

Резюме

на дипломна работа на тема

Уеб базирани системи, представящи големи масиви от данни

Факултет по математика и информатика

Катедра: Информационни Технологии

Дипломант: Атанас Кънчов Гегов, факултетен номер M21704

Научен ръководител: доц. д-р Силвия Илиева, ИТ, ФМИ

Ключови думи: дизайн, данни, шаблони за проектиране, потребителски интерфейс, филтри, сортиране, сървърни процеси, визуализация на данни, странициране, бързодействие, скорост на изчисление, анализ, кодиране, тестване, профайлинг, алгоритми, структури от данни, I/O;

В тази дипломна работа, системите, които стоят на централно място и върху които се фокусира цялото внимание, са уеб базирани. Основни черти, които ги характеризират са големия брой конкурентни потребители, представянето на големи масиви от данни (които могат да се разглеждат като дълги списъци от обекти) и постоянната натовареност или големия брой активни сесии.

Направен е обзор на шаблони за проектиране, даващи актуални решения за скелета на софтуер, представящ големи масиви от данни и са предложени препоръки за употребата им. Представените шаблони за проектиране, в зависимост от характерните черти на разглежданите системи са разпределени на три типа – подходящи за управление на сложни, многослойни системи, подходящи за управление на комуникацията с носителите на данни и подходящи за постигане на висока производителност.

Изложени са техники, свързани с проектирането на визуализацията на данните във вид, който е максимално удобен за масовия потребител. Направен е анализ на принципите на дизайн на потребителския интерфейс. Като са взети предвид условията за представяне на огромното количество информация и конструиране на заявки към сървъра, пестящи машинни ресурси. Изложени са също и техники, свързани с проектиране на процесите на сървъра, грижещи се за извличането и подготовката на данните за визуализация.

Едно от най-важните свойства на софтуера, представящ големи масиви от данни е неговото бързодействие. В дипломната работа са систематизирани стратегии за подобряване на бързодействието, ориентирани към високите нива на софтуера. Стратегиите засягат процесите на анализ на изискванията на системата, обектно-ориентирания дизайн, тестване и профайлинг. Систематизирани са също и тактики за постигане на високо бързодействие, ориентирани към най-ниското ниво на софтуера – нивото на програмния код. Тактиките са представени чрез добри практики за писане на качествен високопроизводителен програмен код.

Представено е уеб базирано приложение, демонстриращо три различни подхода при извличане на големи масиви от данни – Традиционен, Page-by-Page Iterator и Value List Handler. Направен е сравнителен анализ и са обобщени резултатите от тестове при различни условия (в зависимост от броя активни сесии на сървъра и размера на данните за визуализация)