

Глава 2.

Базовая инфраструктура будущего проекта

Если принято решение размещать кластер на чужой площадке или воспользоваться уже работающей кластерной системой, то этот раздел можно пропустить. В данном разделе мы обсудим основные требования к инженерной инфраструктуре кластера: помещению и его оборудованию, системам охлаждения, электропитания и другим не менее важным ее составляющим.

Итак, решено установить кластер на своей территории. В этом случае потребуется продумать и составить систему подготовительных мероприятий, а также выполнить расчеты необходимых параметров инфраструктуры. Если нужных данных в распоряжении не оказалось, от расчетов не отказывайтесь, а просто отложите их на некоторое время. Не исключено, что к расчетам придется вернуться не один раз, если не удастся сразу, например, подобрать помещение с достаточной площадью или же обеспечить подводку электропитания нужной мощности.

Настоятельно рекомендуем **выделить для кластера отдельное помещение**, изолированное от рабочих мест пользователей. В нашей практике не было ни одного случая, когда размещение кластера в одной комнате с людьми было бы действительно обоснованным и оправданным. Причин этому несколько.

- Даже кластер начального уровня из 2-3 узлов может производить шум, при котором трудно разговаривать, не повышая голоса. Например, уровень шума сервера HP DL145 (1U, 2 процессора AMD Opteron) составляет 63,2 децибел, что сравнимо с шумом хорошо заполненного ресторана. Если поставить 10 таких серверов, то уровень шума будет около 70 децибел, что сопоставимо с нахождением в пяти метрах от оживленной автотрассы.

- Скорее всего, для кластера потребуется кондиционирование, а его правильно организовать и контролировать можно только в отдельном помещении.
- Для своего охлаждения вычислительные узлы протягивают сквозь корпус внушительный поток воздуха, а вместе с ним – пыль и грязь. Отдельное помещение, где не будет открываться окно, а доступ посторонних будет ограничен, гарантирует, что количество пыли, попадающее в узлы, будет минимальным.
- Физический доступ к кластеру желательно ограничить небольшим кругом лиц, подготовленных и сертифицированных соответствующим образом. Имея отдельное помещение, сделать это будет намного проще.

Заранее подумайте о том, **какое место в помещении займет кластер**. Не забудьте о расстояниях до стен и вентиляции, возможности подойти к стойкам со всех сторон, открыть их двери, о рациональном подводе электропитания.

Чем больше **объем помещения**, тем лучше условия для нормального функционирования кластера, поскольку риск резкого повышения температуры из-за отключения систем кондиционирования меньше. Если есть мониторинг температуры, то в случае возникновения аварийной ситуации будет достаточно времени для автоматического выключения кластера до того момента, пока температура в помещении поднимется выше критической отметки. С этой точки зрения, при прочих равных условиях под кластер лучше выделить комнату с высокими потолками.

Для кластера потребуется оперативный набор запасных частей; к узлам, к системному и прикладному программному обеспечению будут прилагаться многочисленные диски, документация и гарантийные листы – все это необходимо где-то хранить. Ящика стола и даже одной полки в шкафу, скорее всего, будет мало, поэтому желательно предусмотреть **специальное место для хранения вспомогательных материалов**.

Как и любой электроприбор, кластер и связанная с ним аппаратура выделяют тепло, поэтому проведите расчет мощности необходимой **системы охлаждения воздуха**. Не стоит опираться на ощущения, прикидывать “на глаз” или надеяться, что солидного по внешнему виду кондиционера хватит на создаваемый комплекс. Неверный расчет параметров температурного режима в помещении может доставить большие неприятности в ходе эксплуатации кластерной системы.

Посмотрите на web-сайте фирмы-производителя характеристики вычислительных узлов, которые планируется установить. Найдите мощность тепловыделения. Если не удалось найти эту характеристику, найдите мощность блока питания в ваттах. Если в ваттах нет, то возьмите мощность в вольт-амперах. Полученное число можно принять за тепловую мощность (реально она будет чуть ниже). Ту же процедуру проведите со всей периферией кластера: управляющим узлом, файл-сервером, сетевым оборудованием и другими компонентами.

Рассмотрим пример. Потребляемая мощность упоминавшегося выше сервера HP DL145 составляет 500 Вт. Не забудьте о сетевых коммутаторах, поскольку они потребляют не так уж мало, как может показаться на первый взгляд. Тепловыделение коммутаторов 3Com SuperStack 3 Switch серии 4400 может составлять от 100 до 275 Вт. В результате, кластер из десяти серверов DL145 и двух коммутаторов 3Com потенциально может выделять до 5,55 кВт тепла.

Потребуется кондиционер, мощность охлаждения которого (но не потребляемая мощность, не путайте эти два понятия!) будет не меньше суммарной мощности тепловыделения всех узлов и всего набора периферии кластера. Обратите внимание, что сам кондиционер тоже может выделять тепло.

Кроме оборудования, необходимого для работы кластера, помните о других источниках тепла. Первый – это солнце. Если в помещении окна выходят на южную сторону, и на них нет ни отражающей пленки, ни штор или жалюзи, то комната будет нагреваться одним только солнцем до

внутренней температуры.

Второй источник тепла – это центральное отопление. В некоторых случаях проще вовсе убрать из комнаты батареи. Это не всегда реально, поэтому имеет смысл предусмотреть возможность явного отключения батарей в комнате. Помните, что при отключении батарей нужно предусмотреть слив оставшейся в них воды. Еще одна потенциальная опасность центрального отопления состоит в губительном для аппаратуры **увеличении влажности** при возникновении протечек соединений или разрывов труб. Установите **систему мониторинга с набором датчиков**, которая будет следить за температурой и влажностью в помещении. В случае необходимости она оповестит по электронной почте или мобильной связи о возникновении нештатной ситуации либо автоматически выключит оборудование.

Предупреждая почти уже традиционное предложение, хотим подчеркнуть сразу: **охлаждение “открытым окном” равносильно медленной казни кластерного оборудования**. Этого нельзя делать по многим причинам. Во-первых, как правило, этого явно недостаточно и не стоит сравнивать кластер с рабочими машинами в офисе, поскольку они работают в абсолютно разных режимах. Во-вторых, контроль за температурой в этом режиме просто невозможен. В-третьих, пыль и грязь с улицы за короткий промежуток времени приведут вычислительные узлы в полную негодность.

Немаловажен **выбор места установки кондиционера**. Ни в коем случае нельзя ставить кондиционер в непосредственной близости от стойки кластера. Лучшее место – на потолке или у потолка на расстоянии 3-4 или более метров от стоек (рис. 2.1).



Рис. 2.1. Пример установки кондиционеров

Заметим, что эффективность воздушного охлаждения сильно зависит от высоты помещения над уровнем моря. Например, национальная лаборатория в Лос Аламосе (США) расположена на высоте более двух километров над уровнем моря, и эффективность охлаждения воздухом там примерно в 2 раза ниже по сравнению с охлаждением на нулевой отметке.

Не забывайте о том, что любой кондиционер выделяет конденсат, который нужно будет куда-то отводить. Если внешний блок кондиционера установлен на первом этаже во дворе здания, то это не проблема. А если на 5-м этаже и над парадным входом, то придется обдумать этот вопрос специально.

Предмет особого внимания – это работа кондиционера зимой. Если мы сами, как правило, зимой им не пользуемся, то на кластер смена времен года, в этом смысле, почти никак не влияет. Основное требование состоит в том, чтобы кондиционер был рассчитан на работу зимой. Учтите, что на холоде выделяемый конденсат может замерзнуть, делая кондиционер неработоспособным, поэтому отвод должен быть спроектирован так, чтобы подобная ситуация не возникала.

Кроме традиционных кондиционеров можно воспользоваться еще

одним решением – водяным стоечным охлаждением. Такое решение предлагает, например, компания APC (<http://www.apc.com/>): система охлаждения ARAC15000U. Это полностью готовая стойка с комплектом водяного охлаждения мощностью до 18 кВт, которую нужно только правильно подключить.

Подчеркнем еще раз, что правильно организованное охлаждение крайне важно. За считанные минуты температура в помещении с работающим кластером даже менее чем из десятка узлов может подняться до 50-60 градусов, при этом температура внутри системных блоков будет 90 градусов и выше. В этих условиях оборудование может попросту перегореть, и по гарантии его в таком случае не заменят.

Следующим необходимым для аккуратного расчета параметром является **электропитание**. Подсчитайте мощность, потребляемую всеми узлами кластера и сопутствующими компонентами проекта (источниками бесперебойного питания, файловым хранилищем, коммутаторами и т.п.). Проконсультируйтесь с главным инженером, как обеспечить соответствующее электропитание размещаемого оборудования. Скорее всего, придется сделать дополнительную проводку. Если будете устанавливать мощные источники бесперебойного питания, то для них потребуются специальная проводка, так как в обычные розетки они не включаются. Проверьте характеристики устанавливаемого оборудования: оборудование ставится мощное, стандартные и привычные решения могут уже и не работать. Например, обычная электрическая розетка рассчитана примерно на 16 ампер, т.е. примерно на 3,5 кВт, больше оборудования на нее вешать нельзя.



Рис. 2.2. Короб для укладки кабелей с открытой крышкой

Обратите внимание, что, меняя компоновку вычислительных узлов, их архитектуру, состав, степень отказоустойчивости, производителей, можно изменить энергопотребление всей кластерной установки в несколько раз. Стоит обратить особое внимание на новые модели процессорных платформ, которые при большей производительности по сравнению с существующими вариантами, как правило, имеют пониженное энергопотребление. Например, стоечный сервер C1000 от компании Rackable Systems, представленный в формате 1 U на базе двух четырехядерных процессоров Intel Xeon E5345 с частотой 2,33 ГГц, потребляет всего 290 Вт при пиковой производительности 74,5 Гфлопс. В последнее время характеристика “производительность на ватт” многими принимается в расчет, и свобода выбора существует большая.

Важна не только подводимая мощность электропитания, но также его надежность и гарантированное качество, особенно для систем, работающих в круглосуточном режиме. Во многих случаях все диктуется существующими возможностями здания и его инженерных коммуникаций. Очень хороший вариант – это обеспечить поступление питания к кластеру от разных электроподстанций. Установите силовую автоматику, которая обеспечит автоматическое переключение с одной подстанции на другую при возникновении аварийной ситуации. Подключите кластер через источники бесперебойного питания (UPS). Они сгладят незначительные

скачки напряжения, обеспечат стабильную работу вычислительной системы при кратковременных перебоях питания и позволят кластеру нормально завершить работу при отсутствии питания в течение длительного времени.

Ставить или нет источники бесперебойного питания, каждый решает для себя сам. Это сильно зависит от места установки, конкретных условий эксплуатации кластера, предъявляемых к нему требований. Расскажем лишь один пример из нашей практики. Организация поставила у себя вычислительную систему, но от UPS отказались, аргументируя высокой стабильностью питания. Какое-то время работали без проблем, но через некоторое время система стала останавливаться по непонятным причинам. Начали разбираться и сразу обратили внимание на то, что аварийные остановки происходят приблизительно в одно и то же время. Выяснили причину остановок – сильные скачки напряжения. Провели расследование происходящего и поразились нелепости возникшей ситуации. В организации открыли столовую, в которой мощные электроплиты для приготовления пищи оказались в той же сети, что и вычислительная система. Утром плиты включали, а вечером выключали приблизительно в одно и то же время, отсюда и регулярность бросков напряжения. Пока выясняли источник проблем, несколько недель использовать систему было невозможно.

Тщательно продумайте способ прокладки кабелей и их подведения к розеткам, UPS и другому оборудованию (рис. 2.2). Как правило, обычного розеточного удлинителя для питания кластера недостаточно. Пучки подводимых кабелей будут весьма внушительными, без должной укладки провода будут путаться, спотыкаться об них будет и неприятно, и опасно. Вопрос внешнего вида этого клубка оставим в стороне и обсуждать не будем. Обратите внимание на сечение используемой проводки, площадь которого зависит от мощности проектируемой кластерной системы.

Если помещение позволяет, то **рекомендуем использовать фальшпол** для подведения электропитания. Кроме фальшпола есть

вариант использования навесных корзин для проводки, но выглядят они не так аккуратно. Если ставите фальшполы, посоветуйтесь с инженерами относительно их прочности, поскольку мощные кластерные системы могут иметь значительный вес. В ноябре 2006 года для проектирования болидов Формулы-1 компания BMW установила кластер из 10 стоек стандартного формата, вес которого с полным комплектом источников бесперебойного питания составил 21 тонну. Вес кластерной системы СКИФ Cyberia со всем набором необходимого силового электрооборудования и устройств климат-контроля составляет 16 тонн.

Учтите, что **не стоит прокладывать рядом коммуникационную и силовую проводку**, поскольку циркулирующие по силовым кабелям токи могут создавать значительные наводки, приводя к ошибкам при передаче данных по сети кластера.

Вне зависимости от того, используются источники бесперебойного питания или нет, кластерную стойку и ее электропитание **необходимо заземлить**.

Уделите особое внимание покрытию пола в помещении. Это должен быть специальный **антистатический линолеум**, паркет или плитка (последние варианты нежелательны, так как могут просто не выдержать веса стойки, а паркет еще и пожароопасен). Не допускайте даже временной установки кластера в помещении с обычным линолеумом или ковровым покрытием.

Желательно иметь возможность обесточить все оборудование в помещении из одной точки, лучше всего – у входа. Например, с единого электрического щитка, расположенного в этом же помещении. Это очень важно для обеспечения безопасности, прежде всего, в случае возникновения экстренных ситуаций.

Необходимость специальных мер для повышения **вибрационной устойчивости** возникает не часто, однако задуматься об этой потенциальной проблеме стоит. Если рядом проходит железная дорога или трамвайные пути, то проходящие составы вполне могут служить причиной

неустойчивой работы отдельных компонент кластера, среди которых наиболее чувствительными к подобного рода помехам являются диски.

Кроме обязательной защиты кластера от хакеров, надо позаботиться и о его **физической защите** и ограничении доступа в помещение. Безусловно, помещение должно быть надежно укреплено, и стоимость соответствующей подготовки нужно включить заранее в проект. Ключи от помещения должны быть только у ответственных лиц. Возможно, стоит поставить сигнализацию или организовать систему видеонаблюдения с возможностью видеоконтроля за помещением через Интернет.

Требования к **пожарной безопасности** в случае с круглосуточно работающим кластером выше требований, предъявляемым к местам, где на компьютерах работают лишь в рабочие часы, следовательно, может потребоваться соответствующая доработка помещения. Как минимум, там должен быть порошковый огнетушитель и пожарная сигнализация, для серьезных систем лучше поставить автоматическую систему пожаротушения.

Будет ли кластер расширяться или как-то модифицироваться в будущем? Очень желательно решить этот вопрос сразу. Сейчас помещение подготовлено под текущую конфигурацию кластера и никаких проблем с обеспечением его работы не ожидается. А что будет при добавлении стоек, новых кластеров, расширении дисковых подсистем? Если планы развития вычислительной базы вполне реальны, то нужно сразу оценить имеющиеся возможности по его администрированию и сопровождению, площади для размещения нового оборудования, резервы по электропитанию, по мощности систем климат-контроля. По-хорошему, нужно еще раз аккуратно пройти по всем вопросам данного раздела, сопоставив масштабность будущих изменений с теми последствиями и затратами, которые они вызовут.