

## 脱炭素向けワークロード再配置方式の提案と評価 Proposal and evaluation of workload replacement schemes for decarbonization

金子 聡  
Satoshi Kaneko

### 1. はじめに

AI 活用により知識を生み出すデータセンタ (DC) は同時に大量の電力を消費し、CN2050 に向けた社会課題となっている。DC の脱炭素化に向け、IT 機器や冷却設備の効率化などの施策により消費電力の削減が継続的に進められている一方で、消費電力に、CO<sub>2</sub> を排出しない再生可能エネルギー (再エネ) を採用する動きが活発化している。再エネの導入方法について、昨今、先進的な企業は社会全体の再エネ導入拡大のために、再エネの発電と消費を 1 時間単位で合致させる取り組みを始めている。

一方、DC は、一日を通じて大量の電力を必要とするため、不安定かつ地理的制約を有する再エネの効率利用が課題である。そこで、異なる場所にある複数 DC の中から、再エネが豊富な地域及び時間帯に応じて、計算負荷 (WL: Workload) を実行する DC を選択する技術が注目されている。このとき、実行中の WL を別の地域で実行させるには、WL により構成されるアプリケーション(アプリ)のダウンタイムやアプリが利用するデータのコピーに伴う電力消費といったオーバーヘッドの削減が課題となっている。

そこで、本報では、アプリ再配置のオーバーヘッドを最小化する移行 WL 選出手法[1]についてアプリの稼働情報などを用いて評価を行い、方式の有用性について考察を行う。

### 2. 移行 WL 選出手法

評価対象の WL 選出手法の概要を図 1 に示す。本手法では、サイト間毎の電力需要調整量を推定し、当該電力量に相当する WL を、WL と WL が利用するデータの情報を用い、移行オーバーヘッドであるアプリダウンタイム、データコピーによる消費電力を最小化するように図の手順で選出する。サイト間の電力需要調整量とは、Site1 が再エネ不足となり、Site2 が再エネ過剰となった場合に、Site1 から Site2 へ調整する電力量を示す。

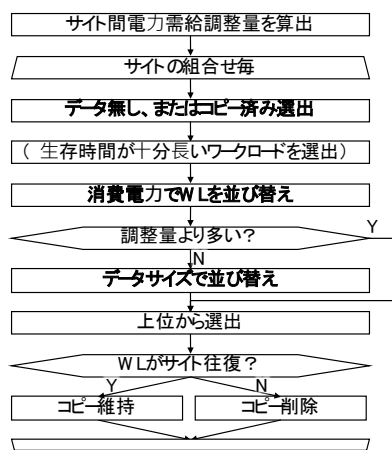


図 1 WL 移行手法概要

### 3. 提案手法の評価と考察

#### 3.1 評価方法

提案手法により選出された WL 移行時の、アプリの総ダウンタイムと、アプリが利用するデータのコピーにかかる消費電力を計算し、提案手法によるそれぞれの削減率と削減量を測定する。まず評価の前提について説明する。

##### 3.1.1 対象のワークロード

実際のデータセンタにおける有用性を測るべく、実際のデータセンタにおける WL のデータを利用した。対象のデータは HPC 系のデータセンタのジョブの実行ログデータであり、WL の開始時刻、終了時刻、CPU/Memory の利用率の情報を持つ。

提案手法は WL の属性情報として、消費電力と利用するデータの有無、データのサイズの情報を用いて WL の選出を行う。よってこれらの属性情報が必要となるが、これらの属性情報は対象の WL データに含まれていなかった。そこで、不足している情報についてはそれぞれ仮の値を用いた。WL の消費電力については、WL のリソースの利用率と WL が実行されるサーバの消費電力から WL の消費電力を推定する技術 [2]を用いて、前述の WL データから WL 消費電力の推定を行った。データの有無については、半数弱の WL がデータを用いた処理を行っているとして想定し、WL 全体の三分の一として 33% の WL をランダムに抽出し、データ領域をマウントして利用しているものと仮定した。更にこの WL が持つデータのサイズについては 10GB から 500GB の間でランダムに決定した。

##### 3.1.2 電力需要調整シナリオ

評価において、実際のデータセンタの CO<sub>2</sub> 排出量削減に必要な電力需要調整量について想定を置く必要がある。そこで、先進的な DC で取り組まれているデータセンタの電力消費と再生可能エネルギーの発電を時間単位のマッチングの方法を参考とし、電力需要を調整することとした。

太陽光や風力といった再エネの発電量は時間変動するため、時差により再エネの発電量の多寡が逆転するサイトの組合せを想定し、本評価では、13-16 時間の時差が存在する日本とアメリカの 2 拠点のデータセンタ間で移行すると仮定した。データセンタへの再エネ供給量は、実際は電力供給契約により決まると想定されるが、本評価ではその地域の気候のデータに基づいた推定値を用いた。再エネは太陽光発電と風力発電と対象とした。太陽光発電と風力発電の発電量の推定式は Geo らの従来技術[3]の式(21)と(22)を参考にした。これらの式によると、発電設備の規模と気候データである日射量と風速によって発電量を計算できる。

また、日本の気候データについては気象庁の公開データを利用し、東京のデータを用いた。アメリカの気候データについては国立再生可能エネルギー研究所の公開データを利用し、Denver のデータを用いた。電力需要調整の期間は、図 1 の WL 往復判定処理をシミュレーションすることを目

的とし、再エネ供給量が時間によって逆転するケースをカバーすべく、12 時間のデータを用いた。

発電設備の規模については、前項で述べた WL 消費電力推定結果を用いて、12 時間分の WL 消費電力の総和と一致するように、12 時間分の再エネ発電量が一致するように、設定した。これにより、発電量の時間的特性は各地域の自然状況に基づきながら、データセンタの規模に応じた再エネ発電規模をシミュレートした。

### 3.1.3 評価指標

評価指標はアプリの総ダウンタイムと、データコピーにかかる消費電力である。其々の計算方法について述べる。

アプリのダウンタイムはアプリの処理が移行元から移行先への切り替えに要する時間である。処理の切り替え時間はアプリのデータのコピー要否によって異なる。

データコピーが不要なアプリの場合、移行先でアプリ起動後に処理を移行先のアプリに切り替えるだけで良い。具体的には、クライアントから処理を受け付けるサーバアプリが例示でき、移行元と先の間でロードバランサ (LB) を設定し、移行先でアプリ起動後に移行元を削除して処理を切り替えることで移行可能である。このとき切替時間は LB の切替時間であり、長くとも 1 分未満と想定される。

データコピーが必要なアプリの場合、データコピー完了のタイミングで、データ同期を解除し、移行先アプリを起動する。よって、データ同期解除とアプリ起動の時間を要する。データ同期解除とは、移行元のアプリを削除しストレージサービスに対してデータ同期解除指示を出す、方法が挙げられる。アプリ削除・起動の時間は、アプリによっても異なるが数秒から数分と考えられる。データ同期解除指示については、ストレージサービスの仕様依存する。

以上に基づき、ダウンタイムはデータコピー不要なアプリは 1 分、データコピー用のアプリは 5 分と設定した。

次に消費電力について、データコピーに要する電力はコピーデータ容量に比例する。その係数はデータセンタの構成やネットワーク機器によっても異なるため一概に定義することはできないが、典型的な係数について過去に議論されている論文[4]に基づいて 0.06kWh/GB の値を用いることとした。よって、この係数とデータ容量を掛け合わせることでオーバーヘッドとなる消費電力を算出した。

## 3.2 比較対象

移行 WL 選出の提案手法について評価を行うが、比較対象として a)ランダム, b)少データ移行量優先, c)低 WL 消費電力優先, d)低 WL 消費電力かつデータ移行不要優先(提案手法), e)低 WL 消費電力かつデータ移行不要かつデータコピー済み優先(提案手法), の 5 つの WL 選出手法とした。

a は WL を無作為に選出する。b は従来研究[5]で論じられているデータ転送量を優先する手法を想定した。従来研究では、仮想マシン (VM) の移行のオーバーヘッドをデータ転送量で評価しており、転送データ総量とピークスループットを下げるように最適化して VM 移行計画を立てている。そこで、b では移行するデータ量を最小化するように VM を選出した。c は WL の消費電力が低いものを優先し、データを考慮しない。d は、データ移行が不要な WL 群の中から消費電力が低い WL を優先選出し、それでも足りない場合にデータ移行が必要な WL 群の中から消費電力が低い WL を優先選出する。e は、データ移行が不要な WL 群の中

から消費電力が低い WL を優先選出し、その中でもすでに移行された WL を優先して選出する。さらに、それでも足りない場合は、d と同様にデータ移行が必要な WL 群の中から消費電力が低い WL を優先選出する。(e)がダウンタイムと電力消費オーバーヘッドの双方を考慮した方式であり、本提案手法だが、データコピー有無を考慮することの効果の評価するために提案手法からデータコピー有無評価を行わない手法として(d)の評価も行った。

以上に示す、a,b,c,d,e の手法について実験を行った。

## 4. 評価結果と考察

表 1 実験結果

	a	b	c	d	e
WL 数	87	100	64	68	68
ダウンタイム(分)	189	143	144	100	100
データ量(GB)	6141	1457	4236	1698	1134

実験結果を表 1 に示す。移行 WL 数は b が最も多いがこれは、データを持たない WL が優先して選ばれているため、WL の消費電力を考慮した方式(c)(d)(e)よりも多くなっている。また、データを保持する WL を全体の三分の一かつランダムに決定した都合上、三分の二の WL が優先的に選ばれ、(a)ランダムよりも WL 数が多くなっている。

WL のダウンタイムの総和 (a)ランダムが最もダウンタイムが長い、これは WL 数が二番目に多い上にデータ有無を考慮していないことが原因と考えられる。アプリのデータ有無によりダウンタイムが異なるためである。最も WL 数が多かった(b)は(a)に比べてダウンタイムが短く、データ有無及びデータサイズを考慮している効果と見られる。(c)は(b)に比べ WL 数が少なかったにも関わらず、ダウンタイムが同等となっており、これもデータサイズの影響と見られる。一方で、WL 数とデータの両方を考慮している(d)(e)は同様にダウンタイムが一番低い、という結果になった。

データの総量は(a)が最も多く、データを持った WL 数が多く、かつデータ量が多い WL も多いと想定される。(c)はデータを考慮していないため(a)に次いでデータ量が多い。(d)は WL 消費電力とデータの双方を考慮しているため、(c)に比べ半分以下データ量が少ないが、(b)よりは大きい。一方、コピー済みのデータ構成を考慮する(e)は、データ有無及びデータサイズのみを考慮した(b)よりも少なく、全手法のうち最小となった。

以上により提案手法は、データサイズのみを考慮した従来手法に比べダウンタイムが 31% 短く、また、コピーデータ量も 22% 少なく、手法の有効性を示せたと考える。

### 参考文献

- [1] 金子 聡, “脱炭素向け再配置アプリケーション選定手法の提案”, 論文誌名, FIT2023 (2023).
- [2] P. J. Chaudhari, S. Kaneko, T. Okamura, “Estimating Power Consumption of Collocated Workloads in a Real-World Data Center,” (*SoftCOM 2023*) 2023
- [3] H. W. H. S. Jiechao Gao, “Smartly Handling Renewable Energy Instability in Supporting A Cloud Datacenter,” (*IPDPS2020*) 2020
- [4] J. M. K. K. J. a. F. C. Aslan, “Electricity Intensity of Internet Data Transmission: Untangling the Estimates,” *Journal of Industrial Ecology*, 22, 4, pp. 785-798, 2017.
- [5] J. S. S. N. S. I. A. B. R. C. S. S. a. S. K. Anup Agarwal, “Redesigning Data Centers for Renewable Energy,” (*HotNets'21*) 2021.