



Ganzheitliches Management von Energie- und Ressourceneffizienz in Unternehmen (MERU)

Rebound-Effekte und Wirkungsdefizite von Effizienzmaßnahmen in Rechenzentren

Problemanalyse und Handlungsoptionen

// Dr. Dieter Thiel

Technische Transformation, Digitalisierung, Künstliche Intelligenz, soziale Medien, Entertainment: Der IT-Markt und damit der Bedarf an Rechenzentren zur Verarbeitung der Daten wächst. Knapp 17 Milliarden kWh Strom sind es aktuell in Deutschland, die die rund 52.000 Serverräume, diversen Rechenzentren und Hyperscaler in Deutschland verbrauchen. Nahezu die gleiche Menge geht als Abwärme ungenutzt in die Umwelt. Das entspricht rund 2 % unseres jährlichen Stromverbrauches. Dabei werden rund 7,1 Mio. Tonnen CO₂ emittiert. Dieser Trend ist ungebrochen: Die Wachstumserwartung liegt bei rund 10 bis 12 % pro Jahr.

Die Energieknappheit ist spätestens mit dem Ukraine-Krieg offensichtlich, die Strompreise 2022 um das 2 bis 3-fache gestiegen, Gas ist teuer und knapp. Damit steigt der ökologische und ökonomische Druck zur Ressourceneffizienz. Energieeffizienzgesetz und Sofortprogramm der Bundesregierung geben erste Leitplanken für die Energieeffizienz vor. Gefordert wird Klimaneutralität neuer Rechenzentren bis 2027, ein öffentliches Effizienzregister sowie eine Abwärmenutzung zumindest eines Teils der Abwärme und ein PUE von < 1,3. (PUE bzw. die Power Usage Effectiveness definiert das Verhältnis von Gesamtenergiebedarf der Rechenzentren im Verhältnis zum Energiebedarf der IT).

Getrieben durch die hohen Energiekosten sind die Energieeffizienzsteigerungen in den letzten Jahren, insbesondere für Klimatisierung und Stromversorgung, immens. Vor 10 Jahren lagen die PUE-Werte noch in der Größenordnung von 2, jetzt liegen sie bei neuen Rechenzentren in der einer Range um 1,3. Nach den Effizienzsteigerungen, die sich bisher im Wesentlichen auf die Infrastruktur konzentrierten, stellt mit fallenden PUE-Werten der Energiebedarf der IT mit rund 75% zukünftig die größte Stell-schraube dar. An ihr muss gedreht werden, um die Energieeffizienz weiter zu verbessern.

Es sind zwischenzeitlich alle Werkzeuge und Komponenten vorhanden, um den Energieverbrauch zu senken. Der steigende Bedarf führt jedoch immer noch zu einem Verbrauchsanstieg. Hintergrund: Der kaum abschätzbare zukünftige IT-Bedarf macht verlässliche Wachstumsprognosen nahezu unmöglich. Allerdings ist eine genaue Prognose die Grundlage eines ressourcenschonenden Energiedesigns. Denn ohne genaue Dimensionierung verpuffen Effizienzmaßnahmen. Werden die Infrastruktur für die Kühlung und die unterbrechungsfreie Stromversorgung (USV) überdimensioniert, sind zu viele Server nicht ausgelastet. Damit wird die Wirksamkeit einzelner, für sich effizienter Maßnahmen, im Gesamtsystem faktisch gemindert.



Die Möglichkeiten, die Effizienz zu verbessern, sind unterschiedlich groß. Bei selbstgenutzten Rechenzentren hat der Nutzer Einfluss auf beide Wirkfaktoren, sowohl auf die Infrastruktur als auch die IT-Architektur. Bei angemieteten IT-Flächen bestimmt der Colocator die Effizienz der Infrastruktur. Der Mieter konzentriert sich auf die Effizienz seiner IT. Bei Cloud-Anwendungen gibt der Endnutzer sogar alle Einflussfaktoren auf die Ressourceneffizienz aus der Hand.

Zentrale Handlungsempfehlungen

- Eine realistische Prognose über den zukünftigen Bedarf ist zwingend erforderlich. Überschätzungen führen zu Überdimensionierung. Diese verhindert die Energieeffizienz von Infrastruktur und IT und führt zu erhöhtem Ressourcenverbrauch.
- Durch eine Skalierbarkeit der Infrastruktur wird eine Anpassung auch bei unsicherer Prognose in der Zukunft ermöglicht.
- Virtualisierung von Servern, Zusammenlegung unausgelasteter IT-Komponenten
- Betriebsüberwachung und Monitoring

Rebound-Effekte im Bereich Rechenzentrums-Infrastruktur

Überhöhte Wachstumsprognosen führen zu überdimensionierten Komponenten der technischen Infrastruktur. Das Auslegungsdesign der energieintensiven Klimatisierung und unterbrechungsfreie Stromversorgung (USV) entspricht aber genau den falschen Prognosewerten. Hohe Ansprüche an die Verfügbarkeit und Betriebssicherheit erfordern zudem redundante Komponenten, die je nach Betriebsweise zu einem Realbetrieb im Teillastbereich führen können. So entsteht ein komplexes System aus hocheffizienten Einzelkomponenten, optimiert für den vorgegebenen Betriebspunkt.

Was passiert jedoch, wenn die reale Betriebs- und Auslastungssituation vom Auslegungsdesign abweicht und die Sollwerte nicht erreicht werden? Die Performance der Komponenten, die ihre höchste Effizienz im Auslegungspunkt haben, nimmt dann rapide ab. Eine Leistungsregelung auf den tatsächlichen Betriebspunkt ist nur mit einem teilweise hohen Verlust von Effizienz möglich. Im ungünstigsten Fall rutschen die Komponenten der Infrastruktur für Klimatisierung und unterbrechungsfreie Stromversorgung in einen Betriebsbereich, der nicht mehr regelbar ist und nur einen taktenden Betrieb zulässt.

Dies hat zur Folge, dass aus dem Auslegungsdesign mit vielen hocheffizienten Einzelkomponenten im tatsächlichen Betrieb ein ineffizientes System wird, weil es in Teillastbereich betrieben wird. Die Effizienz wird zerstört.

Im MERU-Vorhaben konnte gezeigt werden, dass durch die Konsolidierung von drei nicht ausgelasteten Einzelrechenzentren, die im Teillast betrieben wurden, zu einem gut ausgelasteten Rechenzentrum in Verbindung mit einigen technischen Anpassungen 50 % Energie eingespart werden konnte. Auch die Server- und Storage-Systeme wurden am neuen Standort sukzessive durch moderne Komponenten ersetzt.

Die stärkere Nachfrage für IT-Leistungen führt allerdings neuerdings wieder zu einem Anstieg des Bedarfs. Auch wenn dieser noch nicht die Werte von vor der Konsolidierung erreicht hat, signalisiert er, wie der zunehmende Bedarf die Ressourceneffizienz mindert.



Wirkungsdefizite im Bereich IT

Bei gemieteten Rechenzentrums-Bereichen wird die Effizienz der Infrastruktur vom Betreiber bestimmt. Er dimensioniert die erforderlichen Komponenten entsprechend den Anforderungen der Mieter und vermietet ihm IT-Fläche. Der Mieter konzentriert sich auf die Effizienzsteigerung seiner eigenen IT-Komponenten. Hauptverbraucher sind hier die Server.

Genau wie bei der Infrastruktur ist auch hier die Auslegungsprognose für Leistungsfähigkeit und Auslastung entscheidend für den energieeffizienten Betrieb. Vor dem Hintergrund, dass der Energieverbrauch für die IT mehr als 2/3 des Gesamtverbrauches eines Rechenzentrums ausmacht, ist diese Optimierung umso wichtiger.

Ähnlich wie bei der Infrastruktur ist die Entwicklung auch hier in den letzten Jahren weiter gegangen. Die Effizienz der Einzelkomponenten im Auslegungsfall hat sich deutlich erhöht. Das Grundproblem jedoch ist geblieben. Die Effizienzverluste im Teillastbetrieb sind sogar noch höher als bei der Infrastruktur. 20% Auslastung bedeuten bereits rund 75% des maximalen Stromverbrauches unter Vollast.

Virtualisierungssysteme, die teilausgelastete Server in optimale Auslastungsbereiche zusammenlegen, konnten den Bedarf mit vergleichsweise wenig Aufwand verringern. Erweiterte Kundenwünsche lassen den Energieaufwand zwischenzeitlich aber auf die Absolutwerte der Vorjahre steigen.

Reboundeffekte durch verbesserte Rechnerperformance und Konkurrenzdruck

Hochleistungsrechner sind das beste Beispiel, wie die Energieeffizienz von Großrechnern gesteigert wird, Rechenprozesse dadurch schneller ablaufen und sich so neue Anwendungsmöglichkeiten eröffnen – und damit Rebound-Effekte entstehen können. National und international herrscht zwischen den Großrechenbetreibern ein harter Konkurrenzkampf. Größe und Leistungsfähigkeit ist im Top 500-Ranking der Superrechner im Netz für jeden nachlesbar. Der Konkurrenzkampf befeuert einerseits die Entwicklung zur Energieeffizienz. Der Energieaufwand pro Rechenprozess hat sich so in den letzten Jahren um den Faktor 3-5 verbessert. Der Rechenvorgang verläuft schneller und erfordert nur noch 25 % des Energieverbrauchs.

Andererseits eröffnen sich so neue vielfältigere Anwendungsmöglichkeiten. Rechenvorgänge, die früher Tage oder benötigen, werden in wenigen Stunden oder Minuten abgewickelt. Neue Anwendungsmöglichkeiten, Mehrbedarf und Konkurrenzdruck fraßen im Projekt den gesamten Effizienzgewinn auf. Hierbei handelt es sich um „Output-Rebounds“ (Wolff et al. 2023).

Wie sehr die Entwicklung expandiert, zeigt die Planung für die nächste Entwicklungsstufe, die Verdoppelung der Rechnerkapazität in den nächsten 4-5 Jahren, um den explodierenden Bedarf zu decken und wettbewerbsfähig zu bleiben.



Fazit

Rechenzentren sind energieintensive, ausschließlich stromverbrauchende Industriebereiche in einem immer weiter im zweistelligen Prozentbereich jährlich wachsenden Markt, die ihre Abwärme ungenutzt an die Umgebung abgeben.

Die Energieeffizienzbestrebungen im Bereich Infrastruktur und IT-Komponenten haben die Performance in den letzten 10 Jahren bereits um 20 bis 30 % verbessert. Es gibt viele hocheffiziente Einzelkomponenten. Dies führt zu einem relativen Sinken des spezifischen Verbrauches. Absolut gesehen treiben aber neue Anwendungen und die überall präsente Digitalisierung diesen weiter in die Höhe.

Eine der wichtigsten Erkenntnisse aus dem MERU-Projekt zeigt allerdings eine fast grotesk erscheinende Situation: Die Wirksamkeit hocheffizienter Einzelkomponenten, unabhängig ob Infrastruktur oder IT, wird bei Überdimensionierung aufgrund falscher Prognosewerte oder Planungsfehler zunichte gemacht. Das Gesamtdesign muss stimmen.

Skalierbare Technik, die auf unsichere Wachstumsprognosen reagiert, ist zwingend erforderlich.

Der Entwurf des Energieeffizienzgesetzes mit der Forderung nach Klimaneutralität, einer Zielgröße für die Energieeffizienz und der Pflicht zur Abwärmenutzung, geht in die richtige Richtung.

Ergänzend sollte zukünftig ähnlich wie in nordeuropäischen Ländern ein Konzept zu Abwärmenutzung Bestandteil der Baugenehmigung sein.

Literatur

Wolff, F., Gensch, C.-O., Kampffmeyer, N.; Schöpflin, P.; Lautermann, C.; Gebauer, J.; Schaltegger, S.; Norris, S.; Wüst, S.; Thiel, D.; Buda, F. (2022): Ganzheitliches Management von Energie- und Ressourceneffizienz in Unternehmen: Wie können Rebound-Effekte vermindert werden? Handlungsoptionen für die Politik. <https://www.meru-projekt.de/publikationen>

Wolff, F.; Gensch, C.-O.; Kampffmeyer, N.; Schöpflin, P.; Lautermann, C.; Gebauer, J.; Schaltegger, S.; Norris, S.; Wüst, S.; Thiel, D.; Buda, F. (2023): Leitfaden für Unternehmen zum Management und der Vermeidung von Rebound-Effekten. Handreichung im Rahmen des Projekts „Ganzheitliches Management von Energie- und Ressourceneffizienz in Unternehmen“ (MERU). Verfügbar unter <https://www.meru-projekt.de/publikationen>

Wüst, S.; Schaltegger, S.; Wolff, F.; Lautermann, C. & Schöpflin, P. (2023): Konzeptioneller Rahmen zur Erforschung von unternehmens-bezogenen Rebound-Effekten. Konzeptpapier im Rahmen des Projekts „Ganzheitliches Management von Energie- und Ressourceneffizienz in Unternehmen“ (MERU). Unter Mitarbeit von: Gensch, C.-O.; Kampffmeyer, N.; Gebauer, J.; Thiel, D.; Verfügbar unter <https://www.meru-projekt.de/publikationen>



Das Projekt MERU

Das Vorhaben „Ganzheitliches Management von Energie- und Ressourceneffizienz in Unternehmen“ wird vom Bundesministerium für Bildung und Forschung (BMBF) gefördert.

Die Verbundpartner des Vorhabens sind Öko-Institut, Institut für ökologische Wirtschaftsforschung (IÖW), Leuphana Universität (Centre for Sustainability Management), Data Center Group (DCG) und der B.A.U.M. e.V..

www.meru-projekt.de

Kontakt

Franziska Wolff | Öko-Institut e.V. | Projektleitung MERU | +49 30 405085-371 | f.wolff@oeko.de

Max Mustermann | [Partnerorganisation] | +49 ### | m.mustermann@partner.de
