

Energie-Effizienz in Vergabeverfahren

- Wie bewertet man Energie-Verbrauch? Wie geht man mit Idle-Power-Consumption um?
- Wie lässt sich Energie-Verbrauch oder Energie-Effizienz praktikabel in Bewertungs-Funktionen einbauen?
- Welche in der Praxis verwendeten Kriterien gibt es?
- Wie können Nachbesserungsklauseln formuliert werden?
- Wie wird Energie(Strom)-Verbrauch gemessen?
- Wie wird mit möglicher Sekundär-Nutzung von Energie umgegangen?
- Wie gut vorhersagbar sind solche Messungen, insbesondere bei Hardware, die zum Vergabe-Zeitpunkt noch in Entwicklung ist?
- Welchen Einfluss haben BIOS-Einstellungen?
- Welche Benchmarks verwendet man zur Messung?
- Was kann man von einem Vergabeverfahren zum anderen übertragen?
- Was ist rechtssicher im Vergabeverfahren möglich?
- Wie gehe ich mit der Alterung der Hardware um? Wie empfindlich sind die Kennzahlen und Bewertungen gegenüber den technischen Fortschritten?
- Welche alternativen Betreibermodelle für das Gesamtsystem wären denkbar (z.B. TCO)?
- Welche alternativen Betriebskonzepte sind machbar (z.B. Power-Capping)?
- Welche Vereinfachungen sind möglich?
- Wie kann die Risiko-Verteilung zwischen Anbietern und Auftraggebern so gestaltet werden, dass alle Anbieter fair behandelt werden?
- Wie geht man mit begrenzenden Faktoren (maximale Anschlussleistung/Kühlleistung) um?
- Welche Gestaltungsmöglichkeiten haben Anbieter, um den Energieverbrauch zu beeinflussen?
- Wie vermeiden wir Augenwischerei?
- Welche Konsequenzen sind bei schlechter Energie-Effizienz zu erwarten?

Energie-Effizienz in Vergabeverfahren

Messen

- Wie misst und bewertet man Energie-Verbrauch oder Energie-Effizienz?
- Welche Kennzahlen kann man messen?
- Wie geht man mit Idle-Power-Consumption um?
- Wie wird Energie(Strom)-Verbrauch gemessen?
- Wie lässt sich Energie-Verbrauch oder Energie-Effizienz praktikabel in Bewertungs-Funktionen einbauen?
- Wie gut vorhersagbar sind solche Messungen, insbesondere bei Hardware, die zum Vergabe-Zeitpunkt noch in Entwicklung ist?
- Welchen Einfluss haben BIOS-Einstellungen?
- Welche Benchmarks verwendet man zur Messung?

Bewerten und Vergabepaxis

- Welche Vereinfachungen sind möglich?
- Welche Kriterien für Energie-Verbrauch/Effizienz werden in der Praxis verwendet?
- Wie geht man mit begrenzenden Faktoren (maximale Anschlussleistung/Kühlleistung) um?
- Wie wird mit möglicher Sekundär-Nutzung von Energie umgegangen?

- Was ist rechtssicher im Vergabeverfahren möglich?
- Wie können Nachbesserungsklauseln formuliert werden?
- Was kann man von einem Vergabeverfahren zum anderen übertragen?

Chancen, Risiken, Alternativen

- Wie vermeiden wir Augenwischerei?
- Wie kann die Risiko-Verteilung zwischen Anbietern und Auftraggebern so gestaltet werden, dass alle Anbieter fair behandelt werden?
- Welche Gestaltungsmöglichkeiten haben Anbieter, um den Energieverbrauch zu beeinflussen?
- Welche Konsequenzen sind bei schlechter Energie-Effizienz (z.B. aufgrund ungeeigneter Vergabekriterien) zu erwarten?
- Wie gehe ich mit der Alterung der Hardware um? Wie empfindlich sind die Kennzahlen und Bewertungen gegenüber den technischen Fortschritten?
- Welche alternativen Betreibermodelle für das Gesamtsystem wären denkbar (z.B. TCO)?
- Welche alternativen Betriebskonzepte sind machbar (z.B. Power-Capping)?

Messen – Energieeffizienz für Vergabeverfahren

Messverfahren Strom

- Mögliche Verfahren: Ablesung der Werte an PDU, am zentralen Verteiler (Sicherungen), IPMI
 - Erfahrung: IPMI zu ungenau
 - Messung immer “am weitesten weg” von Details der Hardware, bspw. Smart-meter PDUs, Rack-PDUs
 - Aufpunkt am besten für komplettes Rack; aber: Hersteller haben meist kein komplettes Rack
 - Hersteller: Skalierung auf gesamtes Rack auf Basis der Komponentenverbräuche meist ziemlich gut
 - Herstellersicht: Verbrauchsangabe muss gewisse Streuung / Unsicherheit ohne Strafe zulassen
- **Fazit:** Sinnvollster Messpunkt ist der Gesamtstromverbrauch aus der Steckdose (**Strommessung**) /Verteilerkasten; möglichst zeitlichen Verlauf über Benchmark-Laufzeit messen (**elektr. Leistung**)
- Einfluss der Umgebung (bspw. Temperatur des Kühlwassers)
 - Leistungsaufnahme für Side-Cooler vs. Warmwasserkühlung: Leitung für Kühlung abhängig von lokaler Installation; besser raus aus der Bewertung

Messen – Energieeffizienz für Vergabeverfahren

Messverfahren Benchmarks

- Benchmarks; Standardverfahren oder Spezialbenchmarks?
 - Statement: eigener Benchmark-Mix ist sehr aufwändig für alle; Generell Linpack verwenden? HPL?
 - Linpack / HPL decken meist mindestens 70% der Last ab.
 - Entscheidene Größe / Definition Thermal Design Power als erste Orientierung; aber nur **messen** zählt. Beobachtung/Anhaltspunkt: Linpack produziert typisch 70% TDP
 - TDP nur Anhaltspunkt;
 - Weiterer Verbrauch auf dem Knoten: Netzteile, IB-Karten, Lüfter, Speicher,...
 - Weiterer Verbräuche des Gesamtsystems: Netzwerk, Speicher, Fileserver,...
 - Aber Linpack-Problematik u.a. Nutzung von Libraries / Umsetzung; bspw. AVX 512 vs?
 - HPL schwierig; wie definiert man den? Firmen nehmen eigene, nicht nachvollziehbare Optimierungen vor
 - HPCCG als Benchmark verwenden?
 - Vorgaben: für alle zu verwendenden Bechmarks sämtliche BIOS-Einstellungen beibehalten
 - Compiler-Vorgaben? Kann man nicht mehr machen, da wir derzeit mehrere CPU-Hersteller haben
 - FP64-Performance ggf. Nicht mehr entscheidender Maßstab für Jobmix; GPUs mit FP8, FP16...
 - **Empfehlung: Einen Standard-Benchmark für CPU-System (LINPACK), einen für Speicherbandbreite (STREAM), einen für GPU-/KI/ML-Anwendungen (GPU-burn?)**
- Problematik: Projektion auf die nächste(n) noch nicht verfügbare Hardware
- Problematik: Maximale Anschlussleistung vs. mittlerer Energieverlauf auf Basis Benchmarks

Messen – Energieeffizienz für Vergabeverfahren

Wie bekomme ich die Messwerte eingefordert?

- Anfragen bei Anbieter: Maximale Stromaufnahme über HPL
- Statement: LinPACK gut als burn-in; Maximale Stromaufnahme wird mit LINPACK oder auch
 - Erfahrungswert: Alltagsvollast ist typisch 80% vom LINPACK (CPU double precision)
 - Unterschied zu GPUs und LLM Benchmark-Last kann **über** LINPACK-Last liegen
 - Anwendungen gehen Richtung mixed precision / GPU
- Benchmarks
 - TOP 500 -Benchmark
 - Green IT –Benchmark
 - Eigene Applikationen
 - Benchmark für Stresstest (wenn das System steht)

Vorgaben bei der Ausschreibung

- Budget für Invest
- Maximale Stromaufnahme (entweder technisch bedingt oder als Limit für Betriebskosten)
- Mindestperformance pro Euro (“muss besser sein als das, was wir bereits betreiben”)
- Mindestperformance pro Watt (“muss besser sein als das, was wir bereits betreiben”)
 - Mögliche Schwierigkeit: vorgegebenes Strombudget kann bereits die Gesamtinvestition limitieren
 - Weitere Schwierigkeit: Wechselkursschwankungen und Marktpolitik (bspw. Exploriende Komponentenpreise)

Messen – Energieeffizienz für Vergabeverfahren

Hinweis zur Riskikobetrachtung Anbieter vs Kunde

- Strafen für zu späte Lieferung etc. sind eigentlich nicht mehr machbar
- Ausschreibungszeitskalen möglichst kurz gestalten, um Marktschwankung weitestgehend auszuschließen

Energieeffizienz

- Vergabe von Punkten für anteilig effizient gekühlte Komponenten
 - Warmwasser der CPU, bspw. Zusatzpunkte für flüssiggekühlte RAM-Kühlung, Netzteile etc.
 - Aber: Kostenfaktor! Inwieweit sind Unis bereit / in der Lage, für Effizienzoptimierung zu bezahlen?
 - Können / wollen Unis Aufwand für bspw. CO2-Verbrauch bei Herstellung berücksichtigen?

Bios-Einstellungen

- So gut wie nicht zu berücksichtigen
 - Einstellungen der CPU-/Bordhersteller ändern sich in nicht nachvollziehbarer Weise
 - BIOS-Updates machen ggf. überprüfte Detailsinstellungen wieder zunichte
 - Reproduzierbarkeit über mehrere Monate nicht zu gewährleisten

Messen – Energieeffizienz für Vergabeverfahren

Idle-Power-Consumption

- Abfragen, was der Hersteller liefern kann
- Anbieter- /Herstellerverfahren zum Zurückschalten auf Idle-Power-Consumption mit berücksichtigen / bewerten
- Am besten: Abschalten, wenn Idle
 - Aber Achtung, keine Überbewertung. Idle-Verbrauch im Idealfall (bei vernünftiger Auslastung) im Prozentbereich des gesamten Betriebs; wichtiger: „right-sized system“ = Gesamtauslastung > 80%

Fazit bisher:

- Messungen, Vorgaben, Bewertungen nicht besser als bis auf 10% möglich
- Einfache und durchschaubare Mess- und Bewertungsverfahren notwendig
 - nicht zu kleinteilig, da entweder keine Angebote oder Einspruchsgefechte