مجله علمی- پژوهشی "پدافند الکترونیکی و سایبری"
سال دوم، شماره ۴، زمستان ۱۳۹۳، ص ۵۴-۴۱

یک معماری جدید برون‌سپاری پایگاه‌داده در بستر رایانش ابری از منظر چرخه حیات داده

مجتبی رفیعی کرکوندی[1*]، سید کامیار ایزدی[2]، ابوالفضل خوش‌صفت[3]
۱ و ۳- کارشناس ارشد دانشگاه شهید بهشتی تهران ۲- استادیار دانشگاه شهید بهشتی تهران

(دریافت: ۹۳/۳/٤؛ پذیرش: ۹٤/۲/۲۱)

چکیده

رشد روزافزون حجم اطلاعات و نداشتن امکانات کافی محاسباتی و ذخیره‌سازی، سازمان‌ها را با چالش‌های مدیریتی متنوعی روبه‌رو کرده است. وجود این چالش‌ها از یک سو و گسترش روزافزون سرویس‌های ذخیره‌سازی از سوی دیگر، سازمان‌ها را بر آن داشته تا نگهداری و مدیریت داده‌ها و پرس‌وجوهای خود را به ارائه‌دهندگان خدمات فضای ذخیره‌سازی ابری واگذار نمایند. از آنجا که داده‌های سازمان درصورت استفاده از چنین سرویس‌هایی، در قالب برون‌سپاری خارج از محیط سازمان نگهداری می‌شود و داده‌ها تحت نظارت و کنترل مستقیم مالک داده نمی‌باشد، نگرانی‌های امنیتی به‌وجود می‌آید. برای مقابله با این نگرانی‌های امنیتی راه‌حل‌های بسیاری ارائه گردیده است اما بیشتر این راه‌حل‌ها بر روی جنبه خاصی از چرخه حیات داده مانند فازهای ذخیره‌سازی و استفاده، تاکید داشته‌اند. آشنایی با چرخه حیات داده و چالش‌ها و فرصت‌های فراروی سازمان‌ها می‌تواند کمک شایانی در ارائه راهکارهای مناسب برای بهبود این فناوری جدید به همراه داشته باشد. در این مقاله ابتدا به بررسی چالش‌ها و فرصت‌های فراروی سازمان‌ها می‌پردازیم و در ادامه، معماری جدیدی برای برون‌سپاری پایگاه داده با توجه به چرخه حیات داده ارائه شود.

واژه‌های کلیدی: رایانش ابری، برون‌سپاری، چرخه حیات داده، امنیت برون‌سپاری، طبقه‌بندی اطلاعات، معماری برون‌سپاری

۱- مقدمه

با پیشرفت فناوری‌های نو ظهور، حجم اطلاعات سازمان ها روزبه‌روز در حال افزایش است. رشد روزافزون اطلاعات، کنترل و مدیریت داده‌ها را بیش از پیش پیچیده‌تر ساخته و هزینه‌های مربوط به آن را افزایش داده است. سازمان‌ها برای رویارویی با این مسئله می‌توانند تمهیداتی چون افزودن منابع ذخیره‌سازی و به‌کار گیری افراد اجرایی بیشتر و یا واگذاری مدیریت داده‌های خود به یک کارگزار خارجی[1] را در پیش گیرند. راه‌حل اول تا حدودی کنترل و مدیریت داده‌ها را آسان می‌سازد اما به سبب به‌کارگیری منابع ذخیره‌سازی و نیروی انسانی بیشتر، سبب افزایش هزینه‌های سازمان می‌گردد. از این‌رو سازمان‌ها تمایل بیشتری به برون‌سپاری اطلاعات داشته و چنین سرویس‌هایی روزبه‌روز در حال فراگیر شدن است.

برای اولین بار شرکت EDS[2] در دهه ۱۹۶۰ به اجرای خدمات گوناگون پردازش داده، پرداخت. مبلغ قرارداد بین EDS و مشتریانش به مقدار چشم‌گیری پایین بود. در دهه ۱۹۸۰، برون‌سپاری فناوری اطلاعات به یک کسب وکار سودمند تبدیل شد و تعداد عرضه کننده‌گان و سطح تخصص آنها افزایش قابل توجهی یافت. در این زمان، عرضه‌کنندگان، سرویس‌های خدماتی با کیفیت بالا و قیمت پایین تامین می‌کردند و این امر راه را برای برون‌سپاری کامل در سال ۱۹۹۰ هموار کرد. در آن زمان با تصمیم شرکت کوداک مبنی بر برون‌سپاری تمام فعالیت‌های فناوری اطلاعات خود به شرکت IBM[3]، دیگر سازمان‌ها اعم از دولتی و خصوصی به استفاده از این فناوری ترغیب شدند. امروزه شرکت مایکروسافت تقریبا همه بخش‌های خود- از تولید نرم‌افزارهای کامپیوتری گرفته تا توزیع محصولات- را برون‌سپاری می‌کند و خود تنها بر روی

* رایانامه نویسنده مسئول: student.rafiee@gmail.com
1- Service Provider
2- Electronic Data Systems
3- International Business Machines

پردرآمدترین بخش یعنی نوشتن کد نرم‌افزار متمرکز شده است [۱].

با ظهور میزبانی محاسبات ابری و سرویس‌های ذخیره‌سازی، سرویس‌های برون‌سپاری داده در بستر رایانش ابری، شکل تازه‌ای به خود گرفت. با پیشرفت فناوری به‌عنوان سرویس، کاربران فقط به منابعی که برای انجام کارشان نیاز دارند دسترسی پیدا می‌کنند. بنابراین نیازی به پرداخت هزینه برای منابع مصرف نشده شبیه به روش سنتی ندارند. رایانش ابری علاوه بر این‌که باعث صرفه‌جویی در هزینه‌های سازمان می‌گردد به کاربران نیز امکان دسترسی به آخرین نرم‌افزارها و زیرساخت‌ها به‌منظور نوآوری در کسب‌وکار را ارائه می‌دهد.

با وجود این مزایا، به‌علت مدیریت داده توسط یک سازمان خارجی و عدم کنترل مستقیم مالک داده بر آن، چالش‌های امنیتی جدیدی در این زمینه مطرح می‌گردد. برای رویارویی با این چالش‌ها، راه‌حل‌های امنیتی زیادی ارائه گردیده است اما بیشتر این راه‌حل‌ها تنها بر جنبه خاصی از چرخه حیات داده مثل ذخیره‌سازی و استفاده متمرکز شده‌اند. در این مقاله سعی داریم چرخه امنیتی برون‌سپاری داده در بستر رایانش ابری را بررسی کرده و بر طبق آن سناریوی برون‌سپاری جدیدی را ارائه نماییم.

از آن‌جا که ویژگی‌ها و خصایص محیط برون‌سپاری پایگاه داده یکی از مهم‌ترین پارامترهای تاثیرگذار بر کمیت چالش‌های امنیتی بوده و مدل‌های تهدید متنوعی را معرفی می‌نماید، لذا آشنایی با آن ضروری است و به همین خاطر در بخش دوم به مباحث پیرامون رایانش ابری می‌پردازیم. در بخش سوم، مبانی امنیتی داده را مورد بحث و بررسی قرار می‌دهیم. در بخش چهارم، چرخه حیات داده را معرفی می‌کنیم. در بخش پنجم مباحث برون‌سپاری داده در بستر رایانش ابری و چالش‌های موجود در آن را بررسی خواهیم کرد. در بخش ششم، معماری‌های موجود برای برون‌سپاری را مطرح می‌کنیم. در بخش هفتم، اجزاء معماری برون‌سپاری را معرفی و در بخش هشتم معماری جدیدی برای برون‌سپاری ارائه می‌نماییم. در بخش نهم نیز به مقایسه معماری‌های موجود و معماری ارایه شده می‌پردازیم. در آخر نیز نتایج حاصل از این پژوهش آورده شده است.

۲- رایانش ابری

نتایج مطالعات اخیر شرکت‌ها نشان می‌دهد به طور متوسط ۱۸٪ کاهش در بودجه فناوری اطلاعات و ۱۶٪ کاهش در هزینه‌های منابع را از طریق بهره‌مندی از رایانش ابری بدست‌آورده‌اند [۲]. بسیاری از متخصصان تجاری و دانشگاهی سعی بر ارائه تعریف دقیقی از رایانش ابری و خصیصه‌های منحصر به فرد آن داشته‌اند. موسسه ملی فناوری و استاندارد آمریکا رایانش ابری را مدلی برای فراهم کردن دسترسی آسان، بر اساس تقاضای کاربر از طریق شبکه به مجموعه‌ای از منابع رایانشی قابل پیکره‌بندی (مثل شبکه، سرور، فضاهای ذخیره سازی، برنامه‌های کاربردی و ...) می‌داند و از جمله ویژگی‌های بارز آن را دسترسی سریع با کمترین نیاز به مدیریت منابع و یا نیاز به دخالت مستقیم کارگزار برمی‌شمارد [۳]. بی‌شک آنچه که رایانش ابری را به این درجه از محبوبیت رسانده، وجود خصیصه‌ها و ویژگی‌هایی چون مقیاس‌پذیری و قابلیت ارتجاعی، دسترس‌پذیری و قابلیت اعتماد، قابلیت اداره و برهم‌کنش‌پذیری، قابلیت دسترسی و جابه‌جایی، کارایی و بهینه‌سازی می‌باشد [۴].

۲-۱- مزایا و چالش‌های رایانش ابری

منافع رایانش ابری را می‌توان از دو دیدگاه سازمان و کاربران نهایی دسته‌بندی نمود [۵]. از دید سازمان، رایانش ابری سبب کاهش سرمایه‌گذاری اولیه، کاهش هزینه‌های سرمایه ای، بهبود تخصص صنعتی و بهبود بهره‌برداری از منابع در سازمان می‌گردد و از دید کاربران نهایی نیز منافعی چون کاهش قدرت محاسباتی محلی، کاهش قدرت ذخیره‌سازی محلی و تنوع Thin Client[1] در زندگی روزمره را به‌همراه دارد.

با توجه به این‌که رایانش ابری، انبوه متمرکزی از منابع را برای ما مهیا می‌سازد این نکته قابل تامل است که انبوهی متمرکز از خطرات و ریسک‌ها را نیز دربر می‌گیرد چرا که اگر روزنه‌ای از نقص و خطا در بخشی از ابر پدیدار گردد، آسیب بسیار بزرگ و غیرقابل جبرانی را برجای خواهد گذاشت. ابر مانند یک جعبه سیاه بزرگ عمل می‌کند و درون خود را از دید کاربران پنهان می‌سازد، بنابراین مالکان داده[2] هیچ‌گونه کنترل یا ارائه ایده‌ای از خود برای درون ابر نخواهند داشت. با این وجود هرچند ارایه‌دهندگان ابر صادق باشند، اما این خطر وجود دارد که مدیران مخرب سیستم به محرمانگی و جامعیت داده‌ها آسیب‌هایی را وارد نمایند.

۲-۲- مدل‌های استقرار رایانش ابری

موسسه ملی فناوری و استاندارد آمریکا مدل‌های استقرار رایانش ابری را در چهار رده تقسیم‌بندی کرده است [۳]. تفاوت اصلی

۱- به کامپیوتر یا برنامه کامپیوتری اطلاق می‌شود که برای تحقق وظایف محاسباتی خود به کامپیوترهای دیگر وابسته است.

2- Data Owner

مدل‌های استقرا رایانش ابری در دامنه پوششی آن‌ها و همچنین دست‌یابی سرویس‌های ارائه‌شده در ابر می‌باشد.

ابر عمومی: در این مدل، زیرساخت ابر برای استفاده عموم فراهم گردیده است و ممکن است توسط یک سازمان دولتی، دانشگاهی و یا تجاری مالکیت و مدیریت گردد. این نوع ابر توصیف‌کننده رایانش ابری در معنای اصلی و سنتی‌اش می‌باشد و بیشترین سطح چالش‌های امنیتی نیز در این مدل مطرح می‌گردد.

ابر خصوصی: در این مدل، زیرساخت ابر برای استفاده یک سازمان تکی که شامل چندین مشتری است، تهیه می‌گردد و ممکن است توسط خود سازمان، سازمان خارجی و یا ترکیبی از آن‌ها مدیریت و مالکیت گردد. مزیت اصلی ابر خصوصی، امنیت بیشتر آن به‌دلیل استقرار تجهیزات در درون سازمان و عدم برقراری ارتباط با دنیای خارج می‌باشد.

ابر گروهی: در این مدل، زیرساخت ابر برای استفاده انحصاری چندین سازمان که مشابهت‌های یکسانی از لحاظ ماموریت، سیاست‌های کاری و نیازمندی‌های امنیتی دارند فراهم می‌گردد و ممکن است توسط یک یا چند سازمان بهره‌گیرنده از این مدل، سازمان خارجی و یا ترکیبی از آنها مدیریت و مالکیت گردد.

ابر آمیخته: در این مدل، زیرساخت ابر ترکیبی از دو یا چند ابر متمایز عمومی، خصوصی و یا گروهی می‌باشد. این نوع ابر گزینه مناسبی برای بیشتر موسسات تجاری به حساب می‌آید.

۲-۳- مدل‌های سرویس رایانش ابری

زیرساخت به عنوان سرویس : قابلیت ارائه‌شده برای مشتریان، به‌منظور تامین منابع پردازشی، ذخیره‌سازی، شبکه و سایر منابع محاسباتی اساسی می‌باشد. در این سرویس مشتریان قادر به استقرار و اجرای نرم‌افزارهای دلخواه شبیه سیستم‌عامل یا دیگر برنامه‌های کاربردی می‌باشند. مشتریان کنترل و مدیریتی بر روی زیرساخت ابر نداشته و آنچه می‌توانند مدیریت نمایند فضای ذخیره‌سازی، سیستم‌عامل، برنامه‌های کاربردی مستقر در ابر و تا حدودی کنترل بر اجزای شبکه است. نمونه‌هایی از این سرویس عبارت اند از: Amazon EC2, Eucalyputs, OpenNebula [۶].

بستر به‌عنوان سرویس: قابلیت ارائه‌شده، برای مشتریانی است که بر روی زیرساخت‌های ابر استقرار یافته و برنامه‌های کاربردی ابر را تغذیه می‌نمایند. مشتریان، زیرساخت‌های ابر چون شبکه، سرور، سیستم‌عامل و منابع ذخیره‌سازی را کنترل و مدیریت نمی‌کنند بلکه آنچه را که مدیریت و کنترل می‌کنند برنامه‌های کاربردی مستقر بر روی زیرساخت ابر و پیکره‌بندی محیط میزبانی برنامه است. نمونه‌هایی از این سرویس عبارت‌اند از: Microsoft Windows Azure, Google App Engine, Hadoop [۶].

نرم‌افزار به‌عنوان سرویس: قابلیت تامین‌شده برای مشتریان است که اجرای برنامه‌های کاربردی بر روی زیرساخت ابر را ممکن می‌سازد. برنامه‌های کاربردی از طریق یک رابط Thin Client در دسترس مشتریان قرار می‌گیرند. مشتری زیرساخت‌های ابر چون شبکه، سرور، منبع ذخیره‌سازی و سیستم عامل را مدیریت و کنترل نمی‌کند و حتی در مورد تنظیمات برنامه‌های کاربردی نیز تنها تعداد محدودی از کاربران مجاز به پیکره‌بندی آن‌ها می‌باشند و کاربر تنها یک استفاده‌کننده محض از سرویس است. نمونه‌هایی از این سرویس عبارت‌اند از: Gmail, Google Docs, Google sites [۶].

۳- مبانی امنیتی داده

۳-۱- محرمانگی داده

محرمانگی داده به‌عنوان یک نگرانی اصلی در سیستم‌های پایگاه داده محسوب می‌شود. سازمان‌های مختلف تعاریف متنوعی برای محرمانگی ارائه کرده‌اند. بر طبق تعریف ارائه‌شده توسط ISO/IEC [1] محرمانگی به معنای حفظ و حراست اطلاعات به‌منظور جلوگیری از افشا و یا دسترس قرار دادن اطلاعات می‌باشد. هامر و اشنایدر محرمانگی را این‌گونه تعریف کرده‌اند [۷].

۱- اطمینان از این‌که داده‌ها تنها برای افراد مجاز به‌منظور خواندن، ثبت، تغییر و یا حذف فیزیکی، مورد دست‌یابی قرار می‌گیرند.

۲- افراد تنها می‌توانند تا حد مجاز در داده‌ها دخل و تصرف داشته باشند.

به‌طور کلی محرمانه بودن داده‌ها به معنای حفاظت و جلوگیری از افشای اطلاعات خصوصی موجود در رسانه‌های ذخیره‌سازی و یا جریان شبکه است. مفهوم محرمانگی داده با توجه به طبقه بندی

1- International Organization for Standardization / International Electrotechnical Commission

اطلاعات و نوع سازمان می‌تواند متفاوت باشد. به‌عنوان نمونه می‌توان گفت که یک قطعه‌ی خاص از داده می‌تواند در یک سازمان محرمانه تلقی شود و این درحالی است که همین قطعه می‌تواند در سازمان دیگر این‌گونه نباشد.

۳-۲- حریم خصوصی

امروزه رسانه‌های نوین در حال استفاده از اطلاعات افراد هستند و این در حالی است که خود افراد نسبت به این موضوع بی‌خبرند. با پیشرفت فناوری در بخش الکترونیک تجاوز به حریم‌خصوصی نسبت به گذشته رشد چشم‌گیری پیدا کرده است. از دید کاربران، حفظ حریم خصوصی به معنای حق کاربران برای حفاظت از داده‌های شخصی حساس خود در مقابل تقلب، سرقت هویت و یا استفاده غیر مجاز می‌باشد. حفظ حریم خصوصی در حقیقت بیانگر جلوگیری از مشاهده، کنترل، استفاده و توزیع اطلاعات محرمانه افراد توسط اشخاص دیگر می‌باشد. حفظ حریم خصوصی از آنجایی که تمرکز روی داده‌های محافظت‌شده شخصی افراد دارد متفاوت از محرمانگی است. در واقع می‌توان گفت که حفظ حریم خصوصی یک حالت خاص محرمانگی است [۸]. به عنوان مثال در پایگاه داده یک بیمارستان، حفظ حریم خصوصی به معنای حفاظت از اطلاعات شخصی افراد از مشاهده یا تغییر توسط بخش‌های غیر مجاز می‌باشد.

۳-۳- جامعیت داده

جامعیت داده بیانگر اطمینان از عدم تغییر داده توسط بخش‌های غیر مجاز می‌باشد. جامعیت شامل حفاظت اطلاعات در طول ذخیره‌سازی، انتقال، دستکاری و تهیه نسخه پشتیبان می‌باشد. نگهداری جامعیت داده برای حفظ حریم خصوصی، امنیت و قابلیت اطمینان داده ضروری می‌باشد.

۴- چرخه حیات داده

امروزه امنیت داده به یکی از مسائل مهم رایانش ابری تبدیل شده است. در این باره تعداد زیادی راه‌حل‌های امنیتی ارائه گردیده است اما بیشتر آنها تنها بر بخشی از چرخه حیات داده متمرکز بوده‌اند و از آنجا که آسیب‌پذیری امنیتی در هر بخش از چرخه حیات داده می‌تواند سبب اختلال در کل امنیت داده گردد، می‌بایست به کلیه مراحل این چرخه توجه ویژه شود. در یک دسته‌بندی کلی می‌توان کل چرخه حیات داده را در سه وضعیت فعال، نیمه فعال و بایگانی خلاصه کرد. اتحادیه امنیت رایانش ابری[1] مفهوم چرخه حیات داده را در شش فرآیند تولید، ذخیره‌سازی، استفاده، اشتراک، آرشیو و انهدام ارائه داده است [۹].

شکل (۱). چرخه حیات داده

ایجاد: این مرحله بیانگر رویه ایجاد داده می‌باشد. داده ممکن است سمت کلاینت[2] و یا حتی سمت سرور[3] تولید گردد.

ذخیره‌سازی: این مرحله بیانگر بارگیری و ذخیره‌سازی داده در محیط ابر می‌باشد و برای حصول اطمینان در ابر ممکن است داده در چندین گره[4] ذخیره شود.

استفاده و اشتراک: این مرحله بیانگر استخراج داده‌ها از ابر می‌باشد. داده‌ها می‌توانند توسط مالک داده مورد استفاده قرار گیرند و یا بین چندین شخص به اشتراک گذاشته شوند.

بایگانی: داده‌هایی که به‌طور موقت مورد استفاده قرار نمی‌گیرند، توسط سرور به مکان دیگری در ابر انتقال داده می‌شوند.

انهدام: در این مرحله، داده‌ها مطابق نظر مالک داده حذف می‌شوند و در این شرایط برای اطمینان از عدم فاش شدن اطلاعات،

1- Cloud Security Alliance
2- Client
3- Server
4- Node

سرور می‌بایست داده‌های حذفی را در سمت خود نامعتبر و غیرقابل بازیافت نماید.

۵- برون‌سپاری داده

در پی افزایش کاربردهای فناوری اطلاعات در سازمان‌ها، هزینه‌های ناشی از مدیریت و نگهداری داده‌ها نیز افزایش یافته است. رشد سریع فناوری اطلاعات و ارتباطات منجر به رشد ۵۲ درصدی هزینه‌های مدیریت و ذخیره‌سازی اطلاعات شده است [۱۰]. این هزینه‌ها شامل سخت‌افزار، نرم‌افزار و نیروی‌انسانی متخصص بوده و بسیاری از سازمان‌ها و ادارات از عهده پرداخت این هزینه‌ها برنمی‌آیند. این مسئله منجر به گسترش ایده برون‌سپاری داده به‌منظور کاهش هزینه‌های ذخیره‌سازی و مدیریت داده‌ها شده است. برون‌سپاری به واگذاری تمام یا بخشی از وظایف و فعالیت‌های سازمان به سازمان خارجی ارائه‌دهنده خدمات اطلاق می‌گردد. برون‌سپاری مزایای قابل توجهی چون کاهش هزینه‌های مدیریتی، دسترس‌پذیری بالا و کارآمدی بیشتر را به همراه دارد. از مهم‌ترین انگیزه‌های سازمان به‌منظور برون‌سپاری و مدیریت داده‌ها می‌توان به کمبود زیرساخت‌های سخت‌افزاری و نرم‌افزاری مناسب، کمبود نیروی انسانی متخصص، پیچیده شدن ماموریت‌ها و تمایل سازمان‌ها برای تمرکز روی اهداف اصلی اشاره کرد.

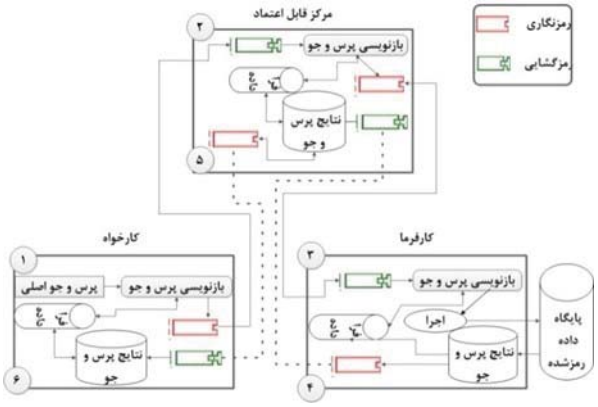
امروزه سازمان‌ها به اهمیت فناوری اطلاعات که اطلاعات در آنها به‌عنوان یک ابزار ضروری و مهم برای پیشرفت کسب وکار محسوب می‌شود، پی برده‌اند. افزایش پیچیدگی نگهداری، توسعه و مدیریت داخلی سیستم‌های فناوری اطلاعات و هزینه‌های راه‌اندازی، مشکلات قابل توجهی را برای سازمان‌ها در زمینه‌های زیرساختی و نیروی انسانی ایجاد کرده است که با برون‌سپاری این مشکلات تا حد قابل قبولی برطرف خواهد شد. اما با این وجود هنوز نیز چالش‌ها و تهدیدهایی در این زمینه وجود دارد. پیرامون طبقه‌بندی چالش‌ها و تهدیدها ی موجود در برون‌سپاری پایگاه‌داده تحقیقات بسیاری صورت پذیرفته است [۱۱]. به‌طور کلی چالش‌ها و موانع برون‌سپاری پایگاه داده را می‌توان در دو رده چالش‌های فیزیکی و چالش‌های منطقی تقسیم‌بندی نمود. این چالش‌ها می‌تواند در هر دو سمت سرور و کلاینت وجود داشته باشد. چالش‌های فیزیکی بواسطه مسائلی چون خطاهای سخت‌افزاری و نرم‌افزاری، سرقت رسانه داده، بلایای طبیعی و ... پدید می‌آیند و دسترس‌پذیری و محرمانگی داده را تحت تاثیر قرار می‌دهند.

چالش‌های منطقی که در برون‌سپاری پایگاه داده مطرح می‌گردد عبارت‌اند از:

- **احراز اصالت :** این چالش در دو سطح احراز اصالت کاربر و داده مطرح می‌گردد. در سطح احراز اصالت کاربران، ارائه‌دهنده سرویس در واقع می‌خواهد اطمینان پیدا کند که داده یا پرس‌وجوی دریافتی، از سمت کاربران مجاز صورت گرفته است یا خیر. در سطح احراز اصالت داده، ارائه‌دهنده سرویس می‌خواهد از منشا تولید داده و زمان تولید آن اطمینان حاصل نماید. برای رویارویی با این دسته از چالش‌ها می‌توان از روش‌های موجود برای امضای دیجیتال بهره‌مند گردید.
- **محرمانگی داده:** در این سطح باید داده‌های برون‌سپاری شده برای ارائه دهندگان خدمات و کاربران غیرمجاز نامفهوم باشد. برای رویارویی با این سطح از چالش ها می‌توان از الگوریتم‌های رمزنگاری بهره‌مند گردید.
- **حریم خصوصی:** این چالش نیز در دو سطح حریم خصوصی کاربر و داده مطرح می‌گردد. در سطح حریم خصوصی کاربر، ارائه‌دهنده خدمات نباید درباره پرس‌وجوهای کاربر و نتایج بازگشتی حاصل از آن اطلاعی پیدا کند و در سطح حفظ حریم خصوصی داده نیز کاربران نباید اطلاعاتی بیش از آنچه که از سرور درخواست نموده‌اند دریافت نمایند. برای رویارویی با این چالش می‌توان از رمزنگاری و اعمال کنترل دسترسی بهره‌مند گردید.
- **اطمینان از پرس‌وجو:** این سطح، توانایی کاربر برای بررسی صحت، تمامیت و تازگی پرس‌وجو را شامل می شود. صحت پرس‌وجو در واقع بیانگر این است که نتایج پرس‌وجو همه شرایط پرس‌وجو را پوشش داده و همچنین داده‌های بازگشتی همان داده‌های مالک، بدون هیچگونه دست‌کاری در آن باشد. تمامیت پرس‌وجو نیز بیانگر اطمینان از این است که نتایج بازگشتی تمام مجموعه داده‌های درخواستی مشتری را دربر گرفته و حاوی اطلاعات اضافه‌تر یا کمتر نباشد. همچنین تازه‌گی پرس‌وجو نیز بیانگر این مفهوم است که نتایج پرس‌وجو باید حاوی آخرین تغییرات اعمال‌شده در پایگاه داده باشد.

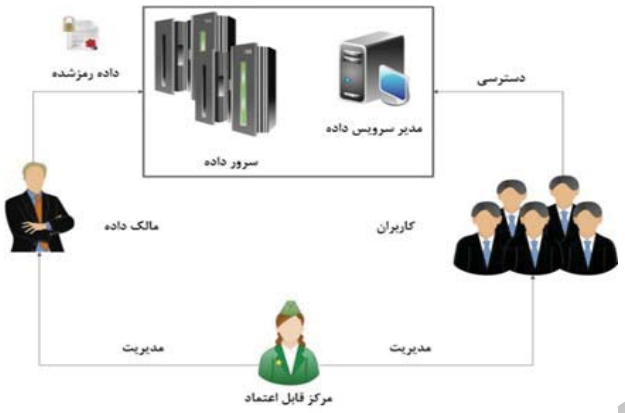
۶. بررسی معماری‌های موجود

کادهم و همکاران[۱۲] یک معماری جدید برون‌سپاری پایگاه داده بر پایه حفظ محرمانگی پایگاه ارائه داده‌اند. نقش‌های موجود در این معماری عبارت‌اند از : کلاینت، سرور و شخص سوم قابل اعتماد. در این معماری، پرس‌وجوهای درخواستی کلاینت با توجه به فراداده‌های موجود در این بخش به پرس‌وجوی جدیدی برای اجرا در سمت سرور بازنویسی می‌شود، سپس این پرس وجو به شخص سوم قابل اعتماد ارسال گردیده و در این بخش نیز با توجه به سیاست‌ها و کنترل‌های موجودی، بازنویسی دیگری بر روی پرس‌وجو صورت می‌پذیرد و نهایتا پرس‌وجو تولیدی به سرور ارسال می‌گردد. در سمت سرور نیز بر روی پرس‌وجوی دریافتی با توجه به فراداده، تمهیدات لازم اتخاذ شده و پرس‌وجوی نهایی بر روی پایگاه داده رمزشده اجرا می‌گردد. داده‌های حاصل از اجرای پرس‌وجو به شخص سوم قابل اعتماد و سپس به کلاینت بازگردانده شده و فرآیند پرس‌وجو به اتمام می‌رسد.



شکل (۲). معماری برون‌سپاری کادهم و همکاران

هور و همکاران [۱۳] معماری برون‌سپاری جدیدی را بر پایه کنترل دسترسی ارائه دادند. این معماری شامل مولفه‌هایی چون تامین کننده خدمات، مالک داده، کاربر و شخص قابل اعتماد صادر کننده اختیارات (TA)[1] می‌باشد. TA یک عنصر کلیدی در این معماری محسوب شده و پارامترهای عمومی و خصوصی مورد نیاز در سیستم را تولید می‌نماید. TA مسئول صدور، لغو و به‌روزرسانی کلیدهای کاربران بوده و با توجه به ویژگی‌های کاربران حقوق دسترسی متفاوتی را به آنها ارائه می‌نماید. مالک داده، مسئولیت تعریف سیاست‌های دسترسی، کنترل‌ها و تمهیدات امنیتی مورد نیاز قبل از برون‌سپاری پایگاه داده را بر عهده دارد. کاربر در واقع عاملی است که می‌خواهد داده برون‌سپاری‌شده را مورد دست‌یابی قرار دهد. اگر کاربری مجموعه تمامی خصیصه‌های مورد نیاز برای ارضا کردن سیاست‌های دسترسی داده مورد نظر را داشته باشد، قادر خواهد بود تا به آن‌ها دسترسی داشته و داده مورد نظر را رمزگشایی نماید. تامین کننده خدمات در واقع سرویس‌های برون‌سپاری را در اختیار کلاینت‌ها قرار می دهد و مسئولیت مدیریت و کنترل داده‌های برون‌سپاری‌شده را بر عهده دارد. در اغلب سناریوها فرض می‌شود که سرور از لحاظ نگهداری اطلاعات و پاسخ به پرس‌وجوها صادق بوده و نسبت به کسب اطلاعات در مورد داده‌های اصلی کنجکاو است [۱۶-۱۴].



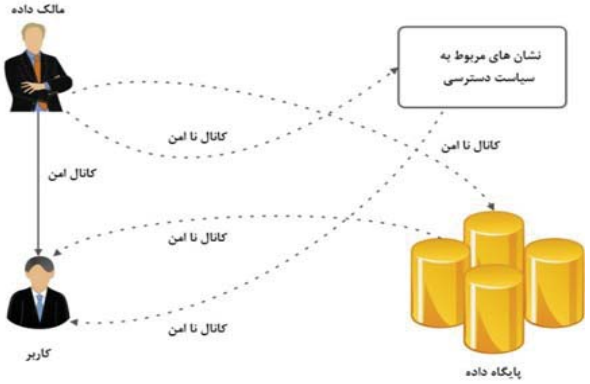
شکل (۳). معماری برون‌سپاری هور و همکاران

وایرمکاتی و همکاران [۱۷] معماری دیگری بر پایه حفظ محرمانگی و کنترل دسترسی ارائه کرده‌اند. در معماری ارائه‌شده دو نوع پایگاه داده به قرار زیر وجود دارد :

- پایگاه داده رمزشده حاوی داده‌های برون‌سپاری‌شده مشتریان.
- پایگاه داده محتوی سیاست‌های کنترل دسترسی به‌منظور مدیریت دستیابی اطلاعات سمت سرور.

پایگاه داده دوم در واقع به‌منظور تضمین دسترسی کاربران مجاز به پایگاه داده رمزشده مورد استفاده قرار می‌گیرد. ایده پایه لحاظ‌شده در این معماری آن است که ابتدا یک کلید محرمانه توسط مالک برای کاربر تولید شده و سپس کاربر با استفاده از این کلید و مجموعه‌ای از نشان‌ها می‌تواند کلیدهای دیگری را نیز بدست آورد. فهرست نشان‌ها تنها برای کاربرانی که از قبل به کلید محرمانه دسترسی دارند، قابل دستیابی می‌باشد. مسئولیت فهرست

1- Trusted Authority



شکل (۴). معماری برون‌سپاری وایرمکاتی و همکاران

نشان‌ها[1] بر عهده مالک داده بوده و می‌بایست ارایه صحیحی از سیاست‌های دسترسی که مالک داده می‌خواهد بر روی داده‌های خود پیاده‌سازی کند، داشته باشد. کاربران مجاز، به‌منظور دسترسی به منابع داده مورد نظرخود می‌بایست از فهرست نشان‌ها، تمامی نشان‌های مورد نیاز را بازیابی نمایند.

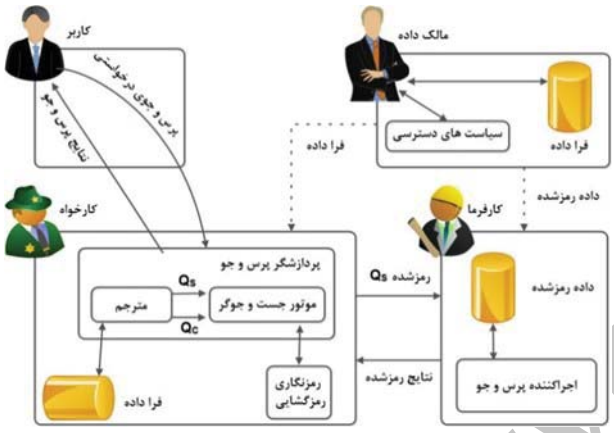
فورستی [۱۸] یک معماری کلی با توجه به معماری‌های موجود برای برون‌سپاری پایگاه داده ارائه داده است که شامل چهار مولفه زیر می‌باشد:

- **مالک داده:** مالک وظیفه تولید و برون سپاری داده به‌منظور قابل دسترس نمودن داده‌ها تحت یک آزاد سازی کنترل‌شده برای کاربران را بر عهده دارد.
- **کاربر:** درخواست‌های خود را در قالب پرس‌وجو به سیستم تحویل می‌دهد.
- **کلاینت:** پرس‌وجوی ارائه‌شده توسط کاربر را به یک پرس‌وجوی معادل برای اجرا شدن روی داده رمزنگاری‌شده سمت سرور تبدیل می‌کند.
- **سرور:** داده‌های رمزشده توسط یک یا چند مالک را دریافت و ذخیره‌سازی کرده و برای کاربران مجاز قابل دسترس می‌نماید.
- به‌طور خلاصه مهم‌ترین گام‌های مورد نیاز برای اجرای یک پرس‌وجوی مطرح‌شده از طرف کاربر در این معماری نیز به قرار زیر است:

گام اول، تحویل پرس‌وجوی کاربر(Q) به کلاینت می‌باشد. در این گام کاربر نیازی به آگاهی از برون‌سپاری داده‌ها ندارد.

گام دوم، کلاینت پرس‌وجوی کاربر را به دو پرس‌وجوی Q_S و Q_C نگاشت می‌کند و سپس پرس‌وجوی Q_S را به سرور می‌فرستد.

پرس‌وجوی Q_S روی داده رمزشده عمل می‌کند و پرس وجوی اضافی Q_C روی نتایج حاصل از Q_S اعمال‌شده و نتایج برگشتی را پالایش نموده و جواب خالص را تولید می‌نماید.



شکل (۵). معماری برون‌سپاری فورستی

گام سوم، سرور، پرس‌وجوی Q_S دریافتی را بر روی داده رمزشده اجرا نموده و نتایج حاصل از آن را که مجموعه ای از داده‌های رمزشده است به سمت کلاینت ارسال می‌نماید.

گام چهارم، کلاینت نتایج برگشتی از سرور را رمزگشایی نموده و در نهایت با اجرای Q_C بر روی آن، داده های خالص (داده‌های هدف) را به‌دست‌آورده و به کاربر می‌فرستد.

کلیه معماری‌های بیان‌شده در بالا تنها بر بخشی از چرخه حیات داده متمرکز هستند و از آن‌جا که آسیب‌پذیری امنیتی در هر بخش از چرخه حیات داده می‌تواند سبب اختلال در امنیت داده گردد می‌بایست به کلیه مراحل این چرخه توجه ویژه شود. طراحی یک معماری خوب برای سازمان‌ها این امکان را فراهم خواهد کرد تا

1- Token

خط مشی‌های مدیریتی، نیازهای کاربران و استراتژی های درونی سازمان با کمترین نیاز به تغییرات اعمال گردد.

۷- اجزاء معماری برون‌سپاری پایگاه داده

۷-۱- پایگاه داده

سطوح ریزدانگی در رمزنگاری پایگاه داده می‌تواند بر اساس تعدد دست‌یابی و نوع داده‌های مورد استفاده، متفاوت باشد.

- **رابطه**[1]: هر رابطه در یک پایگاه داده پس از اعمال تابع رمزنگاری به عنوان یک مقدار داده‌ای واحد در پایگاه داده رمزشده در نظر گرفته می‌شود . بنابراین تاپل‌ها و خصیصه‌ها در داده برون‌سپاری‌شده غیرقابل تمیزند و نمی‌توانند در یک پرس‌وجو روی پایگاه داده، لحاظ گردند.
- **خصیصه**[2]: هر ستون(خصیصه) در پایگاه داده به‌عنوان یک ارزش تکی در رابطه رمزشده تبدیل می‌شود.
- **تاپل**[3]: هر سطر از پایگاه داده به‌عنوان یک ارزش تکی در رابطه رمزشده تبدیل می‌شود.
- **عنصر**[4]: هر سلول در پایگاه داده به‌عنوان یک ارزش واحد در رابطه رمزشده تبدیل می‌شود.

با توجه به تعاریف بالا، رمزنگاری در سطح رابطه و خصیصه بر این موضوع دلالت دارند که برای یک پرس‌وجو بایستی کل رابطه شامل‌شده در پرس‌وجو به سمت کلاینت بازگردانده شود و بنابراین استخراج زیرمجموعه‌ای از تاپل‌ها در این نوع از ریزدانگی غیرممکن است. از طرف دیگر، ریزدانگی در سطح عنصر نیز به کار اضافی برای مالک داده و کلاینت جهت رمزگذاری و رمزگشایی منجر می‌شود. برای ایجاد یک تعادل بین بارکاری سرور و کارایی اجرای پرس‌وجوها، بیشتر ارائه‌ها فرض می‌کنند که پایگاه داده در سطح تاپل رمزنگاری می‌شوند. در حالی رمزنگاری پایگاه داده، یک سطح امنیتی نسبتا خوبی را برای محافظت از داده‌ها تامین می‌نماید، اما این کار اجرای مستقیم پرس‌وجوهای کاربر بر روی داده‌های رمزشده سمت سرور را غیرممکن می‌سازد. در اصل، سرور به محض دریافت یک پرس‌وجو، تنها می‌تواند رابطه رمزنگاری‌شده درگیر در پرس‌وجو را به درخواست کننده بفرستد و کلاینت مجبور است که کل رابطه برگشت‌داده‌شده را رمزگشایی نموده و پرس‌وجو را بر روی آن اجرا نماید. برای این‌که سرور قادر باشد تا مجموعه‌ای از تاپل‌های یک رابطه را به عنوان پاسخ یک پرس‌وجو انتخاب نماید، مجموعه‌ای از شاخص‌ها می‌تواند به رابطه رمزگذاری‌شده اضافه گردد. در این صورت سرور، یک رابطه رمزگذاری‌شده را به‌همراه یک شاخص برای هر خصیصه ذخیره می‌نماید. برای سادگی کار، ما فرض می‌کنیم که یک شاخص برای هر خصیصه در هر رابطه از پایگاه داده وجود دارد. انواع متفاوتی از شاخص‌ها می‌تواند برای ویژگی‌های یک رابطه تعیین گردند و انتخاب این شاخص‌ها بستگی به عبارت و شرط‌های مورد نیاز در پرس‌وجو دارد.

فرض کنید یک پایگاه داده به نام R داریم. هر رابطه R_i با شمای $R_i (A_{i,1}, \dots, A_{i,n})$ در R به یک رابطه R_i^k دارای شمای R_i^k (counter, Etuple, $I_{i,1}, \dots, I_{i,n}$) در پایگاه داده رمزشده نگاشت می‌شود. در اینجا Counter یک خصیصه عددی است که به عنوان کلید اصلی رابطه رمزنگاری‌شده در نظر گرفته می‌شود. خصیصه Etuple نیز حاوی تاپل رمزشده می باشد و محتوای آن به‌وسیله اعمال یک تابع رمزنگاری چون E_k بر روی

Emp-Id	Name	YoB	Dept	Salary
P01	Ann	1980	Production	10
R01	Bob	1975	R&D	15
F01	Bob	1985	Financial	10
P02	Carol	1980	Production	20
F02	Ann	1980	Financial	15
R02	David	1978	R&D	15

شکل (۶). رابطه Employee

Counter	Etuple	I_1	I_2	I_3	I_4	I_5
1	ite6Az*+8wc	π	α	γ	ε	λ
2	8(Xznfeua4!=	ϕ	β	δ	θ	λ
3	Q73gnew321*/	ϕ	β	γ	μ	λ
4	-1vs9e892s	π	α	γ	ε	ρ
5	e32rfs4aS+@	π	α	γ	μ	λ
6	r43arg*5[)	ϕ	β	δ	θ	λ

شکل (۷). رابطه رمزشده Employee

تاپل اصلی به‌دست آمده است. در اینجا k در واقع همان کلید است و $I_{i,j}$ نیز بیانگر شاخص نسبت‌داده‌شده به j امین خصیصه رابطه R_i یعنی $A_{i,j}$ می‌باشد. در اینجا ما فرض کرده‌ایم که تاپل‌های

1- Relation
2- Attribute
3- Tuple
4- Element

رمزگذاری‌شده و شاخص‌ها در یک رابطه یکسان قرار دارند، اما شاخص‌ها می‌توانند در یک رابطه مجزا نیز قرار داشته باشند. برای روشن شدن مسئله، کار را با یک مثال ادامه می‌دهیم. رابطه Employee در شکل (۶) را در نظر بگیرید. رابطه رمزنگاری‌شده مطابق آن در شکل (۷) آورده شده است.

مقادیر شاخص به طور قراردادی با حروف یونانی ارائه شده است و همچنین رابطه رمزنگاری‌شده دارای تعداد یکسانی تاپل با رابطه اصلی می‌باشد. به خاطر خوانایی بیشتر، ترتیب تاپل‌های رابطه رمزگذاری‌شده و تاپل‌های رابطه اصلی و همچنین ترتیب شاخص در هردو رابطه اصلی و رابطه رمزگذاری‌شده حفظ گردیده است. البته در سیستم‌های واقعی برای امنیت بیشتر خصیصه‌ها و تاپل‌ها و همچنین تطابق بین خصیصه‌ها و شاخص‌ها، آنها در رابطه‌های مربوط به فراداده‌ها[1] نگهداری می‌گردد و این رابطه ها نیز تنها توسط بخش های مجاز قابل دسترسی می‌باشند. از آنجایی که قدرت محاسباتی ذخیره‌سازی مشتری ممکن است محدود باشد، یکی از اهداف اصلی فرآیند اجرای پرس وجو حداقل کردن بار کاری سمت کلاینت و حداکثرسازی بار کاری محاسباتی سمت سرور می‌باشد.

۷-۲- فراداده

کاربران، مالکان و تامین‌کنندگان خدمات به‌منظور مدیریت و دستیابی بهتر پایگاه داده‌های برون‌سپاری شده مبادرت به ذخیره‌سازی برخی اطلاعات اضافی به نام فراداده می‌کنند. کلاینت و

جدول (۱). رابطه Tabuser

Tabuser		
Iduser	Surname	Name

جدول(۲). رابطه Access_Matrix

Access_Matrix		
Iduser	ERelation	Counter

سرور این فراداده را برای تفسیر و اجرای پرس‌وجوهای درخواستی، مورد استفاده قرار می‌دهند. به طور کلی فراداده‌ها به‌صورت جدول‌های رابطه‌ای ذخیره شده و می‌توانند همانند داده‌های اصلی مورد پرس‌وجو واقع شوند. به طور اساسی سه نوع فراداده (اختیارات، توصیفی و مدیریت کلید) وجود دارد [۱۹].

1- Metadata

فراداده اختیارات شامل اطلاعات مربوط به سیاست‌های کنترل دسترسی تعریف‌شده توسط مالک داده می‌باشد. رابطه های مربوط به این فراداده در جدول‌های (۲ - ۱) آورده شده است.

خصیصه‌های زیرخط‌دار در هر رابطه بیانگر کلید آن رابطه می‌باشد.

رابطه Tabuser اطلاعات پیرامون کاربران سیستم را نگهداری می‌کند. شمای این جدول وابستگی زیادی به اطلاعات مورد نیاز مالک داده دارد. برای سادگی کار، ما فرض کرده‌ایم که هر کاربر توسط یک شناسه منحصر به فرد (Iduser) ، یک نام (Name) و یک نام‌خانوادگی (Surname) مشخص می‌شود. رابطه Access Matrix نیز اطلاعاتی پیرامون این‌که چه کسی (Iduser) اجازه

جدول(۳). رابطه TabRelation

TabRelation	
Relation	EncryptedRel

جدول (۴). رابطه Tabindex

Tabindex			
Relation	Attribute	Index	Idmethod

جدول (۵). رابطه TabMethod

TabMethod			
Idmethod	Function	IdParameter	Value

جدول (۶). رابطه EncyptAlgo

EncyptAlgo		
Algorithm	IdParameter	Value

دسترسی به چه رابطه‌ای (ERelation, Counter) را دارد، دربر می گیرد. از آنجایی که این جدول‌ها خیلی حساس هستند، توصیه می‌شود که در سمت کلاینت نگهداری شوند.

فراداده‌های توصیفی در حقیقت توصیف‌کننده اطلاعات بوده و شبیه به بروشور سیستمی عمل می‌کنند. به‌طور اساسی فراداده‌های توصیفی، ساختار پایگاه داده‌های رمزشده را توصیف می‌کنند. رابطه‌های مربوط به این فراداده در جدول‌های (۶ - ۳) آورده شده است.

و نام شاخص متناظر با آن (Index) با توجه به روش شاخص‌گذاری (Idmethod) می‌باشد. رابطه Tabmethod اطلاعات مربوط به تابع درهم‌ساز (Function) استفاده‌شده برای یک روش شاخص‌گذاری (Idmethod) به همراه مقدار (Value) متناظر با پارامترها (IdParameter) را در برمی‌گیرد. رابطه EncrypAlgo اطلاعاتی پیرامون الگوریتم رمزنگاری (Algorithm) به همراه مقادیر (Value) متناظر با پارامترها را دربر می‌گیرد. فاش شدن این جدول‌ها امکان دسترسی به پایگاه داده‌های رمزشده را برای کاربران متخاصم مهیا می‌سازد. بنابراین نباید فراداده‌های توصیفی را در سمت سرور ذخیره‌سازی نمود.

فرا داده‌های مدیریت کلید: شامل اطلاعاتی پیرامون روش اشتقاق کلید و مقادیر کلید تبادل‌شده بین مالکان داده و کاربران می‌باشد. استراتژی‌های متفاوتی برای ذخیره‌سازی این نوع فراداده وجود دارد؛ به‌عنوان مثال می‌توان آن‌ها را به‌طور کامل در سمت کلاینت و یا سرور و یا به طور ترکیبی، بخشی از آن را در سمت سرور و بخش دیگر را در بخش کلاینت ذخیره نمود. استراتژی ذخیره‌سازی فراداده مدیریت کلید در سمت سرور سبب کاهش میزان حافظه مصرفی سمت کلاینت می‌شود اما پهنای باند بیشتری را مصرف می‌کند. با توجه به دلایل ذکرشده در بالا معمولا از حالت ترکیبی برای ذخیره‌سازی این نوع فراداده استفاده می‌شود.

۷-۳- بخش انهدام

این بخش به‌منظور کسب اطمینان کلاینت از حذف کامل داده درخواستی طراحی گردیده است. به سبب خصیصه‌های فیزیکی رسانه‌های ذخیره‌سازی، داده حذف‌شده هنوز هم امکان بازیابی دارد و این امر ممکن است سبب فاش شدن اطلاعات گردد. از این رو تامین‌کننده سرویس می‌بایست تصدیقی را مبنی بر حذف داده و غیر قابل بازیافت بودن آن به کلاینت ارائه نماید. روش‌های موجود برای چنین کنترل‌هایی، وابسته به نوع سرویس فراهم شده و همچنین زیرساخت‌های موجود در سمت فراهم‌کننده متفاوت است. لازم به‌ذکر است که کلیه متدولوژی‌های مربوط به انهدام و بایگانی داده می‌بایست در توافق‌نامه سطح سرویس بین کلاینت و سرور لحاظ گردد.

۷-۴- بخش بایگانی

به‌منظور بهره‌گیری بیشتر از حافظه ذخیره‌سازی، افزایش کارایی و دنبال کردن سیاست‌های سازمان، داده‌های ذخیره‌شده در سمت سرور و یا حتی کلاینت بر اساس پارامترهایی چون تازه گی درج، تازه‌گی پرس‌وجو، تازه‌گی تغییر، تعدد پرس وجو و ... وارد مرحله بایگانی می‌گردند. ورود به مرحله بایگانی ممکن است همراه با انتقال داده‌ها به مکان دیگری نسبت به مکان فعلی باشد.

۷-۵- بخش طبقه‌بندی حساسیت اطلاعات

دسته‌بندی اطلاعات تولیدشده توسط سازمان با توجه به میزان حساسیت، از دست رفتن و یا افشای اطلاعات را طبقه‌بندی اطلاعات گویند. در واقع این رویکرد ما را قادر می‌سازد تا کنترل‌های امنیتی را به درستی اجرا و در طرح طبقه‌بندی خود لحاظ کنیم [۹]. با توجه به جایگاه ویژه طبقه‌بندی اطلاعات، می‌توان اطلاعات یک سازمان را به‌طور کلی به‌صورت زیر طبقه‌بندی نمود.

- **عمومی:** عموم به‌راحتی به این دسته از اطلاعات دسترسی داشته و این آزادی دسترسی باعث نقض اصل محرمانگی در داده‌ها نمی‌شود. مانند اطلاعات معرفی شرکت‌ها که به‌صورت عمومی در اختیار افراد برای برقراری روابط تجاری قرار می‌گیرد.
- **حساس:** اطلاعاتی که نیازمند سطح دسترسی بالاتری می‌باشند و افشای آن‌ها باعث از بین رفتن اصل محرمانگی و یکپارچگی داده‌ها می‌شود در این دسته جای می‌گیرند.
- **خصوصی:** اطلاعاتی است که به‌عنوان ماهیت خصوصی افراد در نظر گرفته می‌شود. هر چند که افشای آن ممکن است آسیب جدی به سازمان نرساند ولی افراد به‌صورت خصوصی تمایلی به افشای آن ندارند. مانند میزان حقوق کارکنان، نوع قرارداد افراد، اطلاعات پزشکی شخص.
- **محرمانه:** اطلاعاتی است که در سطح بسیار حساس در نظر گرفته شده است و صرفا برای استفاده درون‌سازمانی و معاف از قانون دسترسی آزاد به اطلاعات است. افشای غیرمجاز آن جدی تلقی شده و بر روابط سازمان تاثیرگذار می‌باشد؛ مانند اطلاعات مربوط به توسعه یک محصول جدید و اسرار تجاری.

۷-۶- ترجمه پرس‌وجو

این بخش برای ترجمه پرس‌وجوهای کاربر به پرس‌وجوهای قابل اجرا در سمت سرور استفاده می‌شود. این ترجمه مطابق با سیاست‌های موجود در جدول های فراداده سمت کلاینت صورت می‌پذیرد و خروجی حاصل از این مرحله می‌تواند یک یا چند

پرس‌وجو باشد که برخی از آن‌ها سمت سرور و بر روی داده رمزشده و پاره‌ای دیگر سمت کلاینت و بر روی داده رمزگشایی‌شده اجرا می‌گردد.

۸- معماری جدید برون‌سپاری پایگاه داده

مسائل امنیتی فراوانی از آغاز تولید داده تا انهدام آن، امنیت داده را تهدید می‌کنند. در این قسمت سعی داریم تا با توجه به اجزای معرفی‌شده در بخش قبل، فرآیند امنیتی معماری ارایه‌شده در شکل (۸) را مورد بحث و بررسی قرار دهیم.

گام اول: چرخه حیات به محض تولید داده توسط مالک داده آغاز می‌گردد و از همین نقطه ریسک‌های امنیتی مختلفی مطرح می‌گردند. در این گام می‌بایست داده تولیدشده با توجه به سیاست‌های سازمانی از لحاظ طبقه‌بندی اطلاعات برچسب‌گذاری شده و همچنین به‌منظور احراز اصالت داده در گام‌های بعدی برچسب‌های منشا تولید داده و زمان تولید نیز به آن اضافه شود. داده‌ها می‌توانند حتی توسط سرور نیز تولید شده و چرخه حیات خود را آغاز نمایند، به‌عنوان مثال فرض کنید یک تابع محاسباتی

شکل (۸). سناریوی برون‌سپاری داده با توجه به چرخه حیات داده

جدول(۷). شبه‌کد تولید داده

Data Owner
1. Input: ……
2.Output: the relation $R(F_1, F_2, \ldots, F_n)$ along with meta data labels related to field
3.Begin
4. For each field (F_i) created by data owner
5. Labeling F_i in terms of sensitivity.
6. Labeling F_i in terms of creation time.
7. Labeling F_i in terms of creation orgin.
8. Labeling F_i in terms of organizational policies.
9. End For
10.End

سنگین ریاضی به‌عنوان یک سرویس نرم‌افزاری در بستر رایانش ابری ارائه شده باشد. کاربران داده‌های خود را به‌عنوان ورودی به این سرویس تحویل داده و حاصل محاسبات را دریافت می‌کنند. در واقع حاصل محاسبات، داده جدیدی است که توسط سرور تولید شده است. واضح است که اگر روند، این‌گونه که بیان شد پیگیری شود، سرویس‌دهنده می‌تواند از نتایج و دست آوردهای کاربران استفاده نماید. از این رو کاربران قبل از ارسال داده‌های خود، آنها را توسط الگوریتم‌های رمزنگاری هومومورفیک خاص منظوره، رمز نموده و سپس آنها را به‌عنوان داده‌های ورودی به سرور انتقال می‌دهند. شبه‌کد مربوط به این بخش در جدول (۷) آورده شده است.

گام دوم: داده‌های تولیدشده در گام قبل می‌بایست به فضای ذخیره‌سازی سرور انتقال یابد. این انتقال می‌تواند از طریق ایجاد

جدول (۸). شبه‌کد برون‌سپاری داده

Client / relation outsourcing
1. Input: the relation of labeled $R(F_1, F_2, \ldots, F_n)$ recived of data owner
2.Output: the encrypted relation $R'(F'_1, F'_2, \ldots, F'_n)$ for outsourcing
3.Begin
4.For each field (F_i) available in relation R
5. Encrypt field name F_i
6. Encrypt field data F_i
7. End For
8.End

جدول (۹). شبه‌کد ذخیره‌سازی سمت سرور

Server / storage of recived encrypted data from client
1. Input: submitting encrypted relation $R'(F'_1, F'_2, \ldots, F'_n)$ from client
2.Output: the message containing the success or failure of storage of recived data from client
3.Begin
4.Boolean flg=false
5.For each tuple (T_i) from relation (R')
6. For each field (F'_i) related to (T_i)
7. IF (privacy-data(F'_i) ≠ True and privacy-
8. user(F'_i) ≠ true) then
9. Flg=True
10. End IF
11. End For
12. IF (flg=True) then
13. Failure insertion label should be added to T_i
14. Else
15. Save (T_i)
16. End IF
17. End For
18.End

یک کانال امن و یا از طریق یک شبکه نا امن صورت پذیرد. در هر دو صورت سرور نباید از محتوای داده کلاینت باخبر شود. بنابراین کلاینت قبل از برون‌سپاری داده، به‌منظور حفظ امنیت و برقراری محرمانگی، داده‌ها را رمزنگاری می‌نماید. لازم به ذکر است که الگوریتم‌های رمزنگاری استفاده‌شده می‌بایست اجرای سریع و موثر پرس‌وجو روی داده‌های برون‌سپاری‌شده را پشتیبانی نماید. سرور پس از رمزنگاری و اثبات احراز اصالت کاربر، مبادرت به ارسال اطلاعات می‌کند. اگر در این فرآیند هیچ‌گونه خطایی رخ ندهد، سرور مبادرت به ذخیره‌سازی داده‌های دریافتی در ابر می‌کند و در غیر این‌صورت اگر نیازی احساس شود، تبادلات از سرگرفته می شود. شبه‌کدهای مربوط به این بخش در جدول‌های (۹-۸) آورده شده است.

جدول (۱۰). شبه‌کد ترجمه پرس‌وجو

Client / translate query
1. Input: user requested query (Q) 2.Output: creation of two queries Q_s and Q_c 3.Begin 4.For each field (F_i), relation (R), database (C) available in query Q 5. Encrypt field name F_i 6. Encrypt relation name R 7. Encrypt database name C 8. IF (there is a condition on F_i in query Q) then 9. encrypt field value F_i 10. End IF 11.End For 12.End

جدول (۱۱). شبه‌کد اجرای پرس‌وجوی سمت سرور

Server / Query runing.
1. Input: reciving query (Q_s) from client 2.Output: dataset corresponding with query (Q_s) 3.Begin 4. IF (requested query was valid in terms of user's 5. and data authentication) then 6. For each (encrypted relation (R'_i) available 7. in server's storage) 8. For each tuple (T_i) related to (R'_i) 9. IF (query (Q_s) and tuple (T_i) 10. are correspondent) then 11. Add tuple (T_i) to dataset 12. End IF 13. End For 14. End For 15. End IF 16.End

گام سوم: در این گام به‌منظور بهره‌گیری بیشتر از حافظه ذخیره‌سازی، افزایش کارایی و دنبال کردن سیاست‌های سازمان، داده‌های ذخیره‌شده در سمت سرور و یا حتی کلاینت بر اساس پارامترهایی چون تازه‌گی درج، تازه‌گی پرس‌وجو، تازه‌گی تغییر، تعدد پرس‌وجو و ... وارد مرحله بایگانی می‌گردد. ورود به مرحله بایگانی ممکن است همراه با انتقال داده‌ها به مکان دیگری نسبت به مکان فعلی باشد.

گام چهارم، این گام فاز بهره‌برداری از داده می باشد. داده‌ها می‌توانند توسط مالک داده مورد استفاده قرار بگیرند و یا حتی بین چندین شخص به اشتراک گذاشته شوند. از آنجا که داده‌ها بین کابران به اشتراک گذاشته می‌شوند، سرور وظیفه دارد تا سیاست‌های کنترل همروندی را به‌منظور جلوگیری از بروز هرگونه خطای جامعیتی داده اتخاذ نماید. از طرف دیگر هر کاربر به‌منظور استخراج داده‌های مورد نیاز خود از ابر، می‌بایست یک سری اطلاعات اضافی مبنی بر قانونی بودن هویت خود و داده درخواستی به سرور ارسال نماید. شبه‌کدهای مربوط به این بخش در جدول‌های (۱۱-۱۰) آورده شده است.

گام پنجم، هنگامی که دیگر ضرورت وجود بخش خاصی از داده‌ها و یا حتی کل داده‌های سازمان توسط مالک داده احساس نشد، مالک داده از طریق کلاینت درخواست انهدام داده را به سرور ارسال می‌نماید. پس از تصدیق و احراز هویت کاربر درخواست دهنده، داده درخواستی توسط سرور حذف گردیده و پروتکل توافق-شده اثبات انهدام داده بین سرور و کلاینت اجرا می‌گردد. اگر پروتکل اثبات با موفقیت به کار خود پایان دهد، در واقع چرخه حیات داده مورد نظر به پایان رسیده است.

۹-تحلیل و بررسی

تهدیدات مختلفی از آغاز تولید داده تا انهدام آن، امنیت داده را تحت تاثیر قرار می‌دهد. بنابراین می‌بایست در هر مرحله از چرخه حیات داده، تهدیدات را شناسایی و تمهیدات لازم را برای آنها اتخاذ نمود. در این بخش سعی داریم تا معماری‌های ارایه‌شده را از نظر چالش‌های امنیتی مطرح در برون‌سپاری پایگاه داده مورد بحث و بررسی قرار دهیم. معیار مقایسه هر روش را بر پایه ۸ چالش امنیتی مطرح در بخش پنجم و ششم وضعیت چرخه حیات داده قرار داده‌ایم. در جدول های مربوط به این مقایسه‌ها، قسمت‌های تیره‌رنگ بیانگر این است که چالش امنیتی مد نظر در این وضعیت مطرح نمی‌باشد. همچنین علامت‌های $\otimes, \oplus, \ominus$ به‌ترتیب از چب به

شکل (۹). مقایسه معماری‌های برون‌سپاری پایگاه داده

راست بیانگر پشتیبانی معماری از چالش مطرح، پشتیبانی جزئی و عدم پشتیبانی می‌باشد. شکل (۹) معیارهای فوق را برای معماری‌های مختلف نشان می‌دهد.

توجه به پیوستگی مراحل چرخه حیات داده به هنگام اتخاذ تمهیدات امنیتی امری ضروری است، چرا که اگر نقص امنیتی تنها در یک مرحله از چرخه حیات داده صورت پذیرد در حالی‌که مراحل دیگر بهترین تدابیر امنیتی را لحاظ کرده باشند، می‌تواند امنیت کل چرخه حیات داده را تحت تاثیر قرار دهد. لذا در مقایسه‌های فوق پرواضح است که تمام قسمت‌های چرخه حیات داده در نظر گرفته نشده و تنها به بخشی از آن اتکا شده است. لذا ما در روش خود سعی کردیم معماری را با توجه به معماری‌های موجود ارائه دهیم که برکلیه چرخه حیات داده متمرکز باشد. لازم به ذکر است کارایی استفاده از هر معماری منوط به الگوریتم ها و پروتکل‌های استفاده‌شده در اجزاء معماری برون‌سپاری آن دارد. ما در مقاله تمرکزمان بر روی معماری امن برون‌سپاری فارغ از الگوریتم ها و پروتکل‌های موجود برای اجزاء است و در واقع هدف، ایجاد بستری مناسب و فراگیر برای توسعه امن برون‌سپاری می‌باشد.

۱۰- نتیجه‌گیری

رشد روزافزون اطلاعات سازمان‌ها و نیاز به کاهش هزینه‌های ذخیره‌سازی و مدیریت داده‌ها سبب شده تا تمایل به برون‌سپاری داده‌ها روزبه‌روز در حال افزایش باشد. هر چند برون‌سپاری سبب کاهش هزینه‌های مدیریت داده‌ها می‌گردد اما مشکلات و چالش‌های جدید امنیتی برای داده‌های برون‌سپاری‌شده ایجاد می‌نماید. طراحی یک معماری خوب برای سازمان‌ها این امکان را فراهم خواهد کرد تا خط مشی‌های مدیریتی، نیازهای کاربران و استراتژی‌های درونی سازمان با کمترین نیاز به تغییرات در کل سیستم اعمال گردد. از این رو ما در این مقاله به ارائه یک معماری جدید برون‌سپاری مطابق با چرخه حیات داده پرداخته‌ایم که این معماری با توجه به ویژگی‌های لحاظ‌شده در مقایسه با معماری‌های قبلی به‌لحاظ در نظر گرفتن چرخه حیات داده، امن‌تر است. همچنین با مطالعه این مقاله، سازمان‌های مشتاق به برون‌سپاری پایگاه داده در بستر رایانش ابری قادر خواهند بود با محیط برون‌سپاری، ابعاد داده از لحاظ حساسیت و چرخه حیات، چالش‌ها و تهدیدهای پیش روی و راه‌حل‌ها و کنترل‌های امنیتی برون‌سپاری پایگاه داده در بستر رایانش ابری آشنا و تمهیدات لازم برای ارتقاء قابلیت اعتماد و اتکاپذیری بر خدمات برون‌سپاری را اتخاذ نماید.

۱۱- مراجع

[1] J. N. Lee, M. Q. Huynh, R. C. W. Kwok, and S. M. Pi, “IT outsourcing evolution: past present and future,” Communications of the ACM, pp. 84-89, 2003.

[2] Mcfredries, “Technically speaking: The cloud is the computer,” Spectrum, IEEE, pp. 20-20, 2008.

[3] P. Mell and T. Grance, “The NIST definition of cloud computing (draft),” NIST special publication, pp.1-8, 2011.

[4] J. Q. Anderson and H. Rainie, “The future of cloud computing,” Washington, DC: Pew Internet & American Life Project, 2010.

based encryption and its application," In Information Security Applications, pp. 309-323, 2009.

[17] S. D. C. Di Vimercati, S. Foresti, S. Jajodia, Paraboschi, S., & Samarati, P. "A data outsourcing architecture combining cryptography and access control." In Proceedings of the ACM workshop on Computer security architecture, pp. 63-69, 2007.

[18] Foresti, S. "Preserving privacy in data outsourcing", Springer 2010.

[19] E. Damiani, S. D. C. Di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi and P. Samarati, "Metadata management in outsourced encrypted databases," In Secure Data Management, pp. 16-32, 2005.

[20] E. Mykletun, M. Narasimha, and G. Tsudik, "Authentication and integrity in outsourced databases," ACM, 2006.

[5] M. Miller, "Cloud computing: Web-based applications that change the way you work and collaborate online," Que publishing, 2008.

[6] R. Buyya, J. Broberg, and Goscinski, "Cloud computing: Principles and paradigms," John Wiley & Sons, 2010.

[7] R. Saint-Germain, "Information security management best practice based on ISO/IEC 17799," Information Management Journal, pp. 60-66, 2005.

[8] L. Bouganim and P. Pucheral, "Chip-secured data access: Confidential data on untrusted servers," In Proceedings of the 28th international conference on Very Large Data Bases, 2002.

[9] G. Brunette and R. Mogull, "Security guidance for critical areas of focus in cloud computing v2. 1," Cloud Security Alliance, pp. 1-76, 2009.

[10] C. Dong, R. Giovanni, and D. Naranker, "Shared and searchable encrypted data for untrusted servers," Journal of Computer Security, pp. 367-397, 2011.

[11] E. Ferrari, "Database as a Service: Challenges and solutions for privacy and security," Services Computing Conference, 2009.

[12] H. Kadhem, T. Amagasa, and H. Kitagawa, "A novel framework for database security based on mixed cryptography," In Internet and Web Applications and Services, Fourth International Conference IEEE, pp. 163-170, 2009.

[13] J. Hur and D. K. Noh, "Attribute-based access control with efficient revocation in data outsourcing systems," Parallel and Distributed Systems, IEEE Transactions, pp. 1214-1221, 2011.

[14] S. Yu, C. Wang, K. Ren, and W. Lou, "Attribute based data sharing with attribute revocation," In Proceedings of the 5th ACM Symposium on Information, Computer and Communications Security, pp. 261-270, 2010.

[15] S. D. C. Di Vimercati, S. Foresti, S. Jajodia, S. Paraboschi, and P. Samarati, "Over-encryption: management of access control evolution on outsourced data," In Proceedings of the 33rd international conference on Very large data bases, pp. 123-134, 2007.

[16] L. Ibraimi, M. Petkovic, S. Nikova, P. Hartel, and W. Jonker, "Mediated ciphertext-policy attribute-

Journal Of Electronical & Cyber Defence
Vol. 2, No. 4, 2015, Serial No. 8

A novel architecture for database outsourcing in cloud computing with regard to data life cycle

M. Rafiee karkavandi[1*] , S. K. Izadi[2], A. Khoshsefat[3]

1, 3- Master Of Science, Shahid Beheshti University

2- Assistant Professor, Shahid Beheshti University

(Received: 25/05/2014, Accepted: 11/05/2015)

ABSTRACT

The increasing amount of information as well as lack of existence of sufficient computational facilities and storage in organizations have caused various management problems. These problems on the one hand and the rapid expansion of storage services on the other hand have made different organizations to use cloud storage service providers in order to store and manage their organizational information. Using such services, causes organizational information to be stored outside of the organization environment and therefore the owner have less control over its information. Therefore, security concerns will be raised. Many security solutions are proposed to deal with these security concerns, but most of these solutions have focused on a particular aspect of data life cycle such as storage phases. Understanding and considering the data life cycle as well as the challenges and the opportunities facing organizations leads to provide appropriate solutions to overcome security concerns. This paper aims at discussing and analyzing the challenges and opportunities facing organizations using data outsourcing services, and then a new architecture for the database outsourcing with regards to the data life cycle will be presented.

Keywords: Cloud Computing, Outsourcing, Data Life Cycle, Outsourcing Security, Data Classification, Outsourcing Architecture.

* Corresponding Author Email: student.rafiee@gmail.com