

# SYSTEMY RÓWNOLEGŁE I ROZPROSZONE

2015-2016

semestr letni

## PROJEKT

## 1. Wprowadzenie.

Kto z twoich znajomych ma smartfona, w którym procesor ma tylko 1 rdzeń? Odpowiedź: prawie na pewno żaden. Jaki procent programów, które używasz na co dzień NIE jest systemem rozproszonym? Odpowiedź: bardzo niewiele. Dlaczego usługi np.: wyszukiwarka google, bing, etc. są dostępne *cały czas* ? Odpowiedź: nie wiesz. Ale możesz się dowiedzieć – np. szukając w google. Jak zdefiniować system rozproszony i po czym go poznać? Odpowiedź: pewnie było coś na wykładzie, ale nie pamiętasz. A powinieneś, bo korzystasz z nich *cały czas*.

Specyfika studiów informatycznych – póki co – jest taka, że sporo tego, czego się uczysz na studiach, nie jest tym, z czego będziesz korzystał w pracy, bo to już zdąży się zmienić przez lata. Dlatego bardzo ważne jest by rozumieć nie tylko technologie, ale i algorytmy, które są przez nią wykorzystywane.

W ramach projektu z przedmiotu Systemy Równoległe i Rozproszone będziesz miał okazję poznać pewne teoretyczne podstawy jak i praktyczne realizacje, wykorzystujące zarówno dojrzałe jak i całkiem nowe technologie rozproszone.

Niniejszy projekt ma dwojaki charakter: z jednej strony będzie wymagał pracy o charakterze naukowym, a z drugiej zrealizowania wybranych zagadnień z wykorzystaniem technologii rozproszonych – czyli implementacji rozumianej jako zaprojektowanie i zaprogramowanie programu realizującego konkretne rozwiązania.

## 2. Zadania

Jakich algorytmów dla systemów rozproszonych używamy cały czas, nie zdając sobie z tego sprawy? Dlaczego wiodące firmy ciągle tworzą nowe technologie dla systemów rozproszonych? Czy zapewnienia o ich przełomowości są coś warte? Jak dziś mają się technologie ogłaszane 5 lat temu jako przełomowe? Z którymi technologiami, które zdominują rynek, w czasie gdy będziesz pracować, już się zetknąłeś, a o których nie wiesz że istnieją? Jakie algorytmy „odwalają całą brudną robotę”, za każdym razem gdy korzystasz z internetu?

W ramach zadania należy, dla **wybranego zagadnienia z zakresu systemów rozproszonych\***, odnaleźć dotychczas opracowane algorytmy, przeprowadzić analizę porównawczą i wybrać najlepszy z nich. Następnie należy zidentyfikować praktycznie wykorzystywane **technologie\*\*** na rynku i zaimplementować ww. algorytm, po czym na tej podstawie przeprowadzić analizę porównawczą tych technologii.

**Korzystajcie tylko z wiarygodnych źródeł!**

**\*\*technologia – nie język programowania, a konkretny paradygmat realizujący komunikację w modelu pamięci rozproszonej**

Grupa	*zagadnienie w ujęciu rozproszonym (na wielu komputerach)
1	elekcja
2	synchronizacja
3	uzgadnianie
4	synchronizacja czasu rzeczywistego
5	Zegar logiczny
6	zegar wektorowy
7	wzajemne wykluczanie
8	konsensus
9	replikacja danych
10	odnajdywanie drogi (routing)
11	wykrywanie zakleszczania (deadlock)
12	transakcje
13	kolorowanie grafu
14	przechodzenie grafu
15	rozgłaszanie
16	zapisanie stanu - „migawka” (global snapshot)

### 3. Proces

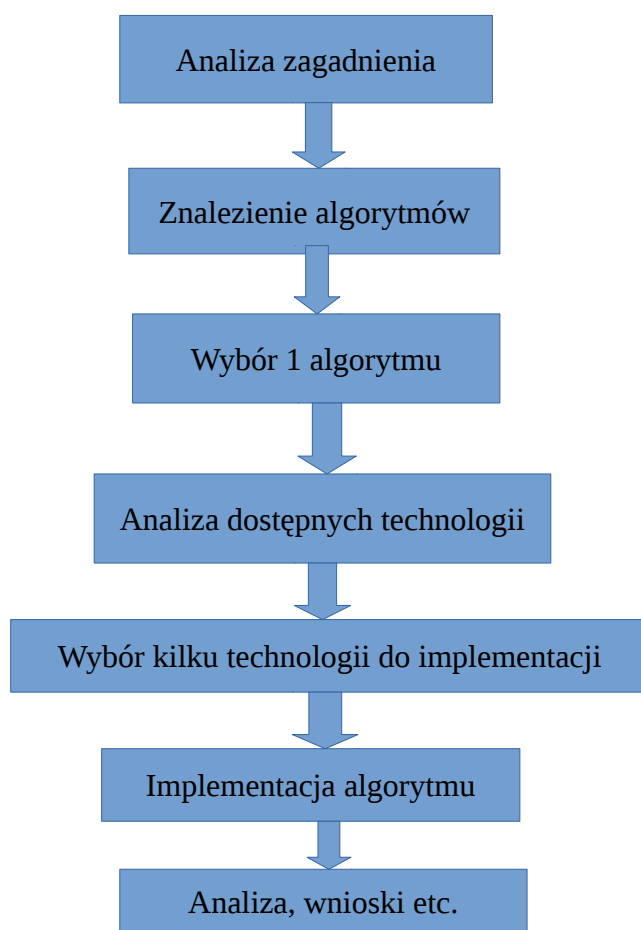
Realizacja zadania odbywa się grupowo. Każda grupa projektowa liczy 4-5 osób z jednej grupy laboratoryjnej. Należy wybrać lidera, który będzie się kontaktował z prowadzącym w imieniu grupy.

Projekt zaczyna się 28-29.04.2016, a kończy się 9-10.06.2016.

Na potrzeby projektu można korzystać z pracowni, gdy jest wolna – należy zgłosić się do prowadzącego.

Efektem waszej pracy powinno być:

1. tabelaryczne zestawienie algorytmów praktycznie rozwiązujących *wskazane zagadnienie z zakresu systemów rozproszonych\** wraz ze wskazaniem twórcy algorytmu (podać źródło; najlepiej oryginalne)
2. analiza porównawcza (w formie nie-tekstowej) algorytmów pod kątem (wskazać źródła informacji):
  - 2.1. złożoności obliczeniowej
  - 2.2. zakresu stosowalności
  - 2.3. trudności implementacyjnych
  - 2.4. znanych realizacji/wdrożeń
3. opis sposobu oceniania algorytmów i wyboru najlepszego
4. tabelaryczne zestawienie technologii rozproszonych obecnych na rynku (wskazać źródła)  
(Uwaga: Np. Java RMI i C# RMI to jedna technologia – RMI. Ale sockety w C i RPC w C to już dwie technologie, mimo że ten sam język programowania.)
5. implementacja wybranego algorytmu w 4-5 różnych technologiach
6. analiza porównawcza implementacji (w formie wykresów) z uwzględnieniem:
  - 6.1. nakładu (czasu) waszej pracy wymaganego do zaimplementowania
  - 6.2. porównania szybkości działania tego samego algorytmu w różnych technologiach
7. Prezentacja multimedialna (ok. 10 min) podsumowująca waszą pracę.
8. Kompletna bibliografia z podaniem wiarygodnych źródeł zdobytych informacji.
9. Płyta CD/DVD zawierająca elementy z pkt. 1-8.



**Podział pracy wewnątrz grup:**

Każda osoba jest odpowiedzialna za:

- znalezienie przynajmniej 1 algorytmu dla wskazanego zagadnienia; zanalizowania i zrozumienia go oraz przedstawienia jego oceny (w ramach opracowanego sposobu oceniania) w celu umożliwienia analizy porównawczej
- za znalezienie przynajmniej 1 technologii i zaimplementowanie wybranego (przez całą grupę) algorytmu oraz dokonanie pomiarów w celu umożliwienia analizy porównawczej
- ponadto poszczególne osoby są odpowiedzialne za:

Osoba 1	Stworzenie prezentacji (7) [i bibliografii(8) gdy 4 osoby w zespole].
Osoba 2	Analiza porównawcza algorytmów i ich oceniania (2,3)
Osoba 3	Tabelaryczne zestawienia (1,4).
Osoba 4	Analiza porównawcza implementacji (6)
(Osoba 5)	Stworzenie wysokiej jakości bibliografii (8).

#### 4. Ewaluacja.

Na ocenę waszej pracy, będą miały wpływ następujące czynniki:

1. sposób gromadzenia danych, ich analiza, selekcja, hierarchizacja, uporządkowanie i wnioskowanie
2. zawartość merytoryczna analiz teoretycznych
3. poprawność językowa, stylistyczna i schludność pracy
4. poprawność implementacji
5. dobór i sposób cytowania źródeł
6. terminowość wykonania pracy

Za wykonanie zadania można zdobyć od 0 do 18 punktów.

PUNKTACJA	0 pkt	1 pkt	2 pkt	3 pkt
1. sposób gromadzenia danych, ich analiza, selekcja, hierarchizacja, uporządkowanie i wnioskowanie	<ul style="list-style-type: none"><li>• Zebrane wiadomości są niekompletne i nieuporządkowane</li><li>• treści zostały wprost skopiowane</li><li>• brak bibliografii</li><li>• brak odnośników</li><li>• wnioski są niespójne i nielogiczne</li></ul>	<ul style="list-style-type: none"><li>• Zebrano i uporządkowano podstawowe informacje</li><li>• zebrany materiał ma pewne luki</li><li>• podana bibliografia z niewielkimi błędami</li><li>• wnioski poprawne, lecz brak logicznego uzasadnienia</li></ul>	<ul style="list-style-type: none"><li>• Zgromadzono kompletne dane, uporządkowane w należyty sposób</li><li>• poprawnie opracowana bibliografia oraz odnośniki do źródeł literaturowych</li><li>• zamieszczono logiczny wywód zakończony trafnymi wnioskami świadczącymi o właściwym doborze i zrozumieniu treści</li></ul>	<ul style="list-style-type: none"><li>• Zebrano informacje z różnych źródeł, znacznie wykraczające poza założone cele</li><li>• niekonwencjonalnie i wnikliwie opracowane dane służą opracowaniu wniosków</li><li>• poprawnie i z dużą starannością opracowana bibliografia oraz odnośniki do wszystkich istotniejszych informacji i założeń</li><li>• trafne, wzorowo skonstruowane wnioski o wysokim stopniu przenikliwości i świadczące o głębokim</li></ul>

				zrozumieniu problemu
2. zawartość merytoryczna analiz teoretycznych	<ul style="list-style-type: none"> <li>Wybrany algorytm jest niewłaściwy dla wskazanego zagadnienia</li> <li>Wybrana technologia lub algorytm jest wykorzystywana przez więcej niż 1 osobę w grupie</li> <li>Analiza porównawcza skupia się na cechach nieistotnych</li> <li>Tabele zawierają nieistotne dla problemu informacje</li> </ul>	<ul style="list-style-type: none"> <li>Analiza obejmuje algorytm i technologię poprawnie dobraną do problemu</li> <li>Zakres poruszanych kwestii jest teoretyczno-podręcznikowo-modelowy i nie wkracza w rozwiązania używane w rzeczywistości</li> <li>Tabele zawierają wszystkie istotne informacje</li> </ul>	<ul style="list-style-type: none"> <li>Analiza dotyczy algorytmów i technologii także używanych w zagadnieniach praktycznych</li> <li>Wybór został dokonany spośród szerszego wachlarza algorytmów i technologii</li> </ul>	<ul style="list-style-type: none"> <li>Analiza obejmuje znacznie więcej algorytmów i technologii niż jest to wymagane</li> <li>Zestawienia zawierają wszystkie istotne dane, z pominięciem danych nieistotnych dla problemu</li> <li>Zakres poruszanych kwestii dotyczy aktualnego stanu wiedzy naukowej (state-of-the-art) oraz rozwiązań praktycznych</li> </ul>
3. poprawność językowa, stylistyczna i schludność pracy	<ul style="list-style-type: none"> <li>Tabele i wykresy są nieopisane lub opisane niepoprawnie</li> <li>Tekst jest niechlujny, niegramatyczny</li> <li>Słownictwo jest niejasne, pojęcia używane niekonsekwentnie</li> <li>Niewłaściwie użyte określenia z dziedziny zagadnienia</li> </ul>	<ul style="list-style-type: none"> <li>Tabele, wykresy etc. są opisane poprawnie</li> <li>Tekst jest czytelny i gramatyczny</li> <li>Struktura dokumentów jest poprawna</li> <li>Występuje poprawnie używane słownictwo dziedzinowe</li> </ul>	<ul style="list-style-type: none"> <li>Słownictwo z dziedziny zagadnienia jest zdefiniowane i używane precyzyjnie</li> <li>Tekst jest klarowny i jasny</li> <li>Brak „pseudo naukowego bełkotu”</li> </ul>	<ul style="list-style-type: none"> <li>Tekst klasy artykułu/sprawozdania naukowego</li> <li>Struktura dokumentu logiczna i przejrzysta</li> <li>Ze struktury tekstu jednoznacznie wynika co jest podawanym faktem, a co wnioskiem</li> <li>Brak subiektywnych opinii autora, wszystkie wnioski jednoznacznie wynikają z podanych faktów</li> </ul>
4. poprawność implementacji	<ul style="list-style-type: none"> <li>Implementacja niekompletna lub</li> </ul>	<ul style="list-style-type: none"> <li>Implementacja działa poprawnie na 2</li> </ul>	<ul style="list-style-type: none"> <li>Implementacja działa poprawnie na 3-4</li> </ul>	<ul style="list-style-type: none"> <li>Implementacja działa poprawnie na 5+</li> </ul>

	<ul style="list-style-type: none"> <li>niedziałająca</li> <li>Implementacja niewłaściwego algorytmu</li> <li>Implementacja z użyciem niewłaściwej technologii</li> <li>brak projektu implementacji</li> </ul>	<ul style="list-style-type: none"> <li>komputerach</li> <li>projekt implementacji jest ad-hoc</li> <li>brak systemu kontroli wersji</li> </ul>	<ul style="list-style-type: none"> <li>komputerach</li> <li>Implementacja wykrywa podstawowe awarie</li> <li>projekt zawiera diagram komunikacji UML</li> <li>Grupa korzystała z systemu kontroli wersji Git/SVN etc.</li> </ul>	<ul style="list-style-type: none"> <li>komputerach</li> <li>Implementacja ma opcje trwałej konfiguracji ustawień</li> <li>projekt zawiera diagramy UML przepływu danych i sekwencji</li> <li>kody źródłowe zawierają dokumentację w formacie Doxygen lub podobnym</li> </ul>
5. dobór i sposób cytowania źródeł	<ul style="list-style-type: none"> <li>Cytowania z niemożliwą do sprawdzenia wiarygodnością (wikipedia, fora internetowe etc.)</li> <li>Cytowania niezwiązane z treścią</li> </ul>	<ul style="list-style-type: none"> <li>Cytowania wiarygodnych źródeł ze sprawdzonym autorem (nazwisko/firma)</li> <li>Cytowania źródeł pośrednich (opracowania, podręczniki, skrypty, blogi)</li> <li>Cytowania ogólne, bez wskazania konkretnych fragmentów cytowanych dzieł</li> </ul>	<ul style="list-style-type: none"> <li>Cytowania posiadają datę dostępu</li> <li>Cytowania źródeł oryginalnych</li> <li>Cytowania źródeł obcojęzycznych</li> </ul>	<ul style="list-style-type: none"> <li>Wszystkie cytowania odnoszą się do wysokiej jakości źródeł o uznanej międzynarodowej renomie</li> <li>Cytowania podane w profesjonalny sposób, ze wskazaniem autora, tytułu, daty, wydania, wydawcy, numeru ISBN</li> <li>wskazanie konkretnej strony/działu</li> </ul>
6. terminowość wykonania pracy	<ul style="list-style-type: none"> <li>Praca oddana 7 dni po terminie lub później</li> </ul>	<ul style="list-style-type: none"> <li>Praca oddana do 6 dni od daty ukończenia projektu</li> </ul>	<ul style="list-style-type: none"> <li>Praca oddana do 24h po terminie oddania projektu</li> </ul>	<ul style="list-style-type: none"> <li>Praca oddana przed lub w terminie</li> </ul>

Ocena końcowa	2,0	3,0	3,5	4,0	4,5	5,0
Punkty	0-8	9-10	11-12	13-14	15-16	17-18



## 5. Źródła.

Poniżej podane są przykładowe rzetelne źródła. Nie obejmują one całości zagadnienia i stanowią tylko przykład i punkt zaczepienia. Zachęcam do korzystania z:

- Herb Sutter, „Welcome to the Jungle”, artykuł: <https://herbsutter.com/welcome-to-the-jungle/>, video: <https://herbsutter.com/2012/06/21/talk-video-welcome-to-the-jungle-60-min-version-qa/> (dostęp 27.04.2016)
- Lista aktualnych konferencji i wydarzeń min. dot. Systemów Rozproszonych: <http://lanyrd.com/topics/distributed-systems/> (dostęp: 27.04.2016)
- Aktualna lista punktowanych publikacji naukowych: <http://www.nauka.gov.pl/aktualnosci-ministerstwo/juz-sa-nowe-listy-punktowanych-czasopism-na-2015-rok.html> (dostęp 27.04.2016)
- A. S. Tanenbaum, “Distributed Systems. Principles and Paradigms” (2nd ed.), Prentice Hall 2002
- G. Coulouris et al., “Distributed Systems. Concepts and Design” (4th ed.), Addison Wesley, 2005 (wydanie polskie “Systemy rozproszone. Podstawy i projektowanie”, WNT 1998)
- Sukumar Ghosh, „Distributed Systems: An Algorithmic Approach”, 2007 by Taylor & Francis Group, LLC
- „Systemy Równoległe i Rozproszone”, strona z materiałami dydaktycznym dr hab. inż. K. Banasia, <http://www.metal.agh.edu.pl/~banas/SRR/SRR.html> (dostęp 27.04.2016)
- Baza publikacji wydawnictwa Springer: <http://link.springer.com/>, (dostęp 27.04.2016) – może wymagać dostępu z IP AGH'owskiego w celu zalogowania
- Baza publikacji Web Of Science: <http://webofscience.com> (dostęp 27.04.2016) – może wymagać dostępu z IP AGH'owskiego w celu zalogowania
- Douglas C. Schmidt, „Software Technologies for Developing Distributed Systems: Objects and Beyond”, <https://www.dre.vanderbilt.edu/~schmidt/PDF/CSI-article.pdf>, (dostęp 27.04.2016)

## 6. Konkluzja

Projekt stanowi niezłą okazję do zmierzenia się z nieco poważniejszym, trudniejszym – ale bardziej rzetelnym podejściem do postawionego zagadnienia. Mam nadzieję, że w czasie pracy nad nim uda się Wam odkryć zagadnienia i treści, o których dotychczas nie wiedzieliście. Wierzę, że zaprezentowana formuła zachęci Was do szerszego spojrzenia na koncepcję i wykorzystanie systemów rozproszonych.