Załącznik nr 3 do Studium Wykonalności

# Metoda wymiarowania infrastruktury wykorzystywana w ramach SIG

## Wprowadzenie

Przy realizacji SIG wykorzystywana jest dedykowana dla potrzeb GUGiK metoda wymiarowania infrastruktury sprzętowo – programowej, metoda jest używana na etapie projektowania systemów informatycznych. Szacowanie infrastruktury dla uruchomienia systemu uwzględnia wolumetrię systemu, co najmniej w wymiarach wymaganej pojemności przestrzeni dyskowej oraz wymagań wysokiej dostępności, które przekładają się na stosowną redundancję w infrastrukturze.

Niezależnie od wyników wymiarowania infrastruktury na etapie projektowania, w ramach utrzymania funkcjonuje proces zarządzania pojemnością który zajmuje się zapewnieniem wymaganego poziomu pojemności infrastruktury. Funkcjonowanie procesu zarządzania pojemnością zapewnia korygowanie zapotrzebowania na infrastrukturę bazując aktualnych informacji pochodzących z monitorowania wykorzystania (utylizacji) komponentów infrastruktury sprzętowo – programowej.

Przy tworzeniu metody uwzględniono następujące, główne założenia (pryncypia):

1. Szacowanie dotyczy usług biznesowych realizowanych przez GUGiK
2. Wynik wymiarowania powinien być referencyjny i łatwy do weryfikacji
3. Preferowane jest podejście kompleksowe nad podejściem szczegółowym

Pierwsze pryncypium oznacza, że punkt wyjścia oraz wynik działania metody powinien wprost dotyczyć usług biznesowych określonych w architekturze biznesowej SIG, na którą składają się architektury biznesowe poszczególnych projektów inwestycyjnych, tj. Geoportal2, Teryt3, GBODT, ISOK, ZSIN, GESUT oraz SZTEMAP. Metoda powinna umożliwiać przełożenie celów usługi biznesowej na wymagania sprzętowo – programowe.

Pryncypium 2 oznacza, że wynik metody wymiarowania będzie wyrażony w jednostce uznawanego testu wzorcowego (tzw. benchmarku). Dzięki temu można odnosić się do poziomów wydajności referencyjnych platform sprzętowych przetestowanych w ramach metody powiązanej z testem wzorcowym. Pierwotnie dla opisywanej metody wymiarowania przyjęto test wzorcowy - *SPECjEnterprise2010*. Jednostka wynikowa testu *SPECjEnterprise2010* to EOpS (Enterprise Operations per Second). Jest to miara wydajności testowanego sprzętu mówiąca ile wysoko-poziomowych operacji biznesowych jest w stanie obsłużyć testowany sprzęt w 1 sekundzie. Po zdobyciu własnych doświadczeń w ramach SIG odnośnie potrzeb infrastrukturalnych wynikających z realizowanych procesów biznesowych, wyniki uzyskiwane dzięki wymiarowaniu są odnoszone do posiadanej infrastruktury, dzięki czemu dobór komponentów infrastruktury jest adekwatny do sposobów wsparcia procesów biznesowych realizowanych w SIG.

Pryncypium 3 oznacza, że nie powinno się badać do najmniejszego szczegółu złożoności i różnorodności przyjętych pojęć (np. usług, scenariuszy, właściwości użytkowników), tylko preferować cechy dominujące, np. uśredniając lub pomijać nieznaczące właściwości pojęć. Celem nadrzędnym jest podejście kompleksowe obejmujące wszystkie badane cechy.

## Przedmiot metody

Metoda obejmuje usługi świadczone przez projekty Geoportal2, GBDOT, ISOK, TERYT3, GESUT, ZSIN oraz SZTemMAP. W szczególności zakres przedmiotowy metody został wyznaczony na podstawie ogólnego diagramu architektury systemów informatycznych GUGiK.



Rysunek 1 Diagram ogólny architektury systemów informatycznych GUGiK

Na podstawie diagramu ogólnego architektury GUGiK określono 4 obszary tematyczne.

1. Obszar systemów dziedzinowych
2. Obszar geoportali
3. Obszar węzłów IIP
4. Obszar szyny usług i węzła centralnego IIP

Zakres analizy każdego z obszarów tematycznych określają scenariusze, tj. przypadki ich wykorzystania przez klientów (użytkowników lub inne systemy).

Poszczególne obszary w znacznej części pokrywają się ze sobą, ponieważ w większości przypadków scenariusze nie ograniczają się jedynie do danego obszaru, lecz ich wykonanie wymaga interakcji z większością usług SIG. Na przykład, scenariusze z obszaru systemów dziedzinowych wymagają interakcji z usługami szyny usług i węzła centralnego. W konsekwencji, diagramy ilustrujące poszczególne obszary przedstawiają nie tyko usługi z danego obszaru, lecz również pozostałe elementy uczestniczące w scenariuszach.

## Procedura wymiarowania

Końcowy wynik wymiarowania wylicza się zgodnie z następującą procedurą.

Rysunek 2 Procedura wymiarowania

Podstawowe wyniki uzyskane dzięki zastosowaniu metody to:

* Liczba operacji biznesowych w rozkładzie na usługi biznesowe SIG
* Liczba wolumenu danych w rozkładzie na usługi biznesowe SIG
* Wskaźniki wyrażające wydajność warstwy aplikacji oraz bazodanowej wyrażone w jednostkach stanowiących standard przemysłowy

## Lista parametrów metody

Podstawową operacją wpływającą na uzyskiwane wyniki jest modyfikacja wartości parametrów metody. Niniejszy rozdział przedstawia listę podstawowych parametrów metody w postaci tabelarycznej.

Tabela 2 Lista parametrów metody

|  |  |  |  |
| --- | --- | --- | --- |
| Nazwa parametru | Opis parametru | Źródło | Przeznaczenie |
| Lista użytkowników | Zestawienie typów interesariuszy korzystających z usług systemu | Architektura biznesowa | Identyfikacja użytkowników wpływających na obciążenie SIG |
| Grupy użytkowników | Zestawienie grup użytkowników  | Mały Rocznik Statystyczny Polski, GUS | Identyfikacja grup charakteryzujących się zbliżonymi właściwościami (w zakresie wykorzystania SIG) |
| Liczność populacji użytkowników | Zestawienie ilościowe poszczególnych elementów listy użytkowników w kontekście danej grupy | Analiza na etapie projektowania systemów (odniesiona do danych referencyjnych) | Identyfikacja docelowej liczności grup użytkowników |
| Współczynnik rocznego wykorzystania szczytowego | Procentowy udział reprezentantów grupy korzystających z usług systemu | Analiza na etapie projektowania systemów | Identyfikacja rzeczywistej liczby użytkowników wykorzystujących SIG |
| Założenia wolumetryczne | Zbiór założeń dotyczących wolumetrii danych wykorzystywanych głównie w szacowaniu pasma cechujących się znaczącą wrażliwością na wyniki metody | Analiza na etapie projektowania SIG | Szacowania wolumetryczne |
| Lista scenariuszy | Zestawienie scenariuszy realizowanych w systemie | Na podstawie scenariuszy przyjętych podczas opracowania SIG dla kluczowych przepływów komunikatów w ramach SIG | Badanie złożoności usług biznesowych |
| Wskaźnik „Int-Scen-Proc”  | Procent użytkowników wykorzystujących dany scenariusz | Analiza na etapie projektowania systemów | Wskazanie realnego wykorzystania poszczególnych usług biznesowych przez daną grupę użytkowników |
| Lista usług biznesowych | Zestawienie usług biznesowych udostępnianych przez system | Architektura biznesowa  | Podstawowa jednostka wymiarowania |
| Wskaźnik „ZłożonośćOper”, „ZłożonośćWol” | Wskazanie złożoności usługi biznesowej i mapowanie, który scenariusz przypada na daną usługę biznesową | Na podstawie scenariuszy przyjętych podczas opracowania SIG dla kluczowych przepływów w ramach SIG | Mapowanie wskazujące jakie usługi biznesowe biorą udział w realizacji wybranego scenariusza |
| Wskaźnik „BSpYear” | Wskazanie jak często pojedynczy przedstawiciel danej grupy korzysta z danej usługi biznesowej | Analiza na etapie projektowania systemów | Wskazanie stopnia rocznego wykorzystania usługi biznesowej przez pojedynczego reprezentanta grupy użytkowników |
| Współczynniki technologiczne | Współczynniki określające stopień potencjalnej rozbieżności pomiędzy biznesowym szacunkiem wydajności a techniczną implementacją potencjonalnego wykonawcy systemu |
| Współczynniki rozkładu ruchu | Współczynniki określające właściwości rozkładu ruchu z uwzględnieniem wartości szczytowych (np. obciążenia z powodu jednoczesnego wykorzystania systemów przez znaczną liczbę użytkowników, w szczególności przez liczbę wyższą niż średnia) |

## Lista danych wynikowych

Dla zadanego zbioru założeń otrzymuje się zbiór wyników, który zawarty jest w kilku arkuszach. Podstawowymi arkuszami wynikowymi są:

* Wskaźniki – wskaźniki operacji biznesowych wpływających na wymaganą moc obliczeniową
* Warstwy – rozkład wymaganej mocy obliczeniowej w warstwach technicznych systemu
* RozkładEO – rozkład operacji biznesowych
* RozkładDV – rozkład przepływu danych

Wyniki zawarte są również w arkuszach pośrednich metody, które niosą oddzielną wartość i mogą być wykorzystywane do innych celów:

* Użytkownicy i GrupyUżytkowników zawierają liczebność użytkowników oraz grup użytkowników
* Int-Scen zawiera procentowe wykorzystanie scenariuszy
* EOpBS i EOpYear zawierają liczbę operacji biznesowych w skali roku w rozbiciu na poszczególne usługi biznesowe
* DVpBS i DVpYear zawierają wolumen danych przesyłany w skali roku w rozbiciu na poszczególne usługi biznesowe

Tabela 3 Lista danych wynikowych zawiera listę najważniejszych parametrów wynikowych.

Tabela 3 Lista danych wynikowych

|  |  |  |
| --- | --- | --- |
| Nazwa parametru | Opis parametru | Przeznaczenie |
| Liczba użytkowników korzystających z usług | Zestawienie ilościowe poszczególnych elementów listy użytkowników w kontekście danej grupy (urealnione). | Opracowania uwzględniające liczbę użytkowników |
| Wskaźnik Int-Scen | Liczność użytkowników korzystających z danego scenariusza | Wskazanie realnej liczby użytkowników korzystająca z danego scenariusza |
| Wskaźnik EOpBS | Liczba rzeczywistych użytkowników danej grupy korzystających z danej usługi biznesowej | Wskazanie stopnia wykorzystania usługi biznesowej |
| Wskaźnik EOpYear | Liczba operacji biznesowych obsługiwanych przez system w kontekście całej populacji użytkowników w skali roku | Rozkład usług biznesowych na podstawie ich złożoności liczonej operacjami biznesowymi |
| Wskaźnik DVpBS | Wolumen danych przesyłany przez jednostkowe wykorzystanie usługi biznesowej przez użytkowników danej grupy | Wskazanie stopnia wykorzystania usługi biznesowej |
| Wskaźnik DVpYear | Wolumen danych przesyłany w skali roku dla usług biznesowych | Rozkład usług biznesowych na podstawie wolumeny przetwarzanych danych |
| Wolumen danych | Miara całkowitego przepływu danych | Wynik końcowy szacowania wolumetrii danych |