



**UNIVERSIDAD DE LAS CIENCIAS INFORMÁTICAS**

**Centro de Soporte**

**COMPONENTE DE SOFTWARE PARA SALVAS AUTOMÁTICAS UTILIZANDO UN  
ENTORNO WEB.**

**TRABAJO FINAL PRESENTADO EN OPCIÓN AL TÍTULO DE MÁSTER EN  
INFORMÁTICA APLICADA**

**Autor:** Ing. Frank David Avalos Palomo

**Tutor:** Dra. C. Marely Del Rosario Cruz Felipe

*La Habana, Cuba. Junio 2012*

*“Año 54 de la Revolución”*

## **Declaración jurada de autoría**

Yo Frank David Avalos Palomo, con carné de identidad 83111521744, declaro que soy el autor principal del resultado que expongo en la presente tesis titulada Componente de software para salvadas automáticas utilizando un entorno web, para optar por el título de Máster en Informática Aplicada.

Este trabajo fue desarrollado durante 2 años en colaboración con mi colega de equipo Yulio Seriocha García Gallardo. El mismo me reconoce la autoría principal del resultado expuesto en esta tesis.

Finalmente declaro que todo lo anteriormente expuesto se ajusta a la verdad, y asumo la responsabilidad moral y jurídica que se derive de este juramento profesional.

Y para que así conste, firmo la presente declaración jurada de autoría en La Habana a los \_\_\_\_ días del mes de \_\_\_\_\_ del año \_\_\_\_\_.

---

Dra. C. Marely Del Rosario Cruz Felipe

---

Ing. Frank David Avalos Palomo

## **Resumen.**

Para insertarse en el mercado internacional del software más allá de su producción y comercialización, se deben crear las herramientas internas capaces de garantizar una producción continua, estable, de alta calidad y sobre todo con una base de control múltiple y segura.

Entre los procesos que se llevan a cabo en la producción de los proyectos de una empresa de desarrollo de software se encuentra el proceso de salvadas de información. Actualmente este proceso se realiza de forma manual y es considerado una tarea complicada, debido primeramente a la gran cantidad de información que se maneja, la importancia de esta y a lo complicado que se le hace a los administradores configurar los servicios de salvadas con un mayor grado de calidad.

Este trabajo consiste en la creación de un sistema informatizado para la configuración y control de las salvadas automáticas de la universidad, en respuesta a la necesidad de mejorar y perfeccionar la manera en que actualmente se lleva a cabo esta tarea. En este trabajo se crea una nueva arquitectura de ficheros de la herramienta de respaldo seleccionada y el versionado para distintos sistemas operativos. La utilización de un sistema informatizado para la configuración y control de salvadas de los proyectos productivos representaría una mejora considerable de este proceso en cuanto a eficiencia y sencillez, contribuyendo favorablemente al desarrollo de software en la universidad.

**Abstract.**

In entering the international market beyond software production and marketing, you must create internal tools capable of ensuring continuous production, stable, high quality and above all with a base of multiple control and safe.

Among the processes carried out at the production company projects a software development process is the bursts of information. Currently this is done manually and is considered a difficult task, due primarily to the large amount of information handled, the importance of this and how hard it is made an administrator to configure the services saved with a higher degree quality.

This work involves the creation of a computerized system for configuration and control of automatic salvos of the university, in response to the need to improve and perfect the way it currently performs this task. The contributions of this work are a new architecture files from the backup tool selected and versioning for different operating systems. With these contributions and the use of a computerized system for configuration and control of productive projects saved would mean a considerable improvement of this process in terms of efficiency, time and organization, contributing favorably to software development in college.

## ÍNDICE

INTRODUCCIÓN .....	1
CAPÍTULO 1. TENDENCIAS DE DESARROLLO DE SOLUCIONES PARA IMPLANTAR SALVAS AUTOMÁTICAS. ....	8
1.1 ¿Qué son las copias de seguridad? .....	8
1.2 Desarrollo de las copias de seguridad .....	9
1.3 Ventajas y desventajas de hacer un respaldo .....	9
1.4 Procedimiento para hacer un respaldo de información .....	10
1.5 Características para hacer un respaldo seguro .....	13
1.6 Estrategias para hacer respaldo de información .....	13
1.7 Copia de información (Backups) .....	15
1.7.1 Tipos de respaldo .....	16
1.8 Software de respaldo y respaldo "En Línea" .....	17
1.9 Lista de software para respaldo .....	19
1.9.1 Descripción de la herramienta para realizar respaldos .....	21
1.10 Metodología y tecnologías utilizadas para el desarrollo del software .....	25
1.11 Conclusiones parciales.....	30
CAPÍTULO 2. SOLUCIÓN PROPUESTA. ....	32
2.1 Arquitectura de Bacula.....	32
2.1.1 Bacula-director.....	34
2.1.2 Bacula-storage daemon .....	35
2.1.3 Bacula-file daemon.....	36
2.1.4 Bacula-console .....	36
2.2 Configuración e instalación de Bacula .....	38
2.2.1 Propuesta de componentes y arquitectura de Bacula.....	38
2.2.2 Proceso de instalación de Bacula.....	39

2.3 Integración de la solución propuesta y Bacula.....	40
2.3.1 Instalación de Bacula en el servidor.....	41
2.3.2 Instalación del cliente de Bacula.....	43
2.3.3 Versionado de Bacula.....	44
2.3.4 Despliegue de la solución propuesta .....	45
2.4 Procedimiento de respaldo de información utilizando la solución propuesta.....	46
2.5 Conclusiones parciales.....	47
CAPÍTULO 3. EVALUACIÓN DEL SISTEMA PROPUESTO.....	49
3.1 Método seleccionado .....	49
3.2 Selección de expertos.....	49
3.3 Confección del instrumento.....	50
3.4 Resultados arrojados del método Delphi.....	51
3.5 Validación de la hipótesis.....	53
3.6 Conclusiones parciales.....	54
CONCLUSIONES GENERALES .....	55
RECOMENDACIONES.....	56
REFERENCIAS BIBLIOGRÁFICAS .....	57
BIBLIOGRAFÍA.....	60
GLOSARIO.....	63
ANEXO 1. Procedimiento para el respaldo de información de una empresa.....	65
ANEXO 2. Configuración del fichero Clients.....	66
ANEXO 3. Configuración del fichero Devices.....	70
ANEXO 4: Instalación de Bacula en el servidor .....	71
ANEXO 5: Instalación de Bacula en el cliente .....	72
ANEXO 6: Procedimiento de respaldo de información utilizando la solución propuesta .....	73
ANEXO 7. Encuesta: Diagnóstico sobre las salvas automáticas .....	74

ANEXO 8. Nivel de experiencia de los expertos .....	75
ANEXO 9. Medidores de los resultados de la investigación. ....	76
ANEXO 10. Cálculo del coeficiente de Kendall y Chi cuadrado. ....	77
ANEXO 11. Encuesta: Diagnóstico sobre la solución propuesta. ....	78

## **INTRODUCCIÓN**

Uno de los factores que siempre ha sido clave para toda empresa, es el proceso de toma de decisiones oportunas y acertadas. Hoy en día, el elemento diferenciador entre empresas supervivientes y sobrevivientes, radica en el aprovechamiento de los recursos que la tecnología ofrece, y la manera en que dichos recursos son explotados por cada una de las organizaciones. Todos ellos relacionados con la manipulación de datos para proveer información clara, precisa y confiable que sea utilizada para la toma de decisiones oportuna y acertada.

Actualmente las empresas y organizaciones enfocan gran parte de sus esfuerzos en detectar áreas de mejora que les permitan optimizar su desempeño, con la finalidad de mantenerse en el nivel competitivo deseado.

Uno de los factores que influyen para que el proceso de la administración de la información se lleve a cabo de manera adecuada, es el uso de herramientas tecnológicas que proporcionen el soporte necesario para agilizar esos procesos. Y como consecuencia de ello un incremento en el desempeño dentro de la empresa así como también de la reducción sus costos.

Dentro de las herramientas tecnológicas, se encuentran los programas o software, así como también la infraestructura física necesaria para soportar los programas. Un software se refiere a "las instrucciones electrónicas que van a indicar al ordenador qué es lo que tiene que hacer. También se puede decir que son los programas usados para dirigir las funciones de un sistema de computación o un hardware". [1]

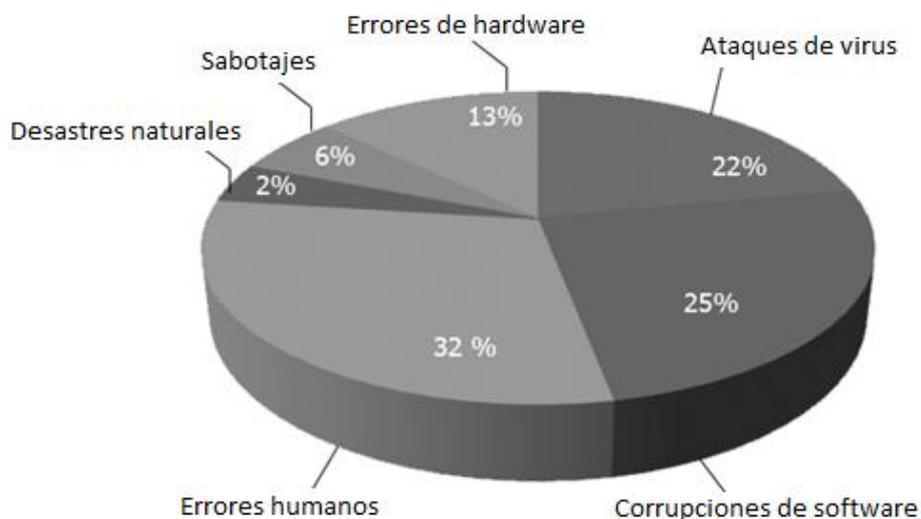
Si el monitor, la memoria e incluso la CPU de una computadora dejan de funcionar, simplemente se reemplaza por otra, y no hay mayores dificultades. Pero si falla el disco duro, el daño puede ser irreversible, puede significar la pérdida total de la información. Es principalmente por esta razón, por la que se debe respaldar la información importante. Si esto le sucediera a una empresa, las pérdidas económicas podrían ser cuantiosas. Los negocios de todos los tipos y tamaños confían en la información computarizada para facilitar su operación. La pérdida de información provoca un daño de fondo:

- Pérdida de oportunidades de negocio
- Clientes decepcionados
- Reputación perdida

La tecnología no está exenta de fallas o errores, y los respaldos de información son utilizados como un plan de contingencia en caso de que una falla o error se presente. [2][3]

Las interrupciones se presentan de formas muy variadas: virus informáticos, fallos de electricidad, errores de hardware y software, caídas de red, hackers, errores humanos, incendios, inundaciones, entre otros. Y aunque no se pueda evitar totalmente cada una de estas interrupciones, la empresa sí puede prepararse para evitar las consecuencias que éstas puedan tener sobre su negocio; la gravedad dependerá del tiempo que tarde en reaccionar una empresa

A continuación la figura 1 muestra el escenario actual de los porcentajes relacionados con la pérdida de información en cualquier ordenador.



*Figura 1. Principales causas de pérdida de la información. [4]*

Como se muestra en la figura 1 son diversas las causas que pueden ocurrir para que se pierda información de cualquier ordenador o red de ordenadores. Como se puede observar, el aspecto con más porcentaje de pérdidas de datos es el error humano, puesto que los trabajadores de una empresa muestran a menudo descuido en la utilización de los procedimientos de seguridad; uso inadecuado de mecanismos de autenticación o contraseñas; transferencia de archivos a dispositivos de memoria USB e imprudencia en la protección de sus portátiles cuando viajan.

Acciones como las mencionadas anteriormente, además de poner en serio peligro los datos confidenciales y sensibles de las empresas, pueden conllevar además a consecuencias muy negativas para la credibilidad de las organizaciones; como la pérdida de confianza de accionistas y clientes, daño a la imagen empresarial, interrupción del negocio y acciones regulatorias.

Anteriormente se explicaba el valor que tiene la información y los datos para las empresas, lo que resulta increíble de esto es la falta de precauciones que se suele tener al confiar el núcleo de los negocios al sistema de almacenamiento, de lo que en la mayoría de los casos resulta ser una computadora pobremente armada tanto del punto de vista de hardware como de software. Sin duda alguna uno de los programas que mayor utilidad representa dentro de una empresa, son los denominados “Sistemas de Respaldo”.

Los sistemas de salvados de los proyectos de software de una empresa se pueden desarrollar de forma manual, lo que provoca deficiencia en la calidad de este procedimiento debido a lo engorroso que puede tornarse este proceso ante la cantidad de información existente.

Los sistemas de salvados o respaldo que existen hoy en el mundo no son del todo eficientes y la configuración de estos sistemas es poco sencilla por la gran cantidad de variables en los ficheros que hay que configurar. [5] Los cuales cuando se configuran de forma manual pueden estar sujetos a errores, pues se debe tener alto dominio del sistema a configurar y cuando hay un cambio a otro servidor de salvado hay que repetir el proceso desde el principio.

El respaldo es una tarea compleja y constituye cada vez más un problema que impide a las empresas circular con fluidez, introduciendo atascos y peligros para garantizar su continuidad vital. En un mundo en permanente cambio, es preciso introducir nuevos servicios de respaldo y recuperación de datos que garanticen a las empresas que su seguridad está en línea con el cambio permanente. Además de que las empresas están expuestas a desastres naturales lo que representa un alto riesgo para la pérdida de información.

Los sistemas de salvados pueden ser implementados para hacer backups a un área específica dentro de la empresa, o bien puede ser diseñado para que uno solo de ellos proporcione soporte a diversas áreas.

En el mundo ya existen varios software, libres y privados, que dan de alguna manera una solución a este tipo de problema, estos sin duda pueden usarse con sistemas en producción, pero aún existen algunos elementos que deben mejorarse.

Si un sistema de respaldo está haciendo peticiones al usuario a cada momento para realizar un backup es mejor hacerlo manualmente. Lo más factible es que exista un servicio que se encargue de todo automáticamente, que no consuma tantos recursos y sea confiable; así que se debe investigar cuáles de estos programas son los más populares y eficaces.

Para utilizar un sistema de respaldo en una gran empresa donde se desarrollen varios proyectos a la vez, es necesario instalar y configurar la herramienta en cada uno de los servidores que se dediquen a almacenar información. Si en uno de los proyectos se determina que se debe cambiar el IP del servidor, dirección de respaldo u otro dato importante para el funcionamiento del sistema, se deben volver a configurar los clientes y servidores para un funcionamiento estable.

Vale destacar que los servidores pueden estar situados en lugares alejados y si se cuenta con poco personal para suplir esta tarea, la misma puede resultar engorrosa para las personas que se encarguen de resolver la configuración y funcionamiento de los servidores de los proyectos productivos de software de la empresa.

Por lo antes analizado el **problema a resolver** queda formulado de la siguiente forma:

¿Cómo lograr que la configuración y control de los servidores de salvadas de los proyectos de una entidad de software se torne más sencilla y eficiente?

El **objeto de estudio** de esta investigación lo constituye el proceso de configuración y control de los servidores de salvadas.

Se plantea entonces como **objetivo general** desarrollar un sistema informático que permita la configuración y control de los servidores de salvadas de manera más sencilla y eficiente, y para darle solución al objetivo general se plantea los **objetivos específicos** siguientes:

1. Realizar un estudio de los diferentes software que agilizan la configuración y control del sistema de salvadas.
2. Aplicar las técnicas, lenguaje de programación y herramientas adecuadas para la implementación de la aplicación.

Para cumplir con el objetivo general propuesto, se proponen las **tareas** siguientes:

1. Investigación del funcionamiento de las herramientas que permiten salvas automáticas para la selección de la herramienta que se le hará el proceso de configuración y control.
2. Investigación de las tecnologías y metodologías para el desarrollo de un software que permita la configuración y control de los servidores y clientes de salvas.
3. Implementación del sistema informático que permita la configuración y control de los servidores y clientes de salvas.
4. Validación de los resultados obtenidos en la aplicación.

El **campo de acción** que abarca esta investigación es la automatización de la configuración y control de los servidores de salvas.

Como **hipótesis** de la investigación se plantea: Al desarrollar un sistema informático que agilice la configuración y gestión de los servidores de respaldo de información de una empresa, el proceso de salvas puede tener mejores resultados en cuanto a la eficiencia y sencillez.

Entiéndase como “eficiencia” la cualidad que facilitaría la reducción, agilización y organización del flujo de trabajo manual de cada empresa, es decir, se trata de la capacidad de alcanzar un objetivo fijado con anterioridad [6] en el menor tiempo posible, con el mínimo uso de los recursos humanos y minimizar el margen de error que el objetivo pueda presentar.

Entiéndase como “sencillez” a reducir la cantidad de procesos que intervienen en el respaldo de información de una empresa.

### **Métodos Científicos usados en la investigación**

- **Teóricos**

**Analítico–sintético:** Este método ha servido para analizar y comprender la teoría y documentación relacionada con el tema de investigación, permitiendo así, extraer los elementos más relacionados e importantes con el objeto de estudio.

**Análisis histórico–lógico:** Este método ha ayudado a entender el surgimiento y la evolución del tema de la investigación, así como otras temáticas estrechamente relacionadas con ella a lo largo de la historia de la informática.

- **Empíricos**

**Observación:** Este método es de vital importancia ya que ha permitido percibir a partir de la situación real que se está investigando cómo se desarrolla a groso modo el proceso que constituye el objeto de estudio.

**Entrevista:** Para el desarrollo de este método se ha entrevistado al jefe del grupo de soporte y administradores centrales de la producción, quienes han aportado elementos significativos a la investigación. Con este método se validarán los resultados de la investigación teniendo en cuenta la aprobación de la variable de eficiencia.

Los aportes del sistema radican en que con su realización, se espera solventar los problemas existentes en el proceso de gestión y control de salvadas automáticas en una empresa de desarrollo de software.

Se tendrá un sistema informático en el que se tengan en cuenta todos los ficheros de configuración del sistema de salvadas para ser manejado, proponiendo una nueva arquitectura de ficheros la cual hace más eficiente y sencillo el proceso que constituye el objeto de estudio.

Por otro lado se incorpora una solución para el versionado del sistema de salvadas escogido, la cual deja abierto el uso de los diferentes tipos de versiones en los sistemas operativos con que cuente la distribución de los servidores de la empresa.

Para mostrar el desarrollo de la investigación y sus resultados, el trabajo se ha estructurado en tres capítulos, además de las Conclusiones, Recomendaciones, Glosario de términos, Referencias bibliográficas, Bibliografía y Anexos.

En el primer capítulo, se detallan algunos temas elementales de salvadas automáticas y se realiza un análisis de las herramientas más empleadas para el resguardo de la información. Por otro lado se elabora un procedimiento para realizar respaldos de información en una empresa. Por último se hace una investigación de las tecnologías y metodologías para desarrollar el sistema propuesto.

El segundo capítulo está dedicado a la construcción y presentación de la solución propuesta. Se elabora el diseño del sistema planteado y su integración con el sistema de salvada automática escogido en el capítulo 1. Por último se detalla un nuevo procedimiento para realizar respaldo utilizando esta solución.

En el tercer capítulo se describe la evaluación de la solución, empleando el método de criterio de experto con varias iteraciones y sin intercambio directo. También en este capítulo se valida la hipótesis del problema planteado.

## **CAPÍTULO 1. TENDENCIAS DE DESARROLLO DE SOLUCIONES PARA IMPLANTAR SALVAS AUTOMÁTICAS.**

### **INTRODUCCIÓN**

Con el objetivo de facilitar la comprensión del alcance de la investigación, en el presente capítulo se introducen una serie de temas fundamentales asociados al problema planteado.

Primeramente se profundizará para qué sirven las salvas automáticas, así como, algunos de los sistemas que se utilizan en el mundo. Se detallarán las ventajas y desventajas que presenta utilizar un sistema de salvas en una empresa. Se hará una investigación del procedimiento utilizado en una empresa de software para hacer copias de seguridad de forma confiable y continua. Además se analizarán los tipos de mecanismos que se utilizan para las salvas y por tanto se valora que ellos sirven para lo mismo para una gran corporación como para una pequeña empresa. Por último se hará un estudio de las distintas tecnologías y herramientas para un posterior desarrollo de la solución propuesta.

Se profundizará además en el objeto de estudio y la situación problemática haciéndose una descripción detallada.

### **1.1 ¿Qué son las copias de seguridad?**

Existen varias definiciones acerca de lo que son las copias de respaldo, aunque todas convergen en la misma idea. A continuación se muestran varias de estas definiciones obtenidas de diferentes fuentes:

Copia de ficheros o datos de forma que estén disponibles en caso de que un fallo produzca la pérdida de los originales. Esta sencilla acción evita numerosos y a veces irremediables problemas si se realiza de forma habitual y periódica [7].

Hacer una copia de seguridad o copia de respaldo (Backus en inglés, el uso de éste está ampliamente extendido) se refiere a la copia de datos de tal forma que estas copias adicionales pueden restaurar un sistema después de una pérdida de información [8].

Copia de seguridad es la actividad de copiar archivos a bases de datos a fin de que puedan ser conservadas en casos de fallo del equipo o de otra catástrofe. La copia de seguridad suele ser una parte rutinaria de la operación de grandes empresas con

supercomputadoras, así como los administradores de las empresas más pequeñas con menos recursos. Para los usuarios de computadoras personales, la copia de seguridad también es necesaria pero no se realiza muy a menudo. La recuperación de los archivos de la copia de seguridad se denomina: restablecimiento [9].

De los conceptos anteriormente planteados el presente trabajo se ajusta al concepto de hacer copias de seguridad mediante backups ya sea local o no, puesto que su idea fundamental es utilizar un sistema de respaldo que realice backups con una configuración de forma eficiente, sencilla y segura.

### **1.2 Desarrollo de las copias de seguridad**

Las copias de seguridad, hoy en día, son un campo rápido en vías de desarrollo. En las nuevas tendencias y soluciones aparecen, los métodos de respaldo y las tecnologías se convierten en más complejas. El respaldo es la copia de un archivo para efectos de su posterior reconstrucción. Se puede enfatizar que los primeros dos aspectos a tener en cuenta son: el soporte lógico informático de almacenamiento para los datos y los depositarios para soporte lógico informático del respaldo. Otro aspecto importante es la necesidad del respaldo causado por el desarrollo tecnológico y la expansión de volúmenes de datos.

Las primeras copias de seguridad estaban hechas de grandes carretes de cinta magnetofónica y hasta en papel, como las cintas perforadas. En la siguiente era se guardaron en su mayor parte en discos flexibles de tamaños diversos. Los respaldos son actualmente escritos en discos compactos, unidades de disco duro, dispositivos flash y hasta por vía remota. Pero algunas tecnologías, como cinta, aún son muy populares continuándose su desarrollo y fabricación.

Con el desarrollo del mundo informático ya hoy en día son muy variadas las alternativas tanto de software como de hardware para hacer respaldos seguros. Es importante destacar que con este avance informático las soluciones pueden ser con costes monetarios muy altos, medios o nulos, estos últimos no son pocos y ya dan un gran salto en la competencia mundial en este campo de acuerdo a sus prestaciones.

### **1.3 Ventajas y desventajas de hacer un respaldo**

Ventajas:

- No es necesario ningún dispositivo de almacenamiento para el respaldo.
- No es necesario separar o subdividir, como tampoco es necesario ordenarlos de ningún modo.

- Todo el proceso es automático.
- No importa si no está en función el ordenador, recupera la información de todos modos.

Desventajas:

- Hacer respaldo absoluto, inclusive los que no son necesarios.
- Respalda todos tipos de datos aunque estén infectados por algún tipo de virus o algo por el estilo.

#### **1.4 Procedimiento para hacer un respaldo de información**

Independientemente del tipo de respaldo que se haga las empresas deben cumplir diferentes procesos para que se haga un respaldo seguro. Por lo tanto una empresa debe realizar periódicamente copias de resguardo de la información. Se debe contar con adecuadas instalaciones de resguardo para garantizar que toda la información de la empresa pueda recuperarse una vez ocurrido un desastre o falla de los dispositivos.

Las disposiciones para el resguardo de cada uno de los sistemas deben ser probadas periódicamente para garantizar que cumplan con los requerimientos de los planes de continuidad de los negocios. Se deben tener en cuenta los siguientes controles:

- Se debe almacenar en una ubicación remota un nivel mínimo de información de resguardo, junto con registros exactos y completos de las copias de resguardo y los procedimientos documentados de restauración, a una distancia suficiente como para evitar daños provenientes de un desastre en el sitio principal. Se deben retener al menos tres generaciones o ciclos de información de resguardo para aplicaciones importantes de la empresa.
- Se debe asignar a la información de resguardo un nivel adecuado de protección física y ambiental consecuente con los estándares aplicados en el sitio principal. Los controles aplicados a los dispositivos en el sitio principal deben extenderse para cubrir el sitio de resguardo.
- Los medios de resguardo deben probarse periódicamente, cuando sea factible, a fin de garantizar su confiabilidad en relación a su eventual uso en casos de emergencia.
- Los procedimientos de restauración deben verificarse y probarse periódicamente para garantizar su eficacia y cumplimiento dentro del tiempo asignado a la recuperación en los procedimientos operativos.

A continuación, en la tabla 1, se muestra el procedimiento para hacer un respaldo de información en una empresa.

*Tabla #1: Procedimiento para el respaldo de información de una empresa.*

<b>Secuencia</b>	<b>Proceso</b>	<b>Responsable</b>
1.0 Asignar responsable que realizará las actividades.	1.1 Asignación de responsable quien realizará las actividades de respaldo de información diariamente.	Área productiva
2.0 Encender equipo diariamente a las 8:00 a.m.	2.1 Encendido del equipo diariamente a las 8:00 a.m. donde se encuentra el programa que realiza el respaldo de información.	Área productiva
3.0 Respalda información en servidor local.	3.1 Respaldo automático de la información a las 8:30 a.m. diariamente en el servidor de salvadas del área productiva.	Área productiva
4.0 Renombrar carpeta.	4.1 Renombrar la carpeta donde se realiza el respaldo diario, el cual tendrá el siguiente nombre, Respddmmaa (dd=día, mm=mes, aa=año) que corresponderá al día, mes y año en que se hace el respaldo.	Área productiva
5.0 Respalda información en servidor central.	5.1 Respaldo en servidor central al finalizar la semana, únicamente se guardarán las carpetas del último día de la semana y en el caso del código de programación cuando se haga alguna modificación. 5.2 Probar que la información grabada esté correcta. 5.3 Eliminar las carpetas de los respaldos contenidas en el servidor local, para evitar su saturación. 5.4 Hacer copia y conservarla para cuando las áreas la soliciten.	Dirección de redes.
6.0 Informar que cuenta	6.1 Informa al área productiva que	Dirección de

**Capítulo 1. Tendencias de desarrollo de soluciones para implantar salvadas automáticas.**

con respaldo en servidor central.	cuenta con un respaldo en el servidor central del último día de cada semana.	producción.
7.0 Solicitar información.	7.1 Solicitar información respaldada. <ul style="list-style-type: none"> <li>• Memorando solicitud de respaldo.</li> </ul>	Área productiva.
8.0 Recibir solicitud.	8.1 Recibe solicitud de información y gira instrucciones a la Dirección de redes. <ul style="list-style-type: none"> <li>• Memorando solicitud de respaldo.</li> </ul>	Dirección de producción.
9.0 Recibir solicitud.	9.1 Recibe solicitud de información y gira instrucciones al área productiva, para la realización del trabajo. <ul style="list-style-type: none"> <li>• Memorando solicitud de respaldo.</li> </ul>	Dirección de redes.
10.0 Recibir solicitud de información.	10.1 Recibe solicitud de información y procede a enviar informe de los datos necesarios para restablecer la información salvada.	Área productiva.
11.0 Respaldo de información.	11.1 Resguarda en servidor local la información respaldada en el servidor central. <ul style="list-style-type: none"> <li>• Memorando con solicitud de respaldo.</li> </ul>	Dirección de redes.
12.0 Fin del proceso	12.1 Confirmación a la dirección de redes del respaldo satisfactorio.	Área productiva.

En la tabla anterior se muestran los procesos que debe seguir una empresa, independientemente de los medios físicos de resguardo que se utilicen o el sistema informático empleado para ello. En el Anexo 1, se muestra un diagrama de flujo con un detalle más ampliado del procedimiento anterior.

### **1.5 Características para hacer un respaldo seguro**

Los sistemas de salvadas automáticas deben tener características horizontales que ayuden a su ejecución. No importa el tipo de sistema de acuerdo a su licencia o costo monetario, si es para redes grandes o pequeñas, lo más importante es que cada sistema debe tener las siguientes características:

- **Confiable:** Minimizar las probabilidades de error. Muchos medios magnéticos como las cintas de respaldo, los disquetes, o discos duros tienen probabilidades de error o son particularmente sensibles a campos magnéticos, elementos todos que atentan contra la información que se ha respaldado allí.

Otras veces la falta de confiabilidad se genera al rehusar los medios magnéticos. Las cintas en particular tienen una vida útil concreta. Es común que se subestime este factor y se reutilicen más allá de su vida útil con resultados nefastos; particularmente porque se va a descubrir su falta de confiabilidad en el peor momento: cuando se necesita recuperar la información.

- **Lugar seguro:** Tan pronto se realiza el respaldo de información, el soporte que almacena este respaldo debe ser desconectado de la computadora y almacenado en un lugar seguro tanto desde el punto de vista de sus requerimientos técnicos: humedad, temperatura, campos magnéticos, seguridad física y lógica. No es de gran utilidad respaldar la información y dejar el respaldo conectado a la computadora donde puede haber un ataque de cualquier índole que lo afecte.
- **Rápida y eficiente:** Es necesario probar la confiabilidad del sistema de respaldo no sólo para respaldar sino también para recuperar. Hay sistemas de respaldo que aparentemente no tienen ninguna falla al generar el respaldo de la información pero que fallan completamente al recuperar estos datos al sistema informático. Esto depende de la efectividad y calidad del sistema que realiza el respaldo y la recuperación.

Estas características de respaldo se tendrán en cuenta en la confección de la solución propuesta del capítulo 2.

### **1.6 Estrategias para hacer respaldo de información**

Se pueden adoptar diferentes estrategias respecto a la forma de la copia, que condicionan el volumen de información a copiar, para ello la copia puede ser:

- Copiar sólo los datos: poco recomendable, ya que en caso de incidencia, será preciso recuperar el entorno que proporcionan los programas para acceder a ellos, lo que influye negativamente en el plazo de recuperación del sistema.
- Copia completa: recomendable, si el soporte, tiempo de copia y frecuencia lo permiten, incluye una copia de datos y programas, restaurando el sistema al momento anterior a la copia.
- Copia incremental: solamente se almacenan las modificaciones realizadas desde la última copia de seguridad, por lo que es necesario mantener la copia original sobre la cual se pueda restaurar el resto de copias. Utiliza un mínimo espacio de almacenamiento y minimiza el tipo de desarrollo, a costa de una recuperación más complicada.
- Copia diferencial: como la incremental, pero en vez de solamente modificaciones, se almacenan los ficheros completos que han sido modificados. También necesita la copia original.
- Tiempo disponible para efectuar la copia: el tiempo disponible para efectuar la copia de seguridad es importante, ya que el soporte utilizado, unidad de grabación y volumen de datos a almacenar, puede hacer que el proceso de grabación de los datos dure horas, y teniendo en cuenta que mientras se efectúa el proceso es conveniente no realizar accesos o modificaciones sobre los datos objeto de la copia, este proceso ha de planificarse para que no suponga un contratiempo en el funcionamiento habitual del sistema de información.
- Frecuencia de realización de copias de seguridad: la realización de copias de seguridad ha de realizarse diariamente. Éste es el principio que debe regir la planificación de las copias, sin embargo existen condicionantes, tales como la frecuencia de actualización de los datos, el volumen de datos modificados, entre otros, que pueden hacer que las copias se realicen con más tiempo.
- Mecanismos de comprobación: se deben definir mecanismos de comprobación de las copias de seguridad, aunque los propios programas que las efectúan suelen disponer de ellos para verificar el estado de la copia, es conveniente planificar dentro de las tareas de seguridad la restauración de una parte de la copia o de la copia completa periódicamente, como mecanismo de prueba y garantía.

- Responsable del proceso: la mejor forma de controlar los procesos que se desarrollan en el sistema de información, aunque estos estén desarrollados en una parte importante por el propio sistema, es que exista un responsable de la supervisión. Para ello se debe designar a una persona que incluya entre sus funciones la supervisión del procedimiento de copias de seguridad. Y el almacenamiento de los soportes empleados en un lugar designado a tal fin e inclusive en de la verificación de que las copias se hayan realizado correctamente.
- Medidas de Seguridad: respecto a las medidas de seguridad, se deben tener en cuenta los siguientes puntos:
  1. Deberá existir un usuario del sistema, entre cuyas funciones esté la de verificar la correcta aplicación de los procedimientos de realización de las copias de respaldo y recuperación de los datos.
  2. Los procedimientos establecidos para la realización de las copias de seguridad deberán garantizar su reconstrucción en el estado en que se encontraban al tiempo de producirse la pérdida o destrucción.
  3. Deberán realizarse copias de respaldo al menos semanalmente, salvo que en dicho período no se hubiera producido ninguna actualización de los datos.

El sistema propuesto en el segundo capítulo debe cumplir con todas estas características, pues este debe ser seguro, confiable, flexible al cambio, con respaldos incrementales y completos, con diferentes frecuencias de salvas y sobre todo de código abierto por las políticas que hoy lleva el país de tender al software libre.

### **1.7 Copia de información (Backups)**

Estos respaldos son sólo duplicados de archivos que se guardan en "Tape Drives" de alta capacidad. Los archivos que son respaldados pueden variar desde archivos del sistema operativo, bases de datos , hasta archivos de un usuario común. Existen varios tipos de software que automatizan la ejecución de estos respaldos, pero el funcionamiento básico de estos paquetes depende del denominado archive bit. [10]

Este archive bit indica un punto de respaldo y puede existir por archivo o al nivel de "Bloque de Información", esto dependerá tanto del software que sea utilizado para los respaldos así como el archivo que sea respaldado. Este mismo archive bit es activado en los archivos (o bloques) cada vez que estos sean modificados y es mediante este archive bit que se llevan a cabo los tres tipos de respaldos comúnmente utilizados.

### 1.7.1 Tipos de respaldo

- Total o completo: es donde cada archivo es escrito a la media de respaldo. Si los datos a respaldar nunca cambian, cada respaldo completo creado será una copia exacta.
- Incremental: primero revisa para ver si la fecha de modificación de un archivo es más reciente que la fecha de su último respaldo. De no serlo, significa que el archivo no ha sido modificado desde su último respaldo y por tanto se puede resaltar esta vez. Por otro lado, si la fecha de modificación es más reciente, el archivo ha sido modificado y se debería copiar. Los respaldos incrementales son utilizados en conjunto con respaldos regulares completos, por ejemplo un respaldo semanal completo, con respaldos incrementales diarios.
- Diferencial: son similares a los respaldos incrementales ya que copian los archivos que han sido modificados. Los respaldos diferenciales son acumulativos, en otras palabras, con un respaldo diferencial, una vez que un archivo ha sido modificado continuo siendo incluido en todos los respaldos diferenciales subsecuentes.[11]

A continuación se muestra una tabla con las ventajas y desventajas de las distintas maneras de copia de la información.

*Tabla #2: Tipos de copia de seguridad*

Respaldo	Archivos en respaldo	Archivo bit	Ventajas	Desventajas
<b>Completo ("Full")</b>	Todos	Eliminado en todos los archivos	Con este respaldo únicamente es posible recuperar toda la información	Tiempo de ejecución
<b>De incremento ("Incremental")</b>	Archivos con archive bit activo.(Aquellos que hayan cambiado desde el último respaldo completo)	Eliminado en los archivos que se respaldan	Velocidad	Requiere del último respaldo completo y de todos los respaldos de incremento que le siguieron para recuperar el Sistema
<b>Diferencial ("Differential")</b>	Archivos con archive bit	Intacto	Sólo requiere del último	Ocupa mayor espacio en

	activo.(Aquellos que hayan cambiado desde el último respaldo completo)		respaldo completo y del último respaldo diferencial	discos comparado con respaldos de incremento
--	--	--	---	--

En la tabla anterior se destacan los diferentes tipos de respaldos a lo cual se une las ventajas y desventajas de los mismos. Con el respaldo completo la primera vez que se haga salva y el respaldo de incremento se puede tener una empresa con un respaldo de información seguro, puesto que se tienen todos los datos salvados, pero para las empresas desarrolladoras de software no sólo hacen falta este tipo de respaldo, pues se deben tener distintas salvas de otras versiones por cambios de negocio que puedan surgir en el desarrollo del/los software de la empresa.

### 1.8 Software de respaldo y respaldo "En Línea"

Existen algunos software y servicios que ayudan a mantener un orden en los respaldos de las empresas, los cuales se pueden clasificar en:

- Software de respaldo tradicional: Con estos productos, se pueden elegir los archivos o carpetas a guardar, seleccionar un dispositivo de almacenamiento, y ejecutar el respaldo sin ayuda. También ofrecen soporte para una gran variedad de dispositivos de almacenamiento, que incluyen cintas y discos duros.
- Software de respaldo de fondo: Ideal para los usuarios que no tienen una "disciplina" en respaldar su información. Estos programas hacen una copia de los archivos en forma automática.
- Servicios de respaldo en Internet: tienen muchas ventajas, pues guardan la información fuera del lugar de trabajo y evitan tener que intercambiar medios.

Para esto, se contrata un plan y la compañía asigna cierta capacidad, (estos son los costos aproximados): [12]

*Tabla #3: Costos de planes de copia de seguridad*

Conectado TLM	Ibackup	Punto respaldo	Disco Web
US\$ 14.95 / mes	US\$ 10 / mes	US\$ 579.95 /mes	US\$ 1500 / mes
4GB	500 MB	1GB	20MB

## Capítulo 1. Tendencias de desarrollo de soluciones para implantar salvadas automáticas.

Plan "más conveniente" de Punto Respaldo (capacidad / precio), pero además existen estas opciones:

Tabla #4: Costos de planes de copia de seguridad

Plan	Precio anual (USD)	Precio semestral (USD)
100 MB	109.95	56.95
200 MB	197.95	99.95
300 MB	273.95	139.95
400 MB	349.95	179.95
500 MB	389.95	199.95
1 GB	579.95	299.95

Precio plan básico Disco Web, además se pueden contratar MB adicionales a los siguientes precios:

Tabla #5: Costos de plan básico Disco Web

Plan Adicional	MB Adicionales	Precio Venta(*)
1	20	\$1230
2	40	\$2460
3	60	\$3690
4	80	\$4920
5	100	\$6000
6	120	\$7200
7	140	\$8400
8	160	\$9600
9	180	\$10800

10	200	\$11550
----	-----	---------

Como se puede observar los servicios de respaldo en internet son algo costosos dados estos últimos precios de venta. El respaldo que se desea con esta solución propuesta no traería ningún costo a los clientes finales, pues se quiere una aplicación de código abierto con un sistema de respaldo con licencia GPL.

Por tanto se tiene claro que hasta ahora se quiere un sistema libre de costo monetario para las copias de seguridad y que cubra las necesidades de continuidad de las empresas.

Estos sistemas deben tener las siguientes características:

- Permitir copias continuas de los datos cambiantes, superando el concepto de respaldo tradicional.
- Permitir la recuperación local o no en caso de emergencia.
- Mejorar los niveles de servicio reduciendo costes.
- Liberar a las empresas de complejidad, gestión no productiva.

### 1.9 Lista de software para respaldo

En el mundo existe una gran variedad de sistemas dedicados al respaldo de datos. En la tabla 6 se reflejan las herramientas más destacadas de este tipo en el software libre así como la licencia bajo la cual fueron liberadas.

*Tabla #6: Programas de respaldo de software libre*

Redes Extensas	Licencias
Bacula	GPL
BackupPc	GPL
AMANDA	BSD
Redes pequeñas	Licencias
RSync	GPL
Duplicity	GPL
Dump	GPL

Tar	GPL
CPIO	GPL
DAR	GPL
Rdiff-Backup	GPL
Rsnapshot	GPL
Dirvish	OSL
<b>Sistemas Locales</b>	<b>Licencias</b>
Areca Backup	GPL
Cobian Backup	MPL
FlyBack	GPL
Mondo	GPL

En la tabla 7 se reflejan las más destacadas en el mundo privativo [13]. Para poder llevar a cabo los respaldos, es necesario tener el software de respaldo apropiado. Este software no solamente debe ser capaz de realizar la tarea de hacer copias en una media de respaldo, también debe tener en cuenta las necesidades y limitantes de su organización.

*Tabla #7: Programas de respaldo de software privativo*

<b>Redes Extensas</b>
Atempo TIMEnavigator
Bakbone NetVault
CommVault Systems Galaxy
EMC Legato Networker
<b>Redes pequeñas</b>
CommVault Systems
EMC Corporation Retrospect
Symantec Backup Exec
Unitrends
VisionWorks Solutions

Sistemas Locales
Mac Backup
Backup4all
Genie Backup Manager
Macrium Reflect

En los últimos tiempos la tendencia de Cuba en el campo de la informática es hacia los sistemas de software libre, por lo tanto de la tabla anterior (tabla 6) se investigarán los sistemas de respaldo que cuentan con licencias GPL.

Debido al carácter que tienen los sistemas de la tabla 6 y teniendo en cuenta que el sistema puede ser desplegado en una empresa con red grande o pequeña, se tomaron en cuenta sólo las herramientas de redes grandes ideales para administradores de sistema, ya que hacen posible la realización de copias de seguridad entre ordenadores con distintos sistemas operativos. Estos sistemas de respaldo se pueden ejecutar en los sistemas operativos: Linux, Solaris, FreeBSD, NetBSD, Windows, Mac OS X, OpenBSD.

### 1.9.1 Descripción de la herramienta para realizar respaldos

**BackupPC** es un software libre suite con una interfaz basada en web. El servidor de la cruz-plataforma puede funcionar en cualquier sistema operativo Linux , Solaris o UNIX basado en servidor. Ningún cliente es necesario, como el propio servidor es un cliente para varios protocolos que se manejan por otros servicios nativos del sistema operativo cliente. En 2007, BackupPC se mencionó como uno de los tres programas más conocidos de copia de seguridad de código abierto [14] a pesar de que es una de las herramientas que son "tan sorprendente, pero, por desgracia, si nadie habla de ellos, muchas personas nunca van a oír hablar de ello " [15].

BackupPC incorpora un Server Message Block (SMB) que puede ser usado para copia de seguridad de cuota de red en equipos que ejecutan Windows. Paradójicamente, en una configuración del servidor de BackupPC se encuentra detrás de un NAT'd firewall de Windows, mientras que la máquina opera en una dirección IP pública. Si bien esto puede no ser aconsejable para el tráfico SMB, es más útil para los servidores web corriendo SSH con GNU tar y rsync disponibles, ya que permite que el servidor BackupPC que se almacena en una subred independiente del servidor web de la DMZ [16].

El sistema BackupPC es uno de los más potentes pero tiene varias desventajas. Primeramente existe muy poca documentación para realizar los volcados sobre cinta y la configuración mediante SAMBA sin tener que usar el rsync para todos los clientes está en proceso de desarrollo.

**AMANDA**, Advanced Maryland Archiver automática de disco de red, es una solución de copia de seguridad que permite al administrador de TI para configurar un servidor maestro de copia de seguridad único para respaldar varios hosts en la red de unidades de cinta / cambiadores o los discos o medios ópticos. Amanda utiliza los servicios nativos y formatos (por ejemplo, descarga y / o GNU tar) y copias de seguridad de un gran número de servidores y estaciones de trabajo con múltiples versiones de Linux o Unix. Amanda utiliza un cliente nativo de Windows para copias de seguridad de Microsoft de escritorio y servidores Windows [17].

Unas de las principales desventajas de AMANDA es la poca documentación que existe, teniendo en cuenta la diversa documentación de otros software de respaldo. Otros de los problemas con que cuenta AMANDA es que dentro de sus configuraciones no se pueden hacer las exclusiones dentro de los ficheros que se le quieren hacer la salva.

Estas tres herramientas son muy potentes en cuanto a la seguridad que brindan en el respaldo de información, pero tienen una particularidad en común que la convierte en desventaja y es la configuración de sus ficheros manualmente, que de una u otra forma le quitan valor agregado al software de copia de seguridad por lo tedioso que puede resultar este proceso.

**Bacula**, es una colección de herramientas de respaldo, capaz de cubrir las necesidades de respaldo de equipos bajo redes IP. Se basa en una arquitectura Cliente-servidor que resulta eficaz y fácil de manejar, dada la amplia gama de funciones y características que brinda; copiar y restaurar ficheros dañados o perdidos. Además, debido a su desarrollo y estructura modular, Bacula se adapta tanto al uso personal como profesional, para parques de ordenadores muy grandes [18].

Todas las tareas que realiza Bacula se han modularizado y repartido, entre varios dominios y servicios. Trabaja bajo Linux, la familia BSD, Solaris y Windows, tiene soporte para base de datos tanto MySQL, PostgreSQL y SQLite [19].

El servicio de almacenamiento de Bacula consiste en un software que permite almacenar y recuperar los atributos de los archivos y datos al backup físico u otro

volumen. En otras palabras, el Bacula storage es el responsable de leer y escribir en las cintas (u otro dispositivo de almacenamiento que se utilice).

Bacula es un software con muchas facilidades para la integración con otros sistemas de software y su arquitectura tan flexible lo hace ser muy poderoso a la hora de desplegarlo, por tanto a continuación se detallará el porqué se escoge este sistema para la integración con el software propuesto.

Varios programas componen Bacula. La mayoría de estos programas corren en modo background y se utiliza alguna consola para interactuar con ellos. Resta ver cómo fluye la información entre tantos programas.

Cuando no es necesario que todas las partes de un sistema corran en la misma máquina, se dice que ese sistema es distribuido. Y siempre, para comunicar cada parte del sistema para que actúe como un todo, se necesita alguna clase de red. En este caso, Bacula corre sobre una red TCP/IP (Que son las más comunes, así que no es razón para preocuparse).

Cada programa espera las órdenes dadas por el director. Por ejemplo, cuando se ejecuta un backup, el Director se lo ordena al cliente correspondiente y al Storage. Luego el flujo de datos es directo entre Storage y cliente, sin pasar por el Director.

Hasta el momento se concluyen dos cosas importantes:

1. El tráfico de datos es directo entre cliente y el demonio de almacenamiento (storage), sin pasar por el director. Esto es importante porque, en definitiva, el segmento de red que se debe optimizar es el que va del cliente al demonio de almacenamiento, y no del cliente al director o del director al storage.
2. La otra cuestión a tener en cuenta es que el storage y el director pueden ejecutarse sobre la misma máquina. En realidad, esta es una decisión de diseño, ya que la arquitectura de Bacula es flexible y se adapta a las necesidades de cada empresa

La red física subyacente debe concordar con el diseño elegido para la implementación de Bacula. De manera explícita, hay que intentar situar "juntas" y sobre enlaces de alta velocidad a las máquinas que transferirán datos, mientras que el director puede estar en otra máquina conectada a una velocidad más baja.

**Capítulo 1. Tendencias de desarrollo de soluciones para implantar salvas automáticas.**

Uno de los aspectos más importantes a tratar con la herramienta seleccionada y las anteriores es que con Bacula se pueden hacer volcados físicos a cintas, se pueden manejar las exclusiones de los ficheros que se le quieren hacer salvas y se pueden configurar mediante SAMBA o SSH los archivos de los clientes y servidores.

A continuación se muestra en la tabla 8 un resumen de las características de Bacula.

*Tabla #8: Características de Bacula*

Características	Disponible	Detalle	Comentarios
Código abierto	Si		
Multiplataforma	Si	Servidor requiere Linux/Unix. Soporta cualquier cliente.	
Compresión	No		
Encriptación	Si	MD5, SHA1	
Filtros	Si	Extensión, subdirectorio	Dispone tanto de filtros de inclusión como de exclusión.
Tipos de Respaldo	Si	Completo, Incrementa, Diferencial.	
Modo de almacenamiento	Si	En volúmenes.	Un volumen es un repositorio simple para la información respaldada.
Lugar de almacenamiento.	Si	Discos, cintas, CDs, DVDs., USBs	
Respaldo de permisos y ACL de archivos	Si	Requiere habilitar VSS en clientes windows.	
Planificador de tareas	Si		
Eficiencia y facilidad para restaurar respaldos	Si		

Utilización de estándares para respaldo (tar, dump, zip).	Si	gzip	
Cliente-Servidor			
Buena documentación	Si		
Soporte para respaldo de bases de datos on-line (hot-backup)	Si	En Windows utiliza vss y en Linux usa scripts previstos por la comunidad Bacula.	

### 1.10 Metodología y tecnologías utilizadas para el desarrollo del software

- **Metodología RUP.**

El Proceso Unificado de Rational (Rational Unified Process en inglés, habitualmente resumido como **RUP**) es un proceso de desarrollo de software y junto con el Lenguaje Unificado de Modelado (**UML**), constituye la metodología estándar más utilizada para el análisis, implementación y documentación de sistemas orientados a objetos. Cuenta con las mejores prácticas del modelo de desarrollo de un software en particular [20].

El RUP está basado en 5 principios clave que son: adaptar el proceso, balancear prioridades, demostrar valor iterativamente, elevar el nivel de abstracción y enfocarse en la calidad.

El ciclo de vida RUP es una implementación del desarrollo en espiral. El ciclo de vida organiza las tareas en fases e iteraciones.

RUP divide el proceso en cuatro fases: Inicio, Elaboración, Construcción y Transición, dentro de las cuales se realizan varias iteraciones en número variable según el proyecto y en las que se hace un mayor o menor hincapié en las distintas actividades. Las primeras iteraciones (en las fases de Inicio y Elaboración) se enfocan hacia la comprensión del problema y la tecnología, la delimitación del ámbito del proyecto, la eliminación de los riesgos críticos, y al establecimiento de una línea base de la arquitectura. En la fase de construcción, se lleva a cabo la construcción del producto por medio de una serie de iteraciones. En la fase de transición se pretende garantizar que se tiene un producto preparado para su entrega a la comunidad de usuarios.

Principales características de RUP:

- Forma disciplinada de asignar tareas y responsabilidades (quién hace qué, cuándo y cómo)
  - Pretende implementar las mejores prácticas en Ingeniería de Software
  - Desarrollo iterativo
  - Administración de requisitos
  - Uso de arquitectura basada en componentes
  - Control de cambios
  - Modelado visual del software
  - Verificación de la calidad del software
- **Programación extrema (XP).**

La programación extrema (XP) es el más destacado de los procesos ágiles de desarrollo de software. Al igual que éstos, la programación extrema se diferencia de las metodologías tradicionales principalmente en que pone más énfasis en la adaptabilidad que en la previsibilidad. Se puede considerar la programación extrema como la adopción de las mejores metodologías de desarrollo de acuerdo a lo que se pretende llevar a cabo con el proyecto, y aplicarlo de manera dinámica durante el ciclo de vida del software.

Los principios básicos de la programación extrema son: Simplicidad, Comunicación, Retroalimentación.

Las características fundamentales del método son:

- Desarrollo iterativo e incremental.
- Pruebas unitarias continuas, frecuentemente repetidas y automatizadas, incluyendo pruebas de regresión.
- Programación en parejas.
- Frecuente integración del equipo de programación con el cliente o usuario.

- Corrección de todos los errores antes de añadir nueva funcionalidad. Hacer entregas frecuentes.
- Refactorización del código, es decir, reescribir ciertas partes del código para aumentar su legibilidad y mantenibilidad pero sin modificar su comportamiento.
- Propiedad del código compartida.
- Simplicidad en el código.

La simplicidad y la comunicación son extraordinariamente complementarias. Con más comunicación resulta más fácil identificar qué se debe y qué no se debe hacer. Mientras más simple es el sistema, menos tendrá que comunicar sobre este, lo que lleva a una comunicación más completa, especialmente si se puede reducir el equipo de programadores.

#### **Justificación de la selección realizada.**

Luego de analizar dos de las metodologías más utilizadas se decide utilizar RUP pues a pesar de que este pudiera parecer un proyecto de corto plazo, representa un subsistema de un proyecto de larga duración por la que necesita una metodología robusta para su desarrollo, además de que debe existir uniformidad entre todos los subsistemas, por lo que no puede utilizarse una metodología ligera en algunos módulos y una pesada en otros.

- **Lenguaje de programación.**

Se utilizará **PHP** como lenguaje de programación para la implementación del sistema, debido a que es un lenguaje "open source" interpretado de alto nivel embebido en páginas HTML, y su código es ejecutado en el servidor. Este lenguaje permite no sólo leer, sino también manipular datos desde diversas fuentes, incluso los datos que pueden ser ingresados por los usuarios desde formularios HTML. La meta del lenguaje es permitir rápidamente a la comunidad de desarrolladores la generación dinámica de páginas Web. [21]

El código PHP es mucho más legible que el de PERL, este último lenguaje consume muchos recursos de la máquina haciéndolo menos ligero. PHP no es el único lenguaje de programación cliente/servidor Web que existe, entre las alternativas más conocidas esta ASP por parte de Microsoft, el cual corre bajo el servidor Web propietario de Microsoft: IIS, debido a esto es de pago, por lo que es muy poco usado y aunque es

muy potente su metodología de programación es poco robusta y no logra superar a PHP como lenguaje de programación Web.

Para desarrollar un sistema de este tipo, más por lo extenso que por la variedad de reglas de negocio que pueda tener se necesita un marco de trabajo. Éste simplifica el desarrollo de una aplicación mediante la automatización de algunos de los patrones utilizados para resolver las tareas comunes; proporciona estructura al código fuente, forzando al desarrollador a crear un código más legible y fácil de mantener; facilitando programación de aplicaciones, ya que encapsula operaciones complejas en instrucciones sencillas. [22]

En el mundo del desarrollo de software existen varios marcos de trabajo, algunos libres y otros privados, pero al tener PHP como lenguaje de programación en el lado del servidor se escoge Symfony; pues su marco de trabajo es diseñado para optimizar, gracias a sus características, el desarrollo de las aplicaciones web. [23] Para empezar, separa la lógica de negocio, la lógica de servidor y la presentación de la aplicación web. Proporciona varias herramientas y clases encaminadas a reducir el tiempo de desarrollo de una aplicación web compleja.

Además, automatiza las tareas más comunes. Y permite al desarrollador dedicarse por completo a los aspectos específicos de cada aplicación. El resultado de todas estas ventajas es que no se debe reinventar la rueda cada vez que se crea una nueva aplicación web.

Symfony está desarrollado completamente con PHP 5. Ha sido probado en numerosos proyectos reales y se utiliza en sitios web de comercio electrónico de primer nivel. Es compatible con la mayoría de gestores de bases de datos, como MySQL, PostgreSQL, Oracle y SQL Server de Microsoft [24].

- **Sistemas gestores de bases de datos.**

### **PostgreSQL.**

PostgreSQL ofrece muchas ventajas para su compañía o negocio respecto a otros sistemas de bases de datos:

- Este software ha sido diseñado y creado para tener un mantenimiento y ajuste mucho menor que los productos de los proveedores comerciales, conservando todas las características, estabilidad y rendimiento.

- En contraste a muchos sistemas de bases de datos comerciales, es extremadamente común que compañías reporten que PostgreSQL nunca ha presentado caídas en varios años de operación de alta actividad.
- El código fuente está disponible para todos sin costo.
- PostgreSQL está disponible en casi cualquier Unix (34 plataformas en la última versión estable), y una versión nativa de Windows está actualmente en estado beta de pruebas.
- Diseñado para ambientes de alto volumen.

PostgreSQL usa una estrategia de almacenamiento de filas llamada MVCC para conseguir una mucha mejor respuesta en ambientes de grandes volúmenes. Los principales proveedores de sistemas de bases de datos comerciales usan también esta tecnología, por las mismas razones.

#### **Manejador de bases de datos. PgAdmin.**

Todas las bases de datos construidas necesitan ser administradas y qué mejor para cumplir con dicho objetivo que un programa con interfaz gráfica. Para PostgreSQL existen muchas herramientas de propósito general que administran sus bases de datos entre las que se encuentra PgAdmin una de las más utilizadas por la comunidad de desarrollo del software libre. PgAdmin: diseña, mantiene y administra fácilmente las bases de datos construidas en PostgreSQL, la cual funciona en distintos sistemas operativos como Windows 95/98, NT, Xp, así como en plataformas libres. Está liberada bajo la licencia Open Source. Es la herramienta más popular y completa, diseñada para responder a las necesidades de los usuarios permitiéndoles escribir desde simples consultas y sentencias SQL, hasta diseñar complejas bases de datos. Algunas de las características de PgAdmin son:

- Entradas SQL aleatorias.
- Pantallas de información y ayudas para bases de datos, tablas, índices, secuencias, vistas, programas de arranque, funciones y lenguajes.
- Preguntas y respuestas para configurar usuarios, grupos y privilegios.
- Control de revisión con mejora de la generación de script.
- Ayudas para importar y exportar datos.

- Ayuda para migrar bases de datos.

## **Oracle**

Oracle es sin duda una de las mejores bases de datos que se tiene en el mercado, es un sistema gestor de base de datos robusto, tiene muchas características que garantizan la seguridad e integridad de los datos; haciendo que las transacciones se ejecuten de forma correcta, sin causar inconsistencias; ayuda a administrar y almacenar grandes volúmenes de datos; estabilidad y escalabilidad.

Oracle está disponible en múltiples plataformas como Windows y Linux. La naturaleza multiplataforma de Oracle, lo convierte en una verdadera solución empresarial.

Aunque su dominio en el mercado de servidores empresariales ha sido casi total, recientemente sufre la competencia de gestores de bases de datos comerciales y de la oferta de otros con licencia Software Libre como PostgreSQL y MySQL. Las últimas versiones de Oracle han sido certificadas para poder trabajar bajo Linux.

### **Justificación de la selección realizada**

Se decide utilizar Postgres SQL por las razones siguientes: posee una instalación ilimitada, es decir, nadie puede demandarlo por violar acuerdos de licencia, puesto que no hay costo asociado a la licencia del software.

Tiene todas las características de los SGBD modernos, como son las llaves foráneas, vistas, procedimientos almacenados, varios tipos modernos de datos, entre otros [25].

Además posee ahorros considerables en costos de operación puesto que ha sido diseñado y creado para tener un mantenimiento y ajuste mucho menor que otros productos, conservando todas las características, estabilidad y rendimiento. Es extensible ya que el código fuente está disponible para todos sin costo. Es multiplataforma puesto que está disponible en casi cualquier Unix (34 plataformas en la última versión estable), y ahora en versión nativa para Windows.

### **1.11 Conclusiones parciales**

En este capítulo se trataron diferentes temas referidos a las salvvas automáticas.

Se definieron las necesidades de las copias de seguridad de una forma segura, rápida y eficiente, quedando expuesto que la herramienta a obtener tendría que cumplir con las exigencias de ser confiable, rápido y eficiente.

Se definieron los distintos volúmenes de información que pueden ser salvados con una herramienta de respaldo, quedando expuesto que la herramienta escogida debe de adaptarse a cualquier tipo de volúmenes de información, pues se pretende que este trabajo cumpla todas las necesidades de cualquier tipo de empresa.

Uno de los aspectos más importantes tratado fue la enumeración de los distintos dispositivos de respaldo, tecnologías, dispositivos de almacenamiento, así como costos a nivel mundial de estos dispositivos y la justificación de la selección de una herramienta con licencia GPL para redes extensas.

Se realizó una comparación de las distintas herramientas de redes extensas con licencia GPL para hacer salvas automáticas de información donde se obtuvo como resultado la herramienta Bacula, fundamentando su aval de ser seleccionada.

Se investigaron las características de configuración de Bacula para que la solución propuesta se integre correctamente con este software de respaldo.

Por último se realizó un análisis de las distintas metodologías y tecnologías para el desarrollo del software que darán solución al problema planteado, donde se obtuvieron las siguientes: PHP – Symfony como marco de trabajo de desarrollo, RUP como metodología para el desarrollo y PostgreSQL como gestor de base de datos.

## CAPÍTULO 2. SOLUCIÓN PROPUESTA.

### INTRODUCCIÓN

El presente capítulo integra la aplicación de software con Bacula, teniendo en cuenta las pautas de instalación y configuración con que cuenta la herramienta de respaldo.

En este capítulo se verán todas las configuraciones de Bacula y se describirán todas las dificultades para su instalación de forma manual. Se mostrará la arquitectura y una investigación detallada de los elementos que componen Bacula, así con la integración de la aplicación con cada uno de los elementos enunciados. Por otra parte se verá la polémica del versionado que tiene la herramienta de respaldo y la solución que muestra el software propuesto.

### 2.1 Arquitectura de Bacula

Bacula es un programa para hacer copias de seguridad de un ordenador, el mismo se basa en una colección de demonios que cooperan entre sí para realizar copias de respaldo de los archivos necesarios, sean de la máquina que sea. Para interactuar con este programa se necesita un elemento más: la consola. Todos estos elementos son independientes entre sí y pueden estar en máquinas distintas, así pues el principal problema a la hora de configurar este sistema de respaldo consiste en hacer que todos estos elementos se comuniquen correctamente entre sí. A continuación la figura 2 muestra la estructura de ficheros que propone Bacula para su funcionamiento.

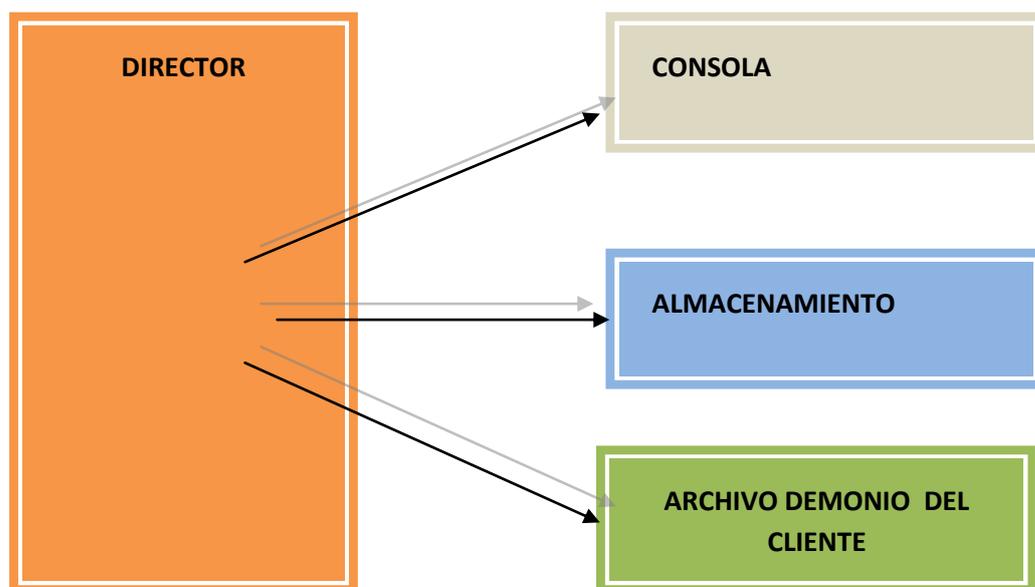


Figura 2: Componentes de Bacula

El servicio de almacenamiento de Bacula consiste en un software que permite almacenar y recuperar los atributos de los archivos y datos al backup físico u otro volumen. En otras palabras, el demonio de almacenamiento es el responsable de leer y escribir en las cintas (u otro dispositivo de almacenamiento que se utilice).

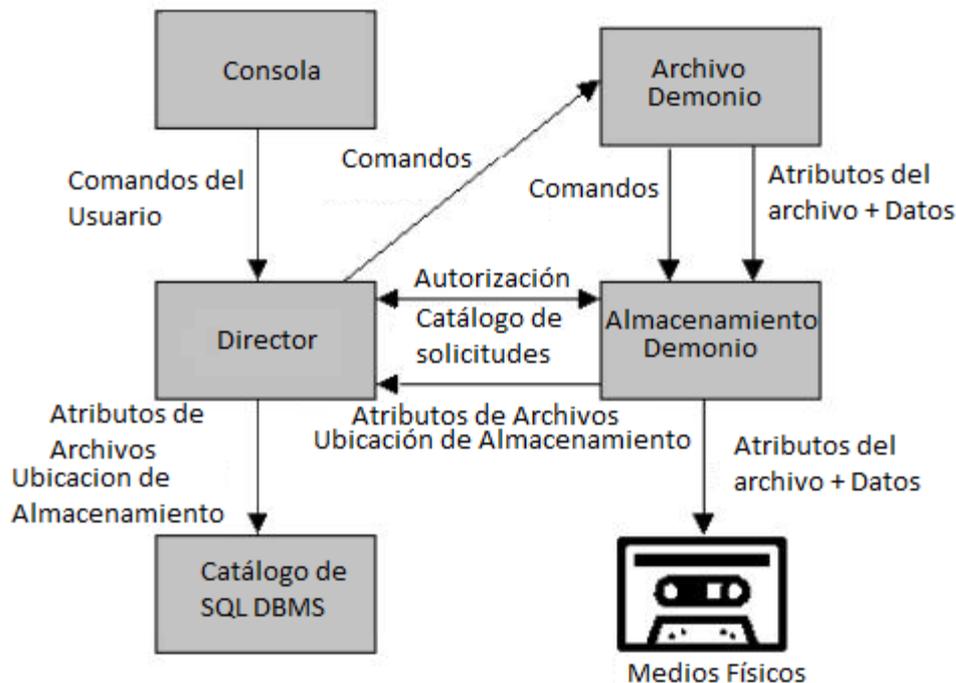


Figura 3: Arquitectura de Bacula

En la figura 3 se muestra totalmente la arquitectura de Bacula. En esta solución propuesta el catálogo SQL, no será utilizado pues se tendrá un gestor de bases de datos conectado al servidor de bases de datos con toda la información de cada uno de los ficheros que intervengan en la red de salvadas.

Es importante destacar que:

- Bacula se ha compilado y probado en los sistemas OpenSuSe Linux, FreeBSD y Solaris, para su compilación requiere GNU C++ versión 2.95 o superior.
- Dependiendo del gestor de base de datos elegido, se necesita como mínimo la versión MySQL 4.1, PostgreSQL 7.4, SQLite 3.
- La consola para GNOME ha sido desarrollada y probada en GNOME 2.x, ya no existe soporte para GNOME 1.4.

- El Tray Monitor se ha desarrollado para GTK+-2.x. Precisa GNOME menor o iguala la versión 2.2, KDE mayor o igual a la versión 3.1 o cualquier gestor de ventanas que soporte el estándar FreeDesktop System tray.

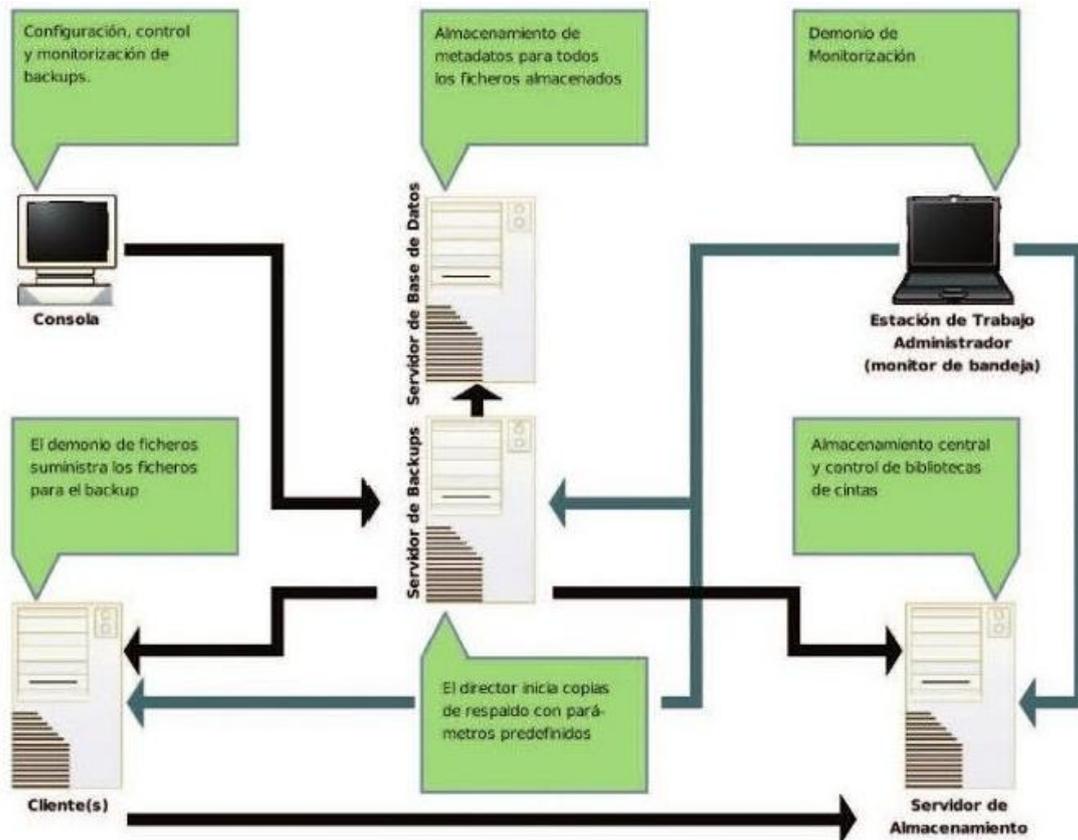


Figura 4: Elementos de Bacula

Los elementos necesarios para que Bacula funcione son:

### 2.1.1 Bacula-director

Es el demonio que gestiona la lógica de los procesos de backup y los demás servicios. El servidor de la base de datos debe estar accesible desde la máquina que ejecuta este demonio (o también puede estar en la misma máquina y escuchar en localhost) [26].

En el archivo de configuración de este demonio se especifica dónde y cómo acceder al resto de demonios y recursos, la contraseña para el acceso mediante Bacula-console y los trabajos o jobs. Ver figura 5.

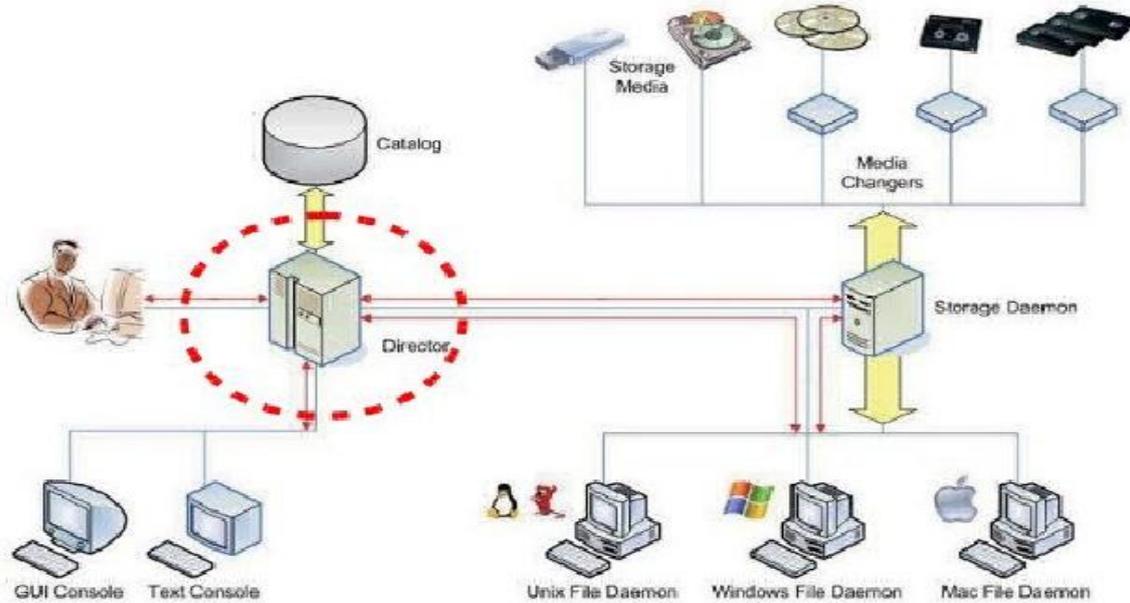


Figura 5: Bacula Director.

### 2.1.2 Bacula-storage daemon

Este demonio es el encargado de manejar los dispositivos de almacenamiento; esto exige que este demonio esté instalado en la máquina que posea la conexión física a los dispositivos de almacenamiento, tales como: discos locales, grabadoras de CD o DVD, unidades de cinta, volúmenes NAS o SAN [27], autocargadores o librerías de cinta.

El fichero de configuración de este servicio define tanto los dispositivos de almacenamiento que maneja, como qué directores pueden utilizarlo [28]. Ver figura 6.

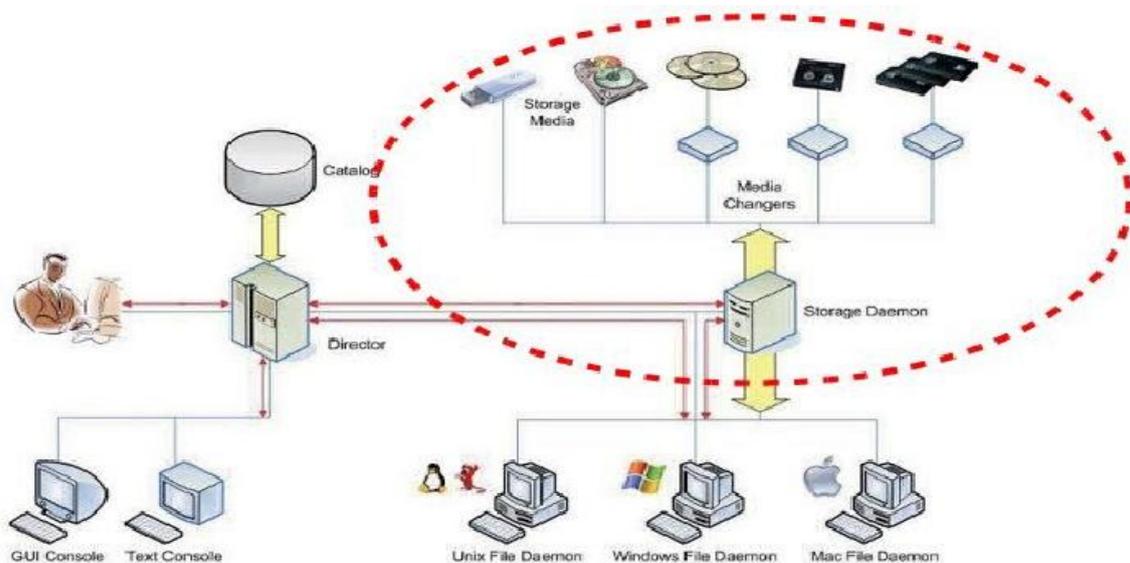


Figura 6: Bacula Storage

### 2.1.3 Bacula-file daemon

Mediante este demonio Bacula obtiene los ficheros que necesita respaldar, así pues éste es el componente que hay que instalar en las máquinas que necesiten respaldo. Realiza la misma función que los "agentes" en otros sistemas de backup.

Este archivo de configuración es el más simple de todos. Pues sólo contiene qué directores pueden realizarle peticiones.

Para poder interactuar con el servicio de backup, se necesitará un cliente [29]. Ver figura 7.

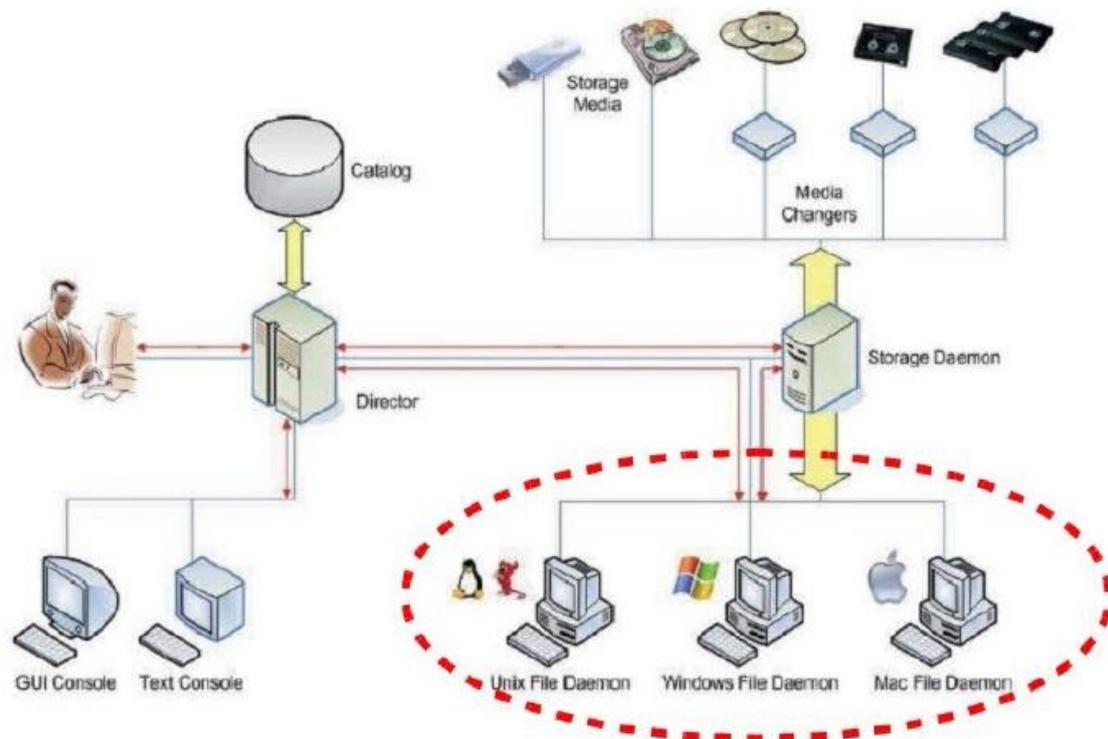


Figura 7: Bacula-file daemon

### 2.1.4 Bacula-console

Bacula-console, a veces llamado el agente de usuario, es un programa que permite que el usuario pueda interactuar con el demonio Bacula Director, mientras que el demonio se está ejecutando.

El actual Bacula-consola viene en dos versiones: una interfaz de shell (TTY estilo), y una interfaz gráfica de GNOME interfaz [30]. Ambos permiten al administrador o los usuarios autorizados interactuar con Bacula. Puede determinar el estado de un trabajo en particular, examinar el contenido del Catálogo, así como realizar determinadas manipulaciones de cinta con la consola del programa.

Además, hay una consola bwx wxWidgets construido con un gráfico que permite restaurar los archivos [31]. A partir de la versión 1.34.1 se encuentra en una fase temprana de desarrollo, pero ya es bastante útil. Lamentablemente, no ha sido mayor desde hace algún tiempo. Desde la consola del programa interactúa con el Director a través de la red, su consola y Director programas no necesariamente tienen que ejecutar en la misma máquina.

De hecho, un cierto conocimiento mínimo de la consola del programa es necesario para que Bacula pueda escribir en más de una cinta, porque cuando Bacula pide una nueva cinta, que espera hasta que el usuario, a través de la consola programa, indica que la nueva cinta está montada. [32] Ver figura 8.

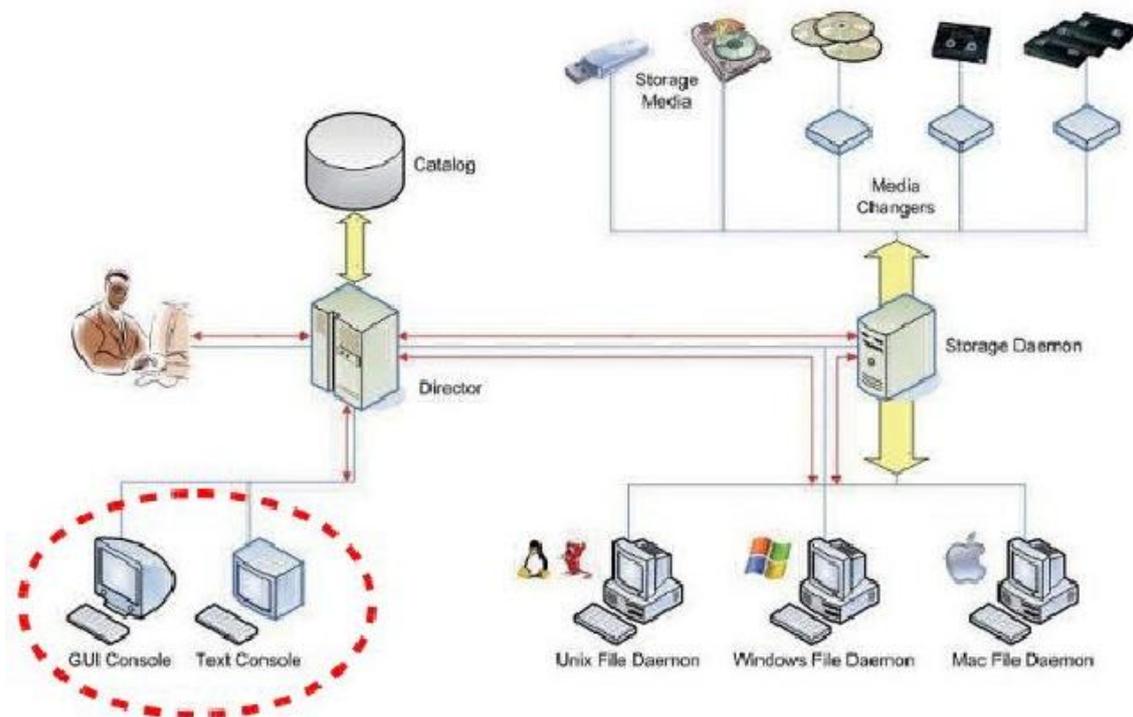


Figura 8: Bacula-console

Todo el conjunto de elementos que forman Bacula trabaja en sincronía y es totalmente compatible con bases de datos como MySQL, SQLite y PostgreSQL [33].

Cuando se desea configurar este software para realizar una salva automática se deben de crear varios ficheros de configuración por el usuario como son el *Bacula-dir.conf*, *Bacula-sd.conf* estos se encuentran en la raíz de la configuración, separando los demás también se deben configurar los *Clients* y *Devices*.

## 2.2 Configuración e instalación de Bacula

Las directivas que se encuentran en el *Bacula-dir.conf* y *Bacula-sd.conf* están relacionadas gracias a la directiva *Job* que define una colección de archivos (FileSet) para un cliente (Client), de acuerdo con una planificación de copias (Schedule) en un conjunto de cintas (Pool) de un determinado dispositivo (Storage). Ver anexos 2 y 3 respectivamente.

### 2.2.1 Propuesta de componentes y arquitectura de Bacula

Anteriormente se explicaba que el director es el fichero más importante dentro de la arquitectura de Bacula pues es el demonio que gestiona la lógica de los procesos de backup y los demás servicios.

En el director se guardan todas las configuraciones de los clientes que se le hará la salva de información como se muestran en los anexos 2 y 3. Las configuraciones de los clientes en el director puede resultar desventajoso a la hora de la integración con la solución propuesta, pues si se tienen todas las configuraciones de los clientes alojadas en un fichero único resulta muy difícil editar las etiquetas o configuraciones de un cliente específico. También puede resultar muy complicado eliminar un cliente del director pues son muchas las etiquetas a eliminar y tienen muchas similitudes con otras configuraciones de otros clientes, lo que podría tener consecuencias poco favorables para la salva de información en general.

La nueva arquitectura de ficheros de Bacula para esta solución quedaría planteada como se muestra la figura 9.

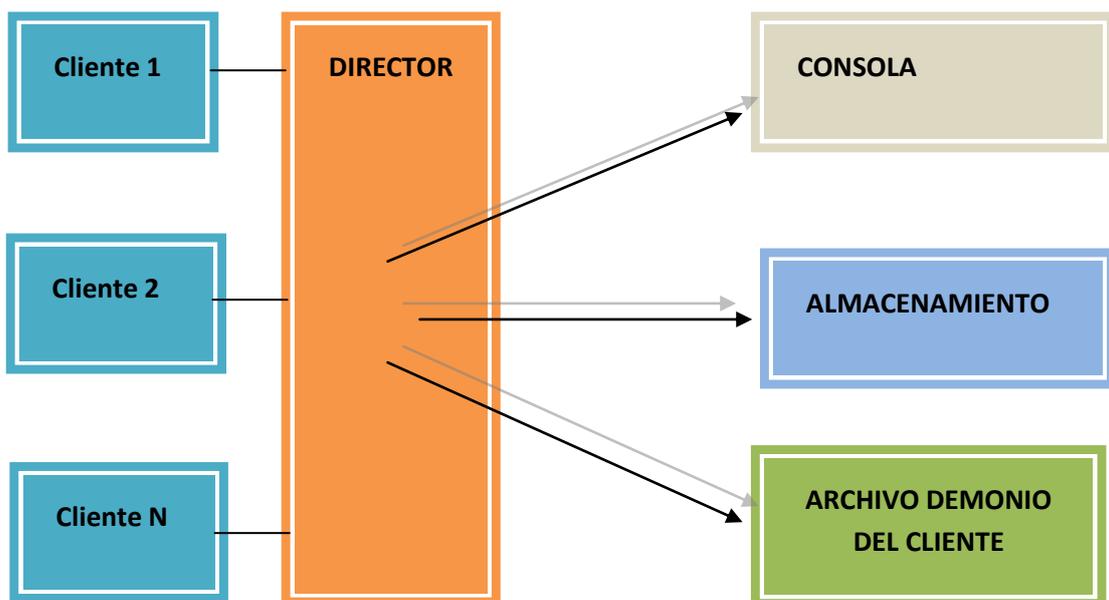


Figura 9: Propuesta de arquitectura de Bacula

Como se muestra en la figura 9 los clientes se comunicarán con el director mediante una configuración muy simple en el director. Por cada cliente que se tenga el director tendrá la ruta del mismo con un apuntador de la siguiente manera: `@/etc/bacula/conf.d/clienteN.conf`.

Se creó un directorio llamado “*conf.d*” donde se tendrán alojados todos los clientes para facilitar la sencillez de la arquitectura de ficheros antes propuesta.

### 2.2.2 Proceso de instalación de Bacula

A continuación se muestran los pasos para instalar Bacula. Este se debe hacer en este orden y sin alterar los pasos. Tener en cuenta también el aspecto de los permisos para ejecutar los mismos pues con un usuario que no es el de *root* será imposible terminar factiblemente la instalación. Los pasos son los siguientes:

1. Configurar los puertos de escucha del servidor.
2. Instalar Bacula-server.
3. Instalar Bacula-client.
4. Configurar archivo Bacula-fd.conf(Cliente de Bacula)
5. Configurar archivo Bacula-sd.conf(Storage Daemon)
6. Configurar archivo bconsole.conf(Console)
7. Configurar archivo Bacula-dir.conf(Director Daemon)
8. Probar los archivos de Bacula.
9. Ejecutar a Bacula-Server por primera vez.
10. Accediendo a Bacula desde bconsole.
11. Ejecutar el primer Respaldo Full y Diferencial.
12. Restaurar archivos con Bacula.
13. Trabajar con volúmenes desde bconsole.
14. Parámetros Extras de Bacula.
15. Notas finales.

Como se puede observar la configuración de cada software de salvadas automáticas se hace poco trivial por la cantidad de datos que hay que configurar (en rojo todos los datos configurables; ver los anexos 2 y 3).

Bacula puede ser complicado para un usuario con pocos o abundantes conocimientos informáticos a la hora de configurar todos los archivos y demonios. La integración con esta herramienta se basa en mantener la misma arquitectura propuesta pero los ficheros o demonios de cada cliente se escriben en la configuración del disco aparte

del Director, esto es importante pues si se desea adicionar, editar, eliminar o listar los clientes de un servidor dado, se hace mucho más sencillo a la hora de consultar los datos. Por otra parte la integración muestra a los clientes y servidores de forma separada, aunque existan clientes que sean servidores o viceversa.

### 2.3 Integración de la solución propuesta y Bacula

Este es el momento de integrar la aplicación propuesta con el Bacula. Las instalaciones, configuraciones, pruebas, salvas, respaldos y notificaciones con que cuenta el Bacula, ahora no se harán de manera manual y por consola.

La arquitectura de la solución propuesta es Cliente – Servidor y al quedar integrada con la de Bacula quedaría de la siguiente manera: Ver figura 10.

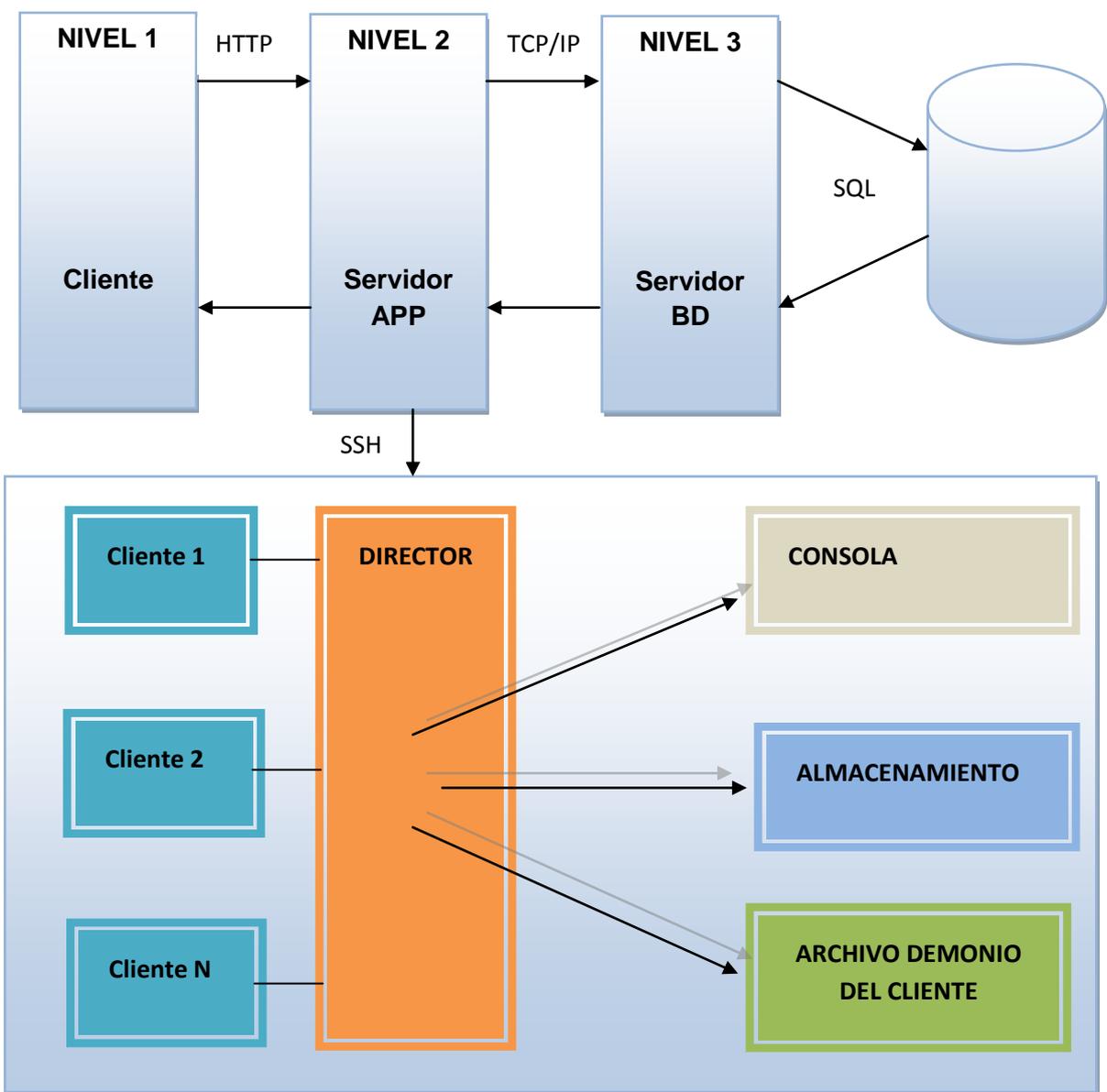


Figura 10: Arquitectura general de la solución propuesta

Como se muestra en la figura anterior la comunicación entre el servidor de aplicaciones y los componentes de Bacula será mediante el protocolo de comunicación SSH.

El protocolo SSH proporciona los siguientes tipos de protección:

- Después de la conexión inicial, el cliente puede verificar que se está conectando al mismo servidor al que se conectó anteriormente.
- El cliente transmite su información de autenticación al servidor usando una encriptación robusta de 128 bits.
- Todos los datos enviados y recibidos durante la sesión se transfieren por medio de encriptación de 128 bits, lo cual los hacen extremadamente difícil de descifrar y leer.

Por las características del protocolo SSH y los métodos soportados del lenguaje de programación PHP: `ssh2_auth_password` y `ssh2_exec`, que autentican a un usuario con los permisos pertinentes y ejecutan comandos simples o bash respectivamente, se hará más sencilla la comunicación de ambas arquitecturas.

### 2.3.1 Instalación de Bacula en el servidor.

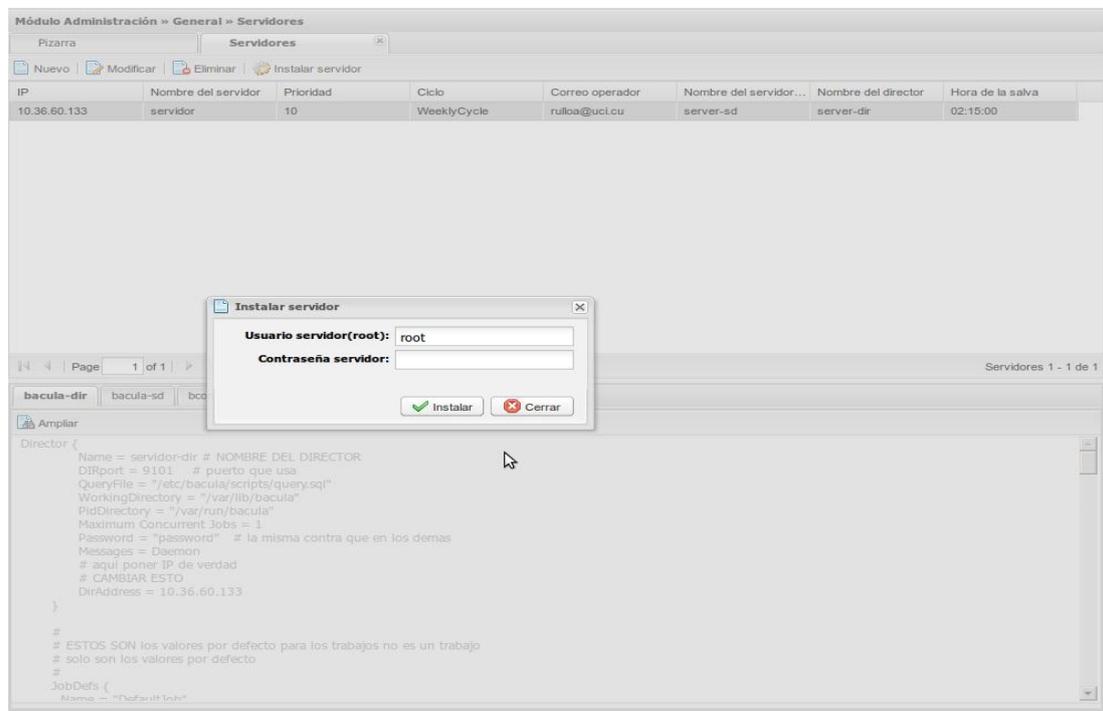


Figura 11: Módulo de administración. Instalación del servidor.

En la figura anterior se tiene una pantalla interfaz de la solución propuesta donde se tiene un listado de los servidores de Bacula guardados en el servidor de base de datos. Un vez seleccionado el servidor que se desea instalar se introduce el usuario y contraseña del ordenador donde se instalará el servidor Bacula.

Primeramente el servidor de aplicaciones hace una petición al servidor de base de datos y él le envía todos los datos necesarios para la correcta instalación del servidor de Bacula. Teniendo en el servidor de aplicaciones todos los datos, este procede a enviarlos por el protocolo SSH con la estructura que tienen los ficheros de configuración. Luego de hacer el envío al servidor de salvos, un comando bash instala y crea los directorios necesarios para la ejecución satisfactoria de los componentes de Bacula en el servidor. Ver figura 12.

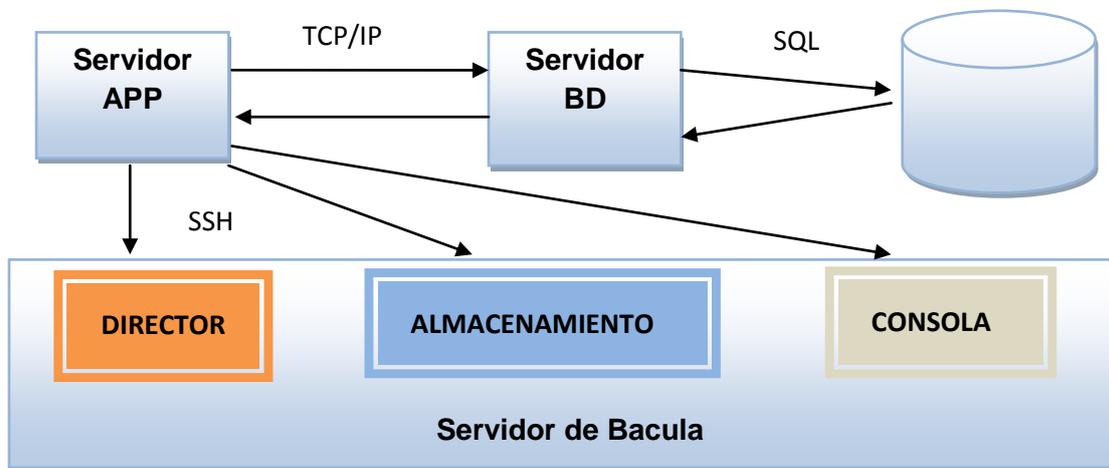


Figura 12: Integración con el servidor de Bacula

En la figura anterior se muestra que los ficheros a configurar en el servidor de Bacula son el *bacula-dir*, *bacula-sd* y *bconsole*.

El *bacula-dir* actúa como el controlador del sistema de copia de seguridad de red y es responsable de programar y coordinar las copias de seguridad a través de la red.

El *bacula-sd* es el encargado de manejar los dispositivos de almacenamiento; esto exige que este demonio esté instalado en la máquina que posea la conexión física a los dispositivos de almacenamiento.

El *bconsole* es el fichero más simple de todos y los datos que tiene son los mismos del director y se utiliza por si se desea acceder por consola al servidor de Bacula.

Teniendo en cuenta que cada cambio en el servicio de Bacula lleva un reinicio del demonio de ese sistema, el comando bash antes mencionado se encarga de reiniciar

el demonio Bacula para su correcta y definitiva instalación. En el Anexo 4 se muestra un diagrama de flujo con el proceso de instalación del servidor de Bacula.

### 2.3.2 Instalación del cliente de Bacula

Para que los equipos a respaldar puedan comunicarse con el servidor Bacula se debe instalar *bacula-fd* en el ordenador cliente, mediante este demonio Bacula obtiene los ficheros que necesita respaldar, éste es el componente que hay que instalar en las máquinas que necesiten respaldo. Realiza la misma función que los “agentes” en otros sistemas de backup.

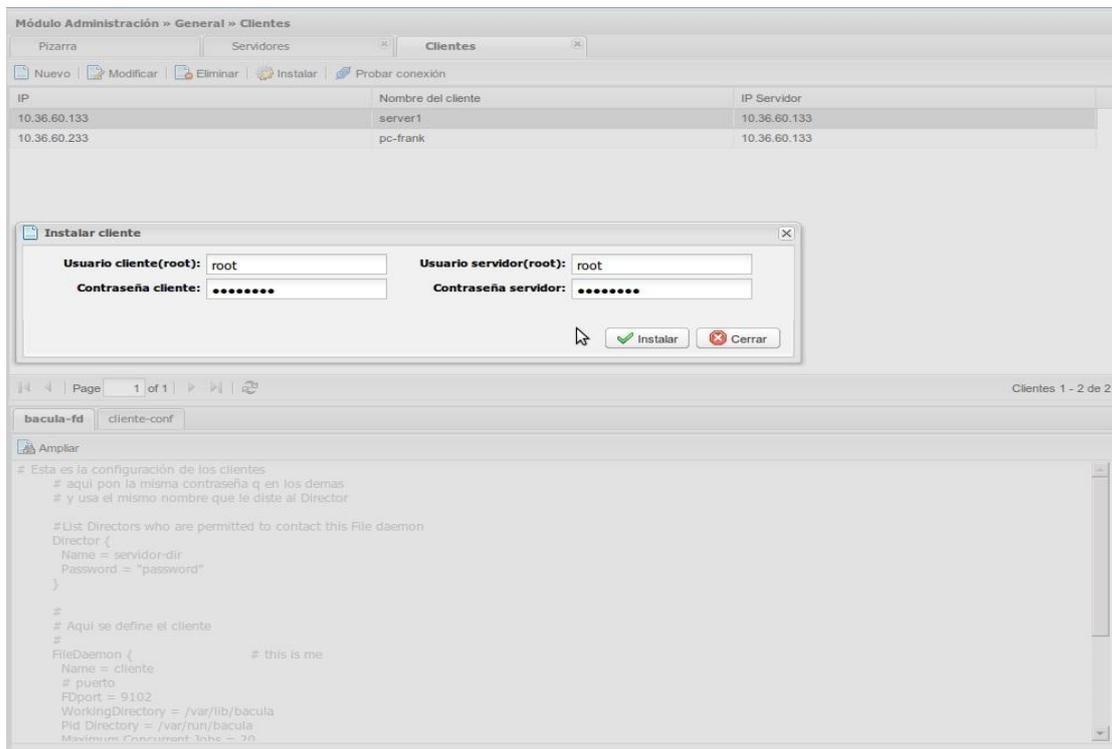


Figura 13: Módulo de administración. Instalación del cliente.

En la figura anterior se tiene una pantalla interfaz de la solución propuesta donde se tiene un listado de los clientes de Bacula guardados en el servidor de base de datos. Una vez seleccionado el cliente que se instalará se introduce el usuario y contraseña del ordenador donde se instalará el cliente de Bacula y el usuario y contraseña del servidor donde se hará el resguardo de la información de dicho cliente.

Teniendo en cuenta el proceso antes mencionado, no deja de ser muy diferente la primera parte en cuanto a la petición de información del servidor de aplicaciones con el servidor de base de datos. Es importante destacar que ahora en la instalación del cliente hay información que viaja a dos servidores, cliente y servidor de Bacula. La información necesaria para el cliente tendrá que viajar; entre otros datos viaja el IP del

ordenador donde se efectúa el resguardo de información. Esta dirección IP se encuentra como clave ajena en la entidad cliente de donde se hace la petición de los datos solicitados por el servidor de aplicaciones. Ver figura 14.

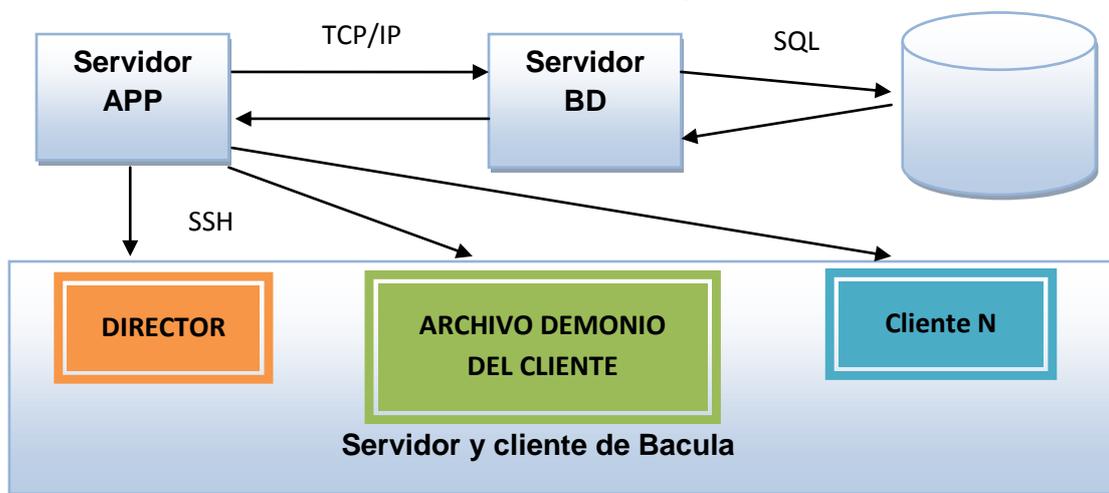


Figura 14: Integración con el servidor y cliente de Bacula

En la figura 14 se muestra el archivo demonio del cliente, y este debe instalarse en el ordenador del cliente de Bacula, ahora el archivo *clienteN.conf* será guardado en el directorio *conf.d* de la ruta */etc/bacula/conf.d/*. El director será actualizado en dependencia del nombre del cliente Bacula que fue guardado, al final del archivo se guardará el apuntador a dicho fichero.

Si se eliminara un cliente, no sólo ocurre en el servidor de base de datos, pues el servidor de aplicaciones luego de recibir la respuesta del servidor de base de datos, envía un comando bash al servidor Bacula actualizando los ficheros director y elimina el cliente que fue seleccionado. Luego del proceso de eliminación el comando bash antes mencionado tiene la responsabilidad de reiniciar el demonio Bacula para su configuración satisfactoria. En el Anexo 5 se muestra un diagrama de flujo con el proceso de instalación del cliente de Bacula.

### 2.3.3 Versionado de Bacula

Hoy en día existen muchas versiones de Bacula, tanto para el servidor como para los clientes. Todas estas versiones se pueden encontrar en una página en internet bien apoyada por la comunidad y muy estable por parte de sus administradores.

Cuando se intentan conectar una versión menor del servidor con una más actualizada en el cliente, se lanza una excepción de error diciendo que la conexión no se hará y que el servidor necesita una versión igual o superior a la del cliente.

Ahora se puede tener una posible solución de acuerdo a este problema del versionado y sería el compilado del código fuente de la versión del servidor y la posterior reconfiguración.

Primeramente hay que añadirle unas entradas al script de configuración para que compile correctamente, por tanto se decidió realizar un algoritmo de configuraciones con las opciones recomendadas.

Terminada la compilación sin errores, sólo queda por hacer la inicialización de la base de datos, así como la ejecución del siguiente script (que está en las fuentes de Bacula en la carpeta /etc/Bacula).

Con esto se concluye que con la aplicación propuesta se soluciona el problema del versionado del Bacula que a muchos usuarios le suele dar muchos errores y problemas de configuración, no sólo por Bacula sino por los servidores que muchas veces no son dedicados a una sola funcionalidad o tarea.

No obstante siguen existiendo problemas a nivel de repositorios pues en una distribución de Ubuntu 9.04 / 9.10 sólo puede instalarse la versión de Bacula 2.4.4 y esta ya tiene un versionado por 5.5.0.

### **2.3.4 Despliegue de la solución propuesta**

El despliegue de la solución propuesta se hará por partes. Primeramente el despliegue de los servidores de aplicación y gestor de base de datos, utilizando la arquitectura cliente /servidor, y luego el de los servidores local y central de Bacula.

La arquitectura cliente/servidor es un modelo para el desarrollo de sistemas de información, en el que las transacciones se dividen en procesos independientes que cooperan entre sí para intercambiar información, servicios o recursos. Se denomina cliente al proceso que inicia el diálogo o solicita los recursos al servidor. Este es el modelo de interacción más común entre aplicaciones en una red.

Uno de los principales aportes de esta arquitectura a los sistemas de información, es la interfaz gráfica de usuario. En esta arquitectura los datos se presentan, editan y validan en la parte de la aplicación cliente.

Todo esto tiene como fin que el usuario de un sistema de información soportado por una arquitectura cliente / servidor, trabaje desde su estación de trabajo con distintos datos y aplicaciones, sin importarle dónde están o dónde se ejecuta cada uno de ellos.

Utilizar esta arquitectura es de vital importancia pues con ella se puede acceder a cada estación de trabajo, desde cualquier sitio conectado a la red, donde se desee configurar un respaldo de información.

Antes de detallar el despliegue de la herramienta de respaldo se debe destacar que a los servidores locales de Bacula se le llaman comúnmente como servidores clientes de Bacula y los centrales como servidores de Bacula que atienden a los clientes.

Los clientes de Bacula estarán alojados en cada área productiva por lo que los mismos harán su respaldo con el/los servidor/es central/es de la empresa. Las empresas que cuenten con este tipo de despliegue garantizan que cada área de trabajo tenga un respaldo eficaz de su información.

La comunicación entre estos despliegues, tanto de la aplicación como del software de respaldo, se hará por un túnel SSH, teniendo una comunicación segura y confiable.

#### **2.4 Procedimiento de respaldo de información utilizando la solución propuesta**

En el capítulo anterior, acápite 1.4, se detallaba el procedimiento para hacer un respaldo de información. El procedimiento previo contaba con 12 procesos, y se llevaba a cabo utilizando cualquier tipo de sistema informático para respaldar la información de una empresa.

A continuación se hace una descripción del procedimiento utilizando la solución propuesta el cual se hace mucho más corto como se muestra en la tabla siguiente. En el Anexo 6 se muestra el diagrama de flujo de este procedimiento.

*Tabla #9: Procedimiento de respaldo de información utilizando la solución propuesta*

<b>Secuencia</b>	<b>Proceso</b>	<b>Responsable</b>
1 Introducir datos del servidor central de respaldo	1.1 Se introducen todos los datos necesarios que necesita un servidor central de respaldo. 1.2 Se comprueban que los datos estén guardados correctamente en el servidor de BD.	Dirección de redes
2 Instalar servidor central de respaldo	2.1 Se instala el servidor central, utilizando usuario y contraseña con los permisos de	Dirección de redes

	administración.	
3 Introducir datos del servidor local de respaldo	<p>3.1 Se introducen todos los datos necesarios que necesita un servidor local de respaldo.</p> <p>3.2 Se comprueban que los datos estén guardados correctamente en el servidor de BD.</p>	Dirección de redes
4. Instalar servidor local de respaldo	<p>4.1 Se instala el servidor local, utilizando usuario y contraseña con los permisos de administración tanto del servidor local como el central al que pertenece el mismo.</p> <p>4.2 Se prueba la conexión entre el servidor local y central para comprobar si la configuración ha sido satisfactoria.</p>	Dirección de redes

### **2.5 Conclusiones parciales**

Con la implementación de la aplicación de la solución propuesta las configuraciones se hacen de manera dinámica, es decir, llenando un formulario por un usuario capacitado se puede resolver que cada proyecto salve su información, quedando transparente para los usuarios finales los archivos de configuración.

Con la puesta en práctica de este sistema de salvas, la información se encontraría en los servidores locales de cada proyecto, sin necesidad de ir a cada uno a instalarlo y configurarlos; también estaría disponible en los servidores centrales donde sólo tenga acceso el personal adecuado.

Con este mecanismo de salva implementado, la información de todos los productos de la empresa se encontraría centralizada y actualizada, lo que traería consigo varias mejoras dentro del proceso de toma de decisiones por parte de los directivos. Se podría conocer el avance real de cada proyecto si se cuenta con la autorización para hacerlo y sin necesidad de ir al laboratorio de desarrollo ni reunirse con los implicados con este. También sería una gran ventaja para el área de calidad pues podrían descargar el producto en cualquier momento, confiando en la actualidad del mismo, y

realizar las pruebas pertinentes incluso antes de que el producto sea liberado por sus desarrolladores.

## CAPÍTULO 3. EVALUACIÓN DEL SISTEMA PROPUESTO.

### INTRODUCCIÓN

En el presente capítulo se realiza la valoración del sistema propuesto. Para ello se utiliza el método Delphi de evaluación por criterio de experto, que permiten dar un pronóstico sobre la efectividad de la solución.

Además se hace una validación de la hipótesis para dejar expuesto que el sistema de salvaguarda de la solución propuesta se hace de esta manera.

#### 3.1 Método seleccionado

Entre los métodos más comunes para evaluar la propuesta se encuentra el de "Test de Turing" [34], el de "Validación" [35] de comportamientos en casos extremos y el método "Delphi". Se seleccionó como método de evaluación para el presente trabajo el método Delphi, pues la aplicación y recolección de datos puede demorar mucho tiempo y en este caso, aun no se tienen suficientes datos sobre los resultados obtenidos al aplicar la propuesta. El método Delphi permite obtener las opiniones de un panel de especialistas interrogando de forma individual, iniciando con una encuesta y evitando el encuentro para garantizar el anonimato e impedir la influencia de unos en otros. Este proceso se repite las veces que sean necesarias para llegar a un consenso.

#### 3.2 Selección de expertos

Para la selección de los expertos se identificaron 11 especialistas (Anexo 7) donde a cada uno se le entrega un cuestionario de autovaloración en cuanto al dominio del tema y las fuentes de argumentación a partir de las cuales ha logrado ese conocimiento.

Se estableció el coeficiente de competencia para cada posible experto a partir de la fórmula:  $K = (kc + ka) / 2$

**K:** Coeficiente de competencia de cada experto (Tabla 10).

**kc:** Coeficiente de conocimiento o información que tiene el experto acerca del problema. Según lo establecido en esta metodología se le pide a cada experto que marquen en una escala creciente de 1 a 10, con una cruz, el valor que se corresponde con el grado de conocimiento o información que tienen sobre el tema de estudio. El coeficiente se calcula multiplicando la valoración del propio experto por 0,1.

**ka:** Este coeficiente se autoevalúa en alto (A), medio (M) o bajo (B) para un grupo de aspectos que influyen sobre el nivel de argumentación o fundamentación del tema a estudiar (Anexo 7).

*Tabla #10: Resultados arrojados del procesamiento para la determinación del coeficiente de competencia.*

<b>Experto</b>	<b>Kc</b>	<b>Ka</b>	<b>K</b>	<b>Valoración</b>
1	0.3	0.7	0.5	M
2	0.4	0.8	0.6	M
3	0.5	0.8	0.65	M
4	0.9	0.6	0.75	M
5	0.7	0.9	0.8	A
6	0.8	0.6	0.7	M
7	0.7	0.8	0.75	M
8	0.8	0.8	0.8	A
9	0.9	0.8	0.85	A
10	0.6	0.7	0.65	M
11	0.7	0.7	0.7	M

La selección final de los expertos tuvo lugar atendiendo a los siguientes criterios de interpretación del coeficiente de competencia (k):

- Si  $0.8 \leq k \leq 1.0$ , el coeficiente de competencia es alto.
- Si  $0.5 \leq k < 0.8$ , el coeficiente de competencia es medio.
- Si  $k < 0.5$ , el coeficiente de competencia es bajo.

### **3.3 Confección del instrumento**

Para determinar los aspectos a valorar se utilizó la experiencia y conocimientos de los expertos. La documentación de la guía estuvo a disposición de los especialistas con la posibilidad de aclarar dudas al respecto. En una primera iteración, se les solicitaron ideas para definir los posibles parámetros de evaluación, a partir de la propuesta que aparece en el Anexo 8 que es el nivel de experiencia de los expertos. Las respuestas debían estar enmarcadas en una escala numérica representando a las calificaciones: Muy bueno, Bueno, Regular, Pobre y Malo. Debido a las sugerencias de la inclusión de nuevos parámetros se elaboró una propuesta más completa (ver Anexo 9), que se envió nuevamente a los expertos para su valoración, los cual fue aceptada por ellos.

### 3.4 Resultados arrojados del método Delphi

Para verificar la consistencia en el trabajo de expertos, se utiliza el coeficiente de concordancia de Kendall (k) y el estadígrafo Chi cuadrado. Dados n el número total de criterios a evaluarse y m el número de expertos involucrados en la evaluación, se realiza el siguiente procedimiento para determinar la consistencia del trabajo de los expertos:

1. Calcular para cada criterio la sumatoria del peso dado por cada experto, mediante

la expresión:  $\sum_{j=1}^m C_j$

2. Determinar el valor de puntuación promedio de cada criterio:  $P_j$ .

3. Se calcula el peso medio dado por cada experto  $P_i = \frac{\sum_{i=1}^n C_i}{n_i}$  y luego  $\bar{C} = \frac{\sum_{i=1}^m P_i}{m}$

4. Determinar la desviación de la media y elevar el resultado al cuadrado para

obtener la dispersión, mediante la expresión:  $\sigma^2 = \sum_{j=1}^n (\sum_{i=1}^m C_i - \bar{C})^2$

5. Conociendo la dispersión se puede calcular el coeficiente de concordancia de

Kendall dado por la expresión:  $k = \frac{12 \sigma^2}{m^2(n^3-n)}$

6. Calcular el Chi cuadrado real a partir del valor del coeficiente de Kendall teniendo

en cuenta la siguiente expresión:  $x^2 = m(n-1)k$

Después de recibir los valores del peso relativo de cada criterio se construye la tabla de los pesos otorgados, para proceder con los cálculos (Anexo 10). De donde se obtuvo que  $X^2 = 0.66$ .

Para que exista concordancia en el trabajo de los expertos, según la Tabla de Distribución Chi Cuadrado debe cumplirse que:  $X_{real}^2 < X_{tabla}^2$ . Como  $\alpha=0.01$  con 4 grados de libertad, entonces:  $X_{tabla}^2 = 2.4701$ .

Al cumplirse que  $0.66 < 2.4701$ , se llega a la conclusión de que existe concordancia entre los expertos y no es necesario realizar otra iteración.

Al conocer el peso relativo de cada criterio y la calificación cuantitativa dada por los expertos a cada criterio en una escala de 1 a 5 (Malo, Pobre, Regular, Bueno y Muy bueno), se determina el índice de aceptación (IA) de la propuesta,

Donde:

**Pr:** peso relativo de cada criterio

**c:** calificación cuantitativa dada por los expertos a cada criterio en una escala de 1 a 5.

*Tabla #11: Tabla de Richman.*

	1	2	3	4	5	<i>Pr</i>	<i>Pr x c</i>
C1				x		0.08	0.32
C2					x	0.08	0.4
C3				x		0.08	0.32
C4					x	0.07	0.35
C5					x	0.06	0.3
C6					x	0.08	0.4
C7					x	0.08	0.4
C8					x	0.07	0.35
C9				x		0.08	0.32
C10					x	0.06	0.3
C11					x	0.08	0.4
<i>Pr * c</i>							3.86
<i>Pr * c/5</i>							0.772

El IA está dado por la expresión  $Pr * c/5$  el cual es de 0.772 por lo que se puede afirmar que existe una alta probabilidad de éxito.

Teniendo en cuenta que:

- $IA > 0.7$ : Alta probabilidad de éxito
- $0.7 > IA > 0.5$ : Probabilidad media de éxito
- $0.5 > IA > 0.3$ : Probabilidad de éxito baja
- $0.3 > IA$ : Fracaso seguro

### 3.5 Validación de la hipótesis

La eficiencia es la relación entre costo y beneficios, enfocada hacia la búsqueda de la mejor manera de hacer y ejecutar las tareas, con el fin de que los recursos se utilicen del modo más racional posible.

Hasta ahora los resultados afirman que la aplicación tiene un alto nivel de aceptación por los expertos y que la misma tiene una alta probabilidad de éxito, sin embargo es importante validar la eficiencia del software teniendo en cuenta el problema científico planteado.

En el Anexo 8 se vieron dos medidores económicos que son muy importantes para tener en cuenta a la hora de obtener resultados de acuerdo a las variables de tiempo y cantidad de hombres que minimiza el sistema. En el Anexo 11 se obtienen resultados mucho más detallados de una forma cuantitativa de estas variables y de la sencillez con que se puede configurar y administrar el Bacula desde el software propuesto.

Los resultados promedios arrojados de la encuesta, realizada a los expertos seleccionados, que se muestra en el Anexo 11, muestran el nivel de eficiencia que puede tener la solución propuesta.

Tabla #12: Resultados de la encuesta para validar la eficiencia de la solución.

Pregunta	Manual	Software
1	30 min	3 min
2	30 min	5 min
3	60 min	20 min
4	2880 min (48 h)	120 min (2 h)
5	15 hombres	1 hombre
6	100 hombres	7 hombres
7	12 procesos	4 procesos

Como se muestra en la tabla 12, los resultados arrojados son muy satisfactorios para el software planteado, puesto que los recursos humanos y de tiempo mejoran radicalmente cuando se utiliza la aplicación. Por otra parte la cantidad de procesos a

configurar o administrar de forma manual supera 3 veces a la cantidad de hacerlo con la solución propuesta. Es importante destacar que los expertos coinciden que la cantidad de procesos de forma manual y utilizando el software es la misma que se detalla en los acápite 1.4 y 2.4 respectivamente. Quedando claro que configurar o administrar estos procesos o pasos con la aplicación propuesta es mucho más sencillo que manualmente.

Vale destacar que la medición de estos indicadores de eficiencia de la solución propuesta, debe ser acompañada por una dinámica de análisis y reflexión periódica con todo el personal de trabajo, para que se generen valoraciones en torno a los problemas y las propuestas de solución, así como para transmitir conocimientos y tomar decisiones.

#### **3.6 Conclusiones parciales**

Al aplicar el método multicriterio para evaluar la factibilidad de aplicar la solución propuesta, se obtuvieron resultados que califican la solución con una alta probabilidad de éxito.

Por último se realizó una validación de la hipótesis de acuerdo a una encuesta realizada a varios especialistas expertos en el tema. La eficiencia fue probada de acuerdo a la medición de los indicadores de hombres y tiempo, los cuales son muy notables cuando se usa el software en cuanto a disminuir cada indicador de acuerdo a las preguntas planteadas en la encuesta. Y la sencillez en cuanto a la cantidad de procesos para configurar el sistema de respaldo.

## **CONCLUSIONES GENERALES**

Se obtuvo un sistema informático en el que se tienen en cuenta todos los ficheros de configuración del Bacula para ser manejado.

En el sistema informático se propone una nueva arquitectura de ficheros la cual ayuda a que la integración del software con Bacula se haga de manera óptima. Además se incorpora una solución para el versionado de Bacula la que no obliga a los usuarios utilizar una misma versión para la integración de sus clientes y servidores.

Por último se aplicó un método multicriterio el cual deja plasmado en el documento la efectividad del software propuesto. También se validó la hipótesis de la investigación de acuerdo a las variables eficiencia y sencillez, obteniendo resultados favorables a disminuir cuando se utiliza el software propuesto.

## **RECOMENDACIONES**

Al concluir esta investigación, se recomiendan los siguientes aspectos:

- Desarrollar, dentro de la misma arquitectura de la solución propuesta, otras funcionalidades como la instalación y configuración de las diferentes herramientas que se utilizan para desarrollo y despliegue de los distintos proyectos productivos de la UCI.
- Desplegar la solución utilizando los servicios de autenticación del LDAP, para tener garantizados las trazas y la sesión única de los usuarios autorizados.
- Desplegar la solución en una VLAN aislada para así garantizar la seguridad del proceso de salvas.
- La propuesta se aplicó con buenos resultados en el Centro de Soporte, Salón de Exposiciones y Centro de Datos de la UCI, pero también sería factible introducirla en todos los proyectos productivos que se desarrollan en la universidad y paulatinamente incrementar el número de estos a medida que se vaya perfeccionando el proceso.
- Socializar la solución propuesta a varias entidades de nuestro país para tener un aseguramiento de las salvas y poder tomar mejores decisiones.

## REFERENCIAS BIBLIOGRÁFICAS

1. Arteaga, Luis. *Nueva economía, internet y tecnología*. Gestio Polis [en línea].  
Disponible de World Wide Web:  
<http://www.gestiopolis.com/canales2/gerencia/1/ddsluisart.htm>
2. Castro Murillo, Iván Jesús. *Políticas de respaldo de información 2009*  
<http://es.scribd.com/doc/20191352/Políticas-de-Respaldo-de-Información>.
3. Llerena, Priscila. *Resguardos información y modos almacenamiento* [en línea].  
Disponible de World Wide Web:  
<http://brandymuse.espacioblog.com/post/2009/09/25/resguardos-informacion-y-modos-almacenamiento>
4. Alonzo Huamán, Max. *Fundamentos de sistemas de respaldo de información. Universidad Continental de Ciencias e Ingeniería. 01-2009*  
<http://www.slideshare.net/maxalonzohuaman/sistemas-de-respaldo-de-informacion-presentation>
5. MANUEL HUIDOBRO, José y Agustín López Raya. *Módulo de herramientas informáticas 2010*.
6. Disponible de World Wide Web: <http://www.definicion.org/copia-de-seguridad>
7. Swedorski, Scoot. *Handy backup the power of simplicity*. Disponible de World Wide Web: <http://www.handybackup.es/>
8. Search Sotorage. *PROVEEDORES DE TECNOLOGÍA DE ALMACENAMIENTO 2010*. [en línea] Disponible de World Wide Web:  
<http://www.searchstorage.es/acerca-de-nosotros/>
9. Osmosis Latina. *Restauración y copia de seguridad en línea* [en línea].  
Disponible de World Wide Web:  
<http://www.osmosislatina.com/soporte/respaldos.htm>.
10. Grzegorz Adam Hankiewicz. *DAR differential backup 2006-10-07*  
Disponible de World Wide Web:  
<http://dar.linux.free.fr/doc/mini-howto/dar-differential-backup-mini-howto.es.html>
11. Andres B. Rosas. *Resguardar la información y sus medios de almacenamiento*  
<http://andreasarely.blogspot.com/2009/10/resguardar-la-informacion-y-sus-medios.html>.
12. IBM. *Open source backup tools Amanda, BackupPC and Bacula compared* [en línea].  
Disponible de World Wide Web:

- <http://searchstorage.techtarget.com.au/news/2240020810/Open-source-backup-tools-Amanda-BackupPC-and-Bacula-compared>
13. W. Curtis Preston. *Backup y Recuperación de O'Reilly Media 2007*, ISBN 978-0596102463
  14. Shawn Powers. *Linux Journal: BackupPC Linux Journal, Marzo 2011*
  15. Falko Timme. *Copia de seguridad de los sistemas Linux y Windows con BackupPC*, Enero 2007. Consultado el 30 de julio 2010.
  16. W. Curtis Preston. *Backup & Recovery Inexpensive Backup Solutions for Open Systems 2007* [en línea]. <http://shop.oreilly.com/product/9780596102463.do>
  17. PRESSMAN, R. S. *Ingeniería del software. Un enfoque práctico, 5ta ed. McGraw-Hill, 2002*. Disponible en: <http://bibliodoc.uci.cu/pdf/reg02689.pdf>
  18. Frances, Juan Luis. *Open source monitoring and reporting tool for Bacula (3 de diciembre de 2006)* [en línea] Disponible de World Wide Web: <http://Bacula-web.dflc.ch/index.php/home.html/>
  19. Ovalle Gahona, Andres Esteban. *Respaldo en otro pc, 2007*. Disponible de World Wide Web: <http://www.mail-archive.com/linux@listas.inf.utfsm.cl/msg25229.html>
  20. PHP 5.4.0 RC6 released <http://www.php.net/>
  21. *librosweb.es, 2008*
  22. *Symfony en pocas palabras* [en línea] Disponible de World Wide Web: [http://www.librosweb.es/symfony/capitulo1/symfony\\_en\\_pocas\\_palabras.html](http://www.librosweb.es/symfony/capitulo1/symfony_en_pocas_palabras.html)
  23. François Zaninotto y Fabien Potencier. *Guía definitiva de Symfony* [http://asteinmetz.com.ar/weblog/wp-content/uploads/2008/11/symfony\\_guia\\_definitiva.pdf](http://asteinmetz.com.ar/weblog/wp-content/uploads/2008/11/symfony_guia_definitiva.pdf)
  24. AVILA MOLINA, Jesús. *Características de PostgreSQL* [en línea] Disponible de World Wide Web: <http://archives.postgresql.org/pgsql-es-ayuda/1998-07/msg00002.php>
  25. ARCO. *Instalar un sistema Bacula en GNU* [en línea] Disponible de World Wide Web: <http://crysol.org/node/400>
  26. *Servidor de almacenamiento - Servidor de ficheros. Openfiler - Appliance NAS/SAN.* [en línea] Disponible de World Wide Web: <http://www.bujarra.com/ProcedimientoOpenfiler.html>
  27. *Bacula manual.* [en línea] Disponible de World Wide Web: [http://www.Bacula.org/manuals/en/install/install/Storage\\_Daemon\\_Configuration](http://www.Bacula.org/manuals/en/install/install/Storage_Daemon_Configuration)
  28. RPM repository on fr2.rpmfind.net. [en línea] Disponible de World Wide Web: <http://rpmfind.net/linux/RPM/fedora/devel/rawhide/i386/b/Bacula-client-5.2.4-4.fc17.i686.html>

29. BACULA SOFTWARE DEAMON, Disponible de World Wide Web:  
<http://infosertec.loquefaltaba.com/tuxinfo20.pdf>
30. SAN DIEGO DIVERS [en línea] Disponible de World Wide Web:  
<http://rpm.pbone.net/index.php3/stat/45/idpl/17442937/numer/1/nazwa/Bacula-bwxconsole/BWX-CONSOLE>
31. Bacula, sistema de backup en red [en línea] Disponible de World Wide Web:  
<http://blog.lebrijo.com/?p=129&language=es>
32. Bacula y los 1000 y un Servidores [en línea] Disponible de World Wide Web:  
<http://www.vafe.com.ar/2008/09/18/Bacula-y-los-1000-y-un-servidores/>
33. STROZ FRIEDBERG [en línea] Disponible de World Wide Web:  
<http://www.velero25.net/pdfs/2005/turing.pdf>
34. HAKING [en línea] Disponible de World Wide Web:  
[www.echalemojo.com/uploadsarchivos/metodo\\_delphi.pdf](http://www.echalemojo.com/uploadsarchivos/metodo_delphi.pdf)
35. SIDNEY SIEGEL and N. JOHN CASTELLAN. *Estadística No Paramétrica: Aplicada a las ciencias de la conducta*. Trillas, 1994. 437 p. ISBN-10: 9682451019 | ISBN-13: 978-9682451010.

## BIBLIOGRAFÍA

1. Anónimo **2009**. *Writings - How To*. 2009. [en línea] Disponible de World Wide Web: <https://www.haught.org/article/Bacula-osx/>.
2. Anónimo **2008**. *WIN32 compile problem with current svn*. 2008. <http://copilotco.com/mail-archives/Bacula-devel.2008/msg00035.html>.
3. Anónimo **WordPress 2008 – 2012**. [Definición] [en línea] Disponible de World Wide Web: <http://definicion.de/eficiencia/>
4. Anónimo **2007**. *Bacula Cluster Howto*. 2007. [en línea] Disponible de World Wide Web: <http://linuxlab.wordpress.com/2007/02/22/Bacula-cluster-howto/>.
5. Anónimo *Bacula Install*. [en línea] Disponible de World Wide Web: [http://wiki.itlinux.cl/data/pdfex/PDF\\_Bacula\\_Bacula\\_install.pdf](http://wiki.itlinux.cl/data/pdfex/PDF_Bacula_Bacula_install.pdf). **2010**. USA : s.n., 2010.
6. Anónimo **Bacula.org**. **2012**. *Getting Started with Bacula*. 2012. [en línea] Disponible de World Wide Web: [http://www.Bacula.org/en/dev-anual/main/main/Getting\\_Started\\_with\\_Bacula.html](http://www.Bacula.org/en/dev-anual/main/main/Getting_Started_with_Bacula.html).
7. Anónimo **2009**. *Bacula: Respaldo Confiable y libre. Software libre vnz*. 2009. <http://softwarelibrevenezuela.blogspot.es/1233179340/>.
8. Anónimo **2010**. *DevToolShed.com*. 2010. [en línea] Disponible de World Wide Web: <http://www.devtoolshed.com/content/how-back-windows-Bacula>.
9. Anónimo **2012**. *El blog de 3-ik*. 2012. [en línea] Disponible de World Wide Web: <http://www.3-ik.com.ar/blog/tag/Bacula-opensource/>.
10. Anónimo *How to Bacula Cliente-Servidor*. **2008**. Mexico : s.n., 2008.
11. Anónimo **2012**. *Howto.gr*. 2012. [en línea] Disponible de World Wide Web: <http://www.howto.gr/wp/Bacula-network-backup/>.
12. Anónimo **2010**. <http://www.cybercanibal.com>. [en línea] Disponible de World Wide Web: <http://www.cybercanibal.com>. 2010.
13. Anónimo **2012**. *online-tutorials.net*. 2012. [en línea] Disponible de World Wide Web: <http://www.online-tutorials.net/backup/Bacula/tutorials-t-125-302.html>.
14. Anónimo **2012**. *Open-Source PHP Web Marco de trabajo*. 2012. [en línea] Disponible de World Wide Web: <http://www.symfony-project.org/>.
15. Anónimo **2001**. *Pixelado.org*. 4 de 2001. [en línea] Disponible de World Wide Web: <http://www.pixelado.org/2011/08/Bacula/>.

16. Anónimo **2011**. *Renich's Blog*. 17 de 1 de 2011. [en línea] Disponible de World Wide Web: <http://blog.woralelandia.com/2011/01/17/howto-como-resetear-Bacula/>.
17. Anónimo **2006 - 2009**. *the Fresh Open Source Software archive*. Member "Bacula-5.2.6/po/es.po". 2006 - 2009. [en línea] Disponible de World Wide Web: <http://fossies.org/unix/misc/Bacula-5.2.6.tar.gz:a/Bacula-5.2.6/po/es.po>.
18. Anónimo **2010**. *The Open Source Network Backup Solution*. The Open Source Network Backup Solution. 2010.
19. Anónimo **2012**. *LinuxLinks.com*. 2012. [en línea] Disponible de World Wide Web: <http://www.linuxlinks.com/article/20071011183702621/Bacula.html>.
20. **Banchoff, Matias**. **2009**. *Guía Práctica de Bacula*. Argentina : s.n., 2009.
21. **Belkin, Sergio**. [CentOS-es] *Guia instalacion Bacula*. **2009**. España : s.n., 2009.
22. **Bleda, Valero**. **2012**. *Sistema de Backup*. España : s.n., 2012.
23. **CrySol**. **2006**. *CrySol*. CrySol. 10 de 09 de 2006. [en línea] Disponible de World Wide Web: <http://crysol.org/node/400>.
24. **ERACC**. **2012**. *The ERACC Web Log*. <http://blog.eracc.com/2011/03/21/linux-Bacula-is-for-everyone-backup-software/>. 2012.
25. **HERNÁNDEZ LEÓN, ROLANDO ALFREDO and SAYDA COELLO GONZÁLEZ**. *El proceso de investigación científica*. Editorial Universitaria del Ministerio de Educación Superior, 2011. 110 p. 978-959-16-1307-3
26. **IBM**. **2003**. *IBM Rational Unified Process (RUP)*. 2003. [en línea] Disponible de World Wide Web: <http://www-01.ibm.com/software/awdtools/rup/>.
27. **Kemp, Juliet**. **2007**. *Backups with Bacula: A cross-platform system to archive your data*. 2007. [en línea] Disponible de World Wide Web: <http://www.linux-mag.com/id/4429/>.
28. **Kern Sibbald**. **2009**. *Bacula Installation and Configuration Guide . Bacula Installation and Configuration Guide .* 2009. [en línea] Disponible de World Wide Web: <http://www.Bacula.org/manuals/en/install/install/index.html>.
29. **Langille, Dan**. **2004**. *SYS ADMIN*. 2004. [en línea] Disponible de World Wide Web: <http://onlamp.com/pub/a/onlamp/2004/01/09/Bacula.html>.
30. **León Rodríguez, Angel Gabriel** *sysadmin.org.mx*. [en línea] Disponible de World Wide Web: <http://sysadmin.org.mx/contenidos/manejo-de-respaldos-con-Bacula.html>.
31. **Ltd, DVL Software**. **2004**. *FreeBSD*. 2004. [en línea] Disponible de World Wide Web: <http://www.freebsdidiary.org/Bacula.php>.
32. **MANCHA, CASTILLA - LA**. **2010**. *CENTRO DE EXCELECNCIA DE SOFTWARE LIBRE*. 2010. [en línea] Disponible de World Wide Web: <http://www.ceslcam.com/>.

33. **Matias D. Banchoff T. 2011.** *Backups con Bacula.* 2011. [en línea] Disponible de World Wide Web: <http://es.scribd.com/doc/65119893/Introduccion-a-Bacula>.
34. **Perez, Julio Carrau - Gustavo. 2009.** *Gestión de Registros y Respaldos.* 2009.
35. **TECHGURULIVE.** 2008. [en línea] Disponible de World Wide Web: <http://techgurulive.com/2009/04/13/how-to-buildg-Bacula-from-source-an-open-source-network-backup-and-restore-solution/>.
36. **Urbina, Cristián Aguilera. 2011** *Gestión de Backups Mediante Solución Open Source..* Chile : s.n., 2011.
37. **Valentine, Nathan. 2005.** *The Bacula Philosophy.* 2005.
38. **Weimer, Hendrik. 2006.** *OS REVIEWS.* 2006. [en línea] Disponible de World Wide Web: <http://www.osreviews.net/reviews/admin/Bacula>.

## **GLOSARIO**

**Proyecto productivo:** Un proyecto productivo es la búsqueda de una solución para un problema o una necesidad humana, con un esfuerzo temporal para la creación de un producto o servicio.

**Software:** Se conoce como software al equipamiento lógico o soporte lógico de un sistema informático; comprende el conjunto de los componentes lógicos necesarios que hacen posible la realización de tareas específicas, en contraposición a los componentes físicos, que son llamados hardware.

**Sistema informático:** Un sistema informático como todo sistema, es el conjunto de partes interrelacionadas, hardware, software y de recurso humano que permite almacenar y procesar información. El hardware incluye computadoras o cualquier tipo de dispositivo electrónico inteligente, que consisten en procesadores, memoria, sistemas de almacenamiento externo, etc.

**Hardware:** corresponde a todas las partes tangibles de un sistema informático sus componentes son: eléctricos, electrónicos, electromecánicos y mecánicos. Sus cables, gabinetes o cajas, periféricos de todo tipo y cualquier otro elemento físico involucrado; contrariamente, el soporte lógico es intangible y es llamado software.

**Sistema informático:** En informática, una aplicación es un tipo de programa informático diseñado como herramienta para permitir a un usuario realizar uno o diversos tipos de trabajo. Esto lo diferencia principalmente de otros tipos de programas como los sistemas operativos (que hacen funcionar al ordenador), las utilidades (que realizan tareas de mantenimiento o de uso general), y los lenguajes de programación (con el cual se crean los programas informáticos).

**Sistema gestor de base de datos:** son un tipo de software muy específico, dedicado a servir de interfaz entre la base de datos, el usuario y las aplicaciones que la utilizan.

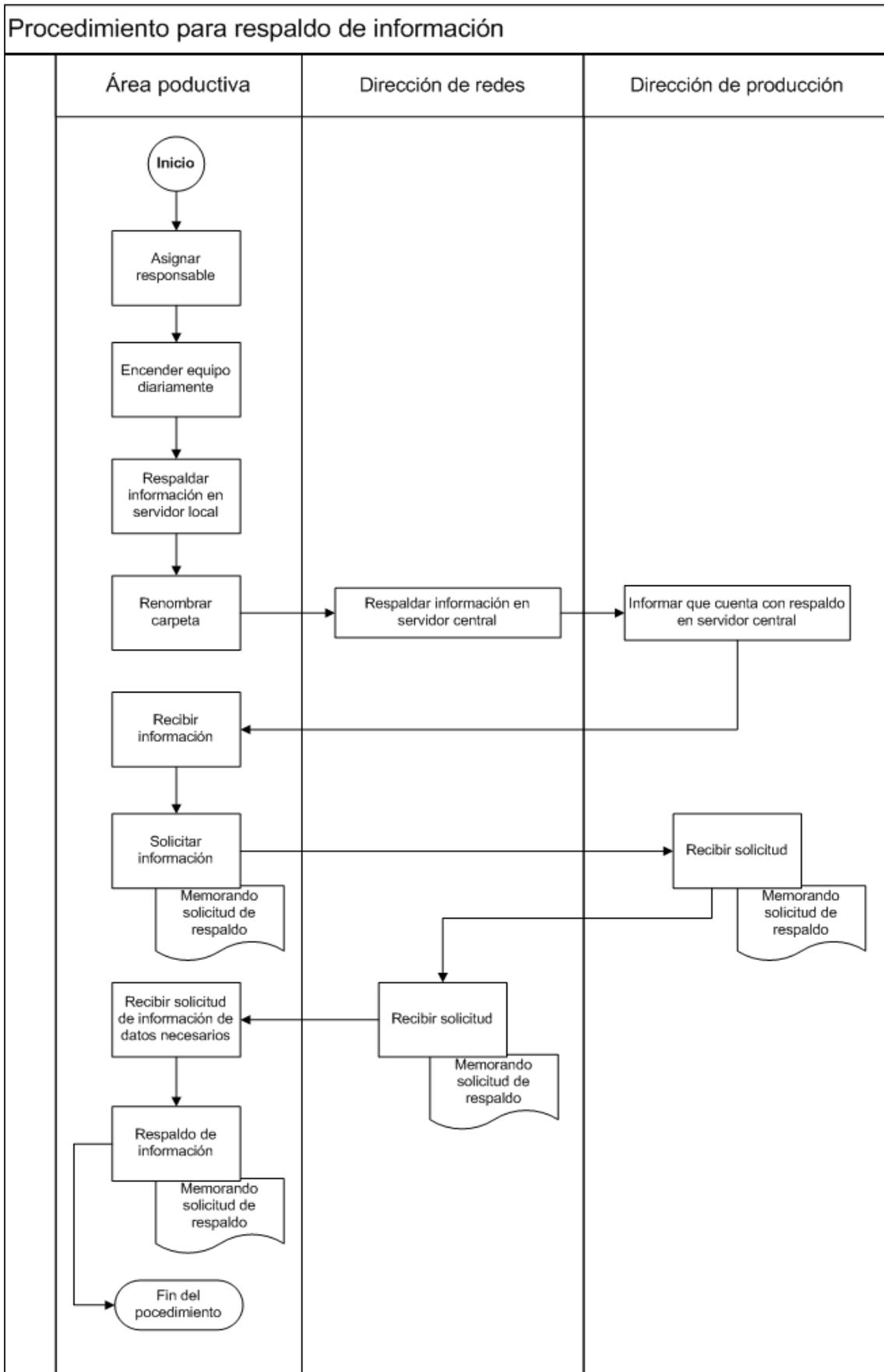
**Respaldo:** El respaldo de información es un proceso muy importante que debe de tener cada usuario de computadora, sea un equipo portátil o un equipo de escritorio. El contar con respaldos permite al usuario en algún momento dado recuperar información que haya sido dañada por virus, fallas en el equipo o por accidentes.

**Ataque:** Una acción que se sirve de una o varias vulnerabilidades para materializar una amenaza.

Disponibilidad: Los recursos deben poder ser accedidos cuando son necesarios y funcionar a un nivel aceptable.

Vulnerabilidad: Una vulnerabilidad es un punto débil que de explotarse con éxito puede llegar a originar la consecución de una amenaza.

**ANEXO 1. Procedimiento para el respaldo de información de una empresa**



**ANEXO 2. Configuración del fichero Clients**

```
# <NOMBRE_DE_LA_PC>
JobDefs {
Name = "DefaultJob_<NOMBRE_DE_LA_PC>"
Type = Backup
Level = Incremental
Client = <NOMBRE_DE_LA_PC>-fd
FileSet = "Full Set <NOMBRE_DE_LA_PC>"
Schedule = "<TIEMPO_DE_SALVA>_<NOMBRE_DE_LA_PC>"
Storage = File_<NOMBRE_DE_LA_PC>
Messages = Standard
Pool = <DIRECCIÓN_IP_DE_LA_PC>
Priority = 10
}
```

```
Job {
Name = <NOMBRE_DE_LA_PC>
JobDefs = "DefaultJob_<NOMBRE_DE_LA_PC>"
Write Bootstrap = "/var/lib/Bacula/<CLIENTE>.bsr"
}
```

```
Job {
Name = "BackupCatalog_<NOMBRE_DE_LA_PC>"
JobDefs = "DefaultJob_<NOMBRE_DE_LA_PC>"
Level = Full
FileSet="Catalog_<NOMBRE_DE_LA_PC>"
}
```

```

Schedule = "<TIEMPO_DESPUÉS_DE_LA_SALVA>_<NOMBRE_DE_LA_PC>"

RunBeforeJob          =          "/usr/bin/awk          -f
/etc/Bacula/scripts/make_catalog_backup_awk -v cat1=<MyCatalog>
/etc/Bacula/Bacula-dir.conf"

RunAfterJob   = "/etc/Bacula/scripts/delete_catalog_backup"

Write Bootstrap = "/var/lib/Bacula/BackupCatalog.bsr"

Priority = 11          # run after main backup

}

```

```

Job {

Name = "RestoreFiles_<NOMBRE_DE_LA_PC>"

Type = Restore

Client=<NOMBRE_DE_LA_PC>-fd

FileSet="Full Set <NOMBRE_DE_LA_PC>"

Storage = File_<NOMBRE_DE_LA_PC>

Pool = IP

Messages = Standard

Where = <RUTA>

}

```

```

FileSet {

Name = "Full Set <NOMBRE_DE_LA_PC>"

Include {

Options {

signature = MD5

Compression = GZIP

}

}

```

```
    }  
    File = <ARCHIVO_1_PARA_INCLUIR>  
    File = <ARCHIVO_N_PARA_INCLUIR>  
}  
  
Exclude {  
    File = <ARCHIVO_1_PARA_EXCLUIR>  
    File = <ARCHIVO_N_PARA_EXCLUIR>  
}  
}
```

```
Schedule {  
    Name = "<TIEMPO_DE_SALVA>_<NOMBRE_DE_LA_PC>"  
    Run = Full daily at <HORA_DE_RESPALDO>  
}
```

```
Schedule {  
    Name = "<TIEMPO_DESPUÉS_DE_LA_SALVA>_<NOMBRE_DE_LA_PC>"  
    Run = Full sun-sat at <HORA_DE_LA_SALVA>  
}
```

```
FileSet {  
    Name = "Catalog_<NOMBRE_DE_LA_PC>"  
    Include {
```

```
Options {  
signature = MD5  
}  
File = /var/lib/Bacula/.sql  
}  
}
```

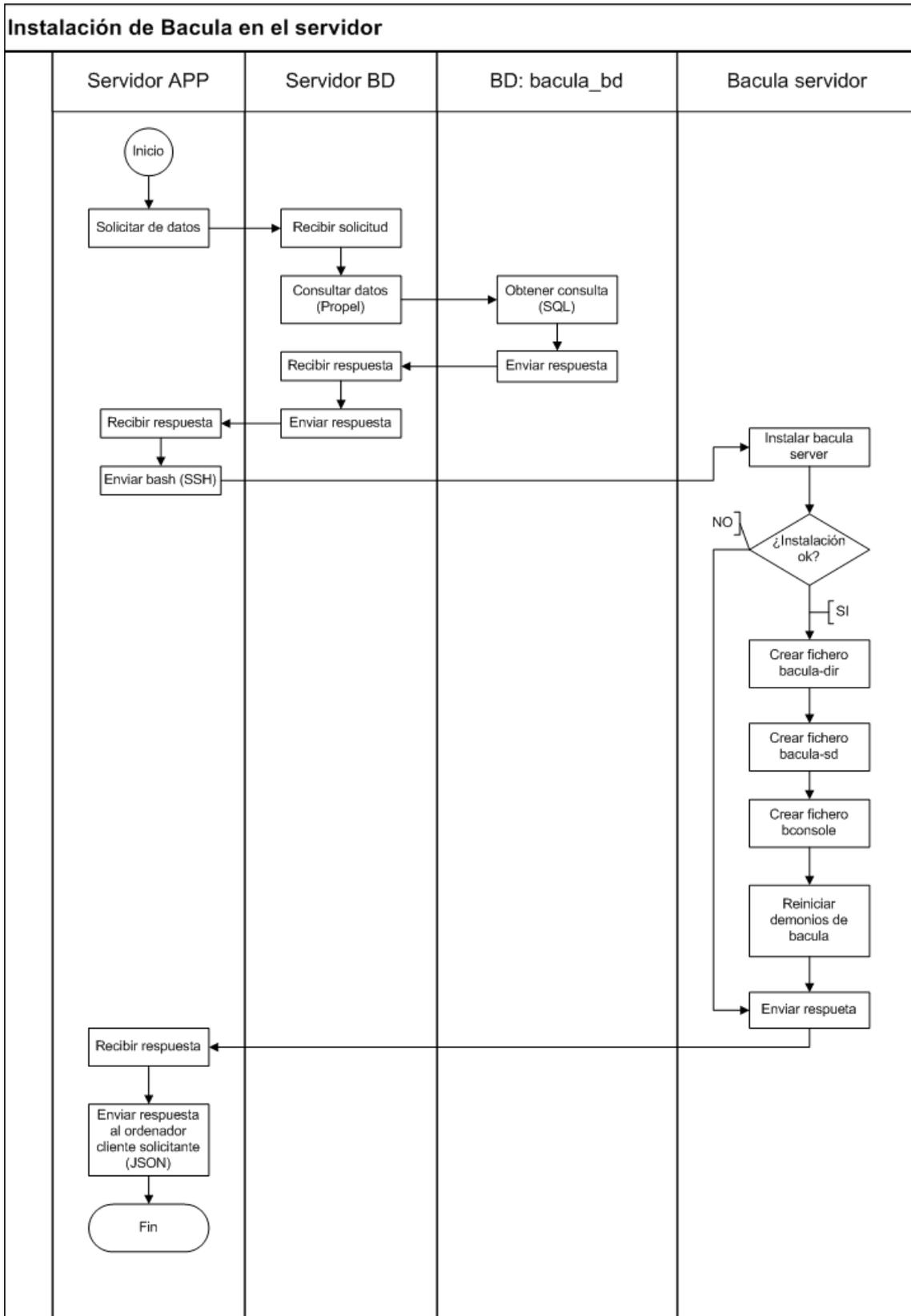
```
Client {  
Name = <NOMBRE_DE_LA_PC>-fd  
Address = <DIRECCIÓN_IP>  
FDPort = 9102  
Catalog = MyCatalog  
Password = <PASSWORD> # password for FileDaemon  
File Retention = 30 days # 30 days  
Job Retention = 6 months # six months  
AutoPrune = yes # Prune expired  
Jobs/Files  
}
```

```
Storage {  
Name = File_<NOMBRE_DE_LA_PC>  
Address = <DIRECCIÓN_IP> # N.B. Use a fully  
qualified name here  
SDPort = 9103  
Password = <PASSWORD>  
Device = FileStorage_<NOMBRE_DE_LA_PC>  
Media Type = File  
}
```

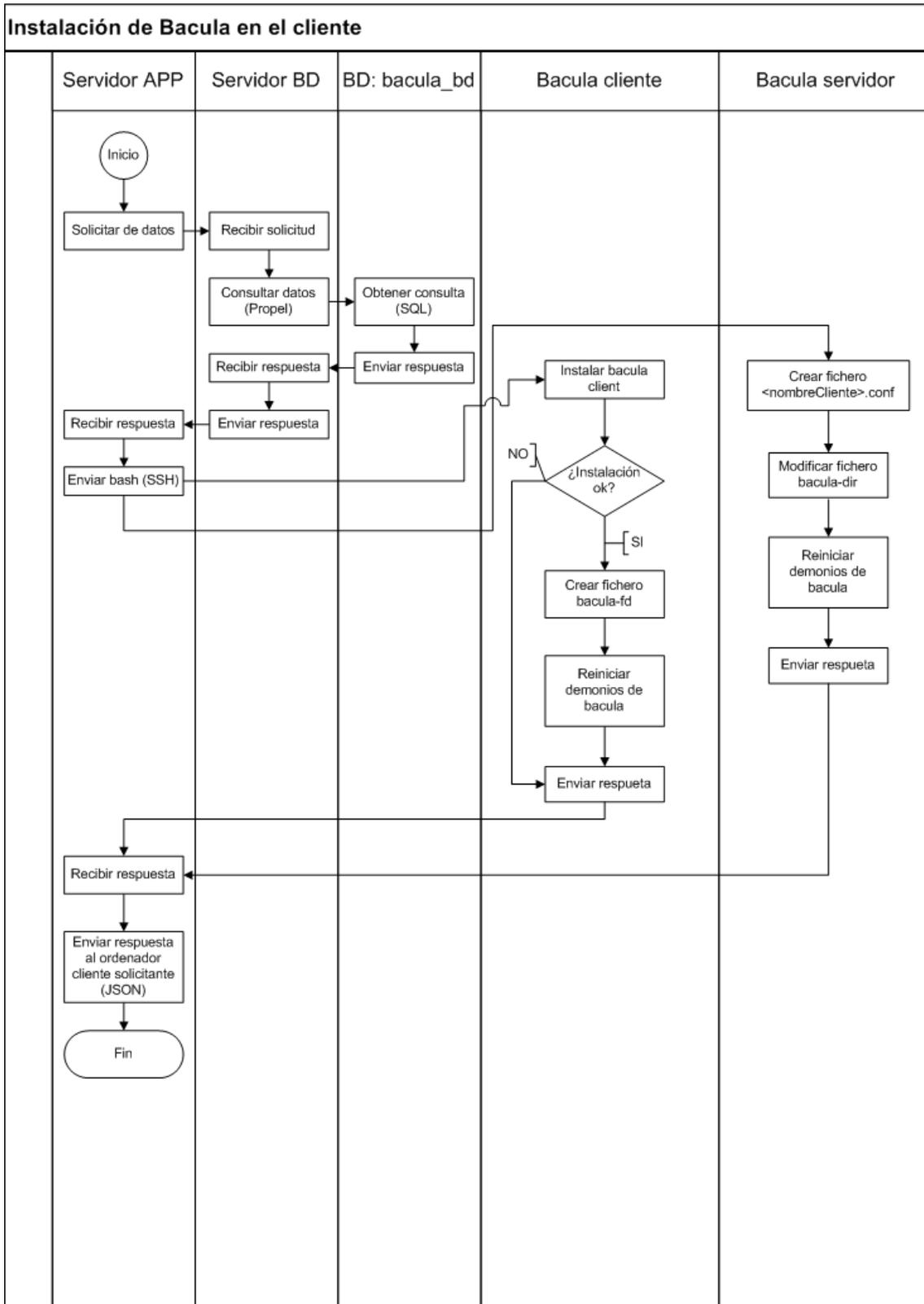
**ANEXO 3. Configuración del fichero Devices**

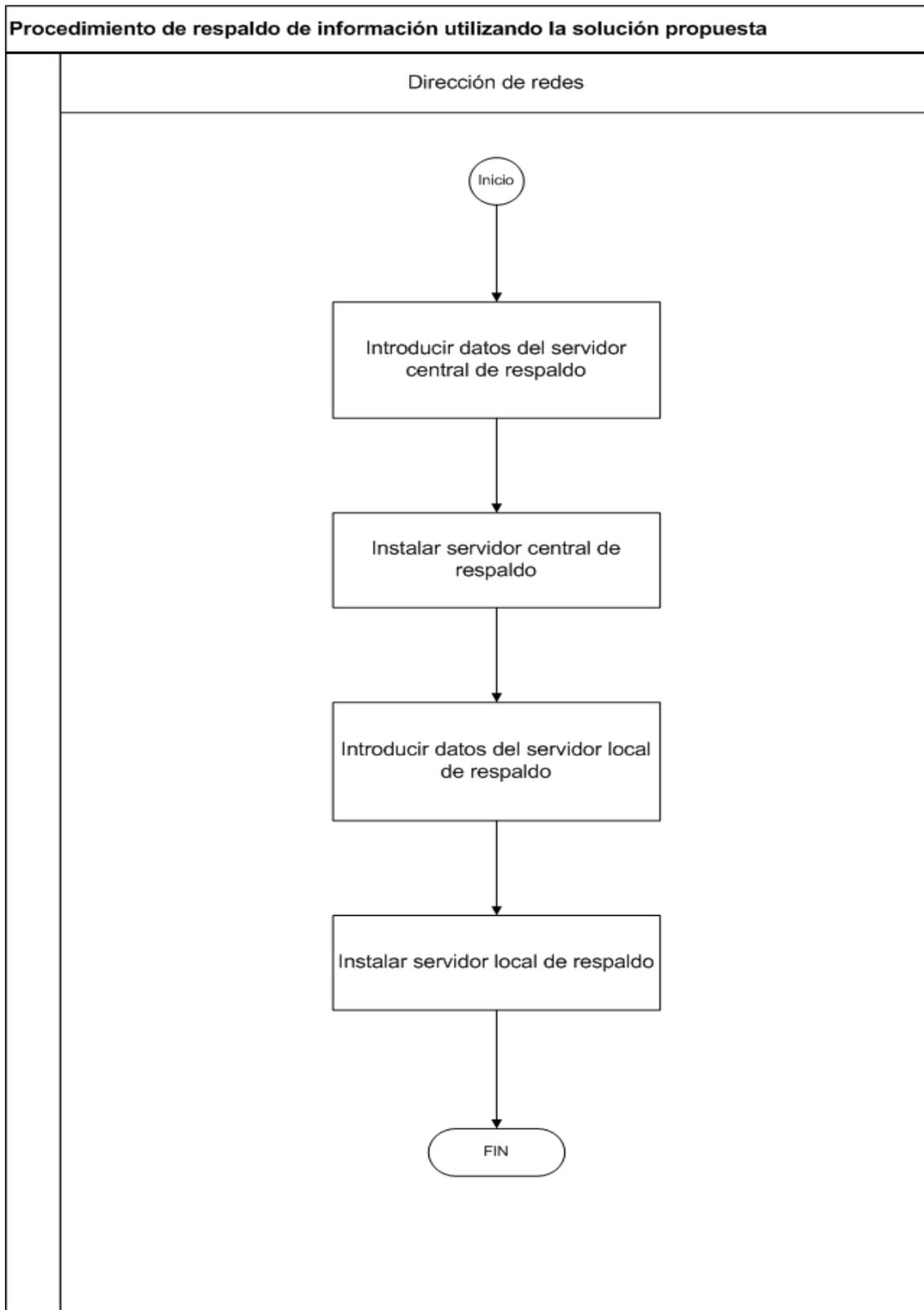
```
Device {
Name = FileStorage_<NOMBRE_DE_LA_PC>
Media Type = File
Archive Device = /etc/Bacula/Bacula_Ficheros/Labels/<NOMBRE>
LabelMedia = yes;                # lets Bacula label
unlabeled media
Random Access = Yes;
AutomaticMount = yes;            # when device opened, read
it
RemovableMedia = no;
AlwaysOpen = yes;
}
```

ANEXO 4: Instalación de Bacula en el servidor



ANEXO 5: Instalación de Bacula en el cliente



**ANEXO 6: Procedimiento de respaldo de información utilizando la solución propuesta**

---

**ANEXO 7. Encuesta: Diagnóstico sobre las salvas automáticas**

**1-** ¿En la configuración de Bacula cuantos archivos de configuración se deben tener en cuenta?

1 \_\_\_ 2 \_\_\_ 3 \_\_\_ +3 \_\_\_

**2-** ¿Es necesario tener un cliente en cada ordenador que se desee la salva automática?

Si \_\_\_ No \_\_\_ No se \_\_\_

**3-** ¿Usted cuando va a configurar el Bacula en un servidor para realizar operaciones de respaldo en qué tiempo logra la configuración correcta?

5min \_\_\_ 10min \_\_\_ 15min \_\_\_ +15min \_\_\_

**4-** ¿Desde su punto de vista el sistema Bacula es un software confiable para salvas automáticas?

Si \_\_\_ No \_\_\_ No se \_\_\_

**5-** ¿Desde su punto de vista si fuera a implantar un sistema de respaldo en una empresa cuál escogería?

Bacula \_\_\_ AMANDA \_\_\_ BackupPc \_\_\_ Otro \_\_\_

**ANEXO 8. Nivel de experiencia de los expertos**

Teniendo en cuenta su experiencia profesional usted ha sido seleccionado para colaborar con una investigación encaminada a proponer un sistema de software que agilice a la configuración de salvas automáticas en la Universidad de las Ciencias Informáticas. Para valorar sobre el nivel de experiencia que Ud. Posee, le pedimos realice la siguiente autoevaluación.

**Nombre y Apellidos:**

**Cargo que ocupa:**

**Grado científico:**

**Número de veces que ha sido seleccionado como experto:**

**Años de experiencia en el desarrollo de software:**

**Años de experiencia en temas de seguridad de aplicaciones:**

Marque con una cruz (x), en la casilla que le corresponda al grado de conocimientos que usted posee acerca del tema de investigación, valorándolo en una escala de 0 a 10 (considerando 0 como no tener absolutamente ningún conocimiento y 10 el de pleno conocimiento de la problemática tratada).

0      1      2      3      4      5      6      7      8      9      10

--	--	--	--	--	--	--	--	--	--	--

Las fuentes presentadas son

Fuente de argumentación	Alto	Medio	Bajo
Experiencia adquirida			
Conocimiento de trabajos nacionales.			
Conocimiento de trabajos exteriores.			
Conocimiento del estado del problema en el mundo.			

**ANEXO 9. Medidores de los resultados de la investigación.**

Usted ha sido seleccionado por la calificación científica-técnica, sus años de experiencia y los resultados alcanzados en su labor profesional, como experto para seleccionar de acuerdo a su opinión, los posibles criterios de evaluación para esta investigación, a partir de la siguiente propuesta. La evaluación del criterio, significa su valoración respecto a ese parámetro como medidor de la investigación que puede ser Muy bueno, Bueno, Regular, Pobre y Malo. Expresar su opinión en un valor numérico.

Categorías	Criterio	Evaluación(1-5)
<b>Científica</b>	Calidad de la investigación	
	Novedad científica	
	Aporte de la investigación	
<b>Necesidad</b>	Necesidad de la aplicación	
<b>Económico</b>	Contribución a disminuir costos de recursos humanos	
	Contribución a disminuir tiempos de configuración.	
	Contribución a disminuir tiempos de control y monitoreo.	
	Contribución a disminuir tiempos de despliegue.	
<b>Aplicabilidad</b>	Facilidad de manejo con la aplicación	
	Documentación y ayuda	
	Seguridad	

## ANEXO 10. Cálculo del coeficiente de Kendall y Chi cuadrado.

	P1	P2	P3	P4	$P_i$	$C$
E1	4	4	5	3	4	45.25
E2	5	3	4	4	4	
E3	4	4	4	4	4	
E4	4	5	4	3	4	
E5	5	3	3	5	4	
E6	5	5	4	4	4.5	
E7	5	4	4	4	4.25	
E8	5	4	3	4	4	
E9	4	4	3	5	4	
E10	5	4	4	5	4.25	
E11	4	5	5	3	4.25	
$\sum_{i=1}^m c_i$	47	45	43	44		
$P_j$						
$\sum_{i=1}^m c_i - \bar{c}$	1.75	0.25	2.25	1.25		
$(\sum_{i=1}^m c_i - \bar{c})^2$	3.1	0.06	5.1	1.6		
$\sum_{j=1}^n (\sum_{i=1}^m c_i - \bar{c})^2$	9.86					
KENDALL	0.02					
Chi cuadrado	0.66					

**ANEXO 11. Encuesta: Diagnóstico sobre la solución propuesta.**

Teniendo en cuenta su experiencia profesional usted ha sido seleccionado para colaborar con una investigación encaminada a proponer un sistema de software que agilice a la configuración de salvas automáticas en la Universidad de las Ciencias Informáticas. De acuerdo a las siguientes preguntas, llene la tabla que aparece después de las mismas.

- 1- ¿Qué tiempo le tomaría configurar un servidor para salvas automáticas?
- 2- ¿Qué tiempo le tomaría configurar un cliente para salvas automáticas?
- 3- ¿Qué tiempo le tomaría configurar un proyecto de desarrollo de la UCI, teniendo en cuenta que cada proyecto tiene 2 clientes y 1 servidor promedio?
- 4- ¿Qué tiempo le tomaría configurar un centro de desarrollo de la UCI, teniendo en cuenta que cada centro como promedio cuenta con 15 proyectos en desarrollo?
- 5- ¿Qué cantidad de personas, a su juicio, le tomaría configurar un centro de desarrollo de la UCI en 1 día?
- 6- ¿Qué cantidad de personas, a su juicio, le tomaría configurar todos los proyectos de desarrollo de la UCI en 1 día?
- 7- ¿Qué cantidad de procesos o pasos, a su juicio, debe configurar para obtener un buen resultado en la salva de 1 proyecto productivo?

Pregunta	Manual	Software
1		
2		
3		
4		
5		
6		
7		