# KERNEL REPORT

LG 전자 김준수
iamjoonsoo.kim@lge.com
js1304@gmail.com

# 개요

- 커널 개발 동향
- v5.4 ~ v5.18
- 주관적인 주제 선정

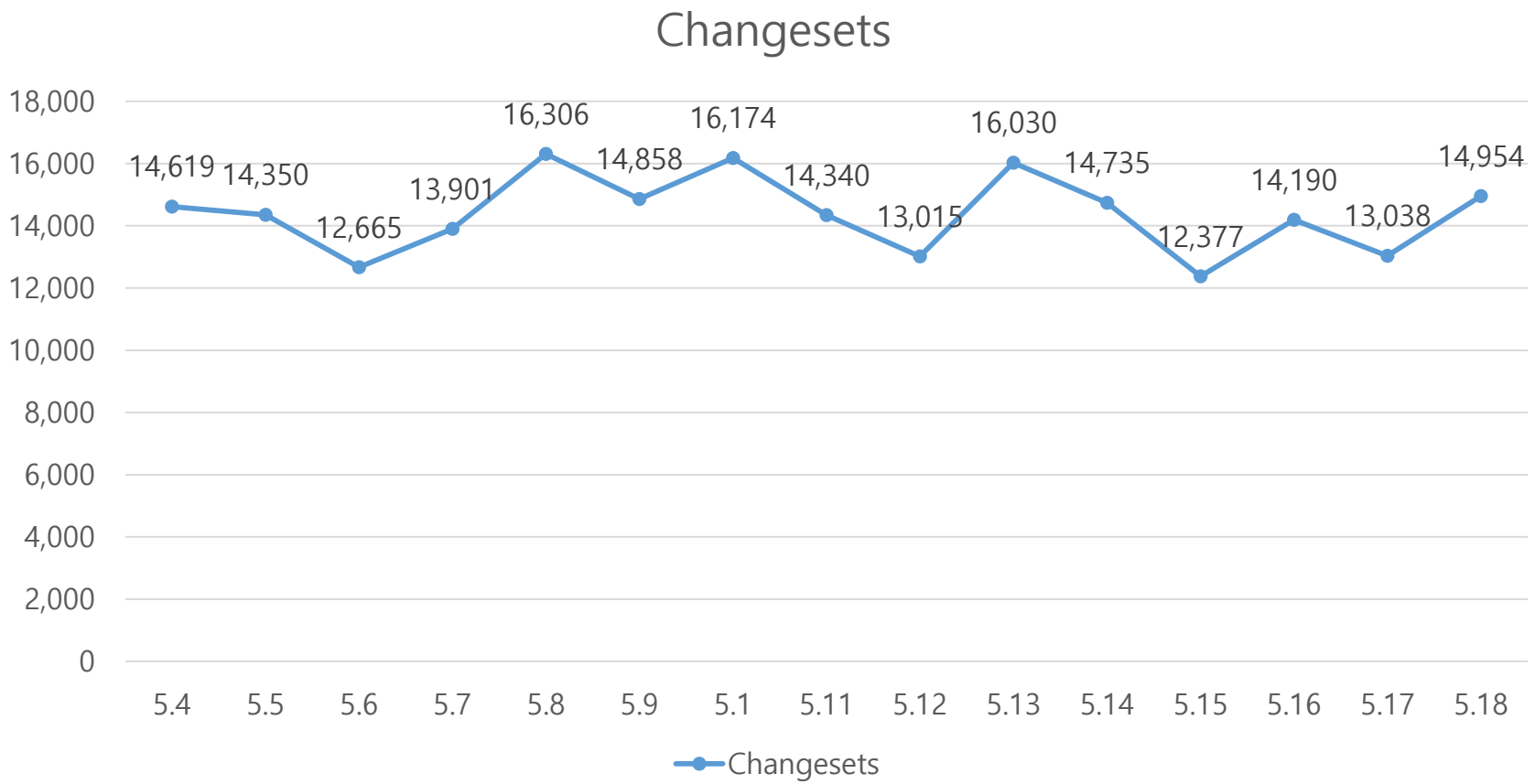
# CONTENTS

- 리눅스 커널 개발 현황
- 개발 동향
  - CORE
  - MM
  - FS
  - SECURITY/DEBUGGING
  - 선택하지 못한 주제들

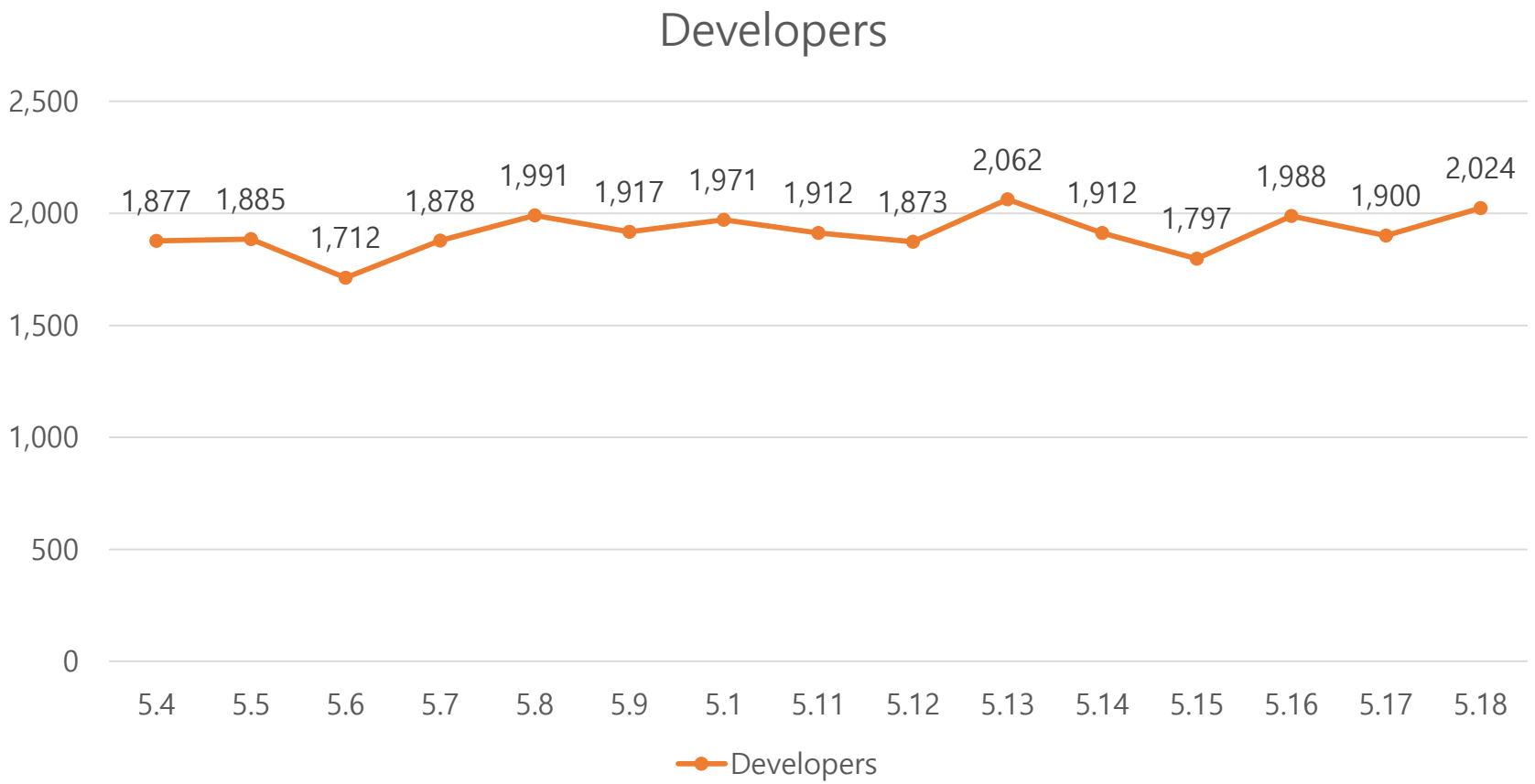
# 리눅스 커널 개발 현황

| Version | Changesets | Developers | New Developers |
|---|---|---|---|
| 5.4 | 14,619 | 1,877 | 266 |
| 5.5 | 14,350 | 1,885 | 285 |
| 5.6 | 12,665 | 1,712 | 214 |
| 5.7 | 13,901 | 1,878 | 281 |
| 5.8 | 16,306 | 1,991 | 304 |
| 5.9 | 14,858 | 1,917 | 306 |
| 5.10 | 16,174 | 1,971 | 252 |
| 5.11 | 14,340 | 1,912 | 280 |
| 5.12 | 13,015 | 1,873 | 262 |
| 5.13 | 16,030 | 2,062 | 329 |
| 5.14 | 14,735 | 1,912 | 261 |
| 5.15 | 12,377 | 1,797 | 251 |
| 5.16 | 14,190 | 1,988 | 296 |
| 5.17 | 13,038 | 1,900 | 268 |
| 5.18 | 14,954 | 2,024 | 289 |

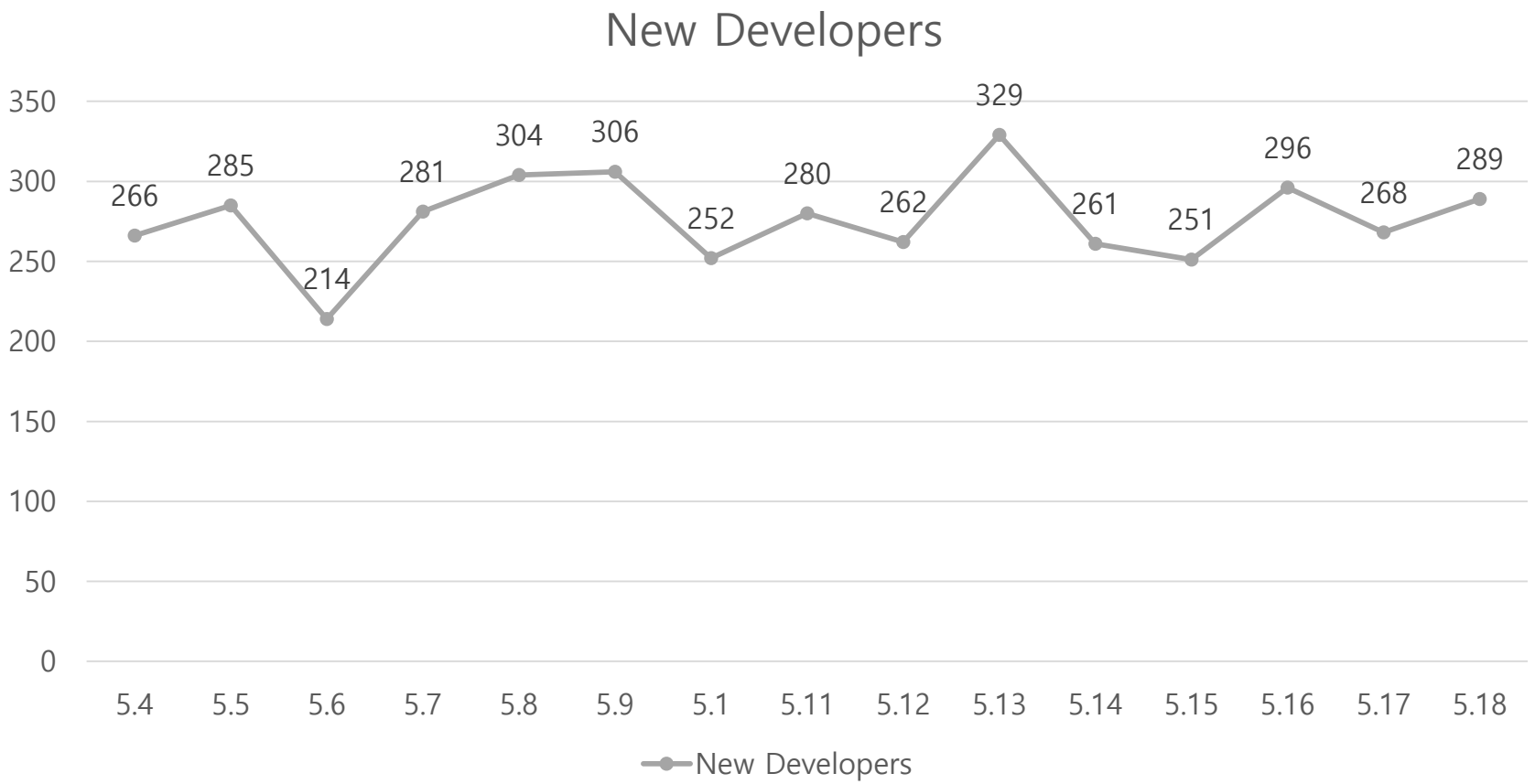
# 리눅스 커널 개발 현황


Changesets
18,000
16,000
14,000
12,000
10,000
8,000
6,000
4,000
2,000
0
14,619
14,350
12,665
13,901
16,306
14,858
16,174
14,340
13,015
16,030
14,735
12,377
14,190
13,038
14,954
5.4
5.5
5.6
5.7
5.8
5.9
5.1
5.11
5.12
5.13
5.14
5.15
5.16
5.17
5.18
Changesets

# 리눅스 커널 개발 현황


Developers
2,500
2,000
1,500
1,000
500
0
1,877
1,885
1,712
1,878
1,991
1,917
1,971
1,912
1,873
2,062
1,912
1,797
1,988
1,900
2,024
5.4
5.5
5.6
5.7
5.8
5.9
5.1
5.11
5.12
5.13
5.14
5.15
5.16
5.17
5.18
Developers

# 리눅스 커널 개발 현황


New Developers
350
300
250
200
150
100
50
0
266
285
214
281
304
306
252
280
262
329
261
251
296
268
289
5.4
5.5
5.6
5.7
5.8
5.9
5.1
5.11
5.12
5.13
5.14
5.15
5.16
5.17
5.18
New Developers

# 개발 동향: CORE

# RT

- RT 정식 지원

- 변경사항
  - v5.3: Introduce CONFIG_PREEMPT_RT
  - v5.4: Preparations for PREEMPT_RT, hr(timer), posix-timers, Enforce interrupt threading
  - v5.7: Percpu-rwsem rewrite
  - v5.8: Introduce the concept of local_locks
  - v5.10: seqlock: Introduce PREEMPT_RT support
  - v5.11: highmem: Preemptible variant of kmap_atomic & friends
  - v5.12: RCU: Some real-time enhancements
  - v5.13: softirq: Real Time awareness
  - v5.15: Real-time locking progress
  - v5.15: SLUB: reduce irq disabled scope and make it PREEMPT_RT compatible
  - v5.16: Clean up might_sleep() and make it RT aware
  - v5.17: New Real-Time Linux Analysis (RTLA) tool

# CGROUP

- cgroup v2 로 전환 완료

- 변경사항
  - v5.0: Support cpuset resource controller in cgroupv2
  - v5.2: Add freezer controller for cgroups v2
  - v5.6: memcg: Port hugetlb controller for cgroupsv2
  - v5.11: memcg: deprecate cgroup v1 non-hierarchical mode
  - v5.12: memcg: add swapcache stat for memcg v2
  - v5.17: memcg: add per-memcg vmalloc stat
  - v5.18: Add per-memcg total kernel memory stat

# 개발 동향: MM

# RECLAIM

- 메모리 회수 알고리즘 개선

- 변경사항
  - v5.3: vmscan: scan anonymous pages on file refaults
  - v5.5: Fix page aging across multiple cgroups
  - v5.8: Better behavior in memory thrashing situations
  - v5.9: Better management of anonymous memory

# MEMORY HINT

- Userspace 에서 메모리 관리의 자유도 증가

- 변경사항
  - v5.4: Two new madvise() flags: MADV_COLD and MADV_PAGEOUT
  - v5.10: Memory hints for other processes
  - v5.14: Two new madvise(2) flags: MADV_POPULATE_READ and MADV_POPULATE_READ
  - v5.15: Introduce process_mrelease(2) system call
  - v5.15: Add MADV_WILLNEED to process_madvise()
  - v5.18: madvise: Add MADV_DONTNEED_LOCKED

# 개발 동향: FS

# MS Windows FILE SYSTEM

- MS Windows 에서 개발한 FS 들에 대한 지원

- 변경사항
  - v5.4: EROFS and exFAT
  - v5.7: New exFAT file system
  - v5.15: New NTFS file system implementation
  - v5.15: ksmbd, a in-kernel SMB 3 server

# ANDROID

- Android 에서 사용할 목적으로 여러 기능들을 추가

- 변경사항
  - v5.4: fs-verity, for detecting file modifications
  - v5.5: fscrypt: In preparation for adding inline encryption support to fscrypt
  - v5.6: F2FS: Support data compression
  - v5.8: Support for Inline Encryption hardware

# 개발 동향: SECURITY/DEBUGGING

# TOOL

- S/W 의 결함을 없애고 보안 취약점을 줄이는 기능

- 변경사항
  - v5.8: Kernel Concurrency Sanitizer
  - v5.11: enable kasan for ARM
  - v5.11: hardware tag-based KASAN for ARM 64
  - v5.12: KFENCE memory error detector
  - v5.13: Support for Clang Control Flow Integrity

# 개발 동향:
# 선택하지 못한 주제들

# 선택하지 못한 주제들

- CORE: 에너지/온도에 대한 기능 강화
- CORE: 이종 CPU SYSTEM 에 대한 개발
- MM: (file) page cache 의 기본단위로 multiple page 지원
- MM: persistent memory 의 안정성/활용성을 높이는 작업
- SECURITY/DEBUGGING: H/W 의 보안 기능 추가 및 커널 지원
- SECURITY/DEBUGGING: H/W 구조로 인한 보안 취약점을 S/W 를 통해 완화, 다양한 보안 관련 기능들 추가
- BLOCK: I/O uring
- TOOL: 다수의 eBPF 기능들

# Q & A

# 참고 자료

- lwn.net
- kernelnewbies.org
- git.kernel.org (source code)

# 부록

# CORE: SCHED/POWER

- 에너지/온도 관리에 대한 기능 강화

- 변경사항
  - v5.7: Thermal Pressure in the task scheduler
  - v5.9: CPU Capacity awareness for the deadline scheduling class
  - v5.9: Power management: Make the Energy Model cover non-CPU devices
  - v5.12: Dynamic Thermal Power Management

# CORE: ASYMMETRIC CPU

- 이종 CPU SYSTEM 에 대한 처리 강화

- 변경사항
  - v5.7: Rework asymmetric CPU capacity wakeup to improve capacity utilization on asymmetric topologies (DynamIQ big.LITTLE systems)
  - v5.9: CPU Capacity awareness for the deadline scheduling class
  - v5.14: Rework CPU capacity asymmetry detection
  - v5.15: Support for asymmetric scheduling affinity

# MM: PAGE CACHE

- (file) page cache 의 기본 단위로 multiple page 를 지원

- 변경사항
  - v5.4: Experimentally enable Transparent Huge Page support for text section (executable code) of non-shmem files
  - v5.4: Change the handling of Transparent Huge Page faults
  - v5.9: Transparent Huge Pages in the page cache, preparation patches
  - v5.10: Remove assumptions of THP size
  - v5.12: Overhaul multi-page lookups for future THP
  - v5.16: Memory folios infrastructure for a faster memory management
  - v5.17: Convert much of the page cache to use folios
  - v5.18: Memory management folio patches (get_user_pages, vmscan, start on the page cache, make readahead use large folios)
  - v5.18: Filesystem conversions to folio structures

# MM: PERSISTENT MEMORY

- PMEM 의 안정성/활용성을 높이는 작업 진행

- 변경사항
  - v5.6: Explicit user-space page pinning via pin_user_pages()
  - v5.6: Fix the access of uninitialized memmaps when shrinking zones/nodes and when removing memory
  - v5.8: memory_hotplug: Interface to add driver-managed system ram
  - v5.13: Allocate memmap from hotadded memory (per device)
  - v5.13: Prohibit pinning pages in ZONE_MOVABLE
  - v5.15: Migrate memory pages to persistent memory in lieu of discard

# SECURITY: H/W

- H/W 제조사가 보안 기능을 지속적으로 추가, 커널은 이를 지원

- 변경사항
  - v5.7: ARM Kernel Pointer Authentication support
  - v5.8: Shadow Call Stack and Branch Target Identification for improved security on ARM64
  - v5.10: Support ARM Memory Tagging Extension
  - v5.18: Support for Indirect Branch Tracking on Intel CPUs

# SECURITY: S/W

- H/W 구조로 인한 보안 취약점을 S/W 를 통해 완화, 다양한 보안 관련 기능들 추가

- 변경사항
  - v5.10: Static calls for improved post-Spectre performance
  - v5.13: Support for randomising the stack address offset in each syscall
  - v5.14: Core Scheduling, for safe hyperthreading
  - v5.14: New memfd_secret(2) system call to create secret memory areas
  - v5.17: Mitigate straight-line speculation attacks
  - v5.18: Stricter memcpy() compile-time bounds checking