

## Глава 8.

### Сопровождение кластерных систем

---

Кластер построен, установлен и настроен. Как поддерживать его работоспособность и как решать возникающие вопросы? **Необходим квалифицированный обслуживающий персонал.** Нужно сразу отбросить иллюзии и четко осознать, что если нет администратора, отвечающего за работу кластера, то нет и самого кластера. Никакой серьезной работы на такой системе выполнить будет нельзя. В нашей практике был пример, когда купленный кластер проработал месяц, но из-за отсутствия администратора и постоянных споров между пользователями был попросту выключен: кто не мог, а кто не хотел брать на себя ответственность за его работоспособность. В таком состоянии кластер простоял два года. Естественно возникает вопрос, зачем покупали или же почему потом не продали? Не все поддается разумному объяснению...

В зависимости от масштабности проекта в состав обслуживающего персонала могут входить следующие категории специалистов:

- системный администратор, отвечающий за настройки кластера, обеспечение его безопасности и текущую работу с пользователями,
- инженер по сопровождению силовой электрики и систем климат-контроля,
- специалист по системному программному обеспечению,
- консультант, способный решать проблемы с прикладным программным обеспечением и помогающий пользователям в работе на кластерной системе.

Желательно иметь штатных операторов, которые могут оперативно решать текущие вопросы. Естественно, что все специалисты

должны иметь соответствующую квалификацию. Персонал в целом должен быстро решать задачи, связанные с безопасностью и администрированием, управлением пользователями и программным обеспечением, разбираться с возникающими аппаратными проблемами, связываться с поставщиками и сервисными службами.

Операторы и администраторы должны быть хорошо подготовлены, и **процесс начального обучения персонала необходимо закончить до того, как кластер будет введен в эксплуатацию.** Минимальный уровень знаний администратора должен включать знание всех используемых на кластере программных пакетов и хорошее понимание их взаимосвязей. Это не предполагает знание тонкостей программирования в математических, физических, химических и других специализированных прикладных комплексах, но предполагает понимание механизмов их установки и настройки.

Кроме подготовки администраторов, **необходима подготовка пользователей.** Пользователи кластера должны хорошо уметь работать с командной строкой UNIX, знать основные технологические приемы, уметь пользоваться компиляторами и программами для удаленного доступа, знать как передать и получить файлы с кластера, уметь запускать параллельные программы, владеть азами параллельных вычислений и разбираться в технологиях параллельного программирования. Признаком не только хорошего тона, но и показателем отношения к пользователям является организация on-line ресурса, посвященного организации работ на кластере. Такой ресурс поможет не только новичку быстро войти в курс дела, он будет полезен и опытному человеку, ранее работавшему в другом окружении. Регламент работы системы в целом, состав программного обеспечения, описание политик и квот в различных разделах кластера, новости, примеры программ, документация, учебные материалы, советы и описание типичных ситуаций, вопросы и ответы – все это, как масса другой полезной информации может быть размещено на подобных страницах.

Говоря о персонале и пользователях, нельзя не затронуть вопрос их взаимоотношений. Сложный вопрос. В некоторых случаях администраторы вводят жесткие меры, заставляя пользователей выполнять чрезмерно строгие правила, ходить на цыпочках и записываться к ним на прием заранее в случае возникновения каких-либо вопросов. В других местах ситуация диаметрально противоположная: пользователи исполнены ощущением собственной значимости, считая невозможным или ненужным общение с администраторами в силу своей высокой миссии. И в том, и в другом случае хороший результат маловероятен. Выяснить кто главнее: администратор или пользователь, столь же бессмысленно, как пытаться отдать первенство пилоту или болиду в гонках Формулы-1. Только вместе, помогая и подсказывая друг другу.

Оставим пафос и вернемся к технической стороне дела. Для повышения эффективности работы системного администратора очень рекомендуем установить пакеты для **мониторинга состояния кластера**. Важно не только уметь получать данные о текущем состоянии кластера и истории его функционирования, но и оперативно реагировать на возникающие нештатные ситуации.

Не все пакеты одинаково хороши. Например, пакет Ganglia (<http://ganglia.sourceforge.net/>) позволяет получать наглядную информацию об истории состояния кластера, но абсолютно не способен ее анализировать.

Пакет Nagios (<http://nagios.org/>) ориентирован на мониторинг доступности компьютеров и сервисов. Он может быть весьма полезен для отслеживания работоспособности кластера и оповещения в случае серьезных сбоев. Однако, он не способен отслеживать оперативную картину загрузки и использования кластера без серьезных дополнительных нагрузок на систему.

Постоянно следите за состоянием жестких дисков и вентиляторов, которые по статистике чаще всего выходят из строя. В этом сильно могут помочь пакеты `smartd` и `lm_sensors`.

Чтобы понимать, насколько эффективно используются ресурсы кластера, не пренебрегайте статистикой. В частности, в пакет Cleo входит программа `cleo-stat`, позволяющая получить информацию об использовании ресурсов кластера как в целом, так и с детализацией по пользователям за любой период времени.

Чтобы не пришлось часто разбираться с множеством мелких проблем на вычислительных узлах, постарайтесь минимизировать эти проблемы заранее. В этом может помочь пакет `logcheck`, а также периодическая (например, с помощью утилиты `cron`) очистка временных каталогов. В частности, команда

```
find /tmp -mtime +3 -exec rm -f \{\} \;
```

удалит из `/tmp` все файлы и каталоги, которые не изменялись за последние трое суток.

Важная деталь, о которой нередко забывают в начале, а потом регулярно откладывают на “завтра” – это **резервное копирование**. О важности сохранения данных говорить излишне. Надеяться на аппаратный RAID можно, но это не панацея. Из-за нелепой ошибки или программного сбоя можно потерять важные данные на логическом уровне, т.е. на уровне файловой системы. Тут никакой RAID не спасет.

Эффективная организация резервного копирования подробно описана во многих специальных изданиях по администрированию, обсуждать этот вопрос в деталях здесь мы не будем. Можно использовать архиваторы `tar` или `cpio`. Или даже `zip` или `rar`, однако имейте в виду, что эти утилиты не сохраняют UNIX-атрибуты файлов, поэтому восстанавливать данные из этих архивов будет сложнее. Можно воспользоваться многочисленными системами для автоматизации резервного копирования, такими как `Vacula`, `Amanda`, `Backupinja`, `Cdbackup`, `af-backup`, `ibackup`, `Tarup` и многими другими.

Производить резервное копирование можно как вручную, так и в автоматическом режиме. Если ведется инкрементальное копирование, при котором сохраняются лишь изменения со времени последнего

копирования, то хотя бы раз в две недели надо делать полный сброс информации. В противном случае, на восстановление будет уходить очень много времени, да и надежность копирования снизится.

Ни в коем случае не делайте резервной копии на том же физическом или логическом диске, что и оригинал. Лучший вариант – это делать ее на другом компьютере либо воспользоваться для копирования ленточным устройством или пишущим приводом CD/DVD. Хранить только одну резервную копию тоже плохо: если были потеряны данные, и такое состояние попало на резервную копию, то восстановить пропавшие файлы не получится.

Часто принимают схему, при которой полные резервные копии делаются каждые одну-две недели (лучше всего во время выходных дней, когда пользователи работают не так активно), а инкрементальные выполняются ежедневно. Если нет специальных требований, то хранить резервные копии имеет смысл в течение периода от месяца до полугода. Вариантов организации резервного копирования и в самом деле существует много, но самое главное – помнить о его необходимости и поставить этот процесс на регулярную основу.

Постарайтесь максимально автоматизировать все процессы и процедуры, связанные с сопровождением и администрированием кластера. Кластерное хозяйство является сложным и динамичным, важных вопросов, требующих личного участия администраторов будет много, поэтому все то, что может выполняться в автоматическом или автоматизированном режиме, так и должно быть настроено. Это в полной мере касается только что обсуждавшихся систем мониторинга и резервного копирования, за которыми должен быть оставлен лишь минимальный “ручной” контроль, позволяющий быть уверенным в их нормальном функционировании. Грамотное использование подобных технологий позволяет повысить надежность и эффективность работы кластера, и одновременно, что также немаловажно, понизить такой важный показатель, как стоимость владения кластерной системой.

Если речь идет о долгосрочной эксплуатации кластерной системы, а на практике так чаще всего и бывает, то нужно позаботиться о выделении специальной строки в бюджете организации. Какие-то расходы незаметны, например, замена одного локального диска на узле, однако в целом текущие расходы могут составить весьма значительную сумму. Оплата электроэнергии и аренды помещения, продление лицензий на компиляторы, прикладные пакеты, расширение набора используемого программного обеспечения, оплата сервисного обслуживания кластера, систем электропитания и охлаждения, замена вышедшего из строя оборудования – все подобные детали должны приниматься в расчет, чтобы организовать эффективную текущую эксплуатацию кластера и минимизировать перерывы в его работе.

Чтобы пользователи могли планировать свою деятельность, заранее определите и объявите **план проведения профилактических работ**. Именно в это время имеет смысл вносить изменения в настройки работающих программ, устанавливать новые или обновлять уже существующие пакеты, выполнять какие-то действия по резервному копированию, установке нового оборудования. Ко времени профилактики разумно привязать и многие обязательные регламентные работы, например, очистку фильтров систем кондиционирования или проведение контрольного тестирования компонентов системы. Как правило, многие текущие дела по сопровождению кластера можно немного отложить и затем спокойно выполнить во время профилактики: это и администраторам проще, и пользователям спокойнее.

В целом же, чем четче, понятнее и прозрачнее будет организована работа на вычислительной системе, тем дольше и эффективнее она будет использоваться.