

智算中心相关的10类知识点

1、和传统数据中心一样吗？

从数据中心发展趋势来讲，智算中心和我们经常所说的的数据中心，算力中心，超算中心等，其实并没有本质区别，换了一个时髦名字而已。

智算中心作为新型数据中心在业务特点，算力种类，算力规模，机柜密度，PUE要求（液冷应用），网络带宽，存储种类等有新的变化。

网上非常多的智算中心的文档，只是视角不同，重点强调的技术领域有差距而已，比如做服务器强调算力，做网络的强调组网等。

2、更多元的算力（智算-ai需求增多）

类型：智算中心规划除了通用算力（cpu）外，主要新增了ai算力。各类ai芯片成为主角，gpu，dpu，npv，fpga等，英伟达一家独大，国产卡百花齐放。

需求：chatgpt为代表的aigc产业极大的推动了智算中心的建设。尤其是对ai算力卡的需求，大模型的训练和推理都依赖强大的ai算力支撑。

现状：受米国芯片禁令影响，nv卡的供应存在不确定性，以华为昇腾为代表国产卡迎来重大机遇，但受限于技术和产能的影响，还需突破。

3、更快速的网络（IB成为主流）

趋势：大模型需要大算力，大算力需要超高速，低延时，无阻塞的网络。IB的优势明显，mellanox一家独大，200G，400G成为主流，组网成本非常高。

方案：除了IB之外。Roce组网也逐步被越来越多客户所接受。同样可以支持Rdma，但是价格比ib低，产品品牌和方案选择多（华为，华三，锐捷等都可以），组网和带宽和ib基本一致，但是延迟方面还有差距。

网卡：服务器需要配对应网卡，pcie 4.0的服务器平台支持最高到200G单口，最新的服务器spr和genoa 支持到了pcie 5.0，x16的网卡支持单口400Gb/s。

4、更海量的存储（并行，gds，海量，全闪）

算力和网络的升级，配套的存储同样需要，新建算力中心除了关注性能（IO）和容量外，还需要重点考虑GDS，通过GDS可以实现GPU直连存储的操作，性能达到最优，另外就是并行支持，和HPC场景类似，AI大模型由于数据量和过程中的数据流，通常需要存储支持并行操作。

市场上常见的分布式存储包括Lustre、GPFS、Ceph等，由于需要支持GDS和存储并行，Lustre和GPFS应用较多，从开源和商业化的角度，Lustre都支持更灵活，适合于大文件的连续读写，GPFS比较成熟，源于IBM，更适合于海量小文件，Ceph在通用场景中使用多（开源投入小，运维难度高）。

PUE是数据中心的总用电量除以IT设备的总用电量，这个值是恒大于1的（目前常见的数据中心往往在1.4-1.8的范围），随着双碳政策的推出，对不同地区的新建数据中心PUE有了明确要求，PUE要控制在1.2-1.25或以下范围，这个对于越来越高的单芯片功耗（CPU-400W、GPU-700W）越来越高，风冷将遇到瓶颈的情况下，液冷未来将是必然。

5、更低的PUE要求（双碳和高密算力的均衡）

液冷加速落地，三个典型液冷方案冷板式、浸没式和喷淋式，最标准化的是冷板式，从新服务器实施交付到旧服务器升级改造门槛最低，英伟达最近也放出消息新的DGX将采用液冷，无疑给液冷市场加了一把火。

6、智算中心的“投建运”概述（什么意思）

投：建智算中心的钱从哪来，几种情况，有钱-甲方全资（需求方）、钱不够-引入资方（金融类机构和公司）共同投资、没钱-找资方垫资。

建：谁来负责建设，主要是的方案的规划，设备的选项及实施交付等，确保先进性、可用性、易用性、易维护性等，可以是甲方、厂商和大集成商；

运：建成以后怎么运营（运营、维护、商业模式等），价值发挥或实现盈利，可以是自用、算力出租和混合模式（自用为主，富余的资源进行出租）；

7、智算中心-产业链上中下游

上游（技术制高点）：AI芯片类等配套公司，也是最活跃的，英伟达、AMD、Intel和国内各个芯片，同时以液冷、微模块等数据基础设施的公司也是迎来业务增长爆发。

中游（产品、方案和资金主要玩家）：厂商和集成公司为主，像浪潮、曙光、超聚变等厂商，因为计算占比高。当然还包括网络、存储、安全和配套的软件公司，拥有垫资实力的集成公司或者金融类机构公司。

下游（买单者和使用方）：除了头部的BBAT（玩得大）外，更多的各地算力公司（云端集成公司）如雨后春笋爆发，鱼龙混杂，比如莲花味精大批量采购GPU服务器（专门智算公司），真正的使用者除了AI相关领域公司，也面向传统行业，尤其是对采购NVIDIA有风险的企业或者机构；

8、旧转新智算中心技术会遇到哪些问题？

硬件：旧服务器对GPU等AI芯片的扩展能力差，利旧困难，基本用不了，包括网络、存储以及配套的供电、散热都需要重新规划。

软件：智算中心往往需要增加AI相关软件和平台以适配新的业务，比如对GPU资源的管理调度，需要适配各类AI框架、架构和算法模型的环境。

人员：因为涉及管理和运维AI相关软硬件，智算中心的运维人员专业知识的赋能，面临运维成本的增加（培养和新招）。

9、旧转新智算中心需要考虑涉及哪些投入？

改造成本（TC）可以大致分为以下几部分：TC=硬件成本（HC）+软件成本（SC）+运维成本（OC）+能耗成本（EC）+其他成本（MiscC）

硬件成本（HC）：HC=(服务器成本×服务器数量)+(存储设备成本×存储设备数量)+(网络设备成本×网络设备数量)+其他硬件成本：

软件成本（SC）：SC=操作系统许可证成本+AI平台和工具许可证成本+安全软件许可证成本+其他软件成本

运维成本（OC）：OC=人员培训成本+人员工资成本+日常维护成本

能耗成本（EC）：EC=(每千瓦时成本×总功耗×小时数)

其他成本（MiscC）：MiscC=网络带宽费用+安全设施投资+不可预见成本

10、智算中心未来预测，有哪些风险？

对于以租赁为目的智算中心或者自用率占比不高的，回本周期非常关键，往往需要在3-5年，会面临中途断租的风险。

Nvlink机器的保修风险，众所周知的原因，随着设备的使用，故障率会逐步提升，正常质保通道不畅通，风险很大。

国产AI卡真正的前景不清晰，厂商角度都在大吹特吹，但是市场未充分验证，对原厂的技术调优依赖强，风险不可控。

竞争激烈利润低，现在遍地开花，大投入，快建设，强竞争，市场优胜劣汰，不可控AI降温、芯片升级（A800可能不值钱了），1-2年内谁能坚持下来，不好说，洗牌是必然。