

Zastosowanie chmur obliczeniowych dla aplikacji naukowych

Maciej Malawski

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie
Wydział Informatyki, Elektroniki i Telekomunikacji
Katedra Informatyki

Seminarium Katedry Inżynierii Oprogramowania (K7/W8)
Politechniki Wrocławskiej
i Sekcji Inżynierii Oprogramowania Komitetu Informatyki PAN
5.04.2017



Wykształcenie i zatrudnienie

» Wykształcenie

- 2009: stopień doktora nauk technicznych w dyscyplinie: Informatyka, Wydział Elektrotechniki, Automatyki, Informatyki i Elektroniki (WEAliE), AGH
tytuł: „Component-based Methodology for Programming and Running Scientific Applications on the Grid”, promotor: prof. dr hab. inż. Jacek Kitowski
- 2004: tytuł magistra fizyki, specjalność: fizyka teoretyczna, Wydział Fizyki, Astronomii i Informatyki Stosowanej Uniwersytet Jagielloński, Kraków
- 2001: tytuł magistra inżyniera informatyka, specjalność: systemy komputerowe; WEAliE, AGH

» Zatrudnienie

- 2001 – 2009: asystent, Katedra Informatyki, WEAiE, AGH
- 2011 – 2012: post-doc, Center for Research Computing, University of Notre Dame, Indiana, USA
- 2009 – obecnie: adiunkt, Katedra Informatyki, WIEiT, AGH
- 2002 – obecnie: współpraca z ACK Cyfronet AGH w zakresie realizacji projektów

Zainteresowania naukowe

- » Programowanie równoległe i rozproszone
- » Modele programowania: komponentowe, zorientowane na usługi, systemy workflow
- » Systemy obliczeniowe wielkiej skali typu klastry, grid i chmury
- » Zarządzanie zasobami (obliczeniowymi, magazynowymi)
 - Optymalizacja przydziału zasobów
 - Wydajność i skalowalność aplikacji
- » Narzędzia wspierające aplikacje naukowe
 - Obliczenia dużej mocy (HPC)
 - Obliczenia dużej przepustowości (HTC)
 - Zastosowania w fizyce, biomedycynie



Wskazanie osiągnięcia

- » Tytuł osiągnięcia naukowego

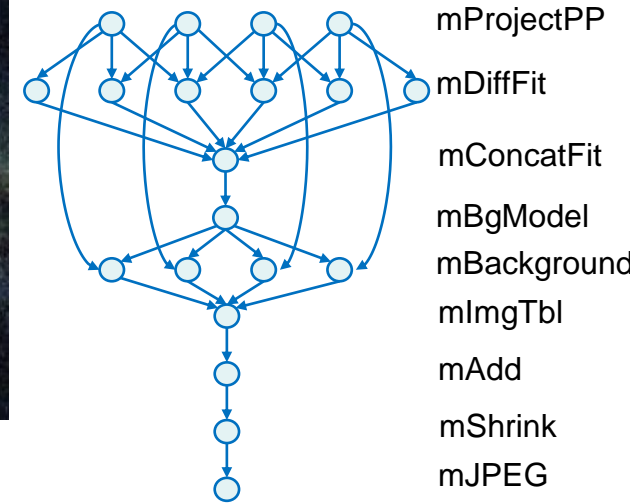
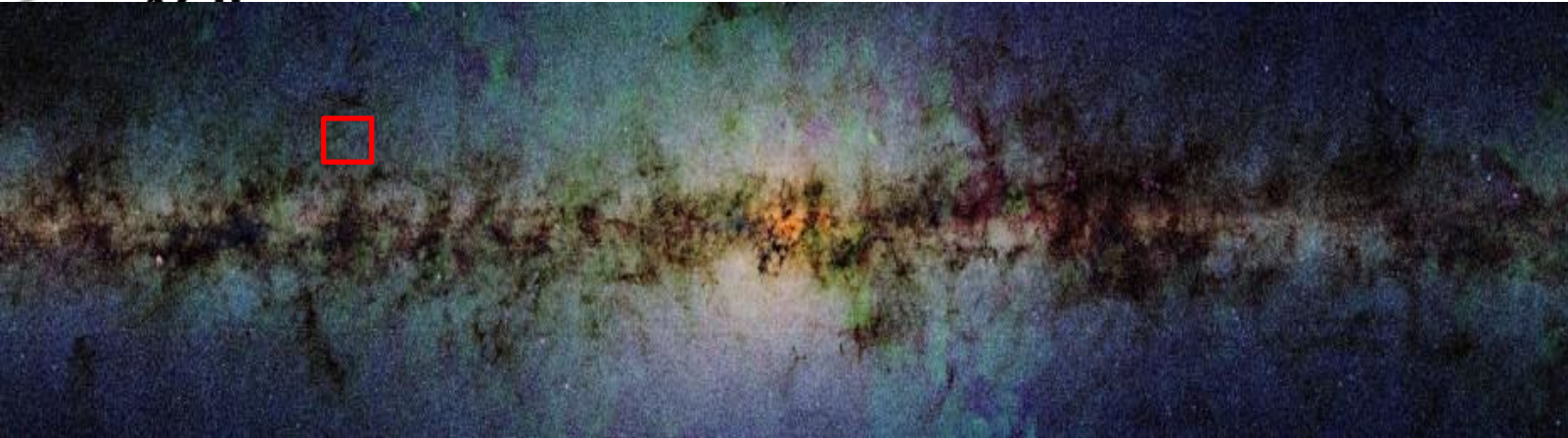
Zarządzanie zasobami dla aplikacji naukowych w chmurach obliczeniowych

- » Cykl publikacji powiązanych tematycznie dokumentujący osiągnięcie zawiera 13 pozycji opublikowanych w latach 2011-2016,
- » Prace te są opublikowane:
 - w czasopismach posiadających wskaźnik wpływu IF (wg JCR), 6 pozycji,
 - w materiałach konferencji międzynarodowych
 - indeksowanych przez Web of Science, 6 pozycji
 - open access, 1 pozycja

Motywacja

- » Złożone aplikacje naukowe: komponentowe, bag-of-tasks, grafy zadań (*workflow*) i ich zbiory (*ensemble*)
- » Nowe rodzaje infrastruktury: chmury obliczeniowe, publiczne, prywatne, hybrydowe, *Infrastructure-as-a-service (IaaS)*, *Platform-as-a-service (PaaS)*, *Function-as-a-Service (FaaS)*:
zasoby zwirtualizowane, dostarczane na żądanie (*on-demand*), rozliczane za użycie (*pay-per-use*)
- » Problemy badawcze:
 - Minimalizacja kosztów monetarnych wykonania aplikacji przy ograniczeniu czasowym
 - Maksymalizacja wykonanej pracy przy ograniczeniach czasowych i budżetowych

Workflow – grafy zadań



» Montage Galactic Plane Workflow

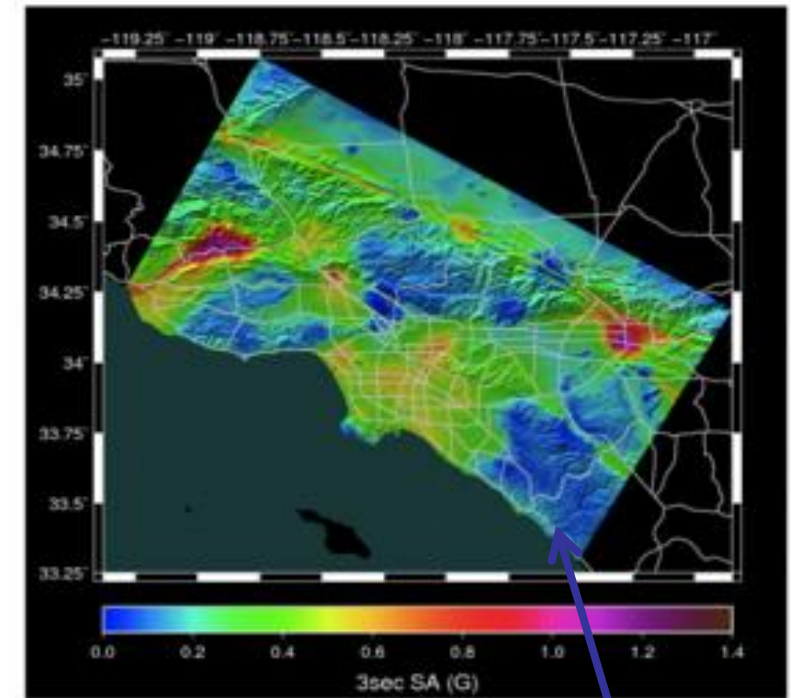
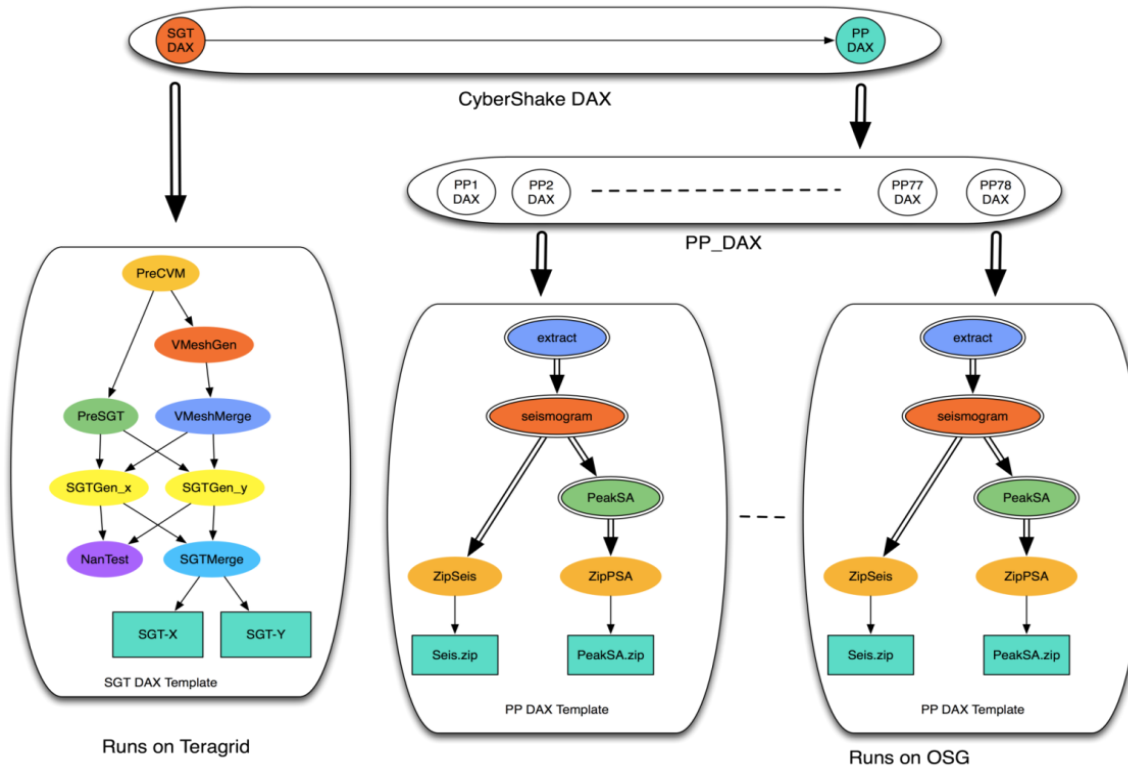
- 18 million input images (~2.5 TB)
- 900 output images (2.5 GB each, 2.4 TB total)
- 10.5 million tasks (34,000 CPU hours)

} × 17

- » Workflow = graf zadań i ich zależności, często graf acykliczny skierowany (directed acyclic graph - DAG)
- » Aplikacja często wymaga uruchomienia wielu takich grafów – zespół grafów (*workflow ensemble*)



Przykład – CyberShake (University of Southern California)



- Obliczenie mapy zagrożeń trzęsieniem ziemi
- ~415,000 scenariuszy
~ 850,000 zadań



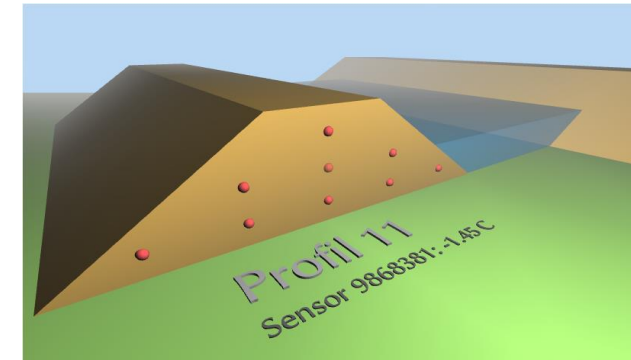


Aplikacja typu Bag-of-tasks

» ISMOP: informatyczny system monitorowania obwałowań przeciwpowodziowych

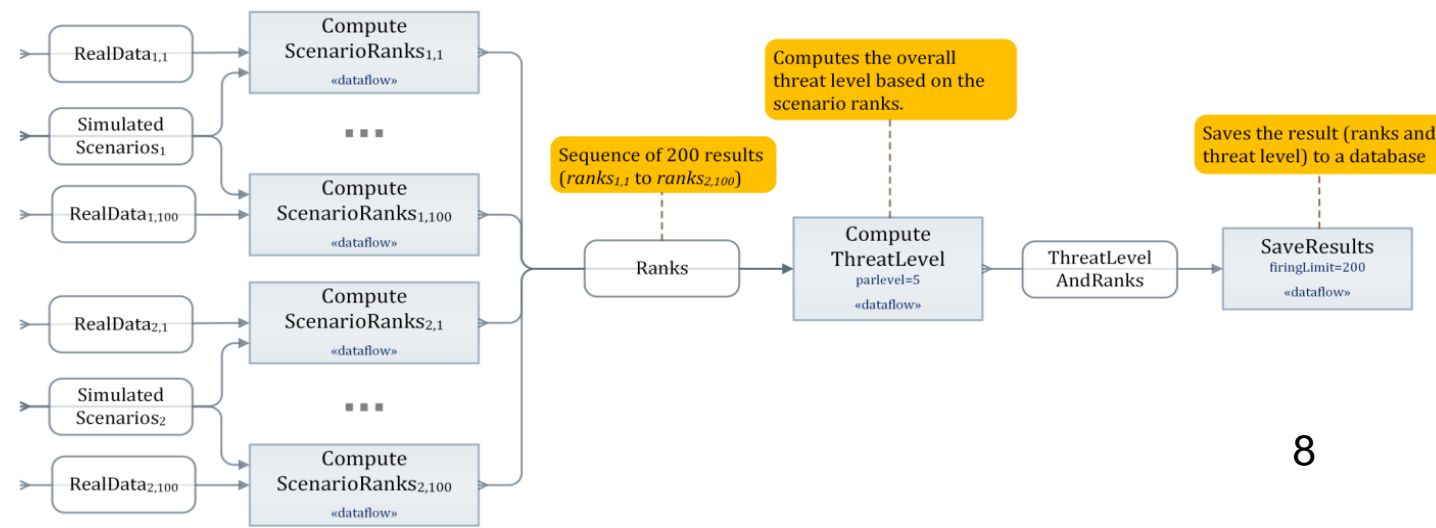
» Obliczanie ryzyka powodzi:

- Graf zadań
- Porównanie bieżących odczytów sensorów z wyliczonymi wcześniej scenariuszami
- Cel: identyfikacja scenariusza w przyszłości
- Oszacowanie ryzyka przerwania wału



» Wymagania

- Deadline
- Docelowa duża długość wału (50+ km)



Infrastruktura – od klastrów do chmur

» Tradycyjne klastry HPC (KDM) w centrach obliczeniowych

- Systemy kolejkowe
- Równoległe systemy przechowywania danych



» Chmury

- Infrastructure as a service (IaaS)
- Globalnie rozproszone
- Maszyny wirtualne (VMs)
- Na żądanie (On-demand)
- Koszt w \$ na jednostkę czasu





Cennik Amazon EC2 45 typów instancji

	vCPU	ECU	Memory (GiB)	Instance Storage (GB)	Linux/UNIX Usage
General Purpose - Current Generation					
t2.nano	1	Variable	0.5	EBS Only	\$0.0065 per Hour
t2.micro	1	Variable	1	EBS Only	\$0.013 per Hour
t2.small	2	3	2	EBS Only	\$0.026 per Hour
t2.medium	4	6	4	EBS Only	\$0.052 per Hour
t2.large	8	12	8	EBS Only	\$0.104 per Hour
m4.large	4	16	17	EBS Only	\$0.105 per Hour
m4.xlarge	8	32	34	EBS Only	\$0.209 per Hour
m4.2xlarge	16	64	68	EBS Only	\$0.419 per Hour
m4.4xlarge	32	128	136	EBS Only	\$0.838 per Hour
m4.10xlarge	80	320	340	EBS Only	\$2.095 per Hour
m4.16xlarge	128	512	512	EBS Only	\$3.352 per Hour
Compute Optimized - Current Generation					
c4.large	2	8	3.75	EBS Only	\$0.105 per Hour
c4.xlarge	4	16	7.5	EBS Only	\$0.209 per Hour
c4.2xlarge	8	32	15	EBS Only	\$0.419 per Hour
c4.4xlarge	16	64	30	EBS Only	\$0.838 per Hour
c4.8xlarge	32	128	60	EBS Only	\$1.676 per Hour
Storage Optimized - Current Generation					
i2.xlarge	4	14	30.5	1 x 800 SSD	\$0.853 per Hour
i2.2xlarge	8	27	61	2 x 800 SSD	\$1.705 per Hour
i2.4xlarge	16	53	122	4 x 800 SSD	\$3.41 per Hour
i2.8xlarge	32	104	244	8 x 800 SSD	\$6.82 per Hour
d2.xlarge	4	14	30.5	3 x 2000 HDD	\$0.69 per Hour
d2.2xlarge	8	28	61	6 x 2000 HDD	\$1.38 per Hour
d2.4xlarge	16	56	122	12 x 2000 HDD	\$2.76 per Hour
d2.8xlarge	36	116	244	24 x 2000 HDD	\$5.52 per Hour

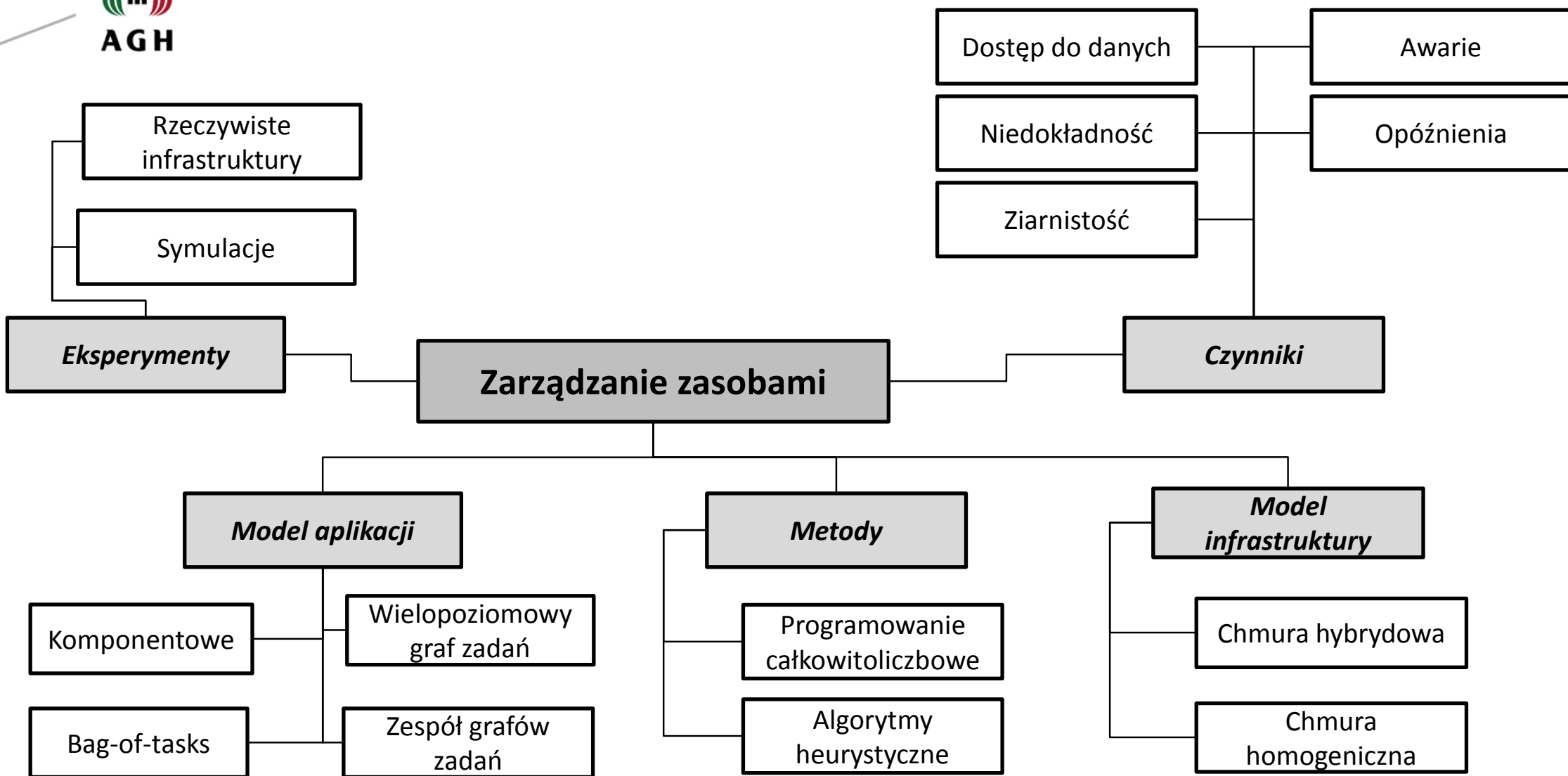
850 x (!!!)

10

Cele naukowe

- » Opracowanie metod zarządzania zasobami dla aplikacji naukowych w chmurach, z uwzględnieniem problemów
 - dostarczania zasobów (*resource provisioning*),
 - szeregowania zadań (*workflow scheduling*),traktowanych jako problemy optymalizacyjne z ograniczeniami.
- » Cele szczegółowe:
 - Eksperymentalne zbadanie przydatności chmur obliczeniowych dla aplikacji naukowych w celu poznania ich charakterystyk kosztowo-wydajnościowych.
 - Stworzenie modeli infrastruktury i aplikacji dla wybranych klas aplikacji naukowych w celu rozwoju metod zarządzania zasobami.
 - Ocena efektywności i wydajności zaproponowanych metod przez symulację i eksperymenty.

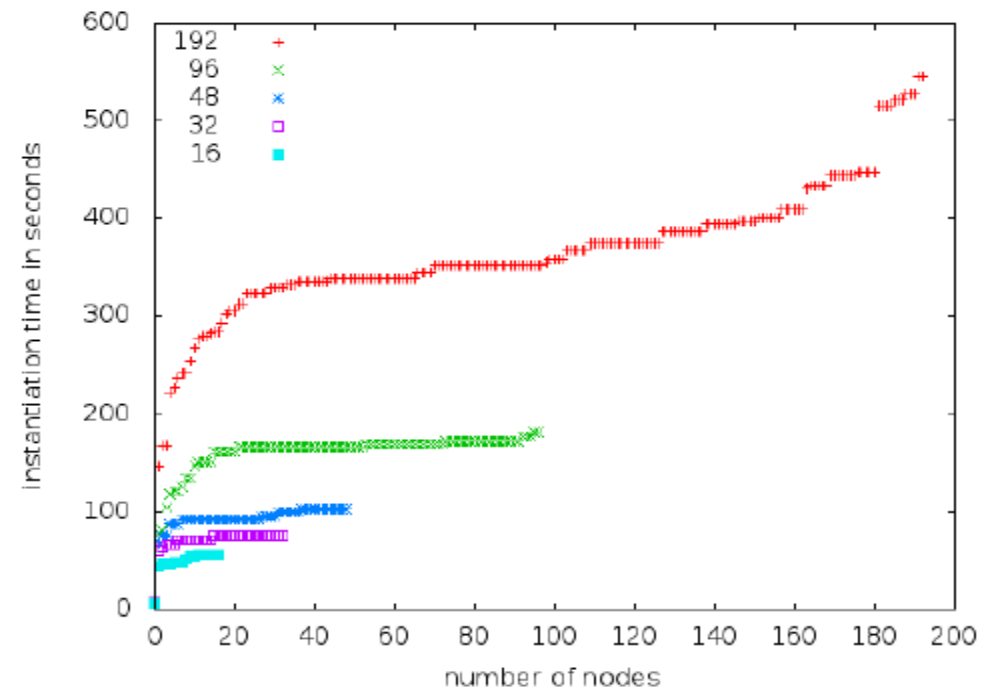
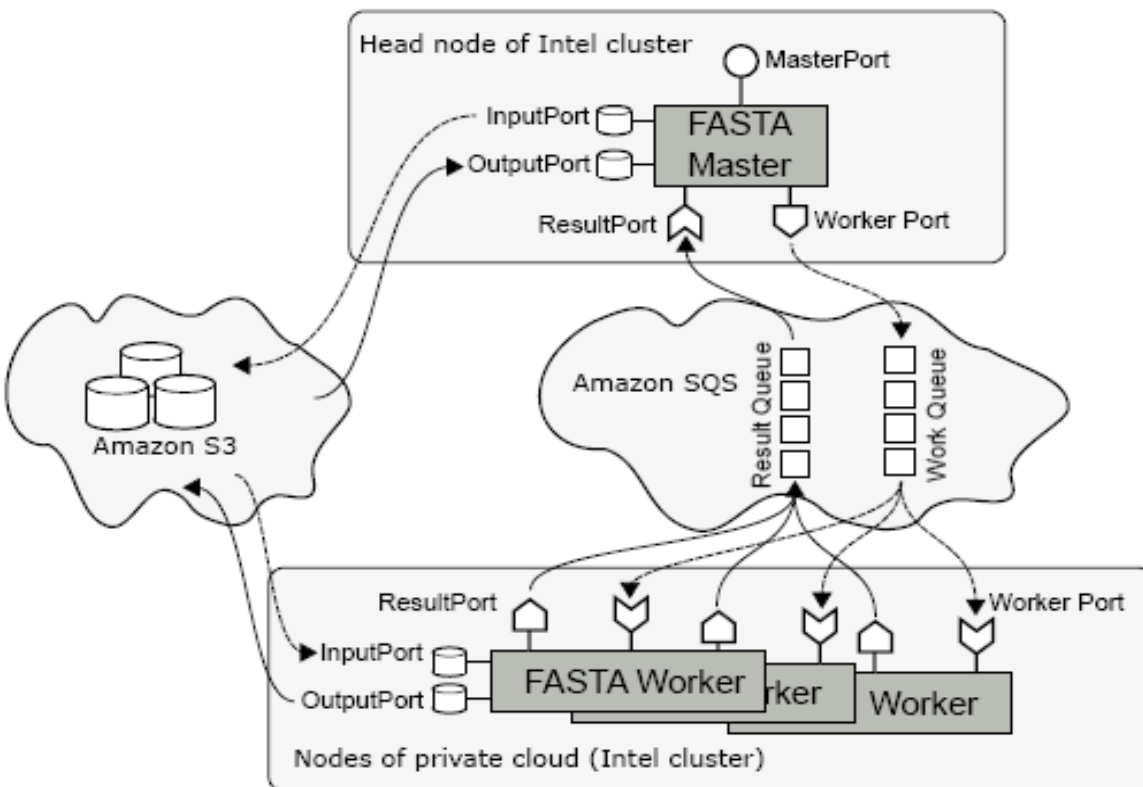
Zakres tematyczny prac



Osiągnięcie habilitacyjne

1/5

- » Opracowanie modelu komponentowego do obliczeń naukowych w chmurach [H12, H13], w tym hybrydowej chmury (Amazon EC2, Eucalyptus)



Osiągnięcia habilitacyjne 1/5 - wyniki

- » Potwierdzenie efektywności i wydajności takiego podejścia przy pomocy eksperymentów z użyciem przykładowej aplikacji bioinformatycznej
- » Znaczenie wyników:
 - uzyskanie wiedzy oraz danych użytecznych do optymalizacji kosztów czasowych i monetarnych użycia chmury.
 - doświadczenie w technologiach chmur wykorzystane przy opracowaniu platformy Atmosphere dla projektu EU FP7 VPH-Share

Publikacje 1/5

[H12] **M. Malawski**, T. Gubala, and M. Bubak: “*Component-based approach for programming and running scientific applications on grids and clouds*”, International Journal of High Performance Computing Applications, vol. 26, no. 3, pp. 275–295, 2012

IF: 1,295, CI_{WoS} : 5, CI_{GS} : 17

[H13] **M. Malawski**, J. Meizner, M. Bubak, and P. Gepner: “*Component approach to computational applications on clouds*”, International Conference on Computational Science, ICCS, Procedia Computer Science, vol. 4, pp. 432–441, 2011

CI_{WoS} : 6, CI_{GS} : 32

Legenda:

IF = Impact Factor czasopisma, brak oznacza konferencję

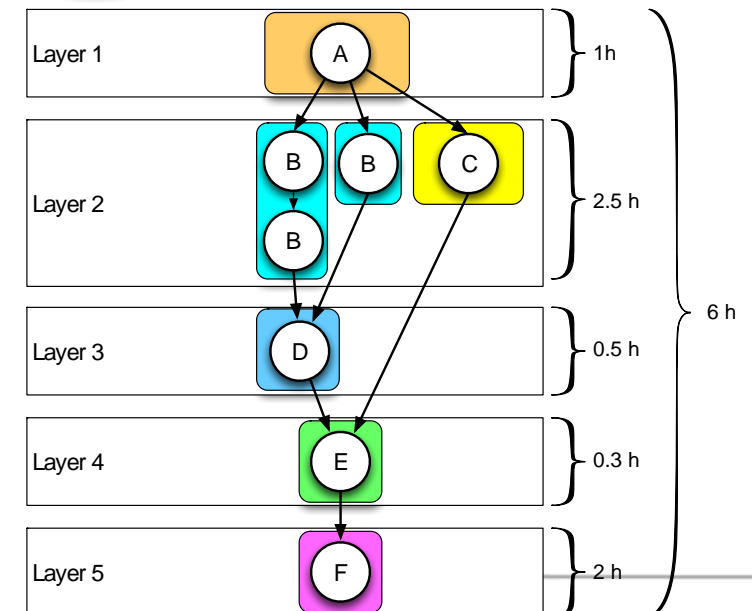
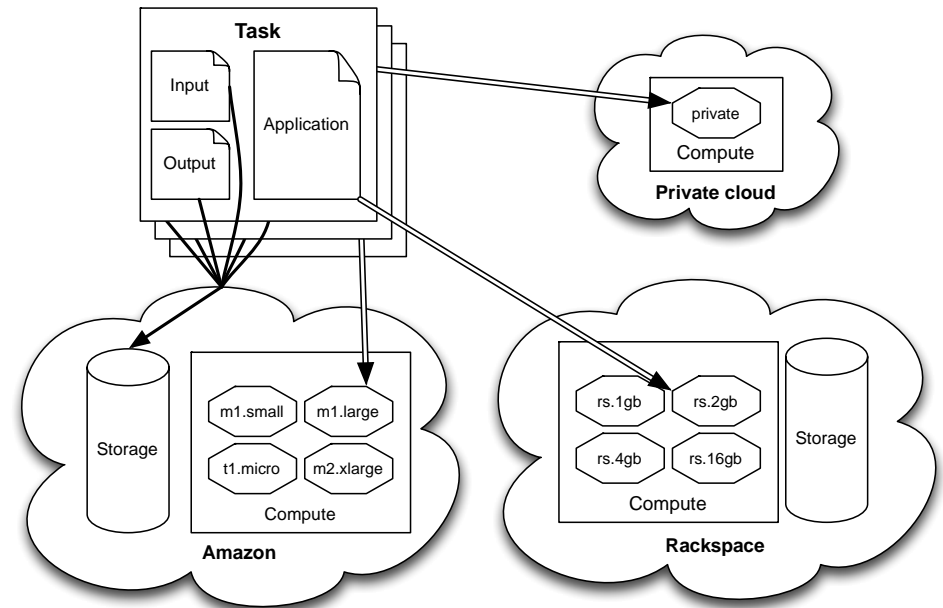
CI_{WoS} = liczba cytowań w WoS (bez autocytowań)

CI_{GS} = liczba cytowań w Google Scholar

Osiągnięcie habilitacyjne 2/5



- » Opracowanie modeli infrastruktur dla hybrydowych chmur obliczeniowych
 - modelu heterogenicznej chmury
 - modelu homogenicznego z uwzględnieniem dostępu do danych
- » Opracowanie modeli aplikacji dla aplikacji naukowych
 - aplikacji typu bag-of-tasks,
 - wielowarstwowych grafów zadań,
 - zespołów grafów,
 - modeli wydajności dla konkretnych przykładów aplikacji

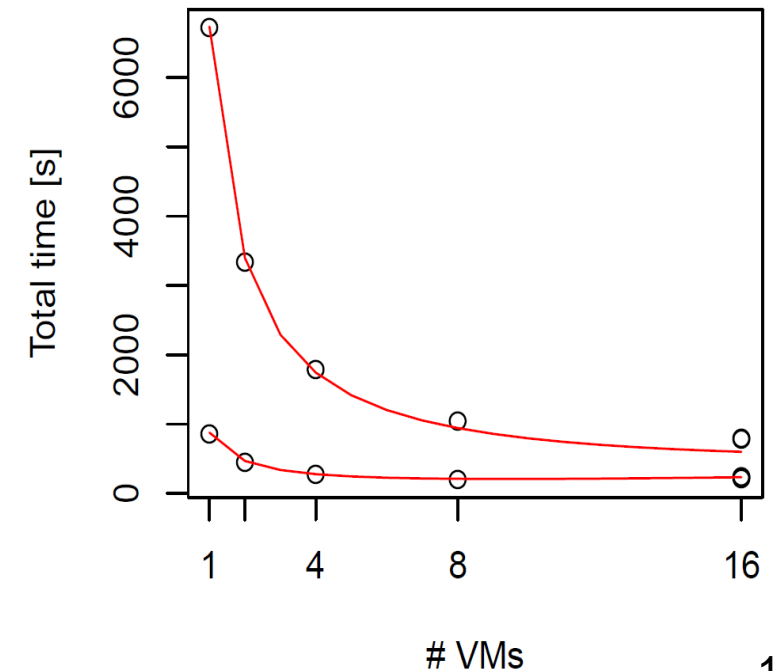
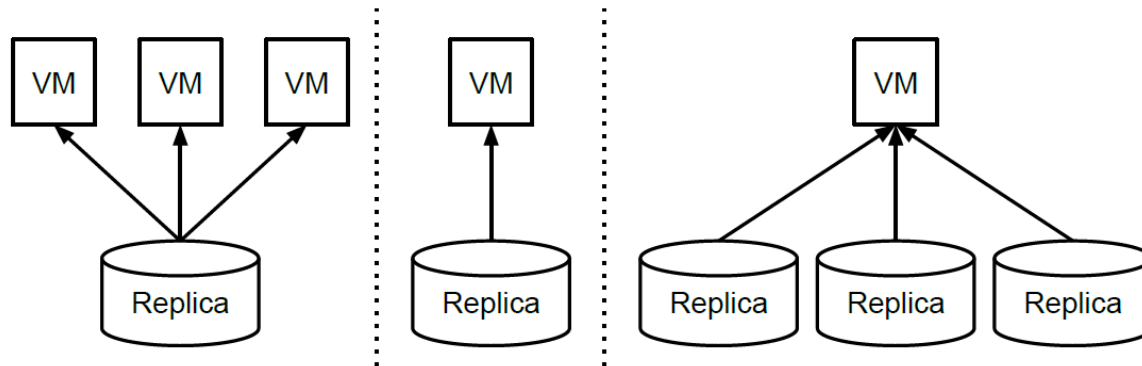


Osiągnięcie habilitacyjne

2/5 - wyniki

» Znaczenie wyników:

- Podstawa dla prac nad nowymi metodami zarządzania zasobami i optymalizacji
- Zaproponowanie nowego modelu dostępu do danych zgodnego z architekturą chmur
- Zastosowanie modelu wydajności dla aplikacji wyliczającej poziom zagrożenia powodziowego

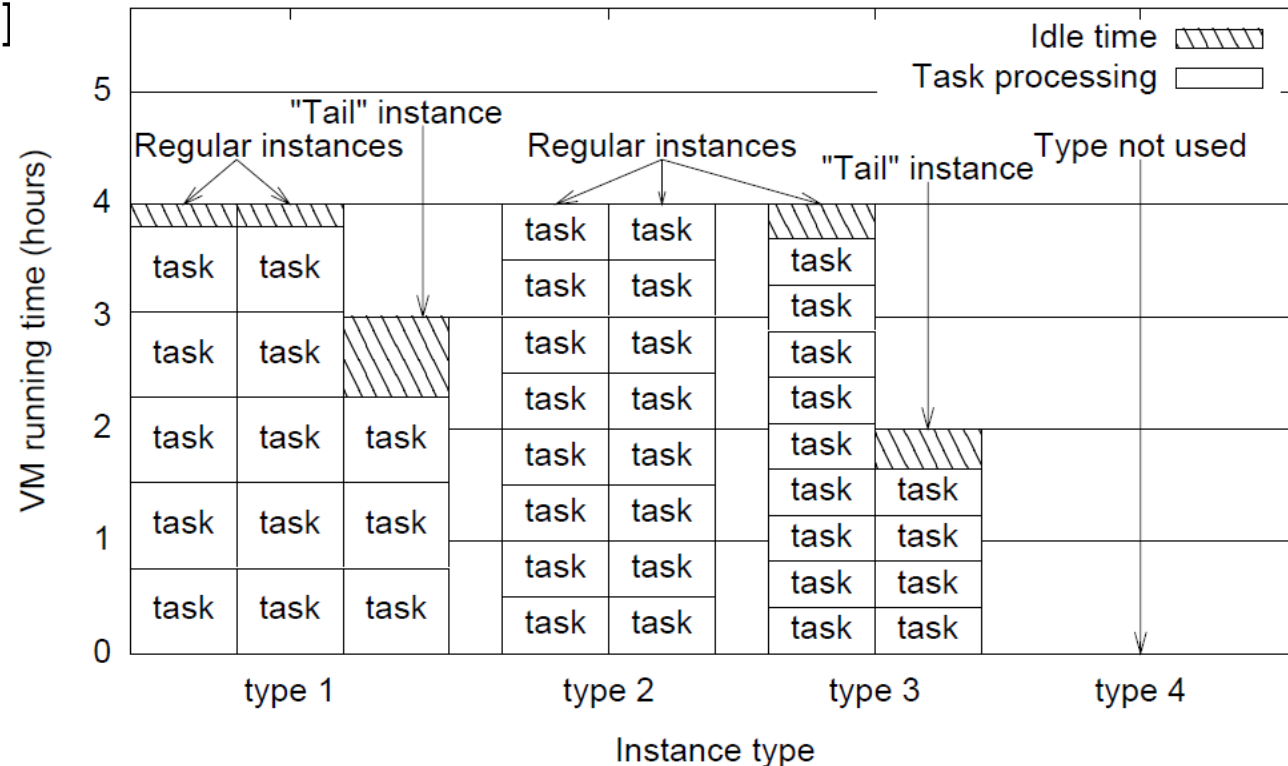


- [H2] P. Bryk, **M. Malawski**, G. Juve, and E. Deelman: “*Storage-aware Algorithms for Scheduling of Workflow Ensembles in Clouds*”, Journal of Grid Computing, vol. 14, no. 2, pp. 359–378, 2016
IF: 1,507, CI_{WoS} : 0, CI_{GS} : 2
- [H4] B. Balis, M. Kasztelnik, **M. Malawski**, P. Nowakowski, B. Wilk, M. Pawlik, and M. Bubak: “*Execution Management and Efficient Resource Provisioning for Flood Decision Support*”, International Conference on Computational Science, ICCS, Procedia Computer Science, vol. 51, pp. 2377–2386, Elsevier, 2015
CI_{WoS} : 1, CI_{GS} : 3
- [H7] **M. Malawski**, K. Figiela, M. Bubak, E. Deelman, and J. Nabrzyski: “*Cost optimization of execution of multi-level deadline-constrained scientific workflows on clouds*”, Parallel Processing and Applied Mathematics - 10th International Conference, PPAM 2013, Warsaw, Poland, September 8-11, 2013, Revised Selected Papers, Part I, Lecture Notes in Computer Science, vol. 8384, pp. 251-260, Springer, 2014
CI_{WoS} : 1, CI_{GS} : 10
- [H8] **M. Malawski**, K. Figiela, and J. Nabrzyski: “*Cost Minimization for Computational Applications on Hybrid Cloud Infrastructures*”, Future Generation Computer Systems vol. 29 no. 7, pp. 1786-1794, 2013
IF: 2,639, CI_{WoS} : 1, CI_{GS} : 24
- [H11] **M. Malawski**, G. Juve, E. Deelman, and J. Nabrzyski: “*Cost- and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds*”, SC '12: Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, pp. 1-10, IEEE/ACM 2012
CI_{WoS} : 14, CI_{GS} : 131

Osiągnięcie habilitacyjne

3/5

- » Opracowanie modeli minimalizujących koszt wykonania aplikacji z użyciem programowania całkowitoliczbowego
 - w chmurze heterogenicznej przy ograniczeniu czasowym,
 - dla aplikacji bag-of-tasks [H8],
 - wielopoziomowych grafów [H6],
 - modelu adaptacyjnego [H3]
- » Zmienne:
 - Typ, liczba maszyn (instancji)
 - Przypisanie zadań do maszyn
 - Podział na poziomy (sub-deadline)
- » Ograniczenia:
 - Czas
 - Limity dostawców

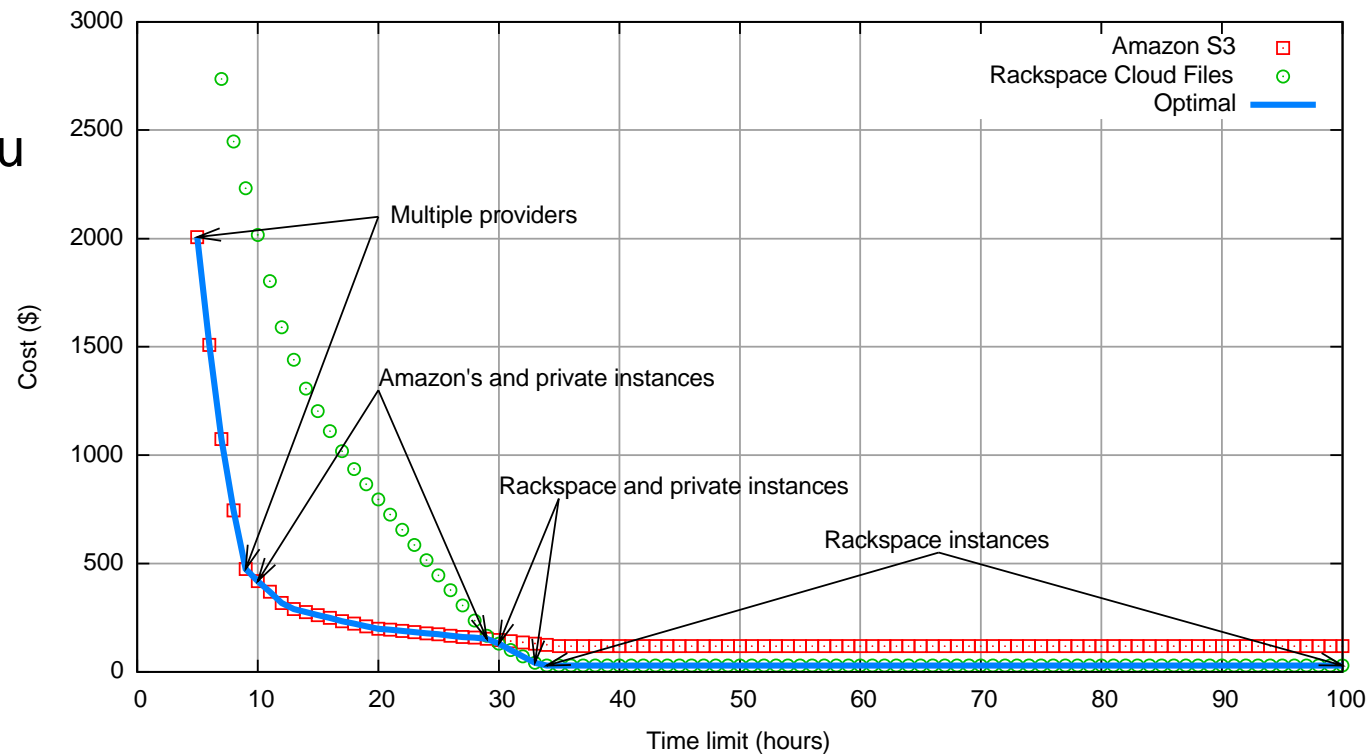


Osiągnięcie habilitacyjne

3/5 - wyniki

20000 tasks, 512 MiB input and 512 MiB output, task execution time 0.1h @ 1ccu machine

Zależność kosztu od czasu dla chmury hybrydowej



» Znaczenie wyników

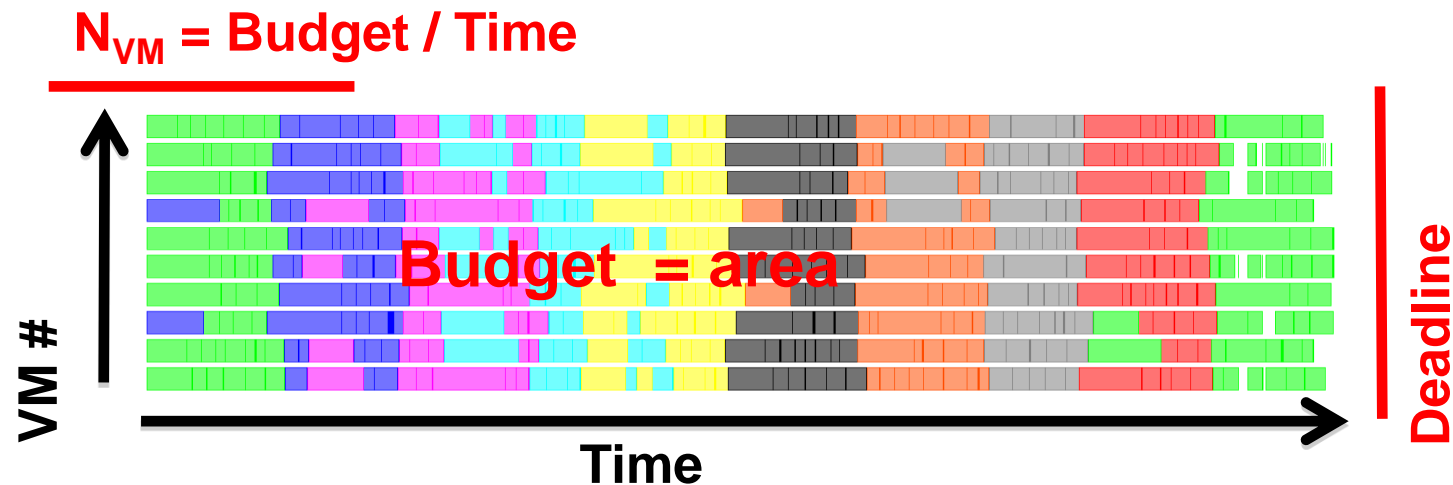
- Pokazanie w jaki sposób konstruować efektywne modele optymalizacyjne korzystając z własności grafów zadań (podział na poziomy, optymalizacja lokalna i globalna)

- [H3] T. Dziok, K. Figiela, and **M. Malawski**: “*Adaptive multi-level workflow scheduling with uncertain task estimates*”, Parallel Processing and Applied Mathematics - 11th International Conference, PPAM 2015, Krakow, Poland, September 6-9, 2015. Revised Selected Papers, Part II, Lecture Notes in Computer Science, vol. 9574, pp. 90–100, Springer, 2016
CI_{WOS} : 0, CI_{GS} : 0
- [H6] **M. Malawski**, K. Figiela, M. Bubak, E. Deelman, and J. Nabrzyski: “*Scheduling multilevel deadline-constrained scientific workflows on clouds based on cost optimization*”, Scientific Programming, vol. 2015, 2015
IF: 0,559, CI_{WOS} : 0, CI_{GS} : 10
- [H8] **M. Malawski**, K. Figiela, and J. Nabrzyski: “*Cost Minimization for Computational Applications on Hybrid Cloud Infrastructures*”, Future Generation Computer Systems vol. 29 no. 7, pp. 1786-1794, 2013
IF: 2,639, CI_{WOS} : 1, CI_{GS} : 24

Osiągnięcie habilitacyjne

4/5

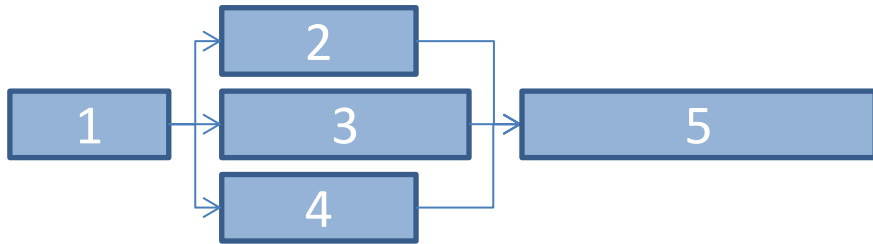
- » Opracowanie nowych algorytmów dotyczących dostarczania zasobów i szeregowania zadań
 - dla aplikacji typu zbiory grafów zadań, przy ograniczeniach czasowych i kosztowych
 - dynamiczne algorytmy heurystyczne:
 - przydziału zasobów i szeregowania [H5, H11],
 - uwzględniające przechowywanie danych i ich lokalność [H2].



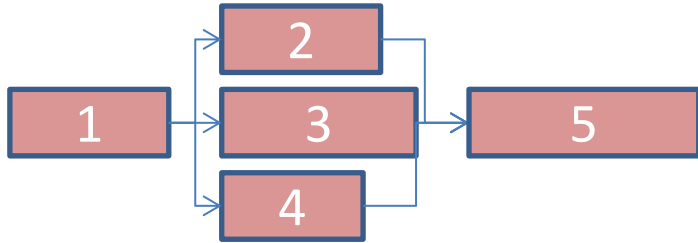
Heurystyki - przykład

Priority:

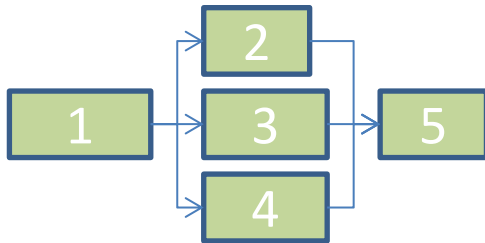
1



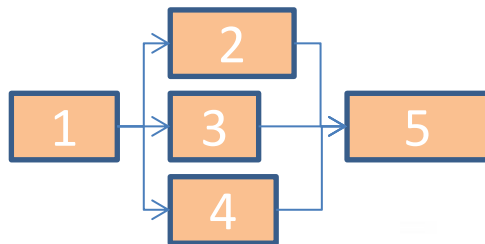
2



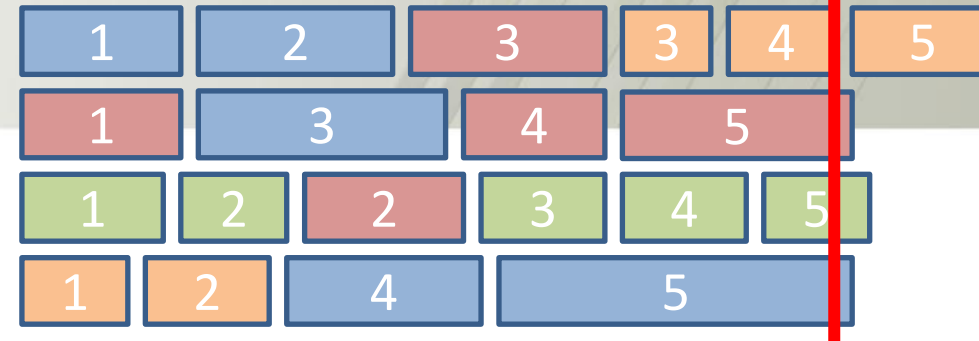
3



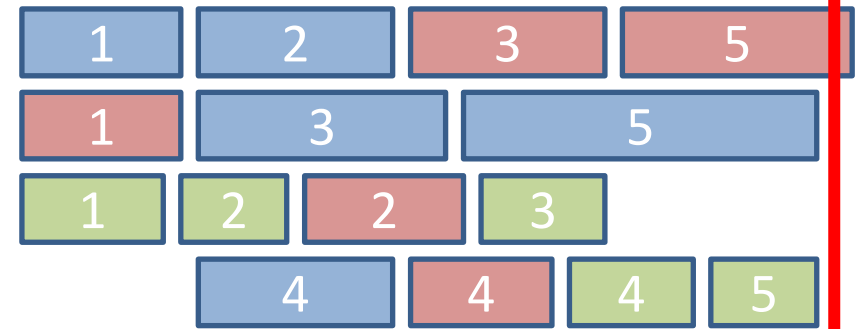
4



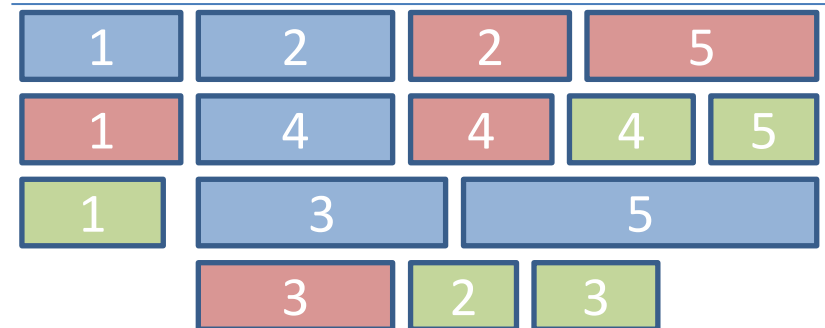
DPDS



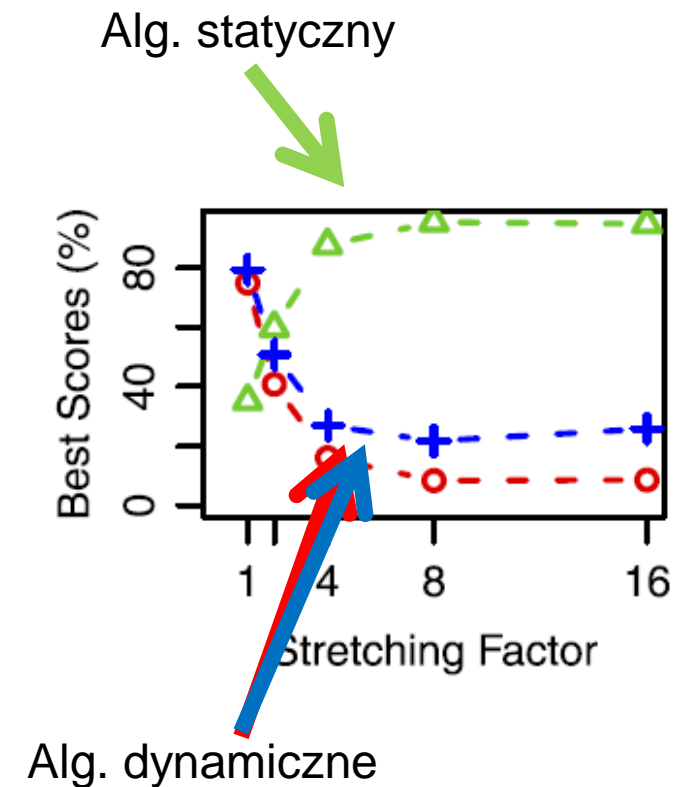
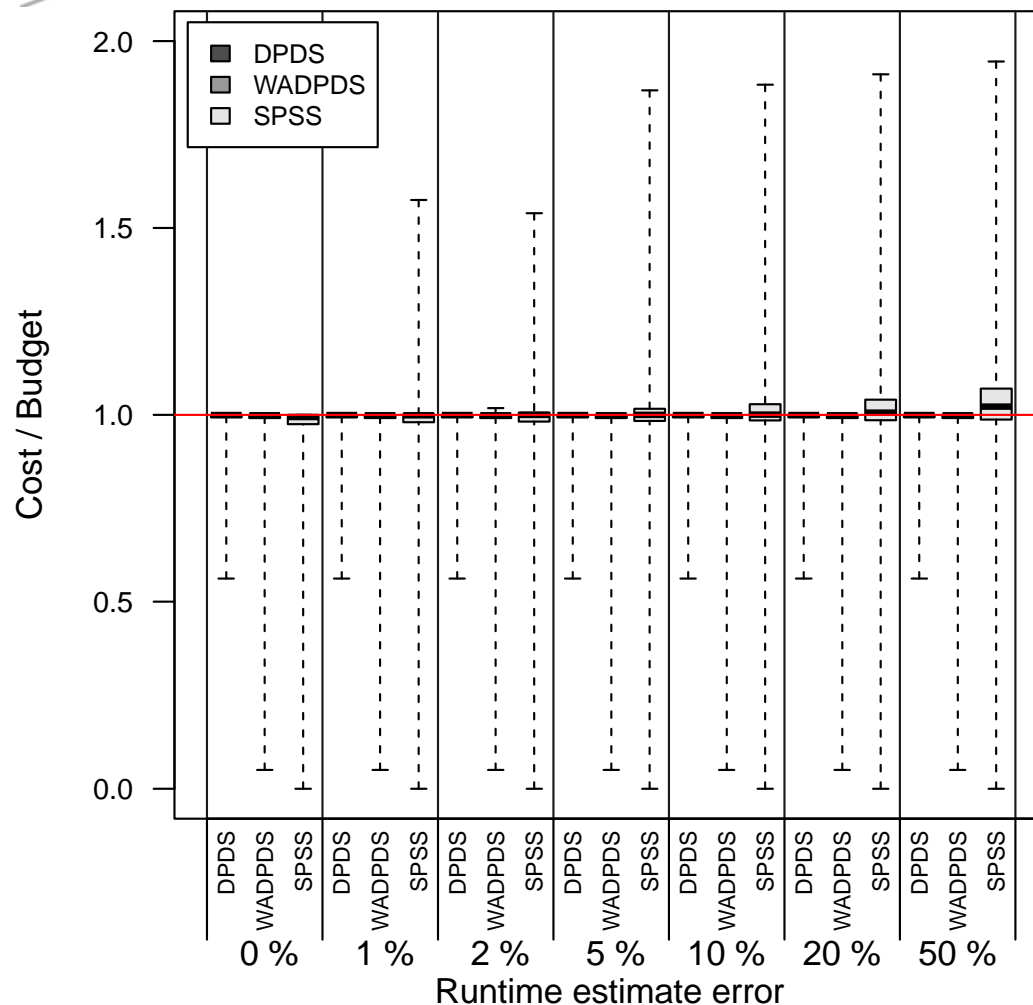
WA-DPDS



SPSS



Osiągnięcie habilitacyjne 4/5 - wyniki



» Znaczenie wyników

- Pokazanie zalet dynamicznych algorytmów korzystających z informacji o strukturze grafu i dostępie do danych w środowisku chmury



Publikacje 4/5

[H5] **M. Malawski**, G. Juve, E. Deelman, and J. Nabrzyski: “*Algorithms for Cost- and Deadline-Constrained Provisioning for Scientific Workflow Ensembles in IaaS Clouds*”, *Future Generation Computer Systems*, vol. 48, pp. 1-18, 2015
IF: 2,786, CI_{WoS} : 8, CI_{GS} : 19

Także w [H2], [H7], [H11]

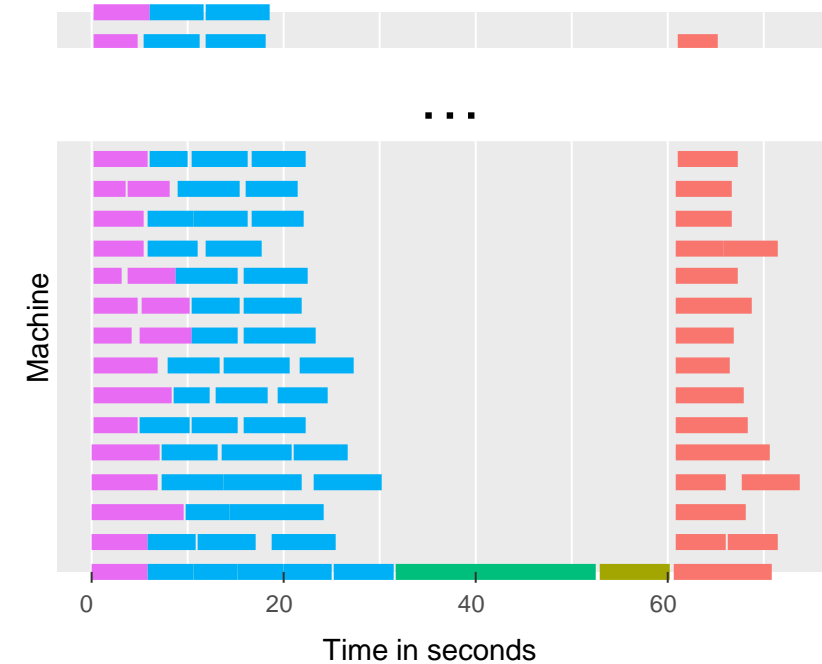
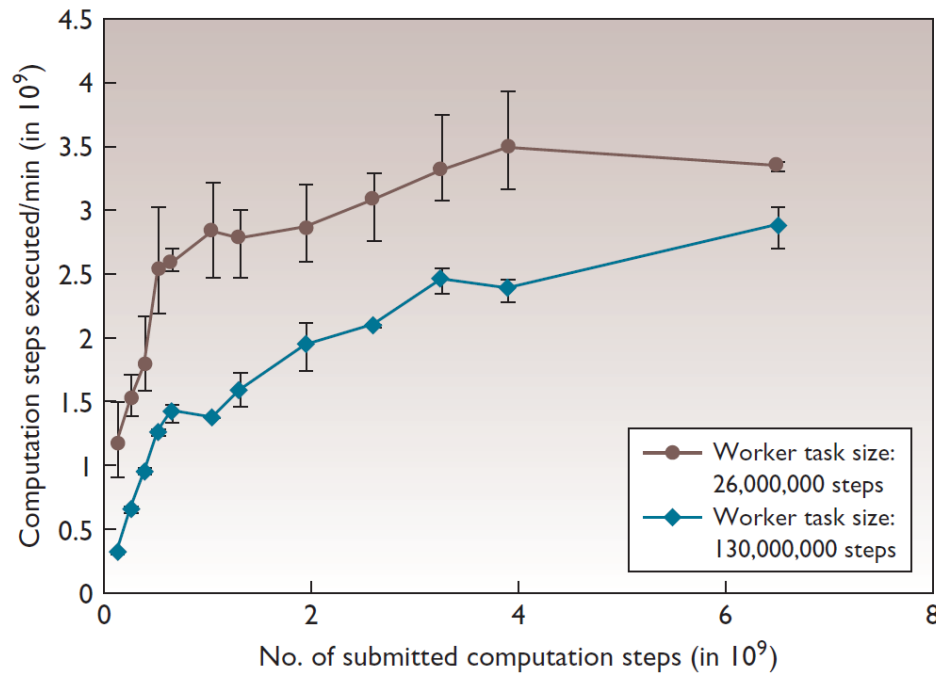
Osiągnięcia habilitacyjne

5/5

- » Badania eksperymentalne:
 - Pomiar wydajności aplikacji w chmurach [H9, H10, H13], (Amazon EC2, Google App Engine, Rackspace, IBM Softlayer, Google Cloud Functions, chmury prywatne OpenStack),
 - Opracowanie i wykorzystanie symulatora dla zbiorów grafów zadań [H2, H5],
- » Potwierdzenie efektywności i wydajności modeli optymalizacyjnych [H8, H6, H3] oraz algorytmów w szerokim zakresie parametrów (budżet, czas, wielkość zadań, wielkość i struktura grafu, rozmiary danych)
- » Pokazanie istotnego wpływu czynników takich jak:
 - opóźnienia w dostarczaniu zasobów [H4, H13],
 - awarie [H5],
 - ziarnistość obliczeń [H5, H6, H10]
 - zaburzenia i niedokładność estymacji [H3, H5, H8].

Osiągnięcie habilitacyjne

5/5 - wyniki



» Znaczenie wyników:

- Pokazanie zalet i zakresu stosowalności metod i algorytmów.
- Uzyskanie charakterystyk kosztowo-wydajnościowych.
- Zaobserwowanie nietrywialnego związku między ziarnistością zadań a wydajnością.
- Poszerzenie wiedzy o użyteczności chmur dla aplikacji naukowych.

Publikacje 5/5

- [H1] **M. Malawski**, “*Towards Serverless Execution of Scientific Workflows – HyperFlow Case Study*”: WORKS 2016 Workshop, Workflows in Support of Large-Scale Science, November 2016, Salt Lake City, Utah, USA, CEUR-WS.org, 2016
CI_{WoS} : 0, CI_{GS} : 0
- [H9] M. Bubak, M. Kasztelnik, **M. Malawski**, J. Meizner, P. Nowakowski, and S. Varma: “*Evaluation of cloud providers for VPH applications*”, 13th IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid 2013, Delft, Netherlands, May 13-16, 2013, pp 200-201. IEEE, 2013
CI_{WoS} : 1, CI_{GS} : 11
- [H10] **M. Malawski**, M. Kuzniar, P. Wójcik, and M. Bubak: “*How to Use Google App Engine for Free Computing*”, IEEE Internet Computing, vol. 17, no. 1, pp. 50-59, 2013
IF: 2,0, CI_{WoS} : 11, CI_{GS} : 37

Także w [H2], [H3], [H5], [H6], [H8], [H11], [H12], [H13]

Inne osiągnięcia naukowe

- » Zastosowanie chmur dla aplikacji bio-medycznych
 - Rozwój platformy Atmosphere (OpenStack, EC2, GCE, Azure) dla projektów VPH-Share, PL-Grid, EurValve
 - Narzędzia bioinformatyczne na infrastrukturze PL-Grid
- » Wdrażanie aplikacji workflow na chmurach
 - Wkład w rozwój silnika HyperFlow
 - Rozwój platformy PaaSage w projekcie EU FP7
- » Skalowalne architektury
 - Współpraca z Lufthansa Systems

PaaSage: użycie technik Model Driven Software Engineering (MDE)

Platforma PaaSage dla chmur:

- » CAMEL: Cloud Application Modeling and Execution Language
 - Model wdrożenia: komponenty, połączenia
 - Model wymagań
 - Model skalowania
- » Uruchamianie aplikacji na wielu chmurach
- » Autoskalowanie, adaptacja



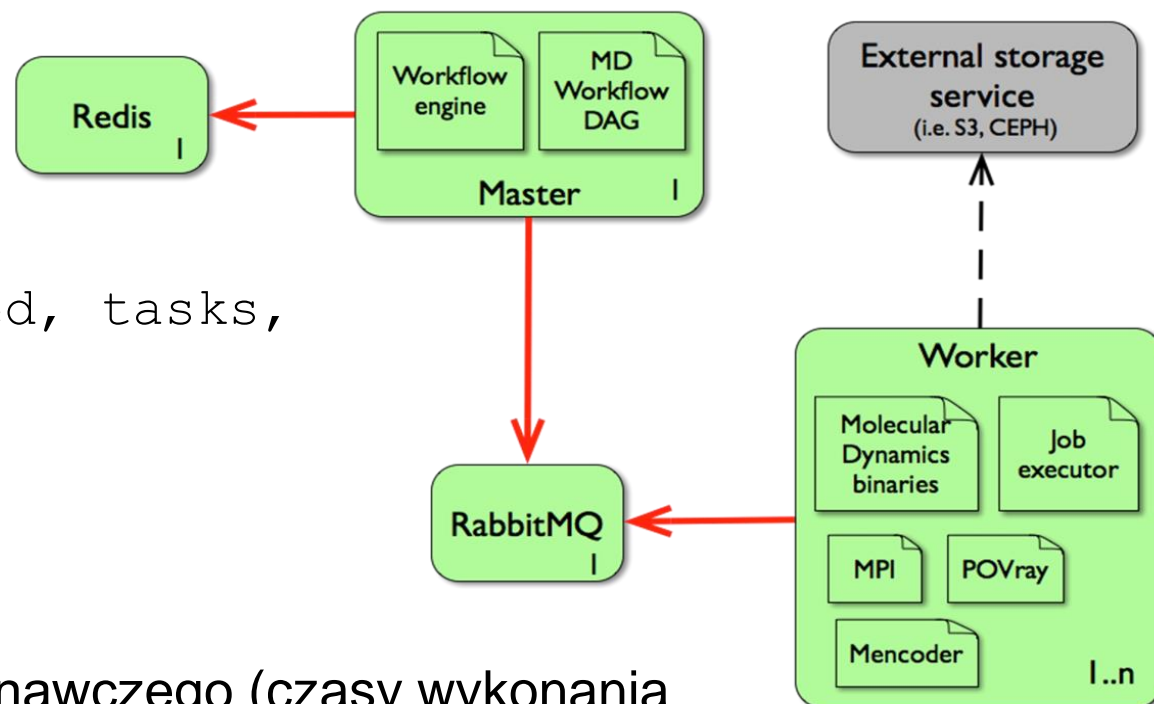
Integracja silnika HyperFlow z PaaSage:

- » Połączenie modelu aplikacji CAMEL z modelem grafu zadań HyperFlow
- » **Monitorowanie aplikacji** w celu autoskalowania

Uruchamianie i skalowanie HyperFlow w PaaSage

Metryki dla aplikacji:

- » Parametry silnika: `tasksLeft`, `outputsLeft`, `tasksProcessed`, `tasks`,
- » `stage` – aktualny poziom w grafie
- » `nConsumers` – liczba workerów
- » Długość kolejki
- » Szybkość przetwarzania zadań
- » Dodatkowe metryki z modułu wykonawczego (czasy wykonania zadań, czasy przesyłu plików)



Reguły skalowania

- » Ogólne (np. zależne od obciążenia procesora)
- » Specyficzne dla HyperFlow (np. czas wykonania zadań)

Skalowalna architektura oparta na wzorcach CQRS i Event Sourcing

- » Wzorce „reaktywne”
 - Command Query Responsibility Separation
 - Event Sourcing
- » Aplikacja do planowania lotów (Lufthansa Systems)
- » Nowe technologie: Akka, Cassandra, Kafka, Neo4J
- » Eksperymenty pokazujące skalowalność dla zapisów i odczytów
- » Andrzej Dębski, Bartłomiej Szczepanik, Maciej Malawski, Stefan Spahr, Dirk Muthig: „A Scalable, Reactive Architecture for Cloud Applications”, in IEEE Software, 2016 (accepted)



FEATURE: CLOUD COMPUTING

A Scalable, Reactive Architecture for Cloud Applications

Andrzej Dębski, Bartłomiej Szczepanik, and Maciej Malawski, AGH University of Science and Technology

Stefan Spahr and Dirk Muthig, Lufthansa Systems

// To create a prototype flight-scheduling application, researchers employed an architecture based on the Command Query Responsibility Segregation and Event Sourcing patterns. The architecture adheres to the Reactive Manifesto's goals: it's message-driven, elastic, resilient, and responsive. //

CLOUD TECHNOLOGIES, OWING to their capability for dynamic on-demand provisioning of computing resources, create new opportunities for developing scalable applications. However, to fully benefit from this capability, application architectures must be designed with scalability as a main objective. This lets developers deliver low-latency solutions that handle a high rate of requests per second and adjust resource usage to current needs.

In this article, we focus on the

Reactive Manifesto together with the Command Query Responsibility Segregation and Event Sourcing patterns and domain-driven design (DDD),³ a software design paradigm originating from the Reactive Manifesto. Our main objective was to determine how possible to implement a completely reactive, scalable CQRS-based application using the Reactive Manifesto's principles. We built a prototype flight-scheduling application using the Reactive Manifesto's principles. It provides a subset of the Reactive Manifesto's principles in a real-world application on scalability, elasticity, and responsiveness as a way to evaluate these new architectural principles.

Our experience with the application's architecture is embodied by an abstract that's available as a project at <https://github.com/andrewd/endeavour-cqrs-end>. We believe this framework can help solve scalability problems. Many similar applications of the CQRS and Event Sourcing patterns, implemented in various technologies, and evaluation of their present here.

Core Concepts

The Reactive Manifesto defines the need to build scalable systems. It defines a system's basic traits: autonomy, elasticity, responsiveness. It also defines techniques for building systems that are more flexible, more scalable, and failure-resilient.

The CQRS pattern separates operation state (commands) from query state (queries). This creates useful p

Publikacje a tematy

		[H1]	[H12]	[H13]	[H8]	[H6]	[H7]	[H3]	[H11]	[H5]	[H2]	[H4]	[H10]	[H9]
Model aplikacji	model komponentowy		■	■	■									
	bag-of-tasks		■	■	■							■	■	■
	wielopoziomowy graf zadań	■				■	■	■						
	zespół grafów zadań								■	■	■			
Model chmury	hybrydowa		■	■	■	■	■							
	homogeniczna	■						■	■	■	■	■	■	■
Metody	programowanie całkowitoliczbowe				■	■	■	■						
	heurystyki							■	■	■	■	■	■	
Czynniki	opóźnienia			■					■	■	■	■	■	
	niedokładność	■			■			■	■	■	■	■		
	awarie	■						■	■	■	■	■		
	ziarnistość	■				■	■			■			■	
	dane	■		■	■	■	■				■			
Eksperymenty	rzeczywiste infrastruktury		■	■			■					■	■	■
	symulacje				■	■		■	■	■				

Cykl publikacji

- » Czasopisma i materiały konferencyjne z listy JCR
 - 2 Future Generation Computer Systems, IF z roku wydania = {2,786 2,639}
 - 1 Journal of Grid Computing, IF = 1,507
 - 1 Scientific Programming, IF = 0,559
 - 1 IEEE Internet Computing, IF = 2,0
 - 1 International Journal of High Performance Computing Applications, IF = 1,295
- » Materiały konferencyjne
 - 1 SC: International Conference on High Performance Computing, Networking, Storage and Analysis (Supercomputing)
 - 2 International Conference on Computational Science, ICCS, Procedia Computer Science
 - 2 Parallel Processing and Applied Mathematics, LNCS, Springer (WoS)
 - 1 WORKS 2016 Workshop, Workflows in Support of Large-Scale Science at SC16
 - 1 IEEE/ACM International Symposium on Cluster, Cloud, and Grid Computing, CCGrid
- » Sumaryczny CI dla cyklu publikacji
 - według bazy WoS: 48 (bez autocytowań)
 - według Google Scholar: 296 (w tym publikacja [H11] 131 cytowań)

suma IF,
z roku
opublikowania:
11,482

Dane bibliometryczne

Statystyka (na koniec 2016)	Web of Science		Google Scholar	
	wszystkie	po doktoracie 2009-2016	wszystkie	2011-2016
Liczba prac	54	25	131	36
Liczba cytowań (bez autocytowań)	243 (199)	111 (96)	1288	671
Indeks Hirscha	10	7	19	13

- » publikacje naukowe w czasopismach - lista JCR: 11 pozycji,
- » publikacje w materiałach konferencyjnych (Web of Science): 43 pozycji,
- » publikacje w innych czasopismach: 9 pozycji,
- » rozdziały w książkach: 10 pozycji,
- » materiały międzynarodowej konferencji CGW (Cracow Grid Workshop): 16 pozycji,

Sumaryczny IF publikacji (po doktoracie), według JCR, zgodnie z rokiem opublikowania wynosi: 15,452



Projekty badawcze

- » Resource management for scientific applications in highly elastic cloud infrastructures
 - (NCN OPUS-11) 2017-2019 (**kierownik**)
 - » Projekty europejskie (FP5,6,7, H2020)
 - EURValve: Personalised Decision-Support for Heart Valve Disease, 2016-2019
 - PaaSage: Model-based Cloud Platform Upperware, 2012-2016,
 - VPH-Share: Virtual Physiological Human: Sharing for Healthcare - A Research Environment, 2012-2015,
 - MAPPER: Multiscale Applications on European e-Infrastructures, 2010-2011,
 - ViroLab: Virtual Laboratory for Infectious Diseases (IST-27446), 2006-2009,
 - Wcześniej: GREDIA, 2006-2009, CoreGRID Network of Excellence, 2005-2008, EuChinaGRID, 2006-2007, K-WfGrid, 2005-2007, CrossGrid, 2002-2005
 - » Krajowe
 - PL-Grid, PL-Grid Core (POIG), ISMOP (NCBiR)
 - Clusterix, SGI-Grid (KBN)
 - » AWS in Education z firmy Amazon Web Services \$30 000 na dostęp do zasobów (**kierownik**)
 - Research Grant: 2010-2015, 2015-2017
 - Teaching Grant, 2013-2015
- Rola habilitanta:
kierownik grantu (2 x),
kierownik zadania (4 x),



Staże zagraniczne

- » University of Notre Dame, Indiana, USA, Center for Research Computing, Prof. Jarek Nabrzyski
 - Czerwiec 2015 – sierpień 2015, Adjunct Research Assistant Professor
 - Czerwiec 2013 – wrzesień 2013, Adjunct Research Assistant Professor (CSE Department)
 - Luty 2011 – sierpień 2012, Postdoctoral Research Associate
- » Department of Math and Computer Science, Emory University, Atlanta, Prof. Vaidy Sunderam 2004 - 2005
 - 6-miesięczny staż naukowy,
- » INRIA, Sophia Antipolis, OASIS, Prof. Denis Caromel
 - 2006 miesięczny staż w ramach program CoreGRID Researcher Exchange Program
 - 2006 - 2008, trzy krótkie wizyty naukowe w zespole
- » Krótkie wizyty
 - University of Southern California, Information Sciences Institute, projekt Pegasus, Prof. Ewa Deelman, 2011, 2012, 2013
 - Section Computational Science, University of Amsterdam, Prof. Peter Sloot 2009, 10 dni
 - University of Westminster, Prof. Vladimir Getov 2005, 10 dni, projekt CoreGRID,
 - Argonne National Laboratory, Prof. Ian Foster 2002, USA: 10 dni, Globus Team,
 - CERN, 2002, 10 dniowa wizyta, spotkania robocze projektów EU DataGrid oraz CrossGrid

Czynny udział w konferencjach

- » Ważniejsze konferencje międzynarodowe (indeksowane przez WoS):
 - SC (Supercomputing): IEEE/ACM International Conference on High Performance Computing, Networking, Storage and Analysis 2011 (Seattle), 2012, 2016 (Salt Lake City) (CORE: A)
 - CCGrid: IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing 2012 (Ottawa), 2013 (Delft) (CORE: A)
 - UCC: IEEE/ACM International Conference on Utility and Cloud Computing, 2015 (Limassol)
 - HPCS: International Conference on High Performance Computing & Simulation, 2016 (Innsbruck) (CORE: B)
 - IPDPS: IEEE International Symposium on Parallel and Distributed Processing, 2004 (Denver), 2009 (Rzym) (CORE: A)
 - PPAM - International Conference on Parallel Processing and Applied Mathematics - 2013 (Warszawa), 2015 (Kraków) (CORE: C)

Zaproszone wykłady i seminaria

- » Wykład *Tech Talk* “Scientific workflows on clouds and on cloud functions” w **Google** Polska, Warszawa, grudzień 2016
- » Wykład **Keynote Lecture** “Scientific workflows in clouds and serverless infrastructures” na konferencji 5th International Young Scientists Conference in HPC and Simulation, Kraków, październik 2016
- » Wykład “Workflow scheduling and optimization on clouds” na seminarium **Dagstuhl Seminar 15281** – Algorithms and Scheduling Techniques to Manage Resilience and Power Consumption in Distributed Systems, Dagstuhl, Niemcy, lipiec 2015
- » Seminarium “Cloud Platform for VPH Applications” w Information Sciences Institute, **University of Southern California**, Marina Del Rey, USA, kwiecień 2013
- » **Zaproszony** wykład “Grid Computing – the Last Decade?” na konferencji Grid Computing: the Next Decade Workshop, Zakopane, styczeń 2012
- » Seminarium “Component Approach to Computational Applications on Grids and Clouds” na **University of Notre Dame, USA**, kwiecień 2011
- » Seminarium “GridSpace – environment for programming and executing complex applications on the Grid” w **Universiteit van Amsterdam, Holandia**, listopad 2009

Działalność organizacyjna 1/2

Członkostwo w komitetach organizacyjnych konferencji:

- » ACM/IEEE Supercomputing - SC16
Workshops Vice-Chair (A):
37 workshopów, 2500+ uczestników, SC16 10000+ uczestników
- » IEEE Cluster 2016
Posters Vice-Chair (A)
- » HPCS 2015, 2016,
Tutorials Co-Chair (B)

Członkostwo w komitetach programowych konferencji:

- » ICCS 2005-2017 (A)
- » ACM/IEEE SC 2014 (A)
- » IEEE IPDPS 2017 (A)
- » IEEE CCGrid 2014-2017 (A)
- » IEEE UCC 2015-2016
- » WORKS 2011-2016

(klasa konferencji wg. CORE)

Członek komitetu redakcyjnego

- » Future Generation Computer Systems od 2015 (IF 2.430)
- » Computing and Informatics – guest handling editor (IF 0.504)

Recenzowanie publikacji w czasopismach z listy JCR

- » Future Generation Computer Systems, Elsevier (IF 2.430)
- » Transactions on Parallel and Distributed Systems, IEEE (IF 2.173)
- » Transactions on Computers, IEEE (IF 1.723)
- » Journal of Grid Computing, Springer (IF 1.561)
- » Journal of Parallel and Distributed Computing, Elsevier (IF 1.320)
- » Computing, Springer (IF 0.872)
- » Scientific Programming, IOS, Hindawi (IF 0.455)

Nagrody i wyróżnienia

- » Stypendium Naukowe dla Wybitnych Młodych Naukowców Ministra Nauki i Szkolnictwa Wyższego na lata 2013-2015
- » I nagroda w konkursie Executable Paper Grand Challenge, Elsevier 2011 - dla zespołu za system „Collage Authoring Environment” wspierający łączenie interaktywnych wyników obliczeniowych z artykułami naukowymi.
- » Nagroda za najlepszy artykuł (Best paper award) na konferencji 6th Workshop on Large Scale Computations on Grids and 1st Workshop on Scalable Computing in Distributed Systems (LaSCoG-SCoDiS'10) , International Multiconference on Computer Science and Information Technology, 2010, IEEE
- » Rektorskie Nagrody Naukowe AGH za osiągnięcia naukowe
 - I, II i III stopnia za lata 2001, 2002, 2003, 2004, 2005, 2006, 2007, 2008, 2009 i 2013



Osiągnięcia dydaktyczne

- » **Promotor pomocniczy pracy doktorskiej** (Kamil Figiela, otwarcie przewodu 2016)
- » **Diamenty AGH**
 - I miejsce 2014/2015: Andrzej Dębski, Bartłomiej Szczepanik
 - Wyróżnienie 2012/2013: Kamil Figiela
- » **Konkurs na Najlepsze Prace Magisterskie z Informatyki**, Polskie Towarzystwo Informatyczne PTI
 - III miejsce 2015 Andrzej Dębski, Bartłomiej Szczepanik
- » **Opieka nad studentami**
 - 9 prac magisterskich (większość zakończonych publikacjami)
 - 6 projektów i prac inżynierskich (15 studentów)
- » **Dydaktyka**
 - wykłady:
 - Systemy gridowe / Metody i systemy obliczeń wielkiej skali (Inf. II st. niestac), Obiektowe metody projektowania systemów (Inż. akustyczna I st.), Inżynieria oprogramowania (Inf. Stosowana I st.)
 - ćwiczenia, laboratoria, projekty, seminaria:
 - Konceptcje i modele programowania (Inf. II st), Introduction to quantum computing (Inf. II st), Introduction to research methods (Inf. II st), Algorytmy równoległe (Inf. II st), Teoria współbieżności (Inf. I st), Metody obliczeniowe w nauce i technice (Inf. I st)



Dziękuję za uwagę

- » Grupa Systemów Komputerowych (CSG)
 - Prof. Jacek Kitowski
Dr Marian Bubak
- » Zespół DICE KI AGH & Cyfronet
 - Dr Marian Bubak, Piotr Nowakowski, Dr Bartosz Baliś,
Dr Katarzyna Rycerz, Tomasz Gubała, Maciej Pawlik,
Marek Kasztelnik, Bartosz Wilk, Tomasz Bartyński,
Jan Meizner, Daniel Haręźlak
- » Doktorant:
 - Kamil Figiela
- » Magistranci
 - Piotr Bryk, Tomasz Dziok, Andrzej Dębski, Bartłomiej
Szczepanik, Adam Zima, Adam Gajek
- » Studenci:
 - Dr Maciej Kuźniar, Piotr Wójcik
- » University of Southern California, Information Sciences Institute (USA)
 - Prof. Ewa Deelman (Research Professor and Director),
Dr Gideon Juve & Pegasus Team
- » Notre Dame, Center for Research Computing (USA)
 - Prof. Jarek Nabrzyski (Director)

» Referencje:

- Zespół DICE:
<http://dice.cyfronet.pl>
- Grupa Systemów Komputerowych (CSG)
Katedry Informatyki AGH:
www.icsr.agh.edu.pl/



DISTRIBUTED
COMPUTING
ENVIRONMENTS
T E A M

