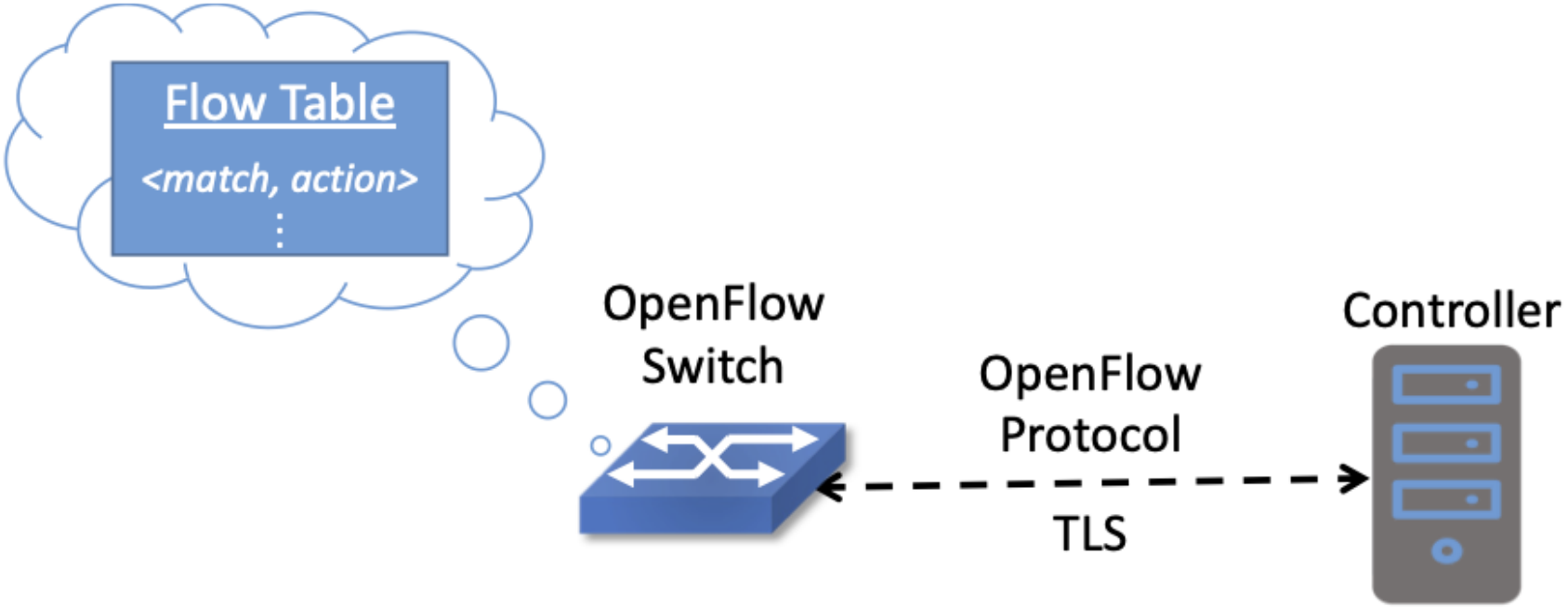
SDN- Estudos de Caso

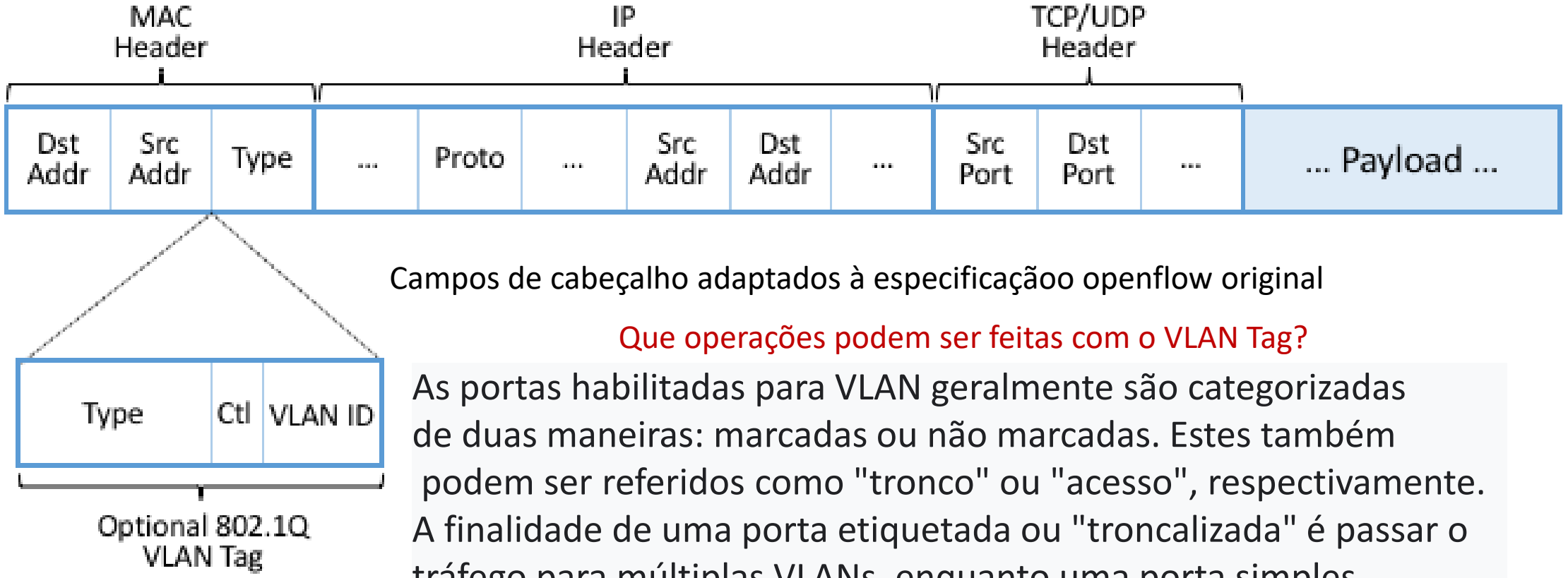
Aula 2

Estudos de Caso

- Nesta aula queremos apresentar os diversos estudos de caso de utilização das Redes Definidas por Software.
- Basicamente o que vamos analisar são benefícios que podem ser trazidos pelas SDNs às diversas aplicações e equipamentos tradicionais



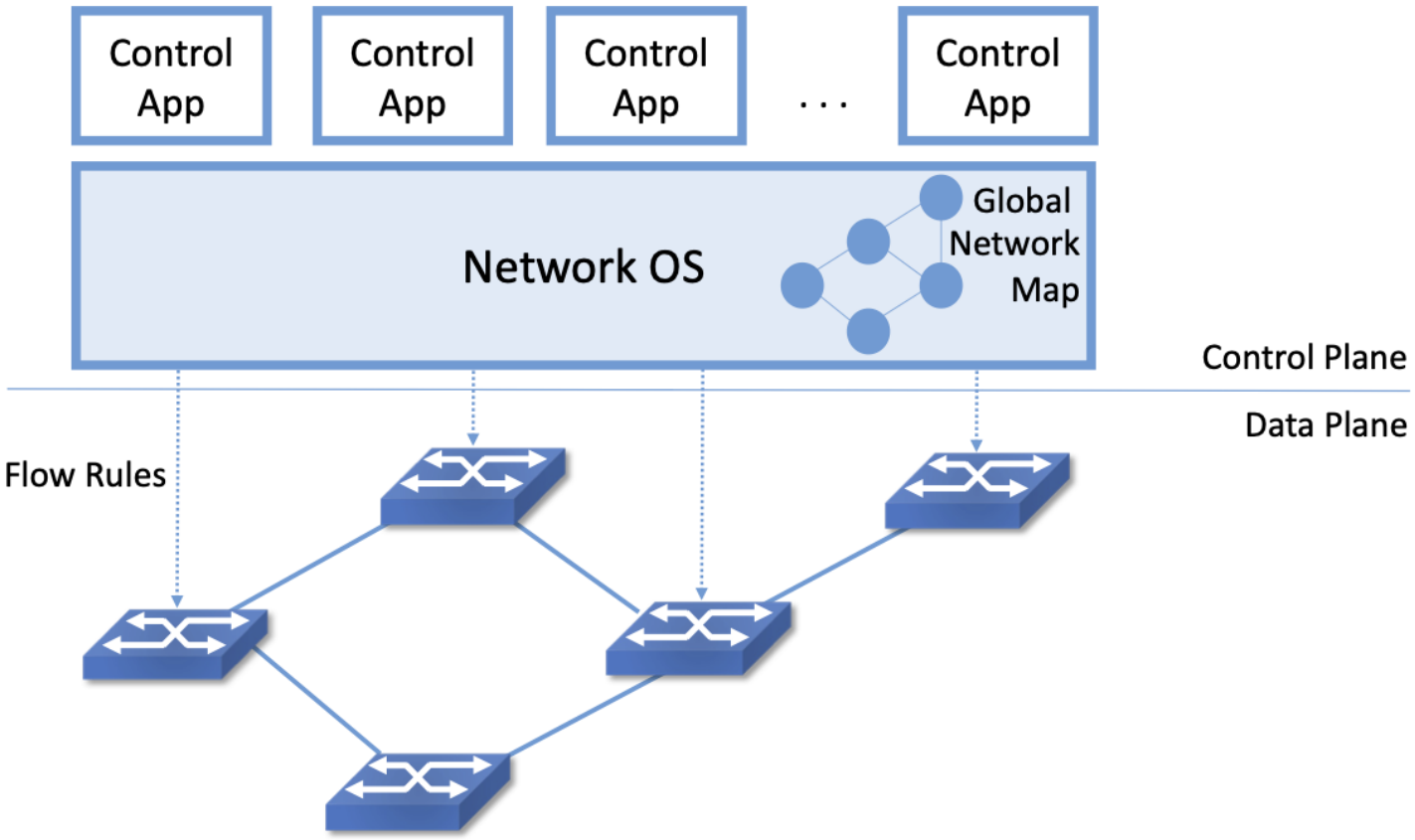
O controlador passa com segurança regras de fluxo para um switch OF, que mantém a Tabela de Fluxo



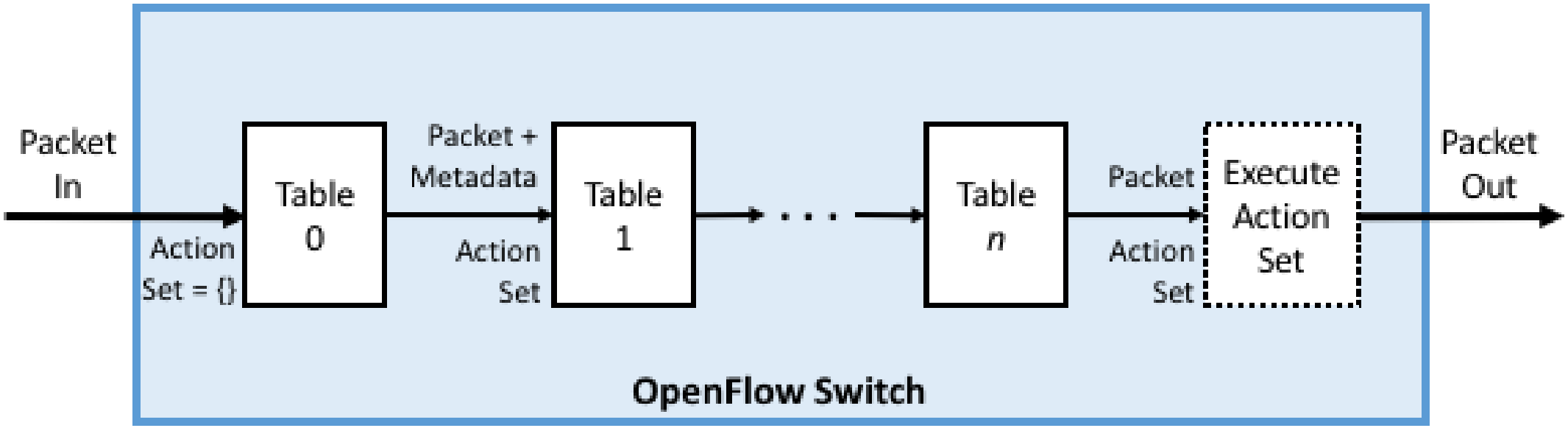
Campos de cabeçalho adaptados à especificaçãoo openflow original

Que operações podem ser feitas com o VLAN Tag?

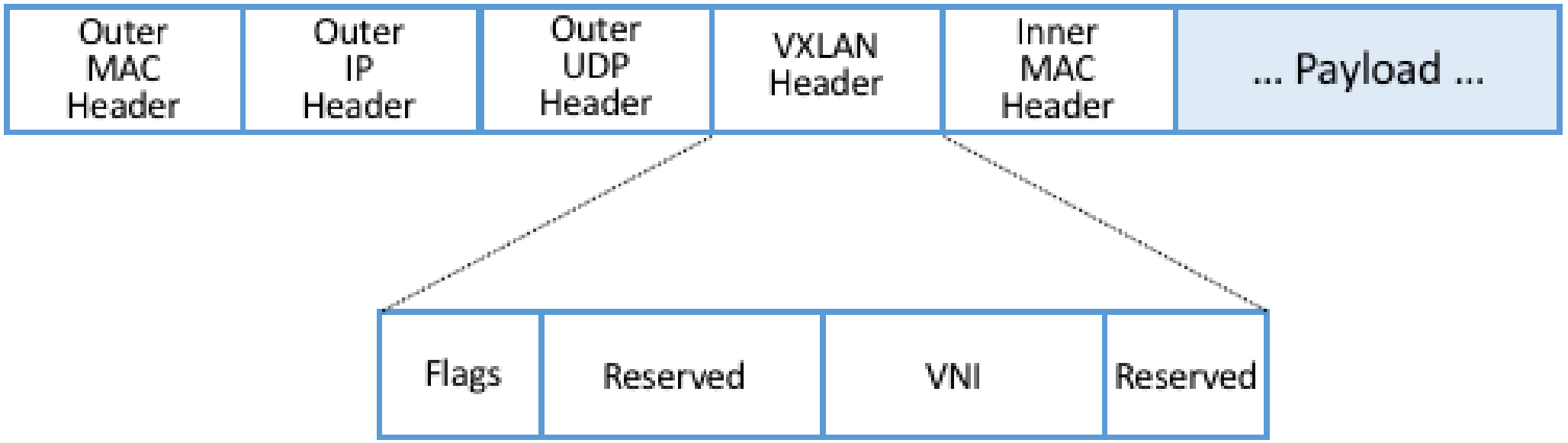
As portas habilitadas para VLAN geralmente são categorizadas de duas maneiras: marcadas ou não marcadas. Estes também podem ser referidos como "tronco" ou "acesso", respectivamente. A finalidade de uma porta etiquetada ou "troncalizada" é passar o tráfego para múltiplas VLANs, enquanto uma porta simples ou de "acesso" aceita tráfego para apenas uma única VLAN.



Network Operating System (NOS) que hospeda um conjunto de aplicativos de controle e fornece um ponto de controle logicamente centralizado para um plano de dados de rede distribuído subjacente.



Esquema simples de um pipeline de encaminhamento OpenFlow.



Cabeçalho VXLAN encapsulado em um pacote UDP/IP.

VXLAN

- **O que é o VXLAN?**
- Virtual eXtensible Local-Area Network, ou VXLAN, é um padrão de tecnologia de virtualização de rede da Internet Engineering Task Force (IETF).
- Ele permite que uma única rede física seja compartilhada por várias organizações diferentes, ou "locatários", sem que nenhum locatário seja capaz de ver o tráfego de rede de nenhum outro.
- Dessa forma, os VXLANs são análogos a unidades individuais em um prédio de apartamentos: cada apartamento é uma habitação privada e separada dentro de uma estrutura comum, assim como cada VXLAN é um segmento de rede dedicado e privado dentro de uma rede física compartilhada.
- Tecnicamente falando, um VXLAN permite que uma rede física seja segmentada em até 16 milhões de redes virtuais ou lógicas. Ele funciona encapsulando os quadros de Ethernet de camada 2 em um pacote de protocolo de datagrama de usuário (UDP) de camada 4 ao lado de um cabeçalho de VXLAN.
- Quando combinado com uma rede virtual privada (EVPN) Ethernet — que transporta o tráfego Ethernet em redes virtualizadas usando protocolos de WAN — o VXLAN permite que as redes de Camada 2 sejam estendidas em uma rede IP ou MPLS de Camada 3.

VXLAN: O futuro das Redes de Centros de Dados

Hoje em dia, a emergência e o rápido desenvolvimento de novas tecnologias,como a computação em nuvem, big data e IA, estão a levar os centros de dados a adotar tecnologias de virtualização e a aumentar a virtualiza dos seus servidores para fornecer serviços em nuvem.

Os links VLAN tradicionais têm se mostrado insuficientes para lidar com os requisitos dos data centers em nuv para grande escala e flexibilidade.

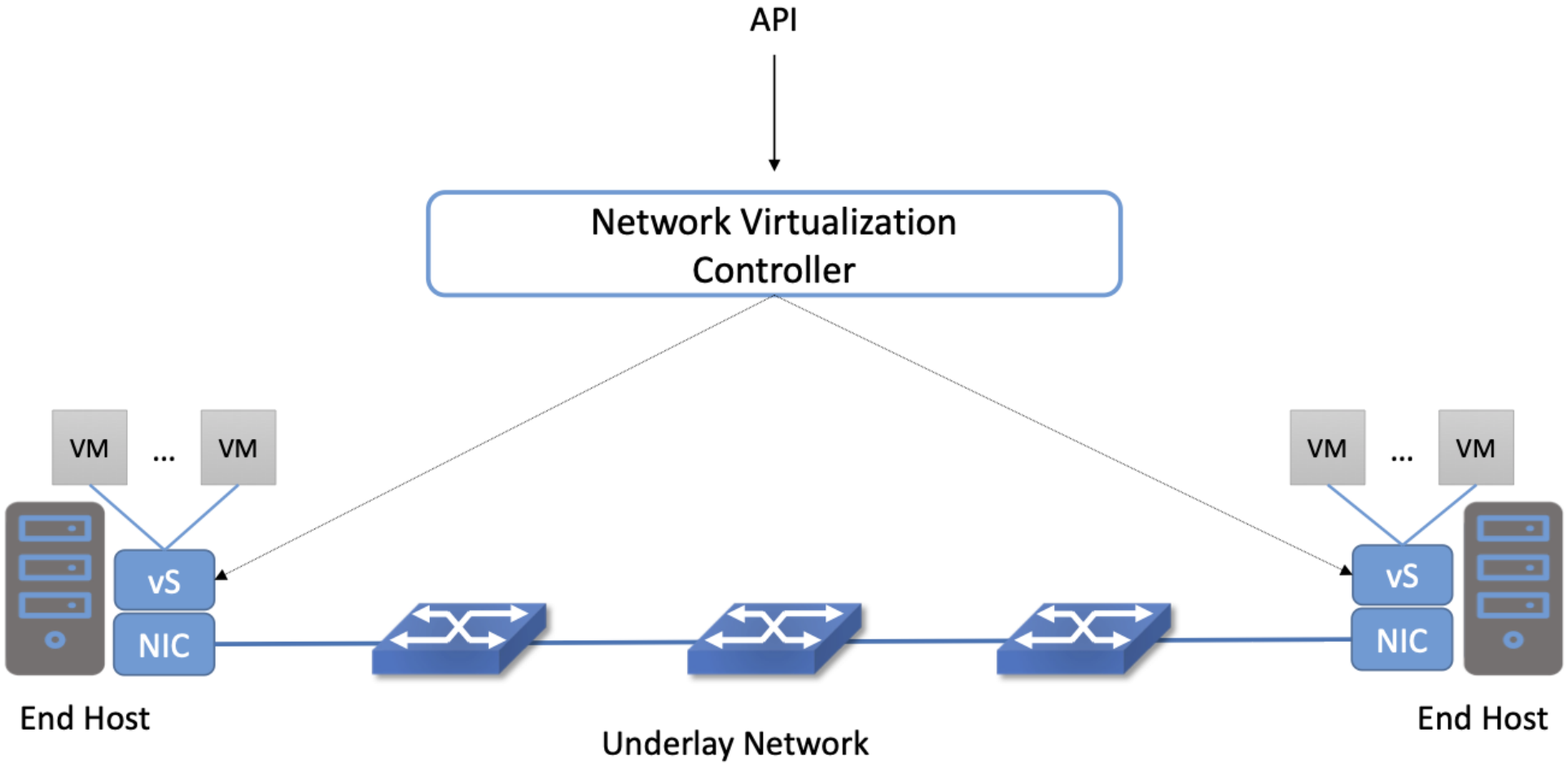
Portanto, a VXLAN entrou em cena e se tornou uma parte significativa da arquitetura de rede de data centers modernos.

O que é VXLAN?

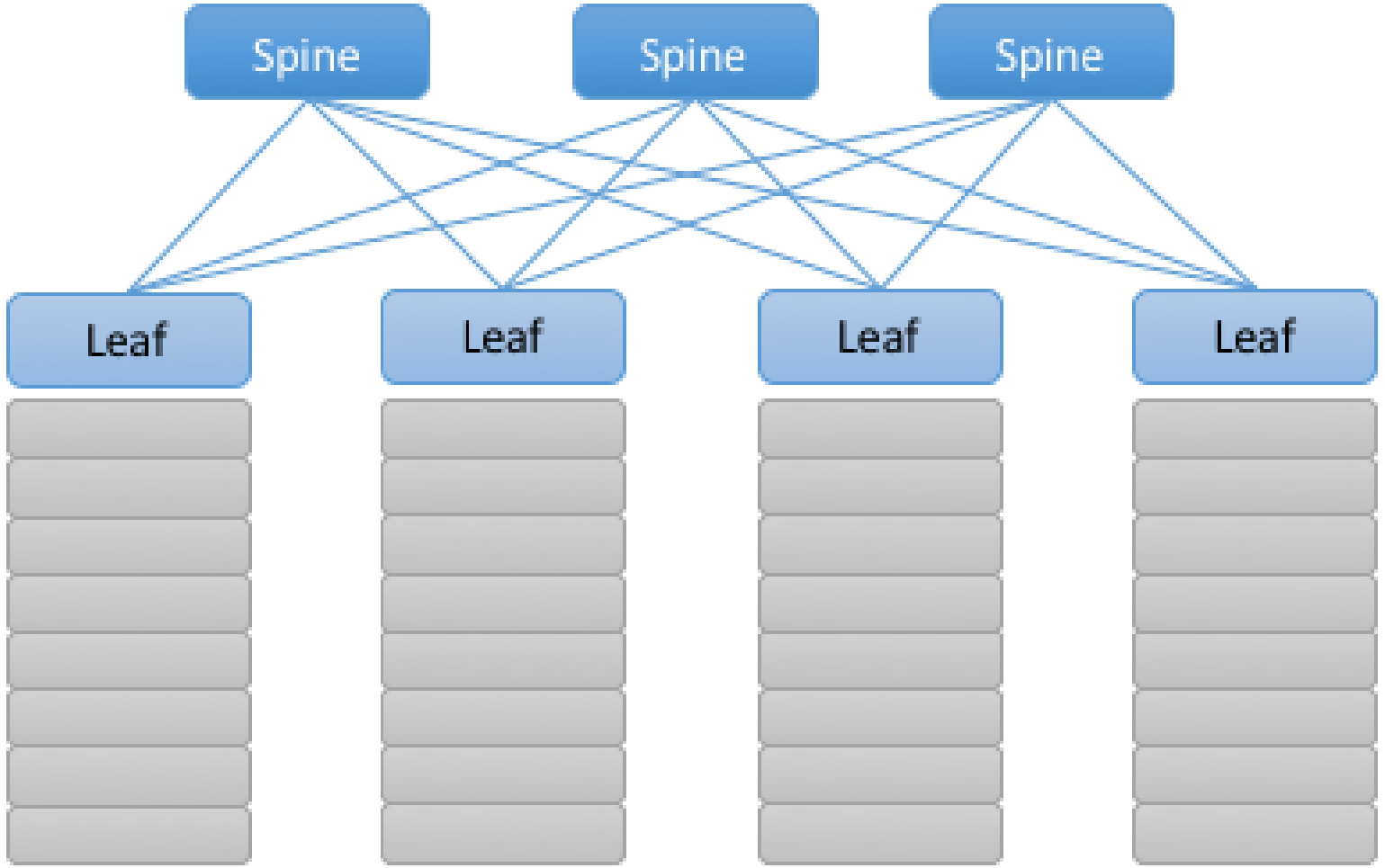
VXLAN (Virtual Extensible Local Area Network) é uma tecnologia de sobreposição para virtualização de rede, que estabelece um túnel lógico na rede IP para estender a rede da Camada 2 sobre uma rede subjacente da Camada 3 existente.

VXLAN usa VXLAN Tunnel Endpoint (VTEP), que podem ser hosts finais ou switches de rede, ou routers, para encapsular e desencapsular o tráfego da camada 2.

A VXLAN foi projetada para fornecer serviços de rede de data center confiáveis e escaláveis para clientes de serviços geridos e é uma tecnologia para construção de data centers de próxima geração.

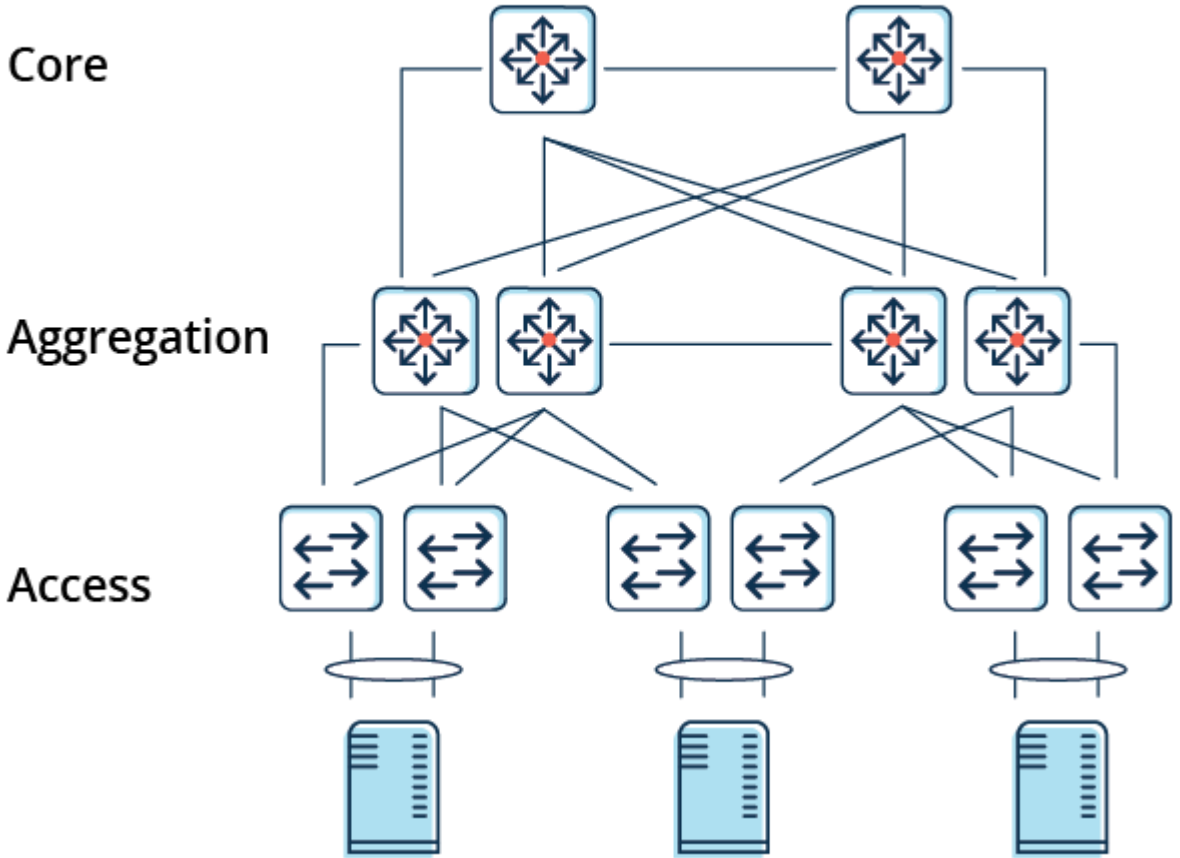


Um exemplo de sistema de virtualização de rede

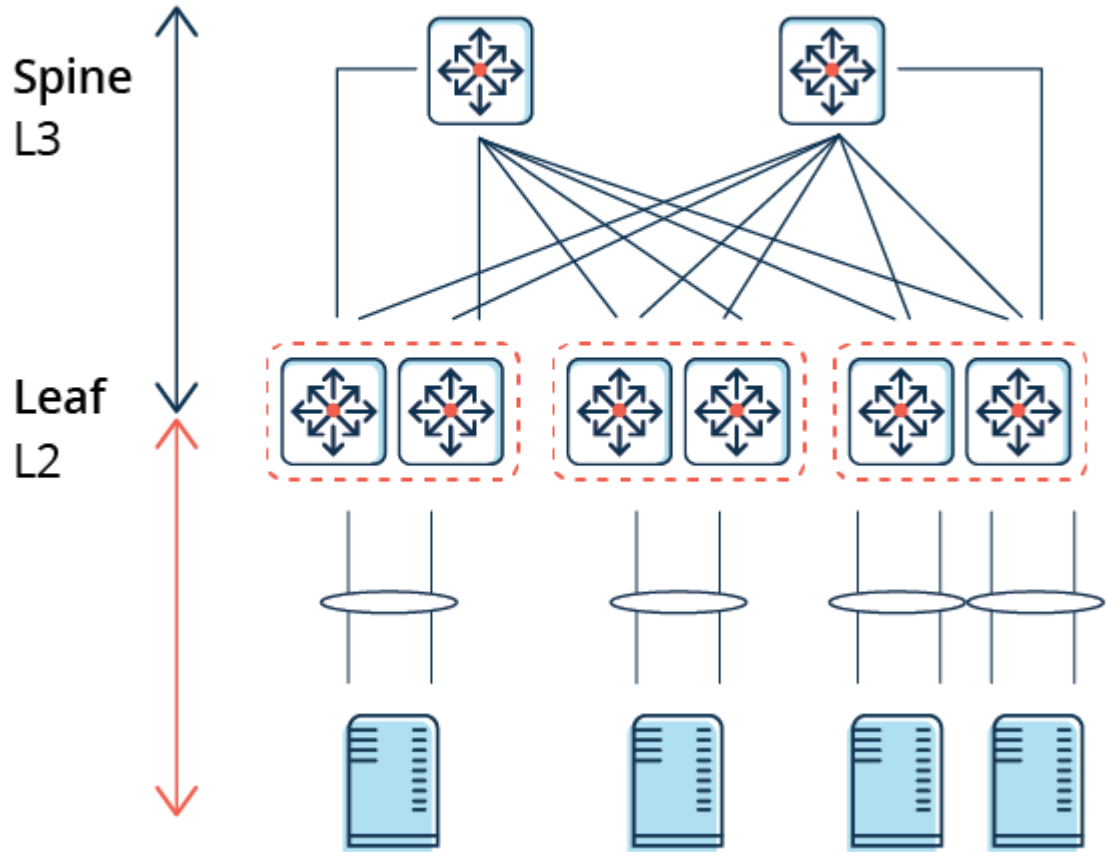


Exemplo de uma estrutura de comutação leaf-spine comum a datacenters em nuvem e outros clusters, como nuvens de borda locais

Traditional 3-Tier Architecture



2-Tier Spine-Leaf Architecture



Tecido Folha-Espinha

No nível folha desempenha três funções

1- Primeiro, disponibiliza uma fábrica de switches os servidores e as VMs a correr nestes servidores num cluster multi-rack.

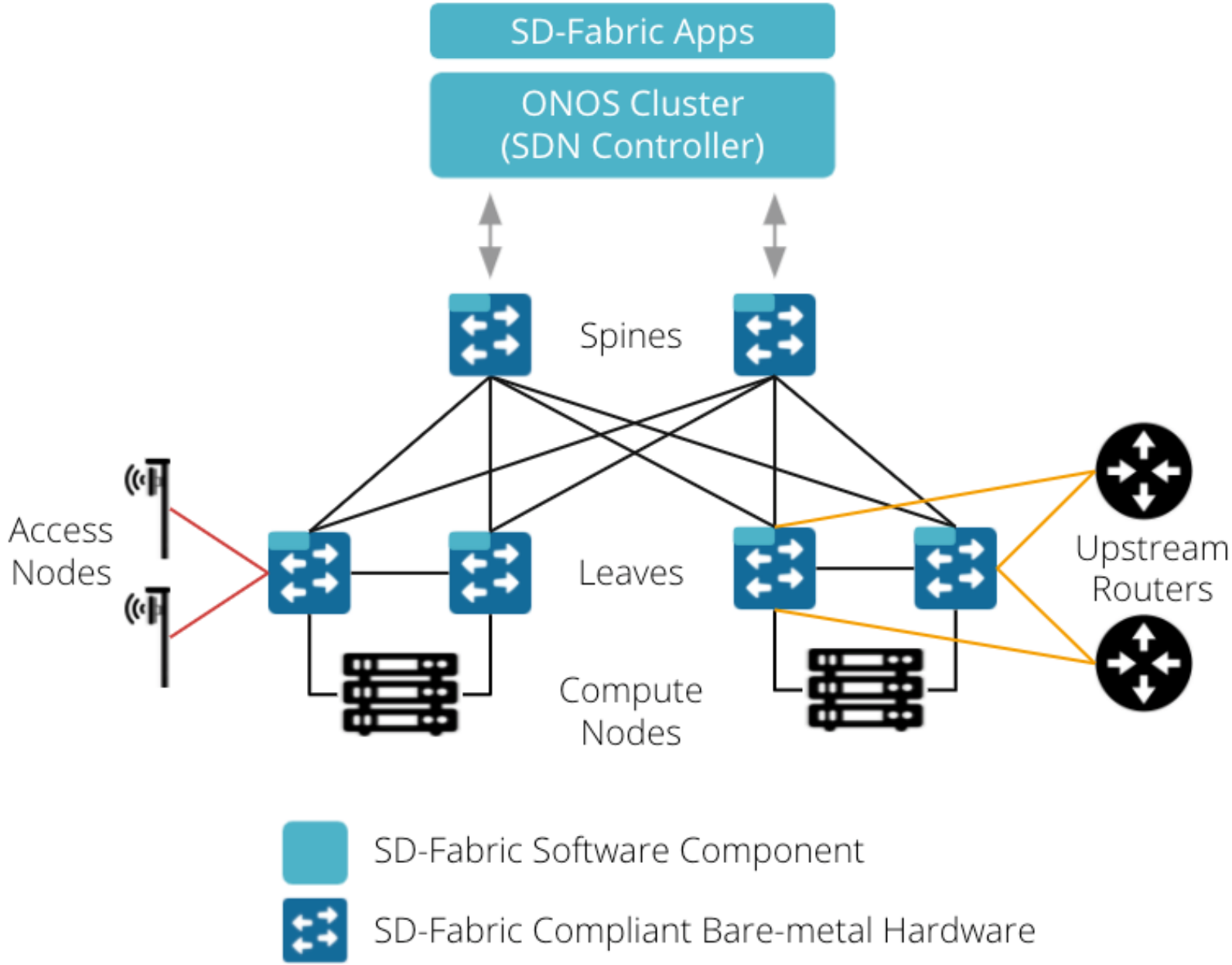
2-Em segundo lugar, ele conecta o cluster como um todo upstream para redes peer, incluindo a Internet, usando BGP (isto é, ele se comporta como um router).

3-Terceiro, ele conecta o cluster como um todo às redes de acesso downstream (termina a rede de acesso com tecnologias como PON e RAN).

Por outras palavras, em vez de pensar na SD-Fabric como uma malha leaf-spine convencional fechada num datacenter, a SD-Fabric é melhor visualizada como um interconexão em execução à borda da rede, ajudando a conectar nuvens de borda a clouds de borda com clouds IP em datacenters.

SDN-Fabric

O que é SDN-Fabric? SD-Fabric é um exemplo de caso de uso para SDN. É um conjunto de aplicações de controle executados em um sistema operacional de rede, que por sua vez é executado em uma coleção de switches programáveis organizados em uma topologia leaf-spine, onde cada switch executa um sistema operacional de switch local.


SD-Fabric Apps
ONOS Cluster
(SDN Controller)
Spines
Access
Nodes
Leaves
Upstream
Routers
Compute
Nodes
SD-Fabric Software Component
SD-Fabric Compliant Bare-metal Hardware

Engenharia de Tráfego

Outro caso inspirado na nuvem é a engenharia de tráfego aplicada aos links WAN entre DCs

B4--- backbone privado da Google construído inteiramente usando switches bare-metal e SDN.

SWAN --- abordagem similar da Microsoft para interligação dos seus data-centers.

O componente central tanto do B4 como SWAN é um programa de control de TE que disponibiliza os recursos de rede de acordo com a necessidade das várias classes de aplicação.

A ideia do TE para redes de comutação de pacotes foram experimentadas pela Arpanet.

O TE só se tornou popular com a chegada do MPLS que disponibiliza um conjunto de ferramentas para balancear a carga entre diversos percursos.

Todavia, uma deficiência notável do MPLS é que o cálculo de percursos é um processo completamente distribuído.

TE para WANs

B4 e SWAN reconhecem essa falha e movem o cálculo do caminho para um controlador SDN logicamente centralizado.

Quando falha um link, o controlador calcula um novo mapeamento do tráfego nos links disponíveis e programa os switches para expedir o tráfego de forma a não sobrecarregar nenhum link.

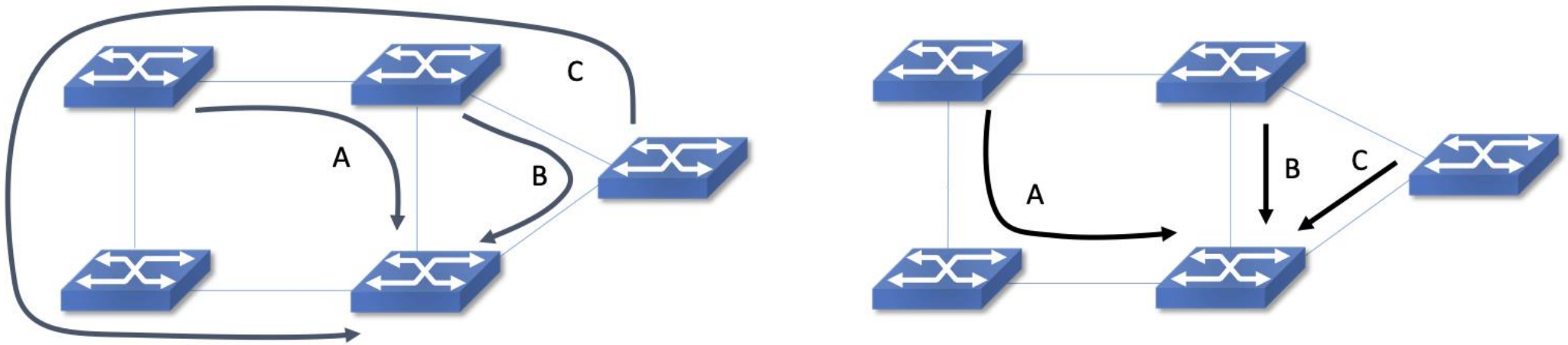
Ao longo de muitos anos de operação, essas abordagens ficaram mais sofisticadas. Por exemplo o B4 evoluiu desde tratar todo tráfego igualmente até suportar uma gama de classes com diferentes classes de tráfego com vários níveis de atraso e requisitos de disponibilidade.

Exemplos de classe de tráfego incluem:

(1) copiar dados do utilizador (por exemplo, e-mail,documentos, áudio/vídeo) para datacenters remotos para aumentar a disponibilidade.

(2) Aceder a armazenamento remoto por cálculos efectuados sobre fontes de dados distribuídas.

(3) Enviar dados em larga escala para sincronizar o estado em vários datacenters.



Exemplo de engenharia de tráfego não ideal (esquerda) e posicionamento ideal (direita).

TE Distribuído e Sub-Óptimo

Isso significa que é quase impossível alcançar qualquer tipo de otimização global, como os algoritmos de cálculo de caminho – que ativam a qualquer momento um link muda de status, ou conforme as cargas de tráfego mudam – a fazer escolhas locais sobre o que parece melhor.

Consideremos o exemplo da Figura. Vamos assumir que todos links são de capacidade um e estamos a tentar encontrar percursos para 3 unidades de tráfego.

Figura da esquerda:

Fluxo A - é colocado primeiro e escolhe um dos percursos mais curto; Fluxo B- é colocado a seguir e escolhe o percurso seguinte mais curto já que o percuso de um salto já foi atribuído ao fluxo A; Fluxo C- Neste caso o único percurso disponível é o mais longo.

Figura da direita:

Usamos um algoritmo central que analisa os 3 fluxos duma vez que acaba com muito menos desperdício. Embora com 1 exemplo artificial, resultados abaixo do ideal são inevitáveis quando não há uma visão central do tráfego.

TE para WANs

Dividindo o tráfego nessas classes com propriedades diferentes e executando um algoritmo de cálculo de caminho para cada um, a equipe conseguiu melhorar consideravelmente a eficiência do rede, enquanto atende aos requisitos das aplicações mais exigentes.

Por meio de uma combinação de centralização do processo de tomada de decisão, limitação de tráfego de forma programática nos remetentes e diferenciando classes de tráfego, o Google conseguiu direcionar suas utilizações de link para quase 100%.

Isso é duas a três vezes melhor do que a utilização média de 30 a 40% que os links de WAN são normalmente aprovisionados, o que é necessário para permitir que essas redes lidem com rajadas de tráfego e falhas de link/switch.

As experiências reportadas com o SWAN pela Microsoft são similares.

Estas experiências de hiperescala com a SDN mostra tanto a capacidade para personalizar a rede e o poder de controle centralizado para alterar as abstrações de rede.

Software-Defined WANs

Outro caso de uso para SDN que se popularizou para empresas foi o SD-WAN.

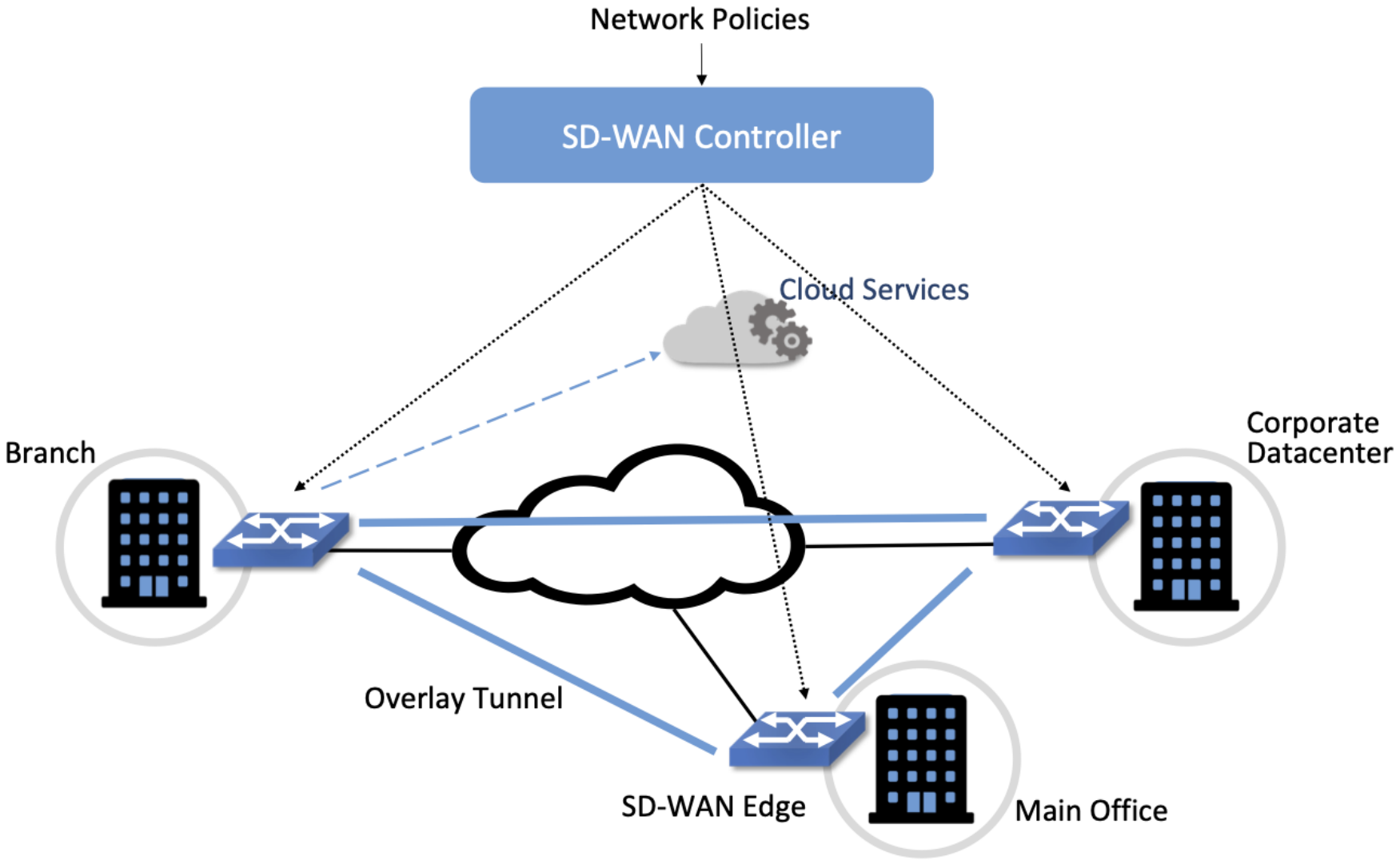
Durante muitos anos as empresas compraram serviços WAN de empresas de telecomunicações, principalmente para obter serviços de rede privados para interligar escritórios principais, filiais e centros de dados corporativos.

Na maior parte do século 21 a abordagem técnica mais popular para construir essas redes foi o MPLS, usando uma técnica chamada MPLS-BGP VPNs.

A rápida ascensão do SD-WAN como alternativa ao MPLS é outro exemplo do poder da centralização do controle.

O aprovisionamento da VPN usando o MPLS embora menos complexo que as opções anteriores, ainda requer alguma configuração em cada site do cliente, e o router do Provider Edge (PE) ao qual o site vai ser conectado.

Para além disso, normalmente exigiria o aprovisionamento dum circuito do lado do cliente ao POP mais próximo da Telco apropriada.



Um controlador SD-WAN recebe políticas centralmente e as envia para switches de borda em vários locais. Os switches constroem uma sobreposição de túneis pela Internet ou outras redes físicas e implementam políticas, incluindo a permissão de acesso direto a serviços em nuvem.

Redes de Acesso

Redes de acesso que implementam a última milha conectando residências, empresas e dispositivos móveis à Internet são outra oportunidade de aplicar os princípios SDN.

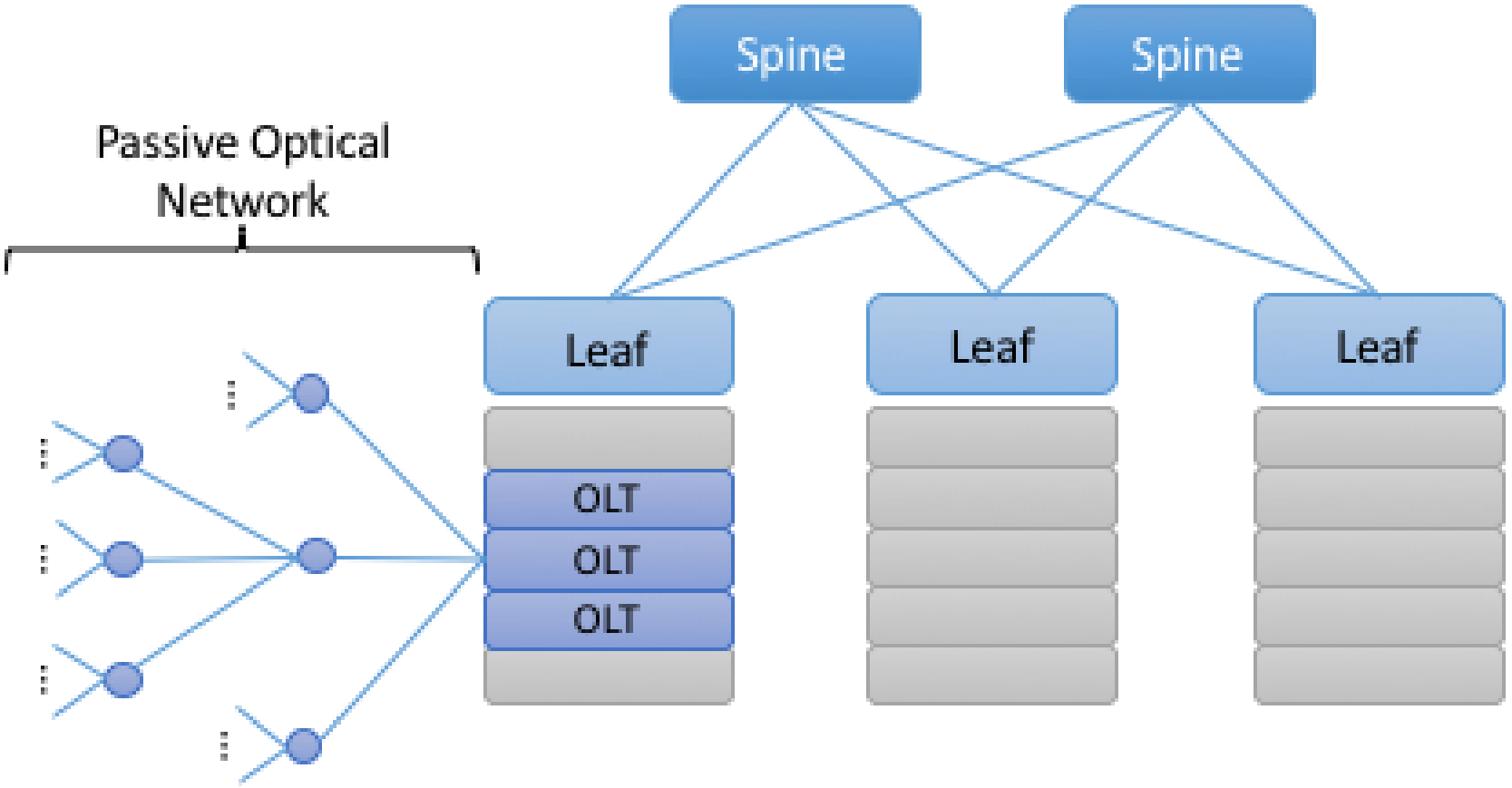
Exemplos de tecnologias de acesso incluem as PON (Passive Optical Networks), conhecidas como fibra para casa e as Radio Access Network (RAN), no coração das redes celulares 4G/5G.

O interessante nestes casos de uso é que ao contrário de outros que são switches de uso geral para controlo programado são construídos de dispositivos de hardware de propósito especial.

O desafio é transformar esses dispositivos nos suas contrapartes comerciais de silício/bare-metal.

No caso das PONs temos as Optical Line Terminals (OLT) and Broadband Network Gateways (BNG).

No caso de redes celulares há dois componentes legados relevantes: eNodeB (the RAN base station)e o the Enhanced Packet Core (EPC).



General hardware architecture of SEBA: SDN-Enabled Broadband Access.

SEBA (SDNEnabled Broadband Access)

Como esses dispositivos são construídos fechados e proprietários, são pior caso para casos de uso dos princípios SDN.

Também significa que sejam a melhor oportunidade para uma maior recompensa, e é por isso que grandes operadoras estão a tentar ativamente encontrar soluções SD para Redes PON (Passive Optical Network) e RAN (Radio Access Network).

Esta iniciativa é algumas vezes referida como CORD (Central Office Re-architected as a Datacenter) e tem sido sujeita a profunda análise de negócio.

O principal desafio de iniciativas como a CORD é desagregar os dispositivos legados de forma a isolar o engenho de expedição de pacotes (central no plano de dados) do plano de controlo.

Assim é possível empacotar o primeiro como hardware e implementar o último em software. O progresso na desagregação das redes de acesso baseadas em PON está bastante avançado, com uma solução conhecida como SEBA em produção.

Virtual Network Function (VNF) e SD-RAN

Tal como o Open Compute Project (OCP) tem switches bare-metal certificados agora também certificam dispositivos OLT bare-metal.

Os switches de malha e os dispositivos de acesso são controlados por um plano de controle definido por software, com o código que implementa esse plano de controle em execução nos servidores do cluster.

Além disso, quando a malha é construída usando switches com pipelines programáveis, certas funcionalidades originalmente fornecidas pelo hardware legado podem ser programadas nos switches que compõem a malha.

A funcionalidade equivalente à BNG que pode ser empacotada como uma VNF a correr num processador de uso geral, pode ser programado em vez disso num switch programável.

Essa prática às vezes é chamada de descarregamento de VNF porque o processamento de pacotes é movido dos servidores de computação para os switches.

Trata-se dum ótimo exemplo do que acontece quando os planos de dados do switch se tornam programáveis: os software é escrito de forma a tirar proveito do hardware em novas e inesperadas formas.

Telemetria de Rede dentro Banda

Visão geral de conclusão dos casos de uso SDN é um exemplo habilitado pela programação de pipelines de expedição: In-Band Network Telemetry (INT).

A ideia INT é programar o pipeline de expedição colectando o estado da rede conforme os pacotes são processados (in-band). Isto constrasta com a monitorização convencional feito no plano de controlo lendo contadores ou amostrando subconjuntos de pacotes (sFlow).

Na amostragem INT as instruções de telemetria são codificadas nos campos de cabeçalho dos pacotes e processados pelos switches conforme eles fluem no pipeline de expedição

Telemetria de Rede (Network Telemetry)

Essas instruções dizem a um dispositivo com capacidade INT que estado colectar e em seguida como gravar esse estado no pacote conforme transita na rede.

As fontes de tráfego INT (aplicações, pilhas de rede de hosts, hipervisores) podem incorporar as instruções em pacotes de dados normais ou em pacotes especiais de sondagem.

Da mesma forma, os coletores de tráfego INT recuperam e relatam os dados coletados resultados dessas instruções, permitindo que os coletores de tráfego monitorem o estado exato do plano de dados que os pacotes observam durante o encaminhamento.

A ideia está ilustrada na Figura, que mostra o exemplo dum pacote que atravessa do switch fonte S1 para o switch destino S5 via um switch de trânsito S2.

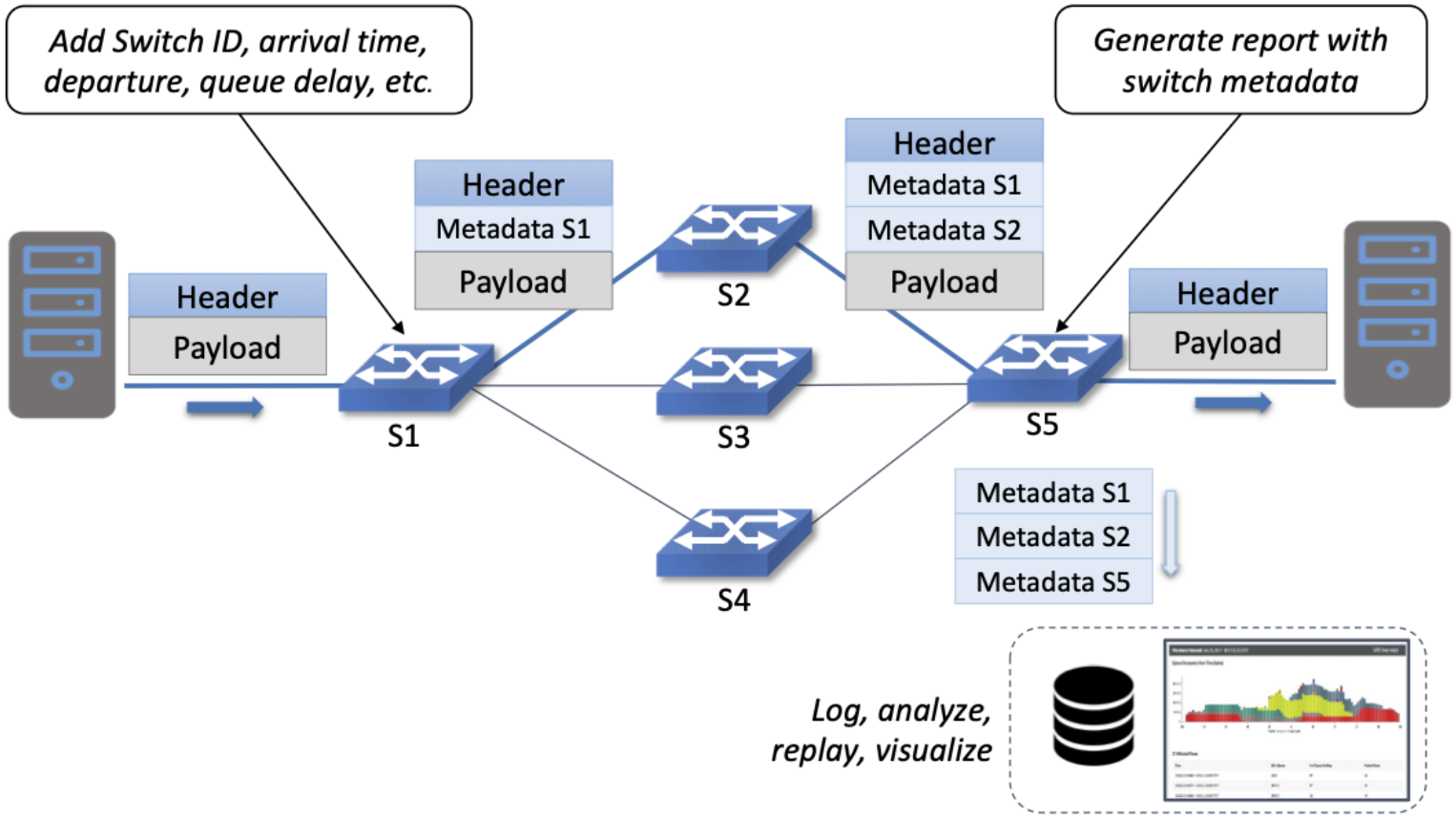


Ilustração da Inband Network Telemetry (INT), em que cada pacote colecta medidas de dados conforme atravessa a rede.