



POST CARD  
CALLE...  
C...  
FOTOGRAFIA  
ANT... TORONI  
N. 36  
ORINO

17.3.1926

Fotografia postale

Umberto Bonivento  
PIOLA  
Via Promontore 4

Stimata signora  
Contessa Obento  
voglio dirle che...

Atta mia amata  
Maurina pare che  
si ricorda qualche  
volta di me. Ah  
che affetto  
l'avevo. Piola

X'CON 2003

嵌入式网络协议引擎(ETA)加速网络应用程序处理

作者：林明峰

日期：2003-12

# 概要

- 今天的网络环境和网络操作系统
- Linux 的tcp/ip stack 的性能
- 嵌入内核的tcp 加速模块
- ETA的现实应用

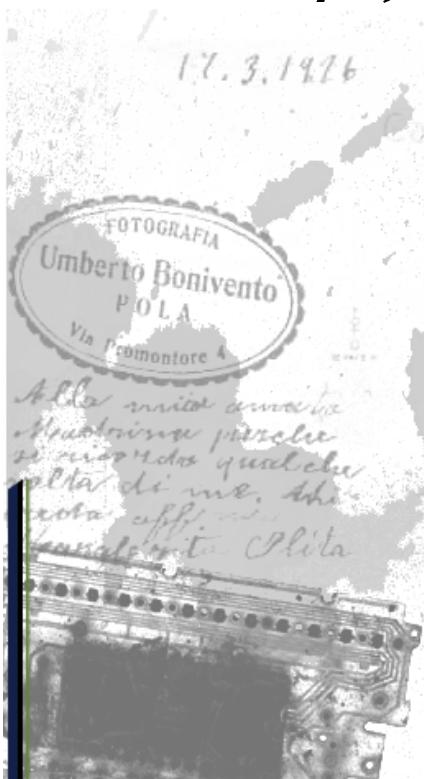


17 3 1976  
Alla mia amata  
Maurina perché  
si ricorda qualche  
volta di me. Ah  
nota aff.  
Umberto Piola

Stimabilissima  
Contessa Monti  
voglio augurarvi



# 今天的网络环境和操作系统



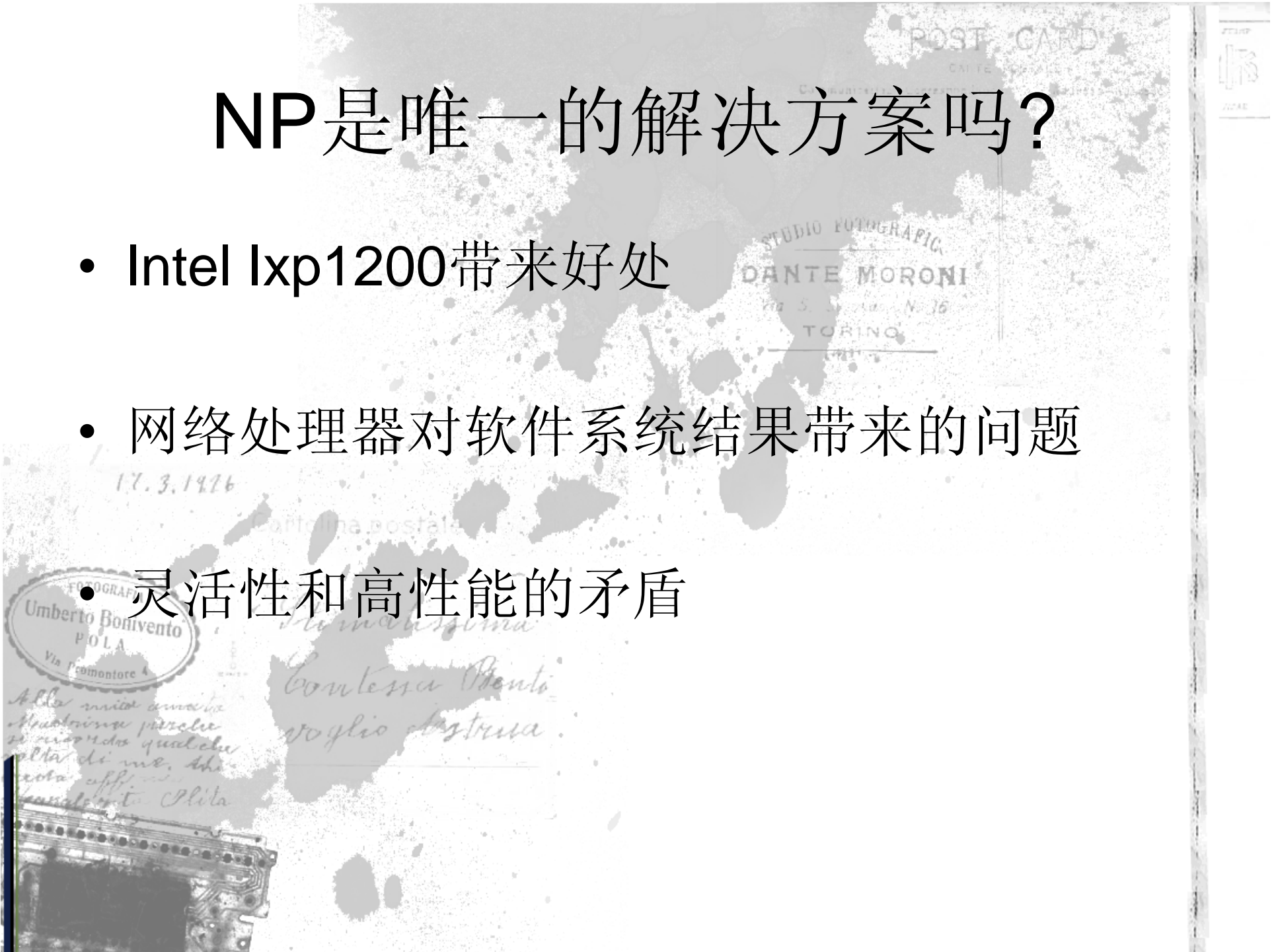
# 10G的以太网带来什么

- 通用处理器必须每67奈秒处理一个数据包
- 超过通用处理器处理极限
- 协议占用太多时间



# NP是唯一的解决方案吗？

- Intel Ixp1200带来好处
- 网络处理器对软件系统结果带来的问题
- 灵活性和高性能的矛盾

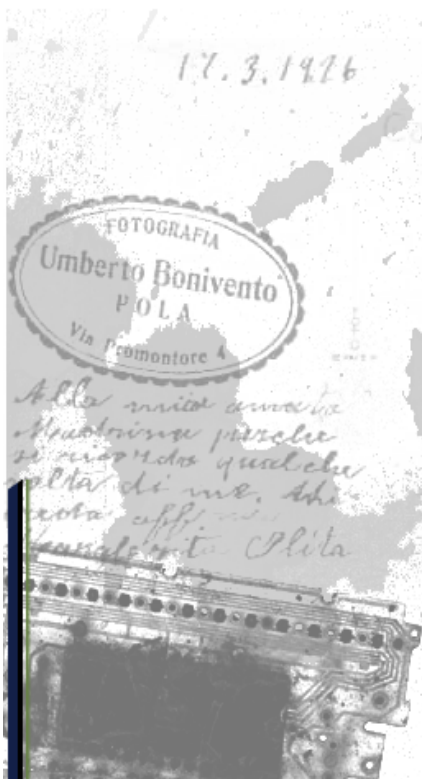


# 谁占用的cpu的时间.

系统中断  
内存拷贝  
缓冲区管理  
系统呼叫  
Tcp/ip协议  
应用程序只能占用3%



# Linux 的tcp/ip协议



# 内核的resource manager (ckrm)

- 控制 diskIO 的资源分配
- 控制网络的IO分配
- 目前还不成熟

• <http://ckrm.sourceforge.net/>



# 我们的解决方案

- 区分出独立的协议处理层
- 固定一个cpu来专门处理协议
- 制定一个异步的通讯接口
- 新的socket库来利用ETA



17.3.1936  
Alla mia amata  
Maurina perché  
si ricorda qualche  
volta di me. Ah  
nota aff.  
Umberto Piola

Stimabilissima  
Contessa Orenti  
voglio dirle.

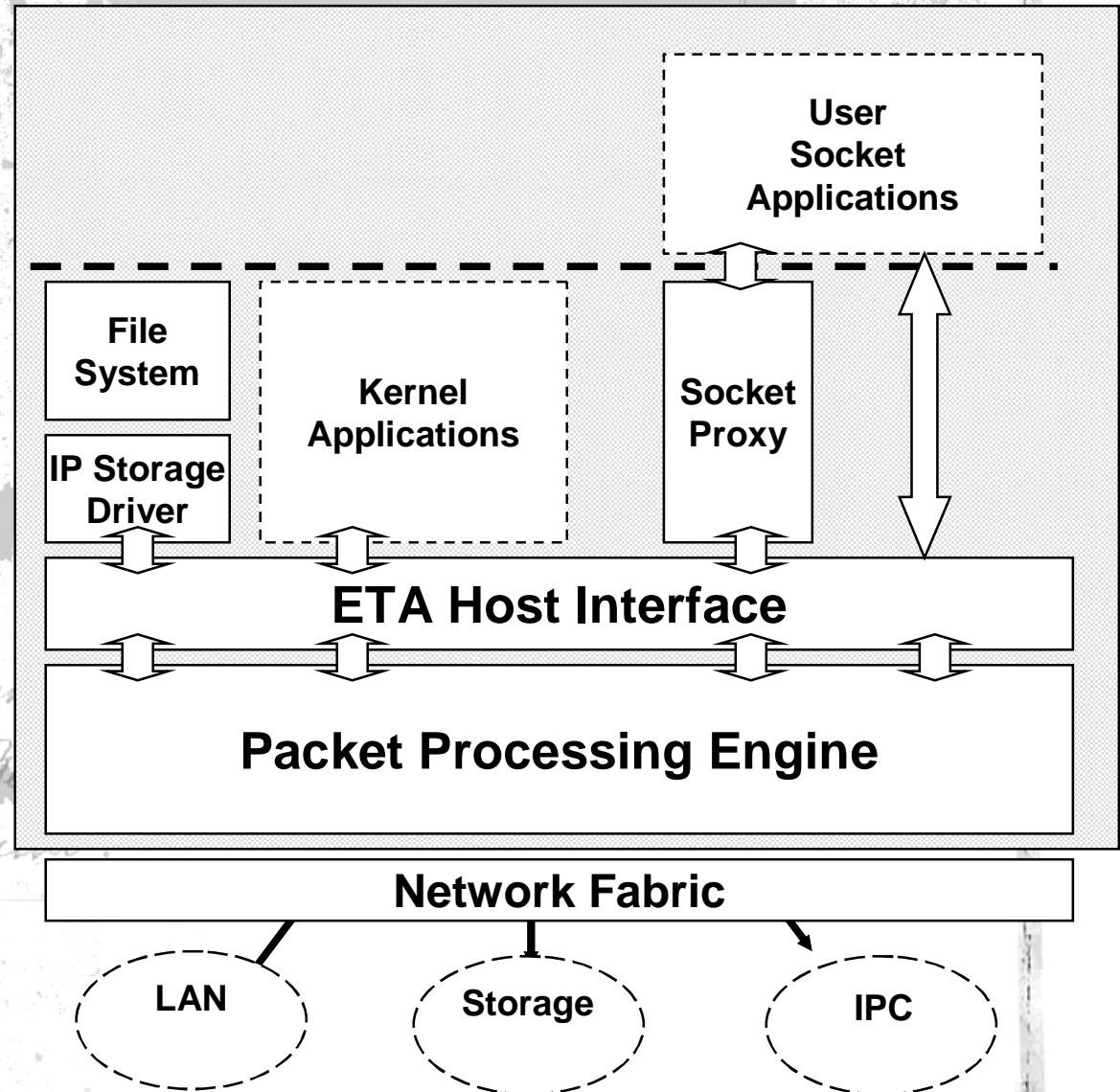


# ETA System Architecture

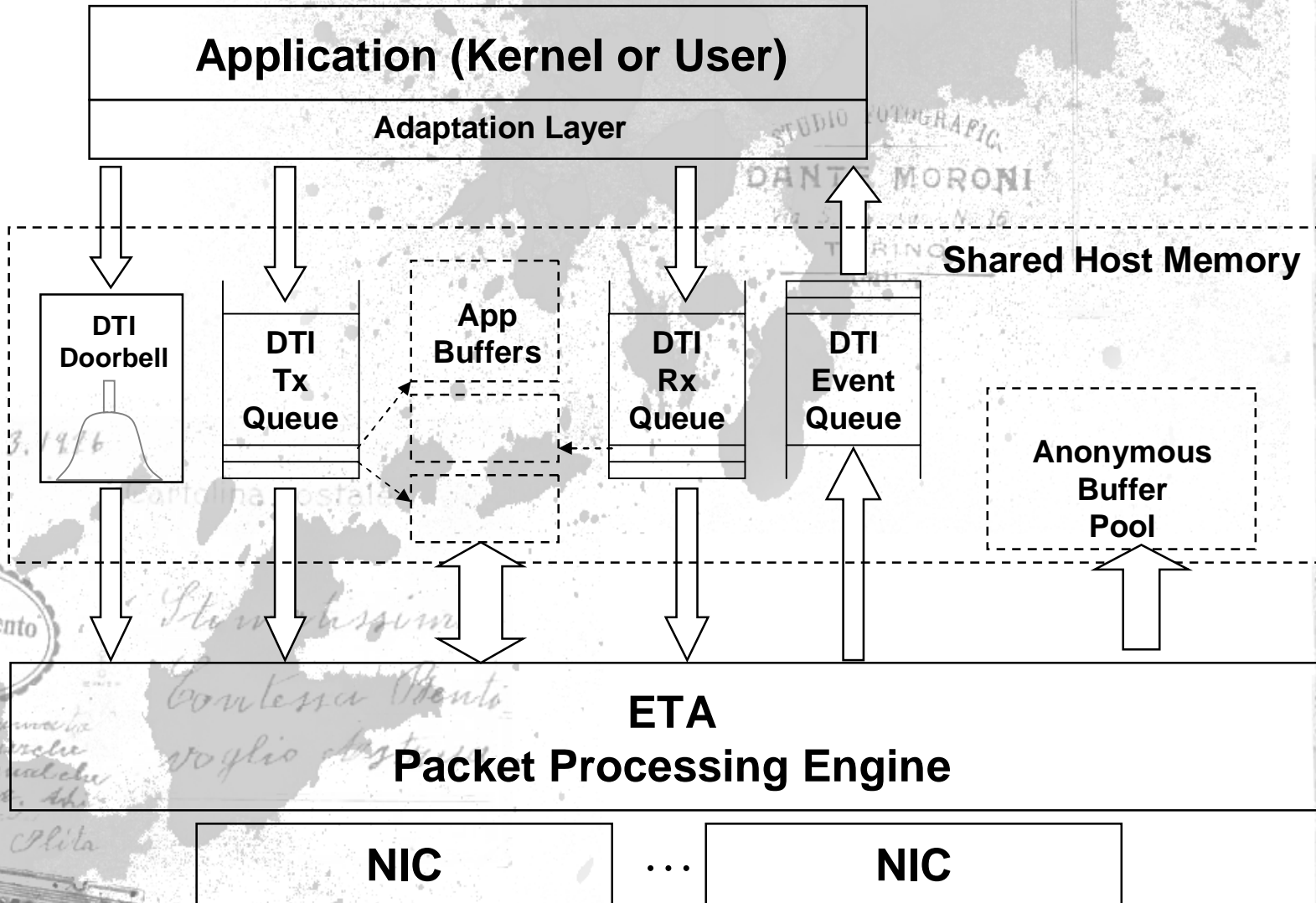
• Network stack

• Virtualized, asynchronous queuing and event handling

• Engine Architecture & platform integration

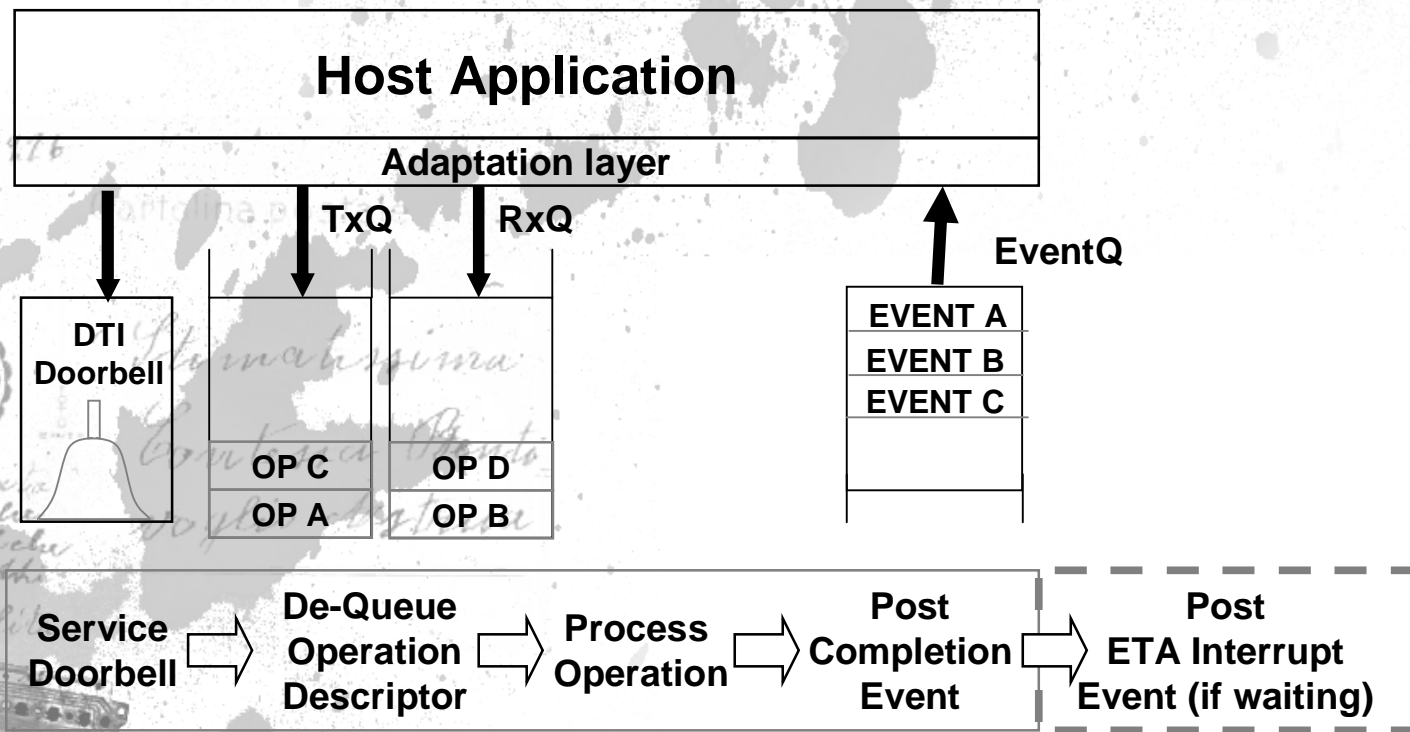


# Direct Transport Interface

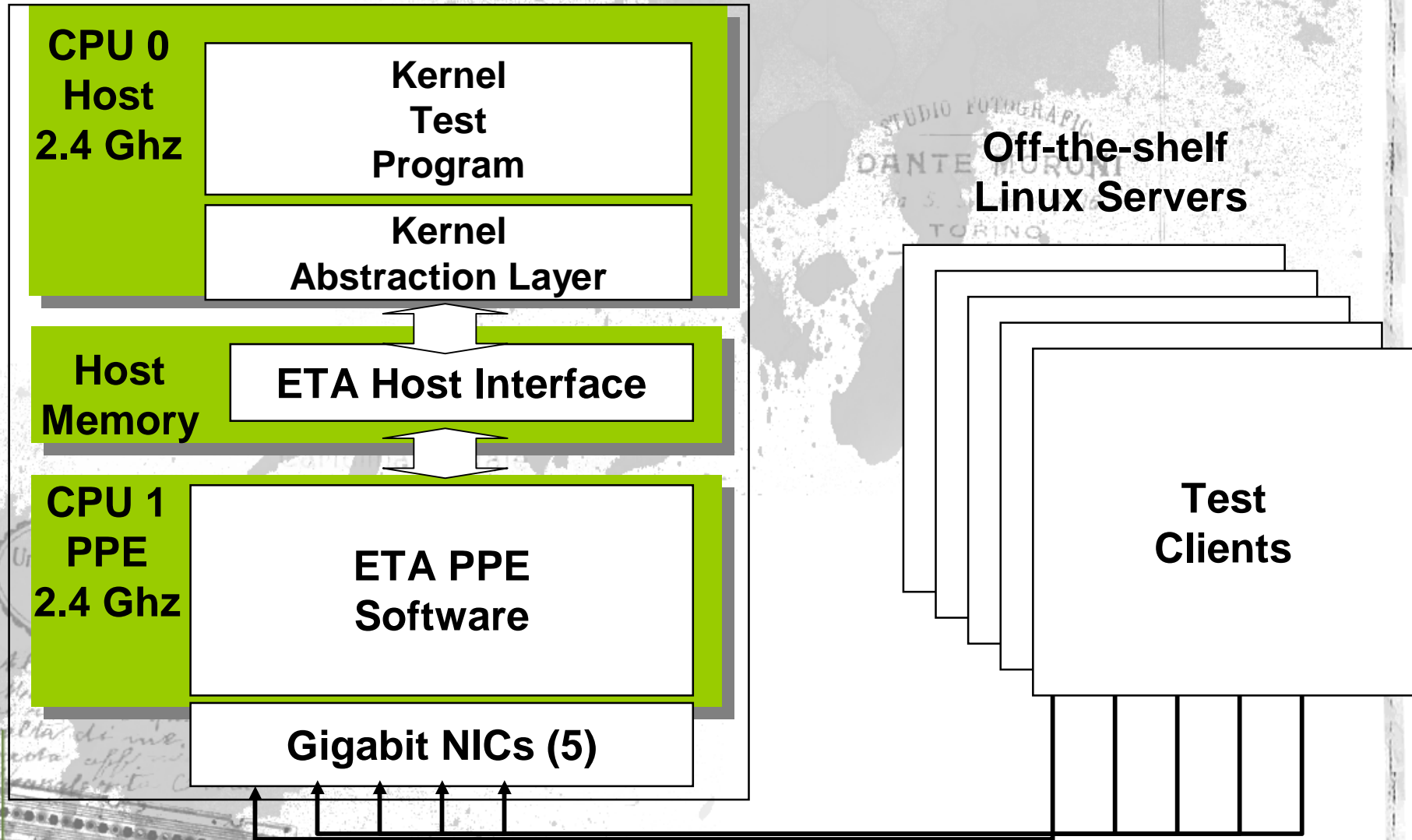


# DTI Operation Model

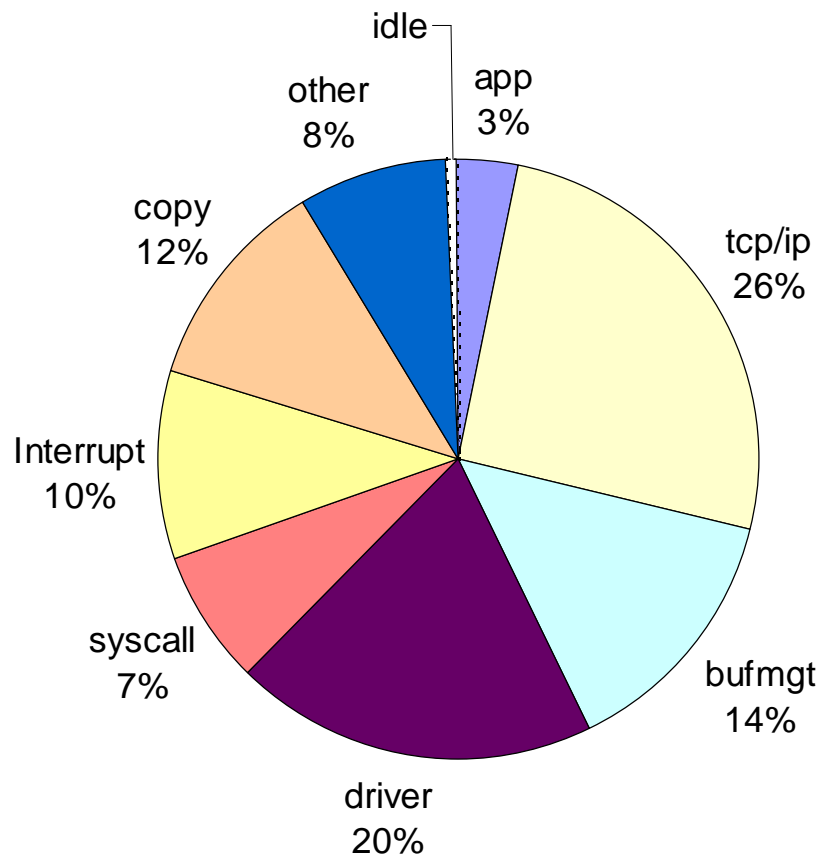
- DTI operations:
  - Connection requests (Connect, Listen, Bind, Accept, Close, ...)
  - Data transfer requests (Send, Receive)
  - Misc. operations (Set/Get Options, ...)



# ETA Test Environment



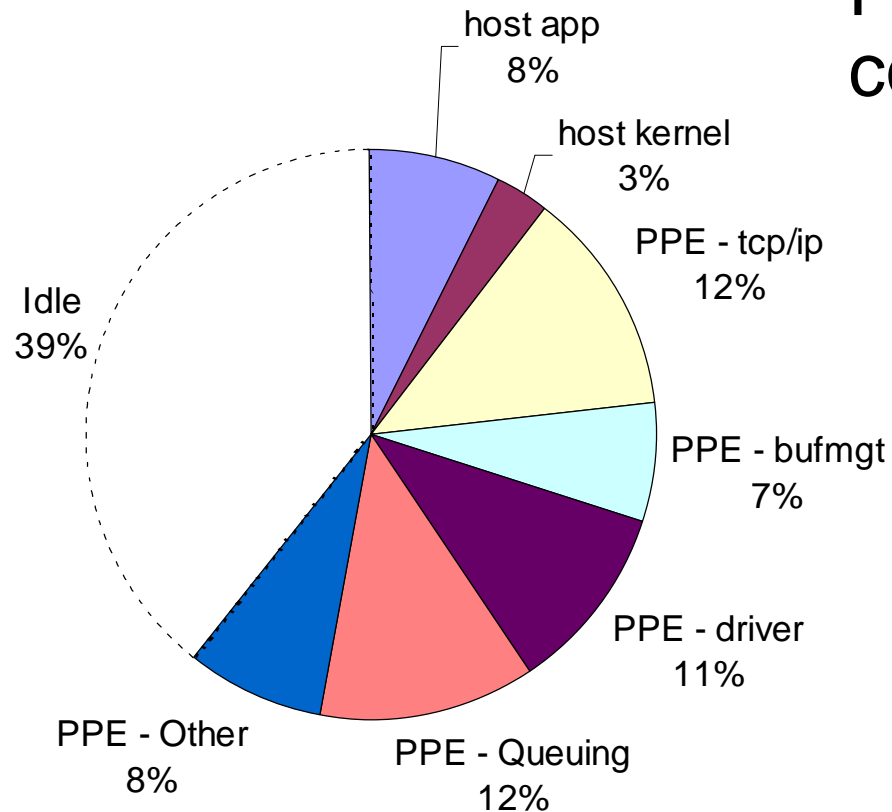
# 2P SMP Profile



Processing requirements in multiple components

TCP/IP is the largest single component, but is small compared to total. The copy overhead is required to support legacy (synchronous) socket semantics. Interrupts and system calls are required in order to time-share the CPU resources.

# ETA Profile (1 host CPU + 1 PPE)

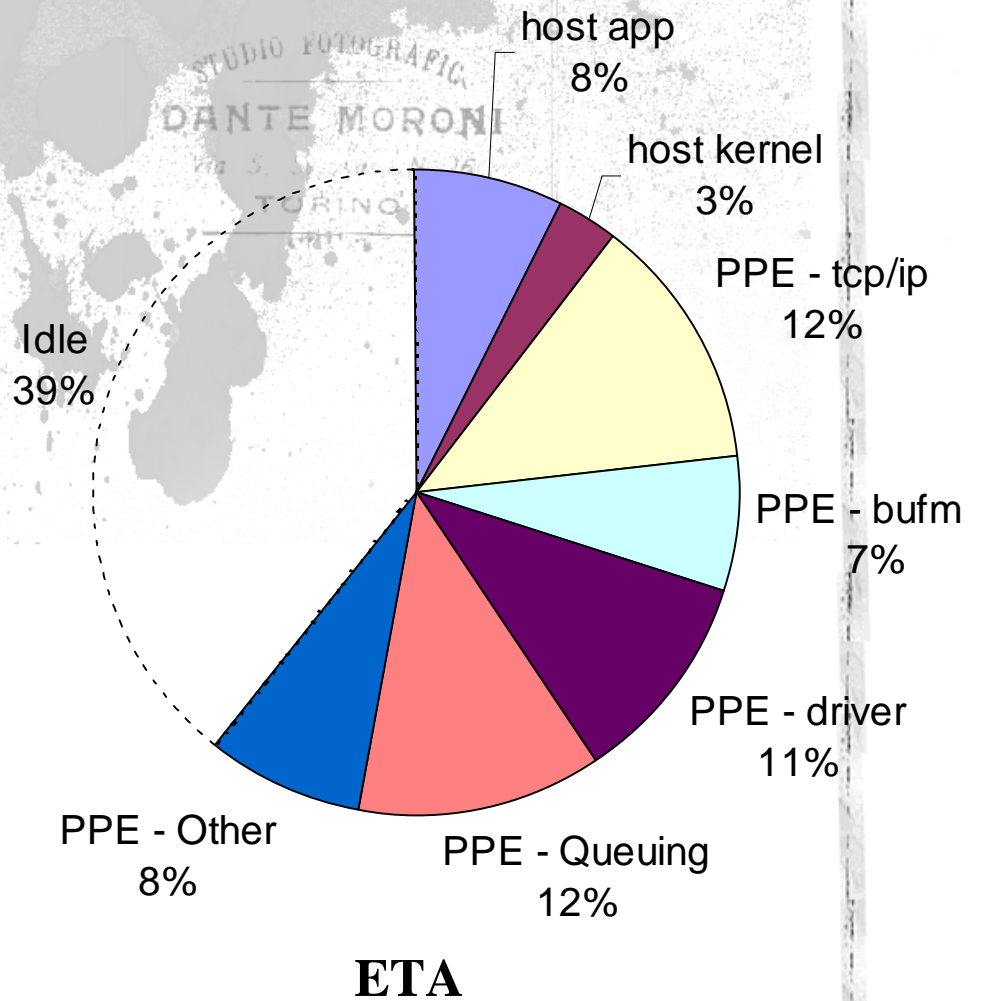
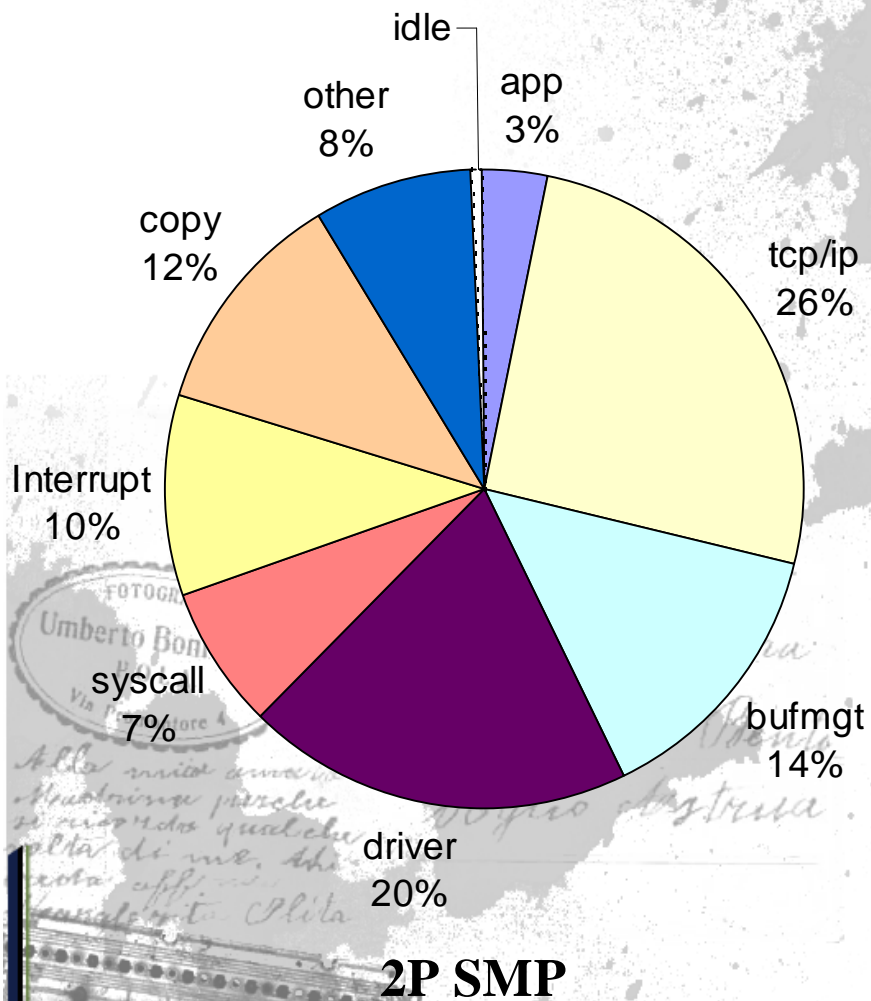


Processing times are compressed

Idle time represents CPU resource that is usable for applications  
Asynchronous queuing interface avoids copy overhead

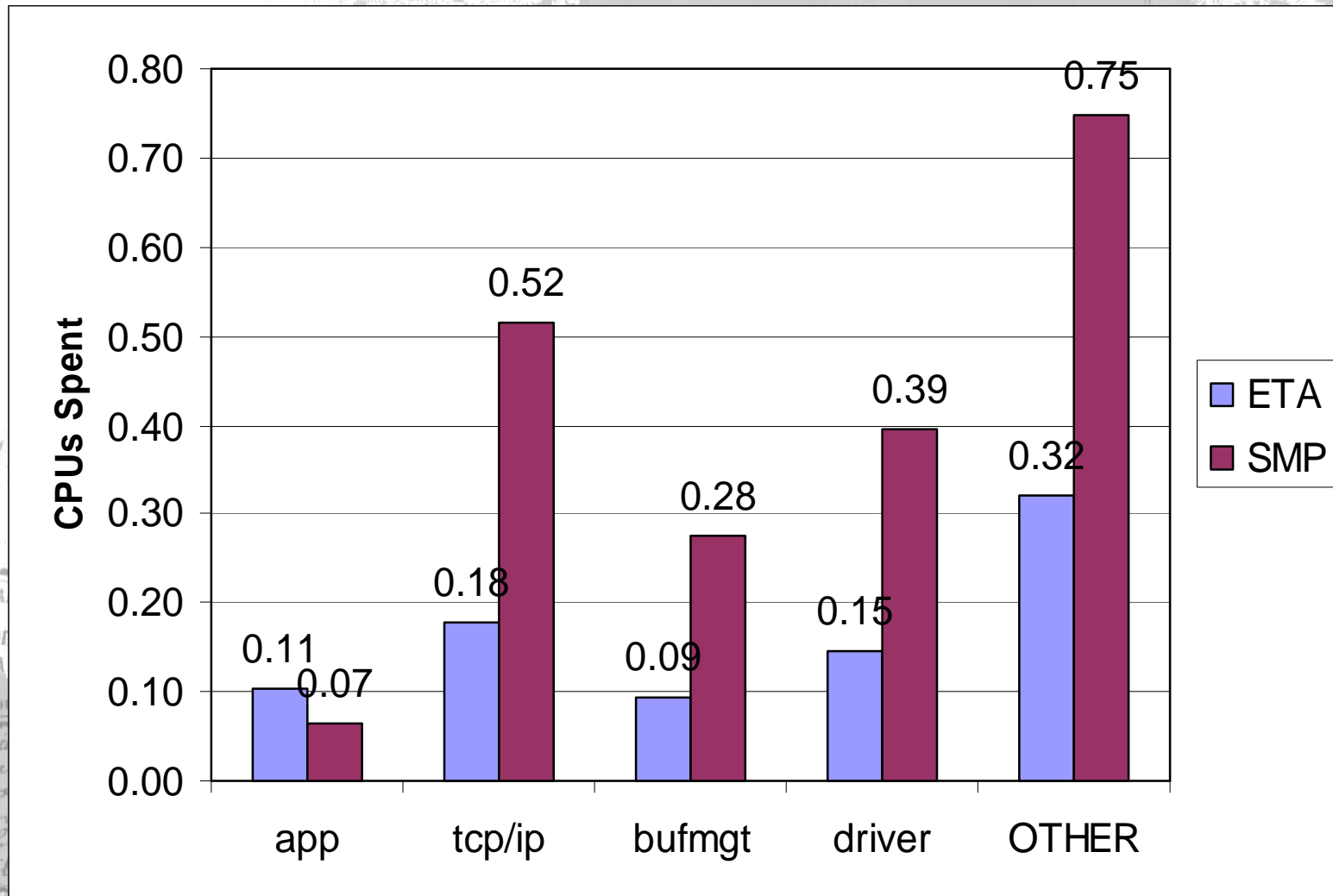
Interrupts avoided by not time-sharing CPU  
System calls avoided by ETA queuing model

# Profile Comparisons





## Normalized CPU Usage



# ETA的应用

- 基于内核级别的ids程序
- 独立于操作系统的包获取
- 接管网卡

