# *Automation of the distribution process of sensitive data processing in a hybrid cloud computing environment*

***Anatoly Tsaregorodtsev, Anna Zelenina***

***Abstract:***

*Use of cloud computing applications and services requires review and adaptation of existing formal models for informational telecommunication systems security. It is necessary to consider the benefits of cloud deployment models and provide the procedure for allocating process among components of cloud computing environment for achieving confidentiality and data protection.*

***Key words:***

*Security model of informational telecommunication systems, cloud computing, public cloud, private cloud, hybrid cloud, security requirements, theory of graphs, data confidentiality.*

***ACM Computing Classification System:***

*Distributed architectures, Cloud computing, Distributed systems organizing principles, Cloud based storage, Information flow control, Security requirements, Formal security models, Distributed systems security.*

## Введение

Принимая во внимание парадигму облачных вычислений, организация отказывается от прямого контроля над многими аспектами безопасности и, тем самым, создаёт беспрецедентный уровень доверия облачному провайдеру [1]. Преимущества облачных вычислений могут позволить существенно сократить сроки и издержки на разработку систем для федеральных агентств и государственных организаций. Особую актуальность принимает гибридная модель развёртывания облачных сервисов, которая подразумевает обработку критичных данных в частной среде облачных вычислений, которая находится под полным контролем организации. Но многие из функций, которые делают привлекательными облачные вычисления, могут вступать

в противоречие с традиционными моделями обеспечения информационной безопасности. При анализе сложных бизнес процессов очень трудно определить факт соответствия текущих полномочий субъекта к объекту доступа с соответствующим уровнем секретности [2].

В связи с этим возникает задача обеспечения безопасности при обработке данных в условиях среды облачных вычислений, которая требует построения гибкой ИТ-инфраструктуры на основе общедоступной среды облачных вычислений. В то же время для обработки конфиденциальной информации в ИТ-инфраструктуру необходимо включать демилитаризованные компоненты, роль которых могут выполнять частные облака, контролируемые внутренними силами организации [3].

## 1. Формализованная модель безопасности процессов обработки конфиденциальных данных

Для построения формализованной модели безопасности рабочих процессов в среде облачных вычислений в качестве основы рассматриваются существующие положения дискреционного, мандатного, ролевого управления доступом, безопасности информационных потоков, в том числе формальная модель безопасности Белла-ЛаПадула с использованием ключевых элементов теории графов.

С целью модификации предполагается расширение классической модели безопасности Белла-ЛаПадула на новые ключевые составляющие с последующей её интерпретацией для описания требований безопасности информационных рабочих процессов, протекающих в среде облачных вычислений [3]. Необходимо отметить, что использование классической теоремы является отправной точкой для описания и формализации модели безопасности рабочего процесса в среде облачных вычислений. Для преодоления ограничений, присущих модели Белла-ЛаПадула, вводятся дополнительные правила по обеспечению ИБ в рамках новой среды обработки информационных потоков.

Основываясь на основных положениях классической модели безопасности, модифицируем её применительно к рабочим процессам, протекающим в среде облачных вычислений.

– Представим облачные сервисы в виде субъектов ($T$), а данные в виде объекта ($O$).

– Определим набор действий ($A$), который субъект ($T$) может совершить с объектами ($O$). Сервис рабочего информационного процесса оперирует с данными путём применения операций чтения и записи. В результате получаем набор действий ($A$): $A = \{r, w\}$.

– Определим множество функций безопасности ($L$) для облачного сервиса.

– Определим матрицу доступов, как: $M : TxO \to A$.

Например, если сервис $t_1$ совершает операцию чтения над данными $O_1$, то запись в матрице будет выглядеть, как: $t_1 xo_1 \to r$. Если сервис $t_1$ совершает операцию записи данных $O_1$, то запись принимает вид: $t_1 xo_1 \to w$.

– Определим уровни конфиденциальности данных, как $B : SxO \to A$.

– Определим текущие уровни доступа сервиса к данным, как $C : S \to L$.

– Введём в модель новый элемент: карта текущих размещений блоков рабочего процесса $l : S + O \to L$.

При рассмотрении классического сценария обеспечения многоуровневого доступа система проходит через множество состояний. Модель принимает различные значения для матрицы полномочий, прав доступа, разрешений и размещения для каждого конкретного случая. Тем не менее, исполнение информационного потока происходит в конкретном состоянии. Обычно сервис ожидает разрешения, значение которого постоянно для всех использований этого сервиса в рабочих процессах. В то же время размещение для каждого рабочего процесса может быть определено в самом начале или определяться динамически при каждом запуске. Рассматриваемая модель является общей и не делает никаких предположений по этому поводу.

Модель Белла-ЛаПадула утверждает, что система безопасна по отношению к вышеописанной модели, если выполняются следующие условия, что

$\forall$ *субъект* $T \in T$ и $\forall$ *оьект* $o \in O$, тогда можно описать авторизацию, как:

$$B_{to} \subseteq M_{to}, \tag{1}$$

а разрешение:

$$l(t) \le c(t). \tag{2}$$

Запрет считывания информации сервисом, имеющей уровень доступа ниже уровня секретности, как:

$$\kappa \in B_{to} \Rightarrow c(t) \ge l(o) \tag{3}$$

Запрет сервису понижать уровень секретности информации, к которой он допущен как:

$$w \in B_{to} \Rightarrow l(t) \le l(o) \tag{4}$$

Для рабочего процесса, протекающего в среде облачных вычислений последствия этих условий можно выразить в виде следующих положений.

(1) Все действия, выполняемые сервисами, должны соответствовать их полномочиям.

(2) Сервисы могут функционировать только на узле облачной инфраструктуры с уровнем конфиденциальности (текущим местом размещения) меньшим или равным его разрешению (текущему доступу).

(3) Сервис не может прочитать данные, которые имеют более высокий уровень секретности, чем его собственное разрешение (текущий доступ).

(4) Сервис не может записать данные на узел с более низким уровнем секретности, чем его собственный.

При этом должны быть соблюдены правила:

$$c(t_1) \ge l(o_1), \tag{5}$$

$$l(o_1) \ge l(t_1). \tag{6}$$

Назначение уровня конфиденциальности для данных в рабочем информационном процессе аналогично присвоению уровня объекту в классической модели Белла-ЛаПадула, но присвоение уровня доступа для сервиса может оказаться менее очевидным. Дело в том, что организация может иметь различные уровни конфиденциальности данных для разных наборов сервисов. Например, внутренний сервис с высокой степенью доверия, или сервис, предоставляемый надежным поставщиком, будет иметь высокую степень защищенности для обработки конфиденциальных данных, в то время сервис как сервис, функционирующий в не доверенной зоне, может привести к утечке этих данных.

Когда специалисты по информационной безопасности организации определяют места размещения для сервисов конкретного рабочего процесса на основании

экспертной оценки, в некоторых случаях может оказаться, что компонент для размещения сервиса имеет уровень безопасности ниже, чем его текущий уровень доступа [3]. Это в свою очередь позволит сервису создавать данные с более низким уровнем конфиденциальности, что противоречит **-свойству* модели управления доступом Белла-ЛаПадула. В связи с возможным появлением такого рода ошибок, необходимо предусмотреть системный метод автоматического размещения рабочего процесса между компонентами облачной среды.

Для решения поставленной задачи рассмотрим, как классическая модель управления доступом Белла-ЛаПадула, примененная к рабочим информационным процессам, может быть расширена для учёта требований безопасности при распределении процессов между компонентами среды облачных вычислений. Поскольку облачная архитектура позволяет выбрать для размещения более одного облака на выбор, необходимо принять решение относительно того, как следует распределить данные и облачные сервисы между компонентами облачной архитектуры с разными уровнями защищенности.

На практике эксперт по безопасности или системный администратор принимает субъективное мнение о разрешённом уровне доступа рабочего процесса и уровне конфиденциальности данных на компоненте облачной среды, на котором эти данные могут быть развернуты.

Таким образом, актуальным становится вопрос расширения модели управления доступом рабочего процесса в целях обеспечения системного принятия решения о том, где сервисы и данные рабочего информационного процесса могут быть развернуты в рамках *гибридной защищенной облачной среды* для обеспечения непрерывности бизнеса и соблюдения требований безопасности.

Для этого в классическую модель добавим следующие новые переменные.

*1. Карта размещений рабочего процесса.*

Карта размещения рабочего информационного процесса должна включать в себя доступные узлы облачных вычислений, которые обозначим, как $P$:

$$l : T + O + P \to L.$$

*2. Карта присвоений сервисов и данных к облаку.*

Показатель $H$ будет использоваться для описания присвоения каждого сервиса и данных в облаке:

$$H : T + O \to P.$$

*Сформулируем новое правило*, говорящее о том, что блок рабочего процесса (сервис или данные) может быть развернут только на облаке только в том случае, если уровень конфиденциальности облака больше или равен текущему уровню доступа сервиса и уровню конфиденциальности данных. Например, для блока $x$ на облаке $y$ справедливо следующее неравенство:

$$l(p_y) \geq l(b_x). \quad (7)$$

Если в $H$ данные $O_1$ располагаются на облаке $p_a$, сервис $t_1$ на облаке $p_b$, а $O_2$ на облаке $p_c$, то должны выполняться следующие условия.

$$l(p_a) \geq l(o_1), \quad (8)$$

$$l(p_b) \geq l(t_1), \quad (9)$$

$$l(p_c) \geq l(o_2). \quad (10)$$

Это позволяет записать формулу (6) в следующем виде:

$$l(p_c) \geq l(o_2) \geq l(t_1) \quad (11)$$

## 2. Автоматизация распределения рабочего процесса в гибридной среде облачных вычислений

Используя полученную модификацию модели доступа Белла-ЛаПадула можно автоматически получить все допустимые варианты распределения рабочего процесса. Это достигается за счёт прохождения двух этапов.

Во-первых, определим новый вводный параметр для учёта разных компонентов облачной среды в виде (*P*) и примем во внимание набор сервисов (*T*), набор данных (*O*) и карту зон безопасности (*l*).

Используя правило (7), определим варианты (*V*) перехода сервисов и данных между облаками.

$$V: T + O \rightarrow P$$

$$V = \{b \rightarrow p \mid b \in T + O, p \in P, l(b) \leq l(p)\}$$

В качестве примера рассмотрим рабочий процесс, выполняющий аналитические вычисления над конфиденциальными данными в гибридной среде облачных вычислений, состоящей из двух типов облаков. На рисунке 1 изобразим граф, который включает в себя последовательное выполнение двух сервисов с разными входными и выходными данными.


Рисунок 1. Граф рабочего процесса с последовательным выполнением двух облачных сервисов

В таблице 1 приведён пример зон безопасности и разрешений с двумя уровнями конфиденциальности 0 и 1. Частное облако с повышенным уровнем безопасности определим, как $c_1$, общедоступное облако с базовым уровнем безопасности, как $c_0$. Конфиденциальные данные $o_1$ должны храниться и обрабатываться только на частном облаке, данные $o_2$, $o_3$ можно хранить и в рамках общедоступного облака. Сервис $t_1$ имеет разрешение на доступ к конфиденциальным данным только в зоне безопасности 1, но его текущее размещение определено в зоне безопасности 0, поэтому в результате его работы создаются выходные данные в зоне безопасности 0 без нарушения *- свойства теоремы Белла-ЛаПадула.

Таблица 1. Размещение и разрешения для примера анализа критически важных данных

| Вершина графа | Зона безопасности ($l$) | Разрешение ($c$) |
|---|---|---|
| $o_1$ | 1 | |
| $t_1$ | 0 | 1 |
| $o_2$ | 0 | |
| $t_2$ | 0 | 0 |
| $o_3$ | 0 | |
| $c_0$ | 0 | |
| $c_1$ | 1 | |

На основании правил переходов блоков рабочего процесса между компонентами облачной среды из разных зон безопасности покажем возможные варианты размещения каждого блока в таблице 2.

Таблица 2 – Возможные переходы блоков рабочего процесса между компонентами среды облачных вычислений

| Вершина графа | Частное облако $c_1$ | Общедоступное облако $c_0$ |
|---|---|---|
| $o_1$ | x | |
| $t_1$ | x | x |
| $o_2$ | x | x |
| $t_2$ | x | x |
| $o_3$ | x | x |

После определения всех допустимых переходов сервисов и данных в облаках, множество всех действительных развертываний рабочего процесса будет определяться по формуле:

$$W: (T + O \to P) \to \{(T + O \to P)\} = \{w \in \|V\|, \forall\, b \in T + O. \exists p \in P. b \to p \in w, |\, w \,| = |\, T + O \,|\}$$

где: $||V||$ – это показательное множество (множество всех подмножеств) $V$, $|W|$ – это количество элементов (мощность множества) $W$.

Алгоритмически $W$ вычисляется путем перекрёстного произведения присвоения блоков на облака, содержащихся в $V$. В результате получаем все возможные действительные развертывания рабочего процесса ($W$) (рис. 2). Облако, на котором

развернуты данные или сервис будем обозначать, как верхний индекс, например: $\boldsymbol{d}_{\boldsymbol{x}}^{\boldsymbol{y}}$ – это данные *x,* развернутые на облаке *y*.

$$o_1^1 \to t_1^0 \to o_2^0 \to t_2^0 \to o_3^0$$

$$o_1^1 \to t_1^0 \to o_2^0 \to t_2^0 \to o_3^1$$

$$o_1^1 \to t_1^0 \to o_2^1 \to t_2^0 \to o_3^0$$

$$o_1^1 \to t_1^0 \to o_2^1 \to t_2^0 \to o_3^1$$

$$o_1^1 \to t_1^0 \to o_2^0 \to t_2^1 \to o_3^0$$

$$o_1^1 \to t_1^0 \to o_2^0 \to t_2^1 \to o_3^1$$

$$o_1^1 \to t_1^0 \to o_2^1 \to t_2^1 \to o_3^0$$

$$o_1^1 \to t_1^0 \to o_2^1 \to t_2^1 \to o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^0 \to t_2^0 \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^0 \to t_2^0 \to o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^0 \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^0 \to o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^0 \to t_2^1 \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^0 \to t_2^1 \to o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^1 \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^1 \to o_3^1$$

Рисунок 2. Множество всех действительных размещений блоков рабочего процесса между компонентами среды облачных вычислений

## 3. Включение компонента безопасной передачи данных в гибридной среде облачных вычислений

Другим не менее важным вопросом при построении формализованной модели безопасности рабочих процессов в среде облачных вычислений является то, что применение метода в текущей постановке в чистом виде является невыполнимой задачей, так как практически реализовать его в рамках распределенной системы затруднительно по следующим причинам:

1) сервис может генерировать выходные данные непосредственно на другом облаке, без предварительного сохранения данных на компоненте, на котором он сам размещен;

2) сервис может использовать в качестве входной информации данные из другого облака без обязательного сохранения их на своём компоненте.

Для преодоления указанных ограничений, во-первых, введем понятие нового вида сервиса – *транспортного*, который будет передавать данные из одного облака в другое. Аналогом транспортного сервиса является оператор обмена в распределенной обработке запросов.

Переход будет осуществляться за счёт добавления в модель новых компонентов, функционирующих на облаке-источнике и облаке-получателе. Транспортный сервис принимает данные на одном облаке и создает её копию на другом. Все рабочие процессы из множества *W* трансформируются и включают в себя транспортные сервисы.

Введём четыре правила преобразования графа информационного потока.

$$o_j^a \rightarrow t_i^a \Rightarrow o_j^a \rightarrow t_i^a \tag{12}$$

$$o_j^a \rightarrow t_i^b \Rightarrow o_j^a \rightarrow \text{передатчик} \rightarrow o_j^b \rightarrow t_i^b \tag{13}$$

$$t_i^a \rightarrow o_j^a \Rightarrow t_i^a \rightarrow o_j^a \tag{14}$$

$$t_i^a \rightarrow o_j^b \Rightarrow t_i^a \rightarrow o_j^a \rightarrow \text{передатчик} \rightarrow o_j^b \tag{15}$$

Преобразования (12) и (14) отражают тот факт, что если оба узла размещены на одном облаке, то включения дополнительных модификаций не требуется. В преобразованиях (13) и (15) вводится новый компонент (*транспортный сервис*) для передачи данных между облаками.

Создание новых копий данных с помощью правил (13) и (15) может привести к потенциальным проблемам раскрытия копируемых данных. При применении правила (13), необходимо удостовериться, что облако *b* имеет уровень конфиденциальности достаточный для хранения копии данных $o_j$, которая наследует уровень конфиденциальности оригинала. В силу этих причин должны быть соблюдены следующие правила:

$$l(p_b) \geq l(o_j) \tag{16}$$

Аналогично, для правила (15):

$$l(p_a) \geq l(o_j) \tag{17}$$

Если происходит нарушение любого приведенного условия, то, варианты, по которым происходит распределение рабочего процесса, перестают отвечать установленным требованиям безопасности, должны быть исключены из набора *W* надёжных переходов. Докажем, что нарушение конфиденциальности может иметь место только в двух конкретных случаях.

Во-первых, рассмотрим (16). По правилу (2) получим:

$$c(t_i^b) \geq l(t_i^b) \tag{18}$$

Сначала рассмотрим случай, когда

$$c(t_i^b) = l(t_i^b), \tag{19}$$

при котором текущий уровень доступа объекта соответствует его расположению на карте зон безопасности. Правила (3) и (4) при этом изменяются на:

$$c(t_i^b) \geq l(o_j) \tag{20}$$

и

$$c(t_i^b) = l(t_i^b) \tag{21}$$

Тогда по правилу (19) получаем, что:

$$l(p_b) \geq l(t_i^b) \geq l(o_j) \Rightarrow l(p_b) \geq l(o_j), \tag{22}$$

что подтверждает выполнение правила (16) и доказывает отсутствие нарушений информационной безопасности.

Однако, если:

$$c(s_i^b) \geq l(s_i^b) \tag{23}$$

т. е. уровень безопасности сервиса выше его уровня безопасности его текущего размещения, то комбинируя результат (23) с (3) и (4) получаем:

$$l(s_i^b) < c(s_i^b) \geq l(d_j) \tag{24}$$

$$l(p_b) \leq l(s_i^b) < c(s_i^b) \tag{25}$$

что не исключает

$$l(p_b) < l(d_j) \tag{26}$$

При таком случае правило (16) нарушается, и текущее распределение рабочего информационного процесса не соответствует требованиям безопасности.

Рассмотрим данные, которые получаются в результате работы сервиса. Правило (17) может быть нарушено путем применения преобразований (15) в случае, когда сервис $t_1$ производит операцию записи данных (4) таким образом, выполняется условие $l(p_a) < l(o_j)$. Переход от $l(p_b)$ до $l(o_1)$ осуществляется по правилу (13), которое позволяет рабочему процессу создавать копию $o_1$. Переход от $l(p_b)$ до $l(o_2)$ совершается по правилу (15), которое делает возможным добавление копии $o_2$. Применение преобразований (16) и (17) к рабочему процессу из рисунка 7 не допускает использование половины возможных вариантов распределения рабочего процесса. Удаление двух дублей, созданных в результате преобразования, оставляет шесть допустимых вариантов, которые показаны на рисунке 3.

$$o_1^1 \rightarrow t_1^1 \rightarrow o_2^1 \rightarrow \text{передатчик} \rightarrow o_2^0 \rightarrow t_2^0 \rightarrow o_3^0$$

$$o_1^1 \rightarrow t_1^1 \rightarrow o_2^1 \rightarrow \text{передатчик} \rightarrow o_2^0 \rightarrow t_2^0 \rightarrow o_3^0 \rightarrow \text{передатчик} \rightarrow o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^1 \to \text{передатчик} \to o_2^0 \to \text{передатчик} \to o_2^1 \to t_2^1 \to o_3^1 \to \text{передатчик} \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^1 \to \text{передатчик} \to o_2^0 \to \text{передатчик} \to o_2^1 \to t_2^1 \to o_3^1$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^1 \to o_3^1 \to \text{передатчик} \to o_3^0$$

$$o_1^1 \to t_1^1 \to o_2^1 \to t_2^1 \to o_3^1$$

Рисунок 3. Варианты распределения рабочего процесса после добавления транспортного сервиса

Альтернативный вариант отображения распределений рабочего процесса с использованием специальных компонентов для передачи данных между облаками показан на рисунке 4.

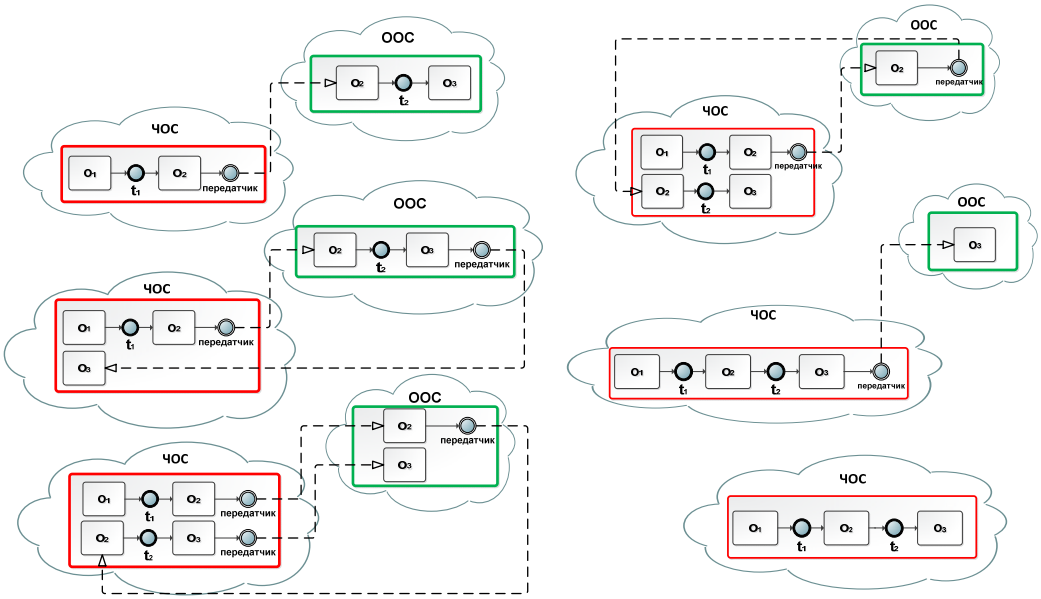


Рисунок 4. Допустимые распределения рабочих процессов в гибридной среде облачных вычислений

Для иллюстрации уровня безопасности 0 для общедоступной облачной среды (ООС), данных и сервисов использованы контуры, выделенные зеленым цветом, на уровне 1 (частной облачной среды, ЧОС) – красным цветом. Построенные диаграммы позволят специалисту ИБ получить представление о возможных вариантах развёртывания рабочих процессов в рамках гибридной облачной архитектуры. Для примера описания работы метода были использованы простые, линейные рабочие процессы, в то же время данный метод может быть применен и для ориентированных графов, независимо от сложности их структуры.

## Заключение

Описанный подход предлагается в качестве основы для автоматизации процесса разделения рабочих процессов в рамках гибридной среды облачных вычисле-

ний. Данный подход должен заменить процесс выбора администратором возможного варианта распределения процессов, который носит субъективный характер и может привести к ошибке. На смену ручному определению предлагается внедрить автоматический механизм, реализующий работу описанного метода, который определит допустимые параметры на основе строгого набора правил, а затем предложит лучший на основании стоимостной модели. Рассмотренный подход, имеет преимущества, которые могут снизить, как потенциальные нарушения безопасности, так и снизить расходы на ИТ-инфраструктуру.

## Литература

[1] ЦАРЕГОРОДЦЕВ А. В., КАЧКО А. К. Обеспечение информационной безопасности на облачной архитектуре организации // *Национальная безопасность*. – М.: Изд. «НБ Медиа», 2011. – № 5. – С. 25-34.

[2] PETER MELL, TIMOTHY GRANCE. NIST Special Publication 800-145. The NIST Definition of cloud computing. – 2011. – 3 p.

[3] ЦАРЕГОРОДЦЕВ А. В. Формализованная модель безопасности рабочих процессов информационно-телекоммуникационных систем, функционирующих на основе технологии облачных вычислений // *Нелинейный мир*. – М.: Изд-во Радиотехника, 2013. – № 9. – Т. 11. – С. 610-621.

## References

[1] TSAREGORODTSEV A. V., KACHKO A. K. Ensuring of information security in the cloud architecture of organization // National security. – Moscow: «NB Media» Publishing, 2011.

[2] PETER MELL, TIMOTHY GRANCE. NIST Special Publication 800-145. The NIST Definition of cloud computing. – 2011. – 3 p.

[3] TSAREGORODTSEV A. V. The workflow security formalized model for cloud computing informational telecommunication systems // *Nonlinear world.* – Moscow: «Radiotekhnika» Publishing, 2013.

# *Автоматизация процесса распределения обработки критичных данных в гибридной среде облачных вычислений*

***Анатолий Царегородцев, Анна Зеленина***

***Аннотация:***

*Широкое распространение и применение облачных вычислений диктует необходимость адаптации и доработки существующих моделей безопасности информационно-телекоммуникационных систем. Для достижения конфиденциальности данных необходимо рассмотреть преимущества моделей развёртывания облачных сервисов и предусмотреть процедуру распределения рабочего процесса между компонентами среды облачных вычислений.*

***Ключевые слова:***

*Модели безопасности информационно-телекоммуникационных систем, облачные вычисления, публичное облако, частное облако, гибридное облако, требования безопасности, теория графов, конфиденциальность данных.*

**Prof. Anatoly Tsaregorodtsev, D. of Sc. (Engineering)**
Moscow State Linguistic University, Russia
AVTsaregorodtsev@lingvanet.ru

**Associate Prof. Anna Zelenina, Ph.D.**
Voronezh Institute of High Technologies, Russia
snakeans@gmail.com