<https://www.manufacturing.net/article/2017/04/scada-data-center>

**Автономная система управления Дата-центром на основе экспертных правил и поддерживающих самообучающихся нейронных сетях.**

**Один из основных вопросов, который состоит перед современными Центрами обработки данных, а также любой инфраструктуры, сочетающей в себе мультивендорность, множественный уровень иерархии, сложность взаимосвязей между компонентоми и неполной информации о внутренних протекающих процессах:** Может ли автоматизированная система управления дата-центра перейти полностью в автономный режим работы, корректируя собственные параметры на основе знаний специальных спроектированных экспертных систем и нейронных системах, обучаемых на основе информации с телеметрии датчиков ДЦ?

**Для уточнения проблематики, необходимо в первую очередь понять, что**

современные дата-центры представляют из себя большую смесь различных технологий, компинация которых непосредственно участвует в реализации конечной цели существования ДЦ. Нередки случаи, когда инфраструктура становится настолько сложной, что люди, ослуживающие один участок установок, совершенно не понимают, что происходит в другой части. Люди, ответственные за общий мониторинг ситуации и реагировании в чрезвычайной ситуации, часто не способны определить правильную последовательность действий, возвращающую систему в стабильное состояние по различным причинам, в группу котороых входят такие проблемы, как отстутсвие общей картины, незнание определенных технологий, отсутсвие необходимого опыта работы с оборудованием и знания того, как поведет себя система в том или ином случае. Немаловажной проблемой является то, что в попытке в кратчайшие сроки приостановить падение уровня сервиса, теряются детали инцидента, что повлекло его, были ли другие возможные варианта исхода. Это приводит нас к вопросу о необходимости гибридной системы мониторинга и контроля состояния инфраструктуры, которая способная помочь операторам принять решение на основе экспертных правил, полученных на основе эмпирических знаний, а также самообучаемых нейронных сетях.

Наиболее проблемным местой является визуализация возникновения инцидента, анализ графа затрагиваемых подсистем, взаимного влияния одной подсистемы на другую. В таких случаях сложно получить одну картину, на которой будет видна история событий.Концепция автоматизированной системы управления дата-центром состоит в возможности системы поддерживать внутреннее состояние на заданном оптимальном уровене, а также реагировать на неизвесткие раздражители, руководствуясь набором экспертной логики, а также выполнять рефлексию принятых решений. Типичным проблемой является потеря соединения с сервером, расположенным физически в Машинном зале – существует как минимум три причины, по которым может произойти подобное событие – потеря электропитания, выход из строя межсетевого оборудования, обрыв на линии связи. Подобные причины человек способен без труда определить, но он не сможет выполнить это хладнокровно в случае крупномасштабной аварии, где счет идет на секунды до перехода на резервное оборудование.

Решение проблемы может быть достигнуто различными способами. На данный момент уже существует как минимум 2 класса программных продуктов, в той или иной степени позволяющих решить заданную проблему – SIEM и SCADA. Но при этом SIEM больше ориентирована на сбор событий информационной безопасности и построение корреляций между ними, а SCADA на визуализацию технических процессов и контроль над ними. Оба класса продуктов подразумевают активное привлечение оператора для подверждения тех или иных операций. Также и SCADA, и SIEM можно рассматривать как пассивные элементы инфраструктуры – они помогают принять управленческое воздействие и даже его назначить на выполнение, однако не могут самостоятельно сделать вывод о дальшейних действиях. Примерами программных продуктов класса SIEM могут послужить такое ПО, как IBM Qradar и PT SIEM – оба их них имеют богатый функционал в обасти сбора событий, но для того, чтобы заложить в него экспертную логику, требуетс большое количество человеко-часов. При этом продукты не имеют возможность выполнять самокоррекцию, рефлексию; также имеются проблемы с визуализацией полученных данных.

В рамках этой проблемной области, перед исследователем стоит несколько крупных задач, решение которых необходимо для успешного функционирования автономной системы:

* **Проблема сбора и хранения событий**. Подсистема сбора событий должна предоставлять необходимый функционал для подключения к разнородным информационныым системам, которые чаще всего имеют собственный формат записей. Также немаловажным аспектом является производительность системы и возможность маштабирования, проблемы резервирования можностей и емкости подсистемы сбора.
* **Подсистема анализа событий**. Система должна иметь возможность обрабатывать множество данных парралельно в единицу времени, обрабатывать в непоследовательном режиме, уметь выстраивать взаимосвязи между разными по времени и месту появления, а также характеру события. Для каждого из наборов событий правила вывода должны применяться максимально быстро и точно.
* **Подсистема хранения знаний**. Экспертные знания сложно формулизованы, еще более сложно их применить в рамках компьютерных проблем, так как невозможно предстазать содержимое экспертного правило – оно может основываться на значениях параметров, которые до сих пор не появлялись в рассматриваемой системе, или же является нормальным параметром в данном случае. События, на которых применяются экспертные знания, должны быть унифицированы, иметь гарантированное число полей, позволяющие минимально реализовать экспертную логику.
* **Подсистема самообучения**. Одна из задач каждой автономной системы – выполнять корректирование собственной логики работы, некоторых системных параметров. Данное обучение должно происходить постоянно, даже если не непрерывно – система должна гарантировать, что в случае нахождения в зоне оптимальных ппараметров, она будет сохранять свое состояние максимально долго, а в случае нарушения максимально активно начнет искать возможные выходы, в том числе постарается предложить несколько дальнейших исходов, для каждого из которых рассчитвает необходимые метрики, которые будут учитываться при переходе на ту или иную ветвь развития событий.
* **Подсистема реагирования**. Система должна знать, как реагировать в той или иной ситуации, знать критерии, при которой она в праве применить конкретное действие, а также те последствия, к которым приведет действие.
* **Подсистема визуализации**. Система не можт обойтись без участия человека. Ему в свою очередь для оперативного анализа ситуации необходим удобный интерфейс, через который он будет интерпретировать происходящие события и принимать корректирующие действия.
* **Подсистема управления**. После проведения анализа событий системы, оператом должен иметь возможность повлиять на работу систему, перевести ее в ручной режим или скорректировать ее поведение. Сделать корректировку он может либо за счет определения новых экспертных правил, либо за счет дополнительного обучения на заданных параметрах, либо за счет изменения глобальных переменных.

Автоматизированная система может функционировать автономно без участия человека в рамках установленных желаемых параметрах среды, если в ней присутствют и работают 5 компонентов из 7, описанных выше – это Подсистема сбора и хранения событий, Подсистема анализа событий, Подсистема хранения знаний, Подсистема самообучения, Подсистема реагизования.

Для подверждения всей сложности инфрастуктуры ЦОДов, проблемы слабости контроля и автоматического реагирования, при котором вся зона ответственности лежит на человеке, необходимо получить журнал всех событий со всей инфраструктуры среднестатического ЦОД Tier3 – размер данного файла может достигать 500-1000 Мб за день.

Проблема автоматизации систем управления ЦОД является подмножеством более крупной проблемы, а именно Проблемы создания автономных систем за контролей состояния подсистемы на заданных уровнях.