

Dostęp do danych - usługa PLG-Data

Krótki opis usługi

Usługa PLG-Data pozwala, za pośrednictwem przeglądarki Web, zarządzać swoimi zbiorami danych, zgromadzonymi na zasobach dyskowych **Infrastruktury PLGrid**. Dzięki niej, posługując się swoim zestawem informacji uwierzytelniających (może to być para login-hasło do Infrastruktury PLGrid, lub certyfikat p12 załadowany w przeglądarce) można **oglądać, pobierać, dodawać, usuwać, współdzielić lub zmieniać nazwy swoich plików i katalogów**. Usługa jest zabezpieczona szyfrowanym protokołem HTTPS (na odcinku połączenia komputera użytkownika z portalem usługi) oraz protokołem GridFTP, pomiędzy portalem usługi oraz klastrem obliczeniowym. Zastosowanie tych technik pozwala użytkownikom korzystać z niej w sposób bezpieczny oraz, z punktu widzenia innych współużytkowników Infrastruktury PLGrid, ogranicza ich możliwości do zakresu przyznanego przez operatorów zasobów PLGrid. Innymi słowy, **osoba korzystająca z usługi PLG-Data nie nabywa żadnych uprawnień w dostępie do zbiorów danych, poza tymi, które już posiada** i może wykorzystać na inne sposoby (zdalne narzędzia SSH, SCP itp.).



W obecnej wersji usługa jest skonfigurowana tak, aby współpracować z przestrzeniami dyskowymi klastrów obliczeniowych **Prometheus**, w ośrodku obliczeniowym **ACK Cyfronet AGH, Kraków**.

Adres usługi:

<https://data.plgrid.pl/>

Usługa przeznaczona jest dla wszystkich użytkowników zarejestrowanych w Portalu PL-Grid, którzy posiadają aktywne konto na przynajmniej jednym z docelowych klastrów obliczeniowych (patrz notka powyżej), oraz chcieliby zarządzać własnymi plikami zdalnie, za pośrednictwem przeglądarki. Przy wykorzystaniu **mechanizmu zespołów PLGrid**, usługa ta także pozwala w prosty sposób współdzielić np. wyniki badań z innymi współpracownikami w obrębie grup badawczych bądź instytutów.

Aktywowanie usługi

Aby móc użyć usługi PLG-Data, należy spełnić kilka warunków:

- posiadać konto PLGrid - w celu otrzymania takiego konta (bezpłatnego dla wszystkich polskich naukowców i ich zagranicznych współpracowników), trzeba udać się na [stronę Portalu PLGrid](#) a następnie **zarejestrować** (instrukcja: [Rejestracja](#)).
- usługa, ze swej istoty, ma sens tylko w przypadku posiadania konta użytkownika (tzw. **konto shellowe**) na wybranym klastrze obliczeniowym; po uzyskaniu potwierdzenia dostępu do PLGrid, należy w [Katalogu Aplikacji i Usług](#) odszukać poniższe usługę/usługi, a następnie je aktywować:
 - "**Dostęp do klastra Prometheus**" aby uzyskać dostęp do klastra Prometheus ([strona podręcznika poświęcona grantom](#))
- w Portalu PLGrid, w [zakładce Certyfikaty](#), należy wygenerować swój własny **certyfikat SimpleCA** (w oparciu o ten certyfikat następuje efekt delegacji uprawnień użytkownika, dzięki któremu usługa PLG-Data będzie mogła działać na właściwym koncie klastra - tzw. mechanizm *X.509 proxy certificates*); więcej danych odnośnie tworzenia i użytkowania certyfikatów znajduje się w [odpowiedniej sekcji podręcznika użytkownika](#); należy zapamiętać hasło klucza, które podaliśmy tworząc certyfikat SimpleCA
 - Alternatywą jest skorzystanie z własnego certyfikatu gridowego (np. PolishGridCA), który trzeba umieścić w LDAP - odpowiednia formatka także w [zakładce Certyfikaty](#)
- uzyskawszy potwierdzenie dostępu do klastra Prometheus (m.in. powinno ono przyjść pocztą elektroniczną na nasz adres), należy w tym [Katalogu Aplikacji i Usług](#) aktywować usługę "**PLG-Data: mini-usługa do zarządzania plikami na klastrze**"; również w tym przypadku aktywacja powinna nastąpić automatycznie, z ewentualnym opóźnieniem kilku minut.

Jeśli powyższe kroki zakończyły się pomyślnie, usługa PLG-Data powinna być dostępna. Należy odwiedzić stronę usługi: https://data.plgrid.pl, gdzie będzie możliwe zalogowanie się z wykorzystaniem wyżej opisanych mechanizmów bezpieczeństwa. Obrazki poniżej pokazują cztery kolejne kroki logowania się - kliknij konkretny obrazek aby uzyskać powiększenie i krótkie wyjaśnienie danego kroku.

Witamy!

Prezentujemy Ci usługę PL-Grid Data, dzięki której uzyskasz dostęp do katalogów i plików, które
nich są zarówno twoje personalne zbiory, jak również współdzielone zasoby zespołów, do który

Poniższy przycisk pozwoli Ci zalogować się do usługi, korzystając z systemu PL-Grid OpenID. W t
Infrastruktury PL-Grid) jak również dodatkowego hasła, użytego podczas generowania certyfika

Zaloguj się przez OpenID PL-Grid

Właśnie korzystasz z systemu OpenID, który pozwoli Ci zalogować się do usługi dziedzicznej PLG-Data: mini-

Logowanie

Login

Hasło:

Zaloguj się [Odzyskiwanie hasła \(Portal PL-Grid\)](#)

Właśnie korzystasz z systemu OpenID, który pozwoli Ci zalogować się do usługi dziedzicznej PLG-Data: mini-usługa

Generowanie proxy certyfikatu wymaganego przez usługę dziedziczną.

Proxy zostanie wygenerowane na podstawie certyfikatu SimpleCA zgłoszonego do LDAP:

DN użytkownika

O=Uzytkownik,O=PL-Grid,C=PL

Data wygaśnięcia certyfikatu 2014-12-11 18:20:10.0

Hasło do klucza prywatnego

Kontynuuj

Logowanie przy pomocy OpenID PL-Grid zakończyło się sukcesem.

Witamy!

Prezentujemy Ci usługę PL-Grid Data, dzięki której uzyskasz dostęp do katalogów i plików, któ
nich są zarówno twoje personalne zbiory, jak również współdzielone zasoby zespołów

Nie zawsze, niestety, pierwsze logowanie do usługi się powiedzie (głównie ze względu na złożony mechanizm delegacji uprawnień). Jeśli na którymś etapie (krok 2 i krok 3) logowania do OpenID PL-Grid, pojawił się błąd, lub w kroku 4 usługa PLG-Data przywitała Cię komunikatem czerwonym, prosimy o:

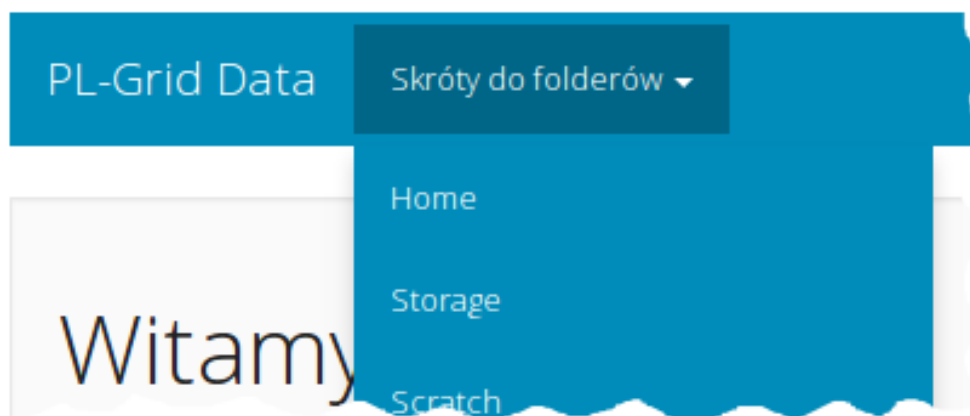
1. (w pierwszej kolejności) upewnienie się jeszcze raz, że wszystkie warunki wymienione na początku tej instrukcji (aktywne konto PL-Grid, aktywna usługa Prometheus, aktywna usługa PLG-Data, posiadanie certyfikatu SimpleCA) zostały spełnione,
2. skontaktowanie się z nami [poprzez system Helpdesk](#).

Męczące logowanie?

Można ułatwić sobie nieco powyższą, dwuetapową procedurę logowania się do usługi PLG-Data. W chwili generowania certyfikatu SimpleCA w Portalu PL-Grid, otrzymujemy (jako plik) kopię naszego nowego certyfikatu w formacie .p12. Taki [certyfikat można umieścić w przeglądarce](#). Wówczas krok 2 procedury logowania będzie przebiegał automatycznie i bez konieczności podania hasła (przeglądarka przedstawi się w OpenID PL-Grid naszym certyfikatem).

Pierwsze kroki

Spójrz na górną belkę ekranu powitalnego usługi PLG-Data. Po lewej stronie znajduje się menu skrótów, po prawej informacje o Tobie, przełącznik języka oraz przycisk wylogowania się⁽²⁾.



Menu skrótów zawiera zawsze kilka elementów. Dla użytkowników klastra Prometheus, skrót **Storage** nie będzie istniał, gdyż na tym klastrze nie ma możliwości długoterminowego składowania plików poza swoim katalogiem domowym lub katalogami grup (poniżej).

[Rozdział instrukcji użytkownika Infrastruktury PLGrid poświęcony przechowywaniu danych](#) zawiera dokładne informacje odnośnie przeznaczenia poszczególnych

katalogów, charakteru plików jakie powinny być w nich przechowywane, limitów objętości etc. Katalogi określone w instrukcji zmiennymi środowiskowymi PLG_* odpowiadają skrótom umieszczonym na pierwszych pozycjach górnego menu usługi.

Dalsze pozycje menu skrótów, ich etykiety i kolejność, zależą od **zestawu zespołów PL-Grid**, do których należysz. Jeśli jesteś pierwszą osobą w twojej grupie badawczej, która korzysta z Infrastruktury PLGrid, czas zostać pionierem i, [korzystając z instrukcji](#), założyć w Portalu PLGrid wasz zespół, oraz zaprosić do niego twoich współpracowników.

Kliknięcie na którejkolwiek pozycji z menu skrótów przeniesie Cię do wskazanego miejsca w strukturze katalogów klastra obliczeniowego (wrócimy do tego za chwilę).



Prawa część górnej belki z kolei zawiera informacje o Tobie, uzyskane przez usługę PLG-Data z serwera OpenID PLGrid. Możesz je obejrzeć najężdżając kursorem myszy na sylwetkę człowieka. Poza twoimi imionami i nazwiskiem oraz loginem, możesz sprawdzić tutaj listę zespołów PLGrid, do których należysz. Jeśli nie ma w tym zestawie twojego zespołu, a uważasz, że taki zespół został założony w Infrastrukturze PLGrid i twoje konto zostało

do niego przypisane, [proszę zgłoś ten fakt przez Helpdesk PLGrid](#).

Przycisk PL|EN pozwala zmienić język z domyślnego języka polskiego, na angielski.

Panel zawartości folderu

Praktycznie rzecz ujmując, jest to jedyny typ panelu (widoku) usługi PLG-Data. Jego zawartość zmienia się wraz ze zmianą aktualnej lokalizacji użytkownika, ale jego funkcjonalność pozostaje taka sama.

ZAWARTOŚĆ FOLDERU: PEOPLE /  BADANIA / LUTY 2014 / Dodaj pliki Nowy folder					
Prawa	Rozmiar	Data modyfikacji	Nazwa	Typ	
-rw-r--r-- 	11.6 MB	Jan 27 22:24	 GSM1019938_NBM6-P.CEL	Plik	Na pewno? Tak Nie
 	3.5 MB	Jan 27 22:31	 GSM933008_US91903683 Dec08_1_1.txt	Plik	Usuń
-rw-r--r-- 	5.11 MB	Feb 5 16:25	GSM980568_MB2009031253.CEL	 	Usuń
drwxr-xr-x 	2 KB	Feb 8 14:47	 aktualne_pliki	Folder	Usuń
-r--r--r-- 	78 KB	25 21:02	 all-...-3class-tr	Plik	Usuń

Mamy nadzieję, że większość informacji zawartych w panelu folderu, jak i operacji, które on umożliwia, jest dość oczywistych. Poniżej spróbujemy pokrótce je omówić.

Zawartość listy plików i folderów

Belka z zawartością folderu pokazuje Ci twoją aktualną lokalizację w drzewie katalogów klastra obliczeniowego. Rzecz jasna, będzie inna w zależności od miejsca, gdzie się znajdujesz. Ułatwia orientację. Poszczególne elementy *ścieżki* (ang. *path*) są tutaj linkami i ich kliknięcie przeniesie Cię właśnie do tego folderu (uwaga: nie radzimy klikać skrótu [PEOPLE](#) - chyba, że chcesz poznać nazwy katalogów domowych wszystkich użytkowników klastra - uprzedzamy, ta operacja może być czasochłonna).

 W powyższym przykładzie: jesteśmy w folderze `/people/....` (**tutaj nazwa konta**)/`badania/luty 2014`, kliknięcie np. skrótu [BADANIA](#) przeniesie nas do folderu nadrzędnego, względem tego, w którym się znajdujemy.

Główna tabela zawiera zestaw dość oczywistych danych: nazwę zbioru, typ (czy jest to plik czy folder), jego wielkość (tak, w systemie plików Linux, który zarządza klastrem obliczeniowym, także foldery mają swój rozmiar, nie jest on jednak związany z objętością plików, które się w nich znajdują), datę ostatniej modyfikacji oraz zestaw uprawnień.

Jak się zapewne domyślasz, kliknięcie na nazwę folderu spowoduje przeniesienie Cię do widoku jego zawartości, natomiast kliknięcie nazwy pliku spowoduje rozpoczęcie operacji pobierania go na dysk twardy twojego komputera. Zatrzymując na chwilę wskaźnik myszy nad nazwą pliku graficznego (o rozszerzeniach .gif, .png, .jpg lub .jpeg) zobaczysz podgląd tego pliku (co np. pozwoli Ci szybciej odszukać odpowiedni plik graficzny w zbiorze wielu takich pozycji).

Operacje podstawowe

Zestaw akcji, które możesz wykonać korzystając z panelu zawartości folderu, to:

- zmiana aktualnego folderu lub pobranie pliku (poprzez kliknięcie na nazwę folderu bądź pliku)
- usunięcie pliku lub pustego folderu (pomarańczowy przycisk [Usuń](#) z prawej strony oraz potwierdzenie czerwonym przyciskiem z etykietą **Tak**); folder zawierający dalsze zbiory jest zabezpieczony przed takim usunięciem; akcję widać w pierwszym wierszu katalogu na powyższym obrazie
- zmiana nazwy pliku lub folderu (w wyniku kliknięcia **szarej** ikony z lewej strony jego nazwy, wpisania nowej wartości i potwierdzenia **zielonym** przyciskiem lub klawiszem Enter); akcję widać w trzecim wierszu na powyższym obrazie; akcja nie jest możliwa, jeśli nie jesteś *właścicielem* danego zbioru
- dodanie nowego folderu (poprzez kliknięcie na zielonym przycisku [Nowy folder](#) w górnej belce widoku, wpisanie żądanej nazwy oraz zatwierdzenie jej).

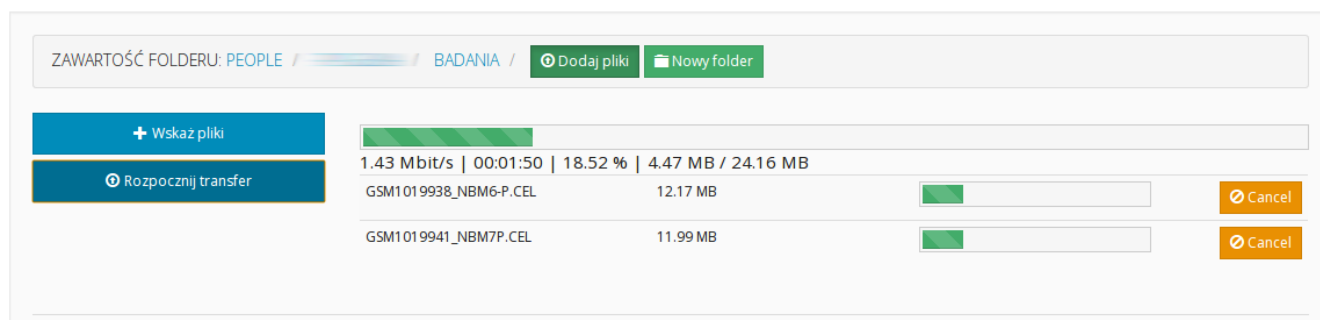
Wiele z tych akcji może się nie powieść, ze względu na ograniczony zakres twoich uprawnień. Dla przykładu, zazwyczaj nie możesz zakładać nowych folderów w folderze, którego nie jesteś właścicielką/właścicielem, często nie możesz kasować plików nie należących do Ciebie, niektórych folderów nie możesz także odwiedzać. W każdym z tych przypadków otrzymasz **czerwony komunikat**, pochodzący wprost z serwera GridFTP nadzorującego operację, wyjaśniający, dlaczego dana akcja nie zakończyła się sukcesem. Poniżej przykład takiego komunikatu, pojawiający się przy próbie skasowania niepełnego folderu.

```
globus_ftp_client: the server responded with an error 500 500-Command failed : System error in rmdir: Directory not empty 500-A system call failed: Directory not empty 500 End.
```

 W obecnej wersji usługi znakomita większość tego typu komunikatów nie jest tłumaczona - usługa PLG-Data przekazuje Ci je dokładnie takie, jakie otrzymała z klastra obliczeniowego (dokładniej, z serwera GridFTP wykonanego w technologii Globus, który odpowiada za zarządzanie dostępem do twoich plików na klastrze). Przepraszamy za nieco nieczytelny format tych komunikatów, jednak w większości przypadków ich lektura oraz niewielka znajomość języka angielskiego powinny pozwolić zrozumieć, co było przyczyną problemu.

Transfer nowych plików⁽¹⁾

Akcja ta pozwala Ci umieścić nowe pliki wewnątrz twoich folderów (lub w folderach zespołów, do których należysz) - pliki te umieszczasz tam poprzez załadowanie ich z dysku twojego komputera. Panel transferu plików otwierasz klikając na zielony przycisk **Dodaj pliki** w górnej części widoku zawartości folderu.



Transfer nowych plików w usłudze PLG-Data odbywa się w sposób nieblokujący (tj. "w tle"). Listę plików, które mają zostać przesłane, tworzysz klikając na niebieski przycisk **Wskaz pliki**. W zależności od przeglądarki i systemu operacyjnego, których używasz, możesz wskazać od razu kilka plików (zamiast wielokrotnie klikać wspomniany przycisk) lub - alternatywnie - możesz *upuścić* (tzw. *drag and drop*) wybrane pliki z dysku twojego komputera wprost na ekran zawartości folderu. W każdym z tych przypadków wybrane pliki pojawią się na "liście transferowej" (patrz obraz powyżej).

Kliknięcie niebieskiego przycisku **Rozpocznij transfer** powoduje rozpoczęcie przesyłania plików z twojego komputera na klastę. Prędkość przesyłu zależy od przepustowości łącza internetowego, jakim dysponujesz oraz (w pewnym stopniu) od aktualnego obciążenia serwera usługi PLG-Data. Postępy są uwidocznione poprzez zielone paski postępu (ogólny, powyżej, i dla każdego pliku osobno).

Zmiana uprawnień dostępu do pliku lub folderu

W systemie operacyjnym Linux, który zarządza uprawnieniami zbiorów twoich i wszystkich innych współużytkowników klastra, zaimplementowano odpowiedni mechanizm strzeżenia dostępu do plików i folderów. Serwery GridFTP oraz PLG-Data, które pośredniczą w dostępie do twoich zbiorów, respektują ten mechanizm (dokładniej: nie mają żadnej możliwości zignorowania tego mechanizmu ze względu na ich ograniczone uprawnienia). Jeśli nie znasz działania tego mechanizmu, [artykuł w Wikipedii](#) Ci go przybliży. Ogólnie jednak rzecz ujmując, jeśli nie chcesz wnikać w tę dość zawiłą materię, doradzamy nie zmieniać praw dostępu do twoich plików, pozostawiając ich (względnie bezpieczne) wartości domyślne.

Zmiana praw dostępu do pliku lub folderu jest możliwa po kliknięciu **szarej** ikonki zlokalizowanej w lewej kolumnie listy. Pozwala ona zmienić dowolnie z **dziwięciu** uprawnień *odczytu*, *zapisu* i *uruchamiania* dla, odpowiednio, *właściciela*, *grupy* oraz *wszystkich pozostałych użytkowników*. Wpisanie znaku **r**, **w** bądź **x** przyznaje dane uprawnienia (odpowiednio: *r-odczyt*, *w-zapis*, *x-uruchamianie*) podczas kiedy wpisanie znaku **-** (myślnik) te uprawnienia likwiduje. Przykładowo, wektor uprawnień następującej postaci:

rw-rw-r--

pozwala właścicielowi (Tobie) plik odczytać oraz zmodyfikować (bądź usunąć). Także członkowie twojego zespołu będą mieli podobne uprawnienia. Natomiast każdy inny użytkownik klastra, tj. nie należący do twojego zespołu, będzie mógł plik odczytać (czyli pobrać na swój komputer) ale nie będzie go mógł zmodyfikować ani usunąć. Akcja zmiany uprawnień dostępu jest możliwa tylko jeśli jesteś *właścicielem* danego zbioru.



Uprawnienia domyślne na klastrach

Polityka domyślnych praw dostępu na klastrach Prometheus blokuje dostęp innym użytkownikom klastra do twojego katalogu domowego (tzw. *Home*). Proponujemy pozostawić to ustawienie praw na tym poziomie - pozwala ono spokojnie zarządzać własnymi plikami w obrębie katalogu osobistego. Natomiast kiedy będziesz chciał udostępnić jakiś plik twoim współpracownikom, proponujemy [założenie w tym celu zespołu PL-Grid](#) i przypisanie do niego wszystkich użytkowników, z którymi chcesz się dzielić danymi. W wyniku tej operacji, po ponownym zalogowaniu się do usługi PLG-Data, każdy z członków zespołu będzie mógł przejść do katalogu zespołu wybierając jego nazwę z menu skrótów. Umieszczone w tym katalogu pliki będą możliwe do odczytu, pobrania, modyfikacji oraz skasowania przez wszystkich członków zespołu. Pozostali użytkownicy PL-Grid nie będą mogli odwiedzać folderu tego zespołu ani czytać jego zawartości.

Zapraszamy do używania

Coś nie działa albo potrzebujesz dodatkowych funkcjonalności: prosimy o [zgłoszenie w HelpDesku](#). Nie masz konta w PL-Grid, a chcesz np. porozmawiać o współpracy - wyślij mail bezpośrednio na operator@plgrid.pl.

Usługa PLG-Data jest ochotniczą ktrybucją do zasobów usług PL-Grid, powstała jako odzew na zgłoszone zapotrzebowanie użytkowników gridu dziedzinowego LifeScience. Ze względu jednak na ogólniejszy charakter usługi, nie ograniczamy dostępu do niej jedynie dla ludzi związanych z jedną konkretną dziedziną. Głównym autorem usługi jest [Tomasz Gubała](#) (ACK Cyfronet AGH, [DICE](#)), który stworzył ją wraz z *pewną pomocą swoich przyjaciół*. Życzymy rychłych transferów!

Podziękowania. Tomkowi Szepeńcowi i Marcinowi Radeckiemu (oni wiedzą, za co) oraz Łukaszowi Flisowi i Tomkowi Bartyńskiemu (za pomoc z konfiguracją usługi). Szczególne podziękowania dla Maćka Pawlika za wspólną pracę nad ruby-gridftp, biblioteką kluczową w działaniu PLG-Data. Orz wszystkim osobom zaangażowanym w proces wdrażania usług w PL-Grid (you know who you are).

⁽¹⁾ Jeśli korzystasz z przeglądarki Internet Explorer 9, dużo z opisanych w tej sekcji funkcji nie będzie działać. Nie będzie można wybrać więcej niż 1 pliku jednocześnie w celu załadowania, nie jest też możliwe "upuszczanie" plików z folderu Windows na panel przeglądarki, w celu przesłania ich na serwer. Podczas przesyłu danych na serwer, nie będą działać paski postępu. Dodatkowo mogą wystąpić problemy z przesłaniem większych plików. Polecamy aktualizację do wersji 10 programu, lub skorzystanie z przeglądarki [Chrome](#) bądź [Firefox](#).

⁽²⁾ Jeśli korzystasz z przeglądarki na urządzeniu mobilnym, górne menu usługi jest dostępne po kliknięciu przycisku w prawym górnym rogu ekranu. Także obrócenie urządzenia do orientacji poziomej (ang. landscape) pomoże w nawigacji. Część opcji widoku zostanie zredukowana aby uzyskać większą przejrzystość prezentowanych informacji.