AI 算力爆发下的美国电力需求：数据中心与电网改造

作者：SVSAC 硅谷战略顾问委员会

核心判断

美国数据中心电力需求值得关注，企业更需要判断哪些需求能转化为可进入的设备与工程环节。对企业而言，值得优先推进的是：按区域电网条件、项目阶段与采购角色拆分机会，核对设备适配、交付能力及本地服务要求。本篇结合公开资料与分析，讨论企业可以如何识别机会、组织能力并安排下一步投入。

关键词：人工智能；数据中心；电力需求；电网改造；负荷预测；风险分配

SVSAC - sv-sac.org

SVSAC | sv-sac.org

美国数据中心电力需求值得关注，企业更需要判断哪些需求能转化为可进入的设备与工程环节。

我们的建议是：按区域电网条件、项目阶段与采购角色拆分机会，核对设备适配、交付能力及本地服务要求。企业可以先据此梳理自己的产品、客户、能力与投入边界，把讨论推进到可以分工、可以验证的具体任务。

对管理层来说，这个议题值得放进经营议程。围绕“先做什么、凭什么做、做到哪一步再扩大投入”形成共识，往往比同时铺开多个方向更有推进价值。下文保留相关资料、比较和推演，供企业判断自身条件与行动顺序。

1 核心问题与分析边界

1.1 从算力叙事转向可检验的电力问题

“算力爆发”首先是需要界定的题目表述，而不是已经完成测量的解释变量。模型调用次数、加速器出货量、机房建设面积与电网接口功率分别描述不同活动。如果把它们合并为一条增长曲线，就可能得到方向明确却无法检验的结论。对电力系统而言，一项计算活动的意义取决于其何时发生、在哪里发生、持续多久，以及供电受到约束时是否允许调整。研究应先回答这些问题，再判断它是否构成扩建需求。

本篇将人工智能理解为可能增加计算活动、改变设备组合和调度方式的一类用途。数据中心则是承载多种计算、存储和网络活动的设施边界。二者既有交集，也不相等。缺少工作负载分类计量时，可以研究数据中心总用电及其与 AI 发展相关的机制，但不能将总量全部标记为 AI 用电。对于同一设施内的通用计算与加速计算，也不应只依据企业名称或宣传定位划分。理想证据是可复核的设备、利用率、设施辅助负荷与工作负载口径；无法获得时，应缩小结论范围。

本篇的核心问题由此分为三项。第一，公开资料实际证明了什么程度的需求增长，它描述的是历史估计、近期预测还是长期情景？第二，在全国总量增长之外，哪些工程与制度条件决定局部电网能否接纳新负荷？第三，当需求实现路径不确定而网络投资具有较长使用期时，怎样安排扩建、灵活接入和风险承担，才能避免将一个预测数值当作不可更改的建设指令？这些是本篇拟回答的核心问题；本篇不提供已经识别的全国因果效应。

1.2 时间边界与证据版本

本篇将 2026 年 10 月 23 日设为检索截止日。截止日表示研究纳入材料的时间边界，不表示所有数据都更新到该日。某份 2026 年发布的报告可能采用此前年度

SVSAC | sv-sac.org

的运行数据，某份较早报告则可能预测 2028 年或 2030 年的需求。研究必须同时记录发布日期、历史基期和预测终点。若只在表格中写一个年份，就无法知道该数是当年观测还是若干年前对当年的预测。

美国能源部 2024 年 12 月 20 日的报告发布说明给出，数据中心 2023 年用电估计为 176 太瓦时，约占全国用电 4.4%；其中 325—580 太瓦时是对 2028 年的预测区间。[1] 这些数字可用于说明当时报告所判断的量级，但不能当作 2028 年已经发生的结果。LBNL 题为“2025 Update”的更新报告网页标注出版时间为 2026 年 6 月，所读报告摘要给出的 2030 年占比估计为 11.8%，情景范围为 9.5%—15.3%。[60] 由于终点不同，本篇不将两份材料的数值差异直接解释为预测上调的幅度。LBNL 更新报告完整方法的分析见第 27–28 章[60]；[2]仅为未独立读取的出版入口，数值与方法论据以已读报告为准。

版本管理的意义不只是保留旧数字。若一个区域预测变化，需要辨别是项目状态改变、宏观假设修订、统计边界调整，还是模型本身发生变化。不同原因对应不同决策：项目延期可能影响近期设备交付节奏，统计修订可能只改变账面比较，持续利用率变化则影响电量与收入。只有将修订拆解到原因，预测更新才能成为工程和经营信息。

1.3 研究对象与决策建议强度

本次分析的观察对象不是某一家设备厂商的订单，也不是某个地方政府公布的招商金额，而是负荷进入电力系统的过程。有关设备企业的市场机会可作为系统扩建的后续问题，但不能从全国用电增长直接推出某一产品必然获得订单。产品是否进入采购，仍需核查具体技术方案、预算、供应商资格、招标与合同。本篇的边界是解释负荷进入电力系统的条件；设备出海或市场准入判断只有在具体采购主体、产品、适用规则和合同证据相连时才成立。全国用电增长本身不能预测某一产品获得订单，本篇也不对未经核验的产品认证状态作结论。

正文按三类陈述约束论证。事实是来源明确支持的历史记录、制度文件内容或预测文件所作陈述；推断是本篇根据这些材料建立的机制解释；建议是有条件的行动主张。例如，“报告预计需求增长”是关于报告内容的事实，“增长集中于受限节点可能抬升扩建压力”是机制推断，“对高不确定项目分期配置设备”则是建议。三类陈述可以共同出现，但不能互相替代。

SVSAC | sv-sac.org

2 需求测量：电量、功率与计算活动

2.1 电量并不等于峰值负荷

电量是一个时间区间内的能量累计，功率是某时刻的能量转换速率。研究中常见的太瓦时与吉瓦不能直接比较。若已知时间序列功率 P(t)，电量可以写成 $E=\int P(t)dt$；若只知道额定接口功率，则仍需知道实际利用率、爬坡过程、停机时间以及计量边界。把接口容量乘以全年小时数，只能得到特定连续满功率假设下的计算值，不能据此宣称实际年用电。

作为纯粹的量纲示例，一个假设的 100 兆瓦接口若全年平均使用率为 70%，则一年 8760 小时的电量为 613200 兆瓦时，即 0.6132 太瓦时。这里的 100 兆瓦、70%和全年稳定使用均为作者设定，不是美国数据中心的观测值。若 100 兆瓦指 IT 设备功率而非电网接口功率，还需要设施辅助能耗；若已是设施总功率，则再乘一个设施能耗系数会重复计算。本示例仅用来展示边界错误怎样进入估算。

对电网改造而言，年电量相同的负荷可以产生不同工程要求。如果一类负荷在系统紧张时能够降低功率，而另一类必须在同一时段维持高负荷，则二者对容量和网络的要求不应由年电量相同而被视为相同。反过来，平稳运行的负荷也不能自动被判断为没有容量压力；其稳定需求仍可能占用受限接口。这里应考察的是负荷与系统条件之间的联合分布，而不是孤立的平均值。

2.2 IT 设备、设施辅助系统和电网接口

完整核算需要明确服务器、存储、网络、冷却及供配电等环节是否包含在内。PUE 常用于表达设施总能耗与 IT 设备能耗的比值，但比值改善并不必然意味着设施总用电下降。如果 IT 用电增长超过辅助系统效率改善，总量仍会增加。Masanet 等的同行评议研究说明，计算服务增长与数据中心能源增长之间受到设备和设施效率变化的调节，不能依靠单一服务指标机械外推。[7] 该研究的早期全球历史分析不等于当前美国 AI 设施的参数估计。

我们提出，在资料允许时分别报告三组变量：第一组为计算或服务活动量及质量边界，第二组为 IT 设备功率与运行电量，第三组为设施或电网接口电量。三组变量之间的转换必须列出系数来源与计量周期。同一设施不同年份的比较也应检查服务质量是否相当；如果模型能力、延迟要求或输出任务发生变化，“每次调用”未必仍是可比较的单位。

一个容易忽略的问题是辅助系统的边界。某些冷却或储能设备可能在设施之外购入服务，某些现场发电可能减少电网购电而未减少总能源消耗。若研究关注电网负荷，应测量接口净取电；若关注能源与环境影响，则需要另外记录现场发

SVSAC | sv-sac.org

电、燃料和储能损耗。我们建议用双账而不是单一指标：一张账跟踪对电网的功率与电量需求，另一张账跟踪设施的全部能源投入。双账之间可以核对，但不宜混为一个“绿色程度”分数。

2.3 全国部门统计的归因限制

EIA 部门统计是区分商业、工业和住宅用电的资料入口，但本篇尚未核实 2026 年 10 月固定版本原表，故不把旧稿转录的 2027 年部门增速作为已核事实。[9] 商业部门包含数据中心，并不等于其全部用电来自数据中心，更不等于来自 AI。即使取得部门合计，仍须核对零售售电、总用电、发电量及自发自用边界，不能用部门变化完成用途归因。

本篇对全国统计设置一条保守规则：若来源只提供部门合计，就在该层级描述变化；只有来源直接分解或研究具有可复核分解方法时，才向更窄用途归因。对于无法识别的部分，保留“未分解”比用一个看似精确的比例填补更有解释力。否则，分类上的不确定性会被隐藏在小数位之后，读者却可能误认为测量已经完成。

3 从接入申请到实际负荷的证据链

3.1 项目不是单一状态

项目公告表明某个主体公开表达建设或投资意图，但不充分证明最终地点、供电条件、设备配置和时间表均已确定。项目进入接入研究，也不等于已具备所有建设条件。本篇据此提出项目成熟度记录框架，依次记录意向公告、场地与关键权利、正式服务申请、工程研究、服务或建设承诺、施工、阶段接电以及持续运行。这是用于组织证据的研究框架，不主张各州或企业均采用同一法定程序，也不保证每个项目完全按此顺序推进。

在这个框架中，一项项目可以在不同维度处于不同阶段。建筑已施工而网络尚未扩建，或者首期接电而后续机房仍是计划，都不能用一个“已落地”标签覆盖。研究应保存每次状态变化的时间、文件、范围和责任主体。如果一份公告只支持首期容量，就不能将整个园区规划容量记为运行负荷；如果多个地点是互斥选址，应记录其关联，防止把同一需求重复计入多个区域。

3.2 PJM 材料提供的方法启示

PJM2026 年负荷预测报告指出，相较 2025 年版，近端至 2032 年的预测降低，原因包括电动车、经济预测及大负荷调整的更新。报告还区分近期需要较明

SVSAC | sv-sac.org

确承诺的项目与远期不确定性更高、需要折减的项目。[3] 本篇从中得到的方法启示是：对项目落实条件进行审核可以改变预测路径；较低的近期预测与长期增长预期并不逻辑冲突。这一启示并不等于已经估计出美国所有地区申请量的虚增比例。

PJM2026 年 5 月准确性报告将大负荷预测与可得运行数据对照，明确区分真实峰值、月内日峰值的平均以及未来预测，并说明部分客户数据因可识别性而不展示。[4] 对研究者而言，缺少展示值不能补为零，PDF 文本抽取中相邻数字也不能未经图像核对便配成一组。一个严谨的比较应先保证时点、区域、客户集合和峰值定义一致，再计算误差。否则，所谓误差可能只是口径差异。

3.3 概率权重与相关性

为组织不确定性，我们提出按项目建立候选负荷路径 L_i(t)，并将状态证据与实现条件并列保存。在说明性模型中，预期路径可写为 Σp_i(t)L_i(t)。这里 p_i 是研究设定或经数据校准的条件实现权重，而不是天然可观察的事实。未完成历史样本校准时，只能将不同权重作为敏感性假设，不能写成已经测得的成功率。

简单加权还存在相关性限制。多个项目可能依赖同一网络扩建、同一供电合同或同一终端客户；若其中共同条件失败，各项目可能同时延期。把每个项目当作互相独立，会低估组合风险。对于互斥选址，直接相加会高估负荷；对于分期投运，直接使用最终园区容量则会抬高近期路径。因此，项目库需要记录关联组、共同约束和阶段容量，而不仅是公司、地点与兆瓦数。

在工程决策中，期望值本身也未必足够。某些设备无法平滑扩容，某些线路建设存在前置时间，规划者可能需要同时比较低、中、高路径及极端共同条件。我们建议将期望情景用于解释平均路径，将压力情景用于检验后果，并将两者与可靠性分析区分。用一个统一概率分数代替所有工程约束，反而可能制造精确性的错觉。

4 电网改造的系统边界与决策顺序

4.1 全国充足与局部受限可以并存

全国发电量足够并不能充分证明某个受限节点能够按时接纳新增负荷。对一个具体接口，需要检查上游电源可获得性、输电通道、变电容量、配电或专用线路、保护控制和运行安排。本篇列出这些环节，是工程问题的分析目录，不是对任何特定项目完成安全研究的声明。局部改造方案仍应由适用的技术研究和审批决定。

SVSAC | sv-sac.org

从需求研究角度，地点并不是事后附加的标签。相同计算活动若落在不同节点，可能具有不同接入日期、网损、可靠性安排和扩建费用。研究因此应把地点嵌入需求路径，而不能先建立一个全国总需求数，再机械按州份额分配。区域分析还应保留网络边界与行政边界的区别，避免将州内销售统计直接等同于区域运营机构的负荷。

4.2 长期规划与近期接入属于不同决策

FERC 关于 1920 系列的官方说明概括了至少 20 年区域输电规划及费用分配方法的要求。[6] 这提供了长期协同规划的制度背景，但不意味着某个数据中心因此自动取得接入权，也不证明全部地区的后续申报与执行已经完成。2026 年 6 月 18 日，FERC 又宣布要求六个区域运营机构论证或改革大负荷相关费率条款。[10] 本篇目前只确认该行动公告，不将其描述为各区域已执行完毕的统一接入制度。

长期规划需要考虑多种需求和资源路径，而近期接入需要更具体的地点、工程和合同证据。二者应相互反馈：近期项目的兑现情况可以修正长期路径，长期网络建设也可能改变项目可选地点。若把二者混合，就可能出现两种误判：用长期增长情景为所有近期扩建背书，或因个别近期项目延期而否定全部长期改造需求。合理研究应明确每项证据服务于哪个决策周期。

4.3 灵活性需要可履约条件

DOE 的数据中心负荷灵活性研讨会总结讨论了储能、现场供能和负荷管理等方式。[8] 研讨会材料说明可研究的技术与组织方向，但不是全国灵活容量已实现规模的测量。本篇因此将“可能调整”与“可签约、可测试、可计量的调整”分开。只有在需要响应时能够交付、响应后仍能满足工作任务，并且恢复过程不造成新的约束，灵活性才可以进入具体规划假设。

我们建议，评估一个灵活方案至少回答响应幅度、通知时间、最长持续时间、年度允许次数、服务质量限制以及恢复时段六个问题。若方案依赖电池，还应单独核查能量与功率约束、备电要求和补能安排；若依赖工作负载迁移，还应说明目标地点是否同时受限及数据传输条件。这里不预设所有 AI 任务都可推迟，也不预设低延迟业务完全没有可调空间，而要求按照可验证的任务类别逐项判断。

灵活接入的比较对象也应明确。对于受限节点，它可能减少某些时段的扩建需求；对于全年持续缺电的情形，短时削峰未必能解决电量供给。若只有少数小时受到限制，需求调整可能具有价值；若约束长期存在，单纯依靠短时储能可能

SVSAC | sv-sac.org

只是推迟而非消除问题。这一判断来自时间尺度与能量守恒的分析，并非对某个美国项目的实证结论。

4.4 建设顺序中的风险分配

当需求尚未完全兑现而基础设施需要提前建设时，研究不能只问“是否会增长”，还应问“谁为提前建设承担风险”。若全部风险由网络用户共同承担，预测过高可能产生闲置成本；若全部风险由新负荷承担，某些具有系统共享效益的投资也可能难以推进。本篇的建议是按成本因果、共享收益和可退出条件分别讨论，而不是以支持或反对数据中心的立场替代费用分析。

分期投资可以是一个待评估选项。例如先建设适合确定首期负荷的设施，再依据里程碑安排后续扩容。但分期并不总是成本最低：重复施工、设备接口限制或土地走廊条件都可能改变比较结果。研究需要在同一需求情景下比较一次建设、分期建设与可履约灵活接入的全周期成本，并说明未满足需求、延误和闲置的后果。当前尚未建立足以支持全国性净收益结论的数据，因此本篇只提出比较框架。

上述分析形成一个论点：AI 相关电力需求研究的关键是把需求规模与兑现条件、空间约束和责任安排放在同一证据链中。区域规划和项目数据仍是检验该框架的依据，不能仅凭概念分析完成政策推荐。

5 情景计算：三个容易被平均数掩盖的问题

5.1 将参数敏感性与需求预测分开

需求研究常把一个最终电量值呈现为模型结果，而不说明它对利用率和设施边界的依赖程度。为了暴露这种依赖，本节构造一个可复算的说明性算例。设某设施的 IT 额定功率为 100 兆瓦，全年 8760 小时，IT 平均利用率 u 分别取 0.5、0.7 和 0.9，设施能耗比 q 分别取 1.15、1.30 和 1.50。此处 u 表示平均 IT 功率与上述额定 IT 功率之比，不是 GPU 的计算利用率；q 按相同统计区间定义为设施电量与 IT 电量之比。两者均为作者选择的敏感性参数，不代表行业均值或美国观测样本。

在这些限定下，设施年电量为 E=100×u×q×8760/10^6，单位为太瓦时。计算结果如下。该表用于量纲和参数检验，不是美国需求预测。

SVSAC | sv-sac.org

平均 IT 功率比例 u	q=1.15	q=1.30	q=1.50
0.50	0.50370	0.56940	0.65700
0.70	0.70518	0.79716	0.91980
0.90	0.90666	1.02492	1.18260

表中首尾值相差较大，并非由额定 IT 容量变化造成，而是输入参数变化的结果。如果某份研究只公布“100 兆瓦项目”，却没有说明该功率的边界、平均使用比例和辅助能耗，读者就无法从该数唯一推出年电量。更多小数位不会弥补输入缺失。进一步说，若 100 兆瓦已经是设施接口额定功率，再乘 q 就不符合本表的定义。这也是第 2.1 节接口算例与本节 IT 算例不能直接横向比较的原因。

对于该简化模型，E 对 u 和 q 分别呈线性，对二者共同变化则呈乘积关系。但这种数学形式不证明实际设施中二者独立。实际辅助负荷可能随气温、设备负载与冷却方式变化；年度平均比值与逐时比值也不一定相同。本篇使用区间一致的年度电量比，使本算例在定义上自洽，却没有据此识别真实的逐时机制。后续若用现场数据估计，应将设备运行与气象、辅助系统负荷共同对齐。

作为反例，一座效率改善的设施可能同时扩大 IT 运行规模，因而出现 q 降低而总电量提高。另一座设施可能因低使用率使电量较少，却仍为高功率运行保留接口。二者均表明，电量、效率与连接容量回答不同问题。不能仅凭其中一个指标给出设施对系统影响的完整排序。

5.2 少量转移电量为什么仍需要恢复约束

第二个算例讨论灵活性。假设一个以 100 兆瓦恒定取电作为比较基线的设施，在若干受限时段累计 50 小时降低 20 兆瓦，且这些小时互不重叠。被减少的基线电量为 1000 兆瓦时，即 1 吉瓦时；相对该基线全年 876 吉瓦时，占比约 0.114%。所有数值仍是作者假设。

这个很小的全年占比不能说明方案没有系统价值，因为容量约束可能集中在少数小时。但它也不能证明方案可执行。若工作只是推迟而没有消失，并且在同一功耗与任务产出关系、忽略损耗的条件下需要补回 1 吉瓦时，则仅给 10 小时恢复窗口，就需要相对基线平均增加 100 兆瓦。这可能超过原接口余量；若不允许增加接口功率，恢复可能需要降低其他任务，或扩大恢复时间。该反例揭示的是时间约束，而不是某种技术已被证明不可用。

因此，一份灵活方案至少需要两条时间序列：受限期减少多少，以及恢复期增加多少。若仅计算受限期削减，却不追踪恢复负荷，就可能将峰值移到另一个

SVSAC | sv-sac.org

紧张时段。若依靠电池在接口侧削峰而 IT 不中断，则补能和储能效率仍需计入。若依靠任务迁移，则应将目标节点增加的负荷纳入同一边界，否则局部改善可能只是地理转移。

比较对象决定结论的含义。在接口受限但系统其他时段仍有余量的条件下，有限次数响应可能有价值；在全时段供电均不足的条件下，同样的响应承诺就不能替代稳定能源供应。研究不能把一个响应百分比无条件转换成可免建的发电或输电容量，因为不同资产服务的约束与时间尺度并不一致。

5.3 相同期望值不意味着相同建设风险

第三个算例保留十个候选项目，每个项目投运后负荷 100 兆瓦，各自实现概率均假设为 0.6。考虑两个极端模型。模型 A 假设项目实现事件相互独立；模型 B 假设它们依赖同一个共同事件，要么全部实现，要么全部不实现。两个模型的期望总负荷均为 600 兆瓦，不能通过平均数区分。

模型 A 的总负荷方差为 $10\times100^2\times0.6\times0.4=24000$ 兆瓦平方，标准差约 154.92 兆瓦；模型 B 的方差为 $1000^2\times0.6\times0.4=240000$ 兆瓦平方，标准差约 489.90 兆瓦。在独立模型中，十个项目全部实现的概率约为 0.6047%；在完全共同事件模型中，该概率为 60%。这些是给定假设的数学结果，不是对美国项目成功率或相关性的估计。

这个对照说明，以相同期望路径制定同一投资方案可能忽略尾部结果。但它不直接告诉规划者应按哪一个分位数建设。建设决策还需要未满足负荷的后果、可追加投资的时间、资产可转用程度与合同约束。对于能够服务多类用户的网络资产和高度专用的连接资产，闲置风险也不相同。简单地认为方差越大就应按最大值一次建成，同样缺少成本与后果依据。

现实项目通常既非完全独立，也非完全同向。共同因素可能来自上游电网扩建、算力采购客户、资金条件或同类设备供给；互斥因素可能来自相同业务的备选地点。后续研究应识别这些关联再设定情景，而不是先选一个相关系数以获得预期结果。由于目前缺少覆盖候选项目、退出项目和真实投运项目的统一样本，本节将两个极端模型作为误差诊断，不作为实际组合分布。

6 全国统计的固定版本核对

6.1 以表级证据约束总量陈述

下表保留旧稿注明来自 EIA《2026 年 10 月短期能源展望》表 7a 的转录，以便后续逐格核对；该固定原表尚未核实，表值不作为本篇已验证的全国统计基线。

SVSAC | sv-sac.org

[9] 数字原样保留不猜改，也不据其计算 AI 增量。单位按旧稿标注为十亿千瓦时，数值等于太瓦时；2026 和 2027 的估计/预测属性及 2025 的统计状态均须由原表确认，不能将这些转录数当作已核全年实绩。

指标	2025	2026	2027
总用电	4195	4288	4356
向最终用户售电	4058	4152	4217
其中商业部门售电	1493	1549	1592
直接使用电量	137	136	140

表中“直接使用”属于官方定义的特定边界，不能简单解释为所有表后能源；分类与来源脚注应保留。分项独立取整可能造成加总差异，不能为追求表面平衡修改官方值。[9] 原表核实前，本表只作待核转录保留；上述分类和取整原则不等于已验证各数值。

固定版本对比较的重要性可由一个研究操作说明。如果研究者在本月从更新网页取得预测值，数月后再从同一网址取得另一数值，网址相同并不能证明对象相同。未保存版本时，很难判断差异来自误引还是更新。固定版本次分析应保留来源文件和哈希，同时记录网页入口；只有确实取得文件后才可声称归档，现有 EIA 转录尚不能凭旧记录完成验证。文件哈希证明本地文件的一致性，不能替代对官方来源、发布日期和内容的核验。

6.2 为什么增长的会计分解不等于 AI 的因果贡献

假设全国总用电变化可在一致口径下拆为各部门变化之和，这只是会计关系。即使某部门同时出现数据中心扩张与电量增长，也不能排除天气、其他商业活动、能效、统计调整等因素。会计上的“来自商业部门”与因果上的“由 AI 造成”处于不同推断层级。前者可由一致部门表计算，后者需要用途分解或反事实设计。

直接归因的另一困难是替代关系。某项 AI 活动可能在既有云设施中发生，也可能替换原有计算任务；新建园区则可能承接从其他地点迁移的业务。新增设施电量不必等于全国净增量。相反，若 AI 使某些服务成本下降并扩展使用，效率提升也可能与活动增加同时发生。这里讨论的是可能机制，当前材料不足以估计它们各自的大小。

SVSAC | sv-sac.org

因此，本篇在后续统计章节采用分层结论：全国表说明总体和部门路径，设施证据说明局部运行，只有经过一致归因的方法才说明 AI 净贡献。三者若不完全衔接，应报告缺口，而不是用全国宏观增长来填补项目数据缺失。

7 因果链、替代解释与可证伪检验

7.1 需要识别的不是一个相关系数

要判断 AI 发展是否造成电网扩建，首先应明确结果变量。它可以是新增变电容量、网络资本支出、工程开工、接入等待时间或用户费用；这些指标不具有相同时间顺序和机制。资本预算增加可能先于实际建设，工程投运则可能服务此前已存在的可靠性需求。将所有结果统一称为“基建热潮”，会掩盖需要解释的差别。

本篇拟检验的具体机制是：可验证计算业务增加，使项目提出新增接口需求；该需求在与现有负荷和供给条件共同研究后，暴露局部约束；运营机构或公用事业形成相应工程与成本安排；随后进入批准、采购和投运。每个箭头都可能被其他因素阻断。项目取消可能阻断需求到施工的联系，现有余量可能使新增负荷无需重大扩建，已有改造计划也可能早于 AI 需求。只有检验这些中间环节，才能判断扩建是否确实由研究对象推动。

一个可证伪的项目层面命题是：若某项工程的必要性主要来自一个确定的新增负荷，那么在其他关键条件不变且该负荷退出后，工程范围、时间或成本归属应出现可解释的重新评估。如果工程完全不变，可能说明它仍有其他需要，也可能说明调整成本或程序刚性过高。不能仅据“不变”就断言原归因错误，需要继续寻找工程说明和决策文件。这种反例分析能帮助区分需求驱动与既有网络改造，而不是将所有新工程都计入 AI。

7.2 选址内生性与比较组污染

简单比较“有数据中心地区”和“没有数据中心地区”的投资增长，容易受到选址内生性影响。企业可能优先进入电网余量大、审批进度较快或已有扩建计划的地点。于是观察到的相关性既可能来自数据中心推动投资，也可能来自基础设施条件吸引数据中心。控制几个地区变量并不能自动消除这种反向关系。

若使用分期建设形成的时间差，还要检查所谓处理日期。公告、许可、施工和接电对应不同机制。以公告日当作负荷冲击日可能提前记录处理，以投运日衡量工程投资又可能漏掉前置建设。比较组也可能受到同一网络工程影响，或者承

SVSAC | sv-sac.org

接本地区未能落地的项目；此时它并非未受影响的对照。电网边界与经济活动边界并不完全重合，因果设计必须体现这种空间联系。

在数据足够时，可以把事件研究或其他准实验方法作为候选，但不能先写方法名称再假定识别成立。需要检查事件前趋势、同期产业或政策变化、样本选择、干预预期，以及对照组受影响程度。若检验不能支持这些条件，论文应退回机制案例和描述统计，并明确限制。使用更复杂估计器不应成为提高结论强度的替代品。

7.3 未观测项目与幸存者偏差

项目数据库若只收集公开宣布的大项目，容易遗漏小型扩容、取消项目或从未进入正式程序的意向。只观察建成设施会高估从意向到运行的兑现率，反过来只观察申请队列又可能高估最终规模。两种偏差可以同时存在于不同数据源，不能用它们相互抵消的想象来消除。

本篇拟采取的最低数据要求，是保留项目进入样本的规则，并对失去跟踪的项目设置“未知”，而非直接归为取消或成功。退出研究的原因也应单独编码：主动取消、位置转移、阶段延期和缺少公开信息并非同一事件。若无法覆盖这些状态，概率权重只能作为情景输入，不应从不完整样本估出貌似精确的实现率。

7.4 机制论证的决策建议边界

证据分工应随问题而定：预测约束有条件路径，运行记录约束已发生负荷，项目文件连接承诺与工程，比较设计排查竞争解释。第 19–20 章据此分析信号误差如何改变分期选择，第 27–28 章追踪出货、存量、平均功率和容量的转换；两者均不能用未说明方法的访谈或宣传替代检验。

当前研究已完成测量框架、说明性计算与部分官方统计核对，尚未形成项目级因果样本。因而不能声称 AI 已经造成某一全国比例的电网投资，也不能据上述算例给出最优建设规模。后续区域案例应围绕这些可检验环节补充正反证据。只有证据支持的范围扩大，最终结论的范围才应随之扩大。

8 Virginia：把无约束需求与可交付供电分开

8.1 案例选择与可迁移范围

Virginia 案例适合检验一个具体问题：当数据中心的商业扩张意向强于基础设施短期交付能力时，需求预测应如何进入建设决策？选择这个案例的目的，是分析证据链中的中间环节，不是把单一区域当作全国缩影。不同区域的网络条

SVSAC | sv-sac.org

件、资源配置和制度边界会改变同一新增负荷的后果。因此，本篇不据 Virginia 情况直接推断 Texas 或其他地区的建设规模。

JLARC2024 年报告区分无约束需求、其一半以及没有新增数据中心需求的模型情景；无约束需求指尚未扣除供能基础设施建设等限制的需求。报告的费用研究还同时指出，当时费率对现有服务成本的分配基本适当，而未来增长仍可能提高其他用户承担的系统成本。[5] 前者是情景定义，后者涉及既有成本分配与新增系统成本的不同问题，均不能直接转述为 2026 年的观察结果。

这组材料使研究可以避免两种相反误读。第一种把无约束需求当作电网必须全部按时供应的确定负荷。第二种则认为只要实际供应受限，需求本身就不值得研究。更合理的处理是保存原始意向路径，再将建设、价格、合同与时间限制映射为可交付路径。两条路径的差额本身是研究对象，可能通过延期、迁移、规模调整或项目退出表现出来；现有材料尚不足以估计各机制所占比重。

8.2 模型中的对照情景与现实中的反事实

设置“没有新增数据中心需求”的模型情景，有助于在固定其他模型假设的条件下观察增量差异。但它仍是模型内对照，不是已经观察到另一个完全相同、唯独没有数据中心的现实 Virginia。模型结果的解释取决于资源可建性、成本、时序、政策边界和需求设定。若这些参数同时变化，差额就不再仅对应一个需求因素。

因此，在阅读区域规划时，应先建立变动清单：哪些变量被固定，哪些允许调整，哪些约束在情景间发生变化。仅比较两个最终投资数而不保存清单，容易把政策假设的变化误称为负荷影响。我们建议对每个区域情景至少保留需求路径、资源可选集、可建时间、网络约束和评价目标五个字段，并注明来源。该建议是可复核研究设计，不是声称 JLARC 或所有规划机构已经统一采用本表。

还需要考虑实际均衡反馈。若供电成本上升导致项目迁移，最终需求就不再等于原先无约束路径；若企业愿意自担部分专用成本，交付顺序也可能变化。对需求采取完全外生假设可以帮助回答“若这些负荷实现，需要什么”，却不足以回答“这些负荷最终会实现多少”。这两个问题应分别建模，不能把工程可行性计算直接命名为商业需求预测。

8.3 现有分配合理与未来费用上升并不矛盾

费用讨论中常把“数据中心是否支付自己的成本”与“其他用户费用是否上升”看作同一个问题。实际上，前者可能针对某一时点已发生的服务成本及其费率分配，后者还可能来自未来新增投资、资源稀缺和价格变化。即使没有直接交叉

SVSAC | sv-sac.org

补贴，边际供给条件变化也可能使其他用户费用增加。反过来，其他用户费用上升本身也不能充分证明存在不当补贴。

研究应把费用变化分成至少三层。第一层是可识别的专用连接资产及其直接费用；第二层是同时服务多类用户的共享资产费用；第三层是需求与供给变化通过市场或采购机制影响的能源成本。三者的证据不同：第一层更接近具体工程和合同，第二层需要规划收益和分摊方法，第三层需要价格形成与采购组合。使用单一“每度电价格”无法揭示这些机制。

例如，在纯粹假设的情形中，一条专用线路全部由新增用户付费，而共享电源的边际成本因系统供需变化上升。此时既可能不存在该线路费用转嫁，又可能出现其他用户电费增加。另一个相反的假设是，短期市场电价下降，但某项专用资产风险被错误分配给其他用户。仅观察平均电价下降同样不能证明费用分配公平。这些反例说明，应直接检查成本归属，而不是用电价方向代替分配分析。

8.4 价格、可靠性与地方收益的评价单位

地方政府可能更关注税基与建设活动，电力系统研究则关注网络、资源与可靠性，家庭用户关注账单。这些结果的空间和时间边界并不相同。不能把地方财政收入直接抵扣跨区域用户的电费增长，也不能把某一类用户的费用变化视作全部社会福利变化。评价至少需要说明受益者、付款者、期间和价格基准。

若以后开展净收益估计，应区分财务转移与实际资源消耗。例如税收在不同主体之间形成转移，而建设材料、燃料、运维和土地机会成本涉及真实资源。就业和投资额也不能无条件相加为福利收益；是否新增、是否替代其他活动以及如何计价，需要独立依据。当前稿件没有这些估值数据，因此不输出一个把全部项目收益成本混合相减的总数。

可靠性也不应只以项目是否接上电衡量。接入时点、持续可用性、紧张时段响应义务以及未履约后果均可能影响价值。若某种加速接入依赖可中断安排，研究应说明它与连续供电产品的服务差异。费用较低或接入较快并不自动意味着同质服务更高效，只有在服务内容可比时，这种比较才成立。

8.5 区域案例尚需补足的证据

第 12 章继续讨论 2024 研究之后的制度材料，但更新链仍需按原命令和适用对象逐项确认，尤其是最新州级资源组合与个案服务文件。2024 年材料可支持历史研究设计，不能单独支持当前具体合同建议；后续制度信息也不能代替项目层核对。尤其不能把报告中的预测、建议或情景设定写成已经实施的政策。

SVSAC | sv-sac.org

PJM 的预测审核材料可用于考察大负荷如何进入区域预测，Virginia 材料可用于考察州内成本与规划问题；二者需按地域、周期和定义对接，不能因都涉及同一地区就直接拼接数值。当前研究已说明这一接口，但尚未完成项目样本及规划案号级核对。因此，本章形成的是关于分析机制的案例解释，而不是对 Virginia 截至 2026 年 10 月全部电力需求与政策状态的最终评估。

9 从购电协议到可用电力：Crane 案例的证据边界

9.1 购电承诺能够证明什么

Constellation 于 2024 年 9 月 20 日宣布与 Microsoft 签署 20 年购电协议，支持重启后来命名为 Crane Clean Energy Center 的 TMI 一号机组；公告称相关能源用于匹配 Microsoft 在 PJM 的数据中心用电，预计恢复约 835 兆瓦容量，当时预计 2028 年上线。[11] 其项目网页现把预计投运时间写为 2027 年。[12] 两个时间点应作为不同版本保留，后者仍是企业预期，不是运行证明。

该公告为“计算基础设施需求与电源商业安排存在具体联系”提供了项目证据，但不能单独确定新增 AI 净负荷。合同涉及的电源容量不等于某个数据中心接口的实际功率，也不说明其所有电量均用于 AI。合同期限不能替代逐时发电曲线，公告的预计投运年也不能替代许可与调试记录。研究若把 835 兆瓦直接加到数据中心需求表，就跨越了供给与需求、额定功率与实际运行的边界。

9.2 商业对应与物理供电不是同一条证据链

一个购电安排可以支持项目融资与收入确定性，其物理交付仍需满足电源可用性、网络条件与适用程序。在一般分析中，合同上的能源或属性对应，不充分证明任一时刻的负荷均由指定电源直接供应。因此，研究应同时保存商业合同链与物理供电链，再检查它们在时间、地点和计量边界上怎样衔接。这是分析原则，不是声称已经取得该合同的非公开全文。

Pennsylvania 环境部门的项目页面说明，重启涉及 NRC 审查及州、地方相关许可，并分别描述联邦与州的职责。[13] 本篇据此不能把长期购电协议描述成全部许可已取得。企业发布的进度、州许可与联邦核安全许可分别支持不同主张；任何一项的进展都不能自动替代其余程序。若要判断投运概率或剩余关键路径，还需查证原始许可档案与具体生效状态。

这个案例还提供了时间修订的实质含义。2028 与 2027 不是两个可以择一装饰文章的年份。对于设备采购、调试与电网计划，提前一年可能影响交付顺序和供电安排；对于需求统计，则应避免在两个年份重复计入同一容量。项目数据库

SVSAC | sv-sac.org

应保留原预期、修订预期和实际投运三个字段，只有最后一项取得运行证据后，才能转入实际供给序列。

9.3 需求归因与供给增加应分账核算

若一个新增负荷促进了某项电源重启，可以研究其商业促成作用；但系统净影响还取决于电源和负荷各自的实现路径。负荷先到而电源后到，可能在过渡期依赖其他资源；电源先到而负荷延后，则可能服务其他需求。将二者在总量表上同时配平，会遮蔽这些过渡状态。

反过来，即使电源与负荷年度电量完全相同，也不能由此推出网络不需改造。发电点与负荷点的位置、逐时出力与取电、检修以及网络约束均会改变交付。这个反例并不否认购电协议的价值，而是说明研究应将合同稳定性、能源匹配和接口可靠性分别评价。当前公开材料不足以对该案例形成完整成本收益估计，本篇不据此给核电、储能或其他技术作普遍优劣排序。

10 输电需求研究与工程方案的层级

DOE 于 2026 年 7 月 9 日发布 2026 年全国输电需求研究草案征求意见，说明需求驱动包括数据中心、制造业和其他大型工业负荷。[14] 本篇把该资料作为截至检索时已识别的官方研究版本，不将“草案”写成最终裁定，也不将多种驱动合并为 AI 单一归因。下一步仍需核实是否存在之后发布的最终版本及其变化。

需求研究识别的是需要解决的问题；某条线路、某种设备或某项投资则是具体方案。两者之间还需要备选方案、成本、可实施时间、收益与风险分配的比较。即使区域需求已经被确认，也不能据此跳过工程选择，更不能直接推导出某一国别供应商的可获得订单。设备市场判断仍需采购制度与合同证据，不能把宏观需求分析当作出口订单证明。

这里存在一个可供后续章节检验的命题：在相同新增负荷情景下，局部扩建、跨区输电和可履约需求调整可能改变工程组合，但其效果取决于原约束。若原问题是接口热容量，增加远方电源未必解除接口瓶颈；若原问题是持续能源不足，仅扩接口也不够。论文需要先识别约束，再讨论方案，而不是按技术名称分别列出优势后把它们相加为总收益。

SVSAC | sv-sac.org

11 区域比较：增长预期怎样转化为可使用的规划输入

11.1 比较对象应当是预测程序，而非两个醒目数字

对区域电力需求的比较，最容易采用的方法是并列各运营机构公布的远期峰值。这样做直观，却往往没有回答数值如何形成。有的数来自大型负荷申报，有的包含兑现率调整，有的采用气象归一化，而另一些则是压力情景。若这些区别未被识别，表格中较大的数可能仅表示纳入了更多未承诺项目，并不表示当地计算活动已经更大，也不证明当地扩建需求按相同比例增加。

本篇比较 PJM 和 ERCOT，是因为已获得的公开资料能显示各自对需求真实性和时间路径的处理。比较不要求二者制度相同，反而要解释为什么输入来源、审核规则和决策用途不同，会改变数字的含义。研究中的关键单位是“某机构在某个日期、为某个用途发布的某一版预测”，而不是一个脱离版本和程序的“地区需求”。

研究据此设置四项对照维度。第一项是新增负荷来自项目申报、既有计量还是宏观模型。第二项是哪些承诺足以进入近期路径，以及尚未达到条件的项目如何处理。第三项是预期误差来自项目未实现、投运延后还是利用率不足。第四项是预测将用于资源充足性、输电规划还是短期运行。四项维度共同决定可比较性。若其中任一项尚未核清，比较应保留缺失而不计算排名。

11.2 ERCOT 初步预测中的大数与官方自我限定

ERCOT 于 2026 年 4 月 15 日发布面向 2026—2032 年的初步长期负荷预测，列出 2032 年 367790 兆瓦，并在同一公告中明确表示该预测可能高于预期实际增长。[15] 这个限定不是可以删掉的公关附言，而是决定如何使用数值的信息。将初步预测当作已确定的建设规模，会使原本用于讨论和修订的输入被误用为不可变需求。

ERCOT 五月市场通知的具体调整方案尚未取得可核原段，本篇不据此复述两种方案或量化兑现率。[16] 已取得的后续文件显示 PUCT 于 6 月 18 日批准 Batch Zero 流程，[24]九月董事会材料预计新版长期预测于 2026 年 12 月中下旬发布。[25] 因而可以确认流程获批、新版数值待发，但不能用流程批准替代最终预测或个别项目接电许可。

这类材料说明，未调整申报路径与可行动规划路径可以属于同一个预测体系的不同层次。研究若只摘录最上端申报数，就可能把审查过程中的不确定项目全部当作建设义务；若只摘录调整后的一个数，又可能隐藏尚未满足条件但未来可

SVSAC | sv-sac.org

能实现的需求。更完整的记录应保留原申报、调整依据和决策采用值，并分别保存每个版本的时间。三者之间的差额不宜一律命名为“虚假需求”。

尤其需要区别“项目不存在”与“项目尚不能在指定年份供电”。前者涉及申报真实性，后者可能涉及网络时序、设备供应或商业安排。若把二者合并，一个真实但推迟的项目会被误当作永久退出，下一次预测又可能将其作为全新增量重复计入。跨版本项目标识和阶段记录因此具有实质统计价值，而不是数据库维护细节。

11.3 PJM 近端修订与长期增长判断

PJM 于 2026 年 1 月发布的官方网页说明，2026 版相较 2025 版下调至 2032 年的近端峰值路径，并将大负荷承诺审核纳入调整。网页表列 2027 年夏峰预测 160451 兆瓦、2030 年 183008 兆瓦。[17] 两个数均为该版本的预测，不是运行观测。本篇不将它们与 ERCOT2032 年初步值直接相除来比较两个市场的增长速度，因为终点、程序和输入成熟度尚不一致。

近端下调并不充分支持“AI 需求泡沫已经破裂”，长期上行也不充分支持“所有近期申报均会兑现”。这两种结论都把一个时间层级的结果投射到另一个层级。对工程采购而言，近端路径可能决定首批设备安排；长期路径则影响走廊预留和更大范围的规划。若同一报告同时给出这两类信息，研究应分别解释它们的用途，而不是寻找一个方向标签概括整份报告。

PJM 网页还说明其审核会考虑项目实现可能性及重复计算风险。[17] 这一做法的研究意义是，单纯汇总企业申报不是唯一的官方预测逻辑。但是，审核具有内容并不证明其结果一定准确。评价程序仍需比较预测与之后的运行，并识别审核规则改变后样本组成是否变化。只要缺少这样的后验验证，程序改进最多支持“有明确筛选机制”，不能直接升级为“预测误差已经消除”。

11.4 用运行记录检验量级时，必须先剥离天气

EIA 于 2026 年 8 月 3 日发布的运行回顾称，ERCOT 在 7 月 22 日小时峰值达到约 91.1 吉瓦，并指出该时段处于热浪背景。[21] 这一运行记录可以更新 4 月公告所引用的旧峰值背景，但不能单独识别 AI 造成的增长。小时峰值同时受气温、既有用户行为和其他负荷影响；把新纪录与数据中心扩张放在同一段落，并不能建立排他归因。

更有辨识力的检验应至少采用两个层次。第一层在相同天气与日历条件下比较总负荷，检查基线是否改变。第二层在可获得的项目或客户计量中核对新增负荷何时进入系统，再与总量变化对接。若只有第一层，可以讨论天气调整后的总

SVSAC | sv-sac.org

体变化；若只有第二层，可以讨论局部项目的已实现需求。只有在两者边界一致且其他增减项可处理时，才有条件将局部项目连接到区域净增量。

反例可以说明为什么这一步不可省略。一个异常炎热的夏季可能使系统创纪录，即使数据中心增量很小；一个气温温和的年份也可能没有创纪录，而新增数据中心电量仍持续增加。因此，“是否创峰”不适合作为 AI 需求是否增长的唯一检验指标。全年电量、非高温时段基荷、不同客户计量和负荷持续曲线需要共同提供证据。

11.5 审核与许可变化使项目时序内生化

Texas 州长办公室于 2026 年 8 月 3 日发布数据中心审核指示，9 月 21 日进一步宣布在审核完成前暂停 TCEQ 相关许可的发放。[22],[23] 这里记录的是官方指示及其发布日期，不是对各机构之后实际执行、个案豁免或司法效力的完整核查。本篇尤其不把暂停指示等同于项目全部取消，也不根据公告自行推定审核已经完成。

这些材料改变了预测分析的一个假设：项目时序不能始终被看作与制度状态无关。若项目需要满足额外信息或许可条件，其供电与施工日期就可能改变。即使最终计算业务需求没有消失，区域内实现的时间与地点也可能重新分配。因而，基于 4 月意向的路径不能在 10 月未经状态核对直接沿用，更不能假定政治或程序变化只会影响投资者情绪而不影响实际交付。

但制度变化也不是所有延期的充分解释。某项目可能在指示之前已因设备或融资问题延后，另一个项目则可能不处于相同程序阶段。识别审核影响需要保留项目的原定时间、已取得条件与实际受影响环节，而不能将同一时期所有延误都归于政策。本篇当前没有这种完整样本，所以将制度变化作为需要显式建模的条件，不估计其导致的取消容量。

11.6 可形成的比较决策建议及其边界

截至当前已核实的材料，能够支持的比较结论是：两个区域都不能被简单表述为“需求申报即需求实绩”，而且预测中的项目成熟度与时间审核会影响规划输入。它不支持一个统一的全国折减率，更不支持把某一区域的审核结果无条件迁移到另一地区。概率模型若要跨区使用，必须重新检查样本、退出原因、供电条件和程序规则。

比较问题	本研究已获得的证据	尚不能推出的结论
ERCOT 原始量级与修订	4 月初步公告、5 月建议、6 月流程批准与 9 月	新版最终数值、全国项目成功率

SVSAC | sv-sac.org

比较问题	本研究已获得的证据	尚不能推出的结论
	发布延期说明	
PJM 近期与长期路径	2026 版网页与承诺审核说明	全部申报均会兑现或全部为泡沫
已发生负荷	2026 年 7 月 ERCOT 小时峰值回顾	峰值增量全部由 AI 造成
Texas 审核与许可	8 月与 9 月官方指示	全部项目取消或所有个案执行一致

ERCOT 同一 9 月规划材料还将正在审核的项目区分为依赖大负荷与不依赖大负荷两类，列示依赖大负荷的审核中项目金额为 213.11 亿美元。[25] 这是项目审核状态及其条件依赖，不是已投运资产、已支出资金或 AI 设备订单。条件性项目需要随着对应负荷的审核结果重新评估，这为“需求状态会回写工程组合”提供了比单纯增长口号更具体的文件证据。

后续区域定量比较应按这张证据矩阵补齐项目与运行资料。若证据仍停留在程序和版本层级，论文就应保留在这个层级作机制解释，而不以假设数据填满一张看似完整的区域投资排名表。

11.7 州级综合资源规划提供另一种决策边界

州级资源规划与区域运营机构预测不能互换。文献 26 保留为 Dominion2024 年规划的历史线索，但所引 5.5%增长和 2039 翻倍等具体陈述尚无本次分析已读原段支持，不承担本篇事实论据。[26] 比较应先识别年份、地域、负荷与电量、规划情景及审批状态；一份规划文件不自动等于具体资源组合的建设许可。

这个例子说明，预测对象与审批对象需要分别记录。IRP 可以列出为满足未来义务而考虑的供需资源，但资源方案进入文件不等于每个项目已获批准。类似地，区域峰值预测可以影响州内资源规划，却不能替代企业服务区、客户集合和需求侧资源的单独核对。如果把公司供电区需求、全州销售和 PJM 总负荷混用，增长率可能在没有真实业务变化时就产生差异。

对于规划分析，我们建议保留一张“义务–资源–程序”对应表：负荷义务采用何种区域和客户边界，供需方案依赖哪些技术和时间假设，哪些决定只在规划层讨论、哪些需要后续程序。表的作用是防止将研究方案误读为可直接采购的建设清单，而不是否定规划文件的价值。2026 年更新版的正文不在本篇可核验范围内，故本篇不判断其是否保留或改变了 2024 版的具体组合；这一限制只影响对最新方案的评估，不影响关于文件层级区别的分析。

SVSAC | sv-sac.org

12 大型负荷费率：从需求预测走向风险契约

12.1 费率不是需求预测的附属变量

在只关注物理容量的模型中，需求路径常被当作外部输入，费率仅影响成本。但对尚未完全落实的数据中心，最低收费、合同期限和担保条件也可能影响容量申请、分期规模和选址。费率因此既是分配成本的工具，也可能改变被预测行为。忽略这种反馈，会使“先按原申请预测，再按预测建设，最后设计费率”成为一个没有校验的单向流程。

这不是说提高收费一定会降低无效申报，也不是说长期合同天然有效。强制承诺可能使真实但资金有限的项目退出，也可能让大型企业保留多个有财务能力的备选地点。要判断筛选效果，需要比较条款变化前后申请、承诺和实际接电的项目集合，并识别融资条件同时变化的影响。条款的存在与筛选的实现是不同事实。

12.2 Virginia GS-5：获批、生效与适用对象

Virginia SCC 于 2025 年 11 月 25 日发布最终命令及新闻说明，批准大型用户 GS-5 分类，生效日期为 2027 年 1 月 1 日。[18] 旧稿所引最终命令条款仍待原段复核：其转述为相关分类包含连续场址上测量或合同需求至少 25 兆瓦、测量或预期负荷率至少 75%的客户；对符合条件的新签服务合同设置 14 年期限。[19] 这些阈值、期限及生效状态未在本篇作为已核具体法律条件；研究必须保留适用条件，不能把任何数据中心、任何年份的合同都自动置于同一安排下。

对于 GS-5 安排，本篇将新旧客户、起始服务时间、最低收费、合同与信用担保作为须按适用文件核定的风险维度，不据此认定这些条件已统一适用于所有客户或合同。[20] 原命令及说明材料需进一步核实，本篇据此只分析待核条款所对应的风险机制，不将 2026 年 10 月表述为新分类已经全面生效。至于之后提交的具体费率文本、客户个案与任何修订，仍须以其文件另行核验。

获批与生效之间的时间差具有研究价值。若在生效前观察到企业改变申请行为，可能是对未来规则的预期响应，而不是现行收费已经作用的结果。相反，生效后短期没有变化，也不必然表示规则无效，因为既有合同、建设周期与提前锁定安排可能延迟观察结果。政策评价应区分宣布、批准、生效和首批受约束合同四个日期。

SVSAC | sv-sac.org

12.3 最低收费首先处理容量占用风险

在一般机制上，需求收费的最低基数使客户即使未充分使用所保留容量，也可能承担一定付款义务。这有助于把部分容量保留风险分配给提出需求的一方。但是，付款义务并不直接创造实际电量，计费需求也不等于计量峰值。研究若将最低收费基数作为真实负荷填入预测模型，会把财务承诺错误地转换成物理运行。

我们提出用四个变量分开记录：合同预留容量 C、计费基数 B、实际峰值 P 和实际电量 E。C 描述服务承诺，B 取决于适用条款，P 描述运行中的最高功率，E 描述期间累计。即使某种条款使 B 不低于 C 的一定比例，也不能推导 P 或 E 同样达到该比例。这个区分对于低利用率设施尤为重要，因为其账单可能稳定而物理负荷仍低。

一个假设例子是：某客户保留 100 兆瓦，但某年实际峰值只有 40 兆瓦；若适用条款使部分需求收费按 85 兆瓦计，财务收入与实际用电就发生分离。这里数值只是说明，不是对任何客户的计费计算，也不代表所有收费项目都使用同一比例。对于规划者，85 兆瓦不能直接当作运行负荷；对于成本次分析者，40 兆瓦也不能直接用来推断公用事业收入不足。两类分析需要各自的变量。

12.4 长期义务、担保与风险并不等价

合同期限可以延长客户承担义务的时间，但长期现金流承诺并不等于同等期限现金已经到位。客户违约、主体结构和可执行担保都会影响最终回收。担保则处理信用风险的一部分，其保障程度依赖形式、额度、触发条件和履约主体。不能仅依据“14 年”三个字就认为所有搁浅成本已被消除，也不能把担保额度直接解释为监管者估计的预期损失。

在财务解释上，预期损失通常需要同时考虑违约概率、违约时敞口和回收程度。某个担保上限可能来自政策设计、合同结构或审慎要求，不一定是这三个因素的实证乘积。若研究要从条款反推风险价格，就需要证明其定价机制和参数，而不能把观察到的美元数当作一个已校准风险模型的输出。

反例同样重要。一个信用较强的客户可能承诺很大容量，但最终只使用一部分；一个信用较弱的客户可能实际高负荷运行，却在行业下行时无法履约。前者主要涉及容量和时序，后者更集中于信用与回收。若用一个笼统“项目成熟度”分数处理两者，就会失去选择对应条款的依据。项目库应将工程成熟度、商业承诺和信用条件分列。

SVSAC | sv-sac.org

12.5 保护既有用户与提高系统效率的双重检验

一项规则可以改善成本承担，却不必然使系统投资组合最优。例如，客户能够支付专用连接成本，不说明该地点是系统最低成本地点；长期最低付款能够改善收入确定性，也不代表一次建成全部远期容量最有效率。费用分配解决“谁付”，工程比较解决“建什么、何时建”，二者需要衔接而不是彼此替代。

另一方面，不能把系统效率简单等同于最少支出。延后供电可能影响计算服务，提前建设可能保留未来选择，灵活接入可能改变服务质量。这些后果需要在同一评价边界下比较。若只计工程资本支出而忽略服务延误，容易偏好少建设；若只计交付速度而忽略闲置和风险，又可能偏好过早扩建。研究应先明确目标和承担者，再讨论最优方案。

我们建议使用两组可审查结果评价大型负荷规则。第一组是分配结果，包括哪些费用由新客户直接承担、哪些仍需共享，以及未兑现需求的剩余风险。第二组是行为和工程结果，包括申请是否减少重复、承诺是否更接近运行、分期安排是否变化、服务和网络成本如何调整。第一组即使改善，也不能代替第二组的验证。

12.6 数据可得性对规则评价的限制

进行上述评价，需要客户或项目层面的合同时间、容量、关键里程碑和运行数据。然而公开文件可能因商业保密或客户可识别性只展示汇总。因此，研究未必能够精确估计某个条款的边际效果。在这种情况下，可以报告制度设计、可观察结果及无法识别的机制，不应通过编造客户样本完成回归。

即使只有汇总数据，也仍可进行有限但有价值的检验。例如保留同一口径下的申请总量、正式承诺总量与实际接电总量，观察它们的跨版本变化；再检查制度改变与这些变化的时间是否一致，并列出同期融资和工程变化作为竞争解释。这种描述不能证明因果，但能排除一些过强说法，比如“规则宣布当天全部风险已经转移”或“申报下调就意味着真实业务需求消失”。

区域预测与风险契约之间的联系在于：项目何时进入预测，不只取决于技术需求，也取决于可验证承诺与制度条件；承诺如何影响投资风险，又不能仅靠财务条款推断物理负荷。工程比较还须检查不同约束对应的技术方案，避免把负荷审核、信用保障和系统可靠性混为同一项工作。

SVSAC | sv-sac.org

13 电量、资源充足性、网络接入与动态稳定性是不同约束

13.1 四类问题对应四种时间尺度

讨论数据中心“缺电”时，至少应先问它究竟指哪种约束。第一类是能量供给：在数小时、季节乃至年度尺度，系统能否提供足够的兆瓦时，尤其是在燃料、储能状态、天气和机组检修共同不利时。第二类是资源充足性：面对负荷峰值与资源故障、天气风险，系统有无足够可依赖的发电、储能、可控负荷和备用能力。充足性不是把铭牌容量相加；资源能否在需要的时段持续运行、爬坡和充能会改变其有效贡献。本篇据上述定义提出：充足性分析须明确时间因素、资源可用性和输电约束；文献 38 仅保留为尚待对应原段的参考线索，不将此建议归为该文件已核结论。[38]

第三类是输配电接入与可交付性：即使某区域的年度能源和资源容量充足，特定并网点上游的线路、变压器、母线、保护和电压条件也可能不允许立刻增加指定功率。静态潮流、故障电流和保护研究回答的是具体电气边界下的工程问题，不能从全国发电量或资源充足率推断某个园区的接入日期。大型负荷可能先受输电走廊限制，也可能在靠近接口的变电站或配电馈线上遇到热容量、短路容量或电压调节限制；这些瓶颈位置不同，替代工程也不同。

第四类是动态稳定与运行安全：扰动发生后，系统在毫秒到数分钟的范围内能否维持电压、频率和控制稳定，以及设备能否安全穿越故障并按预期恢复。规划者即使已经证明稳态潮流可行，也不能据此推出负荷在电压暂降、重合闸、频率偏移或保护动作下会保持同一行为。NERC 的 2026 年大型负荷可靠性指南分别建议提交稳态、动态和短路模型，并建议在需要时使用正序动态与电磁暂态模型；指南同时说明其为自愿、非约束性指导，而非可靠性标准。[31],[39]

把这些约束放在同一张表中，不是将其混为一个“电网容量”数字，而是检查决策先后关系。若年能量短缺，短时削峰不等同于补足能源；若资源可用但节点受限，远端新增电源也未必缓解接口瓶颈；若稳态网络有余量而保护控制导致大量站点同时转供，静态容量仍不足以说明动态风险。工程方案应针对实际约束提出，并说明该方案能解除哪一层、还留下哪些层。

13.2 资源充足不代表本地可交付

资源充足性通常在区域或系统层面评估。其输入包含负荷路径、资源组合、停运与天气情景，以及需求侧响应假设。它回答的是系统在概率性不利条件下满足负荷的能力，不直接回答某个客户能否从某一个节点得到额定功率。相反，接入研究可以发现本地网络瓶颈，却不负责证明整个区域长期拥有足够可调用能

SVSAC | sv-sac.org

源。将两个结论对调，会把区域风险错误投射到单个客户，或把局部工程可行误写成系统长期有余量。

对数据中心而言，电网接口的“可接入容量”也必须附带条件。研究需明确它是稳态正常方式下的最大取电、满足单一故障后限额的容量、可中断服务下的合同功率，还是含现场发电的净接口取电。一个园区可装设比并网容量更大的 IT 设备，但只有在控制方案、现场电源和限制条件能够在系统需要时履约的情况下，较低接口容量才成立。若备用机只是停电时启用、没有燃料保障、排放许可或并网控制验证，就不能把它当作规划中确定的发电能力。

可靠性评估中的同一负荷也不应只用一个年均利用率描述。一个实用的工程诊断表应至少分别记下观察点（电源端、区域、并网点或设施端）、时间步长、所需变量和失败判据：年度能量研究看逐时能量平衡与燃料/储能连续性；资源充足性研究看可用功率及失供概率；接入研究看节点电压、热限、故障电流和拓扑；动态研究看扰动轨迹、保护动作与恢复。不同层级的资料若只通过一个峰值兆瓦数汇总，关键约束会被抹平。夏季高温既可能提高冷却功率，也可能降低某些输电设备的承载余度；冬季燃料或设备条件又可能形成另一类紧张时段。数据中心自身在各时段的计算需求、维护和冷却状态与系统紧张时段是否重合，是可证伪的经验问题。缺少逐时数据时，规划情景应明确同峰假设，并至少比较不同相关性情景，而不把“平稳负荷”当作必然与系统峰值相遇或必然错开的事实。

13.3 输电、变电与配电约束的空间层级

数据中心讨论经常用“输电受限”概括从发电端到服务器插座之间的全部问题。实际链条可能包含区域输电、局部输电、变电站母线、降压变压器、专用馈线、机房中压系统和 UPS 输出。每个环节的额定值、应急限额、设备冗余方式与保护边界不同。NERC 大型负荷指南建议申请方提供并网点、分期投运功率、现场发电和电池、并联馈线及无功装置等数据，这正说明额定兆瓦不是足够的并网描述。[31]

在输电节点，瓶颈可能是热限、稳定限、相角或电压条件，也可能是已有发电和负荷的共同影响。配电侧则需进一步考虑变压器负载、馈线压降、谐波、故障电流水平和保护协调。一个主变压器具有足够铭牌容量，不自动意味着下游能承受快速负荷变化；一条线路满足稳态热限，也不自动保证故障后的电压恢复和相邻设施控制不会相互作用。因此工程报告宜列出约束元件、运行方式、负荷条件、失效假设和解除约束的措施，避免只给“系统有/无容量”的二元结论。

SVSAC | sv-sac.org

费用与工程可交付性也应分开。用户可能支付专用变电设施，但上游共享网络仍有扩建需要；反过来，区域共享投资也可能改善其他用户可靠性而非仅服务单一园区。本篇不据此推出任何具体成本分摊规则，只强调技术需要与财务分配必须用不同证据回答。源头负荷的接入申请可以触发研究，却不能独自证明全部相关建设均为该负荷专用或均应由其承担。

13.4 不同约束间存在反馈，但不可相互替代

四类约束在运行中会相互作用。高温可能提高数据中心冷却需求、压缩线路承载余量并增加发电机组故障概率；负荷转移到备用系统可暂时降低接口功率，却让区域发电和调频出现瞬时过剩；电池可能兼具短时削峰和扰动缓冲，但能量容量耗尽后不能延续服务。分析应把这些反馈放入同一时间序列，而不是将每类资源按其最大铭牌贡献简单相加。

以下是用于说明边界的作者设定情景，并非实测项目：某区域全年能源供应充分，日峰时段仍有资源备用，但园区并网变压器限制其首期接口；一次电压暂降又使园区 UPS 切换，导致其净取电骤降。此时增加年度能源采购不直接解除变压器限制，也不必然解决站内保护的共同动作。可用选项可能包括变压器扩建、分期投运、控制协调、负荷响应或现场资源，但必须分别核验工程时间、故障穿越和能源持续时间。该例说明约束可以同时存在，不能被称作某地区的真实案例。

13.5 工程研究应交付哪些可复核结果

为了避免不同研究报告中的“容量”不可比，我们建议区域规划与项目接入材料分别列明：负荷阶段曲线及其置信边界、逐时资源和能量限制、约束节点与设备、关键运行方式及故障集合、现场资源的燃料或充能假设、动态模型版本、代表性扰动响应、缓解措施的责任方和完成日期。每一项应注明是申请方陈述、系统研究结果、批准条件还是已投运测量。若只展示单一最大需求和单一新建线路方案，读者无法判断替代解释，例如既有老旧设备更新、其他区域负荷增长或系统可靠性标准变更是否同样推动工程。

研究步骤可以迭代，不是每个地区都必须按同一顺序建设。首先用能量和资源情景识别系统级风险，再将负荷映射到并网点并测试网络可交付性；针对关键扰动检查稳定性，最后比较网络扩建、供给组合、分期接入与履约响应。新证据若改变负荷时序、并网地点或站内配置，应回写上游分析，而不是只修改某一张规划图。该流程是作者为可复核性提出的研究设计，适用范围受地方数据权限、市场制度和运营方建模方法约束。

SVSAC | sv-sac.org

14 高密度电力电子负荷的扰动响应与建模

14.1 数据中心接入点背后是多层控制系统

电网看到的数据中心负荷并非一台等效“服务器电阻”。从并网点向内，设施可能包含变压器、整流器和开关电源、UPS、电池、机架 PDU、服务器电源、冷却泵与风机、变频驱动、无功补偿和备用发电。不同园区的拓扑、冗余策略、保护设置及设备组合可以不同。服务器电源和 UPS 中的电力电子控制可能在很短时间内改变有功、无功和电流响应；冷却电机和变频器则加入不同的热惯性和控制时间常数。应将站点单线图与实际参数纳入模型，不宜仅用行业平均 PUE 或一个固定功率因数替代动态行为。

设施能效设计也会改变这些变量，却不能替代电力系统动态模型。作为条件分析，若冷冻水蓄冷能在目标时段代替冷机耗电，它可能提供移峰；但还需储量、充放时序和可靠性测试，不能直接计作 AI 园区可调度 MW。文献 37 尚未取得支持原段，不作该特定设计效果的已核出处。[37] 建模中的“恒功率负荷”尤其需要条件说明。受控电源在一定电压范围内可能尝试维持输出功率，电压下降时会提高输入电流，从而改变电压恢复的负担；但限流、欠压闭锁、UPS 转供、设备保护和控制器优先级可能很快改变此行为。它不是可以无限维持的理想恒功率元件。把这一近似直接外推到深度电压跌落，会忽略电流上限和保护逻辑；反之，将所有设备当成固定阻抗，也可能低估控制负荷在电压变化时的响应。

电力电子负荷的时间变化还不等同于 AI 任务吞吐变化。计算工作负载可能改变服务器平均功率，但设施电表看到的净变化经过电源效率、冷却控制、UPS 损耗、冗余设备待机和站内电池状态调节。分钟级的功率轨迹与毫秒级故障响应需来自不同时间分辨率的数据。只有把服务器调度信号、设施计量、并网点测量和保护事件时间戳对齐，才能讨论某种计算策略是否解释了电网侧的波动。

14.2 电压暂降如何转化为系统有功扰动

NERC 对 2024 年 7 月东部互联系统事件的复盘记载，一次 230 千伏线路故障及其重合闸造成的 82 秒内六次故障序列期间，约 1.5 吉瓦以数据中心为主的电压敏感负荷同时降低了从电网取电；设施开关并非全部跳闸，部分负荷转向备用系统。报告将其描述为客户侧转供导致的负荷减少，并讨论电压敏感性及不同 UPS 架构。[32],[33] 这个区别很关键：从系统功率平衡看，无论是负荷脱网还是转至现场供电，净取电骤降都表现为需求侧阶跃；但设备状态、恢复过程和持续时间并不相同。将这类事件简单称为“电厂跳机”会误述因果，将其说成 IT 负载永久消失也会误述持续性。

SVSAC | sv-sac.org

同一事件不说明所有数据中心对故障有相同行为，也不证明 1.5 吉瓦是任意区域的可信故障规模。本篇将站点间及跨事件响应差异作为需检验的可能性，而不是已核实的同业主或同站点观测结果；文献 33、34 缺对应原段，只保留为待核线索。[33],[34] 电压跌落的幅值、持续时间和相序、保护计数逻辑、UPS 控制、备用电源就绪状态、馈线位置以及多次自动重合闸之间的间隔，都会改变响应。因此，区域模型应使用有证据约束的分布或分组情景，并公开聚合规则；不应把单个测点曲线复制给所有设施，也不应假设各站点完全独立而稀释共同模式风险。

负荷骤降造成的影响也具有位置与系统状态依赖性。有功需求瞬间下降会改变区域功率平衡和频率，局部无功取电同步变化则可能使电压升高或改变无功设备响应。事件复盘和 NERC 研究记录了该事件后的频率偏移、区域控制误差变化及局部过电压处置，但指出该地区电网相对强、频率没有升至引起担忧的水平。[32],[33] 这构成重要反例：一次较大负荷变化未导致失稳，不等于同样规模在较弱互联系统、不同运行方式或不同发生时刻无风险；反过来，也不能因为扰动机制可能产生高后果，就将最坏情景写作通常结果。

14.3 模型层级应由核心问题决定

稳态潮流模型用于检查指定运行点的功率、无功、电压与设备约束；短路研究关注故障电流、保护整定和设备承受条件。正序相量动态模型适于研究相对较慢的机电和控制动态、负荷骤增减、故障清除后的频率与电压恢复。电磁暂态模型则使用瞬时电压和电流，可表示电力电子控制、快速保护动作、相间不平衡与波形细节，计算代价与参数要求也更高。模型越细并不天然越可信：若设备控制和保护参数是默认值，复杂仿真仍可能精确地算错。

NERC 指南建议大型负荷提供站点专属、随设计更新为竣工状态的模型；正序动态模型应涵盖保护控制的动作阈值、恢复阈值和有功爬坡速度、负荷组成、恒功率/电流/电压特性、无功设备及平滑振荡的控制措施；指南还建议与设备供应商协作形成在需要时可用的站点级 EMT 模型。[31] 文献 30 所列模型页面未取得原段，本篇不把 Texas A&M 参与、PSCAD 实现或其通用性声明当作已核事实。[30] 一般模型只有在适用范围和参数可验证时才可用于情景讨论；不能在未核模型说明时将其称作真实设施的数字孪生。

有效研究应说明模型用来回答什么。如果关心热限和稳态电压，静态模型可能足够；若需评估一处大型电力电子负荷的故障穿越和控制耦合，可能需要 EMT；若关心大量站点在同一电压事件下的聚合频率响应，区域正序模型加上经

SVSAC | sv-sac.org

过校准的负荷群体特性可能更适合初筛，重要边界再用 EMT 校验。以超高细节对整个互联系统求解未必可行，以静态等值替代明确的快速动态风险也不足。合理的多层流程应先筛查、后针对关键节点加细，并记录模型转换造成的信息损失。

14.4 从并网数据到可校准模型

并网申请阶段的模型描述的是设计意图，不是已经安装的事实。NERC 指南建议模型随设计施工更新为竣工模型，且在满足其适用程序的范围内开展模型验证、投运前校验及投运后的事件数据分析。[31] 一个有用的数据包需包括并网点逐时/高频有功无功、相电压和相电流、事件标志、保护和 UPS 动作时间、备用发电与电池状态、冷却设备动作、负荷恢复轨迹及时间同步质量。若只给每小时能量，通常无法识别毫秒至秒级的转供过程；若只给设备面板告警而缺少同步电网测量，也难以区分电网事件与站内故障。

校验不应只看拟合均方误差。模型至少应在正常稳态、缓慢负荷变化、受控功率限制、代表性电压扰动、恢复过程以及合理保护边界上与独立记录相符。模型参数应能区分设备容量、控制逻辑和聚合站点数量，避免依靠调一个比例系数去拟合事件总幅度，却掩盖每个站点的动作阈值差别。对数据不足的参数，应给范围并分析敏感性；无法校准的模型只能用于定性探索，不能据其输出推导精确的可靠性裕度。

一个可复核的群体建模办法，是先按电气特征、UPS 类型、保护设定和并网位置将设施分类，再在各类内使用经过事件数据校验的响应参数。对同一扰动，可分别模拟独立动作、部分相关动作和共同保护逻辑触发等假设，报告系统结果区间。这是作者提出的研究程序，而非目前已有的全国数据库。若项目方不披露敏感控制参数，可在保密机制下交由具备授权的规划机构验证，并公布聚合参数范围、测试条件及结果摘要；若无任何可验证访问，则模型可信度限制应直接写出。

15 灵活性必须可履约、可测量、可复核

15.1 把“可调潜力”与可依赖的系统服务分开

数据中心灵活性至少包括四种物理机制：延后可延期计算任务、将任务迁移到其他站点、降低或暂缓部分工作负载、使用电池或现场电源改变净并网功率。它们不能用一个通用百分比代表。批处理训练可能允许等待或分段执行；在线推理、数据库事务、数据传输和有服务等级约束的任务则有不同的延迟与可用率边

SVSAC | sv-sac.org

界。任务迁移只有在另一地点存在可用计算资源、网络路径、许可与供电余度时才是系统可交付的空间灵活性。备用电源切换只能在其容量、燃料、排放许可、同步和运行控制满足条件时计为特定服务。

NERC 大型负荷指南建议资源充足性模型明确区分坚实与灵活负荷，表示最短运行时间、启动/停机、AI 训练时间窗口、可调功率、爬坡速率及表后电池和发电，并建议用真实运行界限、实际负荷测试或经验证的技术文件校准假设。[31] 这同样揭示了灵活性的反例：某个负荷在报价文件中允许下降，并不代表它能在系统最需要的高温傍晚持续下降；某个电池具备高功率逆变器，也不代表具有足够 MWh 在连续数小时支撑相同功率。功率与持续时间须分别核验。

15.2 约定服务产品的物理边界

我们建议在合同与测试计划中将可履约产品写成一组可测变量，而非只写“可削减 X 兆瓦”。第一，基线定义：以并网点净取电为对象，说明基线时间窗、温度和计算量调整、站内发电/充电以及计划检修如何处理。第二，响应规模：在指定净取电基线上最多下降或增加多少兆瓦，并明确功率因数或无功范围。第三，通知和响应时限：服务信号到测量点变化之间允许多长延迟，是否要求自动触发。第四，持续时间、可用时段、每日/年度次数和相邻事件最小恢复间隔。第五，解除约束后的恢复功率、恢复斜率和反弹负荷限制。第六，允许的例外条件，例如冷却安全温度、设备维护、通信中断和重大计算截止任务，以及这些例外怎样申报并留痕。

将需求灵活性纳入备选，与将其计作可靠容量是两回事。本次分析提出，应先测量 IT 与冷却控制在接入点的联合响应，明确调度权、调用时长、违约计量及通信或控制失效时的可用性，再讨论容量认定。文献 35、36 尚缺支持其具体方法或结果的原段，只列为待核政策与研究线索。[35][36] 本段是工程承诺与系统责任对齐的分析建议，不能替代当地费率、市场规则、监管批准或实际可履约合同。

15.3 分阶段验证协议

可履约能力应通过四个阶段验证。第一阶段是资料与模型审查：核对并网点拓扑、保护及 UPS 设置、可控工作负载清单、服务器/冷却功率关系、表后电池和发电机参数；分别保存设计值、竣工值和试验值，并标注版本。第二阶段是投运前的控制器在环或设备试验：在隔离且安全的条件下发送限功率、恢复及通信丢失信号，验证调度命令传递链、保护边界和失效安全状态。此类试验可以发现集成问题，但不能证明真实园区群在电网扰动中表现一致。

SVSAC | sv-sac.org

第三阶段是在并网条件下执行分级响应测试：先测量正常工况下基线波动，再逐级测试小幅削减、目标功率、持续时间、恢复和重复调用。必要时分开测试可延期工作负载、蓄电池、冷却热惯性和现场发电，避免几种手段同时变化时无法识别贡献。任何故障穿越测试须由责任运营机构和设施安全程序批准，不能为验证灵活性而人为制造危险的系统故障。第四阶段是常态监视和事件复核：建立并网点高分辨率计量、时间同步、状态标志和保留期限；每次调用后比对承诺与实际净响应，记录缺额、响应延迟、持续能力和恢复反弹；在设备、软件、主要工作负载或保护设定发生实质变更后重新校验。

测试基线本身也应有审计规则。若采用事件前若干分钟的净取电作为反事实，应检查该时段是否恰逢服务器作业切换、备用电池充电或冷却设备启停；若以历史同类日建立基线，应披露温度、IT 任务量和设备可用状态的匹配方式。基线若被调用方事后选择，或采用园区申请的最大容量作为“本来会用”的负荷，测得的响应会系统性偏高。建议在测试前冻结基线算法、缺失数据处理、时钟误差容忍值和独立电表位置，并用未调用的对照日检查基线偏差。这是测量方案建议，不能替代运营方对安全运行和计量规则的批准。若要将服务用于资源充足性，测试样本还应覆盖不同季节、气象、负荷水平、任务组合和电池荷电状态，并对可用性失败率作保守处理。一次成功测试仅证明在该次条件下完成交付，不证明全年每小时都有相同能力。若服务提供者能够选择不利时段不参加，规划中需显式表示此选择权，而非仍按完全可用折算。对具有相关天气或共同软件平台的多站点，也应评估其共同失效概率。

15.4 防止错误承诺、反弹与重复计算

灵活负荷提供的不是凭空消失的能源。延后任务会在之后重新运行，除非客户取消或降低计算服务；迁移任务会增加另一处设施取电及网络通信；降低冷却负荷可能把能耗移至稍后时段；电池放电后需要充电。核算应报告基线期、调用期和恢复期的完整能量轨迹，区分净能量节省、时间转移和地点转移。若把调用时的峰值减少当作年度节能，会夸大其对电量充足性和排放的影响。

设施侧与系统侧计量还要避免重复记账。如果系统运营方将同一电池作为储能资源调度，数据中心又将其计为负荷削减，两者不能在同一时段都按全额计算。现场发电机若已有容量义务或环境运行限制，也不能仅因安装存在而将其全额列作可信备用。灵活负荷可能降低接口峰值或推迟特定网络升级，但其能否解决持续能量短缺、N-1 输电约束或故障穿越问题，需分别证明。

SVSAC | sv-sac.org

反例同样重要：如果园区负荷响应时，受限线路的上游另一条支路仍拥塞，削峰可能无法解除拥塞；如果同区的所有数据中心同时提高计算以赶回任务，恢复反弹可能构成新的峰；如果园区本身已在低负荷运行，继续削减的绝对量可能很小。另一方向上，若系统稀缺只出现在有限小时，且园区有可核验的延期任务和足够电池，灵活接入可能在建设期间提供过渡支持。上述判断必须依赖网络位置、时间序列和真实合同边界，而非泛泛认为 AI 负荷“天生可调”或“完全刚性”。

15.5 从验证结果到规划可用量

建议将计划容量与可依赖容量分列。计划容量描述设施在声明条件下提出的最大灵活规模；可依赖容量则由独立计量、合格测试、可执行合同、可用性历史和系统位置共同限定。该转换可按服务类型分别进行：例如，事件前快速削减关注响应时间与连续可用功率，长时削减关注任务可延期窗口和电池能量，空间迁移还需检查目的地电网是否同受约束。简单从技术潜力乘以主观折减系数，不能替代这些逐项约束。

可在研究中使用一个说明性计算模板：某服务的可依赖功率，取测试验证功率、合同上限、系统可调用上限和当前设施可用余量中的最小值；可依赖持续时间再由合格试验与能量/工作负载边界中的较小者确定。这个模板只是一种审计性定义，并未被本篇估计为任何真实项目的数值，也不能覆盖所有市场规则。若不同环节的计量点、净值口径不同，应先统一边界再取最小值，否则会把毛发电、IT 削减或馈线级响应错误地相加。

这种验证逻辑不只保护电网，也避免将工程责任模糊地推给数据中心运营商。需要求客户承担的响应范围、设施安全例外、测试标准和保密机制应可预见，且与其获得的接入条件或补偿相关。合理的协议应同时规定系统运营方提供的预警、命令确认、计量标准、争议处理和不可抗力规则。若这些条件尚未形成，就应将灵活性留在敏感性情景中，不应把它当作已交付的基础资源。

16 新增算力负荷的排放归因：年电量相等不代表逐时零排放

16.1 先说清楚“排放归因”回答哪一个问题

一座数据中心新增用电，至少会触及三种不同的核算问题。第一是企业温室气体盘查：按适用规则，把购买电力的排放分配给某个组织。第二是设施的历史足迹：用其所在电网和用电时段估算电力生产端发生的排放。第三是新增负荷的后果：相对于没有该负荷的反事实，电厂调度、跨区交易、清洁电源建设和退役

SVSAC | sv-sac.org

计划怎样改变。三者分别是会计归属、历史足迹和因果后果，不应以一个“碳强度”数字代替。

对历史盘查，年度平均排放因子通常适于回答“这一年平均每单位用电分配多少排放”。EPA eGRID 建议用电力消费所在 eGRID 子区域的年度输出排放率；子区域设计在较大 NERC 区和较小平衡机构之间折中，意在近似电网物理约束并减少跨区输电影响。[41] 这比全国平均更贴近地点，却仍把小时差异压缩成年度数值，也不保证代表新增 1 MWh 所导致的边际调度变化。EPA 另给出的 nonbaseload（非基荷）排放率依据机组运行水平把发电量和排放分类，包含负荷跟随机组和尖峰机组；它可估算效率或可再生项目可能避免的排放，但由于并非按目标时刻的机组承诺、约束和边际出清计算，不应被当成精确的逐时因果排放因子。[41]

新增负荷的因果分析则应明确反事实。若某小时可再生能源本来会被弃发，新增算力可能吸收这部分电量而少增化石发电；若该小时新增负荷由燃气或煤机增发、进口电力增加，短期边际排放可能高于区域年度平均。若负荷促成新电源或输电建成，长期效果又取决于项目额外性、建设时序、退役替代和系统是否因此少建其他资源。以上是一般电力系统机制，不能由单份购电合同或年度总量直接证明。研究设计应分别保存年度盘查、逐时消费足迹及有明确反事实的增量情景；在缺少出清、机组与合同数据时，结论只能是范围或情景推断，而不是实测因果效应。[40],[41],[44]

16.2 位置、时段和电力交换会改变平均值的含义

LBNL《2024 年美国数据中心能耗报告》将 EIA-930 平衡机构间小时交换数据，与 EIA-923 机组发电、EIA 水耗数据和 EPA 温室气体数据结合，估计各平衡机构供应电力的环境强度；报告再以设施所在县对应的平衡机构赋值。[40],[42] EIA 说明其小时需求数据不按行业细分，且不包括直接由分布式发电供给的负荷；发电数据也不覆盖所有不由平衡机构监测的资源。[42] 因而这种估算增加了地理区分能力，但小时交换数据作为输入，并不使下述年度总量成为逐时足迹；它仍不能识别具体数据中心的表后供电或 AI 用电份额。这是一种比国家平均更细的历史足迹估算。其结果显示差异具有明显地域性：报告将 2023 年全美数据中心约 176 TWh 用电乘以本地电网结构，估算约 8000 亿升间接用水和 610 亿千克二氧化碳当量；平均强度分别约为 4.52 升/kWh 和 0.34 千克 CO_2e/kWh。[40] 上述 2023 年总量按本地平衡机构年度平均强度作模型估算，非逐时或逐设施结果；同一平衡机构内所有数据中心被赋予相同电力组合，且方法未纳入个别 PPA 及表后发电，报

SVSAC | sv-sac.org

告也明确说这些缺失可能显著改变估值。[40] 因此这组总量适合揭示间接环境负担的量级和区域差异，不适合当作每座 AI 园区的足迹结论。

逐时足迹的构造也有分辨率代价。Siddik 等人提出 Water IMPACT 框架，以平衡机构为供需单位，组合小时电源构成和跨机构交换，并以可取得的发电、排放和冷却水数据计算水与碳强度。论文展示了不同地区、季节与小时的差异，并提供历史和实时分析工具。[44] 但其实时估计并非直接观测每座设施的电源：机组级水耗数据按月发布、主要覆盖 100 MW 以上热电机组；小型机组缺漏要用燃料类型平均值补足；eGRID 排放强度的机组数据按年，模型假定同一平衡机构/燃料类型的强度在年内不变。[44] 作者明确指出当月机组数据尚未发布时，以最近可用同月强度估计当前月份；所以实时仪表盘是以延迟、代理和固定系数生成的估计值，不是调度实测的小时全生命周期影响。

在边际模型与平均核算间存在真实张力。平均因子便于年度组织盘查、可复算且可与既有年份比较；边际模型用于政策或项目的后果分析，更依赖电厂启停、输电拥塞、储能状态、进口来源和反事实。短期边际与长期边际也不等价：前者反映既有资产的调度，后者可能包含新建、退役与负荷迁移。把边际结果硬塞进盘查会破坏一致性；把年度均值用于预测新增负荷的系统后果则可能遗漏关键变化。可检验的研究路径，是并列报告地点年度平均、可用小时强度和有界的短期/长期反事实情景，并逐项列明资料缺口，而不是把不同问题合并为“零碳供电”标签。[41],[44]

16.3 PPA 与证书是合同及属性证据，不自动构成逐时物理供电证据

长期购电协议（PPA）可能是新电源融资与商业收益安排的一部分；可再生能源证书或其他能源属性证书（EAC）可将环境属性归属到购买者。它们对合同关系、属性核销或企业市场法盘查可能有意义，但年度数量相等本身不能证明设施用电的每一个小时都由同一电源物理供应。电网按当时可达的发电、跨区交换、储能与负荷共同平衡。一个远端风电项目的年度发电量可以抵销全年用电量的市场法账面属性，同时在无风或线路受限的时段，设施仍从所在电网取电；此时“年度匹配”与“该小时的物理供应组合”是两个命题。反过来，不能仅凭合同是虚拟 PPA 便说它毫无系统影响：有长期收入保障的项目可能促成新增投资，但是否额外、是否改变调度，还需项目融资和系统反事实证据。

GHG Protocol 在 2026 年 10 月的进度说明中记载，现行市场法规则仍源自 2015 年 Scope 2 指南；2025–2026 年公开咨询提出过逐时匹配和可交付区域要求，但截至该状态更新，委员会仍在讨论双路径方法与技术选项，不能把方案写

SVSAC | sv-sac.org

成已生效规则。[43] 该机构对现行规则的描述也指出，当前市场法可允许企业报告与其运营所在电网不相连的清洁能源项目所对应的减排。[43] 这恰好说明会计规则中的属性归属可能宽于物理供电边界，不表示企业故意虚报或合同本身无效。采购与核算分析应分别标注：合同电量、证书唯一注销、发电项目所在区域、消费位置、匹配时间粒度，以及供应物理可达性。没有这些字段，市场法“100%可再生”不能直接翻译成“逐小时物理零排放”。

同时，逐时匹配也不是完整的排放因果识别。即使某企业实现逐小时、同一可交付区域的清洁能源匹配，这可以提高采购与负荷在时间、地点上的一致性；是否带来额外减排还受新增性、替代何种机组、网络瓶颈和项目建设时间影响。GHG Protocol 讨论中把新增性也列入一种拟议方法，而 2026 年状态仍保留开放选项。[43] 因此工程报告可把逐时匹配当作更严格的采购/核算设计候选，不应把它等同于已经验证的系统减排结果。

17 冷却用水与电力需求的耦合：地点特定的约束而非固定换算系数

17.1 直接、间接与水务系统的三条路径

数据中心环境足迹至少有三条相互连接但地点不同的水能路径。第一条是设施直接用水：蒸发冷却塔消耗水以排热，取用量中有部分蒸发、部分排放或回用；干式冷却可减少现场蒸发，却可能提高风机、压缩机或冷机电耗。第二条是发电侧的间接用水：燃煤、燃气或核电机组的冷却取水与耗水随冷却方式、效率、热源温度和运行地点而变；水电水库的蒸发分配还存在方法学争议，不能把“非热电”理解为绝对无水影响。第三条是供水与污水处理的嵌入式电力和水耗：抽水距离、提升高度、处理要求及污水处置路径不同，同一立方米水在不同水源地对应的电力消耗可能不一样。[40],[44],[45]

必须把“取水”与“耗水”分开。取水量描述从河流、湖库或含水层取用的水量，其中部分可能回流。取水、回流与排水过程可能影响进水口生态和时段水位，热电排水还可能改变受纳水体温度；耗水强调未回到原水体、通常因蒸发或其他损失而离开该水文系统的部分。单一“升/kWh”或“年用水”不能同时表达流域短缺、季节峰值、水质、回流水温、回用比例和水权。从工程条件推导，低水位、高源水温或排水温度约束可能限制热电冷却；是否降出力或停运须核具体机组。文献 46、49 仍缺支持对应断言的已读原段，不据其作全国现况或事件事实。[46],[49]

SVSAC | sv-sac.org

LBNL 的模拟显示，蒸发型水冷系统通常能以较低冷却电耗排热，但其现场 WUE 高；气冷冷机不用现场水，却需更多电力。报告还明确区分 site WUE 与 source WUE：前者只计设施水，后者把电力生产耗水纳入。[40] 这意味着“无现场耗水”冷却并不自动降低整个电力–水系统的耗水，也不自动改善排放。如果提高电耗发生在高水耗热电机组占比较大的区域，间接用水可能抵消一部分现场节水；若新增电量来自风光或其他低运行耗水资源，替代方案的水结果就可能更有利。该判断取决于机组组合和季节，不能以冷却技术名称单独推出。

同样，液冷不是“零水冷却”的同义词。直接到芯片或浸没式循环可使冷却液在封闭回路内传热，但热量仍须经干冷器、冷却塔或其他设备排向环境；若使用蒸发辅助，仍有现场耗水。LBNL 报告把液冷设计与不同的干式、湿式辅助系统分别模拟，指出液冷较高的工作温度可以扩展自然冷却或减少机械制冷，但水和能效取决于末端热排放设备和气候。[40] 将芯片侧的闭式回路宣传成园区整体无耗水，混淆了设备回路与最终热汇。

17.2 以流域风险与小时负荷联合选址

工程比较不应只问年水量，而要问水从哪里来、什么时候可供、是否与既有居民和生态需求竞争，以及供电侧水足迹落在哪些流域。数据中心用电所在地的平衡机构，与发电机所在流域并非一一对应；电力跨区交换会把间接水耗转移到异地。Siddik 等人的 2018 年美国数据中心运行足迹研究将设施连接到发电厂、水务设施和污水处理设施，估算直接和电力间接耗水，并用水稀缺因子对体积水量加权；其结果显示直接耗水集中在一些已有水压力的流域，而间接足迹分布在电源所在地。[45] 这项研究针对 2018 年、采用模型关联且不含服务器制造等生命周期环节；它可支持空间机制和方法，不可直接作为 2026 年 AI 专用数据中心的当前数量估计。

实际筛选可按站点、流域、气候和电网四个尺度分层。站点层应对比取水许可、峰值日水量、来源比例、回用水条件、排水温度与污水能力；流域层需看枯水季、水位和其他用水竞争，而不是只用全国年均；气候层应将热浪与高湿、干旱联动，因为高温常同时增加制冷负荷和用水压力；电网层则关注供应时段和电源水强度。此框架是基于来源所示机制提出的筛选方法，不是已有统一许可标准。[40],[44],[46],[49]

替代解释也必须纳入：若某区域枯水年同时有较高存量储水、再生水管网富余，或数据中心可采用不依赖淡水的空气侧自然冷却，则新增设施未必挤占居民供水；若一座干旱区设施通过异地电源购电，现场和发电侧风险也可能分别落在

SVSAC | sv-sac.org

不同区域。相反，一个湿润地区也可能因热电机组集中、河流热排放约束和电力高峰重合而面临供应端水风险。全国水资源丰沛度或州级年度总量不能代替流域级、季节性可用水分析。

17.3 以反例检验“节水必省电”和“清洁电力必低水”

反例一：在气候温和、夜间低温足够且设备允许较高供水温度的站点，干式冷却可以同时减少水和冷机用电；但同一设施在炎热季节仍需要机械制冷或短时蒸发辅助。因而年均 WUE 和 PUE 掩盖了峰季性能，比较应包含逐时温湿度、设备负荷曲线和备用运行策略。本篇将设备进风温度、自然冷却、热回收与干式排热列为待逐项评价的设计选项，不能在未取得文献 37 对应原段时称其为该指南的确定排序。[37] 这些选项不是任何气候区可直接照搬的性能保证。

反例二：有充足再生水、排放许可和处理能力的园区，采用蒸发冷却可能减少电力强度，并避免从饮用水源取水；但“再生水”仍有输送与处理能耗、管网资本成本、停供风险与排水盐分管理。若用水服务机构在旱季需要优先保障基本公共用途，合同可用量也不能假设全年固定。反例三：对水电占比较高地区，按设施当地电网年度平均因子得到低碳结论，并不必然意味着低水耗；水库蒸发量的归属、蓄水的多重用途、自然蒸发基线和季节变化都会改变水耗核算。反之，若将水电库区全部蒸发都分配给发电，也可能夸大发电的增量水代价。应报告分配方法和敏感性，而非只给单值。[44],[45]

因此，站点方案至少应并列报告设施直接取水与耗水、对应的季节峰值、发电端水强度及其地域、耗电变化、温室气体足迹和水稀缺加权结果。若只有年总取水或单一 WUE，不能回答“节水是否把负担转移到电厂所在流域”；若只有电网用水因子，也无法回答社区的现场供水约束。披露缺失将使选址比较更不确定；这是一项数据限制，不自动证明任何项目存在绝对水资源冲突。

18 供电组合、现场发电、储能与建设时序的多目标权衡

18.1 不是技术清单，而是约束条件下的组合问题

新建算力设施需要的电力方案不能仅按年电量相加。规划者需要同时审视：可靠容量与故障风险、逐时发电形状、接入和输电能力、碳排放、发电与冷却用水、项目交付周期、燃料和设备供应链、土地及许可、全寿命成本、资产搁浅与社区影响。没有一种资源能在每个维度都占优；组合的价值来自资源之间互补以及建设顺序适配负荷实际爬坡，而非把“清洁”“可调度”“本地”各自当成绝对优点。[48]

SVSAC | sv-sac.org

风电与光伏可以提供低运行排放的能源，但其出力随天气而变；储能可在充电与放电时段间转移电量、提供短时功率支撑，却有能量容量、持续时长、效率损失和充电电源来源约束。把某块电池的 MW 额定功率当成可替代同等 MW 的持续电源会高估其长时间供电能力。核电、水电、地热或其他可调度低碳资源在适用地点可提供较稳定出力，但项目扩建、许可、地理条件和建设周期限制其普适性。天然气机组可被考虑作为快速可调度资源或过渡方案，但化石燃烧排放、燃料供应、管网能力和长期锁定风险需要同时计入。我们建议对现场与电网供电分别评价成本、可靠性、环境、供应链与技术成熟度，并按平坦或可变负荷设情景。这是作者分析框架；文献 48 缺对应原段，不据此归为 DOE 已核政策结论。[48]

18.2 现场供电并不天然独立于电网，也不天然更可靠

表后电源可能通过绕开某段输电排队缩短项目接入时间，也可能因现场燃料设施、排放许可、设备检修与微网保护而增加新的故障模式。若采用燃气轮机或发动机作为连续主供，必须核对燃料可中断性、天然气管压、同期热浪下燃气与电力系统的共因风险、排放许可时数及备用机组规则。若表后资源与大电网并联，故障穿越、保护协调、同步与谐波等接口要求仍然存在；“behind-the-meter”只是计量与所有权边界，不是自动脱离系统约束。本篇据这些边界提出，应比较表后与电网供电的成本、充足性及可靠性；不把尚缺对应原段的文献 48 作为该命题的已核出处。[48]

柴油或燃气备用机组通常按应急用途设计，若改作常态主供或长期桥接，就不能只引用其备用额定功率。运行小时增加会改变燃料、排放、空气许可和维护负担；若用电池替代部分备用需求，则必须校验其持续时长与极端事件补能路径。前述是设备运行的工程核查方向，并非针对某个项目的许可判断。电源合同、设备铭牌、实际可用容量和环境许可证支持的是不同事实，任何一项都不应代替其余证据。[48]

18.3 把时序风险作为组合变量

如果把完整园区、输电扩建和新电源假定在同一天可用，容易得到不存在的“同步交付”。一种更可检验的计划是分期增加计算和供电容量：先部署能效更高的 IT 与热管理设计、可用的现有清洁资源和必要的接口改造；后续负荷爬坡达到门槛后再投建新增模块。该方案需要保证分期建设不破坏冗余设计，也需避免第一期资产被误当作第二期确定负荷的永久解决方案。这里提出的是风险管理建议，并非 DOE 已经发布的强制序列。

SVSAC | sv-sac.org

时间尺度会改变技术排序。若近期瓶颈是变压器或线路工期，短期现场电源可能缓解部分接口压力，但并不能自动增加长期低碳电量；若瓶颈是能源不足而非局部输送，仅增加接入设备也无济于事。若项目负荷爬坡慢于预期，提前采购的大型专用电源或输电设施可能低利用率；若负荷先于资源投运，电网临时供电可能提高短期排放或系统成本。评估应为每项资源写明最早可用年、爬坡期、容量因子假设、检修和审批风险，并对负荷推迟、燃料受限、建设延期和价格变化做敏感性分析。对预测的新增需求，不能把计划投运当成既成供给或负荷事实。

18.4 用帕累托比较保留价值判断，而非伪造单一最优解

建议在候选组合中建立“可靠服务–排放–水–成本–进度–系统风险”多目标表，每一维都标清数据年份、边界、观察/假设和权重来源。若一种组合在可靠性和排放上更好、成本不高，且用水、进度及其他已列目标均不变差，才可在该目标集合下称为帕累托改进；若其他目标变差，就仍是权衡。现实中更常见的是权衡：水冷风光加储能、干式冷却加更多电网电量、现场燃气加后续清洁供电、或分期降低早期冗余投资，往往在不同指标上各有利弊。给这些目标赋单一权重最终会体现决策者的价值选择，应公开权重并保留未加权结果，避免把政策偏好包装成技术事实。

一个反事实检验可帮助识别组合收益：在相同服务器 IT 负荷、气候与可靠性标准下，比较基准设施、提高 IT 与冷却效率、改变冷却方式、增加现场清洁供电/储能、以及改变并网与项目时序的组合。模型需按小时或更细时间粒度处理资源出力、冷却负荷和电网取电；年度匹配只适合年度账面比较，不足以确认极端高温时段可用性或逐时排放。再用不同流域水稀缺权重、不同燃料价格和资源建设延期情景检验排序是否稳健。模型输出是条件性情景，不是项目实绩，也不应伪称经现场测量验证。

特别需要防止“效率反弹”被误读。单位算力能耗下降，若计算服务量随之显著扩大，设施总能耗仍可能增加；但总量增加也不证明效率措施无价值。应把单位服务能耗、设施总用电、服务量及边际用能分开。可用技术效率、调度和选址来减少单位冲击，却不能预先假设节省下来的电力一定被社会其他需求吸收，或一定全部被新增算力抵销。数据中心商业模式、算力利用率与任务迁移都会影响这个反事实；没有服务级工作负载数据时，只能对反弹机制进行敏感性分析。

SVSAC | sv-sac.org

19 将分期选择写成可复算的两阶段问题

19.1 模型要回答的有限问题

前文已说明，项目预期需求、实际交付与风险承担可能分离。但仅说“应分期”仍不足以指导研究。分期可以减少低需求状态下的提前投资，也可能在高需求状态下形成延误和追加费用。本节问一个更窄的问题：在相同候选容量和成本边界下，当规划者只能获得有误差的需求信号时，分期的期望成本优势是否仍然成立？若不成立，是需求概率、信号质量还是追加建设条件改变了比较？

这不是美国全国投资优化，也不是以案例数据估计需求概率。模型舍弃地理网络、逐时发电和融资主体，只保留需求状态、初期容量、后续信号和追加行为。舍弃是为了使每一步推理可追溯，而非断言被省略因素不重要。实际应用必须把这些因素带回相应工程与合同分析，不能把这里的简单模型扩展为真实电网的安全证明。

19.2 状态、行动与成本边界

设一期决策面对三种最终新增需求状态：50、100、200 兆瓦，作者分别赋予概率 0.5、0.3、0.2。这里的新增需求不包含已有客户总负荷，也不表示任何实际园区。假设初期每兆瓦容量成本为 1 个现值单位，未满足新增需求每兆瓦损失为 4 个单位。损失项只是为比较服务不足的后果而设，不是可靠性标准、停电损失的实测估值或允许违反既有合同的费用。

比较五个策略。策略 A 不增加容量；策略 B 立即建设 100 兆瓦且以后不扩；策略 C 立即建设 200 兆瓦；策略 D 先建设 100 兆瓦，在完全知道最终状态后仅在 200 兆瓦状态下扩容；策略 E 也先建设 100 兆瓦，但是否扩容由有误差的信号决定。策略 D 是完全信息基准，不能在没有证据的情况下当作现实可保证策略。策略 A 和 B 若不满足实际服务义务，也只能作为损失比较基线，不能视作合规可选方案。

假设后续增加 100 兆瓦需 125 个现值单位，高需求状态下还产生 50 个单位的过渡服务损失。低需求时错误扩容仍需 125，但不发生这一高需求过渡损失。策略 E 在非高需求状态下误判为需要扩容的概率为 10%，在高需求状态下漏判的概率为 20%。它们是条件错误率，不能与最终高需求概率混用。所有输入已按现值计，不再额外施加折现；模型不含不同策略的差异化运行成本、残值和融资成本。

这些省略构成解释边界。例如若分期使重复施工或采购费用显著上升，应体现在 125 中；若延期导致服务违约的后果更大，应体现在过渡损失或硬约束中；

SVSAC | sv-sac.org

若剩余容量能转用于其他用户，应加入残值或共享收益。不能一方面省略这些项，另一方面声称所得排序已覆盖全生命周期社会成本。

19.3 结果的计算方法

对固定建设策略，状态成本为初期建设成本加未满足需求损失。例如 B 在 200 兆瓦需求状态下成本为 100+4×100=500；C 在三种状态下成本均为 200。对完全信息分期 D，低中状态只支付 100，高状态支付 100+125+50=275。对有误差信号 E，低中状态的条件期望为 100+0.1×125=112.5；高状态为 100+0.8×175+0.2×400=320。

策略	50 兆瓦状态成本	100 兆瓦状态成本	200 兆瓦状态成本	按设定概率的期望成本
A 不建设	200	400	800	380
B 建 100 且不追加	100	100	500	180
C 一次建 200	200	200	200	200
D 完全信息分期基准	100	100	275	135
E 有误差信号分期	112.5	112.5	320	154

E 行前三列是给定需求状态、对信号结果取平均后的条件期望，不是每种状态下唯一可能的实际成本。这一点不能省略：在高需求状态下，正确扩容成本为 275，漏判不扩成本为 500；将 320 称为最坏成本，会把风险低估。复算文件保留了需求与信号的六个联合结果及其概率，避免只展示平均数。

20 期望优势、尾部风险与信息价值

20.1 平均成本较低不代表在每个状态更好

在上述设定下，E 的 154 低于 C 的 200，但在高需求且漏判时，E 仍会达到 500。C 在三种需求状态下始终是 200。若比较实际最大成本，E 的优势消失；若只比较期望值，E 仍较低。两种准则关注不同后果，不能仅因一个数较小，就声称一个策略在所有意义上更有效率。

完全信息基准 D 的 135 又低于 E，但这一差额依赖规划者准确识别最终需求的能力。把 135 当作可实现成本，会把信息误差造成的 19 个期望单位忽略。真实

SVSAC | sv-sac.org

项目的信息也未必在追加投资之前充分揭示，且部分信号来自客户自身申报；应检查它是否对最后真实负荷具有预测力，而不是仅与合同容量重复相关。

用后悔值也要明确基准。本篇将每种真实需求状态下、表中候选策略可达到的最低成本作为比较值，三种状态分别为 100、100、200。C 的最大后悔为 100；E 按实际需求与信号联合分支计算的最大后悔为 300。若先把信号平均，E 最大的条件期望后悔仅为 120，这个量不能替代实际分支最大后悔。统计聚合的顺序因而会改变风险指标。

这不是推荐公用事业采用最坏情景或后悔最小化。它表明，报告必须说明优化目标，不能在论证中悄悄把期望成本换成尾部安全。对于真实服务义务，某些失供结果可能根本不可行；此时应先施加工程与合同约束，再比较剩余策略的成本，而非让一个任意低损失系数把不合规策略“优化”为最优。

20.2 需求概率阈值如何随信号误差变化

为检验概率敏感性，设高需求概率为 h，其余两种状态按 5:3 分配剩余概率。D 的期望成本为 100+175h；E 为 112.5+207.5h；C 保持 200。由此得到，D 相对 C 的交点为 h 约 0.5714，E 相对 C 的交点为 h 约 0.4217。引入信号误差后，分期保持平均成本优势的高需求概率范围缩小。

h	D 完全信息基准	E 有误差信号	C 一次建 200
0.10	117.50	133.25	200
0.20	135.00	154.00	200
0.40	170.00	195.50	200
0.60	205.00	237.00	200
0.80	240.00	278.50	200

这个阈值只在所列成本和状态下成立，不是美国电网投资的经验临界值。若追加成本、过渡损失或漏判损失变化，交点也会变化。因此，研究不能先选择一组使分期占优的输入，再把结论概括成普遍规律。更有价值的做法是公开交点及其条件，让读者知道在哪些设定下原建议会被推翻。

还应注意，概率与制度未必独立。更严格的承诺可能改变高需求实现概率，也可能改善信号质量；高额最低收费则可能改变客户申请容量。在真实应用中，应明确将这些变化置于哪一层：它是需求分布变了、信号条件分布变了，还是损失承担变了？三者若同时调整而不说明，模型结果就无法解释哪项政策起作用。

SVSAC | sv-sac.org

20.3 数据收集的价值取决于它减少哪一种错误

在 E 中，非高状态错误扩容的期望代价为(1−h)×0.1×125，高状态漏判相对于正确扩容的额外代价为 h×0.2×225。基准 h=0.2 时，两项分别为 10 与 9，总计 19，与 E 和 D 的差额一致。这个拆解说明，改善“预测准确率”不能只报告一个总比例；它必须说明减少的是误扩容还是漏扩容，以及两类错误的成本。

例如，将漏判率从 20%降到 10%，其他条件不变，基准期望成本下降 4.5 个单位；将误扩容率从 10%降到零，则下降 10 个单位。两种改进并非因为后一类错误总是更重要，而是本例高需求概率较低和条件成本共同作用。换一组需求概率或损失，排序可能反转。因此，不应把模型中的 10 与 4.5 移用于真实审核预算。

在相同边界下，信息改进的期望收益还应与取得信息的成本比较。现场工程研究、信用审核、合同执行和独立计量都可能有成本，也可能具有模型之外的安全与治理收益。完全信息差额 19 只是在初期 100 兆瓦决策固定、第二阶段行动与成本边界不变且未计信息取得费用时的信息改进基准，不是可以无条件用于购买数据或追加审核的预算。若信息取得太晚，即使准确也未必改变建设决策；若它只能在客户已退出后确认退出，其前瞻价值也很有限。

21 从说明模型转向可验证的应用

21.1 参数需要不同证据，不能由一个行业均值包办

实际应用须分别匹配四类参数：需求状态用项目/运行记录，追加成本用同范围工程报价，延误后果用服务约定与业务边界，误判率用既定规则下完整历史样本。信号价值取决于最后一项与决策损失的联结，不能由一个行业均值包办；由规划容量反推实现概率再验证原规划尤其构成循环。

工程成本也应按范围对齐。一份含土地、许可、设备、施工和网络升级的总报价不能与仅含某台设备的出厂价比较；一次建成与分期建设的报价应保持相同服务水平和期间。若某方案在第二阶段使用不同技术或不同节点，它已经不是本例只增加同一 100 兆瓦模块的行动，需要重新定义可行集和成本，而不能只替换 125 这个数。

需求信号的校准应保留未实现项目。只用最后建成园区来训练模型，会丢失误扩容与退出的样本；只用接入队列，又可能不知道运行结果。需要将同一项目的申报、承诺、研究、施工、接电和运行连接起来，并把无法跟踪的记录留作未知。高质量信息的价值来自它能在决策之前改变判断，而不是事后把成功项目解释得更完整。

SVSAC | sv-sac.org

21.2 可靠性与财务回收应进入不同约束

最低付款和担保可改变投资回收，但不能解除节点电压、设备热限或扰动响应约束。本例若把欠供损失全部视为可货币补偿，容易忽略不可接受的可靠性结果。实际扩展应先规定哪些状态必须供给、哪些服务允许中断、哪些恢复时间可接受，再计算符合这些条件的策略集合。

同样，信用担保降低公用事业现金流风险，并不自动降低全社会资源成本。若只是把损失从一家企业转给另一主体，财务结果改善与资源效率改善需要分开说明。反过来，可靠性更强的工程可能降低服务中断损失，但增加付款压力。论文应同时报告物理履约、真实资源耗费与各主体财务结果，避免以某一方账单最小替代整个系统目标。

在这个意义上，本篇模型与前文 GS-5 机制分析互补：费率条款决定部分资金义务，需求信息影响扩容行动，工程研究决定行动是否可行。三者并非同一个“风险分数”的不同名称。若要形成实际建议，应明确每个参数的责任部门和证据来源，并记录版本变动是否需要重新比较方案。

21.3 稳健性应覆盖模型结构而不仅是参数

概率阈值的变化只检验 h，不检验行动集合。临时设备、其他站点、可履约中断服务或共享容量可能改排序；一次性订货或不能重入的走廊则可能排除分期。因此，信号比较还须报告增删关键行动后的结论，不能把平滑敏感性曲线当作结构稳健性。

损失函数也可能不是线性的。10 兆瓦短缺发生在某个关键计算环节，后果可能不能按 100 兆瓦短缺的十分之一计；连续多日延误与一次短时限制也可能具有不同后果。此时应使用分段损失、期限约束或明确服务等级，说明所选形式的依据。复杂形式仍需数据，不应用非线性符号掩盖缺少估值。

本例还默认信号不会受到客户策略行为影响。真实情况下，申请者可能根据容量分配规则改变申报、担保或阶段承诺。若制度本身改变信号生成过程，过去误判率可能不再适用。研究需要观察新规则下的后续表现，或将策略行为设为敏感性情景，而不能把一次历史校准永久冻结。

21.4 可复核结果的呈现

本次分析为该模型保存输入、脚本、完整联合分支、期望成本、实际最大成本和阈值。复算核对了完全信息 135、有误差信号 154 与联合分支概率加权一致，并将条件期望最大值和实际最大值分列。此处的程序检查只证明设定下的计算一致，不证明参数真实或策略可实施。

SVSAC | sv-sac.org

后续若替换为真实项目参数，应保存替换理由、来源日期、价格基准和范围，并保留此前版本，使读者能区分“结论改变来自新数据”与“来自模型结构变化”。若部分参数受保密限制，可以公布区间与结果敏感性，但不能声称外部读者已能完整复算未公开参数的结果。

该说明性模型的结论因此保持有限：分期价值取决于信息质量和追加建设条件，平均优势与尾部保障可以指向不同方案，信息审核的价值需要按错误类型和决策时点评估。它既不证明应全面提前扩建，也不证明等待永远更好。其作用是把前文的机制主张转为可以被参数和数据推翻的条件命题，为实际区域研究提供更严格的提问方式。

22 能耗核算、设备效率与系统预测为何会得到不同总量

22.1 三类估计回答的并非同一个问题

数据中心电力估计至少有三种起点。自下而上的存量核算逐类估计服务器、存储、网络设备数量和功率，再加制冷、供电等设施开销。它回答既有设备合计消耗多少，但不必然等于电表实测总量，因为设备出货、安装、寿命、利用率和 PUE 仍须用数据或假设补齐。第二种从服务量或工作负载出发，以每单位计算、训练任务或请求能耗乘以服务规模，再叠加设施损耗。它揭示算法、硬件与工作负载强度，却容易把单项基准误作全设施平均。第三种自上而下或系统动力学预测，围绕数据量、用户行为、计算服务增长、技术效率建立反馈，回答给定增长机制时未来总量如何演化。它能显式研究反弹和效率饱和，但依赖增长因果链和参数设定。

三类方法在历史总量上也未必可比。自下而上模型可能按设备安装地归并用电，服务量模型可能按处理任务计能耗，系统动力学研究则可能以全球部门总量为目标；有些研究只计 IT 设备，不含制冷和电力转换。比较前应统一年份、地域、设施边界、终端电量或一次能源、是否包括网络与边缘设备，以及“计算实例”等工作量指标的定义。边界不一致时，估计值相近也不代表相互验证；差异很大也不必然证明某一项错误。

Masanet 等人 2020 年发表于《Science》的政策论坛把服务器存量、工作负荷与虚拟化数据、存储和网络设备能耗趋势、超大规模数据中心位置及运营者报告的 PUE 纳入全球自下而上估算。作者估计全球数据中心用电从 2010 年约 194 TWh 升至 2018 年 205 TWh（约占全球用电 1%），同期计算实例增加 550%。论文把服务器能效与虚拟化、存储密度与设备效率、网络端口效率、小型传统机房

SVSAC | sv-sac.org

向大型云设施迁移联系起来，说明服务增长未按同一比例转化为电量增长。[7] 这是多组行业和实验室数据拼接的历史估算，并非逐站电表普查；关于下一轮计算实例翻倍仍可能只增加少量能耗的说法，条件是近期效率趋势延续，不能读作无条件概率预测。

Koot 与 Wijnhoven 在《Applied Energy》的研究用系统动力学模型连接数据中心使用量增长与技术效率，设置基准、摩尔定律终止、工业物联网扩张及其组合情景。论文摘要报告，全球用电从 2016 年约 286 TWh 在基准机制下到 2030 年约 321 TWh；摩尔定律结束与工业物联网组合情景约 752 TWh，摘要所称 95%置信区间约 343—1,031 TWh。本篇依据该文摘要讨论其变量、情景与结果，未对正式版全文作独立复核，故不引述其区间的统计构造。模型排除移动设备、边缘和雾计算。[51] 这些不是 2030 年观测，也不是与 Masanet 同口径的预测竞赛：前者用既有文献参数构造使用增长及效率反馈，后者回溯设备存量和设备能耗。该研究提供可复用模型与输入数据，但区间宽度显示用户行为及效率假设足以改变趋势。本篇因此仅用摘要支持变量、情景、结果和明示范围，不推断校准或验证细节。

LBNL《美国数据中心用能报告：2025 更新》把范围限定在美国，以行业规划的 IT 设备出货、单设备年用电模型、冷却性能模拟、设施类型及位置构成底层输入。参考情景给出 2030 年 649 TWh；敏感性分析分别调整 IT 设备安装量、专用 GPU 出货、AI 芯片寿命、AI 服务器闲置功率及利用率。复合不确定性情景范围为 521—843 TWh；报告摘要称参考量级约占美国 2030 年总用电 11.8%。[60] 这不是概率置信区间：摘要把不同数据输入和行业意见提示的参数偏离组合为情景边界，而非对误差分布抽样后的覆盖概率。全国设备出货路径比全球系统动力学模型更贴近当前产业设备部署，但不能辨认每台设备最终地点、实际利用率或某并网点送电时间。报告也不是峰值预测；年度 TWh 不能直接折算成区域峰值 MW。

若核心问题是“某年部门用了多少电”，历史存量模型及设施电表、供应商披露的交叉核对更接近观测量；若问题是“算力服务增长如何与能耗脱钩或重新耦合”，强度和工作负载指标更有解释力；若问题是“给定 AI 芯片部署、服务增长和效率变化，未来总量如何”，预测模型应并列展示结构与情景。不能只挑一个符合当前叙事的数字，忽视其他研究回答的是不同问题。

SVSAC | sv-sac.org

22.2 工作负载效率的测量单位决定外推上限

“每次训练更省电”不等于“整个数据中心更省电”。单次训练可能降低单位任务能耗，却因模型规模、重试、并行实验或部署后推理请求增加而提高总量；GPU 单卡性能功耗比提高，也未必降低电表读数，因为 CPU、内存、网络、存储、供电损耗、制冷和闲置容量仍处设施边界内。要从工作负载结果外推部门能耗，至少需要知道任务构成、任务数量及增长率、利用率、设施开销、硬件更新和失败或重复计算。

Henderson 等人 2020 年在《Journal of Machine Learning Research》提出 experiment-impact-tracker，以软件接口收集 CPU、GPU 和内存实时能耗，再以 PUE 估算接口未覆盖的制冷、供电等设施开销；作者指出 PUE 可自定义，默认值只是平均修正因子而非设施实时实测。[54] 其方法要求随时段和地点记录能耗与碳强度。作者随机抽样 2019 年 NeurIPS 论文 100 篇，发现 1 篇以某种方式测量能耗、45 篇报告运行时长、46 篇报告硬件、17 篇报告某种计算复杂度指标、没有一篇报告碳排放；图像分类案例中，浮点运算量这一常用复杂度指标与实测能耗经常不相关。[54] 这不是全部机器学习文献或生产环境普查，而是会议年份、抽样框和工具范围有限的证据，足以说明 FLOPs、参数量、GPU 小时与电量不可互换。

设备功率外推还可能系统性偏误。用 GPU 热设计功率乘运行时长，默认设备全程在额定上限运行；再乘未经验证的利用率系数则可能漏计 CPU、内存、通信和待机。反过来，只读单卡功率可能排除共享服务器的其他用电。Henderson 等人的框架改善了单项研究的能量记录，但它是项目级工具，不是全园区或全云厂商总用电核算。[54] 因此它适合校正每项算法或作业的能耗强度，不能替代[2]这类服务器库存和设施模型。

还须区分设施能效和服务效率。PUE 以 IT 设备能耗为分母，说明设施开销相对 IT 负荷的比例；它不回答每个有效训练步、完成的推理请求或有用结果需要多少总电量。服务器效率提高而请求增加，年度 TWh 仍可能增长。实例数上升但每个实例承载更多任务时，实例也不是恒定服务单位。Masanet 等人使用计算实例、存储容量和 IP 流量等服务代理解释历史效率机制，说明单一代理仅覆盖特定设备类别。[7]

接到系统层面时，应分别记录设备代际功率曲线、任务类别及完成量、时延和质量约束、利用率时间分布、设施电表与 IT 分表、冷却及转换损耗、所在节点逐时电源条件。随后以明确服务量分母分解总电量变化：任务数量、单位服务能耗、设施开销比例和设备闲置率分别如何变化。该分解是研究建议，不是现有来

SVSAC | sv-sac.org

源已经提供的统一数据库；公开数据往往不能把这些层级逐项连接，不能据此杜撰“AI 任务单位能耗”的行业平均值。

22.3 来源追溯性与数字精度并非一回事

Mytton 与 Ashtine 的综述保留为追踪能耗估计来源质量的文献线索。[53] 原始摘要和正文均未获可定位核验，故不采用其样本数量、来源构成或比例作经验论据。下段的来源透明度与数值精度比较属于本篇方法分析，不据此否定全部部门估计。

估计即使精确到个位 TWh，若基础是行业调查、供应商预测、出货数据、PUE 代理和安装率假设的组合，也不因此成为同等精度的观测值。反过来，宽区间可能反映透明承认数据缺口，而非研究质量低。审查时应沿来源链检查：设备数据是否公开、变量定义是否可重复、样本覆盖哪些设施、缺失值怎样处理、预测时点是否匹配项目投运年、情景端点是统计区间还是专家设定。前文考察项目后来是否建成；此处考察估计本身如何产生。

研究	观察或建模单位	反事实/情景	能支持什么	不能据此推出什么
Masanet 等[7]	全球服务器、存储、网络设备及设施历史估算	若效率趋势延续，服务增长可被较低能耗抵消	2010—2018 年服务与电量增长不同步及机制	AI 时代美国峰值或确定未来走势
Koot 与 Wijnhoven[51]	全球部门使用量、技术效率和行为因子的系统动力学	基准、摩尔定律变化、工业物联网及组合	结构参数如何改变情景及敏感性	当前美国设备部署实时预测
LBNL 2025 报告[60]	美国设备出货、单设备能耗、冷却模拟、设施类别	参考、安装/芯片寿命/利用率敏感性	全国年度用电情景及关键驱动	概率置信区间、节点容量或峰值 MW
Henderson 等[54]	机器学习实验计量接口和算法任务	不同设备、能耗及碳核算方式	任务能耗测量及复杂度代理局限	云园区全部 IT 和冷却用电

表中“不能推出”不是否定研究价值，而是把模型输出限制在设计用途内。

SVSAC | sv-sac.org

23 选址、灵活性与电力系统规划的不同证据强度

23.1 联合输电规划的边际贡献来自内生选址

Kim 等人 2026 年 8 月发表于《International Journal of Electrical Power and Energy Systems》的同行评议论文，将数据中心企业选址与输电扩建、市场出清共同建模。作者摘要描述的 Stackelberg 结构由输电公司作为领导者最小化社会成本，并预期多家采取价格响应策略的数据中心企业的纳什反应；模型包含工作负载路由、服务器启停、模块化容量和通信连接。搜索候选由 KKT 与 McCormick 单层混合整数线性规划生成，再经整数最优反应与市场重新出清事后认证。案例为 IEEE 24 节点和 118 节点系统，比较电网驱动选址、跟随预先价格选址及集中规划基准。[57]

摘要称协调设计在所有列示案例中相较电网驱动和价格跟随基准降低社会成本和企业私人成本，社会成本距离最佳已找到的集中规划结果不超过 0.9%。[57] 这为一个机制提供模型内证据：选址若被当作外生输入，可能错过企业因电价和线路拥塞改变地点的反应；只按新增数据中心前的价格布局，也可能诱导集中拥塞。作者机构页提供同行评审标记、摘要和方法概述；本篇未对期刊正式版全文作独立复核，因此不将摘要外的参数、线路情景或敏感性作为已读证据。24 和 118 节点系统是基准测试网络，不是美国某 RTO 实际节点、规则、企业合同或建设周期；协调设计改善只能说明给定网络和策略假设下可能有价值，不能直接估算真实电网的节省比例。

这项研究不是说明某条线路或变电站现有多少容量，而是分析企业选址与输电投资如何共同变化；也不是需求预测，因为模型把选址作为策略行动，负荷容量与通信路由受设定约束。现实研究仍须判定候选项目是否真实、负荷何时出现，以及运营者能采取哪些投资行动。

23.2 容量扩展模型说明灵活性减排并非单调关系

Knittel、Senga 与 Wang 的 2025 年 MIT CEEPR/NBER 工作论文使用 GenX 容量扩展模型，比较德州、大西洋中部和 WECC 三个区域。情景改变可转移工作负荷比例（1%—100%）及时间窗口（1—24 小时），考察容量投资和运行组合对成本、排放的影响。MIT 网页摘要报告，灵活性在德州降低系统成本；当模型预计风光供应 54%发电量、可再生能源具成本竞争力时，高灵活度最多降低 40%排放。在大西洋中部和 WECC，排放最高反而上升约 3%，但成本仍下降。作者将机制归因于清洁资源投资经济性、可再生出力与煤电退役及运行方式的组合。反

SVSAC | sv-sac.org

事实把德州可再生能源投资成本和固定运维成本提高至基准的 1.3 倍，模型中的可再生占比降至 21%、煤电留存，灵活负荷的减排优势消失。[58]

这不是“灵活负荷必减排”或“必增排”的实测因果结论，而是长期投资与运行同时重新优化后的条件结果。与按当期平均碳强度移任务的调度模拟相比，容量扩展模型允许灵活性改变新建资源、机组退役和储能投资，因此反事实更宽；代价是结果对区域资源成本、负荷增长和可调比例敏感。该材料是工作论文，截止时尚未同行评议；本篇依据 MIT 官方摘要，不声称真实区域已经发生相同机组组合变化。

23.3 运行记录与现场试验回答局部问题，不能直接证明系统收益

Google 作者团队的《Carbon-Aware Computing for Datacenters》描述其内部碳智能计算系统。系统依据次日碳强度和需求预测生成虚拟容量曲线，按小时限制可延迟工作负荷资源，同时保留当日完成工作量的能力。作者以生产运行数据展示，VCC 在碳密集时段限制小时容量，并把可延期任务移至较低碳时段。[55] 这类证据强于纯优化模型对“软件控制能够运行”的主张，因为系统已部署在 Google 生产集群；但边界很窄：作者说明搜索、地图、YouTube 等用户服务及客户 VM 工作负荷不受控制；内部批处理任务须在一天内完成。本篇没有取得可独立复算的逐站原始功率和反事实调度数据，不能据运行部署说明推定每次移时都减少边际排放或缓解节点拥塞。平均碳强度、预测误差和异地可达性会影响结果。

Sukprasert 等人在 2024 年 EuroSys 论文中使用痕迹驱动的云工作负荷分析，结合 2020—2022 年 123 个地区的逐时平均碳强度与 Azure、Google 公开工作负荷痕迹，比较批处理和交互式负荷在不同截止期限、运行时长、服务等级及迁移限制下的时空转移上限。作者发现转移可降低任务碳排放，但现实可达上限通常低于理想上限；电力供应更低碳时，碳感知调度相对不感知碳强度的收益收窄。[56] 这是有真实工作负荷痕迹支撑的系统模拟，不是数据中心现场接受电网调度试验。它以平均碳强度为输入，不模拟新增任务如何改变边际机组、潮流或容量投资；适合说明工作负荷及 SLO 限制可移动范围，不能声称电网实际避免了同等排放。

本篇采用 Phoenix 展示的 arXiv v1 预印本，不把该版本标为已核同行评议出版物。研究描述 Oracle 云园区一个 256 张 NVIDIA A100 GPU 集群，在 2025 年 5 月 1 日及 3 日分别配合 APS 与 SRP 系统高峰进行两次测试；每次目标是相对平均基线降功率 25%、维持 3 小时、用 15 分钟爬坡并避免事件后反弹。团队称两次均达到电力公司设定目标，使用功率封顶、作业暂停等控制，并配置不同训练、微

SVSAC | sv-sac.org

调和推理任务。[59] 这是现场工程展示，证明特定硬件、两种作业组合及预设时长内可完成目标；不证明广泛机群、季节性事件、经济补偿或多年设备退化，本次所读 arXiv v1 未被本次分析核验为同行评议正式版，完整可复算的原始事件数据亦未取得。试验成功不能替代普适可履约容量估计。

三类证据须拆开陈述：系统模型比较规划反事实和成本条件；生产运行数据证明实际组织内的调度流程曾运行；现场展示证明有限设备和事件完成指定动作；时间序列观察可揭示实际运行与事件关联，但通常需要额外设计才能归因。任何一种都不能自动替代其他证据。用小型展示校准区域容量信用，需验证代表性、持续时长、任务占比和同时性；用容量扩展模型估算当年逐时减排，则要承认长期重优化的反事实不同于实时运行。

24 决策建议相反的可检验机制与外部效度边界

24.1 相反决策建议可能来自因变量和分母变化

研究单位若是每个计算实例或已完成任务，效率改善可能明显；另一研究以年度设施 TWh 为因变量，只要服务器数量、利用率或冷却负荷增长更快，绝对用电仍可能上升。这不矛盾。可检验的竞争机制是单位服务能耗下降抵消服务扩张，或任务量增长超过效率改善。前者预测同代硬件与标准作业的 kWh/任务持续下降、总任务量增加；后者预测设施电表增长且设备库存、推理请求或训练周期的增量可以解释其中部分变化。识别需同时记录服务量、设备代际、负荷率和设施总电量，而不能只引用芯片峰值性能或总用电一项。

第二种竞争机制是效率上限与效率迁移。Masanet 等人的历史估算中，虚拟化和向超大规模设施转移可降低单位服务服务器及设施用电；作者亦提醒 PUE 接近低值后，靠设施开销继续下降抵消 IT 用电增长的空间有限。[7] LBNL 2025 敏感性情景显示，AI 芯片寿命、安装量、闲置功率和利用率显著改变预测。[60] 可证伪预测是：若主要问题是 IT 效率趋缓，设备级每瓦任务改善应放慢；若问题来自设施结构，PUE 及 IT 负荷占比应随新建/退役类别变化；若问题是利用率低估，实测负荷会高于铭牌功率乘假定负载率的估计。即使观察到效率改善，也须测量任务规模变化才能排除服务扩张造成的总量反弹。

第三种机制是年度电量与可靠容量/峰值功率不同。相同 TWh 可以代表稳定满负荷，也可以代表夜间及低价时段大量可延后任务；对容量、爬坡、备用和节点拥塞的影响不同。检验应以小时或更细粒度记录站点负荷与 IT 任务，分别计算

SVSAC | sv-sac.org

最大负荷、低谷、爬坡、持续时间和季节同步性。年度电量单独进入容量扩展模型不能验证峰值边界。

24.2 “灵活性降低成本”与“灵活性减排”是不同命题

Knittel 等人的区域比较显示，成本下降可与排放上升并存，关键是灵活性如何改变机组出力及新建投资相对经济性。[58] Kim 等人的输电模型以社会成本和企业私人成本为目标，未把实际负荷减排或现场节能作为主要检验。[57] 生产调度论文与工作负荷模拟更关注某一时间或任务的碳强度和能耗，不一定反映长期发电投资变化。[55],[56] 这些研究的目标函数和反事实不同，不能只抽取“有益/有害”的形容词比较。

至少应并列三种碳问题：平均因子下购电归属、逐时边际供电响应下运行排放变化、投资改变后的长期系统排放。平均因子可用于按规则盘点用电相关排放，却不能识别新增一单位负荷由哪台机组增发；长期扩展模型反映新增和退役资源，却不是某小时边际排放测量。若灵活负荷移到低谷但由煤电供电、可再生能源未扩建，负荷更平坦不保证排放下降；若时移落在可再生能源富余、可吸收弃电时段，则可能降低高排放机组出力。每项判断须说明地区、季节、调度边界和新增电量来源。

Google 生产 VCC 使用次日预测、集群级容量约束和日内完成任务条件；Sukprasert 等人模拟多地区批处理与交互负荷，以工作痕迹及平均碳强度衡量移动边界；Phoenix 展示则对特定系统峰值设定功率削减目标，按服务等级控制。[55],[56],[59] 前者关注真实组织如何移时内部批任务，第二类估计更广工作负荷的理想与现实上限，第三类证明短时响应动作。策略可移动工作量、碳核算边界和电网效益指标各异；将削减百分比并为同一承诺能力，会忽略基线分母、任务构成及恢复轨迹。

24.3 区分竞争解释的研究设计

设备预测与现场负荷的联结，需要设施电表、IT/机架功率、任务产出、代际、可用容量/利用率、冷却、地点及投运退役日期，并将申报、购入、通电与稳定运行分栏。第 27–28 章进一步把这套匹配细化为设备队列、运行功率和设施/接入三层验证。可用保密中介、分层汇总或去标识标签保护隐私，同时保留外部聚合复核途径；这是本篇的设计建议，非已完成的数据整合。

其次，以设备代际与服务量分解效率机制。对同类作业匹配模型质量、输入规模、批量大小和时限，测量机架电量而非仅看 GPU 规格；分别用 CPU/GPU/内存功率及机房电表核对设施加成。若每任务 kWh 下降而工作量增长更快，年度用

SVSAC | sv-sac.org

电上升应由两者共同解释，而不是视作效率结果与总用电彼此否定。训练和推理也应分别评估，因为数量、时延约束和负荷曲线不同。

第三，现场灵活性检验应预先定义无响应基线、外生调度信号、作业质量指标及反弹窗口，并同时观察接入点和设施内部负荷。可用同一设施未参与响应的相邻集群或非事件时段作辅助对照，但须检查集群间是否共享冷却、UPS 和排程器，以免对照组被干预污染。此建议与第 15 章履约验证框架不同：本节关注识别，即削减是否由响应命令造成而非正常负荷变化，以及边际排放是否改变。

第四，联合规划论文应在不同网络拓扑与制度区域复核。IEEE 测试系统适于验证算法能否在已知约束下运行，不能代表真实输电瓶颈分布、规划周期、土地许可、融资成本、线路候选集或企业重复申报。外部检验可把负荷、线路、资源成本、候选接入点、历史选址和实际项目状态替换为可审计的运营者数据，并保留未建成项目。若结论仅在特定节点、企业行为或成本假设下成立，应报告扰动到何处排序翻转；模型收敛不是区域实证证据。

24.4 研究决策建议的外部效度边界

全球历史存量核算覆盖多种数据中心和设备，却可能低估当前某地集中式 GPU 集群爬坡；美国设备出货情景贴近近期产业路径但依赖报告年份及行业输入；全球系统动力学揭示技术和行为反馈，却不能替代北美特定节点的工程可交付性。局部负荷记录空间和技术细节高，却不代表全美；覆盖广也不等于站点实测。

工作负荷外部效度受服务等级、云平台管理和业务组织影响。Google 允许日内完成的内部批处理不等于多租户客户可无成本延期；Phoenix A100 集群、预选训练/推理组合及两次短事件不必然代表新一代加速器、推理主导园区或持续热浪时期的灵活余量；Azure/Google 痕迹和碳强度输入也不能推出某设施在 2026 年可向市场提供多少 MW。[55],[56],[59] 这些限制不抹去证据价值，而要求外推说明样本、控制方式、持续时间和服务质量边界。

区域系统差异也不能平均化。灵活性价值随拓扑、燃料价格、可再生建设成本、可调比例、储能约束和市场结算设计变化。Kim 等人的两个标准节点网络主要检验联合选址规划逻辑；Knittel 等人三个区域的长期容量扩展更接近区域资源组合，但仍是模型而非运营者投资记录。[57],[58] 结合测试系统、区域模型和实际运行数据时，应明确每类证据对应哪一项主张，不让模型输出替代所有事实。

SVSAC | sv-sac.org

这些差异要求后续比较研究逐一保留估计对象、模型边界、对照方案和样本来源。本篇所列研究覆盖若干类型，但不是对全部 AI 数据中心研究的穷尽性元分析；本篇对未读取全文的研究不推断其未披露的方法细节。

25 研究设计与可复核性

25.1 证据取得与纳入原则

本篇以 2026 年 10 月 23 日为资料截止日，围绕三个问题选择材料：需求增长在何种口径上得到支持，负荷如何经由项目条件进入系统，以及不确定性如何改变工程与费用决策。检索优先使用 EIA、DOE、FERC、NERC、州监管机构、区域运营机构和项目责任主体的公开文件，并以同行评议研究比较方法与外部效度。企业公告仅用于确认其表达的承诺或预期，不作为独立验证其全部商业和工程结果的依据。

材料按问题相关性和可定位性纳入，不以支持某一增长方向为条件。本篇没有执行覆盖全部数据库、预注册检索式与双人筛选的系统综述程序，因此不宣称文献穷尽，也不将引用数量当作代表性证明。对同一报告，记录实际读取范围；仅核验摘要时，不扩写摘要未明示的方法、样本或结论。对无法获取的文件保留缺口，而不以检索片段、二手转述或推测补齐正文事实。

证据层级按所支持的主张区分，而不是简单认为某机构全部材料均同样权威。机构发布的历史表值可以支持统计记录，预测说明可以支持预测的内容和假设，最终命令可以支持其裁定内容，技术指南可以支持建议实践，研究论文可以支持特定方法和样本下的结果。它们均不能越过自身边界。例如官方预测不因发布者身份就变成未来事实，同行评议模拟也不因审稿就变成当前设施的现场测试。

25.2 版本、口径与引用映射

来源记录保留作者/机构、题名、发布及访问日期、链接、定位、可支持主张和限制。段落中的来源事实与作者解释分开，引用映射及来源台账共同界定支持范围；一个编号不为全段推断背书。

对可下载的固定文件，保存原文件和哈希；对动态网页，尽可能保留访问时的版本。哈希用于识别本地文件是否变化，不能证明内容正确或官方身份。本篇仅引用公开可读取的定位，不声称已归档未获得的副本，也不以检索片段替代正文事实。对只取得检索记录而未读到正文的材料，不引用其中任何实质结论。

SVSAC | sv-sac.org

版本差异单独记录，不把修订隐去。预测终点、历史基期、发布日期与程序状态分别保存。ERCOT 材料中的初步值、调整建议、流程批准及新版发布时间说明构成状态链，而不是四个可任意择用的“最新数字”。[15],[16],[24],[25] 对项目预计投运时间也采用同一办法。若后来取得实际运行记录，追加实际值而不覆盖此前预期，以便研究误差和时序变化。

25.3 计算与比较的复核边界

两组说明计算分别检验功率边界/设施比率/恢复负荷/项目相关性，以及不完全信号下的两阶段建设；均保存输入和代码。复算只核单位、期望、联合分支与阈值，作者设定不变成经验参数。其计算一致性也不同于第 27–28 章的报告复核：公开存量与功率公式仍不能重建未取得的专有出货输入。

区域比较主要采用文件与机制对照，不计算未经口径对齐的跨区增长排名。因果部分提出可证伪的中间环节和竞争解释，没有声称已完成项目级准实验。环境部分区分年度平均足迹、逐时代理估算和边际后果。工程部分区分一般模型、设计状态、竣工状态和现场事件校验。采用这些分层，不是降低研究要求，而是使结论强度可被读者检验。

跨材料核对聚焦四项一致性：统计变量是否同口径，制度版本是否同日期，模型条件是否在解释中保持，以及区域案例是否被错误外推。数字复算、文本核查与外部有效性是不同环节；即使前两项一致，也不能验证尚未建立的全国样本或真实项目效果。

26 有条件的决策含义与后续检验

26.1 对规划机构：把触发条件写入建设决定

本篇的分析支持一个程序性建议：当新增需求仍有明显状态不确定性时，建设决定应注明依据的负荷版本、项目集合和触发条件，而不只注明一个容量总数。触发条件可以是某阶段的可验证承诺、工程完成或运行记录，但其具体形式应由适用制度确定。建议的核心是可追溯，不是要求所有地区采用同一合同或统一概率阈值。

若后续项目退出或延期，应重新检查哪些工程仍由其他负荷、可靠性或资产更新需要支撑。可以继续建设并不自动说明此前决策有误；取消也不自动证明项目原本虚假。决定应解释范围、时间或收益结构怎样改变，使读者能区分合理调整与单纯事后叙事。对于已经形成不可逆费用的资产，还需要披露沉没成本与未来可避免成本，避免为证明过去正确而继续支出。

SVSAC | sv-sac.org

在信息仍不足时，规划者可以并列确定承诺路径与条件路径，并在安全边界内比较分期、扩建和履约响应。不能用一个不透明折减系数替代网络研究，也不能把“谨慎”理解为一律不建。真正需要比较的是不同误差的后果：建设不足造成何种服务和可靠性损失，建设过量留下何种闲置与费用，以及下一次调整需要多长时间。

26.2 对数据中心：把服务需求与供能证据对接

数据中心运营商若希望加快接入，研究材料应说明逐期接口需求、IT 与设施边界、计算任务限制、现场资源及其运行条件。提供这些资料不保证获得某个接入日期，但可以减少把机架、园区规划或合同电源容量混为一谈的问题。尤其应说明哪些容量属于确定首期，哪些只是远期选择，以及多个候选地点是否互斥。

对宣称的灵活性，应提供具体服务产品、基线、响应和恢复证据。一次成功演示不能代表全年能力，装有电池也不代表所有电池能量均可用于电网响应。若商业服务等级、冷却安全或备电义务使某些时段不可用，应将这些边界带入规划，不能先以理论最大值降低工程要求，再在实际调用时提出未披露例外。

购电与环境承诺则需要另一套对接。长期合同可以有商业与建设作用，但合同能源匹配、小时物理交付和边际排放不是同一结果。研究建议分别披露承诺边界与证据，不因某种采购工具具有局限就否认其全部价值，也不因其有价值就免除逐时、地点与新增性问题。涉及水资源时，设施现场节水和电源侧间接用水应分别说明。

26.3 对成本与市场研究：把机会与已实现收入分开

设备机会需要沿工程设计、许可、预算、资格、采购和交付定位。全国负荷情景既不提供某厂商订单，也不识别利润或估值；本篇未作企业收入预测。后续市场研究所缺的是这条交易链的匹配证据，而非更大的宏观需求数。

费用风险同样应拆开。谁为专用资产付款、共享资产如何分摊、市场能源成本如何变化、信用担保能覆盖多少风险，需要分别核实。某项规则目标是防止费用转嫁，不等于其效果已被证明；某类用户账单上升，也不自动证明另一类用户获得补贴。一个专业的市场判断应说明从制度到行为再到收入的链条，以及哪一环目前只有假设。

26.4 可用来推翻本篇机制判断的证据

本篇将自身判断保持为可检验命题，而不是只能被确认的叙述。若经过天气、边界和用途分解后，新增 AI 活动并未提高净设施用电，则应缩小“活动增长

SVSAC | sv-sac.org

转化为能源增量”的范围。若真实项目计量显示承诺审核并未改善路径误差，就需要重新评价筛选机制。若某类灵活服务在不利时段仍稳定履约，并通过独立计量核验，其规划信用可以高于目前仅作敏感性输入的状态。

反过来，若新增电源无法按时交付、现场资源受到运行限制、恢复反弹产生新的瓶颈，或者低现场耗水技术增加了流域外发电侧压力，则应调整相应方案的效果判断。不同反证不应被合并为“AI 需求不存在”，也不应被忽略为增长过程中的小问题。它们分别针对能量、时间、地点、服务质量和环境边界，需要分别修正。

后续实证优先级应由决策敏感性确定。若某个投资排序主要取决于追加建设时长，最有价值的证据可能是工程交付而非再增加一份全国需求预测；若结果主要取决于高峰时段履约，优先级应是现场响应记录；若依赖信用回收，重点则是主体与担保条件。这样安排研究，可以将资源用于最可能改变结论的缺口，而不是追求来源或模型数量。

第 27 章 从出货到年均功率：LBNL 更新报告的存量核算与测量边界

27.1 一条可追踪的计算链，不是一份服务器普查

LBNL 2026 年 6 月发布的《美国数据中心用能报告：2025 更新》继续使用自下而上模型，但其底层链条需要逐级理解：商业数据库记录服务器和加速器的历史/预计出货；设备分类与机架配置把芯片数量换成服务器数量；使用假设的设备寿命把出货转换成各年“安装存量”；额定功率、运行功率比例、空闲功率和利用率计算设备年平均功率；再以 8760 小时换算 IT 年电量，并乘以不同设施类别的 PUE 估计数据中心总电量。[60]

这条链上不同变量的观测性质并不相同。报告将 2024 定为最后一个历史估计年，因为主要 IT 出货数据仅覆盖 2025 年前三个季度，2025 结果已有部分预测；2026—2030 则进一步依赖出货预测和设备配置假设。[60]例如 IDC 服务器数据库含历史与预计出货，2030 服务器出货通过 2024—2029 子类别增长率线性外推；AI 加速器数据来自 IDC 和 Omdia 的商业研究，部分全球 ASIC 单位估计还被直接用于美国情景。LBNL 明确指出，全球 ASIC 不一定全部安装在美国，当前没有数据充分校正非美国部署，参考情景因此采用全部计入美国的假设。[60]这不是一个已实测的美国 ASIC 安装数，而是为全国模型选用的数据口径。

SVSAC | sv-sac.org

报告的可复核性因此分层：方法图和方程公开，读者能理解如何从参数走向TWh；许多关键商业输入却是付费市场追踪资料，LBNL 获得购买或免费访问，而本次分析没有取得该原始数据库。IDC 产品/出货预测还会随版本更新，报告自身比较了 2024 末、2025 年中、2025 年末多个版本，显示 2030 加速器单位预测有过大幅修订。[60]公开方程可以复算代入值，不能重建未公开的每个型号出货记录，也不能独立验证行业专家提供的运行反馈。故“可复算公式”不等于“完全可复现预测”。

27.2 固定寿命的队列模型把退役变成一个参数

报告把各年出货映射为服务器安装存量，公开方法写成：当年存量=上年存量+当年出货−当年退役；退役按固定平均使用寿命 L 建模，分别适用于传统服务器、AI 加速服务器和必要时单独处理的加速器。[60]这一公式透明地显示预测从出货数进入运行设备存量的关键中间步骤，但固定寿命是模型假设而非逐台观察到的实物退役日期。

报告的参考情景把 AI 加速服务器寿命设为 5 年，并另设 4 年较短寿命敏感性；传统服务器寿命则按不同年度假设从约 4.4 年增加至约 5.7 年后维持。[60]报告讨论了经济折旧和物理可用寿命并非一回事：GPU 可能因能效/性能代际变化，在主集群中提前换代，但被转入次级集群继续工作；反过来，高早期故障或维护困难也可能缩短运行周期。公开来源未提供按型号、工作负荷、站点的退役/转用面板，故平均寿命只能作为存量模型参数，不能读成所有 AI 服务器五年后统一报废，也不代表折旧表上的年限。

这个区分对电力预测有实质影响。若设备经济上已不适合承担最新训练任务，却转为低优先级推理或内部计算，它可能从“主力加速集群”消失，却仍占数据中心电表负荷；若报告寿命定义将其视为退役，而实际转用仍耗电，安装存量和年度电量可能被低估。若设备被卖往境外、拆解、关机或被完整替换，影响则不同。要检验寿命参数，应同时追踪主用途退出、园区内转用、关机、售出和拆解，而不是只以财务折旧或新设备采购发布日期代替实际电力状态。

27.3 “服务器”计数取决于架构切分与芯片功能单位

报告将 AI 服务器配置为 0、2、4 或 8 个 GPU，并分别建立 ASIC 服务器的平均配置；报告假设 2026 年前每台平均 6 个 ASIC，2026 年起降至 5 个。对机架级系统，模型把节点而非整个机架计作服务器：例如 NVL72 含 72 个 GPU、18 个节点，因每节点 4 个 GPU，所以按 18 台四 GPU 服务器计数。[60]这在估算设备数量

SVSAC | sv-sac.org

时避免把整机架误当单服务器，但“server count”并不等于独立插电端口数、机架数或电力变压器馈线数；共享电源、冷却和网络等仍需另行进入设施模型。

芯片出货数转换为服务器功率也需要产品世代的功能单位。LBNL 称，高端未来 GPU 预测按 NVIDIA B300 等效的 die 数量和大小建模，原因是市场受 die 供给约束，预计总 die 级产量可保持接近；报告据此认为 die 尺寸与功耗关系可用于估计未来功率。[60]此方法把型号快速更替的市场预测转成可比单元，但它不是宣称未来所有产品都是 B300，也不能保证先进封装、内存带宽、互连和服务器架构的功率按单一 die 单位同比变化。报告也把 ASIC 纳入；GPU 与 ASIC 的功能、典型服务器配置、控制软件和运行方式不完全相同，故二者合计芯片数不应被当作同质算力或同质负载。

报告还单独估计只配高内存和网络的双 CPU AI 优化服务器，并将其与传统双 CPU 服务器区分；CPU-only 服务器在模型中主要承担可用 CPU 完成的推理应用，而不是大多数模型训练。[60]这说明 AI 设备的边界不仅是“有没有 GPU”：内存、互连和用途都会改变单位功率。若地方接入申请只报 GPU 数量，无法由此直接推得园区总峰值；若全国模型数到“AI server”，也必须知道机架节点如何定义、服务器类别有否重叠、计算设备之外的设施用电如何计入。

27.4 额定功率、空闲功率和利用率不是同一开关

LBNL 从制造商规格书或 Omdia 估计构造服务器额定功率，再应用运行功率比例、空闲功率比例和计算利用率，得到年平均功率。对 AI 服务器，报告的描述公式为：运行功率=额定功率×运行功率比例；空闲功率=额定功率×空闲功率比例；年平均功率=运行功率×利用率+空闲功率×(1−利用率)。[60]这是一项设备能耗模型，不是实际电表曲线，也不是按“额定功率×利用率”简化计算。

利用率在报告中指服务器一年内的平均计算强度。训练任务参考情景采用 80%，推理服务器逐步采用类似传统企业数据中心的较低利用率，2030 年为 20%；报告称推理服务须为用户同时查询的峰值预留能力，所以非高峰时段可较空闲。[60]报告将训练/推理的功率份额从 2024 年约 35%训练逐步降至 2030 年约 20%训练，并按功率份额而非服务器台数加权。这个设置让一个反直觉现象清楚可见：平均利用率低不必表示并网容量无需准备，因为服务需满足高峰响应、冗余和可用性；同时，服务器计算利用率也不是“并网申请容量的利用率”。

空闲功率也影响总量。参考模型约用 20%额定功率作为 AI 设备空闲比例；报告同时转述行业专家的高空闲功率反馈，称敏感性分析采用 30%，且有专家提出延迟敏感型推理设备可能高达 50%，但另一位半导体硬件架构师支持原 20%比

SVSAC | sv-sac.org

例。LBNL 因此把推理高能耗列为参数情景，而没有声称 50%已由广泛实测确认。[60]使用同一芯片额定功率，空闲比和时间利用率变化会改变年电量；若把某一训练基准功率直接外推到所有推理集群，可能同时错误处理运行、空闲和业务峰值三个方面。

这些输入来自规格书、付费数据库、研究测量和专家意见，证据类型不同。额定参数可依据特定设备规格核查；实际运转/空闲比需要安装后遥测；工作负载结构和利用率需要运营者数据；PUE 来自设施系统数据或模拟。报告选择的输入是对全国模型有用的估计，不是逐站全年 15 分钟功率实测。没有这些层级的现场日志，读者不能把“近 19 百万 AI 服务器”当作 19 百万台均满负荷运行的事实，也不能反过来因平均计算利用率较低就断定它们不造成峰值供电需求。

27.5 冷却 PUE 和设施能耗也有边界

报告将服务器、独立存储和网络设备纳入 IT 电量，再乘数据中心类别的 PUE 计入冷却与配电损耗。它引用了 AI 设备服务设施的全国平均 PUE 模型值，但也另给整体数据中心设施组合的 PUE 路径；两者不是同一设施样本或相同权重的指标。[60]2024 年 AI 设备服务设施平均 PUE 估计约 1.145，2030 年约 1.136；全行业模型汇总 PUE 则因设备从传统设施转入较低 PUE、配液冷的 AI 园区而另行变化。[60]

PUE 是建筑总电量/IT 设备电量的比例，不是单个 GPU 能效、设施峰值系数或节点并网利用率。即便液冷降低某类设施的相对冷却损耗，若 IT 总量和单位额定功率大幅增加，设施电量仍可增长。报告对多个情景保留了存储、网络、基础设施和 PUE 输入，不随 AI 服务器敏感性同步变化；所以情景间的总量差异主要来自 AI 服务器能耗。[60]这既让参数贡献可分解，也意味着模型没有把高密度液冷扩张对其它系统参数的所有联动都作为情景。读者不应把一个全国平均 PUE 因子当成特定水冷方案、季节和区域的工程性能。

第 28 章 从年电量到接入能力：情景模型可以回答什么

28.1 将 TWh 换算为 GW，只得到平均值

年度电量可除以 8,760 小时，得到全年平均功率；它不直接给出最大需量、申请的并网容量或可靠容量。以 LBNL 报告自身的参考情景为纯算术示例，649 TWh/年÷8,760 小时约为 74.09 GW 平均设施负荷。报告另采用约 50%的“并网容量利用率”假设，将年电量换算为约 148 GW 的并网容量需求。[60] 该数是模型口径转换，依赖其分母定义和利用率设定；它不是实测的美国数据中心容量，也不是某

SVSAC | sv-sac.org

个区域在某小时的峰值，更不等于需要新增 148 GW 发电机组。把两者直接等同，会把能源、网络接入和资源充足性三类问题压成一个数。

最容易混淆的是两种利用率。服务器利用率描述一年中计算工作负载强度，进入设备平均功率公式；并网容量利用率则把年平均设施负荷与申请/核定接入容量联系起来。即使服务器大部分时间空闲，只要服务要求在突发时响应、冗余设备保持带电或多个任务同时出现，接入容量与年平均功率也可能相差很大。反过来，年均负荷较高也不说明本地输电、变压器或发电资源能在同一时点满足峰值。LBNL 用 50%把能量映射到接入容量，提供全国情景的比较刻度，但其正文没有将这一转换校准为逐站峰值分布或电网可靠性标准。[60]

因此，工程规划至少要并列保存四个量：年设施电量（TWh）、逐时或 15 分钟平均功率序列（MW）、接入申请/合同容量（MW），以及经可靠性规则认可的可用资源能力（MW）。它们有联系，却需各自的观测和模型。全国年电量可以提示规模，却无法告诉规划者高峰是否与夏季负荷重叠，园区是否存在备用发电、UPS 或可削减任务，或者其接入申请是否包含冗余余量。把报告的 148 GW 当作峰值会过度推断；把 74 GW 年均值当作所有区域同时承受的负荷则会制造虚假的空间分布。

28.2 高推理能耗情景揭示模型的条件，而非预测矛盾

报告的“高推理能耗”情景提高推理利用率等设备参数，使年度能量上升；但它保持与参考情景相同的设备存量，因而并网容量需求不变，并指出该情景的数据中心容量利用率较高。[60] 若把年度电量和容量视作同一变量，这看似矛盾；按模型结构则是不同参数作用于不同环节：既有设备以更高年平均功率运行，提高 TWh；可用设备总量未变，报告没有据此增加装机/接入容量假设。它说明在既定服务器存量和容量转换规则下，能源可因运行强度而改变，容量未必同比例改变。

但“存量不变即容量不变”属于情景定义，不是普遍工程定律。若高推理活动使更多服务器在同一小时并发运行，改变了原有峰均比，或者需要提高备用比例、冷却泵/风机同时负荷，现场最大需量可能增加；若调度、队列管理、分批运行或电池放电削平尖峰，则高年电量未必带来同幅峰值增长。报告未提供逐站并发和 15 分钟负荷曲线，故不能据其全国年用电情景验证这些替代机制。合适结论是：年能量对利用率假设敏感，而容量结果还依赖服务器峰值并发、冗余、运营控制和并网申请口径。应分别报告这些量，不能把一个情景中的不变值解释成已测得的容量弹性为零。

SVSAC | sv-sac.org

同样，参考情景与复合低/高情景的范围不是统计置信区间。报告把 GPU/ASIC 出货乘数、加速器寿命、AI 服务器闲置功率及训练和推理利用率等假设取不同组合形成敏感性范围，并明确复合极端同时发生的可能性不高。[60] 这适合作压力测试：告诉规划者多个参数若同向偏离会造成怎样的模型范围。它不提供每个参数误差的概率分布或相关矩阵，不能由“低—高”端点推导发生概率。例如，表 1 中的复合高端采用较高 ASIC 出货乘数、5.5 年加速器寿命（参考情景为 5 年）以及较高闲置功率和训练、推理利用率；这些假设的联合一致性仍需单独论证；反方向的相关关系也可能缩小或改变范围。没有联合分布，就应称其为条件情景，而不是预测误差覆盖区间。

28.3 用分层观测让全国估计变得可证伪

若要把这种模型从全国情景推进到可用于局部规划的证据，验证数据应按计算链分层，而不是仅要求运营商披露一个年度总电量。第一层是设备队列：记录按服务器类别、加速器类型和机架/节点配置划分的到货、安装、首次通电、转用、停机、迁出和退役月份，并说明出货数据库版本及美国安装份额的判定。这样才能检验固定寿命是否近似适用，以及全球 ASIC 出货全计入美国的假设是否造成偏差。序列号和具体型号可由独立审计者保管，公开仅报告脱敏汇总，避免把可证伪性错误地等同于公开商业机密。

第二层是功率与业务：在代表性设备队列上同步采集 IT 电表、设施主表、额定功率、空闲/活动状态、训练/推理任务类别与任务并发度。时间戳至少应能与规划所用峰值窗口对齐，并公布采样频率、缺失值规则和备用设备处理。仅报告“平均利用率”无法区分持续低功率与短时高峰，也无法检验报告对推理峰值预留、低谷空闲功率的拆分。若不能公开原始遥测，可由第三方按预注册的聚合口径复算并披露置信区间、排除条件和校准误差；这将提升验证性，但仍不等于公开可复现全部设备数据。

第三层把 IT 电量与设施、电网接口相连：逐月披露 PUE 分子/分母及冷却方式，在不泄露节点安全信息的前提下公开年度能量、分时最大需量、申请容量、合同容量、获准容量、实际通电容量与备用电源边界。容量“获批”与“已送电”必须分别记账。跨州规划可以公布区域汇总或时间滞后指标，但若没有地点、并发与气象共同数据，就不能检验峰值是否同时发生。公开能源总量也无法识别接入瓶颈究竟位于园区变电站、上级输电、资源充足性还是运营商主动保留裕度。

这套数据协议本身是作者建议，不是 LBNL 报告已执行的采样方案，也不是现实中的强制披露要求。它的价值是给出可失败的检验：若固定寿命模型持续高

SVSAC | sv-sac.org

估仍在线设备或低估转用，队列记录会显现；若高推理参数提高年电量但不提高高分位功率，逐时负荷可以验证；若全国误差主要来自美国安装份额，型号/部署地点分层可以定位。相反，若企业只提供年度总表，需求增长、服务器效率提升、转用存量、设施 PUE 变化和备用策略会混在一起，多个相反机制仍能产生相似 TWh，模型便无法被有力区分。对于输电与发电规划，最有用的补充不是再增加一个全国端点，而是能够连接设备队列、运行功率、地点、接入状态及天气的可审计时间序列。

29 决策建议与研究局限

29.1 对三个核心问题的回答

关于需求增长，在本篇核验范围内，全国报告与 PJM 材料显示显著数据中心增长预期，ERCOT 的 2026 年 7 月记录提供已发生小时峰值证据，其规划文件则提供大型负荷相关工程压力的证据；这些不同性质的材料不支持把所有数据中心电量、商业部门增量或峰值变化都归于 AI。[1],[2],[17],[21],[25]历史估计、条件预测和实际运行应并列解释，不能相互替代。AI 对设备组合与工作负载的影响具有明确研究意义，但其净增量仍须依靠一致用途口径和可复核分解识别。

关于局部接入，年度能源与全国容量不充分决定某个节点的交付。资源可用、网络约束、动态行为和设施控制共同决定可行性；商业购电、通用模型和灵活潜力都不能直接代替现场条件与履约证据。本篇将问题拆分为可对应数据和责任主体的层级，目的不是增加术语，而是避免用同一个“缺电”判断推导互不相同的工程方案。

关于不确定性下的投资，文件比较与说明模型共同表明，需求审核、建设时序和费用义务存在反馈。分期可能改善平均成本，但价值受信号质量、追加条件和服务义务限制；财务回收改善也不等于物理风险消失。合理建议应列出触发条件和可被推翻的参数，而不是给所有区域同一个扩建或等待结论。

29.2 本篇没有识别的结果

本篇没有建立覆盖美国全部候选数据中心和退出项目的普查，没有取得园区级完整逐时计量，也没有估计 AI 造成的全国电网投资比例。没有用说明模型的抽象单位预测真实资本开支，没有对任何具体设备作合规认证判断，也没有计算某项采购工具的普遍净减排效果。这些不是可以用更强措辞填补的空白，而是进一步研究所需的数据与设计。

SVSAC | sv-sac.org

EIA 固定原表、部分费率原命令、动态文件、州级最新规划及原始许可链尚待原始证据确认；因此，对截止日制度状态的结论以本篇实际核验的材料为限。官方机构的建议、白皮书、草案与最终裁定具有不同效力，本篇不把它们统一称为强制规则。同行评议研究也有各自数据年份和模拟边界，不因发表时间接近就自动代表当前 AI 设施。

更细的项目与运行数据可以扩展上述结论，但需要重新核对观察边界和版本；在此之前，本篇的贡献是把需求规模、兑现条件、工程约束和风险分配连接为可检验、可复算、可继续补证的研究框架。第 27–28 章对 LBNL 完整更新报告的分析进一步显示，公开公式仍受专有出货数据及寿命、利用率参数限制；报告采用的约 50%并网容量利用率只是年电量到容量的情景换算，不是设施峰值测量。[60]

SVSAC - sv-sac.org

SVSAC | sv-sac.org

给企业的下一步建议

美国数据中心电力需求值得关注，企业更需要判断哪些需求能转化为可进入的设备与工程环节。因此，我们建议把下一步工作收敛为一项有明确交付物的专项任务。

具体起点是：按区域电网条件、项目阶段与采购角色拆分机会，核对设备适配、交付能力及本地服务要求。在此基础上，管理层可以按证据充分程度、资源投入与业务影响确定优先级，先验证关键假设，再安排后续合作与投入。

要把上述判断用于实际经营，可以围绕具体产品或项目、目标市场、现有能力和决策难点，进一步形成电力需求与项目机会图谱、供应链角色筛选、市场进入研究。这类工作需要把行业研究、企业诊断与合作条件核验衔接起来，让分析结果成为管理层可以使用的决策依据。

资料与分析范围

下列说明列明资料范围、分析依据与适用条件，便于读者辨别文中的事实、推演和建议。市场进入、交易安排及适用规则仍需结合具体项目核实。

人工智能基础设施扩张使美国电力需求、局部接入和投资风险之间的关系受到关注，但设备投资、接入申请与已实现电力需求不能相互替代。本篇以截至2026 年 10 月 23 日可核验的官方材料、项目公告、同行评议研究及明确标示的工作论文与预印本为依据，开展政策与市场数据分析、工程机制比较及说明性情景研究。研究将计算活动、设施电量、接口功率、项目兑现、网络约束和费用责任纳入同一证据链，并比较 PJM 与 ERCOT 预测程序、Virginia 大型负荷费率及 Crane 供电协议。材料表明，增长预期与近期路径修订可以并存，流程批准与最终预测、年度能源与节点可交付性、财务承诺与物理运行必须分别解释。工程分析进一步区分资源充足性、网络接入与动态稳定，环境分析区分平均足迹、逐时代理与边际后果。两个可复算说明模块检验量纲、项目相关性和分期投资；在完全为作者设定的两阶段算例中，有误差信号分期策略的期望成本为 154 个抽象单位，低于一次建设的 200，但实际最坏分支达到 500，说明平均优势不等于尾部保障。本篇提出按项目状态、工程约束和履约证据设置建设触发条件。研究未识别 AI 造成的全国电网投资比例，也未建立设施级普查；其贡献是形成边界清楚、可检验和可继续补证的分析框架，而非无条件预测建设规模。对 LBNL 2026 年更新报告指定方法章节的分析显示，其出货−存量−退役−设备功率−PUE 链条由商业数据库与寿命、利用率等参数共同驱动；公开公式可审读，但不能据此独立复现专有输入或视作设施普查。报告用约 50%的并网容量利

SVSAC | sv-sac.org

用率把年电量映射为容量情景，这不是实测峰值，故全国年能量、局部接入与可靠容量仍须分别核验。

参考文献

[1] U.S. Department of Energy. DOE Releases New Report Evaluating Increase in Electricity Demand from Data Centers. 20 December 2024. https://www.energy.gov/articles/doe-releases-new-report-evaluating-increase-electricity-demand-data-centers （访问：2026-10-23）。

[2] Lawrence Berkeley National Laboratory. United States Data Center Energy Usage Report: 2025 Update. Publication page dated June 2026. https://bies.lbl.gov/publications/united-states-data-center-energy-2025 （访问：2026-10-23；出版入口未独立读取；本研究实际依据文献 60 报告原段）。

[3] PJM Interconnection. 2026 Load Forecast Report. January 2026. https://www.pjm.com/-/media/DotCom/library/reports-notices/load-forecast/2026-load-report.pdf （访问：2026-10-23）。

[4] PJM Interconnection. 2026 Load Forecast Accuracy Report. May 2026. https://www.pjm.com/-/media/DotCom/planning/res-adeq/load-forecast/2026-load-forecast-accuracy-report.pdf （访问：2026-10-23）。

[5] Virginia Joint Legislative Audit and Review Commission. Data Centers in Virginia. 9 December 2024. https://jlarc.virginia.gov/landing-2024-data-centers-in-virginia.asp （访问：2026-10-23）。

[6] Federal Energy Regulatory Commission. What State Regulators Need to Know About Order No. 1920-B. 22 April 2025. https://www.ferc.gov/what-state-regulators-need-know-about-order-no-1920-b （访问：2026-10-23）。

[7] Masanet E, Shehabi A, Lei N, Smith S, Koomey J. Recalibrating global data center energy-use estimates. Science, 2020, 367: 984–986. https://doi.org/10.1126/science.aba3758 。作者机构公开文本：https://datacenters.lbl.gov/sites/default/files/Masanet_et_al_Science_2020.full_.pdf 。

[8] DOE Data Center Load Flexibility Workshop Summary. January 2025. https://eta-publications.lbl.gov/sites/default/files/2025-03/final_doe_data_center_load_flexibility_workshop_summary.v0307.pdf （访问：2026-10-23）。

[9] U.S. Energy Information Administration. Short-Term Energy Outlook, October 2026. 6 October 2026. https://www.eia.gov/outlooks/steo/pdf/steo_full.pdf 。章节：https://www.eia.gov/outlooks/steo/report/elec_coal_renew.php （访问：2026-10-23；在线地址后续会更新，需固定版本归档）。

[10] Federal Energy Regulatory Commission. FERC Launches Aggressive Targeted Action to Speed Large Load Integration. 18 June 2026. https://www.ferc.gov/news-events/news/ferc-launches-aggressive-targeted-action-speed-large-load-integration （访问：2026-10-23；行动公告；本文只确认公告内容）。

[11] Constellation. Constellation to Launch Crane Clean Energy Center, Restoring Jobs and Carbon-Free Power to The Grid. 20 September 2024.

SVSAC | sv-sac.org

https://www.constellationenergy.com/news/2024/Constellation-to-Launch-Crane-Clean-Energy-Center-Restoring-Jobs-and-Carbon-Free-Power-to-The-Grid.html （访问：2026-10-23；历史公告）。

[12] Constellation. Crane Clean Energy Center. https://www.constellationenergy.com/about/locations/crane-clean-energy-center.html （访问：2026-10-23；动态项目网页，预计时间不是实绩）。

[13] Pennsylvania Department of Environmental Protection. Crane Clean Energy Center (CCEC). https://www.pa.gov/agencies/dep/dep-regions/southcentral-regional-office/crane-clean-energy-center-ccec （访问：2026-10-23；职责与程序说明）。

[14] U.S. Department of Energy, Office of Electricity. DOE’s Office of Electricity Publishes 2026 Draft National Transmission Needs Study to Strengthen America’s Grid. 9 July 2026. https://www.energy.gov/oe/articles/does-office-electricity-publishes-2026-draft-national-transmission-needs-study （访问：2026-10-23；草案发布公告）。

[15] ERCOT. ERCOT Releases Preliminary Long-Term Load Forecast for Years 2026–2032 for PUCT Discussion. 15 April 2026. https://www.ercot.com/news/release/04152026-ercot-releases-preliminary （访问：2026-10-23）。

[16] ERCOT. M-B051826-01 Proposal to Adjust 2026 Long-Term Load Forecast. 18 May 2026. https://www.ercot.com/services/comm/mkt_notices/M-B051826-01 （访问：2026-10-23；调整建议不是后续通过结果）。

[17] PJM. PJM’s Updated 20-Year Forecast Continues To See Significant Long-Term Load Growth. 14 January 2026. https://insidelines.pjm.com/pjms-updated-20-year-forecast-continues-to-see-significant-long-term-load-growth/（访问：2026-10-23）。

[18] Virginia SCC. In Biennial Review Ruling, SCC Creates New Class for Large-Scale Energy Users. 25 November 2025. https://www.scc.virginia.gov/about-the-scc/newsreleases/release/scc-issues-order-on-dev-biennial-review-2025/scc-rules-in-dev-biennial-review-case.html （访问：2026-10-23）。

[19] Virginia SCC. Final Order, Case No. PUR-2025-00058. 25 November 2025. https://www.scc.virginia.gov/docketsearch/DOCS/89g601!.PDF （访问：2026-10-23；定位 PDF 第 23—25 页）。

[20] Virginia SCC. SCC Data Center Initiatives. Posted 24 February 2026. https://www.scc.virginia.gov/media/sccvirginiagov-home/about-the-scc/fact-sheets/scc-data-center-initiatives-02-2026.pdf （访问：2026-10-23）。

[21] EIA. Hourly peak load in ERCOT set a new record, exceeding 91 GW on July 22. 3 August 2026. https://www.eia.gov/todayinenergy/detail.php?id=67906 （访问：2026-10-23）。

[22] Office of the Texas Governor. Governor Abbott Directs Comprehensive Data Center Audit. 3 August 2026. https://gov.texas.gov/news/post/governor-abbott-directs-comprehensive-data-center-audit （访问：2026-10-23）。

[23] Office of the Texas Governor. Governor Abbott Directs TCEQ To Halt Data Center Permits. 21 September 2026. https://gov.texas.gov/news/post/governor-abbott-directs-tceq-to-halt-data-center-permits （访问：2026-10-23；指示发布与执行结果区分）。

SVSAC | sv-sac.org

[24] ERCOT. PUCT Approves ERCOT’s Batch Zero Process for Connecting Large Electricity Users While Protecting System Reliability for Texans. 18 June 2026. https://www.ercot.com/news/release/06182026-puct-approves-ercots （访问：2026-10-23）。

[25] ERCOT, Kristi Hobbs. Item 8: System Planning and Weatherization Update. Board meeting 14 –15 September 2026. https://www.ercot.com/files/docs/2026/09/07/8-System-Planning-and-Weatherization-Update.pdf （访问：2026-10-23；第 2、4 页）。

[26] Virginia SCC. SCC Seeks Public Comments on Dominion Energy Virginia’s Integrated Resource Plan. 20 November 2024. https://www.scc.virginia.gov/about-the-scc/newsreleases/release/comments-sought-on-dev-integrated-resource-plan/comments-sought-on-dev-integrated-resource-plan.html (访问：2026-10-23；2024 版规划的官方说明，不是 2026 更新正文).

[30] Electric Reliability Council of Texas (ERCOT). “Large Load Modeling.” 页面含 AI 数据中心和加密货币挖矿 EMT 模型的发布日期、模型一般性与非特定设施说明；访问：2026-10-23。https://www.ercot.com/about/grit/large-load-modeling

[31] North American Electric Reliability Corporation (NERC). *Reliability Guideline: Risk Mitigation for Emerging Large Loads*. 2026 年 5 月。尤其第 1 章模型数据/动态模型、第 3 章资源充足性、第 4 章运行与平衡、第 2 章并网研究；自愿、非约束性指南。https://www.nerc.com/globalassets/our-work/guidelines/reliability/RG_Risk-Mitigation-For-Emerging-Large-Loads.pdf

[32] NERC. *Incident Review: Considering Simultaneous Voltage-Sensitive Load Reductions*. 2025 年 1 月 8 日。第 1—6 页，涉及 2024 年 7 月事件、UPS 与站内保护说明。https://www.nerc.com/globalassets/our-work/reports/event-reports/incident_review_large_load_loss.pdf

[33] NERC. *Characteristics and Risks of Emerging Large Loads*. 2025 年 7 月。第 3 章及第 24—25 页附近讨论事件、频率/电压机理和数据限制；属于特征与风险白皮书，不是所有设施的统计样本。https://www.nerc.com/globalassets/who-we-are/standing-committees/rstc/whitepaper-characteristics-and-risks-of-emerging-large-loads.pdf

[34] NERC. *Assessment of Gaps in Existing Practices, Requirements, and Reliability Standards for Emerging Large Loads*. 2026 年 3 月。第 5 章第 16—18 页附近讨论穿越能力、UPS、ITIC 曲线、数据缺口与测试；并非当下全国统一强制标准。https://www.nerc.com/globalassets/our-work/guidelines/reliability/white-paper---assessment-of-gaps.pdf

[35] U.S. Department of Energy, Office of Electricity. “Clean Energy Resources to Meet Data Center Electricity Demand.” 页面讨论需求灵活性、峰值及配电资源；在线页面，访问：2026-10-23。https://www.energy.gov/oe/clean-energy-resources-meet-data-center-electricity-demand

[36] Zuo, Wangda, Yangyang Fu, Xu Han, and Kyri Baker. “Assessments of Data Centers for Provision of Frequency Regulation.” *Applied Energy* 277 (2020), article 115621. DOI: 10.1016/j.apenergy.2020.115621. 研究讨论使用 IT 与冷却系统联合控制的频率调节框架；模拟/方法结果不等于美国当前大型 AI 园区的现场履约证明。https://doi.org/10.1016/j.apenergy.2020.115621

[37] Van Geet, O., et al. *Best Practices Guide for Energy-Efficient Data Center Design*. NREL/DOE，2024 年 7 月。第 5.5 节（印刷页 19）说明冷冻水蓄冷可以降低峰值电力需

SVSAC | sv-sac.org

求并提高冷却系统可靠性；不是数据中心灵活负荷能力实测研究。https://www.energy.gov/sites/default/files/2024-07/best-practice-guide-data-center-design.pdf

[38] NERC. *Technical Reference Document: Considerations for Performing an Energy Reliability Assessment*, Volume 2. 2024 年 12 月。用于能量可靠性建模边界、输电受限区域和时间耦合风险。https://www.nerc.com/globalassets/our-work/technical-reference-documents/technical-reference-document-considerations-for-performing-an-era-v2.pdf

[39] NERC. “Large Loads Frequently Asked Questions.” 2026 年 8 月 19 日。第 3—4 页概述大型负荷指南的非约束地位、NERC 后续工作状态；用于限定制度成熟度，非工程性能证据。https://www.nerc.com/globalassets/initiatives/large-loads-action-plan/large-loads-faqs.pdf

[40] Shehabi, Arman, et al. *2024 United States Data Center Energy Usage Report*. Lawrence Berkeley National Laboratory, December 2024. https://eta-publications.lbl.gov/sites/default/files/2024-12/lbnl-2024-united-states-data-center-energy-usage-report_1.pdf （访问：2026-10-23；冷却系统与 WUE：印刷页 40–47/PDF 页 39–46；水和排放估计：印刷页 56–58/PDF 页 55–57；2023 间接足迹为平衡机构电源组合模型估计，未纳入设施级 PPA 和表后发电。）

[41] U.S. Environmental Protection Agency. *Frequent Questions About eGRID*. 动态网页，含 eGRID2023 方法问答，页面更新至 2026-07-28。https://www.epa.gov/egrid/frequent-questions-about-egrid （访问：2026-10-23；问 11–13：子区域与年度/非基荷排放率；按机组利用率分类，非逐小时因果边际模型。）

[42] U.S. Energy Information Administration. “EIA’s Hourly Electric Grid Monitor provides timely data about electricity usage patterns.” 2020-04-06. https://www.eia.gov/todayinenergy/detail.php?id=43295 （访问：2026-10-23；段落 56–71 说明 EIA-930 按平衡机构采集小时负荷、发电和交换；需求未按行业分解、直接由分布式资源供给的负荷不在需求序列中，部分数据有发布滞后且实时数据可能含异常。不能单独给出设施逐时排放。）

[43] Greenhouse Gas Protocol. “Scope 2 Market-based Method Update: From 2025–2026 Public Consultation to Current Status.” 2026-10-16（网页发布日期）。https://ghgprotocol.org/blog/scope-2-market-based-method-update-2025-2026-public-consultation-current-status （访问：2026-10-23；现行 2015 市场法与拟议小时匹配/可交付性讨论；网页所述技术路线仍在审议，未作为已生效规则处理。）

[44] Siddik, Md Abu Bakar, et al. “Spatially and Temporally Detailed Water and Carbon Footprints of U.S. Electricity Generation and Use.” *Water Resources Research* 60 (2024), first published 2024-12-22. https://agupubs.onlinelibrary.wiley.com/doi/full/10.1029/2024WR038350 （访问：2026-10-23；方法与限制见第 2 节、尤其 2.3.3，结果见第 3 节，讨论见第 4 节；小时估算依赖平衡机构 EIA-930 电源结构、近期可用年度/月份强度；水耗、排放强度存在月度/年度固定代理和小机组缺数限制。）

[45] Siddik, Md Abu Bakar, et al. “The environmental footprint of data centers in the United States.” *Environmental Research Letters* 16 (2021) 064017. https://vtechworks.lib.vt.edu/bitstream/handle/10919/109747/Siddik_2021_Environ._Res._Lett._16_064017.pdf （访问：2026-10-23；方法第 2 节、结果第 3 节；研究年度 2018，数据中

SVSAC | sv-sac.org

心位置及公用事业–机组映射模型，非 2026 年观测且排除设备制造等非运行生命周期阶段。）

[46] U.S. Environmental Protection Agency. *WaterSense at Work, Section 1.5: Water-Energy Nexus*. November 2023. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P1018SQR.txt （访问：2026-10-23；第 1.5 节及“Water Used to Produce Electric Power”部分；关于取水、回流、热排放和水务电耗的机制说明，文内全国历史数值为 2014–2017 等旧数据，不用于表示 2026 现况。）

[48] U.S. Department of Energy. *Powering AI and Data Center Infrastructure: Recommendations for the Secretary of Energy*. 2024-07-30 提交，2024-08-19 发布。https://www.energy.gov/sites/default/files/2024-08/Powering%20AI%20and%20Data%20Center%20Infrastructure%20Recommendations%20July%202024.pdf （访问：2026-10-23；第 7–9 页：电网与表后方案的不确定性、不同负荷形状下技术偏好、技术/供应链评估建议；属于政策建议和议题界定，不是完成的成本或性能比较。）

[49] U.S. Environmental Protection Agency, National Risk Management Research Laboratory. *Research Summary: Water-Energy Nexus*. https://nepis.epa.gov/Exe/ZyPURL.cgi?Dockey=P100WLAF.TXT （访问：2026-10-23；摘要及第 2 节：低水位、高源水温、排水温度限制影响热电机组运行的机制；网页资料含早期研究和旧统计，本文仅作机理依据，未转为当前全国量化判断。）

[51] Koot, Martijn, and Fons Wijnhoven. “Usage Impact on Data Center Electricity Needs: A System Dynamic Forecasting Model.” Applied Energy 291 (2021), article 116798. 2021-06-01. https://doi.org/10.1016/j.apenergy.2021.116798 . 机构记录：https://research.utwente.nl/en/publications/usage-impact-on-data-center-electricity-needs-a-system-dynamic-fo/ （访问：2026-10-23；摘要与书目元数据；本文未对全文作独立复核）。定位：摘要中的模型范围、2016—2030 情景、敏感性区间及边缘/移动计算排除项。本文不讨论详细校准、验证或情景表。

[53] Mytton, David, and Muhammad Ashtine. “Sources of Data Center Energy Estimates: A Comprehensive Review.” Joule 6(9) (2022): 2032–2056. 2022-08-17. https://doi.org/10.1016/j.joule.2022.07.011 . Oxford 机构记录：https://ora.ox.ac.uk/objects/uuid:37af62a4-40ff-4a5c-ab12-778e197af64f （访问：2026-10-23；仅书目线索，摘要与正文未核；不采用样本数量、比例或附表层面的结论）。

[54] Henderson, Peter, Jieru Hu, Joshua Romoff, Emma Brunskill, Dan Jurafsky, and Joelle Pineau. “Towards the Systematic Reporting of the Energy and Carbon Footprints of Machine Learning.” *Journal of Machine Learning Research* 21(248) (2020): 1–43. 2020-11. https://jmlr.org/papers/v21/20-312.html ； 全文：https://jmlr.org/papers/volume21/20-312/20-312.pdf （访问：2026-10-23）。定位：PDF 第 1—3 页框架与 100 篇 NeurIPS 抽样；第 4—6 页 PUE、碳核算及代理估算局限；第 12—14 页实测与部分信息估计比较。实验级工具和案例，不是全数据中心能耗测量。

[55] Radovanović, Ana, et al. “Carbon-Aware Computing for Datacenters.” *IEEE Transactions on Power Systems* 38(2) (2023): 1270–1280. DOI: 10.1109/TPWRS.2022.3173250。公开作者版本：https://arxiv.org/html/2106.11750 （提交：2021-06-11；访问：2026-10-23）。定位：

SVSAC | sv-sac.org

第 II—III 节 VCC、次日预测与工作负荷边界；第 IV 节生产部署影响；讨论节预测误差及平均/边际碳强度。企业作者报告，无外部可复核的逐站原始功率序列；结果只涉及内部可延期作业。

[56] Sukprasert, Thanathorn, Abel Souza, Noman Bashir, David Irwin, and Prashant Shenoy. “On the Limitations of Carbon-Aware Temporal and Spatial Workload Shifting in the Cloud.” *Proceedings of EuroSys ’24*, 2024. https://doi.org/10.1145/3627703.3650079 . 公开全文：https://arxiv.org/html/2306.06502v2 （版本日期 2024-03-10；访问：2026-10-23）。定位：摘要、第 3—6 节及附录 A.1—A.2，使用 123 地区 2020—2022 小时碳强度与 Azure/Google 工作负荷痕迹。同行评议会议论文；调度与电力碳强度仍为模型输入，不含诱发的电网调度或扩建。

[57] Kim, Jinman, Chiwon Seo, Junsang Cha, Minhan Yoon, and Gilsoo Jang. “Coordinated Transmission Expansion Planning Considering Strategic Data Center Companies: A Game-Theoretic Approach.” International Journal of Electrical Power and Energy Systems 181 (2026), article 112173. 2026-08. https://doi.org/10.1016/j.ijepes.2026.112173 . 机构摘要：https://pure.korea.ac.kr/en/publications/coordinated-transmission-expansion-planning-considering-strategic/ （访问：2026-10-23）。定位：机构记录摘要及同行评审信息，第 17—23、86—91 段。本文未对期刊正式版全文作独立复核，细节仅据机构摘要；IEEE 基准网并非美国真实电网实测。

[58] Knittel, Christopher R., Juan Ramon L. Senga, and Shen Wang. “Flexible Data Centers and the Grid: Lower Costs, Higher Emissions?” NBER Working Paper 34065 / MIT CEEPR Working Paper 2025-14, 2025-07. https://www.nber.org/papers/w34065 ; MIT CEEPR：https://ceepr.mit.edu/flexible-data-centers-and-the-grid-lower-costs-higher-emissions/ （访问：2026-10-23）。定位：MIT 页面第 83—105 段，GenX、三区域、1—24 小时转移窗口、1%—100%灵活负荷比例及德州成本反事实。工作论文，截止日未同行评议；结果为模型情景而非实际排放测量。

[59] Colangelo, Philip, et al. “Turning AI Data Centers into Grid-Interactive Assets: Results from a Field Demonstration in Phoenix, Arizona.” arXiv:2507.00909，2025-07-01。https://arxiv.org/abs/2507.00909 ; 全文：https://arxiv.org/html/2507.00909v1 （访问：2026-10-23）。定位：第 2.2、3.3—3.4 节，第 69—75 段试验地点、日期和目标，第 98—104 段负荷敏感性。预印本、未同行评议；单集群两事件，不是机群代表性样本或多季节可靠容量评估。

[60] Sarah J. Smith; Alex Hubbard; Alex Newkirk; Mohan Ganeshalingam; Billie Holecek; Dale Sartor; Michael Mills; Arman Shehabi. United States Data Center Energy Usage Report: 2025 Update. 2026-06-18. https://www.energy.gov/documents/united-states-data-center-energy-usage-report-2025-update

SVSAC | sv-sac.org