Риски ЦОД: резервирование инженерных систем

HOST1993 19 августа 2013 в 12:35 17,2k

Начинать чинить надо, пока не сломалось — сломанное поддаётся ремонту гораздо неохотней.
Юрий Татаркин

После того как обеспечены надежные стены и крыша над головой для ЦОД (статья «Риски ЦОД: выбираем месторасположение»), следующим шагом на пути обеспечения его отказоустойчивости должно стать резервирование инженерных систем. Строя дата-центры более 10 лет, мы убедились, что не все заказчики в полной мере осознают важность дублирования основных коммуникаций. Космические корабли и те падают, а оборудование в ЦОД в идеале должно работать 365 дней в году и 24 часа в сутки. Любая вышедшая из строя или нуждающаяся в профилактике деталь должна быть заменена без остановки работы всех критичных сервисов.

Как справедливо отметили наши читатели, далеко не всем компаниям нужен надежный ЦОД. Для некоторых его бесперебойная работа не предмет переживаний, а многие предпочтут хранить свои данные в публичном облаке. Данный паблик предназначен в большей степени для тех, кто по тем или иным соображениям безопасности или проходимости каналов связи сделал свой выбор в пользу собственного дата-центра и работы сервисов с уровнем доступности не менее трех девяток (простоя не более 1,6 часов в год).

Отказоустойчивость и резервирование: что говорит мировой опыт?

Согласно стандартам Uptime Institute выделяют четыре уровня отказоустойчивости инфраструктуры ЦОДа:

	Уровень отказоустойчивости	Время простоя ЦОД в год	Схема резервирования
Tier I	99,671%	28,8 часа	Схема резервирования отсутствует (N). Ни одна из систем не резервируется и простой каждой единицы оборудования означает простой всего ЦОД.
Tier II	99,749%	22,0 часа	Схема резервирования N+1. К N единицам добавляется 1 резервная, что уменьшает риски выхода из строя ЦОД.
Tier III	99,982%	1,6 часа	Схема резервирования N+1 с возможностью параллельного проведения профилактических работ.
Tier IV	99,995%	0,4 часа	Схема резервирования 2(N+1). Каждый элемент системы N+1 дублируется аналогичным.

Использование классификации Tier подразумевает, что все инженерные системы и компоненты ЦОД, вплоть до запаса топлива для дизель-генератора, воспринимаются как единое целое. Наличие хотя бы одного нерезервированного компонента приводит к снижению уровня отказоустойчивости и увеличению возможных часов простоя ЦОД. Количество таких компонентов, а также статистика по плановым и внеплановым отказам дата-центров в год влияют на допустимое время простоя. Например, для ЦОД уровня Tier I характерно внеплановое отключение 1,2 раза в год. Плюс, из-за отсутствия резервных систем дата-центр не будет работать еще два раза по двенадцать часов во время планового обслуживания. В итоге суммарное время простоя будет рассчитываться как: 12+12+4x1,2=28,8 часов.

Для расчета уровня отказоустойчивости в процентах нужно:

((t работы — t простоя)×100%)/ t работы,

где t работы – максимальное количество часов работы ЦОД в год (24 часа в сутки 365 дней в году).

t простоя – это время планового простоя ЦОД в год.

Классифицируя способы резервирования, принято выделять следующие схемы: N+1, 2N и 2(N+1). Применение схем N+1 и N+2 по сравнению с 2N дают значительную экономию бюджета и при неплохом уровне отказоустойчивости (разом все элементы системы вряд ли выйдут из строя). Однако, нужно помнить, что с ростом числа рабочих единиц (N), согласно теории вероятности, доступность системы ухудшается. В ситуации большого количества элементов (большого N, например, источников бесперебойного питания) уместнее использовать схему 2N, когда каждый компонент системы полностью задублирован. Это позволит в разы увеличить отказоустойчивость и снизить время простоя. В то же время, ни N+1, ни 2N не резервируют систему в целом, а потому не исключают опасность аварии на участке между зарезервированными элементами системы. Поэтому Tier IV рекомендует использовать 2 независимые схемы, каждая из которой полностью задублирована, 2(N+1).

Неиссякаемая энергия

Основой надежной работы ЦОД является электроснабжение: бесперебойное (источники бесперебойного питания – ИБП) и гарантированное (дизель-генераторные установки – ДГУ). В момент исчезновения напряжения городской сети ИБП должны поддержать питание оборудования до полного запуска ДГУ, который сможет обеспечить электроэнергией весь ЦОД.

Для того чтобы ЦОД не встал в отсутствии электроснабжения, крайне важно, во-первых, зарезервировать ИБП, а, во-вторых, проводить регулярные сервисные работы.

К каким рискам может привести наличие только одного ИБП – в целом понятно. В лучшем случае мы не сможем провести тестирование источника, в худшем – получим простой ЦОД. Но порой даже наличие нескольких ИБП не дает свободу действий. Так в одной организации источников в ЦОДе было два, но каждый питал только свою группу серверов, а не служил резервом друг для друга. При проведении технического обслуживания у сервис-инженера прихватило спину. Падая, он каким-то образом умудрился обесточить выход ИБП. И, по закону подлости, выключившийся в разгар рабочего дня источник обесточил группу серверов с наиболее критичными приложениями.

«Боевой» запуск дизель-генератора (ПБ) – проверка возможности запуска дизель-генератора в автоматическом режиме при пропадании внешней сети. Производится с помощью имитации полного отключения внешнего питания ЦОД. Время от отключения питания до запуска дизель-генератора серверное оборудование работает от батарей ИБП (обычно 1-3 минуты).

Запуск дизель-генератора под нагрузкой (ПН) – проверка способности дизель-генератора поддерживать питание подключенного к нему оборудования. Производится ручным переключением нагрузки на генератор (с помощью панели управления) после его запуска и выхода на нормальную работу. На время переключения АВР серверное оборудование работает от батарей ИБП (около 0,3-1 сек.). Кстати, для переключения нагрузки на ДГУ лучше использовать мотор-приводы, они хоть и работают медленнее, но срок службы и надежность у них выше.

Для предотвращения нежелательных простоев нужны регулярные комплексные сервисные работы. В одном из ЦОД проверки проводились только в отношении ДГУ. ИБП исправно показывал 10 минут автономии, но его никто не обслуживал. Возраст батарей к тому времени перевалил за 5 лет, и во время одного из боевых запусков они смогли проработать лишь 29 секунд. В то время как ДГУ завелась и смогла принять на себя нагрузку спустя только 33 секунды. Ко всему прочему, все оборудование было запитано от одного ИБП (от второго было решено отказаться еще на этапе реализации из-за бюджетных ограничений). В итоге – падение ЦОД. Полное восстановление всех вычислительных систем заняло около 12 часов.

Основные ошибки:

• Отказ на стадии реализации от второго ИБП. Трудные времена закончились, но второй ИБП так и не был приобретен.

• Отсутствие комплексного обслуживания всех инженерных систем ЦОД. При регулярном сервисном обслуживании ИБП об их неудовлетворительном состоянии стало бы известно заранее.

• Отсутствие регламентов планового обслуживания ЦОД и хаос при его эксплуатации.

Пути миграции тока

Ваши ИБП зарезервированы и вы регулярно их обслуживаете? Молодцы, но не вздумайте на этом останавливаться! Зарезервируйте еще и кабельные линии электроснабжения ЦОД, и установите 2 АВР, которые полностью резервируют друг друга. В идеале, они должны быть подключены к разным независимым электрощитам. В крайнем случае можно протянуть две линии и от одной щитовой, чтобы не получилось ситуации, как у одного из наших заказчиков.

При внедрении системы диспетчеризации в небольшой, но значимый ЦОД необходимо было поставить трансформаторы тока на основной ввод. Проблема была в том, что ввод был только один, а обесточить дата-центр было нельзя. После всех подготовительных работ питание было отключено. Пока оборудование ЦОД работало от батарей, монтажники трудились не покладая рук, а инженер, вытирая пот со лба, считал минуты на дисплее ИБП.

Основные ошибки:

• Система диспетчеризации была незаслуженно забыта при проектировании.

• Линия питания ЦОД не была зарезервирована.

Стало жарко

Система «чиллер-фанкойл» – система кондиционирования воздуха, в которой теплоносителем между центральной холодильной машиной (чиллером) и локальными теплообменниками (фанкойлами) служит охлажденная жидкость, циркулирующая под относительно низким давлением – обыкновенная вода (в тропическом климате) или водный раствор этиленгликоля (в умеренном и холодном климате).

Не стоит забывать и о резервировании систем кондиционирования. За последние два месяца довелось увидеть два проекта охлаждения ЦОД с использованием системы чиллер-фанкойл без резервирования трассы между чиллерами и сухими охладителями. Использование данного решения в реальной жизни с высокой долей вероятности приводит к простою ЦОД. В случае замены теплоносителя (что не редкость), только резервная трасса может сохранить работоспособность системы охлаждения, а значит и всего дата-центра.

Еще очень важный момент – разделение внешнего и внутреннего контуров охлаждения. Так в одном проекте на кровле седьмого этажа предлагалось установить два двухтонных чиллера, бак аккумулятор холода, мощную подкачивающую насосную станцию. Подача и обратка длиной двести метров была запланирована напрямую с крыши до блоков охлаждения в ЦОД, который находился в цоколе. В итоге, при даже небольшом прорыве трубы или неплотных соединениях внутренних блоков охлаждения все десять тонн этиленгликоля под давлением могли затопить ЦОД и электрощитовую заказчика.