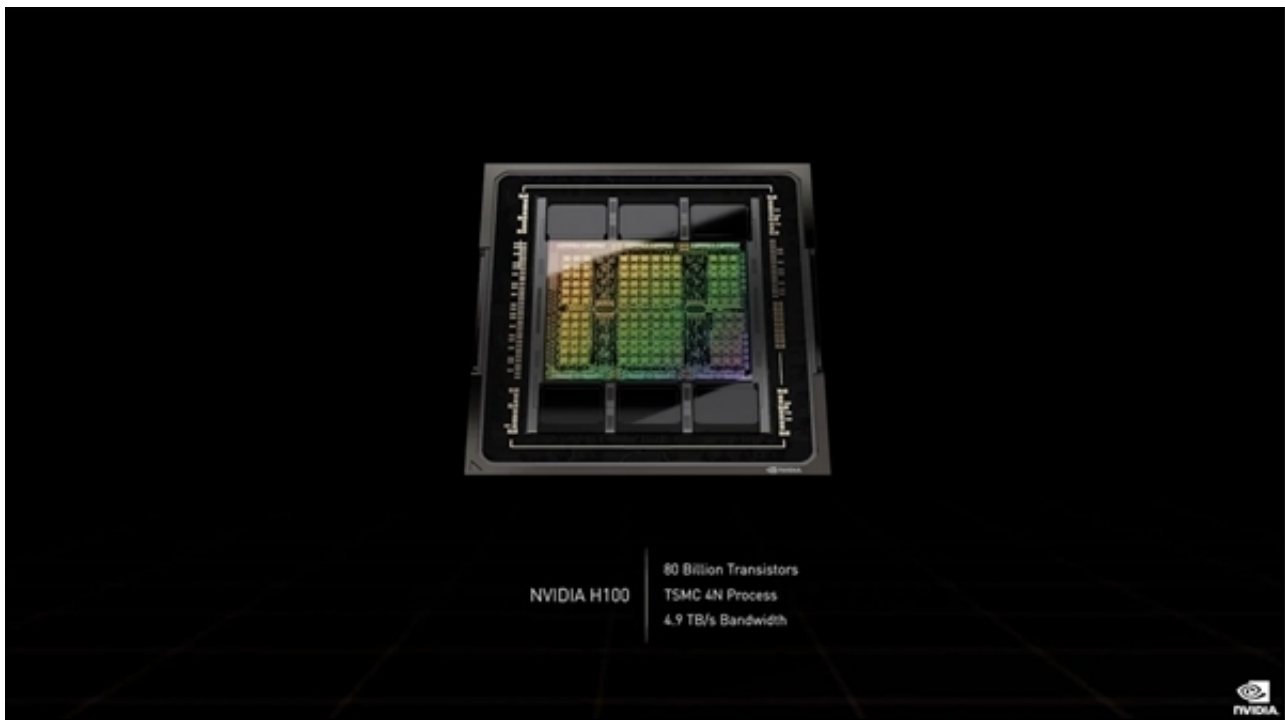


NVIDIA称H100显卡全面投产：700W功耗性能猛兽

本篇将为大家带来《NVIDIA称H100显卡全面投产》内容，在日前的GTC 2023大会上，黄仁勋宣布H100已经全面投产，并部署在多家客户的云计算服务中，包括微软的Azure、谷歌、甲骨文等客户数据中心。感兴趣的小伙伴们一起来看看吧，希望能够帮助到大家。



H100加速卡与2022年3月份的GTC大会上发布，采用Hopper架构，GH100大核心，台积电4nm制造工艺、CoWoS 2.5D封装技术，集成800亿个晶体管，核心面积814平方毫米。

它拥有18432个CUDA核心、576个Tensor核心、60MB二级缓存，支持6144-bit位宽的六颗HBM3/HBM2e，支持PCIe 5.0，支持第四代NVLink总线。

H100计算卡有SXM、PCIe 5.0两种样式，其中SXM版本15872个CUDA核心、528个Tensor核心，PCIe 5.0版本14952个CUDA核心、456个Tensor核心，功耗最高达700W。

性能方面，FP64/FP32 60TFlops(每秒60万亿次)，FP16 2000TFlops(每秒2000万亿次)，TF32 1000TFlops(每秒1000万亿次)，都三倍于A100，FP8 4000TFlops(每秒4000万亿次)，六倍于A100。

售价方面，H100的官方建议价一直没公布，但日本市场之前预售的价格是475万日元，约合人民币24.2万元，比A100的价格至少翻倍了。

以上就是《NVIDIA称H100显卡全面投产》全部内容，想要了解更多相关内容的小伙伴可以持续关注，这里每天会为大家推送最新最全的硬件资讯，一定不要错过哦。

本文链接：<https://dqcm.net/zixun/167945982412906.html>