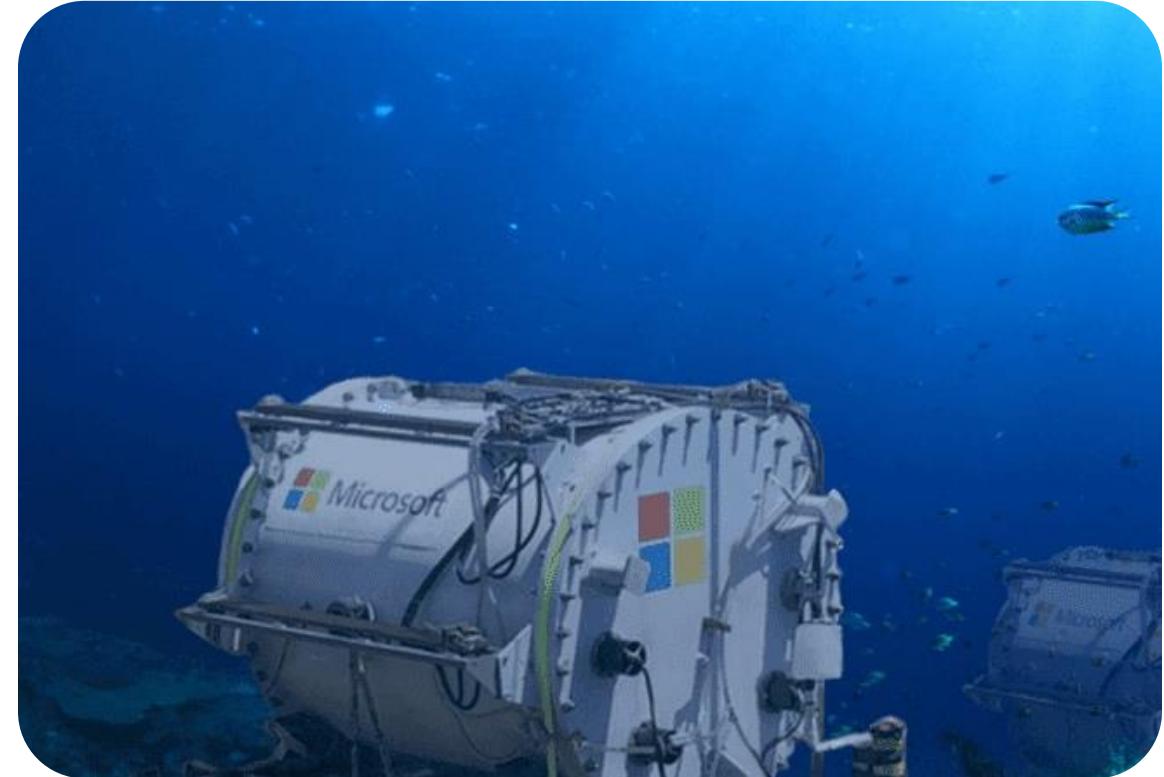
服务器丢到雪山上面

- [Take a close look into Google's Finnish data center](#)



服务器丢到大海里面

- Are underwater data centres the future?



Question?

- 至于吗？给服务器散个热，能花几个钱？

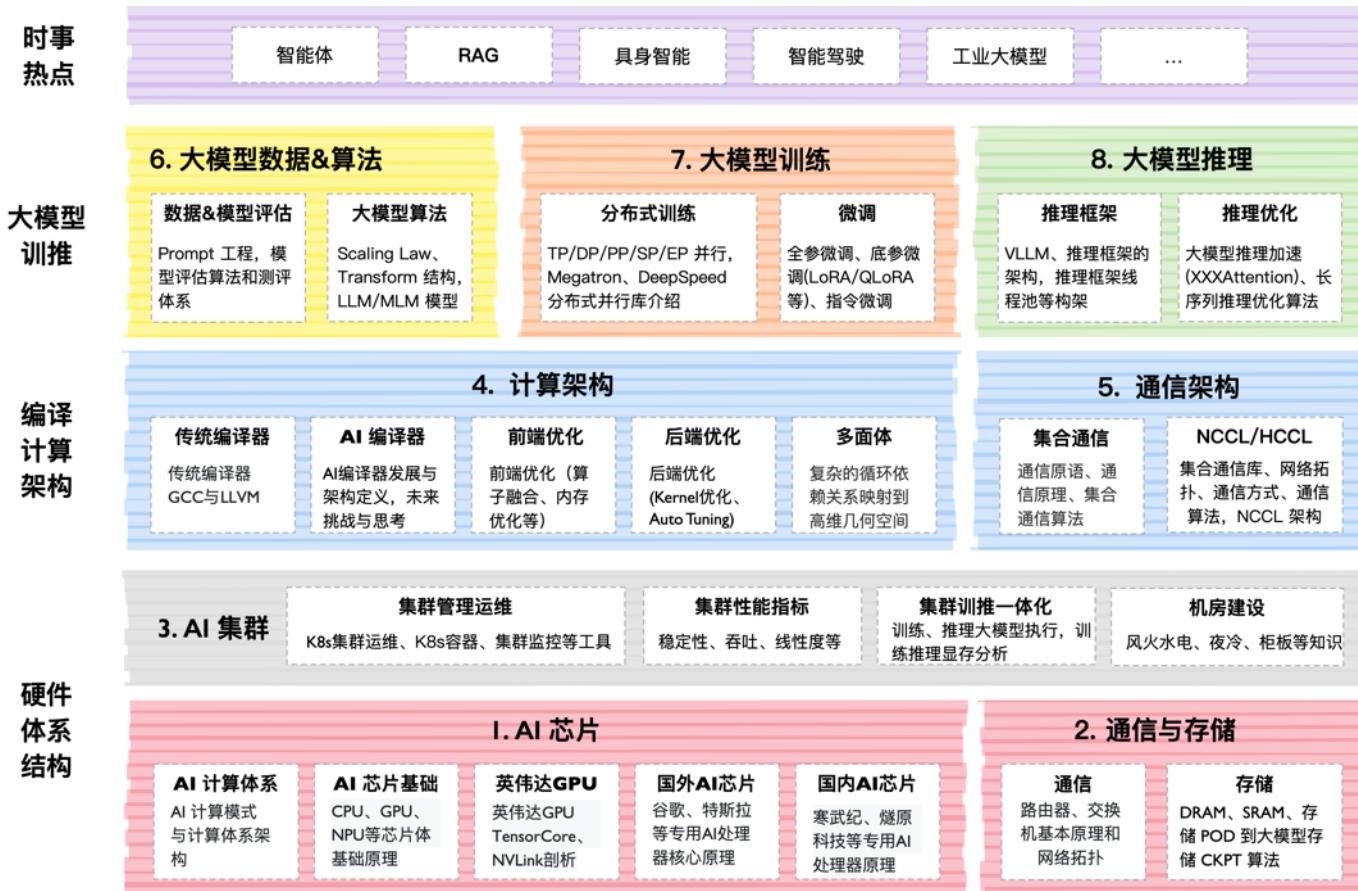


Content

1. 服务器散热有多烧钱
2. 风冷挑战与设计原则
3. 服务器风冷优化实践



Content github.com/Infrasys-AI/AIInfra



Content

L0 集群基础建设



01

服务器散热 有多少烧钱



概念解析

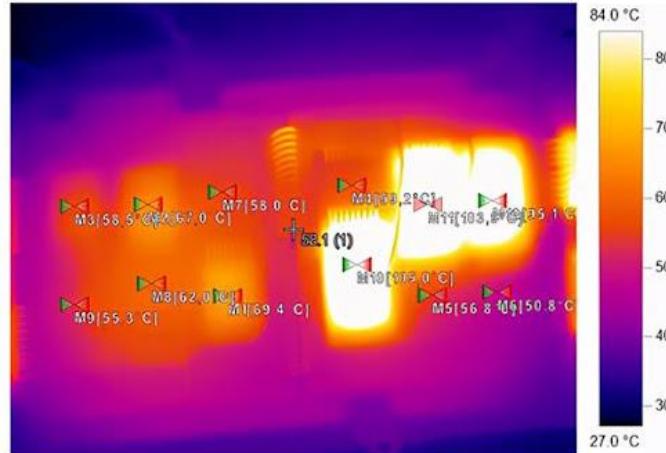
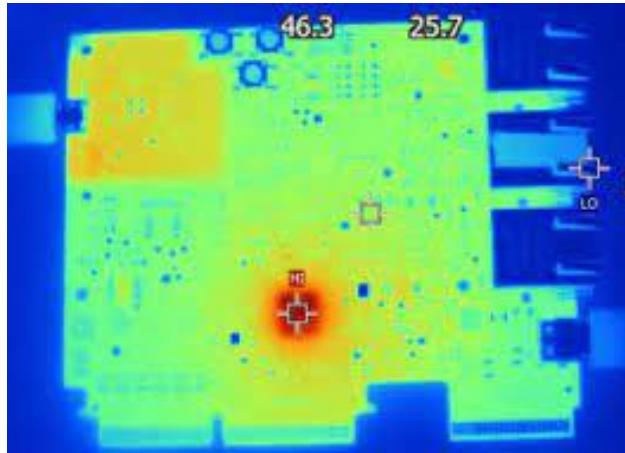
• 什么是 PUE?

- PUE, Power Usage Effectiveness 电能使用效率，衡量数据中心能源效率关键指标。
- $PUE = \text{数据中心总能耗} / \text{IT设备能耗}$
- PUE值越接近1.0，说明DC能源效率越高，用于制冷和供电等非IT设备开销越少。



概念解析

- **什么是结温？**
- 结温 Junction Temperature, 记为 T_j , 半导体器件内部最热点温度, 特别是晶体管PN结处温度。控制结温是电子设备设计和散热管理的核心任务。
- 衡量芯片内部热量积累和散热效果关键参数, 对于确保芯片的稳定运行和可靠性至关重要。



服务器耗电比例

- 数据中心 DC 用能部分为 IT 设备（CPU & GPU & NPU & 交换机）、制冷系统、供配电系统、照明及其它。
- 其中制冷是 IT 设备外最大耗能源。e.g. PUE=1.5 DC 为例，IT 设备耗能约占 67%，制冷系统约占 27%，供配电系统约占 5%，照明及其它约占 1%。



中美数据中心情况

- 截至 2023 年 12 月，全球约有 8000+ 数据中心 DC。其中美国拥有 30%，比世界上任何其他国家都多，这意味着世界上大约一半 DC 都位于美国。在美国，预计 DC 电力消耗量将从 2022 年的 200TWh 增加到 2026 年的 260TWh，约占全国总用电量的 6%。
- 2023 年全球 DC 电力消耗量为 460 太瓦时 TWh，比 2022 增长约 15-19%。平均每个 DC 能源消耗相当于 25,000 户家庭能源。2024 年 DC 约占全球用电量 1.5%，500 太瓦时。



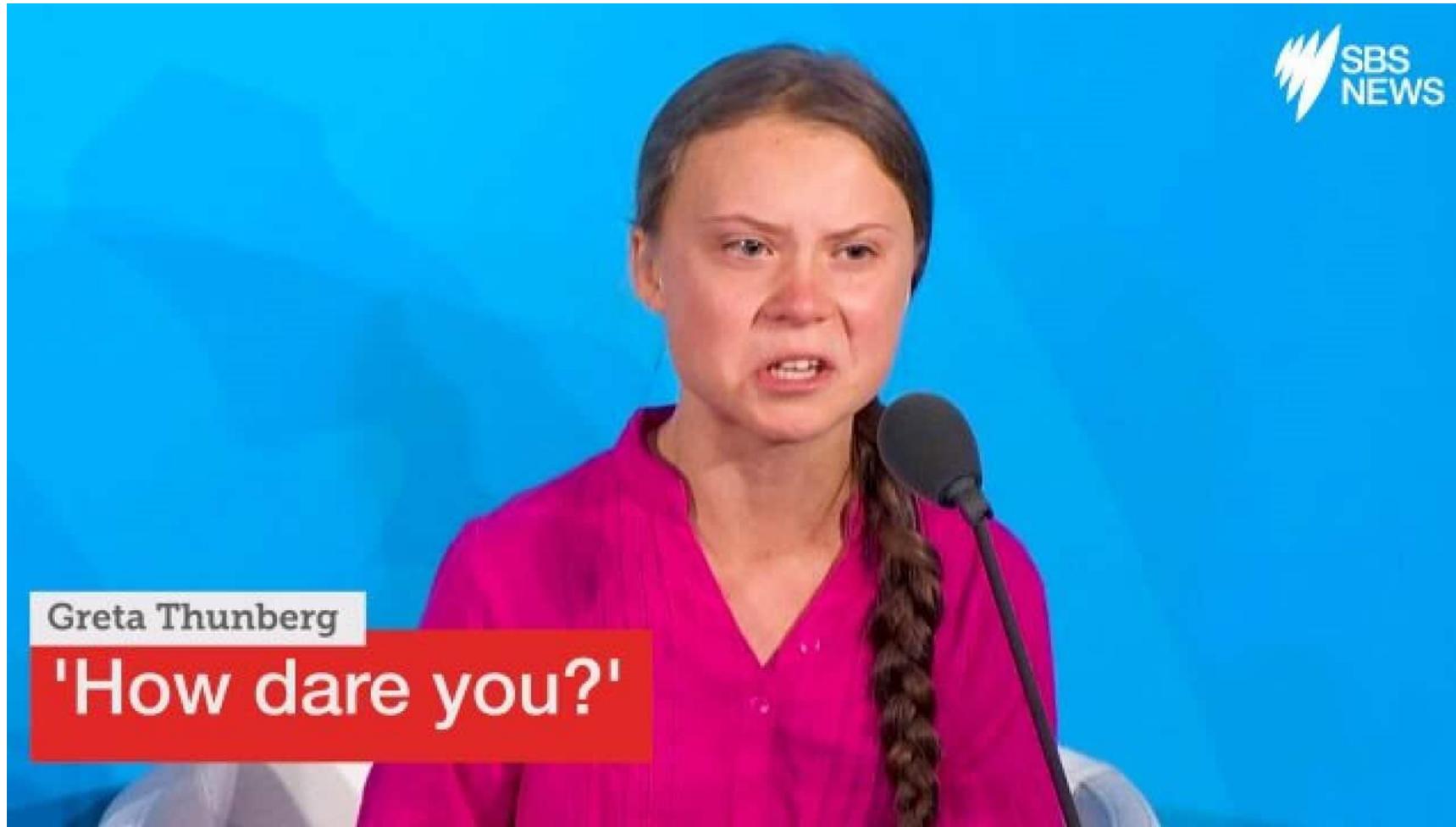
中美数据中心情况

- 中国以 449 个 DC 数量位居亚太之首。国际能源署 (IEA) 最新报告显示，2024 年中国 DC 用电量已占全球 DC 用电总量 25% (1000 亿度电)，仅次美国全球第二大电力消耗国。

PUE=1.5 -> 全年的耗电量 -> 三峡



环保少女



害怕环保少女指责

南贵 vs 北鸟



服务器丢到大山沟里面

- 蘋果中國（貴安）數據中心是蘋果公司在美國和歐洲之後設立的第三個數據中心
- 以“算”提“质” 中国移动内蒙古公司加速培育发展以算力为代表的新质生产力



02

服务器风冷 挑战与设计原则



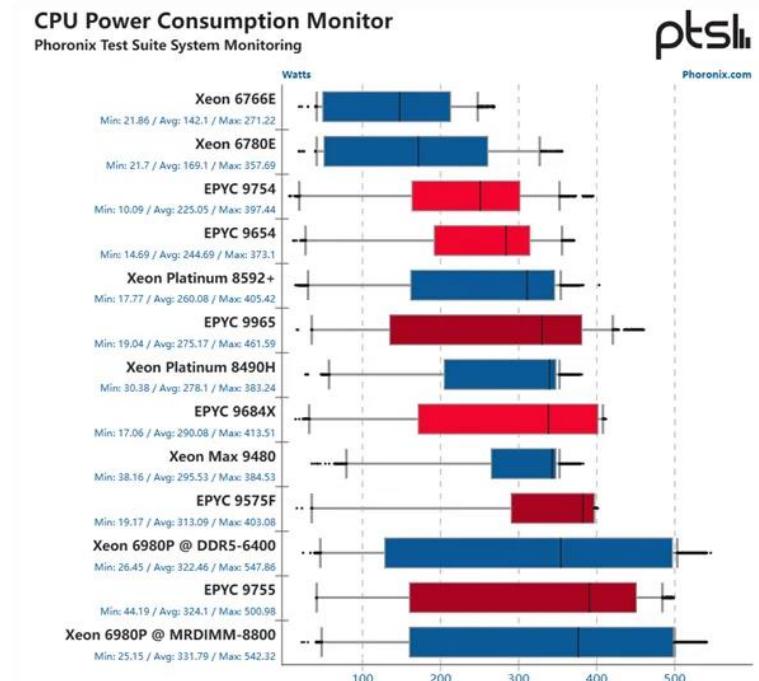
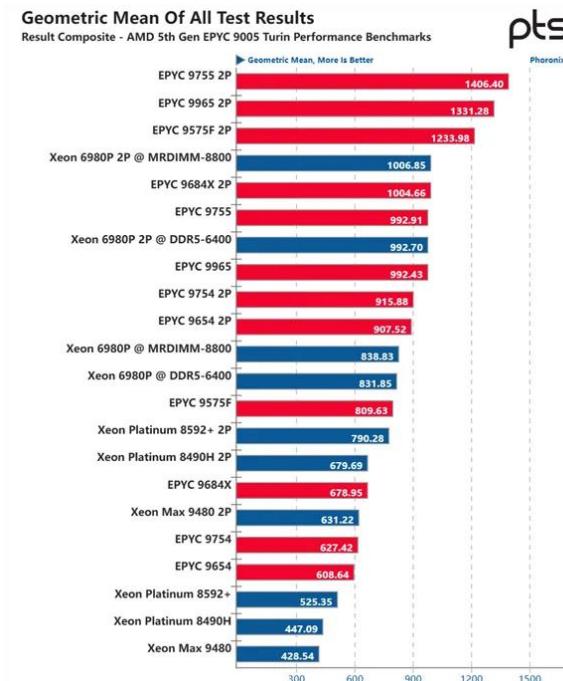
风冷技术优点

- 效率高：**风冷散热利用风扇将空气不断循环，使得热量能够迅速被带走，从而达到降温效果。相对传统的水冷散热技术而言，风冷技术的效率也更高。
- 成本低：**在采用传统的水冷散热技术时，需要额外投资建造冷却塔或冷却设备，大幅增加了系统的成本。风冷只需安装相应风扇、空调等散热设备即可，不需要新增额外投资。
- 可靠性高：**风冷系统中只需要安装风扇、空调等设备，液冷中较为复杂的管路、阀门等设备可减少或免去，从而减少可能存在故障点，提高了系统可靠性。
- 环保：**相比起水冷散热不会产生额外水污染，而且也不需要大量消耗能源。同时，由于无需使用任何化学药品，对环境也不会造成二次污染。
- 易维护：**结构上简单，维护工作主要集中在定期清洁和风扇替换上，减少对专业维护人员依赖。部件易于更换，降低意外故障导致停机时间，对提高服务器的整体可用性和业务连续性具有重要作用。



风冷场景挑战

- **峰值算力:** 释放芯片&服务器的峰值算力，在缺乏先进制程下提升频率，对标友商 xx% + 性能
- **性能达标:** 系统突破 xxxFLOPS/xxxW 算力，功率密度突破 xxxW/cm² 散热
- **系统能效:** 缺乏先进制程，系统存在 xx% 功耗 GAP，能效存在 xx% GAP



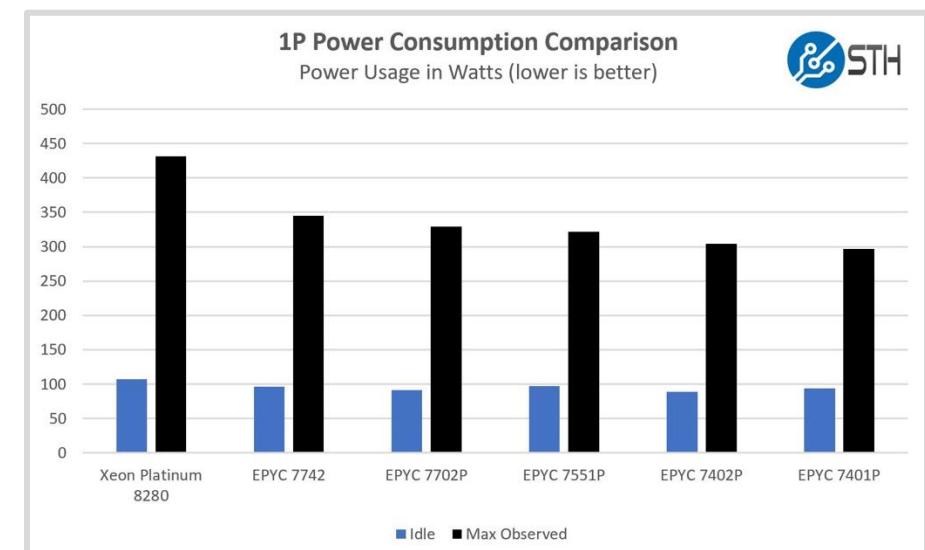
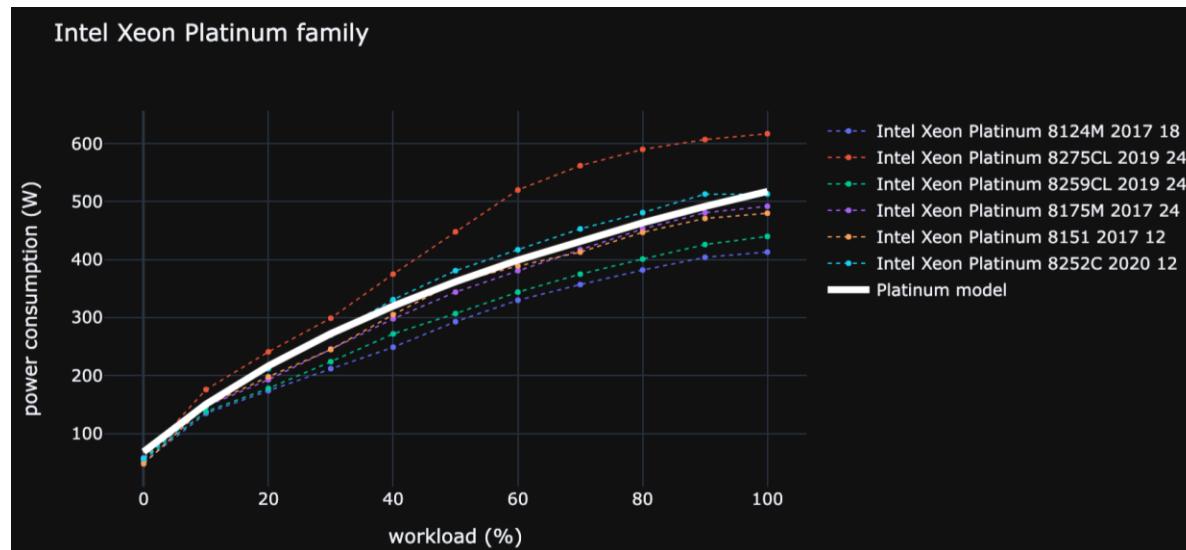
风冷场景设计原则

- **设计原则：**
 - 峰值性能场景以性能为中心，释放 CPU 极致算力；
 - 主流业务场景以能效/性价比为中心，平衡性能/功耗/成本/规格诉求，确保性价最优；
- **设计思路：**基于峰值性能、系统能效及性价比构建商业场景竞争力



风冷场景设计思路

- **设计思路：**基于峰值性能、系统能效及性价比构建商业场景竞争力
 - **峰值性能：**风冷场景通过散热组件（风扇、LAAC/VCE散热器、导热垫）来提升效率，进行风道优化；
 - **系统能效：**结合动态CPU节温和风扇调节，降低系统动态功耗，达到系统能效最优；



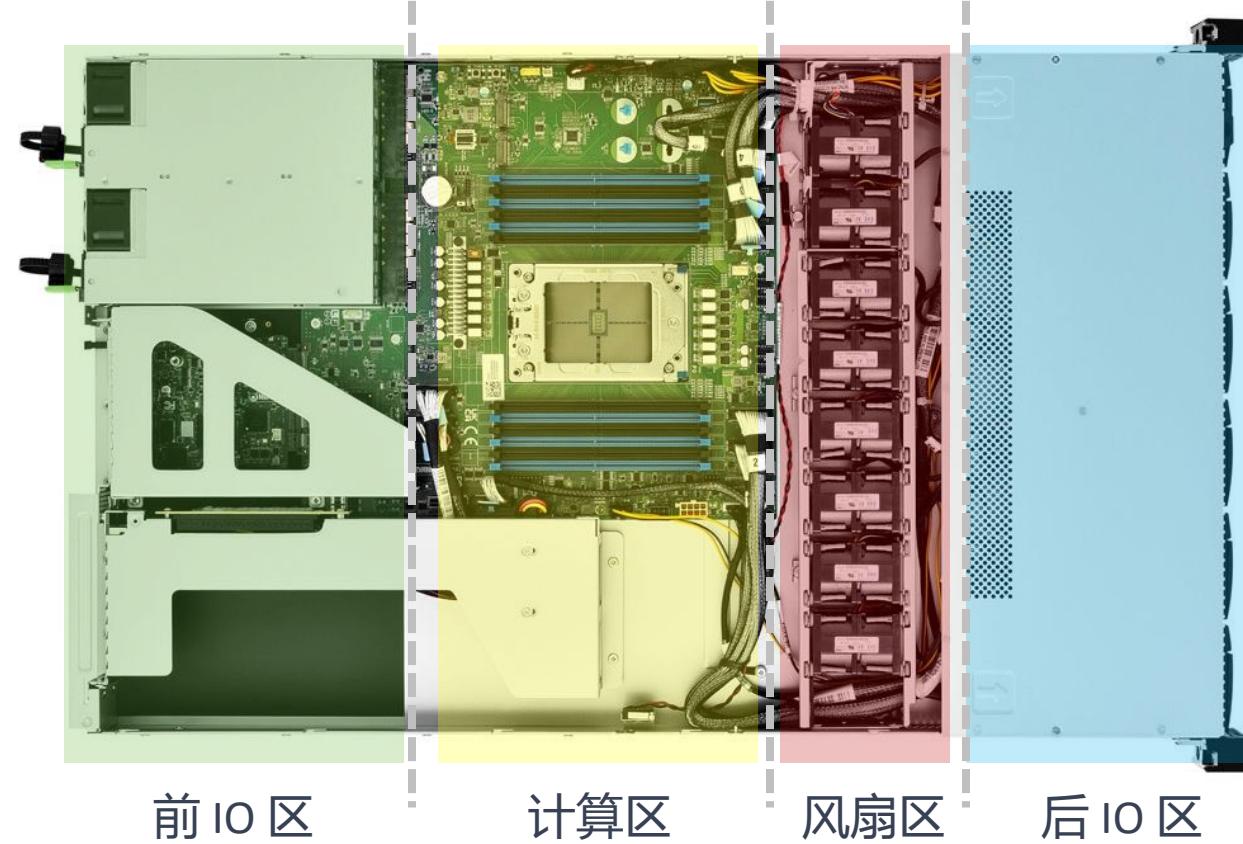
03

服务器 风冷实践



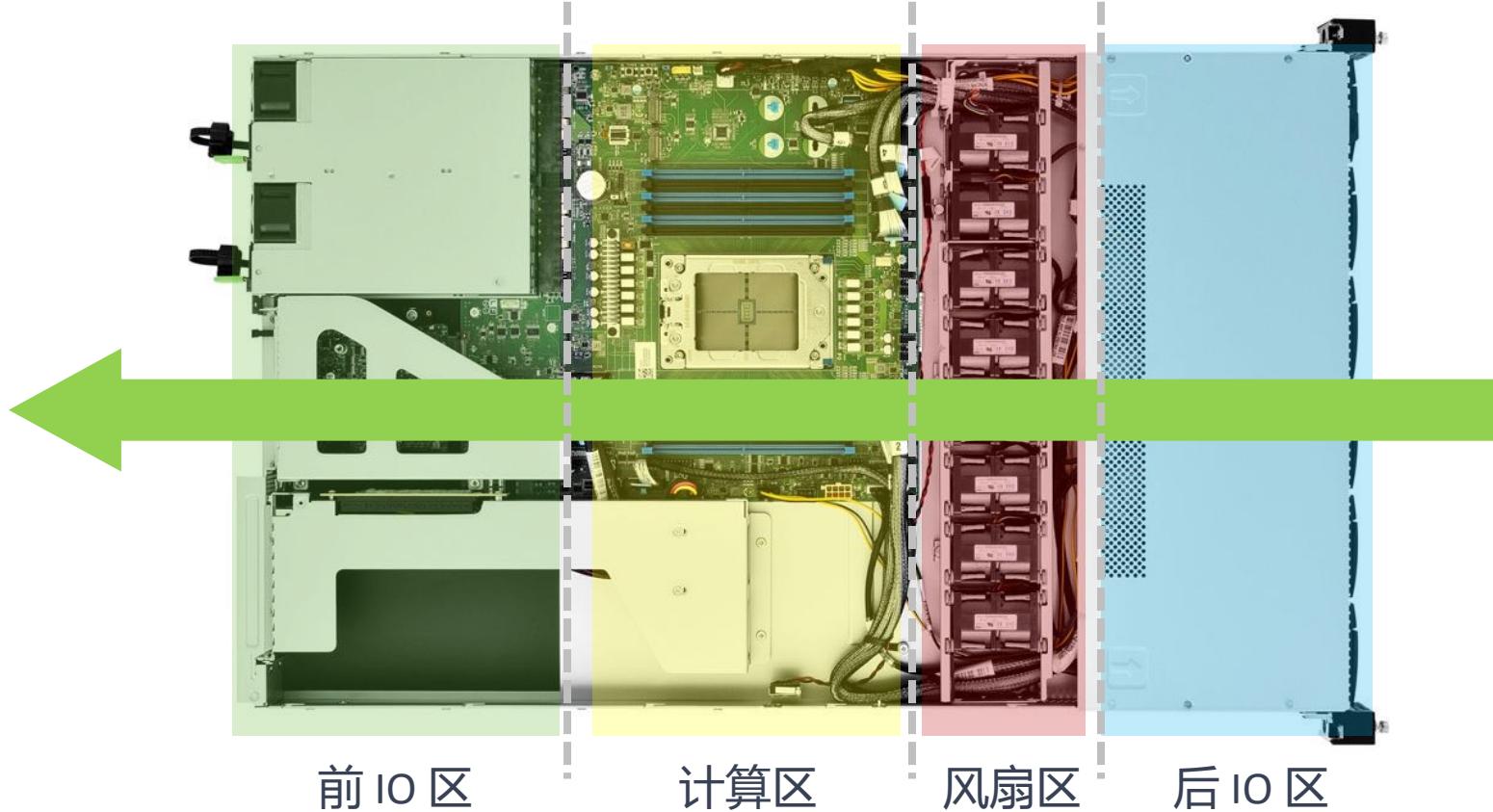
1、系统级风道优化

- **系统级优化：**增加开孔率、风道优化/管理等风阻设计，多段式优化设计，2/4U 整机 xxxxW 散热。
- **提升峰值性能：**通过风道/风量优化，修改散热器/换热器，热从芯片带到系统，支持 CPU/NPU/GPU 算力释放



1、系统级风道优化

- **系统级优化：**增加开孔率、风道优化/管理等风阻设计，多段式优化设计，2/4U 整机 xxxxW 散热。
- **提升峰值性能：**通过风道/风量优化，修改散热器/换热器，热从芯片带到系统，支持 CPU/NPU/GPU 算力释放



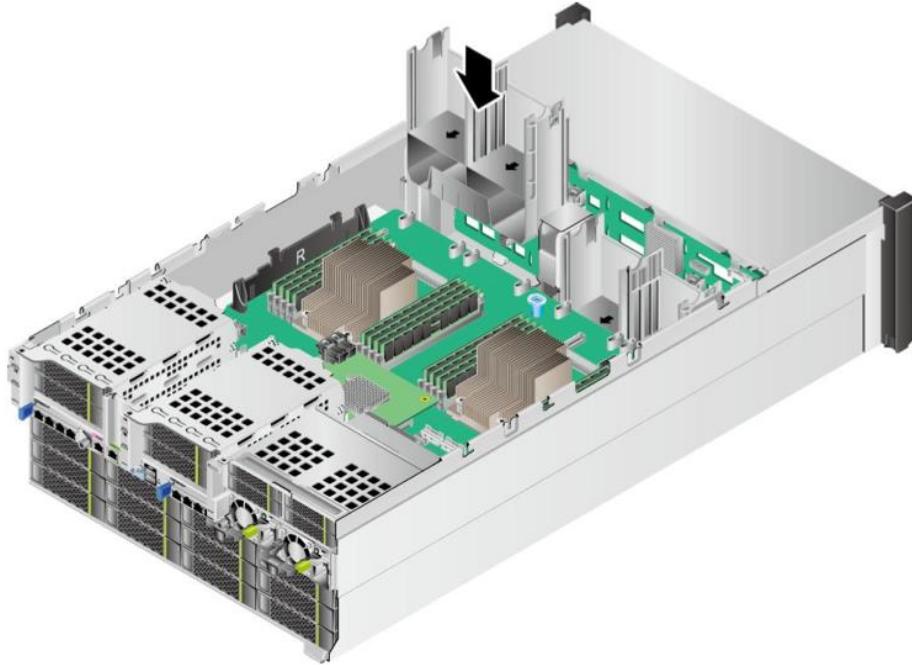
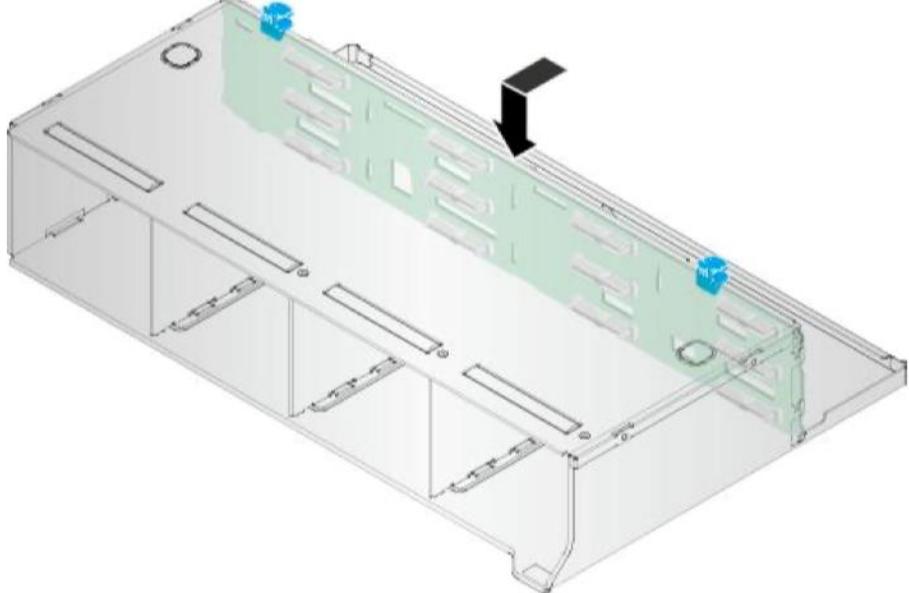
2、背板架构优化

- **挑战：**
 - CPU/MEM/SSD/DPU/GPU 往大功耗演进；
 - 传统背板架构无法提供足够风量支撑大功耗散热；
 - 系统功耗大量浪费在散热上，散热效率低，工艺受限；
- **策略：重构背板优化系统风量/风道，使风进得来出得去，提升系统散热能力和能效**
 - 后置盘入风口温度高散热差，导致系统风扇高转速，系统功耗增加，入风口修改
 - 后置背板竖直背板->水平背板，最大化开孔率，提升等效开孔率，开孔率 XX% +
 - 风扇功耗降 XXW，盘散热能力由 XXW->XXW，噪声降低 X.X dBA，系统散热能力提升 XX%



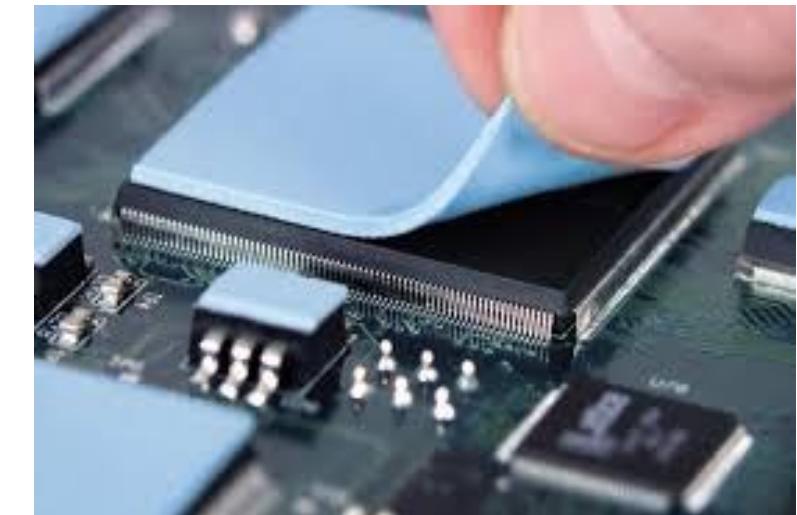
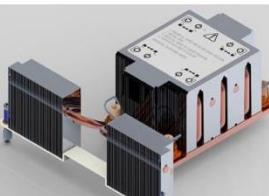
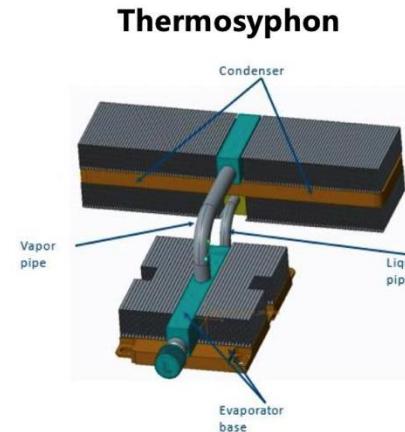
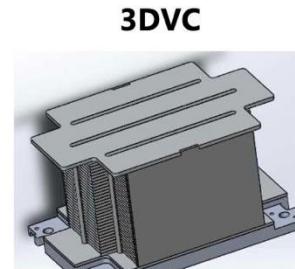
2、背板架构优化

- 策略：重构背板优化系统风量/风道，使风进得来出得去，提升系统散热能力和能效；
 - 后置盘入风口温度高散热差，导致系统风扇高转速，系统功耗增加，入风口修改
 - 后置背板竖直背板->水平背板，最大化开孔率，提升等效开孔率，开孔率 XX% +
 - 风扇功耗降 XXW，盘散热能力由 XXW->XXW，噪声降低 X.X dBA，系统散热能力提升 XX%



3、散热组件动态结温控制

- 挑战：
 - 工艺受限，芯片 & 服务器友商整机存在 XX%+ 能效 GAP
- 策略：工程补工艺，整机散热、供电优化，缩小能效 GAP，做到能效持平/领先
 - 分级散热组件提升系统散热效率，采用高效散热方案
 - 减少系统风扇功耗、提升散热效率，效降低芯片结温



3、散热组件动态结温控制

率密度	部件种类	散热方案
高密度发热器件	CPU/GPU/NPU/DPU	Heat pipe 热管、Vapor Chamber 均热板 EVAC 增强型蒸汽腔空气冷却器 Thermosyphon 热虹吸管 3DVC三维真空腔均热板
低密度发热器件	CPU/DIMM VR, PCH, NIC, Memory chipset(HBM)	铝挤HS
<3W 发热器件	BMC	N/A

Heatpipe HS



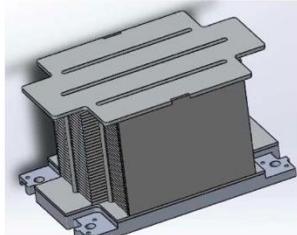
VC HS



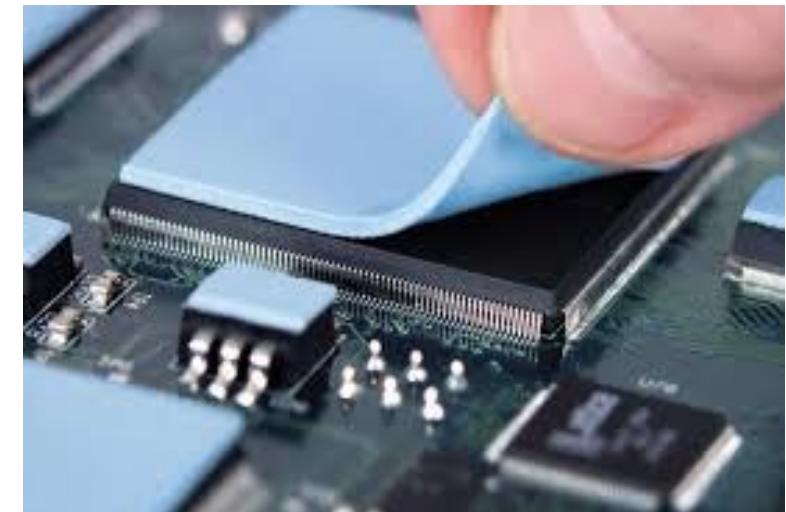
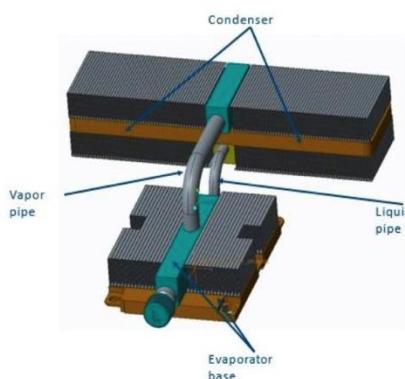
EVAC



3DVC



Thermosyphon



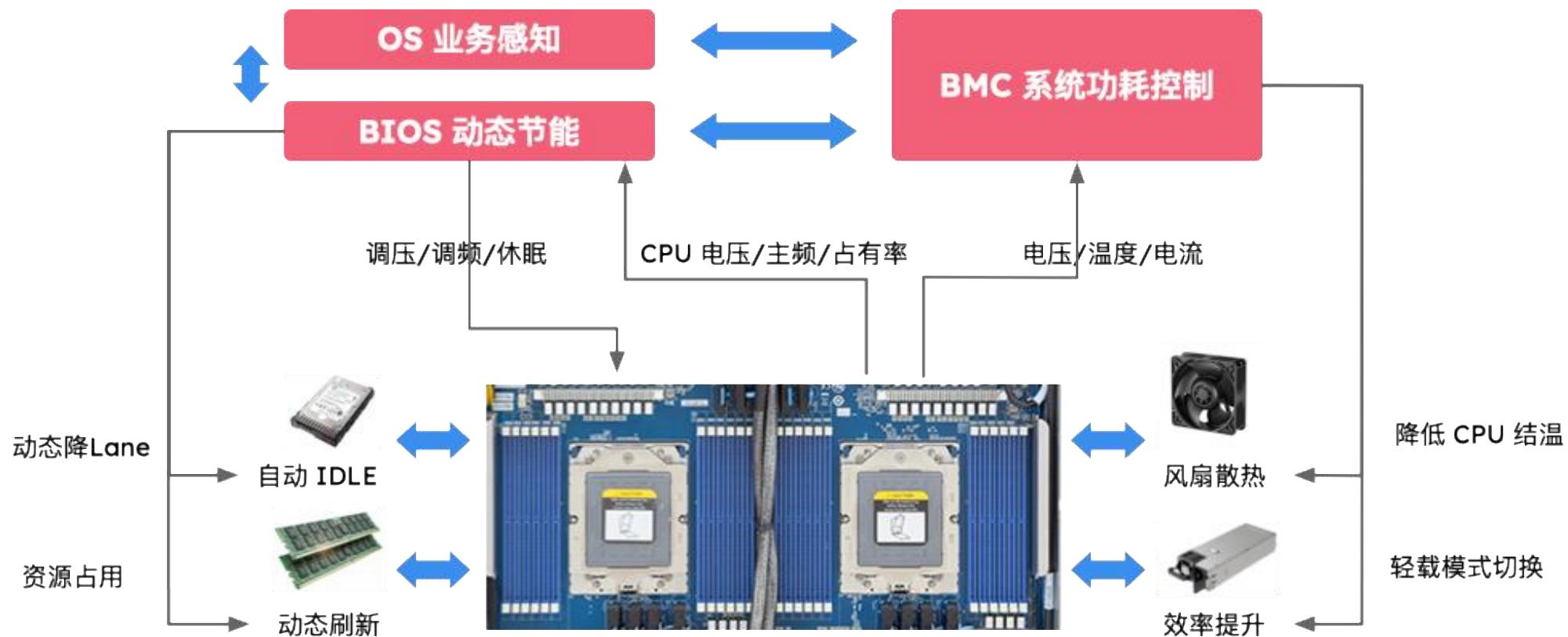
4、场景负载分析

- **设计目标：**分级能效管理，达成系统级最优
- **系统能效：**场景化能效最优
- **业务负载场景：**基于实际业务建模进行能效寻优（包括散热效率提升，系统风道优化，动态 CPU 节温，风扇调节），实现典型场景智能化节能
- **低负载场景：**通过动态休眠、动态调频/调压、降Lane降速、动态关断等技术，达成零负载低功耗 (<5%)；



4、场景负载能效管理

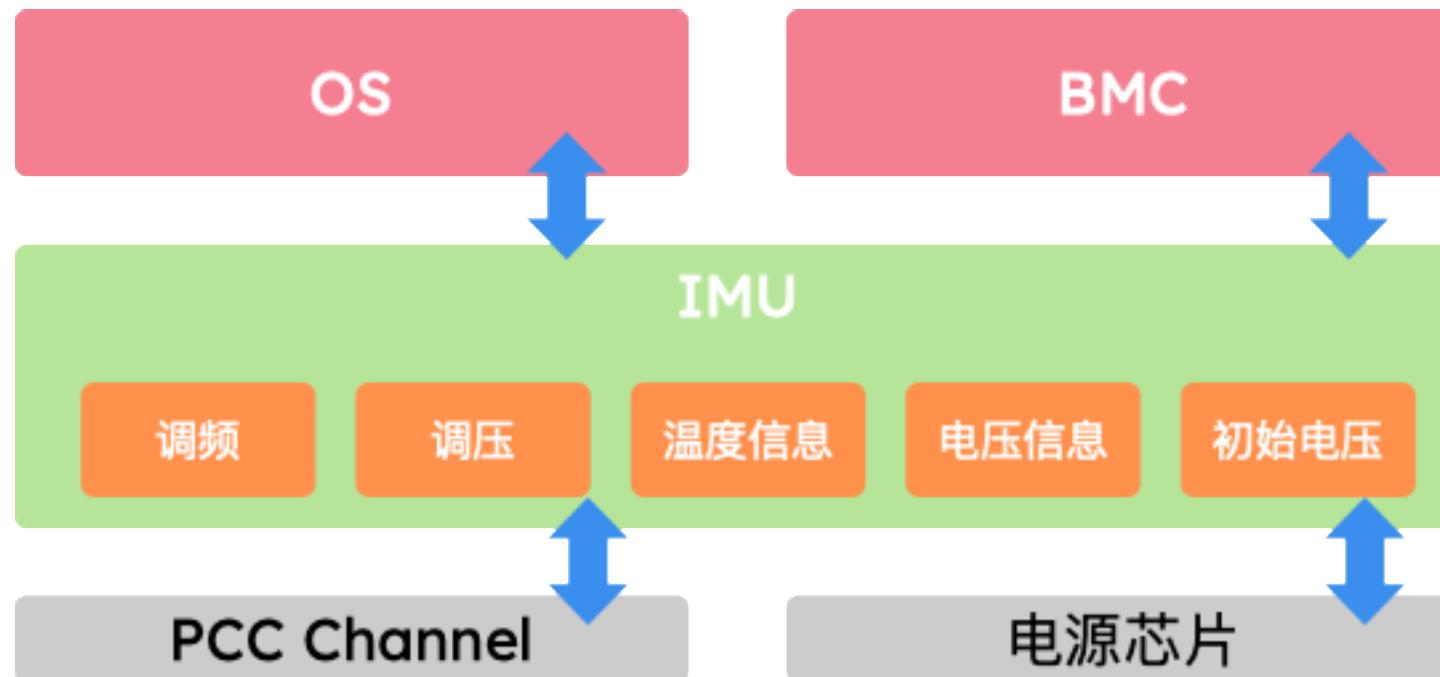
- **低负载场景\业务负载场景：**业务、管理、硬件多级联动进行能效管理



4、软硬件协同智能节能

- **板级节能：**

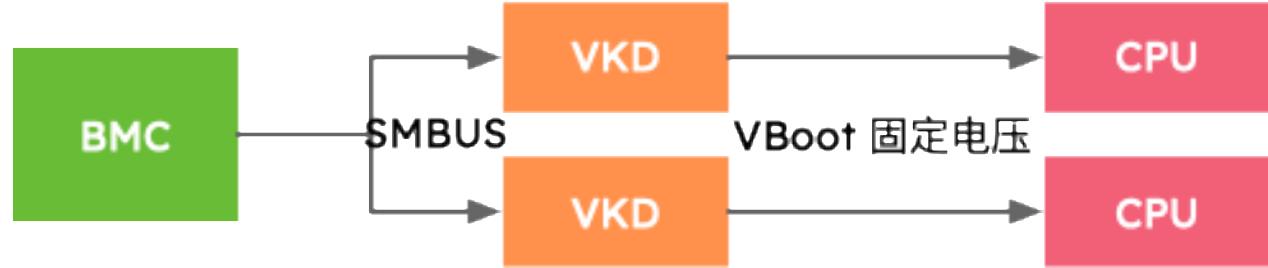
- AVS/Load line：IMU 获取 CPU HPM 信息计算芯片工作电压，一芯一压；
- 板级DVFS：IMU 根据 CPU 实际工作主频，动态调整 CPU 电压，实现动态调压；



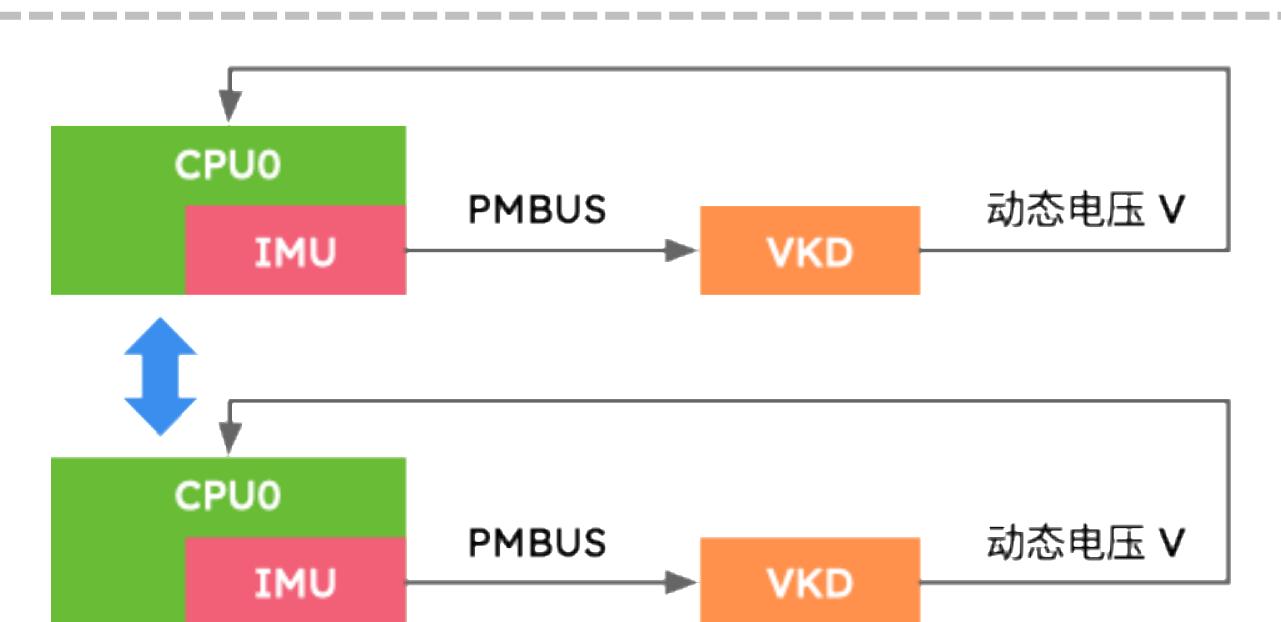
4、软硬件协同智能节能

- 系统节能：

- 联合 OS 优化调频算法，业务感知优化调度算法；
- 系统风扇调速算法优化，MPC 取代传统 PID；
- 软件控制框圈，硬件根据负载调整电压&主频；



传统方案：固定 Vboot 电压供电



优化方案：按工作主频动态调整电压



总结与思考



大型AI DC 机房分区示意图

风冷区（通算业务）

风	风	空	风	风	风	空	风	风	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷
风冷												
风	风	空	风	风	风	空	风	风	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷

风液混合区（通智混合业务）

液	液	空	液	液	液	空	液	液	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷
风液比2:8												
液	液	空	液	液	液	空	液	液	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷

风	风	空	风	风	风	空	风	风	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷
风冷												
风	风	空	风	风	风	空	风	风	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷

液	液	空	液	液	液	空	液	液	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷
风液比4:6												
液	液	空	液	液	液	空	液	液	空	风	风	风
冷	冷	调	冷	冷	冷	调	冷	冷	调	冷	冷	冷



曾经奶奶少的 ZOMI....



Question

- 大功率芯片 + 高密机柜，未来散热走向何方？
- 深挖动态调频/调压、MPC 风扇调速及 OS 调频算法优化等多路径功耗寻优；
- 联合业务进行场景化分析（分析业务负载），软硬协调识别关键功耗点；
- 结合应用精准设计各负载下的调频和调压甜点；
- 说白了，针对特殊业务进行 hardcoded！！！！





Thank you

把 AIInfra 带入每个开发者、每个家庭、
每个组织，构建万物互联的智能世界

Bring AI Infra to every person, home and
organization for a fully connected,
intelligent world.

Copyright © 2025 [Infrasys-AI](#) org. All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. [Infrasys-AI](#) org. may change the information at any time without notice.



ZOMI

GitHub github.com/Infrasys-AI/AIInfra

Book infrasys-ai.github.io



ZOMI

41

引用与参考

1. https://www.bilibili.com/video/BV1TUjepEcd/?spm_id_from=333.337.search-card.all.click&vd_source=26de035c60e6c7f810371fdfd13d14b6

PPT 开源在：<https://github.com/Infrasys-AI/AIIInfra>

