Emmanuel Besluau (Duquesne Group) : «La panne de Delta Airlines contient des zones d'ombre»

Le 8 août, la compagnie aérienne Delta Airlines connaissait la panne la plus importante de son histoire : plus de 2 100 vols annulés, d'innombrables retards, une image sévèrement écornée et des millions de dollars perdus. Selon les explications du Pdg de l'entreprise, Ed Bastian, ce chaos qui a duré 2 jours et demi est dû à la simple défaillance d'un équipement électrique et à une erreur dans l'exploitation du datacenter : 300 serveurs qui n'étaient pas reliés à l'alimentation de secours.

Associé du cabinet de conseils Duquesne Group et spécialiste de continuité d'activité, Emmanuel Besluau revient sur cette panne. Et pointe les zones d'ombre du scénario dessiné par le Pdg de Delta.

Silicon.fr : Quels enseignements tirez-vous de la panne qui a handicapé le fonctionnement de la compagnie aérienne Delta pendant 2 jours et demi ?

Emmanuel Besluau : J'observe d'abord que deux versions ont circulé. La première expliquait que la panne trouvait son origine dans un équipement électrique passif, un power switch. Puis, on a parlé d'un module de contrôle de l'électricité, un équipement qui va répartir les arrivées électriques pour optimiser la consommation ou l'ajuster au besoin. Ce type d'équipements pourrait effectivement constituer un point unique de défaillance dans un datacenter si l'on n'y prend pas garde. Mais ce qui est plus surprenant, c'est qu'on nous explique que ce module a pris feu. Ce qui est assez inhabituel. J'ai certes déjà connu un cas similaire, mais c'était dans une petite salle informatique. Dans un grand datacenter comme doit en posséder Delta, cela soulève quelques questions. Car, avant de prendre feu, la température de l'équipement a dû augmenter, pendant un certain temps ce qui a dû se voir. Or, on a l'impression que Delta prend conscience du problème au moment de l'incendie. Y-a-il eu une défaillance dans le monitoring de la salle, ce qui pourrait expliquer que cette augmentation de température soit restée indétectée ? Difficile de répondre catégoriquement à cette question, mais on peut observer que la panne intervient en pleine nuit, dans le courant du mois d'août, période de l'année où des compétences clefs au sein des entreprises sont habituellement en congés…

Passée cette défaillance de l'équipement électrique, les conséquences paraissent exagérément

élevées. Aller jusqu'à une coupure par Georgia Power (le fournisseur d'électricité de Delta, NDLR), cela signifie l'existence d'une surtension ou d'un autre problème : comment cela a-t-il bien pu se produire ? Est-ce un effet ou une cause ? En tous cas, la réaction de Georgia Power semble, dans le cas présent, normale : le fournisseur d'électricité détecte une surtension, ce qui provoque la coupure de l'alimentation primaire du datacenter.

Au-delà de l'origine électrique de la panne elle-même, Delta explique que ses difficultés opérationnelles sont nées du fait que 300 de ses serveurs n'étaient pas reliés à l'alimentation de secours. N'est-ce pas surprenant ?

E.B. : C'est un autre point intéressant. En effet, dans les datacenters modernes, chaque châssis ou serveur est relié à au moins deux alimentations électriques parallèles, chacune étant capable de faire face à une coupure de sa jumelle. Parfois, comme c'est le cas en France, le fournisseur d'électricité est unique, mais les deux voies sont de toute façon reliées à des onduleurs et des générateurs différents. Dans le scénario Delta, certains serveurs – les fameuses 300 machines évoquées par le Pdg – n'étaient manifestement pas reliés à deux voies, mais à une seule. C'est une pratique courante, à laquelle on a recours afin de réaliser des économies. Il y a alors bien une seule alimentation, mais un switch assure le basculement entre la source d'énergie primaire et une source alternative en cas de panne. Par ailleurs, cette solution est réservée aux serveurs non critiques.

C'est pourquoi le scénario décrit par Delta comporte des zones d'ombre et me surprend quelque peu. Car même sur des serveurs mono-alimentés, les onduleurs et les batteries devaient alimenter toutes les voies et auraient dû offrir aux exploitants un temps de latence suffisant pour pour que les générateurs diesel démarrent et prennent le relais afin d'éviter un arrêt des serveurs. Généralement, cette autonomie varie entre 10 minutes et une heure. La panne décrite n'aurait donc pas dû avoir ces conséquences. Dès lors, comment un tel scénario a-t-il pu se produire ? Difficile de se prononcer. Mais on peut imaginer une proximité trop importante entre certains équipements, le feu causé par le module défectueux ayant alors pu endommager d'autres éléments. On a souvent l'impression que les exploitants ne réfléchissent pas suffisamment aux risques de ce type dans l'agencement de leurs salles. Des équipements redondants sont parfois placés l'un au-dessus de l'autre… Une autre hypothèse, c'est que l'incident n'a pas été bien détecté, et que les générateurs (ou l'un d'entre eux) n'ont pas démarré dans le temps imparti.

Toujours est-il que la séquence d'événements telle que présentée, aboutissant à 300 serveurs non alimentés par le secours, est très curieuse et ne peut pas se produire dans un datacenter de type III par exemple. Ou alors Delta a laissé la situation dégénérer avec des générateurs en panne ou sous-dimensionnés…

Comment expliquer que l'arrêt de ces 300 serveurs, qui devaient être les moins critiques de l'entreprise en théorie, ait provoqué une telle pagaille dans l'ensemble du système d'information ?

E.B. : Le Pdg de Delta évoque la dépendance du système d'information aux systèmes mis à l'arrêt. Peut-être, au fil du temps, les équipes IT ont-elles introduit des dépendances aux applications placées sur ces serveurs mal protégés électriquement, par exemple en déplaçant des VM ? J'ai par exemple déjà vu deux bases de données en réplication active qui s'attendaient l'une l'autre,

annulant tout bénéfice de la protection mise en place. Des choix d'architecture peuvent créer des dépendances qu'on souhaitait éviter à l'origine. C'est pourquoi, quand on déplace une VM, il faut toujours connaître les caractéristiques de continuité du serveur sur lesquels on va les positionner.

Le temps nécessaire pour redémarrer l'activité vous a-t-il surpris ?

E.B. : On est ici dans un monde, l'aérien, où les affectations de vols, celles des personnels, l'embarquement des passagers, etc., sont gérés totalement via des systèmes d'information. Dès que ces systèmes ne sont plus synchronisés avec le réel, avec ce qui se passe sur le terrain, la remise en marche peut devenir d'une complexité diabolique. Surtout si on repart d'une sauvegarde ancienne. Il faut parfois écrire certains champs à la main, directement dans la base de données, pour contourner les blocages applicatifs. Je dirais même que repartir en 2 jours et demi, c'est plutôt bien ! Même si j'ai tendance à penser que la restauration des systèmes n'était pas totalement achevée au moment où la compagnie a déclaré l'incident terminé. Reprendre les activités sur un système qui s'est crashé – comme c'est le cas ici -, cela n'a rien à voir avec le test d'un plan de continuité d'activité (PCA), exercice où on prend la peine de fermer proprement les applications et bases de données avant de tester leur redémarrage…

Par contre, je suis un peu plus surpris de constater que le Pdg de Delta semble découvrir les correspondances entre les exigences métier de son entreprise et les moyens techniques mis en œuvre. Comme si l'analyse de criticité des systèmes d'information (ou BIA) n'avait pas été menée. Quand je lis ses déclarations dans la presse américaine, je relève 5 ou 6 problèmes de non-conformité à la norme ISO 22301 (norme internationale de continuité de l'activité, NDLR).

En dehors de cette étude de criticité, quels facteurs s'avèrent critiques dans le redémarrage de l'activité après une panne de cette nature ?

E.B. : La prise de conscience des conséquences d'un incident d'exploitation et la rapidité de celle-ci sont essentielles pour limiter les dégâts. Pour un exploitant informatique, c'est un exercice très difficile. Dans ce genre de cas, la capacité à dresser un bilan fiable, 15 minutes après la panne, des systèmes qui fonctionnent et de ceux qui ne fonctionnent pas s'avère crucial. Dans le cas de Delta, je serais curieux de savoir qui a été le premier averti du problème sur les systèmes opérationnels de la compagnie. En espérant que ce ne soit pas un passager via les écrans des aéroports…

Emmanuel Besluau est l'auteur de « Management de la continuité d'activité » aux éditions Eyrolles.

A lire aussi :

Un bug logiciel bloque le trafic aérien aux Etats-Unis

Delta équipe ses 11 000 pilotes de tablettes Microsoft Surface 2