

Оптимизация процессов без сожалений

→ → → → → → →
BPMN → очереди → симуляция → деньги

$$L = \lambda \times W$$

$$\rho = \frac{\lambda}{c \times \mu}$$

$$W_q \approx \frac{\rho}{1 - \rho} \times \tau$$

$$\frac{M}{M/C}$$



STORM

$$\sqrt{\rho}$$



Как проверить управленческие решения до внедрения

РАССКАЗЧИК

сегодняшнее путешествие проведёт он

АНАЛИТИК

АРХИТЕКТОР



КОНСТАНТИН ДОНСКОЙ

директор по развитию Stormbpm

МЕТОДОЛОГИИ



ИНСТРУМЕНТЫ



2010

2026



15+ лет: оптимизация и автоматизация производств

ОПТИМИЗАТОР

РАССКАЗЧИК ДНЯ



«В КОПИЛКУ СВОИХ ЗНАНИЙ ВЫ ПОЛОЖИТЕ»

ВРМН
без лишних
квадратиков —
ещё не есть
оптимизация

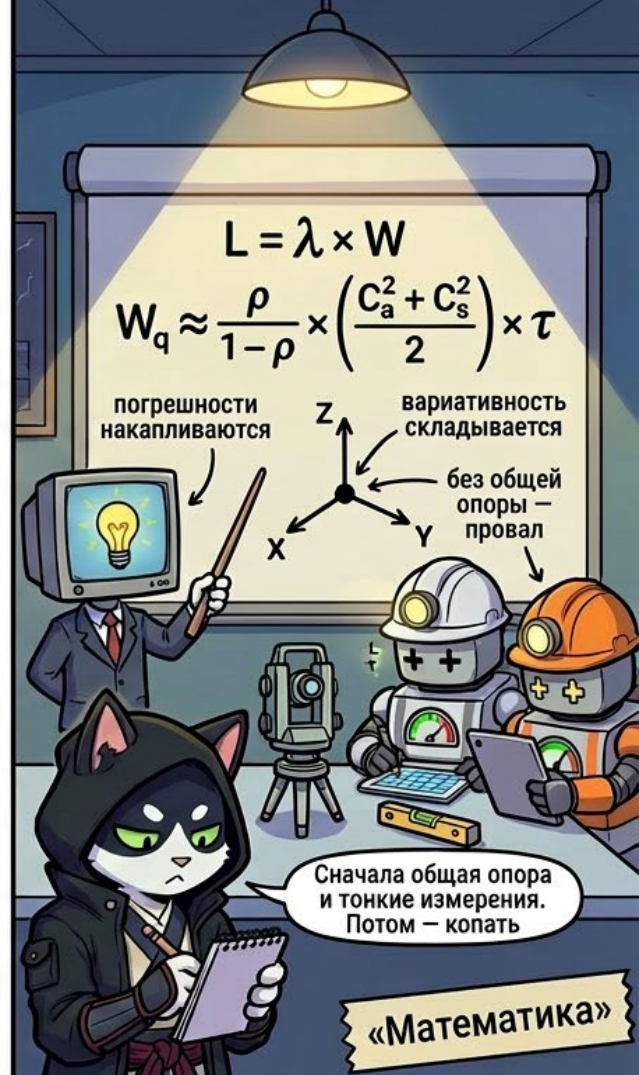
Где зарыта
большая часть
проблем
в процессах



«пропорции мастера-оптимизатора»

Работа только
с операциями или
с автоматизацией
не всегда
эффективно

Как прогнать
кучу гипотез
без капитальных
вложений



более 100 лет

1906 → 2026

ПАНТЕОН ТМО – 100+ ЛЕТ РАЗВИТИЯ

наши и западные коллеги, без которых не было бы этого знания



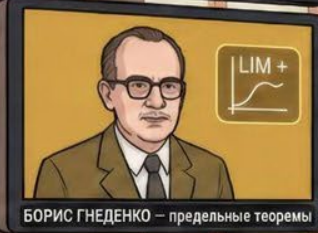
АНДРЕЙ МАРКОВ – цепи Маркова, 1906



АЛЕКСАНДР ХИНЧИН – формула Поллачена-Хинчина



АНДРЕЙ КОЛМОГОРОВ – аксиоматика вероятности



БОРИС ГНЕДЕНКО – предельные теоремы



ЯКОВ СИНАЙ – эргодическая теория



АГНЕР ЭРЛАНГ – старт Теории очередей



ДЖОН ЛИТТЛ – закон Литтла



ДЖОН КИНГМАН – приближение Кингмана



ДЖЕЙМС ДЖЕКСОН – сети Джексона

И ещё масса учёных, без которых не было бы этого знания



$$\rho = \frac{\lambda}{c \times \mu}$$

ρ = Загрузка системы
 λ = Скорость входящего потока (интенсивность)
 c = Число серверов
 μ = Скорость обслуживания (производительность)

ЗАЯВКА (λ)

РАМЕН

ОЧЕРЕДЬ ПЕРЕПОЛНЕНА!

СЕРВЕРЫ ($c = 2, \mu = 6$)

ラーメン

ЛОТОК 1 ($\mu = 6$)

ЛОТОК 2 ($\mu = 6$)

ВЫХОД (ВЫПОЛНЕНО)



Заявки приходят случайно

Очередь растёт, когда $\lambda > \mu$

Сервер обрабатывает по одной

$$\rho = \frac{\lambda}{c \times \mu} = \frac{20}{5 \times 5} = 0,8 = 80\%$$

$\lambda = 20$ обращений/час | $c = 5$ специалистов | $\mu = 5$ обр/час на специалиста

ТЕХПОДДЕРЖКА

$c = 5$ | загрузка системы: 80%

$\mu = 5$ обр/ч

$\mu = 5$ обр/ч

$\mu = 5$ обр/ч

$\mu = 5$ обр/ч

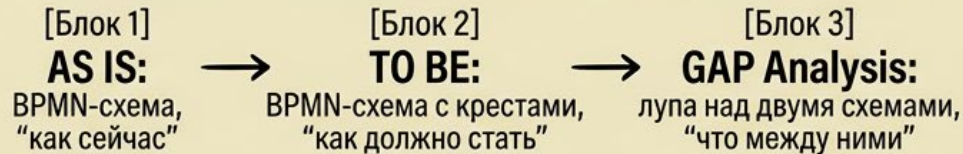
$\lambda = 20$ обращений/час

ОБРАБОТАНО: 20 в час

✓ решено ✓ решено

ЕСЛИ ПОТОК ВЫРАСТЕТ ДО
 $\lambda = 24 \rightarrow \rho = 96\%$
ВРЕМЯ ОЖИДАНИЯ В ОЧЕРЕДИ
ВЫРАСТЕТ В РАЗЫ

КАК ОБЫЧНО ДЕЛАЮТ



Интуиция



Опыт



Бенчмаркинг

- ✓ убрать лишнюю операцию
- ✓ исключить ненужную роль
- ✓ автоматизировать что можно
- ✓ нанять ещё людей
- ✓ нагрузить тех, кто простаивает

ЧТО ОБЫЧНО НЕ СМОТРЯТ

ОЧЕРЕДИ

ВАРИАТИВНОСТЬ
ПРИХОДОВ

ПРИЧИНЫ
ЗАГРУЗКИ

БУФЕРЫ
В СИСТЕМЕ

C_{-a} , C_{-s}

λ во
времени

ВОТ ТУТ
ЗАРЬИТО





Стало в полтора раза быстрее. Завтра отправляем на реальный маршрут



РЕКОРДЫ КРУГА	
8:02	→ 7:14
7:14	→ 6:30
6:30	→ 6:30
5:48	→ 5:48 ✓

Тренировка

0 мин — старт → 2 мин — стоим → 5 мин — перекрёсток → 8 мин — снова стоим → финиш

таймер: 47 сек

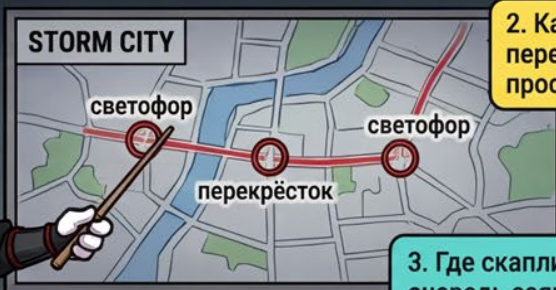


время доставки: 12:30
было до тренировок: 12:34
-4 секунды

Маршрут забрал всё, что выиграла скорость

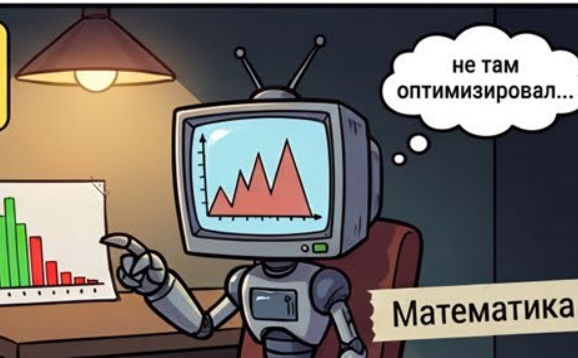
Реальность

1. Не разобрались, где реально тормозит.



2. Какие этапы перегружены, а какие простаивают.

3. Где скапливается очередь заявок.



не там оптимизировал...

Математика



1. При почти 100% —
не успевают
разбирать поток.

«ЧТО НУЖНО ДЕЛАТЬ ДАЛЬШЕ»



2. Заявки приходят
неравномерно.

поток лапши — заявки

Цех 1
Загрузка 70%



Выход:
100%



Цех 2
Загрузка 85%



успеваем-успеваем



Выход:
100%

Цех 3
Загрузка 95%



давайте
побольше!



АВАРИЯ

Выход: 40%
(и падает)



3. Любой сбой
→ мгновенная
очередь.

Чем ближе к 100% загрузке — тем хуже работает система

4. Страдают
сотрудники и
качество продукта.





Что вы
получите

Стабильный
приток денег

Лояльность
клиентов

Счастливые работники,
которые не уйдут

процесс как система обслуживания:
заявки → ожидание → обработка → выход

Если пройти путь

VRMN нужен + его необходимость становится только выше

А нужно ли вообще
это моделирование,
если оно решает только
3% задач
оптимизации?



Конечно, да



VRMN помогает:

- понять процесс
- договориться о структуре
- определить участников
- зафиксировать последовательность операций

Моделирование — это только первый шаг





ЗОНЕ 1: Поток (λ) – интенсивность, пики, изменение во времени



ЗОНЕ 2: Очереди – буферы, накопление, узкие места



ГОТОВАЯ ЛАПША



ЗОНЕ 6: Экономика – стоимость операций, ресурсов, ожидания



ЗОНЕ 5: График работы – сменность, доступность



ЗОНЕ 4: Ресурсы – количество, специализация, правила назначения



ЗОНА 3: Операции – среднее время, вариативность, распределение

Закон Литтла

$$L = \lambda \times W$$

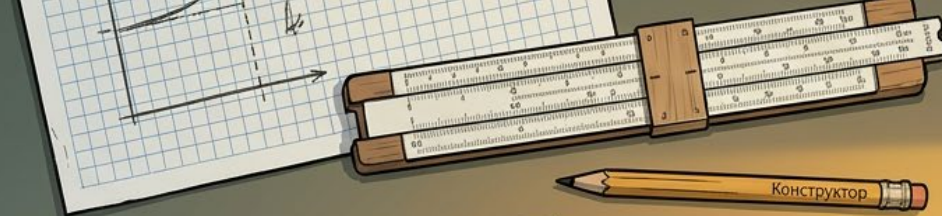
- L — среднее число заявок в системе
 λ — интенсивность входящего потока
 W — среднее время пребывания заявки в системе

John D. C. Little, 1961 — формальное доказательство.

среднее время в
системе W



среднее число в
системе L



Приближение Кингмана

проверить C_a^2 !

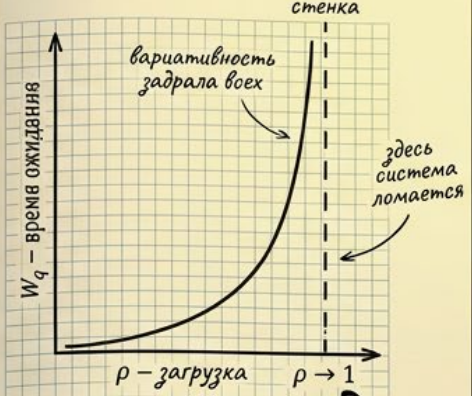
$$W_q \approx \frac{\rho}{1-\rho} \times \left(\frac{C_a^2 + C_s^2}{2} \right) \times \tau$$

W_q – среднее время ожидания в очереди
 ρ – загрузка системы, $\rho = \frac{\lambda}{c \times \mu}$
 C_a – коэффициент вариации поступлений
 C_s – коэффициент вариации обслуживания
 τ – среднее время обработки

¹ John F. C. Kingman, 1961 – оценка верхней границы.

ано, № 286

Ленинград 1940



Square Root Staffing Rule — закон квадратного корня

$$c \approx \rho + \beta \times \sqrt{\rho}$$

c — необходимое количество серверов (операторов, станков, специалистов)

ρ — целевая загрузка, $\rho = \lambda / \mu$

β — коэффициент уровня сервиса (чем выше, тем меньше очередь)

не удваивать!

* Halfin & Whitt, 1981 — режим «многих серверов».

Заявок	Интуиция «удвоить штат»	Расчёт $\sqrt{\cdot}$ -правила
x1	10 операторов	10 операторов
x2	20 операторов	14 операторов
x4	40 операторов	20 операторов

вот это сюрприз

Удвоенный поток не требует удвоенного штата. Корень — это про то, как очередь усредняет случайность.

17

Расчёт штата 1987

совещание
14:00

STORM
CORP.

Модель Q3

Сети Джексона

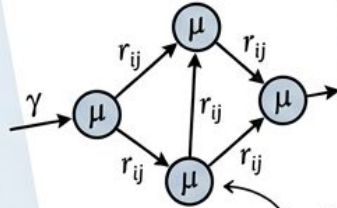
$$\lambda_j = \gamma_j + \sum_i (\lambda_i \times r_{ij})$$

- λ_j — итоговый поток в узел j
- γ_j — внешний поток в узел j
- r_{ij} — вероятность перехода с узла i на узел j
- Σ — суммирование по всем входным узлам i

James R. Jackson, 1957 — теорема о произведении.

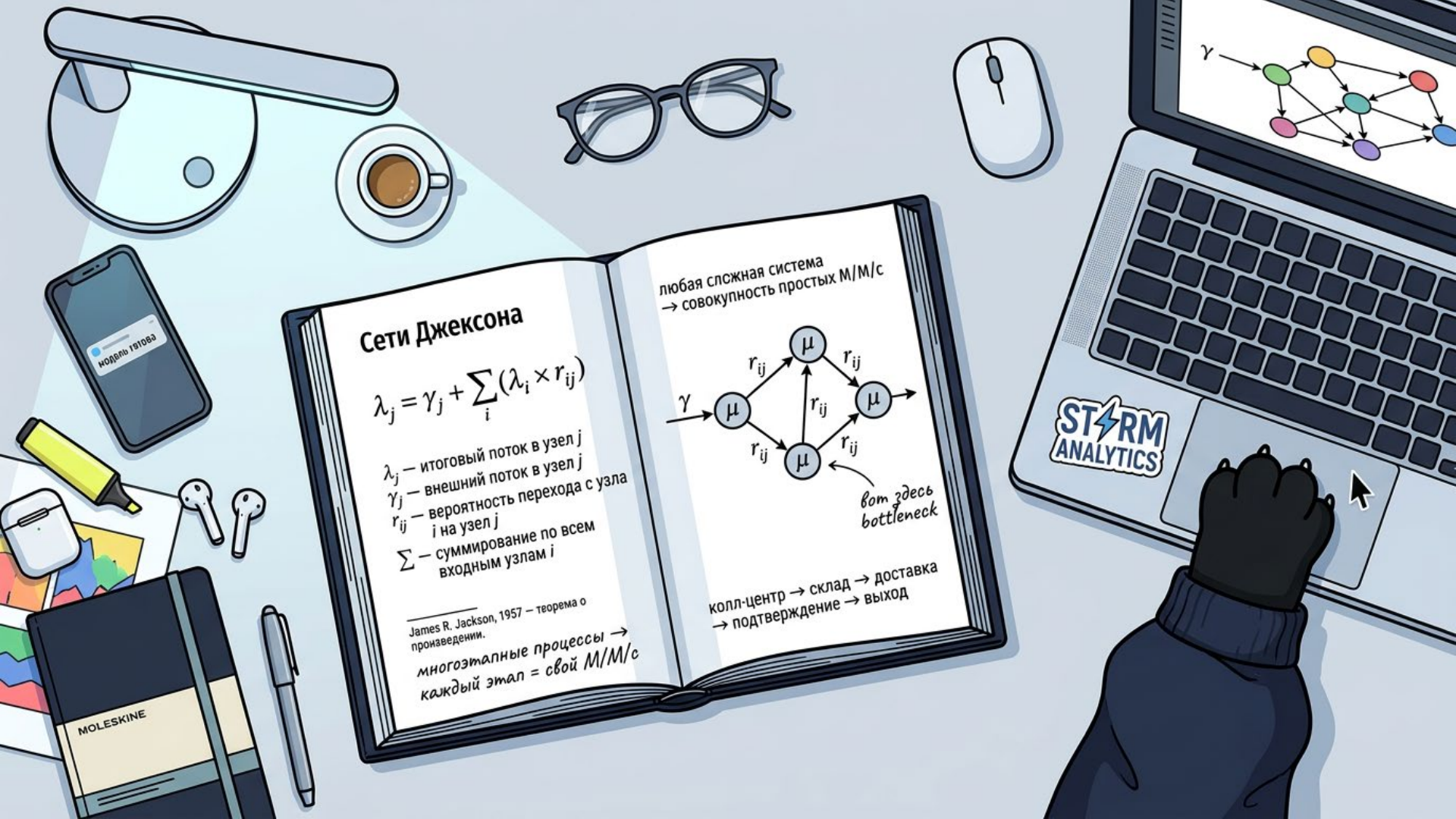
многоэтапные процессы →
каждый этап = свой M/M/c

любая сложная система
→ совокупность простых M/M/c



вот здесь
bottleneck

колл-центр → склад → доставка
→ подтверждение → выход





БИБЛИОТЕКА ОПТИМИЗАЦИЙ

всё открыто для всех, бесплатно

Календарь

- Расширить рабочие часы
- Сократить простои

Мощности

- Добавить ресурсы
- Сохранить ресурсы
- Перебалансировать нагрузку
- Оптимизировать загрузку (%)
- Планирование мощностей

Управление потоком

- SOIЖIP
- Ограничение WIP
- Вытягивающая система

Продвинутые методы

- Теория ограничений
- Пул ресурсов
- Оптимизация партий
- Снижение вариативности

Экспертные методы

- Буферизация
- Оптимальная скорость μ

Фундаментальные законы

- Критический путь
- Формула Бранга
- Закон Литтла

Очереди

- Приоритизация
- Разделение потоков
- Сокращение переключений
- Управление буфером

Входящий поток

- Сглаживание
- Филтрация спроса
- Перенаправление

Развилки

- Обход узкого места
- Усиление фильтра
- Изменение пропорций

Качество

- Показатели
- Контрольные точки
- Проверка с первого раза

Диагностика

- Анализ владений
- Точка бронирования
- Оптимальная загрузка

Планирование

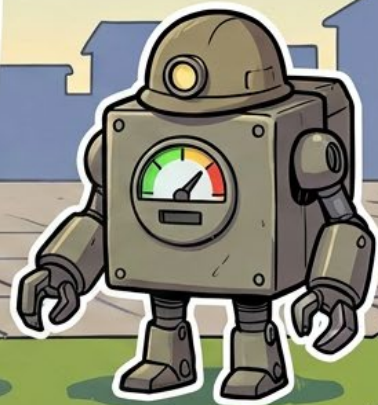
- Сценарии «что если»
- Пулгоих спроса
- Резерв мощностей

Экономика

- Идла
- Размер партии
- Динамическое ценообразование
- SAC
- Коммерция
- Стоимость хранения

Перепроектирование

- Перенос акцентив
- Диверсификация
- Устарение
- Автоматизация



> 60 оптимизаций
> 100 терминов и законов

В ПУТЬ — за данными



ЧТО СОБИРАЕМ — короткий список

#	Параметр	Что измеряется
1	Поток заявок λ	сколько заявок в единицу времени
2	Время обработки	среднее и распределение
3	Вариативность CV	отклонения времени операций
4	Количество ресурсов	сколько роботов на операции
5	График работы	доступность ресурсов
6	Очереди и буферы	где накапливаются заявки

С ЭТИМ МОЖНО НАЧАТЬ ПРИМЕНЯТЬ ФОРМУЛЫ ТМО

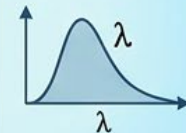
1 — PROCESS MINING

первый и самый трендовый метод

ЧТО СОБИРАЕМ

- ✓ поток заявок λ
- ✓ время обработки операций
- ✓ вариативность операций CV
- ✓ очереди между этапами

PROCESS MINING – журналы событий



CV = 1,8 – высокая вариативность

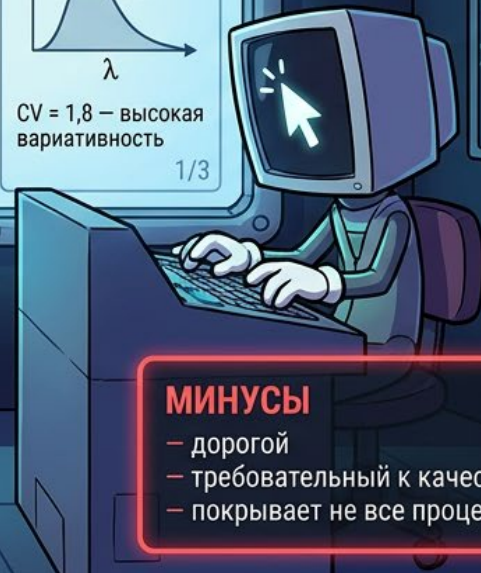
2/3

1/3

МИНУСЫ

- дорогой
- требовательный к качеству данных
- покрывает не все процессы

события → знание



2 – АНАЛИЗ ДАННЫХ ИС

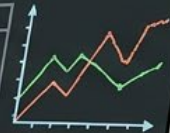
если ИС уже есть – это золотая жила

- ✓ поток заявок λ
- ✓ время выполнения операций
- ✓ количество задействованных ресурсов
- ✓ графики загрузки ресурсов

отдел качества

«АНАЛИЗ ДАННЫХ ИС»

Лог
оп.1 – 12:43 – 8 свк
оп.2 – 12:43 – 23 свк



65%

92%

18%

- не все ИС связаны
- не у всех они вообще есть
- корректность данных нужно проверять

3 — РУКОПИСНЫЕ ЖУРНАЛЫ

классика никуда не делась

- ✓ поток заявок
- ✓ время выполнения операций
- ✓ базовая оценка очередей



- их нужно найти
- желательно перевести в цифровой вид

4 — ФОТОГРАФИЯ РАБОЧЕГО ДНЯ

классика хронометража — никуда от неё

операция	ХРОНОМЕТРАЖЕ			
	Среднев	Макс	Мин	отклонение
~	33,8	46,0	2,0	+13
~	20,0	47,6	7,8	-0,6
~	32	32	17	+0,2
~	13,2	93,2		
~	14,3	24		
~	15,4			
~	15,7			



- ✓ время обработки операций
- ✓ вариативность операций CV
- ✓ загрузку ресурсов

не уйти, пока
не зафиксируем

- долго
- иногда дорого
- очень редко бессмысленно

«5 — ГЕМБА-СЕССИЯ»

выход на процесс своими ногами, 4–8 часов

- ✓ скрытые очереди
- ✓ неформальные правила работы
- ✓ реальные ограничения процесса

не указано в схеме

вот тут реальная очередь

и это тоже



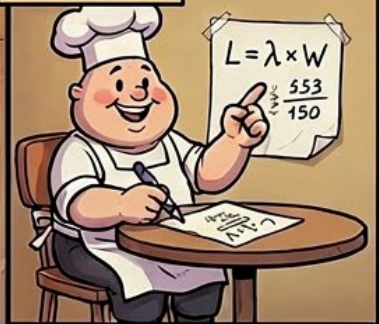
ГЕМБА – стоимость дня в рублях

Уровень / Роль	Своими силами	Средний консалтинг	Big Four
Ton / Strategy / Partner	100 000 – 142 857	85 121	177 487
Старший рук. Senior	57 143 – 114 266	122 792	304 263
Руководитель Lead	45 714 – 76 160	284 341	660 570
Рупьций эксперт Consultant	–	304 263	716 187
Эксперт Analyst	–	166 620	412 418
Младший специалист Junior	–	65 199	139 655
ИТОГО за день	202 857 – 333 333	1 028 326	2 410 628

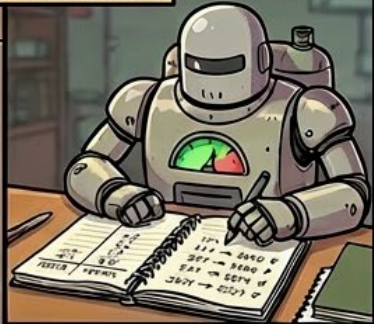
* цифры приведены с учётом среднего дисконта - 30% от ставки

ЧЕМ СЧИТАТЬ?

на бумажке



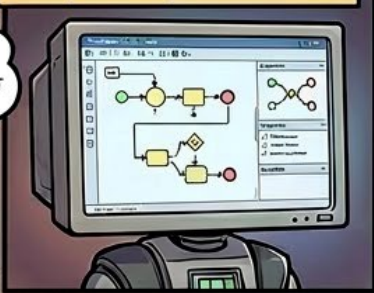
в блокноте



Excel – сеть из формул



специализированное ПО



всё это работает, но...



StormDES – DES-симулятор от StormBPMN

вся теория за 100 лет ТМО



Лин · 6 Сигм · Теория ограничений



прибыль, затраты, ROI

Сценарий	Прибыль	Затраты	ROI
1	1000	500	100%
2	2000	1000	100%
3	3000	1500	100%
4	4000	2000	100%
5	5000	2500	100%

модно и удобно



- ✓ теория за 100 лет
- ✓ Лин, 6 Сигм, ТoS
- ✓ P&L
- ✓ удобный интерфейс

выбор остаётся за вами

DES — ВЫ ТЕПЕРЬ ЗНАЕТЕ ВСЁ



БЕСПЛАТНЫЙ
ИНТЕРАКТИВНЫЙ
КУРС

теперь
я знаю все



ИНТЕРАКТИВНЫЙ
КВЕСТ



СООБЩЕСТВО
В TELEGRAM



YOUTUBE-КАНАЛ



DES
ЭКСПЕРТ



DES
ЧЕМПИОН



DES
АНАЛИТИК



Вперед! К новым горизонтам

STORM CITY ГОРДИТСЯ

ППОШАДКА

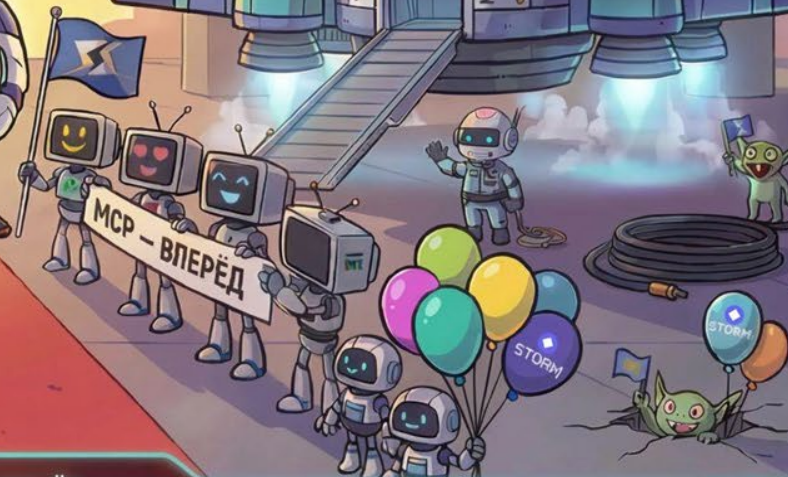
A-7

T-00:08:42



ВЕБИНАР
MCP+AI

STORM
MCP
MCP-001



T-00:08:42 — финальный отсчёт