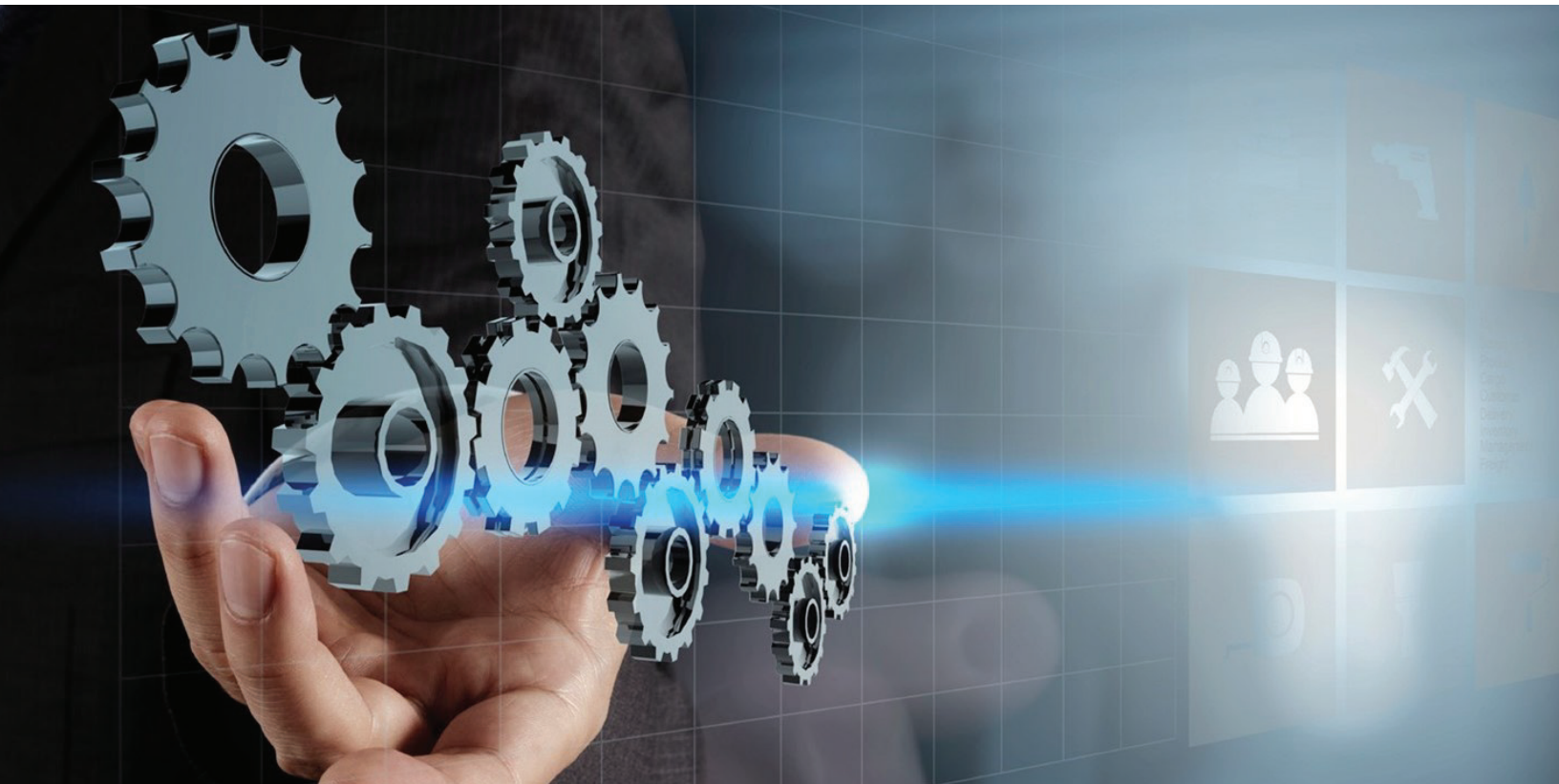


Engaged Engineering

Building trusted partnership to increase service reliability



Mindtree
Welcome to possible



Our customers are aspiring to meet the benchmarks¹ set by high performing IT organizations. Benchmarks include on demand deployments, less than 1 hr. lead-time for changes and less than 1 hr. mean time to repair (MTTR).

While our customers are considering these North Star goals, high downtime costs are a challenge. Cost of yearly downtime² is estimated to be in the range of \$1.25 billion to \$2.5 billion for Fortune 1000 companies, with an average hourly cost of infrastructure failure at \$100k and an average hourly cost of critical application failure at \$500,000 to \$1 million.

Most of the time, these high performing IT organizations have new, jazzier ways of running their production landscape. Our customers are super excited with the way Google, Facebook, and Amazon run their cloud scale production services.

Enterprises are trying to make sense of these cool approaches—DevOps, NoOps, DRI³, Service Engineering, Service Reliability Engineering. Like the elephant and blind men, everyone has their unique definition and a definitive, sure shot, fail-proof approach to adopt these methods.

This white paper details out our experiences in transforming our customer's way of running production and our best practices.

Benchmarks¹ of High Performing IT Organizations

- On demand deployments
- <1 hr. lead time for changes
- <1 hr. MTTR

Cost of downtime²

- Yearly Downtime costs for Fortune 1000 companies - \$1.25 B to \$2.5B
- Hourly cost of infrastructure failure - \$100k
- Hourly cost of critical application failure - \$500k to \$1M

¹ <https://puppet.com/resources/whitepaper/2016-state-of-devops-report>

² IDC Report titled "DevOps and the Cost of Downtime: Fortune 1000 Best Practice Metrics Quantified" and <https://devops.com/real-cost-downtime/>

³ Directly Responsible Individual

1. Organization set-up

Invariably our support customers are responsible for running the production services under an operations leadership. The parallel engineering team owns development of new features and services. Engineering and operations teams have different charters and key performance indicators (KPIs).

2. Two approaches to production support

We can broadly classify production support into the following categories.

- **Conventional way** – Two independent teams responsible for operations and engineering. They interface to mitigate outages, deploy releases and conduct post mortems. Typical challenges in this model are related to active collaboration between teams, trusted relationship based on shared data and workflow, different priorities and KPI objectives
- **New “cool” way** – One team responsible for both engineering and operations. In an ideal state, person who gathers the requirements develops and tests the code, builds and deploys the package and support it in production. In practice, the individual gets to deliver all these pieces, though not in a cyclical way. Engineers within this team are responsible to deliver ticket duty as part of their DRI calendar / shift roster. When they are not resolving tickets, they are responsible for their other roles

We see most production services running either in the conventional way or in between the conventional and the new way.

3. Our experiences with adopting the new “cool” way

We are sharing our experiences with two large distributed services clients and their unique adoption of the new cool way.

| Supported Production Service | Storage Cloud Services | Application Performance Monitoring as a Service |
|-------------------------------|---|--|
| Nature | Existing service. Contains multiple components within the service. | New Service. Contains multiple components within the service. |
| Conventional Way | A centralized production support team to support all cloud services including storage cloud. They mitigate most of the issues; resolve qualified monitoring alerts leveraging scripted resolution. Engineering is engaged only for code level issues and high severity outages | Not applicable |
| New “Cool” Way | A new Service Reliability Engineering (SRE) team was formed within the engineering team. SRE team had to mitigate outages and drive fixes for ageing bugs. Qualified monitoring events were received and resolved directly by the SRE team. Production support team not engaged to support the service. However, SRE team was not responsible for core software development. To fix code level issues, SRE teams still depend on the respective component engineering teams | DRI team formed within engineering. Team comprised of engineers dedicated to ticket duty and engineers responsible for both ticket duty and coding. DRI charter was to mitigate outages and drive fixes for ageing bugs. In a given shift, one of the engineers is staffed as a “Directly Responsible Individual”, based on the shift roster. Qualified monitoring events were received and resolved directly by this DRI team. To fix code level issues, DRI engaged the respective component engineering teams |
| Support Hops | SRE and engineering. No other operations layer | DRI and engineering. No other operations layer |
| Support Volume | High | Low to medium |
| Outcome of the new “cool” way | SRE was choked with high volume of production outages; unable to address all of them. Medium and low severity production incidents were simply ignored. They could not focus on driving fixes for ageing bugs as well. Instead of improving the service reliability, the new way was creating service disruptions. Now, they have transitioned back to their previous model of production support—SRE and engineering. SRE is now primarily focused on driving fixes for ageing bugs. | As the supported service was new, the new support model worked well. It has the construct to provide scalable support for the growing service |

4. Interim model and challenges

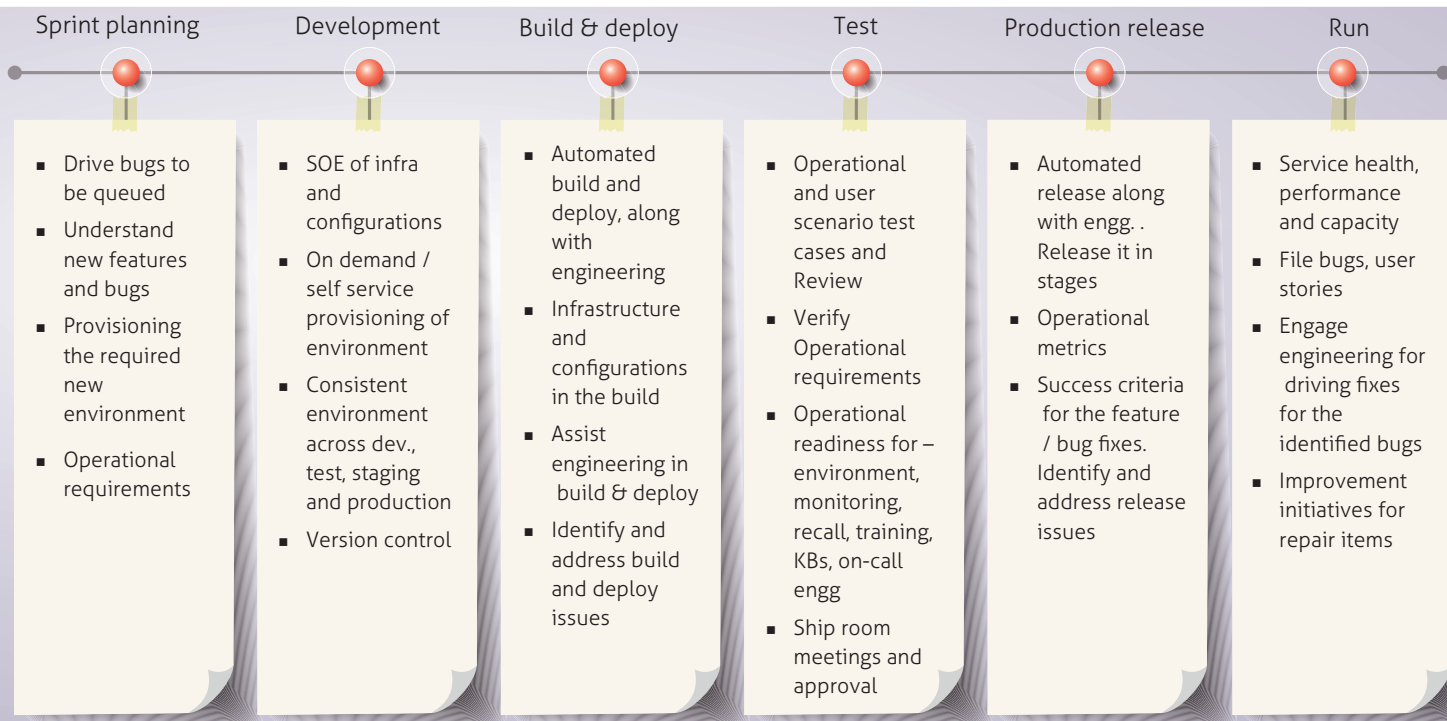
As our customers are trying to transform their production support, in the interim they are required to operate engineering and operations as two independent collaborating functions. Building a trusted partnership with the engineering team during this interim model is a challenge for the operations team. We see the following potential opportunities for the support team to establish a mutually beneficial relationship.

- Adding value during the engineering lifecycle from sprint planning, development, build and deploy, test, production release and run phases
- Influencing future software releases with operational and customer insights
- Improving the service reliability by driving fixes with a vibrant problem management

We will go into further detailing on these aspects in the following section.

4.1 Role of operations in the engineering lifecycle

We trust operations can deliver the following critical value adds during the "sprint planning to run" engineering lifecycle.



4.2 Improving service reliability

Reliability of the supported service directly affects operations. A buggy code can flood the operations team with multiple alerts. Until a permanent fix is developed, operations will need to continue apply workarounds and interim fixes.

We, as managed services partner, have objectively engaged with engineering to drive fixes for ageing bugs. As operations team, we have a wealth of alert data, customer impacting outages, open bugs and their age, social media sentiment analysis, support hours spent to mitigate the outages, and financial penalty paid to end customers. We connect these data points and make a meaningful case to drive fixes for the top bugs. A sample report given below:

| Service Component | Bug # | Bug Title | State | Age (Days) | Support Hrs. burnt | # Customer impacting Outages | Time to Recover (Total/Avg. Median) | Weekly Trend |
|----------------------------|--------|--|--------|------------|--------------------|------------------------------|-------------------------------------|--------------|
| Compute | 123456 | Driver update fails because of too many concurrent operations... | Active | 4 | 300 | 6 | 245/40/35 | NA |
| Reliability Infrastructure | 654321 | Runtime unhandled exception... | Active | 55 | 820 | 25 | 520/30/25 | |
| Fabric | 234654 | Workflow did not create certs... | Active | 30 | 350 | 10 | 450/45/30 | |

We engage the engineering leadership with such reports and present a 360° view connecting the DC level outage to the customer experience. Engineering leadership is able to take tradeoff decisions between building new features and fixing ageing bugs. This exercise benefits the support team, with increased permanent fixes and reduced outages to be mitigated by support team.

4.3 Establishing a vibrant problem management practice

As a key stakeholder in service reliability, we find that operations teams take ownership of problem management practice by collaborating with multiple engineering teams. It is the problem management team's charter to continuously reduce time to detect (TTD) and time to mitigate (TTM). In our experience, most operations and engineering teams see post mortems / RCA as only a process requirement. As they do not derive any value from their problem management exercise, it becomes more of a mandatory exercise. We had an opportunity to transform the problem management practice and deliver business outcomes.

4.3.1 Situation / Context

A large cloud service provider was looking for ideas to improve the overall service reliability and identified TTD and TTM as the core metrics. Mindtree was chosen partner to drive this initiative, with measurable goals to improve cloud service customer availability, minimize impact duration, repeat outages and multi-region failures. To assess the current state and improve, Mindtree's problem management team analyzed ~500 post mortems of their top outages and found 20% missing RCAs, 25% missing repairs, untrustworthy data, inaccurate timelines and other hygiene issues.

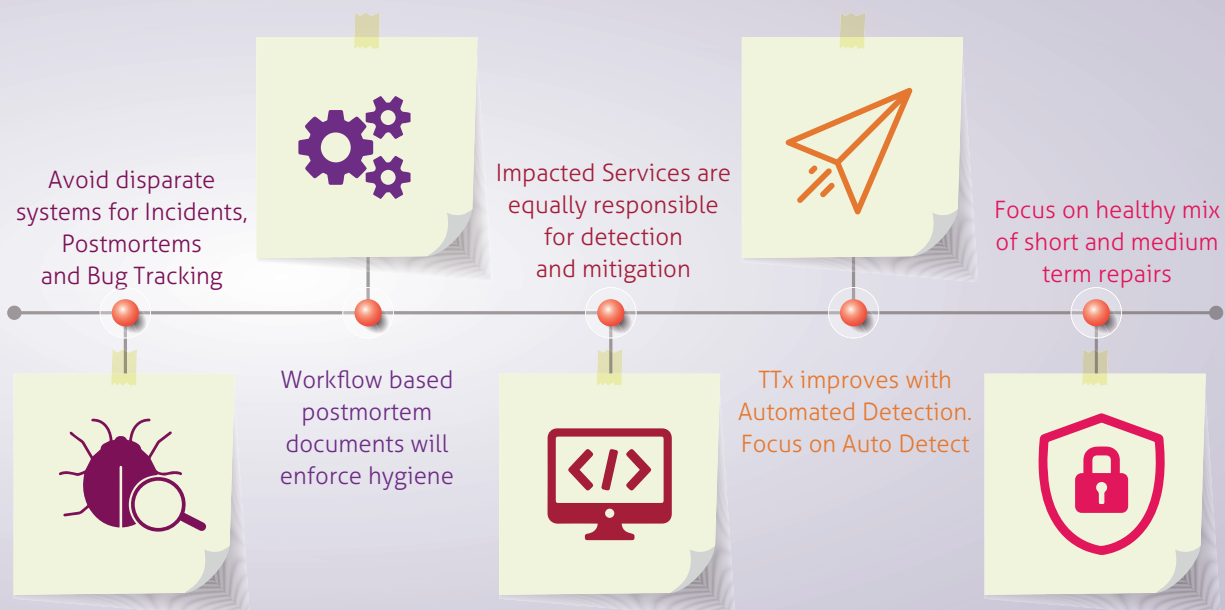
4.3.2 Solution approach

Mindtree engaged the stakeholders, identified their pain areas in developing, reviewing RCAs, driving repair items and connecting them to overall business objectives.

- To make it easy, RCA was moved from word documents to SharePoint portal, enriched automatically with timestamp, severity, impact and other details directly from the ticketing system
- RCA online quality check list was implemented, enabling the customer to assess their RCAs and address gaps proactively
- Post mortem quality dashboard was built to show RCA quality across different services, building peer pressure to improve the overall RCA quality
- Built a production improvement review dashboard that captured incomplete RCAs, TTD and TTM misses, TTD and TTM repair misses. Enabled operations leadership to have a healthy discussion with engineering teams on ageing bugs and making calls on tradeoff between developing new features and fixing ageing bugs
- Developed automatic scheduling of post mortem reviews with stakeholders, enabling them to prepare RCAs on time

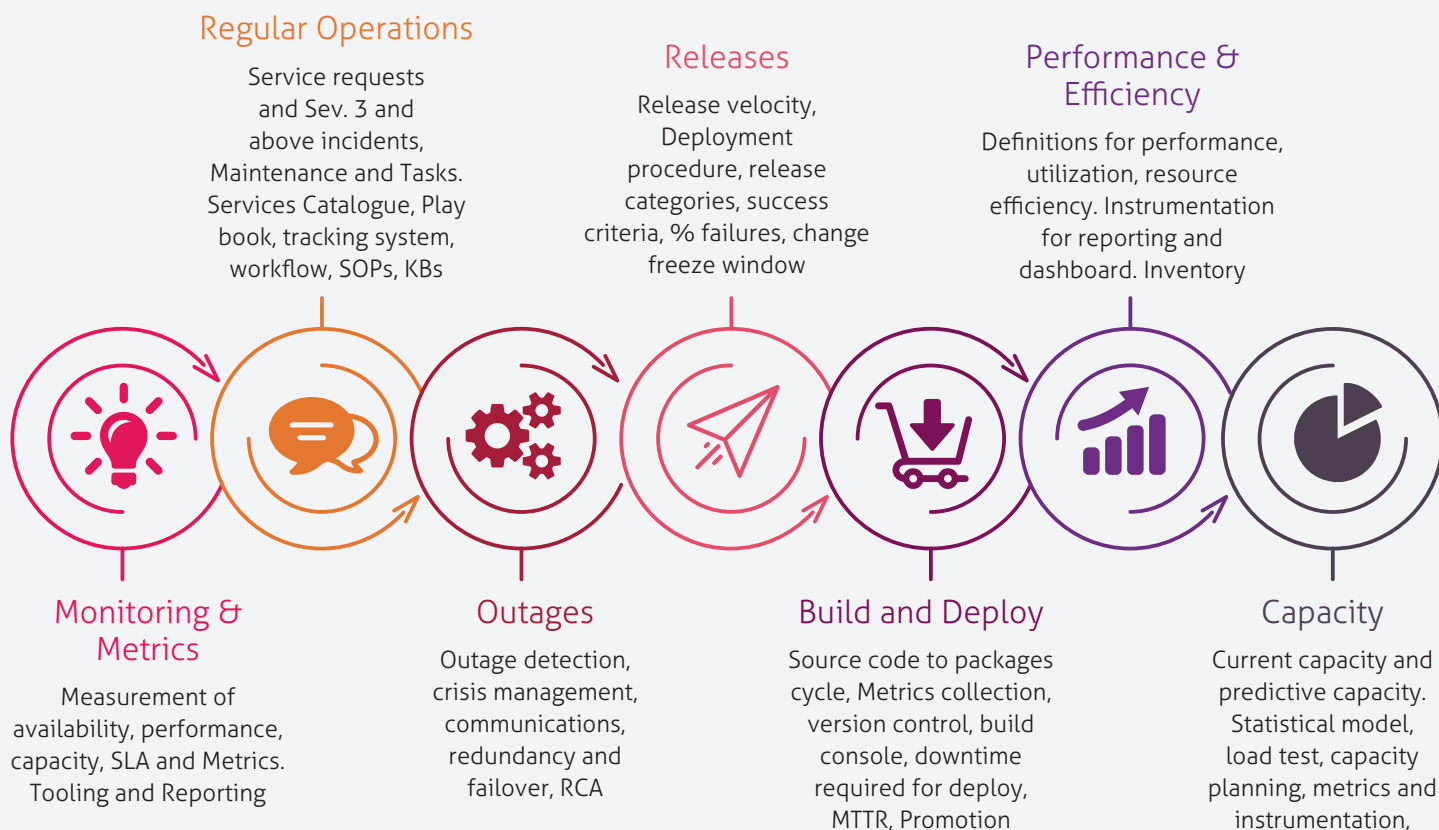
4.3.3 Lessons learnt

We are able to bring in structural changes; stakeholders are now seeing post mortems as an effective tool to improve service availability. We have learnt the following lessons during this journey:



5. Conclusion

Interested in this journey and wondering where to start? You could start by assessing your current state of operations across different parameters. Here is a sample model⁴:



We have deep experience in delivering production services and can be your trusted partner in your production services transformation journey.

About the Author

Ananthanatarajan Muthusamy is a Senior Technical Architect with Infrastructure Management Services (IMS) and is based out of Mindtree's Washington Delivery Center. With over 10 years of experience, he specializes in production support of large distributed computing services.

⁴ We leverage the assessment methodology provided by Thomas A Limoncelli, Strata R. Challup and Christina J. Hogan in their book "The Practice of Cloud System Administration" for assessing the operations maturity of our customers. We adopt the applicable services and metrics from this book to suit our customer's landscape.

About Mindtree

Mindtree (NSE: MINDTREE) delivers technology services and accelerates growth for Global 1000 companies by solving complex business challenges with breakthrough technical innovations. Mindtree specializes in e-commerce, mobility, cloud enablement, digital transformation, business intelligence, data analytics, testing, infrastructure, EAI and ERP solutions. We are among the fastest growing technology firms globally with more than 200 clients and offices in 14 countries.