

Light Weight System Monitoring Tool for Unix Clusters

Matthias Wittgen
SLAC

Outline

- Motivation
- System Monitoring Tool
 - Client/Data Collection
 - What is monitored?
 - Central Server
- Example Plots
- Summary and Outlook

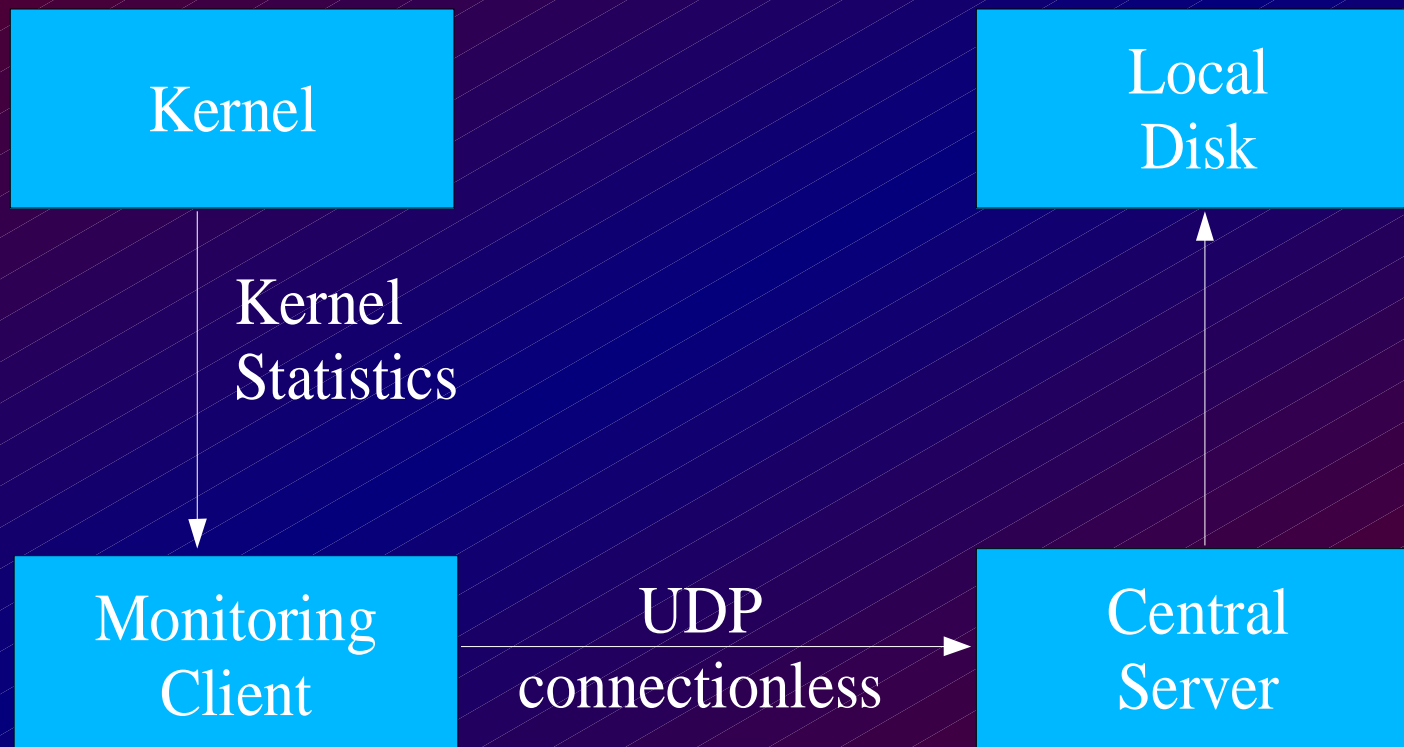
Motivation

- BaBar Online Computing Infrastructure
 - 120 servers, workstations and consoles
- New Level-3 Processing Farm
 - 50 workstations running RedHat7.3
- (File)-Servers
 - 14 servers running SunOS 5.8

Motivation

- Several existing system monitoring tools for Unix
 - Process/CPU: **top**
 - Memory: **vmstat**
 - Disk/File system: **iostat**
 - Network: **netstat**
- **Data are not stored**
- **No central collection of data**

Monitoring System



Monitoring Client

- System dependent part
 - Implementations for:
 - Linux (RedHat 7.x/8.x with kernel $\geq 2.4.18$)
 - ★ Reading file in /proc
 - ★ Plain text files (**easy**)
 - SunOS 5.8 and 5.9
 - ★ C interface to kernel variables
 - ★ Variables hardware dependent (**difficult**)
- Due to modular design other **Unix platform can easily be added**

Monitoring Client

- System independent part
 - Reading kernel statistics every 5 seconds
 - Convert data into network byte order (big endian)
 - Send data to central monitoring server
 - About 30 sec/day of CPU time used

Central Monitoring Server

- UDP used for Client/Server communication
 - Advantage:
 - ★ Connection- and state-less protocol
 - ★ Multicast possible
 - ★ Low overhead
 - Disadvantage:
 - ★ Packet size limited ≤ 1500 bytes
 - ★ Delivery not guaranteed

What is monitored?

- CPU counters (system, user, idle time)
 - 100 Hz counters, recorded for each CPU
- Number of processes, load averages
- Number of bytes/packets sent/received by network interface(s)
 - In addition, number of errors and collisions
- Number of bytes/blocks from/to local disk(s)

Data Structure

-----SYSTEM-----

header	:	4058174404
timestamp	:	1047970804
boottime	:	1047149045
nproc	:	85
ncpu	:	2
nnet	:	2
ndisk	:	2
freemem	:	651328
maxmem	:	1030596
totalswap	:	2096472
freeswap	:	2096472
avenrun_1min	:	70
avenrun_5min	:	59
avenrun_15min	:	40

-----CPU1-----

idle	:	80237586
user	:	1464675
system	:	473179

-----CPU2-----

idle	:	78613349
user	:	3226471
system	:	335888

Data Structure

-----NET1-----

ipackets	:	43731152
opackets	:	57657622
ierrors	:	0
oerrors	:	0
ibytes	:	1478733569
obytes	:	2277120777
collisions	:	0

-----NET2-----

ipackets	:	265919278
opackets	:	5747234
ierrors	:	42
oerrors	:	0
ibytes	:	3024467312
obytes	:	1122866982
collisions	:	0

-----DISK1-----

read	:	359218
write	:	4790658
reads	:	23096
writes	:	358192

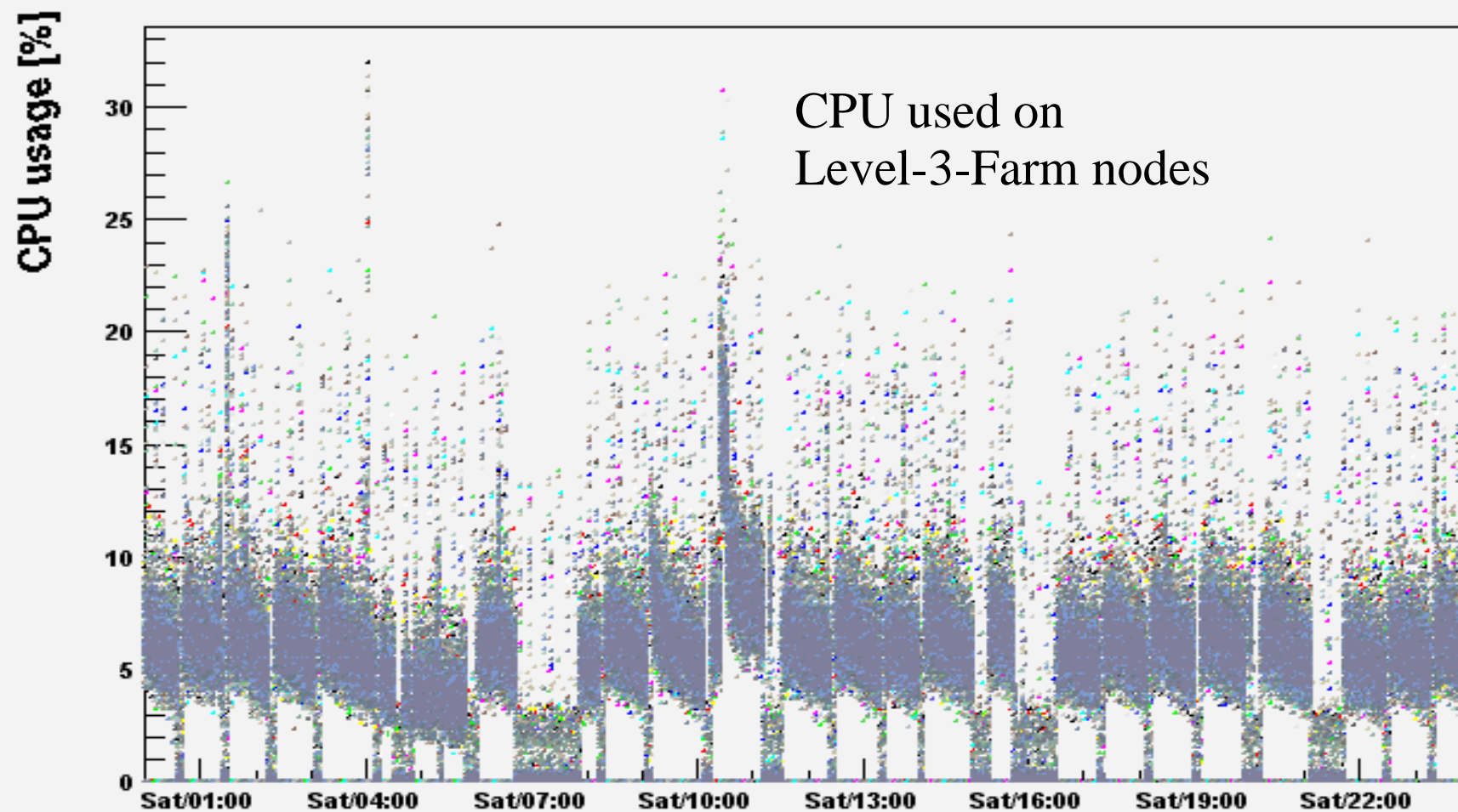
-----DISK2-----

read	:	8
write	:	0
reads	:	1
writes	:	0

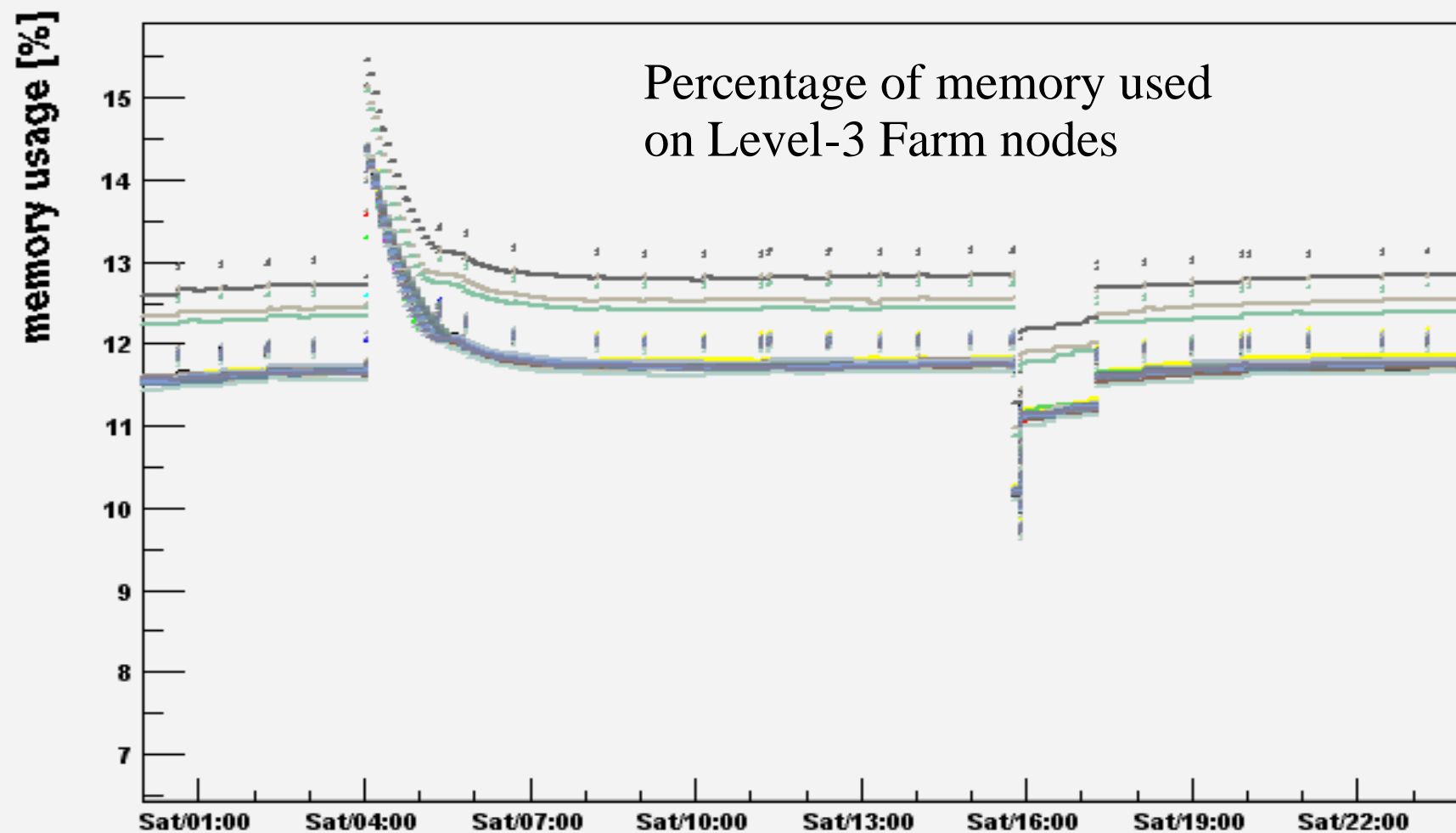
Central Monitoring Server

- 120 clients sending data to server
- Data are stored on local disk
- About 0.3% CPU used
- About 400 MB/day are recorded
- Almost no packet loss seen

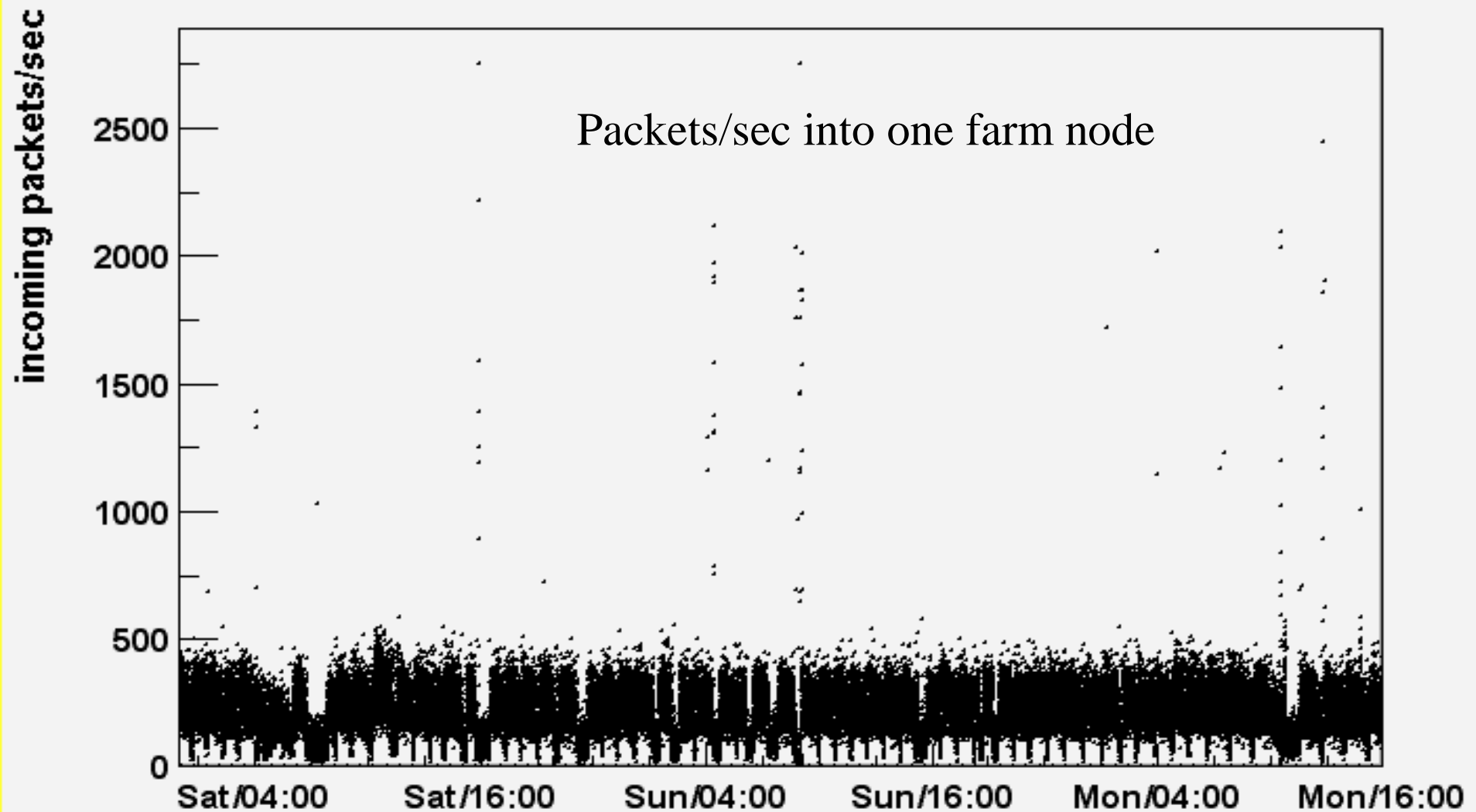
Example Plots



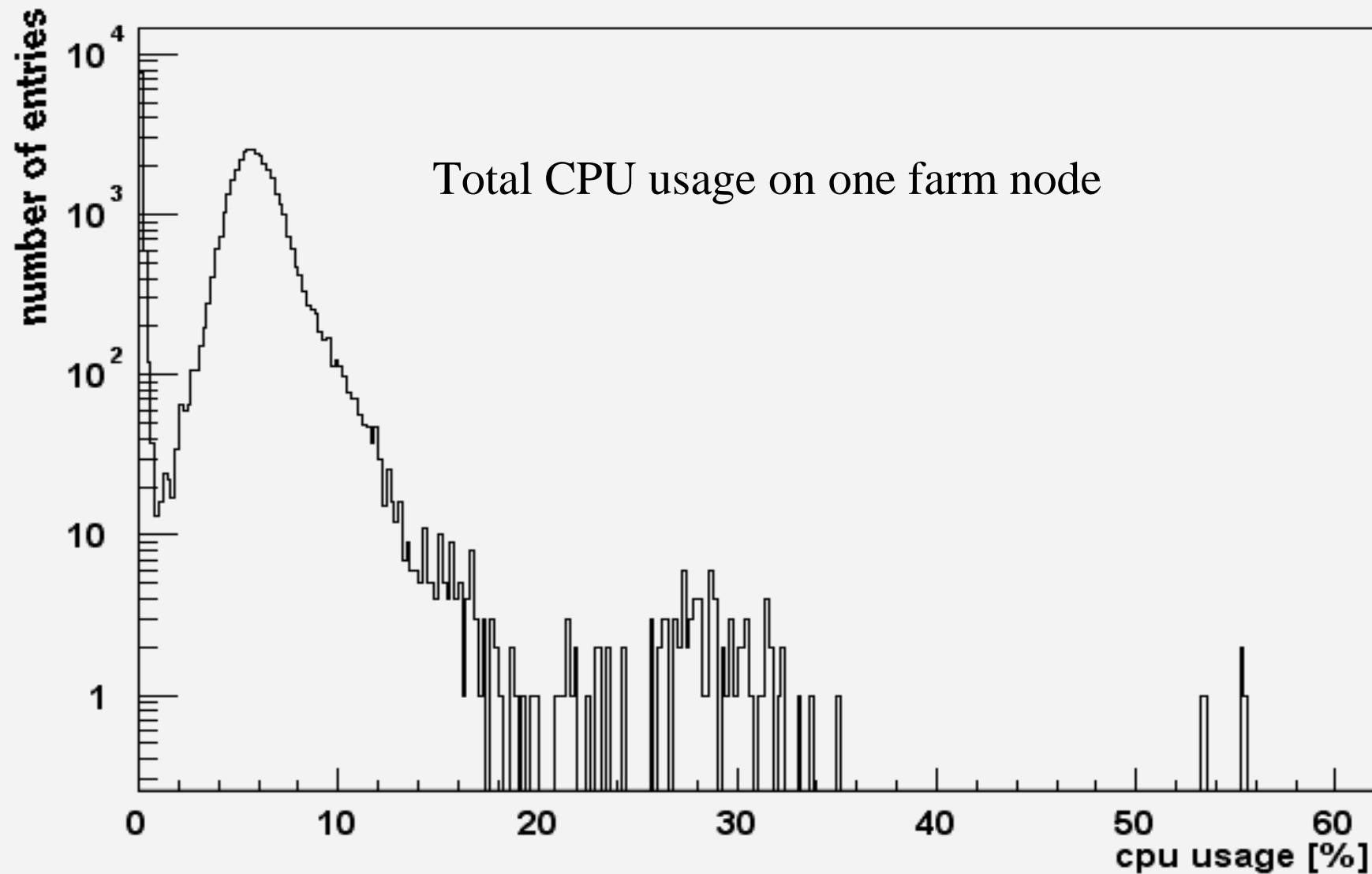
Example Plots



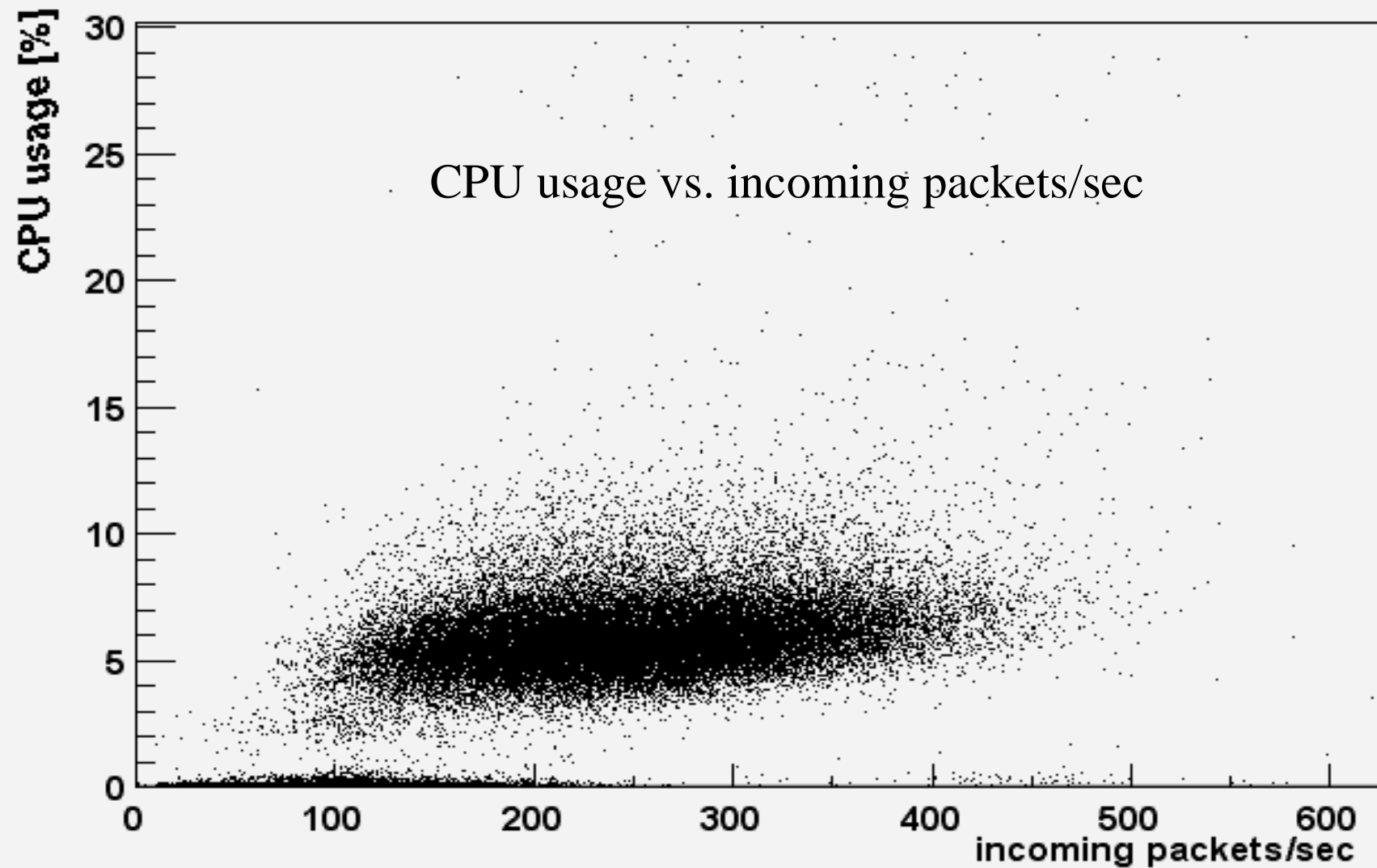
Example Plots



Example Plots



Example Plots



Summary and Outlook

- Light weight system monitoring tool
 - About 100 clients
 - Minimal CPU usage
 - No interference with other processes
- Still to do
 - Test scalability (100 \Rightarrow 1000 nodes)
 - Improve histogram generation (web interface/PHP scripts)