

Hoe betrouwbaar is de Maeslantkering?

De vraag hoe veilig de Maeslantkering is, is in feite een vraag naar de betrouwbaarheid van de gebruikte software. Daar zitten hoe dan ook fouten in.

Op dit moment is er veel controverse over de Maeslantkering en dan met name over de betrouwbaarheid van het belissingsondersteunend systeem, het besturingssysteem en het informatie aanleverende systeem. Er worden allerlei getallen genoemd over de betrouwbaarheid van dit agglomeraat. In het AD stond dat best-in-class benchmarks voor een kleine duizend systeemsoftwareprojecten laten zien dat er zelfs bij 98 procent defect removal efficiency nog zo'n 100 ernstige fouten in dat soort software kunnen zitten. In verschillende publicaties staan schattingen dat de faalkans van de Maeslantkeringssoftware wel 1 op 9 tot 1 op 11 is. Volgens staatssecretaris Schultz is een faalkans van 1 op 100 het hoogst haalbare. Verontrustende cijfers als men kijkt naar het belang van de kering: hij beschermt 1,3 miljoen mensen en honderden miljarden aan economische belangen. Dit is des te verontrustender als men bedenkt dat life-critical software ook gebruikt wordt in kerncentrales en luchtverkeerssystemen. We kunnen ons dan ook nauwelijks conflicterende getallen veroorloven en moeten dus preciezer weten of de burger zich echt zorgen moet gaan maken of hij de voeten droog houdt.

De Maeslantkeringssoftware was een prestigieus project in het ICT-landschap. Bij systemen met een dergelijke impact hoort een cijfermatig onderbouwde verwachting van de betrouwbaarheid. Zelfs het falen van een veiligheidskritisch systeem is nooit uit te sluiten, maar als gemeenschap moet men deze kans wel op een ALARP (As Low As Reasonably Possible)-niveau brengen. Zo ook de Maeslantkering. Blijkens verscheidene publicaties voldoet dit systeem aan het SIL-4 level van de Internationale standaard IEC 61508. Een werkwijze en niveau die ook gebruikt worden voor industriële systemen, kerncentrales en luchtverkeersleidingssystemen. Deze werkwijze wordt door controlerende instanties, wetgever en industrie ook als normgevend geaccepteerd. Volgens deze normering zou men bij SIL-4 uit moeten kunnen gaan van een faalkans van een op tienduizend. Een getal dat, blijkens de acceptatie van het systeem, voor de overheid acceptabel is. De afgelopen weken zijn er verschillende getallen genoemd. Deze getallen wijken enkele factoren af van de hierboven genoemde getallen. We moeten ons dus afvragen wat er echt aan de hand is. De grote vraag is wat nu de werkelijke faalkans is. De feitelijke situatie is dat de eigenaren van de Maeslantke-

ringsoftware dat kennelijk niet kunnen aantonen.

IEC 61508 baseert zijn getallen op een veronderstelde relatie tussen ontwikkelproces en een te verwachte betrouwbaarheid. In de praktijk worden op SIL-4 vele bekende vormen van kwaliteitsverbetering vereist: van reviews tot formeel wiskundige bewijzen. De zwakte van deze benadering is dat de relatie tussen het ontwikkelproces en de betrouwbaarheid van het systeem niet kwantitatief onderbouwd is. Dat er een positief verband zou zijn, wordt nauwelijks bestreden maar over de feitelijke faalkans van een opgeleverd systeem is bij certificering via IEC 612508 niets te zeggen.

Over de feitelijke faalkans is bij certificering via IEC 612508 niets te zeggen

In een benchmark kwamen we er op uit dat de software van de Maeslantkering nog 100 serieus te nemen fouten zou kunnen bevatten. Dit getal is gebaseerd op onderzoek van Capers Jones. We hebben hierbij de volgende redenering gevolgd: De Maeslantkering bevat 450.000 regels C++ code, ontwikkeld in de jaren '90 als systeemsoftware. Omgerekend naar functiepunten met de backfiring power van 53 levert dat 8490 functiepunten op. Zeg maar in de 10K functiepuntenrange. De beste systeemsoftware van die omvang uit die periode wordt opgeleverd met zo'n 800 fouten waarvan 100 ernstig, aldus de benchmarks van Capers Jones. Het is echter goed mogelijk via Capers Jones op een ander, gunstiger getal uit te komen. Allereerst kan men zich afvragen of die 450.000 regels code in een monolithisch project tot een systeem zijn gebouwd. Kijkt men naar gepubliceerde validatierapporten van het Telematica Instituut, dan ziet men dat er 200.000 regels code operationeel worden ingezet en 250.000 regels voor simulaties, tes-

ten en support. Alhoewel dit sterk gecorrleerde systemen zijn, zouden er argumenten kunnen zijn die een berekening over 200.000 regels rechtvaardigen. Maar daarover is niets bekend. Men ziet bovendien dat die 200.000 regels verdeeld zijn over 9 subsystemen waarvan er 3 als kritisch worden beschouwd en er een subsysteem als vangnet dient voor de andere subsystemen. Maar de IEC 61508 moet gelden voor het geheel en niet voor delen van het systeem. Dus de fouten die tot ernstig falen kunnen leiden hoeven niet per se in de kritische subsystemen te zitten.

Feit is wel dat bij oplevering de klant nog 119 fouten aantrof, waarvan 31 in de 3 kritische modules, dus foutvrij waren ze in ieder geval niet bij oplevering. Ook kan men zich ernstig afvragen of het gebruikte werkproces zich als het ontwikkelproces van 'systeemsoftware' laat classificeren. Alhoewel het toepassingsdomein overeenstemt, zijn de gebruikte werkwijzen wellicht meer in

overeenstemming met andere domeinen. Een klassificatie als 'military software' zou dus ook een optie zijn. Maakt men met deze aannames de rekensom overnieuw dan komt men uit op ongeveer 49 serieus te nemen problemen in de software.

De software van de Maeslantkering is een van de weinige projecten die met behulp van formele specificatie en verificatie is ontwikkeld. Door Bowen en Hinchey wordt in 'IEEE computer' van Januari 2006 opgemerkt dat deze methodieken niet of nauwelijks gebruikt worden, uitgezonderd de zeer veiligheidskritische systemen. Maar ze hebben wel een zeer sterk positief effect op de betrouwbaarheid van systemen. Het is aannemelijk dat de statistische gegevens van Capers Jones niet dit soort projecten bevatten en daarmee een negatiever beeld schetsen dan de werkelijkheid. Echter, kosten van de Maeslantkeringssoftware zijn slechts de helft van best-in-classkosten van Jones' benchmarks. Dus als er juist meer ef-

fort in zit, kan dat niet kloppen met de prijs. Een andere aanpak zou het beschouwen van het aantal fouten over de tijd zijn. Het validatierapport van het Telematica Instituut geeft aan dat er softwarefouten zijn ontdekt. Maar er is geen indicatie over het verloop van de fouten door de tijd, waardoor deze gegevens niet bruikbaar zijn voor tijdreeksanalyse van het foutverloop. Wel weten we dat er fouten gevonden zijn door de klant en dat de defect removal efficiency van de leverancier daarmee zo'n 93 procent was. En dat is slechter dan best-in-class military software (99,5 procent). Een conclusie die we veilig lijken te kunnen trekken is dat er vrijwel zeker nog overgebleven softwarefouten in het systeem zitten.

Hoe goed zal de software van de Maeslantkering functioneren? Dat lijkt niemand tot op de dag van vandaag echt te weten. Wat we wel weten is dat een kering zonder software geen optie is vanwege de complexe besturing die zo nauw luistert dat

mensenhanden daar geen plaats hebben. Des te meer reden dus om werkelijk grip te krijgen op deze complexe problematiek. Anders hebben we bij de volgende keuring van de Maeslantkering weer hetzelfde pandemonium. Kern van de problemen is ons insziens ook dat de overheid bereid moet zijn te begrijpen dat er voor haar een taak ligt in het afdwingen hiervan en het faciliteren van het noodzakelijk wetenschappelijk onderzoek. Wij zijn in ieder geval bereid om significant energie te steken in dit soort vragen. Als de overheid openheid van zaken geeft omtrent de Maeslantkering is er alvast een begin.

JAN FRISO GROOTE EN CHRIS VERHOEF

AG · 07-04-'06

Jan Friso Groote is hoogleraar Embedded Systemen aan de Technische Universiteit Eindhoven.

Chris Verhoef is hoogleraar Informatica aan de Vrije Universiteit Amsterdam. Tevens schrijft hij maandelijks een column voor de Automatisering Gids.

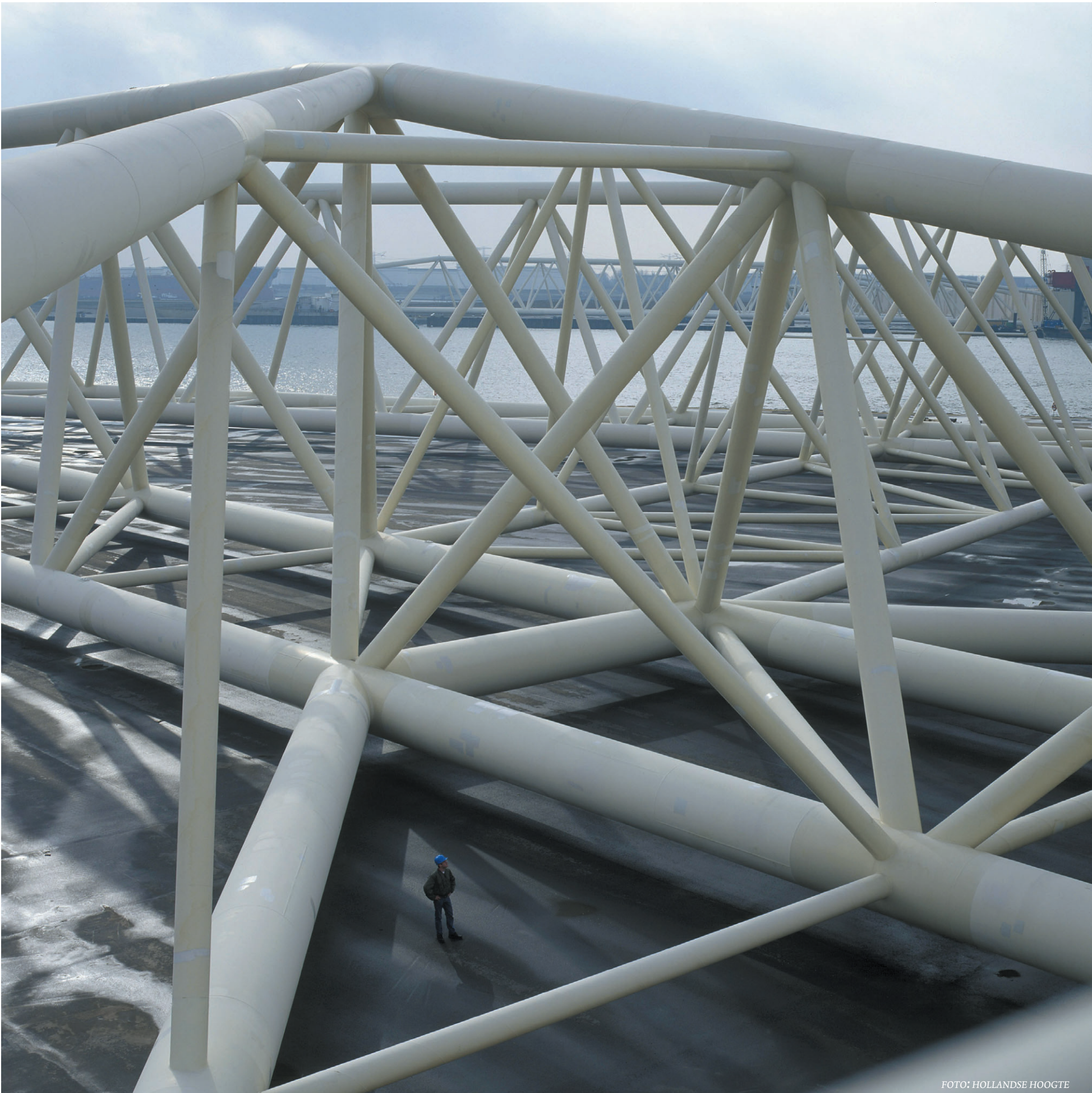


FOTO: HOLLANDSE HOOGTE