命令对照

https://github.com/ClusterLabs/pacemaker/blob/master/doc/pcs-crmsh-quick-ref.md

ocf文件参考

http://www.linux-ha.org/wiki/Documentation

pacemaker官方文档

http://clusterlabs.org/pacemaker/doc/en-US/Pacemaker/1.1/html/

pacemaker-mgmt

https://github.com/ClusterLabs/pacemaker-mgmt

Linux HA中文指南

https://www.novell.com/zh-cn/documentation/sle_ha/book_sleha/data/book_sleha.html

```
# 创建一个VIP
crm configure primitive VirtualIP ocf:heartbeat:IPaddr2 param ip=172.24.10.249 cidr_netmask=32 nic=eth0 op monitor interval=30
# 实现CMS的主备(使用ocf的原因是：lsb服务的service status命令老是没有正确的返回值)
crm configure primitive cloudera-scm-server ocf:cm:rjbgserver op monitor interval=20s timeout=60s on-fail=restart op start tim
# vip位置和cms master的位置绑定
crm configure colocation cms-with-vip inf: VirtualIP  cloudera-scm-server
```

使用crm resource mv 会造成的影响

```
crm_resource -M -r VirtualIP -H master  #移动到master，并且设定偏好，分数设定为正无穷
crm_resource -B -r VirtualIP            #从当前移除，并且当前节点分数设定为负无穷
crm_resource -U -r VirtualIP            #清除Location设定
```

Resource Meta-Attributes和Resource Instance Attributes

前者是关于切换的pacemaker属性配置，后者是ocf文件中定义的配置
参考：http://clusterlabs.org/pacemaker/doc/en-US/Pacemaker/1.1/html/Pacemaker_Explained/s-resource-options.html
修改配置：如crm_resource --meta --resource cloudera-scm-server --set-parameter failure-timeout --parameter-value 20s （好像是立即生效的）
设定默认配置：
crm_attribute --type rsc_defaults --name is-managed --update false

Field	Default	Description
`priority`	0	If not all resources can be active, the cluste stop lower priority resources in order to kee priority ones active.
`target-role`	Started	What state should the cluster attempt to ke resource in? Allowed values: ○ `Stopped:` Force the resource to be stop ○ `Started:` Allow the resource to be star in the case ofmulti-state resources, pror master if appropriate) ○ `Slave:` Allow the resource to be starte only in Slave mode if the resource is mu ○ `Master:` Equivalent to `Started`
`is-managed`	TRUE	Is the cluster allowed to start and stop the r Allowed values: `true` , `false`
`resource-stickiness`	默认值为：crm configure rsc_defaults resource-stickiness=100	resource在前节点的粘性？？
`requires`	`quorum` for resources with a `class` of `stonith` , otherwise `unfencing` if unfencing is active in the cluster, otherwise `fencing` if `stonith-enabled` is true, otherwise `quorum`	Conditions under which the resource can be started *(since 1.1.8)*Allowed values: ○ `nothing:` can always be started ○ `quorum:` The cluster can only start this if a majority of the configured nodes are ○ `fencing:` The cluster can only start thi if a majority of the configured nodes are active *and* any failed or unknown nodes been fenced ○ `unfencing:` The cluster can only start t resource if a majority of the configured active *and* any failed or unknown nodes been fenced *and* only on nodes that hav been unfenced *(since 1.1.9)*
`migration-threshold`	INFINITY 默认配置时，无论server_monitor在节点失败多少次，都不会到其他节点启动？ 似乎启动失败不受这个参数影响？？？除非start-failure-is-fatal是false，这个参数默认是True（意思是只要start的失败了，这个节点就被认为不合格了）。	How many failures may occur for this resour node, before this node is marked ineligible this resource. A value of 0 indicates that thi is disabled (the node will never be marked ineligible); by constrast, the cluster treats I (the default) as a very large but finite numb option has an effect only if the failed opera on-fail=restart (the default), and additional failed start operations, if the cluster proper failure-is-fatal is false.

failure-timeout	0 失败的失效时间！默认失败永远不会失效。cluster-recheck-interval配置是轮询资源限制的时间间隔（默认15分钟），failure超期需要cluster-recheck-interval一次，因此failure-timeout最好设定的比这个参数长	How many seconds to wait before acting as failure had not occurred, and potentially all resource back to the node on which it failec of 0 indicates that this feature is disabled. any time-based actions, this is not guarante checked more frequently than the value of recheck-interval (see Section 3.2, “Cluste Options”).
multiple-active	stop_start 需要配置两个监控角色~~~OK	What should the cluster do if it ever finds tl resource active on more than one node? Allo values: ○ block: mark the resource as unmanage ○ stop_only: stop all active instances ar them that way ○ stop_start: stop all active instances a the resource in one location only
allow-migrate	TRUE for ocf:pacemaker:remote resources, FALSE otherwise	Whether the cluster should try to "live migra resource when it needs to be moved (see Section 9.4.3, “Migrating Resources”)
container-attribute-target		Specific to bundle resources; see Section 10 “Bundle Node Attributes”
remote-node		The name of the Pacemaker Remote guest r resource is associated with, if any. If specifi both enables the resource as a guest node a defines the unique name used to identify th node. The guest must be configured to run t Pacemaker Remote daemon when it is started. WARNING: This value cannot overla any resource or node IDs. *(since 1.1.9)*
remote-port	3121	If remote-node is specified, the port on the used for its Pacemaker Remote connection. Pacemaker Remote daemon on the guest mu configured to listen on this port. *(since 1.1.*
remote-addr	value of remote-node	If remote-node is specified, the IP address hostname used to connect to the guest via Pacemaker Remote. The Pacemaker Remote on the guest must be configured to accept connections on this address. *(since 1.1.9)*
remote-connect-timeout	60s	If remote-node is specified, how long befor pending guest connection will time out. *(sin*

		1.1.10)

Resource Operations

指的是在OCF文件中定义的action，除了monitor、start、stop、meta-data这三个我们还以定义其他的action！在部署资源文件时，通过op 可以让pacemaker定时调用action。
action可以有已下参数：
配置全局默认参数
crm_attribute --type op_defaults --name timeout --update 20s

Field	Default	Description
id		A unique name for the operation.
name		The action to perform. This can be any action suppor agent; common values include monitor , start , an
interval	0	How frequently (in seconds) to perform the operatio 0 means never. A positive value defines a *recurring c* is typically used with monitor.
timeout		How long to wait before declaring the action has fail
on-fail	restart *(except for stop operations, which default tofence when STONITH is enabled and* block *otherwise)*	The action to take if this action ever fails. Allowed v ○ ignore: Pretend the resource did not fail. ○ block: Don’t perform any further operations on resource. ○ stop: Stop the resource and do not start it else ○ restart: Stop the resource and start it again (p different node). ○ fence: STONITH the node on which the resource ○ standby: Move *all* resources away from the nod the resource failed.
enabled	TRUE	If false , ignore this operation definition. This is typ to pause a particular recurring monitor operation; fo can complement the respective resource being unma managed=false), as this alone will not block any con monitoring. Disabling the operation does not suppres of the given type. Allowed values: true , false .
record-pending	FALSE	If true , the intention to perform the operation is re that GUIs and CLI tools can indicate that an operatio progress. This is best set as an *operation default* (se section). Allowed values: true , false .
role		Run the operation only on node(s) that the cluster th be in the specified role. This only makes sense for re monitor operations. Allowed (case-sensitive) values: Stopped , Started , and in the case of mult state resources, Slave and Master .

Cluster Options

Option	Default	Description
`dc-version`		Version of Pacemaker on the cluster's DC. Determined automatically cluster. Often includes the hash which identifies the exact Git chang built from. Used for diagnostic purposes.
`cluster-infrastructure`		The messaging stack on which Pacemaker is currently running. Deter automatically by the cluster. Used for informational and diagnostic p
`expected-quorum-votes`		The number of nodes expected to be in the cluster. Determined auto the cluster. Used to calculate quorum in clusters that use Corosync 1 CMAN as the messaging layer.
`no-quorum-policy`	stop	What to do when the cluster does not have quorum. Allowed values: ○ `ignore:` continue all resource management ○ `freeze:` continue resource management, but don't recover reso nodes not in the affected partition ○ `stop:` stop all resources in the affected cluster partition ○ `suicide:` fence all nodes in the affected cluster partition
`batch-limit`	0 *(30 before version 1.1.11)*	The maximum number of actions that the cluster may execute in par all nodes. The "correct" value will depend on the speed and load of y and cluster nodes. If zero, the cluster will impose a dynamically calc only when any node has high load.
`migration-limit`	-1	The number of migration jobs that the TE is allowed to execute in pa node. A value of -1 means unlimited.
`symmetric-cluster`	TRUE	Can all resources run on any node by default?
`stop-all-resources`	FALSE	Should the cluster stop all resources?
`stop-orphan-resources`	TRUE	Should deleted resources be stopped? This value takes precedence o `managed` (i.e. even unmanaged resources will be stopped if deleted configuration when this value is TRUE).
`stop-orphan-actions`	TRUE	Should deleted actions be cancelled?
`start-failure-is-fatal`	TRUE	Should a failure to start a resource on a particular node prevent furt attempts on that node? If FALSE, the cluster will decide whether the still eligible based on the resource's current failure count and `migra threshold` (see Section 9.3, "Handling Resource Failure").
`enable-startup-probes`	TRUE	Should the cluster check for active resources during startup?
`maintenance-mode`	FALSE	Should the cluster refrain from monitoring, starting and stopping res
`stonith-enabled`	TRUE	Should failed nodes and nodes with resources that can't be stopped value your data, set up a STONITH device and enable this.

		If true, or unset, the cluster will refuse to start resources unless one STONITH resources have been configured. If false, unresponsive node immediately assumed to be running no resources, and resource take nodes starts without any further protection (which means *data loss* i unresponsive node still accesses shared storage, for example). See al the `requires` meta-attribute in Section 5.4, “Resource Options”.
`stonith-action`	reboot	Action to send to STONITH device. Allowed values are `reboot` and value `poweroff` is also allowed, but is only used for legacy devices.
`stonith-timeout`	60s	How long to wait for STONITH actions (reboot, on, off) to complete
`stonith-max-attempts`	10	How many times fencing can fail for a target before the cluster will immediately re-attempt it. *(since 1.1.17)*
`concurrent-fencing`	FALSE	Is the cluster allowed to initiate multiple fence actions concurrently *1.1.15)*
`cluster-delay`	60s	Estimated maximum round-trip delay over the network (excluding ac execution). If the TE requires an action to be executed on another n consider the action failed if it does not get a response from the othe time (after considering the action’s own timeout). The "correct" valu on the speed and load of your network and cluster nodes.
`dc-deadtime`	20s	How long to wait for a response from other nodes during startup. The "correct" value will depend on the speed/load of your network a of switches used.
`cluster-recheck-interval`	15min	Polling interval for time-based changes to options, resource paramet constraints. The Cluster is primarily event-driven, but your configuration can hav that take effect based on the time of day. To ensure these changes t we can optionally poll the cluster’s status for changes. A value of 0 d polling. Positive values are an interval (in seconds unless other SI uni specified, e.g. 5min).
`cluster-ipc-limit`	500	The maximum IPC message backlog before one cluster daemon will d another. This is of use in large clusters, for which a good value is the resources in the cluster multiplied by the number of nodes. The defa also the minimum. Raise this if you see "Evicting client" messages for daemon PIDs in the logs.
`pe-error-series-max`	-1	The number of PE inputs resulting in ERRORs to save. Used when rep problems. A value of -1 means unlimited (report all).
`pe-warn-series-max`	-1	The number of PE inputs resulting in WARNINGs to save. Used when r problems. A value of -1 means unlimited (report all).
`pe-input-series-max`	-1	The number of "normal" PE inputs to save. Used when reporting prob of -1 means unlimited (report all).

`placement-strategy`	default	How the cluster should allocate resources to nodes (see Chapter 12, *and Placement Strategy*). Allowed values are `default`, `utilization`, `balanced`, and `minimal`. *(since 1.1.0*
`node-health-strategy`	none	How the cluster should react to node health attributes (see Section 9 Node Health”). Allowed values are `none`, `migrate-on-red`, `only-green`, `progressive`, and `custom`.
`node-health-base`	0	The base health score assigned to a node. Only used when `node-hea strategy` is `progressive`. *(since 1.1.16)*
`node-health-green`	0	The score to use for a node health attribute whose value is `green`. (when `node-health-strategy` is `progressive` or `custom`.
`node-health-yellow`	0	The score to use for a node health attribute whose value is `yellow`. when `node-health-strategy` is `progressive` or `custom`.
`node-health-red`	0	The score to use for a node health attribute whose value is `red`. On when `node-health-strategy` is `progressive` or `custom`.
`remove-after-stop`	FALSE	*Advanced Use Only:* Should the cluster remove resources from the LR are stopped? Values other than the default are, at best, poorly teste potentially dangerous.
`startup-fencing`	TRUE	*Advanced Use Only:* Should the cluster shoot unseen nodes? Not using is very unsafe!
`election-timeout`	2min	*Advanced Use Only:* If you need to adjust this value, it probably indi presence of a bug.
`shutdown-escalation`	20min	*Advanced Use Only:* If you need to adjust this value, it probably indi presence of a bug.
`crmd-integration-timeout`	3min	*Advanced Use Only:* If you need to adjust this value, it probably indi presence of a bug.
`crmd-finalization-timeout`	30min	*Advanced Use Only:* If you need to adjust this value, it probably indi presence of a bug.
`crmd-transition-delay`	0s	*Advanced Use Only:* Delay cluster recovery for the configured interva additional/related events to occur. Useful if your configuration is se order in which ping updates arrive. Enabling this option will slow dov recovery under all conditions.
`default-resource-stickiness`	0	*Deprecated:* See Section 5.4.2, “Setting Global Defaults for Resource Attributes” instead
`is-managed-default`	TRUE	*Deprecated:* See Section 5.4.2, “Setting Global Defaults for Resource Attributes” instead
`default-action-timeout`	20s	*Deprecated:* See Section 5.5.3, “Setting Global Defaults for Operatio