



بسم الله الرحمن الرحيم

طراحی سیستم ثبت نام دانشجویی توزیع شده

توسط

منصوره شهرکی مقدم

پایان نامه

ارائه شده به دانشکده تحصیلات تکمیلی به عنوان بخشی از فعالیت های

تحصیلی لازم برای اخذ درجه کارشناسی ارشد

~~۱۳۸۰~~

در رشته

مهندسی کامپیوتر - هوش مصنوعی و ریاتیک

از

دانشگاه شیراز

شیراز، ایران

ارزیابی و تصویب شده توسط کمیته پایان نامه با درجه: عالی

امضای اعضای کمیته پایان نامه:

دکتر احمد توحیدی، استادیار مهندسی کامپیوتر (رئیس کمیته)

دکتر محمد هادی صدرالدینی، استادیار مهندسی کامپیوتر (رئیس کمیته)

دکتر منصور ذوالقدری جهرمی، استادیار مهندسی کامپیوتر

آبان ماه ۱۳۷۹

۳۴۹۳۸

تقدیم به :

چشمان مهربان مادرم و دستان زحمت کش پدرم

و

همه آنهایی که در ناملايمات زندگي همراهي ام کردند.

سپاسگزاری

بدین وسیله از اساتید گرانقدر دکتر محمدهادی صدرالدینی و دکتر احمد توحیدی به خاطر زحمات فراوان و بی شائبه ای که در طی مراحل تهیه و تدوین پایان نامه متحمل شده اند و با راهنمایی های دلسوزانه در جهت غنای آن نقش مؤثری داشته اند، کمال تشکر و سپاس را دارم.

همچنین از اساتید بزرگوارم جناب آقای دکتر منصور ذوالقدری جهرمی، عضو کمیته، و جناب آقای دکتر سیروس جوادپور، نماینده محترم تحصیلات تکمیلی، سپاسگزارم.

بعلاوه لازم میدانم از زحمات و کمک های فراوان جناب آقای دکتر شهاب الدین آیت اللهی، دکتر غلامحسین دستغیبی فرد، جناب آقای فتح پور و سرکار خانم نبی زاده که در امر تهیه پایان نامه لطف زیادی به اینجانب داشته اند، تشکر نمایم. در نهایت از همراهی کلیه کادر بخش کامپیوتر دانشگاه شیراز و دوستان عزیزم سپاسگزارم.

چکیده

طراحی سیستم ثبت نام دانشجویی توزیع شده

توسط

منصوره شهرکی مقدم

آنچه در این پایان نامه ارائه شده است، در راستای برآوردن دو هدف اساسی است، مستندسازی سیستم ثبت نام فعلی دانشگاه شیراز و سپس ارائه طرح یک سیستم ثبت نام توزیع شده برای این دانشگاه.

سیستم ثبت نام دانشگاه شیراز شامل یک سری برنامه های کاربردی است که ارتباط کاربر را با آن برقرار میکند. قلب این سیستم شامل یک برنامه اصلی است که همچون برنامه های کاربردی آن، به زبان پایه PLI نوشته شده است و ارتباط داده ای برنامه های کاربردی را با داده های سیستم ثبت نام که در قالب یک سری فایل اطلاعاتی ذخیره شده اند، برقرار میسازد.

طبق آنچه در مستندسازی سیستم ثبت نام دانشگاه شیراز بدست آمد، افزایش تعداد کاربران، پارامتر قابل ملاحظه ای است که در سیستم متمرکز به محدودیت مواجه خواهد شد و بر عکس در یک سیستم توزیع شده استفاده از قابلیت به نام محلی سازی داده ها و اجرای برنامه های کاربردی، هر گونه محدودیتی را در این زمینه رفع میکند. به همین منظور الگوریتمی توزیع شده به فرم Top-Down جهت تطبیق خصوصیات سیستم ثبت نام به گستره ای از گره های شبکه پیشنهاد گردید.

مهمترین بخش در الگوریتم پیشنهادی قطعه قطعه سازی (fragmentation) افقی و عمودی جداول اطلاعاتی میباشد. قطعه قطعه سازی عمودی با تکنیک splitting انجام گرفت و جداول جدید که در قالب ۹ نوع مختلف گروه بندی میشوند، بین گره های

مختلف شبکه ثبت نام که تنها از سه نوع گره، **main-node**، **faculty-node** و **group-node** تشکیل گردیده، توزیع شد. سپس در عمل کارایی این الگوریتم در توزیع داده های سیستم ثبت نام دانشگاه شیراز نسبت به مدل متمرکز فعلی محک زده شد. بر اساس نتایج مشاهده شده، توزیع داده ها و ذخیره سازی هر گروه داده در گره ای که بیشتر به آن نیاز دارد، سرعت کار و عملکرد سیستم را افزایش میدهد. طبق آزمایشات عملی صورت گرفته، هر چه تعداد کاربران درگیر با عملیات ثبت نام در سیستم مربوطه افزایش یابد، برتری مدل توزیع شده نسبت به مدل متمرکز بیشتر نمایان میشود. بعلاوه هر چه تعداد درخواست ها و تقاضاهای بیشتری از سوی کاربران روانه شبکه سیستم ثبت نام گردد، ترافیک اعمال شده در مدل توزیع شده، رشد کمتری نسبت به ترافیکی خواهد داشت که در مدل متمرکز به شبکه اعمال میشود. این دو خصوصیت یافته های بسیار مهمی هستند که پیاده سازی سیستم ثبت نام دانشگاه به فرم توزیع شده را تشویق میکنند.

فهرست مطالب

صفحه	عنوان
نه	فهرست جداول
ده	فهرست اشکال
۱	فصل اول - مقدمه
۴	۱-۱- بررسی سیستم ثبت نام دانشگاه شیراز
۴	۲-۱- طراحی سیستم ثبت نام توزیع شده
۷	فصل دوم - سیستم ثبت نام دانشگاه شیراز
۹	۱-۲- ساختار سخت افزاری
۱۰	۲-۲- ساختار نرم افزاری
۱۰	۱-۲-۲- بخش مدیریت اطلاعات
۱۲	۲-۲-۲- بخش برنامه های کاربردی
۱۴	۳-۲- حوزه اختیارات کاربران
۱۷	۴-۲- شرح برنامه های کاربردی
۵۲	۵-۲- سطوح دسترسی به برنامه های کاربردی
۵۶	۶-۲- چارت ارتباط بین روتین های هر برنامه کاربردی
۵۸	۷-۲- شرح روتین های هر برنامه کاربردی
۵۹	۸-۲- تحلیل سیستم
۶۰	۱-۸-۲- دیدگاه کاربری
۶۲	۲-۸-۲- دیدگاه فنی
۶۵	فصل سوم - بررسی قابلیت های سیستم توزیع شده
۶۸	۱-۳- اجزاء سیستم
۶۹	۲-۳- خصوصیات کلی سیستم

۷۴	۳-۳- مزایای سیستم
۷۵	۴-۳- معایب سیستم
۷۷	۵-۳- قابلیت های سیستم نسبت به مدل متمرکز
۸۱	فصل چهارم - طراحی سیستم ثبت نام توزیع شده
۸۳	۱-۴- معرفی الگوریتم طراحی
۸۵	۲-۴- تحلیل نیازمندیها
۸۵	۳-۴- تشخیص تقاضاهای اصلی سیستم
۸۸	۴-۴- طراحی اولیه جداول داده
۹۱	۵-۴- نرمال سازی جداول داده
۹۶	۶-۴- تشخیص زیرتقاضاهای نهایی
۱۰۵	۷-۴- نیاز به قطعه قطعه سازی و مکان دهی
۱۰۸	۸-۴- قطعه قطعه سازی
۱۱۳	۱-۸-۴- قطعه قطعه سازی افقی
۱۱۴	۲-۸-۴- قطعه قطعه سازی عمودی
۱۲۳	۹-۴- مکان دهی
۱۲۴	۱۰-۴- آزمایش سیستم طراحی شده
۱۲۶	فصل پنجم - بررسی عملی سیستم ثبت نام توزیع شده
۱۲۶	۱-۵- ساختار مدل های متمرکز و توزیع شده
۱۳۰	۲-۵- مقایسه مدل های متمرکز و توزیع شده
۱۳۷	۳-۵- جمع بندی
۱۴۰	پیوست یک - ارتباط روتین های هر برنامه کاربردی
۱۵۵	پیوست دو - خطاهای قابل گزارش سیستم ثبت نام دانشگاه شیراز
۱۵۸	پیوست سه - کدگذاری پارامترهای اصلی برنامه های سیستم
۱۷۲	پیوست چهار - ساختار فایل های اطلاعاتی سیستم
۱۷۶	پیوست پنج - شرح روتین های هر برنامه کاربردی

صفحه

عنوان

۲۵۸

پیوست شش - هزینه اتصالات اصلی شبکه دانشگاه شیراز

۲۶۳

پیوست هفت - پیشنهاد panel های جدید سیستم ثبت نام

۲۷۸

منابع

چکیده و صفحه عنوان انگلیسی

فهرست جداول

صفحه	عنوان
۴۰	جدول ۱-۲- فیلدهای فایل PRECRD CONTROL
۴۴	جدول ۲-۲- فیلدهای فایل SANJESH CONTROL
۸۷	جدول ۱-۴- لیست تقاضاهای اصلی سیستم
۸۹	جدول ۲-۴- ساختار جداول داده سیستم ثبت نام قبل از نرمال سازی
۹۴	جدول ۳-۴- ساختار جداول داده سیستم ثبت نام پس از نرمال سازی
۹۶	جدول ۴-۴- لیست زیرتقاضاهای اصلی سیستم ثبت نام دانشگاه شیراز
۱۰۳	جدول ۵-۴- زیرتقاضاهای استفاده شده در هر تقاضای اصلی
۱۰۹	جدول ۶-۴- فیلدهای استفاده شده از جداول داده در اجرای هر زیرتقاضا
۱۱۲	جدول ۷-۴- زیرتقاضاهای اعمال شده به هر جدول داده
۱۲۲	جدول ۸-۴- جداول نهایی داده های سیستم
۱۲۴	جدول ۹-۴- مکان دهی داده های سیستم ثبت نام دانشگاه شیراز
۱۲۸	جدول ۱-۵- نحوه پاسخ گویی به تقاضاها در مدل توزیع شده
۱۳۸	جدول ۲-۵- نحوه پاسخ گویی به تقاضاها در مدل متمرکز
۱۳۶	جدول ۳-۵- سرعت عملکرد مدل توزیع شده به مدل متمرکز از نظر تئوری
۱۳۷	جدول ۴-۵- تاثیر پارامترهای R و N در سرعت عملکرد مدل توزیع شده به مدل متمرکز در عمل

فهرست اشکال

صفحه	عنوان
۹	شکل ۱-۲- ارتباطات سیستم ثبت نام دانشگاه شیراز
۵۸	شکل ۲-۲- چارت ارتباط بین روتین های برنامه DLT
۵۹	شکل ۳-۲- جدول مشخصات روتین AVGPROC
۶۸	شکل ۱-۳- اجزاء نرم افزاری سیستم بانک اطلاعاتی توزیع شده
۸۲	شکل ۱-۴- ساختار Client / Server دانشگاه
۸۴	شکل ۲-۴- الگوریتم Top-Down
۸۶	شکل ۳-۴- الگوریتم پیشنهادی
۱۱۲	شکل ۴-۴- ارتباط بین جداول داده
۱۱۶	شکل ۵-۴- ماتریس A متناظر با هر جدول داده
۱۱۸	شکل ۶-۴- ماتریس AA برای جدول داده CLS
۱۱۹	شکل ۷-۴- الگوریتم BEA
۱۲۱	شکل ۸-۴- ماتریس CA(CLS)
۱۲۱	شکل ۹-۴- ماتریس CA(CLS) مرتب شده
۱۲۷	شکل ۱-۵- مدل توزیع شده
۱۲۷	شکل ۲-۵- مدل متمرکز
۱۳۱	شکل ۳-۵- تاثیر افزایش پارامترهای R و X و N در هزینه زمانی پاسخ گویی به تقاضاهای ارسالی هر گره در مدل متمرکز
۱۳۲	شکل ۴-۵- تاثیر افزایش پارامترهای R و X و N در هزینه زمانی پاسخ گویی به تقاضاهای ارسالی هر گره در مدل توزیع شده
۱۳۳	شکل ۵-۵- تاثیر متفاوت افزایش پارامتر N در هزینه زمانی پاسخ گویی به تقاضاهای ارسالی در مدل های متمرکز و توزیع شده

صفحه

عنوان

۱۳۴

شکل ۵-۶- بررسی مدل متمرکز برای سیستم ثبت نام

۱۳۴

شکل ۵-۷- بررسی مدل توزیع شده برای سیستم ثبت نام

۱۳۵

شکل ۵-۸- مقایسه مدل های متمرکز و توزیع شده برای سیستم ثبت نام

فصل اول

مقدمه

تا سال ۱۹۶۰ متمرکز کردن داده های هر سازمان در کامپیوتری با ظرفیت حافظه بالا تنها راه پردازش داده های آن سازمان بود. اما امروزه پیشرفت های جدید در زمینه تکنولوژی بانک اطلاعاتی^۱ و شبکه های کامپیوتری و مینی کامپیوترها و میکروکامپیوترها، زمینه ساز پیدایش بانک های اطلاعاتی توزیع شده^۲ گردیده است [۱].

خصوصاً در دهه اخیر حرکت گسترده ای به سمت سیستم های توزیع شده انجام گرفته است و شرکت های زیادی از جمله Oracle در زمینه تغییر سیستم های متمرکز به سیستم های توزیع شده در تلاش و فعالیت هستند [۲].

از جمله کارهای انجام شده میتوان (SDD) System for DDBs محصول Distributed و IBM و R* محصول Computer Corporation of America و INGRESS محصول شرکت Relational Technology را نام برد.

در سیستم متمرکز (centralized)، داده ها و اطلاعات در یک بخش مرکزی مجتمع شده اند و کاربران سیستم از طریق ترمینال هایی به این بخش مرکزی دسترسی دارند. اما در سیستم توزیع شده، داده ها و اطلاعات در سایت (گره) های کامپیوتری مختلفی که از طریق شبکه به هم مرتبط هستند، پراکنده شده اند و هر مکان ضمن کنترل بانک اطلاعاتی خاص خود، قابلیت دسترسی به داده های دیگر سایت ها را هم دارد.

به این ترتیب یک برنامه کاربردی در حال اجرا روی بانک اطلاعاتی توزیع شده

^۱ Database (DB)

^۲ Distributed Database (DDB)

میتواند به داده های ذخیره شده در بیش از یک سایت دسترسی پیدا کند [۱].

در یک سیستم توزیع شده اضافه کردن کامپیوتر به سیستم بدون اینکه اثر نامناسبی روی قابلیت و کارایی سیستم بگذارد، انجام پذیر است. بعلاوه در این سیستم ها میتوان داده ها را به گونه ای توزیع کرد که هر گروهی از داده ها به کاربرانی نزدیک باشد که نیاز بیشتری به آن دارند.

مهمترین مزیت یک سیستم توزیع شده نسبت به سیستم متمرکز مسئله scalability است، به این معنا که رشد اطلاعات در یک سیستم توزیع شده کندتر و قابل تحمل تر اتفاق می افتد [۳].

در مواردی که داده ها در بانک های اطلاعاتی مختلفی وجود دارند و نوع این بانک های اطلاعاتی و سیستم مدیریت آنها متنوع است، از سیستم مدیریت بانک اطلاعاتی توزیع شده^۱ برای دسترسی به بانک های موجود، بدون نیاز به تغییر آنها، استفاده میشود. به این ترتیب میتوان همزمان، با همه این بانک ها ارتباط داشت، بدون اینکه ماهیت اولیه آنها نیاز به اعمال تغییر داشته باشد.

به دلیل همه گیر شدن شبکه برای سیستم های کامپیوتری و اهمیت توزیع اطلاعات روی این شبکه، تلاش های زیادی برای طراحی و پیاده سازی سیستم های بانک اطلاعاتی توزیع شده انجام می گیرد.

به عنوان مثال بسیاری شرکت های تجاری دارای شعبی هستند که در مکان هایی با فاصله زیاد پراکنده شده اند و براساس جمع آوری اطلاعات شعب خود عمل می کنند. به همین جهت استفاده از بانک اطلاعاتی توزیع شده بسیار مورد توجه این شرکت ها قرار گرفته است. طبق مطالعات انجام شده قابلیت پارتیشن کردن (partitioning)، مسائل authority و responsibility در مدیریت اطلاعات دلیل مهمی است که باعث میشود بسیاری از این شرکت های تجاری، سیستم های توزیع شده را ترجیح دهند [۴].

در همین راستا هدف از انجام این پایان نامه نیز بررسی توزیع پذیری داده های سیستم ثبت نام دانشگاه شیراز میباشد. به این منظور شناختی از سیستم متمرکز فعلی و

¹ Distributed Database Management System (DDBMS)

ساختار کلی سیستم های توزیع شده مورد نیاز است، تا بتوان طرح جامع و بهینه ای مناسب سیستم ثبت نام دانشگاه بصورت توزیع شده ارائه داد.

پیش از ورود سیستم های کامپیوتری به حوزه دانشگاه شیراز، کل پروسه ثبت نام، از جمله مراحل پذیرش دانشجویان جدید و ثبت نام ترمی آنان به طریق دستی انجام میگرفت، که کاری بسیار زمان گیر بود، هم برای دانشجویان و هم برای کادر آموزشی دانشکده ها و بخش ها که مسئولیت ثبت نام دانشجویان رشته های مختلف در سال های مختلف را بر عهده داشتند تا اینکه در سال ۱۳۴۷، کامپیوتر به پروسه ثبت نام راه پیدا کرد.

از ابتدای کار سیستم ثبت نام، از نظر تجهیزات کامپیوتری و داده ها و برنامه مربوطه، به فرم متمرکز عمل میکرد و همه دانشجویان جهت ثبت نام به آن مراجعه میکردند. حدوداً از سال ۷۲ سیستم ثبت نام دانشگاه شیراز فرمی نیمه توزیعی به خود گرفت، به این ترتیب که کل داده های مربوط به ثبت نام و برنامه های آن بر روی یک سیستم مرکزی متمرکز شد و دانشکده ها و بخش های مختلف از طریق ترمینال هایی با این سیستم مرکزی ارتباط برقرار کردند. کل ساختار دانشگاه شیراز شامل ۱۱ دانشکده و ۷۵ بخش و چندین ساختمان اداری است که همه در روزهای ثبت نام به سیستم مرکزی متصل شده و انتظار اجرای سریع تقاضا (query) های درخواستی خود را دارند. طبیعی است که ترافیک سنگینی به سیستم مرکزی اعمال میشود که سرعت پاسخ گویی به تقاضا های مربوط به کاربران مختلف را تا حد زیادی کاهش میدهد.

در جهت رفع ترافیک اعمال شده به سیستم مرکزی ثبت نام، خصوصاً در روزهای ثبت نام، فرم توزیع شده برای این سیستم پیشنهاد گردید. در جهت تحقق بخشیدن به این مهم، در این پایان نامه دو کار اصلی دنبال میشود:

- بررسی سیستم ثبت نام دانشگاه شیراز

- طراحی سیستم ثبت نام توزیع شده

هر کدام از این موارد در بخش های جداگانه بررسی شده اند اما در ابتدای کار شرح خلاصه ای از مراحل کاری هر کدام دور از فایده نیست.

۱-۱- بررسی سیستم ثبت نام دانشگاه شیراز

بخش اول این پایان نامه روی مسئله مستندسازی سیستم ثبت نام فعلی تکیه دارد. سیستم ثبت نام فعلی دارای یک بخش مرکزی است شامل یک برنامه اصلی و فایل های داده، که کل اطلاعات دانشگاه را در خود جای داده است و ترمینال های سیستم که از طریق برنامه های کاربردی با این بخش مرکزی ارتباط برقرار میکنند.

در این بخش نحوه عملکرد برنامه های کاربردی و ارتباط روتین های مختلف این برنامه ها (پیوست یک) مورد بررسی قرار میگیرد و بعلاوه شرح مختصری در مورد هر روتین (پیوست پنج) ارائه میشود و همین طور لیست خطاهای قابل گزارش در سیستم ثبت نام فعلی (پیوست دو) و کد مربوط به پارامترهای اصلی برنامه های مختلف سیستم (پیوست سه) و ساختار فایل های داده (پیوست چهار) ارائه میشود. همچنین مشکلات کاری و نقاط ضعف سیستم بررسی خواهد شد. بدین ترتیب ضمن مستندسازی سیستم، جزئیات کاری آن مشخص شده و پس از اینکه درک واضح و روشنی از ساختار و عملکرد سیستم فراهم آمد، در بخش بعدی که تاکید بر طراحی یک سیستم ثبت نام توزیع شده است، مراحل طراحی با اطمینان طی می گردد.

۱-۲- طراحی سیستم ثبت نام توزیع شده

بخش دوم این پایان نامه روی مسئله طراحی توزیع شده تکیه دارد. در این بخش سعی شده ساختاری مناسب جهت ایجاد زیر بنای فیزیکی یا به عبارتی شبکه سازی سیستم ثبت نام ارائه شود. مسائل نرم افزاری مربوط به بانک های اطلاعاتی توزیع شده بررسی گردد و سپس مدلی برای تخصیص داده و عملیات ثبت نام ارائه شود.

در این بخش مسئله بانک های اطلاعاتی توزیع شده مورد بحث قرار گرفته است. در حقیقت ساختار کلی بانک های اطلاعاتی توزیع شده و مزایا و معایب توزیع شدگی بررسی شده است.

برای پیاده سازی یک سیستم بانک اطلاعاتی توزیع شده مثل سیستم مورد بحث در این پایان نامه، انتخاب نرم افزار برنامه نویسی و انتخاب بانک اطلاعاتی مناسب، از اهمیت ویژه ای برخوردار است که این انتخاب بر اساس پارامترهایی انجام خواهد شد