

Techniki programowania równoległego

MPI

Paweł Maczuga

Wybór node'ów

Pliki, zawierające node'y wyglądały następująco:

Dla jednego node'a:

```
vnode-03.dydaktyka.icsr.agh.edu.pl:4
```

Dla 2 node'ów na jednym hoście fizycznym:

```
vnode-03.dydaktyka.icsr.agh.edu.pl:1
```

```
vnode-04.dydaktyka.icsr.agh.edu.pl:1
```

Dla 2 node'ów na różnych hostach fizycznych:

```
vnode-09.dydaktyka.icsr.agh.edu.pl
```

```
vnode-10.dydaktyka.icsr.agh.edu.pl
```

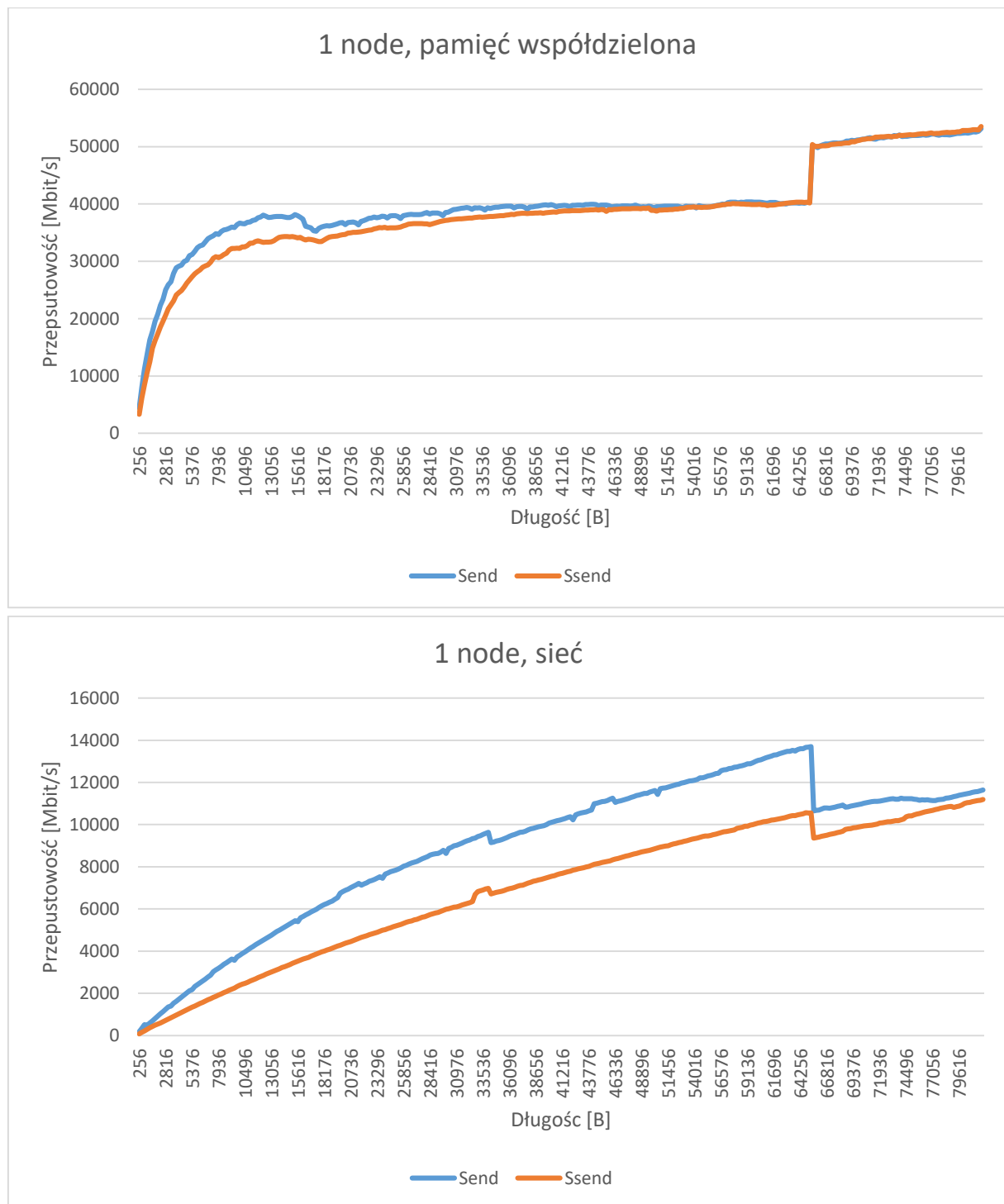
Opóźnienie [ms]

	1 node	1 node, sieć	2 nody, 1 host	2 nody, 2 hosty
Send	0,000295	0,010867	0,151720	0,102250
SSend	0,000583	0,026135	0,453183	0,258753

Użycie operacji synchronicznej Ssend zwiększa opóźnienie w każdym przypadku. Jest to zgodne z przewidywaniami, jako że proces po wysłaniu wiadomości czeka aż cała wiadomość dotrze do odbiorcy.

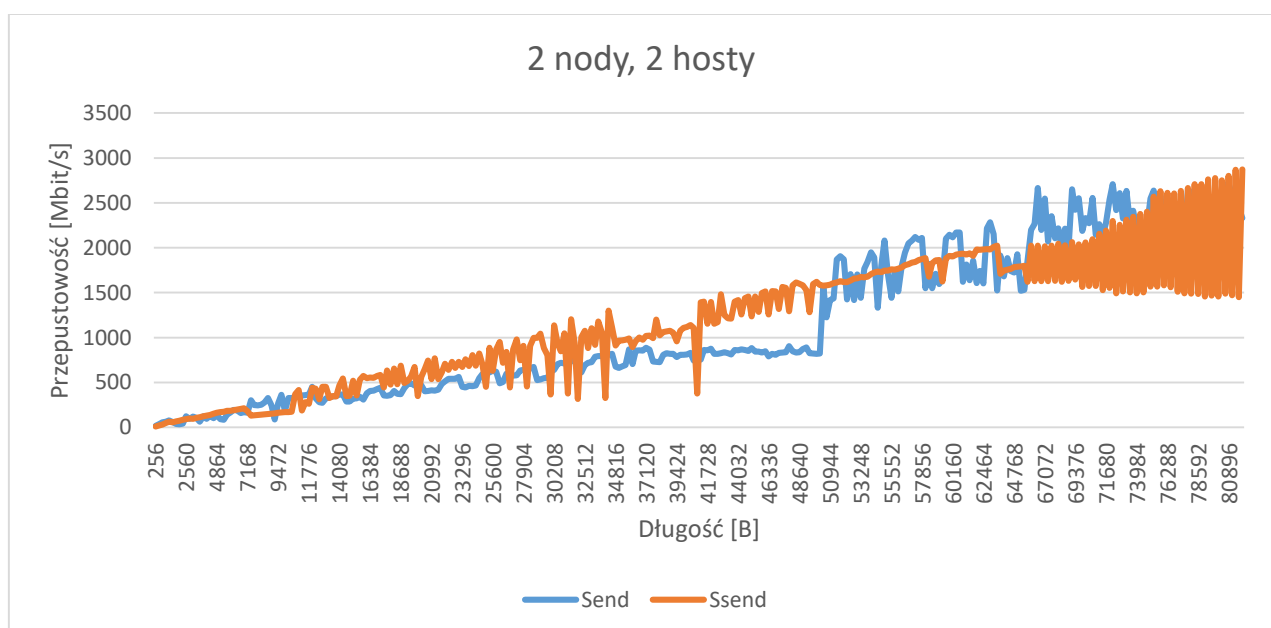
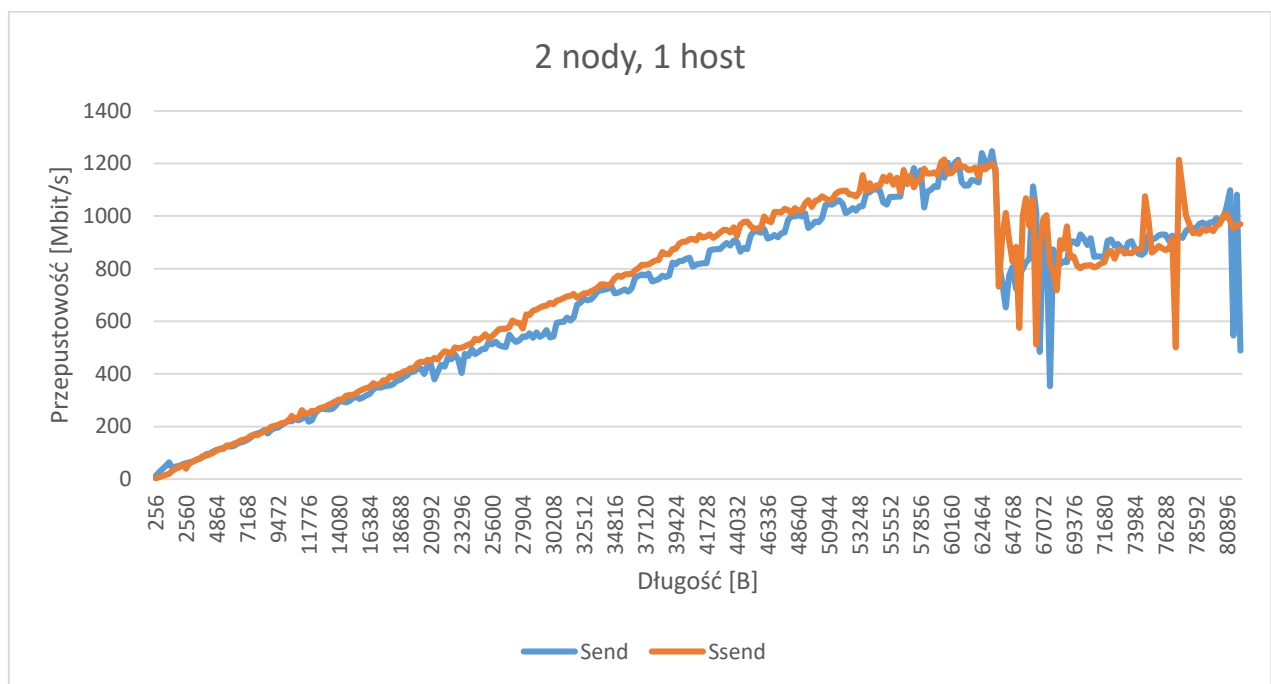
Opóźnienie znacznie zwiększa się po wymuszeniu komunikacji przez sieć w stosunku do pamięci współdzielonej. Kolejny skok następuje po przejściu na 2 node'y. Co ciekawe jest ono większe dla node'ów na jednym hoście fizycznym niż na dwóch. Nie jestem w stanie powiedzieć dlaczego tak się dzieje. Testy były prowadzone dla różnych kombinacji node'ów i w różnych odstępach czasu, za każdym razem czasy były podobne.

Przepustowość



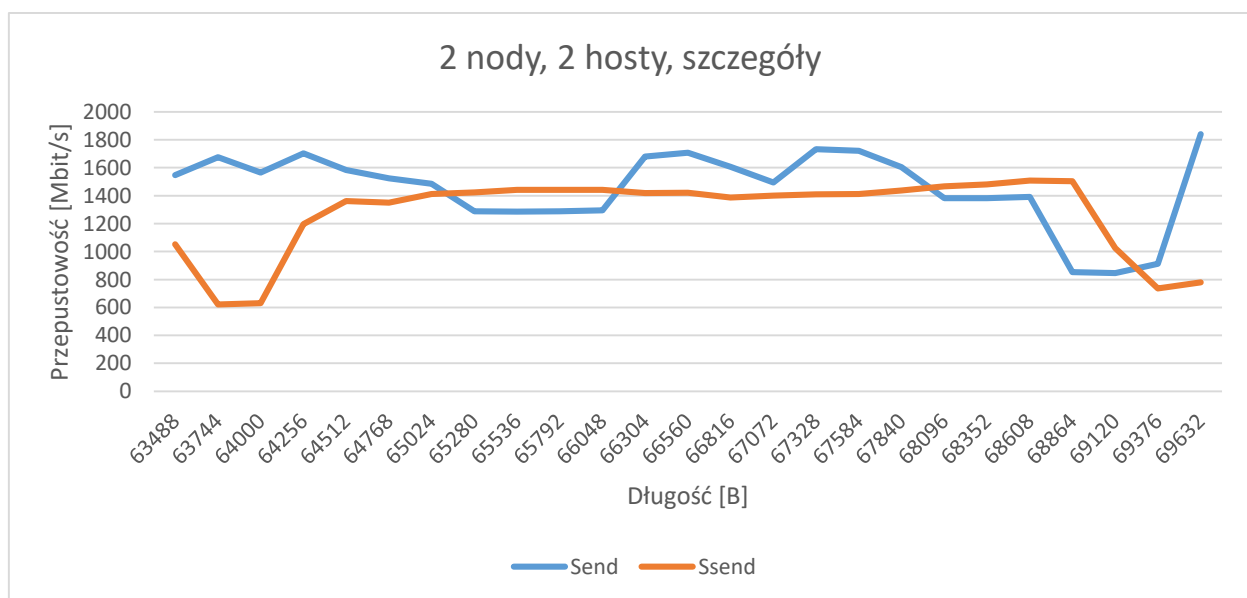
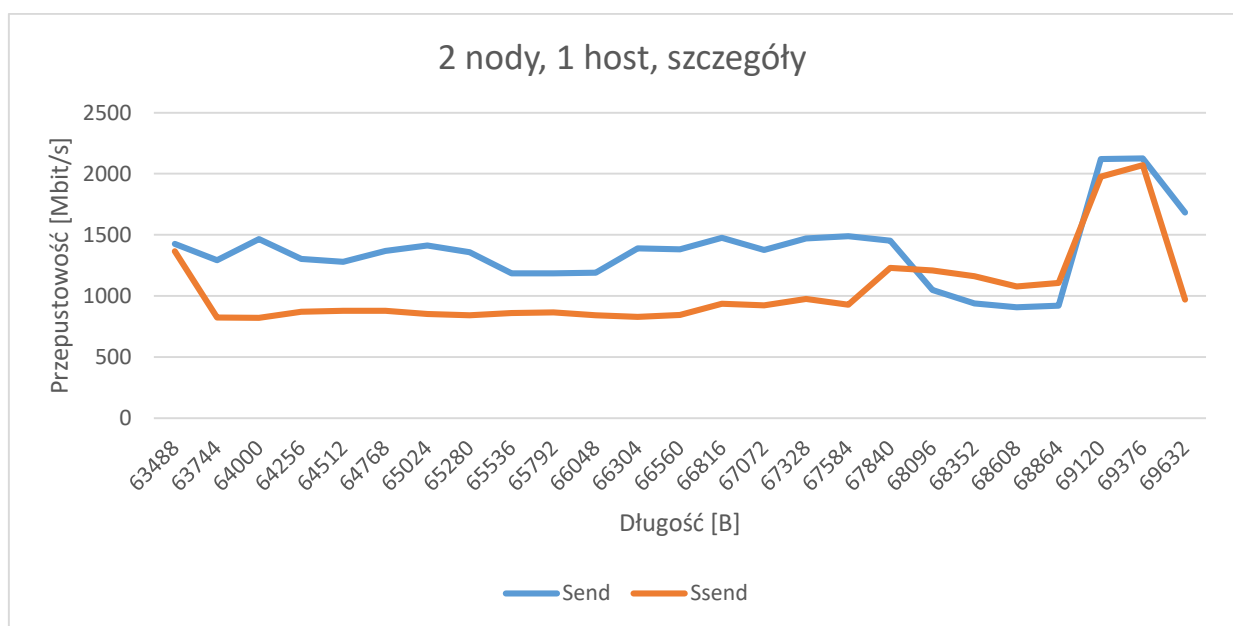
Podczas użycia pamięci współdzielonej różnice między wysyłaniem standardowym i synchronicznym są niewielkie i istotne tylko dla małej wielkości danych. Przepustowość początkowo szybko wzrasta, po czym zatrzymuje się na mniej więcej 40 000 Mbit/s. Co ciekawe gdy rozmiar wiadomości wynosi **dokładnie 64 KB** następuje gwałtowny **wzrost** przepustowości. Program był uruchamiany kilkakrotnie i za każdym razem nagły skok występował dokładnie dla tej samej wielkości komunikatu.

Zmuszenie procesów do komunikacji przez sieć znacząco zmniejsza przepustowość. Wpływa również na kształt wykresu. Wzrost przepustowości jest tutaj bardziej liniowy, aż do długości 64 KB. W tym przypadku następuje jednak znaczący **spadek** przepustowości. Wysyłanie synchroniczne jest również widocznie wolniejsze na całym przedziale.



Dla procesów na 2 node'ach przepustowość jest większa gdy znajdują się na dwóch hostach fizycznych (co pokrywa się z wynikami opóźnienia). Wyniki te są znacznie bardziej zaszumione. Pomiar opóźnienia znaczenie różni się dla sąsiednich wartości. Jest to szczególnie widoczne dla 2 przesyłania synchronicznego na 2 hostach fizycznych. Widać również spadek przepustowości, jednak nie dla długości 64 KB.

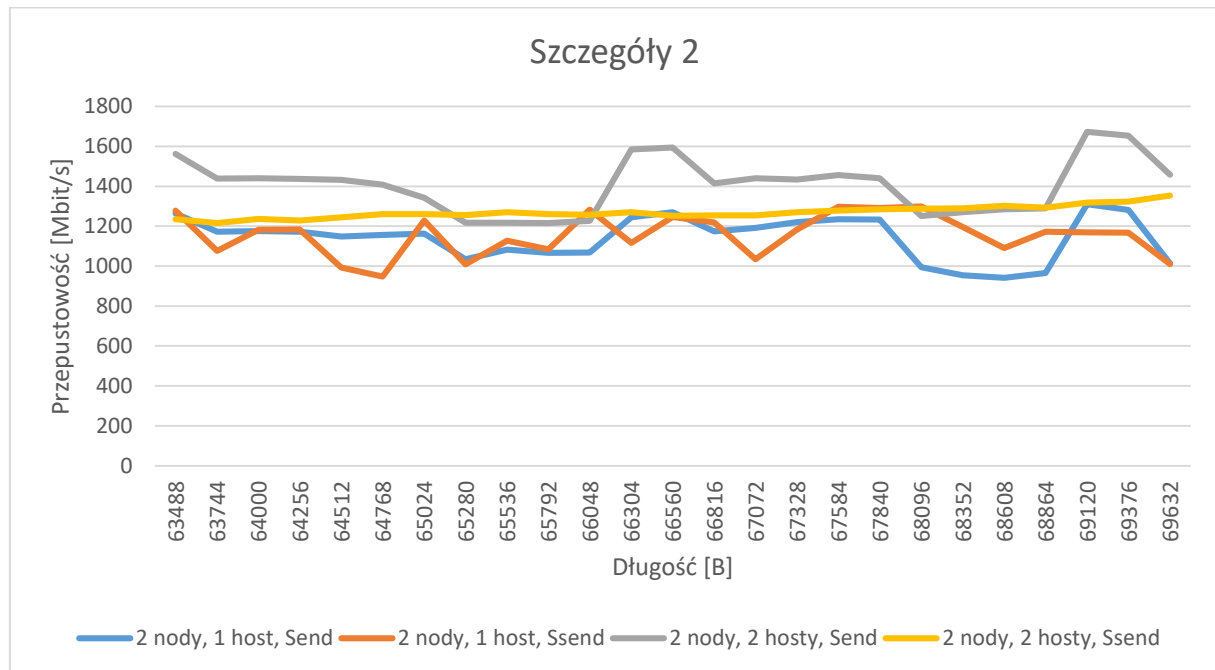
Dane do powyższych wykresów były generowane poprzez wysyłanie wiadomości pomiędzy procesami przez 2 sekundy, a następnie na tej podstawie wyliczana była przepustowość. By ograniczyć szum zdecydowałem się wydłużyć ten czas do 10 sekund dla znacznie mniejszego przedziału długości wiadomości – wokół 64 KB, które wydaje się być najciekawszym punktem. Rezultatem są poniższe wykresy:



Widać, że przepustowość utrzymuję się na mniej więcej stałym poziomie, ze spadkami i wzrostami. Interesujące jest to że przepustowości nie różnią się tak znacząco między procesami na 1, a 2 hostach. Podczas tych testów opóźnienia były jednak bardzo podobne jak z tabeli w rozdziale „Opóźnienia”.

Żadnych zmian w okolicach 64 KB.

Ponowne uruchomienie testów daje podobne rezultaty:



Widać jednak, że nody na 2 hostach fizycznych mają nieco większą przepustowość.

Wnioski

Nie jestem w stanie stwierdzić dlaczego podczas wysyłania wiadomości na 1 node’zie przepustowość znacznie zmienia się po 64 KB. Jest to zapewne związane z implementacją MPI, co może też wyjaśniać wzrost dla przesyłania pamięcią współdzieloną, a spadek dla wysyłania przez sieć. Ta długość zdaje się nie mieć żadnego wpływu dla procesów działających na różnych node’ach.

Wydłużenie czasu wysyłania komunikatów do 10 sekund znacznie zmniejsza szum na wykresie. Nie było możliwe jednak zrobienie tego dla całego przedziału, z taką dokładnością, gdyż po zbyt długim czasie działania procesy były zabijane.

Na wykresie „2 nodes, 1 host” (z pełnym przedziałem) spadek przepustowości w pewnym momencie jest prawdopodobnie związany ze wzrostem obciążenia maszyny, która być może zaczęła być wtedy używana przez inną osobę.