

HBase知识点梳理		
行存储 VS 列存储	各自的优缺点以及弥补措施	
HBase表概念	列族 Column Family	
	行键 RowKey	
	列 Column	
	数据单元 Cell	
	数据历史版本 Version	
HBase基本操作指令	create 'tab1','cf1','cf2'	列族不易过多
	put 'tab1','row1','cf1:name','rose'	行键的设计原则
	scan 'tab1'	
	deleteall 'row1'	
	disable 'tab1' ; drop 'tab1'	
	compact('tab1')-->minor compact	
	major_compact('tab1')-->major compact	
HBase API	基本建表、插入、扫描、删除表等	
	过滤器:①行键正则过滤器 ②列值过滤器	
HBase表物理存储原理	HRegion	HStore(对应列族)
		Memstore(写缓存) LSM-TREE算法可以保证顺序写磁盘
		HFile (index+bloom) DataBlock的大小可以做调优, 小号Block有利于随机找, 大号Block利于顺序扫描
	HRegionServer	BlockCache
		WAL log
	HMaster	
HBase写流程	Client->Zookeeper->RS->WAL->Memstore->HFile->HDFS	

HBase读流程	Client->Zookeeper->Meta RS->RS->BlockCache->Memstore->HFile(index+bloom)	
HBase Compact	Minor Compact	
	Major Compact 选择周末 凌晨 服务器空闲时	
HBase 表设计		
HBase 优化		