

КАЗАХСКИЙ НАЦИОНАЛЬНЫЙ УНИВЕРСИТЕТ ИМЕНИ АЛЬ-ФАРАБИ

А.Т. Агишев

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ ОБРАБОТКИ ДАННЫХ

**Сборник лекций для студентов магистратуры,
обучающихся по образовательной программе
«7М07125 - Электроника и системы управления»**

Алматы, 2025

Лекция 6. Архитектура современных серверных систем и вычислительных кластеров

Цель лекции

Рассмотреть архитектуру современных серверных систем и вычислительных кластеров, принципы их организации, взаимодействие компонентов и подходы к масштабируемым вычислениям, применяемым для обработки больших массивов данных и реализации интеллектуальных систем.

Основные вопросы:

1. Назначение и особенности серверных систем.
2. Многоуровневая структура вычислительной архитектуры.
3. Основные компоненты серверов и их функции.
4. Кластерная организация вычислений.
5. Типы кластеров и принципы их построения.
6. Масштабируемость, отказоустойчивость и производительность.
7. Роль кластеров в реализации интеллектуальных систем обработки данных.

Краткие тезисы:

1. Назначение серверных систем. Современные серверы обеспечивают централизованную обработку, хранение и передачу данных. В отличие от персональных компьютеров, серверы оптимизированы под непрерывную работу, высокую надёжность, параллельную обработку задач и масштабируемость. Они служат базой для вычислительных центров, дата-центров и облачных платформ.

2. Архитектурные уровни серверной системы. Архитектура современных серверов включает три основных уровня:

- **Аппаратный уровень** - процессоры, оперативная память, системы хранения, коммуникационные интерфейсы.
- **Системное программное обеспечение** - операционные системы, гипервизоры, драйверы, системы управления ресурсами.
- **Прикладной уровень** - программные решения для обработки данных, вычислений и обслуживания пользователей.

Такое разделение обеспечивает гибкость и возможность масштабирования при росте вычислительных нагрузок.

3. Основные компоненты сервера.

- **Процессор (CPU):** центральный элемент, выполняющий операции и управляющий потоками данных. Современные процессоры имеют многоядерную архитектуру и поддержку параллельных вычислений.
- **Оперативная память (RAM):** временное хранилище данных для быстрого доступа.

- **Система хранения (Storage):** SSD, NVMe и сетевые хранилища (SAN, NAS) для долговременного хранения данных.
- **Сетевые интерфейсы (NIC):** обеспечивают соединение серверов между собой в кластер или сеть.
- **Системы охлаждения и питания:** поддерживают стабильную работу при высоких нагрузках.

4. Кластерная организация вычислений. Вычислительный кластер - это совокупность серверов (узлов), объединённых высокоскоростной сетью и работающих как единая система. Каждый узел выполняет часть общей задачи, а результаты объединяются для получения общего решения. Кластерная архитектура обеспечивает высокую производительность, масштабируемость и отказоустойчивость.

5. Типы вычислительных кластеров.

- **Вычислительные (Compute clusters):** предназначены для параллельных расчётов и моделирования.
- **Хранилищные (Storage clusters):** оптимизированы для обработки и хранения больших объёмов данных.
- **Гибридные:** объединяют вычислительные и хранилищные узлы в единую систему.
- **Кластеры высокой готовности (High-Availability):** обеспечивают непрерывность работы сервисов при отказе узлов.

Архитектура кластеров может быть симметричной (все узлы равнозначны) или иерархической (наличие управляющего узла и рабочих нод).

6. Принципы масштабируемости и отказоустойчивости.

Масштабируемость означает возможность увеличения вычислительных ресурсов без изменения базовой архитектуры. Отказоустойчивость достигается за счёт резервирования, распределения данных и дублирования критических компонентов. Для управления нагрузкой применяются балансировщики и планировщики задач (например, Slurm, Kubernetes). Кластер может динамически перераспределять задачи, если один из узлов выходит из строя.

7. Высокопроизводительные вычисления (HPC) и их роль.

Современные интеллектуальные системы, использующие нейронные сети, эволюционные алгоритмы и анализ больших данных, требуют огромных вычислительных ресурсов. Кластеры HPC позволяют выполнять такие задачи за счёт:

- параллельного выполнения вычислений;
- эффективного обмена данными между узлами;
- использования GPU и специализированных ускорителей;
- распределённого хранения данных с высокой пропускной способностью.

Такая архитектура обеспечивает сокращение времени обучения моделей, повышение точности анализа и реализацию сложных систем моделирования и прогнозирования.

Вопросы для контроля, изучаемого материал:

- 1) В чём заключается основное назначение серверных систем?
- 2) Перечислите ключевые компоненты серверной архитектуры и их функции.
- 3) Что представляет собой вычислительный кластер и какие преимущества он имеет?
- 4) Назовите основные типы кластеров и области их применения.
- 5) В чём заключается принцип масштабируемости вычислительных систем?
- 6) Почему кластеры являются неотъемлемой частью интеллектуальных систем обработки данных?

Рекомендуемый список литературных источников:

1. Hopgood A. A. Intelligent Systems for Engineers and Scientists. — 3rd ed. — Boca Raton : CRC Press / Taylor & Francis, 2012. — 682 p.
2. Sterling T., Anderson M., Brodowicz M. High Performance Computing: Modern Systems and Practices. — Amsterdam : Elsevier / Morgan Kaufmann, 2017. — 728 p.
3. Russell, S., Norvig, P. Artificial Intelligence: A Modern Approach. 4th Edition. Pearson, 2021.