

SRE : Google 运维解密

孙宇聪

孙宇聪 —— Google SRE 07-15

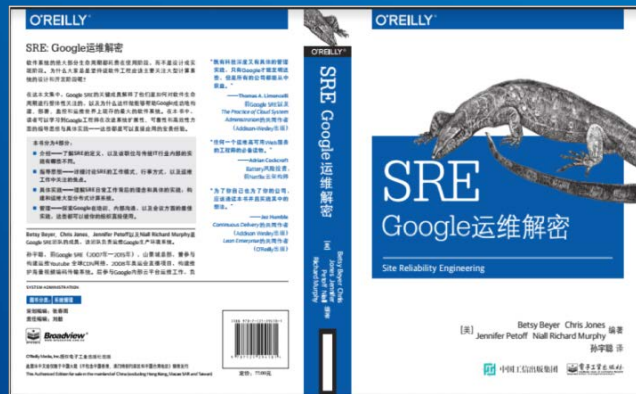
- IT | SysOps
- Youtube Video SRE
- Borg SRE



I have root @ google

The Book

- 近百名 SRE 关键成员的文集，删掉 2/3 内容后共 33 章。
- 全书分为五个部分
 - 概览
 - 指导思想
 - 具体实践
 - 管理
 - 与其他行业的对比
- “SRE is what happens when you ask a software engineer to design an operation team”
—— Ben Treynor Sloss



传统软件项目生命周期



系统管理员 (Sysadmin)

- 组装现成的组件，部署为服务
- 负责应急事件处理与常规更新
- 团队规模一般随着服务部署规模增长而增长

传统运维模式的弱点

- 创新困难
 - 重执行，抗拒变化，组织结构僵硬
 - 积累困难，提高困难，发展困难
- Dev Ops 之间的永恒对立
 - 激励机制中存在根本性问题和冲突
- 成本压力大
 - 团队规模伴随服务规模不断扩大

Google 运维模式：SRE

- Ben Treynor Sloss
 - 资深研发负责人
 - 2003 年加入 Google 组建 7 人“生产运维小组”
- 由软件工程师执行运维工作
 - 从内心里鄙视重复性工作
 - 有能力进行自动化研发
 - 用自动化软件完成运维工作
- SRE 是 DevOps 思想在运维方面的具体实践



SRE 职责分布

工程研发

日常运维

应急响应

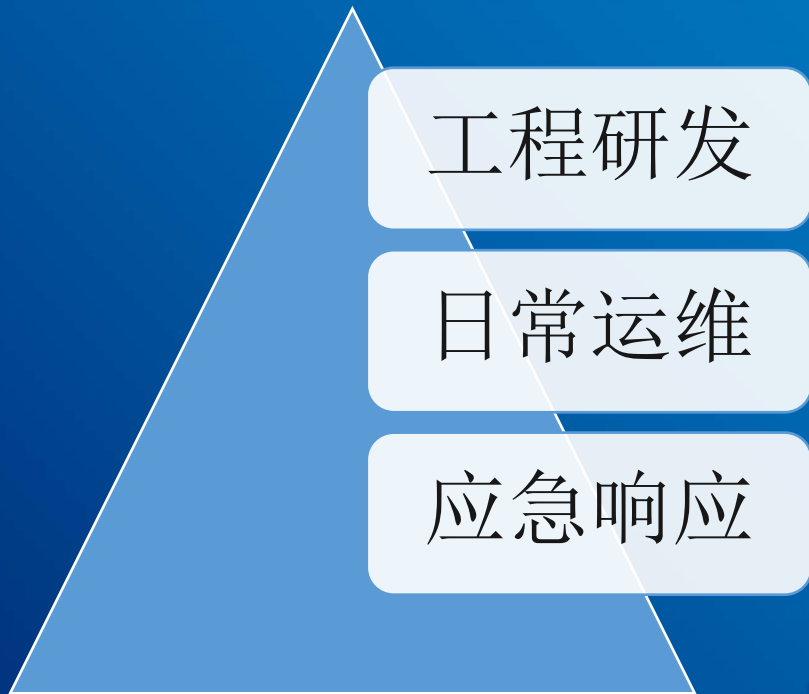
Margaret Hamilton

- 软件研发负责人
 - “软件开发” 还不是一个行业
 - 属于预算外事务，外包给了 MIT
- Apollo 8（载人绕月飞行）
 - “We had been told many times that astronauts would not make any mistakes.”
 - “Do not select P01 during flight,”
 - “Thorough understanding of how to operate system is not enough to prevent human errors”



SRE 工作职责金字塔

- 应急响应
 - 监控
 - 应急事务处理
 - 事后总结
- 日常运维
 - 变更管理
 - 容量规划与置备
 - 性能与效率



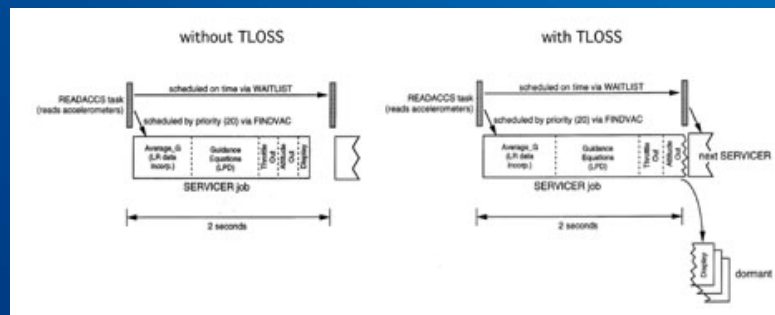
工程研发

日常运维

应急响应

Apollo 11 (人类首次登月)

- 登月着陆的最后关头
 - 警报代码 1201 / 1202
 - 计算单元过载
- “执行清单”中的某个错误导致一个错误的任务占用了过多的CPU资源。
- 能自愈的系统!



SRE 工作职责金字塔顶层 —— 工程研发

- 制定合理的 SLO
- 在 SLO 安全范围内全速前进
- 可靠性是软件的核心功能
- “If we are engineering processes and solutions that are not automatable, we continue having to staff humans to maintain the system. If we have to staff humans to do the work, we are feeding the machines with the blood, sweat, and tears of human beings. Think The Matrix with less special effects and more pissed off System Administrators. “ —— SRE

SRE 模型如何成功?

建立职业化，专业化，拥抱变化的灵活组织

区分责任与指责，坚持推演与演习。

SRE 模型的成功关键点之一

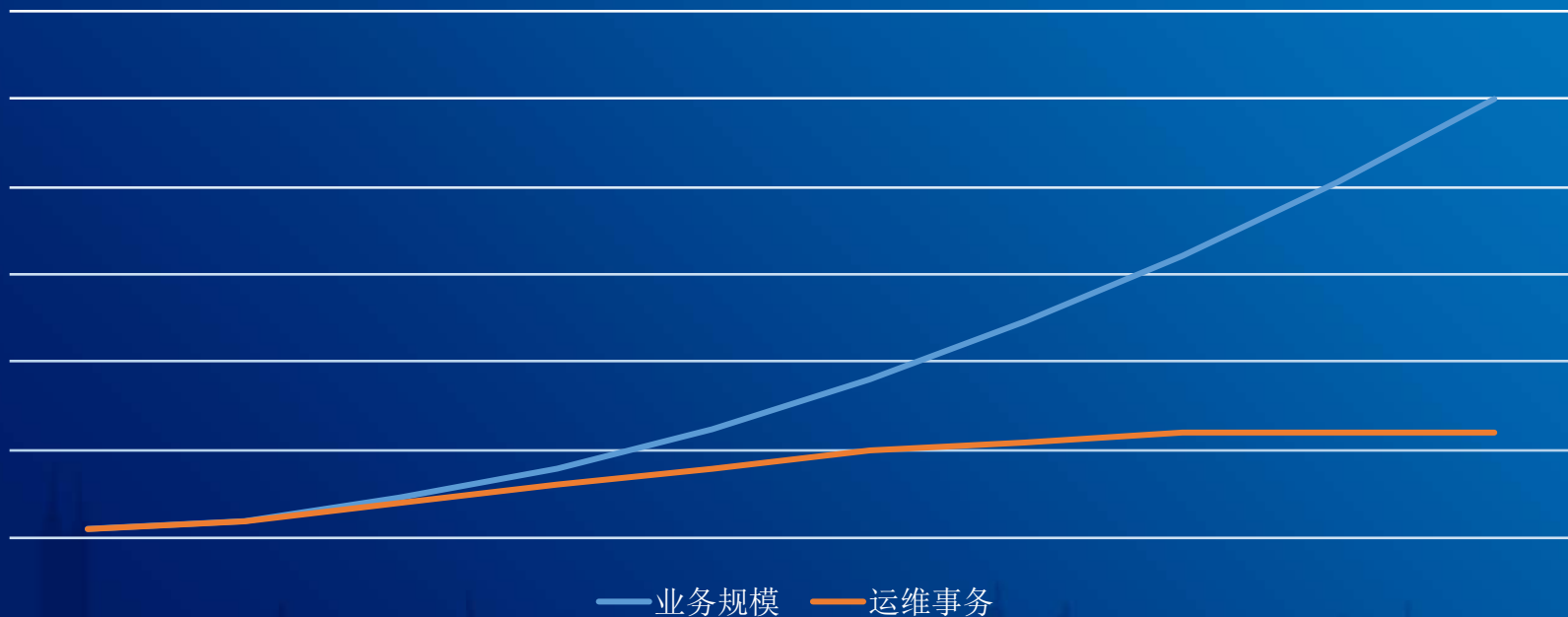
减少琐事，保障 50% 的工程时间

SRE 模型的成功关键点之二

自主决策权 + 产品建言权

SRE 模型的成功关键点之三

The End of Beginning





微信扫一扫