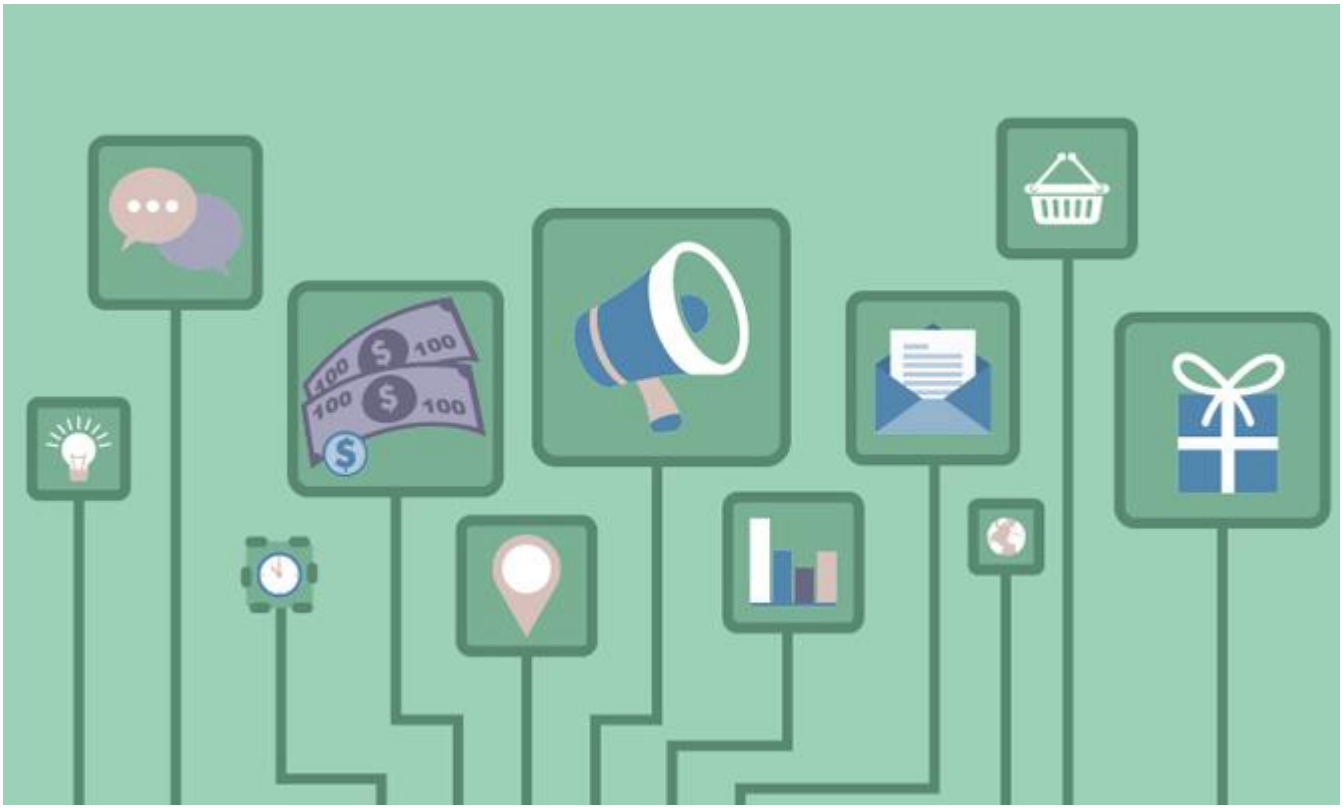
SDN 落地的实践与思考：带着问题找方案，别管定义啦

SDNLAB 君 • 14-12-05

https://www.sdnlab.com/4433.html

编辑按：本文系 InfoQ 中文站对盛科网络软件总监张卫峰的约稿。作者从自身做过的方案和所了解的业界情况为出发点，经过仔细思考，得出本文中对 SDN 当前现状的判断，所以文中难免涉及到作者所在公司之方案、产品，出于参考价值的考虑并未进行删节。



以下为正文：

半年之前写过一篇文章讨论 SDN 的本质，当时就说这会是一个系列文章，后面还要写一下 SDN 的落地。这一等就是半年多，其间有朋友问过我为什么还不写——不是我不想写，是有些问题还没想清楚。直到现在我也不敢说我完全想清楚了，但是毕竟有了更多的实践，实践过程中不停地思考，加上不断有媒体朋友找我约稿，我想还是写一写吧。这篇文章可以看作是对我三年 SDN 工作历程的一个总结和反思。我在这篇文章会回顾一下人们对 SDN 定义的争议，SDN 的落地实践，阻碍 SDN 落地的一些障碍，给出一些对 SDN 落地的建议和看法。

SDN 的定义回顾

现在大多数人对 SDN 的定义是控制跟转发分离+开放的编程接口，包括 Gartner

也是这样的定义。Gartner 的数据中心云计算行业分析总监曾绍清告诉我，他们认为思科的 ACI 不是 SDN，因为 ACI 并非是控制和转发分离，它只是把策略管理的功能分离到了控制器上，控制协议（OSPF、BGP 等）仍然运行在交换机上。但是我曾经在国外著名的通信技术媒体 lightreading 上看到他们综合一些专家的意见给出的定义，该定义里面只提到了开放的可编程接口以及由此带来的业务敏捷性。就我个人的观点来看，我更倾向于 lightreading 的定义（具体请参阅我的第一篇文章）。不过，正如青云 CEO Richard 接受 InfoQ 采访的时候说过的，每个人对 SDN 都有不同的定义，这个并不重要，我深以为然，重要的是 SDN 带来的价值。所以，以后大家不要把精力浪费在讨论 SDN 的定义上吧。

SDN 在网络虚拟化中的应用

总有人说没看到 SDN 有落地案例，但是你去看一些国外专业的咨询机构，总能看到他们的报告中，SDN 的市场份额在逐年增加，且趋势向好。是有人在撒谎吗？No。

咨询报告中说的 SDN 市场份额在增加，主要是强调 SDN 现在最大的一个应用场景——网络虚拟化中的应用。很多人说没看到 SDN 有落地案例，那是因为他们潜意识里面只把控制器+硬件 SDN 交换机的应用认为是 SDN 应用，云平台+虚拟交换机他们认为不是 SDN。而事实上，以 VMware 的 NSX 为代表的网络虚拟化的应用早已经是被广泛认可的 SDN 的典型落地案例。

目前看到的基于 SDN 的网络虚拟化解决方案有以下三种：

1、纯软件方式，以 VMware 的 NSX 为代表。除了 NSX，还有 Juniper 的 Contrail、Midokura 的 MidoNet 以及 Vyatta、Nuage、Plumgrid 等公司的商业网络虚拟化方案。这些公司的实现方式都不太一样，但是都在不同程度上用到了 SDN 技术。有的只是把一些策略管理的东西放在控制器上，转发表项还是由虚拟交换机自己来生成，而有的则是控制器来下发转发表项。而目前影响最广泛的 OpenStack 的网络组件 Neutron，则两种方式都支持，Neutron 更是一种标准的 SDN 架构。由于本文的目的不是介绍技术细节，所以这里就不深入展开来讲了。

2、硬件方式，以思科的 ACI 为代表，即将网络虚拟化在硬件中实现（当然也不排除会用到 vRouter）。具体 ACI 的架构，我之前也写过一篇文章，可以参阅一下。

3、软件+硬件方式。盛科网络推出的 SDN 方案即属于此类（Arista 也有类似方案），本质上它是一个软件方案的思路，只是把部分对性能影响最大的操作 offload 到硬件 SDN 交换机，可以认为是一个超级网卡。并且它为 NSX 之类的软件方案提供了 SDN 交换机作为 Tunnel Gateway 来满足物理服务器跟虚机混合组网的需求。

华为和华三也都相应的都有自己的解决方案，只是目前看到的他们都是推整体云计算解决方案，没有把网络部分整出一个方案来单独卖。

无论纯软件还是硬件的 SDN 解决方案，在云计算数据中心里面，应用的越来越广泛，所以如果要谈 SDN 的落地，这是目前最大的，最不容忽视的一个。

SDN 在别的领域的应用

除了在网络虚拟化领域的应用，SDN 交换机在别的领域也有一些应用，但是从应用广度和影响力来看，都比不上在网络虚拟化中的应用。从我们自己以及国外一些案例来看，落地的 SDN 的应用，其驱动力主要可以归结为两大类：业务层面灵活性的需求，转发层面灵活性的需求。前者的价值要远大于后者。

一、业务层面灵活性的需求

这主要是强调可编程。通过开放的可编程接口，提供给用户原来无法获得的对网络配置管理和策略部署的灵活性控制。前段时间著名的运营商亚太环通（Pacnet）宣布在天津的一个 IDC 正式启用，他们宣称里面使用了 SDN 技术来为用户提供自助调整带宽的功能，其实该功能早就部署在了他们新加坡、澳大利亚，香港等国家和地区的其他数据中心。该功能是通过定制化的 SDN 交换机实现的，其中的千兆交换机是盛科提供的。这是一个很典型的体现业务灵活性的例子。用户可以在他们提供的一个界面上，随时按需修改自己的出口带宽。而且不仅如此，一个用户可能租用了他们多个数据中心，通过 SDN 创建的 MPLS 隧道把用户在多地的数据连通之后，可以通过 SDN 动态调整这些隧道的带宽，一旦出现故障或者拥塞，还可以自动重新选路。没有 SDN，要做到这一步很难。

但是大家更关心的是 SDN 在企业网里面如何用。并不是所有企业网都适合使用 SDN，什么样的企业网需要用 SDN 呢？这个问题后面再分析。国外一个著名设备商 N，他们有挺多的 SDN 案例，特别是有些案例规模还是较大的，不像某些公司挂羊头卖狗肉的宣传，我至少知道他们有一个案例是很值得拿出来讲一讲的（这个案例在国外网站上也有介绍）。他们给某电视台的一个新的网络进行了 SDN 化设计，该网络有一个特点，就是拓扑和策略都是灵活易变的，比如这个星期是为一个大型演出节目准备的，而下个星期就变成为一个体育节目准备，如果没有 SDN，他们需要靠人工去插拔线修改拓扑，重新划分物理和逻辑网络等，非常麻烦，在人工很贵的国外，这个问题特别突出。使用了 SDN 之后，整个物理网络基本不动，每次就依靠 SDN 将网络重构。这个案例还包括无线 AP，也是 SDN 化的。而且值得关注的是，他们并非全部使用 SDN，而是一个混合的网络，既有 SDN，又有传统的。即在需要 SDN 的时候 SDN，不需要的时候就用传统，深得 SDN 的精髓。

在我们碰到的案例中，也有一个复杂度没这么高，但是需要对网络灵活控制的。该网络管理员说他管理了一个较大的实验室，每天都有人在里面做不同的实验，对这些不同的人，网络中都需要有一些不同的安全控制策略，每次都去找他该配置，他不胜其烦。而这个时候，如果建立一套用户权限体系，用户可

以自行登录申请，一旦认证通过，根据他的权限，控制器可以自行下发安全控制策略到交换机上，SDN 的业务灵活性充分体现出来。

二、转发层面灵活性的需求

这主要是针对一些非常特定的场景，主要是为了匹配或者修改特定字段，通常是传统交换机不支持的（其实芯片也许能支持，只是交换机系统没做）。这些场景我们也碰到不少，比如用来做 DDoS 防攻击（日本 Sakura Internet 的应用），用来做负载均衡+NAT，用来做 TAP 应用（价格是专业的 TAP 设备的至少 1/5），用来将 PPPoE 跟 IP 区分开并灵活控制等等。还有一些用户提出来过，但是需要辅助 FPGA 或者 NP 才能做到的。这类应用主要的灵活性体现在转发面上而不是控制面。

SDN 落地的障碍

硬件 SDN 的落地进展并不顺利。虽然现在慢慢有了一些更多的案例，但是离规模部署还很远。我跟 Gartner 的曾绍清一起探讨过原因，曾说 Gartner 经过调查，形成了一些他们的看法。

Gartner 的观点认为，以下几个问题阻碍了 SDN 的落地：

1、厂商的直接支持而欠缺传统渠道的支持

2、SDN 的革命性变革而使销售难度大增，传统厂商偏向销售 Ethernet Fabric 等容易接受的产品

3、SDN 的用户价值较难从单一产品成本分析中体现

4、用户的开发团队开发的东西，运维团队不接受

我个人觉得 Gartner 的观点都很有道理，相对来说看得比较宏观。我根据我们的客户交流和项目实践中碰到过的一些问题，也谈谈我的看法。我的观点其实跟 Gartner 有不少相通之处，算是一枚硬币的两个面。我认为一个用户要想把 SDN 在他的网络中落地，必须同时满足这三个条件，缺一不可。而现实中，这三个条件同时满足的不多，这也导致了 SDN 的落地缓慢。

1、用户必须清楚地知道自己网络中存在的问题，然后带着这些问题来寻找方案。我经常碰到一些人问我，你帮我看看 SDN 能用在我们网络中什么地方？这种用户是没办法让 SDN 落地的。SDN 是用来实现用户的业务敏捷性的，不是用来全面替代传统网络的，如果你都不知道自己有什么问题，怎么引入 SDN？我碰到的最终能落地的，都是明确知道自己网络中的问题，迫切想找方案来解决的。

2、用户做决策的人必须要足够有魄力，而且能够协调开发部门（或者第三方开发）和使用部门之间的关系。某互联网大厂告诉我，他们的自研交换机项目之所以能成功，全面在自己的网络中替换商用交换机，就是因为他们的研发和运维都归一个领导管，这个领导要求运维部门必须用自己研发的交换机，有问题也在所不惜。而其它大厂之所以进展不顺，则恰好相反，研发部门和运维部门

彼此独立。SDN 这里也是如此，如 Gartner 所言，SDN 的革命性变革，必然导致传统运维使用上的不适用，人都有使用自己习惯的东西的惰性，如果没有强制命令来保证运维人员使用新的工作方式，确实会比较难推。盛科就给一个互联网大厂做过一个 SDN 项目，该项目很顺利地解决了一些核心技术需求，但是反倒是在推到运维那里的时候碰到了障碍。其实那些障碍可大可小，如果严格按照传统运维规则去要求，那就会阻碍重重，但是如果愿意给与新生事物足够的耐心，让它在使用中慢慢完善，那就可以顺利推行下去。这都取决于决策人员的魄力和权责范围。

3、用户的研发部门或者第三方研发人员必须有足够的研发能力，能够有充分的理性选择合适的技术。整个 SDN 体系中的核心是什么？是交换机吗？是控制器吗？都不是！核心是应用程序。在 SDN 中，用户自己或者用户委托的第三方必须有足够的能力去研发上层应用软件，必须知道这些应用软件如何去通过控制器控制交换机。很多人通常会问 SDN 交换机厂商：你们除了交换机，还有控制器卖吗？我假设我们有，你拿去就能用吗？不能！因为设备商提供的控制器不知道用户要用来做什么应用，所以它实际上提供的只是一些基础 API 以及实现这些 API 的内部逻辑，如何用这些 API，那是用户或者用户委托的第三方需要去考虑的事情。国外的 SDN 为什么部署得比国内多？至少我看到的原因之一是，国外有一些独立的第三方的 SDN 应用提供商，他们有能力架设起最终用户和 SDN 设备商之间的桥梁，把用户需求和 SDN 设备以及控制器结合在一起，打包交付给用户。比如前面讲的亚太环通的 SDN 应用，就是一个第三方软件提供商把盛科 SDN 交换机、另外一个厂商的 SDN 光传输设备、开源的控制器加上他们自己的应用程序结合在一起，一起交付给客户。而且他们进行技术选择的时候，非常理性不会刻意地去追求标准，他们追求的是满足客户需求，所以有不少私有化的扩展。盛科推到欧洲、日本、美国去的 SDN 设备，也都是因为有强有力的第三方合作伙伴或者客户本身有强有力的研发能力。否则 SDN 交换机只能在实验室玩玩。我也很遗憾地看到过一些反面例子，本来他们或者他们的客户确实有 SDN 的需求，但是他们自己不愿意或者没有能力去针对控制器做二次开发，也不愿意花钱去请第三方开发，而在没有量的保证的时候，设备商也不愿意去做太多定制开发，最终导致落地受阻。

OpenFlow 的局限性

OpenFlow 是最广为人知的 SDN 技术，但是并非唯一。而且实践证明，仅仅靠 OpenFlow，很多事情做不了，OpenFlow 可以应用的场景非常狭窄。

关于 OpenFlow 本身的技术缺陷，很多文章都提过，我之前的书和文章里面也都详细分析过，诸如当前交换芯片支持的流表数量都有限，报文匹配和动作都不够灵活，都无法支持很多级流表等等。这些分析都是对的，但是我要告诉大家的是，这些限制根本不足为惧。为什么这么说？因为这些都是从 OpenFlow 技术规范出发得出来的结论，而不是从 SDN 应用的角度得出来的结论，换句话

说，SDN 的应用，未必真的需要 OpenFlow 规范里面提到的所有技术，所以就算有限制，问题也不大。OpenFlow 真正的限制在别的地方。

运维管理的缺失

还是以我们给那个互联网大厂做的项目为例，该厂所要求的一个核心技术点别的交换机都做不到，只有盛科的能做到（因为盛科是用自己的芯片，恰好支持该功能），而且该技术也能按照客户要求使用 OpenFlow 配置出来，一时皆大欢喜。但是当该产品要转运维的时候，问题来了，运维部门要求所有入网的设备，都要满足他们的运维要求，诸如 SNMP 管理、能够查看统计、能够 ping 通该设备、能够 telnet 该设备、能够通过 Radius 到远程服务器进行管理员身份认证等，这些对交换机来说是再正常不过的基本需求了，但是所有这些东西在 OpenFlow 上都没有定义。当然你可以辩论说这属于管理面的，不属于 OpenFlow 的定义范畴，OpenFlow 只定义转发面和控制面功能，但是管理层面的不少功能依赖于转发面，比如管理员想通过带内口 ping 通交换机以便检验路径的可达。还有运维人员希望交换机能支持基本的 LLDP 协议来进行邻居发现。另外一个互联网公司也给我们提出过类似的要求。
运维管理功能的缺失导致了传统运维人员的抵触是可想而知了。这光靠 OpenFlow 是无法解决的，需要引入传统的东西。

跟传统网络的交互

用户网络中通常都是存在很多传统设备的，不太可能为了引入 SDN 而把这些设备都抛弃，所以这就涉及到一个问题，需要 SDN 设备跟传统设备互通。比如有一个三层汇聚交换机，该交换机会向下发送 ARP 获取下联设备的 Mac，如果下面是个传统的主机或者三层设备，它会自动回复 ARP 请求，但是 OpenFlow 交换机没这能力，它只能把报文发送到控制器，让控制器回复，但是很多用户不想在控制器上进行开发来支持这种非核心业务。而且，实事求是的说，最高效的做法肯定是在交换机上进行回复。
还有更复杂的例子。曾经有一个软件开发商，使用盛科的交换机给一个电商开发 WAN 网的流量调度，它需要跟传统交换机进行路由协议交互，如果不在交换机上运行路由协议，就要在控制器上运行。在控制器上运行路由协议，会让控制器很复杂，而且效率低下，况且，这也并非该软件提供商的核心价值，他们也没这个能力去在控制器上做一个路由协议并把它做稳定，所以他们希望交换机上做。SDN 交换机对他们来说，核心价值是让他们可以控制报文的转发路径，从而动态调度流量，至于跟传统网络的交互，他们不希望重复发明轮子，而是希望借用交换机的传统能力。
以上两个问题，并非是说靠纯 OpenFlow 交换机完全无法满足，如果在控制器上做得足够复杂且不考虑效率，也是可以做到的。我们就有一个云计算的客户，使用我们的纯 OpenFlow 交换机，完全靠自己开发的控制器去进行必要协

议报文交互（主要是 ARP）和各种其它必要的控制，他们之所以能做到这一点，就是因为他们本身有很强的研发团队。对于大多数人来说，要走这条路是很难的，那解决方案是什么呢？就是同时支持 OpenFlow 和传统二三层处理的混合型交换机。

SDN 落地的建议

根据以上的分析，为了加快 SDN 落地，对用户、对 SDN 提供商、对整个行业，我有如下建议。

1、要清晰地认识到 SDN 并非适合所有场景。什么样的场景适合 SDN？前面说过，SDN 应用的两大驱动力：业务层面灵活性的需求和转发层面灵活性的需求，如果你的网络足够复杂（复杂并非是规模大），一些配置管理、安全策略、流量调度策略、拓扑等经常需要变化，那非常适合 SDN，最典型的就是网络虚拟化（频繁的虚机增删、虚拟网络的变化）以及 Google B4（路径经常需要随着带宽的变化而变化）。或者你的网络中某个特定功能，在转发面上需要灵活的报文匹配或者报文编辑，传统网络的固定模式无法满足，那也可以考虑 SDN。

2、对于普通企业来说，我的建议是不要追求 SDN 设备接口的标准化，而是要追求接口的开放性和灵活性，因为你想需要的不是技术标准，而是要能解决你的实际问题。对于运营商或者必须要求引入多家设备提供商的大型互联网公司，如果你要引入 SDN，不要去追求南向接口的标准化，而把精力放在北向接口的标准化上。通过让每个厂商提供插件来适配北向接口的做法，来屏蔽各个厂商的差异，这是最现实的做法，否则推动起来会阻力重重，因为各个厂商都不愿意提供跟别人完全一样的设备编程接口。这一点上可以借鉴 OpenStack 的网络组件 Neutron 的做法。

3、正确认识 OpenFlow 的作用。不要指望纯 OpenFlow 能够解决你的所有问题。真正能给复杂网络带来价值的 SDN 设备必定是混合型设备，而且这种混合不是简单的部分端口支持 OpenFlow，部分端口支持传统路由交换，而一定是在报文处理流程中，OpenFlow 和传统二三层处理混合在一起。让 OpenFlow 去控制你想控制的部分，其它部分对你来说无业务价值，就让它们走传统处理流程就可以了。这样跟传统网络互通性的问题，运维管理的问题也都很容易就解决了。

4、不要期望整个网络全部都 SDN 化。SDN 的价值不在于让用户控制一切，而在于让用户去控制他需要控制的地方，无业务价值的部分，完全不需要 SDN 的参与。无业务价值的部分，有的时候存在于边缘，有的时候存在于汇聚和核心，完全看场景而定。

5、如果有人要进行 SDN 创业，创业的着眼点一定不要放在 SDN 交换机和 SDN 控制器，而是要放在 SDN 应用上。控制器你可以找一个开源的拿过来修改一下就行，比如 OpenDayLight, Ryu 等，SDN 交换机可以跟专业的 SDN 交换机厂商合作，但是应用部分是离最终客户最近的，最能体现价值的部分。SDN 的落地需

要这样专业的第三方 SDN 应用提供商。

6、不要动辄问设备商你的设备是否支持匹配 12 元组，是否支持多级流表，否则我会反问你，你为什么需要匹配 12 元组？为什么需要多级流表？不要只把支持 OpenFlow 的交换机认为是 SDN 交换机，没支持 OpenFlow 就认为是忽悠你，你要问的是，它开放的可编程接口是否能满足你的需要。同理，不要以为控制器就应该是支持 OpenFlow 的，不支持 OpenFlow 的控制器你就认为是忽悠，你要看的是它是否能通过开放的接口去控制交换机。对于 OpenFlow 交换机，不要认为只有使用 ACL 表实现的 OpenFlow 才是 OpenFlow，使用传统表项组合出来的流表就不是 OpenFlow，就是忽悠，你要问的是，使用传统表项组合出来的流表是否能满足你的需求。

7、无论是用户，还是 SDN 设备、方案提供商，一定不要期望你可以做一个批量复制的东西出来，SDN 必然意味着定制化。这是一把双刃剑，一方面它可以通过定制给用户提供真正的灵活性，但是另外一方面，太多的定制导致它难以被快速推广，大型设备商的规模优势无法体现，无法依靠传统渠道去推广而不愿意去定制，而小的设备商限于人力，也没法去做太多定制。所以需要专业的第三方提供商的出现。

8、在没有规模部署的前提下，不要去期望 SDN 设备会有成本优势，相反，因为定制化的研发投入，SDN 整体方案的成本反而会增加。对于用户来说，要关注的是 SDN 所带来的运维成本的下降。

9、如果你要部署 SDN，必须打消买过来就能用的不现实的期望值——至少目前是这样。在你立项或者购买 SDN 设备之前，你必须要问自己，Am I ready? ready 的意思就是你需要自己有懂业务的研发团队或者愿意购买第三方的服务，或者愿意付钱让设备商给你做定制开发（如果设备商愿意的话）。

总结

我们要正确地认识 SDN，不要过高估计 SDN 的能力，也不要对 SDN 丧失信心。SDN 不会取代传统网络，甚至看不到它有占据垄断地位的可能，但是它肯定会是现有网络的一个强力补充。SDN 落地不要太在乎标准化，要着眼于开放性。SDN 落地不仅呼吁第三方应用提供商的出现，更重要的是，SDN 用户企业中的决策者，要有足够魄力，敢于承担风险，愿意在使用中完善 SDN，要勇于拍板。国外的 Google，Facebook 有这个魄力，NTT 有这个魄力，Verizon 有这个魄力，Pacnet 有这个魄力，国内的公司没理由在这方面落后于他们。我们欣喜地看到国内某些公司已经在赶上，腾讯就是一个很好的榜样。

最后也要给所有要学习 SDN 的朋友，特别是学生朋友一个建议：学习 SDN，必须要有基本的网络知识作为基础，不懂网络就想学习 SDN 这是不现实的。

关于作者

张卫峰（@盛科张卫峰），盛科网络软件总监，数据通信和芯片设计领域资深专家，有十几年的网络实践经验，对 SDN、传统二三层交换机、数据传输设备

（PTN 和 IPRAN），从管理面到协议控制面一直到芯片转发面都有着深刻的理解。

转载自：http://www.infoq.com/cn/articles/sdn-practice-and-thinking-problem-plan#0-tsina-1-5746-397232819ff9a47a7b7e80a40613cfe1