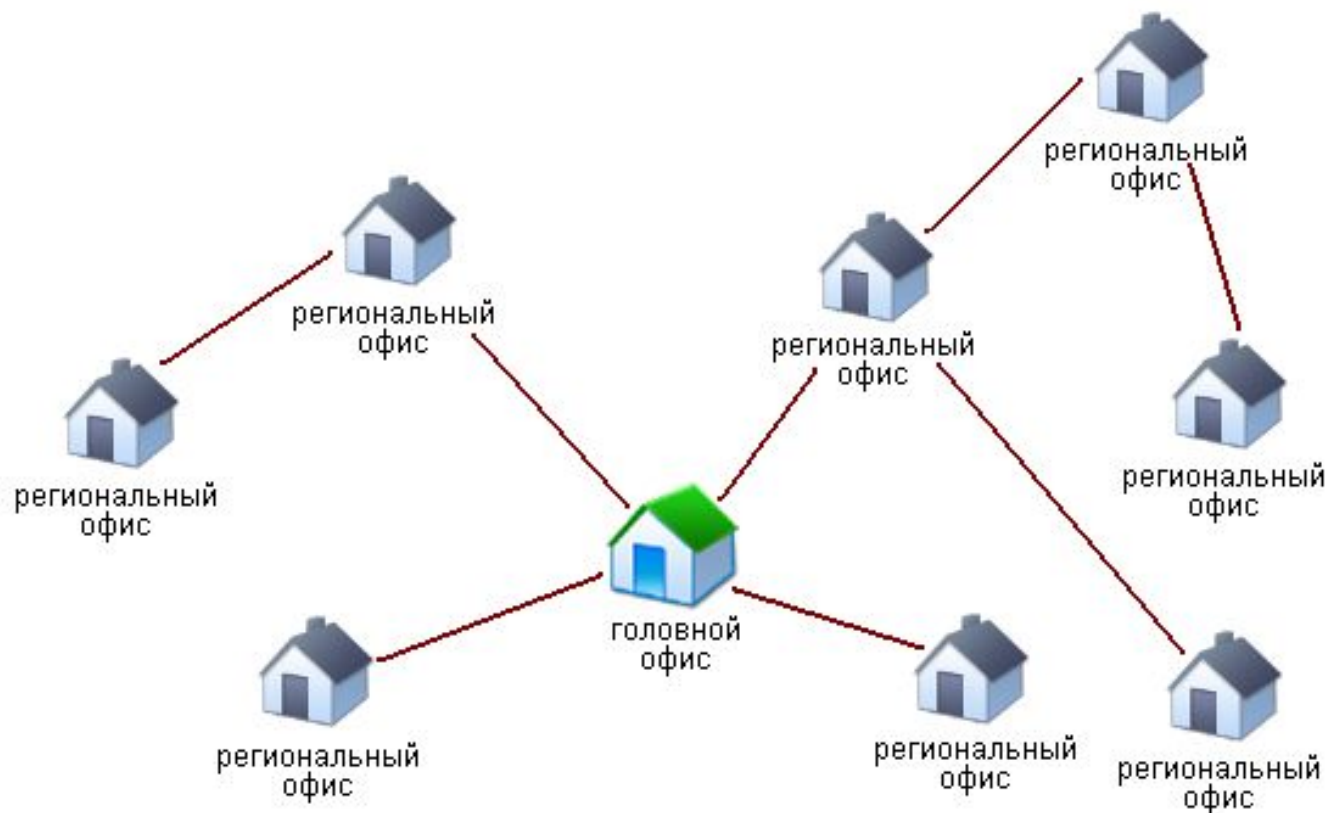


Территориально распределенная ПОИСКОВАЯ СИСТЕМА

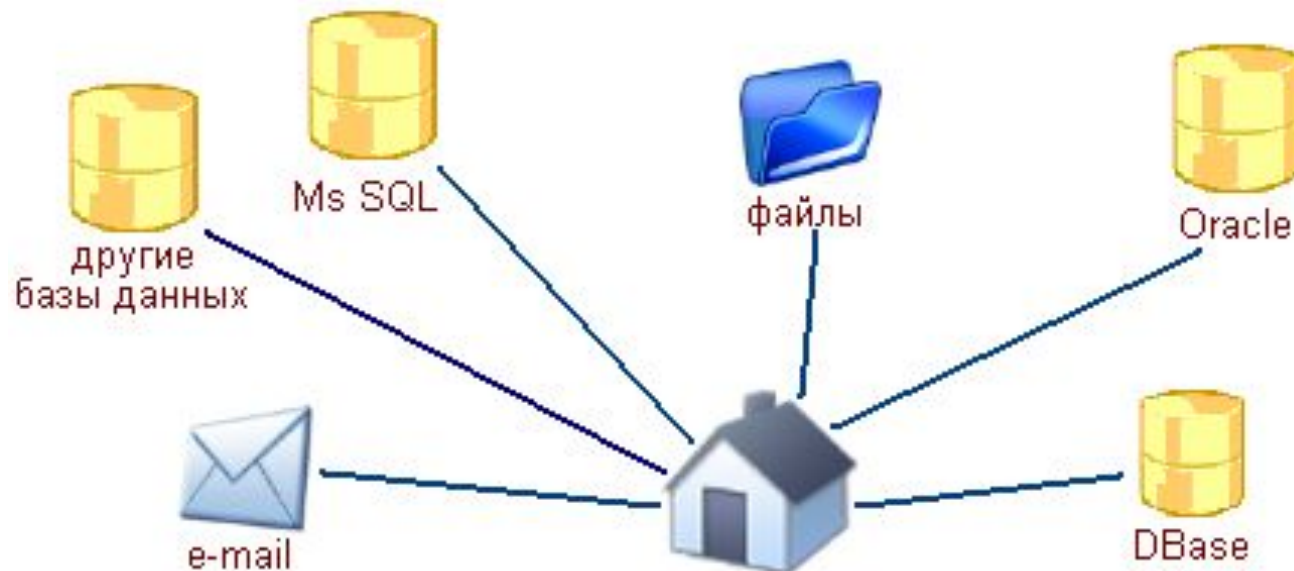
Трудно переоценить роль поисковых систем в корпоративных сетях. Количество информации в них с каждым годом увеличивается, и поиск нужной информации становится проблемой.



А если это не просто предприятие, пусть и огромное, а целая структура, причём территориально распределённая?



В территориально распределенной структуре компоненты рассредоточены по огромной территории, и чёткой связи между информационными составляющими нет.



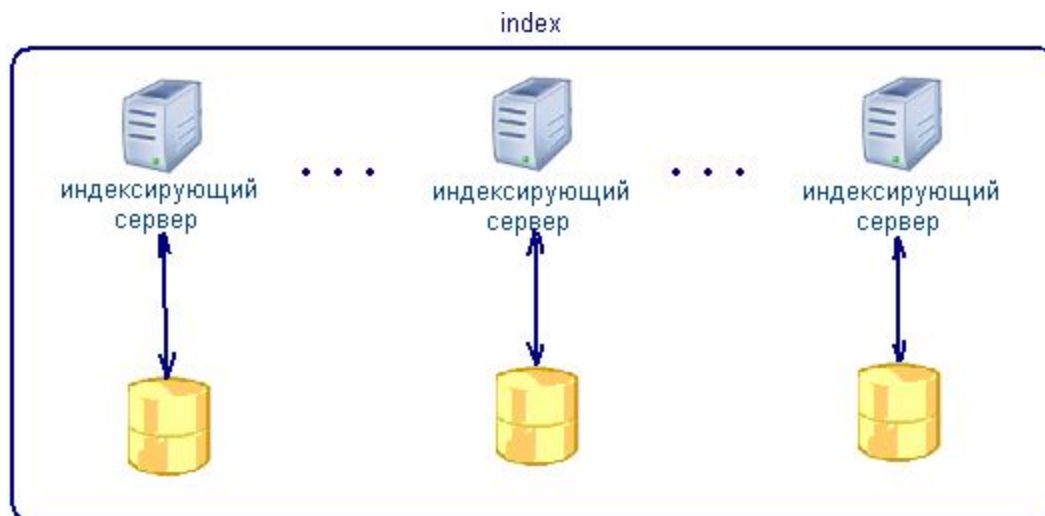
Информация, находящаяся в каждом сегменте упорядочена по-своему и хранится в своём формате. При создании этих информационных баз не было чётких стандартов, и как результат – множество разных баз данных от разных разработчиков.

Типовые проблемы, возникающие у крупных компаний и структур:

- ➔ Растущее кол-во обрабатываемой информации
- ➔ Увеличение кол-во запросов и сотрудников

Увеличение количества данных

- ➔ Количество информации на предприятии растет каждый день. Появляются новые документы, растёт объём баз данных.



Один сервер может проиндексировать 1-5 терабайт данных в зависимости от их формата. Это на порядок больше, чем у большинства конкурентов. Увеличивая число индексирующих серверов мы можем обработать любое количество информации.

Рост количества запросов

- ➔ Если предприятие расширяется, и появляются новые сотрудники, то системе придётся обрабатывать большее количество запросов.

Для решения этой проблемы достаточно увеличить количество поисковых серверов в сети. Основным преимуществом масштабирования в поисковой системе от СофтИнформ является то, что оно осуществляется без изменения структуры существующей сети.

Трёхуровневая иерархия сети

Для примера мы приводим трёхуровневую структуру, хотя количество уровней может быть и большим. Условно обозначим уровни как:



Центр

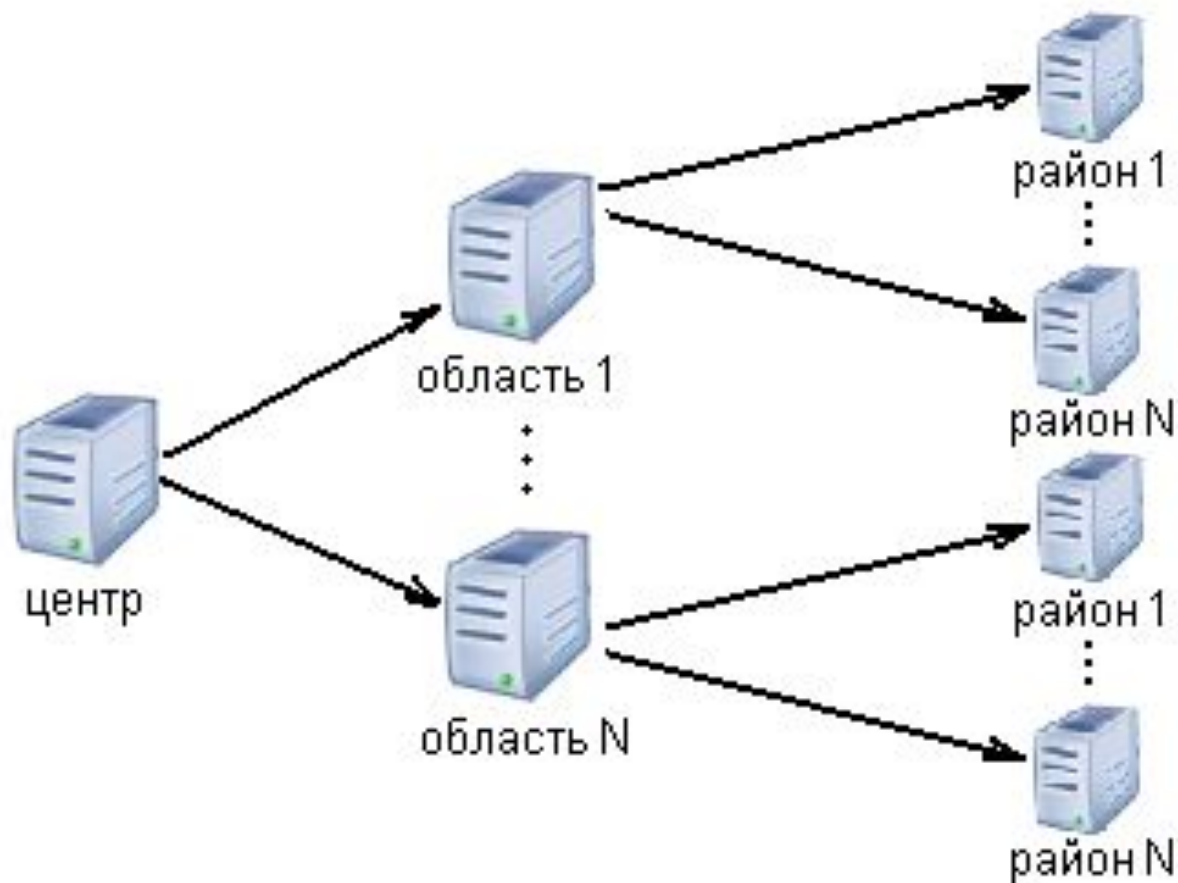


Областной уровень

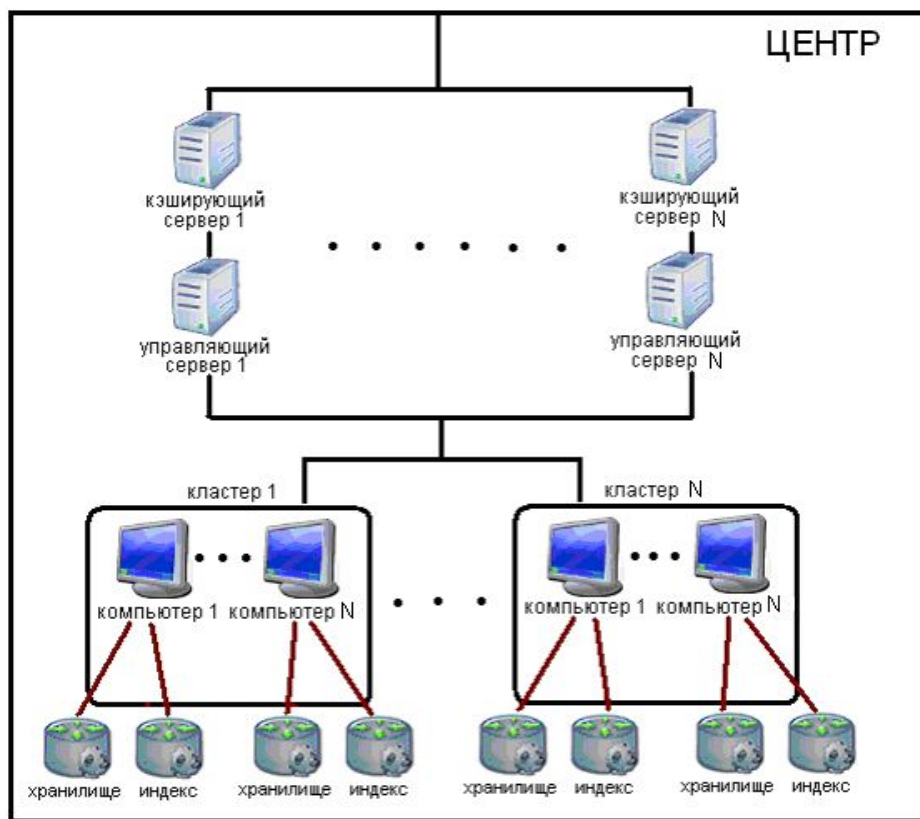


**Районный
уровень**

В одном из проектов, структура состояла из 4000 территориально распределённых офисов. Данные были расположены в разных базах данных, примерное количество которых оценивается в 16000. Проблема поиска по такой структуре нашими технологиями с успехом решается.

Схематичное изображение сети

Головной центр отвечает за общий сбор данных. Он проводит сбор измененной информации со всех областных центров.

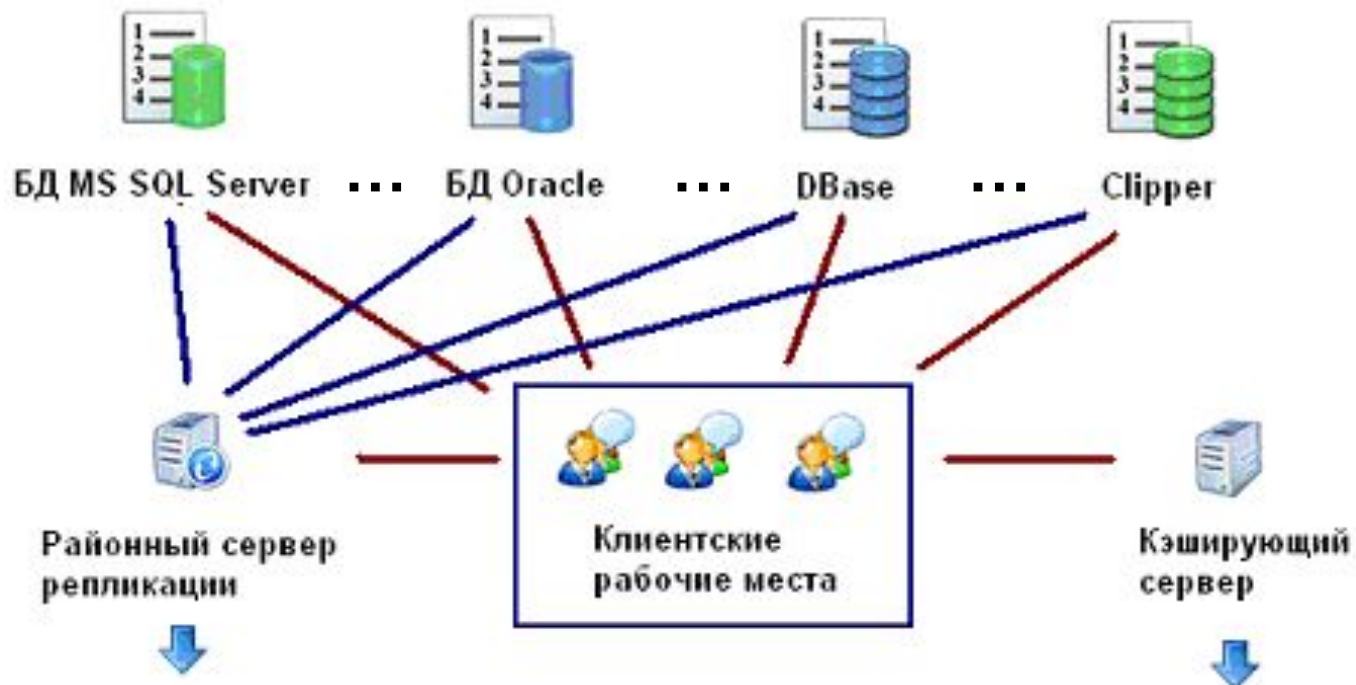


Все поисковые запросы так или иначе направляются к нему, если кэширующий сервер не может предоставить нужной информации.

Областной центр собирает всю информацию с районов и отправляет её в центр.

Основной компонент – система поиска, которая обрабатывает запросы пользователей на поиск информации.





Районный сервер репликации обеспечивает доступ к базам данных из районного центра. Он имеет связь с областным центром для передачи данных в областной центр репликации данных. Также есть клиентские места, с которых идут поисковые запросы к областной базе данных.

Обновление данных

Система репликации отвечает за обновление информации по всей сети от районного уровня и выше. Во всех узлах сложной структуры по заданному графику обновляется информация, синхронизируя между собой все источники данных.

И если даже в имеющихся базах данных нету нормального механизма обновления, с которым при обновлении пересылается только новая информация, то мы берем устранение этой проблемы на себя.

Причем пересылаться обновлённая информация будет автоматически, в заданное время. Пересылка же всей базы отнимет слишком много времени.



Разграничение прав доступа – соблюдение политики информационной безопасности и управление доступом пользователей.



Система кэширования – лишь 20% запросов являются уникальными, поэтому возможность кэширования снижает нагрузку на поисковые серверы.



Система мониторинга и диагностики – инструментарий позволяет производить гибкую настройку и перенастройку центров обработки информации.



Компания **SearchInform Technologies Inc** является разработчиком поисковой системы, на базе которой можно организовать легко масштабируемую информационно-поисковую систему любой сложности. Подтверждением являются многочисленные поисковые проекты, реализованные и реально работающие.

Многолетний опыт успешной работы на российском рынке является лучшим гарантом того, что система не подведет, обеспечив быстрый и чёткий поиск по огромным массивам данных. Опыт в консолидации информации из различных источников решит проблему разрозненности данных, собрав их в единую систему.

Приплюсовав сюда уникальную технологию поиска документов похожих по содержанию на текст запроса и корректную поддержку русской морфологии, мы получаем в итоге **лучшее поисковое решение для корпоративного сектора.**

Спасибо за внимание
ВОПРОСЫ?

