

	2.自动扩缩容机制	4.指标类型	1.Master的kube-controller-manager服务持续监标,以计算是否需要调整副本数量 2.Pod资源使用率: Pod级别的性能指标,通常是一率 3.Pod自定义指标: Pod级别的性能指标,通常是一数量 4.Object自定义指标或外部自定义指标: 通常是一种方式提供, 1.Autoscaler控制器从聚合API获取到Pod性算出目标Pod的副本数量,与当前运行的Pod要扩缩容操作	个比率值,例如CPU使用 个数值,例如接受请求的 个数值,需要容器应用以某 能指标数据后,基于以下算法计	
		5.扩缩容算法详解	2.当前副本数*(当前指标值/期望指标值),	将结果向上取整	
				皮删除(设置了删除时间戳):将不会计入目标Pod副本数量	
			/	前指标值无法获得:本次探测不会将这个Pod纳入目标Pod副本数 探测会被重新纳入计算范围	
			· //	类型是CPU使用率,则对于正在启动但还未达到Ready状态的 时不会纳入目标副本数量范围	
			些Pod在需	缺失指标的Pod时,系统将更保守地重新计算平均值。系统会假设这需要缩容时消耗了期望指标的100%,在需要扩容时,消耗了期望指6,这样可以抑制潜在的扩缩容操作在未达到Ready状态的Pod,并且系统原本会在不考虑缺失指标或ly的Pod情况下进行扩展,则系统任然会保守地假设这些pod 消耗期的0%,从而进一步抑制扩容操作	
			3.几种Pod异常的情况 5.如果存在完 NotReady的		
			全部结果中国	6.如果在HPA中设置了多个指标,系统就会对每个指标都执行上面的算法,在 全部结果中以期望副本数的最大值为最终结果。如果这些指标中的任意一个都 无法转换为期望的副本数(例如无法获取指标的值),系统就会跳过扩缩容操 作	
			· ·	器执行扩缩容操作前,系统会记录扩缩容建议信息。控制器会在操 中考虑所有的建议信息,并从中选择得分最高的建议	
				1.scaleTargetRef:目标作用对象,可以是Deployment、 ReplicationController或ReplicaSet	
				2.targetCPUUtilizatiionPercentage: 期望每个Pod的CPU使用50%,该使用率基于Pod设置的CPU Request值进行计算,例如200m,那么系统将维持Pod的实际CPU使用值为100m	
			4.HorizontalPodAutoscaler配置详情	3.minReplicas和maxReplicas:Pod副本数量的最小值和最大值 个范围内进行自动扩缩容操作,并维持每个Pod的CPU使用率为	室为50% 5 的类型;通过参
				4.metrics:目标指示值。在metricas中通过参数type定义指标的 数target定义相应的指标目标值,系统将在指标数据达到目标值 操作	
					直,可以设置的资源为CPU和内存
				5.metrics的type类型 2.Pods:基于Pod的指标,系统	花将对全部Pod副本的指标值进行平均值计算
				3.Object:基于某种资源对象的	D指标或应用系统的任意自定义指标