

非商业网络设备的新浪应用之路

新浪-研发中心 李晓栋
<http://weibo.com/3rdday>

一切由你开始

You're The One

新浪·研发中心



我所在的部门



新浪-研发中心

负责新浪网基础架构、平台及底层技术研发

负责新浪网数据中心建设&服务器运维



一切由你开始 You're The One

新浪·研发中心¹

我自己

- 研发中心基础架构部 高级技术经理，7+years in Sina
- 负责:全站自动化运维、网络设备及Linux内核 研发
- 集团讲师(>1200人次)，别名：“李老湿”
- 集团高级工程师评审委员会委员
- 2008年度集团创新奖、优秀团队奖

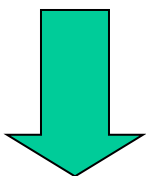


一切由你开始 You're The One

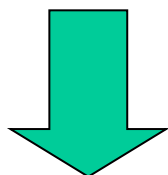
新浪·研发中心²

要讲什么？

新浪非商业网络设备的应用情况



新浪非商业负载均衡的趋势



万兆设备研发情况

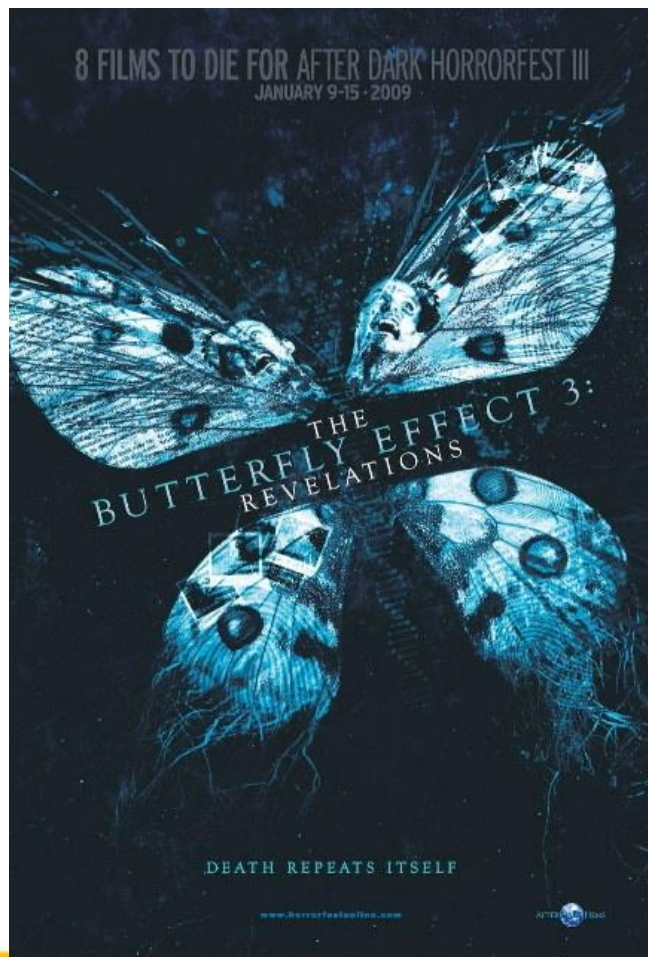


一切由你开始 You're The One

新浪·研发中心³

一、非商业设备的应用情况

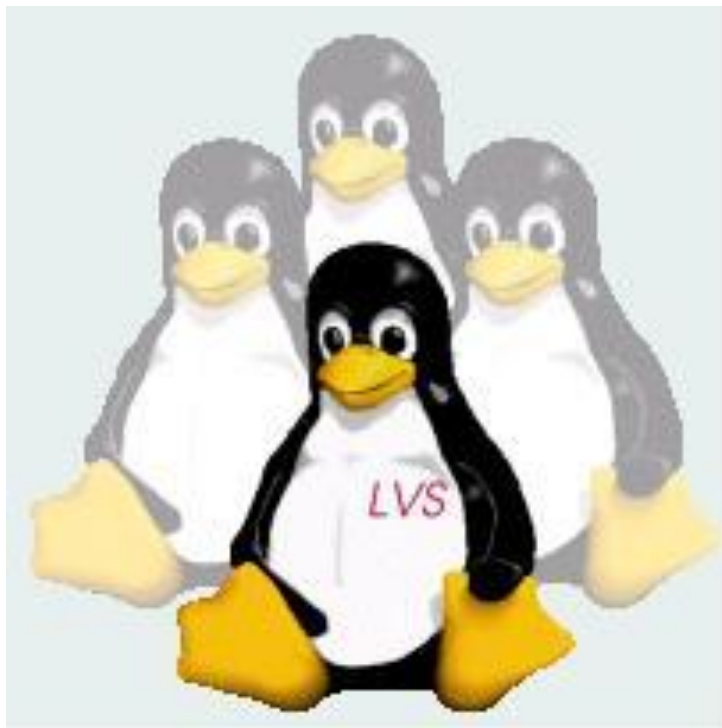
● 由“蝴蝶效应”谈起



一切由你开始 You're The One

一、非商业设备的应用情况

- 开源界的“两只蝴蝶”



网络设备也可以DIY!



一切由你开始 You're The One

一、非商业设备的应用情况

- 网络设备DIY?(2004)

--太难了吧？有必要吗？

- 网络设备DIY?(2007)

--靠谱吗？稳定吗？维护成本高吗？

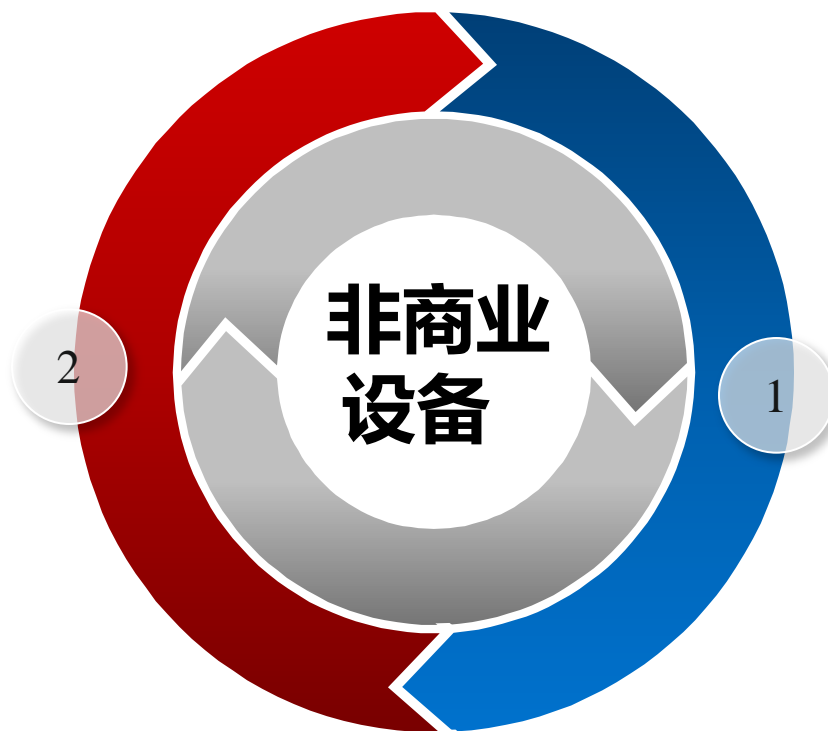
- 网络设备DIY?(2010)

--除了二、三层交换机外，
还有神马不能DIY的？



一、非商业设备的应用情况

自主研发



LVS

Haproxy

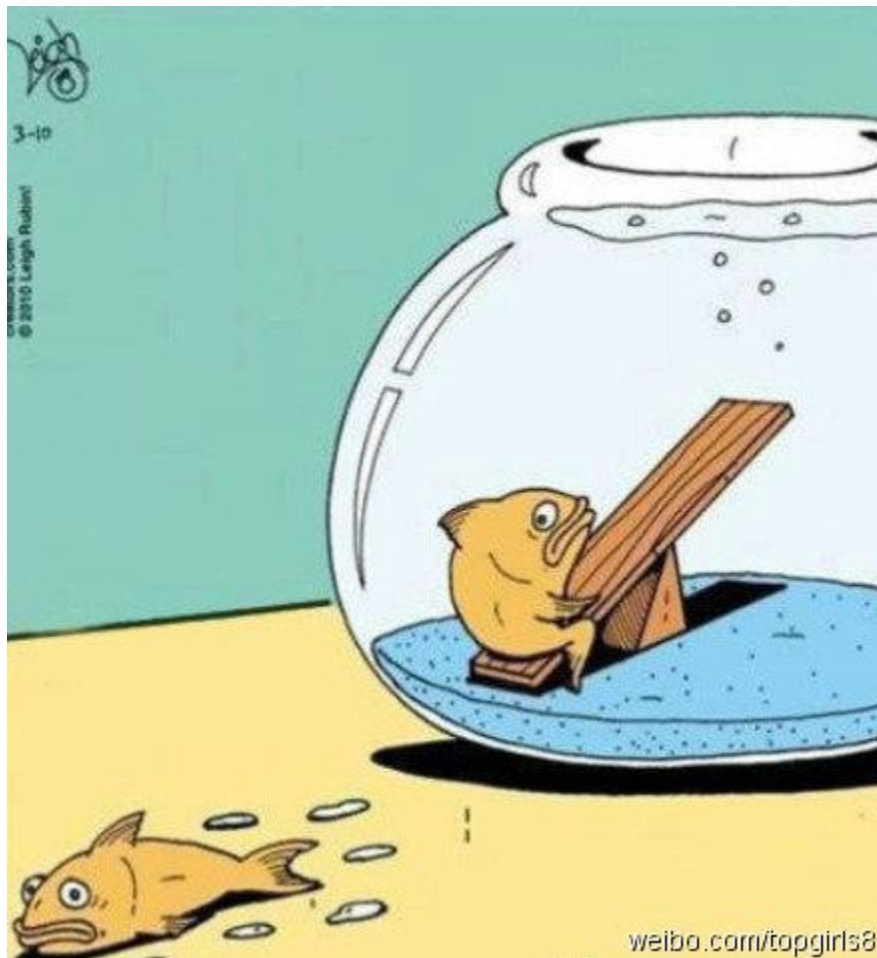
使用开源软件



一切由你开始 You're The One

一、非商业设备的应用情况

● 商业 & 非商业 设备



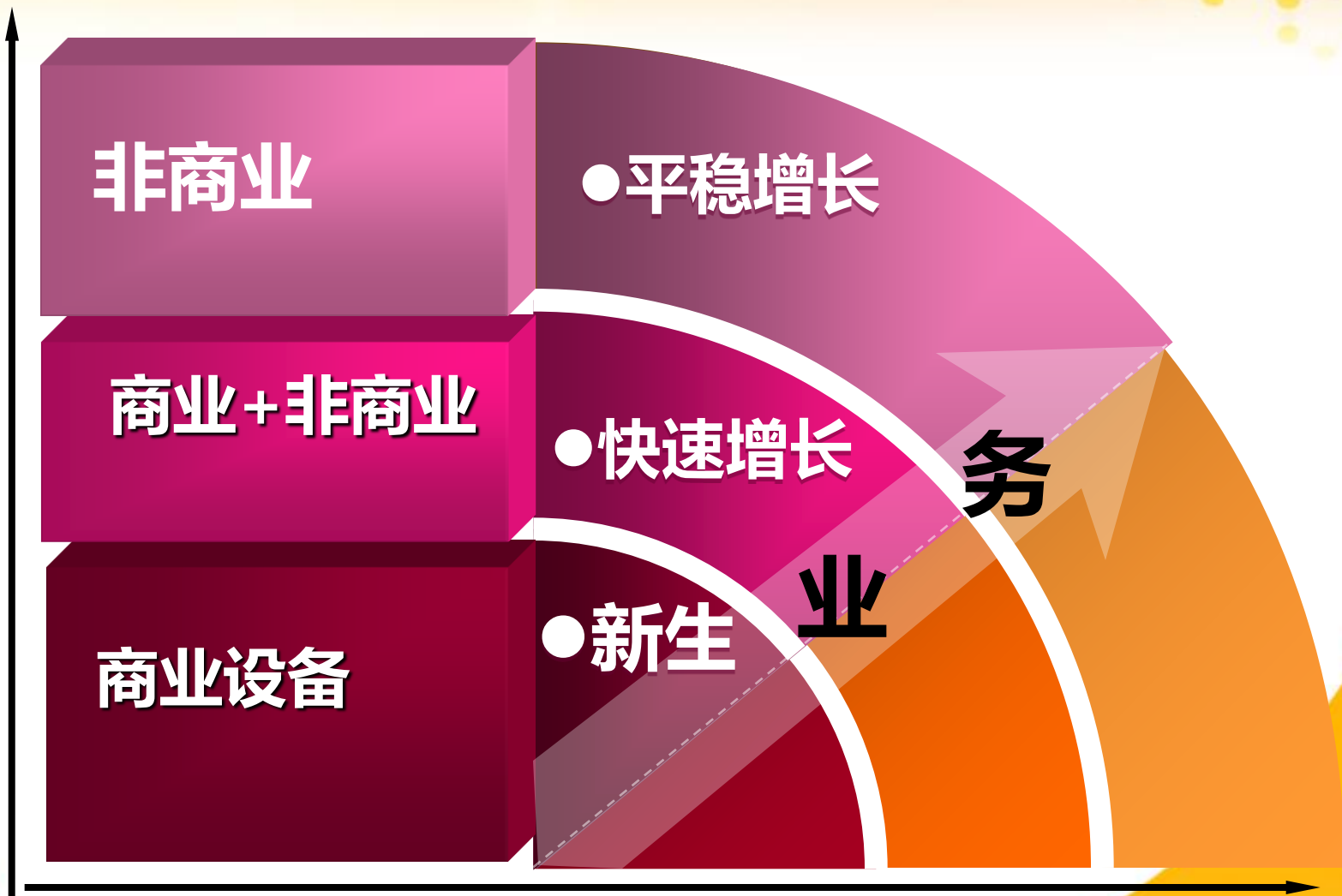
并非完全不用商业设备



一切由你开始 You're The One

新浪·研发中心⁸

一、非商业设备的应用情况



一切由你开始 You're The One

新浪·研发中心⁹

一、非商业设备的应用情况

- 如何用开源软件构建非商业设备？



./configure
make
make install

仅此简单“拼装”吗？

--No!No!No!



一切由你开始 You're The One

一、非商业设备的应用情况

● 如何用开源软件构建非商业设备？

01

性能优化

- 硬件选型
- 软件优化

02

稳定性

- 经验
- 教训

03

维护成本

- 自动化系统

持续化过程



一切由你开始 You're The One

新浪·研发中心¹¹

一、非商业设备的应用情况

- 如何用开源软件构建非商业设备

仅仅做到以上几个方面就足够了吗？

--或许是 (2008年)

--还不够 (2011年)

下面以负载均衡为例说明。。。。。



一切由你开始 You're The One

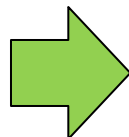
新浪·研发中心¹²

二、非商业负载均衡的趋势

- 2004至今，负载均衡格局发生巨大变化

负载均衡承担角色：

四层交换为主



七层交换为主

七层规则越来越复杂



一切由你开始 You're The One

二、非商业负载均衡的趋势

- 2004至今，负载均衡格局发生巨大变化

承载的业务流：

新闻/论坛 → 博客/播客 → 微博

流量越来越大

对服务质量要求越来越高

对故障诊断和处理效率要求越来越高



二、非商业负载均衡的趋势



非商业负载均衡走向何方？



一切由你开始 You're The One

二、非商业负载均衡的趋势

我们需要更加复杂的架构吗？



一切由你开始 You're The One

新浪·研发中心¹⁶

二、非商业负载均衡的趋势



< ¥20万



> ¥100万

差距在哪里？



一切由你开始 You're The One

新浪·研发中心¹⁷

二、非商业负载均衡的趋势



满足
高端
乘客
需求



一切由你开始 You're The One

18
新浪·研发中心

二、非商业负载均衡的趋势

- 只满足L4/7 规则需求就OK吗？
- 仅提高单机性能就满足了吗？
- 我们还可以为应用服务做些神马？
 - 主动报告应用服务器的潜在异常
 - 提供用户访问质量分析报告



二、非商业负载均衡的趋势



帮你看到你看不到的...

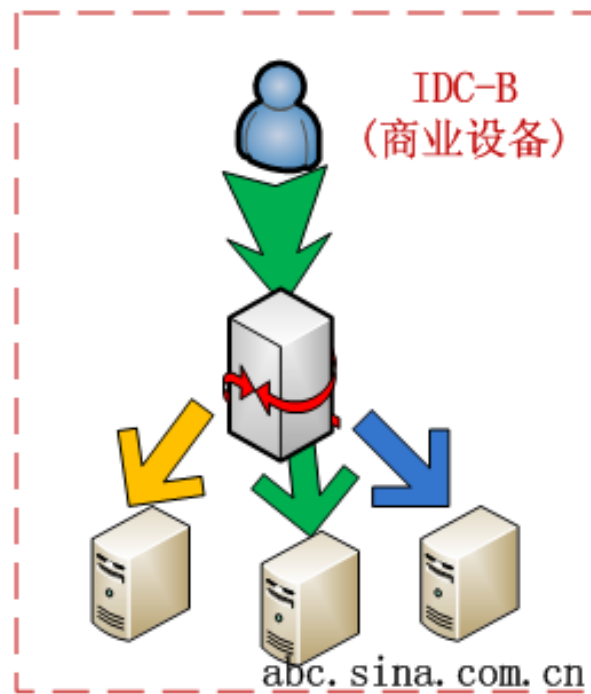
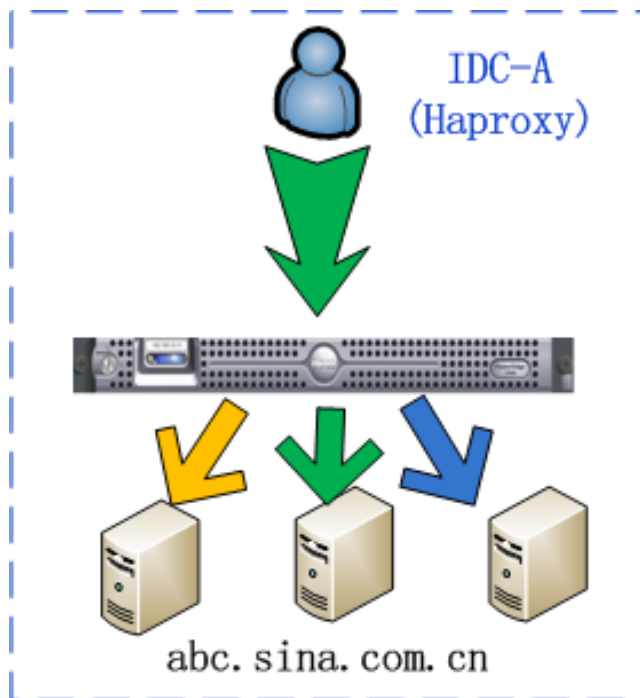


一切由你开始 You're The One

新浪·研发中心²⁰

二、非商业负载均衡的趋势

- 负载均衡系统中的微小隐患都能看到吗？



对于相同服务

使用不同设备

如何保证L7规则一致性

```
if { [HTTP::host] == "abc.sina.com.cn" and [HTTP::uri] starts_with "/zw/" }
```

```
acl zw_url url_beg -i /zw/
```

```
acl t_host hdr(Host) -i abc.sina.com.cn
```

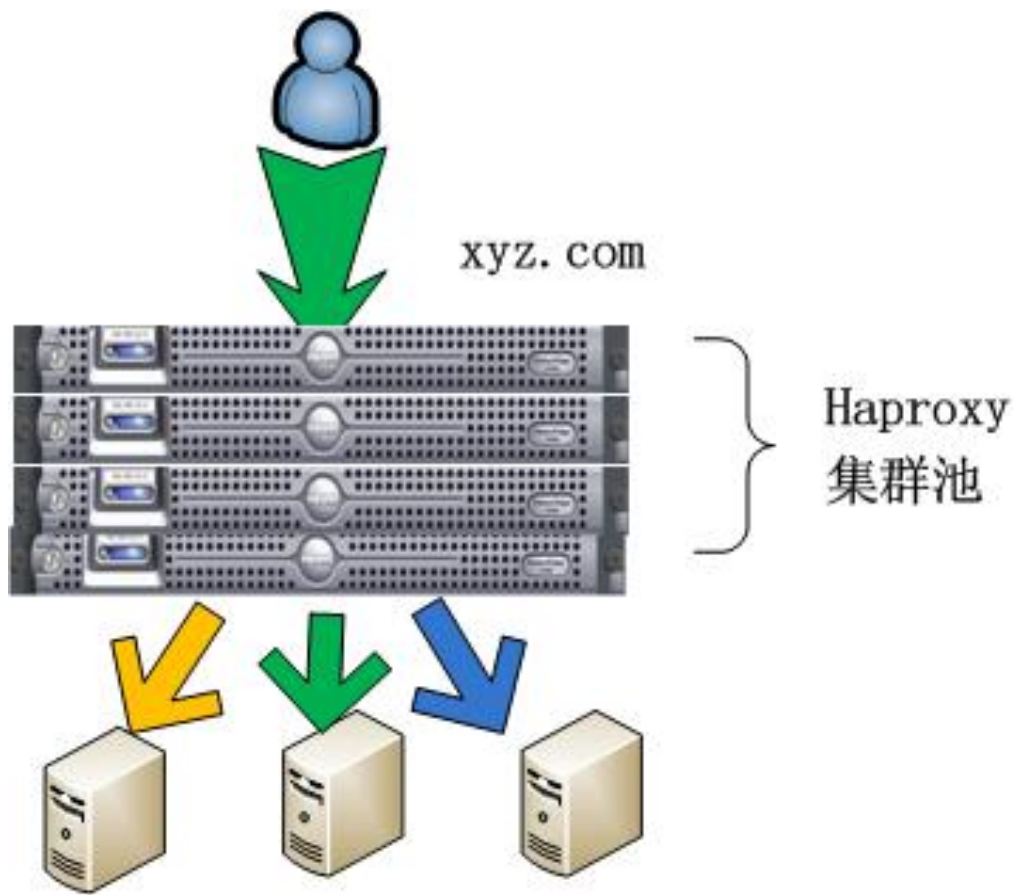


一切由你开始

You're The One

二、非商业负载均衡的趋势

- 负载均衡系统中的微小隐患都能看到吗？



二、非商业负载均衡的趋势



纠正你疏忽的错误...

一切由你开始 You're The One

新浪·研发中心²³



二、非商业负载均衡的趋势

- 即便有自动化运维系统，有严格的运维规范真的可以不再出人为故障吗？

```
acl pubic_stat_url url_beg -i /pub/stat
```

```
acl active path_reg ^/(\w+)/active[^\&\?]?$
```

复杂的七层规则面前，一不留神就“杯具”鸟



二、非商业负载均衡的趋势

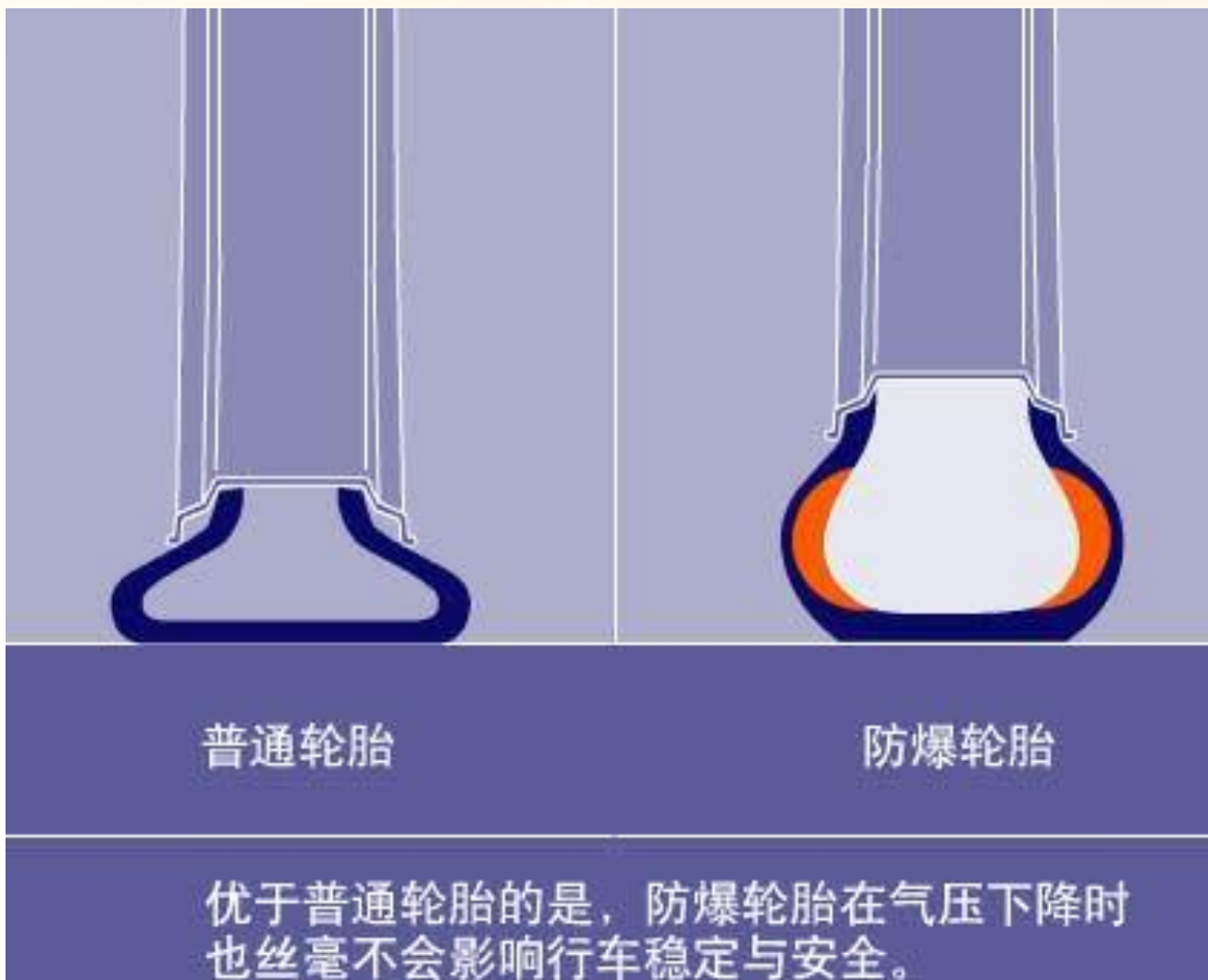
- 即便有自动化运维系统，有严格的运维规范真的可以不再出人为故障吗？

任务：将realserver-A 加到VIP -A 下

结果：将realserver-A 加到了VIP -B 下



二、非商业负载均衡的趋势



抵御各种不预见的钉子



一切由你开始 You're The One

新浪·研发中心²⁶

二、非商业负载均衡的趋势

- 如何让“负载均衡”抵御各种外界意外状况？

Try Best



一切由你开始 You're The One

二、非商业负载均衡的趋势



轻松检查各部件运转是否正常



一切由你开始 You're The One

新浪·研发中心²⁸

二、非商业负载均衡的趋势

- 如何加快“负载均衡”环节的故障诊断效率？



一切由你开始 You're The One

二、非商业负载均衡的趋势



更强劲的动力

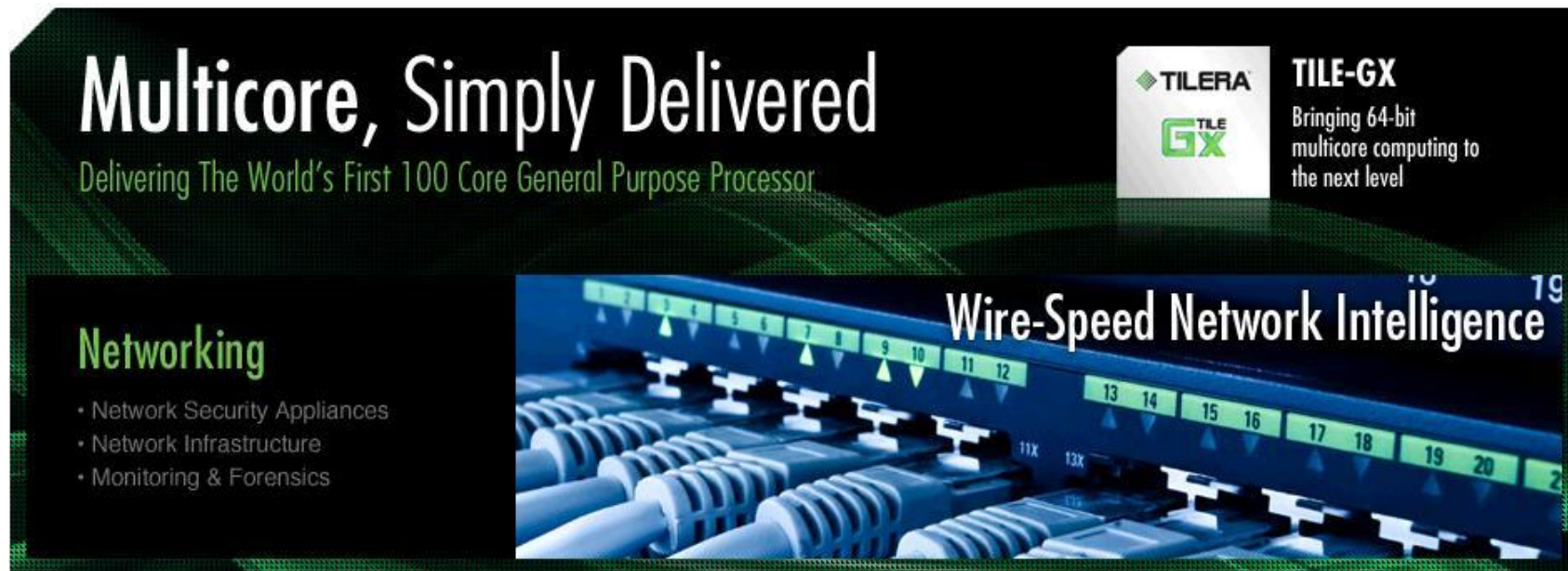


一切由你开始 You're The One

新浪·研发中心³⁰

三、万兆设备研发情况

● 单台万兆线速的四层处理能力



Multicore, Simply Delivered
Delivering The World's First 100 Core General Purpose Processor

TILERAX
TILE
GX

TILE-GX
Bringing 64-bit multicore computing to the next level

Networking

- Network Security Appliances
- Network Infrastructure
- Monitoring & Forensics

Wire-Speed Network Intelligence

The graphic features a background image of network switch ports with blue cables plugged in. The text is overlaid on a dark, textured background with green and blue accents.

基于TILERA平台



一切由你开始 You're The One

新浪·研发中心³¹

三、万兆设备研发情况

Bypass Linux Kernel

User Space “LVS” & relative tools(ipvsadm,keepalived...)

User Space “ARP & Routing”

Performance tuning



一切由你开始 You're The One

总结

● 杭州名菜-叫花鸡



草根也能登大雅之堂



一切由你开始 You're The One

新浪·研发中心³³

谢谢大家!

QA

微博 <http://weibo.com/3rdday>
Email: xiaodong2@staff.sina.com.cn

一切由你开始

You're The One

新浪·研发中心

