

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

课程论文

THESIS OF CURRICULUM



论文题目：数据中心的能源与水资源
——未来的环保聚焦点

学生姓名：_____杨乐天_____

学生学号：_____521030910353_____

专 业：_____人工智能_____

指导教师：_____殷 杉_____

学院(系)：_____电子信息与电气工程学院_____

数据中心的能源与水资源——未来的环保聚焦点

摘要

在数据科学的时代，数据中心是整个时代前进的基石，而数据中心的能源与水资源消耗也就自然成为了环保事业的重大话题。本文尽力以相对全面的视角看待并了解数据中心的能源与水资源消耗问题的重要性及其评价指标，并且简要分析了其引发的社会、环境问题。阻碍环保事业的是人们对经济发展无休止的渴望，但是我们将引入“解耦”的概念，论证数据中心与数字化信息产业如何帮助人们找到未来环保的解决方案。

关键词：能源 水资源 数据中心 经济发展

Energy And Water Consumption In Data Centers

--The Focus Of Environmental Protection In The Future

ABSTRACT

In the era of data science, data centers stand as the cornerstone of the entire epoch, and thus the energy and water consumption of which naturally emerge as pivotal topics in the realm of environmental conservation. This article strives to adopt a comprehensive perspective, gaining insight into the significance of energy and water consumption in data centers, along with their evaluation metrics. A brief analysis of the social and environmental issues stemming from such consumption is also presented. It is people's unrelenting desire for growth that hinders environmental initiatives, but we would introduce the concept of decoupling, and demonstrate how data centers and the ICT industry contributes to a path towards the future of environmental conservation.

Key words: Energy, Water Resource, Data Center, Economic Development

目录

- 1 绪论5
 - 1.1 数据时代的基石：数据中心.....5
 - 1.2 数据中心：能源与水资源.....5
- 2 现状5
 - 2.1 数据中心能源消耗.....6
 - 2.2 数据中心水冷系统.....6
 - 2.3 数据中心水消耗.....6
 - 2.4 衡量指标.....7
 - 2.4.1 PUE7
 - 2.4.2 WUE8
- 3 数据中心与环境的未来8
 - 3.1 数据中心解耦经济增长与能源消耗.....8
 - 3.1.1 解耦及其现状8
 - 3.1.2 数字化帮助绝对解耦8
 - 3.1.3 数据中心是数字化的未来.....9
 - 3.2 中心化：新技术的试验田.....10
 - 3.3 PUE 和 WUE 尚且不足11
 - 3.3.1 计算架构能效11
 - 3.3.2 从能源生产处减少水消耗.....12
- 4 前沿案例12
 - 4.1 脉冲神经网络.....12
 - 4.2 氢能发电.....12
 - 4.3 数据中心供热.....13
- 5 总结13

1 绪论

1.1 数据时代的基石：数据中心

这是一个数据驱动的时代。从**微观角度**看，在每一个日常且平凡的一天，我们获取电子设备上以各种模态展示的数据，悄无声息地产生关乎我们偏好与行为的数据，然后再获取数据、生产数据……如此循环。可以说，获取数据、理解数据、利用数据已经成为主导现代人生活的经济范式。从**宏观角度**看，数据经济的发展与全球化的加速、科学技术的迅猛发展高度耦合。全球化催生跨国巨头企业，企业则需要数字化的方式进行管理和决策，数字化的世界则进一步削弱了国界对经济活动的隔绝效果。

将来，数据与经济之间的千丝万缕将只增不减，无论将来的企业家们是否认可数据驱动的经济发展模式，数字经济都将在长时间内存续并繁荣。

抽象的数据需要物理的载体。驱动数据时代一路高歌的正是当今属于各个科技巨头的众多数据中心。在这里服务器中堆叠的 CPU、GPU 和磁盘时刻保持高效运转，借助人類的智慧将数据处理、分析、重构，产生具有经济价值的新信息。数据中心既是**真实物理世界与抽象的网络拓扑世界之间的桥梁**，也是**执行至关重要的人类决策的代理**。

1.2 数据中心：能源与水资源

任何能创造大量价值的事物都有对应的成本。尽管数据中心在现代社会中扮演着重要甚至不可或缺的角色，甚至可能会成为我们今后进行革新与创造的基础，我们仍然应该意识到它带来了大量的副产物——尤其是在环境方面。供给任何一个数据中心的电力成本都非常高昂，并且由于数据中心集中化的特点，其散热也是一个非常重大的问题。当下的主流解决方案是使用水冷系统，而这其中的水资源消耗也是一个巨大的数字。

本文的剩余部分将探讨数据中心的能源消耗和水资源利用问题。其中文章组织结构如下：第 2 部分描述当前世界上数据中心对能源和水资源的利用状况和其中的评价指标；第 3 部分尝试理论方面理解、论证数据中心资源利用对人类环境保护工程的整体重要性，指出当前可见的一些问题；第 4 部分给出了一些数据中心前沿技术的进展，第 5 部分给出一个对于全文的总结。

2 现状

Amdahl 定律：在计算机架构中，理论的性能提升 T' 可以用优化某个部件带来的性能提升 s 和该部件的执行时间比例 p 描述： $T' = \frac{1}{(1-p)+p/s}$ 。

作为任何复杂系统优化的普遍原则，Amdahl 定律告诉我们，不要在不重要的部件中花费大量时间，即使能够将该部件优化到极致，提升也会受到该部件重要性 p 的制约而显得很小。我们首先提出：**在环境领域关注数据中心相关的问题是颇具价值的**。

2015 年的研究指出，当时数据中心在全球能源消耗总量中大约占据 1.5%，而到了 2030

年将显著地提升到8%，并且这样的增长还会持续⁰。在 AI 大模型取得空前成就的今天，则有估测认为到 2025 年，数据中心的能源消耗将达到20%以上^[5]。可见，人们对未来的规划中，数据中心占有弥足轻重的地位，并且这一地位还在增长。假定我们认同这一观点，在接下来的数据中心建设中，如果我们能采取更加环境友好的方式，我们就可以显著地减少潜在的全球能源消耗和污染排放。

2.1 数据中心能源消耗

数据中心的目的是进行计算任务，执行信息处理，因此其中主要的能源消耗在于供给 CPU、GPU 和磁盘的电力。如果要减少计算用电对环境的影响，我们需要考虑两个方面：**让计算用电变得更清洁和节省计算用电。**

我个人认为计算任务使用的电力如何达到清洁与数据中心本身并没有太大关联，这应当取决于今后的能源行业发展，如何将核能、氢能等清洁能源投入到超大规模的发电设备中才是这个问题真正的答案。随着能源行业的变革，数据中心应对停电用的后备柴油发电机也自然会被清洁能源替代。相对地，节省计算用电就是数据中心可以考虑的问题。尽管计算是刚需，但是我们有义务提出更合理、更高效的计算架构和存储架构，节省尽可能多的计算用电。

除去这一最主要的能源消耗之外，剩下的**能源用于冷却和其他外周设备**（照明、安全设施等）。冷却是我们最主要的关注，因为这是一项纯粹的副产物——它本身并不对数据中心产生任何价值，但是需要可观的资源来维持运作。激进地说，整个冷却设备使用的资源都是一种浪费，而我们要做的就是尽可能减少冷却所需的资源。

2.2 数据中心水冷系统

准确地说，数据中心使用“液冷”而非狭义的水冷，只不过人们通常都选择使用水，便捷且便宜。在日常生活中，水冷系统并不常见，人们使用空调。水冷和空调的区别是显著的，空调制冷的本质是使用热机，通过做功推动高温热源与低温热源的热交换；而水冷则更为直接，用冷的东西让热的东西变冷。于是，一个自然的疑问是：**为什么使用水冷而非空调？**

数据中心的散热与我们通常所能够想象的使用空调冷却一个房间存在本质上的不同。由于数据中心的计算架构极其密集，其单位体积内产生的热量与使用空调冷却的房间可能相差多个数量级，而液冷系统大约拥有50至1000倍于空气散热系统的导热能力^[8]。另一方面，比起大量的商业用电，各国政府（主要指中国和美国，也是大多数数据中心的所在地）对于商业用水的收费价格是更加低廉的。因此**数字化企业往往在用电上多加考量，但是却敢于更“慷慨”地使用水资源。**

现今的数据中心，液体冷却系统有两种类型：直接芯片冷却(direct-to-chip cooling)、沉浸冷却(immersion cooling)，前者在计算机架构设计上就考虑散热问题，而将液体管道接到芯片附近；后者则将整个计算机系统浸入液体，以更高的成本实现更大的冷却范围。在循环端（也就是与外界连接处）冷水系统(chilled water system)用冷却设备对计算设备产生的热水进行冷却^[9]。

2.3 数据中心水消耗

也许与读者的预期不同，数据中心的水消耗问题远非一个简单的数字，列出年耗水量即为截止，事实是这牵扯到许多相关的社会问题。在美国的达拉斯，谷歌经历了 13 个月的法律斗争，试图争取不公开自己在当地用水数据的权利，但是最终，谷歌公布自己使用了当地 25% 的总供水量；在 Los Lunas，针对 Facebook 数据中心的用水，大量农民进行了示威游行，

举起了“农民需要水，人们不吃数据”的标牌。

我们首先要认识到：**数据中心的用水量远远小于其他产业的用水量**，在总用水量上并非极其紧迫亟待改善的领域。有研究提出，2015 年美国的水消耗是每天12180亿升，其中4460亿升用于灌溉，5030亿升用于发电，1470亿升用于饮用水；与之相对的是，在整个2014年，美国的数据中心使用了总计6260亿升的水^[2]。即使在 AI 大模型横行的今天，各大数据公司给出的环境报告也揭示出用水量并没有数量级的增长。

但是数据中心的水资源消耗仍然必须进入人们的视野，其中主要的问题就在于**数据中心往往建立在水资源稀缺的地方**。其中部分原因是科技企业需要在电力价格便宜的地方，而这些地方往往又是经济相对不发达、水资源相对稀缺的地方。正是因此，我们才能看到如前所述由数据中心的水消耗引发的社会问题。当然，在这一方面，中国的情况有所不同。一方面，我国数据中心在环境方面的数据公开性仍然相对不足，另一方面，我国的数据中心建设受到政府的严格审查，并且我国经济发展区域特性使得数据中心几乎都建设在华东地区，这些地区的水资源相对来说并不稀缺。

2.4 衡量指标

至此，我们已经对数据中心的环境影响有一个大概的认知。但是在迈向新的解决方案之前，我们还需要对当前数据中心的环境指标进行相对深入的理解。

大部分的科技公司都采用了如下两个指标衡量能源利用效率和水资源利用效率。

电力使用效率(Power Usage Effectiveness, PUE): 衡量数据中心有多大部分的能量进入计算系统。

$$PUE = \frac{\text{Total Energy Needed}}{\text{Energy Used for Computing}}$$

用水效率(Water Usage Effectiveness, WUE): 衡量使用单位能量需要多少冷却水。在公式中，标明的加湿和冷却实际上是 2.2 部分中的两种不同冷却原理，本质上都是用于冷却。

$$WUE = \frac{\text{Liters used for Humidification and Cooling}}{\text{Annual Energy to Power (kWh)}}$$

2.4.1 PUE

在这一语境下，根据 Uptime Institute 的调查估计 (Figure 1)，当下全世界的数据中心普遍能达到1.55左右^[10]，但是世界均值已经在该值附近停留很久，几乎没有进一步的下降。其原因可能在于更前沿的能效手段往往需要在数据中心建设之初就进行规划，故而无法应用到已有的数据中心。

另一方面，1.5左右也并非数据中心 PUE 值的瓶颈。新建的数据中心通常能够达到1.3甚至更低的 PUE 值，举例来说，Google 声称其数据中心在过去 12 个月中的 PUE 值已经能够降低到1.09，而 Microsoft 则宣称其数据中心的 PUE 值约为1.22，并且新建的数据中心设计标准能够达到1.12。从能源利用的角度来说，这已经是相对而言比较高效的结构了。

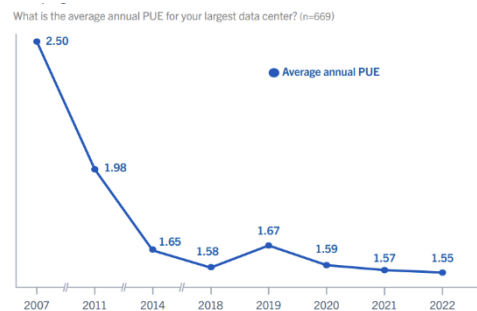


Figure 1 世界 PUE 均值发展趋势

2.4.2 WUE

根据当前的统计，世界范围内数据中心的 WUE 均值约能达到1.8L/kWh。实际的 WUE 值与周边环境关联很大，例如在微软发布的报告中（Figure 2），欧洲地区的用水几乎可忽略，而亚太地区的用水效率则相对较低，并且微软用水效率的地区差异对于其他的数据中心也基本适用。

我认为除了地理因素导致的环境不同之外，原因可能是欧洲和美洲国家对于 IT 企业的环保规定更加严格，而亚太地区在数据科学上相对薄弱，人们对于其环境影响也不够敏感。

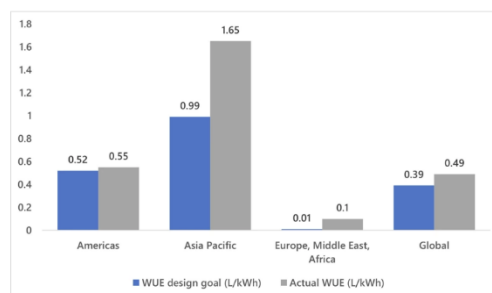


Figure 2 微软在不同地区的用水效率

3 数据中心与环境的未来

在第 2 部分中，我们已经了解了数据中心的能源与水资源消耗的现状，在这一部分中，我们将论证数据中心在接下来的数年、数十年中对环保，特别是能源行业的巨大重要性，并且指出一些将来人们需要关注的重点。

3.1 数据中心解耦经济增长与能源消耗

3.1.1 解耦及其现状

在中低收入地区，经济与能源使用是强相关的。一般而言，人们使用多少能源，就有多少钱。在这样的经济环境下，人们毫无疑问、也必须是以经济发展优先的；即使是在今天，人们也依旧会优先考虑经济发展，其次兼顾环境保护。因此在相对贫穷的过去，人们毫无顾忌地使用自然资源来获取经济增长，而现在，人们就必须考虑到如何提高单位能源能够创造的经济价值，甚至更进一步保持经济发展并降低总体能源消耗。经济学家和环境学者们将这个过程的称之为“解耦” (decoupling)。

从 2023 年的视角来看，已经有较多发达国家实现了部分甚至完全（几乎不增加能源消耗的同时执行经济增长）的解耦。但是我们不难想到，发达国家目前为止实现的解耦有相当一部分是因为制造业向海外转移。先前是中国，现在是东南亚地区，利用当地的低成本劳动力的同时还能利用当地的资源，仅剩的能源消耗就在于原材料供应链和产品运输的供应链。事实也确实如此，在经济发达的 OECD 成员国，上述供应链已经成为使用自然资源最多的行业^[3]。因此，不难发现，距离全球范围内的解耦我们还有很长的路要走。

3.1.2 数字化帮助绝对解耦

我认为数据中心的信息技术能够实现真正的解耦。在环境保护领域，解耦被视为有两种类型，一是**相对解耦**(relative decoupling)，让能源变得更高效率，也就是单位能源产生更高的经济价值；二是**绝对解耦**(absolute decoupling)，在经济增长的同时实现能源消耗的减少。

所谓相对解耦，单纯地提高能效，是非常自然的，并且从理论出发是简单轻松的，因为

随着科技发展我们总是有更好的能源解决方案,能够更充分地利用能源、更轻易地获得能源。但是绝对解耦就并非如此。目前为止宏观上实现绝对解耦的可能性和进展都还没有定论,并且得出结论认为已经实现绝对解耦的统计研究中,大多数都采用的是诸如碳排放这样的环境影响指标而非资源利用指标(注意到,环境指标绝对解耦是可持续发展的理论最低条件)[5]。

我们首先必须承认,在可见的未来,信息与通信技术在两个方面会导致能源的增加:一方面,IT 产业本身的增长会导致能源消耗的增长;另一方面,IT 产业会进一步促进经济的发展,而这会间接带来更多的能源消耗(弹跳效应, rebound effect)。因此在当下我们很难探讨信息技术产业绝对解耦的进展和证据。

但是数字化、信息技术产业在促进绝对解耦的目标上具有弥足轻重的作用。

首先,信息产业在自身具有相对较高能效的同时能够有效提高其他行业的能效。在其他行业中,数字化往往是一种管理手段,而这一手段往往是用于企业降低自身的运营成本。数字化意味着人们能够获取更多的统计信息,这些经济相关的统计信息毫无疑问会隐含能源成本、能源效率、用水成本等信息。即使我们承认企业未必会刻意为了节省这些成本而投入大量更新、修缮成本对生产设施更新换代,我们也不难想象,在运营中它能促进减少无谓的浪费、在新的企划中它能节省未来的长期成本。有统计表明,信息相关产业产值大幅增长的同时,能源消耗和碳排放几乎没有显著的增加(Figure 3) [6]。

另一方面,我们知道数字化与信息产业在产业结构的变化中(尤其是第三产业的增长, tertiarization)扮演了至关重要的角色。在一个典型经济体的发展历程中,农业经济的能耗很低,在工业化进程中能耗大幅增加,但是随着经济主体从工业向服务业转化,其能耗也会显著降低。根据柏林生态经济研究所的研究[4],大多数国家,哪怕是在工业化发展仍然不充分的中国,从 1995 年到 2016 年的 21 年中服务业在信息技术产业中的占比都有了较大增加(Figure 4)。

总体来说,上述四个 IT 产业实现解耦的积极因素和消极因素构成了当下关于 IT 产业对绝对解耦影响的主要争执。在当下所有的研究中,人们都只能通过一种相对合理的指标去估计最终的趋势,形成自己的观点,距离达成共识还有较大距离,但现有的证据(如上所示)相对倾向于支持 IT 产业成为绝对解耦的必要路径。

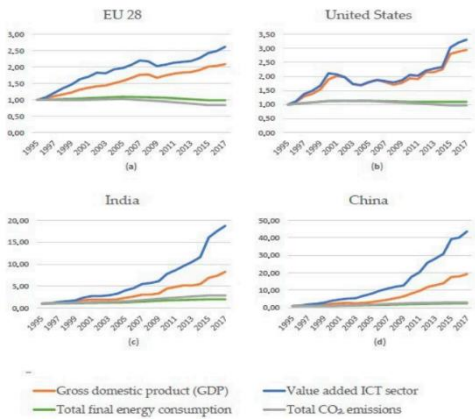


Figure 3 各国信息通讯行业与能耗统计

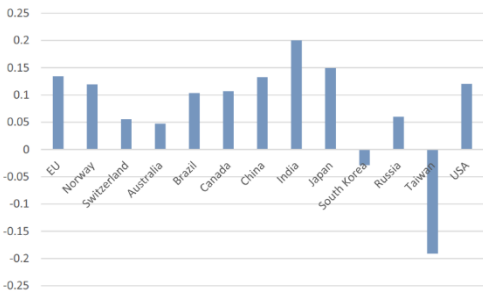


Figure 4 ICT 服务业在 ICT 产业占比

3.1.3 数据中心是数字化的未来

从技术的角度来说,数字化的趋势是中心化、集中化,也就是向数据中心发展。这一观点与我们的主题能耗与环保相对来说没有紧密的关联,因此我们仅仅对其作简单的阐述。

在未来的数字化产业中,随着智能决策、信息挖掘的需求量增加,人们对计算设备的硬件要求是显著增加的,而消费端的设备性能增长则远远跟不上需求的膨胀。举例来说,使用 ChatGPT 要求至少 80GB 的显存,但是目前为止最强大的消费级显卡 RTX 4090 也仅有 24GB

显存，因此这决定了所有大模型相关的服务需求都必须在云端进行。

另一方面，数据中心的单位算力制造成本远低于个人电脑，这一点我们可以类比公共交通和私家车的油耗对比，大规模的需求总是带来更低廉的单价。

总结 3.1 部分的观点和论证，我们得出了数据中心对于解耦能源与经济的总体重要性。数据中心本身可能能够达成解耦，而数据中心也能够帮助其他产业降低能耗、提高经济效益，它将很可能在未来的环境保护行业中产生越来越大的作用。

3.2 中心化：新技术的试验田

在这里，我将明确一个观点：**数据中心是名副其实的“中心化”管理结构**。众人皆知，数字化/IT 相关的公司数量远远大于数据中心的数量，那么这意味着那些相对小的公司就不需要数据中心吗？并非如此。几乎所有国内 IT 企业在“whois”页面中几乎都能指向阿里云、华为云这类高度知名的云数据平台，指向少数几个集中的物理地址。也就是说，少数几个数据中心是大量数字化企业的运转载体。事实上，数据中心所承担的计算任务远比其所有者的业务更加广泛，它们进行着全世界绝大部分的信息处理工作。而我个人认为，**中心化的信息处理架构将成为高能效、高环保的新技术最良好的试验田**。

我的观点是：这是一种**市场行为**。

我们先给出新能源车辆的例子：在欧美地区，新能源车普遍比汽油车贵上万美金^[1]、充电昂贵且不便、无法进行长距离出行、由于技术的发展而存在很大的贬值^[2]，因此客户群体很小；但与之相反，在今天的中国，政府通过各种政策手段让新能源车成为市场的更优解，包括但不限于给予车辆企业相关优惠、充分进行基建解决充电难题、在牌照上让消费者省心。

在新能源车这个**非中心化的**例子中，环保主义者需要让所有汽车的客户为环保买单，而显然一般消费者并不愿意为了环保这个概念而花更多的钱。中国的成功之处在于它让环保成为了价格上的优选，因此人们愿意接受电车相对落后的性能而为此买单。但是在数据中心的例子中，普通用户购买的并非数据中心的算力，而是借助算力提供的服务——而服务是环保的。这句话的含义应该被理解为，**人们本质上是在为服务付费，而数据中心的计算成本几乎对用户是不可见的**。事实上，从计算机操作系统的诞生起，人们就在贯彻“将有限的计算资源抽象成为无限的计算资源接口”这一理念，将人们需求的服务与背后的算力相分离，从而使得计算机行业能够专注于提供服务而不需要对计算成本进行管理。这一理念在云计算中进一步发扬光大，使得经济价值的创造几乎完全来源于服务而非算力。

实际上在数据中心市场中，已经建成的数据中心几乎无法大动干戈地进行改动，而目前几乎所有的数据中心都还在使用寿命中，因此我们仅需要考虑为全新的服务需求所建造的数据中心。

在数据中心进行环保技术尝试的主要考量在于云计算服务的高经济效益导致人们主要考虑数据中心服务器的运营成本即可，而建造成本能够在设计寿命的很小一部分时间内回收，从而实现净收益。从运营成本出发，浪费热量就是在浪费电费，因此选择节约能源是无害的，**高成本的新技术投入并不会在很大程度上影响企业的盈利**。次要的考量是，运营数据中心的科技企业需要社会名声来获取董事会的支持，进而得到资金投入。也就是说，**科技企业需要在环保方面满足公众、政府、特别是环保主义者的期待**。我们在谷歌、微软等巨头企业的年度环境报告中不难发现，这些企业正在以巨大的努力尝试向世人证明自己正在尽力实现环保的目标，而这一场面在任何当下拥有巨大利润的非中心化的制造业都是不可想象的一——大多数这些企业都在通过外包的方式声称自己正在努力，然后把皮球从自己脚下踢开。

3.3 PUE 和 WUE 尚且不足

根据我的了解,迄今为止人们针对能源使用效率和水资源利用效率的衡量指标几乎无一例外的是如前所述的 PUE 和 WUE。但是我想说,这两个指标是远远不够的,并且我需要声明这并非理想主义者的强迫症发作,人们迫切地需要引入新的指标来衡量自身的能源利用。

3.3.1 计算架构能效

如前所述,PUE 的定义是总能耗除以进入计算设备的能耗。总能耗的定义是明确的,但是如果我們仔细思考分母,也就是进入计算设备的能耗,可能会感到困惑——它到底包括什么。对于一个通常的数据中心,其架构是层级的,接入互联网的几个节点连接了下一层级的多个节点,这些节点又连接其下的更多节点,最终达到各个计算节点所在的服务器 (Figure 6 数据中心架构 Figure 6 给出了一个相对常见且普通的架构模式,其他架构虽然有所不同,但是总体上说层级的结构是统一的)。

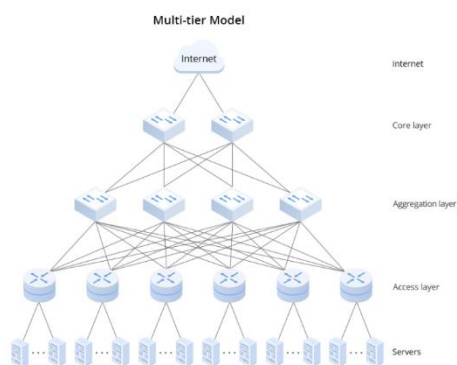


Figure 6 数据中心架构

的能效,这一点并没有被反映到 PUE 数值中。另一个缺点是,虽然电力分配器本身的功率被算作电力设备消耗而不算入 PUE 的分母,但是电力分配器的决策失误导致的电力资源浪费会显著地影响实际的能源消耗,并且这一浪费并不会计入 PUE 的计算中。

由此可见,我们急需一个或者多个 PUE 之外的指标对数据中心的能源效率进行一个更全面的评估。目前仅仅依赖单个 PUE 指标的评价体系很可能让人们提升计算架构的意愿变得更低,而这意味着能源在人们对其了解不足的情况下产生浪费。

当然我们必须承认,这件事情并不是在文章中提一句就能解决的问题。一个重要的挑战在于信息通讯的工作量是非常难以量化衡量的,也就没有办法建立一个基准^[10]。另一方面,衡量计算架构的效率实际上是一个远比想象困难的任务,人们对计算架构的研究几乎全部都集中在提升数据中心的反馈时间、负载均衡等方面,而几乎从来没有人探究过计算架构的能源效率,仅有的研究也是提出了一个单独的软件协议用于相关测量^[13],而这显然对服务器的运营人员来说是少有价值的,带来更大工作量的同时几乎不产生经济价值。

因此我们应该呼吁,也可以预期各个数据中心会采取一些相对简单但不够全面的指标衡量 PUE 之外的其他能效信息。我们期待将来能够有新的受认可的评价标准和一个可供参考的基准,帮助数据中心统计能耗信息、增进能效。

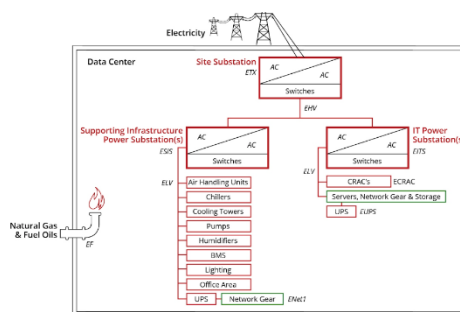


Figure 5 谷歌 PUE 计算方法

显而易见,高等级的节点作用是分配计算工作量到低等级的节点,而所有的任务指令都需要经过这些节点的分流,因此这些节点本质上与电力分配器一样属于管理节点而非真正的计算节点。但是在谷歌的 PUE 计算方法中 (Figure 5, 红色框出的被认为是其他能耗,而绿色框出的被认为是计算能耗),我们发现它把整个网络架构都视作计算设备计算 PUE,而实际上针对计算架构的优化已经有很多探索,这些不同的计算架构显然具有非常不同的

3.3.2 从能源生产处减少水消耗

我们已经知道当今数据中心的水资源利用效率普遍在1.8L/kWh，但是 WUE 对用水的衡量还不充分。其主要原因实际上在于：能源生产端，也就是发电站的水资源消耗也很高，即使是技术较新的燃气火力发电厂，预估用水效率也在1.5L/kWh左右。这意味着简单地对数据中心的用水进行优化是不够的，即使我们能完全不消耗水，相同的能耗也会让我们仍然需要消耗一半左右的水资源。

4 前沿案例

在这一部分中，我们从几个角度理解当下数据中心领域的前沿发展。学术届在这一方面的探索是多方向的，从芯片到计算架构再到清洁能源，我们能够从中窥见当下的人们对数据中心未来能源模式的想象和探索。

4.1 脉冲神经网络

从 2023 年的视角看待数据中心的核​​心任务，其中一个很大的部分就是大模型的推理和训练。GPT 系列模型的能耗可能是此前所有的软件任务都无法相提并论的，能够达到每次推理0.001kWh到0.01kWh之间，这大约是每次 Google 搜索的数十倍到数百倍，更不用说 ChatGPT 单词训练就需要消耗10GWh的电力^[14]。

因此，当下的 AI 领域一个重要关注点就是如何让我们的模型变得更高效。人们的思路是人脑智能实际上并不需要精确的计算，因此使用相对不精确的脉冲神经网络代替精确矩阵计算进行的人工神经网络架构（Figure 7）。

这一架构能够节约能源的主要原因在于它不需要大量矩阵乘法，而只需要进行加法操作，前者对于 $O(n^2)$ 的输入规模需要 $O(n^3)$ 的计算量，而后者则是 $O(n^2)$ 的^[7]。目前由于 SNN 相关硬件技术尚不成熟，其能耗估计尚且没有一个相对合理有效的上下界，但即使最悲观的估计也能有2.5倍左右的能效提升^[15]。

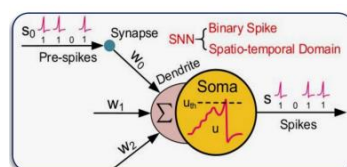


Figure 7 脉冲神经网络架构

4.2 氢能发电

在数据中心的供电系统中，最重要的任务是后备能源，包括不断供电系统(uninterruptible power supply system)等模块。这些模块当下采用的主要能源还是柴油发电，但是已经有新的尝试在于使用氢能发电(hydrogen fuel cells)。这一发电系统的优点在于彻底的零排放，产生的副产物只有水，并且还能够在一定程度上相比先前的柴油发电机降低维护成本。但是目前，氢能发电技术的主要问题——或者说几乎所有新能源发电的普遍问题，就在于它们无法产生传统化石能源火力发电所需的那么大量的能源^[16]。

在常规供电中，我们知道数据中心几乎已经成为新能源电力供应协议(Power Purchasing Agreement)的最大买家，包括太阳能、风能等多种电力来源。同时，有37%的数据中心承诺在短期内引入氢能发电的后备能源模块^[16]。

4.3 数据中心供热

根据热力学第二定律，我们知道再强大的技术都无法做到百分百的能源利用，无论如何我们的数据中心都会产生大量不被使用的热量。同样秉持着浪费热量就是浪费钱的原则，有公司和数据中心进行了合作，尝试使用数据中心对周边建筑进行供热。举例来说，荷兰公司 Switch Datacenters 将它的天然气供能设备替换为了数据中心服务器的外周热源，由此减少了它对天然气的依赖^[17]。

我们可以想象这一技术的前景有限，因为它对地理环境有较高要求、实现这一技术的价格也是非常高昂的，可以说以目前的设备成本来看，它只能说是一项实验技术。但是如果随着相关硬件设备的成熟，它很可能能够成为能源的一种补充，同时在相当程度上减少总体热量排放，降低全球变暖的速率。

5 总结

至此，我们已经对数据中心的能源消耗与水消耗有所了解。我们论证了数据中心本身对经济的重要性和其对环境保护行业，特别是能源行业未来的重要性，并且指出了当下对数据中心能效、用水效率指标的不足之处。最后，从一些前沿技术的案例研究中，我们也能大致窥见当下人们对于环保的努力方向指向何处。

“我们是感受到气候变化的第一代，也是能为之努力的最后一代。”美国前总统奥巴马如此总结我们在环境保护中面临的处境。“宁要绿水青山，不要金山银山。”这是习总书记为我们指明的道路。

在论文中，我们反复地提到环境、能源与经济的关系。作为一个人本主义者，我眼中的节能环保是与经济发展相冲突的，但是在人类发展的最终角度，这二者却又缺一不可。如何处理这对矛盾，让它们之间达成一种平衡，这需要人们的智慧——又或者，也许某一天人们能够将这两个概念置于某种正反馈的循环之下，让它们相互促进。用一句话总结我对环境问题的态度：在星际穿越遥不可及的时候，请别让《星际穿越》中的沙尘席卷我们的城市。

参考文献

- [1] Andrae, Anders S. G., and Tomas Edler. 2015. "On Global Electricity Usage of Communication Technology: Trends to 2030" *Challenges* 6, no. 1: 117-157. <https://doi.org/10.3390/challe6010117>
- [2] Mytton, David. 2021. "Data centre water consumption". *npj Clean Water* 4, 11 (2021). <https://doi.org/10.1038/s41545-021-00101-w>
- [3] Schandl, H., Hatfield-Dodds, S., Wiedmann, T., Geschke, A., Cai, Y., West, J., ... & Owen, A. 2016. "Decoupling global environmental pressure and economic growth: scenarios for energy use, materials use and carbon emissions". *Journal of cleaner production*, 132, 45-56.
- [4] Lange, S., Pohl, J., & Santarius, T. 2020. "Digitalization and energy consumption. Does ICT reduce energy demand? ". *Ecological economics*, 176, 106760.
- [5] Vadén, T., Lähde, V., Majava, A., Järvensivu, P., Toivanen, T., Hakala, E., & Eronen, J. T. 2020. "Decoupling for ecological sustainability: A categorisation and review of research literature". *Environmental science & policy*, 112, 236-244. <https://www.man-es.com/energy-storage/campaigns/energy-supply-for-data-centers>
- [6] Santarius, Tilman, Johanna Pohl, and Steffen Lange. 2020. "Digitalization and the Decoupling Debate: Can ICT Help to Reduce Environmental Impacts While the Economy Keeps Growing?" *Sustainability* 12, no. 18: 7496. <https://doi.org/10.3390/su12187496>
- [7] Wu, Dengyu, Xinping Yi, and Xiaowei Huang. 2022. "A little energy goes a long way: Build an energy-efficient, accurate spiking neural network from convolutional neural network." *Frontiers in neuroscience* 16: 759900.
- [8] <https://www.techtarget.com/searchdatacenter/feature/Liquid-cooling-vs-air-cooling-in-the-data-center>
- [9] <https://dgtlinfra.com/data-center-cooling/>
- [10] <https://uptimeinstitute.com/resources/research-and-reports/uptime-institute-global-data-center-survey-results-2022>
- [11] <https://money.com/why-americans-not-buying-electric-cars/>
- [12] <https://www.bettersafe.com/blog/electric-car-depreciation-rate>
- [13] <https://www.nrel.gov/docs/fy17osti/68576.pdf>
- [14] <https://www.washington.edu/news/2023/07/27/how-much-energy-does-chatgpt-use/>
- [15] <https://danielmuellerkomorowska.com/2021/03/04/spiking-neural-networks-for-a-low-energy-future/>
- [16] <https://www.information-age.com/trends-in-data-centre-sustainability-123500760/>
- [17] <https://www.datacenterdynamics.com/en/news/switch-datacenters-heat-homes-and-offices-using-residual-server-heat/>

谢辞（含意见与建议）

非常感谢殷老师这一个学期的授课，从您的课中我学到的“尺度”、“从社会角度思考环境问题”这些思考方式都是知识本身所不可比拟的，包括您最后两节课所提到的关于学术研究的方法和规范，我想这方面的经验传授是一个真正为学生着想的好老师才能够带来的。可以说，这是我在交大的三年中最满意的课之一（姑且加上之一吧），如果说这样一节通识课还不能称之为通识课的标杆的话，那我想通识教育一定是出了什么差错。

其实在我最开始选课的时候，我所理解的“全球变化”并不是气候、环境意义上的全球变化，我本以为我们所探讨的是一些地缘政治、全球经济发展之下的自然与生态发展。如果一定要说这门课有什么建议，那么我会说：政治、社会学视角下，我认为学生需要知道如何去看待自然与生态——说到底，在自然与生态方面真正付出努力、做出贡献的人和组织，往往在社会上是受到忽视的，他们的意见在政治角度下也往往是不受待见的。当然我理解在这样一门生态环境学的通识课上加入这些元素可能会在多方面受限，因此它也仅仅是我个人的一点小小期望而已。

当然，我个人坚信生态与环境的其中一个出发点就是人文社会，因此在这样一篇课程论文中，我贯彻了自己的观点，试图从经济、社会的角度去理解数据中心相关的自然与生态问题。我本人只是一个纯粹的“理工直男”，作为人工智能专业的学生研究计算机与人工智能技术，在经济、政治、哲学、历史以及生态环境方面兴趣大于学识，因此论文中可能有很多谬误和糟糕的观点，这方面还请老师谅解、不吝您的批评指正。

总而言之，这一门通识课对我的大学生活帮助远比选课时所预期的大，我也因此无比感谢殷老师。我今后毫无疑问会向自己的同学、学弟学妹推荐这一门课，我相信不只有我能够从这节课上得到这么多收获。