LevelDB Study
Week 5

2022. 07. 05

Presented by Min-guk Choi

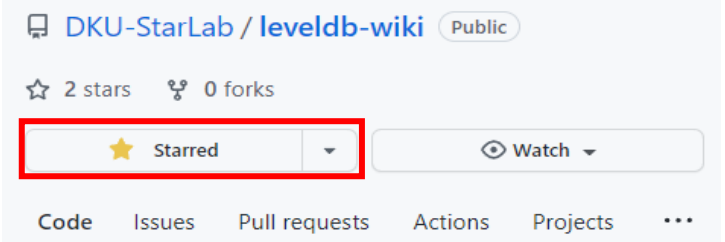
koreachoi96@gmail.com

1. Experiment Report

2. How to analyze LevelDB

- Experiment Report
  - ✓ leveldb-wiki (kor)
  - ✓ Notice

leveldb-wiki (kor)



https://github.com/DKU-StarLab/leveldb-wiki

README.md

leveldb-wiki (kor)

Background

1. What is a key-value store?
2. Why open source?
3. Job market

Analysis

0. Overall
1. Key-Value Interface
2. WAL
3. Memtable
4. Compaction
5. SSTable
6. Bloom Filter
7. Cache
8. Manifest

Benchmarks Experiments

- WAL
- Memtable
- Compaction
- SSTable
- Bloom Filter
- Cache

main leveldb-wiki / benchmarks /

korea-choi readme update

..	
bloomfilter.md	readme update
cache.md	readme update
compaction.md	readme update
memtable.md	readme update
sstable.md	readme update
wal.md	readme update

11 lines (6 sloc) 96 Bytes

WAL

Hypothesis

Design

Experiment Enviornment

Result

Discussion

Notice

- Write your experiment report into markdown documents.
  - Don't forget to cite the source of the referenced figures.
  - Don’t forget to add a watermark on your figures and graphs.

- Upload your documents through pull-request.
  - You can also change front read page if you need.
  - Pull request until **Tuesday, 8/2 11AM.**
    - https://github.com/DKU-StarLab/leveldb-wiki

- It will be finally uploaded at https://sslab.dankook.ac.kr/leveldb-wiki

1. Experiment Report

2. How to analyze LevelDB

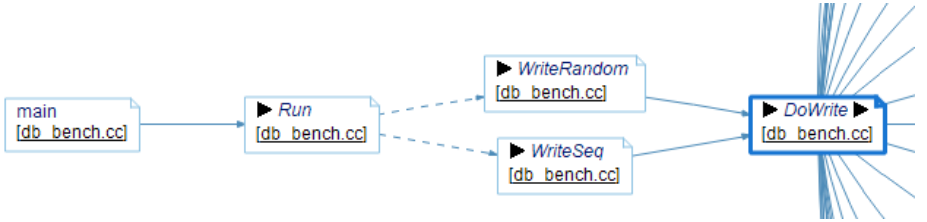
- How to analyze LevelDB
  - ✓ What/When to use
  - ✓ Where to start?
  - ✓ Source Code
  - ✓ Code Flow
  - ✓ Watermark
  - ✓ PPT

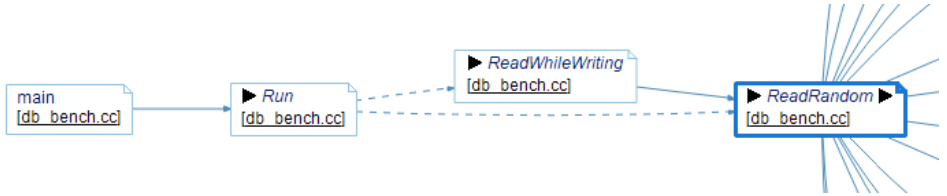
What/When to use?

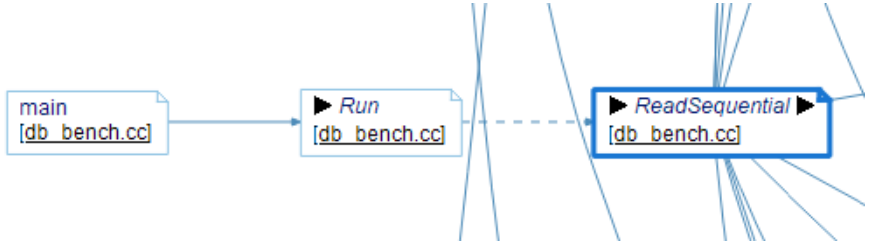
	1	2
Static Analysis Tool **- Remarks & Code**	VS code - Go to Definition - Go to References - Search - File Explorer	Understand - Class Diagram ~~- Object References~~ ~~- Butter Fly~~
Dynamic Analysis Tool **- Code Flow/Tracing**	Uftrace - Record - Replay - Tracing	GDB - Break point - Step into / Next - Print / Display - Line by Line - Argument, Variable

Where to start?

- Start with basic operations
  - Put(Write) operations
    - db_bench –benchmarks=“fillseq, fillrandom”
    - db_bench.cc:DoWrite
  - Get(Read) operations
    - db_bench --benchmarks=“readrandom”
    - db_bench.cc:ReadRandom
  - Seek(Scan) operations
    - db_bench –benchmarks=“readseq”
    - db_bench.cc:ReadSequential







Source Code

Topics

Topic	User API	Internal Operation	Contents	Source Code
WAL/Manifest	Put	Open Flush Compaction	Version Control Log Format Manifest Format	log_.*h* *version*_.h db_impl.h repair.cc
Memtable	Put Get Iterator	Flush	Skiplist Arena Batch Write	wirte_batch_internal.h skiplist.h memtable.h db_impl.h arena.h
Compaction	Put	Compaction	Compaction Policy Merge Iterator	db_impl.h merger.h version_set.h
SSTable	Get	Flush	SST Format	table/

https://github.com/DKU-StarLab/leveldb-study/blob/main/analysis/README.md

Function

WAL/Manifest	Memtable
leveldb::log::Writer::AddRecord	
leveldb::_GLOBAL__N_1::PosixWritableFile::Append	leveldb::Version::PickLevelForMemTableOutput
leveldb::log::Writer::EmitPhysicalRecord	leveldb::DBImpl::CompactMemTable
leveldb::VersionSet::PickCompaction	leveldb::DBImpl::WriteLevel0Table
leveldb::VersionSet::LogAndApply	leveldb::SkipList::RandomHeight
leveldb::VersionSet::AppendVersion	leveldb::SkipList::NewNode
leveldb::VersionSet::NeedsCompaction	leveldb::SkipList::Equal
leveldb::VersionEdit::VersionEdit	leveldb::SkipList::Node::SetNext
leveldb::Version::OverlapInLevel	leveldb::SkipList::Iterator
leveldb::DBImpl::MaybeScheduleCompaction	leveldb::Arena::MemoryUsage
leveldb::DBImpl::InstallCompactionResults	leveldb::Arena::AllocateNewBlock
leveldb::_GLOBAL__N_1::PosixWritableFile::SyncFd	leveldb::Arena::~Arena
leveldb::_GLOBAL__N_1::PosixEnv::RemoveFile	

Function

Compaction	SSTable
leveldb::DBImpl::BGWork leveldb::DBImpl::DoCompactionWork leveldb::DBImpl::FinishCompactionOutputFile leveldb::DBImpl::InstallCompactionResults leveldb::DBImpl::CleanupCompaction leveldb::NewTwoLevelIterator leveldb::NewMergingIterator leveldb::_GLOBAL__N_1::MergingIterator::SeekToFirst leveldb::_GLOBAL__N_1::MergingIterator::key	leveldb::ReadBlock leveldb::TableBuilder::Add leveldb::TableBuilder::Flush leveldb::TableBuilder::WriteBlock leveldb::port::Snappy_Compress leveldb::_GLOBAL__N_1::HandleTable::Lookup leveldb::_GLOBAL__N_1::PosixMmapReadableFile::Read leveldb::_GLOBAL__N_1::TwoLevelIterator::InitDataBlock leveldb::_GLOBAL__N_1::PosixWritableFile::Append leveldb::_GLOBAL__N_1::PosixWritableFile::Sync

Function

Bloom Filter	Cache
leveldb::Table::ReadFilter leveldb::FilterBlockBuilder::Finish leveldb::FilterBlockBuilder::StartBlock leveldb::FilterBlockBuilder::GenerateFilter leveldb::FilterBlockBuilder::FilterBlockBuilder leveldb::FilterBlockBuilder::AddKey leveldb::FilterBlockReader::KeyMayMatch leveldb::_GLOBAL__N_1::BloomFilterPolicy::KeyMayMatch leveldb::FilterBlockReader::FilterBlockReader leveldb::_GLOBAL__N_1::BloomHash leveldb::_GLOBAL__N_1::BloomFilterPolicy::CreateFilter leveldb::_GLOBAL__N_1::BloomFilterPolicy::Name	leveldb::TableCache::FindTable leveldb::DeleteCachedBlock leveldb::_GLOBAL__N_1::LRUCache::Release leveldb::_GLOBAL__N_1::LRUCache::LRU_Remove leveldb::_GLOBAL__N_1::LRUCache::LRU_Append leveldb::_GLOBAL__N_1::LRUCache::FinishErase leveldb::_GLOBAL__N_1::LRUCache::Unref leveldb::_GLOBAL__N_1::ShardedLRUCache::Shard leveldb::_GLOBAL__N_1::ShardedLRUCache::Lookup leveldb::_GLOBAL__N_1::ShardedLRUCache::Insert leveldb::_GLOBAL__N_1::PosixMmapReadableFile::Read

Code Flow

```
========== Back-trace ==========
  5.188   s : (100000) leveldb::SkipList::Insert
  5.188   s : (100000) leveldb::MemTable::Add
  5.188   s : (100000) leveldb::_GLOBAL__N_1::MemTableInserter::Put
  5.188   s : (100000) leveldb::WriteBatch::Iterate
  5.188   s : (100000) leveldb::WriteBatchInternal::InsertInto
  5.188   s : (100000) leveldb::DBImpl::Write
  5.188   s : (100000) leveldb::Benchmark::DoWrite
  5.188   s : (100000) leveldb::Benchmark::WriteRandom
  5.188   s : (100000) leveldb::Benchmark::ThreadBody
            :
========== Call Graph ==========
  5.188   s : (100000) leveldb::SkipList::Insert
  4.824   s :  ├─(100000) leveldb::SkipList::FindGreaterOrEqual
 13.450 ms :  │  ├─(100000) leveldb::SkipList::GetMaxHeight
            :  │  │
473.144 ms :  │  ├─(2495692) leveldb::SkipList::Node::Next
            :  │  │
  4.121   s :  │  └─(2495692) leveldb::SkipList::KeyIsAfterNode
  3.903   s :  │     (2404364) leveldb::MemTable::KeyComparator::operator()
993.576 ms :  │     ├─(4808728) leveldb::GetLengthPrefixedSlice
            :  │     │
  2.544   s :  │     └─(2404364) leveldb::InternalKeyComparator::Compare
492.141 ms :  │       (2404364) leveldb::_GLOBAL__N_1::BytewiseComparatorImpl::Compare
            :  │
 25.035 ms :  ├─(100000) leveldb::SkipList::RandomHeight
 16.364 ms :  (133759) leveldb::Random::OneIn
  4.255 ms :  (133759) leveldb::Random::Next
            :
 11.891 ms :  ├─(100026) leveldb::SkipList::GetMaxHeight
            :  │
 30.790 ms :  ├─(100000) leveldb::SkipList::NewNode
  5.508 ms :  │  ├─(100000) leveldb::Arena::AllocateAligned
  1.455 ms :  │  (890) leveldb::Arena::AllocateFallback
  1.384 ms :  │  (890) leveldb::Arena::AllocateNewBlock
```

```
Help: (press any key to exit)

 ARROW          Navigation
 PgUp/Dn
 Home/End
 Enter          Fold/unfold graph or Select session
 G              Show (full) call graph
 g              Show call graph for this function
 R              Show uftrace report
 r              Show uftrace report for this function
 s              Sort by the next column in report
 I              Show uftrace info
 S              Change session
 O              Open editor
 c/e            Collapse/Expand direct children graph
 C/E            Collapse/Expand all descendant graph
 n/p            Next/Prev sibling
 u              Move up to parent
 l              Move to the longest executed child
 j/k            Move down/up
 z              Set current line to the center of screen
 /              Search
 </>/N/P        Search next/prev
 v              Show debug message
 f              Customize fields in graph or report mode
 h/?            Show this help
 q              Quit
```

Example Figure

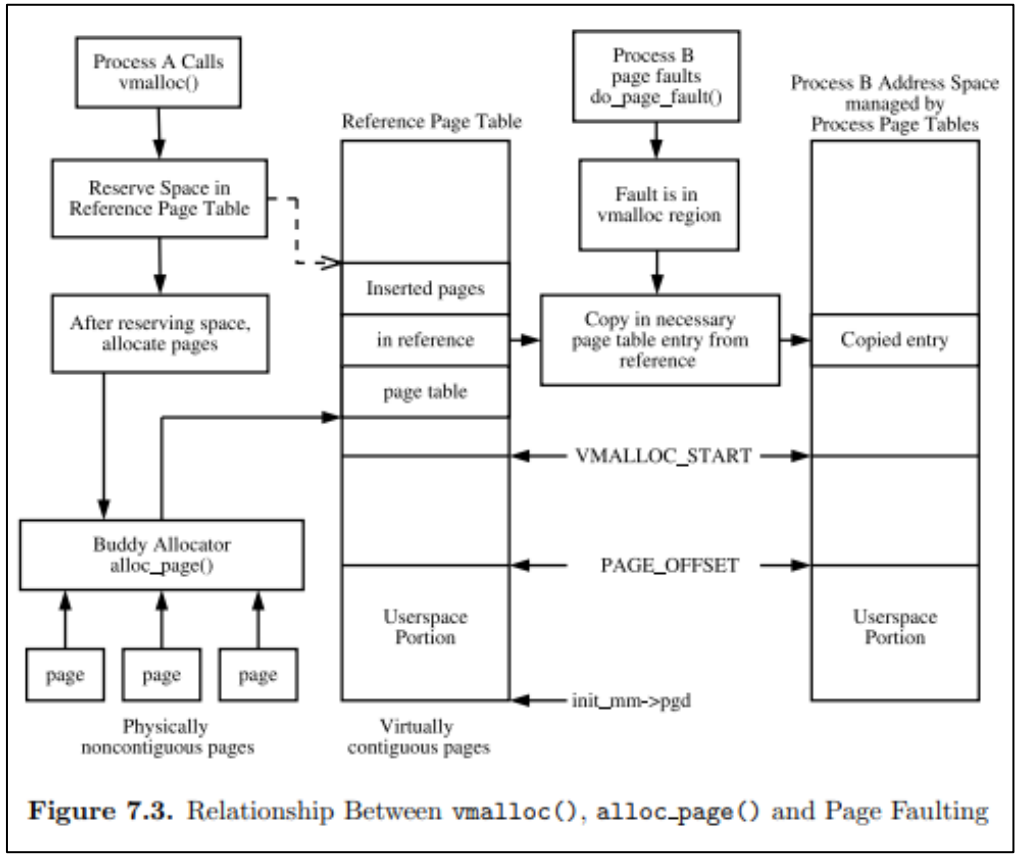


Figure 7.3. Relationship Between `vmalloc()`, `alloc_page()` and Page Faulting

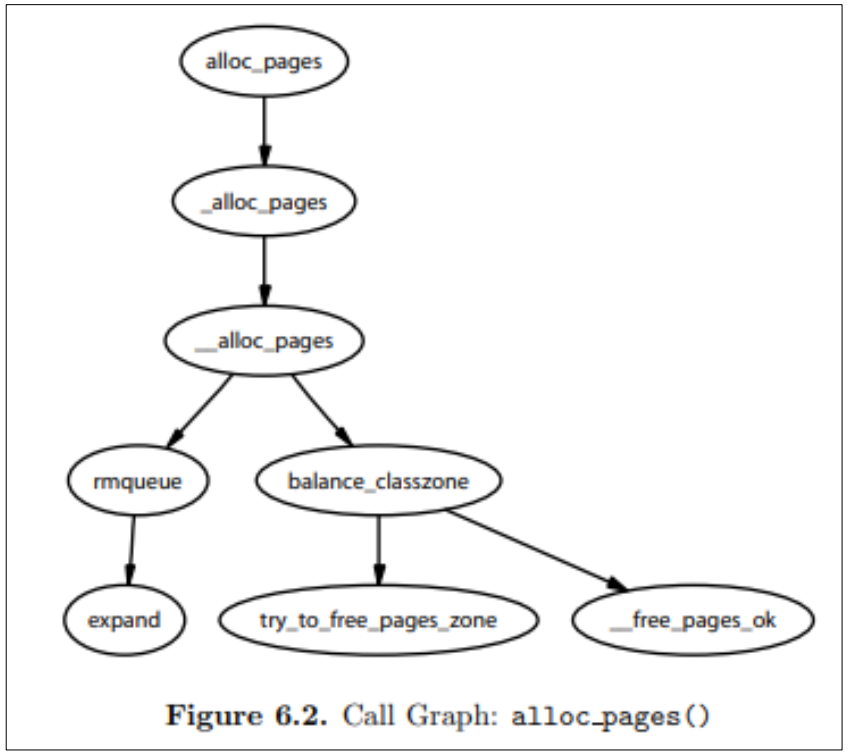
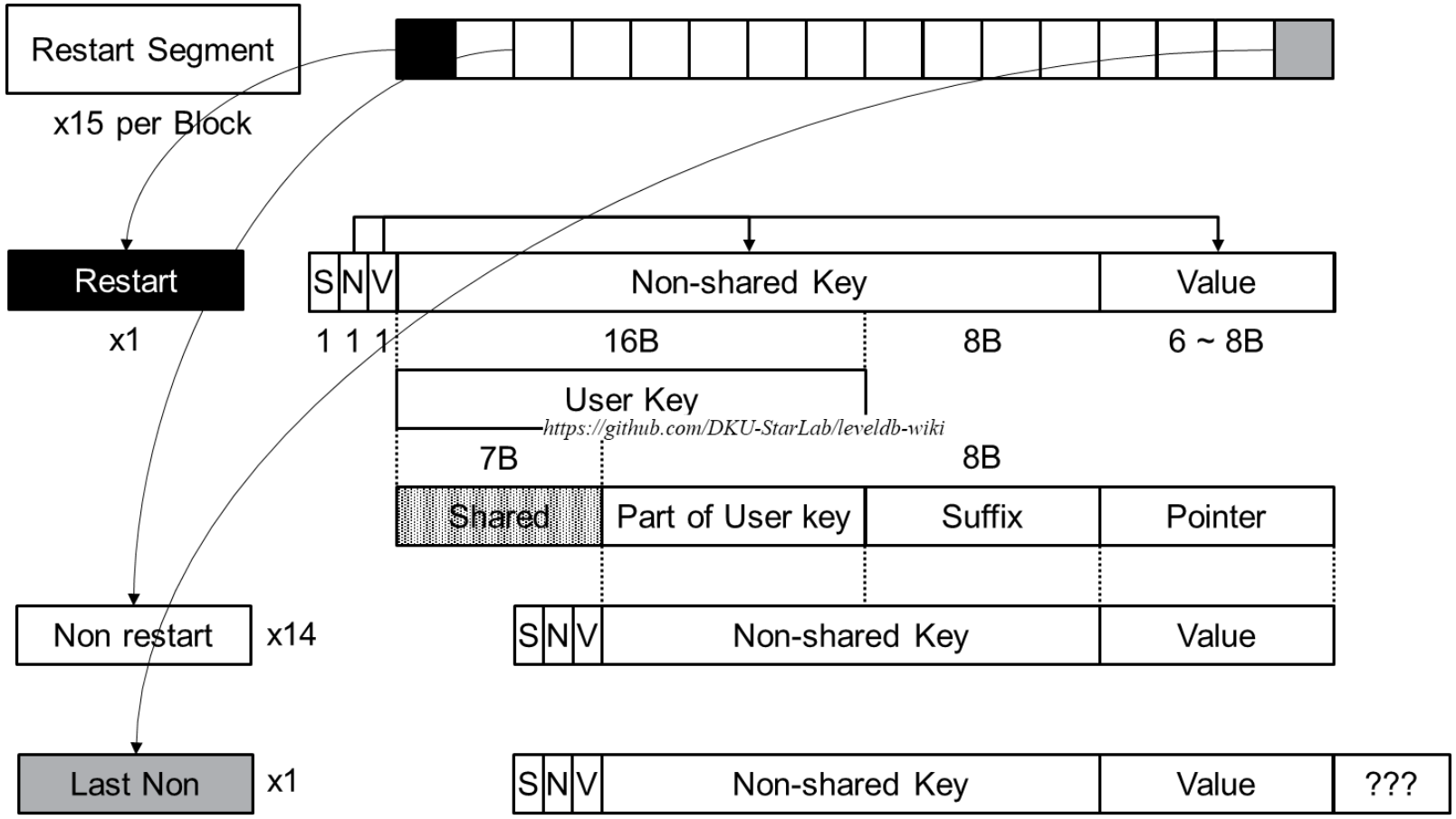


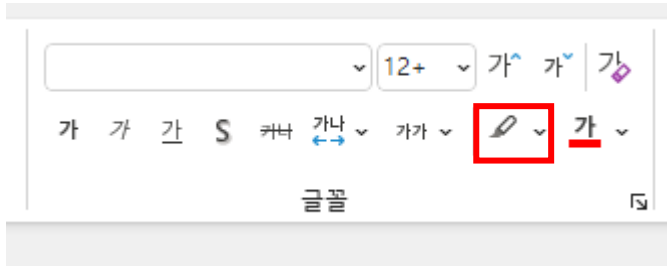
Figure 6.2. Call Graph: `alloc_pages()`

Understanding the Linux® Virtual Memory Manager Mel Gorman

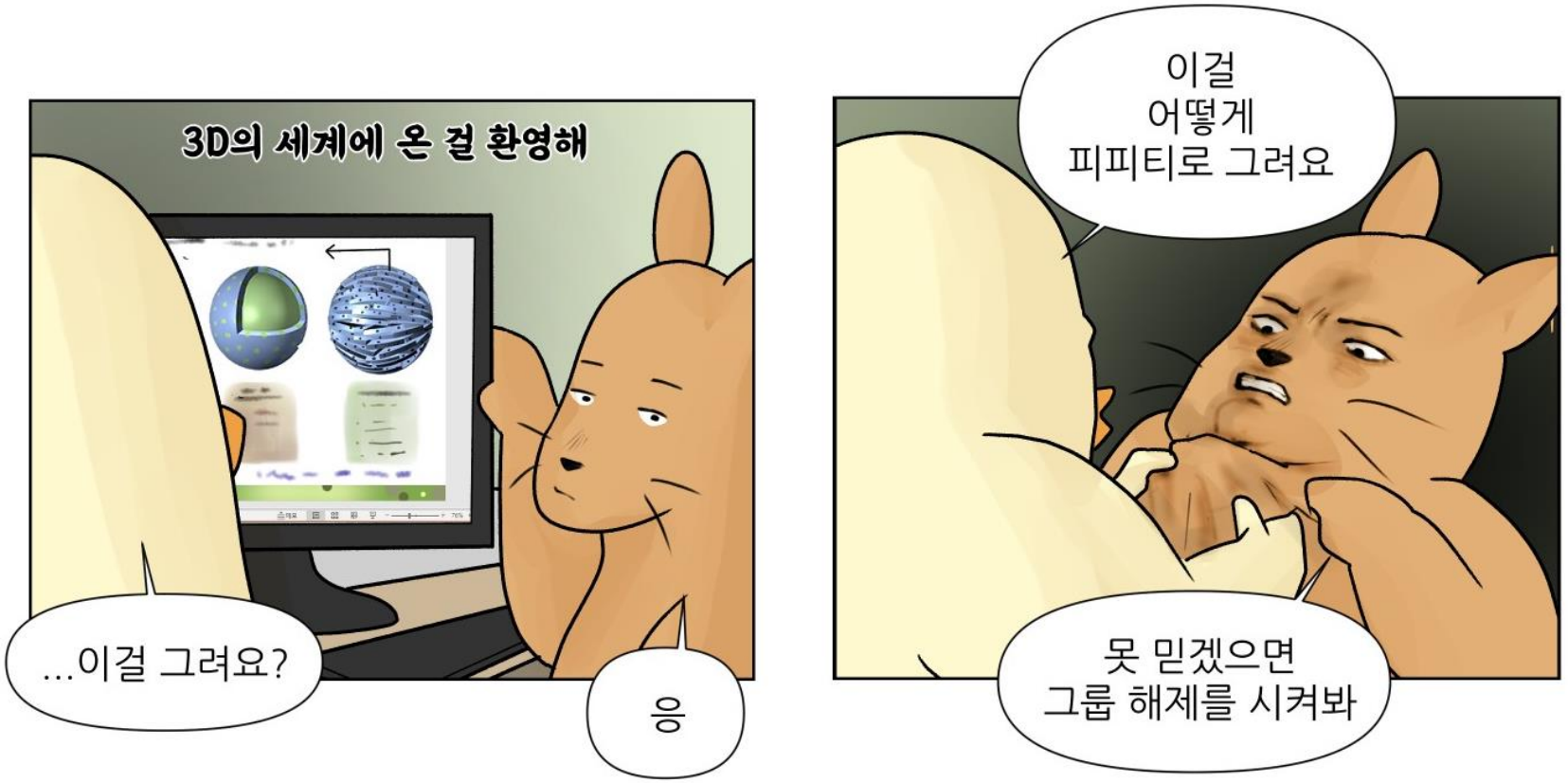
Watermark



- Watermark
  - https://github.com/DKU-StarLab/leveldb-wiki
  - *Use powerpoint or other programs/websites.*
  - *Add a background color on the watermark.*

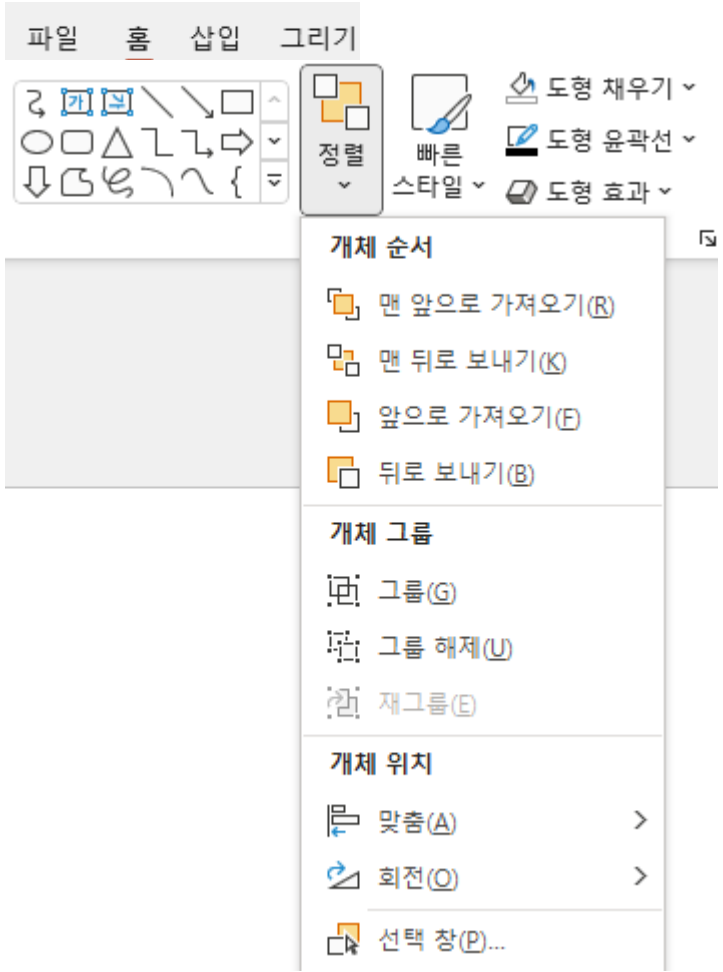


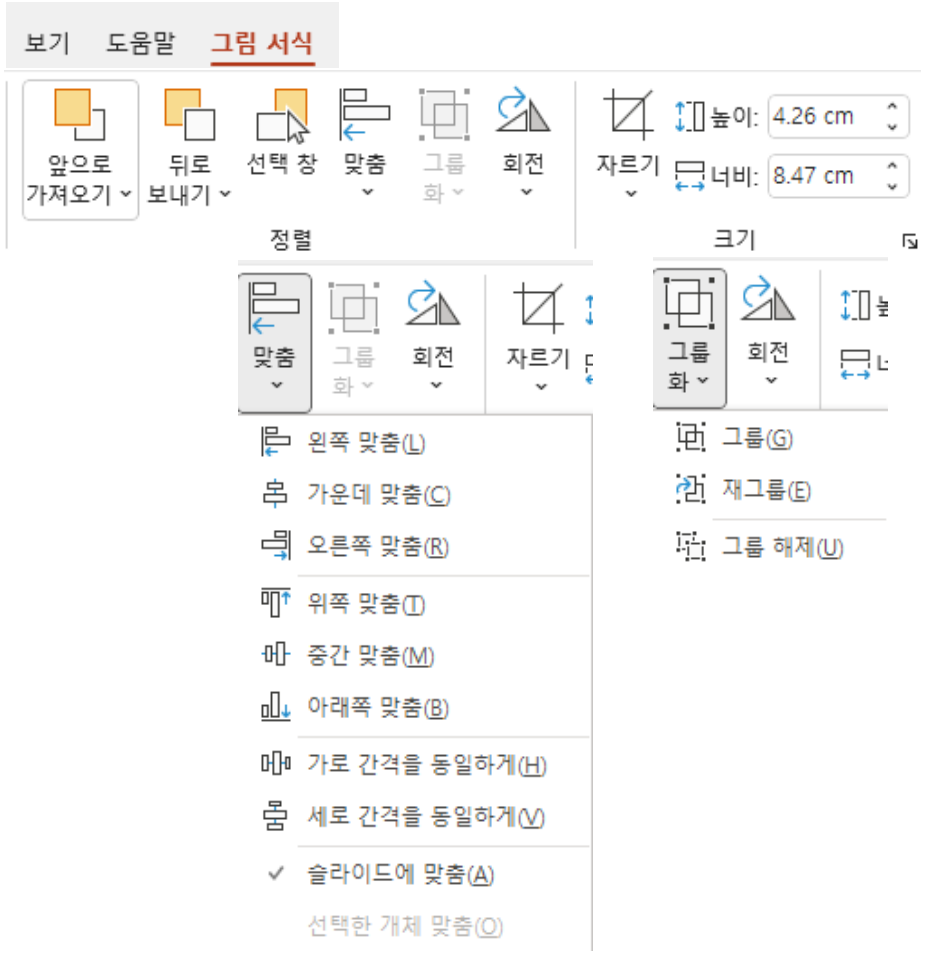
PPT



대학원 탈출일지 12화 - 피피티
https://comic.naver.com/webtoon/detail?titleId=790713&no=13&weekday=tue

PPT


파일
홈
삽입
그리기
정렬
빠른 스타일
도형 채우기
도형 윤곽선
도형 효과
개체 순서
맨 앞으로 가져오기(R)
맨 뒤로 보내기(K)
앞으로 가져오기(F)
뒤로 보내기(B)
개체 그룹
그룹(G)
그룹 해제(U)
재그룹(E)
개체 위치
맞춤(A)
회전(O)
선택 창(P)...


보기
도움말
그림 서식
앞으로 가져오기
뒤로 보내기
선택 창
맞춤
그룹화
회전
자르기
높이: 4.26 cm
너비: 8.47 cm
정렬
크기
왼쪽 맞춤(L)
가운데 맞춤(C)
오른쪽 맞춤(R)
위쪽 맞춤(T)
중간 맞춤(M)
아래쪽 맞춤(B)
가로 간격을 동일하게(H)
세로 간격을 동일하게(V)
슬라이드에 맞춤(A)
선택한 개체 맞춤(O)
그룹(G)
재그룹(E)
그룹 해제(U)

Thank you