kubernetes #55262 issue #50916

commit a366e6ced0943a46d517f71e45bbd71c22239dc4

问题：pod在assume和add时使用了不同的名字，导致update的时候发生fatal错误

复现：无

原因：scheduler在bind pod的时候，为了效率，在watch到pod bind node成功之前就开始下一轮调度，此时会assume该pod已经bind成功。当assumed pod真的bind成功时，就会正式add到cache中，如果assume一直没有add成功，就会在一定时间后forget/expired。

在以前的实现中，add一个assume pod时不会更新该assume pod的nodeName（仍然使用assume pod的nodeName），在下一次update时，如果发现update的pod信息不一致，就会造成scheduler退出。

add pod和assume pod的nodeName产生不同会发生在podStates没有及时更新时。（It was just the podStates' version of the pod that wasn't getting updated. ）

```
   +-------------------------------------------+  +-----+
   |                            Add            |  |     |
   |                                           |  |     | Update
   +      Assume                Add            v  v     |
Initial +--------> Assumed +------------+---> Added <---+
   ^                +   +               |       +
   |                |   |               |       |
   |                |   |           Add |       | Remove
   |                |   |               |       |
   |                |   |               +       |
   +----------------+   +----------> Expired    +----> Deleted
```

```
currSfunc (cache *schedulerCache) UpdatePod(oldPod, newPod *v1.Pod) error {
      ...
      switch {
      // An assumed pod won't have Update/Remove event. It needs to have Add event
      // before Update event, in which case the state would change from Assumed to Added.
      case ok && !cache.assumedPods[key]:
            if currState.pod.Spec.NodeName != newPod.Spec.NodeName {
                  glog.Errorf("Pod %v updated on a different node than previously added to.", key)
                  glog.Fatalf("Schedulercache is corrupted and can badly affect scheduling decisions")
            }
            if err := cache.updatePod(oldPod, newPod); err != nil {
                  return err
            }
```

修复：

```
diff --git a/plugin/pkg/scheduler/schedulercache/cache.go
b/plugin/pkg/scheduler/schedulercache/cache.go
index e37ef233a6..df7ac6601a 100644
--- a/plugin/pkg/scheduler/schedulercache/cache.go
+++ b/plugin/pkg/scheduler/schedulercache/cache.go
@@ -241,6 +241,7 @@ func (cache *schedulerCache) AddPod(pod *v1.Pod)
error {
                }
                delete(cache.assumedPods, key)
                cache.podStates[key].deadline = nil
+               cache.podStates[key].pod = pod
        case !ok:
                // Pod was expired. We should add it back.
                cache.addPod(pod)
```