# How Much Can Data Compressibility Help

# to Improve NAND Flash Memory Lifetime?

问题：尽管数据压缩可以提高闪存的生存期，但是很少有人认真研究利用数据可压缩性来提高内存生存期的全部潜力。

分析：在传统实践中，数据压缩的唯一目标是提高存储效率(即显式地增加一个闪存页面中可存储的数据扇区的数量)，称为显式数据压缩。由于数据可压缩性的运行时变化，显式数据压缩在闪存页面之间造成了每个页面扇区数量的异质性，这可能会使FTL和/或文件系统设计复杂化。因此，基于闪存的商业存储设备根本不使用数据压缩是很常见的。如果先进的FTL和/或文件系统(可以使用显式压缩来提高存储效率)确实可用，人们可以简单地期望存储平均压缩比为1/α的数据可以直接提高闪存寿命1/α。显式压缩，从本质上说，它忽略了两个因素。首先，闪存的损坏与所存储的内容相关。一旦数据压缩之后数据块有剩余的空闲空间，我们就可以以一种损害友好的方式操作它们的数据内容，以减少物理损害。

解决方法：使用隐式数据压缩作为显式数据压缩的补充。使用隐式数据压缩时，我们压缩每个数据扇区，但不增加每个闪存页面的数据扇区数量。隐式存储指的是通过修改MLC上下两个page中的数据分布来使得存储单元中物理损伤最小，如图1-1中所示，STATE从左到右‘Erase’->‘P3’物理损伤逐渐减小。基于闪存单元损坏的内容依赖性，提出了一组设计策略，可以利用闪存页面中未使用的存储空间来减少整体内存损坏，用于显式和隐式数据压缩。文中推导出一套数学公式，该公式可以用于定量估计通过提出的设计策略获得的闪存损伤减少。

创新点：提出了隐式的数据压缩方法，通过充分利用空闲区域来修改压缩后数据的布局来达到减少数据存储所带来的物理损伤。

思考：该隐式存储利用剩余的空闲部分来实现延长使用寿命，但我觉得文中说的一个页面的扇区数量只有一个的话就没有很好的利用数据压缩所带来的存储数据量提升这一特性。文章中隐式压缩中的两种数据布局方法都不可避免地会导致存储单元中间部分的寿命相较于头尾两端来得更低，是否有更好的数据分布方式或者通过修改存储单元的大小来降低这种寿命不均衡所带来的影响。文章中涉及到的MLC的一些相关概念和公式推导我暂时不是很理解，应当在今后进行相关知识的补充。

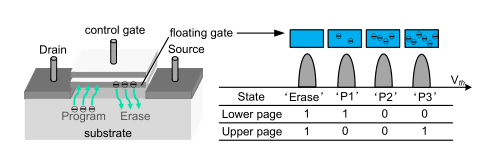


图1-1