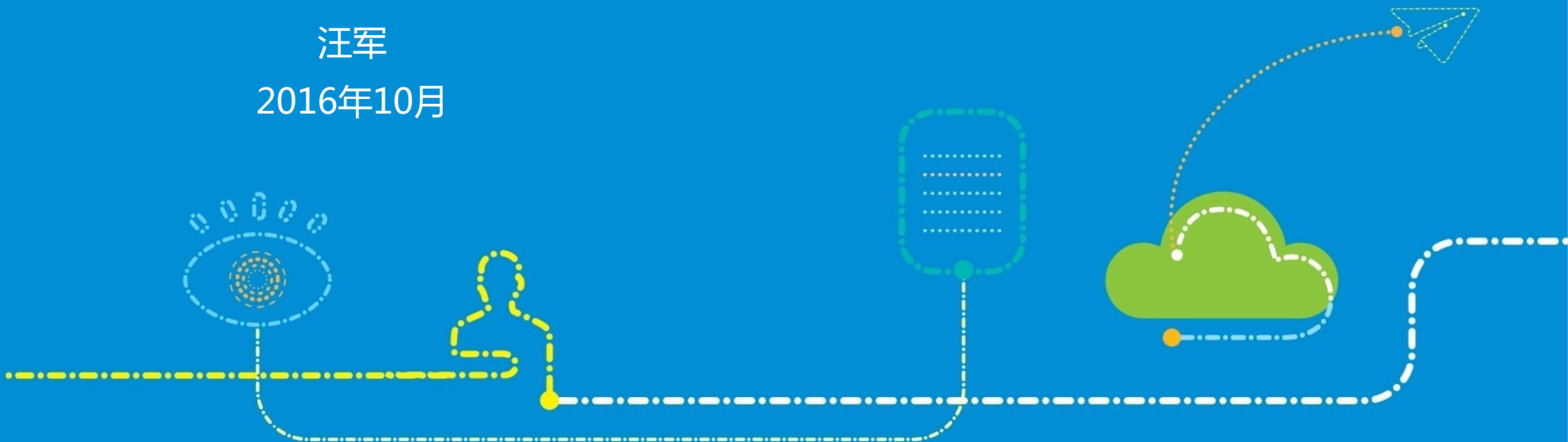


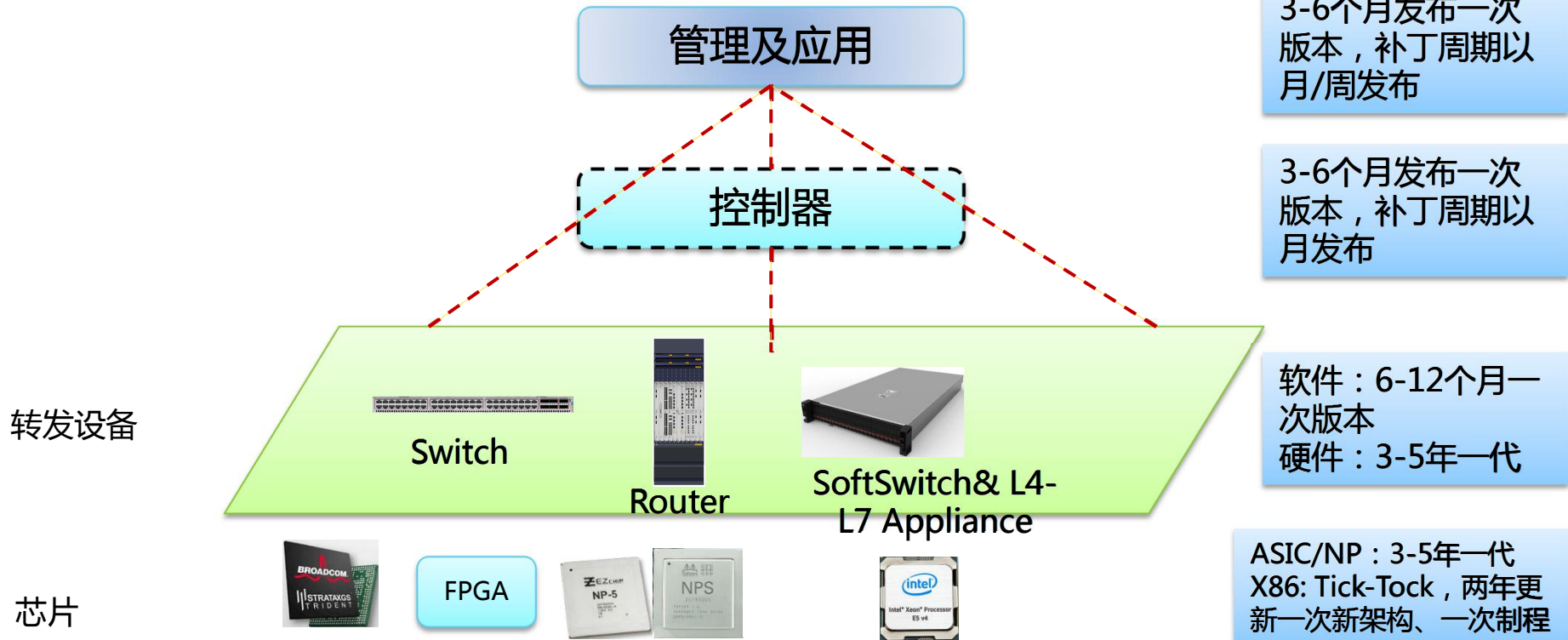
P4应用场景及价值

汪军

2016年10月



网络产品研发现状

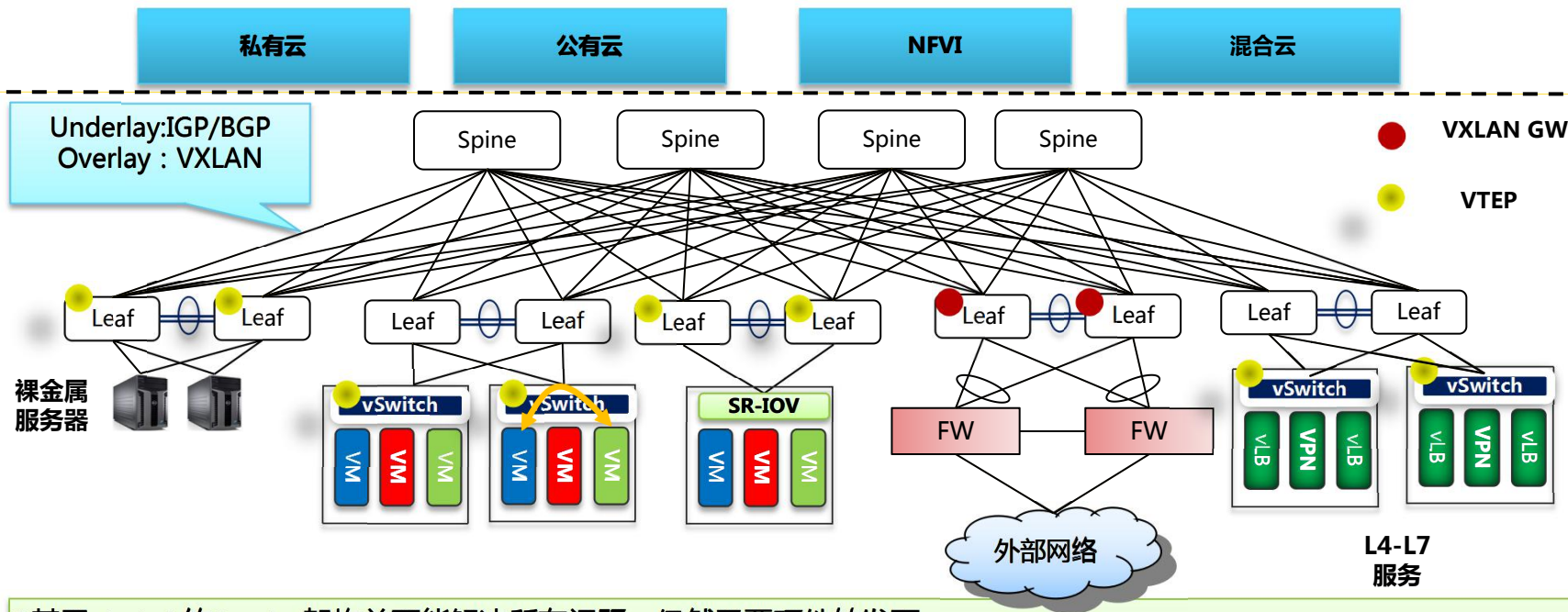


ASIC: 固化逻辑, 简单转发, 最佳每bit转发成本和功耗; 功能和芯片绑定, 新功能需要升级芯片

FPGA/NP: 可编程, L2-L4足够灵活性, 性能是主要升级驱动力

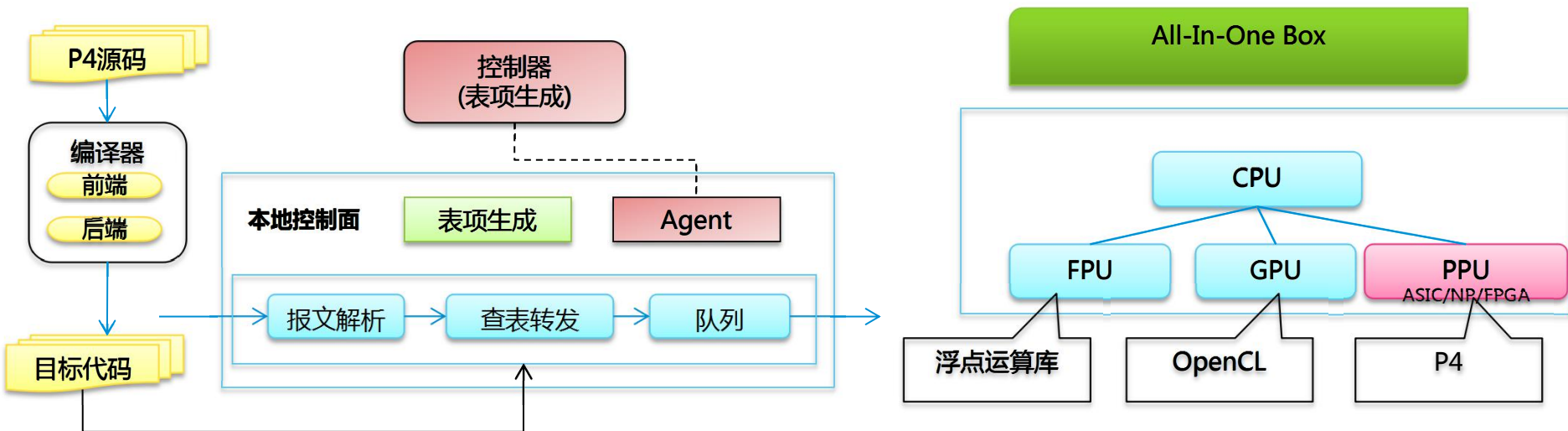
X86: 灵活的L2-L7处理, 单位功耗和成本较高, 更适合低密度转发和高复杂业务处理

数据中心网络对基础网络的需求



- 基于vSwitch的Overlay架构并不能解决所有问题，仍然需要硬件转发面
 - ✓ 裸金属服务器的统一接入；高性能NFV转发需要采用SR-IOV方式
 - ✓ 高流量租户，比如私有云中的视频业务、NFVI，出口软件化困难
- 新的网络业务对新的转发面格式、能力的需求，SFC for NFV；网络的可视化
- 高性能、灵活性和低功耗设计的矛盾

P4的价值



□ Portability, Platform/Hardware Independent

✓ 一次编程、处处可用

□ Protocol Independent，或许叫“重构流水”更准确

□ 以计算为中心设计网络设备，自适应转发，有专用PPU，则硬件转发，否则软转发。(控制面携带多个P4 Target代码，自适应安装转发代码，或者JIT编译、下载代码)

□ 语言+工具链+标准库，构成生态链

对P4/可编程芯片的需求

新的网络业务引入

✓引入新的业务能力和协议

- ✓ E.g. Service-Chaining , NSH ;
Parse/Push/Pop/Modify New header
- ✓ Mirror Over VXLAN or any other encapsulation
- ✓ 有限的基于Session的业务处理

✓优化网络实现

- ✓ 携带Policy/Security Group Id , 避免N平方表项问题
- ✓ 报文携带时间戳 , 方便时延测量
- ✓ 携带其它Metadata , 便于实现网络优化/可视化

✓构建异构网络互通的能力

- ✓ 可升级设备构建不可升级设备之间的互通网关。
E.g. OVS/老硬件和新设备在新架构下的互通

通用网络编程

✓一套代码 , 适应不同的目标芯片 , 包括各种

✓降低异构开发的额外投入

✓让网络成为计算的协处理器

✓X86+网络协处理器/加速卡

✓P4编译成不同的目标代码 , 目标硬件平台上有加速芯片 , 则下载硬件转发逻辑 , 否则启用软转发

P4+可编程芯片的价值

最终客户

- ✓保护投资，新的业务需求不必更换交换机硬件，而是升级软件来实现
- ✓更快的业务创新周期，不依赖于芯片的更新换代
 - ✓E.g. TridentII只支持VXLAN L2通信，T2+才支持VXLAN L3通信，可以支持分布式路由

厂商

- ✓更少的芯片能力限制，加快对客户需求的响应速度
- ✓通过P4工具链，归一化芯片的开发方法和API，减少供应商转换的成本
- ✓在X86上验证ASIC转发能力，缩短研发周期

谢谢！



未来，不等待

