

多様なシステムの安定運用ノウハウ ～トラブルの予防と早期発見～

谷口麻梨香^{A)}、○福田優子^{A)}、宇佐美賢子^{A)}、林麗美^{A)}、安原裕子^{B)}

^{A)} 大阪大学 レーザーエネルギー学研究センター

^{B)} 静岡大学 情報学部 技術部

はじめに

大阪大学レーザーエネルギー学研究センター高性能計算機室では、大型計算機システムをはじめ、メール、ネットワークシステムなど多様なシステムを運用しており、管理機器は120台以上である。台数が増えるとトラブルも増えるが、そのほとんどを未然に防ぐことや、ユーザが気付く前に対処することが出来ている。監視サーバや自動ツールによる情報収集などを駆使し、極力安価に効率よく、且つサービスレベルを落とさないようにシステムを運用する工夫を長年行ってきた。今回はその中でも、トラブルの早期発見と原因切り分けについての工夫を中心に報告する。

1 管理システムの概要

我々が運用管理を行っているシステムの概要を図1に示す。SX9 やワークステーションから構成される実験用コンピュータシステム、ネットワーク機器、実験データベースシステム、メール・WEB・業務支援データベースシステムのサーバなど、合計120台以上の機器の運用管理を行っている。



図1 高性能計算機管理室の運用管理システム概要

管理すべきOSや機器の種類は年々増える傾向にある。さらに、バージョンによる操作性や設定ファイルの違いなど、トラブル対応を含む運用管理は複雑さを増している。大学では、利用者の環境も様々であり、OSの種類やバージョンによって設定方法や、トラブルの症状が異なる場合などもあり、利用者支援も難しさを増している。

2 トラブルデータベース

システム運用をしていると必ずトラブルが発生する。一度しか起こらないものもあれば、何度も起こるものもある。復旧方法が分かっているものもあれば、原因不明のものもある。トラブルやクレームは宝の山と

きたものだということがわかった。

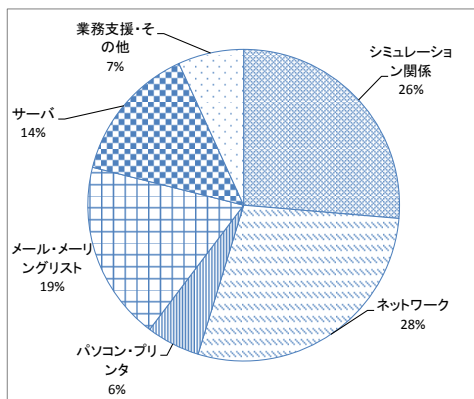


図4 トラブルの種類と割合

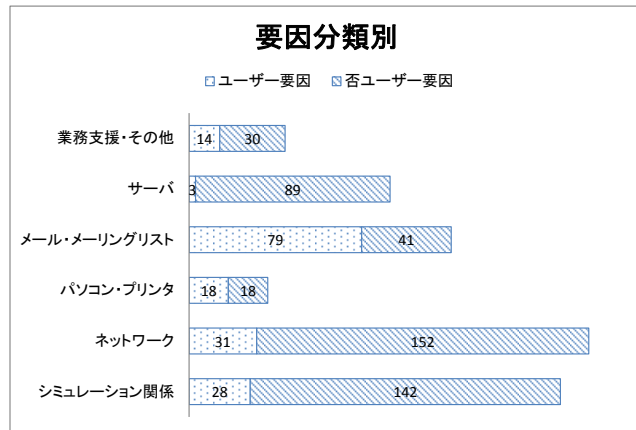
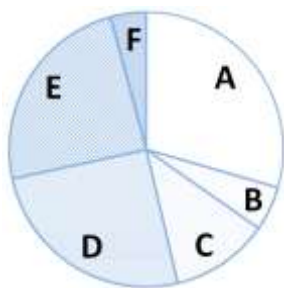


図5 要因別分類

これは、経験と毎日の定型作業に加え、長年工夫してきた定型ツールからの通知メールやログ、監視サーバからの警告メールによって発見できたものである。この自作定型ツールと監視サーバについて、4章と5章で説明する。



A	ユーザーからの質問や相談	28	29%	24%
B	ユーザーからの連絡により発覚したトラブル	5	5%	
C	ユーザーからの連絡により発覚したシミュレーション系トラブル	11	12%	76%
D	CMPが発見して対処したトラブル	24	25%	
E	監視サーバや自動通報により発見したトラブル	23	24%	
F	ODINSからの連絡	4	4%	
合計		95		

図6 1年間のトラブル発見の傾向

4 自作定型ツール

1996年3月にUNIXによる分散処理システムを導入して以来、管理する計算機の台数が増加し、利用状況や通信状況などを自動的に監視、採取するツール^{[2][3]}を作成してきた。自作定型ツールには下記のようなものがある。

- ping 死活監視
- ジョブ情報採取
- ディスク情報採取、及び使用容量監視
- ログイン情報採取
- CPU 使用状況採取
- プロセス情報採取
- quota 情報採取、及び使用容量監視

この自作定型ツールは、我々が管理しているLINUX(UNIX)マシンのほぼ全てで動いており、ping 死活監視は30分毎、ログイン情報採取は一日一度などツールの特性によりcronで実行する間隔を変えている。また、ツールやログは、可能な限りNFSで共有しているファイルサーバ上に集め、変更や確認の際に、個々のサーバにログインしなくても、一カ所で作業できるように工夫している。

ディスク使用量や quota 使用量については、サーバ毎に閾値を設定し、超えた場合はメールで管理者に通知するとともに、ログイン時に管理者および利用者に警告メッセージを表示するなど、早めに通知することで、実際にディスクがパンクする前に対処できるように工夫している。例えば、ディスク容量が閾値を超えたと通知が来た際には、容量の大きいファイルが無いかなどを確認するが、同時に、徐々に使用量が増えて閾値を超えたのか、それとも急激な増加により超えたのかという確認も重要である。毎日ディスク情報を採取して残すことで、どのような増え方をしたのかを確認し、状況に応じて対処を行っている。また、この他

にもシステムログに特定のメッセージが出た場合は、警告メールを送るなど、必要に応じてツールを作成している。

日次作業	メールサーバ	バックアップログの確認 システム負荷チェック メール配送状況チェック メール容量チェック
	STARSSシステム	ディスク使用量確認 ユーザ利用制限確認 自動実行ツールのログ確認 バックアップログの確認 ジョブ実行状況確認 システム負荷チェック
	CMCシステム	ディスク使用量確認 ジョブ実行状況確認 バックアップログの確認
	スパムメール対策機	パターンファイル更新の確認 エンジン更新の確認
	監視サーバ	各サーバでエラーが出ていないかのチェック
	DHCPサーバ	IPアドレス使用状況の確認
	ファイアウォール	ログバックアップ確認
	webサーバ	バックアップログの確認
	ライセンスサーバ	使用ライセンス数の確認
	プリンタ	消耗品状況の確認
空調	マシン室空調や異常ランプ点灯などの確認	
安全衛生	安全衛生管理チェックシートの記入	

図7 日時定型作業一覧

月次作業	メールサーバ	ログイン情報採取 メールログ集計
	STARSSシステム	ログイン情報採取 ジョブ情報採取、グラフ化 稼働時間チェック バックアップログの確認
	CMCシステム	ジョブ情報採取、グラフ化 フェアシェア利用状況採取 稼働時間チェック
	スパムメール対策機	集計情報プリント
	監視サーバ	バックアップログの確認
	DHCPサーバ	IPアドレス、MACアドレス使用状況集計、グラフ化
	認証サーバ	利用ユーザー集計、グラフ化
	共同研究者認証サーバ	利用ユーザー集計、グラフ化
	ログインサーバ	ログイン情報採取
	ネットワーク	登録マシン集計、グラフ化
	ファイアウォール	ログイン情報集計 connファイルバックアップ
	webサーバ	アクセスログ集計 情報採取、グラフ化
	業務データベースサーバ	データベースバックアップ
	CMFファイルサーバ	ファイルバックアップ
	実験データベースサーバ	ログイン情報採取、グラフ化
	空調など	加湿器、換気扇などのメンテナンス

図8 月次定型作業一覧

自作定型ツールで採取しているログ情報について、警告の通知はメールで来るからと、安心するのはよくない。何らかのエラーにより、cron やツールが動いていないという可能性もあるため、曜日ごとにサーバを決め、週に一度はサーバにログインしてログを確認することになっている。図7と図8に、日時と月次の定型作業の項目を示す。月次作業としては、ジョブ集計やログイン集計、メールログの集計など、主に利用状況の集計を行っている。基本的には、集計は、自作のツールにより自動的に行われるので、問題がないかの確認や、システムの運用状況や利用状況の確認に利用している。

5 監視サーバ

2009年3月にSGI社製のPMT (Performance Monitoring Tool)^[4]を導入した。監視ツールは、Nagiosなどオープンソースのものも色々あるが、英語表示であったり、SNMPによる監視やエージェントのインストールが必要であったりすることと、自作定型ツールによる監視を行っているため必要性を強く感じていなかったこともあり、導入を踏みとどまっていた。

PMTは、日本語表示であること、導入や保守などのサポートがしっかりしており、エージェントのインストールなどは不要であり導入が容易であること、NQS バッチジョブ実行状況の採取も出来ること、CygwinをインストールすればWindowsも同様に監視できることなど、我々のニーズに合致していると判断し、導入することとなった。

PMTの監視イメージを図9に示す。監視対象ホストへは監視サーバからSSHログインできればよい。また、監視サーバから直接ログインできない場合は、ヘッドクラスタ経由で監視することもできる。SSHログインできないネットワーク機器については、ping 死活監視が可能である。

PMTでは以下のような監視が可能である。監視対象ホスト毎に、同じ手順で細かく設定することができる。

- ping 死活監視

- ディスク・メモリ・CPUなどのシステムリソースの監視
- 監視サーバの時刻と監視対象ホストの時刻のずれがないかを監視
- プロセスが動いているか、いくつあるかの監視
- プロセスの使用 CPU 時間や使用メモリの監視
- 監視対象サーバから別のサーバに対して POP や FTP や HTTP ポート応答の監視
- システムログに、特定のメッセージが出たら通知
- 任意のファイルの有無や、容量の上限下限の監視

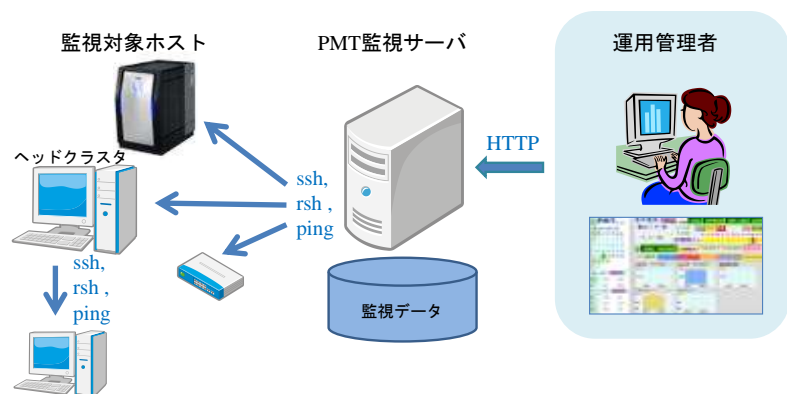


図9 PMT 監視イメージ

監視項目ごとに異常検知の閾値を設定でき、緊急レベル、警告レベル、報告レベルなどの通知レベルを設定することができる。通知先についても、監視対象ホスト毎に設定できるので、他の機器管理担当者にも直接メール通知することが可能である。また、平日の9時～17時のみに通知を送る時間帯も設定できる。

root 権限での監視は危険を伴うので、監視対象ホストへは PMT 専用ユー

ーザで SSH ログインし、この PMT 専用ユーザの権限で監視が行われる。従って、例えばシステムログなど root 権限でないと見えないログを監視したい場合は、PMT 専用ユーザでも見えるように権限を変更する必要がある。

監視画面を図 10 に示す。ジョブやリソースの状態がグラフ化されており、一日、一週間、一か月、一年の単位で表示できるため、傾向をつかみやすい。日付を選択すると、遡ってその日の状態を見ることもできる。トラブル時には表示が赤く変わり、通知メールも送信されるため、異常発生時に速やかに気が付くことができる。また、保守などで停止中の場合は、ステータスを保守中に変更すると図 10 下部に示すように、黄色い表示に変わり、通知メールを抑止することができる。復旧後には、ステータスを元に戻すとすぐに監視を再開できる。図 11 は、あるサーバのルート領域のディスクがあふれた時の状況を表示したものであるが、午前 4 時に突然 CPU 使用率、ロードアベレージ、メモリ使用量が増え、ディスクも一気に 100% になった様子を視覚的に確認することができ、早急にトラブル対応を行うことができたときの例である。図 12 は、あるサーバの 1 か月間のメモリ使用状況を表示したものである。徐々にではあるが、あるタイミングで段階的にメモリ



図10 PMT 監視画面



図 1.1 ディスクがあふれた時の表示



図 1.2 1 か月間のメモリ使用量の状況表示

使用量が増えていることを確認することができ、トラブルの原因究明に役立てることができたときのものである。このように、サーバの状況を把握する作業は、経験と時間のかかる作業であるが、PMT 導入により効率化されたことも、大きな導入効果であった。

さらに、停電明けの全システム起動後などの、全てのサービスの確認作業や、あるプロセスが止まったことなどを確認するのは時間のかかる作業であるが、PMT に設定しておけば、動いているべきプロセスがないということも自動通知してくれるので、効率化された作業である。特に、ユーザの利用頻度があまり高くないサービス

は止まっていることに気付くのが遅れることがあったが、PMT 導入後はスムーズに発見できるようになった。

現在、サーバやプリンタ約 50 台、ネットワーク機器約 40 台を監視している。サーバには Linux、Windows、MacXserve などがある。監視対象ホストの設定は、Web インターフェースで簡単に行うことができる。監視サーバ導入の一年後の 2010 年 3 月に計算機システムの更新があり、新しいサーバの監視設定をする必要があったが、新規の監視対象ホストの設定作業は、自分達だけで随時行うことが出来たため、新システムのサーバの監視も速やかに開始することができた。図 10 に示した保守中のマシンは、前システムのサーバであり、移行が完了しサービスを終了したマシンは、随時保守中に設定するだけでよかったので、移行作業もスムーズであった。

6 自作定型ツールと監視サーバを活用した原因の切り分け

2 章で示したように、自作定型ツールと監視サーバの監視によって、トラブルのほとんどをユーザからの報告を受ける前に、発見出来るようになった。監視サーバを導入した際には、監視項目が自作定型ツールでの情報収集と重複するため、自作定型ツールの運用は廃止しようかと考えていたが、実際に運用してみると二重チェックになり、トラブルの際に色々な通知が来ることで原因の切り分けにも役立つことがあったため、どちらからも監視する運用を続けることにした。監視サーバが接続されているネットワークセグメントとは異なるセグメントに接続されている監視対象ホストもあり、特にネットワーク通信トラブルの場合に、複数から検知メールが届くことで、トラブル箇所の特定が以前より短時間で行えるようになった。ホスト単独のトラブルなのか、ネットワークのトラブルなのかの切り分けの参考にもなっている。

しかし、運用管理者は毎日多数のシステムからのメールを受け取っているため、その中で重要なメールを見落とさないことも重要である。これには、毎日どのようなメールが来ているかを確認しておき、いつもと異なるメールが届いた時に早急に気付くことができるように、経験と注意力が必要である。

また、来なければいけないはずのメールが届いていないことに気付くことも重要である。監視サーバ自体にトラブルが起これば、通知メールが送れなくなる可能性もあるため、PMT では毎週一度「通知試験」メールが送られてきている。自作定型ツールの動作確認は、保守などでサーバを止めたときなどに、来るべきメー

ルがすべて届いているかを確認するようにしている。

最後に、原因切り分けのために、正常時の状態を保存しておくことと、システム概要図を必ず書くことを心掛けている。運用を開始すると、サーバごとのディスクのパーティションや使用量、プロセスの状態、クローンの設定情報、ネットワーク設定などのシステムの情報を極力電子的に保存するようにしている。我々はこれをカルテと呼んでいる。正常時にランプはどんな点灯をしているかもトラブル時に有益な情報であり、デジカメで撮影するなどして保存しておきたいと思っているが、すべてを残すことはまだできていない。テキストデータや、詳細なスペックの数値を示した表のみではなく、システム概要図を作成し、絵の形で全体の把握を誰でも容易にできるようにしておくことは、トラブル時の切り分けや、対処方針の決定に役立っている。この経験を生かして、最近ではユーザにも、居室や実験室のネットワーク配線図を作成しておくことを推奨し、トラブル時には、配線図を見て切り分けを行うように教育している。

7 最後に

長年の経験に基づき、自作のツールや市販の監視ソフトを用いて、安価で効率が良くなるように工夫を凝らして運用管理を行っている。トラブルデータベースに蓄積したデータを、トラブル発見の傾向によって集計してみることで、監視サーバや自作ツールからの通知で発見したトラブルの割合が多いことに改めて気が付いた。我々のこのシステム運用管理の経験は、他の機関の運用管理にも参考になるのではないかと考えている。

ともすると、「何をしているのかわからない」と思われがちな運用管理の仕事だが、ユーザが気付く前に対処しているトラブルがたくさんあり、その予防と早期発見のために努力していることを理解してもらうことも必要ではないかと感じている。

8 謝辞

本報告の機会を与えていただいた、自然科学研究機構分子科学研究所の皆様に感謝いたします。

9 参考文献

[1] 谷口麻梨香、福田優子、宇佐美賢子、安原裕子、“システム更新に伴うアカウント統一”、平成 21 年度高エネルギー加速研究機構技術研究会、5-009 <http://www-eng.kek.jp/meeting09/proceedings/pdf/h21g5009.pdf>

[2] 田村篤和、岡本匡代、福田優子、斉藤昌樹、“分散処理システムにおける運用の自動化の試み”、1998 年度 高エネルギー加速研究機構技術研究会報告集、279

[3] 広崎結里、福田優子、谷口麻梨香、斉藤昌樹、長友英夫、西原功修 ” 計算機システムについて” 平成 16 年度 大阪大学総合技術研究会報告集、P5-13 (2005)

[4] <http://www.sgi.co.jp/products/pmt/> PMT(Performance Monitoring Tool) 日本 SGI 株式会社