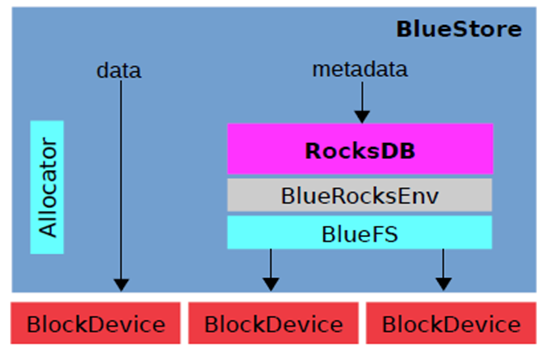
**BlueStore源码分析-架构**



1. Bluestore实现了直接管理裸设备的方式，抛弃了本地文件系统，BlockDevice实现在用户态下使用 aio直接对裸设备进行I/O操作，去除了本地文件系统的消耗，减少系统复杂度，更有利于Flash介质盘发挥性能优势；
2. 为了管理裸设备就需要一个磁盘的空间管理系统，Bluestore采用Allocator进行裸设备的空间管理，目前支持StupidAllocator和BitmapAllocator两种方式；
3. Bluestore的元数据是以KEY-VALUE的形式保存到RockDB里的，而RockDB又不能直接操作裸盘，为此，bluestore实现了一个BlueRocksEnv，继承自EnvWrapper，来为RocksDB提供底层文件系统的抽象接口支持；
4. 为了对接BlueRocksEnv，Bluestore自己实现了一个简洁的文件系统BlueFS，只实现RocksDB Env所需要的接口，在ceph系统启动挂载这个文件系统的时候将所有的元数据都加载到内存中，BluesFS的数据和日志文件都通过BlockDevice保存到底层的裸设备上；
5. BlueFS和BlueStore可以共享裸设备，也可以分别指定不同的设备，比如为了获得更好的性能Bluestore可以采用 SATA SSD 盘，BlueFS采用 NVMe SSD 盘。