Srinivas Mishra Int. Journal of Engineering Research and Application *www.ijera.com*
ISSN : 2248-9622, Vol. 8, Issue 5, (Part -V) May 2018, pp.103-115

RESEARCH ARTICLE OPEN ACCESS

Probabilistic Scheduling Guarantees for Fault-Tolerant Real-Time Systems

Srinivas Mishra, Sambit Mishra
Gandhi Institute of Excellent Technocrats, Bhubaneswar, India
Suddhananda Engineering & Research Centre, Bhubaneswar, Odisha, India

ABSTRACT
Hard real-time systems are usually required to provide an absoluteguarantee that all tasks will always complete by their deadlines.Inthis paper we address fault tolerant hard real-time systems, and intro-ducethenotionofaprobabilisticguarantee.Schedulabilityanalysisisused together with sensitivity analysis to establish the maximum faultfrequency that a system can tolerate. The fault model is then used toderivea probability(likelihood)that,duringthelifetimeofthesystem,faults will not arrive faster than this maximum rate.The frameworkpresented is a general one that can accommodate transient 'software' faults,toleratedbyrecoveryblocksorexceptionhandling;ortransient'hardware'faultsdealtwithbystaterestorationandre-execution.

I. INTRODUCTION

Scheduling work in hard real-time systems is traditionally dominated by the no-tion of absolute guarantee.Static analysis is used to determine that all deadlinesare met even under the worst-case load conditions.With fault-tolerant hard real-time systems this deterministic view is usually preserved even though faults are,by their very nature, stochastic. No fault tolerant system can, however, cope withan arbitrary number of errors in a bounded time. The scheduling guarantee is thuspredicated on a fault model. If the faults are no worse than that defined in the faultmodel then all deadlines are guaranteed.The disadvantage of this separation ofscheduling guarantee and fault model is that it leads to simplistic analysis; eitherthesystemis schedulableoritisnot.

In this paper we bring together scheduling issues and errors to justify the notionof a probabilistic guarantee even for a hard real-time system.By 'probabilisticguarantee'wemeanaschedulingguaranteewithanassociatedprobability.Hence,
a guarantee of 99.95% does not mean that 99.95% of deadlines are met. Rather itimpliesthattheprobabilityofalldeadlinesbeingmetduringagivenperiodofoper-ationis99.95%.Insteadofstartingwiththefaultmodelandusingschedulingteststo see if this is feasible, we start with the scheduling analysis to derive a thresholdinterval between errors that can be tolerated and then employ the fault model toassignaprobabilitytothisthresholdvalue.

To provide the flexibility needed to program fault tolerance, fixed priority pre-emptive scheduling will be used [13]. The faults of interest are those that are tran-sient.Castilloatal[6]intheirstudyofseveralsystemsindicatethattheoccurrencesof transient faults are 10 to 50 times more frequent than permanent faults. In someapplicationsthisfrequencycan bequitelarge;oneexperimentonasatellitesystemobserved35transientfaultsina15minuteintervalduetocosmicrayions[5].

We attempt to keep the framework as general as possible by accommodating'software' faults tolerated by either exception handling or some form of recoveryblock, and 'hardware' faults dealt with by state restoration and re-execution. Errorlatencieswillbeassumedtobeshort.

Other authors have studied the probability of meeting deadlines in fault-tolerantsystems.However, only some facets of this problem have been considered.Forinstance,HouandShin[9]havestudiedarelatedproblem,theprobabilityofmeet-ingdeadlineswhentasksarereplicatedinahardware-redundantsystem.However,they only consider permanent faults without repair or recovery. A similar problemwasstudiedbyShinetal[18].Kimetal[12]consideranotherrelatedproblem:theprobability of a real-time controller meeting a deadline when subject to permanentfaultswithrepair.

Therestofthepaperisorganisedasfollows.Section2brieflydescribestheschedulinganalysisthatisapplicabletonon-fault-tolerantsystems.Section3presentsthe faultmodel andthe frameworkfor the subsequent analysis.InSection4thescheduling analysis for a fault tolerant system is presented. This enables the thresh-oldfaultinterval(TFI)tobederived.Section5thenusesthefaultmodelandtheTFItoassignaprobabilitytothethreshold.ConclusionsarepresentedinSection6.

Standard Scheduling Analysis

Forthestandardfixedpriorityapproach,itisassumedthatthereisafinitenumber(N)oftasks($_{1..N}$).Each taskhastheattributesofminimuminterarrivaltime, T,worst-case execution time, C, deadline, D and priority P . Each task undertakes apotentiallyunboundednumberofinvocations;eachmustbefinishedbythedeadline(which is measured relative to the task's invocation/release time). All tasks aredeemed to start their execution at time 0. We assume a single processor platformand restrict the model to tasks withD T .For this restriction,an optimal setof priorities can be derived such that $D_i<D_j$) $P_i>P_j$for all tasks i; j[15].Tasksmaybeperiodicorsporadic(aslongastwoconsecutivereleasesareseparatedby at least T).Once released, a task is not suspended other than by the possibleactionofaconcurrencycontrolprotocolsurroundingtheuseofshareddata.Atask, however,maybepreemptedatanytimebyahigherprioritytask.Systemoverheadssuch as context switches and kernel manipulations of delay queues etc can easilybeincorporatedintothemodel[11,4]butareignorededhere.

The worst-case response time (completion time) $R_i$for each task (i) is obtainedfromthefollowing[10,1]:

$$R_i = C_i + B_i + \sum \left\lceil \frac{R_i}{T_j} \right\rceil C_j \qquad (1)$$

The most common and effective concurrency control protocol assigns a ceilingpriority to each shared data area. This ceiling is the maximum priority of all tasksthat use the shared data area. When a task enters the protected object that containsthe shared data, its priority is temporarily increased to this ceiling value.As aconsequence(onasingleprocessorsystem):

1. Mutualexclusionisassured(bytheprotocolitself).
2. Eachtaskisonlyblockedonceduringeachinvocation.
3. Deadlocksareprevented(bytheprotocolitself).

ThevalueofB$_i$issimplythemaximumcomputationtimeofanyprotectedobjectthathasaceilingequalorgreaterthanP$_i$andisusedbyataskwithaprioritylowerthanP$_i$.

Table 1 describes a simple 4 task system, together with the response times thatare calculated by equation (2).Priorities are ordered from 1, with 4 the lowestvalue,andblockingtimeshavebeensettozeroforsimplicity.Schedulinganalysisis independent of time units and hence simple integer values are used (they can beinterpretedas milliseconds).

To illustrate how these values are obtained consider$_4$; $r^0$is given the initialvalueof30,r^1isthenjusttheadditionofallthecomputationtimes(30+35+25+30=120),sor^{2}isassigned120.Withthisvalue$_1$givesrisetoanotherhit(of30)

Task	P	T	C	D	B	R	Schedulable
$_1$	1	100	30	100	0	30	TRUE
$_2$	2	175	35	175	0	65	TRUE
$_3$	3	200	25	200	0	90	TRUE
$_4$	4	300	30	300	0	150	TRUE

Table1:ExampleTaskSet

andhencer^{3}is150.Thisvalueisthenstableandhenceistherequiredresponse time.

All tasks are released at time 0. For the purpose of schedulability analysis, wecan assume that their behaviour is repeated every LCM, where LCM is the leastcommon multiple of the task periods. When faults are introduced it will be neces-sary to know for how long the system will be executing. Let L be the lifetime ofthe system. For convenience we assume L is an integer multiple of the LCM. Thisvaluemayhoweverbeverylarge(forexampleLCMcouldbe200ms,andLfifteenyears!).

FaultModel

We assume that a single transient fault will cause just one error, and that thiserrorwillmanifestitselfinjustasingletask.With'software'faultsthisisareason-ableassumption.With'hardware'faultsweareconcernedwitherrorsthatmanifestthemselves in the processing unit (including internal busses, cache etc) rather thanin memory where the error latencies may be very large. We assume that only theexecuting task is affected[1]. Faults that affect the kernel must either be masked orlead to permanent

damage that can only be catered for by replication at the systemlevel. To make the subsequent analysis simpler we assume perfect error recognition coverage; a probabilistic (non zero) measure of coverage could be used with astraightforwardeffect upontheanalysis.
WemakethecommonhomogeneousPoissonprocess(HPP)assumptionsthatthefaultarrivalrateisconstantandthatthedistributionofthefault-countforanyfixedtimeintervalcanbeapproximatedusingaPoissonprobabilitydistribution.Thisis an appropriate model for a random process where the probability of an eventdoesnotchangewithtimeandtheoccurrenceofonefaulteventdoesnotaffecttheprobability of another such event. A HPP process depends only on one parameter,viz, the expected number of events,, in unit time; here events are transient faultswith=1=MTBF,whereMTBFistheMeanTime BetweentransientFaults[2].
PerthedefinitionofaPoissonDistribution,

$$Pr_n(t) = \frac{e^{t}(t)^n}{n!}$$

givestheprobabilityofneventsduringanintervalofdurationt.IfwetakeaneventtobeanoccurrenceofatransientfaultandYtobetherandomvariablerepresentingthe number of faults in the lifetime of the system (L), then the probability of zerofaultsis givenby

$$Pr(Y=0)=e^{L}$$

andtheprobabilityofatleastonefault

$$Pr(Y>0)=1\ \ e^{L}$$

Otherusefulvaluesare:

$$Pr(Y=1)= e^{L}L$$

$$Pr(Y<2) = e^{L}(1+ L) \quad (3)$$

Weareconcerned,inthispaper,withtheprobabilityofthesystembeingschedu-lable. We shall write P r(S) and P r(U) to denote the probability of schedulabilityand unschedulability.OfcoursePr(S)=1Pr(U).
Theanalysisgiveninthenextsectionwilldeterminethethresholdfaultinterval.This gives the sustainable frequency at which faults can occur and the system stillmeet all its deadlines.Let this frequency be represented by the minimum timeinterval allowed between faults, T_F.It follows that if Wis the shortest intervalbetweenfaultarrivalsduringamissionthen[3]

$Pr(U)=Pr(Ujnofaults):Pr(nofaults)$
$+Pr(UjWT_Fandtherearefaults):Pr(W$
$T_Fandtherearefaults)$
$+Pr(UjW<T_Fandtherearefaults):Pr(W<T_Fandtherearefaults)$

Since we are dealing with systems which are schedulable 'under no faults' we canassume Pr(Ujnofaults)iszero.Also$T_F$hasbeendefinedsothatPr(UjWT_F)iszero.Hence

$$Pr(U)=Pr(UjW<T_F):Pr(W<T_F)$$

In this paper we will make the conservative assumption that P r(U jW< T_F) isone.AndhencewearleftwiththeevaluationofPr(W<T_F),i.e.theprobability that at least two faults arrive so close together in time that they cannot both betolerated.This is done in Section 5.Although this assumption is conservative(and hence safe) it is clearly possible to give less pessimistic values.The aboveformulation will allow such values to be combined with the estimates of P r(W <T_F)giveninSection5.
Issues concerned with implementing the features suggested by the Fault Modelare welladdressedbyFetzerandCristian[8].

TypicalValuesofKeyParameters

Before proceeding with the analysis it is worth noting the ranges in value of thekey parameters of the model. In most applications of interest, the "lifetime" overwhich a probability of failure is required is the duration of one mission. Missiontimes for civil aircraft are typically 3-20 hours, but for satellites 15 years of ex-ecution may be expected.The iteration periods for control loops are as short as20ms, other loops and signals may have T values of a few seconds. Precise valuesfor MT BF are not generally known, but in a friendly operating environment per-haps100hoursisnotunreasonable.Inmorehostileconditions,20secondsmaybemore typical. Although $T_F$is derived from the characteristics of the task set underconsideration, it is worth noting that very small values are unlikely (as a task willnot make progress if it suffers repetitive faults), and faults spaced out beyond theLCM of the task periods will easily be catered for; hence: 200ms < T_F<5 Sec-onds.Table2summarizestheseviablerangesforthekey parameters(inhoursandhours[1]).

Parameter	Range
L	3 10^5
T	10^6 10^2
	10^2 10^2
T_F	10^5 10^2

Table2:TypicalValuesofKeyParameters

SchedulabilityAnalysisforFaultTolerantExecution

Let $F_k$be the extra computation time needed by$_k$if an error is detected duringitsexecution.Thiscouldrepresentthere-executionofthetask,theexecutionofanexceptionhandlerorrecoveryblock,orthepartialre-executionofataskwithcheckpoints. In the scheduling analysis the execution of task$_i$will be affected byafaultin$_i$oranyhigherprioritytask.Weassumethatanyextracomputationforataskwillbeexecutedatthetask's(fixed)priority.

Henceifthereisjustasinglefault,equation(1)willbecome[16,2][4]:

R_i=C_i

+B_i *i*

+

j2hp(i)

R_i C

+ maxF_kk2hep(i)

(4)

wherehep(i)isthesetoftaskswithpriorityequalorhigherthan$_i$,thatishep(i)=hp(i)+$_i$.

This equation canagain be solved for$R_i$by forming arecurrencerelation.Ifall $R_i$values are still less than the corresponding $D_i$values then a deterministicguaranteeis furnished.

Giventhatafaulttolerantsystemhasbeenbuiltitcanbeassumed(althoughthiswouldneedtobeverified)thatitwillbeabletotolerateasingleisolatedfault.Andhencethemorerealisticproblemisthatofmultiplefaults;atsomepointallsystemswillbecomeunschedulablewhenfacedwithanarbitrarynumberoffaultevents.

To consider maximum arrival rates, first assume that $T_f$is a known minimumarrivalintervalforfaultevents.Alsoassumetheerrorlatencyiszero(thisrestrictionwillberemovedshortly).Equation(4)becomes[16,2]:

R_i=C_i

X

+B_i+

R_i C

&

+ $\frac{R_i}{}$ maxF

(5)

j

j2hp(i) T_j

T_f k2hep(i) *k*

Thus in interval (0 R] there can be at most $\frac{R_i}{T_f}$fault events, each of which caninduce $F_k$amount of each computation. The validity of this equation comes fromnotingthatfaulteventsbehaveidenticallytosporadictasks,andtheyarerepresentedin the scheduling analysis in this way [1]. Note the equation is not exact (but it issufficient):faultsneednotalwaysinduceamaximumre-executionload.

There is a useful analogy between release jitter and error latency. If a fault canlie dormant for time A_f, then this may cause two errors to appear to come closertogetherthanT_f.Thiswillincreasetheimpactofthefaultrecovery.Equation (5) can be modified to include error latency in the same way that release jitter isincorporatedintothestandardanalysis[1]:

R_i=C_i

X

+B_i+

R_i C

&

+ $\frac{R_i+A_f}{}$ maxF

(6)

j2hp(i) T_j
T_f k2hep(i)

As before, this equation can be solved for $R_i$by forming a recurrency relationship.Table3givesanexample ofapplyingequation(6).Herefullre-executionisrequiredfollowingafault.Twodifferentfaultarrivalintervalsareconsidered.Foronethesystemremainsschedulable,butfortheshorterintervalthefinalta

j

skcannotbeguaranteed.Inthissimpleexample,blockinganderrorlatencyareassumedtobe zero.Notethatforthe first *k* threetasks, thenewresponsetimes areless thantheshorter$T_f$value, andhencewillremainconstantforall$T_f$values greaterthan200.Theabove analysishasassumedthatthetaskdeadlines,Ds,remainineffectevenduringafaulthandlingsituation.Some systemsallowarelaxeddeadline

Task	P	T	C	D	F	R T_f=300	R T_f=200
1	1	100	30	100	30	60	60
2	2	175	35	175	35	100	100
3	3	200	25	200	25	155	155
4	4	300	30	300	30	275	UNSCH

Table3:ExampleTaskSet-T_f=300/200

whenfaultsoccur(aslongasfaultsarerare).Thisiseasily accommodatedintotheanalysis.

LimitstoSchedulability

Havingformedtherelationbetweenschedulabilityand T_f,itispossibletoapplysensitivityanalysistoequation(6)tofindtheminimumvalueof$T_f$thatleadstothesystem being just schedulable. As indicated earlier, let this value be denoted as T_F(itis thethresholdfault interval).

Sensitivity analysis [19, 14, 13, 17] is used with fixed priority systems to inves-tigate the relationship between values of key task parameters and schedulability.Foranunschedulablesystemitcaneasilygenerate(usingsimplebranchandboundtechniques) factors such as the percentage by which all Cs must be reduced for thesystemtobecomeschedulable. Similarlyforschedulablesystems,sensitivityanalysis canbeusedtoinvestigatethe amount by which the load can be increased without jeopardising the deadlineguarantees.Hereweapplysensitivityanalysis to$T_f$toobtainT_F.

When the above task set is subject to sensitivity analysis it yields a value of $T_F$of 275. The behaviour of the system with this threshold fault interval is shown inTable4.Avalueof274wouldcause$_4$tomissitsdeadline.

Task	P	T	C	D	R T_F=275
1	1	100	30	100	60
2	2	175	35	175	100
3	3	200	25	200	155
4	4	300	30	300	275

Table4:ExampleTaskSet-$T_F$setat275

2 EvaluatingPr(W<T_F)

Weneedtocalculatetheprobabilitythatduringthelifetime,L,ofthesystemnotwo faults will be closer than T_F.Two approaches are considered. The attractionof the first is that it shows that a relatively intuitive and uncomplicated approachyields upper and lower bounds on P r(W <T_F) which, for a wide range of parame-tervalues,provideamaximumapproximationerrorwhichcannotbemuchgreater thanafactorof3(since$\frac{\text{upperbound}}{\text{lowerbound}}$3).Withthesecondapproachamorecum-bersome but exact formulation is derived. Despite the inclusion of this latter exactformulation, we believe that, given that it is often rather the order of magnitude ofthe failure probability that is the primary concern (rather than an exact value), themathematicallysignificantlyeasierreasoningofthefirst,boundingapproachretainssomeimportance.

5.1 UpperandLowerBoundsforEvaluating Pr(W<T_F)

WeareconcernedwithtwofaultsbeingcloserthanT$_F$overthemissiontime L.Since in practice L T$_F$we can assume, without loss of generality, that L isan even integer multiple of T$_F$. Let mishap be the undesirable event of two faultsindeedoccurringcloserthanT$_F$i.e.
Pr(mishapduringL) Pr(W<T$_F$)
Wederivetherequiredupperandlowerboundsviathefollowingtheorems:

Theorem 1IfL=(2T$_F$)isapositiveintegerthen

Pr(mishapduringL)<1+ e TF F
(1+T$_F$
L
)TF l
1 2 e2 T
(1+2T$_F$
L
)2TF l

Theorem 2IfL=(2T$_F$)isapositiveintegerthen

Pr(mishapduringL)>1 e TF(1+ T F
ProofofTheorem1
L
)TF l

Letthemissiontimebesplitintoaseriesof'even'timeintervalswithboundaries0,2T$_F$,4T$_F$,:::,L,asshowninFigure1.Similarlyasetof'odd'intervalsstartingat times T$_F$, 3T$_F$, 5T$_F$, : : : , LT$_F$can be defined (extending the lifetime slightlyto L+T$_F$, the end point of the last odd interval, by continuing the same HPP faultmodel). Each set has L=2T$_F$intervals. Let a mishap be said to lie in an interval ifboth of its faults occur during that interval. It follows from the geometry of theseintervalsthat
mishapduringL) mishapinsomeeveninterval[s];
ormishapinsomeoddinterval[s]

Thispropertycomesdirectlyfromthedefinitionoftheintervals;ifamishap(twofaultscloserthanT$_F$)occursit mustlieineitheranevenoranoddinterval[5].
Pr(mishapduringL)<Pr(mishapinsomeeveninterval[s]
ormishapinsomeoddinterval[s])

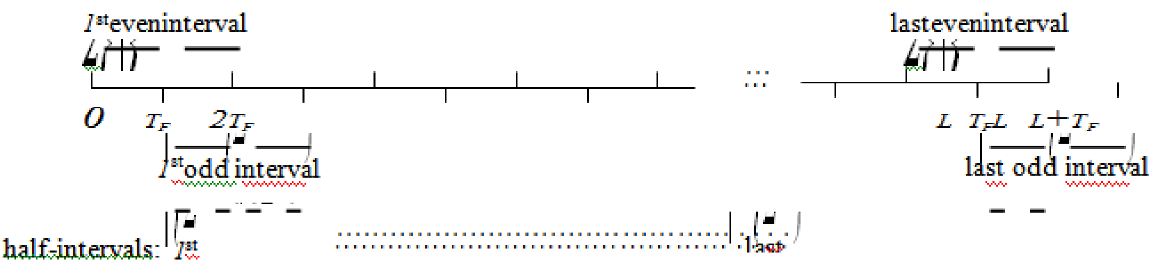

Figure1:Definitionsof'evenintervals','oddintervals'and'half-intervals'

Actually the intersection of the two events on the right hand side has non-zeroprobability.Onewaythattheycanoccurtogetherist hatasinglemishapcouldlie in the overlap between an even and an odd interval. Call these overlaps 'half-intervals': they are of length T, and there are $^6\underline{L}$ 1 of them, respectively startingat timesT_F,$2T_F$,:::,LT_F.So from thebasicaxiomsofprobability

Pr(mishapinsomeeven interval[s];or mishapin someodd interval[s])<
Pr(mishapinsomeeveninterval[s])+P
r(mishapinsomeoddinterval[s])Pr(mishapinsome
half-interval[s])
Now,giventhesymmetryoftheconstructionandtheHP Pprocessassumption,
Pr(mishapinsomeeveninterval[s])=Pr(mishapinsom eoddinterval[s]) F F F
Hence

Pr(mishapduringL)<2Pr(mishapinsomeeveninterval [s])
Pr(mishapinsomehalf-interval[s]) (7)
Theevent"mishapinaparticulareveninterval"isindependentofeventsinallotherevenintervals,andithasthesa meprobabilityforeveryeveninterval.Thus
$\underline{L}$Pr(mishapinsomeeveninterval[s])=1Pr(nomishapi n2T_F)^{2T_F}:
(8)

Foranintervaloflength2$T_F$nottocontainamishap,itiss ufficient(butnotneces-sary)thatitcontain0or1fault.Hence,fromequation(3)
Pr(nomishapin2T)>e^{2T_F} (1+2T): F
(9)Combiningequations(8)and(9)yields

Pr(mishapinsomeeveninterval[s])<1 e^2 T_F(1+2 T
Byasimilarargumentforthehalf-intervals
$\underline{L}$
)^{2T_F}: 2

(10)

Pr(mishapinsomehalf-interval[s])=1 e T_F

(1+T_F

$\underline{L}$ 1 T_F
)^{T_F} ;

(11)
andnowcombiningequations(7),(10),and(11)delivers thetheoremstatement.
2

ProofofTheorem2
In a similar way to the previous proof, consider the series of intervals of lengthTstarting at times 0, T, 2T, 3T, : : :, L T_FT. There are $\underline{L}$ of these, a mishapinanyoneofwhichimpliesamishapduring L(but notviceversa).Hence
Pr(mishapduringL)>Pr(mishapinsomeinterval[s])
but
Pr(mishap insomeinterval[s])=1 Pr(nomishap inanyinterval)
Prooffollowsdirectly(asinproofofTheorem1).
2
Both the upper and lower (exact) bounds are in mathematically non-intuitiveforms, but simple approximations can be derived for most of the parameter rangewithinwhichtheprobabilityofmishapis smallenoughtobeofinterest.

Corollary3
AnapproximationfortheupperboundonPr(W<T_F)givenbyTheorem1is
^{32}LT_F,providedthatT_F, ^{2}L$T_F$aresmall,andL T_F.

Corollary4
AnapproximationforthelowerboundonPr(W<T_F)givenbyTheorem2is
^{12}LT_F,providedonlythatT_F, ^{2}L$T_F$aresmall.

ProofofCorollary3

Theterme ^{T_F} $(1+T_F)$canbeapproximatedbyaTaylorseries,whereterms $(T_F)^3$andbeyondareignored.Thus

$$e^{T_F}(1+T_F) \approx 1 - \frac{T_F^2}{2}$$

Anotherapproximationcomesfromnotingthatforsmallxz^2

$$\frac{1}{z^2} \approx x\,\frac{1}{2}\,xz^2$$

wheretermsz^{4}andhigherpowersofzcanbeignored.Hence,underassumptions T_F,^2LT_Fsmall,andL T_F,wecanwrite

$$\left[e^{T_F}(1+T_F)\right]^{\frac{L}{2T_F}-1} \approx 1 - \frac{L\,T_F}{2} \qquad (12)$$

Under justthefirstassumptions thatT_F,^2LT_Faresmall wehaveequally

$$\left[e^{2T_F}(1+2T_F)\right]^{\frac{L}{2T_F}} \approx 1 - L^2T_F^2 \qquad (13)$$

Applying(12)and(13)totheconclusionofTheorem1,Corollary3isproved.

Corollary4followsbyasimilarargument.

Strictly, the bounds in Theorems 1 and 2 have only been proved here for L aneven multiple of T_F. However, the realistic assumption L$T_F$allows the approximationsgiveninthetwocorollariesstilltoextendtootheervaluesof$T_F$andL.

Infact,where $\frac{L}{T_F}$ isanexactinteger,thisL $T_F$assumptionisnotactuallyre-quired, either for the derivation of the exact bounds in Theorems 1 and 2, nor forthe lower-bound approximation of Corollary 4. For high accuracy in Corollary 4,we need only the assumption of smallT_F,^2LT_F.Corollary 3 is the exception,relying on the L$T_F$assumption at one place in its derivation: the exponent to thesquare-bracketed term in (12) is 'out by 1' and we needed $\frac{T_F}{L}$ to be small in orderto justify effectively ignoring this fact. For very short mission times, such that wedonothaveL T_F,wecaninfact'retreatslightly'toaslackerupperboundforPr(mishapduringL)byusing1asanupperboundforthissquare-bracketedterm inamodifiedversionofTheorem1,thusavoidingtheawkwardexponent $\frac{L}{T_F}$1.

Then,forpositiveintegral $\frac{L}{T_F}$,the resultingequivalentofCorollary 3produces an accurateapproximation2^2LT_FtothisslackerupperboundonPr(W<T_F),withoutanyrequirementthatL T_F;i.e.,undertheassumptionsonlythatT_F, ^2LT_F aresmall,andthat $\frac{L}{T_F}$ isapositiveinteger,butwithoutnowtherequirementthat L T_F,themethodsofthissubsectionareabletoprovideboundsonPr(W<T_F) whichareapproximatelyintheratio $\frac{\text{upperbound}}{\text{lowerbound}}$ 4.

The important upper bound approximation of Corollary 3 can be written in theform ^{3}(L)(T_F).It will often be the case thatT_F< 10^2; indeed this con-straint allowed the approximations to deliver useful values. ButL can vary quiteconsiderablyfrom10^2orlessinfriendlyenvironmentsto10^3ormoreinlong-life,hostiledomains.Clearly,lowprobabilitylevelsforthislattercasewillbeextremelydifficulttoachievebytheschedulingapproachdefinedinthispaper.

L	1	10^2	10^4
1	1.110^4	1.110^8	1.110^{12}
10^1	1.110^3	1.110^7	1.110^{11}

10^2	1.110^2	1.110^6	1.110^{10}	
10^4	1	1.110^4	1.110^8	

Table5:UpperboundonNon-SchedulabilityduetoFaults.

TheexampleintroducedinSection4hadaT$_F$valueof27 5ms.Table5givestheupperboundontheprobabilityguaranteeforvariousvaluesofandL.

WhenL<10^2,L approximates the probability of any fault happening duringthe mission of duration L. So, ^{2}(T$_F$)^1 represents the gain that is achieved by theuseoffaulttolerance,undertheotherassumptionsstated.So,forexampleinTable5,when=10^2andL =1thegainisapproximately10^6.

ExactFormulationforEvaluatingPr(W<T$_F$)

Unlikethebounding argumentusedinthelastsection,ourexactderivationofthe probability P r(W <T$_F$) proceeds in two stages, first conditioning on the totalnumber n of faults seen in the lifetime L of the system. It is a well known propertyoftheHPPprocess[7]thatifweconditionon the numbernofeventsoccurringwithinaspecifiedtime intervalandthendefine$X_1;X_2;:::;X_n$asorderedpositions ofthesenpointswithinthatinterval,expressedasproportionsofitslength,thenthe$X_i$are (conditionallygivenn) jointlydistributedastheorderstatisticsofan i.i.d. random sample from a uniform distribution on the unit interval[0; 1].Thisbeing accepted, we now first fix u with 0u1 and ask the question 'What is theprobability,Psay,thatnotwoofthesepointsarecloser thanu(conditionallygivenn)?'. We can obtain the answer by n-dimensional integration. This is reported inan extended version of this paper available as a technical report [3], which saysessentially that Pis just the nth power of the total amount of 'slack' remainingwithintheunitintervalafterouru-separationconstraintisimposed.

Thissolutionconditionalgivennenablesustocomplete theexactderivationof the final, unconditional P r(W <T$_F$) relatively straightforwardly by using the'chain rule' of conditional probability to 'uncondition on n'. Another fundamentalpropertyofthehomogeneousPoissonprocessisthatthedistributionofn,thecountof the number of events occurring within a fixed time interval, upon which theprobabilitiesareconditioned,isPoissonwithparameterequaltoitsmean,whichinourcaseisL.Then,workingforconveniencewiththe'probabilityofnomishap inatimeintervalof length L',wehave

Pr(W T$_F$) =

P
n=0

n;(T$_F$=L)
:e^{L}:(L)^n / n!

=e^{L} 1+ L+

X
n=2

n
1 (n 1) F T
L

(L)^n
:
n!

=e^{L} 1+ L+

X
n=2
(L (n 1)T$_F$)^n
n!
(14)

A few remarks about this exact expression We remark firstly that (14) is essentially afunction of just two arguments,L,T$_F$, rather than three (as are the bounds de-rived in Section 5.1). Thinking now of the function mathematically in these terms,withoutmuchconcerningourselvesaboutphysicalinterpretationofthearguments,if we agree to confine ourselves to the ranges 0 <L <1, and $0T_F$<1,thenweremarkthattheexpression(14)continuestogivethecorrectmathematicalPoissonprocessprobabilityatallpointsofthisdomain,includingthevalueof1obtained atT_F =0.(This is on the understanding that the d1e occurring as theupper limit of a sum denotes a sum to infinity in the usual sense of a mathematicallimit.) The purpose of stating this last point about the argument domain now to beassumed for this function is related to the practical computation problem associ-ated with (14) which we address briefly in [3].Note that, apart from thisT_F =0case, the expression (14) represents a finite sum throughout the domain identified,although, for certain argument values, the number of terms summed can be astro-nomically large, which can make a simple-minded numerical computation ratherslow. Moreover, some of these awkward parameter ranges may be of real practicalinteresttousinourapplication(seeendofSection3).

Notethatwecanusethecommonnotationforthe‘positivepartfunction’h+, associatedwithanyreal-valuedfunctionh,toobtainthefollowingslightlydifferentexpression,validthroughouttheargumentdomainwehavejustspecified(including T_F=0).

(1 $(L(n\ 1)T_F)^n$)

Pr(W T_F

)=e^L

1+ L+

n=2

+ (15)

n!

SomeNumericalResultsonPr(W<T_F)

Wedecidedtotesttheaccuracyofournumericalapproximationsexperimentally,and found that, over the physically realistic parameter ranges of concern to us, theapproximationsdefinedareextremelyaccurate,evenatverylowordersintheTaylorseries.This enabled us to produce Figure 2, a contour plot indicating the depen-dence of the exact value of P r(W <T_F) on its two argumentsL,T_F. The func-tionplottedis,infact,thelogoddsofPr(W<T_F),choseninordertoensurethat

there are some contours near each extreme, P r(W <T_F) = 0 and P r(W <T_F) = 1.In the top right hand corner the contours bunch too closely as the probability of a‘mishap during L’ becomes extremely close to absolute certainty. (It is difficult toimagineasituationinwhichtheprecisevaluesofthese largeprobabilitieswouldbeof practical interest.) The rectangular box indicates a subdomain of the argumentsoverwhichwehavealsoplottedtheaccuracy ofourTaylorseriesapproximationtothis exact P r(W <T_F) function. The technical report[3] contains plots of the per-centage inaccuracy that results from the truncation of the approximation after oneortwoterms.

Λ

Contoursof $\log_{10}[P(W<T_F)/P(W>T_F)]$

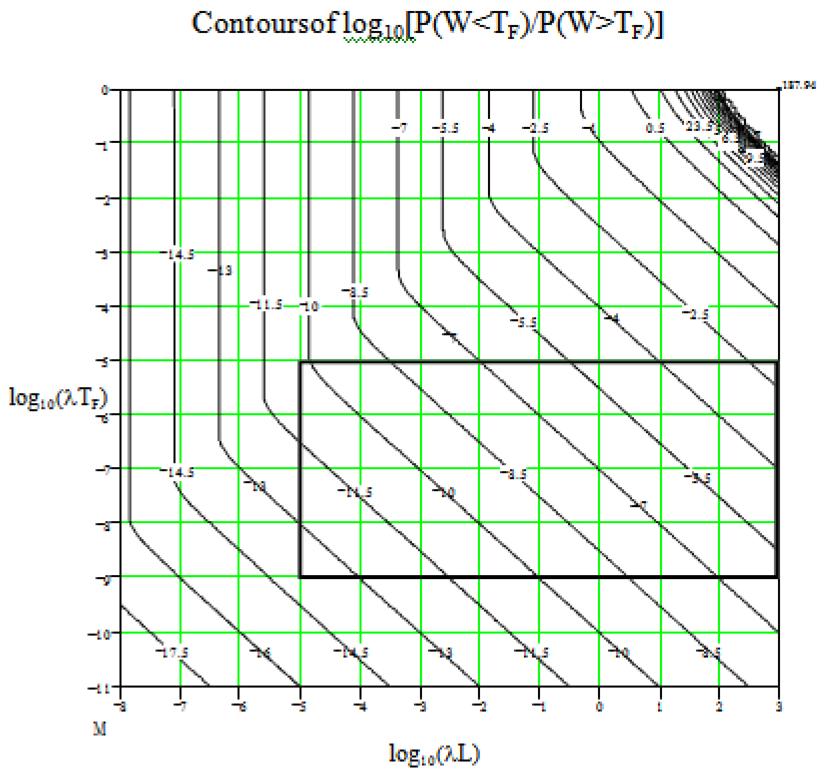

Figure2:PlotsofExactValue.Noticethelog-logscale.

We can illustrate in more detail the interpretation of our numerical results andplots briefly by examining one particular case. Assume L= 10^2 and T_F =10^5.Thatis,oursystemencountersfaultswithanMTBFof100timesitslifetime.It is guaranteed to be schedulable provided that it does not, during its lifetime, ex-perience two faults separated by less than one thousandth of the lifetime duration.In such circumstances, we would clearly expect the system to be schedulable withahighprobability,Psay.Thisisalog-oddscontourplot,sotheproximitytothe-7contour indicates that the odds in favour of a system being schedulable with theseparameter values are approximately 10^{7}to 1. In fact, the bounds on the probabilityof schedulability, in this situation, obtained by the ‘order-of-magnitude’ argumentof Section 5.1, are 0:499996710^7 and 1:50047710^7. The approximations tothese bounds, obtained in the two corollaries in Section 5.1, are 0:510^7 and1:510^7, exactly. The Taylor series approximation allows, in this case, almost ar-bitrarilyaccuratecalculationofthetruevalueofPwithcomparativelyfewtermsoftheseries.Infactweprovedthatalleven-orderpartialsums,uptothe1000th-order sum, are lower bounds on P , and all odd-order partial sums, up to the 1001th-ordersum,areupperbounds.Withtheseparticularargu ments,themodulusofthefourthorder term in the series is less than 10^{19}, so the sum to only three terms wouldgive an accuracy guaranteed to be better than approximately 11 or 12 significantdecimal figures.To eight significant figures, the value of P is :999484961 0^7,correspondingtoalog-oddsveryclosetо7inFigure2(atcoordinates(2;5)).The seriesapproximation gives first andsecond order Taylor approximations of10^7 and :999484951 0^7, respectively. (These numbers are both exact.) See [3]foradetailedderivationoftheseresultsconcerninghighnumericalaccuracy.

II. CONCLUSION

Wehavedevelopedthenotionofaprobabilisticschedulingguaranteeandshownhowitcanbederivedfromthestochasticbehaviouroffaultevents.Itisreasonableto assume that a fault tolerant system will be designed so as to remain schedulablewhendealingwithasinglefault.Themainresultofthepaperisthusthederivationof a probabilistic guarantee for systems experiencing multiple faults. To do this ithas been necessary to formulate a prediction of the likelihood of faults occurringcloser together than some specified distance in time. It has also been necessary touse sensitivity analysis to determine the limits to schedulability; that is, the minimumtolerableintervalbetweenfaults.

Although exact analysis is given for the likelihood of faults occurring quickerthan the rate obtained from the sensitivity analysis, perhaps the main result of thispaperisasimplederivedupperboundforthisprobability(asgiveninCorollary3).Atypicaloutcomeofthisanalysisisthatinasystemthathasalifetimeof10hourswith a mean time between transient faults of 1000 hours and a tolerance of faultsthatdonotappearcloserthan1/100ofanhour,theprobabilityofmissingadeadlineis upper bounded by 1.510^7.A lower bound is also derived (Corollary 4) andthisyieldsavalueof0.5 10^7.Fortheseparameterstheexactanalysisproducesavalueveryclosetol.0$10^7$.

Interestingly(and perhapsnottotallyintuitively)theupper, lowerand exact for-mulations for the probabilistic scheduling guarantee all indicate that the thresholdvalue derived from the scheduling and sensitivity analysis has a linear relationshiptotheprobabilisticguarantee.IfthethresholdvalueT$_F$ishalved,theprobabilityofmissing a deadline is halved. Similarly the length L of execution of the system hasalinearimpact.

Themainobstacletotheuseofsomeoftheanalysisgiveninthispaperisthelackofempiricaldataconcerningfaultarrivaltimes.Inthefutureweaimtoaddressfaultclusteringandlessfavourablefaultprocessmodels.Wealsoaimtomoveawayfromthe conservative assumption that the system is unschedulable (with probability 1)whenfaultsarrivecloserthanthethresholdvalue.

REFERENCES

[1] N. C. Audsley, A. Burns, M. Richardson, K. Tindell, and A. J. Wellings. Ap-plying new scheduling theory to static priority pre-emptive scheduling. Soft-wareEngineeringJournal,8(5):284–292,1993.

[2] A. Burns, R. I. Davis, and S. Punnekkat. Feasibility analysis of fault-tolerantreal-time task sets.Euromicro Real-Time Systems Workshop, pages 29–33,June1996.

[3] A.Burns,S.Punnekkat,L.Strigini,andD.R.Wright.Probabilisticschedulingguarantees for fault-tolerant real-time systems.Technical Report YCS.311,DepartmentofComputerScience,UniversityofYork,1998.

[4] A. Burns and A. J. Wellings.Engineering a hard real-time system:Fromtheorytopractice.Software-PracticeandExperience,25(7):705–26,1995.

[5] A.Campbell,P.McDonald,andK.Ray.Singleeventupsetratesinspace.IEEETransactionsonNuclearScience,39(6):1828–1835,December1992.

[6] X. Castillo, S.P. McConnel, and D.P. Siewiorek. Derivation and Calibrationof a Transient Error Reliability Model.IEEE Transactions on Computers,31(7):658–671,July1982.

[7] D.R.CoxandP.A.W.Lewis.StatisticalAnalysis ofSeriesofEvents.Methuen'sMonographsonAppliedProbabilityandStatistics,London,1966.

[8] C. Fetzer and F. Cristian. Fail-awareness: An approach to construct fail-safeapplications.InProceedingsofthe27thInt.Conf.onFault-TolerantComputerSystems(FTCS),pages 282–291,1997.

[9] C.-J. Hou and K. G. Shin.Allocation of periodic task modules with prece-dence and deadline constraints in distributed real-time systems. IEEE Trans-actionsonComputers,46(12):1338–1356,1997.

[10] M. Joseph and P. Pandya. Finding response times in a real-time system. BCSComputerJournal,29(5):390–395,1986.

[11] D.I. Katcher, H. Arakawa, and J.K. Strosnider.Engineering and analysis offixedpriorityschedulers.IEEE Trans.Softw.Eng.,19,1993.

[12] H. Kim, A.L.White, and K. G.Shin.Reliability modeling of hard real-timesystems.InProceedings28thInt.Symp.onFault-TolerantComputing(FTCS-28),pages 304–313.IEEEComputerSocietyPress,1998.

[13] M.H.Klein,T.A.Ralya,B.Pollak,R.Obenza,andM.G.Harbour. APracti-tioner'sHandbookforReal-TimeAnalysis:AGuidetoRateMonotonicAnal-ysisforReal-TimeSystems.KluwerAcademicPublishers,1993.

[14] J.P.Lehoczky,L.Sha,andV.Ding.Theratemonotonicschedulingalgorithm:Exactcharacterizationandaveragecasebehavior.Techreport,DepartmentofStatistics,Carnegie-Mellon,1987.

[15] J.Y.T.LeungandJ.Whitehead. Onthecomplexityoffixed-priorityschedulingofperiodic,real-timetasks.PerformanceEvaluation(Netherlands),2(4):237–250,1982.

[16] S. Punnekkat. Schedulability Analysis for Fault Tolerant Real-time Systems.PhDthesis,Dept.ComputerScience, UniversityofYork,June1997.

[17] S. Punnekkat, R. Davis, and A. Burns.Sensitivity analysis of real-time tasksets.In Proceedings of the Conference of Advances in Computing Science - ASIAN'97,pages72–82.Springer,1997.

[18] K.G.Shin,M.Krishna,andY.H.Lee.Aunifiedmethodforevaluatingreal-time computer controllers its application.IEEE Transactions on AutomaticControl,30:357–366,1985.
[19] S.Vestal.FixedPrioritySensitivityAnalysisfor LinearComputeTimeModels.IEEETransactionsonSoftwareEngineering,20(4):308–317,April1994.