

什么是 CPU 和 CPU 占用率

文档版本

01

发布日期

2020-11-13



版权所有 © 华为技术有限公司 2020。保留一切权利。

非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

商标声明



HUAWEI和其他华为商标均为华为技术有限公司的商标。

本文档提及的其他所有商标或注册商标，由各自的所有人拥有。

注意

您购买的产品、服务或特性等应受华为公司商业合同和条款的约束，本文档中描述的全部或部分产品、服务或特性可能不在您的购买或使用范围之内。除非合同另有约定，华为公司对本文档内容不做任何明示或默示的声明或保证。

由于产品版本升级或其他原因，本文档内容会不定期进行更新。除非另有约定，本文档仅作为使用指导，本文档中的所有陈述、信息和建议不构成任何明示或暗示的担保。

华为技术有限公司

地址： 深圳市龙岗区坂田华为总部办公楼 邮编： 518129

网址： <https://www.huawei.com>

客户服务邮箱： support@huawei.com

客户服务电话： 4008302118

目 录

1 什么是 CPU 和 CPU 占用率? 1

1.1 简介..... 1

1.2 CPU 和 CPU 占用率定义..... 1

1.3 CPU 处理报文原理.....3

1.4 CPU 占用率高造成的影响.....4

1.5 CPU 占用率高属于正常现象的场景..... 5

1.6 CPU 各任务名称及功能说明..... 6

1.7 相关信息.....7

1 什么是 CPU 和 CPU 占用率?

- 1.1 简介
- 1.2 CPU和CPU占用率定义
- 1.3 CPU处理报文原理
- 1.4 CPU占用率高造成的影响
- 1.5 CPU占用率高属于正常现象的场景
- 1.6 CPU各任务名称及功能说明
- 1.7 相关信息

1.1 简介

本文档简要介绍了CPU和CPU占用率的定义、CPU处理报文的原理、以及CPU占用率高造成的影响。

1.2 CPU 和 CPU 占用率定义

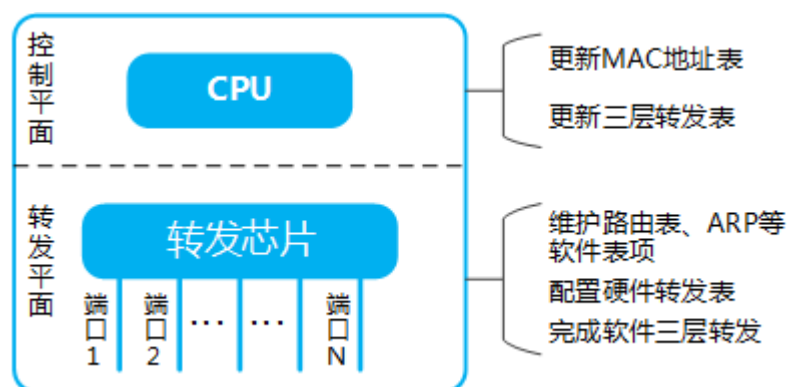
交换机的核心--CPU

交换机采用分布式架构，主要包括转发平面和控制平面。其中转发平面用于实现二三层转发；而控制平面主要用于实现转发的控制。

如**图1-1**所示，交换机的控制平面采用通用嵌入式CPU实现，转发平面采用转发芯片实现：

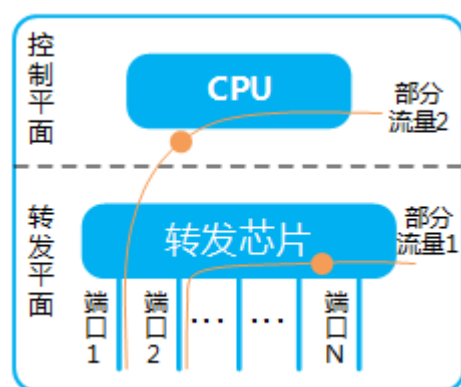
- 转发芯片完成主要的二三层转发功能，如更新用于二层转发的MAC地址表以及用于IP转发的三层转发表。转发芯片的特点是能够实现大吞吐量的数据转发。
- CPU主要维护软件表项（如路由表、ARP表等），并根据软件表项的转发信息来配置转发芯片的硬件三层转发表。同时，CPU本身也可以完成软件三层转发。CPU的特点是数据处理能力低。

图 1-1 交换机的分布式架构



在网络中，可将报文按照功能分为控制报文和数据报文。当交换机上还未建立任何硬件转发表项时，如果报文到达交换机，首包由 CPU 转发并建立三层转发硬件表项，后续包（即非首包）流量由入端口进入转发芯片，如图 1-2 所示：

图 1-2 交换机对非首包流量的处理



- 后续包的**部分流量1**（一般为数据报文）直接由转发芯片转发出去这部分流量不经过 CPU，所以该流量的处理也不消耗 CPU。这部分流量一般为数据报文。
- 后续包的**部分流量2**（一般为控制报文和部分数据报文）经由转发芯片上送 CPU，由 CPU 来决定是否需要将其转发出去或直接在 CPU 终结。该流量需要消耗 CPU，不能进行高速转发。

总的来说，真正决定交换机高速交换转发的是转发芯片中的二三层硬件表项，而转发芯片的硬件表项来源于 CPU 维护的软件表项。可以看出，CPU 是交换机的核心。

CPU 占用率

交换机正常启动后，为了维持系统的正常运行，CPU 上有多个活跃的任务用于完成对设备的管理、监控和三层表项学习。通常交换机支持的特性越多，系统运行的任务也越多。

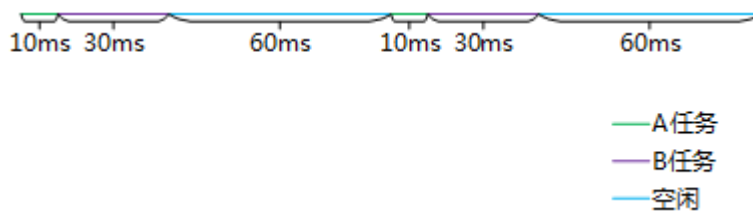
设备的 CPU 占用率指一段时间内系统中非空闲任务占用 CPU 处理的时间比率，是对设备 CPU 使用情况的一个宏观统计，具有以下几个重要特点：

- 持续变化性：系统的 CPU 占用率不是保持不变的，它是随着系统的运行和外部环境的变化而持续变化的。

- 非实时性：系统的CPU占用率反映的是一个CPU统计周期内的CPU使用情况，并不是特指某一个时间点的实时占用情况。
- 实体强相关性：CPU占用率是以物理CPU为粒度进行统计的，通常而言，设备上的每块业务板均有一个单独的物理CPU，因此它们的CPU占用率都是相互独立的。

CPU占用率表示交换机在某个时间点的运行任务情况。如图1-3所示，A任务占用10ms，B任务占用30ms，然后空闲60ms，再又是A任务占10ms，B任务占30ms，空闲60ms。如果在一段时间内都是如此，那么这段时间内的占用率为40%。CPU占用率越高，说明交换机在这个时间上运行了很多任务，反之则很少。

图 1-3 任务占用 CPU 运行时间



可以看出，CPU占用率的高低与CPU的强弱有直接关系，因此，CPU占用率是衡量设备性能的重要指标之一。

1.3 CPU 处理报文原理

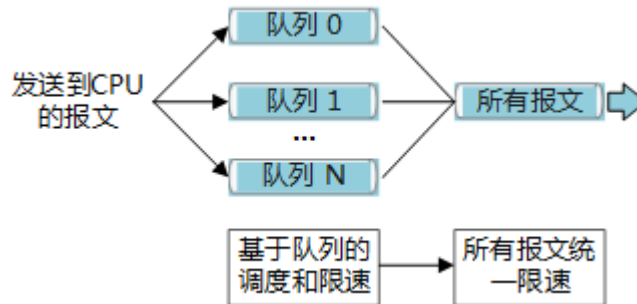
交换机由转发芯片转发普通数据报文，无需CPU参与。以下场景会将报文发送给CPU处理：

- 需要交换机终结的协议报文
所有目的地址为本机的报文均需要上送CPU处理：
 - 各种协议控制报文，如STP、LLDP、LACP、DLDP、EFM、GVRP、VRRP等
 - 路由更新报文，如RIP、OSPF、BGP、IS-IS等
 - SNMP、Telnet、SSH报文
 - ARP、ND回应报文
- 需要特殊处理的报文
 - 带option选项的ICMP报文
 - 带hop-by-hop选项的IPv6报文
 - TTL小于或等于1的IPv4/IPv6数据报文
 - 目的IP地址为本机的数据报文
 - ARP/ND/FIB Miss报文
- 应用了ACL，需要CPU处理的报文
 - 开启logging功能后，通过ACL deny动作丢弃的报文
 - 流策略重定向到CPU的报文
- 组播特性相关的报文
 - IGMP协议报文

- 未知IP组播流
- 其他特性的相关报文
 - DHCP协议报文
 - ARP、ND广播请求报文

交换机根据报文的权重，将上送CPU的不同类型的报文划分到优先级不同的多个队列，确保重要报文优先处理。同时，还可以针对上送CPU的报文进行限速，使单位时间内上送CPU报文的数量限制在一定的范围之内，从而保证CPU对业务的正常处理。

图 1-4 将不同类型报文分发到 CPU 队列



在稳定的网络环境下，上送CPU的报文数量控制在适当的范围内，CPU占用率也稳定在一个合理的区间。如果一段时间内上送CPU的报文数量过大，则CPU会因为忙于处理这些报文而表现为CPU占用率过高。

1.4 CPU 占用率高造成的影响

当设备转发面上送CPU的报文速率过快（如因网络环路导致CPU短时间内收到大量报文）或者某任务长时间占用CPU时，CPU将高负荷运行，可能无法及时调度其他任务，进而引发业务异常。

CPU 占用率过高会影响系统处理能力，导致网络业务表现不如预期，可能导致出现的网络故障现象有：

- 交换机不能响应正常的管理请求。
 - Telnet或SSH会话不能建立，导致无法管理设备或者设备反应慢，命令执行有延迟等。
 - SNMP超时。
 - MAC/IP Ping耗时很长甚至超时。
- 交换机不能及时转发或回应客户端请求，导致DHCP失败或IEEE 802.1x认证失败。
- STP拓扑改变甚至出现网络环路。

交换机通过CPU周期性的接收BPDU报文维持其Root/Alternate端口角色，如果因上游设备CPU繁忙导致BPDU报文不能及时发出或本机CPU繁忙不能及时处理收到的BPDU报文，交换机会认为到根桥的原路径故障而重新选择Root端口，引起网络重新收敛；如果交换机原来同时存在Alternate端口，则将Alternate端口作为新的Root端口，这时就可能导致网络出现环路。

- 路由拓扑改变。

动态路由协议的保活由CPU完成，如果因CPU繁忙不能及时接收和发送hello报文，就会导致路由震荡，如OSPF震荡、BGP震荡、VRRP震荡。

- 可靠性检测协议震荡。
802.3ah、802.1ag、DLDP、BFD、MPLS OAM等检测协议均由CPU完成定时保活，如果因为CPU繁忙不能及时接收和发送协议报文，将会导致协议震荡，进而影响相关业务流量转发。
- LACP类型的Eth-Trunk链路震荡。
LACP的保活由CPU完成，如果因CPU繁忙不能及时接收和发送LACP报文，Eth-Trunk会将链路关闭，产生链路震荡。
- 通过CPU软转发的报文被丢弃或转发时延增大。
- 交换机内存消耗增加。

1.5 CPU 占用率高属于正常现象的场景

在网络运行中，CPU占用率过高常常会导致业务异常，例如BGP震荡、VRRP频繁切换甚至用户无法登录交换机。但某些情况下，CPU占用率高并不会导致网络问题，例如，交换机在某一时刻集中读取光模块信息、瞬间流量增多等各种具体情况，导致CPU占用率暂时性高的现象是正常的、可接受的，所以不能简单的将CPU占用率高当作故障处理。只有当设备长时间不能正常处理业务时，才需要定位是否由于CPU占用率高而引起的。

如下一些场景可能导致CPU占用率高，属于正常现象，而不是故障场景。如果过了一段时间后，CPU占用率恢复到正常值，则可以不需要处理：

- 网络中瞬间流量增多。
- 交换机单板刚启动。
- 网管系统在频繁操作交换机。
- 批量读取光模块信息（执行命令**display interface transceiver**或者在网管读取光模块信息）。
- 交换机在执行**copy flash:/**或输出信息量大、执行时间长的命令，如**debugging**、**display diagnostic-information**命令。
- 交换机在进行生成树的计算。

对于MSTP，CPU占用率同实例数和活跃端口数成正比。对于VBST，由于每个VLAN独立运行一个实例，因此在相同VLAN和端口数目下，VBST比MSTP占用更多的CPU资源。

- 交换机接收到路由更新信息，大规模更新路由表。

当接收到路由更新消息时，设备需占用CPU资源将路由信息更新到转发面。对于堆叠系统，路由信息还需要同步到其他成员交换机。

在路由表更新过程中影响CPU占用率大小的因素有：

- 路由表项的规模
- 更新的频率
- 接收更新的路由协议进程数
- 堆叠系统成员交换机数量

- 导致CPU占用率高的其他事件：

- 将大量端口同时加入大量VLAN（如通过端口组操作，将大批端口加入大量VLAN、修改大批端口的链路类型等）。
- 频繁或大量的IGMP请求。

- 大量并发的DHCP请求（如交换机作为DHCP服务器时，恢复与大量用户的连接）。
- ARP广播风暴。
- 以太网广播风暴。
- 软转发大量并发协议报文（如短时间内L2PT透传大量BPDU报文，DHCP Relay/Snooping软转发DHCP报文等）。
- 大量不能由转发芯片直接转发的数据报文上送CPU（如ARP-Miss）。
- 端口频繁Up/Down。

1.6 CPU 各任务名称及功能说明

任务名称	任务描述
AAA	认证鉴权计费
AM	地址管理
ARP	地址解析协议
BGP	边界网关协议
CMF	配置管理框架
CSPF	最短路径优先算法
DEVICE	设备管理
DHCP	动态主机配置协议
ETRUNK	跨框TRUNK协议
EUM	以太用户管理
EVPN	以太虚拟网络
FEA	功能实体动作
FEC	功能实体控制
FIBRESM	资源管理
IFM	接口管理
IGMP	因特网组管理协议
IP STACK	协议栈
ISIS	ISIS路由协议
L2VPN	二层虚拟私有网络
LDP	标签分发协议
LLDP	链路层发现协议
LOCAL PKT	主机收发

任务名称	任务描述
MACM	静态MAC管理
MSTP	多实例stp
ND	ICMPv6邻居发现
NETSTREAM	网络流采样
OAM	操作管理维护
OSPF	开放最短路径转发
PEM	节能管理
PIM	协议无关组播
PNP	即插即用服务
RBS	远端备份服务
RGM	冗余网关管理
RM	路由管理
SFLOW	采样流
SLA	服务等级承诺
SMLK	smart link协议
STACKMNG	堆叠管理
SYSTEM	系统管理
TNLM	隧道管理
TRILL	TRILL
TUNNEL	隧道
VLAN	虚拟LAN
VRRP	虚拟路由器冗余协议
VXLAN	虚拟扩展Lan

1.7 相关信息

[CloudEngine 12800, 12800E, 8800, 7800, 6800, 5800系列交换机 CPU占用率高技术专题](#)