HIGHMEM API and future
HIGHMEM review

- 32bit kernels
- Large amounts of memory
- Kernel maps pages in ‘highmem’ as needed
- User space can use additional memory
  - Especially with PAE
Protection Keys Supervisor

- Overlays additional protections on kernel mappings.
- Thread/CPU local protections
- No TLB flushes
- PMEM Stray write protection abused kmap interface
  - Presents issues for long term transient direct map uses
Issues

• `page_address()` can’t work for PKS
• `kmap()/kmap_atomic()` are deprecated
  – See `kmap_local_page()`
• `kmap_*` vs `page_address` are confusing
• PKS is not the first user to want this
kmap_local_page()

- Can be used in any context
- Disables migration but not preemption
- Maps can only be used in the local context
- Always suitable to replace kmap_atomic()
  - kmap() calls need to be evaluated
Do we even need HIGHMEM?

- Great LWN article by Arnd Bergmann
  - https://lwn.net/Articles/838807/
- What is the need for large memory on the remaining 32bit Archs?
- Are HIGHMEM systems large core counts?
How much more should HIGHMEM API change?

- The use of page_address() will be an issue for new direct map requirements
- API Conversion on going
  - https://tinyurl.com/yc4u9vtf
Status of HIGHMEM work

- Many call sites converted by Fabio M. De Francesco
- Always looking for more help
  - Email me or Fabio to coordinate
    - “Ira Weiny” <ira.weiny@intel.com>
    - “Fabio M. De Francesco” <fmdefrancesco@gmail.com>
Additional complications

• flush_*cache_*()

• How do we test with more distros dropping 32 bit support?

• Folios
Future

- Stop adding HIGHMEM support to new archs
- New interface to the direct map?
  - Preserve/redefine kmap?
  - Lightweight vmap?
- Alternatives to the direct map?
  - Stop mapping all of memory?
  - Performance?
Thank you

- Thank you to all the reviewers of conversions so far
- Questions, comments, and conversions welcome
References

• End to HIGHMEM – https://lwn.net/Articles/813201/
• Future of 32bit Linux – https://lwn.net/Articles/838807/
• Kmap replacement tracking sheet – https://docs.google.com/spreadsheets/d/1i_ckZ10p90bH_CkxD2bYNi05S2Qz84E2OFPv8zq__0w/edit#gid=1679714357
• Memory size is going down – https://lwn.net/ml/linux-kernel/CAK8P3a3pzgVvwyDhHPoiSOqyv+h_ixbsdWMqG3sELenRJqFuew@mail.gmail.com/
References

- Folios – https://lore.kernel.org/linux-fsdevel/Yv1DzKKzkDjwVuKV@casper.infradead.org/T/#m79ce4caba197172979faa68944d4cf4cf7941f49
- Don’t add HIGHMEM architectures – https://lore.kernel.org/lkml/CAK8P3a3LokurC0n9XiwtPQh9ZgQcswMKY4b+TEsQh1VgYDNeWA@mail.gmail.com/
References


• ARM LPAE –
  https://elinux.org/images/6/6a/Elce11_marinas.pdf

• PKS patch sets
  – https://lore.kernel.org/lkml/20220419170649.1022246-1-ira.weiny@intel.com/
  – https://lore.kernel.org/lkml/20210830235927.6443-1-rick.p.edgecombe@intel.com/