N-Version programming method of Software Fault Tolerance: A Critical Review

Bharathi V

Abstract- **As many critical applications in modern society are dependent on computers, Fault Tolerance is evolved as a technique to increase the dependability of computing systems. Because of limitations with producing of error free software, Software Fault Tolerance (SFT) has become an important consideration. Majority of software errors are design faults. The root cause for software design errors is the complexity of the problem domain. A number of Fault Tolerance techniques aimed at minimizing the effect of software faults are being investigated. One such method based on d*esign diversity* technique is N-Version programming (NVP). In this paper, a critical review of NVP is presented. The advantages, current challenges, and further research areas of NVP are discussed.**

Keywords- **Design Diversity, N-Version Programming, Software Complexity, Software Fault Tolerance.**

I. INTRODUCTION

SERVICES in today's computation based society must be highly dependable. With the nearest competitor just a mouse click away, unplanned service downtime causes revenue loss and, in some cases, contractual penalties. Hence design of fault tolerant systems has gained significant attention. Ensuring shared resources are available despite the failure of certain hardware or a software component is a tremendous challenge for IT specialists. The concept of Fault Tolerance techniques through redundant hardware components was conceived in the early 1950s[5] [6]]. The 1950-1970 period was a time of evolution in both theoretical and practical aspects of hardware Fault Tolerance. The 1980s saw the successful implementation of several system designs with the emergence of product line of Tandem computer [7]. The computer industry has historically been concerned about hardware failures. Today, however, with amazing high hardware availability, the emphasis is shifting to software because software faults are the root cause in a high percentage of operational system failures [8].

Software Fault Tolerance is the ability of software to detect and recover from a fault that is happening or has already happened in either the software or hardware in the system in

which the software is running in order to provide service in accordance with the specification.

Bharathi V is with Wipro Technologies as a Senior Software Engineer, Bangalore. Karnataka. E-mail: bharathi_vkumar@rediffmail.com

Extending hardware reliability theory to software has a number of limitations as software does not exist in a physical state and it does not degrade in the same manner as physical hardware components do [1]. The consequence of software failures depends on the application and the particular characteristic of the fault. The immediate effect can range from minor inconvenience to catastrophic events. The main cause for software faults is design faults. This class of errors originates due to mistakes and oversights of humans that occur while they specify, design, operate, update and maintain the hardware and software of computing systems. The probability of design errors increases as the complexity of the problem rises. According to Woodfield [9], 25 percent rise in the problem complexity can lead to a 100 percent rise in program complexity. The increasing complexity of software is causing growing concern about the robustness of system software components. Even when exceptional caution is taken, simple logic errors can result in catastrophic failures. Examples are, midair explosion of Arian-5 Rocket, massive failure of North American signaling systems (SS7).

As a result of research efforts to apply Fault Tolerance to software design faults, a number of techniques have evolved. The following sections give a brief introduction to various techniques and a critical review of N-Version Programming approach.

II. SOFTWARE FAULT TOLERANCE TECHNIQUES

Software Fault Tolerance can be broadly classified into two groups. Single version software and Multiversion software techniques. Single version techniques focus on improving the Fault Tolerance of a single piece of software by adding mechanism into the design, targeting the detection, containment, and handling of errors caused by the design faults. Some of the key attributes of single version techniques are modularity, system closure, atomicity of actions and exception handling. Multiversion Fault Tolerance is based on the use of two or more versions of a piece of software executed either in sequence or in parallel. The modularity, system closure, atomicity of actions and exception handling attributes are desirable and advantageous in each version of the multiversion techniques too. Some of the classical techniques of multiversion Software Fault Tolerance are Recovery blocks (RB) and N-Version programming. Both of these techniques are based on *design diversity*.

Design diversity is the approach in which components of a system are built through independent designs but deliver the same service. The fundamental conjecture of design diversity is that components build differently will fail differently. Thus, if anyone of the redundant version fails, at least one of the others will provide an acceptable output.

A. Recovery Block technique

This technique was evolved as a result of first long term systematic investigation of multiversion technique initiated by Brian Randell in early 1970s [4]. In this technique, alternate software versions are organized in a manner similar to the dynamic redundancy (standby) technique in hardware. It's objective is to perform runtime Software Fault Tolerance detection by an acceptance test performed on the results delivered by the first version. If the acceptance test is not passed, state is restored to what existed prior to the execution of that algorithm and execution of an alternate version on the same hardware is followed. Recovery is considered complete when acceptance test is passed. Checkpoint memory is needed to recover the state after a version fails, to provide a valid starting operational point for the next version (Fig 1).

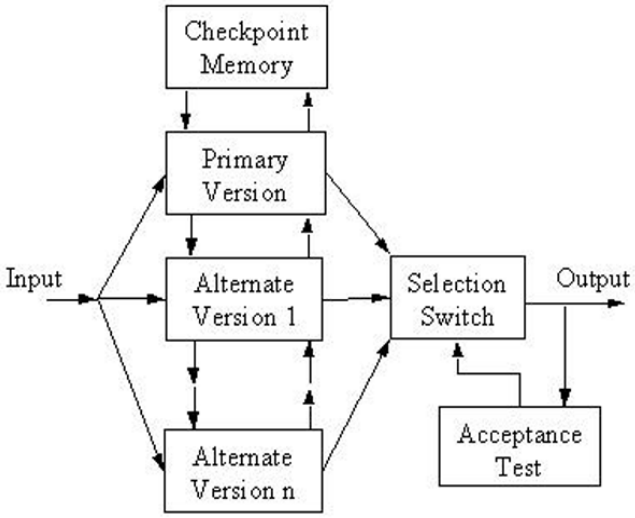

Figure 1. Recovery Block model

B. N-Version Programming

The NVP investigation project was started by A. Avizienis in 1975[2]. In this method, N-fold computation is carried out by using N independently designed software modules or "versions" and their results are sent to a decision algorithm that determines a single decision result [10].

The fundamental difference between the RB and NVP approaches is the decision algorithm. In the RB approach, an acceptance test that is specific to the application program being implemented must be provided. In the NVP approach, a decision algorithm that delivers an agreement/disagreement decision is implemented. The N-Version programming is defined as the independent generation of N>=2 software modules, called "versions", from the same initial requirements [10]. "Independent generation" refer to the programming effort by individual or groups that do not interact with each other with respect to programming process. As the goal of NVP is to minimize the probability of similar errors at decision points, different algorithms, programming languages, environments and tools are used wherever possible.

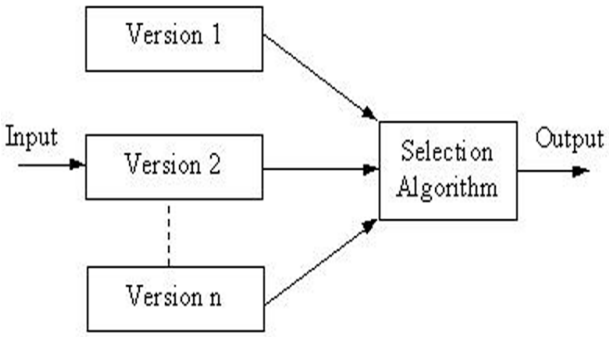

Figure 2. N-Version programming model

In NVP, since all the versions are built to satisfy the same requirements, it requires considerable development effort. But the complexity is not greater than inherent complexity of building a single version.

Comparison of outputs and declaration of single result is carried out by output selection algorithm or voting algorithm (Figure 2). The output selection algorithms should be capable of detecting erroneous version outputs and prevent the propagation of bad values to main output. The output selection algorithm should be developed considering the application attributes like safety and reliability.

For applications where safety is a main concern, algorithm should be capable of detecting erroneous outputs and prevent the propagation of bad values to the main output. Also, the algorithm should be capable of declaring an error condition or initiate an acceptable safe output sequence, when it can not achieve a high confidence of selecting a correct output. For increased reliability, algorithm should be developed such that output is correct with a very high probability.

Some of the generalized selection algorithms are Formalized majority voter, Generalized median voter, Formalized plurality voter and Weighted averaging techniques [11]. Other voting techniques that are being investigated are based on Neural network and Genetic algorithm techniques [12]. They are implemented such that their performance is related to the application and the particular characteristic of the software versions.

As an example to demonstrate the NVP, consider a simple program that counts the number of digits in an input text. The program reads strings, calls the procedure count_digit for each input string, adds up all the counts and prints the result. The module specifications [15] are as shown.

```
module main{
uses:count_digit(string) returning integer
Implementation: main.o
}

module string_function_package{
NVP module
Interface:count_digit(string) returning integer
Implementation: "C_string_function.o@system1"
Implementation: "P_string_function.o@system2"
Implementation: "F_string_function.o@system3"
Voter: "vote.o"
Error_handler: "handler.o"
}

application example{
import main
import string_function_package
bind main.count_digit string_function_package.count_digit
}
```

The specification defines a main module which calls the procedure count_digit. The main module does not have interface definition as this module is not used by any other module. The module string_function_package has an interface definition of count_digit used by module main. Specification for the module string_function_package also defines a voter and different versions of count_digit written in C, Pascal and Fortran. Object code for different versions of count_digit are C_string_function.o, P_string_function.o and F_string_function.o for C, Pascal and Fortran respectively. The module also specify the target machine on which the version has to execute.

Different language versions of function count_digit are as shown.

```
int digit(s)
{
--C function--
}

function digit(s:str):integer;
begin
--Pascal function--
End


INTEGER digit(string)
--Fortran function--
END FUNCTION digit
```

Voting procedure has three parameters, which contain number of digits as returned by three versions.
Pseudo code for voting procedure is as shown.

```
int* vote(output1, output2,output3)
int *output1, *output2, *output3
{
int *return_value
   if ( output1 ==output2) OR ( output1==output3)
   return_value=output1
   elseif (*output2==*output3)
   return_value = output2
   else
   call error_handler
 return return_value
}
```

The object file of the voter procedure as given in Voter clause is executed after all versions have terminated or aborted. Parameters are automatically passed to procedure vote.o at execution. The error_handler clause is for the case when no consensus can be reached by the versions though such a case is rare with versions with different languages. Designer can decide the approach through error_handler clause under no-consensus scenario.

III. ADVANTAGES OF NVP

- As NVP is based on design diversity technique, the built program will fail independently and with low probability of coincidental failures. This ensures that one of the other versions will continue to provide the required functionality
- Especially in VLSI circuits which is growing complex due to advancements in chip technologies, probability of design fault is more since a complete verification of the design is very difficult to achieve. Use of N-versions of VLSI circuits allow the continued use of chips with design faults as long as their errors at decision points are not similar.
- Software verification and validation time is reduced by executing two independent versions in an operating environment thereby completing verification and validation with production operation concurrently.
- Given a formal and an effective specification, different versions of software can be written by programmers working at their time and location using their own personal computing equipment. This "mail-order" approach [3]will drastically bring down the cost of programming that accrues in highly controlled professional programming environments.

IV. CHALLENGES OF NVP

Some of the key challenges of NVP are:

- The important condition for success of NVP is accurate specification of requirements. A series of experiments have been conducted and significant progress has occurred in the development of specification languages. Current goal of research is to compare and assess the ease of use of these methods by application programmers.

- NVP is based on the conjecture that software designed differently will cause very few similar errors at decision points. Though some researchers have developed guidelines and methodologies to achieve design diversity, implementation has remained as a complex issue and evaluation is based on qualitative arguments. Large-scale experiments need to be carried out to statistically evaluate the usefulness of these methods.
- Cost of using NVP is an other important issue. Generation of N versions of a given program instead of a single one increases the cost of software owing to escalated cost of development and supporting environment to complete the implementation. Peter Bishop[13] has argued that development and production cost can be reduced by applying design diversity only to critical paths. Effectiveness of this method however still needs to be quantitatively verified.

V. CONCLUSIONS

- With amazing advancements with hardware technology, the focus of Fault Tolerance is shifting from hardware to software. Research on Software Fault Tolerance is gaining momentum. Hardware reliability theory can not be directly applied to Software, owing to the complexity of Software
- N-Version programming approach of Software Fault Tolerance is based on *design diversity* conjecture. Independence of design and implementation effort with diverse programming languages, algorithms and environment will result in very low probability of similar errors at decision points, thereby increasing the Fault Tolerance capability of software
- Advantages of NVP are:
  - In areas where complete verification of design faults is difficult to attain due to problem complexity, implementation of NVP allows the continuos use of components with design faults, as long as the errors at decision points are not similar
  - Cost reduction can be achieved through "mail-order" approach and by carrying out verification and validation in production environment

VI. SCOPE FOR FURTHER RESEARCH

- Assessment and comparison of specification methods in terms of ease of use by application programmers
- Statistical evaluation of design diversity techniques based on real life data and metrics
- Cost investigations based on data derived out of real life deployment of NVP technique.

REFERENCES

[1] "Software Complexity and Its Impact on software Reliability", Ken S. Lew, Tharam Dillon, Kevin Forward, IEEE –Software Eng., vol 14, No 11, Nov 1988, pp 1645 – 1655.
[2] "Fault Tolerance and fault intolerance. Complimentary approaches to reliable computing", A.Avizienis, Proc. 1975 Int. Conf. Reliable Software, Los Angels, CA, Apr 21- 27, 1975, pp 458 - 464
[3] "N-Version Approach to fault tolerant Software", A.Avizienis, IEEE-Software eg., vol- SE11, No12, Dec 1985, pp.1491 -1501
[4] "System structure for Software Fault Tolerance," B.Randell, IEEE-Software Eng.,vol. SE-1,pp.220-232, June 1975.
[5] "Information processing systems-Reliability and requirements", Proc. East. Joint Comput. Conf., Washington, DC, Dec. 8-10, 1953.
[6] "A self correcting computer", J.Oblonsky, Digital Information processors, W.Hoffman, Ed. New York:Interscience, 1962, pp533-542
[7] "A NonStop operating system",J.F.Barlett, Proc. Hawai Int. Conf. Syst. Sci, Honolulu, HI, Jan. 5-6, 1978, pp 103-119. Reprinted in Theory and Practice of reliable System Design. Bedford, MA: Digital press, 1982, pp.453-460
[8] "Beyond Fault Tolerance", Timothy C.K.Chou, IEEE Computer, Apr 1997, pp 47-49.
[9] " An experiment on unit increase in program complexity", S.N. Woodfield, IEEE-Software Eng., vol-SE5, no 2, pp. 76-79, 1979
[10] "On the implementation of NVP for Fault Tolerance", A.Avizienis and L.Chen, Proc. COMPSAC 77, 1st IEEE-CS Int. Comput. Software. Appl. Conf., Chicago, IL, Nov 8-11, 1977, pp 149 – 155.
[11] "A Theoretical Investigation of Generalized Voters for Redundant Systems", Lorczak, Digest of Papers FTCS-19:The Nineteenth International Symposium on Fault-Tolerant Computing, 1989, pp.444 – 451
[12] "Dependable, Intelligent Voting for Real-Time Control Software, Engineering Applications of Artificial Intelligence, vol 8, no 6, Dec 1995, pp. 615 – 623
[13] "Software Fault Tolerance by Design Diversity", Peter Bishop, Software Fault Tolerance, John Wiley & Sons, 1995
[14] "Software Fault Tolerance: A Tutorial", Wilfredo Torres-Pomales, NASA Technical Memorandum, Oct 2000.
[15] "An Environment for Developing Fault-Tolerant Software", James M.Purtilo and Pankaj Jalote, IEEE-Software Eng., Vol 17, No2, Feb 1991