RSA CONFERENCE
CHINA 2012
RSA信息安全大会2012

THE GREAT CIPHER

MIGHTIER THAN THE SWORD

伟大的密码胜于利剑

最大化多核技术在网络智能与安全中的应用

Fu Lizheng
Wind River China

RSA CONFERENCE CHINA 2012
RSA信息安全大会2012

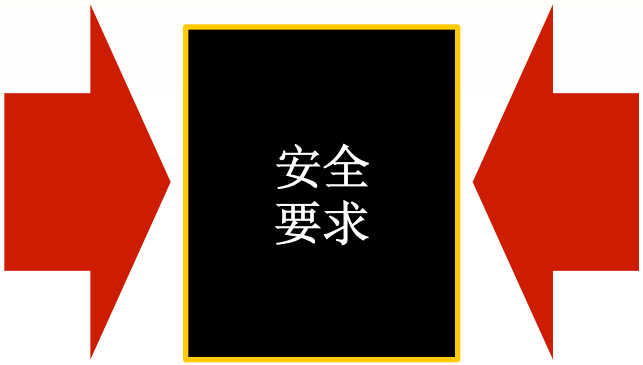
Session ID:

Session Classification:

日益严峻的网络安全挑战

身负重任的安全功能

- 恶意软件
- 恶意入侵/攻击
- 病毒

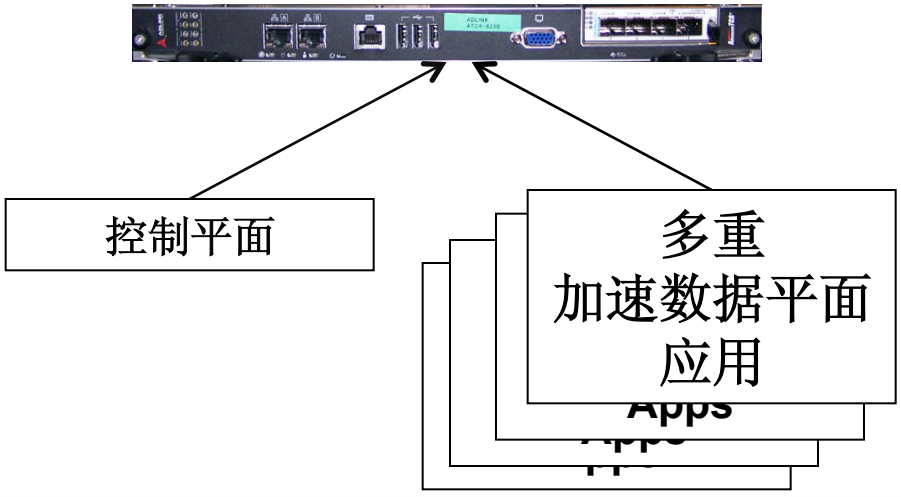


- 更多检测
- **CPU** 负荷超载
- 性能瓶颈
- 日益增加的成本

更智能化的解决方案

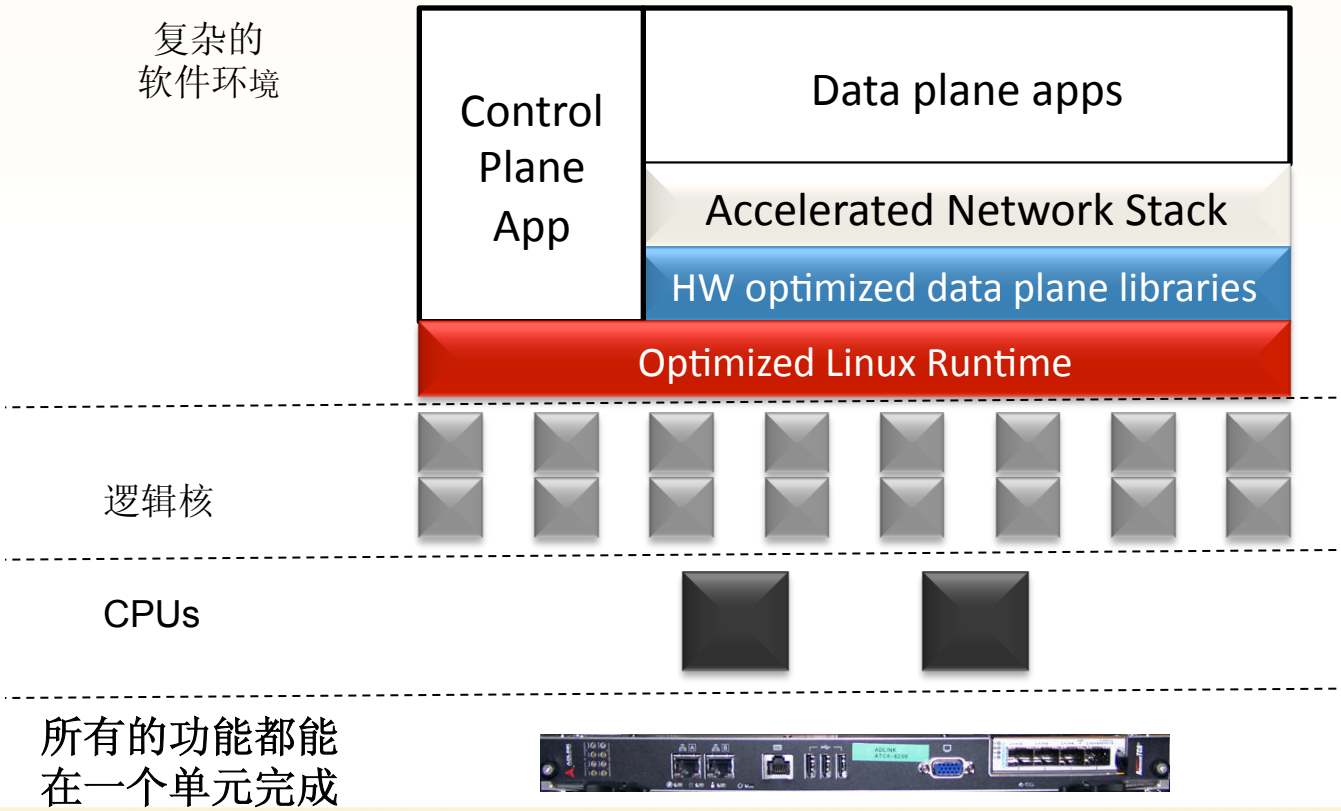
传统的方案：
在一个网络内需要多种网络安全设备

下一代的方案：
借助先进多核技术将多重安全应用整合到同一硬件设备中



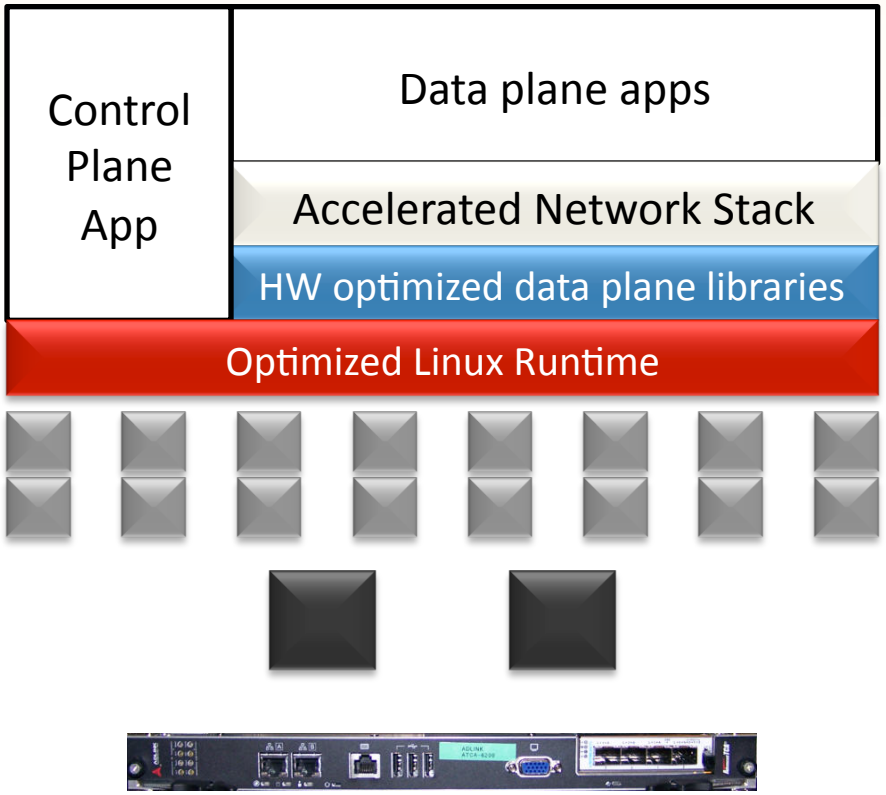
基于先进多核技术的下一代网络设备

实现网络的智能化，先进的多核技术需要提供对网络数据进行加速– 分析 – 保密的功能

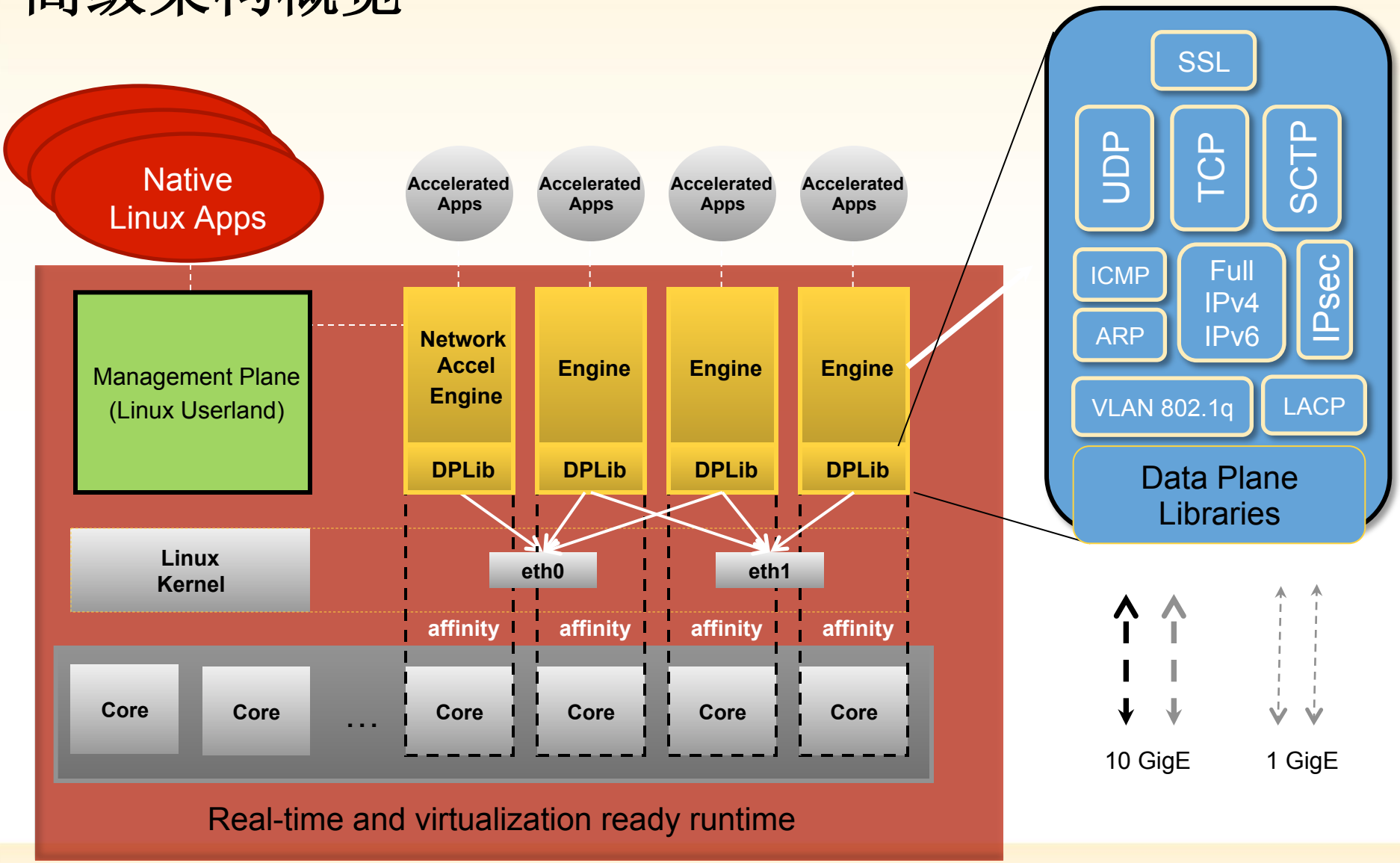


多核系统的优势

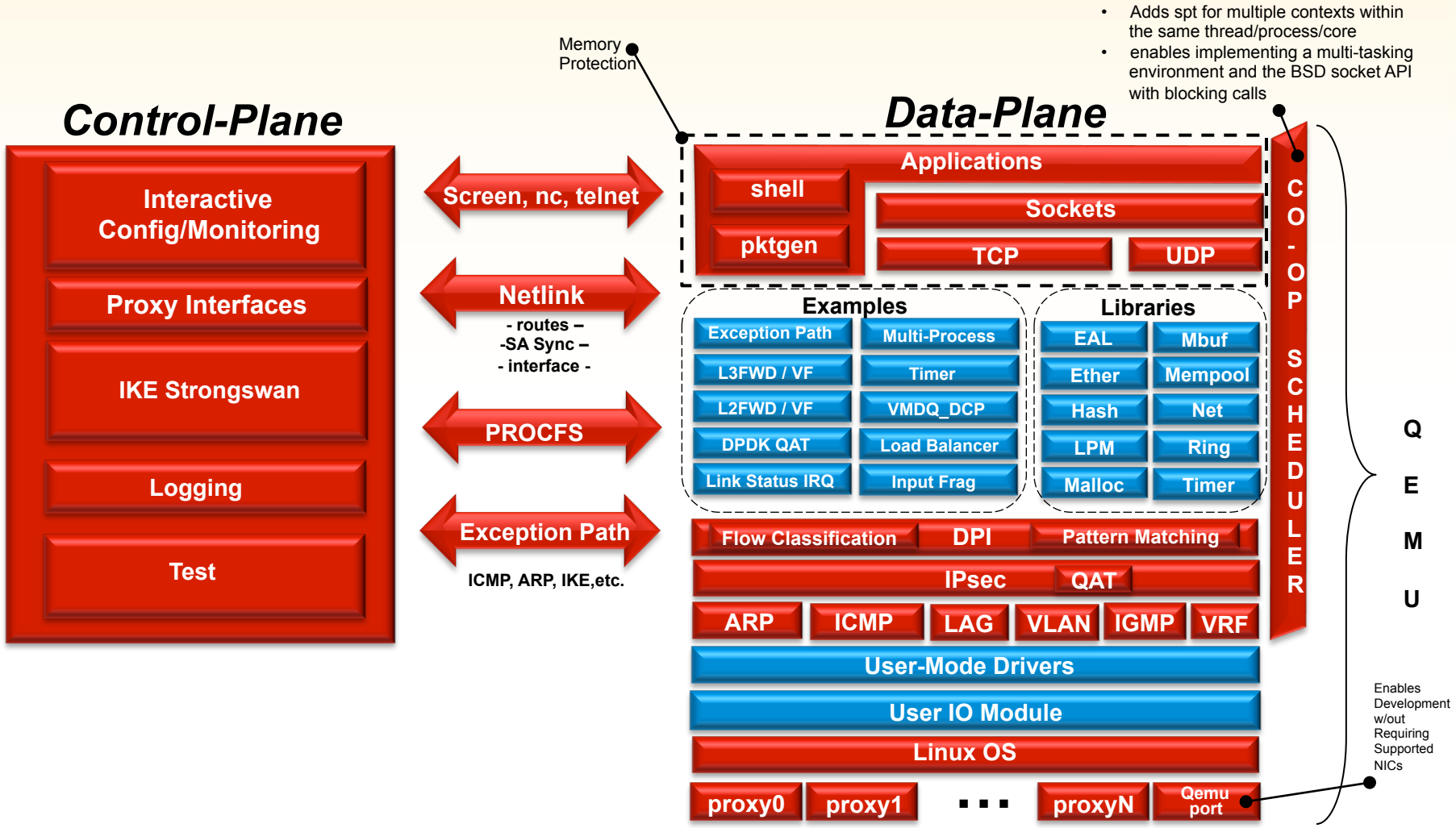
- 更高的性能
- 控制面和数据面共存在同一环境
  -更易管理及调试
- 以线速实现多重深度包检测应用
  -模式匹配
  -流整形
- 可以提供多种功能
  -IPS/IDS, FW, AV, UTM
- 可以形成同一架构下的系列化产品



高级架构概览


Native Linux Apps
Accelerated Apps
Accelerated Apps
Accelerated Apps
Accelerated Apps
Management Plane (Linux Userland)
Network Accel Engine
Engine
Engine
Engine
DPLib
DPLib
DPLib
DPLib
Linux Kernel
eth0
eth1
affinity
affinity
affinity
affinity
Core
Core
...
Core
Core
Core
Core
Real-time and virtualization ready runtime
SSL
UDP
TCP
SCTP
ICMP
ARP
Full IPv4 IPv6
IPsec
VLAN 802.1q
LACP
Data Plane Libraries
10 GigE
1 GigE

多核加速软件系统剖析


Memory Protection
Adds spt for multiple contexts within the same thread/process/core
enables implementing a multi-tasking environment and the BSD socket API with blocking calls
Control-Plane
Interactive Config/Monitoring
Proxy Interfaces
IKE Strongswan
Logging
Test
Screen, nc, telnet
Netlink
- routes –
-SA Sync –
- interface -
PROCFS
Exception Path
ICMP, ARP, IKE,etc.
Data-Plane
Applications
shell
pktgen
Sockets
TCP
UDP
Examples
Exception Path
Multi-Process
L3FWD / VF
Timer
L2FWD / VF
VMDQ_DCP
DPDK QAT
Load Balancer
Link Status IRQ
Input Frag
Libraries
EAL
Mbuf
Ether
Mempool
Hash
Net
LPM
Ring
Malloc
Timer
CO-OP SCHEDULER
Flow Classification
DPI
Pattern Matching
IPsec
QAT
ARP
ICMP
LAG
VLAN
IGMP
VRF
User-Mode Drivers
User IO Module
Linux OS
proxy0
proxy1
proxyN
Qemu port
QEMU
Enables Development w/out Requiring Supported NICs

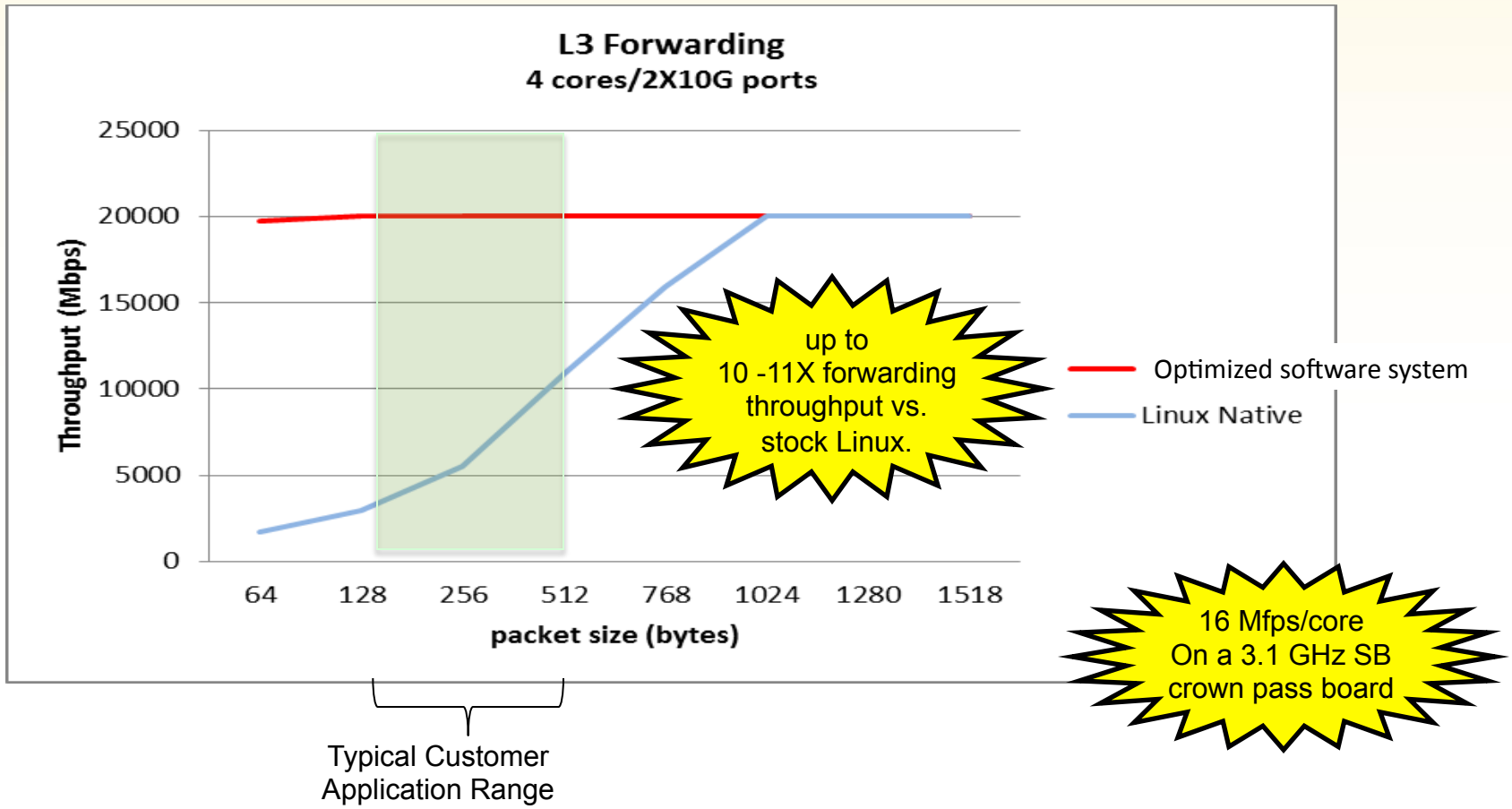
网络协议的加速

- IP 报文的校验和验证
- 支持IPv4 及 IPv6
- 代理接口
- 例外路径
  - 在加速面上可支持：ARP, ICMP, SSH 等
- 4层协议加速：UDP, TCP, SCTP
- IPsec/IKE

采用软件加速解决方案的优势

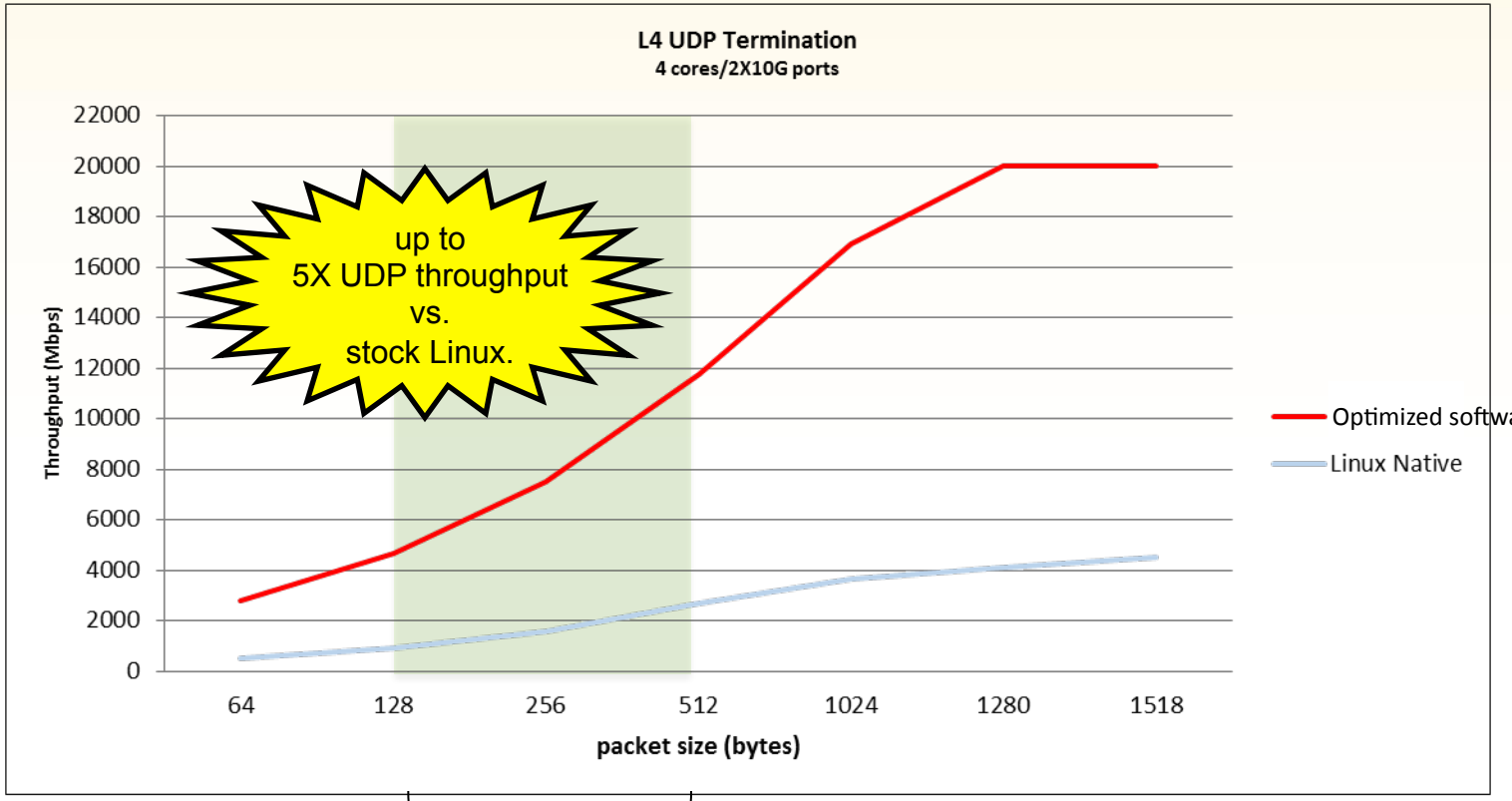
- 在多核CPU中预留部分核用于实现加速平面数据包处理
  - 被加速面各核使用优化TCP/IP 协议栈
  - 数据面的处理能力与处理器核数线性匹配
- 超出数据面库之外的能力
  - 集成了被优化的数据面库
  - 支持现有的Linux应用
  - 采用通用Linux架构及工具
- 更高效的系统
  - 达到更高的吞吐率(~16Mpps/core)
  - 使用更少资源完成加速，以便使更多资源用于控制面的应用

与Linux协议栈的性能对比(3层转发能力)



This benchmark is comparing L3 forwarding throughput performance of Linux native vs. optimized software system. The setup is using 2X10G Ethernet ports and 4 hyperthreads on 4 separate cores on a Sandy Bridge based platform.

与Linux协议栈的性能比对(UDP)



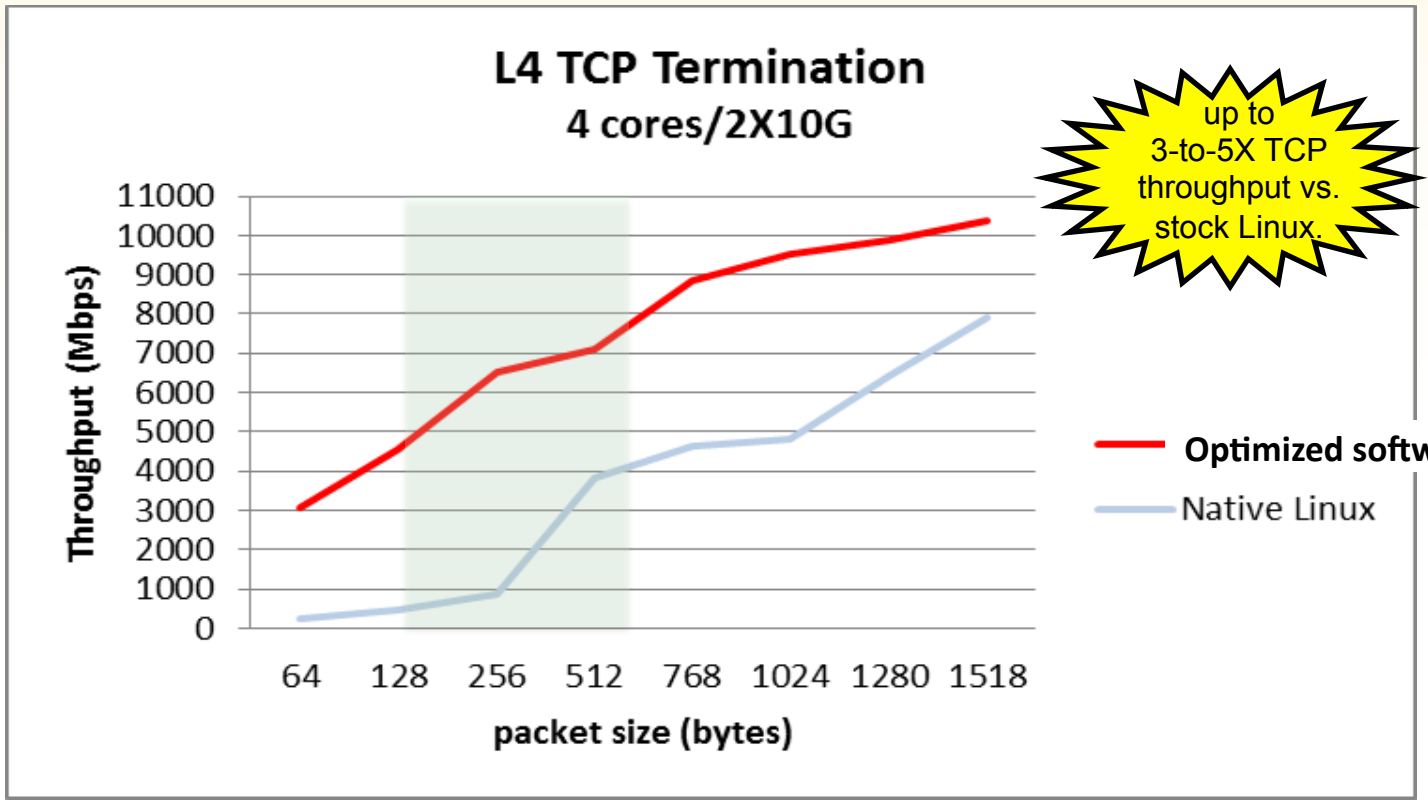
Typical Customer Application Range

This benchmark is comparing L3 forwarding throughput performance of Linux native vs. optimized software system. The setup is using 2X10G Ethernet ports and 4 hyperthreads on 4 separate cores on a Sandy Bridge based platform.

Why this approach is faster?

1. Software system requires fewer cycles than native Linux
2. Faster context switching
3. No interrupts)
4. Much fewer misses in L1 cache than Linux

与本地Linux协议栈的性能比对(TCP)



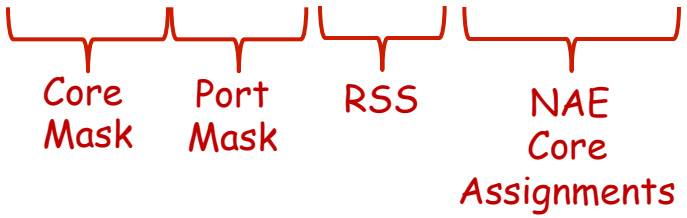
This benchmark is comparing L3 forwarding throughput performance of Linux native vs. optimized software system. The setup is using 2X10G Ethernet ports and 4 hyperthreads on 4 separate cores on a Sandy Bridge based platform.

网络加速引擎（NAE）的工作机制

- 每个NAE作为单独的Linux进程或线程运行
- 采用轮询方式
  - 全互联 (缺省方式) - 每个NAE轮询所有端口
  - 部分互联- NAE轮询的端口数限定到-x个端口
  - 手动互联 - 根据-a选项手动分配NAE
- 突发模式下，轮询多个接收队列
- 针对其轮询的端口建立发送队列

例子：手动模式下的初始化配置

- 网络加速引擎初始化 ：
  - Edit /etc/unap/unap.conf – set UNAP_INSTANCES=0
  - 1 engine per port example:
    - unap-nae –c 0x1f -- -p 0x3 –f 0x5 –a 2/3
  - 2 engines per port example:
    - engine –c 0x1f -p 0x3 –f 0x5 –a 1,2/3,4



- a 2/3 means "assign the core 2 to NAE 1 and core 3 to NAE 2".
- a 1,2/3,4 means "assign core 1 and 2 to NAE1, core 3 and 4 to NAE 2".

- 分配IP给代理接口:
  - ifconfig proxy0 –inet 10.1.1.1 up
  - ifconfig proxy1 –inet 10.2.1.1 up

网络加速实现深度包检测

高性能

- 吞吐率提升至 100Gbps

高伸缩性

- 支持同一产品线中各类高中低档处理器
- 将高速DPI整合到同一产品线的全线产品中仅需一个开发周期

低延时及低开销

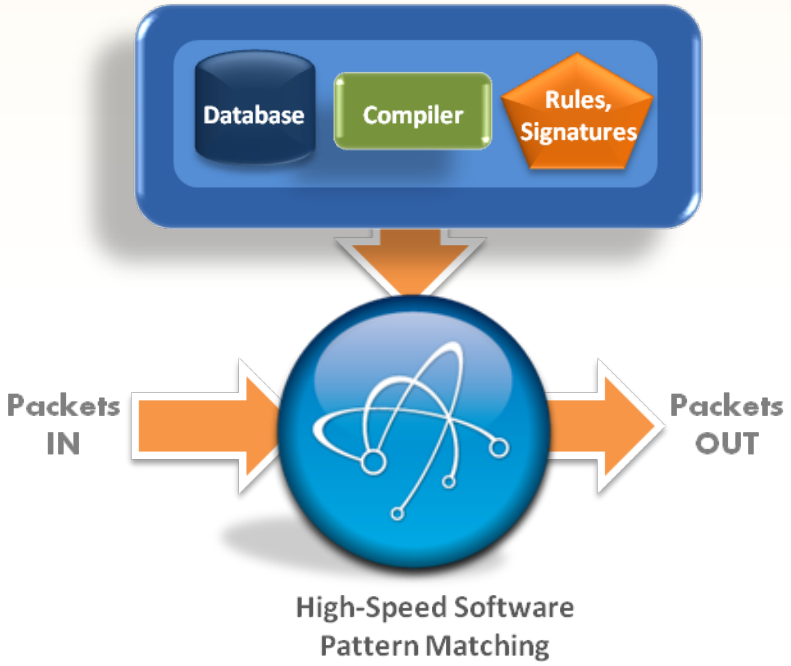
- 数据直接在CPU上进行处理: 低延时，低开销，低编译时间

易于管理

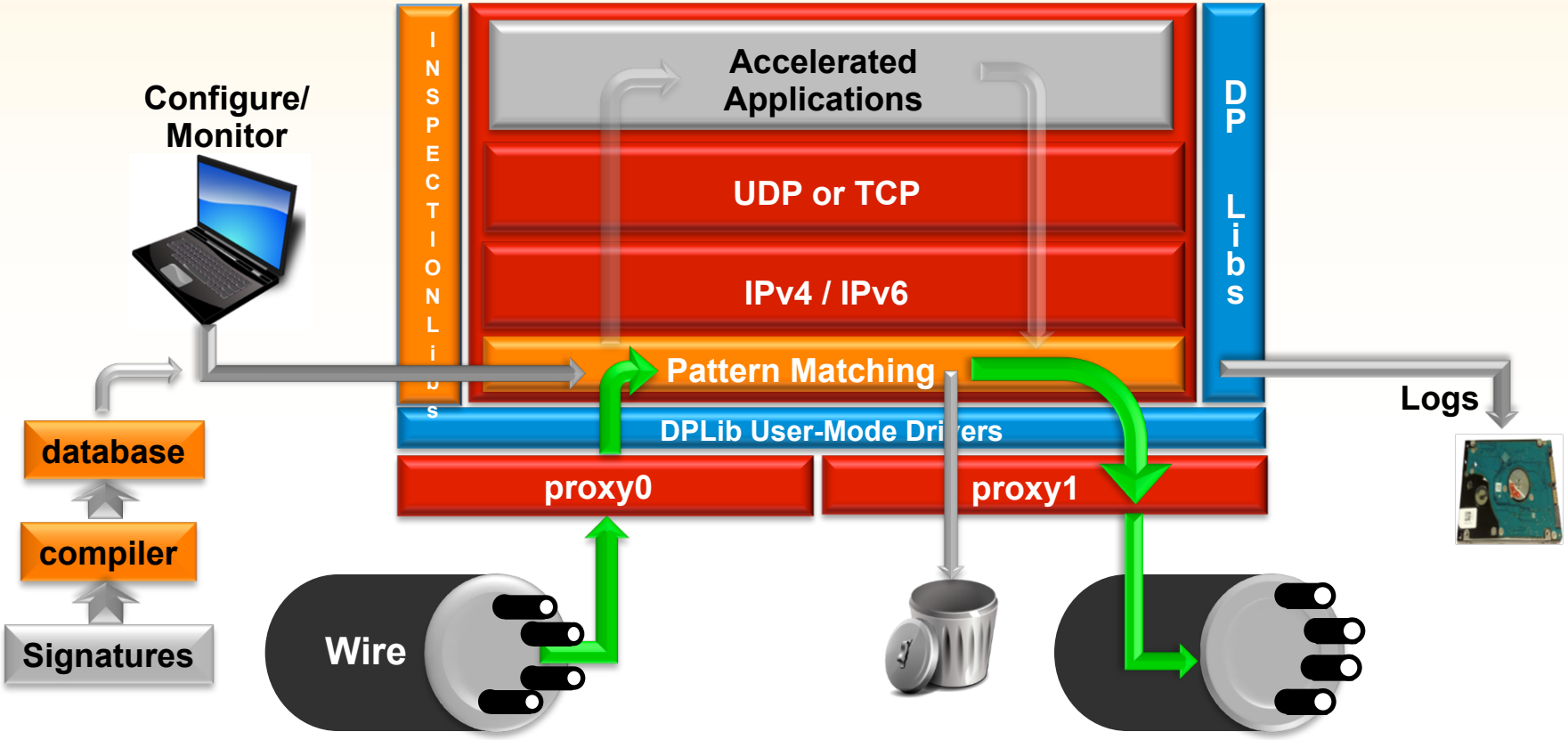
- 便易的集成处理
- 通过软件升级实现新特性、新功能，缩短产品上市时间

被加速的模式匹配

- 软件DPI /模式匹配库
  - 高性能, 比同类HW解决方案更为综合全面
  - 正则表达式库
  - 可移植
  - 高伸缩性(低端产品—高端产品)
- 大规模并行匹配
  - 可支持数十万并行匹配同时进行
- 多路千兆模式匹配
  - L7 DPI; 支持大多数CPU架构
- 低延时及开销
  - 尤为与硬件匹配比对
- 广泛的适用性
  - 广泛支持多种架构/平台 (应用终端，路由器，交换机，服务器)
  - 广泛支持多种应用(IPS/IDS, FW, AV, UTM)



被加速的模式匹配



- Pattern Matching (PM) is a **CPU intensive** operation
- High-speed Pattern matching is achieved using the **compiler** and **optimized libraries**

总结摘要

1. 先进多核技术引领各类网络设备革新浪潮
2. 实现网络加速是迈向高智能化网络时代的第一步
3. 深度包检测软件为您进一步提高分析保密网络数据功能带来无限可能

Thank You

RSA CONFERENCE
CHINA 2012
RSA信息安全大会2012