

ASR5x00 MME超载保护特点

目录

[简介](#)

[MME保护](#)

[网络超载保护：限制的附上速率](#)

[网络超载保护：限制的传呼](#)

[配置示例](#)

[网络超载保护：限制的DDN \(服务网关功能，保护MME\)](#)

[网络超载保护：EGTP路径失败限制](#)

[配置示例](#)

[增强版拥塞控制](#)

[拥塞状况阈值](#)

[阈值和容限级别](#)

[服务控制CPU阈值](#)

[系统CPU阈值](#)

[系统内存阈值](#)

[相关信息](#)

简介

本文突出显示多种移动性管理实体(MME)超载保护方法和功能可用在5000系列Cisco聚合服务的路由器(ASR)。在5000系列的ASR中，Cisco给客户多种平均值达到控制，并且此条款说明功能和相关CLI命令。

MME保护

网络超载保护：限制的附上速率

限制的附上速率保护邻接网元例如家庭用户服务器(HSS)，策略和正在充电规则功能(PCRF)和联机正在充电服务器(OCS)和内部MME资源例如imsimgr和sessmgr。限制的附上速率处理到达对imsimgr的新的呼叫，这样在随员和相互MME/跟踪区域更新(TAUs)的服务GPRS支持节点(SGSN)。

此图片显示呼叫和限制的队列的消息流。

为了保护MME (向前imsimgr和sessmgr)，应该定义限制的速率、队列等待时间和队列大小时间。因为MME产能依靠呼叫型号，限制的速率依靠MME呼叫型号。

对于限制的速率计算相对简单的MME，请采取标准呼叫事件每秒(CEP)在网络加上容差。并且，如果HSS保护是需要的，您也许需要也考虑HSS数据库产能。

示例

在忙碌小时，MME处理170到200呼叫每秒(Attaches+相互TAU)。在一站点失败的情况下，350到370呼叫每秒也许达到一MME。在此呼叫费率下，MME利用率上升近80%和400呼叫每秒是限制限制的速率的一个最佳的级别为了避免额外的信令负载在MME方框里面。

队列等待时间默认情况下是5秒。它为客户是最佳的。队列大小默认情况下是2500。它为客户是最佳的。

配置命令如下。

```
asr5k(config)#network-overload-protection mme-new-connections-per-second
new_connections action attach { drop | reject-with-emm-cause
{ congestion | network-failure | no-suitable-cell-in-tracking-area}
tau { drop | reject-with-emm-cause { congestion | network-failure
| no-suitable-cells-in-tracking-area | no-sec-ctxt-in-nw} fwd-reloc
{ drop | reject} }{wait-time <wait-time>} {queue-size <queue-size>}
```

new_connections

定义了将接受的新的MME连接数量每秒。必须是从50的一个整数到5000。默认是500。

操作

当定步队列变得全双工时，定义了应采取的措施。每当新连接接收在MME，他们在定步队列排队，并且imsimgr处理从队列的消息以配置速率。当队列溢出(由于流入速率)，根据“操作”配置，数据包丢弃或拒绝。

队列大小

定义了缓冲数据包使用的定步队列的最大大小。必须是从250的一个整数到25000。默认是2500。

配置示例

```
network-overload-protection mme-new-connections-per-second 400 action attach
reject-with-emm-cause no-suitable-cell-in-tracking-area tau
reject-with-emm-cause no-suitable-cell-in-tracking-area fwd-reloc drop
```

现在呼叫费率每秒设置到400，并且操作是与原因#15的智能拒绝制作用户设备(UE)重新连接对不同的无线电访问技术(汇率)。等待时间设置为默认(5秒)，并且队列大小是2500。

注意：操作“拒绝”与EMM原因#15 “NO-适当信元在跟踪区域”更喜欢，当呼叫拒绝与#15主要没有rearrive对MME，并且请去不同的RATS层(3G，2G)。操作“丢弃”服务的无线电网子系统(SRNS)拆迁的是以后使用和防止快速再依附对MME在拒绝以后。

网络超载保护：限制的传呼

限制的传呼保护内部MME资源(mmemgr)作为eNodeB /radio资源(若需要)。此速率限制阈值将是可适用的对与一个给的ASR 5000机箱的MME产生关联的所有eNodeB。S1对eNodeB的传呼请求将是速率被限制在此阈值。S1超出此阈值对eNodeB的传呼请求将丢弃。

对于限制的速率计算相对简单的MME，请采取在网络的标准的出口分页率加上容差。(这根据设计小组的决策纯粹地。)

示例

在忙碌小时，每MME处理35000个传呼消息每秒。在一站点失败的情况下，70000个页每秒也许从一MME去。在此分页率下，MME利用率(mmemgr)上升近80%和70000个到80000个页每秒是限制限制的速率的一个最佳的级别为了避免发信号在mmemgr的额外的S1。

然而，速率每平均的eNodeB被限制。平均速率每eNodeB (在6500 eNodeB的情况下)是10个页每秒。然而，跟踪的区域(TA)不是相等的在用户和多种TA/member eNodeB编号不同地装载与传呼。一旦两倍在TA大小的差异与用户平均数每个TA，速率每eNodeB是20。一旦20倍在TA大小的差异与用户平均数每个TA，速率每eNodeB是200。这意味着功能变得最高效在案件，当TA (总数用户)时均匀地装载。

应该平行采取的另一行动是激活智能传呼。参考“TAI mgmt db和LTE传呼”部分在ASR 5000 MME管理指南。

配置命令如下：

```
asr5000(config)# network-overload-protection mme-tx-msg-rate-control enb s1-paging  
<rate in messages per second>
```

- 网络超载保护识别网络超载保护
- MME tx MSG速率控制enb识别MME消息速率控制每平均的eNodeB
- s1-paging识别S1传呼的消息速率控制
- <rate>指定在消息的速率阈值每秒每eNodeB -范围(1到65535)

配置示例

```
network-overload-protection mme-tx-msg-rate-control enb s1-paging 200
```

注意：

-速率限制是进一步调整的主题，在一个越来越少的方向。为调整的基本类型是用户(传呼编号数量)在TA (TA级别统计信息要求)。

-功能变得最高效在案件，当TA (总数用户/传呼每个TA)时均匀地装载。

网络超载保护：限制的DDN (服务网关功能，保护MME)

下行(DDN)限制的数据通知是控制速率的功能DDN请求对从服务网关(SGW)侧的MME。它保护MME资源例如mmemgr和sessmgr以防止DDN (即入口传呼请求)电涌。

有两部分对此功能，一个Rel-10兼容MME的和其他Rel-10固执的MME的：

- 对于Rel-10兼容MME，请设置DDN限制的分配和挽留优先级(ARP)水印在SGW服务中为了启用功能。
- 对于Rel-10固执的MME，一些其他参数需要与ARP水印一起设置(例如限制的要素，限制时间，稳定时间，投票间隔，等等)在SGW服务中。

当此功能在SGW时启用，发送在DDN Req的—ARP水印对MME。在回复，MME发送限制的延迟元件、限制的延迟值和限制的要素。延迟值和延迟元件的组合计算限制的时间。收到这些值后，SGW下降特定的ARP DDN Req，直到限制的计时器超时。

对于非Rel-10使用本地配置的兼容MME，SGW节流与特定ARP水印的DDN Req。

思科ASR5x00 MME发布16和17不支持限制的自动DDN，因此工作，非Rel 10兼容根据DDN限制。

注意：DDN限制提供进一步粒度在入口侧限制的MME传呼顶部(S11)而不是在输出侧(S1)。思科不要求您实现限制的DDN，如果传呼Throttling配置，但是提供更早的超载检测和清除。

技术Specifications(TS) 23.401，MME的参考：

限制DDN请求

在不寻常的情况下(例如，当MME负载超出操作员配置的阈值)时，MME也许限制其SGWs生成对此的信令负载，若被设定如此执行。

MME能拒绝DDN要求UEs的低优先级的流量在空闲模式或进一步卸载MME。MME能请求SGWs选择性地减少DDN编号请求它为为限制的延迟接收的为在空闲模式的UEs符合一个限制的要素和下行低优先级的流量发送指定在DDN Ack消息。

SGW确定是否持票人是为低优先级的流量根据持票人的ARP优先级和操作员策略(即在作为优先级或非优先的流量将考虑的ARP优先级的SGW的操作员的配置)。MME确定是否DDN请求是为低优先级的流量根据从SGW和操作员策略接收的ARP优先级。

如果发信号减少(ISR)的空闲状态为UE不是活跃的，在限制的延迟期间SGW丢弃在叫作不是用户平面的UEs的所有其低优先级的持票人接收的下行数据包连接(即SGW上下文数据不指示下行用户平面通道末端标识符(TEID))服务由该MME以限制的要素的比例，和传送DDN信息对仅MME非被节流的持票人的。

如果ISR为UE是活跃的在限制的延迟期间，SGW不发送DDN对MME和只发送DDN对SGSN。如果MME和SGSN请求装载减少，SGW丢弃在叫作不是用户平面的UEs的所有其低优先级的持票人接收的下行数据包连接(即SGW上下文数据不指示下行用户平面TEID)以限制的要素的比例。

SGW恢复正常操作在限制的延迟的终止。为时接收限制的要素的值，并且限制的延迟取代从该MME接收的所有上一个值。限制的延迟的接收重新启动SGW计时器关联与该MME。

对于SGW与MME，限制的速率计算相对简单。采取是1100个消息每秒每个MME方框的最大允许入口分页率。

配置命令如下：

```
#configure

#context saegw-gtp

#sgw-service sgw-svc

#ddn throttle arp-watermark <arp_value> rate-limit <limit> time-factor <seconds>
throttle-factor <percent> increment-factor <percent> poll-interval <second>
throttle-time-sec <seconds> throttle-time-min <minutes> throttle-time-hour <hour>
stab-time-sec <seconds> stab-time-min <minutes> stab-time-hour <hour>
```

节流孔ARP水印arp_value

如果ARP水印配置，并且，如果MME/SGSN发送限制的要素和延迟的DDN ACK消息，比配置值有一个ARP值极大的所有DDN由节流孔要素节流指定的延迟的。

arp_value是整数从1至15。

速率限制限制

配置速率限制(请使用此和随后的令牌对速率限制，只有当MME是非版本10 MME)。
限制是整数从1至999999999。

时间因素秒钟

配置期间SGW做出限制的决策的时间持续时间在。
秒钟是从1的一个整数到300。

节流孔要素百分比

配置DDN限制的要素。输入将丢弃的DDN的百分比当查出DDN电涌时。
百分比是整数从1至100。

增量要素百分比

配置DDN限制的增量要素。输入限制的DDN应该增加的百分比。
百分比是整数从1至100。

轮询间隔秒

配置在限制的DDN的轮询间隔。
秒钟是整数从2至999999999。

节流孔时间SEC秒钟

以秒钟配置DDN限制的时间。以DDN被节流在SGW的秒钟进入时间。
秒钟是整数从0至59。

节流孔时间MIN分钟

配置DDN限制的时间以分钟。进入时间以DDN被节流在SGW的分钟。
分钟是整数从0至59。

节流孔时间小时小时

以几小时配置DDN限制的时间。以DDN被节流在SGW的几小时进入时间。
小时是整数从0至310。

刺时间SEC秒钟

以秒钟配置DDN限制的稳定时间。以的秒钟进入一个时间，如果系统被稳定，限制将禁用。
秒钟是整数从0至59。

刺时间MIN分钟

配置DDN限制的稳定时间以分钟。进入一个时间以的分钟，如果系统被稳定，限制将禁用。
分钟是整数从0至59。

刺时间小时小时

以几小时配置DDN限制的稳定时间。以的几小时进入一个时间，如果系统被稳定，限制将禁用。
小时是整数从0至310。

配置示例

```
ddn throttle arp-watermark 1 rate-limit RATE time-factor 120 throttle-factor 50  
increment-factor 10 poll-interval 30 throttle-time-sec 0 throttle-time-min 1  
throttle-time-hour 0 stab-time-sec 0 stab-time-min 2 stab-time-hour 0
```

- 1100页/秒钟是最大允许进入速率(包括DDN)
- 1100页/秒钟在DDN电涌的情况下对应于1100个DDN/seconds
- 有4xSGW的地区每MME站点>速率=每允许的SGW最大数量275个DDN/second
- 有3xSGW的地区每MME站点>速率=每允许的SGW最大数量366个DDN/second
- 有2xSGW的地区每MME站点>速率=每允许的SGW最大数量550个DDN/second

- 有1xSGW的地区每MME站点>速率=每允许的SGW最大数量1100个DDN/second

网络超载保护：EGTP路径失败限制

此功能保护MME资源(sessmgr, mmemgr) 4G资源以防止增强版GPRS隧道协议(EGTP)路径失败电涌在IP骨干网和IP回程以及侧网元故障/重新启动的发射失败的情况下。功能enable (event)每sessmgr限制EGTP路径故障事件检测并且定义了进一步粒度对用户管理，在限制的S1传呼顶部。从属在空闲和已连接用户之间的已分解，限额将定。它非常网络细节并且要求调整在关系以eUTRAN和UE状态。

示例

用户拆分大约对已连接的80:20 IDLE。在最坏的情况下，IDLE用户的EGTP PF的一系列导致也许导致mmemgr超载的一系列的传呼，最缩小的瓶颈。这样EGTP传呼要素(PF)电涌(为IDLE用户)首先导致传呼电涌，并且此电涌点击mmemgr瓶颈，因此您需要首先保护mmemgr以防止此。因而IDLE的EGTP PF也许考虑作为允许是最大数量1100个页/第二的一意外的入口传呼电涌。

- 推荐的限制的限制是1000信息/第二IDLE用户的。
- 已连接sub数量是| 5到7次较少比IDLE。
- 传呼电涌不发生与已连接用户，因此2000年推荐信息/sec为已连接用户安全应用。

注意：EGTP PF限制提供进一步粒度在入口侧限制的MME传呼顶部(S11, Sv)而不是在输出侧(S1)。思科不要求您实现限制的EGTP PF，如果传呼Throttling配置，但是提供更早的超载检测和清除。

此配置运用对EGTP有一种接口类型“接口MME”的服务。

配置命令如下：

```
asr5000(config)# network-overload-protection mme-tx-msg-rate-control egtp-pathfail ecm-idle  
< rate in sessions per second > ecm-connected < rate in sessions per second >
```

- 网络超载保护识别网络超载保护
- MME tx MSG速率控制识别MME消息速率控制
- egtp-pathfail识别EGTP路径失败的消息速率控制
- ecm IDLE识别MME UE会话的速率ECM IDLE模式的
- ecm连接识别MME UE会话的速率ECM连接的模式
- <速率在会话上每second>在会话上指定速率阈值每秒，范围是1到5000

配置示例

```
network-overload-protection mme-tx-msg-rate-control egtp-pathfail ecm-idle  
1000 ecm-connected 2000
```

增强版拥塞控制

使用增强版拥塞控制功能，MME能发信号到连接为了重定向流量到在MME池的其他MME的eNodeBs。这用S1接口超载步骤(TS 36.300和TS 36.413)完成。

当超载控制配置时，并且拥塞门限达到，MME可以配置传送S1AP接口超载启动消息到MME连接的百分比eNodeBs。为了反射MME希望减少的负载数量，此百分比可配置。在超载响应信息元素(IE)发送对eNodeBs，MME能请求eNodeB拒绝或允许会话的特定类型，包括：

- 拒绝非紧急事件会话
- 拒绝新建的会话
- permit紧急会话
- permit高优先级会话和莫比尔终止的服务
- 拒绝迪莱宽容访问

拥塞控制功能允许您设置策略和阈值和指定系统如何起反应，当面对与重载加载条件。拥塞控制监视器可能潜在降低性能的情况的系统，当系统在重载下。一般，这些情况是临时的(例如，高CPU或存储器利用率)和迅速被解决。然而，在特定时间时间间隔内的连续或很大数量的这些条件也许有影响系统的能力服务用户会话。拥塞控制帮助识别这样情况并且调用策略论及情况。

拥塞状况阈值

- 系统CPU使用情况
- 系统服务CPU使用情况(Demux卡德CPU使用情况)
- 系统内存使用情况
- 许可证使用情况
- 最大会话每服务

阈值和容限级别

当您配置阈值和容差关键，主要和较小拥塞程度的时，阈值范围和容差不应该交迭。考虑这些配置示例，阈值范围不交迭：

- 在95%的关键拥塞触发和清除在90%
- 在90%的主要拥塞触发和清除在85%
- 在85%的较小拥塞触发和清除在80%

服务控制CPU阈值

此阈值从系统的Demux CPU计算。阈值根据五分钟平均数CPU使用情况计算。

最高的CPU使用情况值Demux CPU的两个CPU核心考虑。例如，如果CPU核心0有五分钟CPU使用情况40%，并且CPU核心1有五分钟CPU使用情况80%，然后CPU核心1为阈值计算考虑。

系统CPU阈值

此阈值计算使用所有CPU五分钟CPU使用情况平均值(除了待机CPU和SMC CPU)。

最高的CPU使用情况值所有CPU两个CPU核心考虑。

系统内存阈值

此阈值计算与所有CPU五分钟内存使用平均值(除了待机CPU和SMC CPU)。

配置拥塞操作配置文件

拥塞操作配置文件定义了可以被执行的一套操作，在对应的阈值被超过后。

连结拥塞操作配置文件与拥塞控制控制政策

必须关联每项拥塞控制控制政策(重要，专业，较小)与拥塞控制配置文件。

配置超载控制

当超载状态在MME时检测，系统可以配置报告情况到eNodeBs的指定的百分比和采取在流入的会话的配置的动作。

这些超载操作也是可用的(除拒绝新塞申斯之外)：

- permit紧急塞申斯和莫比尔终止服务
- permit海伊优先级塞申斯和莫比尔终止服务
- 拒绝迪莱宽容访问
- 拒绝非紧急塞申斯

配置示例说明

这启用拥塞控制功能：

```
congestion-control
```

```
This monitors the overall CPU Utilization including the sessmgr and demux mgrs
```

```
congestion-control threshold system-cpu-utilization critical 90
```

```
congestion-control threshold system-cpu-utilization major 85
```

```
congestion-control threshold system-cpu-utilization minor 80
```

```
Memory utilization thresholds:
```

```
congestion-control threshold system-memory-utilization critical 85
```

```
congestion-control threshold system-memory-utilization major 75
```

```
congestion-control threshold system-memory-utilization minor 70
```

```
CPU utilization on DEMUX card:
```

```
congestion-control threshold service-control-cpu-utilization critical 85
```

```
congestion-control threshold service-control-cpu-utilization major 75
```

```
congestion-control threshold service-control-cpu-utilization minor 70
```

```
Defining tolerance margins:
```

```
congestion-control threshold tolerance critical 5
```

```
congestion-control threshold tolerance major 5
```

congestion-control threshold tolerance minor 5

定义拥塞操作配置文件(重要, 专业和迈纳)

lte-policy

congestion-action-profile criticalCogestionProfile

reject s1-setup time-to-wait 60

drop handovers

drop combined-attaches

drop service-request

drop addn-brr-requests

drop addn-pdn-connects

exclude-voice-events

exclude-emergency-events

report-overload permit-emergency-sessions-and-mobile-terminated-service enodeb-percentage 50

congestion-action-profile majorCogestionProfile

report-overload permit-emergency-sessions-and-mobile-terminated-service enodeb-percentage 50

congestion-action-profile minorCogestionProfile

report-overload permit-emergency-sessions-and-mobile-terminated-service enodeb-percentage 30

end

运用拥塞策略

configure

congestion-control policy critical mme-service action-profile criticalCogestionProfile

congestion-control policy major mme-service action-profile majorCogestionProfile

congestion-control policy minor mme-service action-profile minorCogestionProfile

end

.

相关信息

- [思科ASR 5000移动性管理实体管理指南](#)
- [技术支持和文档 - Cisco Systems](#)