



Opracowanie strategii zarządzania zasobami dla systemów chmurowych wykorzystujących źródła zielonej energii

Zofia Wrona

opiekun projektu: dr hab. Marcin Paprzycki

Przyczyna: Duża popularność chmury obliczeniowej + „zielone” chmury obliczeniowe

Problem: Jak skutecznie i wydajnie zarządzać alokacją zasobów obliczeniowych?

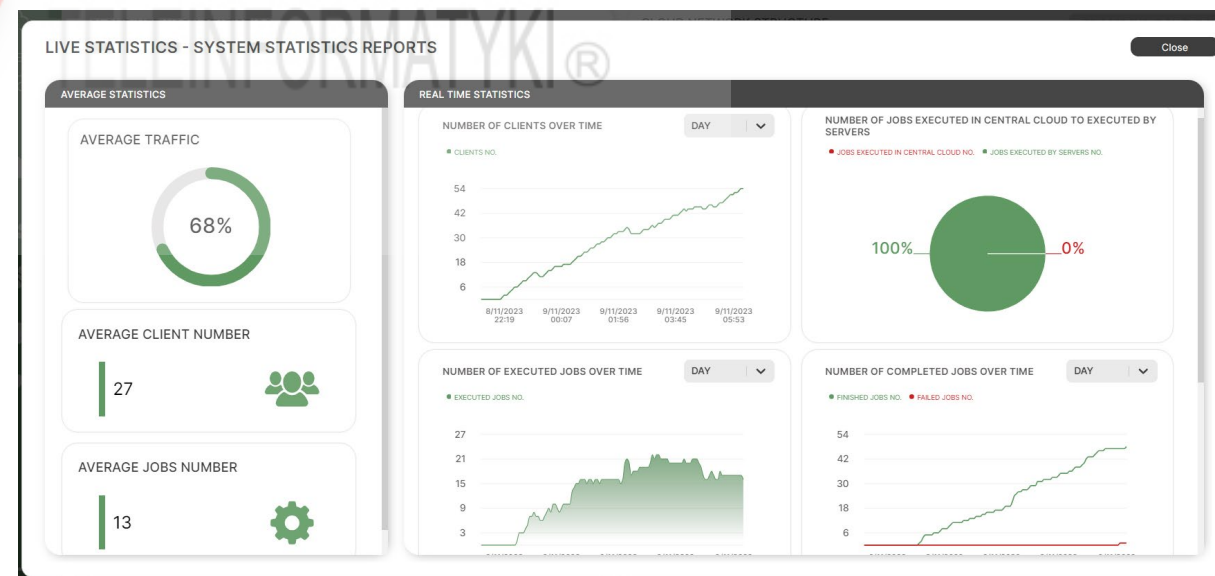
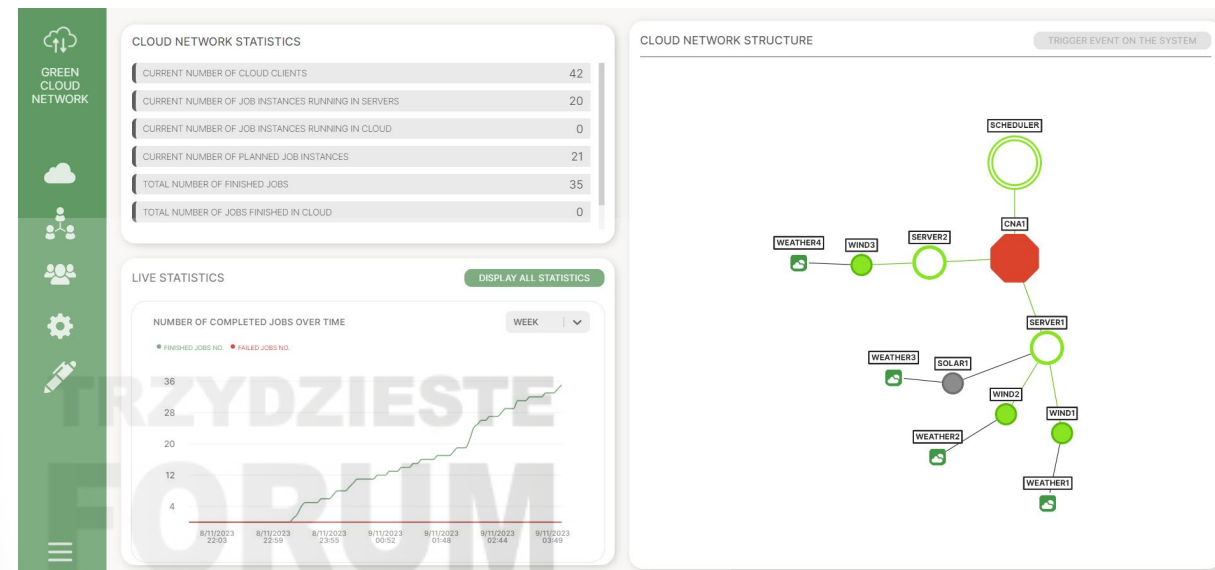
Obecne badania:

- brak algorytmów wykorzystujących **rzeczywiste dane** z chmury obliczeniowej
- niewystarczająca liczba testów przeprowadzonych dla „zielonej” chmury obliczeniowej
- algorytmy oparte na **centralizacji**

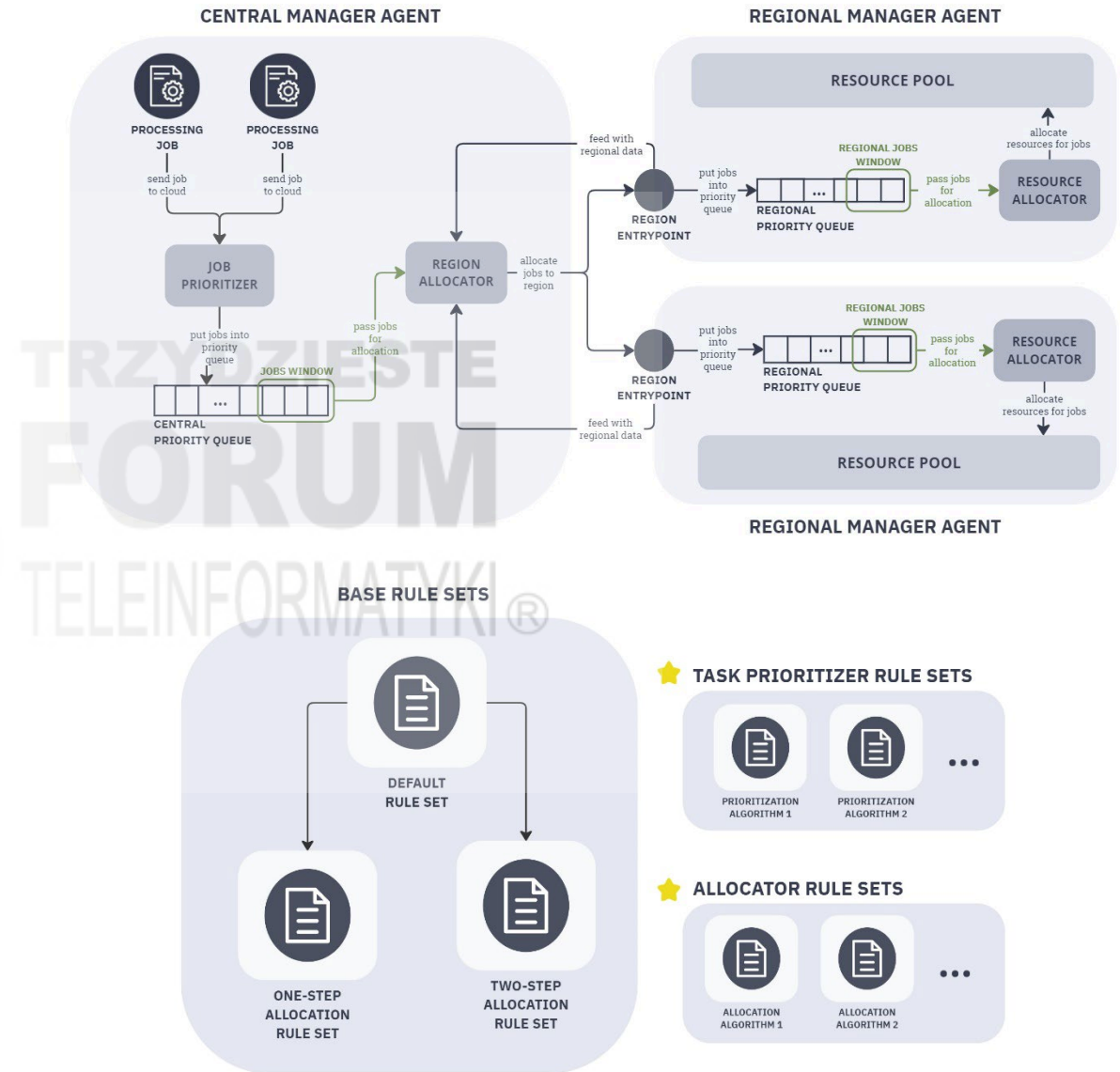
Cele projektu:

1. **Weryfikacja wpływu** jaki ma uwzględnienie wyników analizy danych rzeczywistej chmury obliczeniowej na algorytmy alokacji zasobów
2. **Przetestowanie i porównanie** różnych algorytmów alokacji zasobów w „zielonej” infrastrukturze chmurowej

- **cyfrowy bliźniak** infrastruktury chmurowej opracowana przez CloudFerro (EGC)
- **system wieloagentowy**
- modelowanie strategii z wykorzystaniem:
 - systemu ekspertowego opartego na regułach
 - języków wyrażeń



- oparty na **dwustopniowej alokacji**
- wyodrębnia problemy związane z ustalaniem priorytetów zadań i alokacją zasobów
- umożliwia rozwiązanie problemów związanych z **ograniczeniami udostępniania danych**
- implementacja z wykorzystaniem zestawów reguł → **modularność + adaptowalność**



Algorytmy alokacji:

1. Algorytm oparty na intencjach klientów
(**bazujący na klasteryzacji**)
2. Algorytm oparty na priorytetach zadań
(**bazujący na klasteryzacji**)
3. Algorytm oparty na ograniczeniu kosztów
i terminu wykonania zadań
(**heurystyczny**)
4. Algorytm mrówkowy wykorzystujący czas
wykonania zadań
(**inspirowany naturą**)

Najważniejsze miary jakości:

1. Czas procesu alokacji (AT)
2. Wskaźnik sukcesu wykonania zadań (JSR)
3. Wskaźnik satysfakcji klienta z terminu
wykonania zadania (JDSR)
4. Wskaźnik satysfakcji klienta z kosztu
wykonania zadania (JCSR)
5. Procent wykorzystania zielonej energii
(EU)

- klasteryzacja danych CloudFerro z 2022 i 2024 roku (**51 500 rekordów**)
- wykorzystana metoda: **K-Means**
- pokazanie zależności między typami zadań a zużyciem zasobów
- wnioski:
 - typy zadań uległy zmianie na przestrzeni lat
 - **grd_cog, senzcor** and **card_bs** → niskie zużycie zasobów
 - **download, card_coh, sentinel** → wysokie zużycie zasobów

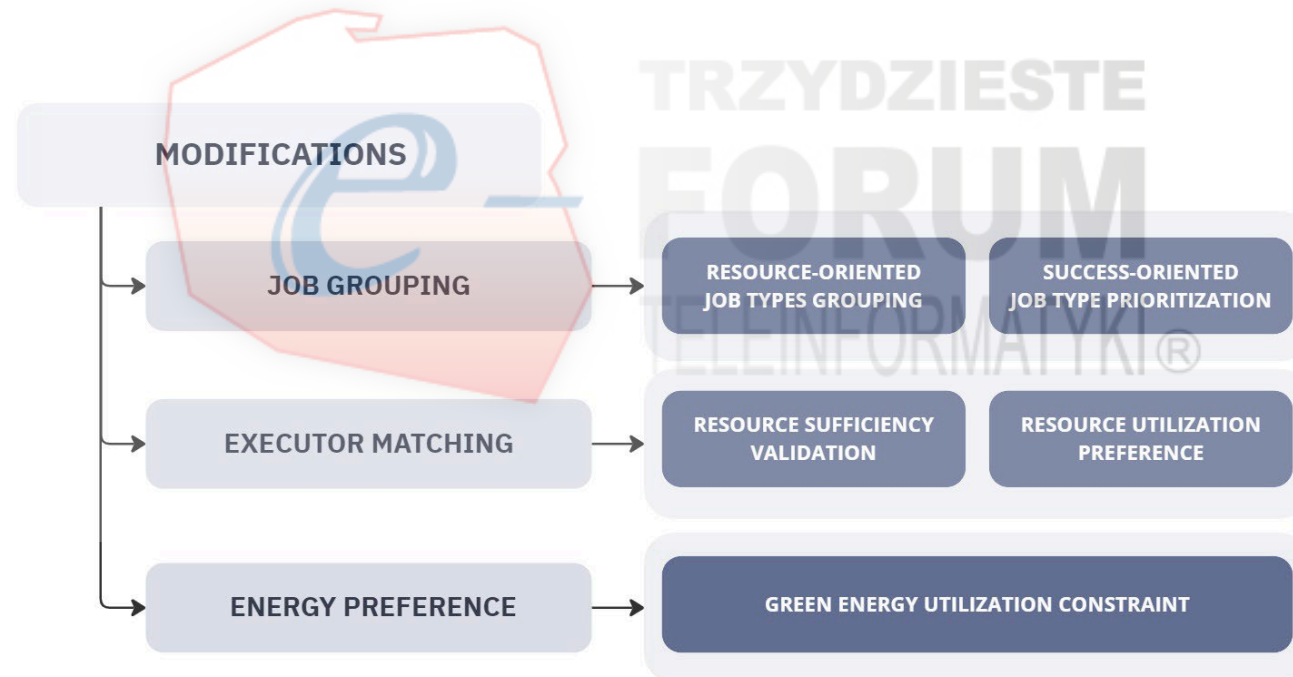
Rezultaty klasteryzacji z 2022

Cluster name	Cluster Size	Avg. CPU	Avg. memory	Avg. duration	Avg. storage	Avg. ephemeral storage	Avg. processed size
PROCESSED WITHOUT ERRORS							
grd_cog	18.5K	1K	30K	13 MIN	5	38	1 BLN
card_bs	1.1K	6K	250K	16 MIN	80	16	5 BLN
senzcor	10.7K	7K	173K	44 MIN	20	1K	0.8 BLN
card_coh	1.6K	91K	11M	1 H	200	52	0.75 BLN
download	6	1M	12MM	6 D	0	0	3 BLN

Rezultaty klasteryzacji z 2024

Name	Size	Δ CPU	Δ Memory	Δ Storage	Δ Duration
grd_cog	19 462	1k (LOW)	18k (LOW)	0 (LOW)	5min (SHORT)
card_coh	59	60k (AVG)	8M (HIGH)	200 (HIGH)	1h (AVG)
card_bs	77	4k (LOW)	260k (HIGH)	80 (AVG)	10min (SHORT)
sentinel	20	170k (HIGH)	55k (AVG)	200 (HIGH)	2days (LONG)
drl_sentinel	2	414k (HIGH)	133k (AVG)	160 (HIGH)	5days (LONG)

- Cel: **Poprawa wydajności alokacji zasobów**
- 5 typów modyfikacji opartych na wynikach analizy i charakterystyce EGC



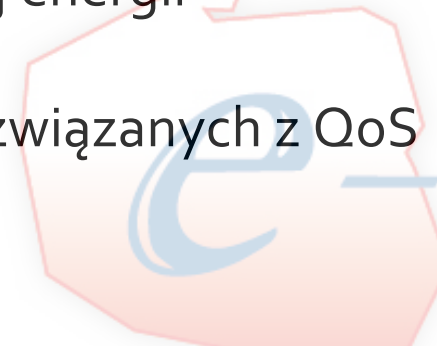
- testy przeprowadzone w EGCS
- oparte na próbkach danych reprezentujących **rzeczywisty rozkład typów danych i rozkład jednolity**
- różne rozmiary próbek danych (**50, 250, 1000**)
- 3 scenariusze testowe:
 - **Scenariusz prosty** (1-stopniowa alokacja)
 - **Scenariusz z ograniczeniem przekazywanych danych** (2-stopniowa alokacja)
 - **Scenariusz zorientowany na odnawialną energię** (1-stopniowa alokacja)

Testowane algorytmy

Algorithm Id.	Allocation Alg.	Prioritization Alg.	Modifications
Intent-Based 1	intent-based	deadline-based	–
Intent-Based 2	intent-based	deadline-based	<i>resource_sufficiency</i>
Intent-Based 3	intent-based	deadline-based	<i>resource_sufficiency</i> <i>green_energy_utilization</i> <i>resource-oriented_grouping</i>
Priority-Based 1	priority-based	duration-based	–
Priority-Based 2	priority-based	duration-based	<i>resource_utilization</i> <i>resource_sufficiency</i>
Priority-Based 3	priority-based	duration-based	<i>green_energy_utilization</i> <i>resource_sufficiency</i> <i>resource_utilization</i> <i>success-oriented_grouping</i>
Budget-Deadline 1	budged-deadline	execution-time-based	–
Budget-Deadline 2	budged-deadline	execution-time-based	<i>resource_utilization</i> <i>resource_sufficiency</i>
Budget-Deadline 3	budged-deadline	execution-time-based	<i>green_energy_utilization</i> <i>resource_sufficiency</i> <i>success-oriented_grouping</i>
ACO-Based 1	aco-based	enhanced-execution-time-based	–
ACO-Based 2	aco-based	enhanced-execution-time-based	<i>green_energy_utilization</i> <i>resource_sufficiency</i>

Po zastosowaniu modyfikacji:

- efektywność czasowa znacząco wzrosła
- nastąpiła poprawa procentu wykorzystania zielonej energii (Scenariusz 3)
- poprawa miar jakości związanych z QoS



Scenariusz 1: rezultaty próbki o rzeczywisty rozkładzie, rozmiar 250 (2022)

Algorithm Id	Δ AAR	Δ ASR	Δ AT	Δ JSR	Δ JDSR	Δ EU	Δ JCSR	Δ CU
Intent-Based 1	100%	94%	9.6s	74%	55%	81%	75%	20%
Intent-Based 2	100%	93%	9s	82%	62%	82%	80%	15%
Intent-Based 3	100%	95%	5s	80%	63%	87%	82%	16%
Priority-Based 1	100%	86%	7.2s	84%	73%	84%	95%	12%
Priority-Based 2	100%	85%	7.8s	87%	76%	90%	96%	12%
Priority-Based 3	100%	91%	3.3s	97%	75%	91%	99%	15%
Budget-Deadline 1	100%	43%	1.9s	60%	72%	98%	82%	24%
Budget-Deadline 2	100%	42%	1.8s	67%	75%	96%	84%	22%
Budget-Deadline 3	100%	42%	1.9s	65%	74%	94%	84%	15%
ACO-Based 1	100%	93%	6.2s	87%	73%	85%	84%	17%
ACO-Based 2	100%	90%	6.4s	90%	76%	89%	85%	15%

Scenariusz 3: rezultaty próbki o rzeczywisty rozkładzie, rozmiar 250 (2022)

Algorithm Id	Δ AAR	Δ ASR	Δ AT	Δ JSR	Δ JDSR	Δ EU	Δ JCSR	Δ CU
Intent-Based 1	100%	78%	10.1s	80%	60%	79%	63%	26%
Intent-Based 3	100%	91%	5.2s	85%	67%	88%	71%	14%
Priority-Based 1	100%	82%	8.1s	66%	68%	89%	71%	19%
Priority-Based 3	100%	88%	3.6s	72%	74%	90%	76%	11%
Budget-Deadline 1	100%	41%	6.7s	55%	70%	91%	68%	22%
Budget-Deadline 3	100%	46%	2.3s	60%	75%	100%	72%	18%
ACO-Based 1	100%	90%	9.1s	77%	71%	81%	74%	21%
ACO-Based 2	100%	95%	8s	82%	76%	92%	73%	19%

Scenariusz 2: rezultaty próbki o rzeczywisty rozkładzie, rozmiar 250 (2022)

Algorithm Id	Δ AAR	Δ ASR	Δ AT	Δ JSR	Δ JDSR	Δ EU	Δ JCSR	Δ CU
Intent-Based 1	100%	86%	15.6s	73%	58%	82%	74%	28%
Intent-Based 2	100%	85%	15.2s	75%	62%	87%	79%	19%
Intent-Based 3	100%	84%	9s	78%	62%	87%	90%	18%
Priority-Based 1	100%	88%	9.3s	82%	74%	84%	85%	25%
Priority-Based 2	100%	85%	9.7s	83%	77%	90%	88%	15%
Priority-Based 3	100%	90%	5.3s	86%	78%	91%	95%	14%


Dziękuję za uwagę!

TRZYDZIESTE
FORUM
TELEINFORMATYKI®