

АГРЕГИРОВАНИЕ КРИТЕРИЕВ ПРИ ВЫБОРЕ ВЫЧИСЛИТЕЛЬНОГО КЛАСТЕРА

В.Н. ЛОБАНОВ¹
А.Б. ПЕТРОВСКИЙ²

¹ОАО НИИ вычислительных комплексов им. М.А.Карцева, г. Москва

²Институт системного анализа Российской академии наук, г. Москва

e-mail:
fisher_1980@mail.ru
pab@isa.ru

В работе описан подход к многокритериальному выбору вычислительного кластера, в котором строится иерархическая система составных критериев, агрегирующих характеристики кластера, и используются методы вербального анализа решений.

Ключевые слова: конфигурация, вычислительный кластер.

В настоящее время при решении различных прикладных и научных задач все большее применение находят относительно дешевые высокопроизводительные вычислительные кластеры, которые выступают альтернативой дорогостоящих суперкомпьютеров. Современные стандартные и относительно недорогие микропроцессоры, сетевые технологии и периферийные устройства позволяют использовать их для построения разнообразных конфигураций кластеров, которые можно гибко и последовательно изменять и наращивать за счет добавления новых вычислительных модулей, обеспечивая требуемую производительность и соответствие потребностям пользователей.

Сравнение разных вариантов конфигураций кластеров, выбор наиболее предпочтительных аппаратных и программных решений представляет собой достаточно сложную слабо структурируемую и плохо формализованную задачу. Это обусловлено тем, что вычислительные кластеры, как и другие сложные технические системы, характеризуются большим числом показателей, а выбор осуществляется по многим критериям, среди которых могут быть и количественные, и качественные. Вместе с тем, как правило, вариантов конфигураций бывает немного. Поэтому такие варианты обычно оказываются несравнимыми друг с другом по своим показателям, что не позволяет применять известные методы принятия решений для выбора лучшего варианта кластера.

Для выбора наиболее предпочтительного вычислительного кластера из имеющегося перечня конфигураций, обладающего требуемыми для прикладных применений параметрами, использована многоэтапная технология ПАКС (Последовательное Агрегирование Классифицируемых Состояний) [4], в которой последовательное сокращение размерности признакового пространства проводится с помощью нескольких разных методов вербального анализа решений и/или их комбинаций, исходя из предпочтений лица, принимающего решение (ЛПР). Подходы к решению задачи многокритериального выбора вычислительных кластеров изложены в работе [5], процедура снижения размерности признакового пространства в задачах многокритериальной классификации описана в работе [3], методология вербального анализа решений, которая ориентирована на слабо структурируемые задачи, где доминируют качественные и субъективные факторы, представлена в книгах [1, 2].

В технологии ПАКС большое число исходных характеристик (базовых показателей) вычислительного кластера агрегируется в небольшое число критериев, имеющих порядковые шкалы оценок (количественные и качественные). Такое сокращенное описание объектов дает возможность упростить процедуру решения первоначальной задачи выбора, сформировать разные наборы промежуточных критериев и воспользоваться разными способами конструирования их шкал. Сопоставление результатов, получаемых для разных иерархических систем критериев, позволяет сравнить их между собой и выбрать как наиболее предпочтительную систему критериев, так и оценить качество сделанного выбора.

Рассмотрим в качестве примера один из этапов решения задачи выбора вычислительного кластера, где строятся несколько иерархических систем критериев путем агрегирования исходных характеристик. Пусть имеется три варианта вычислительных кластеров: ВК₁, ВК₂ и ВК₃. В качестве исходных характеристик кластеров были выбраны следующие группы показателей.

ХМ. Технические характеристики модуля (частота ядра процессора; разрядность ядра процессора; количество потоков; количество ядер процессора; поддержка процессором объема оперативной памяти; количество процессоров в модуле; объем оперативной памяти модуля; наличие ускорителя универсальных вычислений; дисковая память модуля; наличие в модуле оптического накопителя данных).

ВХ. Вычислительные характеристики кластера (число модулей в кластере; скорость обмена между модулями; наличие встроенных средств ввода-вывода; наличие бесперебойного питания; программные характеристики кластера; возможность модернизации технических и программных средств кластера).

КХ. Конструкционные характеристики кластера (размеры кластера (высота, глубина, ширина); масса кластера; защищенность от помех).

ЭХ. Эксплуатационные характеристики кластера (энергопотребление; уровень шума; тепловыделение; условия эксплуатации (температура, влажность); наработка на отказ).

ПК. Производительность кластера.

СИ. Стоимость изготовления кластера.

Для каждого исходного показателя была сформирована вербальная шкала оценок с 3 градациями. Например, производительность кластера оценивалась как ПК0 – высокая (>2000 Гфлопс); ПК1 – средняя (2000-500 Гфлопс); ПК2 – низкая (<500 Гфлопс).

Иерархическая система критериев строится на основе предпочтений и знаний ЛПР. Прежде всего, определяется, какие исходные показатели будут рассматриваться в качестве конечных критериев для сравнения вариантов кластеров и выбора наилучшего. В данном примере такими показателями стали производительность кластера и стоимость изготовления кластера.

Остальные исходные характеристики объединяются в составные критерии. Процедура агрегирования состоит из несколько этапов:

- разбиение всех базовых показателей на группы критериев, в которых критерии близки по смыслу друг другу или являются неотъемлемыми характеристиками какого-либо составного критерия;
- построение дерева агрегирования для каждой сформированной группы критериев;
- формирование шкалы оценок для каждого составного критерия с указанием диапазона изменения для каждой градации на шкале.

Дерево технических характеристик модуля представлено на рис. 1.



Рис. 1. Технические характеристики модуля

В результате агрегирования все исходные характеристики были сведены к следующим пяти критериям, по которым осуществлялись сравнение и выбор лучшего варианта кластера: ПК. Производительность кластера; СИ. Стоимость изготовления кластера; ВХ. Вычислительные характеристики кластера; КХ. Конструкционные характеристики кластера; ЭХ. Эксплуатационные характеристики кластера.

Каждый из трех вариантов кластеров независимо оценивался тремя экспертами по 5 критериям. Результаты оценки кластеров, представленные в виде кортежей, приведены в табл. 1. Как видно из таблицы, кластеры по-прежнему остаются несравнимыми между собой, но их различие по пяти критериям становится более понятным и наглядным. Однако такое представление результатов экспертной оценки не позволяет выбрать лучший вариант.

Таблица 1

Оценки кластеров, представленные как кортежи

	ВК1	ВК2	ВК3
Эксперт 1	(ПК ¹ , СИ ¹ , ВХ ² , КХ ⁰ , ЭХ ⁰)	(ПК ² , СИ ⁰ , ВХ ¹ , КХ ² , ЭХ ¹)	(ПК ⁰ , СИ ² , ВХ ⁰ , КХ ² , ЭХ ²)
Эксперт 2	(ПК ¹ , СИ ¹ , ВХ ² , КХ ⁰ , ЭХ ⁰)	(ПК ² , СИ ⁰ , ВХ ⁰ , КХ ² , ЭХ ¹)	(ПК ⁰ , СИ ² , ВХ ⁰ , КХ ² , ЭХ ²)
Эксперт 3	(ПК ¹ , СИ ¹ , ВХ ² , КХ ⁰ , ЭХ ⁰)	(ПК ² , СИ ⁰ , ВХ ¹ , КХ ² , ЭХ ¹)	(ПК ⁰ , СИ ² , ВХ ⁰ , КХ ² , ЭХ ²)

Для дальнейшего сравнения кластеров и выбора наилучшего варианта экспертные оценки кластеров были представлены в виде мультимножеств:

$$A_i = \{kA_i(x_{11}) \circ x_{11}, \dots, kA_i(x_{1h_1}) \circ x_{1h_1}, \dots, kA_i(x_{51}) \circ x_{51}, \dots, kA_i(x_{5h_5}) \circ x_{5h_5}\}$$

над множеством $X = X_1 \cup \dots \cup X_5$ всех градаций оценок по шкалам X_s критериев ПК, СИ, ВХ, КХ, ЭХ. Здесь число $kA_i(x_{ses})$ указывает, сколько раз градация оценки $x_{ses} \in X_s$, $s=1, \dots, 5$, $es=0, 1, 2$ встречается в описании варианта A_i , $i=1, 2, 3$; знак \circ обозначает кратность градации оценки x_{ses} . Запись результатов оценки кластеров в виде мультимножеств дана в табл. 2.

Таблица 2

Оценки кластеров, представленные как мультимножества

	ПК ⁰ ПК ¹ ПК ²	СИ ⁰ СИ ¹ СИ ²	ВХ ⁰ ВХ ¹ ВХ ²	КХ ⁰ КХ ¹ КХ ²	ЭХ ⁰ ЭХ ¹ ЭХ ²	$d(A_i, A^+)$
ВК1	0 3 0	0 3 0	0 0 3	3 0 0	3 0 0	18
ВК2	0 0 3	3 0 0	1 2 0	0 0 3	0 3 0	22
ВК3	3 0 0	0 0 3	3 0 0	0 0 3	0 0 3	18

Выбор наиболее предпочтительного варианта кластера проводился с помощью трех методов группового многокритериального выбора: АРАМИС, лексикографического упорядочивания по градациям критериальных оценок и взвешенных сумм суждений экспертов [2]. Метод АРАМИС (Агрегирование и Ранжирование Альтернатив около Многопризнаковых Идеальных Ситуаций) упорядочивает многокритериальные варианты по показателю близости к идеальному варианту, имеющему лучшие оценки по всем критериям по суждениям всех экспертов. Метод лексикографического упорядочивания ранжирует варианты сначала по общему числу лучших оценок, затем по числу средних оценок, и так далее. В методе взвешенных сумм рангов место варианта в ранжировке определяется по сумме произведений числа экспертных оценок на вес ранга. Этими методами были построены следующие три ранжировки кластеров:

$$VK1 \approx VK3 > VK2, VK1 > VK3 > VK2, VK1 > VK2 \approx VK3.$$

Итоговое обобщенное упорядочение вариантов кластеров было получено при помощи процедуры Борда [2] и имеет вид: $VK1 > VK3 > VK2$. Вариант $VK1$ оказался предпочтительнее $VK3$, а $VK3$ предпочтительнее $VK2$.

Итак, несмотря на первоначальную сложность задачи выбора, обусловленную большим числом характеристик кластеров, удалось снизить размерность признакового пространства, выполнив агрегирование исходных показателей, а применив методы группового многокритериального принятия решений, найти лучший вариант вычислительного кластера, который обладает конкретными и понятными значениями признаков.

Продланное построение системы критериев показало, что результаты агрегирования сильно зависят от суждений экспертов и предпочтений ЛПР. Более того, даже если задача выбора решается одним человеком, то и в этом случае, в зависимости от суждений человека на разных этапах процедуры агрегирования, как обусловленных взглядом на проблему с разных сторон, так и продиктованных жизненным опытом, результаты агрегирования могут в большей или меньшей степени отличаться друг от друга.

Разработанный подход предполагается использовать в дальнейшем при создании компьютеризированной системы поддержки принятия решений, которая позволит ЛПР с меньшими временными и трудовыми затратами решать многокритериальные задачи выбора.

Литература

1. Ларичев О.И. Вербальный анализ решений. / Под ред. А.Б. Петровского. – М.: Наука, 2006.
2. Петровский А.Б. Теория принятия решений. – М.: Издательский центр «Академия», 2009.
3. Петровский А.Б., Ройзензон Г.В. Интерактивная процедура снижения размерности признакового пространства в задачах многокритериальной классификации // Поддержка принятия решений. Труды Института системного анализа Российской академии наук. / Под ред. А.Б.Петровского. Т.35. – М.: Изд-во ЛКИ, 2008, С.43-53.
4. Петровский А.Б., Ройзензон Г.В. Многокритериальный выбор с уменьшением размерности пространства признаков: многоэтапная технология ПАКС. // Искусственный интеллект и принятие решений, 2012 (в печати).
5. Ройзензон Г.В. Многокритериальный выбор вычислительных кластеров. // Методы поддержки принятия решений. Труды Института системного анализа Российской академии наук. / Под ред. С.В. Емельянова, А.Б. Петровского. Т.12. – М.: Едиториал УРСС, 2005, С.68-94.

AGGREGATING CRITERIA IN SELECTION OF COMPUTING CLUSTER

V.N. LOBANOV¹

A.B. PETROVSKY²

¹M.A. Kartsev Scientific Research Institute for Computer Complexes Inc., Moscow

²Institute for Systems Analysis, Russian Academy of Sciences, Moscow

e-mail: fisher_1980@mail.ru pab@isa.ru

The paper describes an approach to multiple criteria selection of computing cluster, in which a hierarchical system of complex criteria aggregating cluster characteristics is built and methods of verbal decision analysis are used.

Keywords: configuration computer cluster.