

Распределенные системы и технологии.

Введение в предмет

Дмитрий Юрьевич Чалый
декан факультета ИВТ,
зав. кафедрой информационных и сетевых технологий



22 мая 2016 г.

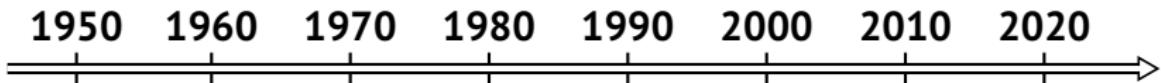
О чем этот курс?

Мы будем изучать основные концепции распределенных систем, которые используются в:

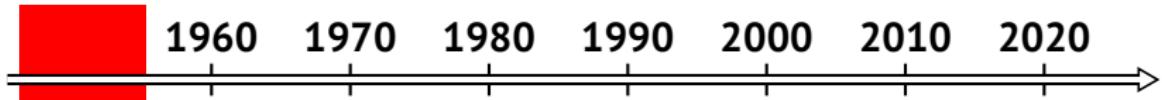
- облачных системах;
- распределенных системах хранения данных;
- распределенных вычислительных системах (но это курс не только о производительности, сколько о корректности).



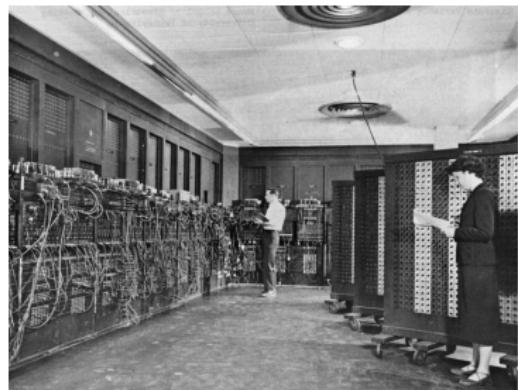
История развития распределенных систем



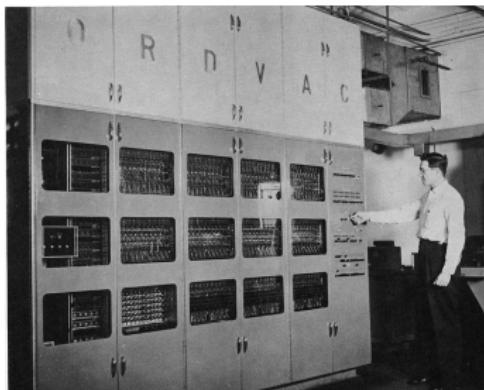
История развития распределенных систем



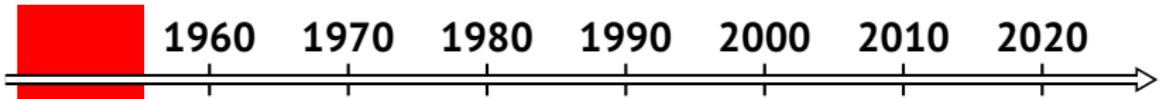
Первые датацентры:



ENIAC



История развития распределенных систем



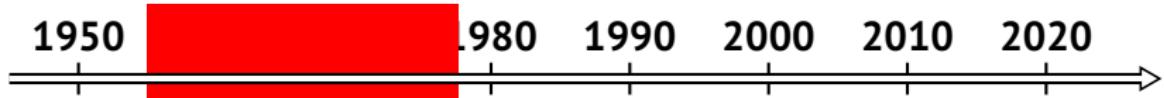
Первые dataцентры:



БЭСМ-1



История развития распределенных систем



Индустрия обработки данных:

- 1965 г.: \$2,4–2,8 млрд.;
- БЭСМ-6 (10 МГц, около 1 млн. операций в секунду), ОС ДИСПАК;
- Time-sharing: аренда терминала, оплата времени соединения, процессорного времени, Кб-месяцы дискового пространства;
- Honeywell, IBM, DEC, Xerox.



История развития распределенных систем



Новые виды распределенных систем:

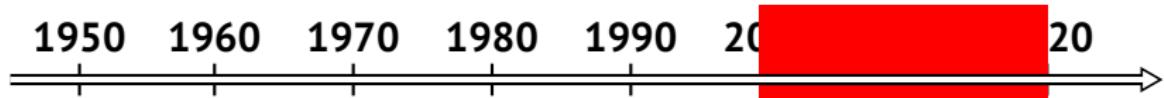
- суперкомпьютеры;
- кластеры;
- grid-системы.



Суперкомпьютер «Ломоносов»
Основан в 1803 году

ИМ. П.Г. ДЕМИДОВА

История развития распределенных систем



Новейшие классы распределенных систем:

- p2p-системы;
- облачные системы.



Технологические тренды

- Экспоненциальный рост производительности:
 - удвоение скорости процессоров: каждые 18 месяцев;
 - удвоение объема дискового пространства: каждые 12 месяцев;
 - удвоение полосы пропускания каналов связи: каждые 9 месяцев.
- Новые вычислительные задачи и потребности пользователей.



Характеристики современных распределенных систем

1 Масштаб (massive scale):

- Amazon: 21 тыс. (июль, 2010), 450 тыс. (июль, 2015);
- Facebook: 180 тыс. (август, 2012);
- Google: около 1 млн.;
- Akmai Technologies: 138 тыс. (ноябрь, 2014);
- Intel: 86 тыс. (ноябрь, 2014);
- eBay: 64 тыс. (ноябрь, 2014).

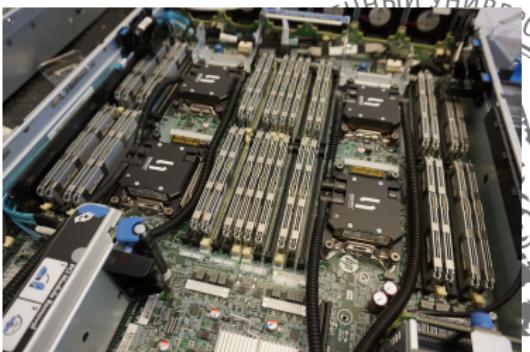
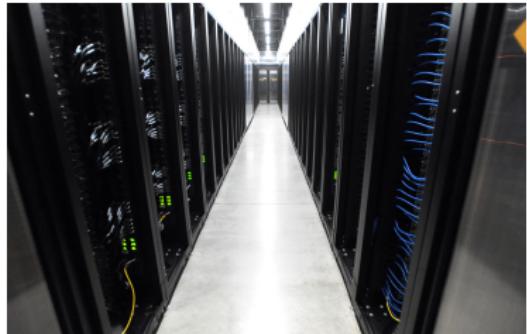
2 Доступ по требованию (on-demand access).

3 Фокус на обработку больших массивов данных (data-intensive).

4 Новые архитектурные парадигмы: MapReduce/Hadoop, NoSQL/Cassandra.



Датацентры



САНКТ-ПЕТЕРБУРГСКИЙ УНИВЕРСИТЕТ ИМ. П.Г. ДЕМИДОВА

Основан в 1803 году

Датацентры – охлаждение и питание

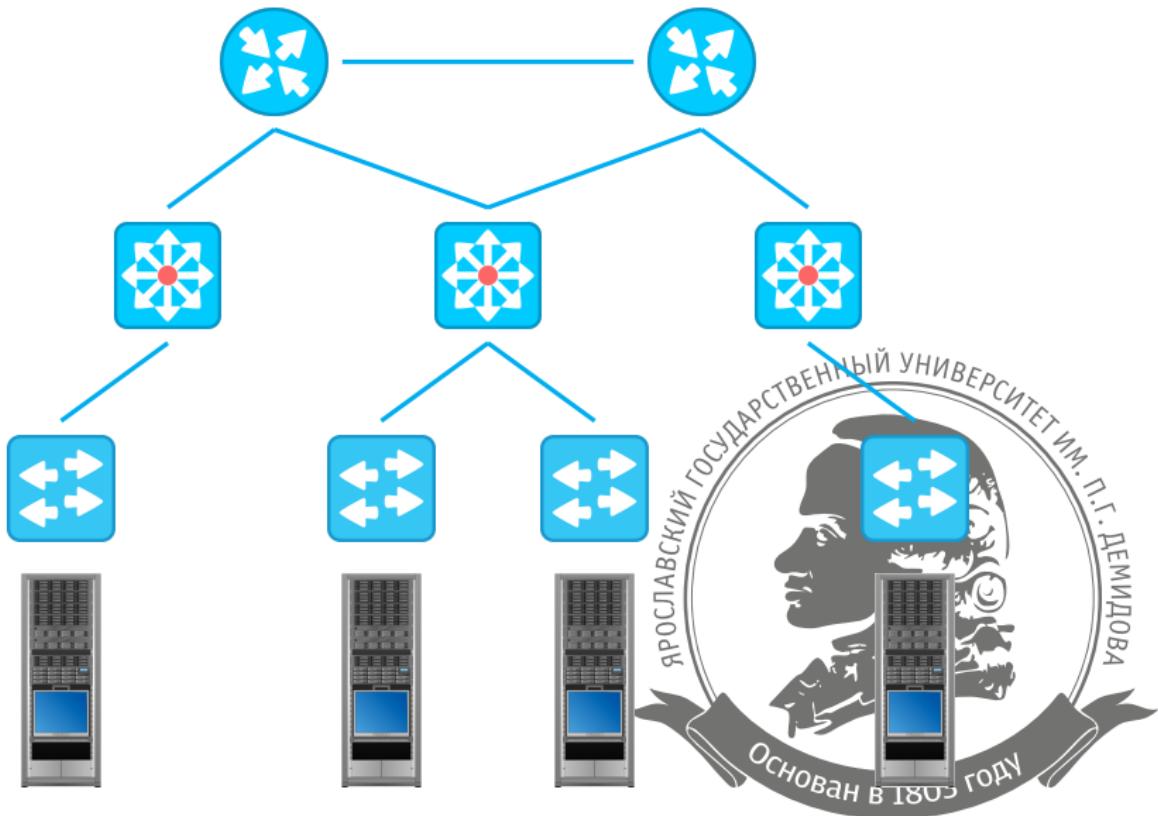


- Power Usage Effectiveness (PUE);
- $PUE = \frac{\text{Total Facility Energy}}{\text{IT Equipment Energy}}$;
- off-site и on-site;
- резервные источники.

- Water Usage Effectiveness (WUE);
- $WUE = \frac{\text{Annual Water Usage}}{\text{IT Equipment Energy}}$;
- охлаждение, влажность.



Пример топологии датацентра



Дополнительные материалы

- ➊ The ryg blog. Networks all the way down ([часть 1](#) и [часть 2](#)).
- ➋ Microsoft GFS Datacenter Tour.
- ➌ Time Lapse of Data Center Construction - Inside.

