



# Introduction to OpenSolaris kernel

**Jakub Jermar**

**Revenue Product Engineer –  
Kernel and Virtualization**

**Sun Microsystems Czech**

# Old-New OpenSolaris Kernel

- contains old bits from AT&T
- contains old bits from BSD
- contains even some bits from Microsoft
- most of old and new bits are from Sun
  - > lots and lots of features, improvements and modifications added by Sun

# Interesting bits by Sun

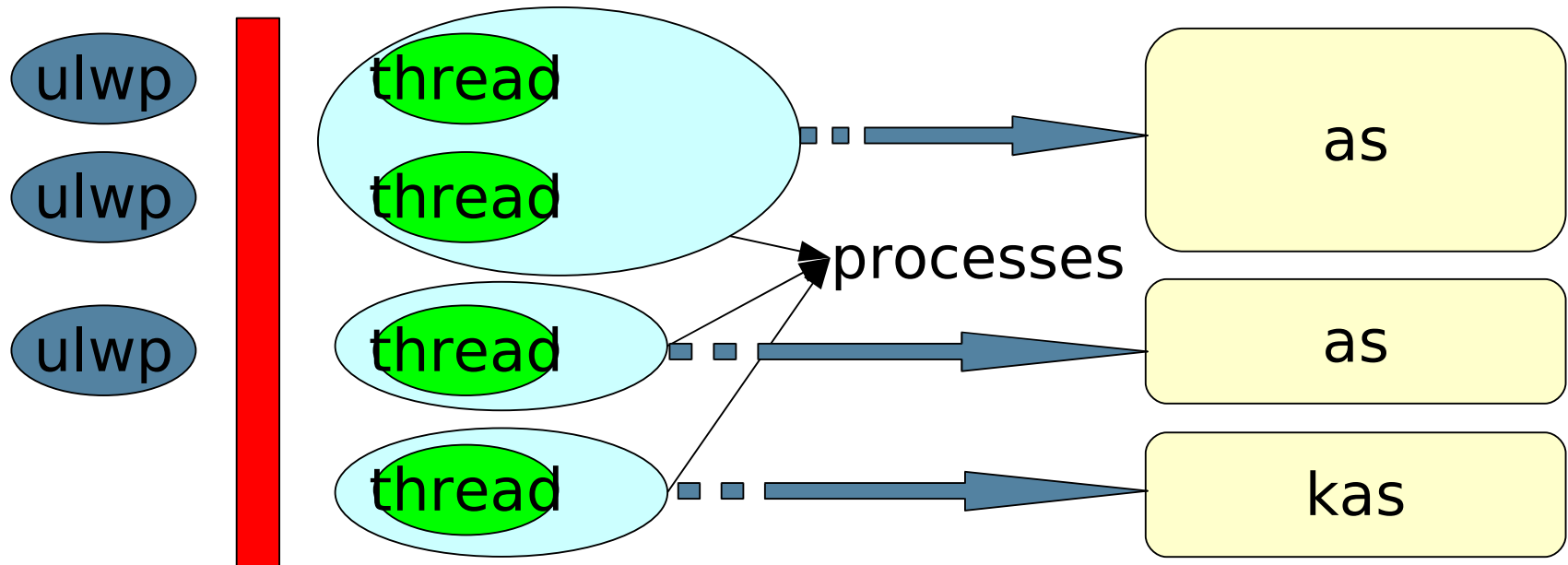
- Everyone knows DTrace and ZFS, right?
  - > let's focus on smaller interesting things
- modular and dynamically linked kernel
- thread model and important kernel citizens
- adaptive mutexes and interrupts
- SMP and preemptive scheduling
- memory management

# Modular Kernel

- everything in OpenSolaris is dynamically linked, no static binaries
- kernel is no exception
- lots of modules
  - > /kernel/\$ISADIR/genunix
  - > /platform/i86pc/\$ISADIR/kernel/unix
  - > /kernel/fs/\*
  - > /kernel/drv/\*
  - > /kernel/dtrace/\*
  - > /kernel/kmdb/\*

# Main Kernel Players

- threads (kthread\_t)
- processes (proc\_t)
- address spaces (as\_t)



# Adaptive mutexes

- traditionally, spinlocks implement active waiting, mutexes implement passive waiting
- OpenSolaris has spin mutexes and adaptive mutexes
  - > spin mutexes used by high priority interrupts and drivers; always spin
  - > adaptive mutexes spin if the mutex owner is running or block otherwise
- low-priority interrupts ( $< \text{PIL } 10$ ) -> interrupt threads

# Scheduling

- kernel is fully preemptible
- schedulable entity: `kthread_t`
- every CPU has its own dispatch queue
  - > `disp()`: kernel preempt queue, CPU-local queues, other CPUs' queues, idle thread
- scheduling classes: timeshare (TS), interactive (IA), fair share (FSS), fixed priority (FX), real-time (RT), system (SYS)

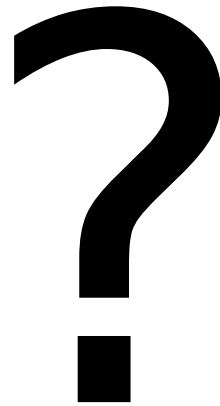
# Memory Management

- physical memory
  - > list of contiguous physical memory memsegs
  - > slab allocator
- virtual memory
  - > address spaces
  - > segments – contiguous virtual memory areas
  - > segment drivers: handle page faults
    - seg\_vn, seg\_map, seg\_kpm, seg\_dev, ...

# Referneces

- /usr/src/uts
- <http://src.opensolaris.org/source>
- McDougall, R., Mauro, J.: Solaris Internals
- mdb/kmdb

# Questions and Answers





# Introduction to OpenSolaris kernel

**Jakub.Jermar@Sun.COM**