

ИННОВАЦИОННОЕ ПРИМЕНЕНИЕ РАСПРЕДЕЛЕННЫХ КЛАСТЕРОВ БОЛЬШИХ ДАННЫХ И ТЕХНОЛОГИИ MAPREDUCE ЛИ ХУНЯН

Белорусский государственный университет, г. Минск
a870406667@gmail.com

Annotation. Cloud computing is a crucial technology trend that offers cost-effective virtualization of computational resources. It's ideal for processing large datasets, but traditional parallel computing methods like MapReduce were designed for static environments. This study explores how implementing MapReduce in cloud computing can significantly boost performance and efficiency, improving big data analysis capabilities in the cloud.

Разнородные и сложные знания в индустрии облачных вычислений создают сложности в их согласовании между членами альянса, влияя на точность и полноту этого согласования. Для решения этих проблем мы предлагаем процесс сопоставления знаний для отраслевых альянсов облачных вычислений и разрабатываем усовершенствованный метод сопоставления, основанный на технологии MapReduce, для всесторонней оценки семантической близости. Наш метод включает неиерархические связи в расчет семантической близости, обеспечивая более полное соответствие знаний в отраслевых альянсах облачных вычислений.

С развитием цифровых технологий и интернета, объем данных стремительно растет, что приводит к дефициту ресурсов для их обработки. Управление таким огромным объемом данных на ограниченном количестве серверов хранения и вычислительных серверов становится невозможной задачей. Поэтому актуальной проблемой является разработка механизмов для распределенного совместного использования ресурсов и вычислительных мощностей для справления с быстрым ростом объема данных в интернете. В этом контексте возникла концепция облачных вычислений.

Облачные вычисления развиваются из параллельных, распределенных и Grid-вычислений, и одной из ключевых технологий в этой области является MapReduce. Она предоставляет эффективное решение для обработки данных в параллельных системах и позволяет работать с большими объемами данных. Были проведены исследования и улучшения метода MapReduce, включая модель HPMR для высокопроизводительных вычислений. Однако остаются пробелы в исследованиях, особенно в оптимизации работы алгоритма MapReduce.

Список использованных источников

1. Орехов, С. Е. Технологии облачных вычислений в интегрированных системах управления / С. Е. Орехов, Д. П. Артамонов, С. А. Иванов // Инфокоммуникационные технологии. – 2020. – Т. 18. – № 4. – С. 477–484.
2. Холод, И. И. Интеллектуальная обработка данных / И. И. Холод, М. С. Куприянов // Мягкие измерения и вычисления. – 2019. – № 3. – С. 55–75.

3. Жаткина, Д. О. Платформа hadoop для транспортной компании: преимущества, ограничения и экосистема / Д. О. Жаткина // Корпоративное управление экономической и финансовой деятельностью на железнодорожном транспорте. – 2019. – С. 118–122.

深入浅出 ZOOKEEPER 之 ZOOKEEPER 架构以及特性

李洪阳

白俄罗斯国立大学

a870406667@gmail.com

Annotation. Driven by the development of science and technology, applications and system architectures are changing: from single-machine single architecture to multi-machine distributed architecture. This article introduces distributed systems to solve load and faults. ZooKeeper maintains data consistency, monitors status, and assists cluster management.

分布式系统是一个硬件或软件组件分布在不同的网络计算机上,彼此间仅仅通过消息传递进行通信和协调的系统,一群互相独立计算机集合共同对外提供服务(图 1)。

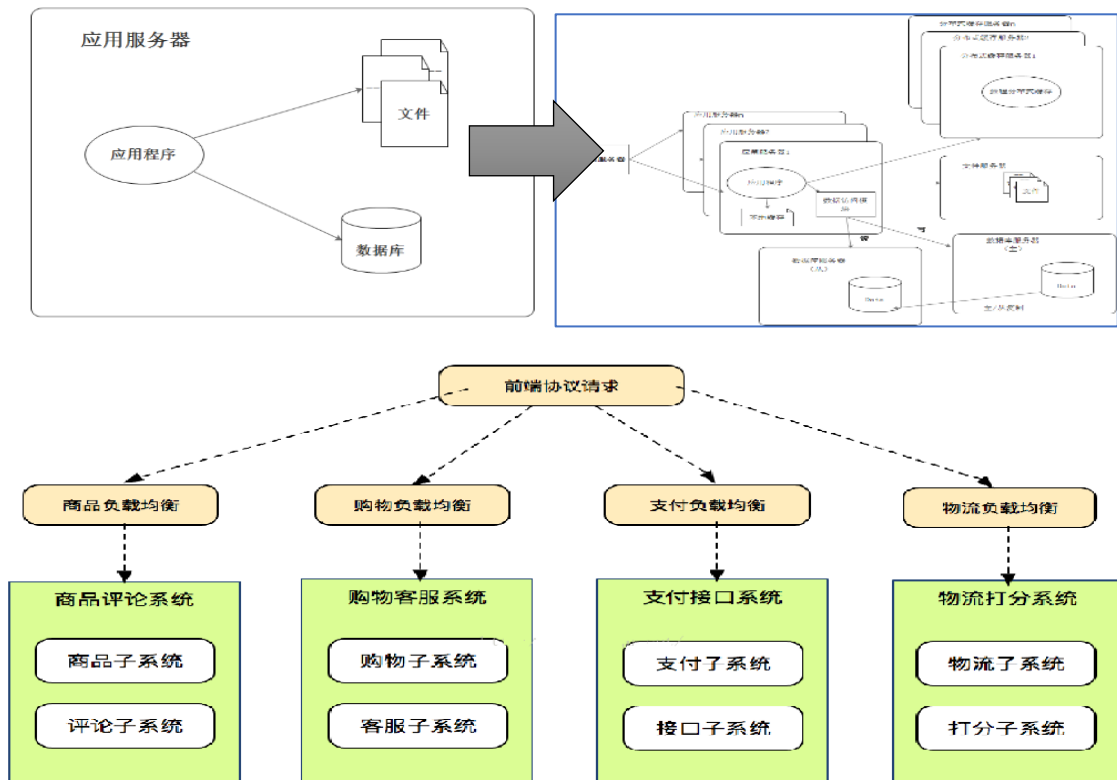


图 1 – 分布式系统是