

КЛЪСТЕРНИ СИСТЕМИ. АВТОМАТИЗИРАНО ИНСТАЛИРАНЕ И ДОБАВЯНЕ НА ВЪЗЛИ В СЪЩЕСТВУВАЩИ КЛЪСТЕРИ

Росен Христов^{1,*}, Ангел Голев²

^{1,2} Факултет по математика и информатика,
ПУ „Паисий Хилендарски“, гр. Пловдив, бул. „България“ № 236
Emails: hristev@uni-plovdiv.bg, angelg@uni-plovdiv.bg

* Автор за кореспонденция: hristev@uni-plovdiv.bg

Резюме. Клъстерните системи интегрирани в една ИТ инфраструктура решават проблеми свързани с мащабируемост на системата, скорост на обработка на данните, наличност и управление на ресурси. Целта на статията е да се прецени възможността за автоматизирано инсталиране и добавяне на нови възли във вече съществуващи клъстери, което би довело до оптимизиране на човешките ресурси, необходими за поддържането и развитието на ИТ инфраструктурата на една организация.

Ключови думи: *клъстерни системи, автоматизирано инсталиране на системи, клъстери с балансирано зареждане, отказоустойчиви клъстери*

Въведение

Клъстерните системи намират все по-голямо приложение в различни отрасли на индустрията. Множество организации са се сблъскали с редица проблеми, свързани със забавяне в изчисленията при дадени услуги и критичност на съдържанието. Клъстерните системи, интегрирани в една ИТ инфраструктура, могат да решат проблеми, свързани с мащабируемостта на системата, скоростта на обработка на данните, наличността и управлението на ресурси.

Едно от основните предимства на клъстерите е мащабируемостта. Това спомага на организациите, използващи клъстерни системи, да оптимизират разходите си, като постепенно добавят нови изчислителни мощности тогава, когато са необходими. Инсталирането, конфигурирането и добавянето на нови възли в съществуващи клъстери е енергоемко за ИТ

специалистите в организациите [1]. Ще разгледаме възможността за автоматизиране на този еднотипен процес по инсталиране и добавяне на нови възли във вече съществуващи клъстери.

Клъстерни системи

Компютърен клъстер е термин с широка употреба, който се отнася до група от поне два компютъра, свързани чрез мрежова инфраструктура [2]. В идеалния случай клъстерът работи като интегриран изчислителен ресурс, управляван от едно място, което обхваща всичките му възли. По този начин потребителите виждат клъстера като единна система, а потребителските процеси могат да се изпълняват на всеки възел от клъстера [3]. Възлите в клъстера се свързват чрез хардуер, софтуер и високо скоростни мрежови канали, използвайки стандартни или специализирани протоколи за комуникация помежду си. Следенето за наличност на възлите и цялостност на клъстера се постига обикновено с техниката heartbeat [4], като всеки възел непрекъснато съобщава за наличността си на останалите, чрез специалната мрежова свързаност.

На практика клъстерите могат да бъдат разгледани като няколко вида:

- клъстери с висока наличност (High-availability clusters, HA);
- клъстери с балансирано зареждане (Load balancing clusters);
- изчислителни клъстери (Computing clusters).

Клъстерните системи с **висока наличност** обединяват повече от един изчислителен ресурс, като приложенията, които се стартират, могат да работят на повече от един сървър едновременно. Целта е при отпадане на сървър, друг автоматично да поеме работата и натоварването с минимален downtime [5].

Основен проблем при използването на техниката heartbeat при клъстерите с висока наличност е, когато всички частни вътрешни връзки са прекъснати, но възлите на клъстера все още функционират. Така произволен възел може погрешно да реши, че всички останали възли са изключени и да стартира услуги, работещи на други възли. Това може да наруши целостта на данните в клъстера.

Съществуват два вида концепции при клъстерите с висока наличност: асиметрични клъстери и симетрични клъстери [4].

При асиметричните клъстери, за да се гарантира целостта и достъпността на данните, е препоръчително винаги да има наличен допълнителен сървър, т.е. сървърите в клъстера да бъдат $n+1$ [4]. Ако възел отпадне от клъстера, на допълнителния сървър в режим на готовност веднага ще се присвои IP-адреса на отпадналия и работата ще продължи [7].

При симетричните кълстери съществуват два или повече възела с еднакви конфигурации, които са активни директно за достъпване от клиенти [7]. Ако един възел се провали, клиентите автоматично се свързват с другия.

Ако в симетричен кълстер имаме система за балансиране на натоварването, разпределяща заявките от клиенти към различните възли, тогава наричаме кълстера с *балансирано зареждане* [7].

Основната цел на тези кълстери е разпределянето на изчислителните задачи на отделни изчислителни мощности [3]. Това е лесно постижимо чрез добавяне на входни възли, които разпределят задачите между останалите възли в кълстера. В зависимост от целта, към която се стреми мрежата, има различни стратегии за изграждането на кълстера. Този вид кълстери могат да помогнат за намаляване на натоварването на изчислителите ресурси при по-големи приложения или да ускорят услугите при позициониране на възлите на отдалечени места.

Изчислителните кълстери се използват за паралелна/разпределена обработка, като потребителят вижда кълстера като една машина. Често се използват с научна цел, като съществен показател за производителността се явява работата с числа с плаваща запетая. Изчисленията се разпределят равномерно на различните възли от кълстера, като по този начин се постига висока скорост на обработка.

Кълстерните системи имат множество предимства пред единичните компютри, като най-значимите от тях са:

- Мащабируемост – Могат да бъдат създавани големи кълстери, които превишават производителността на най-големите самостоятелни машини. Обикновено в кълстера няма ограничение на броя на добавяните машини. Голямо предимство е, че добавянето на нови изчислителни мощности може да става, когато е необходимо.
- Висока наличност – Всеки възел в кълстера е отделна физическа машина, което при правилно конфигуриране предполага, че отпадането на възел няма да доведе до загуба на услуга или загуба на данни.
- Кълстерите обикновено се създават, за да бъде подобрена производителността или да бъде добавена висока наличност. Цената, която се плаща за изграждането на кълстер, обикновено е по-ниска от цената на единичен компютър със сравнима мощност и наличност.
- Централизирано управление – Възлите в кълстера могат да се управляват от едно приложение, като част от операциите могат да

се прилагат и върху няколко възела едновременно. От тази гледна точка, възлите в един клъстер са със сходни хардуерни параметри и сходни операционни системи.

Автоматизирано инсталиране, конфигуриране и добавяне на възли в съществуващи клъстери

Тъй като едно от основните предимства на клъстерите е мащабируемостта, често се налага добавяне на допълнителни мощности към вече съществуващ клъстер. Това е операция, която оптимизира разходите за ИТ инфраструктурата, като по този начин се правят инвестиции в нов хардуер и софтуер тогава, когато те са наистина необходими.

Инсталирането и добавянето на нови възли към клъстер е времеотнемащ процес за екипите от ИТ специалисти на дадена организация. Ще разгледаме възможността за автоматизиране на следните операции при добавяне на нов възел:

1. Конфигуриране на локалното дисково пространство на възела

Въпреки че повечето клъстерни системи използват мрежово достъпни ресурси за съхранение, може да има и локални дискове или дискови масиви. На тези дискове се съхраняват операционната система, софтуера за клъстеризация, отделено е локално дисково пространство за всеки възел и други.

Различават се два вида дискови масиви – софтуерни и хардуерни. Хардуерните дискови масиви са допълнителен хардуер, разширителна платка или чипове на дънната платка, които разпределят информацията по различните физически твърди дискове. При софтуерните масиви разпределянето на данните по твърдите дискове се извършва от операционната система, а необходимите изчисления за това се поемат от процесора на компютъра.

При хардуерните RAID контролери дисковите масиви се конфигурират предварително в софтуера на масива, който се зарежда непосредствено след BIOS на машината. Поради наличието на множество производители на хардуерни масиви и различни варианти за конфигурация, не може еднозначно да се отговори дали операцията може да бъде автоматизирана. Част от производителите на RAID контролери предлагат софтуер за управление, предназначен за конкретна операционна система. При такъв вид контролери е възможна автоматизацията на конфигурирането на дисковите масиви [8].

Софтуерните RAID масиви се конфигурират на ниво операционна система. За Linux основният софтуерен пакет, който се използва за създаването на дисков масив, е mdadm, за FreeBSD – GEOM, а за Windows

могат да се използват вградените функции, които са достъпни през CMD и PowerShell. Могат да бъдат създадени параметризирани скриптове, които автоматично да създават и конфигурират дисковия масив.

2. Инсталиране на базова операционна система

Софтуерните системи, с които може да създадем клъстер в по-голямата част от случаите, се базират на някоя от най-разпространените операционни системи – Linux, FreeBSD и Windows. В зависимост от софтуера, част от тях се предлагат в прекомпилирани инсталации или такива, които трябва да бъдат инсталирани върху базова операционна система или се предлагат и двата варианта. Например, ако разгледаме две от най-използваните системи, с които можем да създадем виртуализационни клъстери с висока наличност – Proxmox VE и VMWare ESXi, то Proxmox VE може да бъде открит в прекомпилирана операционна система или тъй като е базиран на Debian, може да бъде инсталиран върху Debian. VMWare ESXi може да бъде инсталиран само от преконфигурирано копие.

Ако инсталираме Proxmox VE от преконфигуриран дисков образ, няма да разполагаме с всички файлови системи и може да има други ограничения в ОС. Може да се наложи допълнително инсталиране и конфигуриране на файлови системи, но на сървъра ще бъдат налични всички нужни пакети и софтуера за виртуализация. Ако инсталираме Proxmox VE върху базова операционна система, то предварително трябва да се инсталират различни пакети като postfix, open-iscsi и други.

3. Инсталиране на софтуер за клъстеризация

Ако е избран подход за инсталиране на базова операционна система, то трябва да бъде инсталиран и софтуер за клъстеризация. Инсталирането на софтуера за клъстеризация може да бъде разделено на няколко стъпки. Първоначално трябва да бъдат налични всички зависимости за системата, след което да се премине към инсталирането на софтуера за клъстеризация. Проверката за наличност и инсталирането на липсващите зависимости, както и инсталирането на този софтуер, могат да бъдат описани в изцяло автоматизирани скриптове, в зависимост от операционната система – bash, batch, tcsh или Python.

4. Конфигуриране на мрежовите настройки на възела

При големите ИТ инфраструктури, с цел сигурност и лесна управляемост, добрата практика изисква сегментиране на мрежовата инфраструктура. Сегментирането обикновено се извършва с виртуална частна мрежа (VLAN-и). Това повишава нивото на абстракция на мрежата и може да доведе до минимално забавяне при обмяна на големи обеми данни между възлите. Използването на различна капсулирана мрежа за комуникация между възлите повишава сигурността и намалява забавянето

на мрежово ниво. Конфигурирането на мрежовите настройки на възела може да бъде полуавтоматизиран процес, на който се подават необходимите VLAN-и и статични адреси, след което скриптове могат да конфигурират мрежовите карти с всички необходими интерфейси. Ако това е процес, изпълняван в една организация, използваща сходни комутатори, скриптът би позволил VLAN-ите автоматично да се „разплетат“ по комутаторите до възела. Тъй като всеки производител на мрежови комутатори използва различни видове инструкции и собствен „език“ за конфигуриране през CLI, задачата за автоматизиране на конфигурирането на мрежовите комутатори ще бъде сложна.

5. Добавяне на конфигурирания възел към клъстера

При повечето клъстерни системи при добавяне или изключване на възел от клъстер, стартиране на клъстер и други, се изисква събиране на квота. За да е валидна операцията, трябва $50\% + 1$ от възлите да са налични и да отговорят на заявката. Операцията по добавяне зависи от вида клъстер, като такива операции обикновено се изпълняват от интерфейс за управление или от конзола. Новият възел се опитва да се свърже с клъстера и след добавянето му се конфигурира с общите настройки за клъстера. Тези настройки са за споделените за клъстера мрежови ресурси за съхранение на данни, методите за вписване в клъстерната система и други.

Имайки предвид, че е препоръчително възлите в един клъстер да са еднотипни като хардуер и операционни системи, то процесът по добавяне на нови възли във вече съществуващ клъстер може да бъде автоматизиран с написването на различни скриптове и помощни приложения, които конфигурират софтуерните RAID масиви за съхранение на данни, инсталират базовата операционна система, откриват и премахват зависимости, инсталират софтуера за клъстеризация, конфигурират мрежовите настройки и автоматично присъединяват новия възел във вече съществуващ клъстер.

Описаният метод за добавяне на нови възли е приложим за клъстери, изградени на ниво физическа ИТ инфраструктура. При добавяне на нови възли в облачна ИТ инфраструктура последователността и операциите не съответства на гореописаните.

Заклучение

Клъстерните системи намират все по-голямо приложение в различни отрасли на индустрията поради техните безспорни предимства пред използването на единични компютри. За да бъдат ефективно внедрени, ИТ екипите на дадена организация полагат множество усилия, за да поддържат данните консистентни и да гарантират високата наличност на услугите.

С подходящи програмни средства могат да се автоматизират операциите по инсталиране и добавяне на нови възли във вече съществуващи клъстери. Автоматизирането на тези времеотнемачи и еднотипни операции би довело до оптимизиране на човешките ресурси, необходими за поддържането и развитието на ИТ инфраструктурата на една организация, а специалистите ще могат да се съсредоточат върху по-важни за организацията дейности.

Благодарности

Изследването е частично финансирано по проект МУ21-ФМИ-007 при НПД на Пловдивски университет „П. Хилендарски“ и ФП21-ФМИ-002 „Интелигентни иновационни ИКТ в научни изследвания по математика, информатика и педагогика на обучението“ към НПД на Пловдивския университет „Паисий Хилендарски“.

Литература

- [1] K. Bourne, *Application Administrators Handbook*, Morgan Kaufmann, 2014, ISBN: 978-0123985453.
- [2] E. Kaur, A Review of Computing Technologies: Distributed, Utility, Cluster, Grid and Cloud Computing, *International Journal of Advanced Research in Computer Science and Software Engineering (IJARCSSE)*, Vol. 5, Issue 2, 2015, 144–148, ISSN: 227 128X.
- [3] P. Werstein, H. Situ, Z. Huang, Load Balancing in a Cluster Computer, *Proc. of Seventh International Conference on Parallel and Distributed Computing, Applications and Technologies (PDCAT'06)*, 2006, 569–577, doi: 10.1109/PDCAT.2006.77, ISBN: 0-7695-2736-1.
- [4] S. Vugt, *Pro Linux High Availability Clustering*, Apress, 2014, ISBN: 978-1-4842-0080-3.
- [5] M. Rasman, *CentOS High Availability*, Packt Publishing 2015, ISBN: 9781785282485.
- [6] K. Kopper, *Linux Enterprise Cluster*, No Starch Press, 2017, ISBN: 978-1593270360.
- [7] E. Vargas, J. Bianco, D. Deeths, *Sun Cluster Environment Sun Cluster 2.2*, Prentice Hall, 2001, ISBN: 978-0130418708.
- [8] C. Wolf, E. Haliter, *Virtualization – From the Desktop to the Enterprise*, Apress, 2005, ISBN: 9781590594957.

CLUSTER SYSTEMS. ALGORITHM FOR AUTOMATED INSTALLATION AND ADDITION OF NODES IN EXISTING CLUSTERS

Rosen Hristev^{1,*}, Angel Golev²

^{1, 2} *Faculty of Mathematics and Informatics, University of Plovdiv "Paisii
Hilendarski", Plovdiv 4002, Bulgaria blv, 236, Bulgaria*
Emails: hristev@uni-plovdiv.bg, angelg@uni-plovdiv.bg

* *Corresponding author: hristev@uni-plovdiv.bg*

Abstract. Cluster systems, which are integrated into the IT infrastructure, are able to successfully solve problems related to system scalability, data processing speed, as well as availability and resource management. This paper describes the possibility of automated installation and integration of new nodes into already existing clusters. This process leads to optimization of human resource allocation for maintenance and development of IT infrastructure within organizations.