

# Linux Kernel Developers Summit Report

Yasunori Goto  
Sep. 11. 2006



All Rights Reserved, Copyright FUJITSU LIMITED 2006

## Who am I ?



- Yasunori Goto
- I'm developer for Memory Hotplug
  - ◆ NUMA Node's memory can be hot-added in kernel 2.6.18
  - ◆ Hot Remove needs more time
- I was a engineer of old Mainframe OS (in 199x)
  - ◆ Too different cultures against UNIX or Windows
  - ◆ To escape, I started studying about Linux or FreeBSD
  - ◆ Moved to division for Linux development
  - ◆ Attend LKST development at the beginning a bit
  - ◆ After a temporary knock down by sickness, start development of Memory Hotplug

- Pick up of Kernel Summit discussion
- Enthusiasm of Summit

- CPU Panel
- Mini Summit Summaries
- Kernel Quality and development process
- The ioctl() interface
- Kernel ABI
- Software suspend
- Documentation

- Real time
- Embedded systems
- Security
- Hypervisors
- Containers
- Automated testing
- The VFS layer
- Scalability
- DMA and IOMMU issues
- Development process 2

## Mini Summit Summaries (1/2)

- There are some Mini Summits before LKS for each topics
- 5 minutes summaries of conclusion for them
  - ◆ Here are some pieces
- Memory Management
  - ◆ New page reclaim algorism "Clock-pro"
    - Current implementation is LRU
      - ✓ But, recent used page might not be often used.
    - Clock-pro consider "often used" page
    - Current implementation should be fixed before this
  - ◆ Anti-Fragmentation by Mel Gorman
    - List base (divide free list) vs. Zone base (new easy reclaim zone)
    - No conclusion

### ■ Power Management

- ◆ Len Brown reports summary of power consumption
  - Linux is better on idle status than windows XP
  - But it becomes worse when DVD is playing
- ◆ If a little of power consumption is reduced, work time of PC can be increase sharply

## Kernel Quality And Development Process (1/2)

- Quality deterioration? (by Andrew Morton in May)
- He reported surveys of lwn.net, or bugs number in LKML and bugzilla
  - ◆ 60% hit a bug
  - ◆ 1/7th think its deteriorating
  - ◆ 1/4 think lkml response is inadequate
  - ◆ 3/5ths think bugzilla response is inadequate
  - ◆ 2/5thn think we have features-vs-stability wrong
  - ◆ 1/3rd of bugs remain unfixed
  - ◆ 1/5th of users are presently impacted by a kernel bug

### ■ Survey is valid?

- ◆ These value might include bias
- ◆ This might include rare hardware drivers
- ◆ Users and developers are increasing, but many developers think bug report is not increase

### ■ In the end, no big change...

- ◆ Big system change is not necessary
- ◆ -mm or other sub tree maintainer would take care about including new patches
- ◆ Lack of review bandwidth

## Kernel & User Space

### ■ ioctl() is not so bad?

- ◆ Easy to use (especially for application), But...
  - Hard to use from shell script
  - No compatibility between 32 and 64 bit
  - Cause of poor design
  - Hard to enumerate request number of ioctl()

### ■ Others?

- ◆ Ex) sysfs
  - Easy to see just what is available
  - hard to use sysfs for complicated
  - But dir / files might be changed yet.....

### ■ Hard to make conclusion about this

- ◆ Guideline is necessary...

### ■ There are too old documentations

#### ◆ /Documentation of kernel source

- “There are some buggy motherboards which cannot properly deal with the memory above 16MB”

✓ /Documentation/memory.txt

- Should remove it? Move somewhere?

#### ◆ Man

- Ability to write the source code doesn't equal communication ability.
- Needs to employ a special person to write documents.

## Embedded System topics

### ■ Real-time session and embedded system topics

- ◆ Boot time reducing, Dynamic tics, Kernel code size, etc...

### ■ Stock kernel should include real-time patches?

- ◆ Total patches are too big?

### ■ Real-time kernel option should be on?

- ◆ Too big impact?
  - Sleeping lock can be accepted as default?
  - Performance of desktop will be worse by it?

### ■ Many requirements

- ◆ Laser Cutter, Stock Trader, voice processing, etc...

- **Should LSM be deleted?**
  - ◆ **Because, not so many users**
    - **Only SE Linux?**
      - ✓ **AppArmor is coming**
- **AppArmor should be included?**
  - ◆ **"Path name" is used for security policy**
    - **Attacker can avoid it by making new link**
  - ◆ **Inode?**
    - **User can't specify it**
  - ◆ **Both?**
  - ◆ **No conclusion about AppArmor**
- **Conclusion about LSM**
  - ◆ **Users will not allow to remove it**

- **The interface for paravirtualization**
- **There are 2 options**
  - ◆ **Binary interface**
    - **Proposed by VMware**
  - ◆ **Source code interface**
    - **Proposed by Rusty Russell**
- **Binary interface is not popular in open source**
  - ◆ **Everybody seems to support Rusty's one.**

- There are some requirement that resources of OS should be divided for user
  - ◆ Resource Manager (CKRM)
  - ◆ Security (ex. Jail)
  - ◆ ...etc.
- Comprehensive namespace would be desirable than each independent definition
- Seems no objection in Summit
  - ◆ It would be useful, but code and interface (like /proc) should be clean up before it
- Seems high goal
  - ◆ Much time may be necessary...

- Presented by SGI (Christoph Lamer)
  - ◆ Total tone
    - Publicity of their scalability up activity
      - ✓ 1024 CPU is only special hardware like SGI's Altix...
      - ✓ Current big box feature will be on future desktop machine
    - There is good summary of scalability effort
      - ✓ <http://ftp.kernel.org/pub/linux/kernel/people/christoph/pmig/numamemory.pdf>
      - ✓ But its detail is not mentioned in this session
  - ◆ Every kernel hacker already knows his great works
    - No people objected current his activity.

### ■ Kernel version

- ◆ Release vs. Stable version.
- ◆ In conclusion, keep current style of version number

### ■ Next year's kernel summit

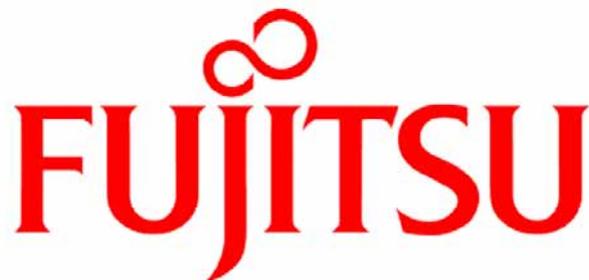
- ◆ Summit Place may be moved to other!
- ◆ Current candidates.
  - Cambridge (UK) or Vancouver (Canada)
- ◆ Followings are necessary
  - Place, enough restaurant, ease of immigration check, and well educated staff
- ◆ Not decided yet...
  - Relationship between OLS and LKS (Same place?)
  - If enough place can be reserved, more people may be able to attend it.

## Photo

### ■ Last session of Summit.



- **Every attendees are very active.**
  - ◆ No fear, No restraint
  - ◆ They have pride and confidence
- **Truth of Linus**
  - ◆ **In ML**
    - He seems to say very harmful expression many times
    - You might think he is very nervous
  - ◆ **F2F impression.**
    - He talk with full of expression, and cheerful
    - Harmful expressions are some kind of joke
      - ✓ Don't mind his word
    - But he has good leadership at important point.
      - ✓ This must be reason that he can lead Linux.



**THE POSSIBILITIES ARE INFINITE**