ЛЕКЦИЯ 6

Проверка технической гипотезы

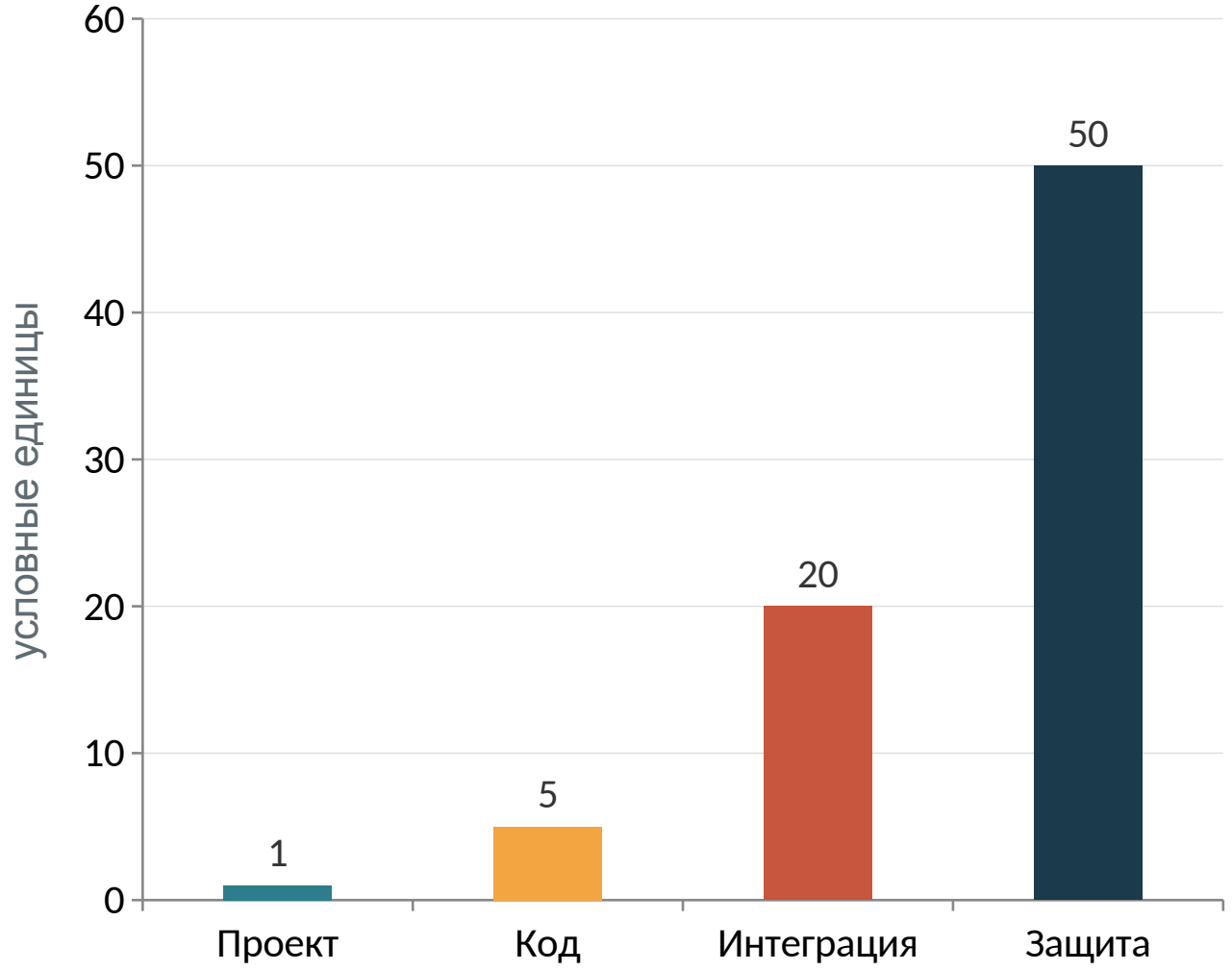
CSE842 Capstone Project
ОП 6B06102 Computer Science



План лекции

1. Зачем
2. Выбор риска
3. Гипотеза
4. Виды
5. План
6. Метрики
7. Статистика
8. Протокол

Стоимость исправления растёт



- на бумаге — часы
- после интеграции — недели
- спираль Боэма: риски раньше
- эксперимент на неделе 6

Какое решение проверять



Неопределённость

команда этого не делала



Последствия

без него нет сценария



Чужие слова

документация, форум, статья

LabBook: проверка пересечений при 100 тыс. броней и параллельных запросах

Что проверять в разных проектах

Тип проекта	Что проверить
Веб-сервис	скорость запросов, параллельность
Мобильное	офлайн, синхронизация, батарея
ML-система	метрика на реальных данных
IoT	дальность связи, питание
Чат-бот	поиск ответа, лимиты API

Шаблон технической гипотезы

Если	При	Метрика	Порог	Критерий
составной индекс	100 тыс. броней, 500 приборов	p95 времени проверки	< 1 мс	до запуска

H0: решение не обеспечивает порог; H1: обеспечивает

Виды технических экспериментов

Spike

1–3 дня, код в архив

Proof of concept

работает ли в принципе

Бенчмарк

время операции

Нагрузочный тест

Locust, k6, JMeter

A/B-сравнение

одни данные и условия

Юзабилити

поведение людей

План эксперимента

Независимая
наличие индекса

Зависимая
p95, мс

Контроль
данные, ПО, seed

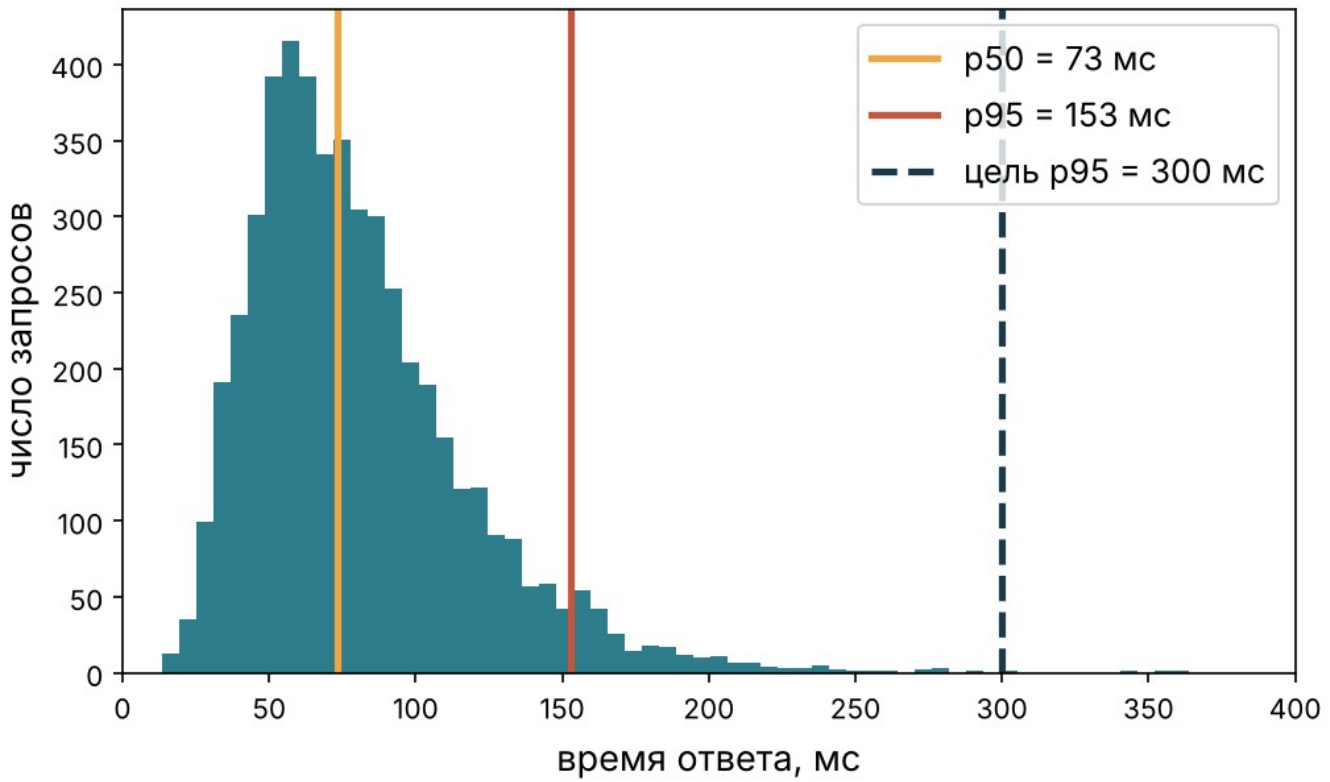


Данные
100 тыс. броней

Процедура
прогрев, 2000 повторов

Критерий
p95 < 1 мс

Перцентили вместо среднего



- p50 — медиана
- p95 — 95% быстрее
- хвост виден на графике
- среднее скрывает выбросы

Правила измерения времени



perf_counter

монотонный таймер



Прогрев

первые запуски дольше



Повторы

сотни и тысячи



Оборудование

как у целевой среды

Эксперимент: индекс SQLite

l06_spike.py

```
import sqlite3, random, time, statistics as st
db = sqlite3.connect(":memory:")
db.execute("CREATE TABLE b(eq INT, s INT, e INT)")
random.seed(1)
rows = [(random.randint(1, 500), t, t + 2)
        for t in random.sample(range(10**6), 100_000)]
db.executemany("INSERT INTO b VALUES (?,?,?)", rows)
Q = "SELECT 1 FROM b WHERE eq=? AND s<? AND ?<e LIMIT 1"

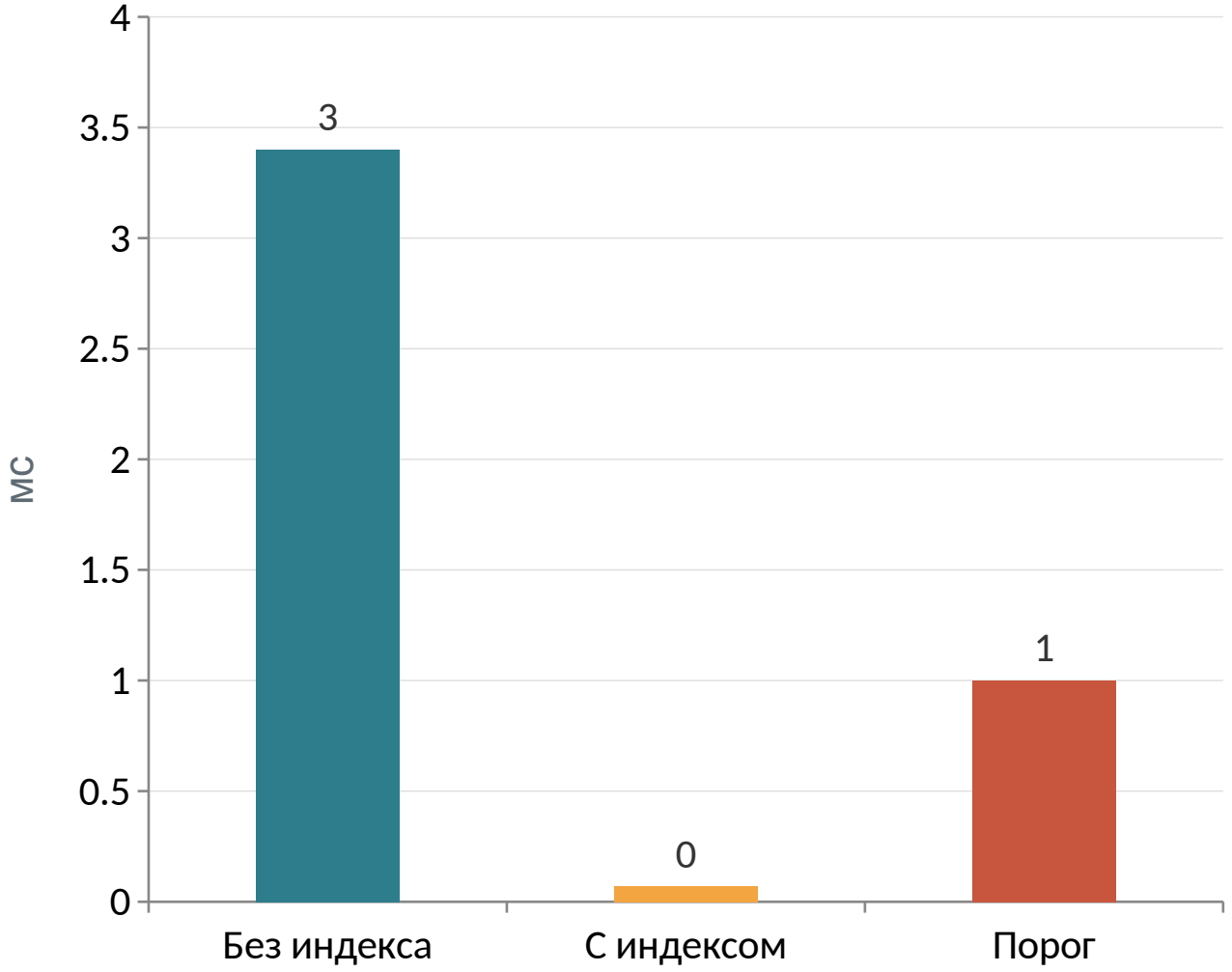
def p95(n=2000):
    ts = []
    for _ in range(n):
        eq, s = random.randint(1, 500), random.randint(0, 10**6)
        t0 = time.perf_counter(); db.execute(Q, (eq, s+2, s)).fetchall()
        ts.append((time.perf_counter() - t0) * 1000)
    return st.quantiles(ts, n=20)[18]

print(f"без индекса p95 = {p95():.3f} мс")
db.execute("CREATE INDEX ix ON b(eq, s)")
print(f"с индексом  p95 = {p95():.3f} мс")
print("H1: с индексом p95 < 1 мс:", p95() < 1)
```

```
$ python l06_spike.py
без индекса p95 = 3.377 мс
с индексом  p95 = 0.068 мс
H1: с индексом p95 < 1 мс: True
```

- без индекса ≈ 3,4 мс
- с индексом ≈ 0,07 мс
- H1 подтверждена
- запас ≈ 14 раз

Результат: p95, мс



- одна переменная — индекс
- данные и seed одинаковы
- база в памяти — угроза
- второй эксперимент: 20 потоков

Угрозы достоверности



Внутренняя

- фоновые процессы
- кэш, нет прогрева
- ошибка измерений



Внешняя

- 100 записей вместо 100 тыс.
- мощный ноутбук
- нет сети и потоков

Сводка экспериментов LabBook

Эксперимент	Результат	Вывод
E1: индекс, p95	0,07 мс < 1 мс	подтверждена
E2: 20 потоков	ровно 1 бронь	подтверждена
E3: SMTP и outbox	ответ не ждёт	подтверждена
E4: хостинг 50 RPS	p95 = 410 мс	опровергнута

Опровержение E4 — ценный результат: смена сервера или NFR

Протокол эксперимента

Решение	Гипотеза	План	Данные	Вывод	Решение
и связь с риском	и критерий	условия, версии	сырые и графики	и угрозы	для проекта

Вопросы для самопроверки

1. Зачем проверять рано
2. Выбор решения
3. Шаблон гипотезы
4. Spike и PoC
5. Переменные
6. Перцентили
7. Угрозы
8. Протокол



Неделя 6: протокол эксперимента (3 балла)